Habitat Association Predicts Genetic Diversity and Population Divergence in Amazonian Birds

Michael G. Harvey,[1,2,*] **Alexandre Aleixo,**[3] **Camila C. Ribas,**[4] **and Robb T. Brumfield**[1]

1. Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana 70803;
2. Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48105;
3. Museu Paraense Emílio Goeldi (MPEG), Caixa Postal 399, Belém, Pará 66040-170, Brazil;   4. Instituto Nacional de Pesquisas da Amazônia (INPA), Avenida André Araújo 2936, Manaus, Amazonas 69060-001, Brazil

abstract: The ecological traits of organisms may predict their genetic diversity and population genetic structure and mediate the action of evolutionary processes important for speciation and adaptation. Making these ecological-evolutionary links is difficult because it requires comparable genetic estimates from many species with differing ecologies. In Amazonian birds, habitat association is an important component of ecological diversity. Here, we examine the link between habitat association and genetic parameters using 20 pairs of closely related Amazonian bird species in which one member of the pair occurs primarily in forest edge and floodplains and the other occurs in upland forest interior. We use standardized geographic sampling and data from 2,416 genomic markers to estimate genetic diversity, population genetic structure, and statistics reflecting demographic and evolutionary processes. We find that species of upland forest have greater genetic diversity and divergence across the landscape as well as signatures of older histories and less gene flow than floodplain species. Our results reveal that species ecology in the form of habitat association is an important predictor of genetic diversity and population divergence and suggest that differences in diversity between floodplain and upland avifaunas in the Amazon may be driven by differences in the demographic and evolutionary processes at work in the two habitats.

Keywords: population genetics, phylogeography, habitat selection, ultraconserved elements, trait-dependent diversification, Amazon rain forest.

Introduction

Genetic and phenotypic variation within species determines how they respond to environmental change (Willi et al. 2006), their propensity to form new species (Riginos et al. 2014; Harvey et al. 2017), and their susceptibility to extinction (Keller and Waller 2002). Differences among species in their genetic and phenotypic diversity are therefore of fundamental interest. Comparative phylogeographic studies examining spatial patterns of genetic diversity, however, have often treated species-specific differences as noise in the quest to unearth the geological events that have broadly affected the histories of biotas (Avise 1992; Bermingham and Moritz 1998). As improved phylogeographic and population genetic estimates become available for more species, widespread observations of rampant discordance (e.g., Soltis et al. 2006; Smith et al. 2014b) have renewed interest in evaluating whether differences across species in patterns of genetic diversity are deterministic.

The ecological and life-history traits of organisms may be important predictors of differences in patterns of genetic diversity across species. Associations between standing genetic diversity within species and organismal traits have received attention because of interest in the adaptive and evolutionary potential of levels of genetic polymorphism and mutation rates (Nevo et al. 1984; Leffler et al. 2012; Romiguier et al. 2014; Miraldo et al. 2016). Trait dependence in the spatial patterning of diversity among populations is also of interest, in part due to its potential evolutionary importance—divergent populations represent potential incipient species. Although relatively few studies are available, they have found that phylogeographic or population genetic structure is predicted by growth form, breeding system, floral morphology, pollination mechanism, seed dispersal mode, phenology, life cycle, and successional stage in woody plants (Loveless and Hamrick 1984; Duminil et al. 2007; Gianoli et al. 2016); microhabitat association (branch circumference) and elevation in Costa Rican orchids (Kisel et al. 2012); larval dispersal mode in a variety of marine organisms (Palumbi 2003; Hellberg 2009); an association with forest canopy or understory in Neotropical birds (Burney and Brumfield 2009); and body size and reproductive mode in frogs (Pabijan et al. 2012; Paz et al. 2015). More investigations into the association be-

* Corresponding author; e-mail: harveymi@umich.edu.
  ORCIDs: Ribas, http://orcid.org/0000-0002-9088-4828.

tween traits and genetic diversity are warranted (Papada-poulou and Knowles 2016).

Measuring associations between organismal traits and patterns of genetic diversity is useful in that it allows us to predict diversity from readily available trait information. If we are able to determine the mechanisms responsible for these correlations, however, they may provide insight into how organismal traits mediate important evolutionary processes like adaptation and speciation. Assessing the mechanisms of trait dependence in genetic variation requires measurement of evolutionary processes, which can be challenging. For example, traits such as r/K-selected ecological strategies that predict genetic diversity have been hypothesized to operate via their effect on population stability through time (Romiguier et al. 2014), but data on long-term population stability for testing this mechanism are difficult to come by. Traits linked to population divergence are often hypothesized to operate by mediating dispersal patterns across space (e.g., Palumbi 2003; Burney and Brumfield 2009), but few data on realized dispersal or migration rates at large scales are available. New data and methods, however, may facilitate improved investigation of the processes that might act as mechanisms linking traits to evolutionary patterns.

Genome-wide approaches to genetic sampling may provide more accurate estimates of genetic diversity and also provide information on processes that are potential mechanisms for differences in genetic diversity among species. Methods for sequencing reduced representation libraries of genomic DNA can be used to obtain information from many independent parts of the genome and many samples (e.g., Davey et al. 2011; Faircloth et al. 2012). Increasing the number of loci under investigation provides more precise estimates of patterns and processes that are less subject to biases resulting from coalescent stochasticity (Edwards and Beerli 2000; Carling and Brumfield 2007). Sampling hundreds of loci is equivalent to sampling an entire population at a few loci, and with enough loci many parameters can be reliably estimated even when populations are represented by only a single diploid individual (e.g., Willing et al. 2012). Data sets with many independent loci may provide sufficient power to evaluate parameter-rich models of population history that include such processes as migration between populations, changes in population size, and selection in addition to divergence (Carstens et al. 2013). Finally, processes like migration between populations and selection may be evident only in subsets of the genome (Counterman et al. 2004; Wall et al. 2009) and are best identified using dense genomic sampling. Improved estimates of genetic diversity, population divergence, and other evolutionary processes may improve our understanding of the evolutionary effects of species traits.

The avifauna of the Amazon basin in northern South America provides an excellent system in which to investi-gate the effect of traits on genomic diversity and population history. The Amazonian avifauna is the most diverse in the world (Pearson 1977) and comprises species with a variety of ecological traits (Parker et al. 1996) and, based on the few species with data, differing levels of genetic diversity (Bates 2000; Smith et al. 2014b). Many species are habitat specialists (Rosenberg 1990; Kratter 1997; Alonso et al. 2013), and closely related species (e.g., congeners) often partition space by associating with different habitats. Two habitats in particular, floodplain forest (*várzea* and *igapó*) and upland forest (*terra firme*), are widespread and are inhabited by a suite of pairs of congeneric species that segregate by habitat (Remsen and Parker 1983) and sometimes exhibit interspecific aggression (Robinson and Terborgh 1995). Floodplain forest, particularly várzea, has an open, edge-like structure as a result of disturbance during floods (Prance 1979; Wittmann et al. 2004), and many floodplain species occur outside of floodplains in other edge habitats, such as the borders of savanna or human-made clearings. Upland forest, conversely, is typified by a high proportion of tall trees, a dark interior, and open understory (Campbell et al. 1986; Gentry and Emmons 1987), and many upland forest species avoid open areas.

The habitat associations of Amazonian birds may be important predictors of patterns of genetic diversity and population divergence across the landscape. Some evidence already exists for higher levels of population divergence in upland forest species than floodplain and edge species, based on a greater number of subspecies within species inhabiting upland forest (Salisbury et al. 2012). In addition, several mechanisms exist that may lead to elevated genetic diversity and divergence in upland forest species. Upland habitats are larger in area than floodplain habitats and have been less subject to periodic flooding over geological time-scales (Irion et al. 2009; Hess et al. 2015), which might lead to larger and more stable populations and the maintenance of genetic diversity. We therefore expect population genetic diversity and estimates of population size to be higher and population size change through time to be reduced in upland bird species. Upland habitats are subdivided by major Amazonian rivers, whereas linear stretches of floodplain habitats along rivers are largely continuous, such that migration might limit opportunities for population divergence (Patton and da Silva 1998; Aleixo 2006). We therefore expect higher population genetic structure and lower estimates of gene flow in upland forest. In addition, the greater stability through time of upland forest may produce greater population divergence because populations have had more time for isolation, in which case we expect deeper population histories in upland species. Other possible mechanisms exist, some of them detailed in "Discussion," but those listed here provide a useful set of hypotheses to evaluate with genetic data.

In this study, we examine 40 species or species complexes (all of which are hereafter referred to as "species," for brevity) of broadly codistributed Amazonian birds that differ in habitat association. The 40 species represent 20 pairs in which one species is found in upland forest and the other is a closely related species found in floodplains and edge habitats. We use genomic sequence data from populations distributed across the Amazon to estimate parameters that reflect genetic diversity and population divergence as well as effective population size, gene flow, and the depth of population history in each species. We then compare these variables between the floodplain/edge and upland forest species. We discuss differences in patterns of genetic diversity and divergence between the habitats and evaluate their potential mechanisms and significance.

Methods

Sample Design

We designed a sampling strategy to minimize the potential effects of sampling bias across species on comparisons of genetic parameters. We first selected genera that contained a pair of species or species complexes that have been found to segregate between floodplains and upland forest using published survey data on Amazonian birds (Remsen and Parker 1983; Terborgh 1985; Robinson and Terborgh 1995), compilations of ecological trait information (Parker et al. 1996; del Hoyo et al. 1992–2011; Schulenberg et al. 2010), expert knowledge (B. M. Whitney and L. N. Naka, personal communication), and personal observation. Although detailed quantitative data on the ecological niches of most Amazonian bird species are lacking, environmental data from where they occur can provide a rough approximation for further assessment of ecological differences. Tuanmu and Jetz (2015) found that dissimilarity in the enhanced vegetation index (EVI) between adjacent pixels in MODIS satellite imagery effectively distinguished edge habitats, such as along rivers, from closed forest. Using the R packages maptools (Bivand and Lewan-Koh 2017), raster (Hijmans 2016), rgdal (Bivand et al. 2017), and spThin (Aiello-Lammens et al. 2014), we combined georeferenced specimen localities (see below) with records from the eBird database (Sullivan et al. 2009; April 2017 version), filtered out eBird records representing long survey periods (6 h or more) or large spatial areas (5 km or farther), clipped records to the Amazon basin (Mayorga et al. 2012), thinned records occurring within 2 km of one another, and extracted EVI dissimilarity values (Tuanmu and Jetz 2015). We compared EVI dissimilarity between floodplain and upland species in each genus.

