DR. LOREN CASSIN-SACKETT (Orcid ID : 0000-0002-6000-4789)

Article type : Original Article

Parallel evolution of gene classes, but not genes: evidence from Hawai'ian honeycreeper populations exposed to avian malaria

Loren Cassin-Sackett*[1,2], Taylor E. Callicrate[1,3], Robert C. Fleischer[1]

1 – Center for Conservation Genomics, Smithsonian Conservation Biology Institute, National Zoological Park, Washington DC 20008

2 – Current Address: Department of Integrative Biology, University of South Florida, Tampa, FL 33620

3 – Current Address: Species Conservation Toolkit Initiative, Department of Conservation Science, Chicago Zoological Society, Brookfield IL

*Corresponding author: cassin.sackett@gmail.com

Key words: adaptation, disease tolerance, population genomics, 'amakihi, rapid evolution

Running title: Genetic variation mirrors malaria extent

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14891

bstract

Adaptation in nature is ubiquitous, yet characterizing its genomic basis is difficult because population demographics cause correlations with non-adaptive loci. Introduction events provide opportunities to observe adaptation over known spatial and temporal scales, facilitating the identification of genes involved in adaptation. The pathogen causing avian malaria, *Plasmodium relictum*, was introduced to Hawai'i in the 1930s and elicited extinctions and precipitous population declines in native honeycreepers. After a sharp initial population decline, the Hawai'i 'amakihi (*Chlorodrepanis virens*) has evolved tolerance to the parasite at low elevations where *P. relictum* exists, and can sustain infection without major fitness consequences. High-elevation, unexposed populations of 'amakihi display little to no tolerance. To explore the genomic basis of adaptation to *P. relictum* in low-elevation 'amakihi, we genotyped 125 'amakihi from the island of Hawai'i via hybridization capture to 40,000 oligonucleotide baits containing SNPs, and used the reference 'amakihi genome to identify genes potentially under selection from malaria. We tested for outlier loci between low- and high-elevation population pairs and identified loci with signatures of selection within low-elevation populations. In some cases, genes commonly involved in the immune response (e.g., major histocompatibility complex) were associated with malaria presence in the population. We also detected several novel candidate loci that may be implicated in surviving malaria infection (e.g., beta defensin, glycoproteins, and interleukin-related genes). Our results suggest that rapid adaptation to pathogens may occur through changes in different immune genes, but in the same classes of genes, across populations.

Introduction

Introduced pathogens have inflicted devastating consequences on wildlife around the globe, leading to extensive population declines and extinctions (e.g., Biggins & Schroeder 1987, Thorne & Williams 1988, Cunningham & Daszak 1998, Cully & Williams 2001, Daszak et al.

2003, Frick et al. 2015). Despite widespread extinctions, some naïve species have evolved resistance (Best & Kerr 2000, Bonneaud *et al.* 2011, Rocke et al. 2012) or tolerance (Atkinson et al. 2013) to introduced diseases over very short time scales (e.g., <50 generations; Decaestecker *et al.* 2007). Recent adaptation to introduced diseases thus provides a useful model for the genomics of rapid adaptation to novel conditions (Epstein et al. 2016). Moreover, explaining the genetic drivers of disease resistance and tolerance is central to understanding disease dynamics (Langwig et al. 2017), the evolution of virulence (Mackinnon & Read 2004; Barclay *et al.* 2012, 2014), coevolutionary dynamics (Lively & Apanius 1995) and the prediction of outbreaks (Acevedo-Whitehouse et al. 2005).

Often, studies of adaptation in nature are limited by the long time scales over which adaptation occurs and the multiple sources of selective pressure on populations, making it difficult to associate particular alleles with a focal source of selection (Lewontin & Krakauer 1973; Lotterhos & Whitlock 2014). Recent biological invasions overcome these obstacles (Lucek et al. 2014; Colautti & Lau 2015), and introduced pathogens can exert strong selection and therefore elicit rapid and detectable adaptation (Burdon & Thompson 1995; Hochachka & Dhondt 2000). As a result, the signatures of selection on host genes involved in recent adaptation are expected to differ from the rest of the genome (e.g., selective sweeps will fix large regions of the genome that have not been disassociated by recombination) (Yeaman & Whitlock 2011). In turn, the geographic distribution of alleles responding to selection by pathogens should mirror the distribution of those pathogens (Haldane 1948; Lewontin & Krakauer 1973; Fumagalli et al. 2011).

In many systems, parallel genetic changes have occurred during convergent phenotypic adaptation (Jones *et al.* 2012; Kooyers and Olsen 2012; Keller et al. 2013), but in others, convergent adaptation has arisen through divergent genomic mechanisms (Hoekstra and Nachman 2003; Rosenblum et al. 2010; Roda *et al.* 2013). In yet other cases, independent genomic adaptations across populations have been replaced by subsequent gene flow (Caprio

and Tabashnik 1992). We still lack a complete understanding of when molecular adaptation should be parallel versus divergent, although some patterns are beginning to emerge (Rosenblum et al. 2014). For phenotypes with multiple physiological or molecular mechanisms, such as tolerance to a pathogen, there are likely multiple genomic solutions that confer adaptation (Pfeifer et al. 2018). In such cases, the first variant to appear in a population should rapidly increase in frequency, independent of the alleles in other populations, leading to differences among populations in the genomic basis of convergent phenotypes. Alternatively, if only one gene can produce an adaptive phenotype, parallel genetic changes would be predicted to underlie convergent adaptation (Colosimo et al. 2005; Chan et al. 2010). Similarly, if gene flow was sufficiently high during initial adaptation, the first variant to confer tolerance should spread to other populations. These two latter scenarios can be distinguished by signatures of different mutations in the same gene (parallel adaptation) versus identical SNPs across all populations (homogenizing effect of gene flow). The genomic basis for convergent evolution can be studied in nature via replicated evolutionary experiments such as species introductions.

Avian malaria, caused by the haemosporidian parasite *Plasmodium relictum*, has been introduced globally (Beadell *et al.* 2006). The parasite causes fitness declines even in asymptomatic hosts (Asghar et al. 2015), and has resulted in vast population declines and extinctions in naïve species (Atkinson & LaPointe 2009). Especially susceptible are the Hawai'ian honeycreepers, an adaptive radiation of at least 55 species (James & Olson 1991; Fleischer et al. 1998, Fleischer & McIntosh 2001) that diversified from Eurasian rosefinches after colonization of the Hawai'ian Islands 5.8 - 7.2 million years ago (Lerner et al. 2011). Honeycreepers are an emblem of the negative consequences of species introduction: As a result of anthropogenic threats, at least 17 species have gone extinct since the arrival of Europeans to the islands (van Riper 1986, Atkinson & LaPointe 2009). At least two introductions of a mosquito vector of avian malaria, *Culex quinquefasciatus*, occurred by the 1930s (Fonseca et al. 2006). Since then, avian malaria has decimated the remaining honeycreepers, contributing to at least 7 extinctions (van Riper III et al. 1986) and population declines in every surviving species

(Atkinson & LaPointe 2009, Paxton *et al.* 2016). Most species have been forced into high-elevation refugia where disease transmission is reduced due to temperature limitations on mosquito larval development, adult mosquito feeding rate, and *Plasmodium* development within the mosquito (van Riper et al. 1986; LaPointe et al. 2010; Samuel et al. 2015). Mosquitoes are currently found at elevations up to 1650m (Goff & van Riper 1980; van Riper *et al.* 1986), but as the climate is predicted to warm on Hawai'i, the existence of susceptible honeycreeper species is in danger (Fortini et al. 2015, Paxton et al. 2016).

Despite the catastrophic consequences of avian malaria in many species, populations of a few native species have begun to recover. In the last several decades, populations of Hawai'i 'amakihi (*Chlorodrepanis virens*) have expanded in size and distribution in low elevation forests (Foster et al. 2007; Eggert et al. 2008) despite high prevalence of *P. relictum* in both mosquitoes (Woodworth et al. 2005) and 'amakihi (Woodworth et al. 2005, Kilpatrick et al. 2006). Challenge experiments have demonstrated higher survivorship of low-elevation 'amakihi—which have evolved in the recent presence of avian malaria—than high-elevation 'amakihi, which are naïve to the disease (Atkinson et al. 2013), and surviving individuals are immune to reinfection (Atkinson et al. 2001). Juvenile 'amakihi disperse farther within the same elevation than up or down in elevation (Lindsey et al. 1998); thus, there is limited gene flow across elevations (Foster et al. 2007; Eggert et al. 2008). As a result, alleles conferring tolerance to avian malaria are likely to spread within low-elevation populations more rapidly than they expand upwards to high-elevation populations, presenting an opportunity to identify genes involved in adaptation.

In this study, we use samples that were collected from 1987–2005 to implement comparative genomics across low, mid and high elevation populations of Hawai'i 'amakihi, analyzing patterns of genomic diversity within and among populations to infer the response to malaria-induced selection. We aim to: 1) identify signatures of selection to determine the genes associated with malaria tolerance in low-elevation populations, 2) evaluate the degree to which the same candidate genes are implicated in adaptation among populations, and 3) assess

whether candidate genes are more often involved in known immune function versus other cellular processes that were coopted to defend against malaria infection.

Materials and Methods

Sampling Design and Genomic Library Preparation

From 1987 – 2005, ‘amakihi blood or tissue was sampled from 21 sites along elevational gradients on the island of Hawai’i (Fig. 1; Tarr & Fleischer 1993; Foster et al. 2007, Eggert et al. 2008). For this study, we used a random subset of samples from each site. Sampling protocols have been described in detail elsewhere (Tarr & Fleischer 1993; Woodworth et al. 2005). All sampling occurred in concordance with IACUC approvals. DNA was extracted from blood using a Qiagen DNeasy Blood and Tissue Kit following the manufacturer’s protocol, or using phenol:chloroform (Tarr & Fleischer 1993).

Using the ‘amakihi genome (Callicrate et al. 2014) as a reference, we designed a custom 40,000 bait in-solution hybridization assay containing baits with SNPs distributed randomly throughout the genome (at least 80 bp apart); the baits were generated by MYcroarray (now Sage Science, Ann Arbor, MI). On average, this design resulted in a bait every 275,000 bp. The genome is available in NCBI's BioProject repository (Accession PRJNA252695), and the baits have been placed in a GitHub repository (https://github.com/CassinSackett/SNPcapture/). For each individual, 1-2 μg DNA in 25 μL nuclease-free water was first sheared in a Q800R sonicator (QSonica LLC, Newton CT) for 4.5 – 6.5 minutes, depending on initial sample quality, to a target size of 400 bp. A customized Nextera-style library preparation was then performed, tagging each individual with a unique combination of two barcoded Illumina primers. To maximize hybridization efficiency, equal amounts of DNA from 8 individuals were pooled to hybridize with the baits. After hybridization for 48 hours, all pools were combined in equimolar ratios and sequenced with Nextera-style adapters on either an Illumina MiSeq or HiSeq paired end, 150bp

read run. We genotyped 121 birds (53 low-elevation, 8 mid-elevation, 52 high-elevation, and 8 from a captive family group to validate Mendelian inheritance in each SNP) at the 40,000 target loci as well as ~160,000 off-target loci recovered as bycatch during hybridization and amplification. Off-target loci were retained in analyses to increase the proportion of the genome represented (beyond the ~6 million base pairs afforded by 150bp reads).

