


Genotype–environment associations reveal genes potentially linked to avian malaria infection in populations of an endemic island bird

Eleanor C. Sheppard¹  | Claudia A. Martin^{1,2}  | Claire Armstrong¹  |
Catalina González-Quevedo^{1,3} | Juan Carlos Illera⁴ | Alexander Suh^{1,5,6}  |
Lewis G. Spurgin¹  | David S. Richardson¹ 

¹School of Biological Sciences, University of East Anglia, Norfolk, UK

²Terrestrial Ecology Unit, Biology Department, Ghent University, Ghent, Belgium

³Grupo Ecología y Evolución de Vertebrados, Instituto de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia UdeA, Medellín, Colombia

⁴Biodiversity Research Institute (CSIC-Oviedo, University-Principality of Asturias), University of Oviedo, Mieres, Asturias, Spain

⁵Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change, Bonn, Germany

⁶Department of Organismal Biology – Systematic Biology, Evolutionary Biology Centre, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Correspondence

Eleanor C. Sheppard and David S. Richardson, School of Biological Sciences, University of East Anglia, Norwich Research Park, Norfolk NR4 7TJ, UK. Email: e.sheppard@uea.ac.uk and david.richardson@uea.ac.uk

Funding information

Natural Environment Research Council, Grant/Award Number: NE/L002582/1 and NE/S007334/1; Regional Government of Asturias, Grant/Award Number: AYUD/2021/51261; European Regional Development Fund, Grant/Award Number: PGC2018-097575-B-I00

Handling Editor: Camille Bonneaud

Abstract

Patterns of pathogen prevalence are, at least partially, the result of coevolutionary host–pathogen interactions. Thus, exploring the distribution of host genetic variation in relation to infection by a pathogen within and across populations can provide important insights into mechanisms of host defence and adaptation. Here, we use a landscape genomics approach (Bayenv) in conjunction with genome-wide data (ddRADseq) to test for associations between avian malaria (*Plasmodium*) prevalence and host genetic variation across 13 populations of the island endemic Berthelot's pipit (*Anthus berthelotii*). Considerable and consistent spatial heterogeneity in malaria prevalence was observed among populations over a period of 15 years. The prevalence of malaria infection was also strongly positively correlated with pox (*Avipoxvirus*) prevalence. Multiple host loci showed significant associations with malaria prevalence after controlling for genome-wide neutral genetic structure. These sites were located near to or within genes linked to metabolism, stress response, transcriptional regulation, complement activity and the inflammatory response, many previously implicated in vertebrate responses to malarial infection. Our findings identify diverse genes – not just limited to the immune system – that may be involved in host protection against malaria and suggest that spatially variable pathogen pressure may be an important evolutionary driver of genetic divergence among wild animal populations, such as Berthelot's pipit. Furthermore, our data indicate that spatio-temporal variation in multiple different pathogens (e.g. malaria and pox in this case) may have to be studied together to develop a more holistic understanding of host pathogen-mediated evolution.

KEYWORDS

adaptation, avian malaria, birds, genotype–environment association, immune defence, pathogen-mediated selection

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Pathogens are – by definition – harmful to hosts, influencing both reproduction and survival (Anderson & May, 1979; Daszak et al., 2000), thereby acting as a strong selective pressure on genes encoding for host defence (Sommer, 2005). Reciprocal selection pressure from hosts on infectivity-related genes also drives counter-adaptation in pathogens, resulting in strong coevolution (Dawkins & Krebs, 1979; Paterson et al., 2010). However, pathogen-mediated selection does not always lead to the most beneficial host variants becoming fixed in the host population; multiple variants at a locus can be maintained via balancing selection (e.g. Ferrer-Admetlla et al., 2008; Hawley & Fleischer, 2012; Loiseau et al., 2009). Different mechanisms of balancing selection (reviewed in Charlesworth, 2006; Spurgin & Richardson, 2010) include heterozygote advantage (Doherty & Zinkernagel, 1975), rare allele advantage (Slade & McCallum, 1992) and spatio-temporally fluctuating selection (Hill et al., 1991). Importantly, these mechanisms and divergent host–pathogen co-evolutionary cycles across a host's geographic range may lead to spatial differences in the variants conferring defence against a given pathogen at any given time. Investigating the distribution of genetic variation within and among host populations facing a given pathogen can help to identify the loci and mechanisms that underlie the host's evolutionary response to that pathogen.

It can, however, be challenging to identify the underlying host loci involved in modulating host–pathogen interactions in natural systems. Analyses of vertebrate pathogen defences tend to have focused on well-known resistance-related genes (i.e. those that prevent or limit infection) (Aguilar et al., 2004; Brouwer et al., 2010; Davies et al., 2021; Hawley & Fleischer, 2012; Quémeré et al., 2015; Tschirren et al., 2011). By comparison, the genes involved in tolerance mechanisms which act to mitigate fitness costs by minimising direct or immune-mediated damage have been largely overlooked, despite evidence that tolerance is an important host defence strategy (Råberg et al., 2007, 2009). The candidate gene approach, as well as other approaches such as genome-wide association studies (GWAS), also requires reliable individual-based phenotypic datasets to establish robust links between genetic variants and pathogen infection. However, limited knowledge of individual-level exposure to, or recovery from, infectious diseases in wild host systems can confound or conceal associations with genotypes. In such complex scenarios, bottom-up genome scanning methods therefore offer a valuable alternative to uncover potentially adaptive loci and identify candidate genes for testing in future in-depth studies. These methods involve systematically screening thousands of variants across the genome, operating without the constraints of pre-existing knowledge about candidate genes or specific individual host phenotypes (Pardo-Díaz et al., 2015).

Conducting genome scans to investigate genomic variation linked to divergent and spatially variable selection entails two primary methods: genetic differentiation-based approaches and genotype–environment association (GEA) analyses (Hoban et al., 2016; Schoville et al., 2012). While differentiation-based methods (such as

those that involve F_{ST} outlier tests or alternative statistics) identify loci that exhibit extreme values of differentiation among populations compared to the genome-wide average (Beaumont, 2005; Lewontin & Krakauer, 1973), GEA analyses identify loci that show a strong association between the geographic pattern of allele frequencies and spatial variation in an environmental variable (Rellstab et al., 2015). GEA methods can therefore help to elucidate the environmental selective pressure driving the allele frequency patterns, yet they do not explicitly incorporate individual phenotypic data. Instead, they operate on the assumption that population-level variation in an environmental variable (e.g. in pathogen prevalence) is representative of the strength of selective pressure acting on phenotypes in that population. Importantly, a number of model-based GEA tools exist that also account for the effects of neutral genetic structure across populations (e.g. latent factor mixed models, Bayenv, BayPass, Günther & Coop, 2013; Gautier, 2015; Frichot et al., 2013). Thus, the application of GEA in host–pathogen systems can help to identify the genetic variants that may have accumulated in different populations of the same host species due to spatially variable selective pressure from a specific pathogen (e.g. Fumagalli et al., 2011; Garroway et al., 2013; Mackinnon et al., 2016; Zueva et al., 2014, 2018).

Avian malaria (here defined widely as infection by *Plasmodium* and *Haemoproteus* spp.) is a well-studied disease of wild birds due to its near-global distribution, wide host range and relevance for host health and conservation (Valkiunas, 2005). Infected birds typically undergo an acute phase, whereby high parasite loads are briefly experienced soon after infection, followed by a chronic phase with low parasitaemia (that may be undetectable in the blood) which can last throughout an individual's lifetime with potential relapses (Lapointe et al., 2012). Evidence from experimentally infected birds has shown reductions in condition and survival as a result of high parasitaemia (Atkinson et al., 1995; Palinauskas et al., 2008; Yorinks et al., 2000). Wild-caught birds are more likely to have chronic levels of malaria rather than acute infections because the probability of capturing an individual depends on its survival and mobility (Mukhin et al., 2016); however, numerous studies have also shown negative effects on fitness including survival, lifespan (Asghar et al., 2015; Lachish et al., 2011; Marzal et al., 2008) and reproductive success (Asghar et al., 2011; Pigeault et al., 2020). Collectively, evidence suggests that malaria has the potential to drive host variation in both immune genes and other parts of the genome related to health and homeostasis.

Berthelot's pipit (*Anthus berthelotii*) is endemic across three archipelagos in the North Atlantic – the Canary Islands, the Madeira archipelago and the Selvagens. It has only relatively recently dispersed across this island system (c. 8500–50,000 years ago) (Illera et al., 2007; Martin et al., 2023; Spurgin et al., 2014) since its origin in the Canary Islands c. 2.1–2.5 million years ago (Martin et al., 2023; Voelker, 1999). Previous genomic work has shown an absence of subsequent gene flow between populations (Martin et al., 2021), thus potentially permitting population differentiation within and among archipelagos through local adaptation. The pipit therefore provides a useful system for studying the initial evolutionary

responses to differential selective pressures in a wild vertebrate, including pathogen-mediated selection (e.g. Armstrong et al., 2018, 2019; González-Quevedo et al., 2015, 2016). Patterns of avian blood pathogen pressure are due to the combined effect of climatic, biogeographic and anthropogenic factors (González-Quevedo et al., 2014; Illera et al., 2017; Padilla et al., 2017; Sehgal, 2015; Spurgin et al., 2012) which are highly heterogeneous within and among these island archipelagos. Importantly, prevalence of pathogen infection in Berthelot's pipits, that is, avian malaria and pox (*Avipoxvirus* species), differs markedly between populations, aligning closely with biogeographical expectations based on island size and isolation (Spurgin et al., 2012). Simultaneously, prevalence demonstrates relative consistency within populations, as evidenced by studies conducted over a time frame ranging up to 15 years (Illera et al., 2008; Sheppard et al., 2022; Spurgin et al., 2012). Different populations are therefore likely to have been exposed to consistently different levels of pathogen-mediated selection leading to variation in disease resistance/tolerance evolving across the Berthelot's pipit geographic range.

Analytical approaches based on individual-level assessments of infection, including candidate gene and GWAS methods, have previously been employed to identify pathogen-related loci in Berthelot's pipit (Armstrong et al., 2018, 2019; González-Quevedo et al., 2016). Polymorphisms in key genes such as Toll-like receptors (TLRs) and the major histocompatibility complex (MHC), as well as a variant near to interleukin-16 (an inflammatory regulator), were found to be associated with malaria (Armstrong et al., 2018, 2019; González-Quevedo et al., 2016). However, it remains uncertain whether the variants underlying individual trait variation exhibit patterns of local adaptation across varying spatial scales, given the potential influence of distinct coevolutionary processes. Observations within populations or archipelagos may not necessarily extrapolate to broader spatial contexts (Armstrong et al., 2018; Bonneaud et al., 2006; Kloch et al., 2010). Genome scans using measures of population differentiation (EigenGWAS; Chen et al., 2016) have also revealed other loci affecting immunity, homeostasis and metabolism that are highly divergent between and/or within archipelagos (Armstrong et al., 2018; Martin et al., 2021). However, such an approach does not incorporate environmental or disease data, thus revealing little about how these drivers of spatially variable selection affect specific sites in the genome. Furthermore, differentiation-based methods primarily identify the most divergent SNPs, whereas GEA methods are more likely to detect loci with gradual allele frequency changes in response to spatially varying environments (Hancock et al., 2010). There is also evidence that GEA tests possess higher power than differentiation outlier methods in detecting adaptive loci within island models (Lotterhos & Whitlock, 2015). A broad, multi-population, multi-gene assessment of associations between pathogen prevalence and genetic variation would therefore be beneficial within this species to gain further insights into the genetic basis of disease.

We recently utilised a Bayenv method with restriction site-associated DNA sequencing (RAD-seq) data across the Berthelot's pipit range to uncover candidate genes involved in the host's

response to avian poxvirus (Sheppard et al., 2022). Here, we use RAD-seq data and population-level measures of avian malaria prevalence from Berthelot's pipit, in conjunction with the Bayenv method, to identify loci that may have evolved in response to avian malaria infection. We then explore the overlap between previously identified avian pox-associated single nucleotide polymorphisms (SNPs) and malaria-associated SNPs in an attempt to disentangle the effects of both diseases. We hypothesise that there could be a limited degree of overlap in SNPs linked to genes governing resistance, primarily due to the specificity of pathogen-targeting mechanisms. In contrast, SNPs linked to genes influencing tolerance are expected to exhibit a higher degree of overlap, given the shared mechanisms aimed at mitigating harm during infections.

2 | MATERIALS AND METHODS

2.1 | Blood sampling

We used 1574 Berthelot's pipits blood samples obtained between 2005 and 2020 as part of other studies (Armstrong et al., 2019; González-Quevedo et al., 2014; Illera et al., 2007; Sheppard et al., 2022; Spurgin et al., 2012) (details in Table S1). Sampling included the 12 islands of the species' range across Macaronesia (Figure 1). Individuals inhabiting the plateau of Mount Teide (>2000m above sea level) were considered a 13th population, separated from the lowland Tenerife population by the dense forest vegetation at moderate elevations which the pipit does not inhabit. Individuals were caught using spring nets baited with *Tenebrio molitor* larvae, and blood samples (c. 25 µL), collected from the brachial vein, were stored in absolute ethanol (800 µL) at room temperature. Every bird was ringed with a numbered metal band from the relevant Spanish or Portuguese authorities to prevent resampling.