Some of the genera selected have since been split into multiple genera (Remsen et al. 2015), but the species pairs examined are still closely related (1%–8% mitochondrial distance and 0.1%–0.5% average genetic distance across nuclear loci; fig. 1). Species pairs are not necessarily sister taxa. For each species, we examined all populations within the Amazon basin. Some of our study species have populations outside the Amazon basin, generally in the Atlantic Forest of southeastern South America or the humid forests of Central America and the Chocó region of northwestern South America, and we included samples from these areas when available, but we focused comparative analyses on Amazonian populations only.

We made lists of vouchered tissue samples from the Amazon basin collected during our fieldwork and available from existing natural history collections. From an initial set of 57 species pairs fitting our taxonomic and ecological criteria, we removed any pair containing a species for which fewer than 20 tissue samples were available. The result was a list of 20 species pairs from 15 avian families (fig. 1). We selected a set of samples for each species that would minimize differences in the spatial dispersion of samples across species. We plotted a random set of 40 localities across the Amazon basin using the genrandompoints function in Geospatial Modelling Environment software (ver. 0.7.1.0; Spatial Ecology, Toronto, Canada). We then georeferenced all genetic samples with locality information more precise than department or state and sufficient precision to determine on which side of any major biogeographic barriers (rivers or mountains) the sample originated. Locality records were plotted using ArcMap (ver. 10.0; ESRI, Redlands, CA) with the WGS84 projection. For each species, we selected samples that were closest to each of the semirandom localities using the spatial join function in ArcMap. The vagaries of sample distribution sometimes resulted in the same sample being selected for multiple localities or in a strong clustering of samples. We thinned sampling to 20 localities by removing clustered samples on the basis of proximity to the nearest sample.

Laboratory Methods

We extracted whole genomic DNA from tissue samples using DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA) and quantified extracts using a QuBit fluorometer (Thermo-Fisher, Waltham, MA). We excluded samples with extracts containing less than 1 μg of total DNA and thinned the remaining samples on the basis of proximity as described above, to arrive at a final set of 11 samples for each species.

Due to the comparative nature of our study, it was important to obtain genetic data that would not bias estimates of genetic parameters across species. Results are generally not comparable across species if different loci are examined because the process of orthology assessment among sequence reads leads to the recovery of subsets of loci that are biased contingent on the amount of diversity in each species (Harvey et al. 2015). Sequence capture of conserved
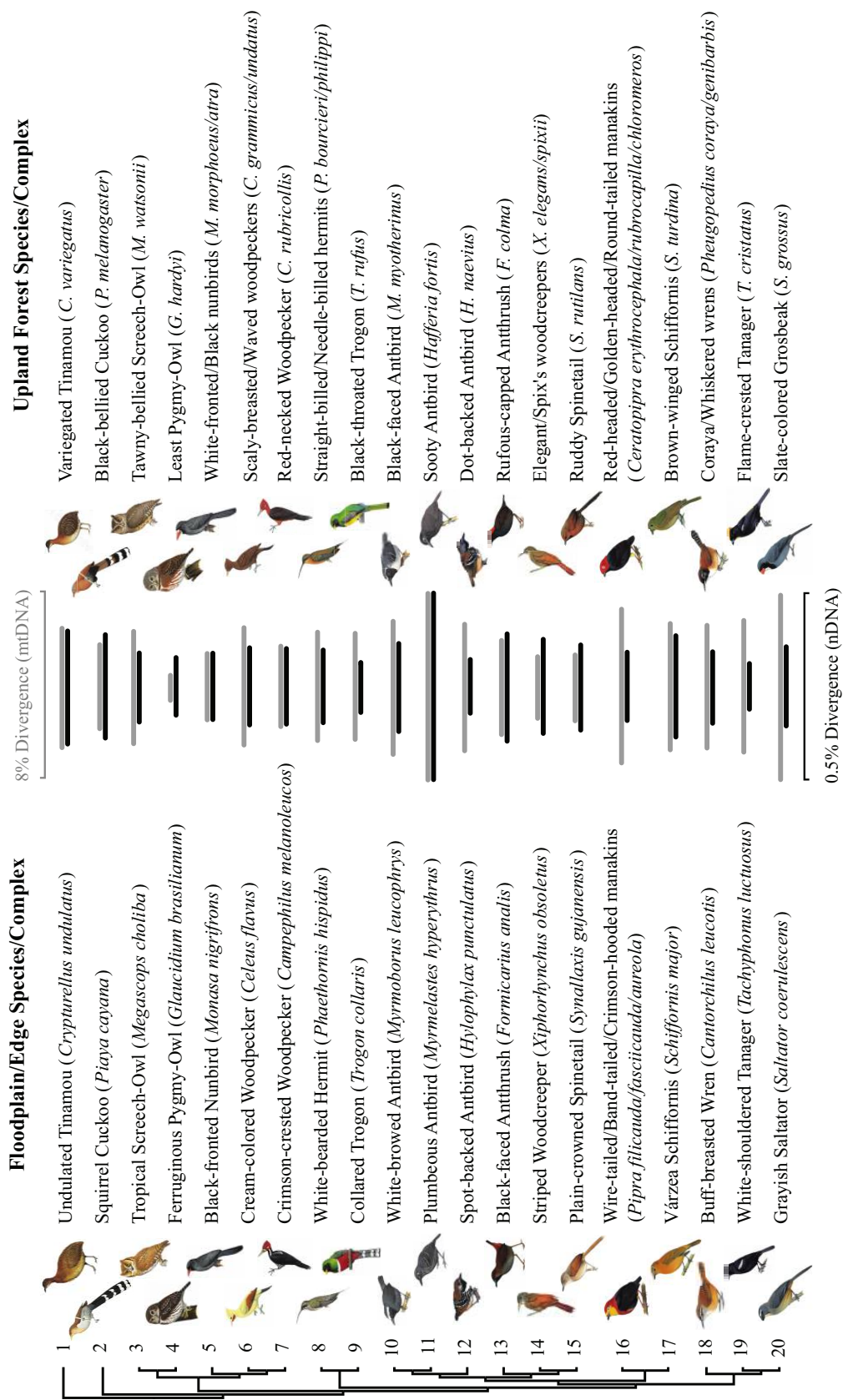
Figure 1: Pairs of study species or species complexes examined. A phylogenetic tree based on a MrBayes analysis of concatenated nuclear loci depicts the relationships among the pairs, and horizontal bars in the middle of the figure depict the mean pairwise sequence divergence (d_{xy}) in mitochondrial genomes (gray) and nuclear loci (black) between members of a pair. Bird images courtesy del Hoyo et al. (2017). mtDNA = mitochondrial DNA; nDNA = nuclear DNA.

genomic regions permits the interrogation of the same loci across divergent species (Bi et al. 2012; Faircloth et al. 2012; Hedtke et al. 2013), and orthology assessment in the assembly of sequence capture data sets is straightforward and has relatively little effect on allelic diversity (Harvey et al. 2016). Despite the conservation of the loci targeted with sequence capture, the flanking regions contain high levels of variation sufficient to obtain many thousands of variable sites, a sufficient number to estimate population genetic parameters with high precision (Smith et al. 2014*a*).

We used sequence capture to target ultraconserved elements (UCEs) and exons from across the genome. We modified existing sequence capture probe sets for UCEs (Faircloth et al. 2012) to obtain additional sequence from the more variable UCE flanking regions that might be useful for inferring shallow population histories. The modified probe set is described in Zucker et al. (2016). In brief, in UCE loci targeted with a single probe, we designed two probes extending 30 bp farther into the UCE flanks in each direction. We also targeted conserved exons adjoining variable introns that have been used in previous avian phylogenetic studies (Kimball et al. 2009; Smith et al. 2012; Wang et al. 2012). Probe sequences were based on the chicken (*Gallus gallus*) genome release ICGSC Gallus_gallus-4.0 (Hillier et al. 2004). The final probe set included 4,715 probes targeting 2,321 UCEs and 96 exons.

We sent all samples ($n = 454$) to Rapid Genomics (Gainesville, FL) for sequence capture and sequencing following the general protocol described in Faircloth et al. (2012) and Smith et al. (2014*a*). Samples were multiplexed at 160 samples per lane on a 100-bp paired-end HiSeq 2500 run (Illumina, San Diego, CA). Rapid Genomics demultiplexed raw reads using custom scripts and strict barcode matching.

Bioinformatics

We cleaned reads with Illumiprocessor software (Faircloth 2013). We processed and assembled data sets following Zucker et al. (2016) using the seqcap_pop pipeline (https://github.com/mgharvey/seqcap_pop). In brief, we used Velvet (Zerbino and Birney 2008) and the wrapper program VelvetOptimiser (Gladman 2009) to assemble reads across all individuals into contigs de novo. We mapped contigs to UCE probe sequences using PHYLUCE (Faircloth 2015) and then mapped cleaned reads to on-target contigs using BWA (Li and Durbin 2009), allowing four mismatches per read. We used samtools (Li et al. 2009) and PICARD (Broad Institute, Cambridge, MA) to process BAM files, soft-clip pileups outside the reference, and add read groups for each individual. We realigned reads and indels to minimize mismatched bases and then called single-nucleotide polymorphisms (SNPs) and

indels in the Genome Analysis Toolkit (GATK; McKenna et al. 2010). We used the GATK to mask indels, remove SNPs with a quality score below Q30, and phase SNP alleles on the basis of their presence on the same reads and mate pairs. We used add_phased_snps_to_seqs_filter.py from seqcap_pop to insert SNPs into reference sequences and produce alignments for each locus across individuals. We collated sequences and produced final alignments using MAFFT (Katoh et al. 2005).

