Received: 17 August 2019 | Revised: 4 June 2020 | Accepted: 18 June 2020

DOI: 10.1111/mec.15534

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Paleoclimatic evolution as the main driver of current genomic diversity in the widespread and polymorphic Neotropical songbird *Arremon taciturnus*

Nelson Buainain[1] | Roberta Canton[2] | Gabriela Zuquim[3] | Hanna Tuomisto[3] | Tomas Hrbek[4] | Hiromitsu Sato[5] | Camila C. Ribas[1]

[1]Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, AM, Brazil

[2]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA

[3]Department of Biology, University of Turku, Turun Yliopisto, Finland

[4]Departmento de Genetica, Universidade Federal do Amazonas, Manaus, AM, Brazil

[5]Department of Earth Sciences, University of Toronto, Toronto, Ontario, Canada

Correspondence
Nelson Buainain, Instituto Nacional de Pesquisas da Amazônia (INPA), Programa de Pós Graduação em Ecologia, Manaus, AM, Brazil.
Email: nnbuainain@gmail.com

Funding information
Collection Study Grant (AMNH); National Science Foundation, Grant/Award Number: NSF DEB 1241056, 2012 and 50260-6; Fundação de Amparo à Pesquisas do Estado de São Paulo; Academy of Finland; Finnish Cultural Foundation

Abstract

Several factors have been proposed as drivers of species diversification in the Neotropics, including environmental heterogeneity, the development of drainage systems and historical changes in forest distribution due to climatic oscillations. Here, we investigate which drivers contributed to the evolutionary history and current patterns of diversity of a polymorphic songbird (*Arremon taciturnus*) that is widely distributed in Amazonian and Atlantic forests as well as in Cerrado gallery and seasonally-dry forests. We use genomic, phenotypic and habitat heterogeneity data coupled with climatic niche modelling. Results suggest the evolutionary history of the species is mainly related to paleoclimatic changes, although changes in the strength of the Amazon river as a barrier to dispersal, current habitat heterogeneity and geographic distance were also relevant. We propose an ancestral distribution in the Guyana Shield, and recent colonization of areas south of the Amazon river at ~380 to 166 kya, and expansion of the distribution to southern Amazonia, Cerrado and the Atlantic Forest. Since then, populations south of the Amazon River have been subjected to cycles of isolation and possibly secondary contact due to climatic changes that affected habitat heterogeneity and population connectivity. Most Amazonian rivers are not associated with long lasting isolation of populations, but some might act as secondary barriers, susceptible to crossing under specific climatic conditions. Morphological variation, while stable in some parts of the distribution, is not a reliable indicator of genetic structure or phylogenetic relationships.

KEYWORDS

Arremon, biogeography, genomics, oscines, South America

1 | INTRODUCTION

Several hypotheses have been proposed to explain Neotropical species diversification (for syntheses see Antonnelli et al., 2018; Cracraft et al., 2020). In Amazonia, an old hypothesis that has attracted much recent attention suggests that the development of the drainage system was a major biodiversity generator, promoting vicariant speciation in some groups of organisms (Boubli et al., 2015; Ferreira, Aleixo, Ribas, & Santos, 2016; Ribas, Aleixo, Nogueira, Miyaki, & Cracraft, 2012; Wallace, 1895). However, it is expected that riverine systems vary in their permeability as barriers depending on the river's physical properties (e.g., width; Pomara, Ruokolainen, & Young, 2014); the effect of historical environmental changes (e.g., precipitation; Pupim et al., 2019); as well as on the organisms' intrinsic dispersal capacity and ecology (Cracraft et al., 2020; Haffer, 1974). Moreover, rivers may also act as secondary barriers for divergent

populations that responded to other diversification drivers (Naka & Brumfield, 2018).

Alternatively, Haffer (1969, 1974, 2001) proposed that historical climatic fluctuations were the major drivers of speciation in the region. According to the Refugia hypothesis, dry periods caused major reduction and fragmentation of the forest cover, resulting in isolation of forest patches by savanna like vegetation and consequent divergence or vicariant speciation of forest adapted organisms, while wet periods caused forest expansions and reconnections. This would explain not only the diversity and distribution patterns within Amazonia, but also the occurrence of shared species or sister taxa in neighbouring biomes, such as the Atlantic Forest (Batalha-Filho, Fjeldså, Fabre, & Miyaki, 2013; Haffer, 1974; Ledo & Colli, 2017). An alternative although related hypothesis is that, during glacial periods, the vegetation cover was still forest throughout most of Amazonia, but with distinct floristic composition and/or significant variation in canopy density (Arruda, Schaefer, Fonseca, Solar, & Fernandes-Filho, 2018; Cowling, Maslin, & Sykes, 2001; Wang et al., 2017). In this case, finer scale environmental variation may have played a role in changing connectivity among populations and in shaping species distributions and abundance patterns.

Current habitat heterogeneity, climate, soil composition, and vegetation structure have also been reported to be a driver of the distribution of forest-adapted organisms (Pomara, Ruokolainen, Tuomisto, & Young, 2012; Tuomisto et al., 2016). Thus, historical and current habitat heterogeneity may synergistically shape the diversification and distribution of the Amazonian biota (de Abreu, Schietti, & Anciães, 2018; Ortiz, Lima, & Werneck, 2018; Tuomisto et al., 2016; Tuomisto & Ruokolainen, 1997).

Historically, most phylogeographic studies focused on understory organisms (mostly birds) with low vagility and strictly associated with terra firme forests because these organisms are considered better models to test the traditional diversification hypotheses (e.g., rivers and refugia; Fernandes, Gonzales, Wink, & Aleixo, 2013; Ribas et al., 2018). Thus, studying organisms with different ecological requirements may reveal complementary aspects of Neotropical diversification.

The Pectoral Sparrow, *Arremon taciturnus* (Hermann 1783), is a widespread small sedentary forest songbird (suborder Passeri) that occurs in most of Amazonia, as well as gallery and seasonally dry forests of the Cerrado in central Brazil, humid forest enclaves in the deciduous Caatinga (locally known as *Brejos de altitude*) in northeastern Brazil, and the Atlantic Forest on the eastern coast of Brazil (BirdLife International, 2019; Buainain, Pinto, & Raposo, 2017) (Figure 1). It inhabits the interior, edge and clearings of primary and secondary forests (Cavalcante, 2015; Dantas, 2017), but seems to be absent from open vegetation habitats, Amazonian flooded forests (*várzea* and *igapó*) and white-sand habitats (*campina and campinara*) (Buainain et al., 2017). Therefore, these birds are relatively tolerant to habitat heterogeneity and disturbance (and thus not strictly associated to terra firme) but yet, they have some habitat specificity and low vagility similar to other small understory birds.

The species has geographically structured but complex phenotypic variation (Buainain et al., 2017). Populations from the Guyana Shield (northern Amazonia) and Atlantic Forest are mostly

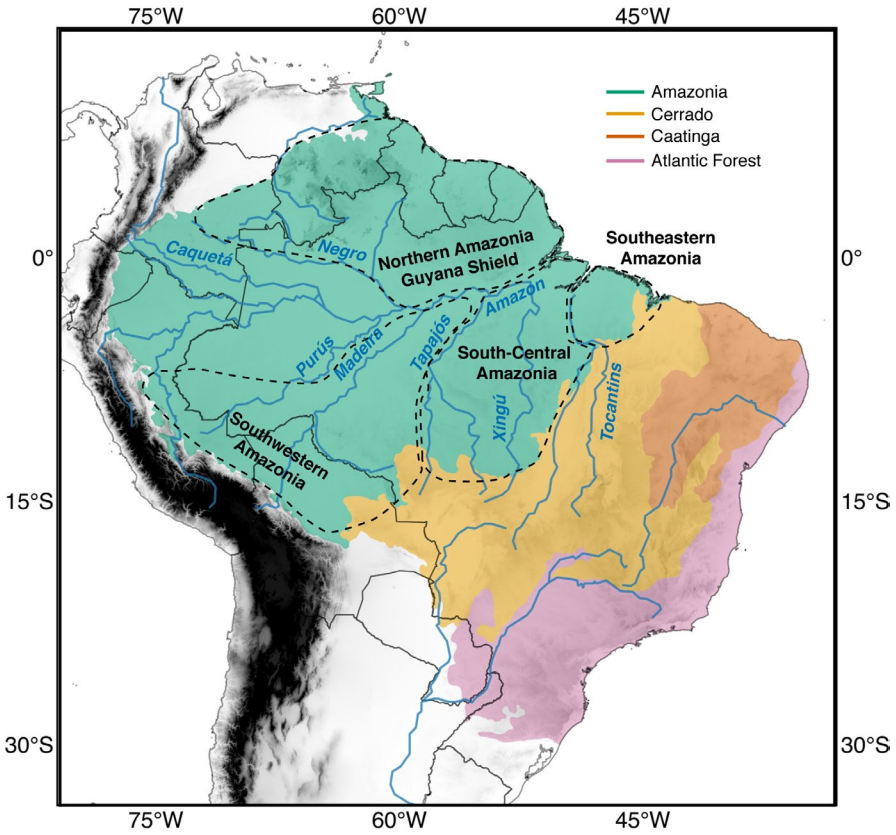

FIGURE 1 Map of South America showing some of the general geographic areas, rivers and biomes mentioned in the text. Coloured regions correspond to the different biomes, rivers and their names are in blue, and darker shading corresponds to higher elevation [Colour figure can be viewed at wileyonlinelibrary.com]

