7-1-2014

DIVERSIFICATION AND ADAPTATION IN THE ANDES: INSIGHTS FROM PHYLOGEOGRAPHY, MALARIA, AND HEMOGLOBIN OF THE HOUSE WREN (TROGLODYTES AEDON)

Spencer Galen

Spencer Galen
Candidate

Biology
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Christopher Witt , Chairperson

Eric Loker

Ben Hanelt

DIVERSIFICATION AND ADAPTATION IN THE ANDES: INSIGHTS FROM PHYLOGEOGRAPHY, MALARIA, AND HEMOGLOBIN OF THE HOUSE WREN (TROGLODYTES AEDON)

BY

SPENCER C. GALEN

B.S., UNIVERSITY OF DELAWARE, 2010

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science in Biology

The University of New Mexico
Albuquerque, New Mexico

July 2014

ACKNOWLEDGEMENTS

I would like to thank my advisor, Chris Witt, for his excellent guidance and support. My thesis benefitted greatly from his endless enthusiasm. I also thank my thesis committee, Eric Loker and Ben Hanelt, for their insight and assistance along the way. For intellectual discourse, I would like to thank Matt Baumann, Libby Beckman, Phred Benham, Andrea Chavez, Shane DuBay, Ariel Gaffney, Andy Johnson, Sabrina McNew, Iris Olivas, Jonathan Schmitt, Ashley Smiley, Cole Wolf, and Natalie Wright. In addition to the aforementioned individuals, I thank Fernando Angulo, Emil Bautista, Daniel F. Lane, Sheila Figeroa, L. Monica Flores, Willy Ñáñez, Jano Nuñez, Alessandra Quiñonez Z., Donna C. Schmitt, Dora Susanibar, Thomas Valqui, Walter Vargas, and many associates of Centro de Ornitología y Biodiversidad (CORBIDI) for assisting with fieldwork to obtain key specimens that greatly improved the research presented here. Lastly, this thesis would not have been possible without funding from NSF DEB-1146491, NSF DEB-0543556, NCRR-NIH P20RR018754, the UNM Center for Evolutionary and Theoretical Immunology (CETI), UNM-RAC, Robert W. Dickerman, and the Lynn A. Hertel Graduate Research Award.

DIVERSIFICATION AND ADAPTATION IN THE ANDES: INSIGHTS FROM PHYLOGEOGRAPHY, MALARIA, AND HEMOGLOBIN OF THE HOUSE WREN (*TROGLODYTES AEDON*)

by

Spencer C. Galen

B.S., University of Delaware, 2010

M.S., University of New Mexico, 2014

ABSTRACT

The tropical Andes of South America represent a highly complex landscape characterized by physical barriers to dispersal and stark ecological gradients. The study of intraspecific genetic diversity in widespread Andean species has the potential to elucidate the historical biogeographic events that have driven diversification as well as the processes by which organisms have adapted to their environments in this region. The house wren (*Troglodytes aedon*) provides a convenient system for the study diversification and adaptation in the Andes, as it is distributed widely across both biogeographic barriers and heterogeneous environments.

In chapter one, I utilize house wren and malaria parasite mtDNA to investigate host-parasite phylogeography and diversification. I also examine the bioclimatic correlates of malaria abundance in order to understand how the host and external environment interact to shape malaria biogeography. I found that the house wren and its

Haemoproteus parasites co-diversified during the Pleistocene within the Andes, but that the diversification of the parasites was diffuse with respect to the regional avifauna. House wren-specific *Haemoproteus* parasites were polyphyletic and appeared to have speciated by switching hosts from other South American bird species. Malaria distributions were structured with respect to climate, with parasites in the genus *Leucocytozoon* significantly associated with high elevation habitats above 2,000 m.

In chapter two I use a population genetic approach to study the evolution of hemoglobin in response to high altitude hypoxia in Andean house wrens. I discovered five amino acid substitutions in the genes that encode the subunits of the major hemoglobin isoform that segregated with respect to elevation. One substitution, β(55)Val⇒Ile, was significantly structured between low (<1,000 m) and high (>3,000 m) elevation house wren populations. The genetic structure associated with β(55) was supported by coalescent analyses that revealed reduced migration at the βA locus. These findings suggest that the β(55)Ile allele confers increased hemoglobin-O_2 affinity and is adaptive at high altitudes.

TABLE OF CONTENTS

Chapter 1

Diverse avian malaria and other haemosporidian parasites in Andean house wrens: evidence for diffuse co-diversification by host-switching

Galen, S.C and Witt, C.C. 2014. *Journal of Avian Biology*. doi: 10.1111/jav.00375

INTRODUCTION

Avian blood parasites of the genera *Leucocytozoon*, *Haemoproteus*, and *Plasmodium* (Phylum Apicomplexa, Order Haemosporida) are a globally abundant and diverse group that may harbor as many as ~10,000 evolutionarily independent mtDNA lineages worldwide (Bensch et al. 2004). However, despite a steadily increasing number of surveys of avian haemosporidians, our understanding of the biogeography of these parasites is still in its infancy.

The diversity and distribution of avian haemosporidian parasites is heavily influenced by interactions with hosts over evolutionary time. The importance of host-parasite evolutionary history is illustrated by the observations that avian haemosporidians are non-randomly distributed across host species (Ricklefs et al. 2005, Medeiros et al. 2013), and that haemosporidians do not develop equally well in different, closely related host species (Palinauskas et al. 2008, Palinauskas et al. 2009). Several processes can determine how the diversification of host and parasite clades occurs over evolutionary time. First, diversification of haemosporidian parasites and their hosts can occur by co-diversification, the concurrent diversification of interacting host and parasite clades (Johnson and Stinchcombe 2007; see Supplementary material Table A1 for glossary of

terms used in this paper). Co-diversification can involve host-specific or non-host-specific interactions; host-specific parasites are likely to diversify by co-speciation, the reciprocal and interdependent diversification of host and parasite lineages that are ecologically associated (Page 2003). If co-speciation is the exclusive mechanism of haemosporidian diversification, host and parasite phylogenies are predicted to be topologically concordant and contemporaneous. The tendency for haemosporidian clades to be specific to host genera or families (Beadell et al. 2004, Pérez-Tris et al. 2007, Beadell et al. 2009) provides circumstantial evidence of co-speciation, but examples of matched branching patterns in host and haemosporidian phylogenies have proved rare (Ricklefs and Fallon 2002, Ricklefs et al. 2004, Jenkins et al. 2012). Moreover, co-speciation cannot be the only mechanism of haemosporidian diversification because host conservatism diminishes rapidly at deeper levels of the avian haemosporidian phylogeny (Ricklefs and Fallon 2002, Ricklefs et al. 2004).

Several evolutionary processes can cause the decoupling of phylogenetic patterns in coevolving hosts and parasites, even when parasites are exclusive to a single host (Johnson et al. 2003). One example is within-host duplication, in which parasites diversify without host population structure. Ricklefs et al. (2004) found support for duplication using event-based phylogenetic approaches, although the frequency of inferred duplication events was sensitive to the model parameters of the analysis. Duplication has also been suggested in the *Haemoproteus* fauna of European *Sylvia* warblers based on analysis of biogeographic and phylogenetic patterns (Pérez-Tris et al. 2007).

The sparse evidence for co-speciation of avian haemosporidians and their hosts has been most often attributed to host-switching (Bensch et al. 2000, Ricklefs and Fallon 2002). Ricklefs et al. (2004) found that host-switching was the most prominent type of diversification event in haemosporidian phylogenies. Host-switching can be most readily inferred when closely related haemosporidian lineages infect distantly related host species, and this has been observed frequently (Bensch et al. 2000, Beadell et al. 2004, Fallon et al. 2005). Although such patterns provide strong evidence of host-switching, the frequency, phylogenetic distribution, and temporal occurrence of host-shifts during the evolution of avian haemosporidians remain largely unknown.

In addition to host-parasite evolutionary history, avian haemosporidians are affected by the external environment, particularly during development in ectothermic vectors (Valkiūnas 2005). The effect of the external environment on haemosporidian distributions has been most thoroughly studied in the genus *Plasmodium*, where both avian and mammalian parasites have been shown to be sensitive to climatic variables, especially temperature and precipitation (Rees 1994, LaPointe et al. 2010, Sehgal et al. 2011, Loiseau et al. 2012). However, despite the clear potential for the environment to structure haemosporidian distributions, the nature of this interaction is unknown for the vast majority of haemosporidian lineages. This is due in part to the fact that few studies have surveyed avian haemosporidians across climatically heterogeneous regions.

To assess the roles of host-parasite evolutionary history and current environmental niche on haemosporidian diversity and distribution, we studied the haemosporidian fauna of a host species that shows both intraspecific divergence and broad ecological tolerance, the house wren (*Troglodytes aedon* Vieillot 1809; *sensu*

Remsen et al. 2013). The house wren includes substantial intraspecific diversity, with 31 subspecies currently recognized (Dickinson 2003). Fifteen of these subspecies are found in South America, and in Peru alone there are at least five resident subspecies that are distributed nearly continuously across the heterogeneous Andean landscape, from sea level to as high as ~4,750 m elevation (Chapman and Griscom 1924, Fjeldså and Krabbe 1990, Schulenberg et al. 2007, Benham et al. 2011). These subspecies are thought to have diversified within the last ~3 million years following the formation of the Panamanian land bridge and subsequent colonization of South America by the house wren (Smith and Klicka 2010).

Haemosporidian parasites are expected to exhibit different biogeographic and phylogenetic patterns depending on their evolutionary history with the house wren. If co-speciation has been important during the evolutionary history of the haemosporidian assemblage of Andean house wrens, we expect to find the following patterns within haemosporidian genera: (1) haemosporidian clades that are exclusive to house wren clades and monophyletic with respect to the regional haemosporidian fauna; (2) co-phylogenetic structure between haemosporidian lineages and the house wren clades that they infect; and (3) contemporaneous diversification between house wrens and their haemosporidians. If within-host duplication is an important mechanism, we expect to find haemosporidian clades that are exclusive to the house wren but exhibit no co-phylogenetic structure, and are no older than the host clades they infect. In contrast, non-host-specific processes such as host-switch diversification would lead to house wren haemosporidians that: (1) are polyphyletic with respect to the parasites infecting other avian host species; and (2) comprise a monophyletic group along with other

haemosporidian lineages from the same region. Finally, if environmental factors influence the distribution of avian haemosporidians across the complex landscape of the Peruvian Andes, we should find geographic structure of haemosporidians across environments that have distinct bioclimatic profiles.

The primary objectives of this study were: (1) to describe the diversity and abundance of avian haemosporidians that infect an intraspecifically diverse host species across the environmentally heterogeneous Peruvian Andes; (2) to evaluate the extent to which phylogenetic and biogeographic patterns are consistent with contemporaneous diversification of interacting haemosporidian and house wren clades (co-diversification); and, (3) if there is temporal evidence of co-diversification, to ask whether phylogenetic patterns of host associations are consistent with host-specific processes such as co-speciation or within-host duplication or non-host specific processes such as host-switching within regional avifaunas.

MATERIAL AND METHODS

DNA extraction, PCR amplification, and sequencing

We extracted host and parasite DNA from frozen muscle and liver tissue from 140 house wren specimens from the Andes and adjacent lowlands (elevational range: 129 –4,454 m; Museum of Southwestern Biology specimens; Appendix A). All tissue samples were collected by flash-freezing in liquid nitrogen within a few hours of death, and were subsequently stored at -80° C. DNA extraction was conducted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Polymerase chain reaction (PCR) was used to amplify 1,041

base pairs (bp) of the house wren mitochondrial gene *ND2* using the primers L5219 and H6313 (Sorenson et al. 1999). Each 15 µl PCR reaction contained 1 µl template DNA, 1X PCR buffer (Invitrogen), 2.5 mM $MgCl_2$, 0.5 µm of each primer, and 0.75 units of Taq DNA polymerase. The thermal profile consisted of an initial 8 minute step at 94 °C, followed by 35 cycles of 94° C for 30 sec, 50° C for 30 sec, and 72° C for 45 sec with a final 10 minute extension at 72° C. All host samples successfully amplified, thereby verifying that the parasite screen would not be subject to false negatives due to poor template quality.

We used three nested PCR protocols that target a 478 bp fragment of the haemosporidian mitochondrial gene cytochrome b (*cytb*) to screen all samples for *Plasmodium*, *Haemoproteus*, and *Leucocytozoon* infection. PCR protocols were performed according to Hellgren et al. 2004 and Waldenström et al. 2004. All samples were screened for *Haemoproteus* and *Plasmodium* using the outer primer pairs HaemNFI/HaemNR3 and HaemNF/HaemNR2 with the nested primer pair HaemF/HaemR2. All samples were also screened for *Leucocytozoon* using the outer primer pair HaemNFI /HaemNR3 and the nested primer pair HaemFL/HaemR2L. The only difference between our PCR protocol and the original references was an initial 8 minute denaturation step at 94 °C.

Positive and negative controls were included in each PCR to check for successful amplification and contamination, respectively. No contamination was detected during the course of the study. All PCR products were stained with SYBR Safe (Invitrogen) and visualized on a 2% agarose gel to check for the presence of a PCR product of the expected length. All successful amplifications were purified using ExoSap-IT (USB,

Cleveland, Ohio) and sequenced in both directions using dye terminator cycle sequencing (BigDye, ABI) on an ABI 3130 automated sequencer (Applied Biosystems, Foster City California).

Microscopic examination

Blood smears were air dried in the field and stained with Giemsa according to the protocol described in Valkiūnas (2005), except blood smears were not fixed with methanol and were stained one or more years after they were made. We used the microscopy protocol of Valkiūnas et al. (2008) to search for blood parasites in the 26 specimens for which adequate blood smears were available. We scanned the entire smear at 1,000X for the seven of the 26 specimens that we found to be PCR positive.

Phylogenetic, genetic diversity, and dating analyses

We edited and aligned host and parasite sequences using Sequencher 4.7 (Gene Codes Corporation, Anne Arbor, Michigan). We calculated pairwise and between group p-distances for each alignment in MEGA 5.05 (Tamura et al. 2011). To determine the optimal model of nucleotide substitution for each alignment, we used jModelTest2 (Darriba et al. 2012). The GTR + I + Γ model was selected for both host and parasite alignments. Host and parasite phylogenies were estimated using Bayesian inference implemented in MrBayes 3.2.1 (Ronquist and Huelsenbeck 2003) and maximum likelihood implemented in PhyML (Guindon and Gascuel, 2003). All likelihood analyses utilized 100 bootstrap replicates to assess nodal support. For the house wren *ND2* dataset we conducted two simultaneous runs in MrBayes for 10 million generations, sampling

every 1,000 generations and using a burn-in of 10% as determined in TRACER v1.5 (Rambaut and Drummond 2007). The winter wren (*Troglodytes hiemalis,* GenBank accession no. AY460280.1) was set as the outgroup. To characterize the genetic diversity of the house wren, we calculated the number of haplotypes (*h*), haplotype diversity (*hd*), and nucleotide diversity (π) for each major clade using DnaSP v5.10.1 (Librado and Rozas 2009). To assess genetic differentiation between house wren populations within clades, we estimated F_{st} values in DnaSP and used 1,000 permutations to evaluate statistical significance.

To date the divergence times of house wren clades, we used the program BEAST 1.7.2 (Drummond et al. 2012). We tested for a strict clock model using likelihood-ratio tests and visual inspection of the distribution of the ucld.stdev parameter in TRACER. We implemented a strict clock model, Yule speciation prior, and GTR + I + Γ substitution model. The substitution rate was estimated from a normal distribution with a mean of 0.0125 substitutions/site/lineage/million years (s/s/l/my) and standard deviation of 0.0025, which corresponds to a divergence rate of 2.5% per million years (sd = 0.5%). This distribution reflects several studies that suggest that the divergence rate for avian *ND2* is faster than the widely used 2.1% per million years estimated from *cytb* (Smith and Klicka 2010, Lerner et al. 2011), and accounts for some uncertainty in this estimate. We ran this analysis for 10 million generations, sampling every 1,000 and utilizing a burn-in of 10%. We used TRACER to confirm that the effective sample size for all parameters was greater than 200.

For the haemosporidian *cytb* data we conducted two simultaneous runs in MrBayes for five million generations, sampling every 500 generations and implementing

a burn in of 10%. We set a mammalian haemosporidian parasite, *Plasmodium falciparum*, as the outgroup as justified based on the haemosporidian topology favored by Outlaw and Ricklefs (2011). We determined the genus of each parasite lineage by conducting BLAST searches within the MalAvi database (Bensch et al. 2009). To visualize genetic diversity within each haemosporidian genus, we constructed haplotype networks using statistical parsimony in TCS v.1.21 (Clement et al. 2000).

Defining parasite lineages

Species limits are poorly understood within avian haemosporidians, but speciation appears to have occurred rapidly relative to mtDNA divergence (Fallon et al. 2005). Previous research has demonstrated that avian haemosporidian *cytb* haplotypes that differ by less than 0.5% sequence divergence can undergo recombination at nuclear loci, but that haplotypes above this divergence threshold seem to be evolving independently (Bensch et al. 2004). In addition, intraspecific *cytb* variation is very limited in *Plasmodium* parasites that infect some lizard populations (Schall and St. Denis, 2010). Accordingly, we followed Ricklefs et al. (2011) in grouping haplotypes into putatively independent evolutionary lineages if they differed by ≤0.2% (1 base pair) in the amplified sequence and had similar host associations and geographic ranges.

Regional haemosporidian analyses

Because the *Haemoproteus* fauna that we detected exhibited greater host and regional specialization than *Leucocytozoon* or *Plasmodium*, with lineages largely restricted to the house wren in Peru, we tested for monophyly and regional conservatism

of house wren *Haemoproteus* lineages with respect to the New World *Haemoproteus* fauna. To do this, we estimated a time-calibrated phylogeny of all *Haemoproteus* cytb haplotypes that have been recorded in North and South America that are available on GenBank and MalAvi. Including the novel lineages found in this study, the final alignment included 184 *Haemoproteus* lineages (Appendix B). We used BEAST for this analysis with the uncorrelated lognormal relaxed clock model, Yule speciation prior, and the GTR + I + Γ substitution model. The substitution rate was estimated from a normal distribution with a mean of 0.006 s/s/l/my and standard deviation of 0.001 (1.2% divergence per million years, sd = 0.2%, Ricklefs and Outlaw 2010). We ran this analysis for 50 million generations, sampling every 5,000 generations and discarding the first 10% as burn-in. We examined the output for all analyses in TRACER, and confirmed that the effective sample size for all parameters was greater than 200.

We tested for regional phylogenetic signal in this phylogeny of New World *Haemoproteus*. First, we assigned each lineage to a geographic zone based on sampling records: continental North America (NA), the Antilles (ANT), South America outside of the Andes (SA), the Andes (AND), or any observed combination of the previous four. We then used parsimony to calculate the number of steps required to trace this character over the maximum clade credibility tree from the BEAST analysis. We tested for significant phylogenetic signal by comparing this value to the null distribution of parsimony steps generated by randomizing the tips of the phylogeny 1,000 times in Mesquite (Maddison and Maddison 2011). In addition, to test whether house wren specific *Haemoproteus* lineages are more likely to be related to other Andean lineages than expected by chance, we used the probability approach of Ricklefs and Fallon (2002).

Briefly, we first calculated the probability of drawing a pair of lineages or clades from the Andes from the maximum clade credibility tree. We then used the binomial distribution to calculate the probability of observing *n* pairs from the Andes out of *k* attempts, with *n* equal to the number of Andean house wren lineages that were sister to an Andean lineage or clade and *k* equal to the total number of Andean house wren lineages in the tree. If an Andean house wren lineage was sister to a clade of lineages from multiple geographic regions, we estimated the ancestral state of that clade using maximum likelihood estimation in Mesquite (Maddison and Maddison 2011).