We also assembled partial mitochondrial genomes for each sample from off-target reads using a similar pipeline. We obtained existing complete or nearly complete mitochondrial genome sequences from the most closely related taxon to each study species for which they were available (table S1; tables S1–S21 are available online). We mapped reads to the mitochondrial genomes and sorted and indexed the BAM file using samtools. We then called variant sites and output VCF files containing variant and invariant bases using FreeBayes (Garrison and Marth 2012) and used these to assemble sequences using freebayes_vcf2fa_mt.py (https://github.com/mgharvey/misc_python). Only sites with a read depth of five or greater were included in sequences. We conducted final alignment with MAFFT.

We searched for potential sample identification errors or signs of contamination by building exploratory trees of concatenated SNPs from the UCE/exon data using MrBayes (ver. 3.2.2; Ronquist et al. 2012) and scrutinizing any long branches and by mapping mitochondrial sequences to existing sequence data in GenBank (Benson et al. 2014) using Blastn (Altschul et al. 1997). We counted the reads in BWA assemblies using samtools.

Summary Statistics

We calculated basic population genetic summary statistics for each species using DendroPy (ver. 3.10.0; Sukumaran and Holder 2010). These included the raw number of variable sites; nucleotide diversity (π; Tajima 1983), a metric of genetic diversity across all individuals in each species; the mutation-scaled effective population size, or Watterson's θ (Watterson 1975), across all individuals; and Tajima's D (Tajima 1989), a ratio of genetic diversity statistics that can reveal signals of population expansions or selection. We also calculated the average observed heterozygosity within each individual of every species as a measure of the standing genetic diversity within populations. Genetic diversity may differ among genomic regions, including between sex-linked chromosomes and autosomes, and this can reveal the action of varied processes, such as differences in effective population size between sexes (Counterman et al. 2004). We determined the genomic location of each locus by mapping it to the Zebra Finch (*Taeniopygia guttata*) genome (Warren et al. 2010). We then com-

pared levels of nucleotide diversity on loci mapping to the Z chromosome to those mapping to the autosomes.

Phylogenetic Analyses

We estimated gene trees for each nuclear locus in every species using RAxML (ver. 8; Stamatakis 2014). We also estimated a phylogeny for the mitochondrial genome in each species using BEAST2 (Bouckaert et al. 2014). Benefits of mitochondrial DNA include its relatively clocklike evolutionary rate and the availability of detailed information on substitution rates from many groups, which permit dating haplotype divergences with some degree of accuracy. We estimated time-calibrated trees using substitution rates on the basis of the formula of Nabholz et al. (2016), which accounts for rate differences associated with differences in avian body mass. We obtained the average body mass of each study species from Dunning et al. (2008). We integrated across possible models of nucleotide substitution on the basis of their probabilities using bModelTest (ver. 0.1.3; Bouckaert and Drummond 2017). We used the total tree depth of nuclear gene trees and the mitochondrial tree as estimates of the overall age of extant populations in each species.

Population Genetic Structure

We examined multiple strategies for estimating population divergence. We first estimated simple summaries of overall population genetic structure in each species using d_{XY} and F_{ST}. We estimated pairwise sequence divergence (d_{XY}) between both haplotypes in each pair of individuals correcting for multiple substitutions using the method of Jukes and Cantor (1969); d_{XY} was estimated separately for each locus. We estimated F_{ST} between all pairs of individuals in each species using the statistic developed by Reich et al. (2009), which has been shown to be unbiased and effective even when dealing with sample sizes as small as two alleles per population (Willing et al. 2012). For each species, we examined mean values of F_{ST} and d_{XY} across all individuals.

We next examined methods to infer population clustering across individuals and assign individuals to populations. Various methods are available to infer population structure, and they can produce different results (Latch et al. 2006; Chen et al. 2007). We therefore examined results from three alternative methods: STRUCTURE (Pritchard et al. 2000), Bayesian analysis of population structure (BAPS; Corander et al. 2003), and discriminant analysis of principal components (DAPC; Jombart et al. 2010). We also used the first two methods to determine whether any of the individuals sampled were assigned with high probabilities to multiple populations, suggestive of admixture between populations. STRUCTURE is a model-based clustering method that simultaneously infers population structure and assesses the probability of individual assignment to a cluster or combination of clusters. We ran STRUCTURE using the linkage model and provided phase information for each site in each individual as well as distances in base pairs between linked sites. Sites mapping to different loci were treated as unlinked. We conducted analyses at k values ranging from 1 to 6, with 10 replicate runs at each value. Each run included a 50,000-iteration burn-in followed by 200,000 sampling iterations, and we assessed convergence by examining α, F, D_{ij}, and the likelihood within and across runs at each k value. We estimated the best value of k using the method of Evanno et al. (2005) implemented in STRUCTURE HARVESTER (Earl 2012). In some cases, the results at the best k value included clusters to which no individuals were assigned. In these situations, we used the largest k value in which at least one individual was assigned to each cluster (Gao et al. 2007). We also explored longer runs of 1,000,000 iterations (following 100,000 burn-in iterations) in two species to ensure that run lengths were sufficient for accurate inference of k. We combined results across replicates runs with the best k value using CLUMPP (Jakobsson and Rosenberg 2007).

BAPS is a model-based clustering method that jointly infers the number of populations and population assignment of individuals, which can then be used in a subsequent analysis of admixture for each individual. Because BAPS requires complete phasing information for linked sites and phasing had failed for some individuals at most linked sites in our data sets, we used the unlinked model and examined only a single randomly selected SNP from each locus for this analysis. We conducted mixture clustering with the maximum number of populations (k) set at 10. We estimated admixture in each individual on the basis of mixture clustering using 50 simulation iterations, 50 reference individuals, and 10 iterations to estimate admixture coefficients in the reference individuals.

DAPC is a fast, nonparametric method for inferring the number of genetic clusters and cluster assignments in large data sets. We inferred the number of clusters and cluster membership in DAPC using the maximum number of principal components available for each species and selected the best value for cluster number by choosing the value at which the Bayesian information criterion reached a low point (Jombart et al. 2010). Unlike STRUCTURE and BAPS, DAPC does not allow for admixture estimation.

Demographic Modeling

We estimated demographic parameters using a coalescent modeling approach in G-PhoCS (ver. 1.2.3; Gronau et al. 2011). We ran analyses using all population assignments inferred in STRUCTURE, BAPS, and DAPC to determine population membership. Admixed individuals were placed in the population with the highest assignment probability.

We specified the population topologies in situations where more than two populations were present using the MrBayes trees of concatenated SNPs. For each species, we examined both a model with no migration between populations subsequent to divergence as well as a model allowing for migration between terminal populations. We used gamma priors of $(1, 5{,}000)$ for θ and τ and $(1, 3)$ for migration and ran analyses for a minimum of 500,000 iterations (sampling every 100). We also explored the effect of θ and τ priors of $(1, 50)$. Convergence was assessed by examining parameter traces and effective sample size values in Tracer (ver. 1.5; Rambaut and Drummond 2007). G-PhoCS implements a multipopulation model and cannot be run in the study species with a single population. For comparative analyses, we used the species-wide θ values from DendroPy and divergence time (τ) values of zero for single-population species.

Comparative Analyses

The analyses described above produced 19 genetic parameters that can be broadly categorized as estimates of data set attributes, genetic diversity, population divergence, population size and stability, rates of gene flow across the landscape, and time in the landscape (table 1). These categories are somewhat contrived because individual parameters may reflect both pattern and process or multiple different processes, an issue we return to in "Discussion." We also expect many genetic parameters to exhibit correlations with each other. We estimated Spearman's correlations between all pairs of variables and significance using the R package Hmisc (Harrell 2016) and grouped highly correlated variables using the ClustOfVar R package, following the developer recommendations (Chavent et al. 2012). Another strategy to account for correlations among variables is to reduce multiple variables to linearly correlated axes using a principal component analysis (PCA). We conducted PCAs on all 16 variables (excluding those considered data attributes) as well as on the subset of variables within each category provided in table 1. Although PC axes may be difficult to interpret biologically, they provide an index of the effect of correlations among variables on our results.

Before testing associations between genetic parameters and traits, we tested whether each parameter was associated with the evolutionary relatedness of study species. We estimated a phylogeny for all 40 study species by aligning UCE and exon sequences from one sample of each species in MAFFT. We concatenated alignment columns that contained all 40 individuals and conducted a Bayesian phylogenetic analysis on the complete matrix in MrBayes to obtain a phylogenetic tree. We square-root transformed right-skewed genetic variables to achieve normality and calculated phylogenetic signal in each variable using Pagel's λ in the R pack-

age Phytools (Revell 2011), with 999 permutations to assess whether λ differed significantly from zero. We also tested whether the degree to which each variable differed between members of a species pair was predicted by the overall level of pairwise sequence divergence (d_{XY}) between them at genomic loci. Because some of the study species represented species complexes (see "Sample Design" above), we tested whether species complexes differed in genetic parameters from the other species.

We tested whether habitat predicted population genetic parameters using two strategies to account for shared evolutionary history. We first used generalized linear mixed models (GLMMs) to test for correlations, including genus as a random variable to account for the shared history between study species pairs. This test serves as a nonparametric paired test of the difference in genetic parameters between floodplain and upland species. The generalized linear modeling approach allowed us to examine response variables with diverse error distributions in the same statistical framework. Gaussian error models were used for continuous and large count data, Poisson models for data composed of low count values (<100), and gamma models with a logarithmic link function for continuous data with positive skew. We examined the relationship between habitat and each genetic response variable in one-way tests using functions for GLMMs in the stats R package (R Core Team 2015). Covariance due to shared history can also be modeled using phylogenetic distance. We square-root transformed right-skewed data to achieve normality and used phylogenetic generalized least squares (PGLS) in the R package caper (Orme et al. 2013) to test for associations between habitat and genetic parameters while controlling for relatedness among species with the MrBayes phylogeny of concatenated data. We accounted for multiple comparisons by conducting a permutation test to evaluate the probability of recovering the observed number of associations between habitat and genetic parameters with randomly permuted values for response variables. All analyses were run on the first three PC axes from each PCA as well as on each of the individual genetic parameters presented in table 1.