2.2 | DNA extraction and molecular detection of malaria infection

Genomic DNA was extracted from blood using a salt extraction protocol (Richardson et al., 2001). Samples were screened for *Haemoproteus* and *Plasmodium* spp. based on mitochondrial genome sequences using either the nested PCR protocols described in Waldenström et al. (2004) (samples collected prior to 2019) or the multiplex PCR assay described in Ciloglu et al. (2019) (samples collected in 2019 and onwards). Recent research into the comparative effectiveness of nested and multiplex PCR methods in diagnosing malaria suggests that they exhibit comparable sensitivity for *Plasmodium* and *Haemoproteus* (Ciloglu et al., 2019; Musa et al., 2022). However, the multiplex PCR assay outperforms the nested PCR in detecting co-infections, involving more than one species or strain of pathogen. Negative and positive controls were included in each PCR run. Each sample underwent two separate successful PCR screening events (three if the first two tests were inconsistent) and

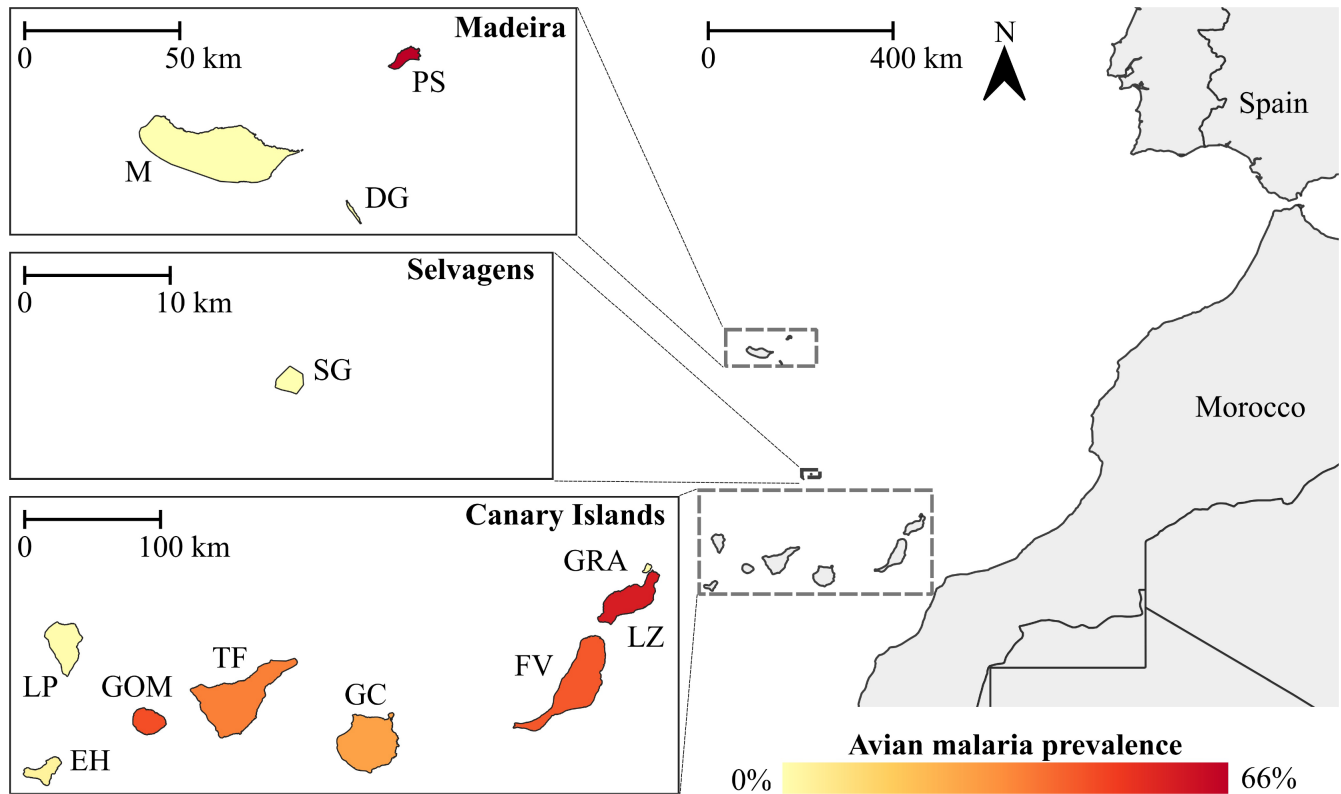


FIGURE 1 The distribution of Berthelot's pipit populations and the mean malaria (*Plasmodium*) prevalence (percentage infected) in those populations across a 15-year sampling period. Canary Island populations: EH, El Hierro; FV, Fuerteventura; GC, Gran Canaria; GOM, La Gomera; GRA, La Graciosa; LP, La Palma; LZ, Lanzarote; TEID, El Teide; TF, Tenerife. Madeiran populations: DG, Deserta Grande; M, Madeira; PS, Porto Santo. Selvagens population: SG, Selvagem Grande.

any samples that were positive twice were assigned as infected. All positive samples collected prior to 2019 ($n=400$) had been previously Sanger sequenced to identify the strain of haemosporidian (see Armstrong et al., 2019; González-Quevedo et al., 2014; Illera et al., 2008; Spurgin et al., 2012). No evidence of *Haemoproteus* spp. infection was detected, but four strains of *Plasmodium* (LK6, LK5, KYS9 and PS1530) were detected with no evidence of mixed-strain infections. The LK6 strain accounted for >95% of infections, therefore testing for differences in strain-specific genetic associations was not possible. Instead, birds were classified by their *Plasmodium* infection status (not infected/infected) only.

2.3 | RAD sequencing

RAD-seq markers have previously been used in GWAS analyses (based on individual assessments of disease) to identify genetic variants associated with LK6 infection status in Berthelot's pipits (Armstrong et al., 2018), and in a GEA to identify genetic variants associated with poxvirus prevalence (population level) across their range (Sheppard et al., 2022). Here, we used the 'Berthelot's' double-digest RAD-seq (ddRAD-seq) library generated in Armstrong et al. (2018) and applied the same filtering steps as in Sheppard et al. (2022). The initial ddRAD-seq library was prepared

and processed according to DaCosta and Sorenson (2014) with modifications detailed in Armstrong et al. (2018) and consisted of 20–22 individuals per population (putatively unrelated individuals and equal numbers of both sexes) sampled in 2006 and 2009. Reads were mapped to the zebra finch (*Taeniopygia guttata*) genome version 3.2.4 (Warren et al., 2010). Loci with ambiguous genotypes in ≤ 3 samples or those that showed an excess of missing/ambiguous genotypes (>10% of samples) were treated as missing data. Multiple SNPs were outputted from each RAD-tag and treated as separate loci (SNPs are denoted by the RAD-tag followed by their bp distance from the beginning of the tag), resulting in 9960 SNPs. We then performed additional filtering steps using Plink version 1.9 (Chang et al., 2015), removing loci with a minor allele frequency (MAF) of <0.05 and one variant from each pair showing strong linkage disequilibrium (LD; $r^2 > .5$) with a sliding window of 50kb (step size of five).

2.4 | Pathogen covariables for association analyses

Population malaria prevalence (Table S2) was calculated using the *Plasmodium* infection status of all birds sampled between 2005 and 2020 to account for year-to-year variability arising from natural fluctuations in infection rates and differences in sample size (cases of

infected individuals/total number of individuals \times 100). We considered utilising a binomial logistic regression to model the relationship between malaria and pox prevalence (previously estimated in Sheppard et al., 2022) across populations because prevalence is a proportion derived from counts of infected and uninfected individuals. However, model diagnostics suggested that a beta distribution provided a more suitable fit for our data. Consequently, we represent prevalence as the proportion of infected birds and employed a zero-inflated beta regression using the glmmTMB package (version 1.1.8; Brooks et al., 2017) in R version 4.3.1 (R Core Team, 2023). A beta regression model is well-suited for variables that are doubly bounded and often exhibit non-normal distributions and heteroscedasticity (Douma & Weedon, 2019). As our results indicated a strong positive correlation between malaria prevalence and pox prevalence across populations (see Section 3), we used both the 'raw' malaria prevalence values and (in a separate model) residuals from malaria prevalence regressed on pox prevalence as inputs for the Bayenv models. This approach allowed us to examine the role of malaria independently of the influence of pox virus, while the 'raw' values encompass more environmental variability and therefore provide greater power. All values were standardised to a mean of 0 and a standard deviation of 1 prior to input into GEA analyses.

2.5 | Genotype–environment association analyses

GEAs were conducted using Bayenv2.0 (Günther & Coop, 2013) to identify associations between the allele frequencies of each SNP across populations and the gradient of malaria prevalence (or residual malaria) observed across populations. Subsequent analyses and visualisation of the output were performed using R version 4.3.1 (R Core Team, 2023).

Bayenv2.0 employs a univariate Bayesian framework and explicitly accounts for shared evolutionary history and population structure by incorporating a covariance matrix of population allele frequencies as the null model (Coop et al., 2010; Günther & Coop, 2013). The matrix was generated previously by Sheppard et al. (2022) by running Bayenv2.0 with the filtered ddRAD-seq data and averaging the last covariance matrices of 10 replicate runs, each of 100,000 iterations.

For each SNP in the dataset, the analysis assigns a Bayes factor (BF) value, which describes the strength of evidence provided by the data for the alternative model – that is, a linear relationship between allele frequency and the pathogen covariable across populations – relative to the null model. Rather than relying solely on the parametric model, we also calculated Spearman's rank correlation coefficients (ρ) and ensured candidates indicated by BFs are not confounded by outlying populations (Günther & Coop, 2013). To further confirm this, we plotted population-level malaria prevalence against population allele frequencies for any candidate SNPs, using Plink version 1.9 (Chang et al., 2015), to estimate minor allele frequencies (MAFs). Following the recommendation of Blair et al. (2014), we averaged estimates across five independent runs, each of 1,500,000

iterations and with different random seeds, because Markov chain Monte Carlo (MCMC) algorithms can generate high run-to-run variability (mean pairwise correlation between the five runs for the main analysis: for BF values, $R = .690 \pm .222$, and for ρ , $R = .997 \pm .0001$). SNPs ranked above both the 99th percentile of averaged BF and the 90th percentile of averaged absolute value of ρ were then considered robust candidates, in line with the manual (Günther & Coop, 2018).

We used the zebra finch genome *Taeniopygia guttata*-3.2.4 (Warren et al., 2010) in the NCBI Genome Data Viewer v.5.1 (www.ncbi.nlm.nih.gov/genome/gdv/browser) to identify candidate genes. These candidates were defined as genes located within a 10 kb proximity either upstream or downstream of a malaria-associated SNP. Given the observed extent of LD within populations of Berthelot's pipit (Armstrong et al., 2018; Martin et al., 2021), we consider candidate loci located at this distance to be physically linked to the respective gene.

3 | RESULTS

Across the Berthelot's pipit geographic range, there was considerable inter-population variation in malaria (*Plasmodium*) prevalence (Table S2). Two populations were completely free of infection (Deserta Grande and Selvagem Grande) and prevalence in infected populations ranged between 1.1% in Madeira and 66.3% in Porto Santo – neighbouring populations in the Madeiran archipelago. Malaria was detected in every population within the Canary Islands, where the general pattern of prevalence is characterised by a negative east–west gradient (Figure 1). At the population level, the proportion of individuals infected with malaria was positively associated with previous estimates (Sheppard et al., 2022) of pox infection within this species (beta regression estimate = 9.16 ± 1.99 , $p < .01$; Figure 2). Estimates of malaria prevalence within populations were largely consistent between 2006 and 2009 (years in which all populations were sampled: Pearson correlation, $R = .64$, $p = .02$; Figure S1).

A total of 2334 SNPs with minor allele frequency >0.05 were retained for analysis with Bayenv following LD filtering ($r^2 > .5$) of the ddRAD-seq dataset (Armstrong et al., 2018). Twenty-three SNPs (c. 1% of the total SNPs) passed the stringent thresholds (Figure 3) and were identified as having an association between allele frequencies and malaria prevalence at the population level following five runs of Bayenv. These associations did not seem to be driven by single outlying populations (Figure 4). The SNPs found in both the top 1% of the averaged BFs (>6.68) and 10% of the averaged absolute values of Spearman's ρ (>0.25) were located in proximity to (within 10 kb upstream or downstream) or within 20 annotated candidate genes (Table 1). Specifically, eight SNPs were located within genes (34.8%), seven were located 10 kb up- or downstream of a gene(s) (30.4%), one was located within 10 kb of a pseudogene (4.3%) and seven were not near genes (30.4%).

