






Phylogenomics of the gray-breasted sabrewing (*Campylopterus largipennis*) species complex in the Amazonia and Cerrado biomes

Jean Carlo Pedroso de Oliveira¹ , Gustavo Sebastián Cabanne²  and Fabrício Rodrigues Santos¹ 

¹Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Departamento de Genética, Ecologia e Evolução, Belo Horizonte, MG, Brazil.

²División de Ornitología, Museo Argentino de Ciencias Naturales “Bernardino Rivadavia” (MACN – CONICET), Buenos Aires, Argentina.

Abstract

The Neotropics are one of the most biodiverse regions of the world, where environmental dynamics, climate and geology resulted in a complex diversity of fauna and flora. In such complex and heterogeneous environments, widely distributed species require deep investigation about their biogeographic history. The gray-breasted sabrewing hummingbird *Campylopterus largipennis* is a species complex that occurs in forest and open ecosystems of South America, including also high-altitude grasslands. It has been recently split into four distinct species distributed in Amazonia (rainforest) and Cerrado (savanna) biomes with boundaries marked by ecological barriers. Here, we investigated the evolutionary dynamics of population lineages within this neotropical taxon to elucidate its biogeographical history and current lineage diversity. We used a reduced-representation sequencing approach to perform fine-scale population genomic analyses of samples distributed throughout Amazonia and Cerrado localities, representing all four recently recognized species. We found a deep genetic structure separating species from both biomes, and a more recent divergence between species within each biome and from distinct habitats. The population dynamics through time was shown to be concordant with known vicariant events, isolation by distance, and altitudinal breaks, where the Amazon River and the Espinhaço Mountain Range worked as important barriers associated to speciation.

Keywords: Hummingbirds, Neotropics, phylogeography, landscape genomics, *Campylopterus*.

Received: March 05, 2024; Accepted: May 29, 2024.

Introduction

Neotropical ecosystems have attracted the worldwide attention of scientists since the dawn of evolutionary biology in the XIX century. The Neotropics hold a large portion of Earth's biodiversity, extending from southern Mexico to as far south as central Argentina and Chile. Besides, four out of the ten most biodiverse countries worldwide are located in South America due to its particular biogeographic history. The South American landscapes have been constantly reshaped, particularly since the Eocene, but underwent significant transformation around 12 million years ago (Ma) due to the rise of the Andes and formation of the current Amazon River. These changes have contributed to the evolution of complex ecosystems within remarkably distinct biomes such as the Amazonia and the Cerrado (Hoorn *et al.*, 2010), both of which shelter a rich and endemic biodiversity.

The Cerrado is a highly heterogeneous Neotropical savanna, and considered a global biodiversity hotspot (Myers *et al.*, 2000; Colli *et al.*, 2020). Besides typical savannas, it also presents grassland fields, woody forests, gallery forests, and patches of Seasonally Dry Tropical Forests (SDTFs) (Prado and Gibbs, 1993; Oliveira-Filho *et al.*, 2006; Neves

et al., 2015). The heterogeneous ecosystems of the Cerrado biome are also connected with the Amazonia biome in central Brazil, fostering important interactions (Marques *et al.*, 2020).

The Amazonia biome includes also many heterogeneous landscapes that are mostly composed of forest ecosystems. The historical environment dynamics in Amazonia have led to high levels of regional diversification and endemism, making it the primary source of Neotropical biodiversity (Antonelli *et al.*, 2018), being considered a relic of a once more extensive environment (Musher *et al.*, 2019). The species richness in the Amazon region has also been associated with high rates of *in situ* speciation and lineage sharing (Hoorn *et al.*, 2010; Smith *et al.*, 2014) due to facilitated emigration (Musher *et al.*, 2019) and the presence of areas of endemism (Ribas *et al.*, 2012; Braga *et al.*, 2022).

Climate change is known to be an important cause of vegetational changes over large areas (Haffer, 1969; Cheng *et al.*, 2013). The Amazonia and Cerrado biomes have experienced intense biogeographic interaction, mostly related to Quaternary Climatic Fluctuations (Werneck, 2011). Also, the Pleistocene Arc Hypothesis (PAH) postulates that dry forests enclaves had a broader and more contiguous distribution during the Pleistocene glacial maxima, but became fragmented and reduced during interglacial periods, as observed in the present (Prado and Gibbs, 1993; Pennington *et al.*, 2000; Neves *et al.*, 2015). These dynamics of expansion and retraction of dry forests during the Pleistocene might have resulted in a heterogeneous landscape of open and forest-

Send correspondence to Fabrício Rodrigues Santos. Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Departamento de Genética, Ecologia e Evolução, Av. Pres. Antônio Carlos, 6627, Pampulha, 31270-901, Belo Horizonte, MG, Brazil. E-mail: fsantos@icb.ufmg.br

savanna biomes between Amazonia and Cerrado (Marques *et al.*, 2020). Consequently, this interaction engendered distinct biogeographic patterns and phylogenetic relationships among species (Werneck *et al.*, 2012; Rocha *et al.*, 2020), including also the sharing of bird taxa between the two biomes (Vasconcelos and D'Angelo, 2018).

Birds drive diverse hypotheses on biota diversification in the Neotropics (Ribas *et al.*, 2012; Musher *et al.*, 2019; Silva *et al.*, 2019; Norambuena and Van Els, 2021). For example, connections between Neotropical forests have been deeply explored and different forest corridors through the dry diagonal biomes (Cerrado, Caatinga and Chaco) promoted gene flow in multiple bird taxa (Batalha-Filho *et al.*, 2013; Cabanne *et al.*, 2016; Trujillo-Arias *et al.*, 2017; Capurucho *et al.*, 2018; Berv *et al.*, 2021). Thus, bird dispersal enables gene flow between distant populations from different biomes and ecosystems (Lima-Rezende *et al.*, 2019; de Freitas *et al.*, 2022). Indeed, the Cerrado biome has been indicated as a provider of corridors for bird dispersal through riparian forests connecting Amazon and Atlantic forests (Batalha-Filho *et al.*, 2013; Ledo and Colli, 2017; Trujillo-Arias *et al.*, 2017; Cabanne *et al.*, 2019), where some related species of rainforest taxa have established within Cerrado. Ledo and Colli (2017) revisited the connections between tropical forests, stating that the central Brazilian riparian network was an important route between Amazonia and Cerrado, but not as much with Atlantic Forest. This strengthens the hypothesis of the ability of bird species to succeed in occupying intermediate or new habitats, such as along the headwaters of Parana River (Tocantins basin), and Jequitinhonha and Doce rivers (Willis, 1992; Capurucho *et al.*, 2018).