Although our primary focus was on the associations between habitat and genetic parameters, we also examined two additional traits previously found to predict population divergence in Neotropical birds. First, whether a bird inhabits the forest canopy or understory has been shown to predict levels of divergence across landscape barriers (Burney and Brumfield 2009; Smith et al. 2014b), so we tested whether canopy and understory species (based on Parker et al. 1996) differed in metrics of population genomic diversity. Second, habitat or microhabitat associations may affect population genetic divergence via differences in dispersal ability among species (Burney and Brumfield 2009). We examined whether Kipp's index, a

Table 1: Relationships between habitat association and genetic parameters

Genetic parameter	Group	GLMM t	GLMM P	PGLS t	PGLS P
Data attributes:					
Mapped reads	1	.25	.80	1.32	.20
Assembled sequence length	2	−.77	.45	−.98	.33
No. variable sites	3	3.47	.001	4.92	<.001
Genetic diversity:					
Nucleotide diversity (π)	2	2.08	.04	3.29	.002
Observed heterozygosity	2	−.02	.98	−.43	.67
Z : A chromosome nucleotide diversity (π)	5	.34	.74	1.08	.29
Population divergence:					
Mean d_{XY} between populations	2	2.10	.04	3.30	.002
No. STRUCTURE populations	4	1.22	.22	2.80	.008
No. BAPS populations	4	.62	.54	1.64	.11
No. DAPC populations	4	1.28	.20	2.88	.007
Population size and stability:					
Watterson's θ	3	3.36	.002	4.81	<.001
Tajima's D	3	−2.03	.05	−2.31	.03
Average θ across populations (G-PhoCS)	6	.10	.92	−.15	.88
Change in θ through time (G-PhoCS)	6	−0.23	.82	−.07	.94
Gene flow:					
Mean F_{ST} between populations	4	2.19	.03	2.89	.006
Average migration rate (G-PhoCS)	7	.08	.94	−.48	.64
Time in the landscape:					
Average gene tree height	2	2.39	.02	3.59	<.001
Crown age of mitochondrial haplotypes	8	.82	.42	1.54	.14
Oldest population divergence or τ (G-PhoCS)	4	.03	.98	1.59	.12

Note: Group numbers indicate assignment to clusters of semi-independent variables based on ClustOfVar. Results are from single-predictor tests. BAPS = Bayesian analysis of population structure; DAPC = discriminant analysis of principal components; GLMM = generalized linear mixed model; PGLS = phylogenetic generalized least squares.

morphological index of dispersal ability that can be measured from museum specimens (Kipp 1959), predicted levels of population genomic diversity across species. Our study design was such that species differing in forest stratum and Kipp's index were not organized into pairs; instead, both members of a pair typically occurred in the same forest stratum and had similar Kipp's index values. We were therefore unable to leverage GLMMs with a random variable to control for covariation across species. Instead, we used PGLS to test for correlations with genetic variables using phylogeny to control for evolutionary relatedness. Finally, we tested whether associations between habitat and genetic parameters involved second-order interactions with forest stratum and/or Kipp's index using multipredictor GLMM and PGLS models.

Results

We found higher dissimilarity in the EVI in localities where the floodplain species occurred (table S2), consistent with an association with edge-like habitats. Fifteen of the 20 comparisons were significant, but all pairs exhibited a higher mean value of vegetation dissimilarity in the floodplain species.

Illumina sequencing produced an average of 2,087,266 (SD, 656,446) raw reads per sample. Raw reads are deposited in the National Center for Biotechnology Information Sequence Read Archive (PRJNA389814). On average, 28.1% (SD, 6.57%) of sequence reads in each species were successfully mapped to target loci after cleaning, and 0.44% (SD, 0.60%) of all reads mapped to the mitochondrion. We obtained data from an average of 2,142 UCEs (SD, 65.5) and 69 exons (SD, 4.8) in each species. We recovered data in at least one species from 2,416 of 2,417 targeted loci. Mean locus length in each species averaged 554 bp (SD, 56.3), and they contained an average of 7,196 variable sites (SD, 1,379) across all loci. Additional summary statistics are provided in tables S3 and S4.

On the basis of MrBayes trees of concatenated SNPs and Blastn results of mitochondrial sequences, we determined that eight samples were likely misidentified or heavily contaminated and removed these from further analyses (table S5). The correct identifications of these individuals, based on BLAST results of mitochondrial fragments, were generally species that are superficially similar in phenotype. Three samples contained large numbers of rare alleles likely resulting from lower levels of contamination or sequencing

errors and were also removed (table S6). Removing these samples resulted in concatenated SNP trees with low to moderate structure based on internal branch lengths (fig. S1; figs. S1–S6 are available online). Three samples failed, with greater than 85% missing data at variable sites, and were removed (table S7). After removing these 14 samples, we were left with 440 samples (plus 24 extra-Amazonian samples) across the 40 study species.

We summarized population genetic summary statistics by obtaining mean values across all loci within a species and present here the averages of those mean values across the 40 study species. Mean nucleotide diversity (π) averaged 1.09×10^{-3} (SD, 2.98×10^{-4}) across study species, mean Watterson's θ averaged 0.79 (SD, 0.22), and Tajima's D averaged -0.79 (SD, 0.36). Mean gene tree depths averaged 3.93×10^{-3} substitutions per site (SD, 7.67×10^{-4}), and the crown ages of mitochondrial haplotypes averaged 4.06 million years ago (SD, 2.34; fig. S2). Across study species, contigs from 2,415 of 2,416 recovered loci successfully mapped to the Zebra Finch genome assembly. Contigs from all species mapped to the Z chromosome for 171 loci, to one of the autosomes for 2,169 loci, and to unplaced scaffolds for 44 loci. For 31 loci, contigs from different species mapped to different chromosomes or scaffolds, resulting in ambiguous positions. Based only on loci mapping to the Z chromosome or autosomes in all study species, the ratio of nucleotide diversity on the Z chromosome to that on the autosomes averaged 1.04 (SD, 0.263).

F_{ST} across loci averaged 0.26 (SD, 0.14), and mean per-locus d_{XY} averaged 1.11×10^{-3} (SD, 3.10×10^{-4}). The number of populations and population assignments inferred from STRUCTURE, BAPS, and DAPC were broadly concordant (figs. 2, S3). The best k value from STRUCTURE analyses based on the Evanno method, after reducing k to remove clusters without assigned individuals, ranged between one and four across study species (median, three). Longer STRUCTURE runs for a subset of two species did not result in different k estimates than the shorter runs used in all species (table S8). The number of populations estimated in BAPS varied from one to three (median, two), and the number of clusters from DAPC varied between one and four (median, two). Many individuals contained mixed probabilities of assignment to different clusters in the STRUCTURE results, potentially indicative of admixture, but no admixture was recovered in the admixture analysis from BAPS. Populations from all three methods were generally partitioned among geographic areas, with boundaries broadly concordant with major rivers (fig. S4).

Because no appropriate calibrations are available, we examined raw parameter values from G-PhoCS and do not present units. Estimates of historical demography from G-PhoCS for the 23 species with multiple populations (fig. 2; table S9) revealed that mean θ across populations within

each species averaged 1.53×10^{-3} (SD, 5.73×10^{-4}) across species. In contemporary populations, θ values were on average 2.68 times larger than the θ inferred for the ancestral population at the root (SD, 1.29). The height of the deepest divergence in the model (τ) varied from 9.72×10^{-5} to 1.13×10^{-3} across species (mean, 4.45×10^{-4}). The average migration rate between populations within a species varied from 0.337 to 4.69 (mean, 0.950). G-PhoCS results were similar in analyses run with different population conformations, so we focused on results from the models using assignments from BAPS.

Genetic parameters compiled for each species from the above analyses and representing data set attributes, genetic diversity, divergence, and demographic history are listed in table 1, and values are presented in table S10. Each genetic parameter was correlated ($P < .05$) with between 1 and 13 others (fig. S5), and we clustered the variables into eight groups containing high within-group correlations. Ten of 19 genetic parameters exhibited phylogenetic signal based on Pagel's λ tests (table S11). The level of overall divergence between the species in a pair, however, was not associated with the degree to which they differed in any genetic variable (table S11). Two parameters, number of mapped reads and number of STRUCTURE populations, were correlated with whether a species represented a single species or species complex (table S11).

