PAUL VAN ELS (Orcid ID : 0000-0002-9499-8873)

DR. CAMILA C RIBAS (Orcid ID : 0000-0002-9088-4828)

Article type : Research Article

Corresponding author mail id: paulvanels@gmail.com

Recent divergence and lack of shared phylogeographic history characterize the diversification of Neotropical savanna birds

Phylogeography of Neotropical savanna birds

Paul van Els*[1,2], Eugenia Zarza[3,4], Lucas Rocha Moreira[5], Valentina Gómez-Bahamón[6,7], Antonita Santana[8,9], Alexandre Aleixo[10, 11], Camila C. Ribas[12], Périces Sena do Rêgo[13], Marco Pérsio Dantas Santos[14], Kristof Zyskowski[15], Richard O. Prum[16], Jacob Berv[17,18]

[1]Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA.

[2]Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, Groningen, 9700 CC, Netherlands.

[3]Departamento de Ciencias de la Sustentabilidad, El Colegio de la Frontera Sur. Carretera Antiguo Aeropuerto Km 2.5, Tapachula, Chiapas, Mexico 30700.

[4]Consejo Nacional de Ciencia y Tecnología (CONACYT), Av. Insurgentes Sur 1582, Col. Crédito Constructor, Alcaldía Benito Juárez, Ciudad de México, Mexico 03940.

[5]Columbia University, 1200 Amsterdam Ave, New York, NY 10027, USA.

[6]Department of Biological Sciences, University of Illinois at Chicago, 845 West Taylor Street (MC066), Chicago, IL 60607, USA.

[7]Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605, USA.

[8]Graduate Program in Zoology, Museu Paraense Emílio Goeldi / Universidade Federal do Pará, Belém-PA, Brazil.

[9] Graduate Program in Environmental Biology, Universidade Federal do Pará, Bragança-PA, Brazil.

[10]Department of Zoology, Museu Paraense Emílio Goeldi, 66040170 Belém, PA, Brazil.

[11]Finnish Museum of Natural History, University of Helsinki, 00014, Helsinki, Finland.

[12]Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/JBI.14065

[13]Universidade Federal do Pará, Laboratório de Biogeografia da Conservação e Macroecologia - BIOMACRO, Belem-PA, Brasil.

[14]Instituto de Estudos Costeiros / Universidade Federal do Pará, Belém-PA, Brazil.

[15]Yale University Peabody Museum of Natural History, 170 Whiney Ave, New Haven, CT 06511, USA.

[16]Department of Ecology & Evolutionary Biology, Yale University, New Haven, CT 06520, USA.

[17]Department of Ecology and Evolutionary Biology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA.

[18]University of Michigan Museum of Paleontology, Biological Sciences Building,1105 North University Ave, Ann Arbor, MI 48109, USA.

*corresponding author

ACKNOWLEDGMENTS

We thank the following institutions and their staff for providing samples: Paul Sweet (AMNH), Nate Rice (ANSP), John Bates and Ben Marks (FMNH), Mark Robbins (KU Biodiversity Institute), Fabricio Santos (UFMG), Edson Guilherme (UFAC), Sharon Birks and John Klicka (UWBM), Brian Schmidt (USNM, Smithsonian), Kimball Garrett (LACM), Jorge Enrique Avendaño at the Universidad de los Llanos. James V. Remsen, Jr. kindly reviewed the manuscript. We also would like to thank Donna Dittmann and Steven Cardiff at LSU for help with permitting and loans. Parts of this research were conducted under U.S.D.A. permit 126158 to LSUMZ. We are grateful to Jordan Satler for help with analyzing Phylogeographic Concordance Factors and to Henk Sierdsema (Sovon) for assistance with map design. Funding to PVE was provided by the LSU Museum of Natural Science Birdathon Fund, the Stichting P.A. Hens Memorial Fund, an American Ornithological Society Alexander Wetmore Memorial Research Award, and the American Museum of Natural History Frank M. Chapman Memorial Fund. MPDS was supported by CNPq research productivity fellowships (#308403/2017-7).

ABSTRACT

Aim

Neotropical savanna birds occur north and south of, but mostly not in the Amazon Basin, except for a few isolated savanna patches. Here, we investigate the phylogeography of 23 taxa of Neotropical savanna birds co-distributed across multiple isolated savanna patches to assess to what extent these species have a shared history of spatial diversification. We explore the role of the forested Amazon Basin as a vicariant barrier separating northern and southern populations, particularly focusing on the

role of the coastal savannas of Amapá as a potential corridor of gene flow between northern and southern populations.

Location

Neotropical savannas.

Taxon

Aves.

Method

We employ 775 mtDNA samples of 24 co-distributed savanna bird taxa from all major savanna patches in South America to infer phylogeographic patterns. For this purpose, we use 24 genomic samples (UCEs) of a subset of 12 taxa in addition to the mtDNA samples to estimate timing of divergence across the Amazon Basin. We use phylogeographic concordance factors (PCF) to assess the level of phylogeographic congruence across co-distributed taxa. Finally, we assess to which level physical distance drives genetic structuring by estimating isolation-by-distance (IBD) effects.

Results

We find that although the study taxa generally do not share similar diversification patterns geographically, many have at least two distinct genetic groups, one north and one south of the Amazon Basin, that have only recently diverged. The timing of divergence between both areas is generally centered in the late Pleistocene, but somewhat variable, indicating there is no single vicariant event responsible for driving diversification.

Main conclusions

Variability in divergence times indicates that landscape processes have not led to shared phylogeographic responses, which indicaties a relatively minor role for vicariance. Shallow divergences suggest that Neotropical grassland habitats may have recently been more connected or that gene flow has played an important role. We did not find evidence of a single dominant corridor of dispersal between savannas north and south of the forested Amazon Basin.

Key Words: Cerrado, grassland biogeography, Llanos, Pampas, Pleistocene, trans-Amazonian dispersal.

INTRODUCTION

The biogeography of Neotropical savanna birds is a mirrored image of that of forest birds. This is caused by the unique geographic configuration of Neotropical savannas: a large semi-continuous block of savannas in south-central South America (Pampas-Chaco-Bolivian savannas-Cerrado), and several smaller enclaves surrounded by forest in the northern half of the continent ('peri-Amazonian distribution'). The birds inhabiting Neotropical savannas are a phylogenetically diverse group, but well represented families include Tyrannidae and Thraupidae. Other species-rich Neotropical families more typical of forests like Thamnophilidae, Cotingidae and Pipridae form a smaller part of the Neotropical savanna species assemblage. Many Neotropical bird taxa adapted to open habitats occur in these habitats south of the Amazon Basin, as well as disjunctly in multiple or all of the savanna fragments north of the Amazon and within the Amazon, in spite of being separated by large distances (Mittermeier, Zyskowski, Stowe, & Lai, 2010; Stotz, Fitzpatrick, Parker, & Moskovits, 1996). Approximately 50 of these are specialized savanna taxa restricted to multiple fragments north and south of the Amazonian forests, whereas many more occur only in southern savannas. Many of the tropical savannas of northern South America receive similar amounts of precipitation as surrounding forest (Furley, 1999). Factors other than climate thus contribute greatly to vegetation structure. Generally, a combination of edaphic, pyric, and seasonal climatic factors is responsible for the existence of these enclaves, whereas the southern savannas likely have mainly climatic origins (Werneck, 2011). Although this may have had an effect on community composition, the relatively recent genesis of northern South American savannas compared to those south of the Amazon Basin (Huber, de Stefano, Aymard, & Riina, 2006) is likely responsible for the discrepancy in species richness between the two areas. Regardless, the broad distribution of savanna birds in pockets of grass-dominated habitat south and north of the forested Amazon Basin (and to a limited extent within) allows us to explore several hypotheses centering on diversification of the group, which may contrast strongly with that of the much better known forest birds.

The distributional limits of Neotropical forest birds often coincide with major rivers (Fernandes, Wink, & Aleixo, 2012; Hayes & Sewlal, 2004; Naka & Brumfield, 2018) and other open habitats, and taxon replacement in forest habitat often occurs over short distances. Rivers generally do not seem to shape the biogeography of Neotropical open-habitat birds and there is evidence that they are uniform phylogeographically over large areas of continuous habitat, even where rivers bisect their distributions

(Bates, Tello, & Silva, 2003; Norambuena, Van Els, Muñoz-Ramírez, & Victoriano, 2018), except for some species of white-sand forests, which are generally a mix of savannas and short-stature forest growing on poor soils in enclaves, particularly in the western Amazon (Capurucho et al., 2013; Borges et al., 2016; Capurucho et al., 2020). Some of the taxa show considerable morphological variation across savanna fragments, whereas others are rather uniform, even across distant fragments (Aleixo & Poletto, 2007; Mittermeier et al., 2010). There is little species-level endemism in these patches, although there are several endemics at the subspecies level (Haffer, 1969; Mittermeier et al., 2010).

The current high degree of geographical isolation between savanna enclaves coupled with relatively low levels of morphological diversification in savanna birds seems to contrast sharply with the replacement of morphologically and genetically distinct taxa across rivers in an otherwise continuously forested Amazon Basin (Haffer, 1969). From a vicariant perspective, this suggests that the savanna fragments (and the bird populations inhabiting them) were recently connected during drier climatic conditions, such as those associated with glacial periods (Capurucho et al., 2013; Matos et al., 2016; Ferreira et al., 2018). However, long-distance dispersal leading to colonization, with or without subsequent gene flow, may cause similar biogeographic patterns. Alternatively, savanna birds may have been isolated for a long time, with strong selective pressures on their phenotype maintaining low levels of morphological divergence.

Compared to the forests that enclose most of the distribution of the savanna biome in South America, grass-dominated communities on the continent are relatively young (Azevedo et al., 2020). Savannas have been continuously present in northern South America since the late Miocene (Cerling et al., 1997; Pennington & Hughes, 2014), but have until the late Quaternary (but see Jones, Mayle, Pennington, & Killeen, 2011) been more widespread than they are today. Several lines of evidence support this view, including palynology (Salgado-Labourieau, 1997; Van der Hammen, 1974; Mayle, Burbridge, & Killeen, 2000; D'Apolito, Absy, & Latrubesse, 2013), climatic models (Markgraf, 1993), and fossil deposits (Webb, 1978). Phytoliths indicate a more ancient (~Eocene) origin of southern South American C3-dominated grasslands (Sánchez, González, & Genise, 2010), but this contrasts with palynological data showing a more recent origin (Barreda & Palazzesi, 2007). Regardless, it is likely that from the late Miocene to Quaternary glaciations savannas were more widespread and connected than they are at present (Haffer, 1969; Cerling et al., 1997; Pennington & Hughes, 2014).

Avian diversification patterns across savanna fragments can inform us about more precise past (and current, through gene flow) patterns of connectivity between fragments. Three different historical corridors between savannas north and south of the Amazon have been proposed (Cardoso da Silva & Bates, 2002), one following the Andes, one across the middle of Amazonia through Monte Alegre and Sipaliwini, and a coastal corridor connecting the Cerrado with Amapá (including Ilha do Marajó) and Sipaliwini. Indeed, Mittermeier et al. (2010) found, based on morphological evidence, that birds occurring in Sipaliwini were more closely related to birds from the Cerrado and Amapá than to populations from the Roraima-Rupununi savannas, which, in turn, shared more affinities with the neighboring Llanos and Gran Sabana (Robbins, Braun, & Finch, 2004; Santos & Silva, 2007). Avian communities in the Monte Alegre savanna showed greater community overlap with the southern Cerrado than with savannas north of Amazonia (Vasconcelos, Dantas & Silva, 2011). Furthermore, limited evidence suggests that a coastal savanna corridor may have been formed recently in response to climatic change (Collevatti et al., 2014). This suggests a coastal route of dispersal ('peri-Atlantic' connection; Cardoso da Silva, 1995) between northern and southern savanna populations, but which needs to be verified using phylogeographic evidence.

