



# Phylogeographic and demographic patterns reveal congruent histories in seven Amazonian White-Sand ecosystems birds

João Marcos Guimarães Capurucho<sup>1,2,3</sup> | Mary V. Ashley<sup>1</sup> | Cintia Cornelius<sup>4</sup> | Sergio H. Borges<sup>4</sup> | Camila C. Ribas<sup>3</sup> | John M. Bates<sup>2</sup>

<sup>1</sup>Department of Biological Sciences, University of Illinois at Chicago, 60607, Illinois, Chicago, 845 W. Taylor Street, USA

<sup>2</sup>Life Sciences Section, Negaunee Integrative Research Center, The Field Museum of Natural History, 60605, Illinois, Chicago, 1400 S. Lake Shore Drive, USA

<sup>3</sup>Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, AM, Manaus, Av. André Araújo 2936, Aleixo, Brazil

<sup>4</sup>Instituto de Ciências Biológicas, Universidade Federal do Amazonas, AM, Manaus, Av. Rodrigo Otávio Jordão Ramos 3000, Bloco ICB01, Setor Sul, Brazil

## Correspondence

João Marcos Guimarães Capurucho, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo 2936, Aleixo, Manaus, AM, Brazil.  
Email: [jcapurucho@fieldmuseum.org](mailto:jcapurucho@fieldmuseum.org)

## Funding information

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; FAPESP; Fundação de Amparo à Pesquisa do Estado do Amazonas; National Science Foundation

Handling Editor: V. V. Robin

## Abstract

**Aim:** The drivers of genetic diversity in Amazonia, the most species-rich set of ecosystems on Earth, are still incompletely understood. Species from distinct Amazonian ecosystems have unique biogeographic histories that will reflect regional landscape and climatic drivers of genetic diversity. We studied bird species from patchy Amazonian white-sand ecosystems (WSE) to evaluate the occurrence of shared biogeographic patterns to better understand the complex environmental and landscape history of Amazonia and its biodiversity.

**Location:** Northern South America; Amazonia.

**Taxon:** Passeriformes.

**Methods:** We sequenced Ultra-conserved Elements (UCEs) from 177 samples of seven bird species associated with WSE that have overlapping ranges. We used the SNP matrices and sequence data to estimate genetic structure and migration surfaces using 'conStruct' and *eems*, performed model-selection to obtain the most probable demographic histories on 'PipeMaster' and implemented analyses of shared demography with *ecoevolity*.

**Results:** Shallow genetic structure patterns varied among species. The Amazon river was the only barrier shared among them. Population structure dates to no more than 450,000 years ago. Nine geographically structured populations showed signals of population size changes and eight of these occur in Northern Amazonia. Population expansion was inferred at two distinct times: ~100,000 and ~50,000 years ago. The timing of co-expanding populations is consistent with differences in habitat preference, as species that prefer dense scrubby to forested vegetation expanded more recently compared to species that prefer open vegetation.

**Main conclusions:** WSE species responded in concert to environmental and landscape changes that occurred in the relatively recent past. Population expansions were likely driven by the genesis of new WSE patches and a return to wetter conditions after glacial periods. Pleistocene climatic cycles affected the distribution and dynamics of open vegetation habitats in Amazonia, especially in the Northern region, driving genetic diversity and demographic patterns of its associated biota.

## KEYWORDS

Campinas, demographic modelling, Neotropics, Pleistocene climatic cycles, population expansion, Quaternary

## 1 | INTRODUCTION

The evolution of Amazonian biota, which harbours the highest biodiversity in the world, has always fascinated evolutionary biologists. The initial debates were largely driven by Haffer's (1969) Refuge hypothesis, which proposed that allopatric divergence through cycles of population isolation during dry Pleistocene climatic periods was the key driver for the megadiverse bird fauna of Amazonia. The Refuge hypothesis, as originally proposed, has been largely dismissed as a general explanation for diversification in Amazonia (Bush, 1994; Rull, 2011). The debate developed into a view that acknowledged that a single process or hypothesis would not be able to encompass all or even most of Amazonian evolutionary complexity (Rull, 2011; Tuomisto, 2007). In the last decade, however, evidence for effects of Pleistocene climatic cycles on habitat distribution and in shaping Amazonia biota has re-emerged, not as the only driver of diversification but by affecting genetic diversity patterns, causing local extinction, and population size fluctuations (Capuruchó et al., 2013; Fouquet et al., 2012; Sato et al., 2021; Silva et al., 2019).

There also has been an increased recognition that taxa specialized on different habitat types or forest strata can have distinct evolutionary histories and respond differently to common regional environmental and landscape characteristics and their changes through time (Burney & Brumfield, 2009; Harvey et al., 2017; Papadopoulou & Knowles, 2016). As such, differences in diversification and genetic structure patterns among species from a regional biota may often be studied from the perspective of species' habitat preferences (Burney & Brumfield, 2009), dispersal capability (Capuruchó, Ashley, et al., 2020), reproductive strategy and other traits. Thus, common drivers of diversification in Amazonia are more likely for species sharing habitat preferences (Papadopoulou & Knowles, 2016), and studies comparing species adapted to specific Amazonian ecosystems may reveal evidence of shared evolutionary history, with deviations being driven by species-specific traits. Indeed, a recent comparative study has found that there is a prevalence of older *terra firme* bird lineages in north and western Amazonia, while most recent diversification events are concentrated in the southeastern region, reflecting the history of this habitat and of the drainage system in Amazonia (Silva et al., 2019).

Interspersed among the *terra-firme* forests of Amazonia are white-sand ecosystems (WSE) that occur patchily over nutrient poor sandy soils and harbour a unique component of Amazonian species diversity, despite their relatively limited area and modest species richness (Alonso & Whitney, 2003; Anderson, 1981; Borges, Cornelius, Moreira, et al., 2016; Costa et al., 2020; Oliveira-Filho et al., 2021; Prance, 1996). The sandy soils that sustain WSE originated from multiple geological processes and patches of various ages are found scattered across Amazonia (Capuruchó, Borges, et al., 2020).

Characteristics including low nutrient availability, seasonal flooding, and varying water table levels have driven the origin and adaptation of a specialized and endemic biota (Anderson, 1981; Capuruchó, Borges, et al., 2020; Damasco et al., 2013; Fine et al., 2004). The lower diversity compared to other Amazonian ecosystems probably is the result of harsh conditions and patchy distribution, which in turn also have generated high endemism, especially in plant groups (Anderson, 1981; Costa et al., 2020; Prance, 1996; Vicentini, 2016).

The distinct bird communities of WSE contribute to beta diversity of the Amazonian avifauna (Alonso et al., 2013; Borges, 2004; Borges, 2013). Borges, Cornelius, Ribas, et al. (2016) identified 35 bird species that are restricted or near-restricted to WSE. These species have overlapping ranges and might share a common biogeographic history in response to WSE dynamics over time. As connectivity of WSE populations may reflect both the past availability of non-forested habitats and changes in the surrounding forested ecosystems, phylogeographic studies of WSE birds can provide insight into Amazonian paleoenvironmental history.

Previous studies show that some WSE birds have experienced demographic fluctuations as a result of climatic changes since the last glacial cycle (~20,000 years ago; Capuruchó et al., 2013; Matos et al., 2016; Ritter, Coelho, et al., 2021). Using single or a few molecular markers, demographic expansions at the end of the last glacial cycle were inferred, coinciding with more suitable climatic conditions and genesis of new WSE patches (Capuruchó et al., 2013, 2018; Horbe et al., 2004; Matos et al., 2016; Ritter, Coelho, et al., 2021; Rossetti, Bertani, et al., 2012; Rossetti, Zani, et al., 2012; Zular et al., 2019). These previous studies investigated four species in total and used a small number of genetic markers, limiting broader generalizations of the observed patterns among WSE birds.

Based on target capture of ultraconserved elements (UCEs), we use a comparative framework to better understand the population genomic patterns of seven WSE birds and their phylogeographic and demographic history. We use this approach to study these patterns and their connection to the landscape and climatic history of Amazonia. Through more extensive sampling of species and their genomes and the application of modelling and comparative phylogeographic analyses, we aimed to answer the following questions: (1) Do these seven WSE species share a similar genetic structure indicating that they have responded to the same barriers through time? (2) How does the timing of population subdivision for WSE birds compare with that reported for birds in other Amazonian ecosystems? (3) Are there shared demographic histories for WSE birds during the Quaternary climatic cycles? (4) Can differences in demographic histories of WSE specialists be explained by differences in their habitat preferences? (5) What can we learn about climatic and environmental changes in Amazonia from the phylogeographic and demographic stories of WSE birds?