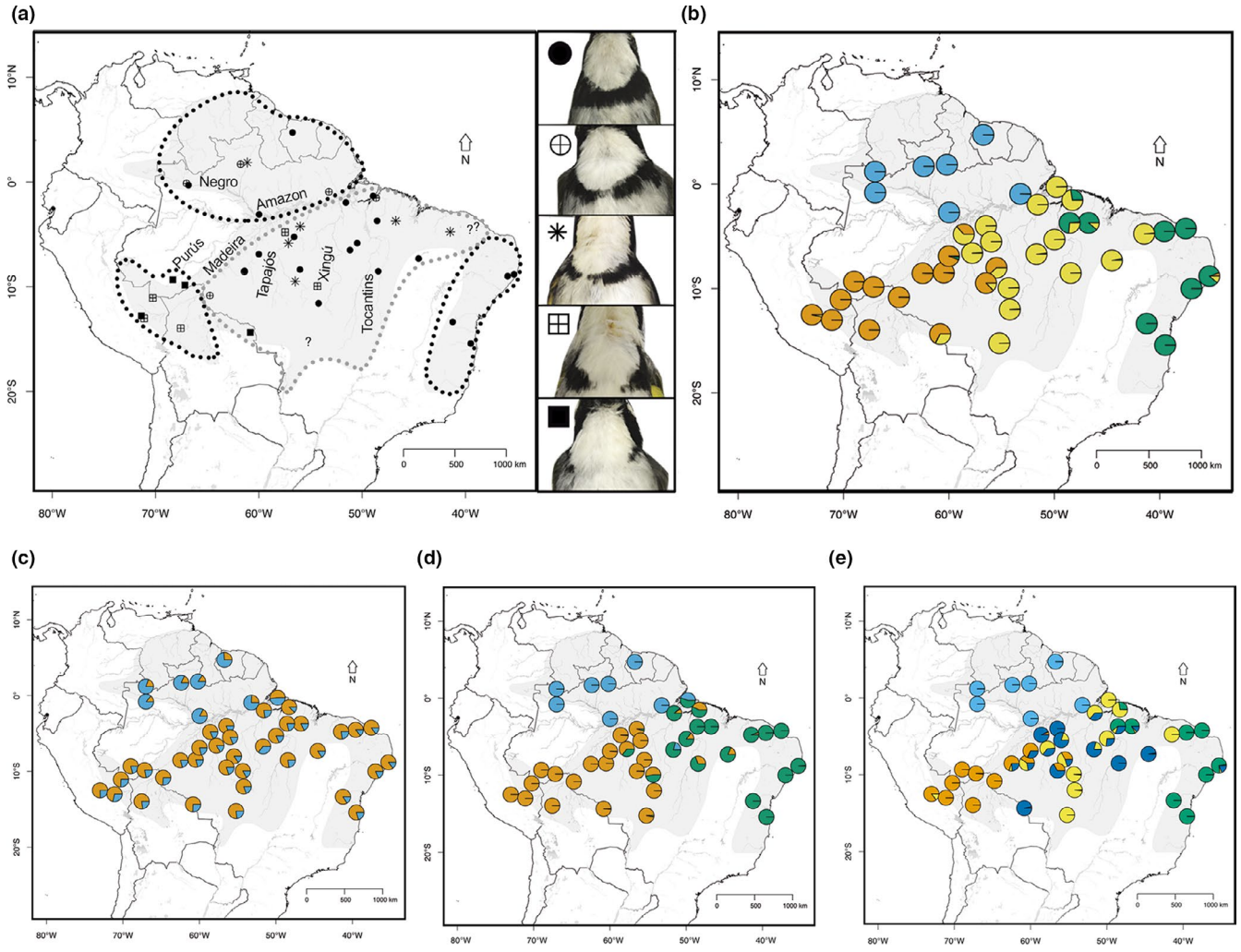

FIGURE 2 Location of sampling points of *Arremon taciturnus* and genetic structure results (STRUCTURE). (a) Tissue samples used according to their morphologic categories (legend on right); samples with unknown category are shown as "?". Polygons of dashed lines represent areas with predominantly pure morphological phenotypes, either complete or incomplete (black) and highly polymorphic pectoral band (grey) based on Buainain et al. (2017). Tissue samples were selected to cover the majority of the morphological variation present in the area, and therefore scores present in (a) may not represent the real frequency of phenotypes in some regions (e.g., Guyana Shield). (b) Genetic structure graphic results for $K = 4$: Guyana (GUY; blue), Southwestern Amazonia (SW; orange), South-Central Amazonia (SC; yellow) and Eastern population (EAS; green). (c) STRUCTURE results for $K = 2$; (d) $K = 3$; and (e) $K = 5$. Samples' positions in (b–e) were slightly geographically adjusted to avoid graphic overlap. STRUCTURE was performed with 2,468 ddrad SNPs. Light grey background represents the species' known distribution according to BirdLife International (2019) [Colour figure can be viewed at wileyonlinelibrary.com]

homogeneous in plumage pattern, with a complete dark pectoral band, while populations from extreme southwestern Amazonia have a predominantly incomplete pectoral band. Populations from central Amazonia through the *Brejos de Altitude* in northeastern Brazil, are highly polymorphic. The observed polymorphism could be due either to primary differentiation or to secondary contact between populations with homogeneous phenotypes (Buainain et al., 2017). In contrast, populations delimited by the large Amazonian rivers and from distinct biomes are vocally discriminated, suggesting that riverine barriers and habitat may have influenced the species' vocalization pattern since there is no evidence of song dialects (i.e., no evidence that neighbouring birds sing more similarly due to a strong learning component; Buainain et al., 2017).

The genetic variation in two mitochondrial DNA genes suggests little differentiation throughout the species distribution (Moura et al., 2020). Phylogenetic analyses revealed low support for most nodes, probably due to very recent history (<0.5 Ma, according to a mtDNA clock rate of 2% divergence per Myr). One of the few clades recovered with moderate support includes individuals from northern (Guyana shield) and southeastern Amazonia (east of the Xingu river; Moura et al., 2020), in an unusual biogeographical pattern that does not reflect the described phenotypic variation. Moura et al. (2020) suggest that isolation by distance (IBD) or changes in habitat distributions attributed to climatic cycles of the late Pleistocene could have shaped the pattern observed, but these hypotheses were not tested due to the small information content in their data set.

As discussed above, currently available information concerning morphological, vocal and mitochondrial genetic variation for *A. taciturnus* seems contradictory. In addition, birds not strictly associated with the understory of upland forests are underrepresented in studies of Neotropical biogeography. Thus, a more robust estimate of the diversification history of *A. taciturnus* has the potential to contribute unique information about the influence of past and present Neotropical landscapes on biotic diversification, and about contrasting patterns of genetic, vocal and plumage variation. Here, we employ genomic data to evaluate which factors associated with current and past landscapes may have contributed to generate current patterns of diversity in this species. In broad terms we expect that: (a) if riverine barriers were important through time, even with varying degrees of permeability, genomic clusters corresponding to current interfluvia should be recognizable; (b) if paleoclimatic changes caused changes in population connectivity in the recent past, then areas of constant climatic suitability over time (refugia) should be congruent with genomic clusters, while genomic breaks or admixed genotypes should correspond to areas of variable climatic suitability; in addition, populations should show signals of recent demographic change; (c) if current habitat heterogeneity is the main driver of observed variation, then there should be an association between current genomic and environmental distances; and (d) if IBD by itself explains the observed variation, there should be a strong association between genetic and geographical distance. In both habitat heterogeneity (c) and IBD (d) scenarios, no signal of recent demographic change is expected. The scenarios above are not mutually exclusive, as, for example, past changes in precipitation patterns may also affect drainage evolution and the strength of riverine barriers (Pupim et al., 2019), but an assessment of relative importance of the different factors is needed for building realistic integrated biogeographical scenarios.

Furthermore, we aim to evaluate the spatial congruence between genomic and plumage variation and the possible mechanisms that might have contributed to the observed patterns. We evaluated two main hypotheses: (a) populations with predominantly homogeneous phenotypes (Guyana shield, Atlantic Forest and extreme southwestern Amazonia) represent historically isolated populations (either by rivers or refugia) and consist of distinct genomic groups, while the area harbouring polymorphic phenotypes corresponds to a contact zone where individuals have admixed genotypes and (b) the populations with predominantly homogeneous plumage represent the extremes of a clinal variation derived from primary differentiation (with no historical isolation) and, thus a strong effect of IBD will be found. Congruence between genomic and vocalization patterns is briefly discussed based on previously published vocal data but are not formally analysed here (Buainain et al., 2017).

2 | MATERIALS AND METHODS

General geographic areas, biomes and rivers mentioned in the text can be seen in Figure 1.

2.1 | Tissue sampling and genomic data acquisition

Forty-four vouchered tissue samples and two blood samples were selected covering most of the geographical distribution and morphological variation of the species (Figure 2a, Supporting information 1; Buainain et al., 2017). One sample of *Arremon semitorquatus* (Swainson 1838) was used as outgroup (Klicka et al., 2014). DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega). A reduced representation genomic library was prepared and sequenced using a double digest RAD sequencing protocol (ddRAD seq) following Peterson, Weber, Kay, Fisher, and Hoekstra (2012). Two restriction enzymes (SdaI and Csp6I) were used to digest each individuals' genomic DNA, and two adapters (Ion Torrent A[with barcode] and P) were ligated to the digested fragments. Fragments were then enriched via PCR. Fragment sizes of 320–400 bp were selected using Pippin Prep (Sage Science). Sequencing was performed on an Ion Torrent PGM (Life Technologies), using an Ion PGM Sequencing 400 kit and a 318 Ion PGM Chip v2 (Life Technologies). The complete protocol is available on GitHub (https://github.com/legalLab/protocols-scripts/blob/master/ddRAD/protocol/ddRADseq_Ion_protocol_en.pdf).

Samples from eight individuals representing the main Amazonian interfluves, Cerrado and Atlantic Forest, plus the outgroup, were sent to RapidGenomics (Gainsville, FL) for sequencing using a probe set targeting 2,321 loci of ultra conserved elements (UCE). These were used as an alternative and independent data source and were used exclusively for a phylogenetic analysis (see Section 2.4). The additional phylogenetic analysis (UCE) was helpful to clarify phylogenetic uncertainties because the ddRAD tree was largely unresolved. This was important to further build the demographic models (see Section 2.8). More information about the capture and sequencing of UCE loci can be found in Faircloth et al. (2012).

2.2 | Bioinformatics

For ddRAD seq, sequences were demultiplexed, filtered, clustered and assembled with de novo method using Pyrad 3.0.6 (Eaton, 2014). The following parameters were modified to obtain the final data sets: Minimum coverage for a cluster = 5; Max individuals with shared heterozygosity sites = 10; phred Qscore offset = 20; filter = Nqual+adapters; trim overhang = 0,1; keep trimmed reads; minimum length = 300. All bases with Q score < 20 were substituted with 'N' and all reads with more than four Ns were discarded, resulting in an average Q score of 28. We relaxed some of the parameters because preliminary results based on initial and more conservative settings outputted a final data set with little information content. The remaining parameters were set as default. Two final data sets were selected by changing the parameter "minimum sample in a final locus" to 17 and four, respectively: a more strict data set with 674 loci (647 unlinked SNPs) and one with a more relaxed data set with 2,563 loci (2,468 unlinked SNPs). We sampled one random SNP per

locus for analyses based on SNPs. The strict data set was used in analyses that are more sensitive to missing data (summary statistics and Fst values). Output statistics from Pyrad can be found with the raw data in the dryad link.

The raw data from UCEs were processed using Phyluce (Faircloth, 2016) incorporating the allele phasing steps from Andermann et al. (2019). Sequences with adapter contamination, and those of low-quality, were trimmed using illumiprocessor (Faircloth, 2013). Trinity RNASeq assembler r201331110 (Grabherr et al., 2011) was used to assemble contigs using a de novo method. UCEs were extracted from assemblies using a 2,321 probe set (Zucker et al., 2016). Alignments were produced and edge trimmed using MAFFT with default settings (Katoh & Standley, 2013). Alignments were then separated into different files for each sample to create individual references. The raw sequences of each individual were then mapped back to the individuals' references and the reads sorted into two separate files producing two alignments (phased alleles) for each individual. The final sequence data matrix was produced with alignments present in at least 80% of samples, giving 2,245 UCEs, and was used solely in the species tree analysis described in the phylogenetic analyses section. All remaining analyses were performed with ddRAD data.

2.3 | Population structure

Population structure and admixture between populations were inferred using STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) and the 2,468 ddRAD SNPs. The number of populations (*K*) was arbitrarily initially tested from 1 to 12, with 10 independent runs per *K*, under the admixture model (allowing mixed ancestry and is capable of dealing with many complexities of real populations), and correlated allelic frequencies model (presuming that frequencies in the different populations are probably similar due to migration or shared ancestry), with a 200,000 burnin followed by collection of 200,000 Monte Carlo Markov Chain proposals. Stationarity of the parameters prior to burnin were visually inspected through the graphic plots provided by the software. CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) was used to produce mean coefficients of ancestry considering all 10 runs. Because the number of populations (*K*) may be problematic and difficult to partition, the ancestry coefficient graphs of all *K*s were visually inspected while looking for the *K* value that represented the finest structure in the data (Evanno, Regnault, & Goudet, 2005; Hahn, 2018; Meirmans, 2015; Pritchard et al., 2000). Essentially, we looked for genetic clusters that were biologically meaningful considering biological and geographical aspects, avoiding samples with admixture ancestry from multiple source populations (>2), which could be a signal of analytical noise and lack of certainty in cluster attribution. Regardless of the "real" number of populations, looking at all Ks is a recommended practice (Cullingham et al., 2020; Meirmans, 2015) and allows us to better explore the structure on the data and look for all possible historical factors responsible for shaping it. We plotted the ancestry coefficients in a map to evaluate spatial congruence with the current position of barriers such as rivers and biomes, and also with the climatic niche modelling to look for possible refugia.