Global and local haemosporidian traits

The global geographic distribution and number of host species that a haemosporidian lineage can infect are important for the assessment of biogeographic and coevolutionary patterns. Accordingly, we searched the MalAvi database (Bensch et al. 2009) and GenBank for haplotypes matching each of the parasite lineages that we detected in house wrens. We used these records to determine whether the lineages we detected have been found outside of the Andes, and the total number of host species that each lineage has been found to infect (i.e., host breadth).

We defined two measures of haemosporidian abundance. We calculated prevalence as the number of infected hosts divided by the total number of screened hosts. In addition, we defined the combined infection rate as the number of infections detected divided by the total number of screened hosts. The combined infection rate is equivalent to the sum of the prevalences for each haemosporidian haplotype plus the prevalence of unsequenced (PCR positive) infections. We were able to distinguish between

Haemoproteus/*Plasmodium* and *Leucocytozoon* in unsequenced infections based on the primers used, as the primer pair FL/R2L only amplifies *Leucocytozoon*. The combined infection rate may be considered more informative than prevalence as it accounts for the total number of infections in a host population (including multiple infections within a single host), while prevalence only accounts for whether a host is infected or not.

To test whether haemosporidians were significantly associated with specific environmental variables, we classified the environmental attributes of each specimen locality using principal components analysis (PCA) on 19 temperature and precipitation variables from the WorldClim dataset (Hijmans et al. 2005) at 30-second resolution. The first three principal components of these variables were used as input for discriminant analysis of principal components (DAPC) as implemented in the R package 'adegenet' (Jombart 2008) to assign sampling sites to environmental clusters. Variables that were not normally distributed were log-transformed prior to analysis.

DAPC determined the best estimate to be three environmental clusters based on the Bayesian information criterion (BIC, Supplementary material Fig. A1). One cluster was composed of warm, dry sites (mean temperature = 18.1 C°, mean annual precipitation = 70 mm) largely below 1,000 m on the west slope, the second cluster was composed of warm, wet sites (mean temperature = 20.8 C°, mean annual precipitation = 1,418 mm) in the eastern lowlands and humid east slope, and the third cluster was composed of cold sites with moderate levels of precipitation (mean temperature = 7.7 C°, mean annual precipitation = 657 mm) predominantly above 3,000 m. We used chi-squared tests or Fisher's exact tests when the chi-square approximation was inappropriate to test for differences in haemosporidian infection rate among environmental clusters.

To test for associations among the number of host species a lineage has been found to infect, the number of house wren clades a lineage infected, and lineage prevalence within house wrens, we used non-parametric Spearman correlations in R version 2.15 (R Development Core Team, 2012). Prevalence was arcsine-square root transformed prior to statistical analysis. To determine the degree of phylogenetic signal in the elevational distribution of haemosporidian lineages, we calculated Blomberg's K and Pagel's λ using the R packages 'picante' (Kembel et al. 2010) and 'geiger' (Harmon et al. 2008).

RESULTS

House wren phylogeographic structure

Bayesian and maximum likelihood analysis of all house wren *ND2* sequence revealed seven well-supported clades that occurred allopatrically or parapatrically across the landscape (Fig. 1). The topologies were largely concordant between methods, differing only in the placement of clade 4. Average between-clade uncorrected p-distance ranged from 3.4 to 5.7%, while average within-clade p-distance was roughly an order of magnitude less, ranging from 0.2 to 0.4%. We found a total of 73 house wren ND2 haplotypes. Within clades that were adequately sampled, haplotype diversity ranged from 0.731 (clade 5) to 0.947 (clade 3), and nucleotide diversity ranged from 0.0015 (clade 5) to 0.0091 (clade 4; Table 1).

Haemosporidian diversity

We successfully sequenced 47 infections, from which we recovered 26 haplotypes. Sixteen of these haplotypes have been previously reported on MalAvi or GenBank. The remaining 10 haplotypes were novel (Appendix 5).

We identified three pairs of haplotypes that differed by one substitution (0.2%), and in each case both haplotypes demonstrated similar host affiliations and geographic range. Accordingly, we collapsed these three closely related pairs, resulting in 23 lineages that we used for diversity and biogeographic analyses (Supplementary material Fig. A2). Due to the possibility that these three haplotypes represented independently evolving units, we repeated pertinent analyses using all haplotypes; however, this did not qualitatively change any results. We found nine *Leucocytozoon* lineages (L1-L9), nine *Haemoproteus* lineages (H1-H9), and five *Plasmodium* lineages (P1-P5). Haemosporidian richness ranged from 1 to 11 lineages detected per house wren clade (Supplementary material Table A2).

The proportion of lineages that have been recorded in hosts other than the house wren differed among the three haemosporidian genera. Our database search revealed that all five detected *Plasmodium* lineages have been recorded in other hosts, while six of nine *Leucocytozoon* lineages and only two of the nine detected *Haemoproteus* lineages are known to infect other host species. Four of five *Plasmodium* lineages have been recorded from distant geographic areas, including two lineages (P2 and P5) that have been recorded in Alaska in Swainson's thrushes (*Catharus ustulatus*; Loiseau et al. 2012), a species that winters in Peru. One lineage of *Leucocytozoon* (L1) has previously been recorded in Chile (Merino et al. 2008), while none of the *Haemoproteus* lineages that we found have been recorded outside of Peru (Appendix C).

Host-parasite phylogeography and divergence date estimates

A visual inspection of the host clade associations for *Leucocytozoon* and *Plasmodium* revealed no evidence of co-phylogeographic structure between the house wren and haemosporidians (Fig. 2). Diversification in these genera was not investigated further because both were comprised largely or entirely of host-generalists and did not exhibit any phylogenetic structure with respect to the house wren. In contrast, *Haemoproteus* parasites exhibited marked host and geographic structure, with seven of nine lineages apparently restricted to the house wren in the Peruvian Andes (Fig. 2).

The large-scale phylogenetic analysis that included the seven host-specific *Haemoproteus* lineages (H1-H6 and H8) and 177 additional New World *Haemoproteus* lineages revealed that the house wren specific lineages are polyphyletic, as several nodes with posterior probabilities of 0.99-1.0 refute the hypothesis of monophyly (Fig. 3). The *Haemoproteus* phylogeny also revealed strong regional phylogenetic signal (Fig. 3). Parsimony reconstruction of region across the phylogeny required significantly fewer steps (58) than observed in the randomized distribution (mean=91.5, SD=3.17, lower bound=80, $p<0.0001$). A binomial test also found significant regional conservatism among the seven *Haemoproteus* lineages that we found to be specific to the house wren. Out of these seven lineages, four were sister to Andean lineages/clades and two were sister to South American lineages/clades found outside of the Andes. Given a probability of selecting a pair of Andean lineages or clades of 0.22, the binomial probability of obtaining four pairs from the Andes out of seven is $p=0.04$. At a broader scale, the probability of obtaining six pairs from South America out of seven is $p=0.02$. One house

wren lineage was sister to a clade of North American lineages that consisted primarily of parasites that infect red-eyed vireos (*Vireo olivaceus*) or Swainson's thrushes (*Catharus ustulatus*), two species that winter in the Andes.

Divergence date estimates indicated that the house wren began diversifying in Peru in the Late Pliocene or Early Pleistocene, with the initial divergence occurring 2.9 million years ago (mya, 95% HPD: 1.6-4.5 mya) and the most recent divergence occurring 1.5 mya (95% HPD: 0.7-2.4 mya, Fig. 4). Divergence estimates for the house wren specific *Haemoproteus* lineages were similar to those of the house wren, but were skewed towards the Late Pleistocene, with five of the seven lineages having mean divergence estimates younger than 1.6 mya (0.3-1.5 mya, 95% HPD 0.002-2.7 mya, Fig. 4).

Haemosporidian abundance and environment associations

We detected an overall haemosporidian prevalence of 30% (42/140) in our sample. Microscopic examination of 26 blood smears detected haemosporidian infections in only three (MSB:Para:19028, MSB:Para:19029, MSB:Para:19030; Appendix C). The three individuals that were positive by microscopy (MSB:Bird:33370, MSB:Bird:34763, MSB:Bird:36574) were also positive by PCR. Thirteen house wrens were infected by multiple haplotypes (identified through PCR amplification and sequencing using different nested primer pairs; multiple infections generally consisted of one *Haemoproteus*/*Plasmodium* infection and one *Leucocytozoon* infection), resulting in 57 total infections for a combined infection rate of 0.41 (57/140). We detected one mixed infection from the primer pair that amplifies *Leucocytozoon* that could not be resolved

and was thus not included in the phylogenetic or diversity analyses, but was included in the combined infection rate calculations.

There were no significant differences in infection rate between sex or between adult and immature birds (both $p > 0.05$). Overall haemosporidian infection rate was lower in warm dry environments than in cold montane or warm wet environments, and this difference was nearly significant ($\chi^2 = 5.52$, df = 2, $p = 0.06$, Supplementary material Table A3). *Leucocytozoon* infection rate differed among environments (Fisher's exact test, $p < 0.001$), as it was far more common in cold montane environments than in warm dry or warm wet environments. *Leucocytozoon* parasites were generally restricted to high elevations, as eight of nine lineages were only found at sites above 2,200 m, though we found no significant phylogenetic signal in lowest recorded elevation over all detected haemosporidian lineages ($K = 0.19$, $p = 0.11$; $\lambda = 0.10$, $p = 0.72$). There were no significant differences in infection rate across environments in *Haemoproteus* ($p = 0.08$) or *Plasmodium* ($p = 0.28$, Supplementary material Table A3).

The number of house wren clades that a haemosporidian lineage infected was strongly positively associated with its prevalence (Spearmans's rho = 0.83, $p < 0.001$). The number of host species a lineage has been found to infect varied from one (10 lineages) to over 30 (lineages L1, H9, and P1, Appendix C). All lineages detected outside of the house wren have been found to infect at least one other family outside of the Troglodytidae. The number of host species a lineage has been found to infect was strongly associated with its prevalence within house wrens (Spearmans's rho = 0.73, $p < 0.001$) as well as the number of house wren clades it infected (Spearmans's rho = 0.72, $p < 0.001$).

Parasite distributions across elevational gradients

Individual haemosporidian lineages exhibited latitudinal and elevational structure within house wren clade 1. This host clade is widespread on the west slope of central Peru and traverses a steep, west-east elevational gradient that spans sea level to >4,000 m. We found no significant host genetic structure between low (below 500 m) and high (above 3,000 m) elevations ($F_{st} = 0.006$, $p > 0.05$), or between populations from the north (Ancash) and south (Lima, $F_{st} = 0.05$, $p > 0.05$). Five haemosporidian lineages were detected more than once in this clade: L1, L5, H1, H9, and P4. Lineages P4 and L1 were found at sites in the Department of Ancash, while H1, H9, and L5 were found at sites in the neighboring Department of Lima. The elevational distributions of lineages did not overlap, and in each case all samples of a given lineage were found at the same sampling location or were found in close proximity within a narrow elevational band (Fig. 5). To test whether haemosporidians were significantly more structured than expected by chance, we determined the exact probability of observing this pattern at random. We randomly simulated 1,000 communities of the same size and composition as the one that we observed. This consisted of assigning 11 total infections from five different lineages to five elevational zones at random. The exact probability of observing complete elevational stratification (i.e. no elevational overlap among individuals of any pair of parasite lineages) was $p<0.0001$.

DISCUSSION

House wren genetic diversity and phylogeographic structure

We found deep phylogeographic structure within Peruvian house wrens, represented by seven genetically divergent clades that appear to have diversified during the Pleistocene. The distributions of at least three of these clades closely matched the distributions of named subspecies (Chapman and Griscom 1924, Bond 1956), even though most distribution boundaries among South American house wren subspecies have never been described in detail. The best-sampled clade (clade 1) was found on the west slope of central Peru and occurred across the entire elevational gradient from sea level to at least 4,123 m. The distribution of clade 1 corresponds closely to that of subspecies *T. a. audax*. A high elevation clade (clade 3) was detected between 1,500 m and 4,454 m in southern Peru. Clade 3 matches the distribution and elevational range of the large-bodied subspecies, *T. a. puna*, known from high elevation habitats of the Peruvian and Bolivian altiplano. Clade 6 was restricted to the dry western slope of southern Peru near the Chilean border, corresponding to the highly distinct, gray-plumaged subspecies *T. a. tecellatus*. The remaining four clades could not be clearly assigned to a single subspecies due to conflicting descriptions of house wren subspecies distributions in Peru (Chapman and Griscom 1924, Bond 1956).

Four of the seven house wren clades appeared to be bounded geographically by physical dispersal barriers or ecotones: clade 6 is isolated by river valleys and wide expanses of desert (Chapman and Griscom 1924), and three widespread clades replace one another across the main Andean ridge in central and south Peru (clade 1, clade 3, and clade 4). While the main Andean ridge comprises a phylogeographic break in many avian taxa that are restricted to low or middle elevations (Chaves et al. 2007, Nyári 2007,

Burney and Brumfield 2009, Milá et al. 2009), it is unclear how the house wren has evolved and maintained such deep mtDNA structure across the Peruvian Andes given its broad elevational tolerance and near-continuous distribution across the landscape. It is possible that local adaptation curtails dispersal of house wrens across ecotones (Hoekstra et al. 2005, McCracken et al. 2009), a hypothesis that is consistent with our observation that three parapatrically distributed clades (clades 1, 3, and 4) are associated with distinct environments (Supplementary material Table A4).

Haemosporidian diversity, host breadth, and global distribution

We detected one of the most diverse haemosporidian assemblages from a single host recorded to date, despite a relatively low number of sequenced infections (26 haplotypes were detected from 47 sequences). This level of diversity equals or exceeds that observed for other host species that are infected by diverse haemosporidian faunas, including the red-billed quelea (*Quelea quelea*, 34 haplotypes from 89 sequences, Durrant et al. 2007), American redstart (*Setophaga ruticilla*, 25 haplotypes from 59 sequences, Durrant et al. 2008), and the blackcap (*Sylvia atricapilla*, 27 haplotypes from 348 infections, Santiago-Alarcon et al. 2011). Increased sampling throughout the range of the house wren in Peru would be expected to detect dozens of additional haemosporidian lineages, particularly of the genera *Haemoproteus* and *Leucocytozoon* (Supplementary material Table A5). One possible explanation for the high diversity of haemosporidians that we detected is our broad sampling across the heterogeneous Andean landscape. Many previous surveys of single host species have been conducted across smaller geographic regions in comparatively homogenous environments that lack the dispersal

barriers and dramatic ecological gradients that are characteristic of the Andes (e.g. eastern United States, Fallon et al. 2006, Durrant et al. 2008; southwest Germany, Santiago-Alarcon et al. 2011).

The host range of haemosporidian lineages was positively associated with lineage prevalence and geographic distribution in Peruvian house wrens. These associations provide support for the observation that avian haemosporidians with the ability to infect multiple host species do not suffer a trade-off in the form of reduced transmission success (Hellgren et al. 2009, Szöllősi et al. 2011), as has been suggested by evolutionary theory (Futuyma and Moreno 1988). In addition, these patterns indicate that avian haemosporidians follow the abundance-distribution relationship described by Brown (1984), whereby the most abundant lineages are also the most widely distributed (Jenkins et al. 2011). This pattern could be driven by the fact that generalist parasites with strong immune evasion capabilities are able to successfully infect a higher proportion of the hosts that their vectors feed on, including migratory hosts with the ability to spread parasites to distant geographic regions (Hellgren et al. 2007). Alternatively, variation in host breadth could be an artifact of sparse sampling of rare lineages; hence, deep sampling will be required to fully understand the nature of the interaction between host specificity and parasite transmission.

Haemosporidian diversification

Given the high degree of host-generalism that we observed among *Leucocytozoon* and *Plasmodium* lineages, it is not surprising that we did not detect co-phylogenetic structure with the house wren in these genera. As more surveys are published with broad

community-level sampling, it is becoming apparent that strict specialization to a single host species is not as common among avian haemosporidians as previously believed (Fallon et al. 2005, Ricklefs et al. 2005, Lacorte et al. 2013, Fecchio et al. 2013). Many haemosporidian lineages infect multiple host species that are phylogenetically distant, a tendency that may be facilitated by generalist dipteran vectors that spread parasites widely throughout a community (Medeiros et al. 2013).

We did, however, detect host and geographic specialization in the genus *Haemoproteus*. Seven of the *Haemoproteus* lineages that we found appear to be specific to the house wren in Peru, and six of these lineages formed a clade in our restricted analysis of the 23 haemosporidian lineages found in this study (Fig. 2). However, a comprehensive analysis that included 184 *Haemoproteus* lineages from throughout North and South America indicated that these seven lineages are actually distantly related to each other. Thus, despite the deep genetic structure of house wren clades, there is no evidence for host-specific co-diversification of haemosporidian parasites. Rather, the phylogenetic and biogeographic patterns that we observed conformed with the predictions of host-switch diversification of *Haemoproteus* parasites. House wren specific *Haemoproteus* parasites were polyphyletic with respect to the regional *Haemoproteus* fauna, rather than monophyletic or paraphyletic as would be expected if co-speciation or within-host duplication had occurred. In fact, no pair of lineages with Troglodytidae hosts were monophyletic or even paraphyletic within the regional analysis. We also observed geographic conservatism of New World *Haemoproteus* parasites, indicating that diversification occurs within regions more than among regions. Most notably, we found that *Haemoproteus* lineages in Andean house wrens were related to

other Andean haemosporidian lineages more often than expected by chance. The latter finding supports the prediction that parasite host-shifts occur between host species that coexist within the same region (Ricklefs et al. 2004). Lastly, we found that the temporal diversification of house wren *Haemoproteus* parasites occurred over roughly the same time period as the diversification of the South American house wrens. The house wren is thought to have colonized South America from North America ~3 mya following the formation of the Panamanian land bridge (Smith and Klicka 2010); this scenario is compatible with our independently derived estimate of ~2.9 mya for the common ancestor of Peruvian house wrens. We estimated that six of the seven *Haemoproteus* lineages diverged within the last three million years, although most appear to be more recent than the house wren clades that they infect (Fig. 4). We acknowledge that there is substantial uncertainty in the *cytb* divergence rate of avian haemosporidians, most notably the possibility that the actual divergence rate may be slower than the 1.2% per million years used here (Bensch et al. 2013). However, even if the divergence rate was decreased by as much as 50%, five of seven house wren *Haemoproteus* lineages would still have speciated during the diversification of Peruvian house wrens.

The finding that host-switching is prominent in the diversification of house wren *Haemoproteus* is not surprising given recent meta-analyses indicating that symbiont diversification occurs most commonly by host-shift (de Vienne et al. 2013). Furthermore, host-switching is more likely to occur in parasites such as haemosporidians that use intermediate hosts (Nieberding and Olivieri 2007). The extensive host-switching of house wren *Haemoproteus* may have been driven in part by the Andean colonization of the

house wren, which created a new, abundant, parasite 'habitat' to be colonized via host-shifts.

The house wren and its *Haemoproteus* parasites appear to have undergone non-host-specific co-diversification because they diversified over the same time period and geographical area, but the parasites were not exclusive to a single host. This pattern conforms to what may be considered "diffuse" or community-level co-diversification, where interacting host and parasite clades diversify without evolutionarily stable host-specificity. In the case of the *Haemoproteus* fauna of the house wren, diversification appears to have been driven by host-switching within Andean bird communities. Diffuse co-diversification can be considered a macroevolutionary analog to diffuse coevolution, in which reciprocal natural selection occurs between entire assemblages of interacting hosts and symbionts (Janzen 1980).