## 2 | MATERIALS AND METHODS

### 2.1 | Species sampling

Following Alonso et al. (2013) and Borges, Cornelius, Ribas, et al. (2016) we selected seven bird species considered to be near-restricted to WSE (sensu Borges, Cornelius, Ribas, et al., 2016): Yapacana Antbird (*Aprositornis disjuncta*, Thamnophilidae; 9 samples), Rufous-capped Elaenia (*Elaenia ruficeps*, Tyrannidae; 21 samples), Saffron-crested Manakin (*Neopelma chrysocephalum*, Pipridae; 23 samples), Yellow-crowned Manakin (*Heterocercus flavivertex*, Pipridae; 21 samples), Black Manakin (*Xenopipo atronitens*, Pipridae; 51 samples), Red-shouldered Tanager (*Tachyphonus phoenicius*, Thraupidae; 39 samples) and White-naped Seedeater (*Sporophila fringilloides*, Thraupidae; 13 samples) (Table S1). These species have overlapping ranges, but they differ in habitat preferences, degree of specialization to WSE and range size, from the range-restricted *A. disjuncta* to the widespread *X. atronitens* and *T. phoenicius*. *Neopelma chrysocephalum* and *H. flavivertex* are more commonly found in forested and *X. atronitens* in dense scrubby habitats, while the range-restricted *A. disjuncta* and *S. fringilloides*, and widespread *E. ruficeps* and *T. phoenicius* prefer the more open vegetation areas with sparse pockets of shrubs and small trees. *Heterocercus flavivertex* is the least restricted to the WSE occurring also in riverine habitats like *igapós* especially those on sandy soils (e.g. Hilty, 2003). Loans of genetic material (tissue and blood samples) from these species were obtained from multiple collections (Table S1).

### 2.2 | UCE data collection and processing

We used an enrichment method with a custom probe set targeting 2321 ultraconserved elements (UCEs) loci (Harvey et al., 2017). Genomic DNA was extracted using DNeasy Blood and Tissue kits (Qiagen®). Genomic DNA was then sent to Rapid Genomics® where UCEs were sequenced following Faircloth et al. (2012). The method relies on Illumina Nextera library preparation kits (Epicentre Biotechnologies) and AMPure XP Beads for PCR product cleaning (Faircloth et al., 2012) and libraries were sequenced using 150bp paired-end reads in Illumina HiSeq 2500.

We used the *phyluce* v1.5.0 pipeline to process the UCEs from the raw reads (Faircloth, 2016; Faircloth et al., 2012). First, raw reads were cleaned from adapters and low quality bases using *illumiprocessor* and *trimmomatic* using *phyluce* v1.5.0 default parameters (Bolger et al., 2014; Faircloth et al., 2012). Next, the data were assembled into contigs using *trinity* v2.5.1 and called from within the *phyluce* pipeline using *phyluce\_assembly\_assemblo\_trinity* with default parameters (--KMER\_SIZE set to 25) (Grabherr et al., 2011). We then used the *phyluce* code to find, extract and align (using *mafft* v7.13; Katoh & Standley, 2013) the UCE loci from the assembled reads. This process was done separately for each species and generated species-specific alignments of every UCE locus.

To generate reference sequences for variant calling and single nucleotide polymorphisms (SNP) extraction, we performed contig

assembly for each species separately using *itern* v1.1.1 (<https://github.com/faircloth-lab/itern>; Faircloth et al., 2012), which performs better with target enrichment data than other commonly used assembly programs. We randomly selected six samples from each species to perform the assembly using *itern* and used the custom probe set as seed. With the assembled data we generated species-specific consensus sequences for every UCE locus. The specific consensus sequences were then used as references for SNP calling. We followed Cooper et al. (2021) protocol with modifications to obtain SNP data (we performed additional comparisons among SNP matrices at the final recalibration rounds and updated flags and functions deprecated in GATK v3.8-1-0). Raw Illumina reads were indexed and mapped to the reference consensus sequences using *bwa-mem* (Li, 2013). Reads were sorted using *samtools* (Li et al., 2009) and duplicates were removed using *picard* (<http://broadinstitute.github.io/picard/>). We used Genome Analysis Toolkit v3.8-1-0 (GATK; Van der Auwera et al., 2013), to realign reads around indels, perform variant calling and indel removal, and employ quality-filters to the data. Next, we used the resulting SNP matrix to recalibrate the BAM files and perform additional rounds of SNP calling. Recalibration rounds were performed up to four rounds or until the total number of different SNPs between matrices was lower than 10. All bioinformatic pipelines and analyses were implemented at the Grainger Bioinformatics Center of the Field Museum of Natural History.

### 2.3 | Genetic structure and migration surfaces

The final VCF file was processed using *vcftools* v0.1.13 (Danecek et al., 2011) following similar procedures as those implemented by Winker et al. (2018) and Thom et al. (2018) studies using UCEs. We applied quality filters (--minGQ 10), removed missing data from the matrix (--max-missing 1) and SNPs not in Hardy-Weinberg Equilibrium (--hwe 0.05), and removed Z-linked UCE loci after blasting the UCEs to the Zebra Finch (*Taeniopygia guttata*; GCF\_000151805.1) genome.

We used PGDSpider v2.1.1.5 (Lischer & Excoffier, 2012) to generate Structure v2.3.4 (Pritchard et al., 2000) input files from VCF SNP matrices. To conduct population structure analyses, we used 'conStruct' v1.0.4 (Bradburd et al., 2018) which implements both non-spatial and spatial analyses by taking into account isolation-by-distance expectations. Before converting the input files, we pruned the SNP matrix using the - mac 2 option on *vcftools* v0.1.13 to exclude singletons, therefore retaining SNPs with a minimum allele count of 2 (two heterozygous or one homozygous individual), and thinned the data based on the per species mean length of UCEs to obtain a single SNP per UCE locus. We compared K values from 1 to 5 (for *A. disjuncta*, *N. chrysocephalum*, *H. flavivertex* and *S. fringilloides*; due their smaller ranges, being mostly restricted to northern Amazonia) and 1 to 10 (for *E. ruficeps*, *X. atronitens* and *T. phoenicius*; which are more widely distributed) from spatial and non-spatial models by performing a cross-validation analysis using 20,000 iterations and 10 replicates per K value. Following Bradburd et al. (2018),

the best K value was selected based on the cross-validation results, layer contributions higher than 5% and observed meaningful geographic structuring.

We used *eems* (Estimated Effective Migration Surfaces; <https://github.com/dipetkov/eems>; Petkova et al., 2016) to estimate migration surfaces for the study species. With the same VCF matrix used for genetic structure analyses we used *plink* v1.9 (Purcell et al., 2007) to generate BED files with *make-bed* function without a chromosome map (*allow-extra-chr* 0). The BED files were used to compute pairwise genetic dissimilarities using *bed2diffs* program available with the *eems* package. After  $2 \times 10^6$  burn-in, all analyses were run for  $10^6$ – $10^8$  steps in the MCMC chain (depending on the number of samples and range size of the species), with parameters saved every  $10^4$  steps. We ran the analyses multiple times for every species to compare results and adjust parameters to improve convergence of the MCMC chain following the developers' recommendations.

## 2.4 | Phylogenetic analyses

For species in which the number of populations was equal to or higher than three, we inferred phylogenetic hypotheses using SNAPP (Bryant et al., 2012), implemented in BEAST v2.6.2 (Bouckaert et al., 2019), based on the SNP matrices trimmed for population structure analyses. Lineages were defined based on genetic clustering results and admixed individuals were assigned to the cluster with which each one had the highest shared ancestry. We ran the MCMC chain for  $2.5 \times 10^6$  generations and parameters were logged every 500 steps. Two independent runs were performed to check for convergence and results were combined after a 10% burn-in.

## 2.5 | Demographic modelling

To generate matrices of phased sequence data, we used the *phyluce* pipeline described in Andermann et al. (2018). The matrices of aligned phased UCEs were used for demographic modelling and for inferring shared demographic events. We included admixed individuals and used matrices with a maximum of 25% missing data (the maximum number of missing samples per UCE locus) and removed gaps using customized commands and *trimal* v1.4.rev15 (Capella-Gutierrez et al., 2009). We double-checked the sequences using Geneious v2019.0.4 (<https://www.geneious.com>) to confirm that all gaps had been removed.

For each study species, we simulated demographic models and estimated model-fit after classification using supervised machine-learning (SML) using the R packages 'PipeMaster' v0.2.3 (Gehara et al., 2020) and 'caret' v6.0.86 (Kuhn, 2020) following recommendations at [https://github.com/gehara/PipeMaster/blob/master/PipeMaster\\_tutorial.md](https://github.com/gehara/PipeMaster/blob/master/PipeMaster_tutorial.md). We used results from population structure analyses to assign individuals to populations (i.e. individuals were assigned to the cluster with which they had more than 50% of admixture proportion) and the SNAPP topology to define the

models that were tested. The models had one (*A. disjuncta*, *H. flavivertex*), two (*N. chryscephalum*, *S. fringilloides*), three (*E. ruficeps*, *T. phoenicius*), and five (*X. atronitens*) populations (for the full set of tested models see electronic SM; Figures S1–S4). We generated observed summary statistics from the data and simulated  $10^5$  sets of summary statistics per model using *msABC* called from within the 'PipeMaster' package. The prior on the mutation rate was fixed at  $2.5 \times 10^{-9}$  substitutions per site per generation (set as a uniform distribution with minimum and maximum values equal to the mutation rate) (Nadachowska-Brzyska et al., 2015), and we assumed generation time to be of one year. Model inference with SML used 75% of the simulated data for training and 25% for testing and evaluating accuracy. In a hierarchical procedure for species with two populations or more, we first inferred the best model comparing those including gene flow with those with complete isolation. Based on the best model, we tested whether models including population size changes would better fit the data than a model with constant population sizes. After finding the most probable model, we simulated  $5 \times 10^5$  sets of summary statistics using 'abc' v2.1 (Csilléry et al., 2012) and estimated parameters by using tolerance rates of 2, 5 and 10% for comparison. The higher number of retained simulations lead to negligible differences or wider confidence intervals and weaker correlations between simulated and estimated values, thus we present the results using the 2% tolerance rate ( $10^4$  sets of simulated data) to estimate parameter values.