2.4 | Phylogenetic analyses

Phylogenetic relationships among individuals were assessed using IQ-TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2015) to produce a maximum likelihood phylogenetic tree based on the 2,563 ddRAD concatenated loci (sequence data with variant and nonvariant sites). The best-fitting substitution model was inferred in ModelFinder (Kalyaanamoorthy, Minh, Wong, Haeseler, & Jermiin, 2017) and node support was assessed using two methods, UltraFast Bootstrap (UFBoot; Hoang, Chernomor, Haeseler, Minh, & Vinh, 2018) and the SH-aLRT branch test (Guindon et al., 2010) with 1,000 replicates each. Typical reliable clades have support ≥95% for UFBoot and ≥80% for SH-aLRT (Nguyen et al., 2015). Because of phylogenetic uncertainties in the tree resulting from the phylogenetic analysis of the ddRAD data set, we alternatively inferred phylogenetic relationships using the UCE data set using a multispecies coalescent (species tree) approach in SVDquartets implemented in PAUP* 4.0a (build 165; Swofford, 2002) with partitioned sequences, which takes into account the history of the different genes independently. All possible quartets were evaluated. Bootstrap was estimated using 10,000 replicates.

2.5 | Population genetics summary statistics

Summary statistics were calculated in Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) using the 647 ddRAD SNPs data set based on the populations inferred using STRUCTURE to obtain metrics of genetic diversity (expected heterozygosity, mean gene diversity and Theta [*H*]) and genetic divergence (F_{ST}) between them. F_{ST} values were calculated between all pairs of population and 1,000 permutations were performed to test for significance level. A pairwise genetic distance (linearized F_{ST}) matrix was also calculated between all individuals and used in the correlation analyses (see Section 2.7).

2.6 | Species climatic niche modelling and resistance to gene flow

Climatic niche modelling was performed to understand the general changes in climatic suitability over the distribution of *A. taciturnus*, which does not necessarily imply the species distribution itself. Climatic suitability over time was then compared to distribution of current genomic and phenotypic patterns in order to evaluate a possible role of climatic fluctuation over the species' evolutionary history. Models were inferred using Maxent (Phillips, Anderson, Dudík, Schapire, & Blair, 2017) using the R package DISMO (Hijmans,

Phillips, Leathwick, & Elith, 2011) to evaluate the species' current and past climatic suitability. Maxent was chosen because it is a user friendly, consistent and popular modelling software (Merow, Smith, & Silander, 2013) producing comparable results between studies, in addition to performing well compared to other methods (Elith et al., 2006) even with limited sampling (Wisz et al., 2008). Occurrence records (n = 1,001) were gathered based on previously examined skins and vocal recordings obtained from zoological collections and online sound archives (Buainain et al., 2017), and the tissue samples used here (Supporting information 1). All records with unknown or poorly defined localities were excluded. To avoid spatial autocorrelation bias, records within a 40 km distance from each other were clustered and considered as a single record. The final database had 289 records (Supporting information 1).

The 19 bioclimatic variables from BIOCLIM (Fick & Hijmans, 2017) in 30 arc-seconds (~1 km^2) resolution were used. Highly correlated variables (>80% correlation and variance inflation factor [VIF] >10) were excluded using the function vifcor in the R package USDM (Naimi, Hamm, Groen, Skidmore, & Toxopeus, 2014). Accordingly, the function first finds a pair of variables with the maximum linear correlation (>80%), and excludes the one with higher VIF. The procedure is repeated until no variable with a high correlation coefficient with other variables remains. Then, we discarded the remaining variables with VIF > 10. The final variables used were: mean diurnal range (Bio2), mean of monthly temperature (max temp–min temp), isothermality (Bio3), temperature seasonality (Bio4), mean temperature of wettest quarter (Bio8), precipitation of wettest month (Bio13), precipitation seasonality (Bio15), precipitation of warmest quarter (Bio18) and precipitation of coldest quarter (Bio19). To limit the potential occurrence area of the species and background points, and to calibrate our models, raster were masked to include areas of Amazon sensu latissimo and some neighbouring areas, including Cerrado, Caatinga and Atlantic Forest thus gathering various accessible habitat types with potential and nonpotential occurrence. Thirty thousand background points were used to generate probability of occurrence maps in Maxent.

The most likely model was selected using the R package ENMEVAL (Muscarella et al., 2014) considering the linear, quadratic and linear + quadratic features, which are simpler and easier to biologically interpret and less susceptible to model overfitting. The beta multiplier parameter was tested from 0.5 to 6 by a 0.5 factor. Performance of the models were evaluated according to two criteria: the minimum value of AICc (a small sample-size corrected version of the Akaike information criterion) and maximum value of AUC (area under the curve). The selected model was then projected to four different times: (a) present; (b) Mid-Holocene (~6,000 kya) with two different climatic models: Community Climate System Model (CCSM) and Model of Interdisciplinary on Climate (MIROC); (c) Last glacial maximum (~21,000 kya) with CCSM and MIROC; and (d) Last Interglacial (~120,000–140,000 kya). These time frames provide an idea of the changes in areas of climatic suitability for the species under different climatic scenarios. Although variables are only available for these recent times, similar climatic conditions might have cyclically happened during the late Pleistocene (Haffer, 1974). A mean model was built by averaging all models in order to look for areas where suitability was overall high/low through time. A climatic niche model of current scenario was also used to generate a pairwise matrix of isolation by resistance (IBR) using Circuitscape v.4.0.5 (Shah & McRae, 2008). This was then used in the correlation analyses with genomic distances (see correlation section below) to evaluate the impact of current environment/climate on genomic distances. Accordingly, areas with lower climatic suitability have lower conductance values and thus higher resistance to gene flow.

2.7 | Correlation between environmental and genomic variation

Correlation and association tests between genomic and, environmental and geographic distances were performed to evaluate the role of current environment and IBD on the species' genomic pattern. Climatic, edaphic and forest reflectance variables were used as descriptors of the environment in the collection localities of samples used in the genomic analysis. Estimates of soil-based cation concentration (Ca + Mg+K) were extracted from a map developed using direct and plant-derived soil data (Zuquim, 2017). Structural and floristic variation in vegetation was represented by reflectance values in a Landsat TM/ETM + satellite image composite that covers the whole Amazon basin (Van doninck & Tuomisto, 2018) at 30 m resolution. The Landsat TM/ETM + composite is based on all Landsat acquisitions from July to September in the 10-year period 2000–2009, and it was produced using a method that minimizes the effect of atmospheric disturbances and other sources of noise. We used 15-by-15-pixel windows centred on each sampling locality, to extract, for each Landsat band separately, the median and standard deviation of the reflectance values. Climate variables included values of the 19 bioclimatic variables from BIOCLIM (present; Fick & Hijmans, 2017) and were included only in the total environmental distance (explained below) because evaluation of present climate alone over genetic distance was assessed using the resistance matrix constructed in Circuitscape based on climatic niche model.

Soil and vegetation variables were only available for the Amazonian samples. In analyses including these variables, samples with missing values were excluded. Environmental distance matrices were calculated with Euclidean distance and standardized variables. Matrices were created first for each class of variables (bioclimatic, soil, and Landsat), separately and then with all variables combined (total environmental distance).

Mantel tests of matrix correlation were performed using the R package VEGAN 2.4–5 (Oksanen et al., 2017). Correlation was inferred between the genetic distance matrix and the matrices of: geographical distance (IBD) calculated with the Geographical Distance Matrix Generator (Ersts, 2018); climatic (present) based on the Circuitscape resistance matrix, and environmental heterogeneity (soil, canopy reflectance and elevation). In cases where environmental variables were significantly correlated with genetic distance,

partial mantel tests were conducted using the geographical distance matrix as the control matrix to assess if the environmental and genetic distances were still correlated after removing the correlation with geographical distances. As Mantel tests have been criticized regarding their application in spatial analysis (Legendre, Fortin, & Borcard, 2015; Mermains, 2015), a distance based Redundancy Analysis (dbRDA) was also employed. This consists of an ordination-based approach, which eliminates some of the linearity problems related to genetic distance and may be better at detecting IBR (i.e., environmental distance) scenarios (Kierepka & Latch, 2015). For that, the genetic distance matrix was transformed using a PCoA and the geographic coordinates were transformed in a third-degree orthogonal polynomial according to the script from Mermain (2015). A stepwise model selection using the ANOVA.cca function in the R package VEGAN 2.4–5 (Oksanen et al., 2017) was performed in each environmental variable data set to select for significant variables. When variables were associated with genetic distances, we performed dbRDAs controlling for geographic distance similarly to the partial Mantel test procedure. Results were significant when $p \leq .05$.

2.8 | Demographic parameters estimations

In order to infer the demographic history of the different populations, G-PhoCS (Gronau, Hubisz, Gulko, Danko, & Siepel, 2011) was used to estimate effective population size, divergence time and migration rates between current and past populations. These parameters can provide useful insight to understand a population's evolutionary history and response to historical events such as climatic fluctuation or introgression. For that, the 2,563 ddRAD concatenated loci was used. G-PhoCS uses a MCMCcoal model, which employs a full-likelihood multispecies coalescent framework using whole sequences. Four samples per population were selected (see Supporting information 1) based on the quality of data (number of shared loci with the remaining samples) due to computational limitation. Individuals were assigned to populations according to their largest ancestry coefficient derived from the population structure analysis with $K = 4$, and the model topology was based on the phylogenetic results. Because of uncertainties related to phylogenetic analyses (UCE tree), two models (topologies) were used to estimate parameters. Migration was evaluated only between geographically adjacent populations. The following parameters were estimated for each phylogenetic node: θ (theta), which was further converted to effective population size using $\theta = 4N_e\mu$; τ (tau), the mutation rate-scaled divergence time, converted into divergence time in years with $\tau = T_\mu$ (where T is the divergence time in generations); and m, the per-generation migration rates between current populations, scaled by mutation rate ($=M_{ST}/_\mu$) where M_{ST} is the per-generation proportion of individuals in the receiving population T having immigrated from the source population S (Gronau et al., 2011). Three independent runs were performed with each model using 1,500,000 MCMC sampling every 50 generations. Convergence of the parameters from the different runs and ESS values were evaluated using Tracer 1.6 (Rambaut, Suchard, Xie, & Drummond, 2014) with 300,000 chains as burnin. Because migration parameters converged to different values, we ran the final model without the migration bands. We applied a mutation rate μ of 2.5×10^{-9} substitution per site per generation (Nadachowska-Brzyska, Li, Smeds, Zhang, & Ellegren, 2015) and assumed two different generation times to interpret temporal results: one year (commonly used in the literature for birds; Oswald, Overcast, Mauck, Andersen, & Smith, 2017) and 2.33 years, estimated for other Neotropical passerine species (Maldonado-Coelho, 2012).

2.9 | Phenotypic data

Because the species' populations vary in frequency of fixation of the different states of plumage pattern (i.e., pectoral band extension), skin specimens whose tissue or blood samples were used for the genomic analyses, were examined and attributed to one of five different categories (scores) according to how complete/incomplete their pectoral bands were, following the same method and categories proposed in Buainain et al. (2017) (Figure 2a). Only adult specimens were used. Because females have faded and sometimes subtle pectoral band, they were only used when the character was unmistakable. Two samples (FMNH 427227, MPEG T12718) did not have their skin examined but had their plumage pattern inferred based on extensive analysis of other individuals from similar collection localities (Buainain et al., 2017). The two blood samples did not have their plumage scores taken due to the lack of photographs of the individuals. Plumage pattern was plotted on the phylogenetic tree and geographically compared with STRUCTURE results in order to assess spatial congruence between genomic and phenotypic variations. Because tissue samples were selected to represent most of the morphological variation present in the areas, phenotypical representativeness does not necessarily reflect the real frequency of phenotypes shown in Buainain et al. (2017).