Sequence Processing and SNP Detection

Populations were classified as low- or high-elevation based on the historical (1900s) occurrence of mosquitoes at the sampled elevation, a proxy for malaria exposure (Keyghobadi et al. 2006, LaPointe et al. 2010). The mid-elevation population is characterized by seasonal malaria transmission; these designations are consistent with the reliance of mosquitoes on temperature and with previous classification of populations (e.g., van Riper et al. 1986, Woodworth et al. 2005, Foster et al. 2007, Eggert et al. 2005). We aligned quality filtered reads to the 'amakihi reference genome (Callicrate et al. 2014) and performed additional quality filtering steps in GATK and vcftools (see Supplementary Materials for details). To maximize the number of SNPs in each downstream analysis, we generated separate files with different subsets of SNPs for (1) the entire spatial dataset after filtering (N=118 birds), (2) all individuals within an elevation after filtering (N=48 low and N=49 high), (3) each low-high elevation population pair after filtering, and (4) each population separately after filtering (single population numbers in Table S1). SNPs that were fixed within populations or within population pairs were removed in the appropriate subset, but retained in other subsets. In each subset, we used vcftools to filter the dataset for missingness through an iterative process of removing individuals and loci designed to maximize both number of individuals and SNPs retained in the final datasets (Supplementary Materials). This filtered dataset resulted in a large number of high-quality SNPs in each subset (mean = 157,546 SNPs, Tables S1-S2). VCF files were converted to other formats for downstream analyses using PGDSpider2 (Lischer and Excoffier 2012). A

detailed description of the pipeline and scripts are available on the lead author's (LCS) GitHub website (https://github.com/cassinsackett/SNP_capture).

Genetic Diversity and Population Structure

We grouped individuals into populations based on sampling locality, prior literature, personal knowledge and correspondence with local ornithologists (Fig. 1; Lindsey *et al.* 1998; Foster et al. 2007; Eggert et al. 2008; E.H. Paxton, personal communication). To understand the genetic background under which recent evolution may have occurred, we estimated genetic diversity and population structure using several methods (Fig. S1). First, we estimated mean heterozygosity within individuals, and performed a two-sample t-test to determine whether heterozygosity was different among low- and high-elevation individuals. We used vcftools to estimate departures from Hardy-Weinberg equilibrium and nucleotide diversity (pi; Nei and Li, 1979) within populations and pairwise F_{ST} between all population pairs (Supporting Information). We performed t-tests in R (The R Foundation for Statistical Computing, http://www.r-project.org/) to determine whether nucleotide diversity differed between elevations and whether this differentiation was greater than within-elevation population pairs. Next, we estimated the number of effective migrants between low- and high-elevation populations using Genepop V4.3 (Rousset 2008). Finally, we performed a Principal Component Analysis (PCA) with the ade4 package (Chessel et al., 2004; Dray and Dufour, 2007) in R on a dataset containing no missing data (N=2816 SNPs). We used the 'bca' function to test whether individual genotypes were differentiated by elevation across all principal components; statistical significance was assessed by a randomization test with 10,000 iterations.

Inference of Selection

We searched for genomic signatures of selection in several ways, in order to minimize false positives arising from a single analysis. First, we performed outlier tests between all pairs of low- and high-elevation populations (N=9 pairs; 3 high-elevation pairings for each of 3 low-elevation populations) in BayeScan 2.1 (Foll and Gaggiotti 2008) using a false discovery rate threshold of 0.1. Second, because few outliers were detected, even under different model parameters (e.g., prior odds 1 – 1000, FDR=0.2), we subsequently designated as 'quasi-outliers' the 0.1% of SNPs with the highest F_{ST} (calculated in vcftools) between low-elevation and high-elevation population pairs, and between all low-elevation individuals and all high-elevation individuals. A threshold of 1% most-differentiated loci is commonly used (e.g., Love *et al.* 2016; Stankowski *et al.* 2016), but our aims were to minimize false positives and to be conservative given the potential effects of genetic drift on differentiation among small populations (and the structure imposed by low dispersal across elevations). Additionally, performing multiple comparisons (e.g., 9 pairwise comparisons among 6 populations) typically requires lowering the critical threshold for significance to minimize false positives. For the outliers and quasi-outliers, we extracted flanking sequence from the ‘amakihi genome around these SNPs (Callicrate et al. 2014), and inferred gene identity using the blastn option of the Basic Local Alignment Search Tool (BLAST, Altschul et al. 1990) and the NCBI nucleotide database (Zhang et al. 2000). To avoid false inference of matching genes that may occur with longer queries, we used the shortest query size for each SNP (at least 300bp on each side of the SNP) that produced significant blast hits. SNPs that returned no hits with the smallest query size were queried with larger flanking regions in a stepwise manner (e.g., 500bp, 1kb, 2kb) up to 10kb on each side of the SNP.

Next, we examined Long Runs of Homozygosity (Auton et al. 2009) using vcftools. Although this method is typically used to infer inbreeding, regions that are homozygous within multiple outbred individuals could be indicative of selection. Therefore, we searched for regions with a high probability of autozygosity in multiple individuals within each low-elevation

population. We examined the resulting regions for overlap with the quasi-outliers in each population. Finally, we calculated Tajima's D in 1kb windows within each low-elevation population using vcftools. Largely negative Tajima's D can be indicative of recent selective sweeps; therefore, we focused on regions with the 25 most negative Tajima's D values in low-elevation populations for follow-up analyses, along with SNPs deemed to be statistical outliers and 'quasi-outliers' (the 0.1% most-differentiated loci) between low and high-elevations. A summary of these methods is presented in Fig. S1.

Gene Ontology Overrepresentation

Rapid adaptation to pathogens may leverage existing immune processes, or it may co-opt unrelated pathways (e.g., the sickling of red blood cells that reduces malaria infection in humans). In order to determine whether inferred loci under selection were more commonly involved in immune function relative to other processes than expected by chance (i.e., relative to the proportion of immune genes versus other genes in the genome), we compared the genes under inferred selection in 'amakihi to the number of known genes for a given functional class in the chicken genome (the most well-annotated avian genome available). To do so, we used Panther (Mi *et al.* 2016) to assess gene ontology on the blast results (up to 3 hits per locus) for multiple analyses: the inferred outliers from BayeScan; the 0.1% of loci with the highest F_{ST} between low- and high-elevation individuals; and the 25 loci with the most negative Tajima's D statistic within low-elevation populations. We pooled the gene lists from the blast results and reduced the list to unique gene entries. Next, we compared the gene lists in 'amakihi with a reference set of genes comprising all genes in the *Gallus gallus* genome (N=15,782 known genes), which we assume equates to all genes in the 'amakihi genome, and performed an overrepresentation test to determine which genes appeared more or less often in the candidate gene set than expected by chance, based on the number of genes in each category. We used the Panther hierarchical classification system for biological processes and for molecular function

(Mi *et al.* 2013). The chicken reference genes were categorized into 247 biological processes and 184 molecular functions, and the number of expected 'amakihi genes in each category was generated based on the total number of unique genes resulting from the blast analyses (N=447 genes). Significance testing included a Bonferroni correction for multiple testing.

Results

SNP statistics

For the entire dataset (all individuals), we recovered 399,197 SNPs passing quality filters, including 359,197 off-target loci. The resulting SNP density in our dataset was ~4 quality-filtered SNPs every 10,000 base pairs. Approximately 8% of SNPs did not follow Mendelian expectations and were discarded. Our final datasets contained the SNPs genotyped in at least 80% of individuals (Table S1).

Genetic Diversity and Population Structure

Individual heterozygosity ranged from 0.0046 – 0.079, and was significantly higher among low-elevation individuals (mean = 0.0512, sd = 0.02) than high-elevation individuals (mean = 0.0417, sd = 0.02, p=0.025, df=93 [t-test]; Table S1, Fig. S3). Heterozygosity was lowest in high-elevation Pu'u La'au (mean H_o = 0.016, sd = 0.009) and highest in low-elevation Hualālai (mean H_o = 0.068, sd = 0.01, Table S1). Within populations, there was a larger proportion of sites that exhibited a significant ($p<0.01$) deficit of heterozygotes (0.06% of SNPs in Hualālai, 0.56% of SNPs in Manukā, and 5.0% of SNPs in Pāhoa) than an excess of heterozygotes (0 - 0.05% per population). Nucleotide diversity (pi) was approximately equal across populations (Table S2); diversity was not significantly different in low- and high-elevation populations ($p > 0.7$).

Average genome-wide differentiation between pooled low- and high-elevation individuals was low but significant (Weir and Cockerham weighted F_{ST} = 0.0065). Pairwise differentiation between populations averaged 0.0124 for low-low population pairs, 0.0182 for high-high population pairs, and 0.0174 for low-high population pairs (Table S2; Fig. 2, Fig. S6). Differentiation was not significantly higher between low-high population pairs than within-elevation pairs ($p > 0.3$). There were no fixed differences between low- and high-elevation populations when pooling all individuals within an elevation. The estimated number of migrants was not significantly different between elevations (m = 2.33) than within elevations ($m_{low\text{-}low}$ = 2.46, $m_{high\text{-}high}$ = 2.18; Table S3). Optimization of the SNP processing pipeline (i.e., performing each set of analyses with different subsets of loci and individuals) demonstrated that several analyses were demonstrably influenced by the amount of missing data. Estimates of F_{ST} decreased and the effective number of migrants (N_m) increased with less missing data.

Principal Component Analysis indicated partial genetic overlap between elevations, with segregation between low- and high-elevations (Fig. 3, Fig. S5). The first two components explained 1.9 and 1.8% of the genetic variance among individuals. The randomization test for the between-class (i.e., between-elevation) analysis along all components demonstrated significant differentiation between low- and high-elevation individuals ($p = 0.0002$).

Loci Under Potential Selection

In each of nine low-high population pairs, BayeScan inferred 0 – 6 outliers at a FDR threshold of 0.1 (Table S4). Of the 10 total statistical outliers, six had blast hits with no inferred relationship to malaria infection (Table S4), while four potentially served a role in surviving infection. First, a region on chromosome 2 blasted to Cytoplasmic FMR1 interacting protein 2, a gene with functions in T-cell adhesion. Second, a region on chromosome 5 blasted to Toll-like receptor 5, a gene with known immune function. This gene was also inferred by Tajima's D to be under selection in Hualālai (Table 1). Third, a region on chromosome 10 blasted to both

Attractin and beta-defensin (Table 1); both genes function in the immune response. Beta-defensin was also a match to several quasi-outlier SNPs. Finally, a region on chromosome 12 blasted to Contactin-3, a gene encoding an immunoglobulin protein that mediates cell-surface interactions.

Low-high elevation population pairs contained from 61 to 138 quasi-outlier SNPs each (878 total, Fig. 2). Of these, 56 SNPs were found in two (N=52 SNPs) or three (N=4 SNPs) population pairs. In all but three of these instances, at least one population was common to both pairs, suggesting that in most cases, a mutation occurred once in a single population. In addition, 145 SNPs were situated within 10kb of a SNP in another population pair—79 SNPs were within 200bp of another—which could indicate different mutations across populations in the same gene region (including regulatory elements). Although the remaining quasi-outlier sites differed among populations, there were only 9 private quasi-outlier alleles within low-elevation populations, indicating primarily shared variation among low-elevation sites.

Of the 818 unique quasi-outliers, 271 blasted to multiple loci that included repetitive DNA (e.g., LINEs, microsatellites, or repeat domains in known genes). These were inferred to be either non-coding DNA that drifted or hitchhiked with other mutations or repetitive regions common to many proteins. An average of 56 quasi-outlier SNPs per population pair were localized in repeat regions detected by RepeatMasker. However, 37 of these 271 loci blasted to beta defensin as one of the top three hits in addition to other matching sequences. Seven quasi-outlier SNPs from one population pair were localized to the mitochondrial cytochrome b gene.