Habitat association predicted ($P < .05$) seven genetic parameters from three semi-independent groups in single-comparison GLMM analyses (fig. 3; table 1). The number of significant comparisons is greater than expected by chance, accounting for multiple comparisons ($P = .001$). Three measures of species-wide genetic diversity, the number of variable sites, nucleotide diversity (π), and the mutation-scaled effective population size (θ) were higher in upland forest species than in floodplain species. Tajima's D was slightly lower in upland forest species than in floodplain species, although this was partly driven by one outlier (without Collared Trogon [*Trogon collaris*] GLMM: $t = -2.02$, $P = .051$). Population divergence across the landscape, measured by both d_{XY} and F_{ST}, was higher in upland forest species. Correlations between habitat and d_{XY} or F_{ST} changed little when corrected for small differences among species in the geographic distances between samples (table S12). Finally, the average height of gene trees was greater in upland forest species. PGLS results were similar to those from GLMMs, with greater nucleotide diversity, higher θ, lower Tajima's D, greater gene tree height, and larger d_{XY} and F_{ST} values in upland than in floodplain species (table 1). In addition, the number of populations inferred using both STRUCTURE and DAPC was greater in upland forest species on the basis of PGLS. The number of significant comparisons was greater than that expected by chance across the same number of comparisons on the basis of a permutation test ($P = .002$).
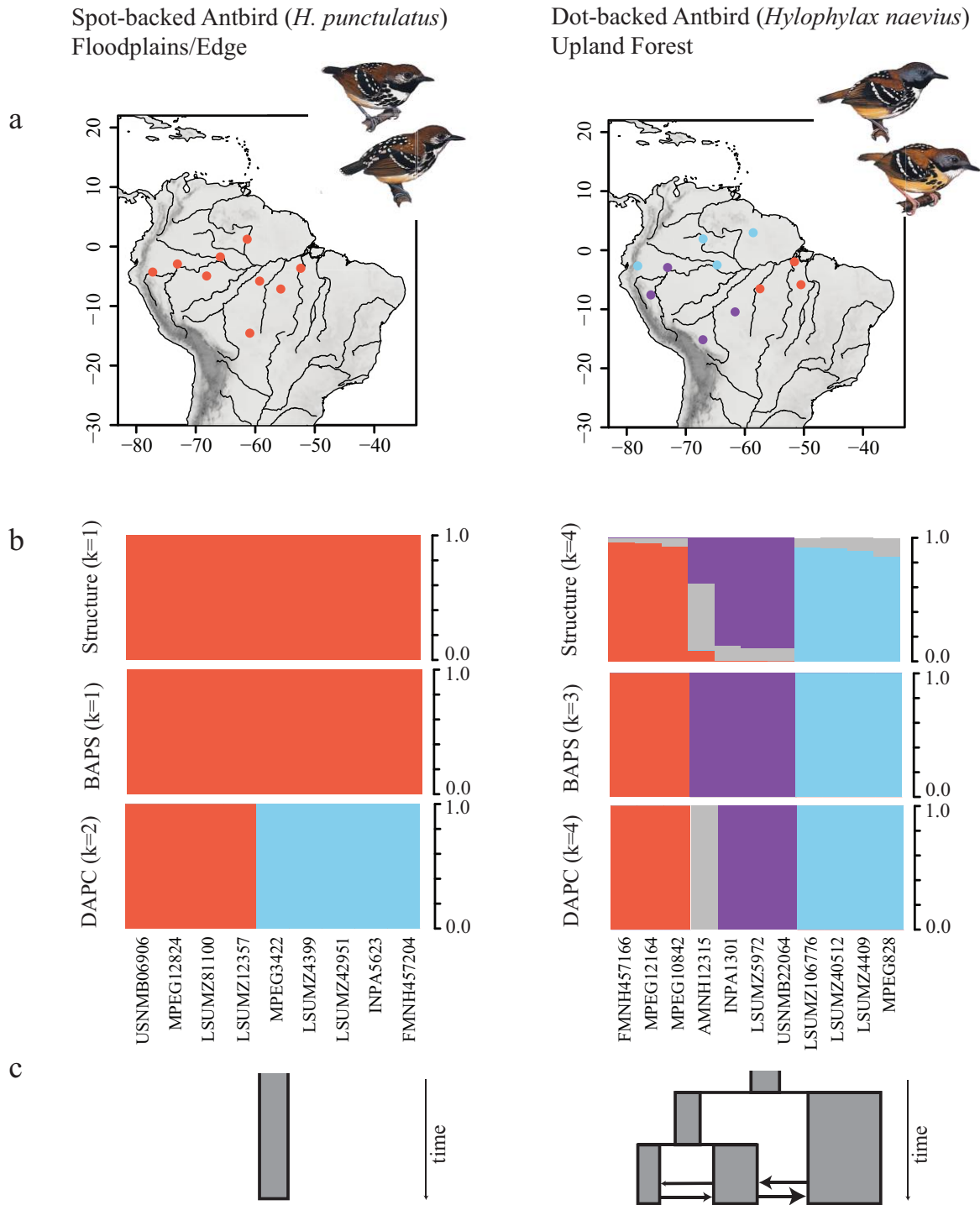
Figure 2: Representative pair of study species depicting (*a*) sample distribution and the distribution of populations inferred with Bayesian analysis of population structure (BAPS); (*b*) individual population assignments based on STRUCTURE analysis, BAPS, and discriminant analysis of principal components (DAPC); and (*c*) demographic models inferred on the basis of the population assignments from BAPS. The individuals are in columns in the population assignment plots. The demographic models depict population history through time, with the width of boxes proportional to their mutation-scaled effective population size (θ), their depth proportional to relative population divergence times (τ), and the size of arrows between them indicating the level of migration between terminal populations. Bird images courtesy del Hoyo et al. (2017).
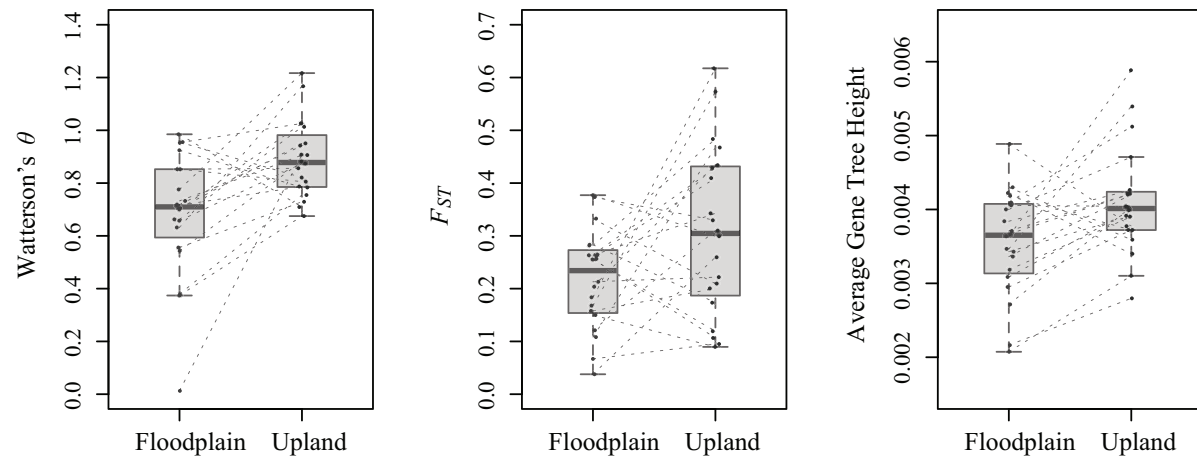
Figure 3: Plots of three representative genetic parameters (each from a different group of correlated variables) that were found to be different between floodplain and upland forest based on generalized linear mixed model analyses. Dashed lines connect members of a species pair. A color version of this figure is available online.

Across response variables, four to seven species pairs showed a difference in the direction opposing the majority of pairs (table S13), with the floodplain species in most of these cases displaying greater diversity or more divergence than the upland forest species.

Relationships between habitat and genetic parameters were similar to those described above in multipredictor models, including forest stratum and/or Kipp's index as well as habitat (tables S14–S20). Analyses using PC axes instead of individual parameters recovered significant associations with habitat (table S21). PC1 from all 16 parameters (excluding only those that reflect data attributes; see table 1) was associated with habitat in a GLMM ($t = -2.44$, $P = .02$). Within each parameter category representing data attributes, genetic diversity, population divergence, and population size and stability, one or two of the first three PC axes was generally correlated with habitat. None of the PC axes reflecting time in the landscape were strongly correlated with habitat, although PC1 was nearly significant ($P = .05$).

PGLS analyses with forest stratum or Kipp's index did not detect strong associations with parameters of population genetic diversity or population history. The only significant relationship was a positive correlation between forest stratum and the relative nucleotide diversity on the Z chromosome versus autosomes ($t = 2.60$, $P = .01$; fig. S6). Results of forest stratum and Kipp's index comparisons were similar between single- and multipredictor PGLS analyses (tables S14–S20).

Discussion

We found that the habitat associations of Amazonian birds predict genome-wide estimates of parameters related to ge-

netic diversity and divergence across the landscape. These results provide further confirmation of the hypothesis that the ecological traits of species can be used to predict levels and patterns of genetic diversity (Loveless and Hamrick 1984; Nevo et al. 1984; Duminil et al. 2007; Burney and Brumfield 2009; Kisel et al. 2012; Leffler et al. 2012; Pabijan et al. 2012; Romiguier et al. 2014; Paz et al. 2015). This is important because it suggests that broad ecological information about species can be used as proxies for diversity. We also found evidence that habitat was associated with parameters such as effective population size and size change, gene flow, and the age of populations, reflecting possible mechanisms responsible for differences in patterns of genetic variation across species. These results highlight the potential of population genomic data to elucidate the mechanisms whereby species ecologies mediate evolutionary processes related to adaptation and speciation.

Six of 16 genetic parameters associated with genetic patterns and processes were associated with habitat in the GLMMs, as were 8 of 16 in the PGLS analyses, greater than what would be expected by chance in permutation tests. Most parameters differing between upland and floodplain species were summary statistics reflecting patterns of variation across all loci. Model-based estimates of population genetic structure and demographic parameters showed fewer associations. The low variance in the number of groups from analyses of population genetic structure may contribute to low power to detect differences in those parameters. Higher-level demographic parameters, such as divergence time and migration rate, are notoriously difficult to estimate accurately (Myers et al. 2008; Strasburg and Rieseberg 2010; Schraiber and Akey 2015), and estimates can be spurious when genetic variation is affected

by unmodeled processes, such as natural selection (Hahn 2008). Purifying selection in the conserved loci examined here may reduce the accuracy of parameter estimates, including those estimated under the multispecies coalescent model in G-PhoCS. Because we used the same conserved markers in each study species, we do not expect purifying selection to result in a bias across species in demographic estimates, but it may reduce our ability to detect patterns in these parameters. However, recent evidence suggests that UCEs provide a better fit to some evolutionary models that assume neutrality than coding loci (Reddy et al. 2017) and that UCEs perform reasonably under the multispecies coalescent, particularly with the addition of more sequence from variable flanking regions (Meiklejohn et al. 2016). Our ability to detect genetic differences between floodplain and upland forest species will surely improve with more complete data sets and better models, but our current data and methods are at least sufficient to identify habitat associations in a modest range of genetic parameters.