3 | RESULTS

3.1 | Population structure

The Structure result for $K = 2$ suggests differentiation north and south of the Amazon River (Figure 2c); $K = 3$ adds a split between east and west, south of the Amazon river (Figure 2d); and $K = 4$ further splits the eastern portion of populations in south-central Amazonia (mostly east of the Tapajos River) + Cerrado, and southeastern Amazonia + Atlantic Forest populations (Figure 2b). Result for $K \geq 5$ shows further splits mainly in central Amazonia, which were very difficult to biologically interpret and were considered as analytical noise due to unrealistic over-splitting and admixture (Figure 2e). Because $K = 4$ represented the finest and most parsimonious and biologically meaningful genetic structure in the data, we adopted four populations in subsequent analyses: Guyana (GUY—blue);

Atlantic Forest and Southeastern Amazonia (Eastern: EAS—green); South-Central Amazonia and Cerrado (South-Central: SC—yellow); and Southwestern Amazonia (SW—orange; Figure 2b). Some individuals in central Amazonia (between Madeira and Xingu rivers) have admixed genotypes between SW and SC, and the area east of the Tocantins river includes individuals with admixture between SC and EAS populations (Figure 2b).

Most individuals from areas with predominantly homogeneous plumage patterns (Guyana, Atlantic Forest and extreme southwestern Amazonia), were not admixed (Figure 2b), while most individuals with admixed genotypes were found in the polymorphic zone. In this large polymorphic area, there was no clear association between genotypes and phenotypes (Figure 2a,b).

3.2 | Phylogenetic analyses

In the maximum likelihood phylogeny from the ddRAD concatenated sequences, the first split within *A. taciturnus* separates all samples from Guyana (blue in Figure 2b) from a well-supported clade containing all the remaining samples (Figure 3). Relationships within the latter were mostly poorly supported, although two internal clades had high support: one including samples from Atlantic Forest and the other including samples from extreme southwestern Amazonia. Thus, individuals from the three areas with predominantly homogeneous plumage (Guyana, Atlantic Forest and extreme southwestern Amazonia) form distinct and well or moderately supported clades. Atlantic Forest and Guyana, despite sharing a similar plumage character (mostly complete pectoral band), are not each other's closest relatives.

The UCEs species tree (SVDquartets) produced an overall well supported phylogeny and clarified the main phylogenetic relationships, which was not possible to observe in the ddRAD tree (Supporting Information 2). The UCE tree confirms a Guyana clade (blue), sister to all other samples. Additionally, it shows that the remaining samples form three major groups, which are congruent with the populations delimited by STRUCTURE (SC, SW and EAS—yellow, orange and green). The relationships among the three southern populations is unclear.

3.3 | Summary statistics

The GUY population has the largest genetic diversity as measured by observed heterozygosities, gene diversity and theta, followed by the EAS, SW and SC populations (Table 1). Genetic differentiation (F_{ST}) is modest but significant among all populations (Table 2). The largest genomic divergence is between GUY and EAS populations, followed by SC and EAS population and SC and SW populations.

3.4 | Species climatic niche

The most likely model (with both the lowest AICc and highest AUC values) was the one with Linear + Quadratic features and a 0.5 Beta-multiplier factor (Model rank and AIC_c values can be seen in

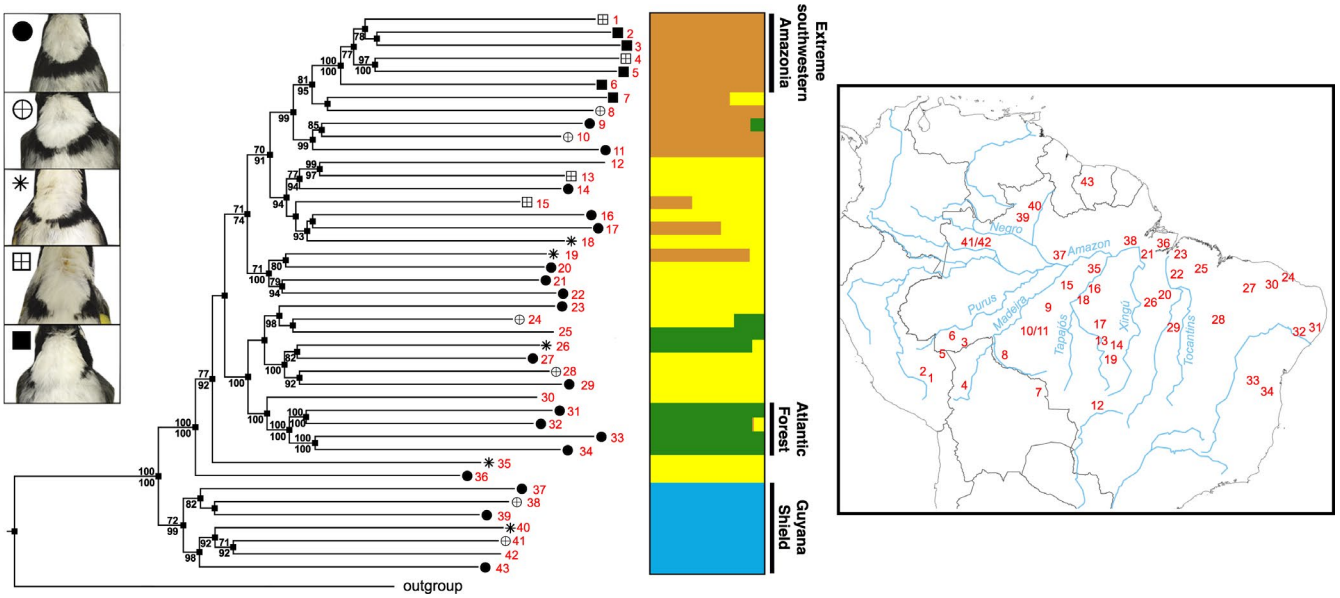

FIGURE 3 Maximum likelihood phylogenetic tree of *Arremon taciturnus* performed with 2,563 ddRAD concatenated sequences in IQtree. Symbols on tips represent the morphological categories (legend on left) of samples. Samples with no symbols have unknown phenotype. Coloured bars are the ancestry coefficient graphs for $K = 4$ in the genetic structure (STRUCTURE) analysis (See Figure 2b). Statistically relevant node support have ≥95 for UltraFast-Bootstrap (above branches) and ≥80 for SH-aRLT (below branches). Node support values below 70 were omitted. Red numbers on the trees' tips correlate with the ones on the map showing their geographic positions. *Arremon semitorquatus* was used as outgroup [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Genetic diversity summary statistics of the different populations of *Arremon taciturnus* delimited by population structure analysis (STRUCTURE; $K = 4$). Summary statistics were calculated using 647 ddRAD SNPs

Population	Expected heterozygosity (mean)	Gene diversity (mean)	Theta *H* (infinite allele model)
Southwestern Amazonia (SW)	0.279	0.037	0.117
Eastern (EAS)	0.383	0.041	0.1
Guyana Shield (GUY)	0.422	0.049	0.139
South-Central Amazonia (SC)	0.257	0.039	0.128

TABLE 2 Pairwise genetic distances based on Fst between populations of *Arremon taciturnus* delimited by population structure analysis (STRUCTURE, $K = 4$). All results were statistically significant ($p \leq .05$). F_{st} values were calculated using 647 ddRAD SNPs

Population	Southwestern Amazonia	Eastern	Guyana Shield	South-Central Amazonia
Southwestern Amazonia (SW)	0.000			
Eastern (EAS)	0.154	0.000		
Guyana Shield (GUY)	0.142	0.189	0.000	
South-Central Amazonia (SC)	0.040	0.075	0.103	0.000

Supporting Information 3). The variables that contributed the most to the model were: precipitation of the coldest quarter (22.4%), precipitation of the wettest month (19.5%), temperature seasonality (16.2%), and precipitation seasonality (15.4%) (Supporting Information 3). Overall, climatic suitability for the species was mostly related to intermediate precipitation seasonality, low temperature seasonality and high precipitation.

Model prediction for the present (Figure 4) shows that the best climatic suitability lies in the Guyana Shield, east and south-central Amazonia, and relatively small regions in extreme southwestern Amazonia in Peru and Bolivia, central highlands in Brazil, and northern Atlantic Forest. The model failed to detect suitability in most localities in Cerrado gallery forests probably due to the microenvironmental condition of these areas. On the other hand, the model predicted low suitability in a large portion of western Amazonia, between the Purus and Negro rivers, from where the species seems to be absent, suggesting a possible climatic influence to this absence.

Overall, comparative visual inspection of the projections suggests general increase in suitability from the past to the present, with the Mid-Holocene and the present being the most suitable, allowing connections and expansion of populations (Figure 4). On the other hand, during the Last Glacial Maximum (CCSM) and notably Last Interglacial, the models predict less suitability thus favour fragmentation and isolation of populations, especially south of the Amazon River. Considering all models, Guyana, extreme southwestern Amazonia and Atlantic Forest always had at least a small area with high climatic suitability. The mean model indicates that two large areas in central Amazonia and one in northeastern Brazil were generally unsuitable for the species over time and models.

3.5 | Association of current environmental variables with genetic distance

The variables associated with genetic distance were geographic distance, total environmental distance (all variables combined), climate (derived from resistance matrix of climatic niche modelling), and soil (Table 3). After excluding the effect of geographic distance, only the total environmental distance and soil were still associated with genetic distance using the Mantel test (although significance test with soil was borderline, $p = .06$), while no associations with environmental variables were detected with the dbRDA test.

Because phylogenetic and structure analyses indicated initial differentiation of the Guyana population (a genomic break related to the Amazon River), the tests were also performed using only samples from areas south of that river. The results were similar, except there was an association with canopy reflectance, but not with soil. After removing the effect of geographic distance, associations with climatic distances was still relevant but weaker and borderline significant in both tests, while association with total environment distance was relevant and significant only in the Mantel test.

Overall, results show a modest effect of geographic distance and environmental heterogeneity (represented by current climate, total environment, and soil conditions) on the current distribution of genetic variation within the species, especially south of the Amazon River. The association (R and R^2), even with multiple variables combined, was never greater than .30.