Our focus in this study is to investigate temporal and spatial diversification patterns in Neotropical savanna birds. First, we test the hypothesis that a recent, late Quaternary increase in forest cover at the expense of savannas, has at best produced very shallow diversification patterns in savanna birds. Given these recent geological events and relatively low amounts of phenotypic divergence across enclaves of open habitat, we hypothesize that divergence of savanna bird populations is lower than that of forest birds across much greater gaps in suitable habitat. To test this, we estimate divergence times between populations from savannas on opposing sides of the forested Amazon Basin.

Further, previous findings indicate that there may be current connections between the southern Cerrado and the Amapá/Pará savannas into the Sipaliwini-Paru savannas, whereas the Llanos-Gran Sabana and Roraima-Rupununi complexes have remained relatively isolated from other northern savannas (Mittermeier et al., 2010). The current distribution of some savanna birds is reflected by this pattern; a relatively arid coastal connection between northeast Brazilian and Guianan populations still present. Other than this, there is little evidence for phenotypic congruence across other savanna fragments. We test the hypothesis that avian community turnover and phenotypic affinities across savanna enclaves reflect phylogeographic patterns in 23 widely distributed Neotropical savanna birds. We expect the relatively remote populations north and south of the Amazon Basin to show signals of

phylogeographic distinctness, which may be linked by populations from Amapá/Pará and the Sipaliwini-Paru savannas through a coastal corridor of gene flow. We explicitly test for shared phylogeographic history across our study taxa using Phylogeographic Concordant Factors (Satler & Carstens, 2016).

If savanna-inhabiting taxa have dissimilar phylogeographic histories, vicariance likely is not the sole variable explaining the distribution of lineages. Rather, dispersal or gene flow may have contributed as well (Smith et al., 2014). We test whether divergence across taxa between populations north and south of the Amazon occurred simultaneously, indicative of a dominant role for vicariance, or at different time periods, congruent with the idea that gene flow and dispersal have driven the diversification of savanna birds.

METHODS

Sampling

We sampled 24 taxa of co-distributed birds from 11 geographically isolated populations with broad ranges across savannas (here defined as grass-dominated ecosystems with variable levels of woody vegetation) of tropical South America. For each taxon, we obtained as many geographically disparate samples as possible, to maximize phylogeographic coverage across multiple isolated savanna enclaves, totaling 775 samples (Fig. 1, Table 1, Table S1). We also focused on sampling the same locations for as many taxa as possible, to enable comparisons across species. For all samples, we amplified one mitochondrial gene (NADH dehydrogenase subunit 2 – *ND2*). Because this gene is fast evolving, it is ideal for phylogeographic study at shallow and deep scales (Zink & Barrowclough, 2008). Although the use of single-locus data is potentially problematic, mainly because of the stochastic nature of gene coalescence (Edwards & Beerli, 2000), we chose to maximize taxonomic breadth and number of samples, rather than sampling multiple loci. For some samples, sequence data were obtained from previous studies through GenBank. Additionally, to allow for inference of better-precision estimates of divergence between populations north and south of the Amazon Basin, we obtained genomic data (Ultraconserved Elements, UCEs) for 24 co-distributed samples pertaining to a subset of 12 taxa (Table S2). Analysis of ND2 showed these samples to pertain to sister lineage pairs or to not have diverged significantly, so that the use of a single genomic sample per lineage for divergence time estimation does not result in problematic estimates (Blair, Bryson, Linkem, Lazcano, Klicka, & McCormack, 2018). We picked as many UCE samples

as possible from identical locations, but this was not always possible due to lack of sampling or failure to yield sufficient DNA for genomic analysis.

Extraction and amplification of NADH2

We extracted total genomic DNA from pectoral muscle using a Qiagen DNeasy tissue extraction kit (QIAGEN, Valencia, California) following manufacturer's protocol. We amplified one mitochondrial gene (NADH dehydrogenase subunit 2 – *ND2*) via polymerase chain reaction (PCR) in 12.5 lL reactions using the following protocol: denaturation at 94 ˚C for 10 min, 40 cycles of 94 ˚C for 30 s, 54 ˚C for 45 s, and 72 ˚C for 2 min, followed by 10 min elongation at 72 ˚C and 4 ˚C soak. We used the primers L5215 (Hackett 1996) and HTrpC (Smithsonian Tropical Research Institution, Balboa, Panama) for the ND2 gene. We used the program Geneious v. 11.1 (www.geneious.com) for alignment. We deposited sequences in GenBank (accession numbers listed in Table S1) and genomic data were deposited in GenBank's Sequence Read Archive under submission number SUB8691362.

Ultraconserved Elements Library Preparation

After quantifying DNA in the extracts with a Qubit 2.0 fluorometer, we cleaned 1,000 ng aliquots with 3× the volume of Sera-Mag Carboxylate-modified SpeedBeads (Rohland & Reich, 2012) and eluted DNA into 30 µl of TE buffer. We mechanically sheared DNA in 2.5-min increments at 17 mA with an Epigentek Episonic sonicator until the average fragment size was ~500 bp, as assessed by eye on an electrophoretic gel.

We prepared DNA libraries using the Kapa Biosystems Hyper Prep Kit for Illumina platforms with dual indexed iTru adapters (Glenn et al., 2016). We used one-fourth of the manufacturer's recommended reagent volume, performing a 1× postligation bead cleanup, and increased the library amplification extension time to 1 min. We combined the resulting DNA libraries in equimolar pools of eight samples and enriched each pool for 5,060 UCE loci using the Tetrapods-UCE-5Kv1 probe set (Faircloth et al., 2012) sold by MYcroarray (Ann Arbor, MI). We followed the manufacturer's instructions for enrichment. We determined the size distribution of enriched libraries with an Agilent Bioanalyzer and removed remaining adapter dimer from pools where present using a 0.8× bead cleanup. We then quantified the enriched libraries with a Qubit 2.0 fluorometer and pooled them in equimolar ratios.

We sent all samples to Rapid Genomics (Gainesville, FL) for sequence capture and sequencing following the general protocol described in Faircloth et al. (2012) and Smith et al. (2014). Samples were

multiplexed at 160 samples per lane on a 100-bp paired-end Illumina HiSeq 2500 run. Rapid Genomics demultiplexed raw reads using custom scripts and strict barcode matching.

Ultraconserved Elements assembly

After sequencing, we checked sequence quality with FASTQC (Andrews, 2012) and then followed the standard PHYLUCE v.1.5.0 pipeline for processing target-enriched UCE data (Faircloth, 2016). Reads were trimmed of adapter contamination and low-quality bases with Trimmomatic (Bolger, Lohse, & Usadel, 2014) implemented in Illumiprocessor (Faircloth, 2013). Cleaned reads were assembled into contigs using Trinity (Grabherr et al., 2011), and the assembled contigs were aligned to the original UCE probe sequences. We created a taxon set containing all samples, and used this taxon set to make a FASTA file containing all data for all taxa. We used MAFFT (Katoh & Standley, 2013) to then create alignments across all loci. After examining the number of UCE loci captured per sample, we created a final concatenated data matrix with up to 75% missing data per locus (meaning each locus contained data for 75% or more of the total number of individuals). There are 2284 UCE alignments containing data for 75% or more of the total number of individuals, resulting in a matrix of 1,268,196 bp. The matrix was analyzed with RAxML v8 (Stamatakis, 2014), with the model GTRCAT and 500 bootstrap replicates.

Estimation of demographic parameters and Ultraconserved Elements

We used the Generalized Phylogenetic Coalescent Sampler (G-PHOCS; Gronau, Hubisz, Gulko, Danko, & Siepel, 2011), a full-likelihood method based on the multispecies coalescent model, to estimate mutation-scaled effective population sizes (N_e), divergence time (T), and migration rates (M_{sx}) across all twelve population/taxon pairs from a set of UCE loci. For each pair, demographic parameters were estimated under two models: pure isolation and isolation-with-migration. Both models assume that an ancestral population splits into two daughter populations, each with different θ parameters ($\theta = 4N_e\mu$) at time T ($T = \tau/\mu$) generations ago. The isolation-with-migration model assumes two additional migration rate per generation parameter M_{sx} ($M_{sx} = m_{sx} \times \theta_x/4$), indicating gene flow in both directions. For the prior distribution of θ and τ, we assumed a gamma distribution with $\alpha = 1$ and $\beta = 300$, and a gamma distribution with $\alpha = 0.002$ and $\beta = 0.00001$ for the prior distribution of m_{sx}. We ran G-PHOCS with default settings for 500,000 MCMC iterations, sampling every 100 generations. We then used TRACER v1.6 (Rambaut, Suchard, Xie, & Drummond, 2014) to evaluate whether runs showed adequate convergence and mixing. Results were converted into biologically informative values assuming an average mutation rate of 4.6×10^{-9}

mutations per site per generation (Smeds, Qvarnström, & Ellegren, 2016) and a generation time of one year (Mooers & Harvey, 1994).

Phylogeographic structure and phylogeographic concordance factors

To infer phylogeographic structure, we constructed haplotype networks per taxon using the R package 'pegas' (Paradis, 2010). We calculated phylogeographic concordance factors (PCF) according to Satler & Carstens (2016) to estimate the level of joint phylogeographic history across species in different isolated grasslands across South America. We first estimated the posterior distributions of species trees in *BEAST v. 1.8.4 (Heled & Drummond, 2009) for each species, including representative samples from geographic areas of interest (Table 2) and then used the consensus tree (based on a tree distribution of 10,000 samples) for PCF analysis. We used mitochondrial data for tree reconstruction, see Satler & Carstens (2016) for rationale. We employed the Bayesian relaxed phylogenetic method used in Naka & Brumfield (2018) and used an uncorrelated relaxed substitution rate (where the rate at each branch is drawn independently from a lognormal distribution) based on an avian molecular clock estimated for the mitochondrial marker used (ND2) by Smith & Klicka (2010). The selected prior distribution for the mean (ucld.mean) was set to 0.0125, with an SD (ucld.stdev) following an exponential distribution with a value of 0.1 and a mean SD of 1. Because we were working at the population level (intraspecific level), we used a population constant size tree prior.

We determined the best-fit finite-sites nucleotide substitution model for each pair of taxa under the Bayesian information criterion, as implemented in the function 'modelTest' of the R package 'phangorn' (Schliep, 2010). In BEAST, we ran analyses for at least 50 million generations, more in some cases with low effective sample size (ESS) values, sampling every 1000 generations. We verified Markov chain Monte Carlo convergence, ESS, and posterior intervals spanning the 95% highest posterior density using TRACER v1.7.1 (Rambaut et al., 2014).