The gray-breasted sabrewing species complex represents a unique case to investigate lineage diversification in the Neotropics, and particularly between Amazonia and Cerrado biomes and ecosystems. Until recently, it was recognized a single widespread species, *Campylopterus largipennis*, that occupied diverse ecosystems of Amazonia and Cerrado. With the recent description of *Campylopterus calcirupicola* (Lopes *et al.*, 2017), the gray-breasted sabrewing species complex taxonomy was revised after a century of uncertainty by the International Ornithological Committee (Gill *et al.*, 2021) and the Brazilian Ornithological Records Committee (Pacheco *et al.*, 2021). Two species are currently recognized in the Cerrado biome: *Campylopterus calcirupicola* Lopes, Vasconcelos & Gonzaga, 2017 that inhabits the SDTFs (*Matas Secas*), and *C. diamantinensis* Ruschi, 1963 endemic to high-altitude grassland rock outcrops (*Campo Rupestre*). Two species are recognized in the Amazonia biome: *Campylopterus largipennis* (Boddaert, 1783) that inhabits southern/southeastern Amazonia, and *C. obscurus* Gould, 1848, that inhabits northern/western Amazonia (Figure 1).

Many competing hypotheses about taxa diversification between Neotropical ecosystems have been reevaluated with genomic evidence (Baker *et al.*, 2014), since the large amount of data generated by massive sequencing increased the robustness of biogeographic analyses. Considering this, the four closely-related hummingbird species of the *C. largipennis* complex represent a unique case to investigate ancient lineage diversification between Neotropical ecosystems and biomes,

as their genetic signatures in a continent-wide distribution may reconstruct their history of successful occupation of Amazonia and Cerrado.

Here, we investigate the population dynamics and biogeographical history of the gray-breasted sabrewing *Campylopterus largipennis* species complex. The broad distribution of these four closely related taxa in distinct Amazonia and Cerrado ecosystems offers a unique opportunity to evaluate the long-term processes of population establishment and divergence in response to diverse environments. We used genome-wide SNPs identified by double digest restriction associated DNA sequencing (ddRAD) method to evaluate population genetic structure, demographic history, gene flow and isolation patterns. Our study aims to elucidate the evolutionary pathways of these populations within the Amazonia and Cerrado biomes through phylogeographic and landscape genetics analyses, thereby addressing hypotheses regarding genetic differentiation, habitat connectivity, and biogeographic implications.

Material and Methods

Sample collection

We assessed a total of 78 individuals distributed along Amazonia and Cerrado biomes (Figure 1, Table S1), covering almost the entire distribution of the *Campylopterus largipennis* species complex, representing all four recognized taxa. All samples were previously available in the public collections of Centro de Coleoes Taxonomicas of Universidade Federal de Minas Gerais (CCT-UFMG), Coleao de Recursos Geneticos of Instituto Nacional de Pesquisas da Amazonia (INPA), Museu de Zoologia of Universidade de Sao Paulo (MZUSP) and Museu Paraense Emilio Goeldi (MPEG).

ddRAD sequencing and SNP filtering

Genomic ddRAD libraries were prepared following the protocol of Thrasher *et al.* (2018), modified from Peterson *et al.* (2012). Genomic DNA was extracted with phenol:chloroform protocol (Sambrook and Russell 2006). Quality and quantifications were made on Nanodrop (Thermo ScientificTM), gel agarose and Qubit (Thermo ScientificTM). The digestion-ligation reactions were made with 400 ng of genomic DNA, digested with restriction enzymes *Sbf*I and *Msp*I (New England Biolabs, MA), then ligated on barcoded adapters for sample identification. We used 20 unique adapters for each library pool and size selected on PippinPrep (Sage Science, MA) for ~400 base pairs fragments. Additionally, samples were randomized between libraries, in order to mitigate technical library effects (O'Leary *et al.*, 2018). We checked the libraries' reliability by qPCR-based quantification using KAPA Library Quantification Kits (KK4604, Kapa Biosystems), before sending them to sequencing facilities. Finally, libraries were sequenced on Illumina HiSeq SE150 in Macrogen (South Korea) and one library on Illumina HiSeq PE150 in GenOne Biotechnologies (Brazil).

The quality of raw reads was checked by FastQC (Andrews, 2010), libraries were demultiplexed using BBMap (Available at <http://sourceforge.net/projects/bbmap/>) and assembled *de novo* (Figure S1) using iPyrad version 0.7.30