3.6 | Population history and demography

Demographic parameter estimates were very similar between both topological models tested indicating that populations diverged very recently and almost simultaneously (from 380 to 287 kya considering generation time of 2.33 years and from 166 to 123 kya considering it one year) (Figure 5, Supporting Information 4). After the north–south divergence across the Amazon river, effective size of the southern population was very low suggesting either a bottleneck or a dispersal event and founder effect. Subsequent divergences resulted in populations with greater effective population sizes indicating population expansion for all populations south of the Amazon River, possibly favouring secondary contact, while population contraction occurred in Guyana. The only substantial difference between alternative topology models is that model B indicates a

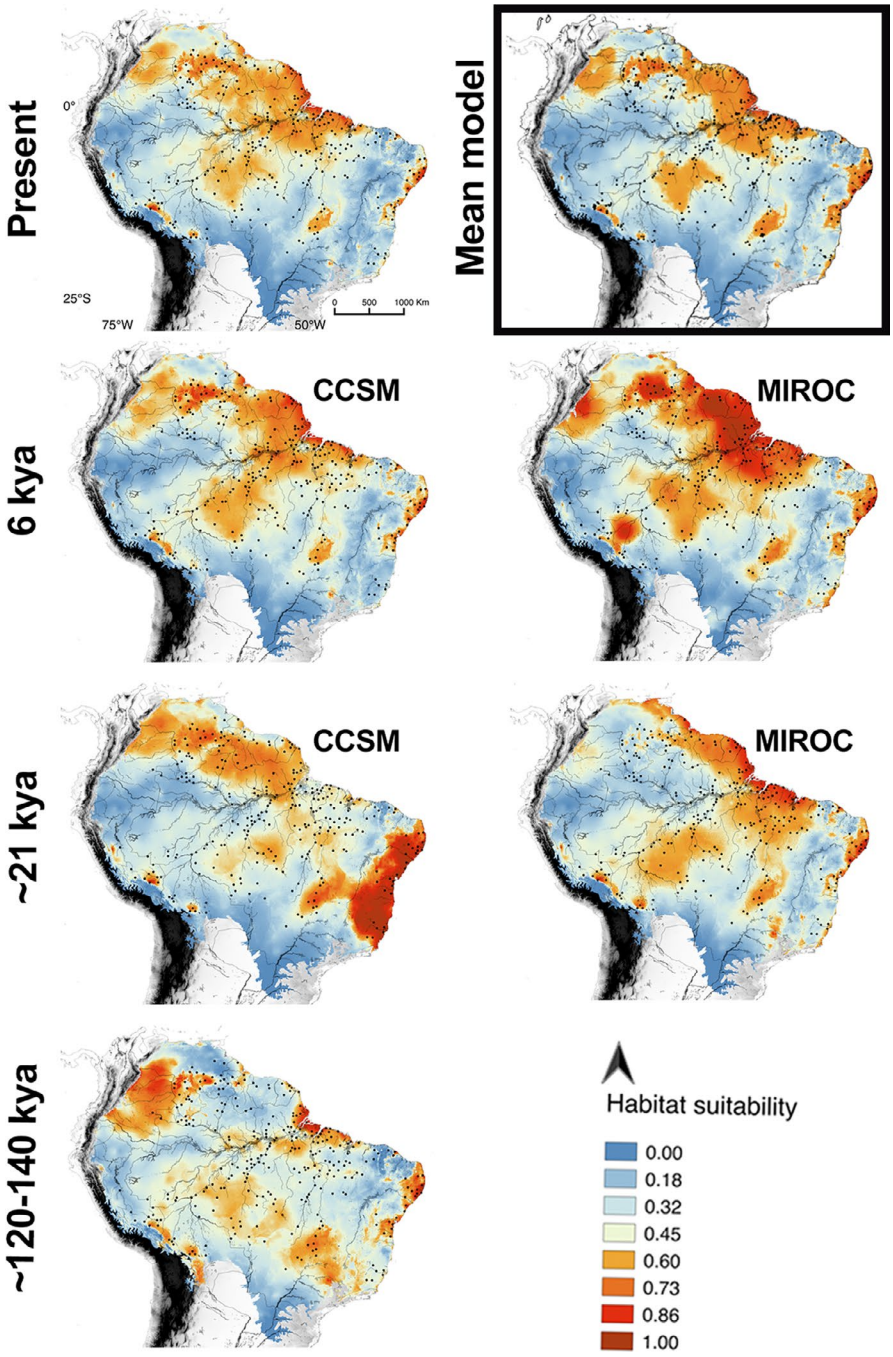

FIGURE 4 Predictions of climatic habitat suitability for *Arremon taciturnus* during four different geological time periods. From top to bottom: Present and mean model, Mid-Holocene (~6 kya), last glacial maximum (LGM) (~21 kya) and last interglacial (LIG) (~120–140 kya). Mid-Holocene and LGM were projected using two different climatic models, CSSM (Community Climate System Model) and MIROC (Model of Interdisciplinary on Climate). Mean model represents the averaged suitability between all models [Colour figure can be viewed at wileyonlinelibrary.com]

second reduction of effective size after the first node in the southern population and before the next split, suggesting a new colonization/bottleneck event. Migration rates could not be confidently estimated from the available data so the models including migration rates were discarded.

4 | DISCUSSION

4.1 | Geographical pattern and associated drivers

Overall, areas with predominantly pure genotypes (Guyana, extreme southwestern Amazonia and Atlantic Forest) correspond to well supported clades (Figure 2) and are congruent with areas of predominantly homogeneous plumage pattern (Figure 1a) and high climatic suitability through time. Therefore, these are candidate refugia areas. Contrastingly, areas where most of the admixed genotypes were found (south-central, southeastern Amazonia and northeastern Brazil) harbour more plumage polymorphism, variable climatic suitability through time, and more gaps of climatic suitability, and thus these areas have probably been more favorable to population fragmentation and secondary contact. Divergence times among populations (last ~380 kya) date to the Late Pleistocene, when glacial cycles were most intense, and most populations effective sizes increased after each divergence event. Thus, patterns of population genomic diversity within *A. taciturnus* fulfill general expectations

TABLE 3 Results of Mantel tests and dbRDA (distance based Redundancy Analysis) showing association between pairwise genetic distances (Fst) of *Arremon taciturnus* with various environmental distances in South American forests and also geographic distance. Values are presented as correlation coefficient (R), coefficient of determination (R^2) and significance test (p)

	Mantel test		dbRDA	
Environmental distance	***R***	***p***	R^2	***p***
All samples				
Geographic	**.23**	**.003**	**.30**	**.001**
Total environmental	**.27**	**.012**	**.16**	**.010**
Total environmental controlling for geographic	**.25**	**.016**	.02	.590
Climatic	.14	.110	**.09**	**.005**
Climatic controlling for geographic	NA	NA	.008	.570
Canopy reflectance (median)	−.12	.850	.008	.340
Canopy reflectance (SD)	−.18	.930	.09	.400
Soil	**.14**	**.047**	−.002	.430
Soil controlling for geographic	.12	.060	NA	NA
Elevation	.03	.350	−.008	.412
Only samples south from the Amazon River				
Geographic	**.22**	**.003**	**.15**	**.001**
Total environmental	**.25**	**.036**	**.26**	**.007**
Total environmental controlling for geographic	**.25**	**.028**	.03	.509
Climatic	**.19**	**.05**	**.27**	**.031**
Climatic controlling for geographic	.18	.083	.16	.060
Canopy reflectance (median)	−.11	.83	.13	.083
Canopy reflectance (SD)	−.12	.796	**.20**	**.02**
Canopy reflectance (SD) controlling for geographic	NA	NA	.20	.272
Soil	.09	.146	.01	.331
Elevation	.07	.45	.04	.351

Statistically significant association was determined based on $p \leq .05$ (values in bold). For variables that were significantly associated with genetic divergence, partial tests were performed to remove the effect of geographic distance on that variable. The first part of the table shows results obtained with all samples, the second part with only samples from areas south of the Amazon river to exclude its effect. Total environmental distances were based on current climate, soil base cation concentration, elevation and canopy reflectance data combined. Climatic distance was calculated using a resistance matrix calculated with Circuitscape and the climatic niche model. NA values show partial tests that were not performed with Mantel or dbRDA because variables were not associated with genetic distances in one of those methods. The genomic distance matrix was constructed using 647 ddRAD SNPs.

derived from a prominent role of paleoclimatic changes in the evolutionary history of the species.

The Amazon river is the only riverine barrier to spatially coincide with sister clades and genetic clusters. IBD and current habitat heterogeneity (represented by soil, total environment and current climatic variables) were associated with genetic distances, but with relatively low coefficients (Table 3), thus suggesting a minor role in explaining current genetic diversity within the species.

Below we detail the most likely diversification scenario based on our combined results, and discuss the potential importance of each driver considering the species' intrinsic ecological preferences and the available information about the landscape history of northern South America.

4.2 | General diversification scenario

The divergence time between north and south populations is estimated as 160–380 kya, a little younger than the estimate based on mtDNA (confidence interval of 430–730 kya) (Moura et al., 2020). Thus, the north-south divergence is younger than the establishment of the transcontinental Amazon river (~10–2 Ma; Bicudo, Sacek, de Almeida, Bates, & Ribas, 2019; Hoorn et al., 2010; Latrubesse et al., 2010), indicating that populations were able to cross the river after its establishment. However, the population structure related to the river remains, indicating that crossing of the barrier occurred along the history of the species, but was not constant enough to avoid genetic differentiation.

The combined results obtained here suggest a north to south dispersal because (a) the Guyana Shield population has the largest within-population effective size and genetic diversity; and (b) the sister taxon of the species, *Arremon axillaris* P. L. Sclater, 1855, occurs on the eastern base of northeastern Andes in Colombia and Venezuela, adjacent to the Guyana Shield (Buainain et al., 2017). After dispersal across the Amazon river, subsequent structuring of populations in the southern portion of the species distribution was probably related to varying habitat availability due to less favorable climatic conditions, as suggested by the species climatic niche models. Populations were probably isolated in at least three regions south from the Amazon River (extreme southwestern and south-central Amazonia and Atlantic Forest). As recent climatic suitability increased, there was an increase in effective population size in all southern populations, favouring secondary contact. Secondary contact is further suggested by extensive polymorphism in the proposed contact zone in opposition to morphological homogeneity found in the areas proposed as refugia (Atlantic Forest and extreme southwestern Amazonia) and the numerous individuals with admixed genotypes.

4.3 | Interplay between paleoclimatic change and riverine barriers

The recent divergence across the Amazon river suggests that individuals either actively crossed the Amazon River when the channel was narrower, or were moved passively through river capture events (as already reported for other rivers; Almeida-Filho & Miranda, 2007; Latrubesse, 2002; Ruokolainen, Moulatlet, Zuquim, Hoorn, &

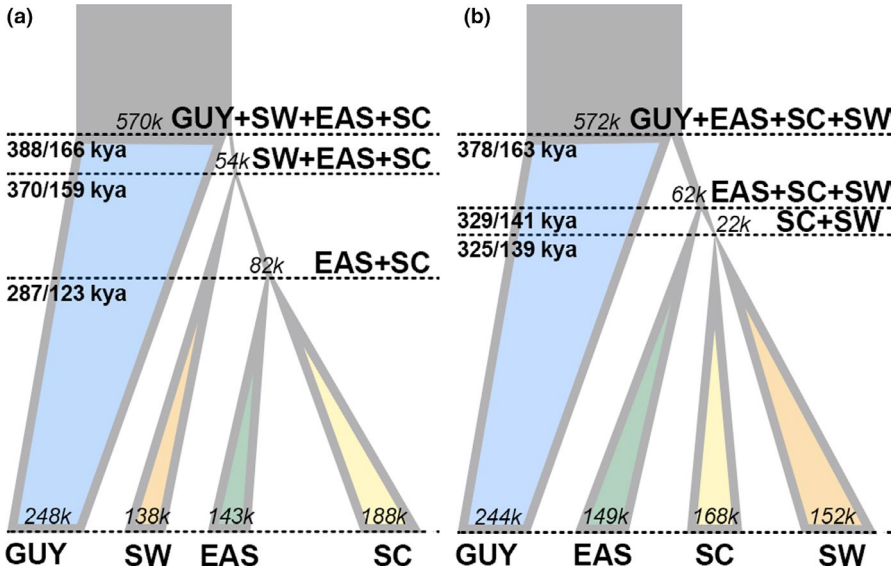

FIGURE 5 Diversification models of *Arremon taciturnus* populations inferred from G-PhocS using the 2,563 ddRAD concatenated loci. The following demographic parameters were estimated for the two most likely topologies (a and b) to accommodate the uncertainties of the phylogenetic tree based UCEs (Supporting Information 2) and the population structure (ddRAD) results: Effective population size (N_e) in italics and illustrated using tree branch width, and divergence time considering generation times of 2.33 and 1.0 years in bold. Migration rate is not shown because results were not consistent between runs. Populations colours correspond to population structure results for $K = 4$ shown in Figure 2b: EAS (Eastern: Atlantic Forest + Southeastern Amazonia), GUY (Guyana), SW (Southwestern Amazonia) and SC (South-Central Amazonia) [Colour figure can be viewed at wileyonlinelibrary.com]

Tuomisto, 2019). This suggests that even the largest Amazonian river can vary in its permeability through time (Pupim et al., 2019). During periods of higher precipitation, increased river discharge in the sedimentary basin caused erosion and river channel incision in river tracts that were previously over unconfined channels and consequently held large floodplain areas. Reduction of floodplains caused reduction of associated seasonally flooded habitats (Varzea) allowing expansion of terra firme forests into these areas (Pupim et al., 2019). This could facilitate crossing the Amazon River by upland taxa, since *terra firme* from both margins would be geographically closer. This idea is supported by the recurrent finding, in birds specialized in seasonally flooded habitats, of genetic and/or phenotypic breaks at the central portion of the Amazon River floodplain (Canton, 2014; Thom et al., 2018, 2020).