Nodal support values of the resulting consensus tree are PCFs, with the topology representing an estimate of joint diversification. We imported log files into TRACER v1.7.1 (Rambaut et al., 2014) to check for convergence. We used nodes separating samples from north and south of the Amazon Basin and their associated divergence times including 95% highest posterior density intervals for divergence time estimation.

We then summarized each species tree distribution using mbsum (Larget, Kotha, Dewey, & Ané, 2010) to count the number of times each unique topology was represented within the posterior distribution. BUCKy (Ané, Larget, Baum, Smith, & Rokas, 2007; Larget et al., 2010) was then used to process the resulting tree summary files, generating a consensus tree with concordance factors. We used a Python pipeline available on https://github.com/jordansatler/PhylogeographicConcordanceFactors for the last two steps. An alpha value of infinity was used, which represents a null hypothesis of no congruence. This approach is conservative, and requires strong support across two or more species to infer a shared phylogeographic history.

To establish if particular subsets of species showed stronger phylogeographic concordance than the full data set, we used the pipeline to iteratively remove taxa from the dataset. To compare the models, we calculated mean nodal support values and ranked each model associated with K taxa according to these values, expecting a strong increase in average PCF values if taxa with discordant phylogeographic histories are removed from the analysis. Inspecting PCF values associated with each combination of K taxa allows us to establish which taxon does not follow phylogeographic congruent patterns, if present.

Because not all taxa have representatives in all major grassland areas or because we lacked some samples, and the PCF analysis requires each taxon to be represented in all geographic areas, we split the analysis in three different combinations ('sets') of co-distributed taxa and populations (Table 2). Sets were chosen so that the number of taxa and areas were maximized.

Finally, we employed the R package 'adegenet' (Jombart & Ahmed, 2011) to calculate the influence of isolation-by-distance (IBD) for 19 taxa with sufficient geographic sampling in the data set. When converting sequence data to a *genind* object, in which only polymorphic loci are stored, we set the *polyThres* argument to 0.01. 'Adegenet' performs a Mantel test to find correlation between two matrices: one containing genetic and another containing Euclidian geographic distances. To establish significance, it permutes distance values and tests the true test value against these.

RESULTS

Timing of divergence and gene flow

Estimates of divergence between populations north and south of the Amazon River (Table S3, Fig. 3) based on mtDNA were mostly Pleistocene in origin (20/23 taxa) and 17/23 taxa did not diverge at all or diverged within the last 1 mya. Mean divergence time across 23 taxa was 0.82 mya. Further, according to the pure isolation model in G-PHOCS divergence estimates were also recent in our subset of 12 genomically sampled taxa, the oldest splits being dated ~0.12 mya in Wedge-tailed Grass-finch (*Emberizoides herbicola* Vieillot) and ~0.05 mya in Masked Yellowthroat (*Geothlypis aequinoctialis* Gmelin). Mean divergence time for all taxa according to the pure isolation model was 0.03 mya.

Accounting for gene flow (Table S4) using the isolation-with-migration model, we found that all taxa still diverged during the Pleistocene, but that there was slightly more disparity in timing of divergence: half of the focal taxa diverged during the last 0.25 mya, whereas others ranged from ~0.31 mya (Fork-tailed Flycatcher *Tyrannus savana*, Daudin) to 1.25 mya in Yellowish Pipit (*Anthus lutescens*). Mean divergence time including gene flow was 0.40 mya.

Phylogeographic patterns

Of the 23 study species, at least 12 show (Fig. 2) separate clades north and south of the Amazon Basin. Northern clades include samples from Central America, the coastal savannas of northern South America, the Llanos, the Guianan Roraima-Rupununi savanna complex, and the Sipaliwini-Paru complex. Southern clades include samples from the Pampas-Chaco complex, the Beni-Bolivian savannas, and the Cerrado. Within northern clades, Sipaliwini-Paru samples shared haplotypes with Roraima-Rupununi samples (five instances), although they also shared with Pará samples (two instances) and once with the Cerrado. Samples from Amapá, Pará, and western Amazonian savanna pockets grouped with either northern or southern clades.

Only in *Sublegatus arenarum/S. modestus* were coastal savanna and Central American populations (considered *S. arenarum*) represented by a separate clade, with other northern populations grouping with southern samples. In all other cases these were part of northern clades or part of a panmictic group. Cerrado and Pampas-Chaco samples shared haplotypes (three instances), as well as Pampas-Chaco samples and Beni-Bolivian samples (three instances).

Seven of the taxa (*Anthus lutescens,* Least Nighthawk *Chordeiles pusillus,* Gould*,* Plain-breasted Ground-Dove *Columbina minuta,* L., Lesser Elaenia *Elaenia chiriquensis,* Lawrence, Plumbeous Seedeater *Sporophila plumbea,* Wied, Red-breasted Blackbird *Leistes militaris,* L.*,* White-browed Blackbird *L. superciliaris,* Bonaparte*,* and Black-faced Tanager *Schistochlamys melanopis,* Latham) show variable patterns of geographic overlap between southern and northern clades, indicating that gene flow may at least potentially take place. Most taxa had dominant haplotypes represented in multiple geographic areas.

Phylogeographic concordance

Satler & Carstens (2016) used simulations to show that PCF values > 0.62 (based on 95% Highest Posterior Density) were considered significantly concordant phylogeographically. In our analysis (Table 3), a maximum of 3 taxa (out of 8) in set 1 (see Table 2) attained significant concordance (*Tyrannus savana, Mimus,* and Campo Flicker *Colaptes campestris*, Vieillot PCF=0.6750), 2 taxa (out of 5) in set 2 (*Anthus lutescens* and Grassland Sparrow *Ammodramus humeralis,* Bosc, PCF=0.7255) and 3 (out of 5) in set 3 (*Emberizoides herbicola, Leistes* spp*., and Elaenia chiriquensis,* PCF=0.6373). This indicates we are unable to reject the null hypothesis of lack of congruence, regardless of geographic and taxon sampling. It is striking that for each of the three sets, a different taxon composition is responsible for maximum concordance. This reveals that adding or removing geographic areas or species does not result in more stable patterns of concordance. One commonality across many taxa is the basal split between areas north and south of the Amazon River (Fig. 4).

Significant ($p<0.05$) isolation-by-distance patterns were only detected in one taxon: *Columbina minuta* ($p=0.03$) (Supplement S5). Inspection of kernel density plots of IBD revealed that other taxa showed no or negative correlation between genetic and geographic distance, indicating that diversification, even over large distances, is minimal, or that there may be less gene flow between nearby patches of savanna than between more distant patches. Additionally, a few taxa do not show signs of clinal variation as expected under an IBD model. These are generally taxa that consist of multiple distinct mtDNA lineages (e.g. *Emberizoides herbicola, Geothlypis aequinoctialis, Sublegatus* spp.).

DISCUSSION

Quaternary population dynamics of savanna birds

Divergence estimates indicated late Quaternary sundering of populations north and south of the Amazon Basin of most taxa examined, using both mitochondrial and genomic data and under scenarios with and without gene flow. A few divergence estimates based on mitochondrial data were Miocene in origin. These are likely overestimates however, because UCE samples for some of these taxa show divergences to be much more recent, which can be explained by the mitochondrial data being a smaller subset of the genome than the UCE data (Edwards & Beerli, 2000). Recent divergences contrast with Neotropical forest birds, which in many cases show higher (early Pleistocene-Miocene) levels of divergence (Capparella, 1987; Smith et al., 2014; Naka & Brumfield, 2018; Silva et al., 2019). The relatively shallow divergence in many savanna birds can be explained by a recent separation of populations (if at all separated) after a long period of connectedness and gene flow. Haffer's (1969) paradigm of Pleistocene Amazonian refugia leading to diversification of Neotropical forest birds may have partially lost ground to alternative hypotheses that explain diversification patterns equally well or better (Colinvaux, Irion, Räsänen, Bush, & De Mello, 2001, but see Silva et al., 2019). Still, a relatively cool and arid Pleistocene promoting (the continued) connection of savanna bird populations, followed by warmer, wetter conditions more recently, may have led to very recent population structuring in some savanna birds. Little if any divergence is present in multiple of our focal taxa, supporting this scenario and indicating that persistent gene flow throughout most of the history of these taxa through high dispersal capability or geographic continuity of populations is a likely explanation.

Vicariance versus ecology as a driver of diversification

Although divergences are mostly late Pleistocene in origin, a temporally convergent signature of divergence across Neotropical savanna birds is absent. This pattern is similar to, but perhaps not as marked as, that found in forest birds distributed on opposite sides of a vicariant barrier (Smith et al., 2014). This likely means that the lineage histories of the studied taxa were not affected solely by vicariance. That for some taxa UCE divergence estimates differ between models with and without gene flow, and that some taxa show no noticeable differentiation north and south of the Amazon Basin at all, indicates that gene flow almost certainly plays an important role in the genetic shaping of sampled populations.

Small populations of savanna birds, including many of the taxa sampled in this study, occur in isolated Amazonian open vegetation enclaves, such as the BX-044 polygon in the Brazilian states of Rondônia

and Amazonas (Aleixo & Poletto, 2007), the campinas of the lower Rio Tocantins (Lees, Moura, Almeida, & Vieira, 2014), and the Alter do Chão savannas (Sanaiotti & Cintra, 2001). Although these enclaves may have been once part of a more extensive savanna landscape, several savanna species appear to quickly colonize recently deforested regions of the Amazon region, including the interior of the Basin far away from other open habitats (Sick, 1997; Vasconcelos, Pacheco & Parrini, 2007; Guilherme & Czaban, 2015; Guimarães, Gomes de Lima, & Pedroza, 2016; Borges et al., 2017; Rutt, Jirinec, Cohn-Haft, Laurance, & Stouffer, 2019).

Savannas are highly seasonal, open environments, factors that are associated with a high hand-wing index in birds (Sheard et al., 2020), indicative of greater dispersal capability. Further, savannas are generally dynamic over short time spans due to the important role of fire (Azevedo et al., 2020), adding to the need for dispersal. Although we acknowledge the taxa selected for this study have broad distributions, another factor correlated with high hand-wing index (Sheard et al., 2020), the ecology of savanna birds undoubtedly contributed greatly to the shaping of their phylogeography and contrasts with the more sedentary and stable lifestyles of forest-based taxa. Many forest birds are represented by multiple clades across much shorter distances, e.g. on opposite sides of Amazonian rivers (Capparella, 1987; Smith et al., 2014; Naka & Brumfield, 2018).

The Amazon as a semi-permeable barrier

Principal Concordance Factors revealed little phylogeographic congruence across most of the study taxa. However, it should be noted that PCF values are averaged across trees, so that a high discrepancy in PCF values across nodes may obscure well-supported, shared phylogeographic breaks, particularly where these happen at the base of a tree and are drowned out by lower support values of more numerous recent splits (Satler & Carstens, 2016). The basal split in our trees in all cases had high support, indicating that a phylogeographic break between areas north (Sipaliwini, Guyana, Llanos, Panama) and south (Cerrado, Pampas, Bolivian savannas) of the Amazon Basin was shared by many taxa. The Amazon Basin, largely dominated by humid climates and associated forests and mostly lacking open habitats, appears thus to act as a barrier at least to some extent between populations north and south of it. Samples from Amapá grouped with either one of these two groups, indicating that a coastal corridor of dispersal between northern and southern populations may have shaped the lineage history of at least some of the studied taxa. However, lack of congruence in phylogeographic patterns across multiple taxa refutes the idea that a single Pleistocene corridor of dispersal produced a community-wide shared

phylogeographic pattern. Rather, the combination of gene flow and a lack of concordance in phylogeography indicate that there have been multiple paths of colonization across the Amazon Basin, be these true corridors (Cardoso da Silva, 1995) or part of long-distance dispersal events. Additional sampling should be done to more robustly assess these patterns across more taxa.