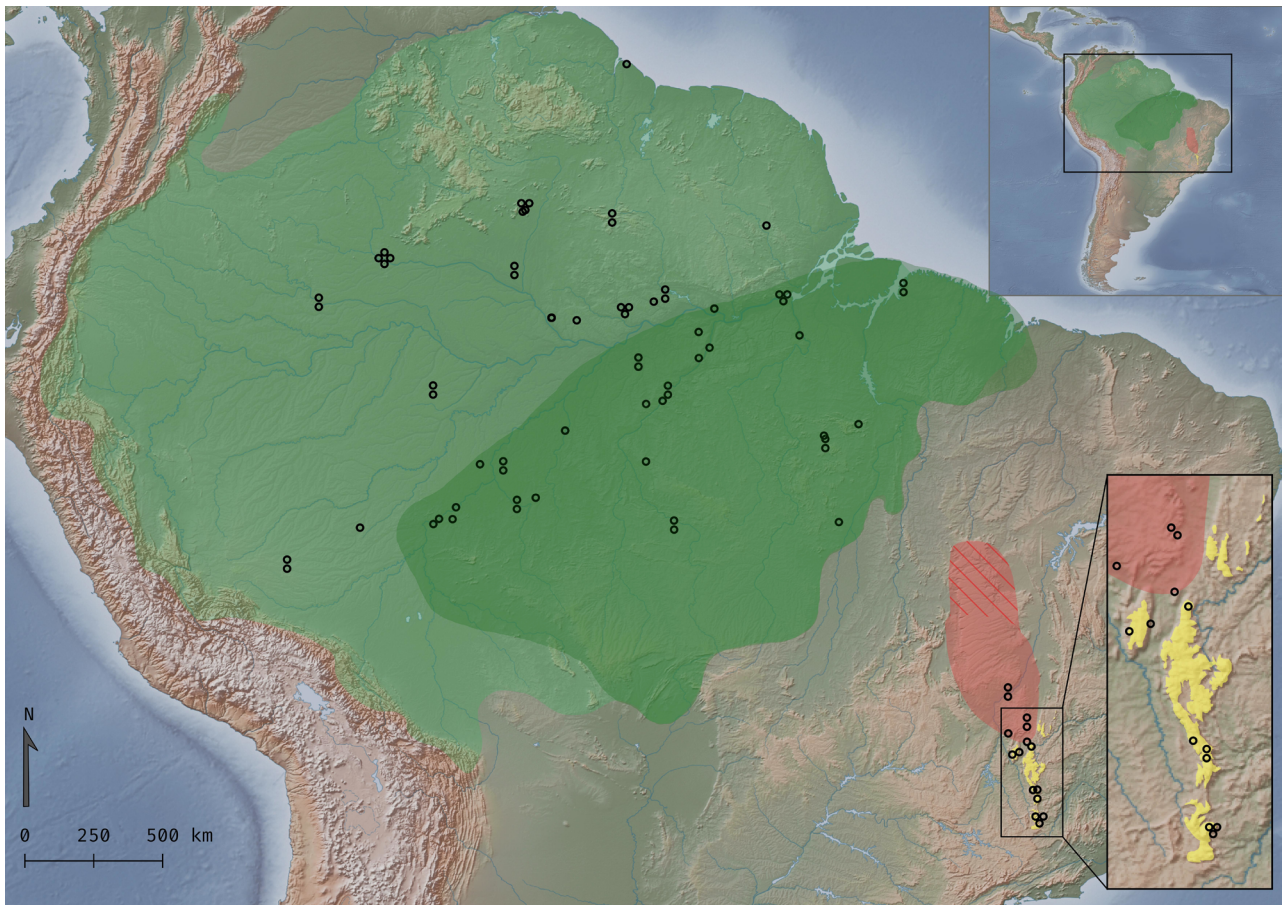


Figure 1 – Geographic distribution of samples used in the phylogenomic analysis of the *Campylopterus largipennis* species complex. Polygons represent species ranges *Campylopterus largipennis* (green), *C. obscurus* (light green), *C. calcirupicola* (red: extant; hatched: possibly extant) and of *C. diamantinensis* (yellow). Distributions adapted from BirdLife International (BirdLife International, 2021) and Lopes *et al.* (2017). The inset indicates the study area in South America and markers indicate sampled individuals.

(Eaton and Overcast, 2020). Most of the parameters were used as default, with a cluster threshold of 0.90 selected after testing the cluster threshold ranging from 0.85 to 0.95, observing the error rates and heterozygosity (Mastretta-Yanes *et al.*, 2015). The minimum samples per locus was set to 8, a stricter filter for adapters/primers was applied and reads were trimmed at 100 base pairs. Finally, the assemblies were evaluated by eye for overall coverage, using Matrix Condenser v.1.0 (Medeiros and Farrell, 2018) (Available at https://github.com/brunoasm/matrix_condenser/). A filter was applied to the Minor Allele Frequency (MAF) ranging from 0.05 to 0.1. This was done to assess the reproducibility of the protocol and to identify any potential confounding structures by conducting Principal Component Analysis (PCA) (Cumer *et al.*, 2021). The main data set was filtered for MAF, indels and missingness using VCFTools version 0.1.16 (Danecek *et al.*, 2011).

Population structure and phylogenetics analysis

The starting point of population structure was inferred using the fineRADstructure package (Malinsky *et al.*, 2018) with 200,000 iterations and burnin of 100,000. PCA was done on *adeget* version 2.1.7 (Jombart and Ahmed, 2011) toolset on R version 4.2.1 (R Core Team, 2022). We also performed an individual-based clustering on STRUCTURE v.2.3.4 (Pritchard *et al.*, 2000), with $k = 2$ to 6, 6 nreps, with 100,000 iterations

and burnin of 50,000. Then, we have chosen the most likely k by the estimated log probability means and ΔK (Evanno *et al.*, 2005). From then on, we considered the dataset separated into four main groups: Northern-Western Amazon (NWA), Southern-Eastern Amazon (SEA), *Campo Rupestre* (CR) and *Matas Secas* (MS). The pairwise F_{ST} was calculated on *hierfstat* version 0.5-11 (Goudet, 2005) using the function *boot.ppfst* with 1000 bootstraps to obtain confidence intervals of 95%. The identity-by-state (IBS) analysis of pairwise distances matrix from SNP data was done with SNPRelate package (Zheng *et al.*, 2012). All packages were run on R version 4.2.1 (R Core Team, 2022). We estimated a maximum likelihood phylogenetic tree using RAXML-NG v.1.0.3 (Kozlov *et al.*, 2019) for all samples, with a GTRgamma substitution model and 1,000 bootstrap replicates.

Demographic history

A direct estimation of divergence times and temporal gene flow between phylogenetic groups were conducted on G-PHOCS version 1.3 (Gronau *et al.*, 2011), on a subset of four samples per group based on cluster analysis and Matrix Condenser, with no indels and only loci shared between all samples. Three independent runs were performed with find-finetunes TRUE, standard priors and mutation rate 2.3×10^{-9} mutations per site per year (Smeds *et al.*, 2016), scaled by 10^4

(Gronau *et al.*, 2011). The runs were set for 3,000,000 MCMC iterations, and approximately 30% of the initial portion of chains were discarded in Tracer version 1.7 (Rambaut *et al.*, 2018), ensuring verification of posterior convergence and parameter values (τ and θ).

Estimating effective migration

We used the Estimated Effective Migration Surfaces (EEMS) method (Petkova *et al.*, 2016) to identify barriers for gene flow over the landscape. We set the parameters for 2,000 demes, 2,000,000 MCMC iterations sampled every 9,999 iterations after a 1,000,000 burn-in. Three independent runs were conducted, the MCMC chain traces were checked, and the proposal variance values were adjusted following the manual. Finally, MCMC runs were combined for final plots using *rEEMSPlots* in R, available with the *EEMS* package.