We found differences between floodplain and upland forest species in some processes that might allow reconstruction of the mechanisms responsible for differences in diversity between the habitats. Greater dispersal over ecological timescales in floodplain species could explain their lower levels of diversity and divergence with respect to upland forest species. Birds of the forest interior are less likely to cross openings than birds of forest edges (Laurance et al. 2004). Seasonal movements are more frequent in birds of edge habitats (Levey and Stiles 1992), and seasonal flooding may annually force some floodplain birds into upland forest, promoting the movement of individuals into new areas (Rosenberg 1990). Rivers—important barriers to dispersal in Amazonia—could be less effective dispersal barriers to floodplain species than to upland species (Capparella 1987; Patton and da Silva 1998; Hayes and Sewlal 2004). Uplands may not occur within several kilometers of the main channel (Hess et al. 2015), potentially augmenting the significance of river barriers for upland bird species. River capture events, in which shifts in river courses result in land moving from one bank to the other, may regularly result in the passive movement of patches of floodplain habitat (Salo et al. 1986; Dumont 1991) and associated organisms (Tuomisto and Ruokolainen 1997; Patton et al. 2000) across river barriers, but river capture events involving upland forest may be less frequent (but see Almeida-Filho and Miranda 2007). We did find that floodplain forest species had lower F_{ST} values than upland forest species, consistent with higher rates of migration under an island model. However, we did not find higher migration rates in floodplain species than in upland forest species in our demographic models (table S9). Genomic data sets including information on linkage and improved population genetic methods for estimating migration may be required to detect concerted differences in patterns of gene flow in these habitats that are associated with the differences in genetic divergence. However, the F_{ST} results provide some evidence that ecological mechanisms are responsible for differences in population divergence between habitats.

Differences in population size, population fluctuations through time, or the time a species has been present in the landscape could also explain differences in diversity and divergence between habitats. Floodplains are currently relatively restricted in the Amazon basin, where they cover about 14% of the lowland area (Hess et al. 2015). The small area in floodplains may constrain population sizes in floodplain species, leading to lower genetic diversity and fewer opportunities for population divergence. Consistent with this hypothesis, we found lower values of Watterson's θ, which scales with effective population size, in floodplain species. However, this estimate includes all individuals across populations in each species and thus could reflect divergence between populations as well as within them. Estimates of within-population θ did not show similar differences between habitats. Sea level rise associated with climatic changes may have reduced the extent of available terrestrial floodplain habitats during the Quaternary period (Latrubesse and Franzinelli 2002; Irion et al. 2009), and recent expansion following these or other events could also help explain lower genetic diversity in floodplain species (Matocq et al. 2000; Aleixo 2002, 2006). Low Tajima's D values are expected under recent population expansion, but Tajima's D values were, if anything, higher in floodplain than in upland forest species. There was no difference between floodplains and uplands in the change in population size between the root population and extant populations in G-PhoCS. We note, however, that signals of expansion may be difficult to detect in conserved loci because purifying selection results in a similar excess of rare alleles to expansion (Hahn et al. 2002). Overall, our evidence for larger or more stable populations in floodplain species is weak, but linkage information combined with better methods of tracking population size changes through time may provide more information.

Population divergence could also be affected by the stability of habitats over geological time and the time that they have been occupied by the species of interest. Recent colonization or expansion across a habitat might lead to low divergence in floodplain species. Consistent with this hypothesis, we found that the average gene tree height of floodplain species was shallower than that in upland species. We did not see similar differences in age on the basis of mitochondrial crown ages or the age of the deepest population divergence from G-PhoCS. However, this result does provide some confirmation of prior evidence that the greater divergence in upland species may be attributable to historical factors, such as the age and stability of populations in that habitat (Smith et al. 2014b).

The evidence described above implicates a combination of ecological and historical mechanisms in differences in diversity and divergence between floodplain and upland birds, but detailed characterization of the mechanisms is still challenging. In addition to the shortcomings listed above for individual parameters, we are limited by issues of identifiability. Many processes affect population genetic variation similarly, making them difficult to distinguish (Myers et al. 2008; Strasburg and Rieseberg 2010). We expect that improvements will come from larger genomic data sets that include linkage information, but we also expect them to come from improved methods that can incorporate diverse processes in a single modeling framework. In addition, nongenomic data may also be needed to address some processes. Developments in animal tracking now permit estimation of dispersal at large spatial and temporal scales (Kays et al. 2015), and large-scale mating assays and observational studies can permit estimation of selection and fitness across a broad area (e.g., Yoder et al. 2014). These estimates could validate or even supplant estimates of these processes from population genomic data. In addition, although substantial progress has been made in understanding the geological history of the Amazon basin (Hoorn and Wesselingh 2011), matching genetic inferences with landscape events is still a challenge (Harvey and Brumfield 2015). This could be addressed with more precise estimates of landscape history combined with better strategies for assigning absolute times to events estimated from genomic data. Although we were able to date events inferred in the mitochondrial tree, improved substitution models and rate parameters need to be developed for UCEs and conserved exons—or perhaps for a subset of them that exhibit clocklike evolution. Finally, the comparative methods used in our study (GLMM and PGLS) are useful, but comparative analyses that allow simultaneous estimation of differences in multiple response variables that covary according to an explicit population genetic or evolutionary model are desirable.

As a result of our focus on comparative analyses, we have barely probed the details of population genetic patterns within individual species (but see figs. S14–S20). Upland forest species, on average, exhibited greater genomic diversity, deeper history, and greater divergence than floodplain species in all significant comparisons. The deep genetic divergences observed in many upland forest species coincided roughly with rivers that represent major putative biogeographic barriers for terrestrial Amazonian species (Cracraft 1985; da Silva et al. 2005). Higher-resolution studies are warranted within particular upland forest species to better characterize intraspecific diversity and determine whether populations merit recognition as full species. In particular, Variegated Tinamou (*Crypturellus variegatus*), Rufous-capped Antthrush (*Formicarius colma*), Spot-backed Antbird (*Hylophylax naevius*), Sooty Antbird (*Hafferia for-*

tis), Black-faced Antbird (*Myrmoborus myotherinus*), and Straight-billed Hermit (*Phaethornis bourcieri*) contained deep divergences within currently recognized species. Studies of some of these species have recently been completed (e.g., Fernandes et al. 2014; Araújo-Silva et al. 2017) or are under way.

However, some species pairs in all comparisons contained higher diversity or divergence in the floodplain species. Genera with consistently higher diversity or divergence in the floodplain species included *Piaya*, *Formicarius*, *Synallaxis*, and *Saltator*. In *Piaya*, there was no notable divergence within the upland forest species (Black-bellied Cuckoo [*P. melanogaster*]) and weak divergence in the floodplain forest species (Squirrel Cuckoo [*P. cayana*]). In the last three cases, however, a single deep divergence was present in the floodplain forest species. In Black-faced Antthrush (*Formicarius analis*) and Grayish Saltator (*Saltator coerulescens*), a highly divergent population was present in the Guianan region, whereas in Plain-crowned Spinetail (*Synallaxis gujanensis*) it was in the southwestern Amazon near the foot of the Andes. These species pairs demonstrate that the trend for greater diversity and divergence in upland species is not universal and support the idea that historical contingency and idiosyncrasy play an important role in determining patterns of intraspecific diversity across Neotropical bird species (Brumfield 2012).

Despite prior evidence that divergence in Neotropical birds is associated with forest stratum or dispersal ability (Burney and Brumfield 2009; Smith et al. 2014*b*), we detected little evidence for relationships between these traits and genomic diversity. The only relationship recovered between genetic parameters and forest stratum or Kipp's index involved higher nucleotide diversity on the Z chromosome relative to autosomes in understory species, a pattern driven largely by the two tinamous. This is surely due in part to low power resulting from our study design. Members of a genus in our study always occurred in the same forest stratum and exhibited similar Kipp's index values, and therefore the number of phylogenetically independent data points with a particular trait value was low. The forest canopy is in many ways analogous to edge habitats like those found in floodplains, and both are thought to harbor higher concentrations of birds that undergo seasonal movements than the interior of tall forest (Levey and Stiles 1992). Both canopy and floodplain bird species have lower subspecies richness than understory and upland forest species, respectively (Salisbury et al. 2012). We expect that a paired study design focused on these traits or larger sample sizes of species would find stronger associations between forest stratum and genetic parameters. Moreover, more ecological studies are needed to provide better characterizations of ecology than the simple binary variables examined here.

In summary, we have demonstrated that an ecological trait, habitat association, predicts variation across species in genetic diversity and divergence and in parameters related to ecological and historical processes. Birds in the interior of upland forest have greater diversity and more divergence across the landscape, and these may be a result of lower levels of gene flow as well as deeper histories in upland forest species. These habitat-associated differences in genetic parameters may reflect different propensities to respond to environmental change, form new species, and succumb to extinction. Interestingly, the upland forest avifauna contains more species (1,058) than the floodplain forest avifauna (154) in the Amazon basin (Parker et al. 1996). Because species proliferation is also tied to trait variation (Stanley 1975) and traits may affect population divergence and speciation similarly (Riginos et al. 2014), the evolutionary differences we detected between upland forest and floodplain/edge species may have played a role in producing their disparate diversities. Different conservation strategies may also be necessary to preserve the divergent patterns of genetic diversity and evolutionary processes observed in upland and floodplain regions. Practically, we have demonstrated that comparative genomic data sets can be used to estimate diverse parameters for testing hypotheses about traits associated with genomic diversity. Studies examining additional taxa, whole-genome data, improved methods for estimating and comparing genetic parameters, and data sources aside from genetic sequences are sure to expand our understanding of the effects of ecological traits on evolutionary patterns and processes in the future.

Acknowledgments

Literature Cited

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2014. spThin: functions for spatial thinning of species occurrence records for use in ecological models. R package version 0.1.0. Accessed April 15, 2017. https://CRAN.R-project.org/package=spThin.

Aleixo, A. 2002. Molecular systematics and the role of the "várzea"–"terra firme" ecotone in the diversification of *Xiphorhynchus* woodcreepers (Aves: Dendrocolaptidae). Auk 119:621–640.

———. 2006. Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). Biological Journal of the Linnean Society 89:383–395.

Almeida-Filho, R., and F. P. Miranda. 2007. Mega capture of the Rio Negro and formation of the Anavilhanas Archipelago, central Amazônia, Brazil: evidences in an SRTM digital elevation model. Remote Sensing of Environment 110:387–392.

Alonso, J. Á., M. R. Metz, and P. V. Fine. 2013. Habitat specialization by birds in western Amazonian white-sand forests. Biotropica 45:365–372.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and J. D. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389–3402.

Araujó-Silva, L. E., L. S. Miranda, L. Carneiro, and A. Aleixo. 2017. Phylogeography and diversification of an Amazonian understory hummingbird: paraphyly and evidence for widespread cryptic speciation in the Plio-Pleistocene. Ibis. doi:10.1111/ibi.12500.