Phylogeographic patterns in savannas versus other habitats

Few taxa have diversified to the point of having geographically non-overlapping divergent lineages within the vast area of open habitat south of the Amazon Basin. Exceptions to this are *Geothlypis aequinoctialis*, *Emberizoides herbicola* and White-fringed Antwren *Formicivora grisea,* Boddaert, which are represented by multiple clades south of the Amazon. Many taxa even share haplotypes across geographically isolated savanna fragments. This agrees with evidence by Cardoso da Silva & Bates (2002) and Bates et al. (2003) that over large distances of (semi-)continuous habitat, savanna birds show little differentiation, even when savannas are bisected by rivers. The rattlesnake *Crotalus durissus* showed a similar pattern of divergence between areas north and south of the Amazon Basin, but little structure within the cerrado (Wüster et al. 2005). However, this pattern does not extend to all widespread inhabitants of Neotropical savannas. Plants (Buzatti, Lemos-Filho, Bueno, & Lovato, 2017; Novaes, Ribeiro, Lemos-Filho, & Lovato, 2013) and squamates (Guarnizo et al., 2016; Santos, Nogueira, Giugliano, & Colli, 2014; Werneck, 2011) can show extensive phylogeographic variation within this biome.

Contact between northern and southern clades

Multiple taxa (*e.g. Chordeiles pusillus, Columbina minuta, Leistes militaris/L. superciliaris*) show divergence between some samples from opposite sides of the Amazon Basin, yet simultaneously have haplotypes with representatives from north and south. This is likely in part caused by (de Paiva & Marini, 2013; Gómez-Bahamón et al., 2020; Marini & Cavalcanti, 1990; Tuero, Jahn & MacPherson, 2019) partly undescribed seasonal migratory patterns, which are highly species- and location-specific, and may depend mostly on food availability in savanna birds (Hockey, 2000; Sanaiotti & Cintra, 2001). Alternatively, they may be caused by the recent incursion of 'alien' haplotypes through deforestation, which may be especially true in some highly dispersive species (Guilherme & Czaban, 2015; Guimarães

et al., 2016). If this holds true, deforestation not only leads to disappearance of forest diversity (Ferraz et al., 2007), but may also lead to secondary contact, gene flow, and possible genetic homogenization of previously isolated savanna bird populations (Rutt et al., 2019).

Table 1. Sample taxa of Neotropical savanna birds with total number of individuals sampled, number of mitochondrial (NADH2) samples, and number of genomic (UCE) samples.

Family	Taxon name	MtDNA (n)	Genomic (n)
Caprimulgidae	*Chordeiles pusillus*	37	0
Trochilidae	*Polytmus guainumbi*	14	0
Columbidae	*Columbina minuta*	20	0
Charadriidae	*Vanellus chilensis*	12	2
Picidae	*Colaptes campestris*	14	0
Furnariidae	*Lepidocolaptes angustirostris*	22	2
Thamnophilidae	*Formicivora grisea*	103	0

	Formicivora rufa	9	0
Tyrannidae	*Elaenia chiriquensis*	17	2
	Sublegatus	24	2
	Tyrannus albogularis	20	0
	Tyrannus savana	58	2
Pipridae	*Neopelma pallescens*	28	0
Mimidae	*Mimus saturninus*	14	2
Motacillidae	*Anthus lutescens*	61	2
Passerellidae	*Ammodramus humeralis*	75	0
Icteridae	*Leistes*	38	2
Parulidae	*Geothlypis aequinoctialis*	18	2
Thraupidae	*Coryphospingus pileatus*	6	0
	Emberizoides herbicola	57	2
	Schistochlamys melanopis	51	0
	Sicalis luteola	6	2
	Sporophila plumbea	53	0
	Stilpnia cayana	18	2

Table 2. Samples used for each combination of taxa and populations in the PCF analysis of Neotropical savanna birds. BOL=Bolivia, PAM=Pampas, CER=Cerrado, AMA=Amapá, SIP=Sipaliwini, GUY=Guyana, LLA=Llanos, PAN=Panama.

Taxon	BOL	PAM	CER	AMA	SIP	GUY	LLA	PAN	Set(s)
Colaptes campestris	x	x	x	x	x				1
Lepidocolaptes angustirostris	x	x	x	x	x				1
Mimus saturninus	x	x	x	x	x				1
Ammodramus humeralis	x	x	x	x	x	x			1,2
Anthus lutescens	x	x	x	x	x	x			1,2
Sublegatus spp.	x	x	x	x	x	x			1,2

Emberizoides herbicola	x	x	x	x	x	x	x	x	1,2,3
Tyrannus savana	x	x	x	x	x	x	x	x	1,2,3
Columbina minuta	x			x		x	x	x	3
Elaenia chiriquensis	x			x		x	x	x	3
Leistes spp.	x			x		x	x	x	3

Table 3. Best scoring phylogeographic concordance factors for all permutations of K levels from two to N total species, for the three different geographic sets of Neotropical savanna birds. For rationale on defining sets, see methods. (see Table 2).

Set	*K*	$PCF_{average}$	Species composition
1	2	0.8133	*Mimus, Colaptes*
1	3	0.6750	*Tyrannus, Mimus, Colaptes*
1	4	0.5990	*Mimus, Ammodramus, Colaptes, Lepidocolaptes*
1	5	0.5577	*Tyrannus, Mimus, Ammodramus, Colaptes, Lepidocolaptes*
1	6	0.5197	*Tyrannus, Mimus, Ammodramus, Colaptes, Lepidocolaptes, Emberizoides*
1	7	0.4473	*Tyrannus, Mimus, Ammodramus, Colaptes, Anthus, Lepidocolaptes, Emberizoides*
1	8	0.3913	*Tyrannus, Mimus, Ammodramus, Colaptes, Anthus, Sublegatus, Lepidocolaptes, Emberizoides*
2	2	0.7255	*Anthus, Ammodramus*
2	3	0.5822	*Anthus, Ammodramus, Tyrannus*
2	4	0.4990	*Emberizoides, Anthus, Ammodramus, Tyrannus*
2	5	0.3997	*Emberizoides, Sublegatus, Anthus, Ammodramus, Tyrannus*
3	2	0.8297	*Emberizoides, Leistes*
3	3	0.6373	*Emberizoides, Leistes, Elaenia*
3	4	0.4913	*Columbina, Emberizoides, Leistes, Elaenia*
3	5	0.3990	*Columbina, Tyrannus, Emberizoides, Leistes, Elaenia*

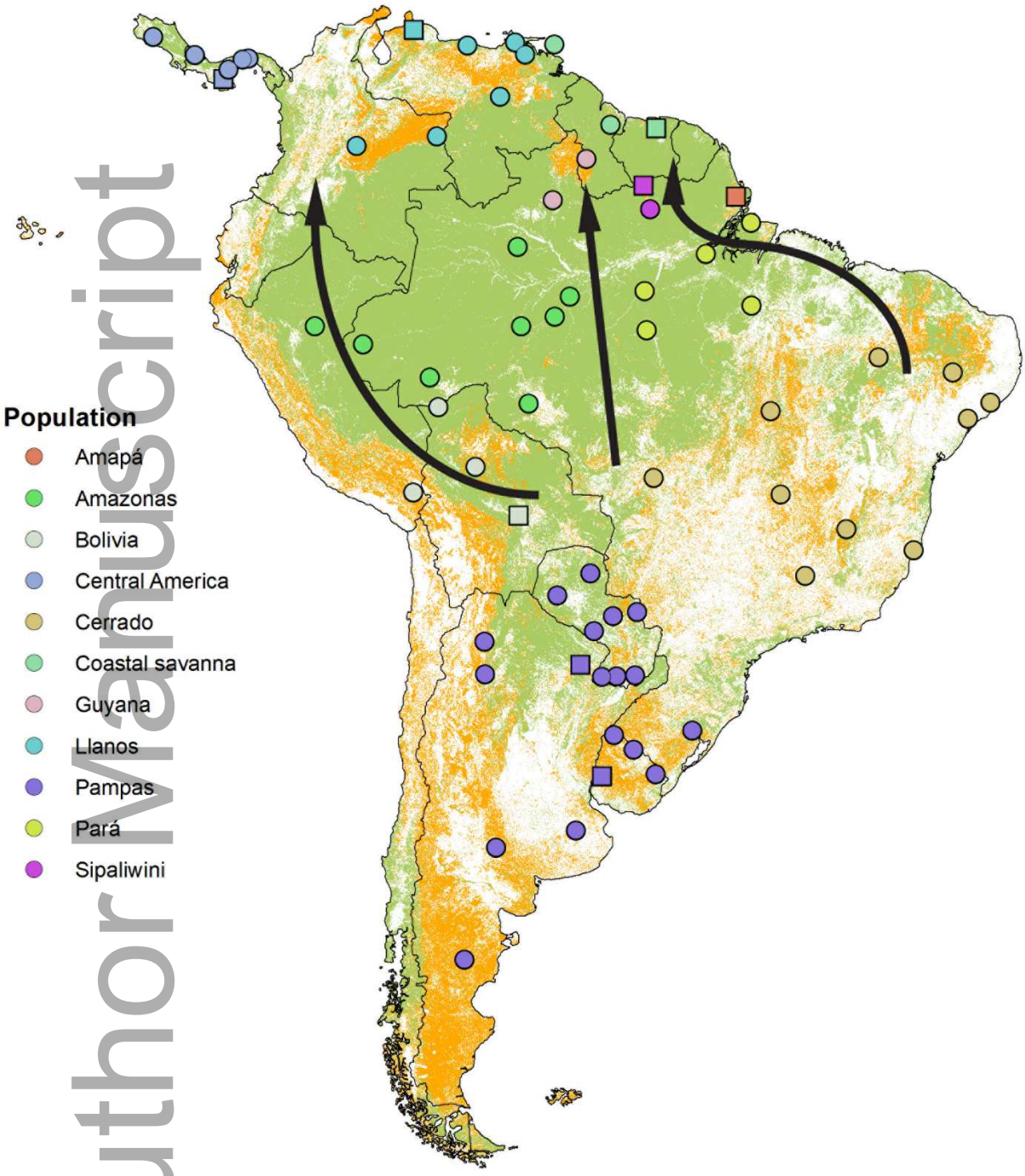


Fig 1. Map of sampling localities of Neotropica savanna birds and corresponding population assignments. Areas in green indicate forests, areas in yellow indicate grass-dominated landscapes such as cerrado, pampa and puna, white areas are other habitats. Squares represent UCE samples (overlapping with mDNA samples), circles are mtDNA samples only. The three arrows indicate putative Pleistocene savanna corridors.