Using the same sequence matrices and the populations for which the best inferred models included changes in population size, we tested for the occurrence of shared demographic events using *ecoevolity* v0.3.2 (Oaks, 2019; Oaks et al., 2020). Individuals with more than 25% admixture were removed from these analyses due to uncertainty in their population assignments. We used a gamma distribution as the hyperprior on the concentration parameter of the Dirichlet process prior (shape=10; scale=0.3) resulting in events with more categories having a higher overall probability. A uniform prior was set for event timings ( $10^3$  –  $3 \times 10^5$ ) and population sizes ( $10^3$  –  $5 \times 10^6$ ) based on the results from the demographic modelling. We used the same mutation rate and generation time from model inference analyses. We performed runs with different priors on the *root\_relative\_population\_size* parameter using an exponential distribution with mean 0.1, 0.5, and 1.0, and a uniform distribution with minimum and maximum values of 0.001 and 10.0, respectively. For each of these analyses we ran ten independent MCMC chains of 150,000 steps, sampling every 100 steps and checked for convergence of results. We combined the ten MCMC chains after applying a 10% burn-in, summarized the results, evaluated Effective Sample Sizes (ESS), and obtained posterior probabilities using *pycoevolity* *pyco-sumchains* and *sumcoevolity* functions.

## 3 | RESULTS

Our sample sizes ranged from nine (*A. disjuncta*) to 51 (*X. atronitens*) samples per species. The total number of UCEs recovered was similar

across species and varied from 2295 to 2309 (Table 1). After filtering the total number of unlinked SNPs varied from 722 (*X. atronitens*) to 1477 (*T. phoenicius*) (Table 1).

### 3.1 | Phylogenetic analyses and genetic structure

Overall, the results from population structure, phylogenetic reconstructions and migration surface analyses showed consistent patterns within each species (Figures S5–S11). Based on the 'conStruct' analyses, the numbers of inferred populations were one in *A. disjuncta* and *H. flavivertex*; two in *N. chrysocephalum* and *S. fringilloides*; three in *E. ruficeps* and *T. phoenicius*; and five in *X. atronitens* (Figure 1). For the four species with distributions including northern and southern Amazonia, there was a genetic break corresponding to the Amazon river (Figure 1). Other clusters were idiosyncratic in relation to geographic limits, revealing distinct structure patterns for each species (Figures S5–S11).

Identified genetic clusters received high support in the SNAPP analyses and the *eems* results showed areas of reduced gene flow in the regions between the genetic clusters. For four species (*A. disjuncta*, *N. chrysocephalum*, *H. flavivertex*, *X. atronitens*; Figures S5, S7–S9), there are additional areas of reduced gene flow identified in *eems* that did not correspond to major genetic breaks identified in the population structure analyses (Figures S5–S11), indicating additional fine scale genetic structuring.

The resulting maximum clade credibility tree topologies from SNAPP analyses in the three species had high node support. In all cases northern (*T. phoenicius*) or southern (*E. ruficeps* and *X. atronitens*) lineages were monophyletic (Figures S6, S9, and S10).

### 3.2 | Divergence times, gene flow and population size changes

Accuracy in identifying the correct model was generally high (Tables S2–S15), but for some species/models the values were lower (e.g. *A. disjuncta*, Table S2), yet the best model had distinctly higher probability. The observed summary statistics were within the parameter space of simulated summary statistics showing that we tested a

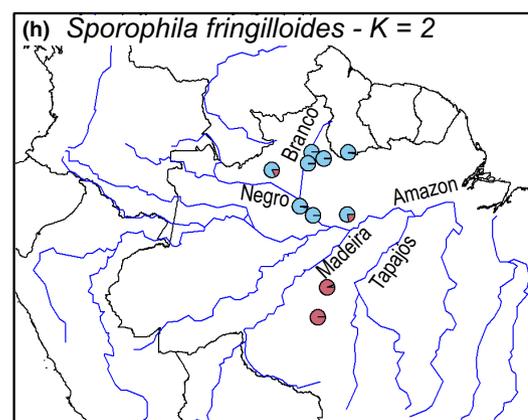
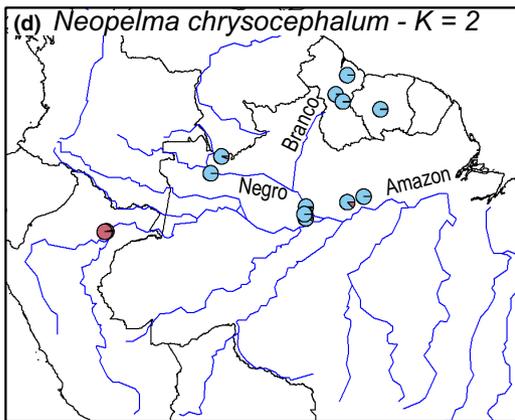
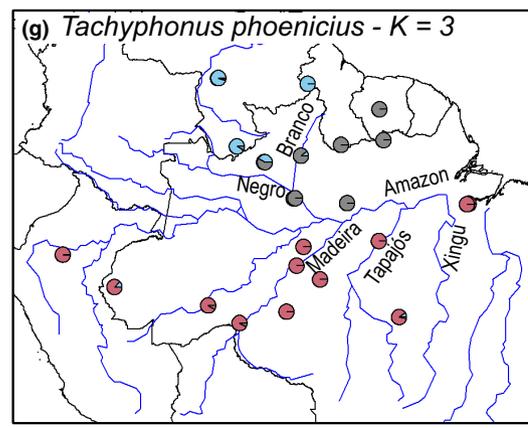
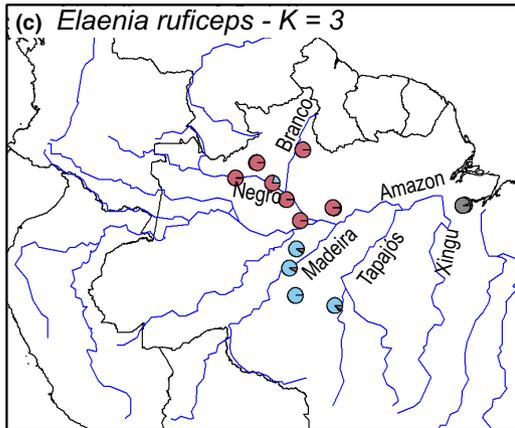
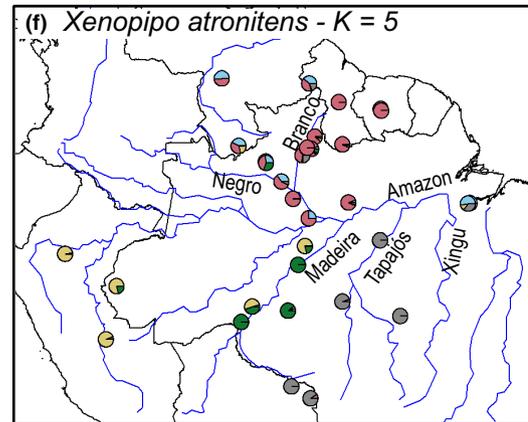
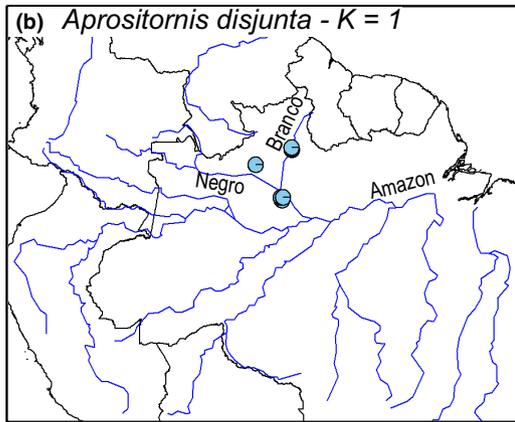
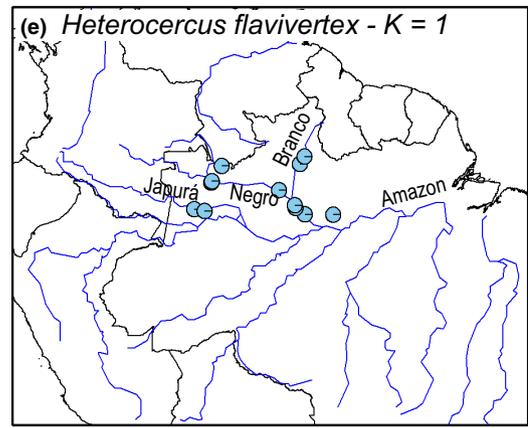
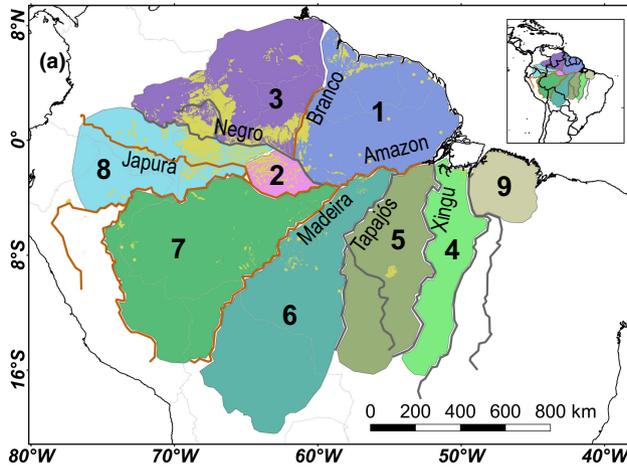
set of plausible models, except for *X. atronitens* which had the most complex model with five populations (Figures S13–S26). Therefore, we reduced the complexity of the *X. atronitens* model by grouping the four southern lineages into southwestern and southeastern clusters (Figure 1a, west and east of the Madeira river, respectively) following the SNAPP topology (Figure S9). This three-population arrangement (north, southwest, and southeast) simplified the models and simulated data had a better fit against the observed data. However, this solution led to downstream issues with parameter estimation that resulted in unreliable and negative parameter values, perhaps due to hidden genetic structure in merged populations. Thus, we discuss results based on the five-populations model.