Avise, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 63:62–76.

Bates, J. M. 2000. Allozymic genetic structure and natural habitat fragmentation: data for five species of Amazonian forest birds. Condor 102:770–783.

Benson, D. A., K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2014. GenBank. Nucleic Acids Research 42:D32–D37.

Bermingham, E., and C. Moritz. 1998. Comparative phylogeography: concepts and applications. Molecular Ecology 7:367–369.

Bi, K., D. Vanderpool, S. Singhal, T. Linderoth, C. Moritz, and J. Good. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13:403.

Bivand, R., T. Keitt, and B. Rowlingson. 2017. rgdal: bindings for the geospatial data abstraction library. R package version 1.2-6. Accessed April 15, 2017. https://CRAN.R-project.org/package=rgdal.

Bivand, R., and N. Lewin-Koh. 2017. maptools: tools for reading and handling spatial objects. R package version 0.9-2. Accessed April 15, 2017. https://CRAN.R-project.org/package=maptools.

Bouckaert, R. R., and A. J. Drummond. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evolutionary Biology 17:42.

Bouckaert, R. R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. Drummond. 2014. BEAST2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology 10:e1003537.

Brumfield, R. T. 2012. Inferring the origins of lowland Neotropical birds. Auk 129:367–376.

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

Campbell, D. G., D. C. Daly, G. T. Prance, and U. N. Maciel. 1986. Quantitative ecological inventory of terra firme and várzea tropical forest on the Rio Xingu, Brazilian Amazon. Brittonia 38:369–393.

Capparella, A. P. 1987. Effects of riverine barriers on genetic differentiation of Amazonian forest undergrowth birds. PhD diss. Louisiana State University, Baton Rouge.

Carling, M. D., and R. T. Brumfield. 2007. Gene sampling strategies for multi-locus population estimates of genetic diversity (θ). PLoS One 2:e160.

Carstens, B. C., R. S. Brennan, V. Chua, C. V. Duffie, M. G. Harvey, R. A. Koch, C. D. McMahan, et al. 2013. Model selection as a tool for phylogeographic inference: an example from the willow *Salix melanopsis*. Molecular Ecology 22:4014–4028.

Chavent, M., V. Kuentz, B. Liquet, and L. Saracco. 2012. ClustOfVar: an R package for the clustering of variables. Journal of Statistical Software 50:1–16.

Chen, C., E. Durand, F. Forbes, and O. François. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Molecular Ecology Notes 7:747–756.

Corander, J., P. Waldmann, and M. J. Sillanpää, M. J. 2003. Bayesian analysis of genetic differentiation between populations. Genetics 163:367–374.

Counterman, B. A., D. Ortíz-Barrientos, and M. A. Noor. 2004. Using comparative genomic data to test for fast-X evolution. Evolution 58:656–660.

Cracraft, J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: areas of endemism. Ornithological Monographs 36:49–84.

da Silva, J., M. Cardoso, A. B. Rylands, A. B. Fonseca, and A. Gustavo. 2005. The fate of the Amazonian areas of endemism. Conservation Biology 19:689–694.

Davey, J. W., P. A. Hohenlohe, P. D. Etter, J. Q. Boone, J. M. Catchen, and M. L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499–510.

del Hoyo, J., A. Elliott, J. Sargatal, and D. Christie, eds. 1992–2011. Handbook of the birds of the world. Lynx Edicions, Barcelona.

del Hoyo, J., A. Elliott, J. Sargatal, D. Christie, and E. de Juana, eds. 2017. Handbook of the birds of the world alive. Lynx Edicions, Barcelona. Accessed June 9, 2017. http://www.hbw.com.

Duminil, J., S. Fineschi, A. Hampe, P. Jordano, D. Salvini, G. G. Vendramin, and R. Petit. 2007. Can population genetic structure be predicted from life-history traits? American Naturalist 169:662–672.

Dumont, J. F. 1991. Cambio del curso de los ríos Marañon y Ucayali en la depresión Ucamara durante el Cuaternario Reciente. Pages 483–487 *in* Volumen de Resúmenes Extendidos, VII Congreso Peruano de Geologia, Lima.

Dunning, J. B. 2008. CRC handbook of avian body masses. 2nd ed. CRC, Boca Raton, FL.

Earl, D. A. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361.

Edwards, S., and P. Beerli. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54:1839–1854.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620.

Faircloth, B. C. 2013. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. Accessed December 20, 2013. http://dx.doi.org/10.6079/J9ILL.

———. 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61:717–726.

Fernandes, A. M., M. Wink, C. H. Sardelli, and A. Aleixo. 2014. Multiple speciation across the Andes and throughout Amazonia: the case of the Spot-backed Antbird species complex (*Hylophylax naevius/ Hylophylax naevioides*). Journal of Biogeography 41:1094–1104.

Gao, H., S. Williamson, and C. D. Bustamante. 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics 176:1635–1651.

Garrison, E., and G. Marth. 2012. Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907.

Gentry, A. H., and L. H. Emmons. 1987. Geographical variation in fertility, phenology, and composition of the understory of Neotropical forests. Biotropica 19:216–227.

Gianoli, E., C. Torres-Díaz, E. Ruiz, C. Salgado-Luarte, M. A. Molina-Montenegro, A. Saldaña, and R. S. Ríos. 2016. Woody climbers show greater population genetic differentiation than trees: insights into the link between ecological traits and diversification. Evolution 70:2736–2745.

Gladman, S. 2009. VelvetOptimiser. Accessed December 20, 2013. http://bioinformatics.net.au/software.velvetoptimiser.shtml.

Gronau, I., M. J. Hubisz, B. Gulko, C. G. Danko, and A. Siepel. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nature Genetics 43:1031–1034.

Hahn, M. W. 2008. Toward a selection theory of molecular evolution. Evolution 62:255–265.

Hahn, M. W., M. D. Rausher, and C. W. Cunningham. 2002. Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. Genetics 161:11–20.

Harrell, F. E. 2016. Hmisc: Harrell miscellaneous. R package version 3.17-4. Accessed November 5, 2016. https://CRAN.R-project.org/package=Hmisc.

Harvey, M. G., and R. T. Brumfield. 2015. Genomic variation in a widespread Neotropical bird reveals divergence, population expansion, and gene flow. Molecular Phylogenetics and Evolution 83:305–316.

Harvey, M. G., C. D. Judy, G. F. Seeholzer, J. M. Maley, G. R. Graves, and R. T. Brumfield. 2015. Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species. PeerJ 3:e895.

Harvey, M. G., G. F. Seeholzer, B. T. Smith, D. L. Rabosky, A. M. Cuervo, and R. T. Brumfield. 2017. A positive association between population genetic differentiation and speciation rates in New World birds. Proceedings of the National Academy of Sciences of the USA. doi:10.1073/pnas.1617397114.

Harvey, M. G., B. T. Smith, T. C. Glenn, B. C. Faircloth, and R. T. Brumfield. 2016. Sequence capture versus restriction site associated

DNA sequencing for shallow systematics. Systematic Biology 65: 910–924.

Hayes, F. E., and J. A. N. Sewlal. 2004. The Amazon River as a dispersal barrier to passerine birds: effects of river width, habitat and taxonomy. Journal of Biogeography 31:1809–1818.

Hedtke, S. M., M. J. Morgan, D. C. Cannatella, and D. M. Hillis. 2013. Targeted enrichment: maximizing orthologous gene comparisons across deep evolutionary time. PLoS One 8:e67908.

Hellberg, M. E. 2009. Gene flow among populations of marine animals. Annual Review of Ecology, Evolution, and Systematics 40:291–310.

Hess, L. L., J. M. Melack, A. G. Affonso, C. Barbosa, M. Gastil-Buhl, and E. M. Novo. 2015. Wetlands of the lowland Amazon basin: extent, vegetative cover, and dual-season inundated area as mapped with JERS-1 synthetic aperture radar. Wetlands 35:745–756.

Hijmans, R. J. 2016. raster: geographic data analysis and modeling. R package version 2.5-8. Accessed April 15, 2017. https://CRAN.R-project.org/package=raster.

Hillier, W. H., W. Miller, E. Birney, W. Warren, R. C. Hardison, C. P. Ponting, P. Bork, et al. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

Hoorn, C., and F. Wesselingh, eds. 2011. Amazonia, landscape and species evolution: a look into the past. Wiley, Chichester.

Irion, G., J. Müller, and J. Morais. 2009. The impact of Quaternary sea level changes on the evolution of the Amazonian lowland. Hydrological Processes 3172:3168–3172.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11:94.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Mammalian Protein Metabolism 3:21–132.

Katoh, K., K. I. Kuma, H. Toh, and T. Miyata. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Research 33:511–518.

Kays, R., M. Crofoot, W. Jetz, and M. Wikelski. 2015. Terrestrial animal tracking as an eye on life and planet. Science 348:aaa2478.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. Trends in Ecology and Evolution 17:19–23.

Kimball, R. T., E. L. Braun, F. K. Barker, R. C. Bowie, M. J. Braun, J. L. Chojnowski, W. A. Cox, et al. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution 50:654–660.

Kipp, F. A. 1959. Der Handflügel-Index als flugbiologisches Maß. Die Vogelwarte 20:77–86.

Kisel, Y., A. C. Moreno-Letelier, D. Bogarín, M. P. Powell, M. W. Chase, and T. G. Barraclough. 2012. Testing the link between population genetic differentiation and clade diversification in Costa Rican orchids. Evolution 66:3035–3052.

Kratter, A. W. 1997. Bamboo specialization by Amazonian birds. Biotropica 29:100–110.

Latch, E. K., G. Dharmarajan, J. C. Glaubitz, and O. E. Rhodes Jr. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics 7:295–302.

Latrubesse, E. M., and E. Franzinelli. 2002. The Holocene alluvial plain of the middle Amazon River, Brazil. Geomorphology 44:241–257.

Laurance, S. G., P. C. Stouffer, and W. F. Laurance. 2004. Effects of road clearings on movement patterns of understory rainforest birds in central Amazonia. Conservation Biology 18:1099–1109.