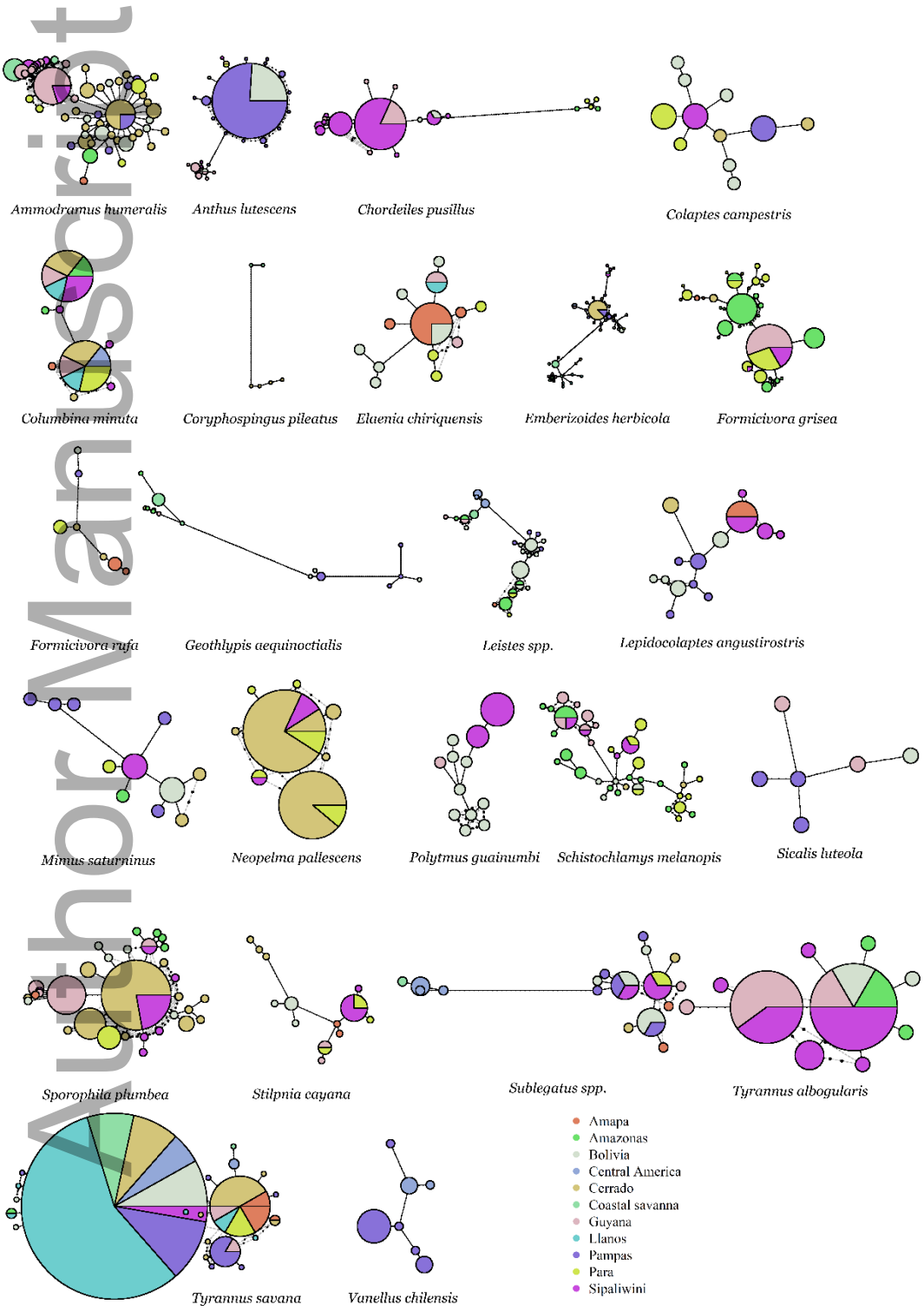


Fig. 2. Haplotype networks of mitochondrial samples of 23 taxa of Neotropical savanna birds. Colors indicate geographic origins of samples, nodes on lines indicate one mutation difference. Size of circles is representative of number of samples.

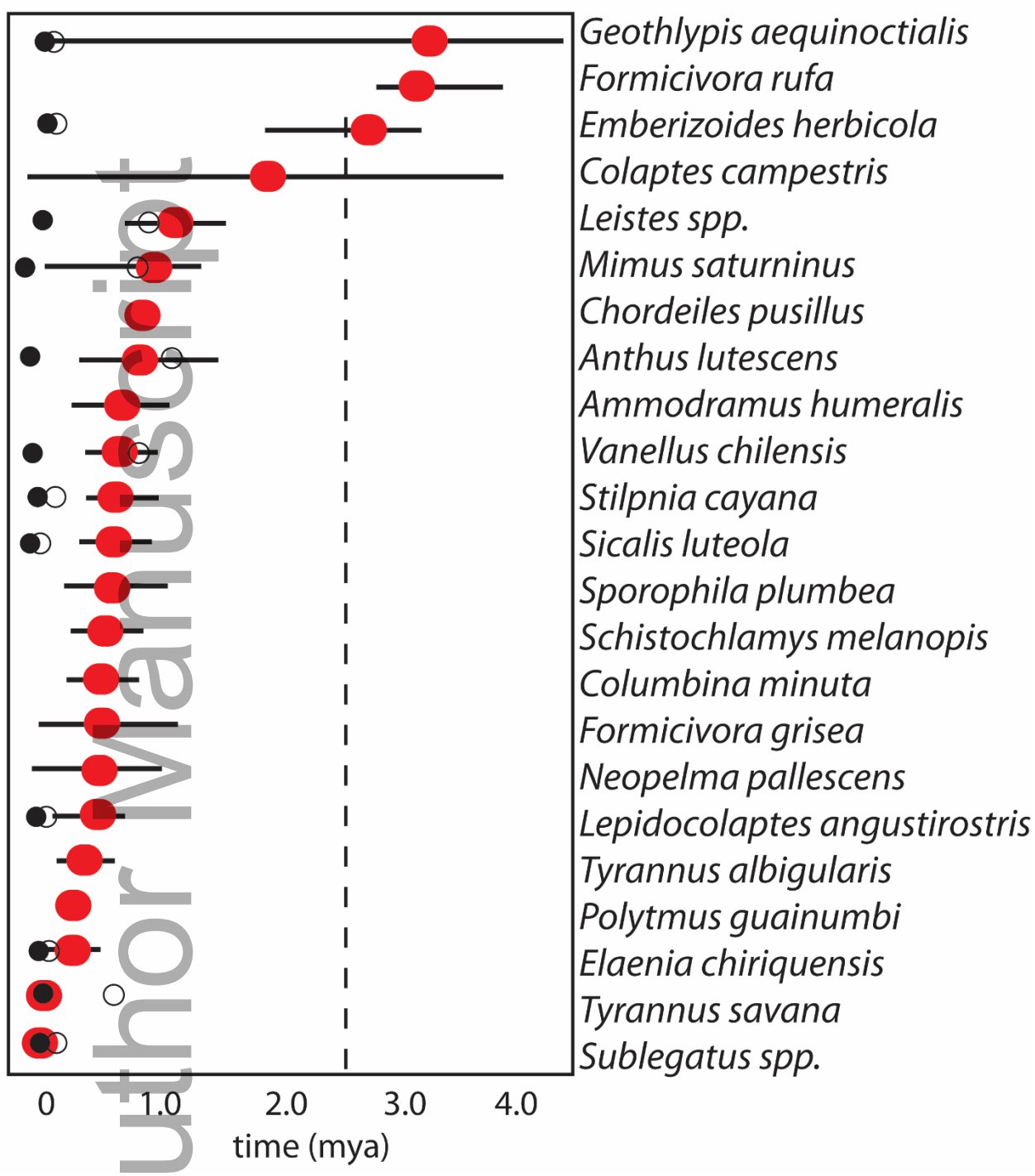


Fig. 3. Divergence times (mya) per taxon between populations of Neotropical savanna birds distributed north and south of the forested Amazon Basin, estimated by Beast based on ND2 (red dots), and estimated by G-PhoCS based on Ultraconserved Elements (UCEs) using a pure isolation model (small black dots) and an isolation-with-migration model (white dots). Taxon order is based on divergence time.

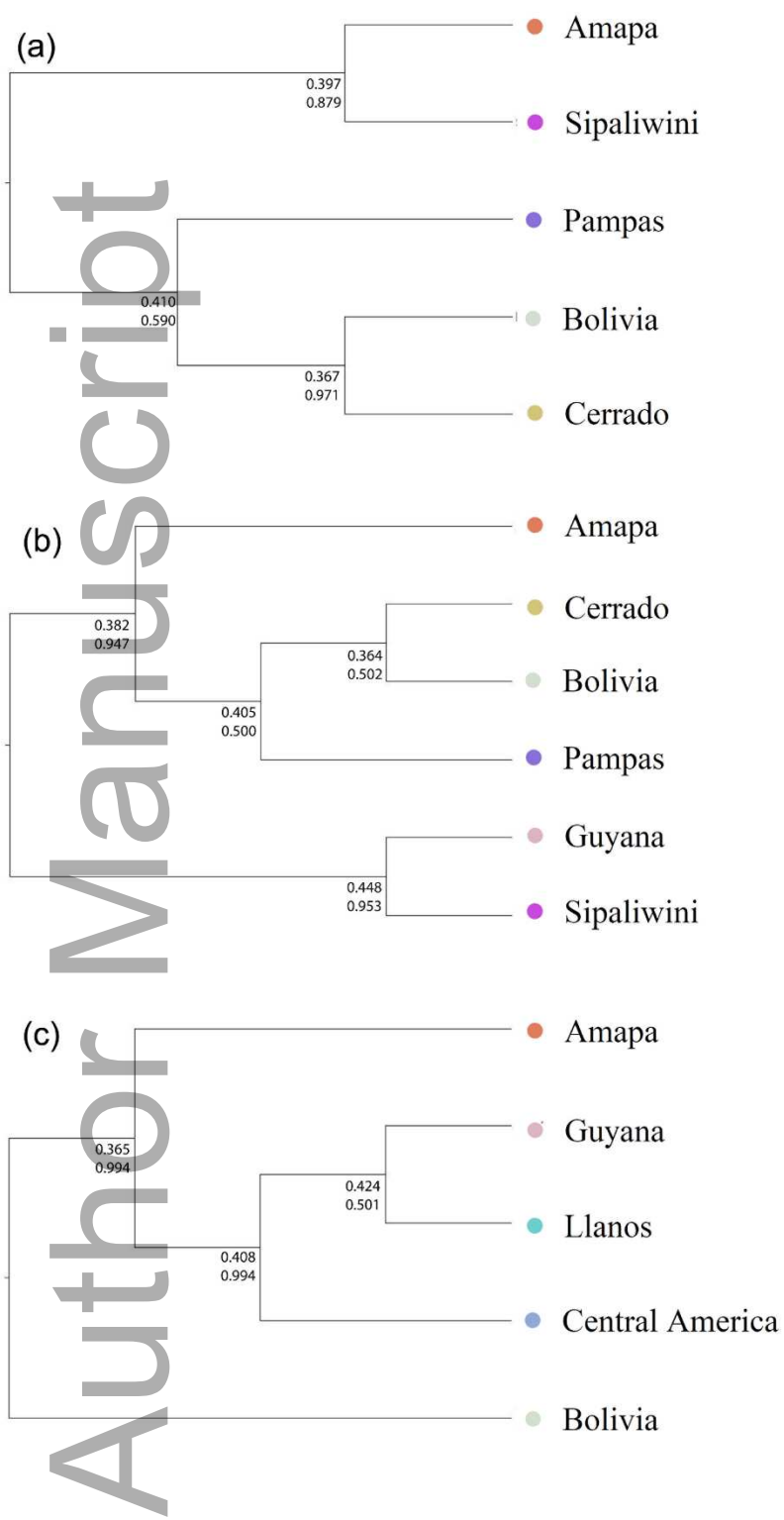


Fig. 4. Joint diversification patterns of Neotropical savanna birds estimated using BUCKy; (a) set 1 (K_{max}=8), (b) set 2 (K_{max}=5), and (c) set 3 (K_{max}=5). Concordance factors on top are when all species are

included, those at bottom represent models where *K*=2. Tip abbreviations: BOL=Bolivia, PAM=Pampas, CER=Cerrado, AMA=Amapá, SIP=Sipaliwini, GUY=Guyana, LLA=Llanos, PAN=Panama.