For the five species with more than one population the models that included gene flow had a higher probability than models without it (Figure 2; Tables S2–S15). *Aprositornis disjuncta* and *H. flavivertex* had best models that included changes in population size. A model with unidirectional gene flow had a better fit in *N. chrysocephalum* (gene flow occurs from central Amazonia to the Iquitos region, Peru) and *S. fringilloides* (gene flow occurs from north to south of the Amazon river), and in both species, the central Amazonian population expanded recently (Figure 2). Gene flow across the Amazon river is also observed in *E. ruficeps*, in which the best model included demographic expansion for the northern population (Figure 2). In *X. atronitens*, the best model has no migration among southern populations, and migration only occurs among northern and southern populations, across the Amazon river (Figure 2). The best model for *T. phoenicius* also included demographic expansion for all populations, in addition to gene flow among all of them (Figure 2). Finally, except for the Peruvian genetic cluster in *N. chrysocephalum*, all populations occurring in northern Amazonia presented signals of population size changes, mainly expansions. Accuracy of the model inference and parameter estimates based on cross-validation varied substantially among species and could be related to both the complexity of the models (high number of parameters), a reduced number of available summary statistics to distinguish relatively similar models in the case of the single population models, and informativeness of the data in each species (Tables S2–S15). Most population subdivisions were estimated to have occurred within the last 300,000 years (nine out of ten; Figure S27).

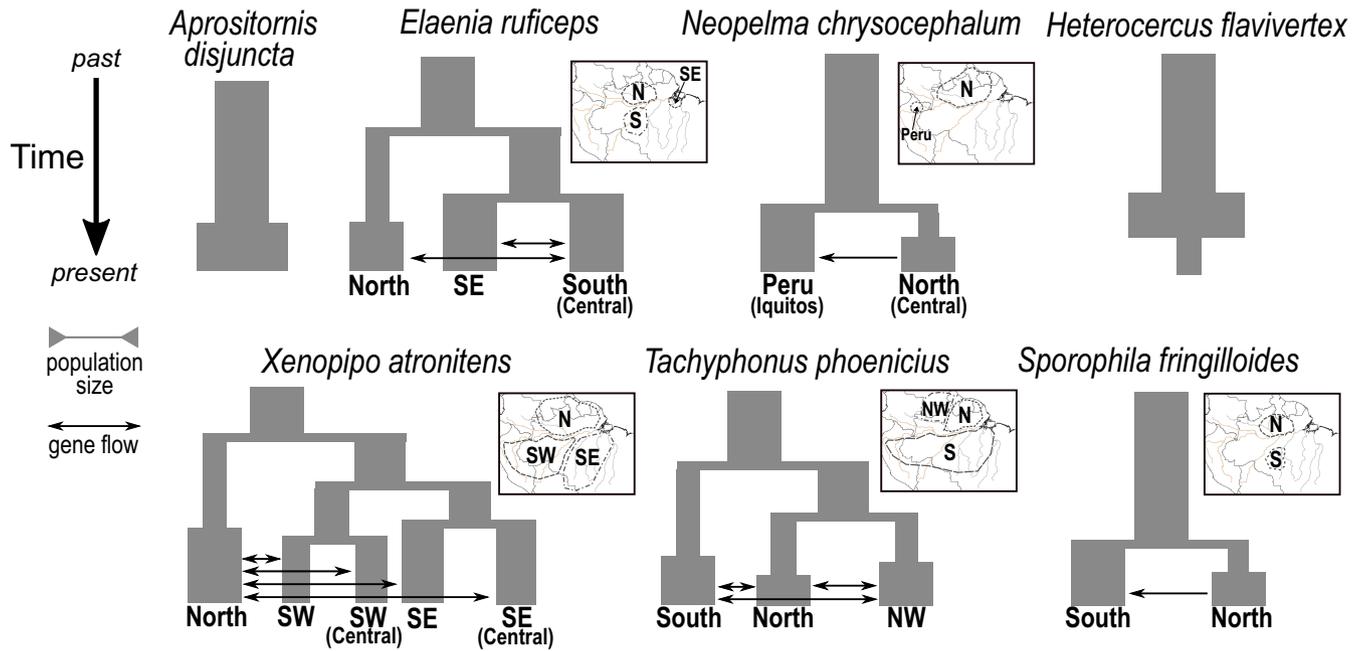
TABLE 1 Total number of samples, SNPs, UCE loci and their mean and total (concatenated) sequence length in base pairs (bp) for each species.

Species	Number of samples	SNPs <sup>a</sup>	Number of UCEs	Mean length of UCEs (bp)	Total sequence length (bp)	Number of informative sites (mean [min-max])
<i>Aprositornis disjuncta</i>	9	883	2299	654	1,504,220	2863 (1.25 [0–22])
<i>Elaenia ruficeps</i>	21	1473	2295	809	1,857,536	9714 (4.23 [0–24])
<i>Neopelma chrysocephalum</i>	23	939	2309	654	1,509,662	5185 (2.25 [0–24])
<i>Heterocercus flavivertex</i>	21	1135	2307	644	1,486,372	5565 (2.41 [0–17])
<i>Xenopipo atronitens</i>	51	722	2308	739	1,707,187	8709 (3.77 [0–26])
<i>Tachyphonus phoenicius</i>	39	1477	2304	760	1,751,694	11,657 (5.06 [0–33])
<i>Sporophila fringilloides</i>	13	946	2307	641	1,478,015	3258 (1.41 [0–13])

<sup>a</sup>Total number of SNPs after filtering the data (see methods).



**FIGURE 1** (a) The distribution of white-sand ecosystems (in yellow) in Amazonia and adjacent areas, main rivers, and currently recognized Areas of Endemism for terra-firme forest birds (Adeney et al., 2016; Sérgio H. Borges & Silva, 2012; Cracraft, 1985): (1) Guiana, (2) Negro, (3) Imerí, (4) Xingu, (5) Tapajós, (6) Rondônia, (7) Inambari, (8) Napo, (9) Belém. Sample distribution with individuals represented by pie plots of admixture proportions according to the best number of genetic clusters (K) obtained from ‘conStruct’ analyses (see methods, Figures S5-S11): (b) *Aprositornis disjuncta*, (c) *Elaenia ruficeps*, (d) *Neopelma chrysocephalum*, (e) *Heterocercus flavivertex*, (f) *Xenopipo atronitens*, (g) *Tachyphonus phoenicius* and (h) *Sporophila fringilloides*. Geographic information was projected using WGS84 (EPSG:4326).



**FIGURE 2** Results from the model selection analyses obtained using the ‘PipeMaster’ pipeline showing the best model that describes the genetic patterns and the occurrence of gene flow and demographic changes (the thickness of the bars indicates populations expansions or contractions) of each white-sand ecosystems species: (from left to right, top row) *Aprositornis disjuncta*, *Elaenia ruficeps*, *Neopelma chrysocephalum*, *Heterocercus flavivertex*, (bottom row) *Xenopipo atronitens*, *Tachyphonus phoenicius* and *Sporophila fringilloides*. The small panels show the geographic location of populations in the Amazonian region (see also Figure 1). The full set of models that were tested can be found in the Supplementary Materials (Figures S1–S4).

**TABLE 2** Posterior, cumulative and prior probabilities for the top ten models in the *ecoevolity* shared demographic events analysis considering a *root\_relative\_population\_size* prior with a uniform distribution (0.001–10.0). Unique numbers refer to groupings of populations sharing the time of demographic change. The model with the highest probability clustered the nine populations into two population size change events and is highlighted in bold.

Model	Number of events	Posterior probability	Cumulative posterior probability	Prior probability
<b>0,0,0,1,1,0,0,0,0</b>	2	0.18	0.18	0.00045
0,0,0,1,1,0,0,2,0	3	0.095	0.275	0.00017
0,1,1,2,2,1,1,1,1	3	0.074	0.35	0.00017
0,1,1,2,2,1,1,3,1	4	0.057	0.40	8.95E-05
0,0,0,1,2,0,0,0,0	3	0.054	0.461	0.00105
0,0,0,1,2,0,0,3,0	4	0.033	0.495	0.00047
0,1,0,2,2,0,0,0,0	3	0.029	0.525	0.00016
0,1,1,2,2,1,1,0,1	3	0.028	0.554	3.64E-05
0,1,1,2,3,1,1,1,1	4	0.026	0.58	0.00046
0,1,1,2,3,1,1,4,1	5	0.02	0.60	0.00026

From a total of 17 populations across all species we identified, nine populations that showed demographic changes and eight of these occur in northern Amazonia (Figure 2). Eight populations showed signals of population expansion, and only one species, *H.*

*flavivertex*, suffered a recent bottleneck that followed a previous population expansion.

We further tested if these populations shared their timing of demographic change using *ecoevolity* (Oaks, 2019; Oaks et al., 2020).