Results

We obtained an average of 1.07 ± 0.7 million reads per sample. The main data set included 5,141 SNP variant positions (one per locus, 5% MAF, no indels and up 10% missing data), shared between 74 individuals of *Campylopterus largipennis* complex, from most of the total geographic distribution, including 25 individuals of *C. largipennis*, 34 individuals of *C. obscurus*, six individuals from *C. calcirupicola* and nine individuals from *C. diamantinensis*. Samples clustered according to geographic correspondence on PCAs using different values of MAF, without detecting library effects between the experiments. Only one sample (INPA_13625) was removed due to a large amount of missing data (see Table S1 for details).

Three major groups were identified in the data set after applying PCA, Structure, and coancestry analysis (Figures 2–4).

All clustering methods showed congruence with the geographical (and biome) location of samples, revealing a closer relationship among samples from Cerrado biome (Figure 4, Figure S2). The genetic structure explained by PCA accounted for 48.5% along axis 1, separating samples from Amazonia and Cerrado, and 14.7% along axis 2, which revealed distinctions among Amazonia samples. Notably, subdivisions within the Cerrado biome only became apparent in the fourth component, indicating a relatively shallow level of differentiation (Figure S3). A hierarchical clustering analysis based on identity-by-state matrix done in *hierfstat* revealed a similar pattern observed with the clustering methods (Figure S4). The Structure analysis also detected three clusters without detecting substructure between populations from Cerrado. Most admixed individuals were located in the Southern-Western Amazon region, along the Madeira River. Although these samples present a greater admixture with the Cerrado, fineRADstructure analysis revealed a higher degree of co-ancestry between Amazonian regions. We designated these groups accordingly, naming them Northern-Western Amazon (NWA) and Southern-Eastern Amazon (SEA) groups from the Amazonia biome, separated from the Cerrado group. Following the established phylogeny (see below), we subdivided Cerrado into *Campo Rupestre* (CR) and *Matas Secas* (MS) for further analyses. Pairwise F_{ST} estimates between geographic groups ranged from 0.23 to 0.70 (Table S2) and showed significant differences between all populations.

The maximum likelihood phylogenetic tree also revealed the same major groups found in clustering methods. Clades from the Amazonia and Cerrado biomes are easily distinguished (Figure 5). Further, in the Cerrado biome, CR and MS groups were both reciprocally monophyletic, even though they were not distinguished in clustering analysis (Figure 2). Besides,

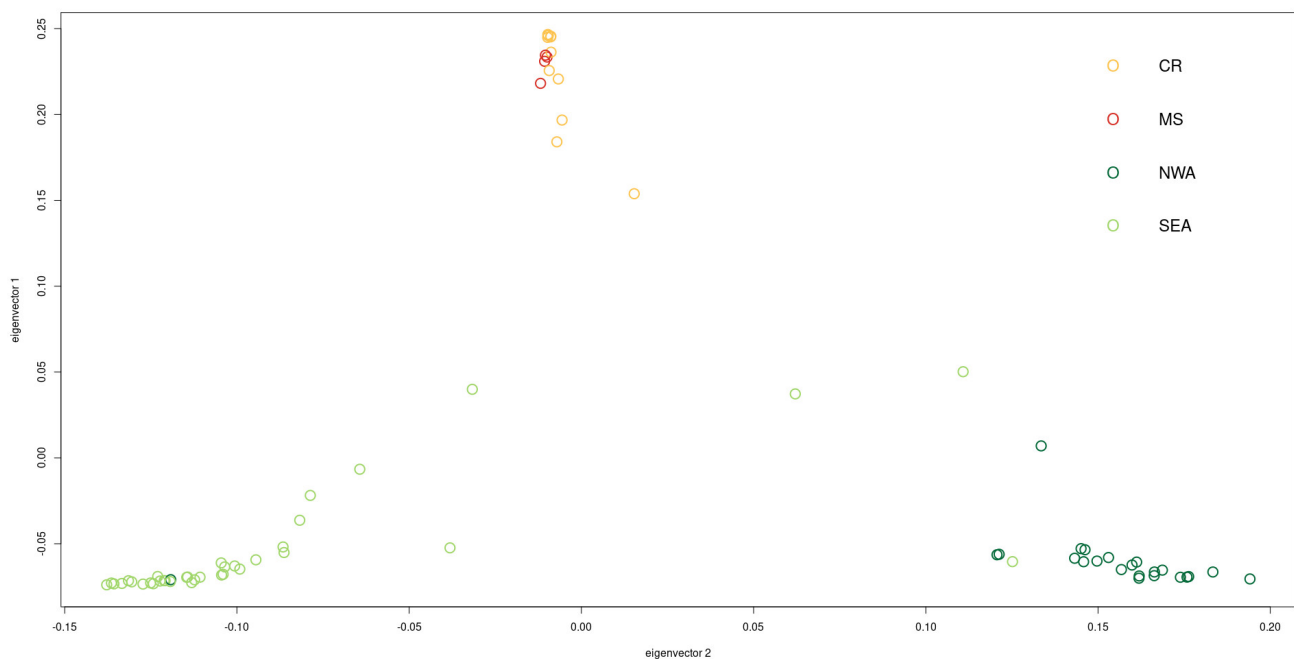


Figure 2 – Principal components analysis (PCA) of *Campylopterus largipennis* species complex based on 3,584 SNPs. The groups are colored as follows: *C. largipennis* in green (NWA: Northern-Western Amazon), *C. obscurus* in light green (SEA: Southern-Eastern Amazon), *C. diamantinensis* in yellow (CR: *Campo Rupestre*), and *C. calcirupicola* in red (MS: *Matas Secas*).