When the same Bayenv analysis was performed using 'residual malaria', 18 SNPs exhibited a strong association. However, several of these associations appeared to be influenced by the most divergent,

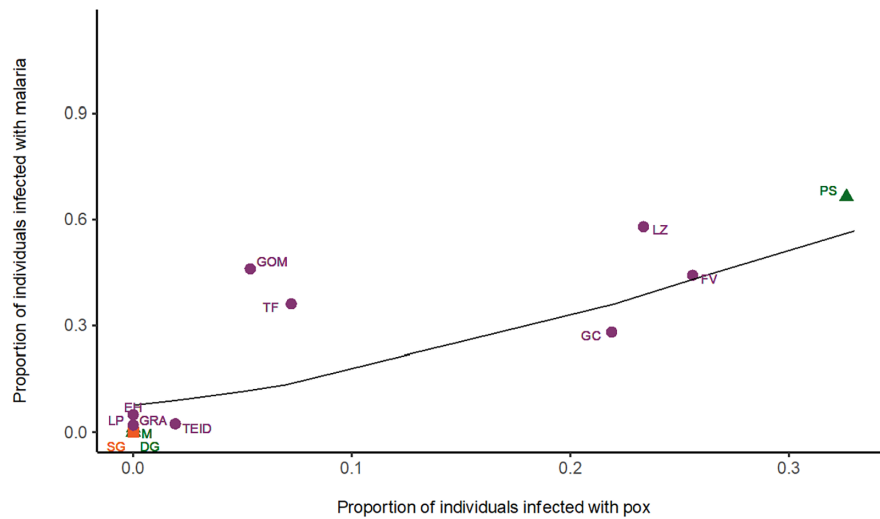


FIGURE 2 Spatial patterns of avian malaria (*Plasmodium*) and avian pox (*Avipoxvirus*) prevalence (proportion of individuals infected) across populations of Berthelot's pipit. Values of pox prevalence previously estimated by Sheppard et al. (2022). The fit line is a zero-inflated beta regression between the x- and y-axes. Grey shading indicates 95% confidence intervals. Points depict raw data. Canary Island populations (purple): EH, El Hierro; FV, Fuerteventura; GC, Gran Canaria; GOM, La Gomera; GRA, La Graciosa; LP, La Palma; LZ, Lanzarote; TEID, El Teide; TF, Tenerife. Madeiran populations (green): DG, Deserta Grande; M, Madeira; PS, Porto Santo. Selvagens population (orange): SG, Selvagem Grande.

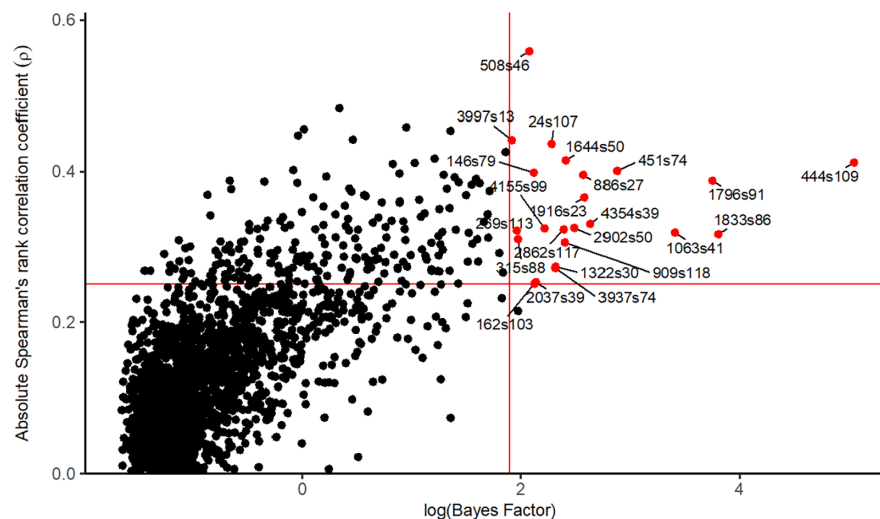
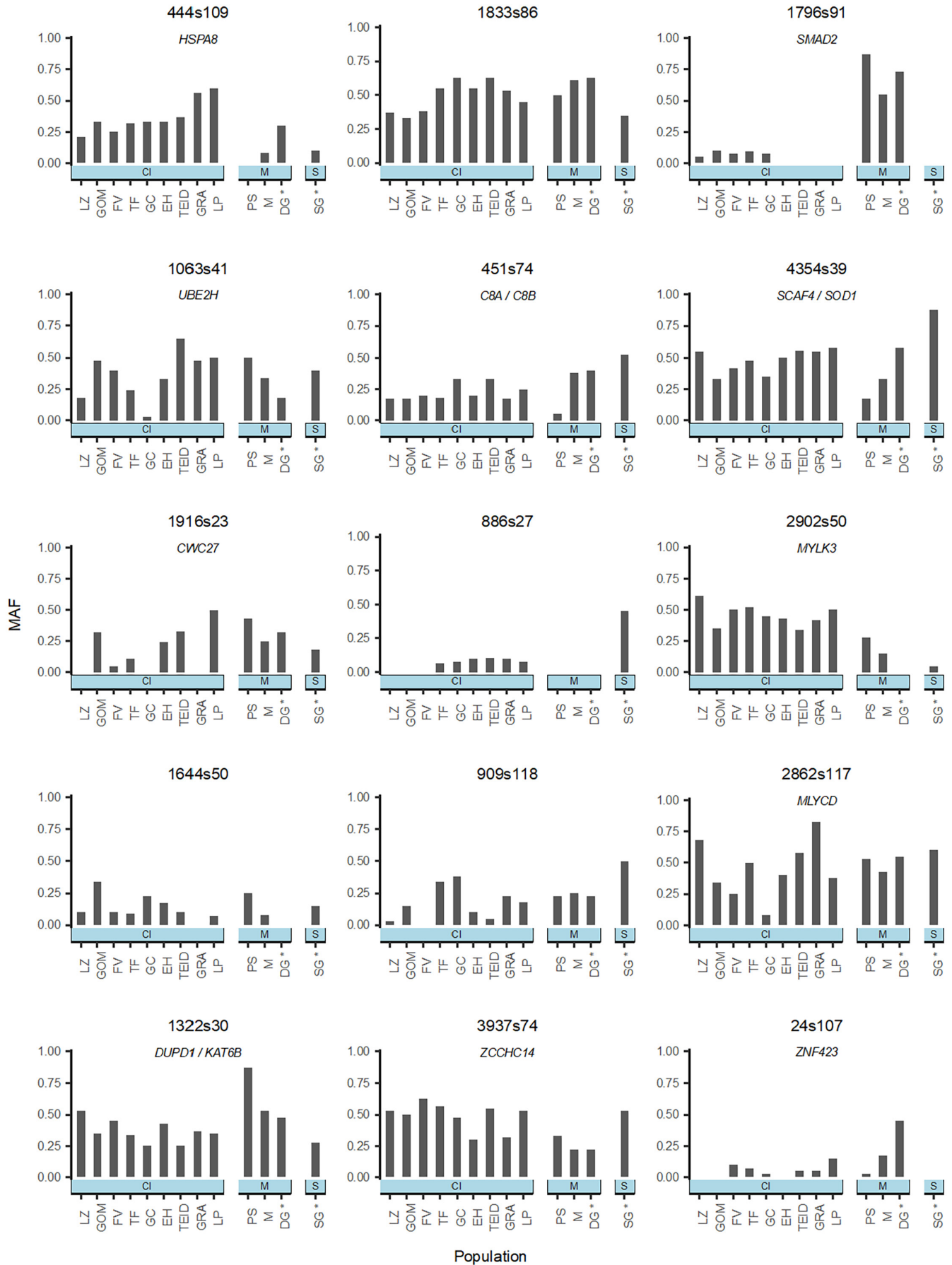


FIGURE 3 The distribution of Bayes factor values and absolute Spearman's rank correlation coefficients (ρ), generated and averaged from five replicate runs of Bayenv, for genome-wide ddRAD SNPs among 13 Berthelot's pipit populations with respect to malaria prevalence. SNPs were considered candidates for associations with population-level malaria prevalence if they ranked in the highest 1% of Bayes factor values (>6.68 , threshold indicated by the vertical red line) and 10% of Spearman's ρ (threshold indicated by the horizontal red line). Twenty-three SNPs were identified as candidates (in red).

malaria-free population(s) (Figure S2). These SNPs were located in proximity to or within 13 candidate genes (Table S3). Notably, two of these SNPs, corresponding to three candidate genes, demonstrated

significant associations that were shared between the analyses conducted with 'raw' prevalence values and 'residual malaria'. Specifically, eight SNPs were located within genes (34.8%), three

FIGURE 4 Minor allele frequency (MAF) of 23 candidate SNPs identified by testing for associations with population-level malaria prevalence in Bayenv. Bayenv corrects for neutral population structure among allele frequencies in the analysis. Nearby genes are noted below the SNP names. Populations are grouped by archipelago (CI, Canary Islands; M, Madeira; S, Selvagens) and then ordered according to malaria prevalence (highest-lowest). Malaria-free populations are indicated by an asterisk. Acronyms: DG, Deserta Grande; EH, El Hierro; FV, Fuerteventura; GC, Gran Canaria; GOM, La Gomera; GRA, La Graciosa; LP, La Palma; LZ, Lanzarote; M, Madeira; PS, Porto Santo; SG, Selvagem Grande; TEID, El Teide; TF, Tenerife.



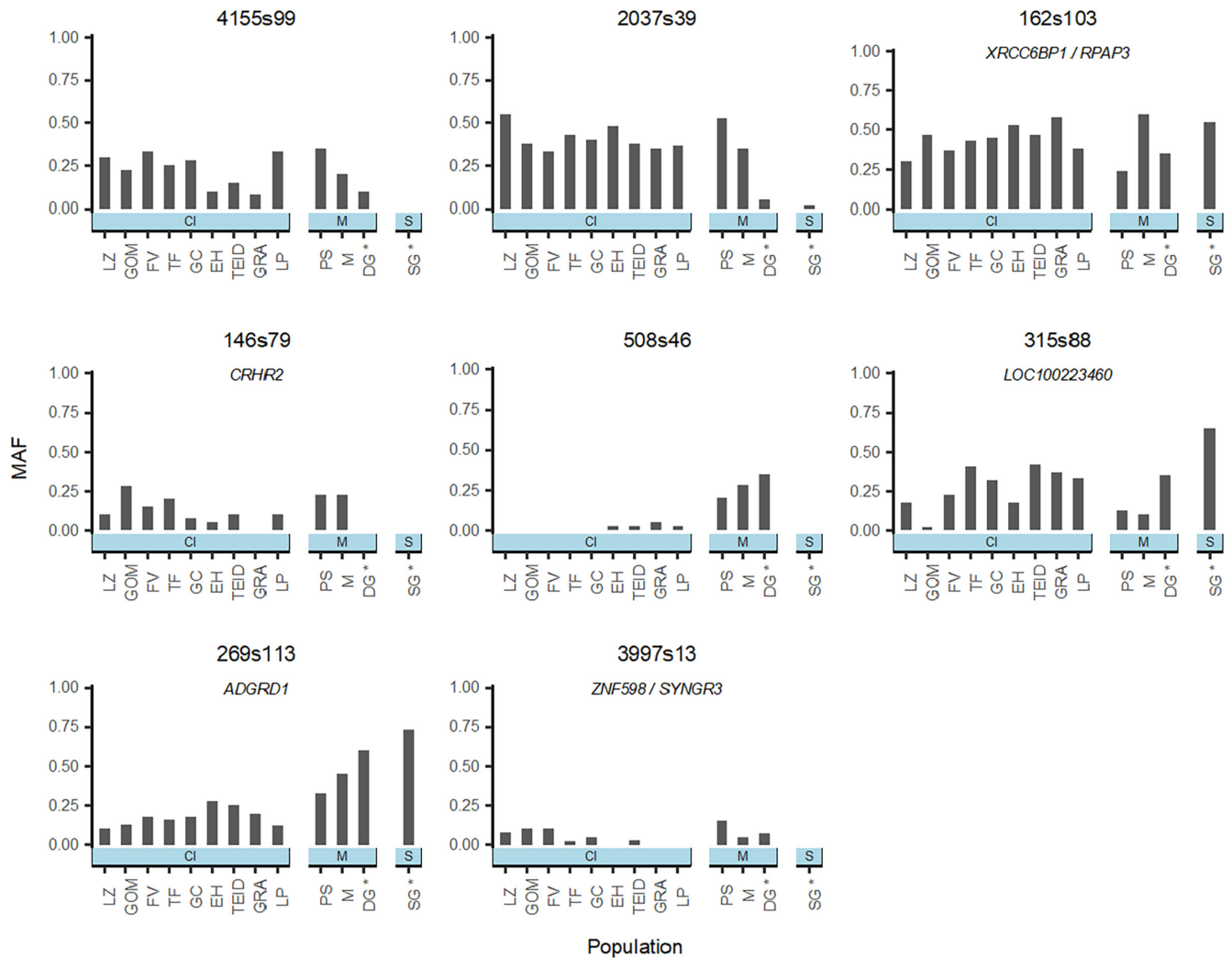


FIGURE 4 (Continued)

were located 10kb up- or downstream of a gene(s) (16.7%), one was located within 10kb of a pseudogene (5.5%) and six were not near genes (33.3%).