Following Oaks et al. (2020), we tested different priors for the *root\_relative\_population\_size* (see Methods) and all led to similar results, except for the most restrictive prior (an exponential distribution with mean=0.1; Table S16), which caused a pronounced change in the time of population size change of *A. disjuncta* (Figure S12). Overall, the models with 2, 3 and 4 shared events of demographic change had higher posterior probabilities. The higher values for the number of events (3 and 4) were related to populations with wider confidence intervals for the timing of demographic changes being included as additional events. We present the results from the analysis with the less restrictive uniform prior on the *root\_relative\_population\_size*. The model with highest probability separated demographic changes into two time periods (Table 2). One event occurred ~50,000 years ago (Figure 3) involving *N. chrysocephalum* and *X. atronitens* (both occurring in central Amazonia, north of the Amazon river). Another event, estimated to have happened ~100,000 years ago (Figure 3), included the remaining populations (*A. disjuncta*, *H. flavivertex*, all three populations in *T. phoenicius*, and the northern *E. ruficeps* and *S. fringilloides* populations).

Estimates of the timing of population expansions based on the best inferred demographic models using 'PipeMaster' were more recent than in the shared co-demographic analysis (Figure S28). While the dates are different, the 'PipeMaster' estimates are very recent and also have overlapping posterior distributions, yet this is not a direct test of shared demography. The methods and models applied within each analysis, *ecoevolity* and 'PipeMaster', are different as the former extracts information from the full concatenated sequence dataset (Oaks et al., 2020) and the latter reduces the complexity of the data by calculating summary statistics so that model inference and parameter estimation become less computationally intensive (Gehara et al., 2020). In addition, *ecoevolity* explicitly implements a simple model that tests for shared demography and divergence times (although we did not implement the latter) (Oaks, 2019), while 'PipeMaster' allows the simulation of data and model inference from simple to more complex model settings describing the evolutionary history of a clade, although it does not perform comparative analyses among lineages (Gehara et al., 2020).

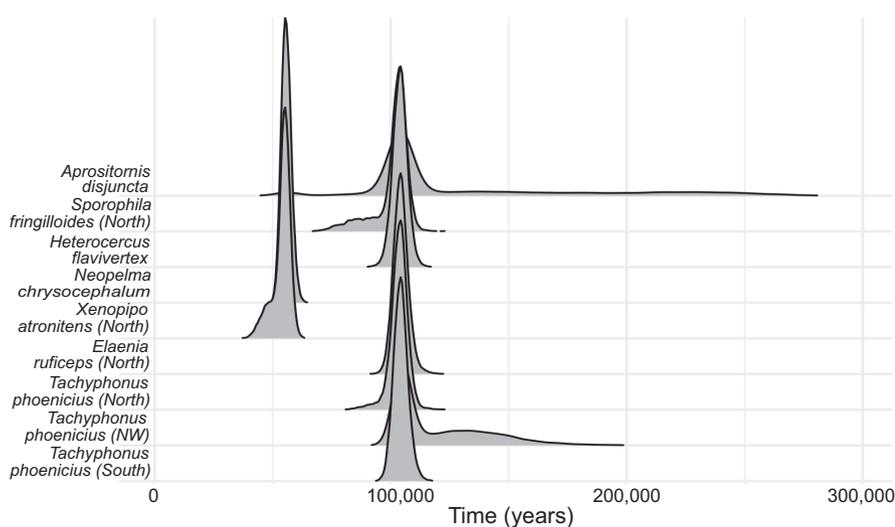
## 4 | DISCUSSION

Access to genomic data coupled with sophisticated model-based methods have contributed to the advance of biogeographic analyses and the testing of shared patterns among lineages resulting from environmental and landscape history (Oaks, 2019; Papadopoulou & Knowles, 2016). Using genomic data obtained through target capture of UCE loci, our study provides a comparative overview of genetic diversity, population structure patterns and demographic history across seven WSE bird species. We focus on the relationship between population history and Amazonian environmental and landscape changes by evaluating the demographic history of the identified genetic clusters, comparing the observed patterns to those of species adapted to other Amazonian ecosystems, and assessing what could be the drivers of common observed patterns. We find concordant genetic structure and population size change patterns among species implying shared responses to climate-driven changes in Amazonian vegetation, including WSE and *terra firme* forests that likely impacted population sizes and connectivity among populations of WSE birds.

Our results are generally consistent with previous studies (*Elaenia ruficeps*, Ritter, Coelho, et al., 2021; Ritter, Ribas, et al., 2021; *Neopelma chrysocephalum*, Capurucho et al., 2018; *Xenopipo atronitens*, Capurucho et al., 2013; and *Tachyphonus phoenicius*, Matos et al., 2016). However, the genomic approach we used provides higher resolution for phylogeographic analyses, allows implementation of complex model testing, and assessment of shared demographic events.

### 4.1 | The Amazon River is an important biogeographic barrier for WSE Birds

Genetic structure across the Amazon river is observed in the UCE data, as previously reported for *E. ruficeps*, *X. atronitens*, and *T. phoenicius* using mtDNA (Capurucho et al., 2013; Matos



**FIGURE 3** Results from *ecoevolity* analysis showing the estimated 95% credible interval for the time of population size change in the nine populations (from top to bottom: *Aprositornis disjuncta*, *Sporophila fringilloides* North, *Heterocercus flavivertex*, *Neopelma chrysocephalum* Guiana Shield, *Xenopipo atronitens* North, *Elaenia ruficeps* North and *Tachyphonus phoenicius* North, Northwest and South) of white-sand ecosystems birds for which the best inferred model included demographic fluctuations (see Figure 2).

et al., 2016; Ritter, Coelho, et al., 2021; Ritter, Ribas, et al., 2021). The north/south structure across the Amazon river also is observed in *Sporophila fringilloides* (this study) and in *Galbula leucogastra* (Ferreira et al., 2018), while the remaining three taxa studied here do not occur south of the Amazon river, which also highlights its importance as barrier to dispersal. Lineage differentiation across the Amazon river is a common pattern in *terra firme* forest birds (Silva et al., 2019) and our results show that the river also limits gene flow among WSE populations. A major turnover in WSE plant communities related to the Amazon river has also been reported, as communities from opposite margins share few species (Costa et al., 2020).

According to parameter estimates there is little overlap in coalescent time estimates for populations on opposite margins of the Amazon river (Figure S27). Our estimates were more recent than previously reported for *E. ruficeps* and *T. phoenicius* (Matos et al., 2016; Ritter, Coelho, et al., 2021), probably because of the different genetic markers and methods (and respective models and assumptions). All populations isolated by the Amazon river diverged within the last 450,000 years ago, with confidence intervals going up to ~600,000 years ago in the case of *X. atronitens* (Figure S27). These dates are very recent in relation to geological estimates of the genesis of the lower Amazon river draining to the Atlantic ocean which vary from ~10 to 2.5 Ma (Campbell et al., 2006; Hoorn et al., 2010). Our divergence values are also more recent than generally found for *terra firme* forest birds separated by the lower Amazon river (Silva et al., 2019). The relatively recent population splits coupled with population size changes, evidences of gene flow among populations, and an old origin of most of the studied taxa (from the mid-Miocene to mid-Pliocene; except for *H. flavivertex* whose origin is dated to the mid-Pleistocene) (Capurucho, Borges, et al., 2020; Harvey et al., 2020) suggest that WSE populations likely have gone through pulses of isolation and reconnection across the Amazon river during the Pleistocene. These could have been driven by local extinctions and recolonizations resulting from past climatic fluctuations and the high dynamism of WSE (Rodrigues et al., 2022). Gene flow across Amazonian rivers may be related to fluctuations in river discharge and corresponding effects on floodplain habitats ultimately affecting the permeability of the river-floodplain system and its effect as a barrier to dispersal (Cremon et al., 2016; Rossetti, Bertani, et al., 2012).

## 4.2 | Recent genetic structure patterns and population expansion

Model inference and parameter estimation showed that genetic structure among populations of WSE birds is rather recent (less than 450,000 years ago; Figure S27). These results are in contrast to an older origin of WSE species (Capurucho, Borges, et al., 2020), as many WSE species reportedly originated in the Pliocene and Miocene. It also contrasts with the timing of within-species lineage

diversification of *terra firme* forest birds which largely date to the late Pliocene and early Pleistocene (Silva et al., 2019). Comparatively recent population genetic patterns are also found in birds occupying flooded forests and river islands, ecosystems that have been dynamic in the late Pleistocene and Holocene (Choueri et al., 2017; Thom et al., 2018).

We found common geographic patterns to population size changes within the WSE species we studied. All populations in northern Amazonia have expanded recently, while expansion south of the Amazon river was found only in *T. phoenicius* (Figure 2). The occurrence of savanna relicts at the southern border of Amazonia and their expansion during the last glacial period (Arruda et al., 2018; Sato et al., 2021; Werneck, 2011), together with the capacity of *T. phoenicius* to occupy Amazonian savanna habitats similar in structure to some WSE habitats, could explain the expansion and absence of genetic structure of this southern population.