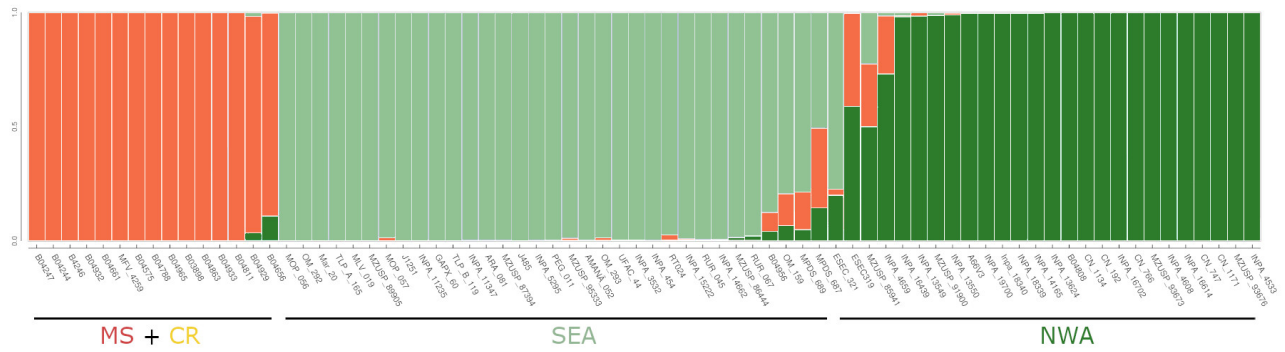


Figure 3 – Population genetic structure of the *Campylopterus largipennis* species complex. Bayesian analysis implemented in Structure v.2.3.4, with $k=3$ populations estimated by ΔK and log probability means. Each bar represents an individual colored proportionally to its probability of assignment to each population. The cluster includes 74 samples of *C. calcirupicola* and *C. diamantiniensis* (orange), *C. obscurus* (green) and *C. largipennis* (light green). Location bars represent *Matas Secas* and *Campo Rupestre* (MS + CR), and Southern-Eastern (SEA) and Northern-Western (NWA) Amazon regions.

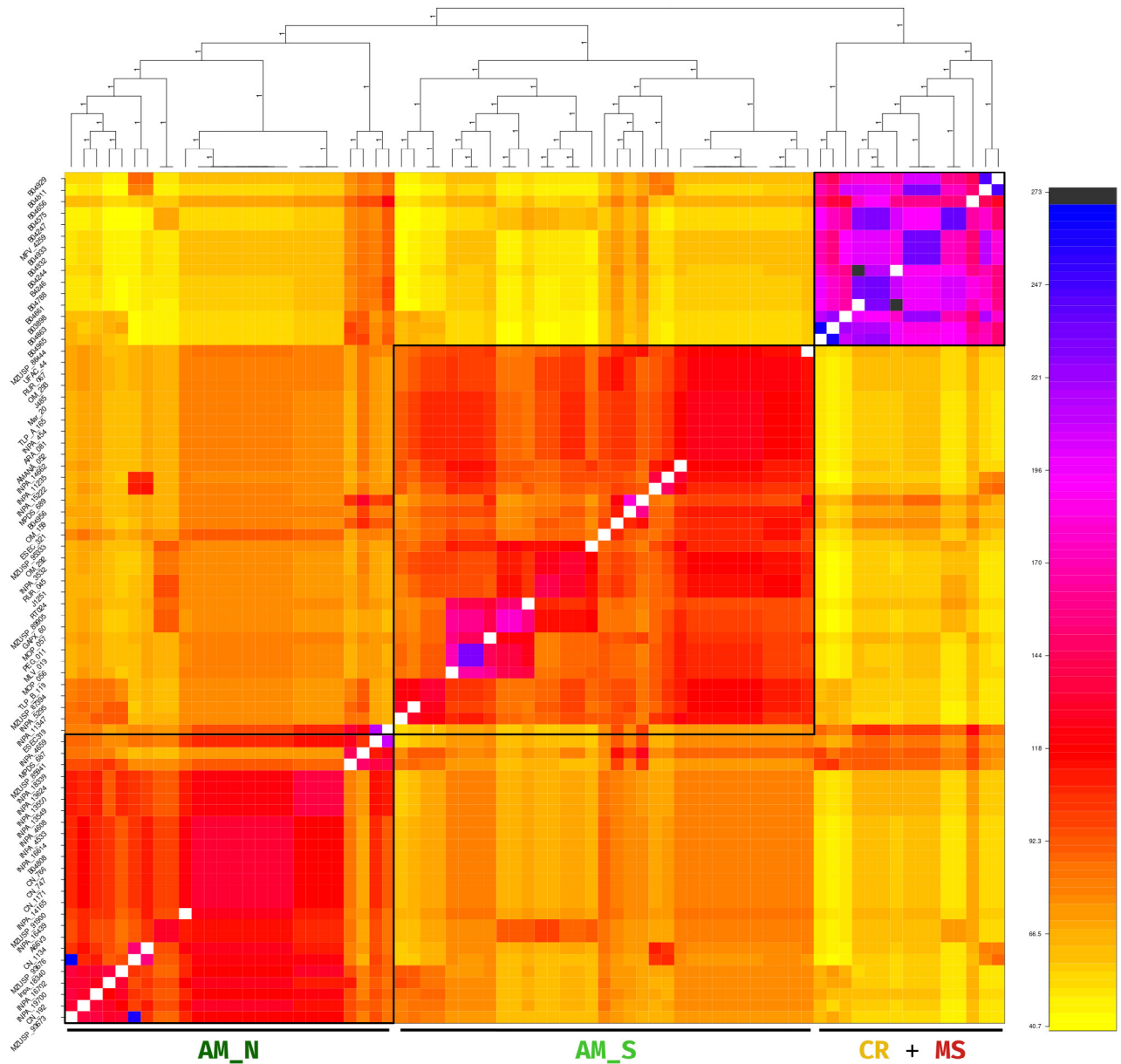


Figure 4 – Averaged co-ancestry matrix of the *Campylopterus largipennis* species complex. Each row represents one of the 74 samples in the heatmap that indicate the degree of co-ancestry, increasing from yellow to blue.