4 | DISCUSSION

Applying a broad multi-population GEA approach, we identify candidate SNPs associated with avian malaria across the range of Berthelot's pipit. Population-level allele frequency variation at these candidate SNPs correlated with the local prevalence of malaria after controlling for genome-wide divergence due to neutral structure. These results reveal genes potentially involved in a host's response to malaria and suggest that local adaptation to spatially variable pathogen pressures, such as malaria, may be an important driver of genetic variation in Berthelot's pipit.

Spatially variable selective pressure from avian malaria has been shown to shape host genetic structure across populations (Foster et al., 2007), drive variation in immunity and health (Names, Schultz, Hahn, et al., 2021) and drive the evolution of

variable tolerance (Atkinson et al., 2013). In the present study, the main GEA analysis revealed 23 loci that showed strong evidence of association with population-level malaria prevalence, likely as a result of selective pressure. While Bayenv corrects for neutral population allele frequency patterns (Coop et al., 2010; Günther & Coop, 2013), the methods used here are correlative so inferences can be confounded by collinearity along the axis of interest, whether with an additional environmental variable (discussed below in relation to *pox*) or the principal axes of neutral population structure. From previous work, malaria prevalence seems to be independent of genetic structure across Berthelot's pipit populations; rather, other biogeographical and environmental factors likely explain the distribution of avian malaria (Spurgin et al., 2012). For instance, Porto Santo experiences the highest levels of malaria prevalence yet other populations within the same archipelago, between which there is little genetic structure, are completely free of avian malaria (Figure 1). Hence, it is unlikely that prior distribution of susceptible alleles through founder effects would drive associations between malaria and specific loci. Instead, the identified genomic associations are consistent with

TABLE 1 SNPs identified as being associated with population-level malaria prevalence using Bayenv, ordered by Bayes factor (BF) values.

SNP	BF	Genomic location ^a	Candidate gene(s) ^b	Distance from gene	Gene name	Putative function
44s109	156.0	24:3526218	<i>hspa8</i>	6319 bp upstream	Heat shock protein family A (Hsp70) member 8	ATP-dependent molecular chaperone, aiding protein folding
1833s86	45.0	19:7248322	-	-	-	-
1796s91	42.6	Z:421273	<i>smad2</i>	In gene (intron)	SMAD family member 2	Transcriptional modulator and downstream effector of the transforming growth factor (TGF)- β signalling pathway
1063s41	30.2	1A:217173	<i>ube2h</i>	In gene (intron)	Ubiquitin-conjugating enzyme E2 H	Catalyses the attachment of ubiquitin to other proteins in the ubiquitin/proteasome degradation pathway
451s74	17.8	8:23610708	<i>c8a</i> <i>c8b</i>	In gene (intron) 8722 bp downstream	Complement C8 alpha chain Complement C8 beta chain	Encodes subunits of the complement system component essential for the formation of the membrane attack complex (MAC). The MAC serves as an important immune effector, creating pores in the membranes of target cells, leading to their destruction
4354s39	14.0	1:98523884	<i>scaf4</i> <i>sod1</i>	440 bp downstream 1606 bp downstream	SR-related CTD-associated factor 4 Superoxide dismutase 1	RNA binding protein, involved in mRNA processing potentially through termination of transcription Copper/zinc binding antioxidant enzyme
1916s23	13.2	Z:50640938	<i>cwc27</i>	In gene (intron)	CWC27 spliceosome-associated cydlophilin	Pre-mRNA splicing factor recruited by the spliceosome
886s27	13.1	11:8338772	-	-	-	-
2902s50	12.1	11:13961995	<i>mylk3</i>	7175 bp upstream	Myosin light-chain kinase 3	Kinase that phosphorylates myosin chains and plays a role in cardiac performance and calcium signalling
1644s50	11.2	18:10178831	-	-	-	-
909s118	11.1	21:5744237	-	-	-	-
2862s117	11.0	11:1265660	<i>mylcD</i>	3598 bp downstream	Malonyl-CoA decarboxylase	Catalyses the conversion of malonyl-CoA to acetyl-CoA during fatty acid metabolism
1322s30	10.2	6:13842170	<i>dupd1</i> <i>kat6B</i>	349 bp downstream 2756 bp downstream	Dual-specificity phosphatase and pro isomerase domain containing 1 Lysine acetyltransferase 6B	Dephosphorylates phospho-serine/threonine and phosphotyrosine residues, modulating downstream cell signalling pathways Catalyses the acetylation of histones and plays a key role in the regulation of transcription of multiple genes
3937s74	10.1	11:10604516	<i>zcchc14</i>	In gene (intron)	Zinc finger CCHC domain-containing 14	May bind nucleic acids, proteins and zinc and play a role in RNA metabolism

(Continues)

TABLE 1 (Continued)

SNP	BF	Genomic location ^a	Candidate gene(s) ^b	Distance from gene	Gene name	Putative function
24s107	9.8	11:12916411	<i>znf423</i>	In gene (intron)	Zinc finger protein 423	DNA-binding, multi-functional transcription factor involved in signal transduction
415s99	9.2	22:2584411	-	-	-	-
2037s39	8.5	9:5918598	-	-	-	-
162s103	8.4	1A:30828068	<i>xrcc6bp1</i> <i>rpap3</i>	986 bp downstream 8853 bp upstream	Mitochondrial inner membrane protease ATP23 RNA polymerase II-associated protein 3	Metalloproteinase involved in ATP synthase assembly and DNA repair Helps the assembly of RNA polymerase II and functions in the regulation of transcription and apoptosis
146s79	8.3	2:4280785	<i>chr2</i>	In gene (intron)	Corticotropin-releasing hormone receptor 2	Binds corticotropin-releasing factor, modulating hormonal release and coordinating physiological changes, thereby mediating the physiological responses to stress
508s46	8.0	2:67296436	-	-	-	-
315s88	7.2	Z:10812737	<i>loc100223460</i>	In pseudogene	-	-
269s113	7.1	15:384062	<i>adgrd1</i>	In gene (intron)	Adhesion G protein-coupled receptor D1	Involved in cell surface receptor signalling
3997s13	6.8	14:9421650	<i>znf598</i> <i>syng3</i>	2333 bp downstream 2904 bp downstream	Zinc finger protein 598 Synaptogyrin-3	Initiates ribosome quality control (RQC) pathway when a ribosome has stalled during translation May play a role in regulated exocytosis

Note: SNPs highlighted in blue text were previously also identified as being associated with pox prevalence (Sheppard et al., 2022) using the same genetic dataset in a Bayesian analysis.

^aChromosome location and base position based on *Taeniopygia_guttata*-3.2.4 genome assembly.

^bWithin 10 kb of focal SNP.

locally divergent selection across the species' range and provide support for the hypothesis that there is adaptation to the local pathogen regime.

A set of 20 malaria-associated candidate genes were located within 10kb of the 23 associated SNPs, many of which have not been previously linked to divergence across this system. The molecular mechanisms that underlie the host response to malaria have predominantly been investigated in humans and murine models, as well as captive and natural populations of birds. We did not identify some of the immune-related genes commonly reported to influence malaria-related phenotypes in these taxa such as *g6pD*, *abO*, *cd40lG*, *tnf*, *atp2b4*, beta globins, interleukins, toll-like receptors or MHC (Bonneaud et al., 2006; Longley et al., 2011; MalariaGEN, 2019; Ravenhall et al., 2018; Rockett et al., 2014; Sepil et al., 2013). This may in part be due to a lack of genomic resolution to evaluate these in our study (see below) or the differences in specific host-pathogen interactions. However, we found a few genes that function in the innate immune system and inflammatory response (*c8A*, *c8B* and *smad2*). For example, all complement pathways culminate in the formation of the membrane attack complex – an important innate immune effector that transverses pathogen cell membranes and facilitates pathogen clearance – of which complement 8 (*c8*) is a key component (Bayly-Jones et al., 2017). Studies in both humans and mice have revealed the complement system is activated against malaria (Kurtovic et al., 2020; Silver et al., 2010) with elevated levels of membrane attack complex formation observed during infection (Roestenberg et al., 2007). In birds, other genes of the complement system have been shown to be differentially expressed during infection with malaria (Videvall et al., 2015) and there is evidence that elevated complement activity might act to control parasitaemia (Ellis et al., 2015). Also, *smad2* encodes an intracellular transcription factor, activated by the transforming growth factor β (TGF- β) signalling pathway, which together regulates inflammation and interacts with other immune system pathways. TGF- β itself was inferred to be under selection from malaria in low-elevation populations of Hawai'i 'amakihi (*Chlorodrepanis virens*) (Cassin-Sackett et al., 2019). Given the dual role inflammation plays in malaria clearance and associated pathological changes (Popa & Popa, 2021), genes involved in its regulation may represent important mechanisms of malaria tolerance. Overall, these genes correspond with previous findings that suggest both the innate immune system and inflammatory response are commonly initiated in response to malaria infection, thus these genes seem appropriate candidates for further study into responses to malaria infection.

Recent transcriptomic studies have uncovered genes that are upregulated or downregulated during infectious disease (e.g. Boštjančić et al., 2022; Chu et al., 2016; Huang et al., 2013; McNew et al., 2022; Rosenblum et al., 2012). Those examining gene expression changes in the host response to malaria have identified genes that were related to a number of functions beyond those strictly within the immune system, including the stress response, cell death regulation, metabolism, cell signalling, transcriptional regulation and telomerase activity (Quin et al., 2017; Videvall et al., 2015, 2020). In

the present study, we detected associations with malaria prevalence within two of the same genes that were previously overexpressed in infected avian hosts (Videvall et al., 2015), *ube2H* and *cwc27*. Both of these genes have been associated with metabolic processes; *ube2H* is involved in ubiquitin-dependent protein catabolism, targeting proteins for degradation and can also affect antigen processing and presentation (Stewart et al., 2016), and *cwc27* is a spliceosome component likely involved in mRNA processing and protein metabolic processes (Busetto et al., 2020). Among the malaria-associated variants identified in the present study, many were also located near genes with links to protein and RNA metabolism (*hspa8*, *sod1*, *mylk3*, *dupd1*, *kat6B*, *zcchc14*, *znf423*, *xrcc6bp1*) and a single variant was near a gene involved in fatty-acid oxidative metabolism (*mlycD*). Immune functions incur high energetic and metabolic costs (Lochmiller & Deerenberg, 2000; Martin et al., 2003; Ots et al., 2001) and there is some evidence that energy costs of infection may impact body condition in Berthelot's pipit (Spurgin et al., 2012), thus metabolic adaptations could influence host fitness.

During malaria infection, the production of reactive oxygen species (ROS) occurs as a result of phagocytosis (which is beneficial for pathogen clearance), as well as host inflammatory signalling, host metabolic changes induced by infection and the degradation of haemoglobin by the pathogen itself (Vasquez et al., 2021). Thus, the production of ROS is a key component of the host defence response, however, high levels can lead to non-specific oxidative damage to host lipids, proteins and DNA (Szabó, 2003). The imbalance between oxidants and antioxidant defences is referred to as oxidative stress and is a common occurrence across a variety of taxa experiencing *Plasmodium* infections (e.g. humans: Bilgin et al., 2012; Narsaria et al., 2012; monkeys: Srivastava et al., 1992; mice: Nneji et al., 2013; Scaccabarozzi et al., 2018 and birds: Delhaye et al., 2016; Jiménez-Peñuela et al., 2023; van de Crommenacker et al., 2012). Antioxidant defences include both limiting oxidation and repairing cellular damage (Pamplona & Costantini, 2011). Among the candidate genes we identified in this study, *hspa8*, an isoform of heat shock protein 70 (*hsp70*), contributes to the protein damage stress response (Sottile & Nadin, 2018). *hsp70* has previously been shown to be highly differentiated between malaria-naïve and -exposed populations of Hawai'i 'amakihi (passerine birds) and therefore inferred to be under selection (Cassin-Sackett et al., 2019). Another candidate gene, *xrcc6bp1*, is potentially involved in the repair of DNA double-strand breaks (Fischer & Meese, 2007), but as far as we are aware, it has not been linked to malaria response previously. A variant near *sod1* was also identified. This gene encodes a well-known antioxidant enzyme that catalyses superoxide radicals, thereby regulating levels of ROS and limiting their potential toxicity (Fridavich, 1995). Superoxide dismutase has been repeatedly linked to malaria infection across a range of taxa, including evidence for lower antioxidant activity in birds (Jiménez-Peñuela et al., 2023), decreasing levels in monkeys (Srivastava et al., 1992) and downregulation of gene expression (Reuling et al., 2018) and links with parasitaemia in humans (Andrade et al., 2010). There is also evidence that polymorphism within *sod1* may be associated with malaria susceptibility (Fernandes

et al., 2015). Therefore, it is plausible that variation within or near these genes may mediate stress, minimise tissue damage and affect the host's ability to tolerate infection by malaria.