In the southern part of the species distribution, the SW genotype (orange, Figure 2b) is present in both banks of the Madeira and Tapajós rivers and the SC genotype (yellow, Figure 2b) is present in both banks of the Tapajós, Tocantins and Xingu rivers. This shows that genomic population structure does not coincide with the position of the current rivers and thus, individuals were/are able to cross these barriers, possibly using gallery forests at their headwaters, as suggested by recent studies in the upper Tapajós basin (Pulido-Santacruz, Aleixo, & Weir, 2018; Weir, Faccio, Pulido-Santacruz, Barrera-Guzman, & Aleixo, 2015).

In contrast to the south-central and southeastern Amazonian rivers, the Purus and Negro (or the Caquetá, located further west, M. Rego personal communication) rivers seem to currently delimit the geographical distribution of *A. taciturnus* (Buainain et al., 2017). Although some specimens examined here are from the right bank of the upper Negro River, it is clear that the species is largely absent or has low abundance in this extensive western Amazonian region. Nutrient-poor soils and large patches of white-sand forests in the Negro basin (Adeney, Christensen, Vicentini, & Cohn-Haft, 2016) may contribute to this pattern. Additionally, climatic niche modelling suggests that the region between Purus and Negro rivers is unsuitable for the species, and edaphic and vegetational conditions of the area are distinct from other regions of the Amazon (Tuomisto et al., 2019) where *A. taciturnus* is known to occur. It is likely that unsuccessful colonization of the area is due to an interplay between riverine barrier effect and unsuitable habitat, a similar scenario to the one proposed by Tuomisto and Ruokolainen (1997).

4.4 | Paleoclimate and historical changes on vegetation structure south of the Amazon River

The late Pleistocene (~2 Ma) was marked by great climatic instability, especially after the last 900 kya when increased orbital forcing triggered more intense cyclical climatic changes (Maslin & Brierley, 2015). Periods of lower temperature and precipitation led to changes in forest structure or fragmentation (Arruda et al., 2018; Haffer, 1969) in Amazonia (especially in eastern portion, Cheng et al., 2013; Wang et al., 2017) and in the Atlantic Forest (Batalha-Filho et al., 2013).

Two genetic breaks were relatively congruent with two areas of past climate instability: central Amazonia between SW and SC populations, and northeastern Brazil between the SC and EAS populations (although admixed individuals containing the

EAS genotype are found in southeastern Amazonia, east of the Tocantins river). These same areas of climatic instability are also recovered in other studies (Batalha-Filho et al., 2013; Bonaccorso, Koch, & Peterson, 2006; Cheng et al., 2013; Prates, Penna, Rodrigues, & Carnaval, 2018; Silva et al., 2019). The historical isolation of populations however, does not corroborate the idea of numerous refugia located in the different interfluves, as initially proposed by Haffer, but instead, two main historically stable forest areas in southern Amazonia (southwestern and south-central) and one in Atlantic Forest.

Despite evidence supporting the presence of savannas in southern Amazonia (Absy, Cleef, D'Apólito, & da Silva, 2014; Rossetti et al., 2019) during the Last Glacial Maximum, some authors are skeptical regarding the extent of such formations due to the general underrepresentation of paleopollen typical of savannas (Colinvaux, de Oliveira, & Bush, 2000). Nevertheless, paleopollen record in Amazonia is largely incomplete, making any current conclusion based on it unreliable (Rocha & Kaefer, 2019). *Arremon taciturnus* is tolerant to forest disturbance, occupying both primary and secondary forests, even in drier vegetation types, gallery forest, and small fragments (Cavalcante, 2015; Dantas, 2017; INPA A 992 in Supporting information 1). Thus, isolation of populations would only be attained due to the expansion of highly unsuitable habitats for the species, such as open vegetation savanna, white-sand and/or floodplain forests. Currently, there are areas of open and dry vegetation adjacent to the Atlantic Forest (Caatinga), SW (savannas at the Parque Nacional Campo Amazônicos near Humaitá and Noel Kempff National Park in Santa Cruz, Bolivia) and SC (*Cangas* at Serra dos Carajás, Pará) populations. These areas may have expanded during dry periods, isolating or reducing gene flow between the three stable areas that harbour distinct populations.

From a biological perspective, there are several open vegetation species currently distributed both in the Brazilian Cerrado and in savannas either in the Marajó island or north of the Amazon river, corroborating past connections between these open habitats that are currently isolated by forest (Quijada-Mascareñas et al., 2007; Resende-Moreira et al., 2019; Silva & Bates, 2002; Vargas-Ramírez,Maran, & Fritz, 2010). A connection through a central Amazonian corridor is further supported by phylogeographic and ecological niche modelling studies (Bonaccorso et al., 2006; Quijada-Mascareñas et al., 2007; Resende-Moreira et al., 2019; Silva et al., 2019; Vargas-Ramírez et al., 2010). A recent study using a combination of Earth Systems models and palynological records suggests that the combination of dry and cold climate, low levels of CO_2 and natural fires during glacial events potentially had a major impact in southern Amazonia and northeastern Brazil, triggering cyclical reductions in tree cover, open vegetation corridor formation and extensive expansion of savannas and dry grass/shrub mosaics during Pleistocene glaciations (H. Sato, personal communication). These could have isolated or drastically reduced connectivity between forest fragments in southwestern and south-central Amazonia, as well as between Amazonia and Atlantic Forests.

Phylogenetic and structure analyses indicate that the Atlantic Forest individuals are closely related to the ones in southeastern Amazonia showing a past connection between those two areas. It is unclear however, if the two populations are currently connected—perhaps by the seasonally dry forests between them—or if the population in southeastern Amazonia is a relict and isolated part of the EAS population, since individuals from that area are polymorphic, while Atlantic Forest individuals have homogeneous phenotype. The current presence of two different genotypes (SC Amazonia and EAS; Figure 2b) in the seasonally dry forests of northeastern Brazil points to a complex scenario of recurrent connections or habitat fragmentation in the area. Other studies have suggested recurrent connection and disconnection between Amazonia and Atlantic Forest through cycles of expansion/retraction of arid vegetation into their adjacent areas (Batalha-Filho et al., 2013; Byrne et al., 2018; Ferreira et al., 2016; Prates et al., 2018; Zhang & Jiang, 2014). Different connection routes have been proposed but our divergence time results are compatible with the "younger connection" established in northeastern Brazil, either through Caatinga, where relict rain forest known as Brejos de Altitude is currently found, or through gallery forests in Tocantins and Bahia States (Batalha-Filho et al., 2013).

4.5 | IBD and habitat heterogeneity

Current habitat heterogeneity and geographic distance seem to have a relevant but minor contribution to the current distribution of genetic diversity within *A. taciturnus*, superimposed to the influence of historical landscape change. The association with geographic distance suggests that there is an effect of IBD (or other spatially autocorrelated environmental variables), but this is not the main factor explaining the distribution of genetic variation due to the low coefficient found (≤0.30) and the presence of sharp genetic breaks instead of gradual changes. Yet, IBD has to be accounted for and explains part of the variation considering such a wide distribution. Similarly, current habitat heterogeneity (represented mainly by climatic variables, soil and total environment variables) possibly helps to maintain current genetic variability through local adaptation. These results highlight the importance of considering both current and past landscapes when studying current patterns of diversity, as the history of populations carries the history of the landscapes within it. Teasing apart the influence of current and past landscapes is key for obtaining information about the environment from patterns of biotic diversity.

4.6 | Contrast between genomic and mitochondrial pattern

Haplotype networks based on mtDNA seem to indicate incipient isolation in portions of the distribution of *A. taciturnus*, which is not recovered with genomic data. However, population structure has not been explicitly tested using mtDNA data. Perhaps the most

striking difference is that in mtDNA, samples from southeastern Amazonia (south of the Amazon, east of the Xingu river) and samples from the Guyana area (north of the Amazon) appear as sister groups with high to moderate support (fig. 3 in Moura et al., 2020), while the genomic data places the SC Amazonia population (yellow in Figure 2b) in a clade that also includes all other samples from localities south of the Amazon river with high support. One possible explanation is an mtDNA introgression from the Guyana area to southeastern Amazonia, establishing a new mitochondrial lineage. It is possible that this introgression happened when SC Amazonia's effective population size was small and the new lineage was fixed due to rapid population expansion. Similar scenarios have been proposed to explain incongruences between nuclear and mitochondrial patterns in other Amazonian bird species (Ferreira et al., 2018), and the accumulation of genomic data will reveal how common these incongruences are.

In general, the comparison of the mtDNA and genomic pattern in *A. taciturnus* suggest that: (a) climatic and habitat distribution changes possibly coupled with riverine barriers resulted in cycles of historical isolation and contact that left unique signals in mitochondrial and genomic DNAs; and (b) after isolation, expansion and secondary contact played an important role in mixing genotypes across rivers and habitats.

4.7 | Genomic diversity and phenotypic variation

Plumage variation in *A. taciturnus* is congruent with the genomic signal of population structure mostly in the geographical extremes of the species distribution where phenotypes are predominantly homogenous. In the remaining areas however, there is no consistent relationship. It is possible that the complete and incomplete pectoral bands were fixed on the extremes of the distribution during isolation, but secondary contact later mixed up the different morphotypes. It is not clear, however, if polymorphism in the SC Amazonia population—which is located in the middle of the polymorphic zone—is derived exclusively from secondary contact or originated in situ. The past climatic niche modelling analyses suggests cyclical unsuitability for populations in the polymorphic zone, supporting the isolation and secondary contact scenario. Accordingly, the Guyana and Atlantic forest populations, despite having predominantly complete pectoral bands, are not each other's closest relatives. It is possible that the SC population also had a complete pectoral band—which seems to be the ancestral character state—but past or current gene flow from the SW population caused introgression of the incomplete pectoral band on most southern Amazonian populations. In any case, it seems like primary differentiation is the less likely scenario due to the signal of isolation evidenced by genomic and mtDNA and the low association with geographic and environmental distances.

Regarding vocal variations in *A. taciturnus*, Buainain et al. (2017) showed that populations from different areas of endemism (mainly Amazonian interfluves, Cerrado and Atlantic Forest) were highly discriminated using multivariate analysis. This is not congruent with genomic data but is somewhat congruent with mtDNA suggesting a genetic component to the vocal variations (Moura et al., 2020). Thus, it remains unclear how much of the species vocal variation is genetic and how much is learned, but the recent development on the evolutionary history of the species should help to design more rigorous tests.

In conclusion, studying the evolutionary history of organisms with more tolerance to different habitats and habitat disturbance can bring new perspectives to the evolution of Neotropical biomes. Our study corroborates that the evolutionary history of *A. taciturnus* is recent (Late Pleistocene) and proposes that it is most likely related to paleoclimatic changes. Evolution of paleoclimate resulted in changes in permeability of rivers to dispersion and substantial modification of habitat configuration—especially in central-eastern Amazonia and northeastern Brazil—that caused genetic differentiation within the species. The overall incongruence between mitochondrial and nuclear DNAs and phenotypic patterns suggests that these changes were complex, and probably result from cycles of isolation and secondary contact among populations.