According to the shared demographic analyses, the expansion of northern populations was estimated to have occurred at two times (Figure 3). An older event occurred around 100,000 years ago involving populations of *A. disjuncta*, *E. ruficeps*, *H. flavivertex*, *T. phoenicius*, and *S. fringilloides* (Figure 3). A more recent expansion event is inferred for *N. chrysocephalum* and *X. atronitens* populations, occurring about 50,000 years ago. The groupings of these species match their habitat preferences, as the species with older expansions prefer more open vegetation with scattered trees and shrubs, and those expanding more recently prefer dense scrubby or forest habitats (Borges, Cornelius, Moreira, et al., 2016; del Hoyo et al., 2019; Hilty, 2003). The estimation of shared events is relatively sensitive to the priors, which is not unexpected (Table S16; Figure S12; Oaks et al., 2020). Two populations, *A. disjuncta* and *T. phoenicius* NW, showed wider posterior distributions of the time of population size change (Figure S12) that consequently affected the posterior probabilities of the different models. Still, it is possible to observe that demographic events are clustered in time in a pattern unlikely to occur due to chance alone (Figure 3) supporting a scenario of shared responses to environmental changes occurring after the end of the last interglacial period.

## 4.3 | An overview of WSE bird Phylogeography

The genetic structure within WSE birds is shallow compared to commonly studied *terra firme* forest birds in Amazonia (Silva et al., 2019). This runs counter to expectations given the patchiness and isolation of WSE. Genetic clusters within each species exhibit recent genetic differentiation (within the last 450,000 years ago) indicating that their distribution, population sizes, and connectivity were dynamic in the Late Pleistocene and Holocene. Climatic oscillations, not necessarily linked only to glacial-interglacial periods, likely played an important role in shaping these genetic patterns and the distribution of WSE birds through extinction and recolonization of available habitats. Differences in geographic distribution and genetic patterns among species are likely driven by the dynamics in WSE availability,

their size and connectivity (Borges, Cornelius, Moreira, et al., 2016; Capurucho, Borges, et al., 2020; Ritter, Ribas, et al., 2021), species traits (e.g. dispersal ability; Capurucho, Ashley, et al., 2020), and habitat preferences (this study).

It is generally expected that the biotas of WSE patches would be isolated by the *terra firme* forest matrix, yet there is no clear evidence of this effect in our analyses. However, Capurucho et al. (2013) and Ritter, Coelho, et al. (2021) found that areas with larger and more connected patches of WSE have higher genetic diversity, indicating that *terra firme* forests limit gene flow for WSE species, otherwise genetic diversity should be independent of area if populations are truly panmictic. In addition, WSE birds seem to reach their geographic ranges more slowly than their sister taxa occupying more connected ecosystems, showing that WSE patchiness might slow down dispersal and colonization of new areas (Capurucho, Ashley, et al., 2020). Changes in *terra firme* forest structure have been empirically shown, modelled, or hypothesized (Arruda et al., 2018; D'Apolito et al., 2013; Sato et al., 2021; Silva et al., 2019). These changes in forest structure might periodically render *terra firme* as a more permeable matrix for WSE specialist birds, connecting WSE patches, while movement through this matrix would still be modulated by each species dispersal ability, habitat preferences, and population size (Capurucho, Ashley, et al., 2020). The absence of genetic structure and restricted distribution of some WSE taxa also could result from local extinctions due to the dynamism of WSE in the past and historical interplay with *terra firme* forest.

There is evidence for dynamism in the distribution of patches of sandy soils, mainly during the Pleistocene glacial periods, including the genesis of new and reworking of older WSE areas in northern (Carneiro-Filho et al., 2002; Horbe et al., 2004; Rossetti et al., 2016; Rossetti et al., 2018; Teeuw & Rhodes, 2004; Zular et al., 2019) and southern Amazonian regions (Hayakawa & Rossetti, 2015; Rossetti et al. 2012). The WSE appear to be fragile due to the low nutrient and water retention capabilities of sandy soils (Damasco et al., 2013). Thus, changes in precipitation levels and seasonality (longer/shorter dry seasons), associated with Milankovitch cycles and climatic anomalies, could impact WSE vegetation and consequently its birds. With the cyclic climatic variation during the Pleistocene, the WSE bird populations would suffer bottlenecks or local extinctions in areas more affected by environmental changes, followed by recolonization or population expansion during wetter periods, and northern Amazonian populations were evidently affected by this dynamic. The shallow structure and generally high levels of gene flow observed among WSE species populations are likely the result of pulses of higher connectivity among WSE patches followed by periods of isolation due to changes in the Amazonian forest matrix modulated by climatic fluctuations.

## 5 | CONCLUSIONS

The model-based approach we used coupled with an explicit test of shared demographic patterns have shown consonant responses

to landscape and environmental changes in a group of characteristic WSE bird species. The dynamics of climate history and impacts on Amazonian ecosystems, together with the new information presented here on demographic history, population genetics, and community structure point to a complex scenario for the evolutionary history of species adapted to different Amazonian ecosystems. Despite harbouring a bird community with lower richness and relatively old species, population structure in the fragmented WSE is counterintuitively shallow, with evidence for recent gene flow and population size fluctuations. We argue that WSE are possibly one of the most dynamic and fragile ecosystems of Amazonia. The abiotic processes generating and reshaping these ecosystems are diverse and have contributed to an ever-changing distribution across Amazonia. The consequences are that WSE bird populations go through cycles of local extinctions and isolation, followed by recolonizations and reconnection of WSE patches. In addition to the environmental filtering due to the restrictive conditions of the WSE, these processes hinder stronger lineage differentiation in these birds, possibly causing the observed lower bird diversity and maintenance of only those species that have been able to cope with this dynamism over evolutionary time. For these reasons, understanding the recent history of WSE specialized biota provides an opportunity to evaluate and anticipate the impacts of future climate change in Amazonia and its biological communities.

## ACKNOWLEDGEMENTS

We thank all the people involved in collecting specimens in the field and curating and managing the Bird Collections of the Instituto Nacional de Pesquisas da Amazônia, Museu Paraense Emílio Goeldi, Field Museum of Natural History, Louisiana State University Museum of Natural Science, American Museum of Natural History, Kansas University Natural History Museum, National Museum of Natural History, Academy of Natural Sciences of Drexel University, and Yale Peabody Museum of Natural History. Many colleagues helped with data processing and analyses, and we greatly appreciate their support. The laboratory work was conducted in the Pritzker Laboratory for Molecular Systematics and Evolution at the Field Museum of Natural History. This project was supported by the grants: Dimensions US-Biota-São Paulo co-funded by the US National Science Foundation (NSF DEB 1241056) and the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP grant #2012/50260-6), FAPESP-FAPEAM (Edital 006/2009), and Universal FAPEAM (Edital 002/2018). We thank ICMBio for providing the necessary permits (#20524, #24597, and #32339) to conduct this research. JMGC was supported by a CAPES Science without Borders scholarship (11881-2013/5). This project was completed in partial fulfilment of a doctoral degree from the Graduate College at the University of Illinois at Chicago to JMGC.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Dryad at <https://doi.org/10.5061/dryad.47d7wm3f9>.