The remaining 521 loci (SNP + flanking region) blasted to identified genes, of which 76 had putative immune-related functions (in addition to the 37 potential beta defensin SNPs) (Table 1, Table S6; E values in Supplementary Files). Of the 56 quasi-outlier SNPs appearing in multiple population pairs, 16 were in repeat regions and 9 were in immune-related regions (Table 1). The 64 regions containing multiple SNPs in close proximity followed a similar pattern: 21 blasted to repeat regions, 3 to UCEs, and 11 to immune-related genes

(Supplementary Materials). Quasi-outlier sites exhibited larger heterozygote deficits than the total set of filtered SNPs by an order of 3.9 – 10.2 (0.6% in Hualālai, 5.5% in Manukā, and 19.8% in Pāhoa). Heterozygote excess within quasi-outliers was negligible (1 site in Hualālai and 1 site in Pāhoa).

The number of long runs of homozygosity (LROH) longer than 3Mb on a chromosome was proportional to the number of baits on that chromosome, suggesting higher density of SNPs results in higher power to detect LROH. In low elevation individuals, 4122 sites were inferred to begin LROH. Although most of these occurred in inbred individuals, 172 sites were located within 10kb of a quasi-outlier SNP; this number of LROH corresponding to quasi-outliers is likely an underestimate because 1838 runs were >10kb. Of these 172 sites, 37 occurred in repeat regions and 26 blasted to genes related to immune function (Tables 1, S6). In some cases, the LROH was inferred in a different population than either population in the corresponding quasi-outlier site.

The distribution of Tajima's D was leptokurtic and centered just below zero (Table S1, Fig. S7) in all but one population from all elevations, consistent with widespread recent population expansions (Foster et al. 2007). The exception was Pu'u La'au (Tajima's D=0.0003), a high-elevation population. Of the 25 1kb regions with the largest negative Tajima's D values in low-elevation populations (indicating possible purifying selection), twelve blasted to genes related to immune function (Tables 1, S6). In addition, three loci related to blood function were characterized by largely negative Tajima's D (Tables 1, S5).

Consistency Across Populations

Of the 5 population pairs with outlier SNPs, none shared outliers, and none of the 10 outliers were within 5 million base pairs of an outlier in another population pair. However, two outliers were quasi-outliers in another population pair; one of these was also situated within

20bp of 3 additional quasi-outliers in the second population pair. In two outliers, all low-elevation populations shared a predominant allele that differed from high-elevation populations, whereas five outliers were marked by only one population with a different predominant allele than the other five populations. There was more consistency among populations in quasi-outliers. Of the 818 quasi-outlier SNPs, 71 contained genotypes passing quality filters in only one low-elevation population. Among the remaining 747 SNPs, 56 were quasi-outliers in multiple populations, and 145 SNPs (in 64 genomic regions) were found within 10kb of a quasi-outlier SNP in the same (70%) or another (30%) population. In addition, 242 (30%) quasi-outlier SNPs were among the top 1% most-differentiated SNPs in another population pair. Allele frequencies at quasi-outlier loci were correlated in low-elevation populations (Spearman's rho = 0.522 – 0.607, all $p << 0.001$), suggesting that loci that were deemed quasi-outliers in only a single population had similar allele frequencies but did not attain the statistical threshold set for quasi-outliers in other populations (Table S7). Only nine quasi-outliers contained alleles that were private to low elevation (four occurring in only one population), and five contained private alleles at high elevation (all occurring in only one population). The remaining SNPs differed in allele frequency across elevation, but alleles were present at all elevations (although not necessarily in all populations). Of the sites beginning LROH, 4.2% were located in close proximity to quasi-outlier SNPs. With specific immune-related genes resulting from blast searches, there was a moderate level of sharing among low-elevation populations (Fig. 4), and when genes were combined into similar classes (e.g., all glycoproteins, all toll-like receptors, all interleukins), the proportion of gene classes shared among populations was high.

Gene Overrepresentation

Of the candidate loci from combined analyses, there were 447 unique genes from the resulting blast searches. Gene ontology categories in Panther were hierarchically classified into 247 biological processes. Of these, there were 25 processes that were under- or over-represented ($p<0.10$, Fig. 5). Six of these were both significantly overrepresented and displayed three-fold enrichment or greater in ‘amakihi with at least two genes recovered (i.e., gene ontology categories appeared more often in ‘amakihi than expected, relative to the number of genes of those categories in the genome, with gene number inferred from the *Gallus gallus* reference). Immune system processes were enriched 1.5-fold. In addition, three biological processes related to mRNA processing and RNA splicing were characterized by significantly lower-fold enrichment (less than one third). Two immune-related categories contained no genes from the candidate loci list: defense response to bacterium and antigen processing and presentation; however, due to the small number of genes in these categories in the chicken genome (<30 out of 15,782), this underrepresentation was not significant.

When genes were categorized according to molecular function, ten functions containing at least two ‘amakihi genes were enriched at least three-fold ($p<0.10$, Supplementary Results). Two additional immune-related functional categories (interferon receptor binding and transforming growth factor beta-activated receptor activity) were enriched threefold but contained only one ‘amakihi gene, so the enrichment was not significant ($p>0.1$). The lack of significance derives from the small number of genes in these respective categories, yet this enrichment may be biologically relevant. No molecular functions were significantly under-represented by the same amount (one third).

Discussion

We use introduced avian malaria as a model of rapid evolution to a novel selection pressure, presenting evidence from multiple approaches that malaria-exposed, low-elevation ‘amakihi possess a suite of genomic differences from high-elevation malaria-naïve ‘amakihi in immune-related genes. In addition, our data suggest that genetic variation of adaptive significance can be maintained despite strong population bottlenecks. Multiple loci displayed signatures of selection and/or unusually high differentiation between malaria-naïve and exposed populations. Mean Tajima's D was below zero in all low- and mid-elevation populations, consistent with widespread recent population expansions following malaria-induced bottlenecks (Foster et al. 2007), a hypothesis that follows from the field observation of an increase in survivorship following malaria infection in recent decades. Genetic diversity was not lower in low-elevation populations, lending support to the idea that malaria survival in low-elevation ‘amakihi has facilitated the maintenance of genetic variation.

Immune genes under potential selection

Among the loci that differed most strongly between low- and high-elevation, a subset were related to pathogen defense and immune response. Some predictable loci were inferred to be under selection in localities with higher rates of malaria transmission, such as the major histocompatibility complex (MHC). However, several different regions of the ‘amakihi genome blasted to the MHC region, suggesting the occurrence of false positive associations resulting from their extensive study (leading to overrepresentation on Genbank) or from gene duplication. Nonetheless, there is evidence that the MHC may play a role in the malaria response. Although MHC is typically subject to balancing selection, particular alleles may be associated with lower incidence of haemosporidian infection (Jones et al. 2015); these alleles would be expected to exert selection on the parasite and decrease in frequency after the parasite adapts. This finding is consistent with the pattern of spatial variation in host MHC

diversity mirroring disease prevalence in some passerines (Jones et al. 2014, 2015; S. Jarvi personal communication). Moreover, host MHC variation has been linked to disease outcome in a variety of vertebrate infectious diseases (e.g., Tarleton et al. 1996; Kaufman 2000; Grimholt et al. 2003; Turner et al. 2008; Savage et al. 2011; Hawley & Fleischer 2012), and avian malaria in passerines in particular (Westerdahl *et al*. 2005; Bonneaud *et al*. 2006). Therefore, the fact that our genome-wide SNP assay recovered MHC as associated with malaria-induced selection is consistent with other patterns of selection on MHC observed in honeycreepers (Jarvi et al. 2004, 2016).

Our dataset also revealed several candidates for adaptation to malaria that are novel to this system but have been documented in other malarial systems. In particular, other infection- and immune-related genes (e.g., toll-like receptors [Coban et al. 2005; Mockenhaupt et al. 2006; Franklin et al. 2009]; interferons and tumor necrosis factors [Grau et al. 1989; De Souza et al. 1997; Franklin et al. 2009]) were invoked in our comparisons as well as other studies of malaria in humans and mice. Interestingly, although beta-defensins are known to play a role in infection and to be under selection in some mammals (Semple et al. 2003; Ganz 2003; van Dijk et al. 2008), they have not, to our knowledge, been linked to malaria response in vertebrates. However, expression of beta-defensin increased after infection with *P. berghei* in *Anopheles* mosquitoes (Richman et al. 1997), and other defensins are toxic to *P. gallinaceum* in *Aedes* mosquitoes (Shahabuddin et al. 1998). In addition, the literature is equivocal on the role of some genes, such as CD1, in combating infection with *Plasmodium* (Schofield et al. 1999; Molano et al. 2000). For non-model species, the utility of the genome-wide approach lies in its identification of previously unknown candidate genes for their putative roles in specific infections.

In some cases, there appears to be a nuanced relationship between malaria and genotype that we did not have the genomic resolution to test here. Immunoglobulins, for example, are related to *Plasmodium* infection in birds (Atkinson et al. 2001) and malaria severity in humans, but the direction of severity changes for different immunoglobulin isotypes and subclasses (Perlmann et al. 1997; Aucan et al. 2000). The loci inferred to be under selection in this study were not identified as a particular type of immunoglobulin, so it is unclear the exact role they may play; follow-up studies using immunoglobulin profiling (Turchaninova et al. 2016) rather than DNA sequencing may prove fruitful. Similarly, certain types of glycoproteins play a protective (Friedman 1983; Ockenhouse et al. 1989; Jakobsen et al. 1994) or facilitative (Egan et al. 2015) role in human malaria, but it is not clear whether distinct avian glycoproteins have an analogous function in defense.

We did not recover a signal of selection in other specific genes known to influence infection with or survivorship from malaria (caused primarily by *P. falciparum*) in humans (e.g., CD55 [Egan et al. 2015], sickled erythrocyte cells [Friedman 1978], G6PD [Hedrick 2011], etc.). However, the pooled low-high elevation dataset and two low-high elevation population pairs were highly differentiated at a genomic region bearing similarity to the CD59 glycoprotein, and this region exhibited highly negative Tajima's D in one low-elevation population (Manuka). The CD59 glycoprotein gene inhibits the membrane attack complex, as does CD55. In addition, tumor necrosis factor and interleukin-10 levels--genes recovered in this study--were associated with malaria severity in humans (Kurtzhals et al. 1998; Othoro et al. 1999) and mice (Kossodo 1997; Li et al. 1999). These genes thus represent exciting candidates for future research on avian malaria.

Several of the genes inferred under selection are involved in the innate immune system and function in the inflammatory response; some of these interact with components of the adaptive immune system. For instance, six candidate genes (beta defensins, CD59, interferon receptor 2, lymphocyte antigen 6E, MHC, and transforming growth factor β) regulate or interact

with T cells or increase the expression of antigens. T cells are involved in the adaptive immune system, which is consistent with previous findings that ‘amakihi individuals that survived initial avian malaria infection were immune to later challenges (Atkinson et al. 2001). T cells capable of efficiently recognizing merozoites could slow the invasion and subsequent replication of *P. relictum*, increasing survival of infected ‘amakihi. One gene inferred under selection, an erythrocyte membrane protein, is not involved in the immune response but is coopted by *Plasmodium* to cause aggregations of cells that increase transmission efficiency among red blood cells. Therefore, this gene in ‘amakihi may be under selection to decrease the tendency to form aggregations.