ACKNOWLEDGEMENTS

We thank the staff of LTBM (Laboratório Temático de Biologia Molecular—INPA) and LEGAL (Laboratório de Evolução e Genética Animal—UFAM), for providing the assistance and structure with lab work; Mateus Ferreira, Leandro Moraes and Marina Maximiano for reviewing earlier versions of the manuscript. We also thank the specimen collectors for their constant hard work, curators and staff of the collections visited in the past to obtain the specimens' records, phenotypic data and tissue loan conceived: Alexandre Aleixo and Fatima Lima (MPEG); Luis Fabio Silveira (MZUSP), Joel Cracraft, Paul Sweet and Lydia Garetano (AMNH), Nate Rice and Jason Weckestein (ANSP), Steve Rogers (CM), Ben Marks and John Bates (FMNH), Brian Schmidt (USNM), Cristina Miyaki (LGEMA/USP), Marlene Freitas (INPA), Robb Brumfield and Donna Dittman (LSUMZ). We thank Romina Batista from MPEG and Glaucia Del Rio, Marco A. Rego and staff of LSUMZ for providing pictures of some museum skins. Jasper Van doninck (University of Turku, Finland) for producing the landsat mosaic. Finally, we thank the editors and four anonymous reviewers for their valuable contributions to the manuscript. CAPES (Coordenação de Aperfeiçoamento Pessoal de Pessoal de Nível Superior) and CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnológico) financed NB; Skin specimens examination was financed by the Collection Study Grant conceived by the AMNH. Part of the project was financially supported by the grant Dimensions US-Biota-São Paulo: Assembly and evolution of the Amazon biota and its environment: an integrated approach, cofunded by the US National Science Foundation (NSF DEB 1241056), the Fundação de Amparo à Pesquisas do Estado de São Paulo (FAPESP grant #2012/50260-6). Academy of Finland (grant to HT) and Finnish Cultural Foundation (grant to GZ).

AUTHOR CONTRIBUTIONS

N.B., and C.R. conceived the project's design and idea, N.B., R.C., T.H. did the laboratory work and performed part of the analyses, N.B. performed the remaining analyses, C.R. provided financial support; G.Z., and H.T. provided part of the environmental data as well as insights on how to analyze and interpret the results. N.B. led the writing with contributions of all authors.

DATA AVAILABILITY STATEMENT

The ddRAD raw and demultiplexed data, the UCEs raw and clean reads, the input for the analyses and the raw environmental variables used in the study can be found at Dryad: Buainain, Nelson et al. (2020), Paleoclimatic evolution as the main driver of current genomic diversity in the widespread and polymorphic Neotropical songbird *Arremon taciturnus*, Dryad, Data set, https://doi.org/10.5061/dryad.280gb5mnd. The ddRAD raw and demultiplexed data are also available at NCBI Sequence Read Archive (SRA) under the BioProject accession PRJNA646310.

ORCID

Nelson Buainain https://orcid.org/0000-0002-6983-165X

REFERENCES

Absy, M. L., Cleef, A. M., D'Apólito, C. D., & da Silva, M. F. F. (2014). Palynological differentiation of savanna types in Carajas, Brazil (southeastern Amazonia). *Palynology*, *38*(1), 78–89. https://doi.org/10.1080/01916122.2013.842189

Adeney, J. M., Christensen, N. L., Vicentini, A., & Cohn-Haft, M. (2016). White-sand Ecosystems in Amazonia. *Biotropica*, *48*, 7–23.

Almeida-Filho, R., & Miranda, F. P. (2007). Mega capture of the Rio Negro and formation of the Anavilhanas Archipelago, Central Amazonia, Brazil: Evidences in an SRTM digital elevation model. *Remote Sensing of Environment*, *110*, 387–392.

Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., … Antonelli, A. (2019). Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology*, *68*(1), 32–46. https://doi.org/10.1093/sysbio/syy039

Antonelli, A., Ariza, M., Albert, J., Andermann, T., Azevedo, J., Bacon, C., … Edwards, S. V. (2018). Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ*, *6*, e5644. https://doi.org/10.7717/peerj.5644

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*, 47–56. https://doi.org/10.1111/geb.12646

Batalha-Filho, H., Fjeldså, J., Fabre, P., & Miyaki, C. Y. (2013). Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithology*, *154*, 41–50. https://doi.org/10.1007/s10336-012-0866-7

Bicudo, T. C., Sacek, V., de Almeida, R. P., Bates, J. M., & Ribas, C. C. (2019). Andean tectonics and mantle dynamics as a pervasive influence on Amazonian ecosystem. *Scientific Reports*, *9*, 16879. https://doi.org/10.1038/s41598-019-53465-y

BirdLife International. (2019). Species factsheet: *Arremon taciturnus*. Retrieved from http://www.birdlife.com (accessed 15 February 2019).

Bonaccorso, E., Koch, I., & Peterson, A. T. (2006). Pleistocene fragmentation of Amazon species' ranges. *Diversity and Distributions*, *12*, 157–164. https://doi.org/10.1111/j.1366-9516.2005.00212.x

Boubli, J. P., Ribas, C., Alfaro, J. W. L., Alfaro, M. E., da Silva, M. N. F., Pinho, G. M., & Farias, I. P. (2015). Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular Phylogenetics and Evolution*, *82*, 400–412. https://doi.org/10.1016/j.ympev.2014.09.005

Buainain, N., Pinto, C. A., & Raposo, M. A. F. (2017). Geographic variation and taxonomy of the *Arremon taciturnus* (Hermann, 1783) species complex (Aves: Passerellidae). *Journal of Ornithology*, *158*(3), 631–650. https://doi.org/10.1007/s10336-016-1421-8

Byrne, H., Lynch-Alfaro, J. W., Sampaio, I., Farias, I. P., Schneider, H., Hrbek, T., & Boubli, J. P. (2018). Titi monkey biogeography: Parallel Pleistocene spread by *Plecturocebus* and *Cheracebus* into a post-Pebas western Amazon. *Zoologica Scripta*, *47*, 499–517. https://doi.org/10.1111/zsc.12300

Canton, R. C. (2014). Análise da variação fenotípica e genotípica do complexo Brotogeris sanctithomae (Aves:Psittaciformes) (MSc thesis). Instituto Nacional de Pesquisas da Amazônia, Manaus.

Cavalcante, E. P. (2015, March 7). WA1911670, Arremon taciturnus (Hermann, 1783) [photograph]. Wiki Aves – A Enciclopédia das Aves do Brasil. Retrieved from http://www.wikiaves.com/1911670

Cheng, H., Sinha, A., Cruz, F. W., Wang, X., Edwards, R. L., d'Horta, F. M., … Auler, A. S. (2013). Climate change patterns in Amazonia and biodiversity. *Nature Communications*, *4*, 1411. https://doi.org/10.1038/ncomms2415

Colinvaux, P. A., de Oliveira, P. E., & Bush, M. B. (2000). Amazonian and neotropical plant communities on glacial time-scales: The failure of the aridity and refuge hypothesis. *Quaternary Science Reviews*, *19*, 141–169.

Cowling, S. A., Maslin, M. A., & Sykes, M. T. (2001). Paleovegetation simulations of lowland Amazonia and implications for neotropical allopatry and speciation. *Quaternary Research*, *55*(2), 140–149. https://doi.org/10.1006/qres.2000.2197

Cracraft, J., Ribas, C. C., d'Horta, F. M., Bates, J., Almeida, R. P., Aleixo, A., Boubli, J. P., … Baker, P. (2020). The origin and evolution of Amazonian species diversity. In V. Rull & A. Carnaval (Eds.), *Neotropical diversification: Patterns and processes*. *Fascinating life sciences*. Champagne, France: Springer.

Cullingham, C., Miller, J., Peery, R., Dupuis, J., Malenfant, R., Gorrell, J., & Janes, J. (2020). Confidently identifying the correct K value using the ΔK method: When does K = 2? *Molecular Ecology*, *29*, 862–869. https://doi.org/10.1111/mec.15374

Dantas, S. M. (2017). WA2583010, Arremon taciturnus (Hermann, 1783) [photograph]. Wiki Aves – A Enciclopédia das Aves do Brasil. Retrieved from http://www.wikiaves.com/2583010

de Abreu, F. H. T., Schietti, J., & Anciães, M. (2018). Spatial and environmental correlates of intraspecific morphological variation in three species of passerine birds from the Purus-Madeira interfluvium, Central Amazonia. *Evolutionary Ecology*, *32*, 191–214. https://doi.org/10.1007/s10682-018-9929-4

Eaton, D. A. R. (2014). PyRAD: Assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*(13), 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Elith, J., Graham, H. C., Anderson, P. R., Dudík, M., Ferrier, S., Guisan, A., … Zimmermann, E. N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Ersts, P. J. (2018). Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Retrieved from http://biodiversityinformatics.amnh.org/open_source/gdmg

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Faircloth, B. C. (2013). Illumiprocessor: A trimmomatic wrapper for parallel adapter and quality trimming. doi: 10.6079/J9ILL.

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 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*, 717–726. https://doi.org/10.1093/sysbio/sys004

Fernandes, A. M., Gonzalez, J., Wink, M., & Aleixo, A. (2013). Multilocus phylogeography of the Wedge-billed Woodcreeper *Glyphorynchus spirurus* (Aves, Furnariidae) in lowland Amazonia: Widespread cryptic diversity and paraphyly reveal a complex diversification pattern. *Molecular Phylogenetics and Evolution*, *66*, 270–282. https://doi.org/10.1016/j.ympev.2012.09.033

Ferreira, M., Aleixo, A., Ribas, C. C., & Santos, M. P. (2016). Biogeography of the neotropical genus Malacoptila (Aves: Bucconidae): The influence of the Andean orogeny, Amazonian drainage evolution and palaeoclimate. *Journal of Biogeography*, *44*, 748–759. https://doi.org/10.1111/jbi.12888

Ferreira, M., Fernandes, A. M., Aleixo, A., Antonelli, A., Olsson, U., Bates, J. M., … Ribas, C. C. (2018). Evidence for mtDNA capture in the jacamar *Galbula leucogastra/chalcothorax* 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

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315. https://doi.org/10.1002/joc.5086

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652. https://doi.org/10.1038/nbt.1883

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*, 1031–1034. https://doi.org/10.1038/ng.937

Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. doi: https://doi.org/10.1093/sysbio/syq010.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, *165*, 131–137. https://doi.org/10.1126/science.165.3889.131

Haffer, J. (1974). *Avian speciation in Tropical South America*, vol. 14. Cambridge, UK: Harvard University. Nuttall Ornithological Club, pp. 1–390.

Haffer, J. (2001). Hypothesis to explain the origin of species in Amazonia. In I. C. G. Vieira, J. M. C. Silva, D. C. Oren, & M. A. D'Incao (Eds.), *Diversidade Biologica e Cultural da Amazonia* (pp. 45–118). Belém, Brazil: Museu Paraense Emílio Goeldi.

Hahn, M. W. (2018). *Molecular population genetics* (p. 334). New York, NY: Oxford University Press.

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2011). Package 'dismo'. Retrieved from http://cran.r-project.org/web/packages/dismo/index.html

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*, 518–522. https://doi.org/10.1093/molbev/msx281

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., … Antonelli, A. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, *330*, 927–931. https://doi.org/10.1126/science.1194585

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kierepka, E. M., & Latch, E. K. (2015). Performance of partial statistics in individual-based landscape genetics. *Molecular Ecology Resources*, *15*, 512–525. https://doi.org/10.1111/1755-0998.12332

Klicka, J., Barker, K. F., Bruns, K. J., Lanyon, S. M., Lovette, I. J., Chaves, J. A., & Bryson-Jr, R. W. (2014). A comprehensive multilocus assessment of sparrow (Aves: Passerellidae) relationships. *Molecular Phylogenetics and Evolution*, *77*, 177–182. https://doi.org/10.1016/j.ympev.2014.04.025

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Latrubesse, E. M. (2002). Evidence of Quaternary palaeohydrological changes in middle Amazonia: The Aripuana-Roosevelt and Jiparana "fans". *Zeitschrift Für Geomorphologie*, *129*, 61–72.