## ORCID

João Marcos Guimarães Capurucho  <https://orcid.org/0000-0002-0817-3243>

Sergio H. Borges  <https://orcid.org/0000-0002-1952-4467>

Camila C. Ribas  <https://orcid.org/0000-0002-9088-4828>

## REFERENCES

- Adeney, J. M., Christensen, N. L., Vicentini, A., & Cohn-Haft, M. (2016). White-sand ecosystems in Amazonia. *Biotropica*, 48(1), 7–23. <https://doi.org/10.1111/btp.12293>
- Alonso, J. Á., Metz, M. R., & Fine, P. V. A. (2013). Habitat specialization by birds in Western Amazonian white-sand forests. *Biotropica*, 45(3), 365–372. <https://doi.org/10.1111/btp.12020>
- Alonso, J. A., & Whitney, B. M. (2003). New distributional records of birds from white-sand forests of the northern peruvian amazon, with implications for biogeography of northern south america. *The Condor*, 105(3), 552. <https://doi.org/10.1650/7159>
- Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., Aleixo, A., Faircloth, B. C., & Antonelli, A. (2018). Allele phasing greatly improves the phylogenetic utility of Ultraconserved elements. *Systematic Biology*, 68(1), 32–46. <https://doi.org/10.1093/sysbio/syy039>
- Anderson, A. (1981). White-sand vegetation of Brazilian Amazonia. *Biotropica*, 13(3), 199–210.
- Arruda, D. M., Schaefer, C. E. G. R., Fonseca, R. S., Solar, R. R. C., & Fernandes-Filho, E. I. (2018). Vegetation cover of Brazil in the last 21 ka: New insights into the Amazonian refugia and Pleistocenic arc hypotheses. *Global Ecology and Biogeography*, 27(1), 47–56. <https://doi.org/10.1111/geb.12646>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borges, S. (2004). Species poor but distinct: Bird assemblages in white sand vegetation in Jaú National Park, Brazilian Amazon. *Ibis*, 146, 114–124.
- Borges, S. H., Cornelius, C., Moreira, M., Ribas, C. C., Cohn-Haft, M., Capurucho, J. M., Vargas, C., & Almeida, R. (2016). Bird communities in Amazonian white-sand vegetation patches: Effects of landscape configuration and biogeographic context. *Biotropica*, 48(1), 121–131. <https://doi.org/10.1111/btp.12296>
- Borges, S. H., & Silva, J. M. C. (2012). A new area of endemism for Amazonian birds in the Rio Negro basin. *The Wilson Journal of Ornithology*, 124(1), 15–23. <https://doi.org/10.1676/07-103.1>
- Borges, S. H. (2013). Bird species distribution in a complex Amazonian landscape: Species diversity, compositional variability and biotic-environmental relationships. *Studies on Neotropical Fauna and Environment*, 48(2), 106–118. <https://doi.org/10.1080/01650521.2013.841627>
- Borges, S. H., Cornelius, C., Ribas, C., Almeida, R., Guilherme, E., Aleixo, A., Dantas, S., Santos, M. P., & Moreira, M. (2016). What is the avifauna of Amazonian white-sand vegetation? *Bird Conservation International*, 26(2), 192–204. <https://doi.org/10.1017/S0959270915000052>
- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Poppinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., ... Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15(4), e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring continuous and discrete population genetic structure across space. *Genetics*, 210(1), 33–52. <https://doi.org/10.1534/genetics.118.301333>
- Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & Roychoudhury, A. (2012). Inferring species trees directly from bi-allelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29(8), 1917–1932. <https://doi.org/10.1093/molbev/mss086>
- Burney, C. W., & Brumfield, R. T. (2009). Ecology predicts levels of genetic differentiation in neotropical birds. *The American Naturalist*, 174(3), 358–368. <https://doi.org/10.1086/603613>
- Bush, M. (1994). Amazonian speciation: a necessarily complex model. *Journal of Biogeography*, 21, 5–17.
- Campbell, K. E., Frailey, C. D., & Romero-Pittman, L. (2006). The pan-Amazonian Ucayali peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon River system. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 239(1–2), 166–219. <https://doi.org/10.1016/j.palaeo.2006.01.020>
- Capella-Gutierrez, S., Silla-Martinez, J. M., & Gabaldon, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>
- Capurucho, J. M. G., Ashley, M. V., Ribas, C. C., & Bates, J. M. (2018). Connecting Amazonian, Cerrado, and Atlantic forest histories: Paraphyly, old divergences, and modern population dynamics in tyrant-manakins (Neopelma/Tyrannetes, Aves: Pipridae). *Molecular Phylogenetics and Evolution*, 127(April), 696–705. <https://doi.org/10.1016/j.ympev.2018.06.015>
- Capurucho, J. M. G., Ashley, M. V., Tsuru, B. R., Cooper, J. C., & Bates, J. M. (2020). Dispersal ability correlates with range size in Amazonian habitat-restricted birds: Correlates of range size in birds. *Proceedings of the Royal Society B: Biological Sciences*, 287(1939). <https://doi.org/10.1098/rspb.2020.1450>
- Capurucho, J. M. G., Borges, S. H., Cornelius, C., Vicentini, A., Prata, E. M. B., Costa, F. M., Campos, P., Sawakuchi, A. O., Rodrigues, F., Zular, A., Aleixo, A., Bates, J. M., & Camila Ribas, C. (2020). Patterns and processes of diversification in Amazonian white sand ecosystems: Insights from birds and plants. In V. Rull & A. C. Carnaval (Eds.), *Neotropical diversification: Patterns and processes* (pp. 245–270). Springer Nature. [https://doi.org/10.1007/978-3-030-31167-4\\_11](https://doi.org/10.1007/978-3-030-31167-4_11)
- Capurucho, J. M. G., Cornelius, C., Borges, S. H., Cohn-Haft, M., Aleixo, A., Metzger, J. P., & Ribas, C. C. (2013). Combining phylogeography and landscape genetics of *Xenopipo atronitens* (Aves: Pipridae), a white sand campina specialist, to understand Pleistocene landscape evolution in Amazonia. *Biological Journal of the Linnean Society*, 110(1), 60–76. <https://doi.org/10.1111/bj.12102>
- Carneiro-Filho, A., Schwartz, D., Tatumi, S. H., & Rosique, T. (2002). Amazonian paleodunes provide evidence for drier climate phases during the Late Pleistocene–Holocene. *Quaternary Research*, 58(2), 205–209. <https://doi.org/10.1006/qres.2002.2345>
- Choueri, É. L., Gubili, C., Borges, S. H., Thom, G., Sawakuchi, A. O., Soares, E. A. A., & Ribas, C. C. (2017). Phylogeography and population dynamics of antbirds (Thamnophilidae) from Amazonian fluvial islands. *Journal of Biogeography*, 44(10), 2284–2294. <https://doi.org/10.1111/jbi.13042>
- Cooper, J. C., Maddox, J. D., McKague, K., & Bates, J. M. (2021). Multiple lines of evidence indicate ongoing allopatric and parapatric diversification in an Afrotropical sunbird (*Cinnyris reichenowi*). *Ornithology*, 138(2), ukaa081. <https://doi.org/10.1093/ornithology/ukaa081>
- Costa, F. M., Terra-Araujo, M. H., Zartman, C. E., Cornelius, C., Carvalho, F. A., Hopkins, M. J. G., Viana, P. L., Prata, E. M. B., & Vicentini, A.

- (2020). Islands in a green ocean: Spatially structured endemism in Amazonian white-sand vegetation. *Biotropica*, 52(1), 34–45. <https://doi.org/10.1111/btp.12732>
- Cracraft, J. (1985). Historical biogeography and patterns of differentiation within the south American avifauna: Areas of endemism. *Ornithological Monographs*, 36, 49–84.
- Cremon, É. H., Rossetti, D., de F., Sawakuchi, A., & Cohen, M. C. L. (2016). The role of tectonics and climate in the late quaternary evolution of a northern Amazonian river. *Geomorphology*, 271, 22–39. <https://doi.org/10.1016/j.geomorph.2016.07.030>
- Csilléry, K., François, O., & Blum, M. G. B. (2012). Abc: An R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution*, 3(3), 475–479. <https://doi.org/10.1111/j.2041-210X.2011.00179.x>
- Damasco, G., Vicentini, A., Castilho, C. V., Pimentel, T. P., & Nascimento, H. E. M. (2013). Disentangling the role of edaphic variability, flooding regime and topography of Amazonian white-sand vegetation. *Journal of Vegetation Science*, 24(2), 384–394. <https://doi.org/10.1111/j.1654-1103.2012.01464.x>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- D'Apolito, C., Absy, M. L., & Latrubesse, E. M. (2013). The hill of six lakes revisited: New data and re-evaluation of a key Pleistocene Amazon site. *Quaternary Science Reviews*, 76, 140–155. <https://doi.org/10.1016/j.quascirev.2013.07.013>
- del Hoyo, J., Elliott, A., Sargatal, J., Christie, D. A., & Kirwan, G. (2019). *Handbook of the Birds of the world alive*. Lynx Edicions, . Retrieved June, 2019. <http://www.hbw.com/>
- Faircloth, B. C. (2016). PHYLUCe is a software package for the analysis of conserved genomic loci. *Bioinformatics*, 32(5), 786–788. <https://doi.org/10.1093/bioinformatics/btv646>
- Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61(5), 717–726. <https://doi.org/10.1093/sysbio/sys004>
- Ferreira, M., Fernandes, A. M., Aleixo, A., Antonelli, A., Olsson, U., Bates, J. M., Cracraft, J., & Ribas, C. C. (2018). Evidence for mtDNA capture in the jacamar *Galbula leucogastra*/chalcophorax species-complex and insights on the evolution of white-sand ecosystems in the Amazon basin. *Molecular Phylogenetics and Evolution*, 129, 149–157. <https://doi.org/10.1016/j.ympev.2018.07.007>
- Fine, P. V. A., Mesones, I., & Coley, P. D. (2004). Herbivores promote habitat specialization by trees in Amazonian forests. *Science*, 305(5684), 663–665. <https://doi.org/10.1126/science.1098982>
- Fouquet, A., Noonan, B. P., Rodrigues, M. T., Pech, N., Gilles, A., & Gemmill, N. J. (2012). Multiple quaternary refugia in the eastern guiana shield revealed by comparative phylogeography of 12 frog species. *Systematic Biology*, 61(3), 461–489. <https://doi.org/10.1093/sysbio/syr130>
- Gehara, M., Mazzochinni, G. G., & Burbrink, F. (2020). PipeMaster: Inferring population divergence and demographic history with approximate Bayesian computation and supervised machine-learning in R. *BioRxiv*, 2020(12) 04.410670. <https://doi.org/10.1101/2020.12.04.410670>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, 29(7), 644–652. <https://doi.org/10.1038/nbt.1883>
- Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, 165, 131–137.
- Harvey, M. G., Aleixo, A., Ribas, C. C., & Brumfield, R. T. (2017). Habitat association predicts genetic diversity and population divergence in amazonian birds. *American Naturalist*, 190(5), 631–648. <https://doi.org/10.1086/693856>
- Harvey, M. G., Bravo, G. A., Claramunt, S., Cuervo, A. M., Derryberry, G. E., Battilana, J., Seeholzer, G. F., Shearer McKay, J., O'Meara, B. C., Faircloth, B. C., Edwards, S. V., Pérez-Emán, J., Moyle, R. G., Sheldon, F. H., Aleixo, A., Smith, B. T., Chesser, R. T., Silveira, L. F., Cracraft, J., ... Derryberry, E. P. (2020). The evolution of a tropical biodiversity hotspot. *Science*, 370(6522), 1343–1348. <https://doi.org/10.1126/science.aaz6970>
- Hayakawa, E. H., & Rossetti, D. F. (2015). Late quaternary dynamics in the Madeira river basin, southern Amazonia (Brazil), As revealed by paleomorphological analysis. *Anais Da Academia Brasileira de Ciencias*, 87(1), 29–50. <https://doi.org/10.1590/0001-3765201520130506>
- Hilty, S. L. (2003). *Birds of Venezuela* (2nd ed.). Princeton University Press.
- Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C. L., Figueiredo, J. P., Jaramillo, C., Riff, D., Negri, F. R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T., & Antonelli, A. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* (New York, N.Y.), 330(6006), 927–931. <https://doi.org/10.1126/science.1194585>
- Horbe, A. M. C., Horbe, M. A., & Suguio, K. (2004). Tropical Spodosols in northeastern Amazonas state. *Brazil. Geoderma*, 119(1–2), 55–68. [https://doi.org/10.1016/S0016-7061\(03\)00233-7](https://doi.org/10.1016/S0016-7061(03)00233-7)
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kuhn, M. (2020). Caret: Classification and Regression Training. R package version 6.0–86.
- Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. arXiv preprint arXiv:1303.3997.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Matos, M. V., Borges, S. H., D'Horta, F. M., Cornelius, C., Latrubesse, E., Cohn-Haft, M., & Ribas, C. C. (2016). Comparative phylogeography of two bird species, *Tachyphonus phoenicius* (Thraupidae) and *Polytmus theresiae* (Trochilidae), specialized in Amazonian white-sand vegetation. *Biotropica*, 48(1), 110–120. <https://doi.org/10.1111/btp.12292>
- Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology*, 25(10), 1375–1380. <https://doi.org/10.1016/j.cub.2015.03.047>
- Oaks, J. R. (2019). Full Bayesian comparative phylogeography from genomic data. *Systematic Biology*, 68(3), 371–395. <https://doi.org/10.1093/sysbio/syy063>
- Oaks, J. R., L'Bahy, N., & Cobb, K. A. (2020). Insights from a general, full-likelihood Bayesian approach to inferring shared evolutionary events from genomic data: Inferring shared demographic events is challenging\*. *Evolution*, 74(10), 2184–2206. <https://doi.org/10.1111/evo.14052>
- Oliveira-Filho, A. T., Dexter, K. G., Pennington, R. T., Simon, M. F., Bueno, M. L., & Neves, D. M. (2021). On the floristic identity of Amazonian vegetation types. *Biotropica*, btp.12932. <https://doi.org/10.1111/btp.12932>
- Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences of the United States of America*, 113(12), 3203–3208. <https://doi.org/10.1073/pnas.1518811113>