Caveats with documenting host viability

Several lines of evidence support our assumption that the seven host-specific *Haemoproteus* lineages represent evolutionarily stable interactions. This is important because we were able to confirm only a small proportion of infections by microscopy and it is difficult to rule out the possibility that the lineages we detected may have been abortive infections that were unable to develop into gametocytes in the blood (Ferrell et al. 2007, Donovan et al. 2008, Olias et al. 2011). First, abortive infections have only been documented in exotic species that have been exposed to novel parasite faunas. Second, abortive infections appear to cause rapid morbidity and death in hosts, reducing the likelihood that a wild host suffering from an abortive infection would be sampled. Third,

abortive infections are associated with severe hemorrhagic damage to the internal organs. We inspected the internal anatomy of every bird we sampled, and no birds included in this study exhibited abnormal or damaged organs. Finally, previous studies have found that most abortive infections result from abundant, widespread lineages (Ferrell et al. 2007, Donovan et al. 2008), which contrasts with our observation that house wren *Haemoproteus* parasites appear to be rare.

Haemosporidian distributions across environments

Our study of avian haemosporidians across a highly complex landscape revealed geographic structure within and among genera. The most striking finding was the clear association of *Leucocytozoon* with high-altitude environments. *Leucocytozoon* was significantly more abundant at high elevations in the Andes than in adjacent lowlands, and eight of nine lineages were restricted to elevations above 2,200 m. This pattern was not driven by elevational sampling bias because nearly half of our samples were from below this elevation. This pattern agrees with a recent study of haemosporidian richness and prevalence across the Iberian Peninsula (Pérez-Rodríguez et al. 2013). Remarkably, two *Leucocytozoon* lineages (L1 and L5) were found at exceptionally high elevations (4,123 m and 4,384 m, respectively), as were single *Haemoproteus* (H9, 4,300 m) and *Plasmodium* lineages (P1, 4,454 m). Indeed, the major vectors for all three avian haemosporidian genera are known from >3,000 m in the Andes (e.g. Wygodzinsky and Coscarón 1979, de Rodriquez and Wirth 1986, Parker et al. 2012), suggesting that avian haemosporidians extend upslope to the limit of available bird habitat.

We observed evidence for environmental effects on the distributions of haemosporidians that infect house wrens found on the west slope of central Peru (clade 1), where five haemosporidian lineages exhibited a significantly non-random pattern of elevational replacement. It is possible that these lineages are tracking the abundance of their preferred vectors (Sol et al. 2000), avian hosts other than the house wren, or specific temperature (Rees 1994) or precipitation (Sehgal et al. 2011) regimes. These hypotheses need to be tested using replicate elevational transects because it remains possible that the apparent elevational pattern was an artifact of highly localized haemosporidian populations.

Conclusions

We demonstrated that a genetically structured host species, the house wren, harbors an unusually diverse haemosporidian fauna that displays marked variation in both host and environmental specificity. Evolutionary history and current ecological conditions both have clear roles for structuring the distributions of these parasites. Avian haemosporidians are geographically structured across the heterogeneous Peruvian landscape with respect to bioclimatic variables and elevation. Despite evidence for contemporaneous diversification of house wren mtDNA clades and host-specific *Haemoproteus* lineages within the Andes over the last three million years, these parasites have not co-diversified exclusively with house wren hosts. Rather, the house wren and its *Haemoproteus* parasites exhibit a pattern that is suggestive of diffuse co-diversification with the regional avifauna, with *Haemoproteus* speciation likely driven by host-switching events.

Chapter 1

REFERENCES

Beadell, J. S., Gering, E., Austin, J., Dumbacher, J. P., Peirce, M. A., Pratt, T. K., Atkinson, C.T., and Fleischer, R.C. 2004. Prevalence and differential host-specificity of two avian blood parasite genera in the Australo-Papuan region. – Mol. Ecol. 13: 3829–3844.

Beadell, J. S., Covas, R., Gebhard, C., Ishtiaq, F., Melo, M., Schmidt, B. K., Perkins, S.L., Graves, G.R., and Fleischer, R.C. 2009. Host associations and evolutionary relationships of avian blood parasites from West Africa. – Int. J. Parasitol. 39: 257–266.

Benham, P.M., Beckman, E.J., DuBay, S.G., Flores, L.M., Johnson, A.B., Lelevier, M.J., Schmitt, C.J., Wright, N.A., and Witt, C.C. 2011. Satellite imagery reveals new critical habitat for endangered bird species in the high Andes of Peru. Endang. Species Res. – 13: 145-157.

Bensch, S., Stjernman, M., Hasselquist, D., Örjan, Ö., Hannson, B., Westerdahl, H., Pinheiro, R.T. 2000. Host specificity in avian blood parasites: a study of *Plasmodium* and *Haemoproteus* mitochondrial DNA amplified from birds. – Proc. R. Soc. B 267: 1583-1589.

Bensch, S., Pérez-Tris, J., Waldenström, J., and Hellgren, O. 2004. Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites: multiple cases of cryptic speciation? – Evolution 58: 1617–1621.

Bensch, S., Hellgren, O., and Pérez-Tris, J. 2009. MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial

cytochrome b lineages. – Mol. Ecol. Resour. 9: 1353–1358.

Bensch, S., Hellgren, O., Križanauskienė, A., Palinauskas, V., Valkiūnas, G., Outlaw, D., and Ricklefs, R.E. 2013. How can we determine the molecular clock of malaria parasites? – Trends Parasitol. 29: 363-369.

Bond, J. 1956. Notes on Peruvian birds II. – P. Acad. Nat. Sci. Phila. 108: 227-247.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. Am. Nat. 124: 255–279.

Burney, C.W., and Brumfield, R.T. 2009. Ecology predicts levels of genetic differentiation in Neotropical birds. – Am. Nat. 174: 358-368.

Chapman, F. M., and Griscom, L. 1924. The house wrens of the genus *Troglodytes*. – Bull. Am. Mus. Nat. Hist. 50: 279-304.

Chaves, J. A., Pollinger, J. P., Smith, T. B., and LeBuhn, G. 2007. The role of geography and ecology in shaping the phylogeography of the speckled hummingbird (*Adelomyia melanogenys*) in Ecuador. – Mol. Phylogenet. Evol. 43: 795–807.

Clement, M., Posada, D., and Crandall, K. 2000. TCS: a computer program to estimate gene genealogies. – Mol. Ecol. 9: 1657-1660.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. – Nat. Meth. 9: 772.

de Rodriguez, M. C., and Wirth, W.W. 1986. A new species of man-biting *Culicoides* from the high Andes of Colombia (Diptera: Ceratopogonidae). – Fla. Entomol. 69: 311–314.

de Vienne, D.M., Refrégier, G., López-Villavicencio, M., Tellier, A., Hood, M.E., and Giraud, T. 2013. Cospeciation vs host-shift speciation: methods for testing,

evidence from natural associations and relation to coevolution. – New Phytol. 198: 347-385.

Dickinson, E.C. 2003. The Howard and Moore Complete Checklist of the birds of the world. – Princeton University Press, Princeton, N.J., USA.

Donovan, T.A., Schrenzel, M., Tucker, T.A., Pessier, A.P., and Stalis, I.H. 2008. Hepatic hemorrhage, hemocoelom, and sudden death due to *Haemoproteus* infection in passerine birds: eleven cases. J. Vet. Diagn. Invest. 20: 304-313.

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST1.7. – Mol. Biol. Evol. 29: 1969–1973.

Durrant, K.L., Reed, J.L., Jones, P.J., Dallimer, M., Cheke, R.A., McWilliam, A.N., and Fleischer, R.C. 2007. Variation in haematozoan parasitism at local and landscape levels in the red-billed quelea *Quelea quelea*. – J. Avian Biol. 38: 662-671.

Durrant, K. L., Marra, P. P., Fallon, S. M., Colbeck, G. J., Gibbs, H. L., Hobson, K. A., Norris, D.R., Bernik, B., Lloyd, V.L., and Fleischer, R.C. 2008. Parasite assemblages distinguish populations of a migratory passerine on its breeding grounds. – J. Zool. 274: 318–326.

Fallon, S. M., Bermingham, E., and Ricklefs, R. E. 2005. Host specialization and geographic localization of avian malaria parasites: a regional analysis in the Lesser Antilles. – Am. Nat. 165: 466–480.

Fallon, S. M., Fleischer, R. C., and Graves, G. R. 2006. Malarial parasites as geographical markers in migratory birds? – Biol. Lett. 2: 213–216.

Fecchio, A., Lima, M.R., Svensson-Coelho, M., Marini, M.Â., and Ricklefs, R.E. 2013.

Structure and organization of an avian haemosporidian assemblage in a Neotropical savanna in Brazil. – Parasitology 140: 181-192.

Ferrell, S.T., Snowden, K., Marlar, A.B., Garner, M., and Lung, N.P. 2007. Fatal hemoprotozoal infections in multiple avian species in a zoological park. – J. Zoo Wildlife Med. 38: 309-316.

Fjeldså, J., and Krabbe, N. 1990. Birds of the High Andes. – Zoological Museum, University of Copenhagen.

Futuyma, D. J., and Moreno, G. 1988. The evolution of ecological specialization. – Annu. Rev. Ecol. Syst. 19: 207–233.

Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. – Syst. Biol. 52: 696–704.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., and Challenger, W. 2008. GEIGER: investigating evolutionary radiations. – Bioinformatics 24: 129-131.

Hellgren, O., Waldenström, J., and Bensch, S. 2004. A new PCR assay for simultaneous studies of *Leucocytozoon*, *Plasmodium*, and *Haemoproteus* from avian blood. – J. Parasitol. 90: 797–802.

Hellgren, O., Waldenström, J., Pérez-Tris, J., Szöllősi, E., Hasselquist, D., Križanauskienė, A., Ottosson, U., and Bensch, S. 2007. Detecting shifts of transmission areas in avian blood parasites - a phylogenetic approach. – Mol. Ecol. 16: 1281–1290.

Hellgren, O., Pérez-Tris, J., and Bensch, S. 2009. A jack-of-all-trades and still a master of some: prevalence and host range in avian malaria and related blood parasites. – Ecology 90: 2840–2849.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.

Hoekstra, H.E., Krenz, J.G., and Nachman, M.W. 2005. Local adaptation in the rock pock mouse (*Chaetodipus intermedius*): natural selection and phylogenetic history of populations. – Heredity 94: 217-228.

Janzen, D.H. 1980. When is it coevolution? – Evolution 34: 611-612.

Jenkins, T., and Owens, I. P. F. 2011. Biogeography of avian blood parasites (*Leucocytozoon* spp.) in two resident hosts across Europe: phylogeographic structuring or the abundance-occupancy relationship? – Mol. Ecol. 20: 3910–3920.

Jenkins, T., Thomas, G. H. Hellgren, O., and Owens, I. P. F. 2012. Migratory behavior of birds affects their coevolutionary relationship with blood parasites. – Evolution 66: 740–751.

Johnson, M.T.J., and Stinchcombe, J.R. 2007. An emerging synthesis between community ecology and evolutionary biology. TREE 22: 250-257.

Johnson, K.P., Adams, R.J., Page, R.D., and Clayton, D.H. 2003. When do parasites fail to speciate in response to host speciation? – Syst. Biol. 52: 37-47.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. – Bioinformatics 24: 1403–1405.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., and Webb, C.O. 2010. Picante: R tools for integrating phylogenies and ecology. – Bioinformatics 26: 1463-1464.

Lacorte, G.A., Félix, G.M.F., Pinheiro, R.R.B., Chaves, A.V., Almeida-Neto, G., Neves, F.S., Leite, L.O., Santos, F.R., and Braga, É.M. 2013. Exploring the diversity and distribution of Neotropical avian malaria parasites – a molecular survey from Southeast Brazil. – PLoS ONE 8: e57770.

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

Lerner, H. R. L., Meyer, M., James, H. F., Hofreiter, M., and Fleischer, R. C. 2011. Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. – Curr. Biol. 21: 1–7.

Librado, P., and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. – Bioinformatics 25: 1451-1452.

Loiseau, C., Harrigan, R.J., Cornel, A.J., Guers, S.L., Dodge, M., Marzec, T., Carlson, J.S., Seppi, B., and Sehgal, R.N.M. 2012. First evidence and predictions of *Plasmodium* transmission in Alaskan bird populations. – PLoS ONE 7:e44729.

Maddison, W.P. and Maddison, D.R. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75, http://mesquiteproject.org.

Merino, S., Moreno, J., Vásquez, R.A., Martínez, J., Sánchez-Monsálvez, I., Estades, C.F., Ippi, S., Sabat, P., Rozzi, R., and McGehee, S. 2008. Hematozoa in forest birds from southern Chile: Latitudinal gradients in prevalence and parasite lineage richness. – Austral Ecol. 33: 329-340.

McCracken, K. G., Bulgarella, M., Johnson, K. P., Kuhner, M. K., Trucco, J., Valqui, T. H., Wilson, R.E., and Peters, J.L. 2009. Gene flow in the face of countervailing

selection: adaptation to high-altitude hypoxia in the βA hemoglobin subunit of yellow-billed pintails in the Andes. – Mol. Biol. Evol. 26: 815–827.

Medeiros, M.C.I., Hamer, G.L., and Ricklefs, R.E. 2013. Host compatibility rather than vector-host-encounter rate determines the host range of avian *Plasmodium* parasites. – Proc. R. Soc. B 280: 20122947.

Milá, B., Wayne, R. K., Fitze, P., and Smith, T. B. 2009. Divergence with gene flow and fine-scale phylogeographical structure in the wedge-billed woodcreeper, *Glyphorynchus spirurus*, a Neotropical rainforest bird. – Mol. Ecol. 18: 2979–2995.

Nieberding, C.M., and Olivieri, I. 2007. Parasites: proxies for host genealogy and ecology? – TREE 22: 156-165.

Nyári, Á. S. 2007. Phylogeographic patterns, molecular and vocal differentiation, and species limits in *Schiffornis turdina* (Aves). – Mol. Phylogenet. Evol. 44: 154–164.

Olias, P., Wegelin, M., Zenker, W., Freter, S., Gruber, A.D., and Klopfleisch, R. 2011. Avian malaria deaths in parrots, Europe. – Emerg. Infect. Dis. 17: 950-952.

Outlaw, D. C., and Ricklefs, R. E. 2011. Rerooting the evolutionary tree of malaria parasites. – Proc. Natl. Acad. Sci. U.S.A 108: 13183–13187.

Page, R.D.M. 2003. Tangled trees: Phylogeny, cospeciation and coevolution. – The University of Chicago Press, Chicago, I.L., USA

Palinauskas, V., Valkiūnas, G., Bolshakov, C.V., and Bensch, S. 2008. *Plasmodium relictum* (lineage P-SGS1): Effects on experimentally infected passerine birds. – Exp. Parasitol. 120: 372-380.

Palinauskas, V., Valkiūnas, G., Križanauskienė, A., Bensch, S., and Bolshakov, C.V. 2009. *Plasmodium relictum* (lineage P-SGS1): Further observation of effects on experimentally infected passeriform birds, with remarks on treatment with Malarone. – Exp Parasitol. 123: 134-139.

Parker, D. M., Zavortink, T. J., Billo, T. J., Valdez, U., and Edwards, J. S. 2012. Mosquitoes and other arthropod macro fauna associated with tank bromeliads in a Peruvian cloud forest. – J. Am. Mosq. Control Assoc. 28: 45–46.

Pérez-Rodríguez, A., Fernández-González, S., de la Hera, I., and Pérez-Tris, J. 2013. Finding the appropriate variables to model the distribution of vector-borne parasites with different environmental preferences: climate is not enough. – Glob. Change. Biol. 19: 3245-3253.

Pérez-Tris, J., Hellgren, O., and Križanauskienė, A. 2007. Within-Host Speciation of Malaria Parasites. – PLoS ONE 2: e235.

R Development Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rees, P.H. 1994. Highland malaria. – East Afr. Med. J. 71: 1.

Rambaut, A., and Drummond, A.J. 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer

Remsen, J.V., Jr., Cadena, C.D., Jaramillo, A., Nores, M., Pacheco, J.F., Perez-Eman, J., Robbins, M.B., Stiles, F.G., Stotz, D.F., and Zimmer, K.J. Version: 7 February 2013. A classification of the bird species of South America. American Ornithologists' Union.

Ricklefs, R. E., and Outlaw, D. C. 2010. A molecular clock for malaria parasites. –

Science 329: 226–229.

Ricklefs, R.E., and Fallon, S.M. 2002. Diversification and host switching in avian malaria parasites. Proc. R. Soc. B 269: 885-892.

Ricklefs, R.E., Fallon, S.M., and Bermingham, E. 2004. Evolutionary relationships, cospeciation, and host switching in avian malaria parasites. Syst. Biol. 53: 111-119.

Ricklefs, R. E., Swanson, B. L., Fallon, S. M., MartÍnez-AbraÍn, A., Scheuerlein, A., Gray, J., and Latta, S. C. 2005. Community relationships of avian malaria parasites in southern Missouri. – Ecol. Monogr. 75: 543–559.

Ricklefs, R.E., Dodge Gray, J., Latta, S.C., and Svensson-Coelho, M. 2011. Distribution anomalies in avian haemosporidian parasites in the southern Lesser Antilles. – J. Avian Biol. 42: 570-584.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. – Bioinformatics 19: 1572–1574.

Santiago-Alarcon, D., Bloch, R., Rolshausen, G., Schaefer, H. M., and Segelbacher, G. 2011. Prevalence, diversity, and interaction patterns of avian haemosporidians in a four-year study of blackcaps in a migratory divide. – Parasitology 138: 824–835.

Schall, J. J., and St. Denis, K. M. 2010. Lack of sequence variation of the mitochondrial cytochrome b gene from a malaria parasite, *Plasmodium mexicanum*. – J. Parasitol. 96: 827–829.

Schulenberg, T.S., Stotz, D.F., Lane, D.F., O'neill, J.P., and Parker, T.A. 2007. Birds of Peru: revised and updated edition. – Princeton University Press, Princeton, N.J.,

USA.

Sehgal, R. N. M., Buermann, W., Harrigan, R. J., Bonneaud, C., Loiseau, C., Chasar, A., Sepil, I., Valkiūnas, G., Iezhova, T., Saatchi, S., and Smith, T.B. 2011. Spatially explicit predictions of blood parasites in a widely distributed African rainforest bird. – Proc. R. Soc. B 278: 1025–1033.

Smith, B. T., and Klicka, J. 2010. The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. – Ecography 33: 333-342.

Sol, D., Jovani, R., and Torres, J. 2000. Geographical variation in blood parasites in feral pigeons: the role of vectors. – Ecography 23: 307–314.

Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T., and Mindell, D. P. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. – Mol. Phylogenet. Evol. 12: 105–114.

Szöllősi, E., Cichoń, M., Eens, M., Hasselquist, D., Kempenaers, B., Merino, S., Nilsson, J.-A, Rosivall, B., Rytkönen, S., Török, J., Wood, M.J., and Garamszegi, L.Z. 2011. Determinants of distribution and prevalence of avian malaria in blue tit populations across Europe: separating host and parasite effects. – J. Evol. Biol. 24: 2014–2024.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – Mol. Biol. Evol. 28: 2731-2739.

Valkiūnas, G. 2005. Avian malaria parasites and other haemosporida. – CRC Press, Boca

Raton, Florida, USA.

Valkiūnas, G., Iezhova, T. A., Križanauskienė, A., Palinauskas, V., Sehgal, R. N. M., and Bensch, S. 2008. A comparative analysis of microscopy and PCR-based detection methods for blood parasites. – J. Parasitol. 94: 1395–1401.

Waldenström, J., Bensch, S., Hasselquist, D., and Östman, Ö. 2004. A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. – J. Parasitol. 90: 191–194.

Wygodzinsky, P. W., and Coscarón, S. 1979. Description of a black fly of the subgenus *Simulium* (Pternaspatha) from the high Andes of Ecuador (Diptera, Simuliidae). Am. Mus. Novit. 2670: 1-9.

Chapter 1

FIGURES

Figure 1. (a) ML majority rule consensus phylogeny of house wrens in Peru based on 1,041 bp of the mitochondrial gene *ND2*. Clade numbers match Table 1 and Figure 2. Values at nodes indicate Bayesian posterior probability/ML bootstrap support (b) Distribution of house wren clades in Peru. Dots represent individual sampling locations and are colored according to the clade(s) present at that site; dots with multiple colors represent sampling locations where individuals from more than one clade were collected. Note that clade seven is represented by a white dot to improve visualization. Elevation is depicted in 1,000 m intervals according to the legend.