Some genes that were recovered in this study were classified according to their primary function, which indicated they were not part of the immune response. Many of these genes may be differentiated as a result of demographic processes, or adaptive differences related to other environmental factors that vary by elevation (e.g., temperature, hypoxia, or other pathogens such as avian pox (Warner 1968, Atkinson et al. 2005)). However, some genes may also play dual roles in the organism. For instance, genes involved in calcium signaling and transport were among the significantly overrepresented genes in our analyses. Although these genes do not function in the immune system, they may be important for combating malaria infection: cellular calcium levels influence infection and transmission success of multiple species of *Plasmodium* (Huff et al. 1958; Tanabe *et al.* 1982; Scheibel *et al.* 1987), as well as other apicomplexan parasites (Donahue *et al.* 2000). These observations underscore the importance of considering all genes—even those seemingly unrelated to a phenotype—as potential candidates for functional phenotypes. As we continue to discover multifarious roles of genes, annotation will improve such that relevant genes can be detected through gene ontology analyses.

Genes in repetitive regions

A large number of differentiated loci occurred in repetitive regions, which may obscure the true identity of the gene containing the SNP. Many loci produced blast results that represented multiple hits of the same gene in different taxa or clones (e.g., matches to attractin in 20 species). A large number of loci, however, resulted in blast hits on a vast diversity of genes, many of which were or contained repeat regions or transmembrane proteins (Supplementary files). For instance, an erythrocyte membrane protein was the result of one blast search; because many proteins have similar domains, the possibility of an actual match to a membrane protein--either in erythrocytes or another cell type--could explain why so many queries had hits in different proteins: They were matching the similar structure of a membrane protein rather than a specific gene. A related explanation is that many genes have repeat motifs, and these can be difficult to distinguish among genes. Because repeat motifs have higher mutation rates than other parts of the genome, the finding that many of these loci were divergent between populations is not surprising. Alternatively, these regions could represent rapidly evolving pathogen-recognition sites such as cell-surface proteins, which are known to contain repeat domains (Katti *et al.* 2000). Many authors have dealt with repetitive regions by removing them from analyses because they impede our ability to detect genes of interest; however, eliminating them from analysis negates the possibility of finding genes with these motifs that are of actual importance (Zhuang et al. 2012). For instance, genes that are important in adaptation may have repeat motifs that influence protein binding or gene expression (Kashi and King 2006; Gemayel et al. 2012; Prentice et al. 2017). In several instances, we detected genes (primarily beta defensin) of potential importance in this system within repetitive regions.

Conclusions

Genome-wide association studies and other inferential approaches such as outlier tests have tremendous potential to reveal novel candidate genes regulating adaptive processes; however, these approaches also bring with them several limitations. In particular, outlier tests may have high false positive rates resulting from the varied demographic history of different genomic regions (Lotterhos & Whitlock 2014; Whitlock & Lotterhos 2015), or false negatives when there are high levels of background differentiation. In the absence of experimental studies, genes inferred with these approaches should be treated not as conclusive genes involved in malaria protection, but as candidates for further study. Nonetheless, using a combination of methods should minimize the number of false positives and elucidate broad patterns. Here, although false positives can be expected from each analysis in isolation (outliers, F_{ST} quasi-outliers, Tajima's D), a neutral SNP or genomic region is unlikely to be inferred as being under selection in multiple analyses. Therefore, although results should be interpreted as candidates for further study, the most likely candidates for genes conferring tolerance to malaria in low-elevation Hawai'i 'amakihi are genes recovered under multiple approaches (Tables 1, S6).

Our results suggest that the early stage of adaptation to novel strong selection such as introduced malaria may occur via changes in multiple genes that each confer tolerance, only some of which are common across populations. Similar patterns have been found in experimental evolution studies (Notley-McRobb and Ferenci 1999; 2000; Elena and Lenski 2003) and in natural populations (Pfeifer et al. 2018). Parallel evolution is less common when multiple traits confer the same phenotypic function (Thompson et al. 2017); it is likely that over time, some of these changes will replace others as a result of gene flow (Caprio and Tabashnik 1992) and variation in fitness of particular mutations in different environments and genetic backgrounds. Indeed, the observation that few alleles at quasi-outlier loci were private to any low-elevation population is consistent with gene flow distributing adaptive variants among

populations. Alternatively, the degree of shared variation could suggest that standing genetic variation in ancestral ‘amakihi populations contained SNPs that were subject to selection in low-elevation populations.

This work is instructive about the consistency and predictability of evolutionary adaptation. Recent work in another emerging disease system, Tasmanian devil facial tumor disease, demonstrated concordant genomic responses to disease-induced selection across populations (Epstein et al. 2016), whereas we detected a only small number of genes that changed in multiple populations. However, our analyses revealed frequent changes in particular classes of genes (e.g., interleukin related; multiple glycoproteins). The diversity of response to selection in Hawai’i ‘amakihi populations could be due to its larger effective population size or lower gene flow among populations than Tasmanian devils. Our finding that certain classes of genes, but not specific genes, are associated with adaptation to malaria among populations within a species is consistent with patterns in the literature at higher taxonomic levels, including across such divergent taxa as mammals and birds. This supports the idea that selection acts on available variation, which differs among independently evolving populations, but that specific types of host genetic variation are targeted by the coevolutionary arms race in host-pathogen systems.

Acknowledgments

We thank Kamehameha Schools, Hawai’i Volcanoes National Park, and Hawai’i Division of Forestry and Wildlife for permission to sample on their land. We are grateful to Scott Derrickson for providing samples from the family group reared in captivity at SCBI aviaries, and to him and many others (including Cheryl Tarr, Bethany Woodworth, and Carl McIntosh) for help obtaining samples from wild birds. This work was funded by Smithsonian Institution fellowships (to LCS and TEC), a Smithsonian Pell Competitive Grant for Science (to RCF, JK and TEC), an Office of the Undersecretary for Science Next Generation Sequencing Small Grant (to

RCF), a Smithsonian Institution Grand Challenges Grant (to RCF et al.), NSF DEB-1547168, and the Center for Conservation Genomics (CCG). We appreciate ongoing feedback and discussions with CCG members and alumni, especially Melissa Hawkins and Courtney Hofman, and to Nancy Rotzel McInerney for extensive logistical and technical support. We are grateful to Nic Kooyers for feedback on an earlier version of this manuscript, and to four anonymous reviewers for insightful comments on the manuscript.

References

Acevedo-Whitehouse, K., J. Vicente, C. Gortazar, U. Hofle, I.G. Fernández-de-Mera, and W. Amos 2005. Genetic resistance to bovine tuberculosis in the Iberian wild boar. *Molecular Ecology* 14: 3209–3217.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403–410.

Antao, T. A. Lopes, R.J. Lopes, A. Beja-Pereira and G. Luikart 2008. LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics* 9:323.

Asghar, M., D. Hasselquist, B. Hansson, P. Zehtindjiev, H. Westerdahl and S. Bensch 2015. Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science* 347: 436–438.

Atkinson, C.T., R.J. Dusek and J.K. Lease 2001. Serological responses and immunity to superinfection with avian malaria in experimentally-infected Hawai'i 'amakihi. *Journal of Wildlife Diseases* 37: 20–27.

Atkinson, C.T., R.J. Dusek J.K. Lease and M.D. Samuel 2005. Prevalence of pox-like lesions and malaria in forest bird communities on leeward Mauna Loa Volcano, Hawaii. *Condor* 107: 537–546.

Atkinson, C.T., R.J. Dusek, K.L. Woods and W.M. Iko 2000. Pathogenicity of avian malaria in experimentally infected Hawai'i 'amakihi. *Journal of Wildlife Diseases* 36: 197–204.

Atkinson, C.T. and D.A. LaPointe 2009. Introduced avian diseases, climate change, and the future of Hawai'ian honeycreepers. *Journal of Avian Medicine and Surgery* 23: 53–63.

Atkinson, C.T., K.S. Saili, R.B. Utzurrum and S.I. Jarvi 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. and R.B. Utzurrum 2010. Changes in prevalence of avian malaria on the Alakai Plateau, Kauai, 1997-2007. Technical Report HCSU-017, Hawai'i Cooperative Studies

Unit, Pacific Aquaculture and Coastal Resources Center, University of Hawai'i at Hilo.

Aucan, C., Y. Traoré, F. Tall, B. Nacro, T. Traore-Leroux, F. Fumoux, and P. Rihet 2000. High immunoglobulin G2 (IgG2) and low IgG4 levels are associated with human resistance to *Plasmodium falciparum* malaria. *Infection and Immunity* 68: 1252–1258.

Barclay, V.C., Sim, D., Chan, B.H., Nell, L.A., Rabaa, M.A., Bell, A.S., Anders, R.F. and Read, A.F., 2012. The evolutionary consequences of blood-stage vaccination on the rodent malaria Plasmodium chabaudi. *PLoS biology*, 10: e1001368.

Barclay, V.C., Kennedy, D.A., Weaver, V.C., Sim, D., Lloyd-Smith, J.O. and Read, A.F., 2014. The effect of immunodeficiency on the evolution of virulence: an experimental test with the rodent malaria Plasmodium chabaudi. *The American Naturalist*, *184*(S1): S47-S57.

Beadell JS, Ishtiaq F, Covas R *et al.* (2006) Global phylogeographic limits of Hawai'i's avian malaria. *Proceedings of The Royal Society B, Biological Sciences*, **273**, 2935–44.

Beaumont M.A. 2005. Adaptation and speciation: what can F_{st} tell us? *Trends in Ecology and Evolution* 20: 435–440.

Beaumont M.A. and R.A. Nichols 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B* 263: 1619–1626.

Best, S.M. and Kerr, P.J., 2000. Coevolution of host and virus: the pathogenesis of virulent and attenuated strains of myxoma virus in resistant and susceptible European rabbits. *Virology*, 267: 36-48.

Biggins, D.E. and M.H. Schroeder 1987. Historical and present status of the black-footed ferret. *Great Plains Wildlife Damage Control Workshop Proceedings*, Paper 50.

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

Bonneaud C, Balenger SL, Russell AF *et al.* (2011) Rapid evolution of disease resistance is accompanied by functional changes in gene expression in a wild bird. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 7866–71.

Burdon, J.J. and J.N. Thompson 1995. Changed patterns of resistance in a population of *Linum marginale* attacked by the rust pathogen *Melamspora lini*. *Journal of Ecology* 83: 199–206.

Callicrate, T., R. Dikow, J.W. Thomas, J.C. Mullikin, E.D. Jarvis, R.C. Fleischer and NISC Comparative Sequencing Program 2014. Genomic resources for the endangered Hawai'ian honeycreepers. *BMC Genomics* 15: 1098.

Caprio, M.A., and B.E. Tabashnik 1992. Gene flow accelerates local adaptation among finite populations: simulating the evolution of insecticide resistance. *Journal of Economic Entomology* 85.3: 611-620.

Chan, Y.F., M.E. Marks, F.C. Jones, G. Jr. Villarreal, M.D. Shapiro, S.D. Brady, A.M. Southwick, D.M. Absher, J. Grimwood, J. Schmutz, R.M. Myers, D. Petrov, B. Jónsson, D. Schluter, M.A. Bell,

and D.M. Kingsley 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327: 302-305.

Chessel, D., Dufour, A.B., Thioulouse, J., 2004. The ade4 package-I-One-table methods. *R News* 4, 5–10.

Coban, C., K.J. Ishii, T. Kawai, H. Hemmi, S. Sato, S. Uematsu, M Yamamoto, O. Takeuchi, S. Itagaki, N. Kumar, T. Horii, and S. Akira 2005. Toll-like receptor 9 mediates innate immune activation by the malaria pigment hemozoin. *The Journal of Experimental Medicine* 201: 19–25.