DATA AVAILABILITY STATEMENT

Sanger sequence data are deposited in GenBank (Table S1) and UCE data were deposited in GenBank's Sequence Read Archive under submission number SUB8691362.

SIGNIFICANCE STATEMENT

This article highlights, for the first time, phylogeographic patterns of Neotropical savanna birds in a comparative framework. We find that populations north and south of the Amazon forest often show some degree of isolation, but no obvious common patterns in each of these two areas. Estimated divergence times between populations north and south of the Amazon are recent, mainly dating back to the Pleistocene, indicating that there likely was connectivity between savannas north and south of the Amazon Basin during that epoch. We did not find evidence of a single dispersal corridor between the two areas.

REFERENCES

Andrews, S. (2012). *FastQC: A Quality Control tool for High Throughput Sequence Data*. Babraham Bioinformatics.

Aleixo, A., & Poletto, F. (2007). Birds of an open vegetation enclave in southern Brazilian Amazonia. *The Wilson Journal of Ornithology*, 119(4), 610–631.

Ané, C., Larget, B., Baum, D. A., Smith, S. D., & Rokas, A. (2007). Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution*, 24(2), 412–426.

Azevedo, J. A., Collevatti, R. G., Jaramillo, C. A., Strömberg, C. A., Guedes, T. B., Matos-Maraví, P., ... & Antonelli, A. (2020). On the young savannas in the land of ancient forests. In Rull, V., & A. C. Carnaval (Eds.), *Neotropical Diversification: Patterns and Processes* (pp. 271-298). Cham, Springer.

Barreda, V., & Palazzesi, L. (2007). Patagonian vegetation turnovers during the Paleogene-early Neogene: origin of arid-adapted floras. *The Botanical Review*, 73(1), 31–50.

Bates, J. M., Tello, J. G., & Silva, J. M. C. (2003). Initial assessment of genetic diversity in ten bird species of South American Cerrado. *Studies on Neotropical Fauna and Environment*, 38(2), 87–94.

Blair, C., Bryson Jr, R. W., Linkem, C. W., Lazcano, D., Klicka, J., & McCormack, J. E. (2019). Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: Crotalus). *Molecular ecology resources*, *19*(2), 349–365.

Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30(15), 2114–2120.

Borges, S. H., Cornelius, C., Moreira, M., Ribas, C. C., Conh-Haft, M., Capurucho, J. M., ... & Almeida, R. (2016). Bird communities in Amazonian white-sand vegetation patches: effects of landscape configuration and biogeographic context. *Biotropica*, 48(1), 121–131.

Borges, S. H., Whittaker, A., Almeida, R. A., Cornélius, C., Santos-Jr, M. A., & Moreira, M. (2017). Bird records in the northwestern and central portions of the Amazon Basin highlight the needs for inventories and long-term monitoring in the region. *Revista Brasileira de Ornitologia*, 25(3), 206–220.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, 10(4), e1003537.

Buzatti, R. S. D. O., Lemos-Filho, J. P., Bueno, M. L., & Lovato, M. B. (2017). Multiple Pleistocene refugia in the Brazilian cerrado: evidence from phylogeography and climatic niche modelling of two *Qualea* species (Vochysiaceae). *Botanical Journal of the Linnean Society*, 185(3), 307–320.

Capparella, A. P. (1987). Effects of riverine barriers on genetic differentiation of Amazonian forest undergrowth birds. Doctoral dissertation, Louisiana State University, Baton Rouge.

Capurucho, J. M. G., Cornelius, C., Borges, S. H., Cohn-Haft, M., Aleixo, A., Metzger, J. P., & Ribas, C. C. (2013). Combining phylogeography and landscape genetics of *Xenopipo atronitens* (Aves: Pipridae), a white sand campina specialist, to understand Pleistocene landscape evolution in Amazonia. *Biological Journal of the Linnean Society*, 110(1), 60–76.

Capurucho, J. M. G., Borges, S. H., Cornelius, C., Vicentini, A., Prata, E. M. B., Costa, F. M., … Ribas, C. C. (2020). Patterns and process of diversification in Amazonian white sand ecosystems: insights from birds and plants. In V. Rull & A. C. Carnaval (Eds.), Neotropical diversification: patterns and processes. Cham: Springer.

Cardoso da Silva, J. M. C. (1995). Biogeographic analysis of the South American Cerrado avifauna. *Steenstrupia*, *21*, 49–67.

Cardoso da Silva, J. M.C., & Bates, J. M. (2002). Biogeographic patterns and conservation in the South American Cerrado: A tropical savanna hotspot: The Cerrado, which includes both forest and savanna habitats, is the second largest South American biome, and among the most threatened on the continent. *BioScience*, *52*(3), 225–234.

Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V., & Ehleringer, J. R. (1997). Global vegetation change through the Miocene/Pliocene boundary. *Nature*, *389*(6647), 153.

Colinvaux, P. A., Irion, G., Räsänen, M. E., Bush, M. B., & De Mello, J. N. (2001). A paradigm to be discarded: geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana*, *16*(3), 609–646.

Collevatti, R. G., Lima-Ribeiro, M. S., Terribile, L. C., Guedes, L. B., Rosa, F. F., & Telles, M. P. (2014). Recovering species demographic history from multi-model inference: the case of a Neotropical savanna tree species. *BMC evolutionary biology*, *14*(1), 213.

D'Apolito, C., Absy, M.L., & Latrubesse, E.M. (2013) The Hill of Six Lakes revisited: new data and re-evaluation. *Quaternary Science Reviews*, 76, 140–155.

Edwards, S. V., & Beerli, P. (2000). Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54(6), 1839–1854

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic biology*, 61(5), 717–726.

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

Faircloth, B.C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32(5), 786–788.

Fernandes, A. M., Wink, M., & Aleixo, A. (2012). Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. *Journal of Biogeography*, 39(8), 1524–1535.

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.

Ferraz, G., Nichols, J. D., Hines, J. E., Stouffer, P. C., Bierregaard, R. O., & Lovejoy, T. E. (2007). A large-scale deforestation experiment: effects of patch area and isolation on Amazon birds. *Science*, 315(5809), 238–241.

Furley, P. A. (1999). The nature and diversity of Neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Global Ecology and Biogeography*, 8(3-4), 223–241.

Glenn, T. C., Nilsen, R., Kieran, T. J., Finger, J. W., Pierson, T. W., Bentley, K. E., … Reed, K. (2016). Adapterama I: universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). *BioRxiv*, 049114.

Gómez-Bahamón, V., Márquez, R., Jahn, A. E., Miyaki, C. Y., Tuero, D. T., Laverde-R, O., ... Cadena, C. D. (2020). Speciation associated with shifts in migratory behavior in an avian radiation. *Current Biology* 30, 1312–1321.

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.

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(10), 1031.

Guarnizo, C. E., Werneck, F. P., Giugliano, L. G., Santos, M. G., Fenker, J., Sousa, L., … Dorado-Rodrigues, T. F. (2016). Cryptic lineages and diversification of an endemic anole lizard (Squamata, Dactyloidae) of the Cerrado hotspot. *Molecular phylogenetics and evolution*, 94, 279–289.

Guilherme, E., & Czaban, R. E. (2015). First record of the yellowish pipit in Acre with notes on other grassland birds in southwestern Amazônia. *Neotropical Biology and Conservation*, 10(3), 169–176.

Guimarães, D. P., Gomes de Lima, A. C., & Pedroza, D. (2016). New record and geographic expansion of *Vanellus chilensis* in the state of Amazonas, Brazil. Atualidades Ornitologicas 190:24.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, 165(3889), 131–137.

Hayes, F. E., & Sewlal, J. A. N. (2004). The Amazon River as a dispersal barrier to passerine birds: effects of river width, habitat and taxonomy. *Journal of Biogeography*, 31(11), 1809–1818.

Heled, J., & Drummond, A. J. (2009). Bayesian inference of species trees from multilocus data. *Molecular biology and evolution*, 27(3), 570–580.

Hockey, P. A. (2000). Patterns and correlates of bird migrations in sub-Saharan Africa. *Emu*, 100(5), 401–417.

Huber, O., de Stefano, R. D., Aymard, G., & Riina, R. (2006). Flora and vegetation of the Venezuelan Llanos: a review. In: Ratter J. A., & R. T. Pennington (Eds.), Neotropical savannas and seasonally dry forests: plant diversity, biogeography, and *conservation* (pp. 95–120). Boca Raton, FL, CRC.

Jombart T, & Ahmed, I. (2011). Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. doi: 10.1093/bioinformatics/btr521.

Jones, H. T., Mayle, F. E., Pennington, R. T., & Killeen, T. J. (2011). Characterisation of Bolivian savanna ecosystems by their modern pollen rain and implications for fossil pollen records. *Review of Palaeobotany and Palynology*, 164(3-4), 223–237.

Katoh, K., & Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30(4), 772–780. https://doi.org/10.1093/molbev/mst010

Larget, B. R., Kotha, S. K., Dewey, C. N., & Ané, C. (2010). BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics*, 26(22), 2910–2911.

Lees, A. C., Moura, N. G., Almeida, A. S., & Vieira, I. C. (2014). Noteworthy ornithological records from the threatened campinas of the lower rio Tocantins, east Amazonian Brazil. *Bulletin of the British Ornithologists' Club*, 134(4), 247–258.

Mayle, F. E., Burbridge, R., & Killeen, T. J. (2000). Millennial-scale dynamics of southern Amazonian rain forests. *Science*, 290(5500), 2291–2294.

Marini, M. Â., & Cavalcanti, R. B. (1990). Migrações de *Elaenia albiceps chilensis* e *Elaenia chiriquensis albivertex* (Aves: Tyrannidae). *Boletim do Museu Paraense Emílio Goeldi. Nova série. Zoologia*, 6(1), 59–67.

Markgraf, V. (1993). Climatic history of Central and South America since 18,000 yr BP: comparison of pollen records and model simulations. *Global climates since the last glacial maximum*, 14, 357–385.

Matos, M. V., Borges, S. H., d'Horta, F. M., Cornelius, C., Latrubesse, E., Cohn-Haft, M., & Ribas, C. C. (2016). Comparative phylogeography of two bird species, *Tachyphonus phoenicius* (Thraupidae) and *Polytmus theresiae* (Trochilidae), specialized in Amazonian white-sand vegetation. *Biotropica*, 48(1), 110–120.