Many of the genes identified in our study are regulatory (e.g. *hspa8*, *smad2*, *c8A*, *c8B*, *sod1*, *mylk3*, *mlycD*, *kat6B*, *znf423*, *rpap3*, *crhr2*, *adgrd1*), which suggests that mechanisms for pathogen tolerance may be an important part of the response to malaria infection in Berthelot's pipit. Furthermore, our findings suggest malaria-mediated selection may drive the evolution of parts of the genome involved in diverse functions related to metabolic processes, stress response and transcriptional regulation, as well as the innate immune system and inflammatory response. This result underscores the importance of genome-wide approaches for the identification of candidate genes that underlie host-pathogen interactions, even when their relevance may not be immediately apparent. Some of the genes we have identified have multifarious roles in broad cellular processes (Table 1), therefore understanding the exact mechanism by which these genes influence the host's malaria response is not possible here. While much of the variation that we have identified seems to be linked with appropriate candidate genes that may putatively confer defence against malaria, we stress that validating an adaptive role requires further exploration through functional, experimental and genomic investigations. If the GEA has identified genes involved in local adaptation in response to malaria, then future research should establish the association among gene-specific variants, infection status and individual host fitness. This can be achieved through extensive within-population studies or – although challenging in non-model wild species like Berthelot's pipit – experimental infection studies.

General signatures of selection in the genome of Berthelot's pipit have previously been investigated at the population and archipelago level using differentiation-based analyses (Armstrong et al., 2018; Martin et al., 2021), with a single overlapping candidate gene region identified between studies at these different scales. In contrast to differentiation-based methods, GEA methods have the capability to identify the specific environmental variables driving selection, providing a more direct insight into the selective pressures at play. Here, we have specifically looked at loci that align with malaria prevalence across the geographic range. One of the genes identified in the current study (*znf598*), which codes for a zinc finger protein, was inferred to be under divergent selection between archipelagos (Armstrong et al., 2018). This family of proteins can interact with DNA, RNA and proteins and are involved in numerous cellular functions including transcriptional regulation, ubiquitin-mediated protein degradation and regulation of apoptosis (Cassandri et al., 2017), but the link to malaria protection is unclear. Thus, the most strongly divergent SNPs among populations of Berthelot's pipit (as identified in differentiation-based analyses) may not be shaped by pathogen pressures because we only detected a signal for one of the same genes using the GEA.

We detected genetic associations with population-level malaria prevalence within our RAD-generated SNP set. While RAD-seq is a cost-effective way to generate genome-wide marker sets, it is a

reduced representation sequencing approach so potentially important genes in unrepresented regions could have gone undetected. This could explain why we did not detect immune genes identified as being relevant for malaria resistance in other taxa (mentioned above), or additional immune genes overall since only a small fraction of avian immune genes can be evaluated with this dataset (Sheppard et al., 2022). Equally, some identified variants were not located within genes, which could either reflect the limited density of the marker set or potentially important functional non-coding regions. Even when associated variants were identified within genes, some lack research (e.g. *zcchc14*) making it difficult to interpret their potential function during infection. It is also important to note that while Bayenv is a powerful tool, it is not without weaknesses. False positives can derive from confounding correlations between environmental factors, which may be especially difficult to separate at the between-population level (see below), or even strong genetic drift. We employed a rather conservative approach, selecting only the candidates that ranked above high percentiles. While this could increase the number of false negatives, the aim of this study was to detect the top candidate gene set that merits future within-population investigation, rather than every possible association with malaria. Future research aimed at detecting additional variants under malaria-mediated selection, or indeed corroborating the candidate genes identified in the current study, may utilise alternative statistical approaches. Specifically, harnessing the capabilities of whole-genome sequencing data would facilitate a comprehensive identification of genetic variants. The broad coverage provided by whole-genome data enhances statistical power, offering a significant advantage in capturing subtle or weak signals of selection. Furthermore, exploiting museum specimens could prove valuable for characterising historical polymorphic SNPs that are close to or have reached fixation in contemporary populations due to positive or purifying selection.

Concurrent infections of avian malaria and poxvirus have been documented in a number of host species (e.g. Alley et al., 2010; Atkinson et al., 2005; Ha et al., 2012; Samuel et al., 2018), including Berthelot's pipits (Spurgin et al., 2012). Although high pathogen prevalence itself might suggest a strong likelihood of coinfection, our previous research demonstrated non-random cooccurrence within the pipit populations of Tenerife and Porto Santo through individual-based analyses (Sheppard et al., 2022). In that same study, we used a Bayenv analysis to identify genetic associations with pox prevalence differences between populations. Here, we show that patterns of both malaria and pox prevalence are similarly structured at the population level, thus making it technically difficult to isolate the effects of either pathogen in this natural system. Due to the univariate nature of Bayenv2.0, the model lacks the capability to account for shared variance among multiple predictors. Therefore, caution must be exercised when interpreting loci correlated with multiple environmental variables. While some identified genes may be associated with pox rather than malaria, and vice versa, there is variability allowing the detection of additional candidates involved in malaria protection. In the

current study, we identified 15 novel candidate genes (17 variants) putatively involved in malaria response in Berthelot's pipit that were not identified in the previous pox analysis. Conversely, seven genes (eight variants) identified in the previous pox analysis were not associated with malaria prevalence in the current study, and only five genes (six variants) were identified in both studies (*hspa8*, *smad2*, *ube2H*, *cwc27* and *mlycD*).

To better discern malaria's individual contribution as a driver of genetic variation, we also tested for genomic associations with 'residual malaria' prevalence. Beyond the three initially identified genes with 'raw' prevalence values (*c8A*, *c8B* and *crhr2*), the analysis identified an additional 10 candidate genes (16 SNPs). As expected, none of these genes, including the overlapping three, were identified in the previous pox study (Sheppard et al., 2022). A literature search did not reveal logical associations with malaria infection for these new genes (Table S3), however, a few have been identified as outliers/candidates in prior pipit studies using differentiation-based outlier tests (*wdr72*) (Armstrong et al., 2018; Martin et al., 2021) or in individual-level associations with other unrelated traits such as bill length (*plekhn1*) (Armstrong et al., 2018). Many of these outliers can be attributed to a combination of factors, including the demographic history of certain populations aligning with the gradient of malaria prevalence across populations. The 'residual' analysis, with reduced variation in pathogen data, has constrained power, potentially resulting in the dominance of drift-driven false positives. Notably, smaller populations like Selvagem Grande and Deserta Grande, having undergone extreme bottlenecks and being the only malaria-free populations, may be driving these new associations when power is limited (refer to Figure S2). Considering the analysis with 'raw' malaria prevalence had significantly greater power, associations with residual malaria did not yield a substantially improved candidate gene set. However, future in-depth analyses examining SNPs associated with malaria in this species should undoubtedly include *c8A*, *c8B* and *crhr2*.

Reasons for the association between malaria and pox in Berthelot's pipit are unknown but could include common risk factors that mediate exposure, such as host or vector densities, behaviours and climatic/habitat conditions (Hellard et al., 2015; Johnson & Buller, 2011). For example, the geographic distribution of malaria and pox infections might be shaped by similar anthropogenic factors at very local scales, as has been shown in other studies (Carrete et al., 2009; González-Quevedo et al., 2014; Padilla et al., 2017) and both pathogens can be transmitted among wild birds via mosquito vectors (Akey et al., 1981; Forrester, 1991). Alternatively, the co-occurrence of these pathogens may reflect pathogen-induced immunosuppression facilitating subsequent infection by a second pathogen (synergistic interaction) (Clark et al., 2016; Schat & Skinner, 2013). Many pathogenic infections, including malaria, are known to induce immunosuppression of the host response, increasing host susceptibility to other pathogens (Cox, 2001; Mabbott, 2018). An investigation of avian pox and malaria in Hawaiian birds by Samuel et al. (2018) concluded from epidemiological data that synergistic interactions or age-related acquisition

were likely explanations for positive associations between these two pathogens. We believe the latter is unlikely in our system since age was not a significant predictor of malaria (Spurgin et al., 2012) or pox infection status (Sheppard et al., 2022).

We have been operating under the assumption that the covariation of SNP frequencies with malaria prevalence reflects pathogen-mediated selection. It is important to note, however, that the causal relationship might also be reversed; populations with high frequencies of specific protective SNPs could potentially be more resistant to malaria. The mechanisms identified in this study, such as the regulation of inflammation, metabolism, stress response and cell death, likely represent broad defence mechanisms that could be activated during infections by various pathogens. If these mechanisms respond similarly to pox infection, then the cooccurrence patterns of avian malaria and pox across Berthelot's pipit populations could, theoretically, be attributed to spatial variation in underlying resistance or tolerance to multiple pathogens. Nevertheless, we consider this scenario unlikely, given that the spatial cooccurrence patterns align with biogeographical rules (Spurgin et al., 2012) and are also likely to be influenced by additional environmental factors (González-Quevedo et al., 2014). Regardless, the confounding factors that exist within the data presented here make it difficult to tease apart the initial drivers of pathogen cooccurrence.

Using the literature across taxa, we categorised genes identified in the GEA pipit studies (Sheppard et al., 2022, and the present study; 27 in total) that might overlap due to the collinearity between pox and malaria prevalence across populations. This categorisation was based on whether these genes have previously been directly linked to malaria, pox or both (Table 2). Among these, five genes exhibited associations with both pox and malaria in Berthelot's pipit (Tables 1 and 2). Two of these had previously been linked with both malaria and pox in other species and taxa, two had been linked with malaria only and none had been linked exclusively to pox only in other species or taxa. For genes identified exclusively in either of our two pipit studies (i.e. for an association with malaria prevalence only or an association with pox prevalence only), only malaria-associated candidates could be linked with malaria and/or pox in other species or taxa. However, a lack of research and annotation for many genes is likely to be an issue. For example, it seems genes with easily identifiable functional pathways, such as the heat shock response, complement pathway and corticosterone stress response, could clearly be linked to both pathogens. The greater number of genes linked to malaria does not necessarily imply a higher involvement of candidates in malaria response compared to pox response. It might reflect an imbalance in research efforts into malaria and pox viruses. Ideally, we would leverage the gradient of malaria and pox prevalence across populations of Berthelot's pipits (Figure 2) to determine the most suitable populations to try and isolate the effect of each pathogen. In many populations where one pathogen appears to be absent, the other is present but at low infection rates (c. 1%–5%). Consequently, selective pressure from that pathogen might be too weak to detect any associations. Outlier populations

TABLE 2 Evidence from the literature for links between the candidate genes identified in Berthelot's pipit and malaria and/or poxviruses in other species/taxa.

Candidate gene	Associated pathogen	Links with poxviruses	Links with malaria
<i>hspa8</i>	Pox & Malaria	<i>hsp70</i> upregulated during poxvirus infection in mouse and human models (Brum et al., 2003; Cheng et al., 2018; Jindal & Young, 1992; Kowalczyk et al., 2005; Sedger & Ruby, 1994)	<i>hsp70</i> highly differentiated between malaria-naïve and -exposed populations of a passerine bird (Cassin-Sackett et al., 2019)
<i>smad2</i>	Pox & Malaria	-	Regulates inflammatory response as part of the TGF- β pathway; TGF- β itself is inferred to be under divergent selection from avian malaria (Cassin-Sackett et al., 2019)
<i>ube2H</i>	Pox & Malaria	-	Overexpressed in malaria-infected passerine, humans and mice (Videvall et al., 2015)
<i>cwc27</i>	Pox & Malaria	Cyclophilins play a role against poxviruses in cell cultures (Castro et al., 2003; Zhou et al., 2021)	Overexpressed in malaria-infected passerine (Videvall et al., 2015)
<i>mlycD</i>	Pox & Malaria	-	-
<i>c8A</i>	Malaria	Complement-related genes differentially expressed in pox-infected passerine birds (McNew et al., 2021). Complement may play a role in poxvirus response in mouse models (Miller et al., 1995)	Complement system is activated against malaria (Kurtovic et al., 2020; Silver et al., 2010) and may act to control parasitaemia in passerine birds (Ellis et al., 2015). Complement-related genes are differentially expressed in malaria-infected passerine (Videvall et al., 2015)
<i>c8B</i>	Malaria		
<i>scaf4</i>	Malaria	-	-
<i>sod1</i>	Malaria	<i>Leporipoxvirus</i> disrupts host <i>sod1</i> activity (Cao et al., 2022; Teoh et al., 2003, 2005)	Lower activity in infected birds (Jiménez-Peñuela et al., 2023). Decreasing levels during course of infection in monkeys (Srivastava et al., 1992). Downregulated in human infections (Reuling et al., 2018) and associated with parasitaemia (Andrade et al., 2010). Polymorphism may be associated with malaria susceptibility (Fernandes et al., 2015)
<i>mylk3</i>	Malaria	-	-
<i>dupd1</i>	Malaria	-	-
<i>kat6B</i>	Malaria	-	-
<i>zcchc14</i>	Malaria	-	-
<i>znf423</i>	Malaria	-	-
<i>xrcc6bp1</i>	Malaria	-	-
<i>rpap3</i>	Malaria	-	-
<i>crhr2</i>	Malaria	-	Modulates the release of corticosterone, the main glucocorticoid stress hormone in birds, for which high levels may result in higher malaria parasite load in a passerine bird (Names, Schultz, Krause, et al., 2021)
<i>adgrd1</i>	Malaria	-	-
<i>znf598</i>	Malaria	Co-opted by poxviruses in cell cultures for replication and the synthesis of viral proteins (DiGiuseppe et al., 2018; Sundaramoorthy et al., 2021)	-
<i>syng3</i>	Malaria	-	-
<i>mthfd1L</i>	Pox	-	-
<i>ablim3</i>	Pox	-	-
<i>Oat</i>	Pox	-	-
<i>mfsd2A</i>	Pox	-	-
<i>Mycl</i>	Pox	-	-
<i>akap13</i>	Pox	-	-

Note: Candidate genes listed here were identified by testing for associations with population-level avian malaria (in this study) or avian pox (Sheppard et al., 2022) prevalence in Bayenv. Genes are first ordered according to the associated pathogen(s) and then by decreasing Bayes factor value.

with significantly higher prevalence of one pathogen compared to the other, such as La Gomera and Tenerife (Figure 2), might provide the best conditions to disentangle these associations and further study the identified candidates. Indeed, extensive within-population investigations focused on the candidate genes identified here, incorporating infection data for both pathogens could help to verify these relationships.