Colautti, R.I. and J.A. Lau 2015. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology* 24: 1999–2017.

Colosimo, P.F., K.E. Hosemann, S. Balabhadra, G. Jr. Villarreal, M. Dickson, J. Grimwood, J. Schmutz, R.M. Myers, D. Schluter, and D.M. Kingsley 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307: 1928-1933.

Cunningham, A.A. and P. Daszak 1998. Extinction of a species of land snail due to infection with a microsporidian parasite. *Conservation Biology* 12: 1139-1141.

Danecek, P., A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, R. Handsaker, G. Lunter, G. Marth, S.T. Sherry, G. McVean, R. Durbin and 1000 Genomes Project Analysis Group 2011. The Variant Call Format and VCFtools, *Bioinformatics* 27: 2156–2158.

Decaestecker, E., Gaba, S., Raeymaekers, J.A., Stoks, R., Van Kerckhoven, L., Ebert, D. and De Meester, L., 2007. Host–parasite 'Red Queen'dynamics archived in pond sediment. *Nature*, 450, p.870.

DePristo,, M.A., E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A.A. Philippakis, G. del Angel, M.A. Rivas, M. Hanna, A. McKenna, T.J. Fennell, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43: 491–498.

De Souza, J.B., K.H. Williamson, T. Otani, and J.H.L. Playfair 1997. Early gamma interferon responses in lethal and nonlethal murine blood-stage malaria. *Infection and Immunity* 65: 1593–1598.

Diaz, H. F., T. W. Giambelluca, and J. K. Eischeid. 2011. Changes in the vertical profiles of mean temperature and humidity in the Hawai'ian Islands. *Global and Planetary Change* 77:217:1.

van Dijk, A., E.J.A. Veldhuizen and H.P. Haagsman 2008. Avian defensins. *Veterinary Immunology and Immunopathology* 124: 1–18.

Donahue CG, Carruthers VB, Gilk SD, Ward GE (2000) The Toxoplasma homolog of Plasmodium apical membrane antigen-1 (AMA-1) is a microneme protein secreted in response to elevated intracellular calcium levels. *Molecular and Biochemical Parasitology*, **111**, 15–30.

Dray, S., and A.B. Dufour 2007. The ade4 package: implementing the duality diagram for

ecologists. *Journal of Statistical Software* 22: 1–20.

Egan, E.S., R.H.Y. Jiang, M.A. Moechtar, N.S. Barteneva, M.P. Weekes, L.V. Nobre, S.P. Gygi, J.A. Paulo, C. Frantzreb, Y. Tani, J. Takahashi, S. Watanabe, J. Goldberg et al. 2015. A forward genetic screen identifies erythrocyte CD55 as essential for *Plasmodium falciparum* invasion. *Science* 348: 711–714.

Eggert, L.S., L.A. Terwilliger, B.L. Woodworth, P.J. Hart, D. Palmer, and R.C. Fleischer 2008. Genetic structure along an elevational gradient in Hawai'ian honeycreepers reveals contrasting evolutionary responses to avian malaria. *BMC Evolutionary Biology* 8: 315.

Elena, S.F., and R.E. Lenski 2003. Evolution experiments with microorganisms: the dynamic and genetic bases of adaptation. *Nature Reviews Genetics* 4: 457-469.

Epstein, B. M. Jones, R. Hamede, S. Hendricks, H. McCallum, E.P. Murchison, B. Schonfeld, C. Wiench, P. Hohenlohe, and A. Storfer 2016. Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nature Communications* 7:12684. doi: 10.1038/ncomms12684.

Fleischer, R.C., C.E. McIntosh and C.L. Tarr 1998. Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawai'ian Islands to estimate molecular evolutionary rates. *Molecular Ecology* 7: 533–545.

Foll, M. and O.E. Gaggiotti 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977–993.

Fonseca, D.M., J.L. Smith, R.C. Wilkerson, and R.C. Fleischer 2006. Pathways of expansion and multiple introductions illustrated by large genetic differentiation among worldwide populations of the southern house mosquito. *American Journal of Tropical Medicine and Hygiene* 74: 284–289.

Fortini, L.B., A.E. Vorsino, F.A. Amidon, E.H. Paxton and J.D. Jacobi 2015. Large-scale range collapse of Hawai'ian forest birds under climate change and the need for 21st century conservation options. *PLoS One* 10: e0144311. DOI:10.1371/journal.pone.0140389 .

Foster, J.T., B. L. Woodworth, L. E. Eggert, P. J. Hart, D. Palmer, D. C. Duffy, and R. C. Fleischer 2007. Genetic structure and evolved malaria resistance in Hawai'ian honeycreepers. *Molecular Ecology* 16: 4738 – 4746.

Franklin, B.S., P. Parroche, M.A. Ataíde, F. Lauw, C. Ropert, R.B. de Oliveira, D. Pereira, M.S. Tada, P. Nogueria, L.H. Pereira da Silva, H. Bjorkbacka, D.T. Golenbock, and R.T. Gazzinelli 2009. Malaria primes the innate immune response due to interferon-γ induced enhancement of toll-like receptor expression and function. *Proceedings of the National Academy of Sciences of the USA* 106: 5789–5794.

Friedman, M. 1978. Erythrocytic mechanism of sickle cell resistance to malaria. *Proceedings of the National Academy of Sciences of the USA* 75: 1994–1997.

Friedman, M. 1983. Control of malaria virulence by alpha1-acid glycoprotein (orosomucoid), an acute-phase (inflammatory) reactant. *Proceedings of the National Academy of Sciences of the USA* 80: 5421–5424.

Frick, W.F., S.J. Puechmaille, J.R. Hoyt, B.A. Nickel, K.E. Langwig, J.T. Foster, K.E. Barlow, T. Bartonicka, D. Feller, A.-J. Haarsma, C. Herzog, I. Horacek et al. 2015. Disease alters macroecological patterns of North American bats. *Global Ecology and Biogeography* 24: 741–749 .

Fumagalli, M., M. Sironi, U. Pozzoli, A. Ferrer-Admetlla, L. Pattini, R. Nielsen 2011. Signatures of environmental genetics adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics* 7: e1002355.

Ganz, T. 2003. Defensins: antimicrobial peptides of innate immunity. *Nature Reviews Immunology* 3: 710–719.

Gemayel, R., J. Cho, S. Boeynaems, and K.J. Verstrepen 2012. Beyond junk- variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. *Genes* 3: 461-480.

Giambelluca, T.W., H.F. Diaz, and M.S.A. Luke 2008. Secular temperature changes in Hawai'i. *Geophysical Research Letters* 35: L12702.

Goff, M.L., and C. van Riper III (1980) Distribution of mosquitoes (Diptera: Culicidae) on the east flank of Mauna Loa volcano, Hawai'i 1. , **22**, 178–188.

Gordon, A & J. Hannon, G. (2010). Fastx-toolkit. FASTQ/A short-reads pre-processing tools (unpublished) http://hannonlab. cshl. edu/fastx_toolkit.

Grau, G.E., H. Heremans, P.-F. Piguet, P. Pointaire, P.-H. Lambert, A. Billiau, and P. Vassalli 1989. Monoclonal antibody against interferon γ can prevent experimental cerebral malaria and its associated overproduction of tumor necrosis factor. *Proceedings of the National Academy of Sciences of the USA* 86: 5572–5574.

Grimholt, U., S. Larsen, R. Nordmo, lP. Midtlyng, S. Kjoeglum, A. Storset, S. Saebo and R.J.M. Stet 2003. MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics* 55: 210–219.

Haldane, J.B.S. 1948. The theory of a cline. *Journal of Genetics* 48: 277–284.

Hawley, D.M. and R.C. Fleischer 2012. Contrasting epidemic histories reveal pathogen-mediated balancing selection on class II MHC diversity in a wild songbird. *PLoS ONE*, **7**, e30222.

Hedrick, P.W. 2011. Population genetics of malaria resistance in humans. *Heredity* 107: 283–304.

Hellgren, O., J. Waldenström, J. Peréz-Tris, E. Szöll Ösi, D. Hasselquist, A. Krizanauskiene, U. Ottosson and S. Bensch 2007. Detecting shifts of transmission areas in avian blood parasites — a phylogenetic approach. *Molecular Ecology* 16: 1281–1290.

Hirayasu, K., J. Ohashi, K. Kashiwase, H. Hananantachai, I. Naka, A. Ogawa, M. Takanashi, M. Satake, K. Nakajima, P. Parham, H. Arase, K. Tokunaga, et al. 2012. Significant association of KIR2DL3-HLA-C1 combination with cerebral malaria and implications for co-evolution of KIR and HLA. *PLoS Pathogens* 8: e1002565.

Hochachka, W.M. and A.A. Dhondt 2000. Density-dependent decline of host abundance resulting from a new infectious disease. *Proceedings of the National Academy of Sciences of the USA* 97: 5303–5306.

Hoekstra, H.E., and M.W. Nachman 2003. Different genes underlie adaptive melanism in different populations of rock pocket mice. *Molecular Ecology* 12: 1185-1194.

Huff, C.G., Marchbank, D.F., and Shiroishi, T. 1958. Changes in infectiousness of malarial gametocytes. II. Analysis of the possible causative factors. *Experimental Parasitology* 7: 399-417

Jakobsen, P.H., V. McKay, S.D. Morris-Jones, W. McGuire, M. Boele van Hensbroek, S. Meisner, K. Bendtzen, I. Schousboe, I.C. Bygbjerg, and B.M. Greenwood 1994. Increased concentrations of interleukin-6 and interleukin-1 receptor antagonist and decreased concentrations of beta-2-glycoprotein I in Gambian children with cerebral malaria. *Infection and Immunity* 4374–4379.

James, H.F. and S.L. Olson 1991. Descriptions of thirty-two new species of birds from the Hawai'ian Islands: Part II. Passeriformes. *Ornithological Monographs* 46: 1–88.

Jarvi, S.I., C.T. Atkinson, and R.C. Fleischer 2001. Immunogenetics and resistance to avian malaria in Hawai'ian honeycreepers (Drepanidinae). *Studies in Avian Biology* 22: 254 – 263.

Jarvi, S. I., Bianchi, K. R., Farias, M. E., Txakeeyang, A., McFarland, T., Belcaid, M., & Asano, A. 2016. Characterization of class II β chain major histocompatibility complex genes in a family of Hawai'ian honeycreepers: ''amakihi (Hemignathus virens). *Immunogenetics*, 68: 461-475.

Jarvi, S. I., Tarr, C. L., Mcintosh, C. E., Atkinson, C. T., & Fleischer, R. C. 2004. Natural selection of the major histocompatibility complex (Mhc) in Hawai'ian honeycreepers (Drepanidinae). *Molecular Ecology*, 13: 2157-2168.

Jones FC, Grabherr MG, Chan YF *et al.* 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484, 55–61.

Jones, M. R., Cheviron, Z. A., & Carling, M. D. 2014. Variation in positively selected major histocompatibility complex class I loci in rufous-collared sparrows (Zonotrichia capensis). *Immunogenetics*, 66: 693-704.

Jones, M. R., Cheviron, Z. A., & Carling, M. D. 2015. Spatially variable coevolution between a haemosporidian parasite and the MHC of a widely distributed passerine. *Ecology and Evolution*, 5: 1045-1060.

Josse, J. and F. Husson, 2012. Handling missing values in exploratory multivariate data analysis methods. *Journal de la SFdS* 153: 79–99.

Kashi, Y., and D.G. King 2006. Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22: 253–259.