Mittermeier, J. C., Zyskowski, K., Stowe, E. S., & Lai, J. E. (2010). Avifauna of the Sipaliwini savanna (Suriname) with insights into its biogeographic affinities. *Bulletin of the Peabody Museum of Natural History*, 51(1), 97–123.

Mooers, A. Ø., & Harvey, P. H. (1994). Metabolic rate, generation time, and the rate of molecular evolution in birds. *Molecular phylogenetics and evolution*, 3(4), 344–350.

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.

Norambuena, H. V., Van Els, P., Muñoz-Ramírez, C. P., & Victoriano, P. F. (2018). First steps towards assessing the evolutionary history and phylogeography of a widely distributed Neotropical grassland bird (Motacillidae: *Anthus correndera*). *PeerJ*, 6, e5886.

Novaes, R. M. L., Ribeiro, R. A., Lemos-Filho, J. P., & Lovato, M. B. (2013). Concordance between phylogeographical and biogeographical patterns in the Brazilian Cerrado: diversification of the endemic tree *Dalbergia miscolobium* (Fabaceae). *PLoS One*, 8(12), e82198.

De Paiva, L. V., & Marini, M. Â. (2013). Timing of migration and breeding of the Lesser Elaenia (*Elaenia chiriquensis*) in a Neotropical savanna. *The Wilson Journal of Ornithology*, 125(1), 116–120.

Paradis, E. (2010). Pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, 26(3), 419–420.

Pennington, R. T., & Hughes, C. E. (2014). The remarkable congruence of New and Old World savanna origins. *New Phytologist*, 204(1), 4–6.

Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). Tracer 1.6.

Robbins, M. B., Braun, M. J., & Finch, D. W. (2004). Avifauna of the Guyana southern Rupununi, with comparisons to other savannas of northern South America. *Ornitologia Neotropical*, 15, 173–200.

Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*, 22(5), 939–946.

Rutt, C. L., Jirinec, V., Cohn-Haft, M., Laurance, W. F., & Stouffer, P. C. (2019). Avian ecological succession in the Amazon: A long-term case study following experimental deforestation. *Ecology and Evolution*, 9(24), 13850–13861.

Salgado-Labouriau, M. L. (1997). Late Quaternary palaeoclimate in the savannas of South America. *Journal of Quaternary Science: Published for the Quaternary Research Association*, 12(5), 371–379.

Sanaiotti, T. M., & Cintra, R. (2001). Breeding and migrating birds in an Amazonian savanna. *Studies on Neotropical Fauna and environment*, 36(1), 23–32.

Sánchez, M. V., González, M. G., & Genise, J. F. (2010). Phytolith analysis of *Coprinisphaera*, unlocking dung beetle behaviour, herbivore diets and palaeoenvironments along the Middle Eocene–Early Miocene of Patagonia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 285(3–4), 224–236.

Santos, M. G., Nogueira, C., Giugliano, L. G., & Colli, G. R. (2014). Landscape evolution and phylogeography of *Micrablepharus atticolus* (Squamata, Gymnophthalmidae), an endemic lizard of the Brazilian Cerrado. *Journal of biogeography*, 41(8), 1506–1519.

Santos, M.P.D. & J.M.C. Silva, 2007 - As aves das savanas de Roraima. Revista Brasileira de Ornitologia, 15(2), 189–207.

Satler, J. D., & Carstens, B. C. (2016). Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the *Sarracenia alata* pitcher plant system. *Evolution*, 70(5), 1105–1119.

Schliep, K. P. (2010). Phangorn: phylogenetic analysis in R. *Bioinformatics*, 27(4), 592–593.

Sheard, C., Neate-Clegg, M. H., Alioravainen, N., Jones, S. E., Vincent, C., MacGregor, H. E., … Tobias, J. A. (2020). Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nature communications*, 11(1), 1–9.

Sick, H., 1997. Ornitologia Brasileira: 1–912. Editora Nova Fronteira, Rio de Janeiro.

Silva, S. M., Peterson, A. T., Carneiro, L., Burlamaqui, T. C. T., Ribas, C. C., Sousa-Neves, T., … Batista, R. (2019). A dynamic continental moisture gradient drove Amazonian bird diversification. *Science advances*, 5(7), eaat5752.

Smeds, L., Qvarnström, A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome research*, 26(9), 1211–1218.

Smith, B. T., & Klicka, J. (2010). The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. *Ecography*, 33(2), 333–342.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., … Faircloth, B. C. (2014). The drivers of tropical speciation. *Nature*, 515(7527), 406.

Stamatakis, A., 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics 30(9), 1312–1313.

Stotz, D.F., Fitzpatrick, J. W., Parker, T. E. III, & Moskovits, D. K. (1996). *Neotropical birds, ecology and conservation*. University of Chicago Press, Chicago.

Tuero, D. T., Jahn, A. E., & MacPherson, M. (2019). Bird Migration in South America: The Fork-tailed Flycatcher (*Tyrannus savana*) as a case study. In J. C. Reboreda, V. D. Fiorini, & D. T. Tuero (Eds.). *Behavioral Ecology of Neotropical Birds* (pp. 133-154). Cham, Switzerland, Springer.

van der Hammen, T. (1974). The Pleistocene changes of vegetation and climate in tropical South America. *Journal of Biogeography*, 1(1), 3–26.

Vasconcelos, M. F., J. F. Pacheco, & R. Parrini. (2007). Levantamento e conservação da avifauna na zona urbana de Marabá, Pará, Brasil. *Cotinga* 28(2007), 45–51.

Vasconcelos, M. F. D., Dantas, S. D. M., & Silva, J. M. C. (2011). Avifaunal inventory of the Amazonian savannas and adjacent habitats of the Monte Alegre region (Pará, Brazil), with comments on biogeography and conservation. *Boletim do Museu Paraense Emílio Goeldi Ciências Naturales*, 6:119–145.

Webb, S. D. (1978). A history of savanna vertebrates in the New World. Part II: South America and the Great Interchange. *Annual Review of Ecology and Systematics*, 9(1), 393–426.

Werneck, F. P. (2011). The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quaternary Science Reviews*, 30(13–14), 1630–1648.

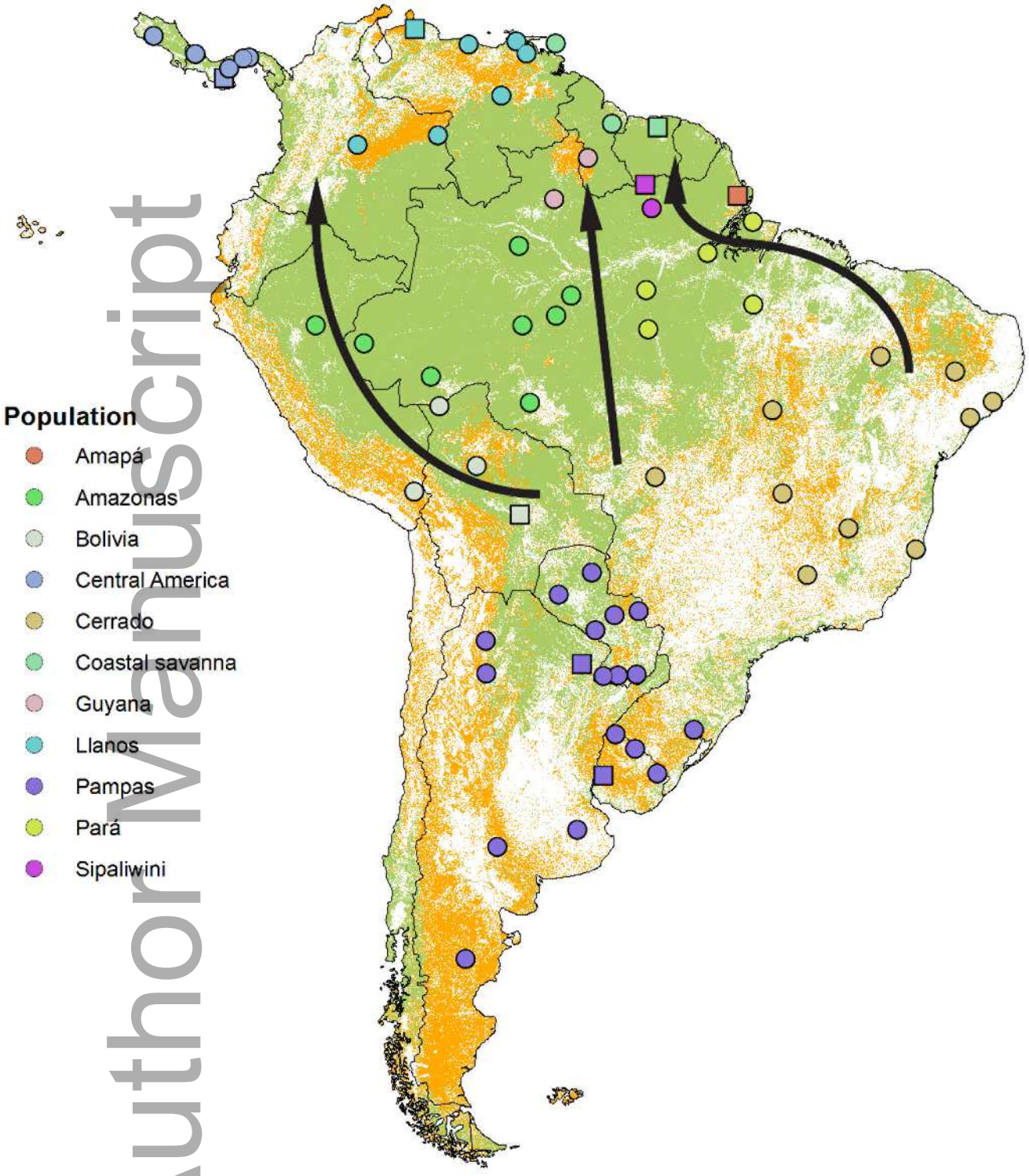
Wüster, W., Ferguson, J. E., Quijada-Mascareñas, J. A., Pook, C. E., Da Graca Salomao, M., & Thorpe, R. S. (2005). Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Molecular Ecology*, 14(4), 1095–1108.

Zink, R. M., & Barrowclough, G. F. (2008). Mitochondrial DNA under siege in avian phylogenetics. *Molecular Ecology,* 17, 2107–2121.