For both pathogens, the molecular identification of strains was limited to a single genetic marker, the partial sequence of either the mitochondrial cytochrome *b* gene (malaria) or the *4b* core protein gene (pox). In some cases, even within a lineage, there can be considerable variation and population structure at other genes throughout the genome, including those potentially linked to infection success (Hellgren et al., 2015; Huang et al., 2019; Lauron et al., 2014). It is therefore possible that undetected population structure in the pathogens among archipelagos or islands could drive the evolution of host defence at a finer geographic scale than considered in the current study, and as such, we would be unable to detect these archipelago/population-specific malaria-associated loci.

5 | CONCLUSIONS

Our study provides correlational evidence that heterogenous pathogen-mediated selective pressure among populations of Berthelot's pipit drives spatial patterns of host genetic variation. It also highlights the methodological challenges involved in teasing apart the separate selective effects of different co-occurring pathogens (avian malaria and pox) in the same natural host system. Despite this, our findings suggest that genes linked to stress response, transcriptional regulation, metabolic processes and the innate immune system may be involved in malaria response, or pathogen response more broadly, in this system. Any variation associated with malaria in this study needs to be followed up with extensive within-population investigations, for which we now have proposed a best-candidate gene set.

AUTHOR CONTRIBUTIONS

ECS and DSR designed the study, with input from CAM. ECS conducted all analyses and drafted the manuscript with input and advice from DSR and AS. ECS, CAM, CA, CG-Q, JCI, LGS and DSR undertook fieldwork and performed molecular laboratory work. CA provided restriction site-associated DNA sequencing data. All authors contributed critically to this work and approved the final manuscript for publication.

ACKNOWLEDGEMENTS

This study would not have been possible without the assistance of fieldworkers, laboratory technicians and students throughout the whole study period. We are grateful to the Regional Governments of the Canary Islands (Ref.: 2019/5555), the Cabildo of Lanzarote (Ref.: 101/2019), the Cabildo of Tenerife (Ref.: 2019-01740) and the

National Parks of El Teide (Ref.: MDV/amp) and Madeira (Ref.: 06/IFCN/2020) for their permission to sample and ring Berthelot's pipits for this research, the Regional Governments of the Canary Islands and Madeira for providing accommodation and the Portuguese Navy for transport to Selvagem Grande and Deserta Grande. Bioinformatics analyses were carried out on the High-Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia. This work was supported by Natural Environment Research Council (NERC) studentships awarded to ECS through the ARIES DTP (NE/S007334/1), and to CAM and CA through the EnvEAST DTP (NE/L002582/1), the former of which is a CASE award with the Institute of Natural Products and Agrobiolgy (IPNA-CSIC). JCI was funded by a research grant from the Spanish Ministry of Science, Innovation and Universities and the European Regional Development Fund (PGC2018-097575-B-I00) and by a Regional Government of Asturias grant from the Regional Government of Asturias (AYUD/2021/51261).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

All data and the code used to perform the analysis within this manuscript are openly available on the Dryad Digital Repository: <https://doi.org/10.5061/dryad.k6djh9w9j>.

ORCID

Eleanor C. Sheppard  <https://orcid.org/0000-0002-4986-7268>
 Claudia A. Martin  <https://orcid.org/0000-0003-2645-0790>
 Claire Armstrong  <https://orcid.org/0000-0003-3874-4243>
 Alexander Suh  <https://orcid.org/0000-0002-8979-9992>
 Lewis G. Spurgin  <https://orcid.org/0000-0002-0874-9281>
 David S. Richardson  <https://orcid.org/0000-0001-7226-9074>

REFERENCES

- Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., & Wayne, R. K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 3490–3494.
- Akey, B. L., Nayar, J. K., & Forrester, D. J. (1981). Avian pox in Florida wild Turkeys: *Culex nigripalpus* and *Wyeomyia vanduzeei* as experimental vectors. *Journal of Wildlife Diseases*, 17, 597–599.
- Alley, M. R., Ha, H. J., Howe, L., Hale, K. A., & Cash, W. (2010). Concurrent avian malaria and avipox virus infection in translocated South Island saddlebacks (*Philesturnus carunculatus carunculatus*). *New Zealand Veterinary Journal*, 58, 218–223.
- Anderson, R. M., & May, R. M. (1979). Population biology of infectious diseases: Part I. *Nature*, 280, 361–367.
- Andrade, B. B., Reis-Filho, A., Souza-Neto, S. M., Raffaele-Netto, I., Camargo, L. M. A., Barral, A., & Barral-Netto, M. (2010). Plasma superoxide dismutase-1 as a surrogate marker of vivax malaria severity. *PLoS Neglected Tropical Diseases*, 4, e650.
- Armstrong, C., Davies, R. G., González-Quevedo, C., Dunne, M., Spurgin, L. G., & Richardson, D. S. (2019). Adaptive landscape genetics and malaria across divergent Island bird populations. *Ecology and Evolution*, 9, 12482–12502.

- Armstrong, C., Richardson, D. S., Hipperson, H., Horsburgh, G. J., Küpper, C., Percival-Alwyn, L., Clark, M., Burke, T., & Spurgin, L. G. (2018). Genomic associations with bill length and disease reveal drift and selection across Island bird populations. *Evolution Letters*, 2, 22–36.
- Asghar, M., Hasselquist, D., & Bensch, S. (2011). Are chronic avian haemosporean infections costly in wild birds? *Journal of Avian Biology*, 42, 530–537.
- Asghar, M., Hasselquist, D., Hansson, B., Zehtindjiev, P., Westerdahl, H., & Bensch, S. (2015). Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science*, 347, 436–438.
- Atkinson, C. T., Lease, J. K., Dusek, R. J., & Samuel, M. D. (2005). Prevalence of pox-like lesions and malaria in forest bird communities on leeward Mauna Loa Volcano, Hawaii. *The Condor*, 107, 537–546.
- Atkinson, C. T., Saili, K. S., Uzzurum, R. B., & Jarvi, S. I. (2013). Experimental evidence for evolved tolerance to avian malaria in a wild population of low elevation Hawai'i 'Amakihi (*Hemignathus virens*). *EcoHealth*, 10, 366–375.
- Atkinson, C. T., Woods, K. L., Dusek, R. J., Sileo, L. S., & Iko, W. M. (1995). Wildlife disease and conservation in Hawaii: Pathogenicity of avian malaria (*Plasmodium relictum*) in experimentally infected liwi (*Vestiaria coccinea*). *Parasitology*, 111, 59–69.
- Bayly-Jones, C., Bubeck, D., & Dunstone, M. A. (2017). The mystery behind membrane insertion: A review of the complement membrane attack complex. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 372, 20160221.
- Beaumont, M. A. (2005). Adaptation and speciation: What can Fst tell us? *Trends in Ecology & Evolution*, 20, 435–440.
- Bilgin, R., Yalcin, M. S., Yucebiligic, G., Koltas, I. S., & Yazar, S. (2012). Oxidative stress in vivax malaria. *Korean Journal of Parasitology*, 50, 375–377.
- Blair, L. M., Granka, J. M., & Feldman, M. W. (2014). On the stability of the Bayenv method in assessing human SNP-environment associations. *Human Genomics*, 8, 1–13.
- Bonneaud, C., Federici, P., Sorci, G., Pérez-Tris, J., & Chastel, O. (2006). Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, 60, 383–389.
- Boštjančić, L. L., Francesconi, C., Rutz, C., Hoffbeck, L., Poidevin, L., Kress, A., Jussila, J., Makkonen, J., Feldmeyer, B., Bálint, M., Schwenk, K., Lecompte, O., & Theissinger, K. (2022). Host-pathogen coevolution drives innate immune response to *Aphanomyces astaci* infection in freshwater crayfish: Transcriptomic evidence. *BMC Genomics*, 23, 600.
- Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9, 378–400.
- Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J., & Richardson, D. S. (2010). MHC-dependent survival in a wild population: Evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, 19, 3444–3455.
- Brum, L. M., Lopez, M. C., Varela, J. C., Baker, H. V., & Moyer, R. W. (2003). Microarray analysis of A549 cells infected with rabbitpox virus (RPV): A comparison of wild-type RPV and RPV deleted for the host range gene, SPI-1. *Virology*, 315, 322–334.
- Busetto, V., Barbosa, I., Basquin, J., Marquet, E., Hocq, R., Hennion, M., Paternina, J. A., Namane, A., Conti, E., Bensaude, O., & Le Hir, H. (2020). Structural and functional insights into CWC27/CWC22 heterodimer linking the exon junction complex to spliceosomes. *Nucleic Acids Research*, 48, 5670–5683.
- Cao, J., Teoh, M., Moon, M., & McFadden, G. (2022). *Leporipoxvirus* Cu-Zn superoxide dismutase homologs inhibit cellular superoxide dismutase, but are not essential for virus replication or virulence. *Virology*, 296, 125–135.
- Carrete, M., Serrano, D., Illera, J. C., López, G., Vögeli, M., Delgado, A., & Tella, J. L. (2009). Goats, birds, and emergent diseases: Apparent and hidden effects of exotic species in an Island environment. *Ecological Applications*, 19, 840–853.
- Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., & Raschellà, G. (2017). Zinc-finger proteins in health and disease. *Cell Death Discovery*, 3, 17071.
- Cassin-Sackett, L., Callicrate, T. E., & Fleischer, R. C. (2019). Parallel evolution of gene classes, but not genes: Evidence from Hawai'iian honeycreeper populations exposed to avian malaria. *Molecular Ecology*, 28, 568–583.
- Castro, A. P. V., Carvalho, T. M. U., Moussatché, N., & Damaso, C. R. A. (2003). Redistribution of cyclophilin a to viral factories during vaccinia virus infection and its incorporation into mature particles. *Journal of Virology*, 77, 9052–9068.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7.
- Charlesworth, D. (2006). Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, 2, 379–384.
- Chen, G. B., Lee, S. H., Zhu, Z. X., Benyamin, B., & Robinson, M. R. (2016). EigenGWAS: Finding loci under selection through genome-wide association studies of eigenvectors in structured populations. *Heredity*, 117, 51–61.
- Cheng, W., Jia, H., Wang, X., He, X., Jin, Q., Cao, J., & Jing, Z. (2018). Ectromelia virus upregulates the expression of heat shock protein 70 to promote viral replication. *International Journal of Molecular Medicine*, 42, 1044–1053.
- Chu, Z. J., Wang, Y. J., Ying, S. H., Wang, X. W., & Feng, M. G. (2016). Genome-wide host-pathogen interaction unveiled by transcriptomic response of diamondback moth to fungal infection. *PLoS One*, 11, e0152908.
- Ciloglu, A., Ellis, V. A., Bernotienė, R., Valkiūnas, G., & Bensch, S. (2019). A new one-step multiplex PCR assay for simultaneous detection and identification of avian haemosporean parasites. *Parasitology Research*, 118, 191–201.
- Clark, N. J., Wells, K., Dimitrov, D., & Clegg, S. M. (2016). Co-infections and environmental conditions drive the distributions of blood parasites in wild birds. *Journal of Animal Ecology*, 85, 1461–1470.
- Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185, 1411–1423.
- Cox, F. E. G. (2001). Concomitant infections, parasites and immune responses. *Parasitology*, 122, S23–S28.
- DaCosta, J. M., & Sorenson, M. D. (2014). Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS One*, 9, e106713.
- Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Emerging infectious diseases of wildlife – Threats to biodiversity and human health. *Science*, 287, 443–449.
- Davies, C. S., Taylor, M. I., Hammers, M., Burke, T., Komdeur, J., Dugdale, H. L., & Richardson, D. S. (2021). Contemporary evolution of the innate immune receptor gene TLR3 in an isolated vertebrate population. *Molecular Ecology*, 30, 2528–2542.
- Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society B: Biological Sciences*, 205, 489–511.
- Delhaye, J., Jenkins, T., & Christe, P. (2016). *Plasmodium* infection and oxidative status in breeding great tits, *Parus major*. *Malaria Journal*, 15, 531.
- DiGiuseppe, S., Rollins, M. G., Bartom, E. T., & Walsh, D. (2018). ZNF598 plays distinct roles in interferon-stimulated gene expression and poxvirus protein synthesis. *Cell Reports*, 23, 1249–1258.
- Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, 256, 50–52.