Leffler, E. M., K. Bullaughey, D. R. Matute, W. K. Meyer, L. Ségurel, A. Venkat, P. Andolfatto, and M. Przeworski. 2012. Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biology 10:e1001388.

Levey, D. J., and F. G. Stiles. 1992. Evolutionary precursors of long-distance migration: resource availability and movement patterns in Neotropical landbirds. American Naturalist 140:447–476.

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

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

Loveless, M. D., and J. D. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. Annual Review of Ecology and Systematics 15:65–95.

Matocq, M. D., J. L. Patton, and M. N. F. da Silva. 2000. Population genetic structure of two ecologically distinct Amazonian spiny rats: separating history and current ecology. Evolution 54:1423–1432.

Mayorga, E., M. G. Logsdon, M. V. R. Ballester, and J. E. Richey. 2012. LBA-ECO CD-06 Amazon River basin land and stream drainage direction maps. Accessed April 15, 2017. http://dx.doi.org/10.3334/ORNLDAAC/1086.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20:1297–1303.

Meiklejohn, K. A., B. C. Faircloth, T. C. Glenn, R. T. Kimball, and E. L. Braun. 2016. Analysis of a rapid evolutionary radiation using ultra-conserved elements: evidence for a bias in some multispecies coalescent methods. Systematic Biology 65:612–627.

Miraldo, A., S. Li, M. K. Borregaard, A. Flórez-Rodríguez, S. Gopalakrishnan, M. Rizvanovic, Z. Wang, C. Rahbek, K. A. Marske, and D. Nogués-Bravo. 2016. An Anthropocene map of genetic diversity. Science 353:1532–1535.

Myers, S., C. Fefferman, and N. Patterson. 2008. Can one learn history from the allelic spectrum? Theoretical Population Biology 73:243–348.

Nabholz, B., R. Lanfear, and J. Fuchs. 2016. Body mass–corrected molecular rate for bird mitochondrial DNA. Molecular Ecology 25:4438–4449.

Nevo, E., A. Beiles, and R. Ben-Shlomo. 1984. The evolutionary significance of genetic diversity: ecological, demographic, and life history correlates. Pages 13–213 in G. S. Mani, ed. Evolutionary dynamics of genetic diversity. Springer, Berlin.

Orme, D., R. Freckleton, G. Thomas, T. Petzoldt, S. Fritz, N. Isaac, and W. Pearse. 2013. caper: comparative analyses of phylogenetics and evolution in R version 0.5.2. Accessed October 20, 2015. http://CRAN.R-project.org/package=caper.

Pabijan, M., K. C. Wollenberg, and M. Vences. 2012. Small body size increases the regional differentiation of populations of tropical Mantellid frogs (Anura: Mantellidae). Journal of Evolutionary Biology 25:2310–2324.

Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. Ecological Applications 13:146–158.

Papadopoulou, A., and L. L. Knowles. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. Proceedings of the National Academy of Sciences of the USA 113:8018–8024.

Parker, T. A., III, D. F. Stotz, and J. W. Fitzpatrick. 1996. Ecological and distributional databases. Pages 113–436 *in* D. F. Stotz, J. W. Fitzpatrick, T. A. Parker III, and D. Moskovits, eds. Neotropical birds: ecology and conservation. University of Chicago Press, Chicago.

Patton, J. L., and M. N. F. da Silva. 1998. Rivers, refuges, and ridges: the geography of speciation of Amazonian mammals. Pages 202–213 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford University Press, Oxford.

Patton, J. L., M. N. F. da Silva, and J. R. Malcolm. 2000. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. Bulletin of the American Museum of Natural History 244:1–306.

Paz, A., R. Ibáñez, K. R. Lips, and A. J. Crawford. 2015. Testing the role of ecology and life history in structuring genetic variation across a landscape: a trait-based phylogeographic approach. Molecular Ecology 24:3723–3737.

Pearson, D. L. 1977. A pantropical comparison of bird community structure on six lowland forest sites. Condor 79:232–244.

Prance, G. T. 1979. Notes on the vegetation of Amazonia. III. The terminology of Amazonian forest types subject to inundation. Brittonia 31:26–38.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rambaut, A., and A. Drummond. 2007. Tracer version 1.5. Accessed February 5, 2015. http://beast.bio.ed.ac.uk/Tracer.

R Core Team. 2015. R: a language and environment for statistical computing. Accessed October 20, 2015. http://www.R-project.org.

Reddy, S. R., T. Kimball, A. Pandey, P. A. Hosner, M. J. Braun, S. J. Hackett, K. L. Han, et al. 2017. Why do phylogenomic data sets yield conflicting trees? data type influences the avian tree of life more than taxon sampling. Systematic Biology. doi:10.1093/sysbio/syx041.

Reich, D., K. Thangaraj, N. Patterson, A. L. Price, and L. Singh. 2009. Reconstructing Indian population history. Nature 461:489–494.

Remsen, J. V., Jr., J. I. Areta, C. D. Cadena, A. Jaramillo, M. Nores, J. F. Pacheco, J. Pérez-Emán, et al. 2015. A classification of the bird species of South America. American Ornithologists' Union. Accessed February 5, 2015. http://www.museum.lsu.edu/~Remsen/SACCBaseline.html.

Remsen, J. V., Jr., and T. A. Parker III. 1983. Contribution of river-created habitats to bird species richness in Amazonia. Biotropica 15:223–231.

Revell, L. J. 2011. Phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Riginos, C., Y. M. Buckley, S. P. Blomberg, and E. A. Treml. 2014. Dispersal capacity predicts both population genetic structure and species richness in reef fishes. American Naturalist 184:52–64.

Robinson, S. K., and J. Terborgh. 1995. Interspecific aggression and habitat selection by Amazonian birds. Journal of Animal Ecology 64:1–11.

Romiguier, J., P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, Y. Chiari, et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261–263.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539–542.

Rosenberg, G. H. 1990. Habitat specialization and foraging behavior by birds of Amazonian river islands in northeastern Peru. Condor 92:427–443.

Salisbury, C. L., N. Seddon, C. R. Cooney, and J. A. Tobias. 2012. The latitudinal gradient in dispersal constraints: ecological specialization drives diversification in tropical birds. Ecology Letters 15:847–855.

Salo, J., R. Kalliola, I. Häkkinen, Y. Mäkinen, P. Niemelä, M. Puhakka, and P. D. Coley. 1986. River dynamics and the diversity of Amazon lowland forest. Nature 322:254–258.

Schraiber, J. G., and J. M. Akey. 2015. Methods and models for unravelling human evolutionary history. Nature Reviews Genetics 16:727–740.

Schulenberg, T. S., D. F. Stotz, D. F. Lane, J. P. O'Neill, and T. A. Parker III. 2010. Birds of Peru: revised and updated edition. Princeton University Press, Princeton, NJ.

Smith, B. T., M. G. Harvey, B. C. Faircloth, T. C. Glenn, and R. T. Brumfield. 2014a. Target capture and massively parallel sequencing of ultraconserved elements (UCEs) for comparative studies at shallow evolutionary timescales. Systematic Biology 63:83–95.

Smith, B. T., J. E. McCormack, A. M. Cuervo, M. J. Hickerson, A. Aleixo, C. D. Cadena, J. E. Pérez Eman, et al. 2014b. The drivers of tropical speciation. Nature 515:406–409.

Smith, J. V., E. L. Braun, and R. T. Kimball. 2012. Ratite non-monophyly: independent evidence from 40 novel loci. Systematic Biology 62:35–49.

Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology 15:4261–4293.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Stanley, S. M. 1975. A theory of evolution above the species level. Proceedings of the National Academy of Sciences of the USA 72:646–650.

Strasburg, J. L., and L. H. Rieseberg. 2010. How robust are "isolation with migration" analyses to violations of the IM model? a simulation study. Molecular Biology and Evolution 27:297–310.

Sukumaran, J., and M. T. Holder. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26:1569–1571.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: a citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282–2292.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460.

———. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Terborgh, J. 1985. Habitat selection in Amazonian birds. Pages 311–338 *in* M. L. Cody, ed. Habitat selection in birds. Academic Press, New York.

Tuanmu, M. N., and W. Jetz. 2015. A global, remote sensing–based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. Global Ecology and Biogeography 24:1329–1339.

Tuomisto, H., and K. Ruokolainen. 1997. The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. Biodiversity and Conservation 6:347–357.

Wall, J. D., K. E. Lohmueller, and V. Plagnol. 2009. Detecting ancient admixture and estimating demographic parameters in multiple human populations. Molecular Biology and Evolution 26:1823–1827.

Wang, N., E. L. Braun, and R. T. Kimball. 2012. Testing hypotheses about the sister group of the Passeriformes using an independent 30-locus data set. Molecular Biology and Evolution 29:737–750.

Warren, W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Künstner, S. Searle, et al. 2010. The genome of a songbird. Nature 464:757–762.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7:256–276.

Willi, Y., J. V. Buskirk, and A. A. Hoffmann. 2006. Limits to the adaptive potential of small populations. Annual Reviews in Ecology, Evolution, and Systematics 37:433–358.

Willing, E. M., C. Dreyer, and C. Van Oosterhout. 2012. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. PLoS One 7: e42649.

Wittmann, F., W. J. Junk, and M. T. Piedade. 2004. The várzea forests in Amazonia: flooding and the highly dynamic geomorphology interact with natural forest succession. Forest Ecology and Management 196: 199–212.

Yoder, J. B. 2014. Genomic signature of adaptation to climate in *Medicago truncatula*. Genetics 196:1263–1275.

Zerbino, D. R., and E. Birney. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research 18: 821–829.

Zucker, M. R., M. G. Harvey, J. A. Oswald, A. Cuervo, E. Derryberry, and R. T. Brumfield. 2016. The Mouse-colored Tyrannulet (*Phaeomyias murina*) is a species complex that includes the Cocos Flycatcher (*Nesotriccus ridgwayi*), an island form that underwent a population bottleneck. Molecular Phylogenetics and Evolution 101:294–302.

Associate Editor: Rebecca J. Safran
Editor: Yannis Michalakis

The western Amazon seen from the foot of the Peruvian Andes. Photo credit: Michael Harvey.