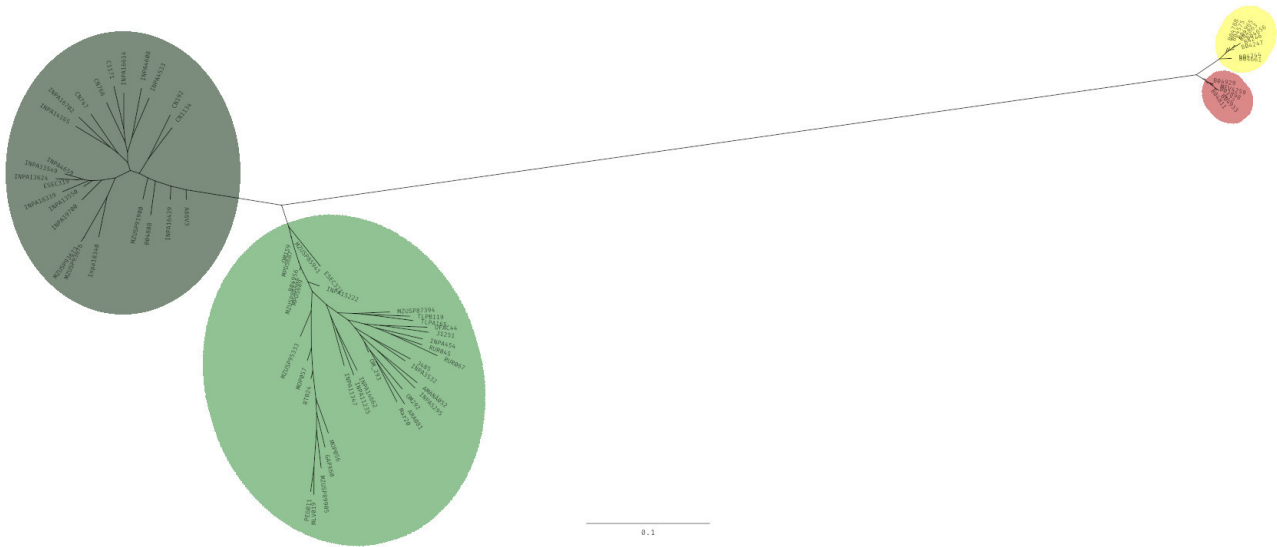


Figure 5 – Maximum likelihood phylogenetic tree of the *Campylopterus largipennis* species complex. RAXML-NG tree generated with 4,917 loci obtained from ddRAD data filtered without indels, MAF 0.05, and missingness 90%. Colors represent *C. largipennis* in green, *C. obscurus* in light green, *C. diamantinensis* in yellow, and *C. calcirupicola* in red.

in Amazonia we can see South (SEA) and North (NWA) groups, even with some level of admixture found between these two clusters.

G-PHOCS estimated all splitting times to different dates of milder interglacial times of the Mid-Pleistocene period (Figure 6, Table S3). The divergence between Amazonia and Cerrado lineages occurred at 1.24 Ma. The subdivisions of NWA and SEA of Amazonia occurred at 622 thousand years ago (Ka) while Cerrado clusters (CR and MS) diverged at 201 Ka. The gene flow (migration) estimates indicate an unbalanced movement of individuals (or genes) from Amazonia to Cerrado, where MS (*Matas Secas*) was the main migrant receiver. The *EEMS* analysis identified two main obstacles for gene flow (Figure 7, Table S4). The drainage system of the Purus-Madeira (upper Amazon River) is moderately permeable, but it becomes a hard barrier when it reaches the lower Amazon River, dividing current species *C. largipennis* (NWA) and *C. obscurus* (SEA). The other visible hurdle to gene flow corresponds to the altitude variance of the Espinhaço Mountain Range, which separates the highlands (CR) and lowlands (limestone SDTF environment of MS) of Cerrado. These distinct habitats, correspond respectively to the localities of occurrence for *C. diamantinensis* and *C. calcirupicola* species.

Discussion

Until the recent description of *Campylopterus calcirupicola*, the gray-breasted sabrewing species complex has carried taxonomic uncertainty for over a century (Lopes *et al.*, 2017). Here we investigated a fine scale diversification of this sabrewing hummingbird species complex, which is broadly distributed across two Neotropical biomes. We assessed samples from both Cerrado and Amazonia populations,

with a SNP data set covering all related taxa and most of the occurrence area. Our findings using a population genomics approach suggest a combination of landscape changes and dispersion leading to taxa diversification. While initial investigations did not identify all taxa, this may be attributed to recent speciation or potential sampling biases. Subsequent explorations have corroborated the existence of the currently recognized taxa *Campylopterus calcirupicola* Lopes, Vasconcelos & Gonzaga, 2017, *C. largipennis* (Boddaert, 1783), *C. obscurus* Gould, 1848, and *C. diamantinensis* Ruschi, 1963, reviewed by the International Ornithological Committee (Gill *et al.*, 2021) and the Brazilian Ornithological Records Committee (Pacheco *et al.*, 2021).

Gray-breasted sabrewing populations were affected by diverse biogeographic processes promoting divergence between and within biomes. The differentiation found between Amazonia and Cerrado clades were consistent with the intermediate dispersal model hypothesis (Norambuena and Van Els, 2021). Dispersal works both ways in the process of speciation. On the one hand, it can contribute to the homogenization of populations through gene flow, but it can sometimes result in the colonization of new areas, and eventually lead to subsequent local adaptation and speciation. Dispersal capacity itself is also related to geographic diversification. While species with greater dispersal capacity tend to occupy widespread areas, generating new subpopulations, species with low dispersal capacity struggle to maintain gene flow even in nearby regions. However, in both cases, the ability to disperse is related to the promotion of the speciation process. The intermediate dispersal hypothesis combines taxa dispersal ability and species diversity, and predicts that the most diverse clades are those with intermediately strong dispersal capacity (Yamaguchi, 2022). This scenario is proposed for wide