- Douma, J. C., & Weedon, J. T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, 10, 1412–1430.
- Ellis, V. A., Cornet, S., Merrill, L., Kunkel, M. R., Tsunekage, T., & Ricklefs, R. E. (2015). Host immune responses to experimental infection of *Plasmodium relictum* (lineage SGS1) in domestic canaries (*Serinus canaria*). *Parasitology Research*, 114, 3627–3636.
- Fernandes, R. C., Hasan, M., Gupta, H., Geetha, K., Rai, P. S., Hande, M. H., D'Souza, S. C., Adhikari, P., Brand, A., & Satyamoorthy, K. (2015). Host genetic variations in glutathione-S-transferases, superoxide dismutases and catalase genes influence susceptibility to malaria infection in an Indian population. *Molecular Genetics and Genomics*, 290, 1155–1168.
- Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., & Muntasell, A. (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. *Journal of Immunology*, 181, 1315–1322.
- Fischer, U., & Meese, E. (2007). Glioblastoma multiforme: The role of DSB repair between genotype and phenotype. *Oncogene*, 26, 7809–7815.
- Forrester, D. J. (1991). The ecology and epizootiology of avian pox and malaria in wild turkeys. *Bulletin of the Society for Vector Ecology*, 16, 127–148.
- Foster, J. T., Woodworth, B. L., Eggert, L. E., Hart, P. J., Palmer, D., Duffy, D. C., & Fleischer, R. C. (2007). Genetic structure and evolved malaria resistance in Hawaiian honeycreepers. *Molecular Ecology*, 16, 4738–4746.
- Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30, 1687–1699.
- Fridavich, I. (1995). Superoxide radical and superoxide dismutases. *Annual Review of Biochemistry*, 64, 97–112.
- Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admetlla, A., Pattini, L., & Nielsen, R. (2011). Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics*, 7, e1002355.
- Garroway, C. J., Radersma, R., Sepil, I., Santure, A. W., De Cauwer, I., Slate, J., & Sheldon, B. C. (2013). Fine-scale genetic structure in a wild bird population: The role of limited dispersal and environmentally based selection as causal factors. *Evolution*, 67, 3488–3500.
- Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, 201, 1555–1579.
- González-Quevedo, C., Davies, R. G., Phillips, K. P., Spurgin, L. G., & Richardson, D. S. (2016). Landscape-scale variation in an anthropogenic factor shapes immune gene variation within a wild population. *Molecular Ecology*, 25, 4234–4246.
- González-Quevedo, C., Davies, R. G., & Richardson, D. S. (2014). Predictors of malaria infection in a wild bird population: Landscape-level analyses reveal climatic and anthropogenic factors. *Journal of Animal Ecology*, 83, 1091–1102.
- González-Quevedo, C., Spurgin, L. G., Illera, J. C., & Richardson, D. S. (2015). Drift, not selection, shapes toll-like receptor variation among oceanic Island populations. *Molecular Ecology*, 24, 5852–5863.
- Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, 195, 205–220.
- Gunther, T., & Coop, G. (2018). *A short manual for Bayenv2.0*.
- Ha, H. J., Banda, M., Alley, M. R., Howe, L., & Gartrell, B. D. (2012). The seroprevalence of avipoxvirus and its association with avian malaria (*Plasmodium* spp.) infection in introduced passerine birds in the southern regions of the North Island of New Zealand. *Avian Diseases*, 57, 109–115.
- Hancock, A. M., Witonsky, D. B., Ehler, E., Alkorta-Aranburu, G., Beall, C., Gebremedhin, A., Sukernik, R., Utermann, G., Pritchard, J., Coop, G., & di Rienzo, A. (2010). Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 8924–8930.
- Hawley, D. M., & Fleischer, R. C. (2012). Contrasting epidemic histories reveal pathogen-mediated balancing selection on class II MHC diversity in a wild songbird. *PLoS One*, 7, e30222.
- Hellard, E., Fouchet, D., Vavre, F., & Pontier, D. (2015). Parasite-parasite interactions in the wild: How to detect them? *Trends in Parasitology*, 31, 640–652.
- Hellgren, O., Atkinson, C. T., Bensch, S., Albayrak, T., Dimitrov, D., Ewen, J. G., Kim, K. S., Lima, M. R., Martin, L., Palinauskas, V., Ricklefs, R., Sehgal, R. N. M., Valkiūnas, G., Tsuda, Y., & Marzal, A. (2015). Global phylogeography of the avian malaria pathogen *Plasmodium relictum* based on MSP1 allelic diversity. *Ecography*, 38, 842–850.
- Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., Bennett, S., Brewster, D., McMichael, A. J., & Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature*, 352, 595–600.
- Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, 188, 379–397.
- Huang, X., Rapševičius, P., Chapa-Vargas, L., Hellgren, O., & Bensch, S. (2019). Within-lineage divergence of avian haemosporidians: A case study to reveal the origin of a widespread *Haemoproteus* parasite. *Journal of Parasitology*, 105, 414–422.
- Huang, Y., Li, Y., Burt, D. W., Chen, H., Zhang, Y., Qian, W., Kim, H., Gan, S., Zhao, Y., Li, J., Yi, K., Feng, H., Zhu, P., Li, B., Liu, Q., Fairley, S., Magor, K. E., du, Z., Hu, X., ... Li, N. (2013). The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nature Genetics*, 45, 776–783.
- Illera, J. C., Emerson, B. C., & Richardson, D. S. (2007). Population history of Berthelot's pipit: Colonization, gene flow and morphological divergence in Macaronesia. *Molecular Ecology*, 16, 4599–4612.
- Illera, J. C., Emerson, B. C., & Richardson, D. S. (2008). Genetic characterization, distribution and prevalence of avian pox and avian malaria in the Berthelot's pipit (*Anthus berthelotii*) in Macaronesia. *Parasitology Research*, 103, 1435–1443.
- Illera, J. C., López, G., García-Padilla, L., & Moreno, Á. (2017). Factors governing the prevalence and richness of avian haemosporidian communities within and between temperate mountains. *PLoS One*, 12, e0184587.
- Jiménez-Peñuela, J., Ferraguti, M., Martínez-De La Puente, J., Soriguer, R. C., & Figuerola, J. (2023). Oxidative status in relation to blood parasite infections in house sparrows living along an urbanization gradient. *Environmental Pollution*, 316, 120712.
- Jindal, S., & Young, R. A. (1992). Vaccinia virus infection induces a stress response that leads to association of Hsp70 with viral proteins. *Journal of Virology*, 66, 5357–5362.
- Johnson, P. T. J., & Buller, I. D. (2011). Parasite competition hidden by correlated coinfection: Using surveys and experiments to understand parasite interactions. *Ecology*, 92, 535–541.
- Kloch, A., Babik, W., Bajer, A., Siński, E., & Radwan, J. (2010). Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole *Myodes glareolus*. *Molecular Ecology*, 19, 255–265.
- Kowalczyk, A., Guzik, K., Slezak, K., Dziedzic, J., & Rokita, H. (2005). Heat shock protein and heat shock factor 1 expression and localization in vaccinia virus infected human monocyte derived macrophages. *Journal of Inflammation*, 2, 1–10.

- Kurtovic, L., Boyle, M. J., Opi, D. H., Kennedy, A. T., Tham, W. H., Reiling, L., Chan, J. A., & Beeson, J. G. (2020). Complement in malaria immunity and vaccines. *Immunological Reviews*, 293, 38–56.
- Lachish, S., Knowles, S. C. L., Alves, R., Wood, M. J., & Sheldon, B. C. (2011). Fitness effects of endemic malaria infections in a wild bird population: The importance of ecological structure. *Journal of Animal Ecology*, 80, 1196–1206.
- Lapointe, D. A., Atkinson, C. T., & Samuel, M. D. (2012). Ecology and conservation biology of avian malaria. *Annals of the New York Academy of Sciences*, 1249, 211–226.
- Lauron, E. J., Oakgrove, K. S., Tell, L. A., Biskar, K., Roy, S. W., & Sehgal, R. N. M. (2014). Transcriptome sequencing and analysis of *Plasmodium gallinaceum* reveals polymorphisms and selection on the apical membrane antigen-1. *Malaria Journal*, 13, 382.
- Lewontin, R. C., & Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74, 175–195.
- Lochmiller, R. L., & Deerenberg, C. (2000). Trade-offs in evolutionary immunology: Just what is the cost of immunity? *Oikos*, 88, 87–98.
- Loiseau, C., Richard, M., Garnier, S., Chastel, O., Julliard, R., Zoorob, R., & Sorci, G. (2009). Diversifying selection on MHC class I in the house sparrow (*Passer domesticus*). *Molecular Ecology*, 18, 1331–1340.
- Longley, R., Smith, C., Fortin, A., Berghout, J., McMorran, B., Burgio, G., Foote, S., & Gros, P. (2011). Host resistance to malaria: Using mouse models to explore the host response. *Mammalian Genome*, 22, 32–42.
- Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24, 1031–1046.
- Mabbott, N. A. (2018). The influence of parasite infections on host immunity to co-infection with other pathogens. *Frontiers in Immunology*, 9, 2579.
- Mackinnon, M. J., Ndila, C., Uyoga, S., Macharia, A., Snow, R. W., Band, G., Rautanen, A., Rockett, K. A., Kwiatkowski, D. P., & Williams, T. N. (2016). Environmental correlation analysis for genes associated with protection against malaria. *Molecular Biology and Evolution*, 33, 1188–1204.
- MalariaGEN. (2019). Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. *Nature Communications*, 10, 1–19.
- Martin, C. A., Armstrong, C., Illera, J. C., Emerson, B. C., Richardson, D. S., & Spurgin, L. G. (2021). Genomic variation, population history and within-archipelago adaptation between Island bird populations. *Royal Society Open Science*, 8, 201146.
- Martin, C. A., Sheppard, E. C., Illera, J. C., Suh, A., Nadachowska-Brzyska, K., Spurgin, L. G., & Richardson, D. S. (2023). Runs of homozygosity reveal past bottlenecks and contemporary inbreeding across diverging Island populations of a bird. *Molecular Ecology*, 32, 1972–1989.
- Martin, L. B., Scheuerlein, A., & Wikelski, M. (2003). Immune activity elevates energy expenditure of house sparrows: A link between direct and indirect costs? *Proceedings of the Royal Society of London B*, 270, 153–158.
- Marzal, A., Bensch, S., Reviriego, M., Balbontin, J., & De Lope, F. (2008). Effects of malaria double infection in birds: One plus one is not two. *Journal of Evolutionary Biology*, 21, 979–987.
- McNew, S. M., Loyola, D. C., Yopez, J., Andreadis, C., Gotanda, K., Saulsbury, A., & Fessl, B. (2022). Transcriptomic responses of Galápagos finches to avian poxvirus infection. *Molecular Ecology*, 31, 5552–5567.
- McNew, S. M., Yopez, J., Loyola, C. D., Andreadis, C., & Fessl, B. (2021). Transcriptomic responses of Galápagos finches to avian pox virus infection. *BioRxiv*, <https://doi.org/10.1101/2021.10.15.46458>
- Miller, C. G., Justus, D. E., Jayaraman, S., & Kotwal, G. J. (1995). Severe and prolonged inflammatory response to localized cowpox virus infection in footpads of C5-deficient mice: Investigation of the role of host complement in poxvirus pathogenesis. *Cellular Immunology*, 162, 326–332.
- Mukhin, A., Palinauskas, V., Platonova, E., Kobylkov, D., Vakoliuk, I., & Valkiunas, G. (2016). The strategy to survive primary malaria infection: An experimental study on behavioural changes in parasitized birds. *PLoS One*, 11, e0159216.
- Musa, S., Mackenstedt, U., Woog, F., & Dinkel, A. (2022). Untangling the actual infection status: Detection of avian haemosporidian parasites of three Malagasy bird species using microscopy, multiplex PCR, and nested PCR methods. *Parasitology Research*, 121, 2817–2829.
- Names, G. R., Schultz, E. M., Hahn, T. P., Hunt, K. E., Angelier, F., Ribout, C., & Klasing, K. C. (2021). Variation in immunity and health in response to introduced avian malaria in an endemic Hawaiian songbird. *Animal Conservation*, 25, 455–466.
- Names, G. R., Schultz, E. M., Krause, J. S., Hahn, T. P., Wingfield, J. C., Heal, M., Cornelius, J. M., Klasing, K. C., & Hunt, K. E. (2021). Stress in paradise: Effects of elevated corticosterone on immunity and avian malaria resilience in a Hawaiian passerine. *The Journal of Experimental Biology*, 224, 1–9. <https://doi.org/10.1242/jeb.242951>
- Narsaria, N., Mohanty, C., Das, B. K., Mishra, S. P., & Prasad, R. (2012). Oxidative stress in children with severe malaria. *Journal of Tropical Pediatrics*, 58, 147–150.
- Nneji, C. M., Adaramoye, O. A., Falade, C. O., & Ademowo, O. G. (2013). Effect of chloroquine, methylene blue and artemether on red cell and hepatic antioxidant defence system in mice infected with *Plasmodium yoelii nigeriensis*. *Parasitology Research*, 112, 2619–2625.
- Ots, I., Kerimov, A. B., Ivankina, E. V., Ilyina, T. A., & Hórák, P. (2001). Immune challenge affects basal metabolic activity in wintering great tits. *Proceedings of the Royal Society of London B*, 268, 1175–1181.
- Padilla, D. P., Illera, J. C., Gonzalez-Quevedo, C., Villalba, M., & Richardson, D. S. (2017). Factors affecting the distribution of haemosporidian parasites within an oceanic Island. *International Journal for Parasitology*, 47, 225–235.
- Palinauskas, V., Valkiunas, G., Bolshakov, C. V., & Bensch, S. (2008). *Plasmodium relictum* (lineage P-SG51): Effects on experimentally infected passerine birds. *Experimental Parasitology*, 120, 372–380.
- Pamplona, R., & Costantini, D. (2011). Molecular and structural antioxidant defenses against oxidative stress in animals. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, 301, 843–863.
- Pardo-Diaz, C., Salazar, C., & Jiggins, C. D. (2015). Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution*, 6, 445–464.
- Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D., Libberton, B., Fenton, A., Hall, N., & Brockhurst, M. A. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature*, 464, 275–278.
- Pigeault, R., Cozzarolo, C., Glaizot, O., & Christe, P. (2020). Effect of age, haemosporidian infection and body condition on pair composition and reproductive success in great tits *Parus major*. *Ibis*, 162, 613–626.
- Popa, G. L., & Popa, M. I. (2021). Recent advances in understanding the inflammatory response in malaria: A review of the dual role of cytokines. *Journal of Immunology Research*, 2021, 1–9.
- Quéméré, E., Galan, M., Cosson, J.-F., Klein, F., Aulagnier, S., Gilot-Fromont, E., Merlet, J., Bonhomme, M., Hewison, A. J. M., & Charbonnel, N. (2015). Immunogenetic heterogeneity in a widespread ungulate: The European roe deer (*Capreolus capreolus*). *Molecular Ecology*, 24, 3873–3887.
- Quin, J. E., Bujila, I., Chérif, M., Sanou, G. S., Qu, Y., Vafa Homann, M., Rolicka, A., Sirima, S. B., O'Connell, M. A., Lennartsson, A., Troye-Blomberg, M., Nebie, I., & Östlund Farrants, A. K. (2017). Major transcriptional changes observed in the Fulani, an ethnic group less susceptible to malaria. *eLife*, 6, 1–19.

- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Råberg, L., Graham, A. L., & Read, A. F. (2009). Decomposing health: Tolerance and resistance to parasites in animals. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *364*, 37–49.
- Råberg, L., Sim, D., & Read, A. F. (2007). Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science*, *318*, 812–814.
- Ravenhall, M., Campino, S., Sepúlveda, N., Manjurano, A., Nadjim, B., Mtove, G., Wangai, H., Maxwell, C., Olomi, R., Reyburn, H., Drakeley, C. J., Riley, E. M., Clark, T. G., & in collaboration with MalariaGEN. (2018). Novel genetic polymorphisms associated with severe malaria and under selective pressure in North-eastern Tanzania. *PLoS Genetics*, *14*, e1007172.
- Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*, 4348–4370.
- Reuling, I. J., de Jong, G. M., Yap, X. Z., Asghar, M., Walk, J., van de Schans, L. A., Koelewijn, R., Färnert, A., de Mast, Q., van der Ven, A. J., Bousema, T., van Hellemond, J. J., van Genderen, P. J. J., & Sauerwein, R. W. (2018). Liver injury in uncomplicated malaria is an overlooked phenomenon: An observational study. *eBioMedicine*, *36*, 131–139.
- Richardson, D. S., Jury, F. L., 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*, 2263–2273.
- Rockett, K. A., Clarke, G. M., Fitzpatrick, K., Hubbart, C., Jeffreys, A. E., Rowlands, K., Craik, R., Jallow, M., Conway, D. J., Bojang, K. A., Pinder, M., Usen, S., Sisay-Joof, F., Sirugo, G., Toure, O., Thera, M. A., Konate, S., Sissoko, S., Niangaly, A., ... Kwiatkowski, D. P. (2014). Reappraisal of known malaria resistance loci in a large multicenter study. *Nature Genetics*, *46*, 1197–1204.
- Roestenberg, M., McCall, M., Mollnes, T. E., van Deuren, M., Sprong, T., Klasen, I., Hermsen, C. C., Sauerwein, R. W., & van der Ven, A. (2007). Complement activation in experimental human malaria infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *101*, 643–649.
- Rosenblum, E. B., Poorten, T. J., Settles, M., & Murdoch, G. K. (2012). Only skin deep: Shared genetic response to the deadly chytrid fungus in susceptible frog species. *Molecular Ecology*, *21*, 3110–3120.
- Samuel, M. D., Woodworth, B. L., Atkinson, C. T., Hart, P. J., & LaPointe, D. A. (2018). The epidemiology of avian pox and interaction with avian malaria in Hawaiian forest birds. *Ecological Monographs*, *88*, 621–637.
- Scaccabarozzi, D., Deroost, K., Corbett, Y., Lays, N., Corsetto, P., Salè, F. O., Van den Steen, P. E., & Taramelli, D. (2018). Differential induction of malaria liver pathology in mice infected with *Plasmodium chabaudi* AS or *Plasmodium berghei* NK65. *Malaria Journal*, *17*, 18.
- Schat, K. A., & Skinner, M. A. (2013). Avian immunosuppressive diseases and immune evasion. In *Avian immunology* (2nd ed., pp. 275–297). Academic Press.
- Schoville, S. D., Bonin, A., François, O., Lobreaux, S., Melodelima, C., & Manel, S. (2012). Adaptive genetic variation on the landscape: Methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 23–43.
- Sedger, L., & Ruby, J. (1994). Heat shock response to vaccinia virus infection. *Journal of Virology*, *68*, 4685–4689.
- Sehgal, R. N. M. (2015). Manifold habitat effects on the prevalence and diversity of avian blood parasites. *International Journal for Parasitology: Parasites and Wildlife*, *4*, 421–430.
- Sepil, I., Lachish, S., Hinks, A. E., & Sheldon, B. C. (2013). MHC super-types confer both qualitative and quantitative resistance to avian malaria infections in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20130134.
- Sheppard, E. C., Martin, C. A., Illera, J. C., Suh, A., Spurgin, L. G., & Richardson, D. S. (2022). Genomic associations with pox virus across divergent populations in Berthelot's pipit. *Molecular Ecology*, *31*, 3154–3173.
- Silver, K. L., Higgins, S. J., McDonald, C. R., & Kain, K. C. (2010). Complement driven innate immune response to malaria: Fuelling severe malarial diseases. *Wiley Online Library*, *12*, 1036–1045.
- Slade, R. W., & McCallum, H. I. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*, *132*, 861–862.
- Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, *2*, 16–34.
- Sottile, M. L., & Nadin, S. B. (2018). Heat shock proteins and DNA repair mechanisms: An updated overview. *Cell Stress and Chaperones*, *23*, 303–315.
- Spurgin, L. G., Illera, J. C., Jorgensen, T. H., Dawson, D. A., & Richardson, D. S. (2014). Genetic and phenotypic divergence in an Island bird: Isolation by distance, by colonization or by adaptation? *Molecular Ecology*, *23*, 1028–1039.
- Spurgin, L. G., Illera, J. C., Padilla, D. P., & Richardson, D. S. (2012). Biogeographical patterns and co-occurrence of pathogenic infection across Island populations of Berthelot's pipit (*Anthus berthelotii*). *Oecologia*, *168*, 691–701.
- Spurgin, L. G., & Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 979–988.
- Srivastava, P., Puri, S. K., Dutta, G. P., & Pandey, V. C. (1992). Status of oxidative stress and antioxidant defences during *Plasmodium knowlesi* infection and chloroquine treatment in *Macaca mulatta*. *International Journal for Parasitology*, *22*, 243–245.
- Stewart, M. D., Ritterhoff, T., Klevit, R. E., & Brzovic, P. S. (2016). E2 enzymes: More than just middle men. *Cell Research*, *26*, 423–440.
- Sundaramoorthy, E., Ryan, A. P., Fulzele, A., Leonard, M., Daugherty, M. D., & Bennett, E. J. (2021). Ribosome quality control activity potentiates vaccinia virus protein synthesis during infection. *Journal of Cell Science*, *134*, 1–12.
- Szabó, C. (2003). Multiple pathways of peroxynitrite cytotoxicity. *Toxicology Letters*, *140–141*, 105–112.
- Teoh, M. L. T., Turner, P. V., & Evans, D. H. (2005). Tumorigenic poxviruses up-regulate intracellular superoxide to inhibit apoptosis and promote cell proliferation. *Journal of Virology*, *79*, 5799–5811.
- Teoh, M. L. T., Walasek, P. J., & Evans, D. H. (2003). Leporipoxvirus Cu,Zn-superoxide dismutase (SOD) homologs are catalytically inert decoy proteins that bind copper chaperone for SOD. *The Journal of Biological Chemistry*, *278*, 33175–33184.
- Tschirren, B., Råberg, L., & Westerdahl, H. (2011). Signatures of selection acting on the innate immunity gene toll-like receptor 2 (TLR2) during the evolutionary history of rodents. *Journal of Evolutionary Biology*, *24*, 1232–1240.
- Valkiunas, G. (2005). *Avian malaria parasites and other Haemosporidia*. CRC Press.
- van de Crommenacker, J., Richardson, D. S., Koltz, A. M., 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: Biological Sciences*, *279*, 1466–1476.
- Vasquez, M., Zuniga, M., & Rodriguez, A. (2021). Oxidative stress and pathogenesis in malaria. *Frontiers in Cellular and Infection Microbiology*, *11*, 768182.
- Videvall, E., Cornwallis, C. K., Palinauskas, V., Valkiunas, G., & Hellgren, O. (2015). The avian transcriptome response to malaria infection. *Molecular Biology and Evolution*, *32*, 1255–1267.
- Videvall, E., Palinauskas, V., Valkiunas, G., & Hellgren, O. (2020). Host transcriptional responses to high- and low-virulent avian malaria parasites. *American Naturalist*, *195*, 1079–1084.

- Voelker, G. (1999). Dispersal, vicariance, and clocks: Historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus: Motacillidae*). *Evolution*, *53*, 1536–1552.
- Waldenström, J., Bensch, S., Hasselquist, D., & Östman, Ö. (2004). A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *Journal of Parasitology*, *90*, 191–194.
- Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Künstner, A., Searle, S., White, S., Vilella, A. J., Fairley, S., Heger, A., Kong, L., Ponting, C. P., Jarvis, E. D., Mello, C. V., Minx, P., Lovell, P., Velho, T. A. F., Ferris, M., ... Wilson, R. K. (2010). The genome of a songbird. *Nature*, *464*, 757–762.
- Yorinks, N., Atkinson, T., & Carter, T. (2000). Effects of malaria on activity budgets of experimentally infected juvenile apapane (*Himatione sanguinea*). *The Auk*, *117*, 731–738.
- Zhou, Y., Gao, F., Lv, L., Wang, S., He, W., Lan, Y., Li, Z., Lu, H., Song, D., Guan, J., & Zhao, K. (2021). Host factor cyclophilin B affects Orf virus replication by interacting with viral ORF058 protein. *Veterinary Microbiology*, *258*, 109099.
- Zueva, K. J., Lumme, J., Veselov, A. E., Kent, M. P., Lien, S., & Primmer, C. R. (2014). Footprints of directional selection in wild Atlantic salmon populations: Evidence for parasite-driven evolution? *PLoS One*, *9*, 91672.
- Zueva, K. J., Lumme, J., Veselov, A. E., Kent, M. P., & Primmer, C. R. (2018). Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (*Salmo salar*). *Marine Genomics*, *39*, 26–38.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sheppard, E. C., Martin, C. A., Armstrong, C., González-Quevedo, C., Illera, J. C., Suh, A., Spurgin, L. G., & Richardson, D. S. (2024). Genotype–environment associations reveal genes potentially linked to avian malaria infection in populations of an endemic island bird. *Molecular Ecology*, *33*, e17329. <https://doi.org/10.1111/mec.17329>