Katti, M.V., R. Sami-Subbu, P.K. Ranjekar, and V.S. Gupta 2000. Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. *Plant Science*

9: 1203–1209.

Kaufman, J. 2000. The simple chicken major histocompatibility complex: life and death in the face of pathogens and vaccines. *Philosophical Transactions of the Royal Society of London B* 355: 1077–1084.

Keller, I., C.E. Wagner, L. Greuter, S. Mwaiko, O.M. Selz, A. Sivasundar, S. Wittwer, and O. Seehausen 2013. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology* 22: 2848-2863.

Keyghobadi, N., D. LaPointe, R.C. Fleischer and D.M. Fonseca 2006. Fine-scale population genetic structure of a wildlife disease vector: the southern house mosquito on the island of Hawai'i. *Molecular Ecology* 15: 3919 – 3930.

Kilpatrick, A.M., D.A. LaPointe, C.T. Atkinson, B.L. Woodworth, J.K. Lease, M.E. Reiter and K. Gross 2006. Effects of chronic avian malaria (*Plasmodium relictum*) infection on reproductive success of Hawai'i 'amakihi (*Hemignathus virens*). *The Auk* 123: 764–774.

Kooyers, N.J. and K.M. Olsen 2012. Rapid evolution of an adaptive cyanogenesis cline in introduced North American white clover (*Trifolium repens* L.). *Molecular Ecology* 21: 2455-2468.

Kossodo, S., C. Monso, P. Juillard, T. Velu, M. Goldman, and G.E. Grau 1997. Interleukin-10 modulates susceptibility in experimental cerebral malaria. *Immunology* 91: 536–540.

Kurtzhals, J.A.L., V. Adabayeri, B.Q. Goka, B.D. Akanmori, J.O. Oliver-Commey, F.K. Nkrumah, C. Behr, and L. Hviid 1998. Low plasma concentrations of interleukin 10 in severe malarial anaemia compared with cerebral and uncomplicated malaria. *The Lancet* 351: 1768–1772.

Langwig, K.E., J.R. Hoyt, K.L. Parise, W.F. Frick, J.T. Foster, and A.M. Kilpatrick 2017. Resistance in persisting bat populations after white-nose syndrome invasion. *Philosophical Transactions of the Royal Society B* 372: 20160044.

LaPointe, D.A., M.L. Goff and C.T. Atkinson 2010. Thermal constraints to the sporogonic development and altitudinal distribution of avian malaria *Plasmodium relictum* in Hawai'i. *Journal of Parasitology* 96: 318–324.

Lerner, H.R.L., M. Meyer, H.F. James, M. Hofreiter, and R.C. Fleischer 2011. Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawai'ian honeycreepers. *Current Biology* 21: 1838–1844.

Lewontin, R.C. and J. Krakauer 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74: 175–195.

Li, C., I. Corraliza, and J. Langhorne 1999. A defect in interleukin-10 leads to enhanced malarial disease in *Plasmodium chabaudi chabaudi* infection in mice. *Infection and Immunity* 67: 4435–4442.

Li, H. and R. Durbin 2009. Fast and accurate short read alignment with Burrows- Wheeler transform. *Bioinformatics* 25: 1754–1760.

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

Lindsey, G.D., E.A. Vanderwerf, H. Baker, and P.E. Baker, Hawai'i 'amakihi (*Hemignathus virens*), in *The Birds of North America*, A. Poole, Editor. 1998, Cornell Lab of Ornithology: Ithaca, NY.

Lischer, H.E.L., and L. Excoffier 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.

Lively, C.M., and V. Apanius 1995. Genetic diversity in host-parasite interactions. In *Ecology of Infectious Diseases in Natural Populations*, Ed. B.T. Grenfell and A.P. Dobson. Cambridge University Press, Cambridge MA.

Lotterhos, K.E., and M.C. Whitlock 2014. Evaluation of demographic history and neutral parameterization of the performance of F_{ST} outlier tests. *Molecular Ecology* 23: 2178–2192.

Love RR, Steele AM, Coulibaly MB *et al.* (2016) Chromosomal inversions and ecotypic differentiation in Anopheles gambiae : the perspective from whole-genome sequencing. *Molecular Ecology*, **25**, 5889–5906.

Lucek, K., M. Lemoine, and O. Seehausen 2014. Contemporary ecotypic divergence during a recent range expansion was facilitated by adaptive introgression. *Journal of Evolutionary Biology* 27: 2233–2248.

Mackinnon, M.J. and Read, A.F., 2004. Immunity promotes virulence evolution in a malaria model. *PLoS biology*, 2: e230.

McKenna, A., E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. DePristo 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20:1297–1303.

Mi H, Muruganujan A, Casagrande JT, Thomas PD (2013) Large-scale gene function analysis with the PANTHER classification system. *Nature protocols*, **8**, 1551–66.

Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD (2016) PANTHER version 10: Expanded protein families and functions, and analysis tools. *Nucleic Acids Research*, **44**, D336–D342.

Mockenhaupt, F.P., J.P. Cramer, L. Hamann, M.S. Stegemann, J. Eckert, N.-R. Oh, R.N. Otchwemah, E. Dietz, S. Ehrhardt, N.W.J. Schroder, U. Bienzle, and R.R. Schumann 2006. Toll-like receptor (TLR) polymorphisms in African children: common TLR-4 variants predispose to severe malaria. *Proceedings of the National Academy of Sciences of the USA* 103: 177–182.

Molano, A., S.H. Park, Y.H. Chiu, S. Nosseir, A. Bendelac, M. Tsuji 2000. Cutting edge: the IgG response to the circumsporozoite protein is MHC class II-dependent and CD1d-independent: exploring the role of GPIs in NK T cell activation and antimalarial responses. *Journal of Immunology* 164: 5005–5009.

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction

endonucleases. *Proceedings of The National Academy of Sciences of the USA*, **76**, 5269–5273.

Nosil, P., L.J. Harmon, and O. Seehausen 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution* 24: 145–156.

Notley-McRobb, L. and Ferenci, T. 1999. The generation of multiple co-existing *mal*-regulatory mutations through polygenic evolution in glucose-limited populations of *Escherichia coli*. *Environmental Microbiology* 1: 45–52.

Notley-McRobb, L. and Ferenci, T. 2000. Experimental Analysis of Molecular Events During Mutational Periodic Selections in Bacterial Evolution. *Genetics* 156: 1493-1501.

Ockenhouse, C.F., N.N. Tandon, C. Magowan, G.A. Jamieson, and J.D. Chulay 1989. Identification of a platelet membrane glycoprotein as a falciparum malaria sequestration receptor. *Science* 243: 1469–1471.

Othoro, C., A.A. Lal, B. Nahlen, D. Koech, A.S.S. Orago, and V. Udhayakumar 1999. A low interleukin-10 tumor necrosis factor-α ratio is associated with malaria anemia in children residing in a holoendemic malaria region in western Kenya. *Journal of Infectious Diseases* 179: 279–282.

Palacios, E.H. and A. Weiss 2004. Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. *Oncogene* 23: 7790–8000.

Park, I. and Y. Yun 2009. Transmembrane adaptor proteins positively regulating the activation of lymphocytes. *Immune Network* 9: 53–57.

Paxton, E.H., R.J. Camp, P.M. Gorresen, L.H. Crampton, D.L. Leonard Jr., and E.A. VanderWerf 2016. Collapsing avian community on a Hawai'ian island. *Science Advances* 2: e1600029.

Perlmann, P., H. Perlmann, B.W. Flyg, M. Hagstedt, G. Elghazali, S. Worku, V. Fernandez, A.S.M. Rutta, and M. Troye-Blombert 1997. Immunoglobulin E, a pathogenic factor in *Plasmodium falciparum* malaria. *Infection and Immunity* 65: 116–121.

Pfeifer, S.P., S. Laurent, V.C. Sousa, C.R. Linnen, M. Foll, L. Excoffier, H.E. Hoekstra, and J.D. Jensen 2018. The evolutionary history of Nebraska deer mice: local adaptation in the face of strong gene flow. *Molecular Biology and Evolution* https://doi.org/10.1093/molbev/msy004

Prentice, M.B., J. Bowman, J.L. Lalor, M.M. McKay, L.A. Thomson, C.M. Watt, A.G. McAdam, D.L. Murray, and P.J. Wilson 2017. Signatures of selection in mammalian clock genes with coding trinucleotide repeats: implications for studying the genomics of high-pace adaptation. *Ecology and Evolution* 00: 1-23.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation

for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raberg, L., D. Sim, and A.F. Read 2007. Disentangling genetic variation for resistance and tolerance to infectious disease in animals. *Science* 318: 812–814.

Remold-O'Donnell E, Rosen FS, Kenney DM (1996) Defects in Wiskott-Aldrich syndrome blood

cells. *Blood*, **87**, 2621–2631.

Richman, A.M., G. Dimopoulos, D. Seeley, and F.C. Kafatos 1997. *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. *The EMBO Journal* 16: 6114–6119.

van Riper, C.I., S.G. van Riper, M.L. Goff and M. Laird 1986. The epizootiology and ecological significance of malaria in Hawai'ian land birds. *Ecological Monographs*, **56**, 327–344.

Rocke, T.E., J. Williamson, K.R. Cobble, J.D. Busch, M.F. Antolin, and D.M. Wagner 2012. Resistance to plague among black-tailed prairie dog populations. *Vector-borne and Zoonotic Diseases* 12: 111–116.

Roda, F., H. Liu, M.J. Wilkinson, G.M. Walter, M.E. James, D.M. Bernal, M.C. Melo, A. Lowe, L.H. Rieseberg, P. Prentis, and D. Ortiz-Barrientos 2013. Convergence and divergence during the adaptation to similar environments by an Australian groundsel. *Evolution* 67: 2515-2519.

Rosenblum, E.B., C.E. Parent, and E.E. Brandt 2014. The molecular basis of phenotypic convergence. *Annual Reviews in Ecology, Evolution and Systematics* 45: 203-226.

Rosenblum, E.B., H. Rompler, T. Schoneberg, and H.E. Hoekstra 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences of the USA* 107: 2113-2117.

Rousset 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103 – 106.

Samuel, M. D., B. L. Woodworth, C. T. Atkinson, P. J. Hart, and D. A. LaPointe. 2015. Avian malaria in Hawai'ian forest birds: infection and population impacts across species and elevations. *Ecosphere* 6:104. http://dx.doi. org/10.1890/ES14-00393.1

Scheibel LW, Colombani PM, Hess AD *et al.* (1987) Calcium and calmodulin antagonists inhibit human malaria parasites (Plasmodium falciparum): Implications for drug design. *Proceedings of The National Academy of Sciences of The United States of America*, **84**, 7310–7314.

Schofield, L., M.J. McConville, D. Hansen, A.S. Campbell, B. Fraser-Reid, M.J. Grusby, and S.D. Tachado 1999. CD1d-restricted immunoglobulin G formation to GPI-anchored antigens mediated by NKT cells. *Science* 283: 225–229.

Semple, C.A., M. Rolfe, and J.R. Dorin 2003. Duplication and selection in the evolution of primate beta-defensin genes. Genome Biol., 4, R31.

Shahabuddin, M., I. Fields, P. Bulet, J.A. Hoffmann, and L.H. Miller 1998. *Plasmodium gallinaceum*: differential killing of some mosquito stages of the parasite by insect defensin. *Experimental Parasitology* 89: 103–112.

Stankowski S, Sobel JM, Streisfeld MA (2016) Geographic cline analysis as a tool for studying genome-wide variation : a case study of pollinator-mediated divergence in a monkeyflower.