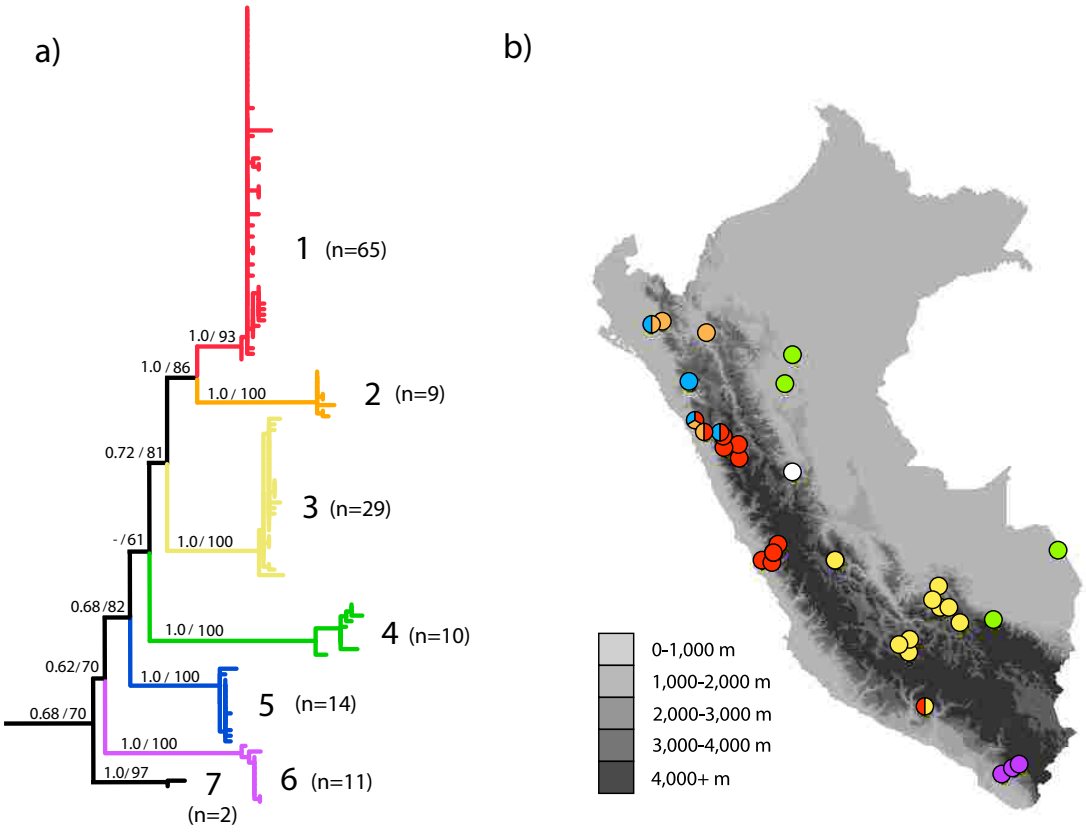


Figure 2. Bayesian phylogeny of 23 haemosporidian *cytb* haplotypes detected in the house wren with respect to host clade and the elevational range over which each lineage was encountered. The house wren phylogeny is shown with clades denoted by color and number as in Figure 1. Values at nodes indicate posterior probability/ML bootstrap support. Lineage names in blue represent parasites that have been detected in host species outside of the house wren; lineage names in red represent parasites that have only been found in house wrens in Peru. Numbers in boxes indicate the number of times a parasite was detected in a given host clade. Elevation is depicted over 1,000 m intervals from the west slope of the Andes (W) to the east slope (E).

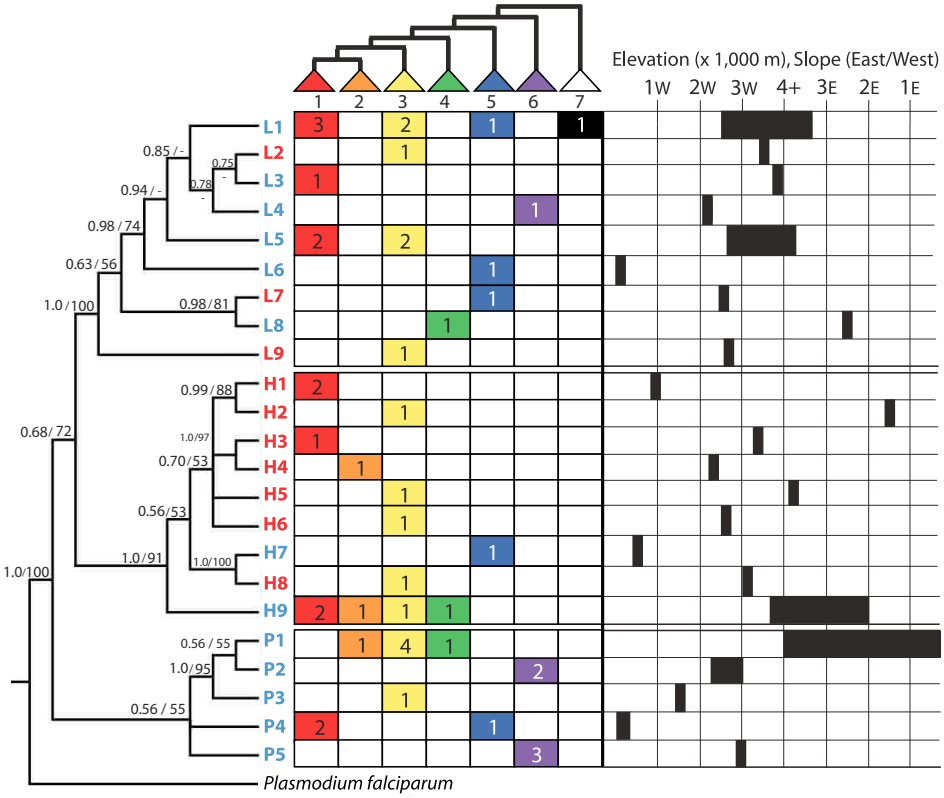


Figure 3. Regional *Haemoproteus* phylogeny for 184 (seven house wren specific lineages plus 177 previously found lineages) cytb lineages from North and South American estimated in BEAST. Branches are colored according to posterior probability support for their descendent node: blue (0.50-0.70), purple (0.71-0.94), red (0.95-1.0). Colored boxes at the tips of the phylogeny represent the geographic region in which a lineage has been found: yellow (Andes), green (South America outside of the Andes), white (North America), and blue (Antilles). Boxes with multiple colors represent lineages that have been found in multiple regions. Lineages specific to the house wren are indicated by asterisks.

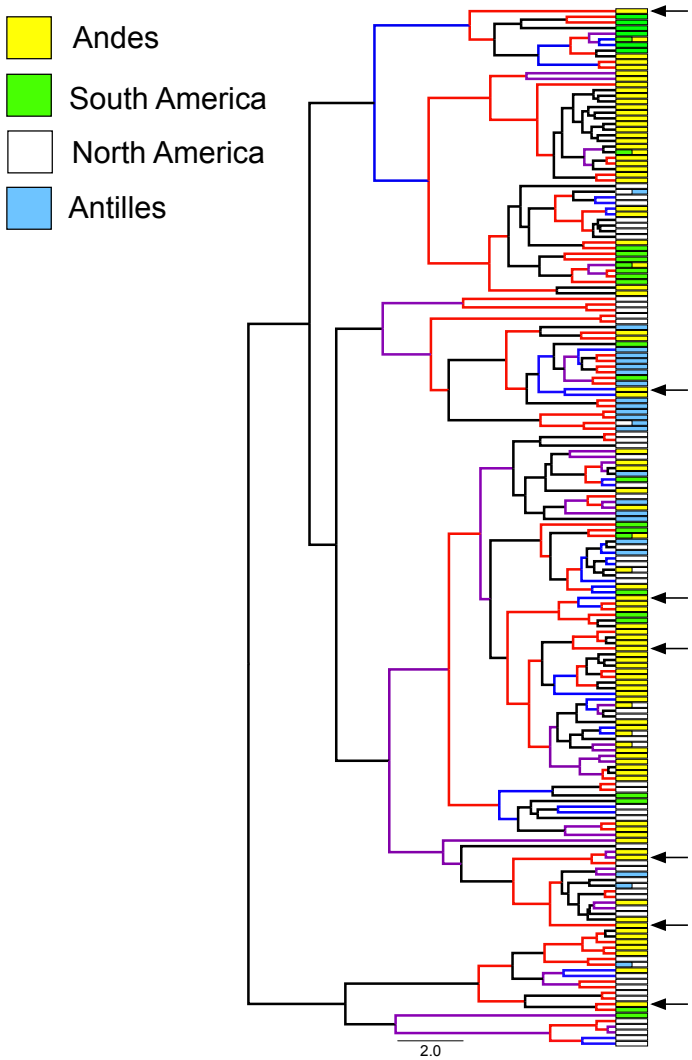


Figure 4. Lineage through time plots for the house wren and its host-specific *Haemoproteus* parasites. Shown in bold are the maximum clade credibility estimates; in gray are the 95% HPD confidence intervals. The estimated point of colonization of South America by the house wren is indicated with a dashed line.

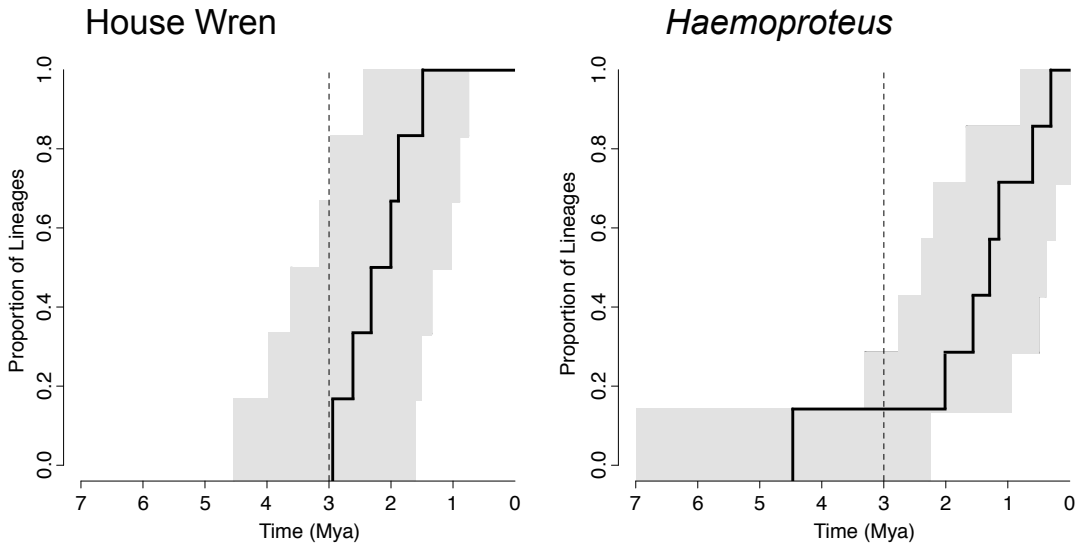
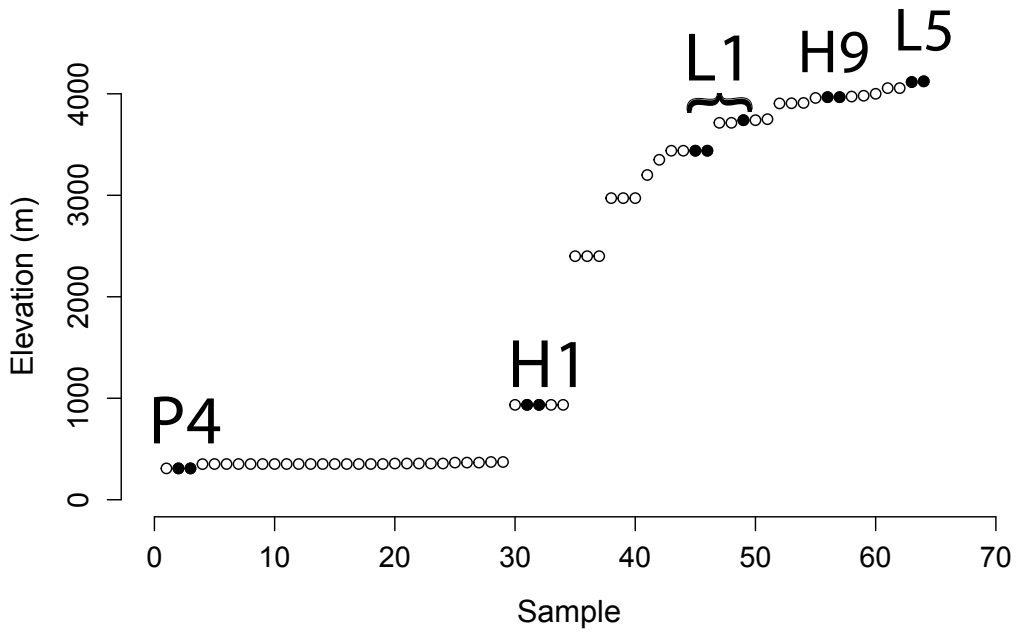
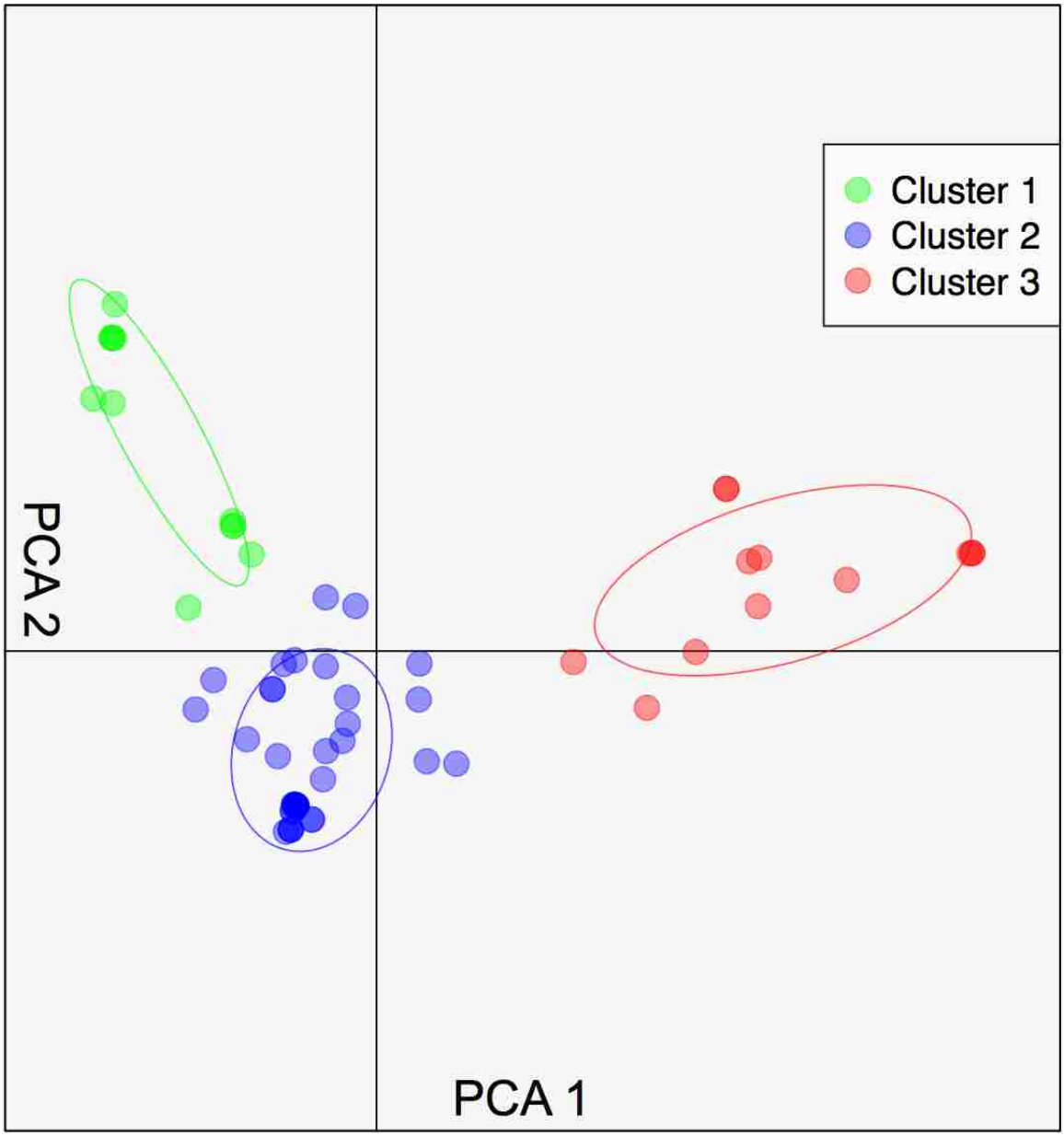


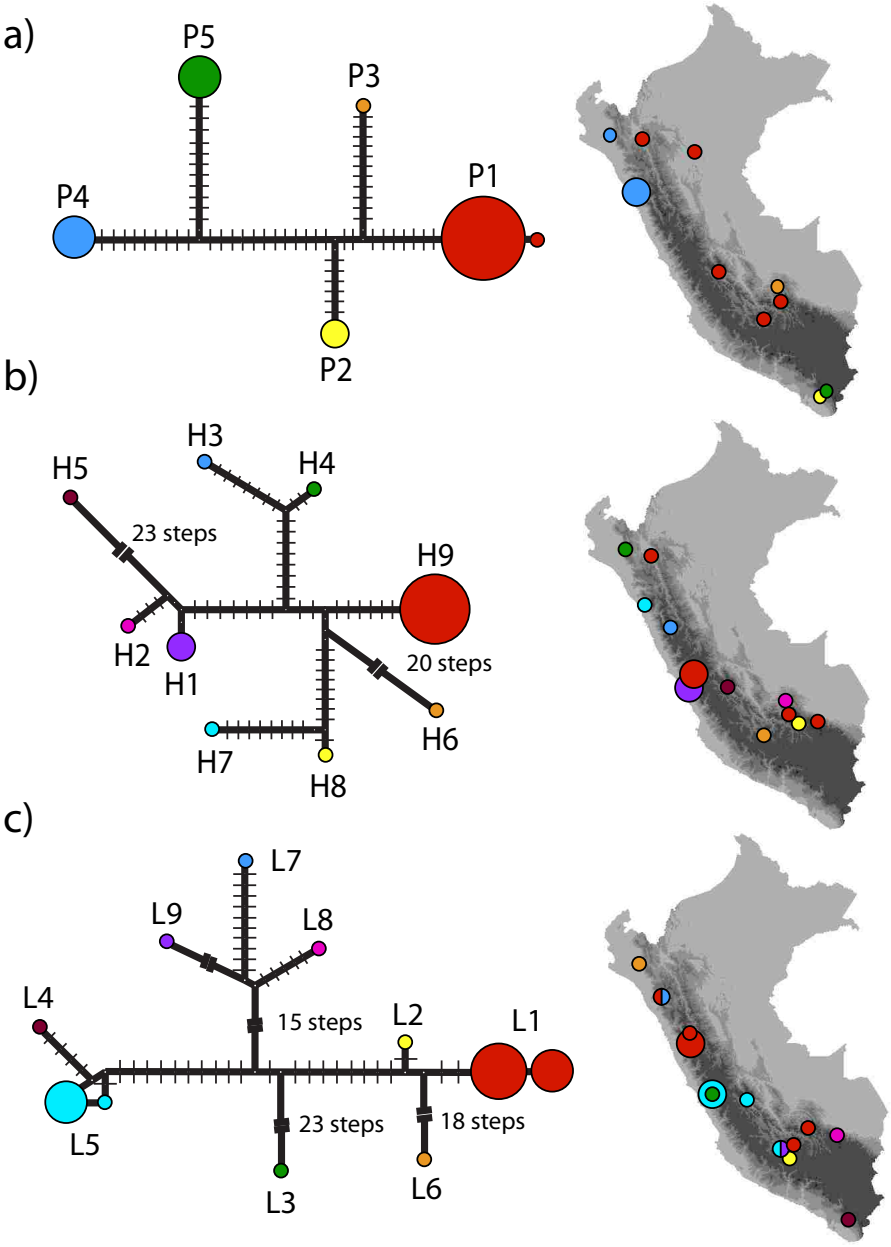
Figure 5. Distributions of five haemosporidian lineages across an elevational gradient within house wren clade 1 on the west slope of central Peru. Empty circles represent uninfected house wrens that were sampled across this gradient; black circles represent infections by the given lineage.