- States of America*, 113(29), 8018–8024. <https://doi.org/10.1073/pnas.1601069113>
- Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, 48(1), 94–100. <https://doi.org/10.1038/ng.3464>
- Prance, G. T. (1996). Islands in Amazonia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351(1341), 823–833. <https://doi.org/10.1098/rstb.1996.0077>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Ritter, C. D., Coelho, L. A., Capurucho, J. M. G., Borges, S. H., Cornelius, C., & Ribas, C. C. (2021). Sister species, different histories: Comparative phylogeography of two bird species associated with Amazonian open vegetation. *Biological Journal of the Linnean Society*, 132(1), 161–173. <https://doi.org/10.1093/biolinnean/blaa167>
- Ritter, C. D., Ribas, C. C., Menger, J., Borges, S. H., Bacon, C. D., Metzger, J. P., Bates, J., & Cornelius, C. (2021). Landscape configuration of an Amazonian Island-like ecosystem drives population structure and genetic diversity of a habitat-specialist bird. *Landscape Ecology*, 36, 2565–2582. <https://doi.org/10.1007/s10980-021-01281-z>
- Rodrigues, F. C. G., Porat, N., Minelli, T. D., Del Río, I., Niyonzima, P., Nogueira, L., Pupim, F. do, N., Silva, C. G., Baker, P., Fritz, S., Wahnfried, I., Kiefer, G., & Sawakuchi, A. O. (2022). Extended-range luminescence dating of central and eastern Amazonia Sandy terrains. *Frontiers in Earth Science*, 10, 888443. <https://doi.org/10.3389/feart.2022.888443>
- Rossetti, D., de, F., Molina, E., & Cremon, É. H. (2016). Genesis of the largest Amazonian wetland in northern Brazil inferred by morphology and gravity anomalies. *Journal of South American Earth Sciences*, 69, 1–10. <https://doi.org/10.1016/j.jsames.2016.03.006>
- Rossetti, D. F., Bertani, T. C., Zani, H., Cremon, E. H., & Hayakawa, E. H. (2012). Late quaternary sedimentary dynamics in Western Amazonia: Implications for the origin of open vegetation/forest contrasts. *Geomorphology*, 177–178, 74–92. <https://doi.org/10.1016/j.geomorph.2012.07.015>
- Rossetti, D. F., Zani, H., Cohen, M. C. L., & Cremon, É. H. (2012). A late Pleistocene–Holocene wetland megafan in the Brazilian Amazonia. *Sedimentary Geology*, 282, 276–293. <https://doi.org/10.1016/j.sedgeo.2012.09.015>
- Rossetti, D. F., Gribel, R., Tuomisto, H., Cordeiro, C. L. O., & Tatum, S. H. (2018). The influence of late quaternary sedimentation on vegetation in an Amazonian lowland megafan. *Earth Surface Processes and Landforms*, 43(6), 1259–1279. <https://doi.org/10.1002/esp.4312>
- Rull, V. (2011). Neotropical biodiversity: Timing and potential drivers. *Trends in Ecology & Evolution*, 26(10), 508–513. <https://doi.org/10.1016/j.tree.2011.05.011>
- Sato, H., Kelley, D. I., Mayor, S. J., Martin Calvo, M., Cowling, S. A., & Prentice, I. C. (2021). Dry corridors opened by fire and low CO<sub>2</sub> in Amazonian rainforest during the last glacial maximum. *Nature Geoscience*, 14(8), 578–585. <https://doi.org/10.1038/s41561-021-00777-2>
- Silva, S. M., Townsend Peterson, A., Carneiro, L., Burlamaqui, T. C. T., Ribas, C. C., Sousa-Neves, T., Miranda, L. S., Fernandes, A. M., D'Horta, F. M., Araújo-Silva, L. E., Batista, R., Bandeira, C. H. M. M., Dantas, S. M., Ferreira, M., Martins, D. M., Oliveira, J., Rocha, T. C., Sardelli, C. H., Thom, G., ... Aleixo, A. (2019). A dynamic continental moisture gradient drove Amazonian bird diversification. *Science Advances*, 5(7) eaat5752. <https://doi.org/10.1126/sciadv.aat5752>
- Teeuw, R. M., & Rhodes, E. J. (2004). Aeolian activity in northern Amazonia: Optical dating of late Pleistocene and Holocene palaeodunes. *Journal of Quaternary Science*, 19(1), 49–54. <https://doi.org/10.1002/jqs.815>
- Thom, G., Amaral, F. R., Do Hickerson, M. J., Aleixo, A., Araujo-Silva, L. E., Ribas, C. C., Choueri, E., & Miyaki, C. Y. (2018). Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species. *Systematic Biology*, 67(4), 700–718. <https://doi.org/10.1093/sysbio/syy004>
- Tuomisto, H. (2007). Interpreting the biogeography of South America. *Journal of Biogeography*, 34(8), 1294–1295. <https://doi.org/10.1111/j.1365-2699.2007.01699.x>
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From fastQ data to high-confidence variant calls: The genome analysis tool kit best practices pipeline. *Current Protocols in Bioinformatics*, 43, 1–33. <https://doi.org/10.1002/0471250953.bi1110s43>
- Vicentini, A. (2016). The evolutionary history of *Pagamea* (Rubiaceae), a white-sand specialist lineage in tropical South America. *Biotropica*, 48(1), 58–69. <https://doi.org/10.1111/btp.12295>
- Werneck, F. P. (2011). The diversification of eastern south American open vegetation biomes: Historical biogeography and perspectives. *Quaternary Science Reviews*, 30(13–14), 1630–1648. <https://doi.org/10.1016/j.quascirev.2011.03.009>
- Winker, K., Glenn, T. C., & Faircloth, B. C. (2018). Ultraconserved elements (UCEs) illuminate the population genomics of a recent, high-latitude avian speciation event. *PeerJ*, 6, e5735. <https://doi.org/10.7717/peerj.5735>
- Zular, A., Sawakuchi, A. O., Chiessi, C. M., D'Horta, F. M., Cruz, F. W., Demattê, J. A. M., Ribas, C. C., Hartmann, G. A., Giannini, P. C. F., & Soares, E. A. A. (2019). The role of abrupt climate change in the formation of an open vegetation enclave in northern Amazonia during the late quaternary. *Global and Planetary Change*, 172, 140–149. <https://doi.org/10.1016/j.gloplacha.2018.09.006>

#### BIOSKETCH

**João M. G. Capurucho** has interests in Neotropical biogeography, evolutionary biology and ecology, mainly studying birds. The authors share interests in phylogeography, population genetics, ecology and their application to understand evolutionary history and species conservation.

**Author contributions:** J.M.G.C., M.V.A., C.C.R. and J.M.B. conceived the study; J.M.G.C., C.C. and S.H.B. collected the data; J.M.G.C. performed the analyses; J.M.G.C. led the writing with input from all authors.

#### SUPPORTING INFORMATION

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

**How to cite this article:** Capurucho, J. M. G., Ashley, M. V., Cornelius, C., Borges, S. H., Ribas, C. C., & Bates, J. M. (2023). Phylogeographic and demographic patterns reveal congruent histories in seven Amazonian White-Sand ecosystems birds. *Journal of Biogeography*, 00, 1–13. <https://doi.org/10.1111/jbi.14611>