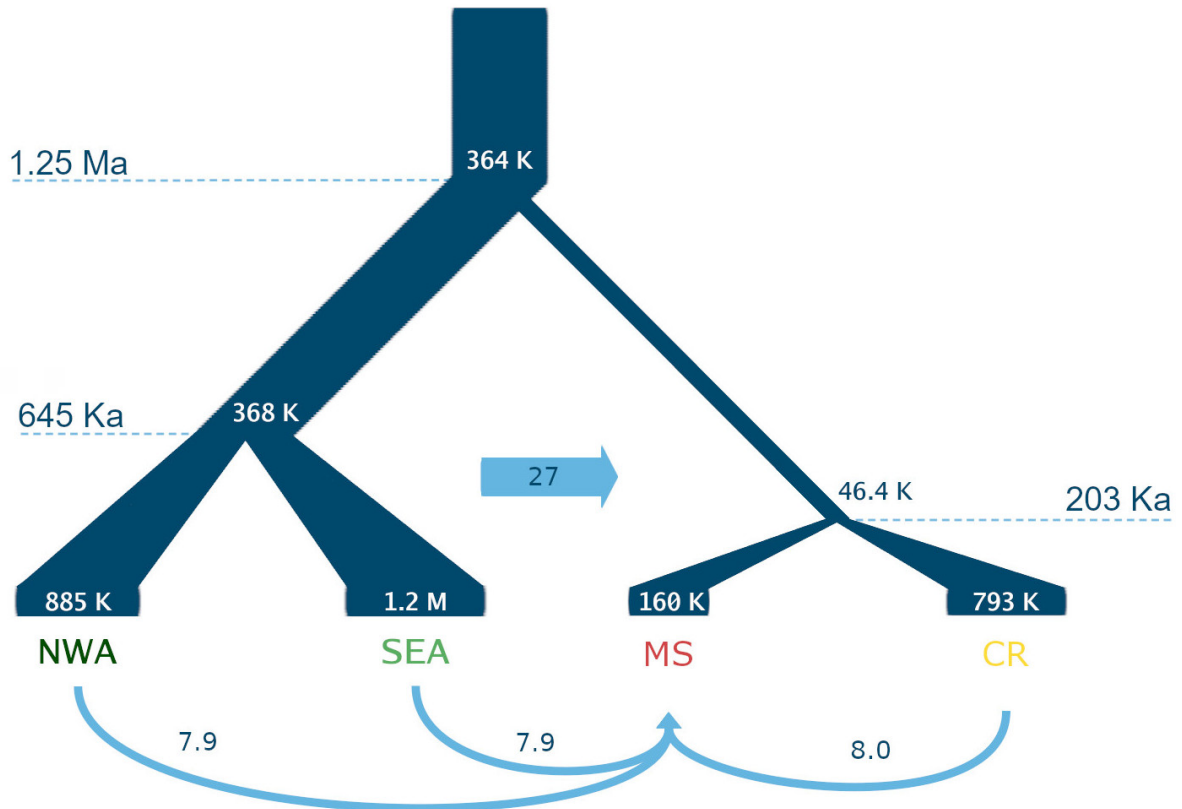


Figure 6 – Demographic histories inferred by G-PhoCS between structured groups of *Campylopterus largipennis* (NWA) and *C. obscurus* (SEA), *C. diamantinensis* from *Campo Rupestre* (CR, montane savanna) and *C. calcirupicola* from *Matas Secas* (MS, dry-forest). Historical effective sizes are inside the tree graph, divergence times are indicated at nodes by dashed lines. Arrows indicate the number of individual migrants per generation ($M_{sx} = m_{sx} \times \theta_{x/4}$).

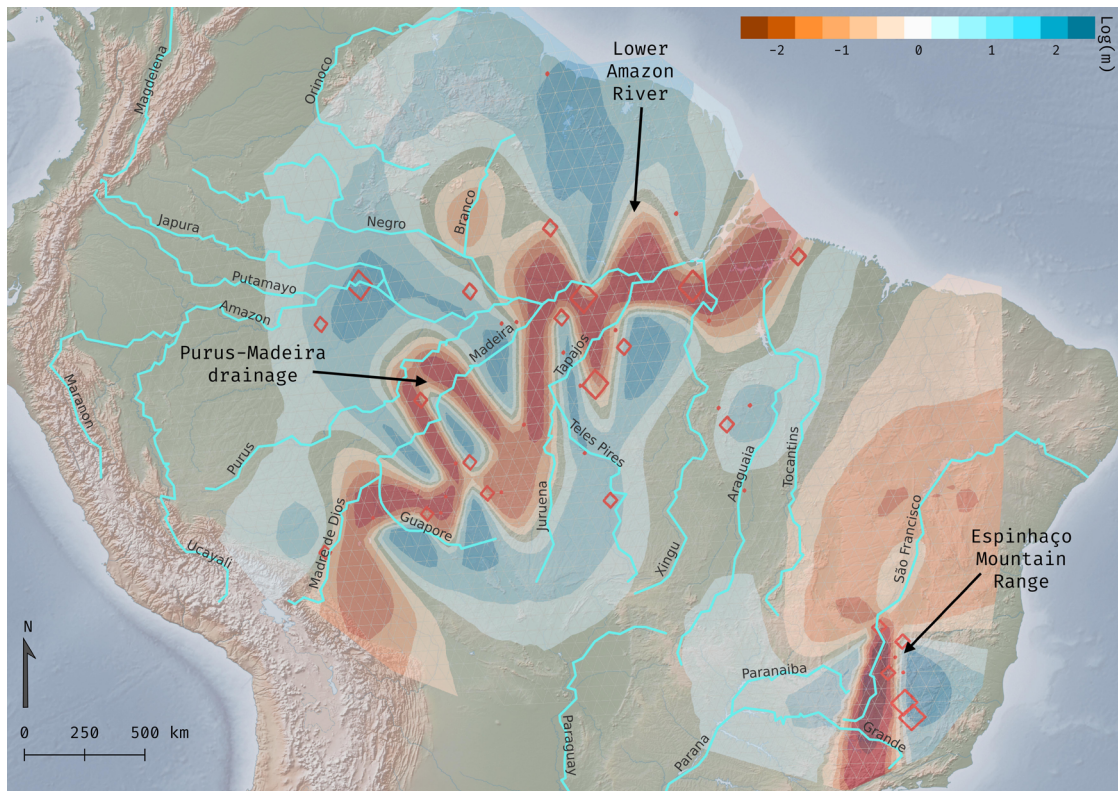


Figure 7 – Estimated effective migration surfaces (EEMS) of the *Campylopterus largipennis* species complex using 2,000 demes. Red diamonds represent populations varying in size according to number of samples in the deme. Reddish areas represent barriers to migration, while bluish regions are corridors for migration. Main regional barriers to gene flow identified by EEMS are indicated with arrows.

distributed grassland birds in the Neotropics (Norambuena and Van Els, 2021), where dispersal is a major factor on the speciation process in a continental scale (Smith *et al.*, 2014).