Supplementary Figure A1. PCA1 vs. PCA2 derived from 19 bioclimatic variables for each sampling site included in this study. The three distinct environmental clusters identified by DAPC are labeled by color: warm dry (red), cold montane (blue), and warm wet (green).



Supplementary Figure A2. Haplotype network and distribution of each haemosporidian genus based on samples in this study. The size of each circle indicates the number of times we detected a given haplotype, from once (smallest circle) to five times (largest circle). (a) *Plasmodium,* (b) *Haemoproteus*, (c) *Leucocytozoon*. The color schemes are specific to each genus and are intended solely for matching nodes on the haplotype network with the sampling map.



Chapter 1

TABLES

Table 1. Elevational range, genetic diversity, and haemosporidian infection rate for each of the major house wren mitochondrial DNA clades (clades numbered as in Fig. 1). The number of individual house wrens (N), number of house wren ND2 haplotypes (h), house wren haplotype diversity (hd), house wren nucleotide diversity (π), and haemosporidian infection rate (IR) in each clade are shown.

Clade	Elevational range (m)	N	*h*	*Hd*	π	IR
1	300 - 4,100	65	28	0.83 ± 0.05	0.0024	0.29
2	100 - 2,200	9	6	0.92 ± 0.07	0.0034	0.33
3	1,500 - 4,400	29	17	0.95 ± 0.02	0.0041	0.6
4	300 - 2,500	10	7	0.91 ± 0.08	0.0091	0.3
5	100 – 3,000	14	7	0.73 ± 0.13	0.0015	0.43
6	700 – 3,000	11	6	0.87 ± 0.07	0.0025	0.72
7	3,000	2	2	1.000 ± 0.50	0.0059	0.5

Supplementary Table A1. Glossary of terms used to describe evolutionary interactions among hosts and parasites in this paper. Definitions are adapted from the references cited in the table.

	Definition	**Reference**
Coevolution	The evolution of reciprocal adaptations in interacting hosts and parasites.	Janzen 1980, Page 2003
Diffuse coevolution	The evolution of reciprocal adaptations in one or more host species with one or more parasite species that are interacting non-exclusively in a community, and generate selective pressure as groups.	Janzen 1980, Johnson and Stinchcombe 2007
Co-diversification	Contemporaneous diversification among interacting host and parasite clades over macroevolutionary time, involving exclusive or non-exclusive interactions. When parasite lineages are host-specific, co-diversification is likely to involve co-speciation.	Johnson and Stinchcombe 2007
Co-speciation	Contemporaneous cladogenesis in a host lineage and its host-specific parasite due to the same underlying mechanism (e.g. joint isolation). Expected to frequently result in matched phylogenetic branching patterns, termed co-phylogeny.	Page 2003

Diffuse co-diversification	Contemporaneous diversification of host and parasite clades in which interactions are non-exclusive over macroevolutionary time, resulting from non-host-specific parasites and/or periodic host-shift events.	This study

Supplementary Table A2. Number of haemosporidian lineages and number of total infections (in parentheses) in each house wren clade. We used rarefaction in EcoSim (Gotelli and Entsminger 2012) to test for differences in haemosporidian lineage richness between host clades 1 and 3. Lineage richness was higher in clade 3 (average richness following rarefaction: 9.47, 95% CI: 8.0 – 11.0) than in clade 1 (7 lineages, $p < 0.001$).

	Leucocytozoon	*Haemoproteus*	*Plasmodium*	Total
Clade 1	3 (6)	3 (5)	1 (2)	7 (13)
Clade 2	0	2 (2)	1 (1)	3 (3)
Clade 3	4 (6)	5 (5)	2 (5)	11 (16)
Clade 4	1 (1)	1 (1)	1 (1)	3 (3)
Clade 5	3 (3)	1 (1)	1 (1)	5 (5)
Clade 6	1 (1)	0	2 (5)	3 (6)
Clade 7	1 (1)	0	0	1 (1)

Supplementary Table A3. Haemosporidian infection rate in house wrens across environments in Peru. Environment descriptions refer to the distinct bioclimatic clusters identified using discriminant analysis of principal components (see methods). A chi-squared test was conducted for overall haemosporidian infection rate; Fisher's exact tests were used for comparisons within individual haemosporidian genera due to one or more low expected values in contingency tables. Because *Haemoproteus* and *Plasmodium* were amplified simultaneously, we could not determine the generic identity of *Haemoproteus*/*Plasmodium* infections that did not successfully sequence. As a result, the combined infection rates of each individual genus do not equal the overall infection rate.

	Warm, dry	Cold, wet	Warm, wet	χ^2	df	p-value
Overall	0.30	0.51	0.44	5.52	2	0.06
Leucocytozoon	0.03	0.26	0.05	-	-	< 0.001
Haemoproteus	0.05	0.11	0.22	-	-	0.08
Plasmodium	0.13	0.06	0.16	-	-	0.28

Supplementary Table A4. Environment associations of three house wren clades that are distributed parapatrically across the main Andean ridge in southern Peru. Environment descriptions represent the three bioclimatic clusters that we identified using PCA (see methods). Clade-environment associations are significantly non-random (Fisher's exact test $P < 0.0001$).

	Warm, Dry	Cold, Wet	Warm, Wet
Clade 1	37	28	0
Clade 3	0	27	2
Clade 4	0	0	10

Supplementary Table A5. Observed and Chao2 estimated richness of haemosporidian genera in this study. We used the program EstimateS v. 8.2 (Colwell 2005) to calculate the Chao2 estimate of species richness separately for *Leucocytozoon*, *Haemoproteus*, and *Plasmodium* in our sample. This statistic estimates the expected species richness of an assemblage given the number of haemosporidian species found only once or only twice in a sample

	Observed richness	Chao2 richness	SD
Leucocytozoon	9	22.1	12.7
Haemoproteus	9	32.5	31.1
Plasmodium	5	4.5	1.3

Chapter 2

Evidence for adaptive evolution of hemoglobin in response to high-altitude hypoxia in Andean house wrens

INTRODUCTION

High-elevation habitats are among the most physiologically challenging environments on earth. Organisms that live at high altitudes must overcome two closely associated, unremitting environmental pressures: low ambient temperatures and reduced partial pressure of oxygen (PO_2). Under hypoxic conditions, O_2 saturation of the arterial blood is reduced, triggering a systematic compensatory response to maintain O_2 delivery to respiring tissues (Storz et al. 2010). The reduced rate of O_2 diffusion across the lung-blood barrier severely limits the capacity for aerobic exercise, thereby exacerbating the stress of metabolic heat production brought on by cold temperatures (Hayes and O'Connor 1999). The strong selection pressures associated with high altitude environments have resulted in the evolution of adaptations that increase the flux of the O_2 transport pathway in many high-altitude vertebrates (Storz et al. 2010).

Potential targets of selection in hypoxic environments include the genes that encode the subunits of the major adult hemoglobin isoform: αA globin and βA globin (Perutz 1983, Weber 2007, Storz and Moriyama 2008). Adult hemoglobin is the principle blood-O_2 transport protein and is a tetramer composed of two α and two β subunits, each of which has a heme group that can bind to a single O_2 molecule. Hemoglobin functions by reversibly binding to O_2 in one of two structural states. In the tense (T) state, the hemoglobin structure has a low affinity for O_2 but a high affinity for allosteric effector

molecules such as H^+, Cl^-, CO_2, inositolpentaphosphate (IPP; found in birds), and 2,3-diphosphoglycerate (DPG; found in mammals). Allosteric effectors reduce hemoglobin-O_2 affinity by stabilizing the tense state structure through the formation of inter-subunit hydrogen bonds and salt bridges. In contrast, relaxed (R) state hemoglobin has a low affinity for allosteric effectors and a high affinity for O_2. Molecular changes to the structure of the hemoglobin molecule can therefore increase hemoglobin-O_2 affinity by: 1) decreasing the stability of the tense state, causing a shift of the allosteric equilibrium in favor of the high O_2 affinity relaxed state; or 2) reducing the sensitivity of the hemoglobin molecule to allosteric effectors (Storz et al. 2010).

The study of hemoglobin evolution has revealed that adaptive changes in hemoglobin-O_2 affinity tend to occur as the result of one or very few mutations of large effect (Perutz 1983, Storz and Moriyama 2008). In birds, comparative studies have found that the same or similar adaptive pathways have evolved multiple times in the hemoglobin of distantly related lineages, suggesting that the number of adaptive "solutions" is limited (Jessen et al. 1991, Weber et al. 1993, McCracken et al. 2009b, Projecto-Garcia et al. 2013). The observation that parallel evolution of specific amino acid substitutions has occurred in the hemoglobin of high altitude waterfowl (Jessen et al. 1991, Weber et al. 1993, McCracken et al. 2009b) and hummingbirds (Projecto-Garcia et al. 2013) strongly suggests that closely related organisms are predisposed to evolve the same hemoglobin adaptations when faced with similar selective pressures. However, hemoglobin has been studied in relatively few species of high-altitude vertebrates to date.

Numerous examples of adaptive hemoglobin evolution have been observed in species that are resident at high altitudes (Piccinini et al. 1990, Jessen et al. 1991, Weber

et al. 1993, Weber et al. 2002), but many species have broad elevational distributions over which temperature and hypoxia vary considerably. Organisms that are distributed broadly across elevational gradients are expected to experience spatially varying selection pressure in response to the environmental extremes of low and high altitudes. One consequence of spatially varying selection is that signatures of adaptation may not be equal throughout the genome; neutral loci not closely linked to those under selection are expected to display different patterns of genetic structure and gene flow relative to those experiencing strong selection. Theoretical (Wu 2001, Via 2009) as well as empirical (McCracken et al. 2009a, 2009c, Wilson et al. 2012) studies show that when selection is strong relative to migration, loci experiencing local selection pressure can exhibit greatly reduced rates of gene flow compared to neutral loci. Therefore, a targeted population genetic approach that is focused on loci likely to be experiencing spatially varying selection pressures has the potential to reveal evidence for local adaptation. This approach has been used recently to identify amino acid substitutions on the αA and βA hemoglobin subunits that may increase hemoglobin-O_2 affinity in high-altitude waterfowl (McCracken et al. 2009a, 2009b, 2009c, Wilson et al. 2012).

Andean populations of the house wren (*Troglodytes aedon*) are ideal for the study of hemoglobin evolution in response to spatially varying selection for hypoxia resistance, as this species occurs continuously across a steep elevational gradient from sea level to over 4,500 m (Chapman and Griscom 1924, Fjeldså and Krabbe 1990, Schulenberg et al. 2007, Benham et al. 2011). The house wren displays deep mitochondrial structure in the Andes, and in Peru alone the house wren is represented by seven divergent mitochondrial clades that are distributed allopatrically or parapatrically across the landscape (Galen and

Witt 2014). These clades appear to have diversified in the late Pliocene and early Pleistocene after the house wren colonized South America following the formation of the Panamanian land bridge (Smith and Klicka 2010, Galen and Witt 2014), indicating that Andean house wrens have been experiencing selection at high altitudes for a lengthy period of evolutionary time.

This study seeks to: 1) address whether the house wren exhibits population genetic signatures of adaptation to high altitude in the genes that comprise the major adult hemoglobin; 2) examine the patterns of hemoglobin evolution across the heterogeneous landscape of the Peruvian Andes and among previously identified mitochondrial clades in order to investigate how landscape features and phylogeographic history have shaped genetic variation in the αA and βA globin genes; and 3) determine whether the house wren exhibits elevationally-segregating amino acid polymorphism at sites previously implicated in high altitude adaptation in other vertebrates.

MATERIALS AND METHODS

Specimen collection and elevational sampling

I utilized 140 house wren specimens from the Andes and adjacent lowlands (elevational range: 129 –4,454 m; Museum of Southwestern Biology specimens; Appendix A). All tissue samples were collected by flash-freezing in liquid nitrogen within a few hours of death, and were subsequently stored at -80° C. Sampling was approximately even between low and high elevations (<2,000 m, n=62; >2,000 m, n=78; <1,000 m, n=59; >3,000 m, n=50). These samples encompass seven divergent

mitochondrial clades (Galen and Witt 2014), although sampling was most thorough in the clade found on the west slope of the central Peruvian Andes (clade 1, n=65; Fig. 1).

DNA extraction, PCR amplification, and sequencing

I extracted DNA from frozen tissue of the 140 house wren specimens. DNA extraction was conducted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. I used polymerase chain reaction (PCR) to amplify six loci (Table 1). I amplified three protein-coding genes: the mitochondrial gene ND2 and the two genes that comprise the major adult hemoglobin, α globin A and β globin A. I also amplified three introns: β globin ρ intron 2 (HBBρ), β globin H intron 2 (HBBH), and myoglobin intron 2 (Myo2). HBBρ and HBBH are paralogs of βA and are positioned 5' of βA on chromosome 1B of the zebra finch genome (UCSC Genome Browser; Kent et al. 2002). It is likely that the house wren maintains this gene order because the structure of the β globin gene family appears to be highly conserved in birds (Alev et al. 2009, Hoffmann et al. 2010). Myoglobin, in contrast, is found on chromosome 1A (unlinked from chromosome 1B) of the zebra finch genome. Primers and PCR conditions are described in Table 2. ND2, the three introns, and αA globin were amplified as a single fragment; however, due to the length of βA this locus was amplified in three overlapping fragments.

Negative controls were included in each PCR to check for contamination. No contamination was detected during the course of the study. All PCR products were stained with SYBR Safe (Invitrogen) and visualized on a 2% agarose gel to check for the presence of a PCR product of the expected length. All successful amplifications were

purified using ExoSap-IT (USB, Cleveland, Ohio) and sequenced in both directions using dye terminator cycle sequencing (BigDye, ABI) on an ABI 3130 automated sequencer (Applied Biosystems, Foster City California).

Population Genetic Analyses

I coded nuclear sequences that displayed double peaks of equal height as heterozygotes according to the IUPAC ambiguity code. For sequences with more than one heterozygous site, I used the Bayesian statistical program PHASE 2.1 (Stephens et al. 2001) as implemented in DNAsp 5.10.1 (Librado and Rozas 2009) to estimate the gametic phase of the component haplotypes using the default settings. I conducted subsequent analyses using both all haplotypes and only the subset of sequences that were resolved with ≥95% probability. No significant differences were found between the two datasets, and so I present only the larger dataset containing all estimated haplotypes.

I used DNAsp 5.10.1 to calculate the number of alleles, the number of segregating sites, nucleotide diversity, haplotype diversity, and Θ for all loci, and I used GenAlEx v.6.5 (Peakall and Smouse 2012) to calculate observed and expected heterozygosity and Hardy-Weinberg equilibrium for the nuclear loci. I calculated these population genetic parameters for the total sample and within separate elevational zones. There is no widely agreed upon elevation that separates "low" and "high" sites, and so I conducted analyses by comparing individuals from below 1,000 m and above 3,000 m because only a small proportion (15%) of our samples were from elevations between these points. To measure the degree of genetic differentiation between populations at different elevations, I calculated Φst and conventional Fst in Arlequin v.3.5 (Excoffier and Lischer 2010) and

used 1,000 permutations to assess statistical significance. I conducted genetic differentiation analyses using entire sequences for all loci, and also on smaller subsets of the αA and βA globin datasets that consisted of just the coding regions or just the sites where a nonsynonymous substitution was observed. Genetic differentiation analyses were also repeated for just those individuals that are in mitochondrial clade 1 (n=65) in order to isolate the effect of elevation from phylogeographic history. Within clade 1 I also calculated Φst and Fst between two low elevation (<1,000 m) sites that are separated by 430 km in the Departments of La Libertad and Lima.

Coalescent estimates of gene flow

I used two Bayesian coalescent programs to estimate migration rates across elevations for each locus. I restricted these analyses to just two populations so as to adhere to the assumptions of the models: individuals sampled below 1,000 m in clade 1 (n=34) and individuals sampled above 3,000 m within clade 1 (n=27). First, I used the program Migrate-N v.3.6.4 (Beerli and Felsenstein 2001, Beerli 2006) to estimate Θ (4Neμ for nuclear loci; Neμ for ND2) for each population and M (m/μ) in each direction between the two populations. I used Bayesian mode and ran preliminary analyses on each locus with default settings to identify appropriate priors for Θ and M. Using these priors I ran one long chain for 100,000,000 steps, recording values every 100 for a total 1,000,000 recorded steps. I discarded the first 100,000 steps as burn-in. I used static heating and four chains with values of 1.0, 1.5, 3.0, and 1,000,000.0 to explore the parameter space more efficiently. I repeated each analysis multiple times to assess

convergence among runs and confirmed that the effective sample size for all parameters was greater than 1,000.

I also used the program IMa (Hey and Nielsen 2007) to estimate Θ and migration rates between the same populations in mitochondrial clade 1. IMa assumes no intralocus recombination, and so I used the program IMgc (Woerner et al. 2007) to return the largest non-recombining block of sequence for each locus. I ran IMa for each locus separately, using appropriate inheritance scalars (mtDNA: 0.25, autosomal nuDNA: 1) and geometric heating with 20-35 chains and heating parameters of 0.8 (g1) and 0.9 (g2). I initially ran the analyses with flat, uniform priors to determine the range over which the entire posterior distribution was contained, and in subsequent analyses altered the upper bound of all parameters accordingly. I ran each analysis twice with different random number seeds, and always ran the program until the effective sample size for each parameter was greater than 100.

Due to the uncertainty regarding the mutation rate for our panel of loci, particularly for the coding sequences of the globin genes, I used the output from Migrate-N and IMa to multiply Θ (xNeμ, where x is the inheritance scalar) for each population by the mutation-scaled migration rate M (m/μ) to calculate the effective number of gene copies that entered each population per generation (xNem). I divided this value by the inheritance scalar (x) to calculate the effective number of migrants per generation.

RESULTS

Elevationally segregating amino acid substitutions detected at the αA and βA globin genes

I identified seven amino acid polymorphisms in the αA and βA globin genes with minor allele frequencies greater than 0.06 in the total sample (range: 0.06-0.47). Five amino acid substitutions were found in the αA globin gene: α(12)Ala⇒Leu, α(15)Gly⇒Cys, α(57)Ala⇒Gly, α(72)His⇒Asn, and α(73)Val⇒Ile. Two amino acid substitutions were found in the βA globin gene: β(55)Val⇒Ile and β(80)Ser⇒Gly. While the substitutions at α(12) and α(15) did not segregate with respect to elevation, the remaining five polymorphisms did demonstrate elevational structure. The three closely linked αA globin polymorphisms at positions α(57), α(72), and α(73) all exhibited very similar patterns of allele frequency change across elevations (Fig. 2). The minor alleles (α(57)Ala, α(72)His, and α(73)Val) were present at frequencies between 0.27-0.44 at elevations below 1,000 m and dropped to frequencies of 0.02-0.03 at elevations above 3,000 m. The major αA globin alleles (α(57)Gly, α(72)Asn, and α(73)Ile) were present at frequencies of 0.55-0.72 at elevations below 1,000 m and increased to frequencies of 0.96-0.97 at elevations above 3,000 m.