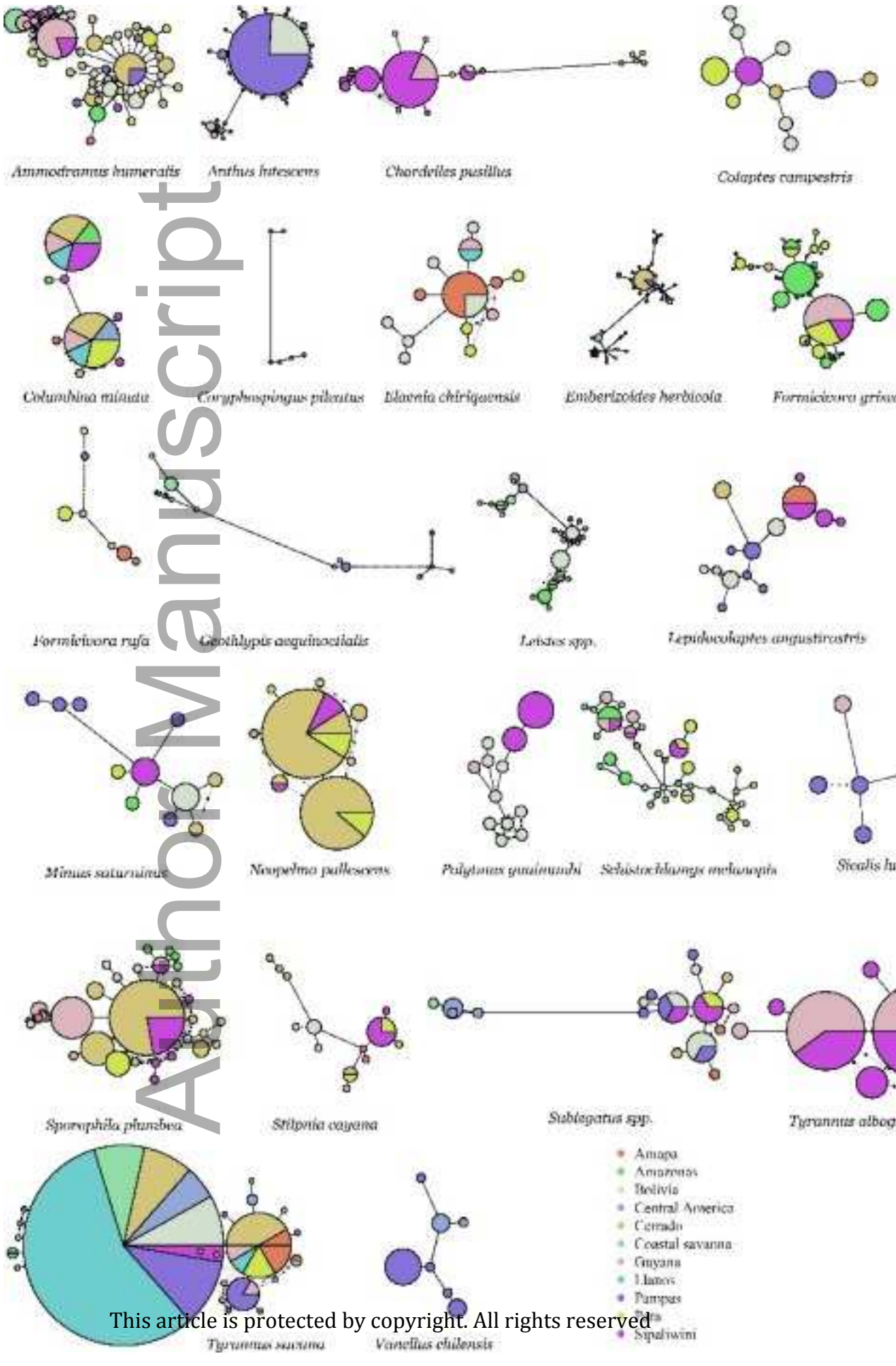
BIOSKETCH

Paul van Els focuses mainly on the ecology and evolution of Neotropical birds, particularly birds of grass-dominated ecosystems. He is especially interested in the timing and spatiality of diversification of this group in relation to the vicariant events that shaped the South American savannas and to the life history characteristics inherent to a life in open habitats with a high degree of seasonal climatic variation.

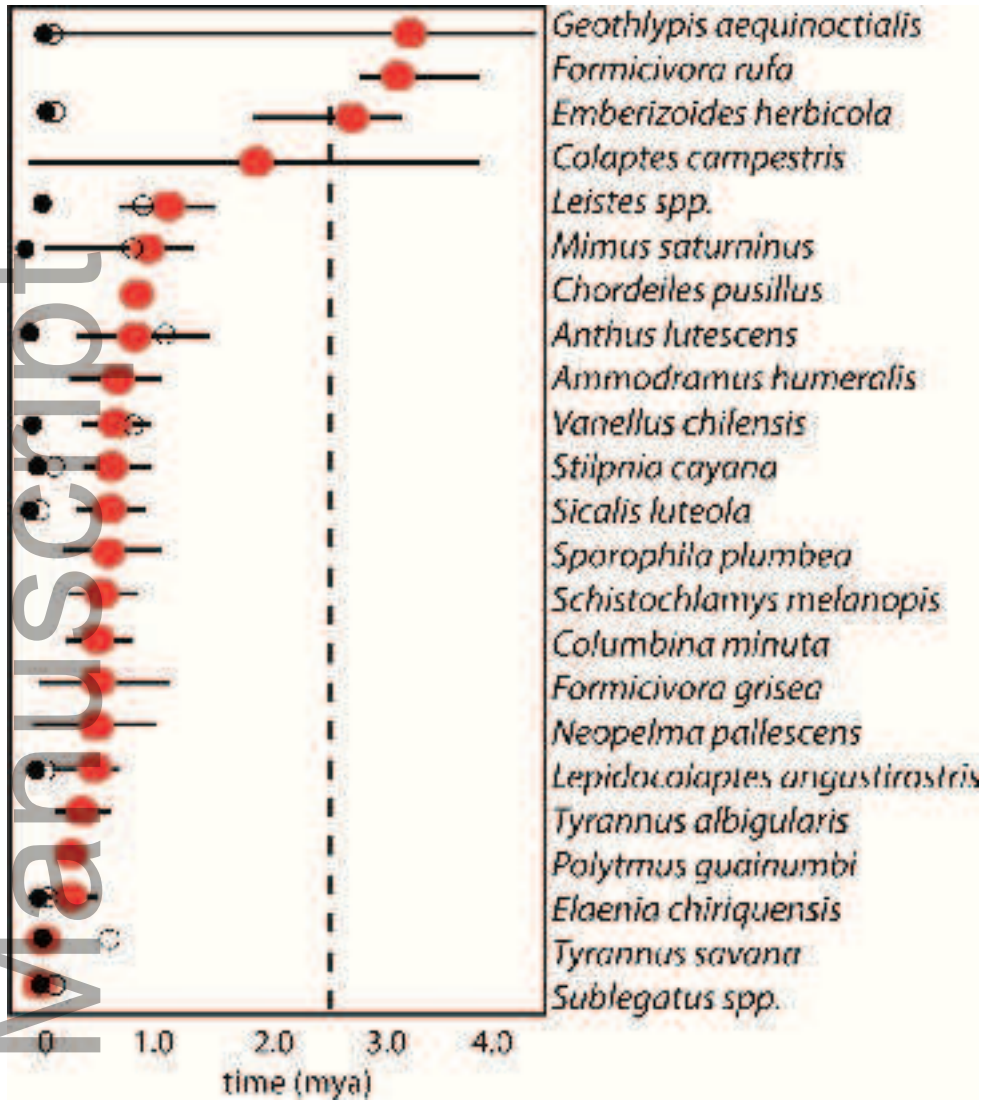
Author contributions: PVE conceived the research idea and performed most sequencing, analysis, and writing. EZ aligned UCE data, LRM performed UCE-based analyses, VGB provided samples of *Tyrannus savana*, AS, AA, CCR, PSDR, and MPDS provided samples of *Formicivora grisea, Neopelma pallescens* and *Schistochlamys melanops*, KZ, ROP, and JB provided samples of *Emberizoides herbicola, Tyrannus albogularis, Sporophila plumbea* and *Chordeiles pusillus*. All reviewed the manuscript.



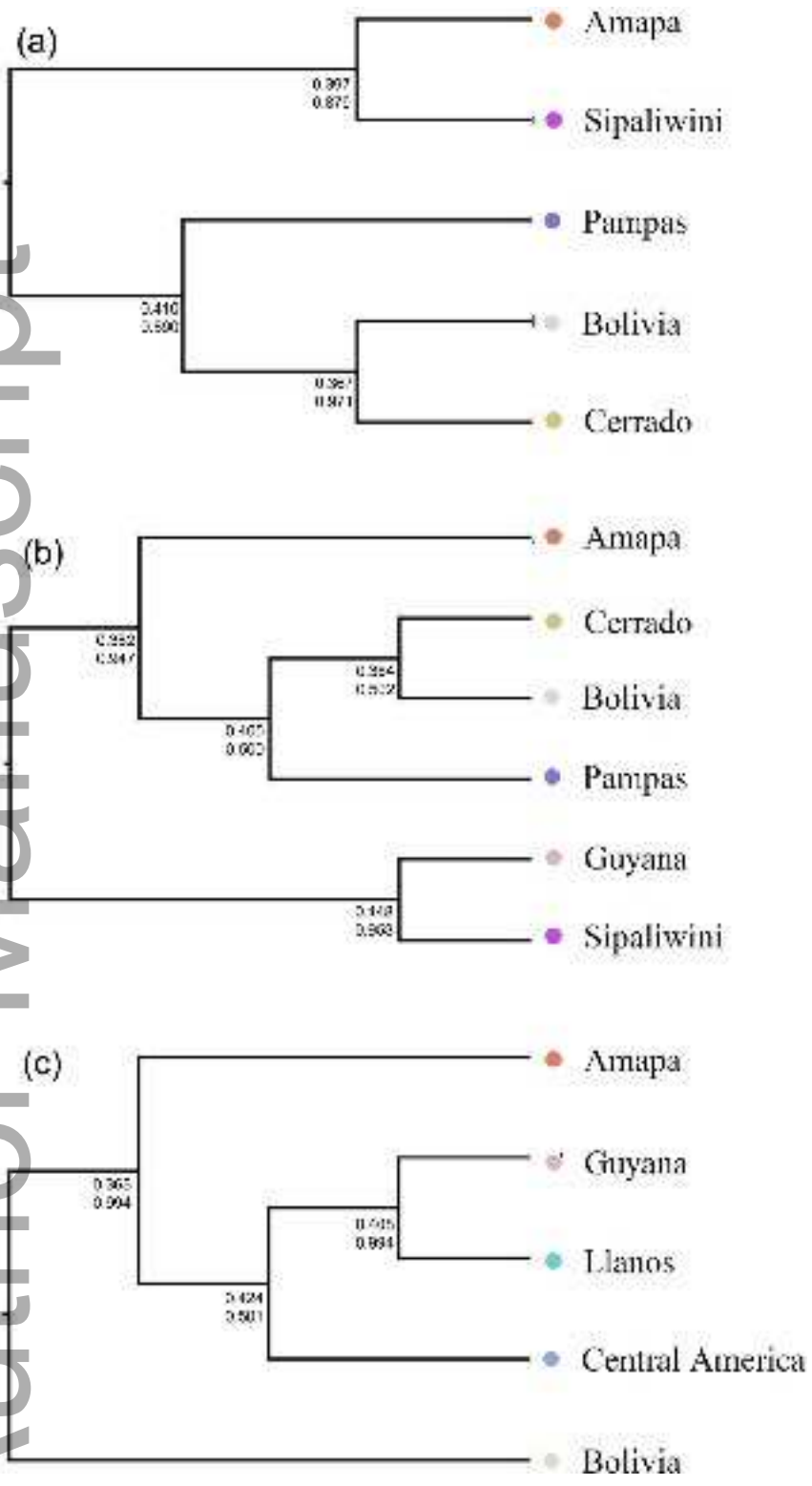
jbi_14065_f1.tif



jbi_14065_f2.tif



jbi_14065_f3.tif



jbi_14065_f4.tif