Tanabe K, Mikkelsen RB, Wallach DFH (1982) Calcium transport of *Plasmodium chabaudi*-

infected erythrocytes. *The Journal of Cell Biology*, **93**, 680–684.

Tarleton, R.L., M.J. Grusby, M. Postan and L.H. Climcher 1996. *Trypanosoma cruzi* infection in MHC-deficient mice: further evidence for the role of both class I- and class II-restricted T cells in immune resistance and disease. *International Immunology* 8: 13–22.

Tarr, C.L. and R.C. Fleischer 1993. Mitochondrial DNA variation and evolutionary relationships in the 'amakihi complex. *The Auk* 110: 825**–**831.

Thompson, C.J., Ahmed, N.I., Veen, T., Peichel, C.L., Hendry, A.P., Bolnick, D.I. and Stuart, Y.E., 2017. Many-to-one form-to-function mapping weakens parallel morphological evolution. *Evolution* 71: pp.2738-2749.

Thorne, E.T. and E.S. Williams 1988. Disease and endangered species: the black-footed ferret as a recent example. *Conservation Biology* 2: 66**–**74.

Turchaninova, M. A., Davydov, A., Britanova, O. V., Shugay, M., Bikos, V., Egorov, E. S., ... & Mamedov, I. Z. 2016. High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nature protocols*, 11: 1599.

Turner, J.K., M.M. McAllister, J.L. Xu, and R.I. Tapping 2008. The resistance of BALB/cJ mice to *Yersinia pestis* maps to the major histocompatibility complex of chromosome 17. *Infection and Immunity* 76: 4092**–**4099.

Van der Auwera, G.A., M.O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan et al. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43: 11.10.1–11.10.33.

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

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. *The American Naturalist*, **186**, S24–S36.

Woodworth, B.L., C.T. Atkinson, D.A. LaPointe, P.J. Hart, C.S. Spiegel, E.J. Tweed, C. Henneman, J. LeBrun, T. Denette, R. DeMots, K.L. Kozar, D. Triglia et al. 2005. Host population persistence in the face of introduced vector-borne diseases: Hawai'i 'amakihi and avian malaria. *Proceedings of the National Academy of Sciences of the USA* 102: 1531–1536.

Yeaman, S. and M.C. Whitlock 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution* 65: 1897–1911.

Zhang, Z., S. Schwartz, L. Wagner, and W. Miller 2000. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* 7: 203**–**14.

Zhuang, X., C. Yang, S.-E. Fevolden, and C.-H.C. Cheng 2012. Protein genes in repetitive sequence--antifreeze glycoproteins in Atlantic cod genome. *BMC Genomics* 13: 293.

Data Accessibility

SNP analysis pipeline and associated files: Available at https://github.com/CassinSackett/SNPcapture

SNP genotypes: Will be made available at the Sequence Read Archive upon acceptance; Project Accession TBD.

Author Contributions

The study was conceived by RCF and LCS, data were generated and pipeline designed by TEC and LCS, data were analyzed by LCS, and the manuscript was written by LCS with contributions from RCF and TEC.

Tables and Figures

Table 1. List of immune-related genes inferred to be under potential selection by various methods. In cases where loci blasted against multiple results, only the top 3 hits are displayed. All population pairs are low-elevation vs. high-elevation unless noted.

Immune-related gene	**How selection was inferred**	**In which population/pair**
Attractin	Statistical outlier	Hualālai – Mauna Loa
Beta defensin gene cluster	Statistical outlier	Hualālai – Mauna Loa
	Top 0.1% highest F_{ST}	Low-high elevation; Hualālai – Mauna Loa; Hualālai – Pu'u La'au; Manukā – Mauna Loa; Manukā – Pōhakuloa; Pāhoa – Mauna Loa; Pāhoa – Pōhakuloa; Pāhoa – Pu'u La'au (all with multiple loci)
		Pāhoa; Manukā; Hualālai (all low)
	25 most negative Tajima's D	Manukā (low); Pāhoa (low); mid-elevation
	LROH	
CD glycoproteins (1b-3, 4, 7, 8b, 59, 99, 101, 180, 200, 276)	Top 0.1% highest F_{ST}	Low – high elevation; Hualālai – Mauna Loa; Manukā – Mauna Loa; Pāhoa – Mauna Loa; Pāhoa – Pōhakuloa
	25 most negative Tajima's D	Manukā (low-elevation); Pōhakuloa (high-elevation)
		Pāhoa (low-elevation)
	LROH	
Contactin 3, 5	Statistical outlier	Manukā – Mauna Loa
	Top 0.1% highest F_{ST}	Manukā – Mauna Loa; Pāhoa – Mauna Loa
	LROH	Pāhoa (low-elevation)
Cytoplasmic FMR1 interacting protein 2 (CYFIP2)	Statistical outlier	Pāhoa – Pu'u La'au
	Top 0.1% highest F_{ST}	Hualālai – Pu'u La'au; Pāhoa – Pu'u La'au

Family with sequence similarity 83H, 174B, 221A (associated with interleukin-8 secretion & viral loads)	Top 0.1% highest F_{ST} LROH	Pāhoa – Pōhakuloa; Pāhoa – Pu'u La'au Hualālai; Pāhoa (both low-elevation)
FK506 binding proteins 1B, 4, 14, 15	Top 0.1% highest F_{ST} LROH	Hualālai – Mauna Loa; Hualālai – Pōhakuloa; Manukā – Pōhakuloa; Pāhoa – Mauna Loa; Pāhoa – Pu'u La'au Manukā (low-elevation)
Heat shock proteins (Hsp40, Hsp70)	Top 0.1% highest F_{ST}	Low – high elevation; Manukā – Pōhakuloa
Hematopoietic lineage cell-specific protein (antigen receptor signaling)	Top 0.1% highest F_{ST}	Pāhoa – Pu'u La'au
Immunoglobulin receptors	Top 0.1% highest F_{ST} 25 most negative Tajima's D LROH	Low – high elevation; Hualālai – Mauna Loa Hualālai (low); Mid-elevation; Pōhakuloa (high) Manukā (low-elevation
Interferon stimulator and α/β receptor 2	Top 0.1% highest F_{ST} 25 most negative Tajima's D LROH	Hualālai – Pōhakuloa; Hualālai – Pu'u La'au; Manukā – Pu'u La'au Hualālai (low-elevation) Mid-elevation
Interleukin regulators/ receptors/ binding & associated proteins (2, 3, 8, 10, 11, 12B, 16, 17B, 18, 23)	Top 0.1% highest F_{ST} 25 most negative Tajima's D	Low – high elevation; Hualālai – Mauna Loa; Hualālai – Pōhakuloa; Hualālai – Pu'u La'au; Manukā – Mauna Loa; Manukā – Pōhakuloa; Manukā – Pu'u La'au; Pāhoa – Mauna Loa; Pāhoa – Pōhakuloa; Pāhoa – Pu'u La'au Hualālai; Manukā; Pāhoa (all low-elevation) Pu'u La'au (high elevation) Mid-elevation

	LROH	Pāhoa (low-elevation); Manukā (low-elevation), Mid-elevation
Lymphocyte antigens (6E, 75)	Top 0.1% highest F_{ST} 25 most negative Tajima's D LROH	Low – high elevation; Hualālai – Pu'u La'au Pāhoa; Hualālai (both low-elevation) (low-elevation)
Major Histocompatibility Complex & NFX1 (regulates expression of MHC II)	Top 0.1% highest F_{ST} 25 most negative Tajima's D LROH	Low – high elevation; Hualālai – Pu'u La'au; Manukā – Mauna Loa; Pāhoa – Mauna Loa; Pāhoa – Pōhakuloa Manukā; Hualālai (both low-elevation) Manukā (low-elevation); Pāhoa (low-elevation); Mid-elevation
Semaphorin (modulates immune response following CNS trauma)	Top 0.1% highest F_{ST} Top 0.1% highest F_{ST}	Manukā - Pu'u La'au Pāhoa - Pu'u La'au
T-cell related (differentiation protein MAL; receptors; activators)	Top 0.1% highest F_{ST} LROH	Hualālai– Mauna Loa; Manukā – Mauna Loa; Manukā – Pōhakuloa; Manukā – Pu'u La'au Manukā (low-elevation)
Toll-like receptor 5, 7	Statistical outlier Top 0.1% highest F_{ST} 25 most negative Tajima's D	Manukā – Mauna Loa Manukā – Mauna Loa Hualālai (low-elevation)
Tumor necrosis factor ligand member 10 (induces apoptosis)	Top 0.1% highest F_{ST} 25 most negative Tajima's D	Pāhoa – Mauna Loa Mauna Loa; Pōhakuloa (both high-elevation) Mid-elevation

Figures

Figure 1: Map of sites from which 125 Hawai'i 'amakihi were sampled across the island of Hawai'i. Circles represent low-elevation sites (0-1200m; mosquitoes present), square represents a mid-elevation site (1200-1400m), and triangles represent high-elevation sites (1450-2400m; mosquitoes absent).

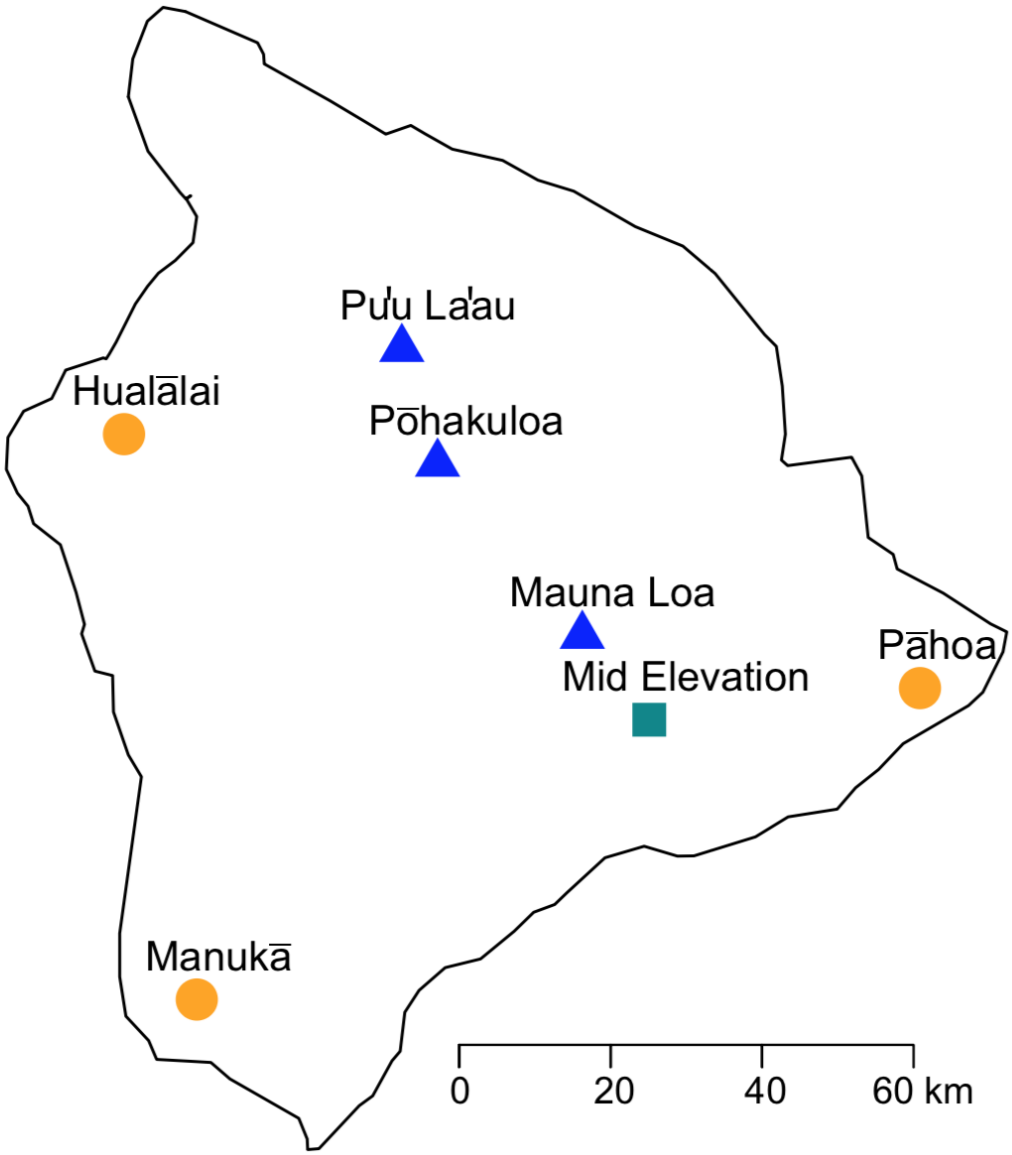

Figure 2. Plot of F_{ST} between all low- and high-elevation population pairs across the genome; alternating colors represent different chromosomes. Dashed lines represent the threshold for denoting quasi-outliers in each population. Outlier genes are denoted with asterisks; genes without asterisks represent the most-differentiated SNPs in each population pair. Population pairs with no gene names had no outliers blast to immune-related genes.