The β(55)Val allele was found at a frequency of 0.79 at elevations below 1,000 m and an allele frequency of 0.18 at elevations above 3,000 m. In contrast, the β(55)Ile allele was found at a frequency of 0.21 at elevations below 1,000 m and a frequency of 0.82 at elevations above 3,000 m (Fig. 2). The β(80)Ser allele was found at a high frequency in each elevational zone, reaching 1.0 at elevations below 1,000 m and 0.81 at elevations above 3,000 m, while the β(80)Gly allele was only found at elevations above 3,000 m where it reached an allele frequency of 0.19 in the total sample.

The elevationally segregating amino acid substitutions in hemoglobin demonstrated geographic structure with respect to the seven well-differentiated house

wren mitochondrial DNA (mtDNA) clades that occur in Peru (Table 3). Several substitutions were associated with specific mtDNA clades: β(80)Gly was exclusive to clade 3, and α(57)Ala and α(73)Val were found almost exclusively in clades 4 and 5. Three mtDNA clades (4, 6, 7) exhibited low allelic diversity at the five amino acid positions, with major allele frequencies of ≥0.90 at all five residues (Table 3).

Genetic diversity and population genetic structure across elevation

The number of alleles per locus varied from 19 (αA) to 78 (HBBH); the number of polymorphic sites and haplotype diversity exhibited similar patterns (Table 4). Tajima's D was nonsignificant for most loci and populations; however, βA and ND2 exhibited a significantly negative Tajima's D in the total population and within each elevational zone. Tajima's D was also significant for myo2 in house wrens sampled below 1,000 m. All globin loci exhibited heterozygote deficiency in at least one population: αA and βA deviated from Hardy-Weinberg equilibrium below 1,000 m, while HBBρ and HBBH deviated from equilibrium above 3,000 m. Myo2 was in Hardy-Weinberg equilibrium in all populations (Table 4).

Estimates of Fst and Φst were uniformly significant but variable across loci between <1,000 m and >3,000 m in the total sample (Fig. 3A). βA and HBBρ showed the highest estimates of Φst (0.28 and 0.31, respectively), though the estimate for HBBH was reduced (0.06). αA showed moderate differentiation (Φst=0.11), while myo2 exhibited very low, yet significant population structure (Φst=0.02). I did not calculate these metrics for ND2 for the total sample due to the deep mitochondrial phylogeographic structure in this species.

I repeated these analyses for only those individuals in clade 1 so as to remove the effects of phylogeographic history and focus on elevation. Within clade 1 the β globin loci again demonstrated high, significant differentiation between elevations, with Φst reaching a peak of 0.15 at HBBA and falling to 0.09 at HBBH and again to 0.05 at HBBρ (Fig. 3B). In contrast, the remaining loci (αA, myo2, and ND2) all showed low and non significant Φst. I detected no population genetic structure within clade 1 between two low elevation (below 1,000 m) sites that were separated by ~430 km in the Departments of Lima and La Libertad on the west coast of Peru (points A and B in Fig. 1), as Φst estimates were uniformly low (0.001-0.04) and non significant across all loci (Fig. 3C).

Migration rates inferred from coalescent analyses

Migrate-N and IMa produced similar estimates of migration rates in both magnitude and direction, although the estimates derived from Migrate-N tended to be slightly larger than those from IMa (Table 5). For all six loci, migration was greater downslope than upslope (Table 5). IMa estimated the migration rate downslope to be lowest for βA (0.44 migrants per generation), which was 2.85 times lower than the next highest rate (HBBH; 1.27), 5.6 times lower than the rate for myo2 (2.51), and 11.3 times lower than the migration rate for αA (5.07). Upslope migration rates from IMa were less than 2.8 migrants per generation for all loci, and the three linked β globin genes exhibited reduced migration rates with respect to the other nuclear loci. HBBH exhibited the lowest migration rate upslope (0.001 migrants/generation), followed by βA (0.013 migrants/generation) and HBBρ (0.335 migrants/generation). The upslope migration rate for βA was 39 times lower than the upslope rate for αA, and 219 times lower than the

upslope rate for myo2. Similarly, Migrate-N found upslope migration for βA to be 14 to 16 times lower than for αA and myoglobin, respectively, but similar in magnitude to HBBH and HBBρ. Migrate-N found that the upslope 95% interval overlapped with zero for HBBρ, HBBH, βA, and myo2, but downslope βA was the only locus with a 95% interval that overlapped with zero.

DISCUSSION

β(55)Ile exhibits signatures of being adaptive at high altitude

I found multiple lines of evidence that suggest that the substitution of valine for isoleucine at position 55 of the βA subunit is adaptive at high altitudes. Although I identified five amino acid substitutions (three on the αA subunit, two on the βA subunit) that showed moderate to large differences in allele frequencies between elevational zones, β(55) showed the most dramatic shift with the valine allele overrepresented at low elevations and isoleucine overrepresented at high elevations. The cline in β(55) allele frequencies was observed in the total sample as well as in the restricted sample of individuals from a single, well-sampled mitochondrial clade that occurs across the entire elevational gradient (clade 1). In contrast, three of the remaining four elevationally segregating substitutions (α(57), α(73), and β(80)) were restricted geographically and found primarily in only one or two mitochondrial clades within a single elevational zone. For instance, β(80)Gly was restricted to clade 3, which is composed almost exclusively of house wrens from high altitudes. Similarly, α(57)Ala and α(73)Val were found almost entirely within clades 4 and 5, which are primarily low elevation clades (Galen and Witt 2014). Therefore the available evidence suggests that the elevational segregation of

β(80), α(57), and α(73) can be attributed primarily to phylogeographic history, and not necessarily selection. α(72) was geographically more widespread than α(57) and α(73) but exhibited a much weaker trend in allele frequency variation across elevations than β(55) in both the total sample and in clade 1 only (Fig. 2).

βA and its closely linked paralogs also exhibited elevated genetic structure between low and high elevations relative to unlinked nuclear loci and mtDNA. This observation was supported by direct estimates of migration rates using coalescent methods, which indicated that βA has experienced the lowest effective migration rate of any locus that I sequenced. This reduction of gene flow at the βA locus strongly suggests that natural selection has acted to eliminate individuals with genotypes mismatched to their environment (i.e. low elevation genotypes at high elevation). The low migration rate of house wren βA globin adds to recent research demonstrating that spatially varying selection can reduce gene flow at loci experiencing strong selective pressure at high altitudes (Storz et al 2008, McCracken et al. 2009a, Storz et al. 2012, Wilson et al. 2012).

Estimates of gene flow strongly supported asymmetric elevational migration, as all five nuclear loci indicated that migration occurred primarily downslope. This pattern was particularly striking for βA, as downslope migration was 34 (IMa) to 57 (Migrate-N) times higher than upslope migration. Reduced upslope gene flow at the βA locus suggests that the effective migration rate of the lowland β(55)Val allele is severely constrained relative to the high altitude β(55)Ile allele, indicating that there is a dramatic difference in fitness between the two alleles when in the "wrong" environment. A sharply reduced migration rate for the lowland βA allele supports the hypothesis that the environmental stresses of high altitudes are a strong selective force for Andean house wrens. The

differential migration rates of the β(55)Val and β(55)Ile alleles may be driven in part by differences in fitness of homozygotes: whereas only three individuals that were homozygous for the low elevation valine allele were found above 3,000 m (as opposed to 10 heterozygotes), nine individuals homozygous for the high elevation isoleucine allele were found below 1,000 m (as opposed to 5 heterozygotes). This could be because heterozygotes and homozygotes for the high altitude allele are able to move up or downslope without severe physiological consequences, but individuals homozygous for the low altitude valine allele suffer from strongly reduced fitness at high altitude. The observation that individuals homozygous for the high-altitude allele are found at low elevations is somewhat surprising considering that in humans it is disadvantageous to have high O_2-affinity hemoglobin at low altitudes due to a reduced ability to unload O_2 at the tissues (humans; Wajcman and Galacteros 2005). However, birds with high O_2-affinity hemoglobin have been successfully raised at low elevations without any apparent reduction in fitness (Scott et al. 2011). Previous population genetic studies of birds have found evidence for both significantly reduced upslope (Wilson et al. 2012) and downslope (McCracken et al. 2009a) migration, and so further study of species with broad elevational distributions will be warranted to determine whether asymmetric migration in one specific direction is characteristic of organisms that are adapted to high altitude.

Parallel evolution of αA and βA globin amino acid substitutions

The evolution of mechanistically similar adaptations in response to the same selective pressure, commonly referred to as parallel evolution, is predicted to occur when

the number of adaptive pathways is limited. Parallel evolution is believed to play an important role in the evolution of hemoglobin due to constraints on the tertiary and quaternary structure of the hemoglobin molecule; structural constraint favors the evolution of a limited number of mutations at key positions as opposed to numerous mutations of small effect (Perutz 1983). The hypothesis that hemoglobin is likely to evolve in a parallel manner has been supported by comparative analyses of hemoglobin evolution in high-altitude birds. McCracken et al. (2009a) found that two substitutions on the αA subunit and five substitutions on the βA subunit have evolved multiple times in Andean waterfowl, including an Ala-⇒Thr substitution at position 77 of the αA subunit that has evolved independently in five lineages. Projecto-Garcia et al. (2013) also documented parallel evolution in Andean hummingbirds, where substitutions at positions 13 and 83 of the βA subunit have evolved independently at least 4 and 13 times, respectively.

Out of the five amino acid substitutions that I observed that showed elevationally segregating allele frequencies, three have been previously documented in other high-altitude species and implicated to varying degrees in influencing hemoglobin-O_2 affinity. Two substitutions identical to those that I observed on the αA chain, α(57)Ala⇒Gly and α(72)His⇒Asn, have been reported from the deer mouse *Peromyscus maniculatus* (Storz et al. 2007). Similar to the house wren, the α(57)Gly allele is associated with high-altitude *Peromyscus* populations but the exact functional effect of this substitution is unknown (Storz et al. 2007). In contrast, the α(72)Ala allele of *Peromyscus* is not strictly associated with high-altitude populations, but segregates on a duplicated α globin gene copy that is believed to result in an alternative hemoglobin isoform with increased O_2

affinity (Storz et al. 2007). The existence of multiple adult hemoglobin isoforms in *Peromyscus* is thought to help mice adapt by fine-tuning saturation of the arterial blood to the ambient partial pressure of O_2. Although it is unclear whether α(57)Ala⇒Gly and α(72)His⇒Asn are adaptive in the house wren, the presence of identical substitutions in *Peromyscus maniculatus* suggests that the functional effects of these mutations warrants further investigation.

In addition, the inference that β(55)Val⇒Ile increases the O_2 affinity of hemoglobin in house wrens is supported by the existence an amino acid substitution at that position in the Andean goose. The Andean goose exhibits a Leu⇒Ser mutation at position β(55) that has been experimentally verified to increase hemoglobin-O_2 affinity by eliminating an α_1-β_1 contact that destabilizes the tense state and shifts the allosteric equilibrium in favor of the O_2-binding relaxed state (Jessen et al. 1991, Weber et al. 1993). Although the Leu⇒Ser substitution in the Andean goose is actually a symplesiomorphy of the entire genus *Chloephaga* (McCracken et al. 2010), including species that occur at low and high altitudes, the experimental evidence demonstrates definitively that it is associated with a dramatic increase in hemoglobin-O_2 affinity. While it is unknown whether the β(55)Val⇒Ile substitution in the house wren also eliminates the α_1-β_1 contact, the population genetic signatures of spatially varying selection that I observed provide independent support for the hypothesis that β(55)Ile is adaptive at high altitudes.

Conclusions

I have demonstrated that a songbird with a broad elevational distribution exhibits population genetic signatures of spatially varying selection at the βA globin gene, and that an amino acid substitution at position 55 of this locus appears to be adaptive at high altitudes. This study adds to a growing body of work showing that vertebrate species that are resident at high altitudes experience strong selective pressure on the O_2 transport system, and that the genes of the adult hemoglobin are targets of this selective pressure. My findings also corroborate previous work that demonstrates that the effects of selection are not homogeneous throughout the genome. Lastly, this study provides evidence for a case of parallel evolution with another high-altitude bird, the Andean goose.

Chapter 2

REFERENCES

Alev, C., Shinmyozu, K., McIntyre, B.A., and Sheng, G. 2009. Genomic organization of zebra finch alpha and beta globin genes and their expression in primitive and definitive blood in comparison with globins in chicken. Dev. Genes Evol. 219: 353-360.

Benham, P.M., Beckman, E.J., DuBay, S.G., Flores, L.M., Johnson, A.B., Lelevier, M.J., Schmitt, C.J., Wright, N.A., and Witt, C.C. 2011. Satellite imagery reveals new critical habitat for endangered bird species in the high Andes of Peru. Endang. Species Res. – 13: 145-157.

Beerli, P. and Felsenstein, J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. USA 98: 4563-4568.

Beerli, P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22: 341-345.

Chapman, F. M., and Griscom, L. 1924. The house wrens of the genus *Troglodytes*. – Bull. Am. Mus. Nat. Hist. 50: 279-304.

Excoffier, L. and Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10: 564-567.

Fjeldså, J., and Krabbe, N. 1990. Birds of the High Andes. – Zoological Museum, University of Copenhagen.

Galen, S.C. and Witt, C.C. 2014. Diverse avian malaria and other haemosporidians in Andean house wrens: evidence for regional co-diversification by host-switching. Journal of Avian Biology, *in press*.

Hayes, J.P. and O'Connor, C.S. 1999. Natural selection on thermogenic capacity of high-altitude deer mice. Evolution 53: 1280-1287.

Hey, J. and Nielsen, R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. USA 104: 2785-2790.

Hoffmann, F.G., Storz, J.F., Gorr, T.A., and Opazo, J.C. 2010. Lineage-specific patterns of functional diversification in the α- and β-globin gene families of tetrapod vertebrates. Mol. Biol. Evol. 27: 1126-1138.

Jessen, T-H., Weber, R.E., Fermi, G., Tame, J., and Braunitzer, G. 1991. Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. Proceedings of the National Academy of Sciences 88: 6519-6522.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. 2002. The human genome browser at UCSC. Genome Res. 12: 996-1006.

Librado, P., and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

McCracken, K.G., Bulgarella, M., Johnson, K.P., Kuhner, M.K., Trucco, J., Valqui, T.H., Wilson, R.E., and Peters, J.L. 2009a. Gene flow in the face of countervailing selection: adaptation to high-altitude hypoxia in the βA hemoglobin subunit of

yellow-billed pintails in the Andes. Molecular Biology and Evolution 26: 815-827.

McCracken, K.G., Barger, C.P., Bulgarella, M., Johnson, K.P., Sonsthagen, S.A., Trucco, J., Valqui, T.H., Wilson, R.E., Winker, K., and Sorenson, M.D. 2009b. Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. Molecular Ecology 18: 3992-4005.

McCracken, K.G., Barger, C.P., Bulgarella, M., Johnson, K.P., Kuhner, M.K., Moore, A.V., Peters, J.L., Trucco, J., Valqui, T.H., Winker, K., and Wilson, R.E. 2009c. Signatures of high-altitude adaptation in the major hemoglobin of five species of Andean dabbling ducks. The American Naturalist 174: 631-650.

McCracken, K.G., Barger, C.P., and Sorenson, M.D. 2010. Phylogenetic and structural analysis of the HbA (αA/βA) and HbD (αD/βA) hemoglobin genes in two high-altitude waterfowl from the Himalayas and the Andes: Bar-headed goose (*Anser indicus*) and Andean goose (*Chloephaga melanoptera*). Molecular Phylogenetics and Evolution 56: 649-658.

Peakall, R. and Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28: 2537-2539.

Perutz, M.F. 1983. Species adaptation in a protein molecule. Mol. Biol. Evol. 1: 1-28.

Piccinini, M., Kleinschmidt, T., Jurgens, K.L., and Braunitzer, G. 1990. Primary structure and oxygen binding properties of the hemoglobin from guanaco (Lama guancoe, Tylopoda). Biol. Chem. 371: 641-648.

Projecto-Garcia, J., Natarajan, C., Moriyama, H., Weber, R.E., Fago, A., Cheviron, Z.A., Dudley, R., McGuire, J.A., Witt, C.C., and Storz, J.F. 2013. Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. Proc. Natl. Acad. Sci. USA 110: 20669-20674.

Scott, G.R., Schulte, P.M., Egginton, S., Scott, A.L.M., Richards, J.G., and Milsom, W.K. 2011. Molecular evolution of cytochrome c oxidase underlies high-altitude adaptation in the bar-headed goose. Mol. Biol. Evol. 28: 351-363.

Schulenberg, T.S., Stotz, D.F., Lane, D.F., O'neill, J.P., and Parker, T.A. 2007. Birds of Peru: revised and updated edition. – Princeton University Press, Princeton, N.J., USA.

Smith, B. T., and Klicka, J. 2010. The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. – Ecography 33: 333-342.

Stephens, M., Smith, N.J., and Donnelly, P. 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68: 978-989.

Storz, J.F., Scott, G.R., and Cheviron, Z.A. 2010. Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. J. Exp. Biol. 213: 4125-4136.

Storz, J.F., Natarajan, C., Cheviron, Z.A., Hoffmann, F.G., and Kelly, J.K. 2012. Altitudinal variation at duplicated β-globin genes in deer mice: effects of selection, recombination, and gene conversion. Genetics 190: 203-216.

Storz, J.F., and Kelly, J.K. 2008. Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. Genetics 180: 367-379.

Storz, J.F. and Moriyama, H. 2008. Mechanisms of hemoglobin adaptation to high altitude hypoxia. High Altitude Medicine and Biology 9: 148-157.

Storz, J.F., Sabatino, S.J., Hoffmann, F.G., Gering, E.J., Moriyama, H., Ferrand, N., Monteiro, B., and Nachman, M.W. 2007. The molecular basis of high-altitude adaptation in deer mice. PLoS Genetics 3: e45.

Via, S. 2009. Natural selection in action during speciation. Proc. Natl. Acad. Sci. USA 106: 9939-9946.

Wajcman, H. and Galacteros, F. 2005. Hemoglobins with high oxygen affinity leading to erythrocytosis. New variants and new concepts. Hemoglobin 29: 91-106.

Weber, R.E., Jessen, T.H., Malte, H., and Tame, J. 1993. Mutant hemoglobins (alpha 119-Ala and beta 55-Ser): functions related to high-altitude respiration in geese. J. Appl. Physiol. 75: 2646-2655.

Weber, R.E., Ostojic, H., Fago, A., Dewilde, S., Van Hauwaert, M., Moens, L., and Monge, C. 2002. Novel mechanism for high-altitude adaptation in hemoglobin of the Andean frog Telmatobius peruvianus. Am. J. Physiol. Regul. Integr. Comp. Physiol. 283: 1052-1060.

Weber, R.E. 2007. High-altitude adaptations in vertebrate hemoglobins. Resp. Physiol. Neuro. 158: 132-142.

Wilson, R.E., Peters, J.L., and McCracken, K.G. 2012. Genetic and phenotypic divergence between low- and high-altitude populations of two recently diverged cinnamon teal subspecies. Evolution 67: 170-184.

Woerner, A.E., Cox, M.P., and Hammer, H.F. 2007. Recombination-filtered genomic datasets by information maximization. Bioinformatics 23: 1851-1853.

Wu, C.I. 2001. The genic view of the process of speciation. J. Evol. Biol 14: 851-865.

Chapter 2

FIGURES

Figure 1. Map of sampling locations colored by mitochondrial clade and elevation. Sampling locations above 2,000 m are shown in the darker shade of the color used to denote the mitochondrial clade found at that site. Sample size for each clade is in parentheses. Points marked "A" and "B" are sampling locations referred to in the Methods ("Population genetic analyses").