Latrubesse, E. M., Cozzuol, M., Silva-Caminha, S. A. F., Rigsby, C. A., Absy, M. L., & Jaramillo, C. (2010). The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. *Earth-Science Reviews*, *99*, 99–124. https://doi.org/10.1016/j.earscirev.2010.02.005

Ledo, R. M. D., & Colli, G. R. (2017). The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of Biogeography*, *44*, 2551–2563. https://doi.org/10.1111/jbi.13049

Legendre, P., Fortin, M. J., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, *6*, 1239–1247. https://doi.org/10.1111/2041-210X.12425

Maldonado-Coelho, M. (2012). Climatic oscillations shape the phylogeographical structure of Atlantic Forest fire-eyes (Aves: Thamnophilidae). *Biological Journal of Linneal Society*, *105*, 900–924.

Maslin, M. A., & Brierley, C. M. (2015). The role of orbital forcing in the Early Middle Pleistocene Transition. *Quaternary International*, *389*, 47–55. https://doi.org/10.1016/j.quaint.2015.01.047

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, *24*, 3223–3231. https://doi.org/10.1111/mec.13243

Merow, C., Smith, M. J., & Silander, J. A. Jr (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, *36*, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Moura, C. C. M., Fernandes, A. M., Aleixo, A., Araújo, H. F. P., Mariano, E. F., & Wink, M. (2020). Evolutionary history of the Pectoral Sparrow *Arremon taciturnus*: Evidence for diversification during the Late Pleistocene. *Ibis*. Accepted Manuscript, https://doi.org/10.1111/ibi.12813

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution*, *5*(11), 1198–1205.

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 Biololgy*, *25*, 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modeling. *Ecography*, *37*, 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Naka, L. N., & Brumfield, R. T. (2018). The dual role of Amazonian rivers in the generation and maintenance of avian diversity. *Science Advances*, *4*(8), eaar8575. https://doi.org/10.1126/sciadv.aar8575

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274. https://doi.org/10.1093/molbev/msu300

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2017). vegan: Community ecology package. R package version 2.4-5.

Ortiz, D. A., Lima, A. P., & Werneck, F. P. (2018). Environmental transition zone and rivers shape intraspecific population structure and genetic diversity of an Amazonian rain forest tree frog. *Evolutionary Ecology*, *32*, 359–378. https://doi.org/10.1007/s10682-018-9939-2

Oswald, J. A., Overcast, I., Mauck, W. M., Andersen, M. J., & Smith, B. T. (2017). Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. *Molecular Ecology*, *26*, 1386–1400. https://doi.org/10.1111/mec.14013

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*, 1–11.

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, *40*, 887–893. https://doi.org/10.1111/ecog.03049

Pomara, L. Y., Ruokolainen, K., Tuomisto, H., & Young, K. R. (2012). Avian composition co-varies with floristic composition and soil nutrient concentration in Amazonian upland forests. *Biotropica*, *44*, 545–553. https://doi.org/10.1111/j.1744-7429.2011.00851.x

Pomara, L. Y., Ruokolainen, K., & Young, K. R. (2014). Avian species composition across the Amazon River: The roles of dispersal limitation and environmental heterogeneity. *Journal of Biogeography*, *41*, 784–796. https://doi.org/10.1111/jbi.12247

Prates, I., Penna, A., Rodrigues, M. T., & Carnaval, A. C. (2018). Local adaptation in mainland anole lizards: Integrating population history and genome–environment associations. *Ecology and Evolution*, *8*(23), 11932–11944. https://doi.org/10.1002/ece3.4650

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Pulido-Santacruz, P., Aleixo, A., & Weir, J. (2018). Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. *Proceedings of the Royal Society B: Biological Sciences*, *285*, 2017–2081. https://doi.org/10.1098/rspb.2017.2081

Pupim, F. N., Sawakuchi, A. O., Almeida, R. P., Ribas, C. C., Kern, A. K., Hartmann, G. A., … Cracraft, J. (2019). Chronology of Terra Firme formation in Amazonian lowlands reveals a dynamic Quaternary landscape. *Quaternary Science Reviews*, *210*, 154–163. https://doi.org/10.1016/j.quascirev.2019.03.008

Quijada-Mascareñas, J. A., Ferguson, J. E., Pook, C. E., Salomão, M. D. G., Thorpe, R. S., & Wüster, W. (2007). Phylogeographic patterns of trans-Amazonian vicariants and Amazonian biogeography: The Neotropical rattlesnake (*Crotalus durissus* complex) as an example. *Journal of Biogeography*, *34*, 1296–1312. https://doi.org/10.1111/j.1365-2699.2007.01707.x

Rambaut, A., Suchard, M., Xie, W., & Drummond, A. (2014). *Tracer v. 1.6*. Edinburgh, UK: Institute of Evolutionary Biology, University of Edinburgh.

Resende-Moreira, L. C., Knowles, L. L., Thomaz, A. T., Prado, J. R., Souto, A. P., Lemos-Filho, J. P., & Lovato, M. B. (2019). Evolving in isolation: Genetic tests reject recent connections of Amazonian savannas with the central Cerrado. *Journal of Biogeography*, *46*, 196–211. https://doi.org/10.1111/jbi.13468

Ribas, C. C., Aleixo, A., Gubili, C., d'Horta, F. M., Brumfield, R. T., & Cracraft, J. (2018). Biogeography and diversification of *Rhegmatorhina* (Aves: Thamnophilidae): Implications for the evolution of Amazonian landscapes during the Quaternary. *Journal of Biogeography*, *45*, 917–928. https://doi.org/10.1111/jbi.13169

Ribas, C. C., Aleixo, A., Nogueira, A. C. R., Miyaki, C. Y., & Cracraft, J. (2012). A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of the Royal Society London Series B: Biological Sciences*, *279*, 681–689. https://doi.org/10.1098/rspb.2011.1120

Rocha, D. G., & Kaefer, I. L. (2019). What has become of the refugia hypothesis to explain biological diversity in Amazonia? *Ecology and Evolution*, *9*(7), 4302–4309. https://doi.org/10.1002/ece3.5051

Rossetti, D. F., Gribel, R., Cohen, M. C. L., Valeriano, M. M., Tatumi, S. H., & Yee, M. (2019). The role of Late Pleistocene-Holocene tectono-sedimentary history on the origin of patches of savanna vegetation in the middle Madeira River, southwest of the Amazonian lowlands. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *526*, 136–156. https://doi.org/10.1016/j.palaeo.2019.04.017

Ruokolainen, K., Moulatlet, G. M., Zuquim, G., Hoorn, C., & Tuomisto, H. (2019). Geologically recent rearrangements in central Amazonian river network and their importance for the riverine barrier hypothesis. *Frontiers of Biogeography*, *11*, e45046. https://doi.org/10.21425/F5FBG45046

Shah, V. B., & McRae, B. H. (2008). Circuitscape: A tool for landscape ecology. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), Proceedings of the 7th Python in Science Conference (pp. 62–66). http://conference.scipy.org/proceedings/SciPy2008/paper_14/, https://circuitscape.org/pubs.html

Silva, J. M. C., & Bates, J. M. (2002). Biogeographic patterns and conservation in the South American Cerrado: A tropical Savanna hotspot. *BioScience*, *52*, 225–233. https://doi.org/10.1641/0006-3568(2002)052[0225:BPACIT]2.0.CO;2

Silva, S. M., Peterson, A. T., Carneiro, L., Burlamaqui, T. C. T., Ribas, C. C., Sousa-Neves, T., … Aleixo, A. (2019). A dynamic continental moisture gradient drove Amazonian Bird diversification. *Sciences Advances*, *5*(7), eaat5752. https://doi.org/10.1126/sciadv.aat5752

Swofford, D. L. (2002). *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sunderland, MA: Sinauer Associates.

Thom, G., Amaral, F. R. D., Hickerson, M. J., Aleixo, A., Araujo-Silva, L. E., Ribas, C. C., … Miyaki, C. Y. (2018). Phenotypic and genetic structure support gene flow generating gene tree discordances in an amazonian floodplain endemic species. *Systematic Biology*, *67*, 700–718. https://doi.org/10.1093/sysbio/syy004

Thom, G., Xue, A. T., Sawakuchi, A. O., Ribas, C. C., Hickerson, M. J., Aleixo, A., & Miyaki, C. (2020). Quarternary climate changes as speciation drivers in the Amazon floodplains. *Science Advances*, *6*(11), eaax4718. https://doi.org/10.1126/sciadv.aax4718

Tuomisto, H., Moulatlet, G. M., Balslev, H., Emilio, T., Figueiredo, F. O. G., Pedersen, D., & Ruokolainen, K. (2016). A compositional turnover zone of biogeographical magnitude within lowland Amazonia. *Journal of Biogeography*, *43*, 2400–2411. https://doi.org/10.1111/jbi.12864

Tuomisto, H., & Ruokolainen, K. (1997). The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. *Biodiversity and Conservation*, *6*, 347–357. https://doi.org/10.1023/A:1018308623229

Tuomisto, H., Van doninck, J., Ruokolainen, K., Moulatlet, G. M., Figueiredo, F. O. G., Sirén, A., … Zuquim, G. (2019). Discovering floristic and geoecological gradients across Amazonia. *Journal of Biogeography*, *46*, 1734–1748. https://doi.org/10.1111/jbi.13627

Van doninck, J., & Tuomisto, H. (2018). A Landsat composite covering all Amazonia for applications in ecology and conservation. *Remote Sensing in Ecology and Conservation*, *4*, 197–210. https://doi.org/10.1002/rse2.77

Vargas-Ramírez, M., Maran, J., & Fritz, U. (2010). Red- and yellow-footed tortoises, *Chelonoidis carbonaria* and *C. denticulate* (Reptilia: Testudines: Testudinidae), in South American savannahs and forests: Do their phylogeographies reflect distinct habitats? *Organism Diversity & Evolution*, *10*(2), 161–172. https://doi.org/10.1007/s13127-010-0016-0

Wallace, A. R. (1895). *Natural selection and tropical nature: Essays on descriptive and theoretical biology* (2nd ed.). London, UK: Macmillan.

Wang, X., Edwards, R. L., Auler, A. S., Cheng, H., Kong, X., Wang, Y., … Chiang, H.-W. (2017). Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature*, *541*, 204–207. https://doi.org/10.1038/nature20787

Weir, J. T., Faccio, M. S., Pulido-Santacruz, P., Barrera-Guzman, A., & Aleixo, A. (2015). Hybridization in headwater regions, and the role of rivers as drivers of speciation in Amazonian birds. *Evolution*, *69*, 1823–1834. https://doi.org/10.1111/evo.12696

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A. & NCEAS Predicting Species Distributions Working Group (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*, 763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Zhang, R., & Jiang, D. (2014). Impact of vegetation feedback on the mid-Pliocene warm climate. *Advances in Atmospheric Sciences*, *31*(6), 1407–1416. https://doi.org/10.1007/s00376-014-4015-5

Zucker, M. R., Harvey, M. G., Oswald, J. A., Cuervo, A., Derryberry, E., & Brumfield, R. T. (2016). The mouse-coloured 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.

Zuquim, G. (2017). Soil exchangeable cation concentration map of Amazonia link to GeoTif. University of Turku, PANGAEA. Available from https://doi.pangaea.de/10.1594/PANGAEA.879542

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Buainain N, Canton R, Zuquim G, et al. Paleoclimatic evolution as the main driver of current genomic diversity in the widespread and polymorphic Neotropical songbird *Arremon taciturnus*. *Mol Ecol*. 2020;29:2922–2939. https://doi.org/10.1111/mec.15534