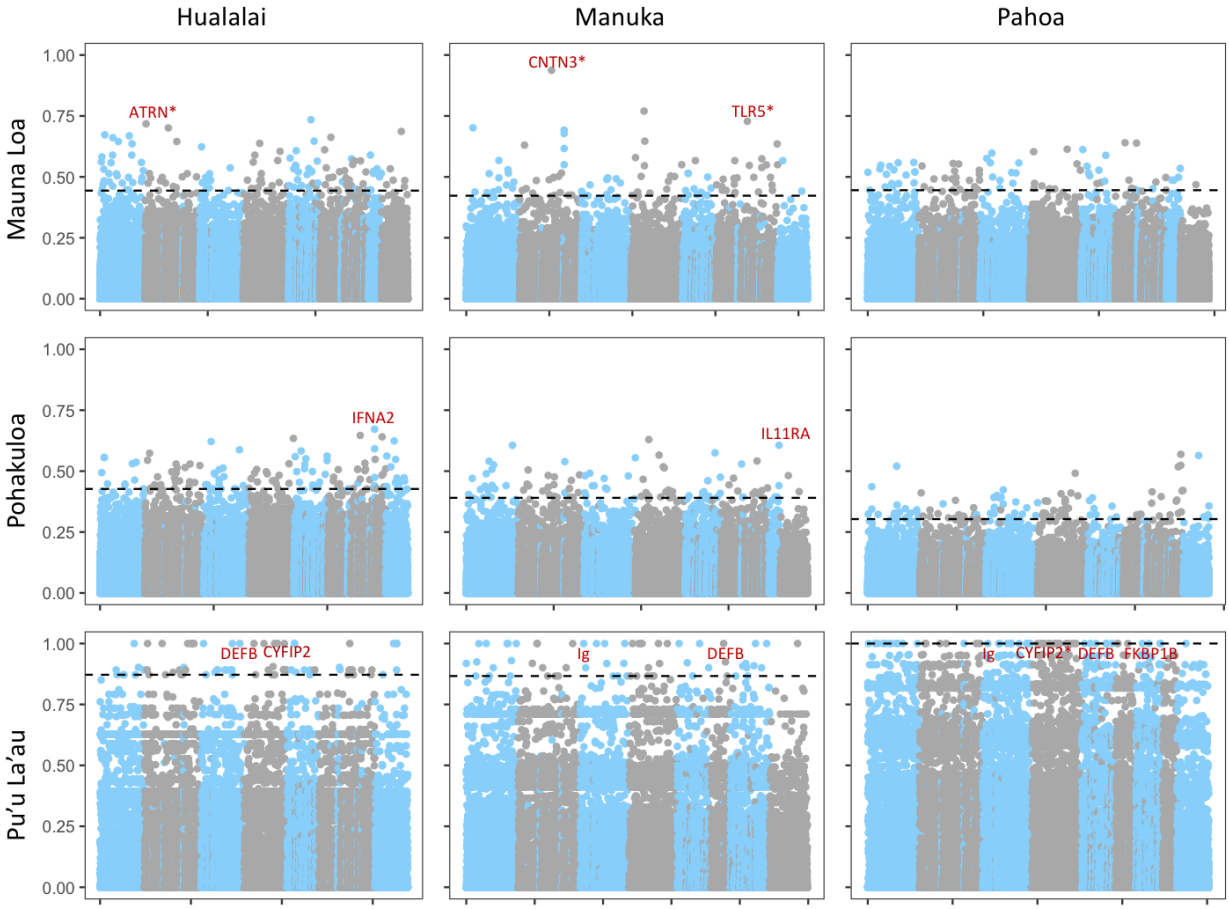

Figure 3. Principal Component Analysis of genotypes in individuals from low (N=46), mid (N=8) and high (N=46) elevations. Plot summarizes genotypes at 2816 SNPs with no missing data. PC1: Principal Component Axis 1 (which explains 1.9% of the variance), PC2: Principal Component Axis 2 (which explains 1.8% of the variance); 96 axes were needed to explain all genetic variance.

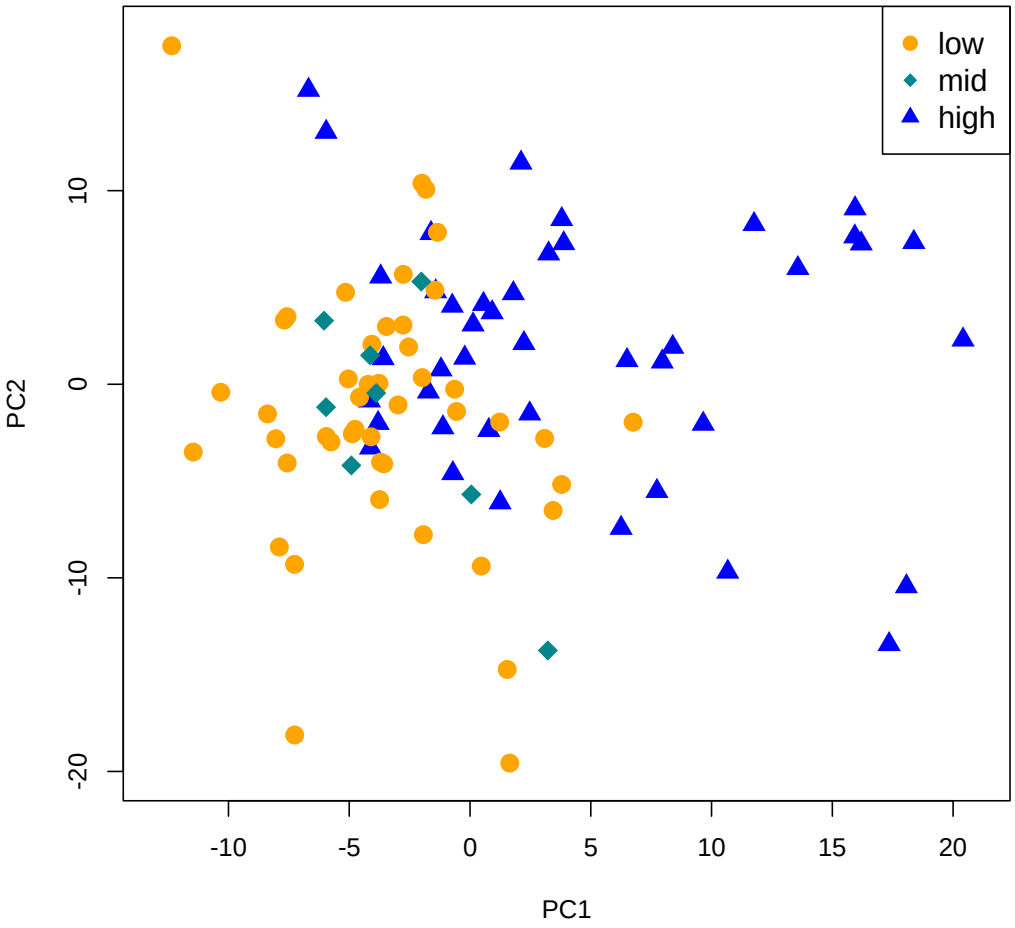

Figure 4. Venn diagrams showing the sharing of immune-related genes, inferred from blast results, among low-elevation populations. a) each specific gene counted uniquely, b) classes of similar genes (e.g., all glycoproteins) combined.

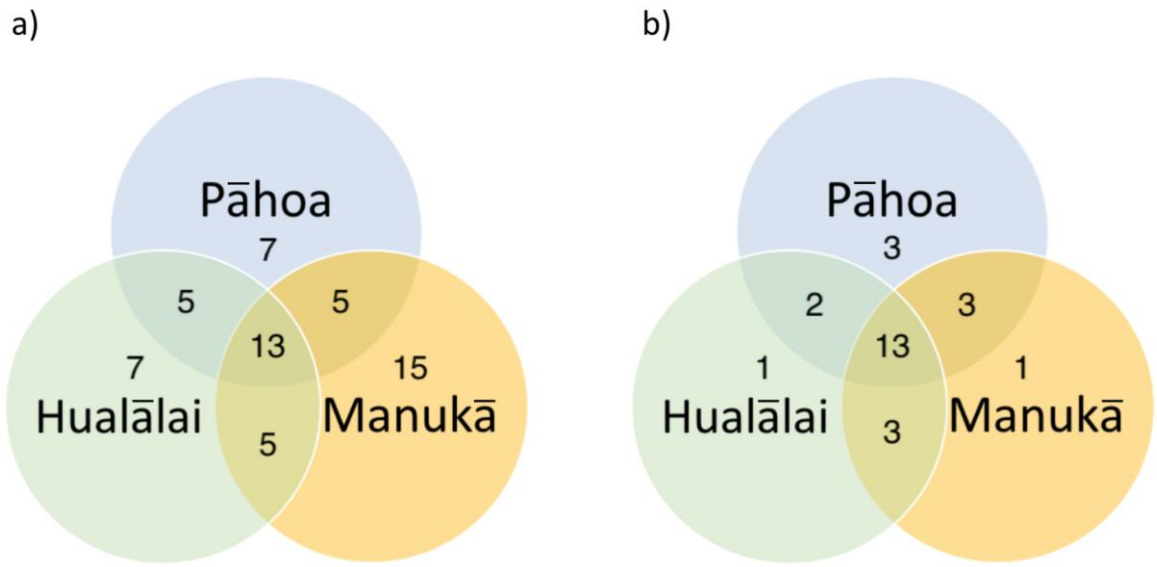

Figure 5. Plot of significantly (p<0.10) under- and over-represented biological process categories (y-axis) in 'amakihi quasi-outlier genes relative to the expected number of genes in each category in the annotated chicken (*Gallus gallus*) genome. 'amakihi genes are derived from a blast search of the 0.1% of loci with the highest F_{ST} between low-elevation and high-elevation 'amakihi. Vertical line represents zero (equally represented between chicken and 'amakihi), and x-axis denotes the degree of under- or overrepresentation (e.g., '3' means a category was overrepresented threefold in 'amakihi).