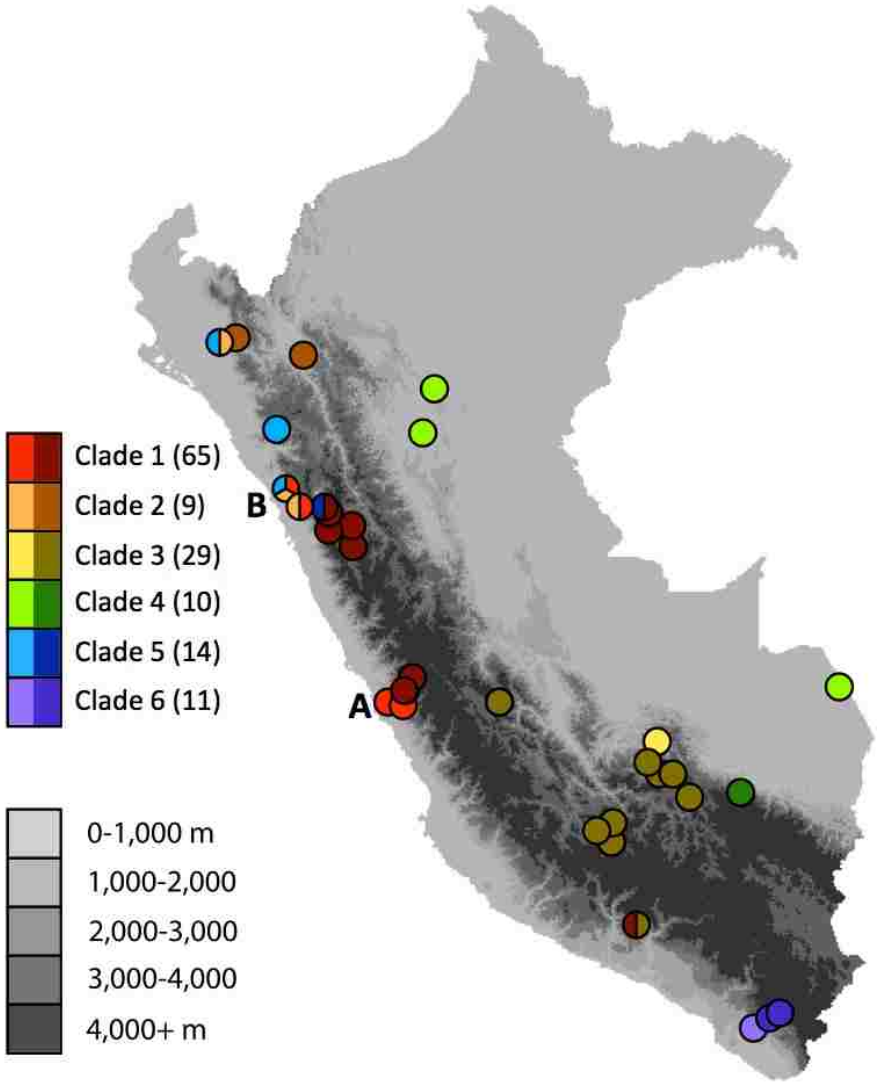


Figure 2. Allele frequencies for each amino acid position that varied between low (<1,000 m) and high (>3,000 m) elevations in the total sample and in clade 1 only. The amino acid alleles in black are: α(57)Ala, α(72)His, α(73)Val, β(55)Val, and β(80)Gly.

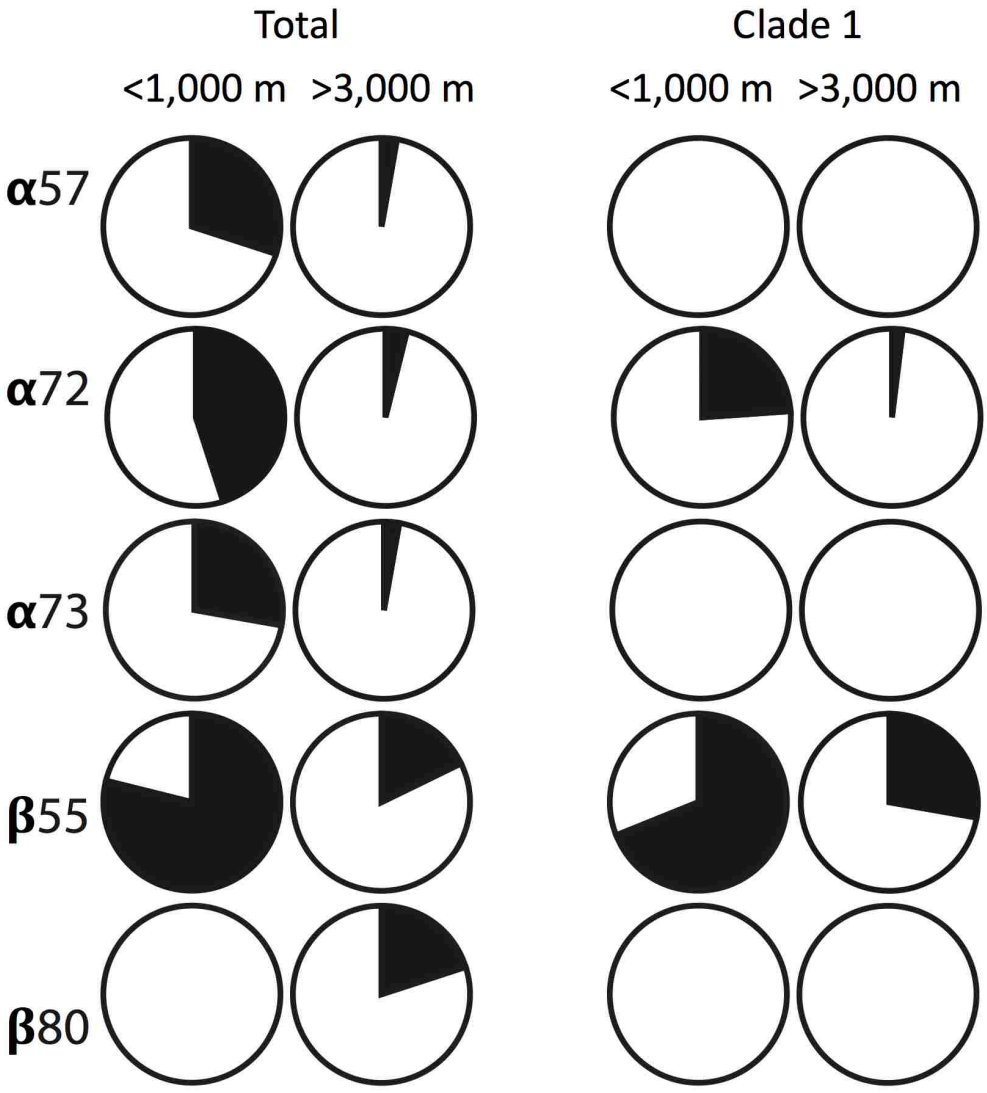


Figure 3. Estimates of genetic structure (Φst) across elevation (A,B) and latitude (C). A) Φst between low (<1,000 m) and high (>3,000 m) elevations in the total sample. B) Φst between low (<1,000 m) and high (>3,000 m) elevations in mitochondrial clade 1. C) Φst between two sites at similar elevations (both <1,000 m) on the west coat of Peru for individuals in clade 1 (see Figure 1 for locations of low elevation sites). β globin loci are shown in the order (5'-3') that they are found in the zebra finch genome.

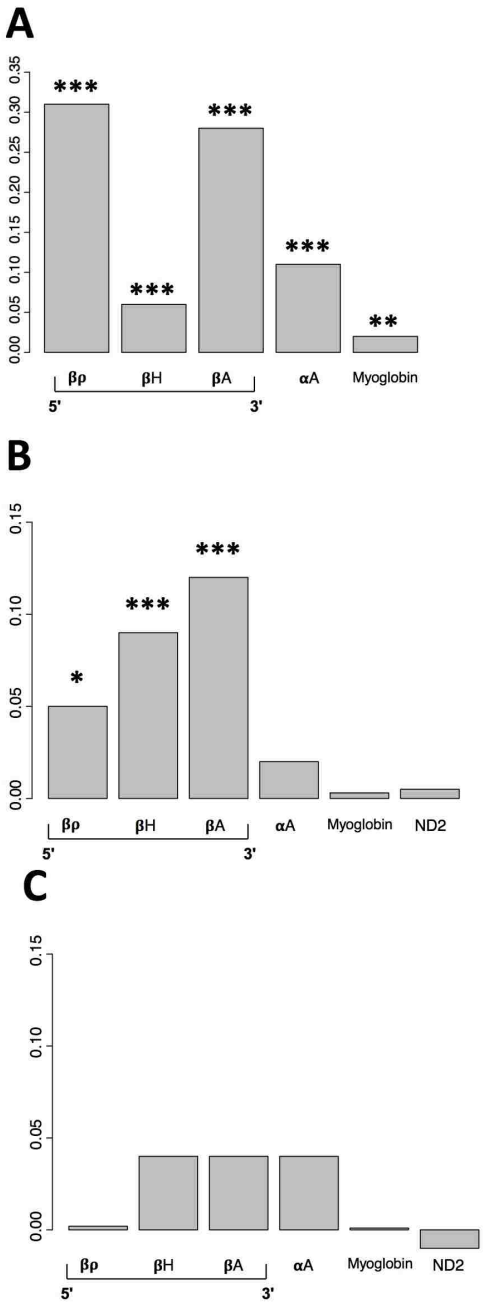


Table 1. Descriptions of the six loci sequenced in this study. Chromosome position based on the zebra finch genome, sequence length in base pairs (length without indels listed first, length with indels in parentheses) and best-fit model of nucleotide substitution used for population genetic analyses.

Locus	Chromosome	Base Pairs	Best-fit Model
HBBρ	1B	595	HKY + I
HBBH	1B	617 (632)	HKY + I + G
βA	1B	1310 (1,318)	-
βA (cds)		435	HKY + I
αA	14	671 (682)	-
αA (cds)		429	HKY + I
Myo2	1A	631	GTR + I + G
ND2 (clade 1 only)	mtDNA	1,041	TrN + I

Table 2. Genes (or gene fragments) amplified in this study, primers used, primer sequence, PCR protocol, and reference for primers (if applicable).

Locus	Primer name	Primer sequence	Cycles	Ann. temp. C°	Ext. time	Source
HBBρ	BetaRhoF	CTGCACTGCGASAAGCTG	40	60	1:30	This study
	BetaRhoR	GGGTGAASTCCTTGGTGAAG				
HBBH	BetaHF	GAACATGGACAGCGTCAAA	45	59	1:20	This study
	BetaHR	ACCACACGGACCAGCTTCT				
βA (exon1-exon2)	BetaAexon1F	AGTGGACAGCCGAGGAGAAGCAGCTC	35	71	1:00	This study
	BetaAexon2R	AGTGCAGCTCGGACAGCTGG GAGAAG				
βA (exon2-exon3)	BetaAexon2F	GACRGCATCAAGAACACCTTCTC	45	63	1:30	This study
	BetaAexon3R	GAGTGAAGTCCTTGCCGAAGTG				
βA (intron2-3' Flank)	BetaAint2F	AACTCAGACCTTTAGAGTAAGAGCAG	40	55	1:00	This study
	HBBT3.flank 1.ZF.R	GGCTTCACAGGATGAGCTTT				
αA	AlphaAF	GGACACCCGTGCTGGGGCG	35	71	1:30	This study
	AlphaAR	TCGCTGTCGGAGGGCTGGTG				
Myo2	E2F1	GAAGATCTGAAGAAACATGGAGCTA	35	59	1:00	DuBay and Witt 2012
	E3R1	CAATGACCTTGATAATGACTTCAGA				
ND2	L5219	CCCATACCCCGAAAATGATG	35	50	1:00	Sorenson et al. 1999
	H6313	CTCTTATTTAAGGCTTTGAAGGC				

Table 3. Allele frequencies for five elevationally segregating amino acid positions within each mitochondrial clade of the house wren. Sample sizes for clades: clade 1 (n=65), clade 2 (n=9), clade 3 (n=29), clade 4 (n=10), clade 5 (n=14), clade 6 (n=11), and clade 7 (n=2).

	α(57)A	**α(57)G**	**α(72)H**	**α(72)N**	**α(73)I**	**α(73)V**	**β(55)V**	**β(55)I**	**β(80)S**	**β(80)G**
Clade 1	0.00	1.00	0.15	0.85	1.00	0.00	0.53	0.47	1.00	0.00
Clade 2	0.14	0.86	0.36	0.64	0.86	0.14	0.72	0.28	1.00	0.00
Clade 3	0.06	0.94	0.06	0.94	0.94	0.06	0.06	0.94	0.67	0.33
Clade 4	0.90	0.10	0.90	0.10	0.10	0.90	0.90	0.10	1.00	0.00
Clade 5	0.54	0.46	0.54	0.46	0.54	0.46	0.73	0.27	1.00	0.00
Clade 6	0.05	0.95	0.05	0.95	0.95	0.05	1.00	0.00	1.00	0.00
Clade 7	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00

Table 4. Population genetic indices for each locus within the total sample and within low (<1,000 m) and high (>3,000 m) elevational zones. Number of individuals in the sample (N), number of segregating sites (*S*), number of alleles, haplotype diversity (*Hd*), nucleotide diversity (π), theta (Θ), Tajima's D, expected heterozygosity (H_e), and observed heterozygosity (H_o). Heterozygosity values in bold deviated significantly from Hardy-Weinberg equilibrium. Values for ND2 represent only individuals from mtDNA clade 1.

Locus	Elevation	N	*S*	No. Alleles	*Hd (Std. dev)*	Pi	Θ (S)	Tajima's D	H_e	H_o
αA (cds)	< 1,000 m	56	15	16	0.819 (0.023)	0.0076	0.0067	0.447 ns	**0.643**	**0.812**
	> 3,000 m	44	10	6	0.534 (0.048)	0.0045	0.0046	-0.027 ns	0.523	0.528
	Total	**130**	**18**	**19**	**0.678 (0.027)**	**0.0061**	**0.007**	**-0.255 ns**		
HBBρ	< 1,000 m	51	36	37	0.766 (0.046)	0.0055	0.012	-1.66 ns	0.784	0.758
	> 3,000 m	43	22	21	0.702 (0.048)	0.007	0.0075	-0.126 ns	**0.535**	**0.694**
	Total	**121**	**54**	**61**	**0.822 (0.020)**	**0.0077**	**0.0157**	**-1.45 ns**		
HBBH	< 1,000 m	38	42	37	0.974 (0.007)	0.0095	0.0144	-1.12 ns	**0.368**	**0.961**
	> 3,000 m	33	29	29	0.946 (0.013)	0.0061	0.0101	-1.20 ns	**0.545**	**0.932**
	Total	**88**	**61**	**78**	**0.974 (0.006)**	**0.0086**	**0.0181**	**-1.57 ns**		
βA (cds)	< 1,000 m	54	17	22	0.676 (0.049)	0.0022	0.0076	-1.94*	**0.463**	**0.662**
	> 3,000 m	44	16	16	0.718 (0.043)	0.0024	0.0074	-1.88*	0.591	0.709
	Total	**129**	**30**	**38**	**0.776 (0.019)**	**0.0028**	**0.0116**	**-2.07***		
Myo2	< 1,000 m	42	26	30	0.881 (0.026)	0.0054	0.0134	-1.83*	0.857	0.871
	> 3,000 m	40	18	23	0.928 (0.013)	0.0061	0.0093	-0.954 ns	0.780	0.918
	Total	**97**	**33**	**51**	**0.913 (0.012)**	**0.0058**	**0.0147**	**-1.72 ns**		
ND2	< 1,000 m	33	20	15	0.816 (0.066)	0.0023	0.0056	-2.01*	-	-
	> 3,000 m	24	24	14	0.833 (0.077)	0.0027	0.0075	-2.34**	-	-
	Total	**65**	**44**	**30**	**0.836 (0.048)**	**0.0024**	**0.0111**	**-2.52*** **	-	-

Table 5. Effective migrants per generation estimated from Migrate-N and IMa. The 95% HPD intervals are in parentheses.

	Migrate-N		**IMa**	
	Low⇒High	**High⇒Low**	**Low⇒High**	**High⇒Low**
HBBρ	0.496 (0-14.42)	32.5 (8.10-169.86)	0.335 (0.063-29.77)	17.1 (3.13-573.5)
HBBH	0.249 (0-4.61)	4.97 (1.01-19.08)	0.001 (0.0007-11.27)	1.27 (0.063-26.16)
βA (cds)	0.08 (0-61.2)	4.34 0-74.8)	0.013 (0.004-110.42)	0.445 (0.036-54.41)
αA (cds)	1.18 (0.034-17.27)	6.44 (0.581-52.1)	0.515 (0.025-113.7))	5.07 (0.316-189.06)
Myo2	1.27 (0-28.94)	15.3 (1.59-137.9)	2.849 (0.001-34.77)	2.51 (0.449-50.92)

APPENDIX A

House wren samples used in this thesis with Museum of Southwestern Biology catalog weblink and sampling details.

Catalog no. with embedded weblink	NK	Day	Month	Elevation (m)	Dept.	Lat.	Long.	Clade	Genbank accession no.
MSB:Bird:27052	159705	30	October	3040	Lima	-11.76	-76.58	1	KF420201
MSB:Bird:27596	162008	3	June	366	Lima	-12.01	-76.92	1	KF386228
MSB:Bird:27606	162022	3	June	366	Lima	-12.01	-76.92	1	KF386218
MSB:Bird:27609	162025	3	June	366	Lima	-12.01	-76.92	1	KF386232
MSB:Bird:31418	162982	8	January	372	Lima	-12	-76.92	1	KF386210
MSB:Bird:31425	162989	8	January	351	Lima	-12	-76.92	1	KF386216
MSB:Bird:31433	162997	8	January	372	Lima	-12	-76.92	1	KF386250
MSB:Bird:31450	163014	9	January	352	Lima	-12.01	-76.92	1	KF420158
MSB:Bird:31454	163018	9	January	352	Lima	-12.01	-76.92	1	KF420165
MSB:Bird:31456	163020	9	January	352	Lima	-12.01	-76.92	1	KF386253
MSB:Bird:31459	163023	9	January	352	Lima	-12.01	-76.92	1	KF386237
MSB:Bird:31469	163033	12	January	3967	Lima	-11.63	-76.43	1	KF386168
MSB:Bird:31482	163046	12	January	3959	Lima	-11.63	-76.43	1	KF386245
MSB:Bird:31489	163053	13	January	3973	Lima	-11.63	-76.43	1	KF386244
MSB:Bird:31498	163062	13	January	3981	Lima	-11.63	-76.43	1	KF420169
MSB:Bird:31503	163067	14	January	3967	Lima	-11.63	-76.43	1	KF386183
MSB:Bird:31739	163411	7	May	3750	Lima	-11.76	-76.55	1	KF386242
MSB:Bird:31756	163428	21	May	2400	Lima	-11.74	-76.61	1	KF420173
MSB:Bird:31766	163438	24	May	2400	Lima	-11.74	-76.61	1	KF420157
MSB:Bird:31767	163439	24	May	2400	Lima	-11.74	-76.61	1	KF420161
MSB:Bird:32902	168074	15	October	935	Lima	-12.03	-76.65	1	KF386249
MSB:Bird:32909	168081	15	October	935	Lima	-12.03	-76.65	1	KF386190
MSB:Bird:32910	168082	15	October	935	Lima	-12.03	-76.65	1	KF420177
MSB:Bird:32922	168094	16	October	935	Lima	-12.03	-76.65	1	KF386239
MSB:Bird:32923	168095	16	October	935	Lima	-12.03	-76.65	1	KF386235
MSB:Bird:32943	168115	18	October	352	Lima	-12.01	-76.92	1	KF420167
MSB:Bird:32950	168122	18	October	352	Lima	-12.01	-76.92	1	KF386257
MSB:Bird:32966	168138	18	October	352	Lima	-12.01	-76.92	1	KF386182
MSB:Bird:32967	168139	18	October	352	Lima	-12.01	-76.92	1	KF386184
MSB:Bird:32969	168141	18	October	352	Lima	-12.01	-76.92	1	KF386230
MSB:Bird:32982	168154	19	October	352	Lima	-12.01	-76.92	1	KF386247
MSB:Bird:32988	168160	19	October	352	Lima	-12.01	-76.92	1	KF386231

MSB:Bird:32991	168163	19	October	352	Lima	-12.01	-76.92	1	KF386246
MSB:Bird:33001	168173	19	October	352	Lima	-12.01	-76.92	1	KF386248
MSB:Bird:33006	168178	19	October	352	Lima	-12.01	-76.92	1	KF386233
MSB:Bird:33008	168180	19	October	352	Lima	-12.01	-76.92	1	KF386211
MSB:Bird:33310	168529	4	Sept.	4056	Lima	-11.77	-76.53	1	KF386180
MSB:Bird:33329	168548	6	Sept.	3910	Lima	-11.76	-76.54	1	KF386189
MSB:Bird:33351	168570	9	Sept.	3907	Lima	-11.77	-76.53	1	KF386191
MSB:Bird:33370	168589	12	Sept.	3905	Lima	-11.77	-76.53	1	KF386188
MSB:Bird:33416	168635	17	Sept.	4056	Lima	-11.77	-76.53	1	KF386205
MSB:Bird:34736	171462	3	July	309	La Libertad	-8.39	-78.65	1	KF386238
MSB:Bird:34739	171465	3	July	309	La Libertad	-8.39	-78.65	1	KF386185
MSB:Bird:34763	171489	4	July	309	La Libertad	-8.39	-78.65	1	KF386186
MSB:Bird:34830	171556	8	July	2972	Ancash	-8.75	-78.05	1	KF386236
MSB:Bird:34832	171558	8	July	2972	Ancash	-8.75	-78.05	1	KF386215
MSB:Bird:34892	171618	11	July	2972	Ancash	-8.75	-78.05	1	KF386167
MSB:Bird:34903	171629	12	July	357	La Libertad	-8.69	-78.38	1	KF420175
MSB:Bird:34907	171633	13	July	357	La Libertad	-8.69	-78.38	1	KF420176
MSB:Bird:34916	171642	13	July	357	La Libertad	-8.69	-78.38	1	KF386187
MSB:Bird:34920	171646	14	July	357	La Libertad	-8.69	-78.38	1	KF420190
MSB:Bird:34921	171647	14	July	357	La Libertad	-8.69	-78.38	1	KF386240
MSB:Bird:34953	171679	16	July	3439	Ancash	-9.34	-77.51	1	KF386220
MSB:Bird:34965	171691	17	July	3439	Ancash	-9.34	-77.51	1	KF386170
MSB:Bird:34966	171692	17	July	3439	Ancash	-9.34	-77.51	1	KF386181
MSB:Bird:34967	171693	17	July	3439	Ancash	-9.34	-77.51	1	KF386171
MSB:Bird:35007	171733	21	July	3714	Ancash	-9.1	-77.87	1	KF386217
MSB:Bird:35018	171744	21	July	3714	Ancash	-9.1	-77.87	1	KF386241
MSB:Bird:35538	172264	8	August	3200	Arequipa	-15.81	-72.67	1	KF420193
MSB:Bird:36014	173845	18	May	3740	Ancash	-8.74	-78.04	1	KF386172
MSB:Bird:36049	173880	23	May	3740	Ancash	-9.02	-77.54	1	KF386219
MSB:Bird:36081	173912	30	May	3350	Ancash	-8.84	-77.93	1	KF386243
MSB:Bird:36568	175520	22	October	4000	Lima	-11.77	-76.53	1	KF420178
MSB:Bird:36573	175525	24	October	4116	Lima	-11.77	-76.53	1	KF386169
MSB:Bird:36574	175526	24	October	4123	Lima	-11.77	-76.53	1	KF386229
MSB:Bird:32351	167523	15	July	2052	Amazonas	-6.1	-78.34	2	KF420174
MSB:Bird:32619	167791	22	July	2066	Amazonas	-6.1	-78.34	2	KF386193
MSB:Bird:32855	168027	28	July	2073	Amazonas	-6.1	-78.34	2	KF386192
MSB:Bird:32862	168034	29	July	2073	Amazonas	-6.1	-78.34	2	KF420170
MSB:Bird:33894	169120	26	Dec.	143	Lambayeque	-5.9	-79.79	2	KF420168
MSB:Bird:34057	169283	22	Dec.	2215	Piura	-5.84	-79.51	2	KF386195
MSB:Bird:34076	169302	23	Dec.	2215	Piura	-5.84	-79.51	2	KF386194
MSB:Bird:34773	171499	4	July	309	La Libertad	-8.39	-78.65	2	KF420182
MSB:Bird:34902	171628	12	July	357	La Libertad	-8.69	-78.38	2	KF420181
MSB:Bird:27066	159722	26	Nov.	3120	Cusco	-13.63	-71.72	3	KF386255

MSB:Bird:27076	159732	27	Nov.	3120	Cusco	-13.63	-71.72	3	KF386254
MSB:Bird:27083	159740	28	Nov.	3120	Cusco	-13.63	-71.72	3	KF420180
MSB:Bird:27131	159789	4	Dec.	4300	Cusco	-13.2	-72.16	3	KF386251
MSB:Bird:27132	159790	4	Dec.	4300	Cusco	-13.2	-72.16	3	KF386226
MSB:Bird:27154	159814	8	Dec.	3380	Cusco	-13.25	-72.17	3	KF420179
MSB:Bird:27174	159834	9	Dec.	3380	Cusco	-13.25	-72.17	3	KF386234
MSB:Bird:27179	159840	9	Dec.	3380	Cusco	-13.25	-72.17	3	KF420172
MSB:Bird:27181	159842	10	Dec.	3380	Cusco	-13.25	-72.17	3	KF420187
MSB:Bird:27203	159868	12	Dec.	3380	Cusco	-13.25	-72.17	3	KF386252
MSB:Bird:27218	159887	13	Dec.	3380	Cusco	-13.25	-72.17	3	KF420171
MSB:Bird:31835	163507	18	June	3710	Junin	-11.98	-74.93	3	KF386223
MSB:Bird:33119	168338	12	March	4030	Cusco	-13.19	-72.23	3	KF420159
MSB:Bird:33650	168876	4	Dec.	3573	Apurimac	-14.41	-73.09	3	KF420164
MSB:Bird:33659	168885	5	Dec.	3548	Apurimac	-14.41	-73.09	3	KF386225
MSB:Bird:34084	169310	9	January	4369	Apurimac	-14.06	-73.01	3	KF386222
MSB:Bird:34085	169311	9	January	4454	Apurimac	-14.06	-73	3	KF420163
MSB:Bird:34089	169315	9	January	4375	Apurimac	-14.06	-73.01	3	KF386227
MSB:Bird:34106	169332	9	January	4384	Apurimac	-14.06	-73.01	3	KF386221
MSB:Bird:34109	169335	10	January	4375	Apurimac	-14.06	-73.01	3	KF386224
MSB:Bird:34114	169340	10	January	4401	Apurimac	-14.06	-73.01	3	KF386260
MSB:Bird:34202	169428	14	January	4363	Apurimac	-14.06	-73	3	KF386258
MSB:Bird:34295	171021	30	May	1500	Cuzco	-12.65	-72.32	3	KF420188
MSB:Bird:34359	171085	3	June	1500	Cuzco	-12.65	-72.32	3	KF420189
MSB:Bird:35522	172248	8	August	3200	Arequipa	-15.81	-72.67	3	KF420192
MSB:Bird:35700	172426	28	Sept.	2671	Apurimac	-14.17	-73.32	3	KF420196
MSB:Bird:35822	172637	5	August	3201	Cusco	-13.08	-72.37	3	KF386259
MSB:Bird:35907	172722	21	Sept.	2672	Apurimac	-14.17	-73.32	3	KF386256
MSB:Bird:35908	172723	21	Sept.	2671	Apurimac	-14.17	-73.32	3	KF386197
MSB:Bird:28029	162535	20	June	322	San Martín	-6.65	-76.07	4	KF386175
MSB:Bird:33581	168807	19	Nov.	2500	Cuzco	-13.56	-70.88	4	KF420200
MSB:Bird:36130	173961	13	June	1673	San Martin	-7.42	-76.29	4	KF420186
MSB:Bird:36909	176089	23	June	292	Madre de Dios	-11.71	-69.21	4	KF386204
MSB:Bird:37014	176194	26	June	297	Madre de Dios	-11.71	-69.21	4	KF386176
MSB:Bird:37084	176264	28	June	297	Madre de Dios	-11.71	-69.21	4	KF420184
MSB:Bird:37149	176329	30	June	290	Madre de Dios	-11.71	-69.21	4	KF420185
MSB:Bird:37202	176382	1	July	297	Madre de Dios	-11.71	-69.21	4	KF386178
MSB:Bird:37318	176498	4	July	297	Madre de Dios	-11.71	-69.21	4	KF386174
MSB:Bird:37341	176521	5	July	297	Madre de Dios	-11.71	-69.21	4	KF386173
MSB:Bird:33720	168946	15	Dec.	133	Lambayeque	-5.9	-79.79	5	KF386201
MSB:Bird:33725	168951	15	Dec.	133	Lambayeque	-5.9	-79.79	5	KF420156
MSB:Bird:33758	168984	16	Dec.	129	Lambayeque	-5.9	-79.78	5	KF420166
MSB:Bird:33778	169004	18	Dec.	133	Lambayeque	-5.9	-79.79	5	KF386200
MSB:Bird:33779	169005	18	Dec.	133	Lambayeque	-5.9	-79.79	5	KF420162

MSB:Bird:33844	169070	21	Dec.	133	Lambayeque	-5.9	-79.79	5	KF386196
MSB:Bird:33889	169115	25	Dec.	143	Lambayeque	-5.9	-79.79	5	KF386199
MSB:Bird:34698	171424	2	July	309	La Libertad	-8.39	-78.65	5	KF420191
MSB:Bird:34699	171425	2	July	309	La Libertad	-8.39	-78.65	5	KF420198
MSB:Bird:34893	171619	11	July	2972	Ancash	-8.75	-78.05	5	KF386198
MSB:Bird:35250	171976	13	July	2500	Cajamarca	-7.4	-78.78	5	KF420183
MSB:Bird:35330	172056	17	July	2550	Cajamarca	-7.4	-78.78	5	KF386202
MSB:Bird:35393	172119	21	July	2550	Cajamarca	-7.4	-78.78	5	KF420197
MSB:Bird:35402	172128	22	July	2550	Cajamarca	-7.4	-78.78	5	KF386203
MSB:Bird:35035	171761	30	July	740	Tacna	-17.56	-70.67	6	KF386206
MSB:Bird:35043	171769	31	July	740	Tacna	-17.56	-70.67	6	KF386213
MSB:Bird:35046	171772	31	July	740	Tacna	-17.56	-70.67	6	KF386208
MSB:Bird:35047	171773	31	July	740	Tacna	-17.56	-70.67	6	KF386212
MSB:Bird:35057	171783	1	August	740	Tacna	-17.56	-70.67	6	KF386207
MSB:Bird:35442	172168	1	August	2200	Tacna	-17.39	-70.35	6	KF420160
MSB:Bird:35476	172202	3	August	2975	Tacna	-17.32	-70.25	6	KF420194
MSB:Bird:35491	172217	4	August	2975	Tacna	-17.32	-70.25	6	KF420195
MSB:Bird:35493	172219	4	August	2975	Tacna	-17.32	-70.25	6	KF386214
MSB:Bird:35507	172233	4	August	2975	Tacna	-17.32	-70.25	6	KF386209
MSB:Bird:35512	172238	5	August	2975	Tacna	-17.32	-70.25	6	KF420199
MSB:Bird:31626	163191	29	January	2798	Huánuco	-9.73	-76.11	7	KF386177
MSB:Bird:35697	172423	27	Nov.	-	Ancash	-	-	7	KF386179

APPENDIX B

Haemoproteus lineages included in the regional analysis of chapter 1, with GenBank accession number, MalAvi name, and geographic region of origin.

Accession No.	MalAvi Name	Geographic Region
EF153646	APSPI02	Andes
EF153648	ZOCAP02	Andes
EF153650	PHALA01	Andes
EF153653	PHFRU02	Andes
EF153654	PHFRU01	Andes
JQ988105	-	Andes
JQ988106	-	Andes
JQ988107	-	Andes
JQ988117	-	Andes
JQ988167	-	Andes
JQ988123	-	Andes
JQ988134	-	Andes
JQ988135	-	Andes
JQ988136	-	Andes
JQ988144	-	Andes
JQ988147	-	Andes
JQ988206	-	Andes
JQ988220	-	Andes
JQ988254	-	Andes
JQ988255	-	Andes
JQ988256	-	Andes
JQ988257	-	Andes
JQ988295	-	Andes
JQ988305	-	Andes
JQ988310	-	Andes
JQ988323	-	Andes
JQ988342	-	Andes
JQ988355	-	Andes
JQ988370	-	Andes
JQ988371	-	Andes
JQ988384	-	Andes
JQ988393	-	Andes
JQ988404	-	Andes
JQ988414	-	Andes

JQ988426	-	Andes
JQ988430	-	Andes
JQ988446	-	Andes
JQ988447	-	Andes
JQ988462	-	Andes
JQ988489	-	Andes
JQ988492	-	Andes
JQ988508	-	Andes
JQ988516	-	Andes
JQ988521	-	Andes
JQ988527	-	Andes
JQ988544	-	Andes
JQ988563	-	Andes
JQ988570	-	Andes
JQ988575	-	Andes
JQ988577	-	Andes
JQ988585	-	Andes
JQ988623	-	Andes
JQ988656	-	Andes
JQ988744	-	Andes
KF767421	TROAED12	Andes
KF767419	TROAED13	Andes
KF767418	TROAED14	Andes
KF767417	TROAED15	Andes
KF767424	TROAED16	Andes
KF767416	TROAED17	Andes
KF767425	TROAED19	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes
Pending	-	Andes

KC480265	-	Andes
KC480266	ZOCAP01	Andes
JQ988150	-	Andes and North America
JQ988222	-	Andes and North America
JQ988571	-	Andes and North America
JN819385	TASCH01	Andes and North America
JX029900	ELALB01	Andes and South America
JX029915	-	Andes and South America
JX029905	-	Andes and South America
DQ241539	CHASPI01	Andes and South America
GQ141568	COFLA06	Antilles
GQ395652	COFLA07	Antilles
GQ141562	COFLA03	Antilles
GQ395638	COFLA05	Antilles
GQ141561	PHAPAL01	Antilles
GQ141565	PHAPAL02	Antilles
GQ141571	ICTLEU01	Antilles
GQ141597	-	Antilles
GQ141579	MAFUS04	Antilles
GQ141599	MIMGIL01	Antilles
GQ141563	SPIDOM01	Antilles
GQ141573	COFLA08	Antilles
GU256263	LOXPOR01	Antilles
GU256261	-	Antilles
GU251992	-	Antilles
AF465579	COFLA01	Antilles
GU251995	-	Antilles
AF465569	LOXNOC01	Antilles
AF465565	MAFUS03	Antilles
GU251991	SALAUR01	Continental North America
GQ141557	ANACRE01	Continental North America
GQ141589	DENPEN02	Continental North America
GQ141575	DENMAG01	Continental North America
GQ141581	CARCAR01	Continental North America
GQ141584	TUMIG06	Continental North America
GQ395668	TABI02	Continental North America
GU256262	-	Continental North America
EU627829	BNOW01	Continental North America
EU627838	BNOW02	Continental North America
EU627830	BNOW03	Continental North America
AF465573	COBRA01	Continental North America
JQ314226	CYGNUS01	Continental North America
GQ141606	BUTJAM12	Continental North America

EU328178	GEOTRI04	Continental North America
AF465590	MALERY01	Continental North America
JX073258	MODO1	Continental North America
GU252004	-	Continental North America
GU252005	-	Continental North America
GU252006	-	Continental North America
AF254977	PARUS1	Continental North America
AF465583	PIOLI01	Continental North America
AF465582	PIRUB01	Continental North America
AF465594	POLPLA01	Continental North America
EU328179	SIAMEX01	Continental North America
AY393806	SISKIN1	Continental North America
JN788938	-	Continental North America
JN819388	TURDUS2	Continental North America
JN819379	-	Continental North America
JN819345	TASCH01	Continental North America
JN819370	-	Continental North America
JN819387	-	Continental North America
JN819369	-	Continental North America
JN819373	-	Continental North America
AF465589	STVAR01	Continental North America
EU627834	STVAR03	Continental North America
JN792136	-	Continental North America
JN792141	-	Continental North America
JN792139	-	Continental North America
JN792142	-	Continental North America
JN792140	-	Continental North America
JN792147	-	Continental North America
JN792143	-	Continental North America
AF465570	TATHA01	Continental North America
AY099034	VIOLI01	Continental North America
JN819399	-	Continental North America
JN819393	COLL2	Continental North America
JN819383	TURDUS2	Continental North America
JN819378	-	Continental North America
GU252009	-	Continental North America
GQ141576	MAFUS02	Continental NA and Antilles
HM222472	-	Continental NA and Antilles
AF465564	TOXRUF01	Continental NA and Antilles
AF465577	VIGRI02	Continental NA and Antilles
JX029902	-	South America
JX029917	-	South America
DQ241546	ICTCAY02	South America

DQ241544	MOLBAD01	South America
DQ241550	MONGUT02	South America
JX029914	-	South America
JX029906	-	South America
JX029911	PACPEC02	South America
JX029907	-	South America
JX029912	-	South America
JX029913	-	South America
JX029919	-	South America
DQ241541	PSAVIR02	South America
JX029910	-	South America
AF465568	TIABIC01	South America
JX029908	-	South America
JX029916	-	South America
JX029918	-	South America
JX029920	-	South America
JX029903	-	South America
GU085190	EMCIR01	South America
JX029901	-	South America
DQ241542	LAMMEL01	South America
DQ241553	ARAPER01	South America

APPENDIX C

Haemosporidian lineages encountered in chapter 1 and their global distributions, number of host species each has been found to infect (host breadth; HB), GenBank accession numbers, and catalog number for lineages that were identified by microscopy on blood smears. Novel lineages are indicated by asterisks.

Lin.	MalAvi Name	Global distribution	HB	Genbank accession no.	Cat. No.
L1	TROAED01, TROAED02	Chile, Peru	34	KF767426, KF767431- KF767436	
L2*	TROAED03	Peru	1	KF767439	
L3	TROAED04	Peru	2	KF767429	MSB:Para:19028
L4	TROAED05	Peru	2	KF767440	
L5	TROAED06, TROAED07	Peru	7	KF767437, KF767438, KF767441, JQ988531	
L6	TROAED08	Peru	3	KF767427	
L7*	TROAED09	Peru	1	KF767430	
L8	TROAED10	Peru	2	KF767442	
L9*	TROAED11	Peru	1	KF767428	
H1*	TROAED12	Peru	1	KF767421, KF767422	
H2*	TROAED13	Peru	1	KF767419	
H3*	TROAED14	Peru	1	KF767418	
H4*	TROAED15	Peru	1	KF767417	
H5*	TROAED16	Peru	1	KF767424	
H6*	TROAED17	Peru	1	KF767416	
H7	TROAED18	Peru	11	KF767423	
H8*	TROAED19	Peru	1	KF767425	
H9	TROAED20	Peru	45	KF767420, JQ988488, JQ988538, JQ988487, JQ988406	
P1	BAEBIC02, TROAED21	Costa Rica, Peru, U.S.A.	33	KF767413, JQ988550, JQ988539, JQ988537, JQ988540, JQ988551	
P2	TUMIG03	Chile, Costa Rica, Peru, Uruguay, U.S.A	13	KF767414, KF767415	
P3	TROAED22	Peru	2	KF767412	
P4	PADOM09	Brazil, Chile, Costa Rica, Peru, Uruguay, U.S.A	14	KF767409-KF767411	MSB:Para:19029
P5	CATUS05	Peru, U.S.A.	3	KF767406-KF767408	