The cladogenesis events of the *Campylopterus largipennis* complex are coincident with major changes in glaciation periods. We estimated that Amazonia and Cerrado lineages separated during the Mid-Pleistocene Transition (Figure 6) at about 1.24 million years ago (Ma), the period when glacial cycles became longer and drier (Tziperman and Gildor, 2003; Willeit *et al.*, 2019). Glacial-interglacial periods during the Quaternary can be inferred from marine oxygen-isotope stages (MIS) obtained from deep sea core samples (Wright, 2013). The subsequent division of Amazonian *C. largipennis* and *C. obscurus* was also during a remarkably drier period approximately 621 thousand years ago (Ka), at the end of MIS 16 (Figure 6), the long-lasting cold phase of the Quaternary, characterized by very low CO₂ atmospheric concentrations (Hughes and Gibbard, 2018). The separation of Cerrado species *C. calcirupicola* and *C. diamantinensis* occurred at the end of MIS 7, an interglacial period around 201 Ka (Figure 6). The MIS 7-6 period was relatively warmer, and the moisture supply may have allowed the formation of extensive glaciers, where the ice volume accumulated in MIS 6 is related to a global water disturbance (Hughes and Gibbard, 2018). Those colder and drier periods possibly made humid forest to shrink and open vegetation and SDTFs (Werneck, 2011) to expand in different parts of Amazonia, as registered in pollen and geochemical data (Anhuf *et al.*, 2006; Reis *et al.*, 2017; Wang *et al.*, 2017; Kern *et al.*, 2023), likely expanding suitable areas for *Campylopterus* dispersal and colonization.

The gray-breasted sabrewings are known to be well adapted to semi-open vegetation, like margins of streams and forest edges (Schuchmann, 1999). A possible hypothesis for sabrewing hummingbird populations to have occupied the Cerrado was using forest edges in open areas as corridors for migration, since long-distance dispersal events are well known for hummingbirds. For example, the Nearctic region was invaded prior to the Panamanian uplift (<3.4 Ma) by Bee and Mountain-gem hummingbirds, followed by a rapid increase in invasions by other hummingbird lineages after the isthmus formation (McGuire *et al.*, 2014). If somehow climate allowed the expansion to new inhabitable areas, geographical barriers worked as maintainers of local diversity. In the *EEMS* analysis (Figure 7) we can see a transition from “soft to hard” constraints for migration along the Purus-Madeira River system to the lower Amazon River, isolating *C. largipennis* (NWA) from *C. obscurus* (SEA). Most of the samples that showed some admixture degrees are from *várzea* (floodplain) areas along the Madeira River. This result also enhances the relevance of interfluves as a biogeographically important suture zone in southernmost Amazonia (Dornas *et al.*, 2022). Similarly, samples from southeast of the Madeira and Tapajós rivers were grouped in *raxml* analysis. The resistance to migration was identified by the *EEMS* analysis between these rivers, showing to be a difficult region to cross, indicating how it has influenced the diversification of the taxon over time. Another important factor that may be contributing to the maintenance of this separation is the variability of precipitation on an orbital scale between the west and east of the Amazonia (Cheng *et al.*, 2013).

During glacial periods, this variability contributed to the greater fragmentation of forests in the eastern portion of the Amazonia.

Most of central Amazonia was supposedly occupied by SDTFs during the Pleistocene (Werneck, 2011) that were likely connected to dry forests (SDTFs) of the Cerrado, where sabrewing hummingbird ancestors could have lived and diverged into current taxa. This hypothesis is supported by a significant historical gene flow detected in G-PHOCS between Amazonia and Cerrado SDTFs (MS).

The more recent speciation of the sister species *C. calcirupicola* and *C. diamantinensis* in the Cerrado biome is likely a parapatric event related to ecological divergence in neighboring populations occupying low-altitude SDTFs/*Matas Secas* (MS) and high-altitude *Campo rupestre* (CR) on the Espinhaço Mountain Range. The occupation of different Cerrado ecosystems by two ecologically distinct sabrewing lineages likely occurred during one of the most recent glaciation-interglacial cycles of the Late Pleistocene. A possible path for the early colonization of Cerrado was through dry forests (SDTFs) and riparian forests of the headwaters of the Paranã River (Willis, 1992; Capurucho *et al.*, 2018), which is supported by recent sabrewing records (WikiAves, 2022).

The Neotropics were impacted by a huge and rapid event known as Great American Biotic Interchange (GABI) that ended 3.4 Ma, breaking up the continental isolation of South America since Gondwanaland split, enhancing the diversification rates (Weir *et al.*, 2009). During the GABI, it was observed an increase in the occupation of the Nearctic region by hummingbird lineages (McGuire *et al.*, 2014), evidencing the capacity for long dispersals by members of this group of birds. Hummingbird dispersal is also largely dependent on floral resources, acquired during a long evolutionary history. Diffuse coevolution with plants and niche conservatism is observed in hummingbird’s diversification, leading to generalists or specialist clades (McGuire *et al.*, 2014). The genus *Campylopterus* belongs to Emerald hummingbirds that are usually considered generalists, being able to visit a wide variety of plants (Rodríguez-Flores *et al.*, 2019). Indeed, the two adjacent Cerrado ecosystems (dry forest and rock outcrop fields) present completely different phytophysiognomies and are located in remarkably divergent landscapes at low (MS) and high (CR) altitudes. This generalist ecology favors distant migration and colonization of new areas, and may be related to its current distribution in a large part of Amazonia, and two particular Cerrado ecosystems.

The historical relationships between taxa from Amazonia and Cerrado are well documented for plant populations (Buzatti *et al.*, 2018). Isolation and limited dispersal contributed to lineage diversification, and the same explain most of endemic taxa of landscapes like the high-altitude ecosystems *Campo Rupestre* and *Pantepui* (Hopper *et al.*, 2021). Niche conservatism is observed in hummingbird mountain species that became restricted to narrow environments after colonization (McGuire *et al.*, 2014; Rodríguez-Flores *et al.*, 2019). Altitudinal clines also play important roles in hummingbird diversity, and genomic signatures of adaptation have been found for elevational gradients (Lim *et al.*, 2021). Local adaptation and speciation events associated to altitude

are commonly found in the Andes mountains, which hold the greatest diversity of modern hummingbirds, including other representatives of the genus *Campylopterus* (McGuire *et al.*, 2014). In sabrewing populations of Cerrado, we have found an altitudinal rift between *C. diamantinensis* and *C. calcirupicola*. The *C. diamantinensis* population is restricted to high-altitude grasslands of *Campo Rupestre* and *C. calcirupicola* to the lowland SDTFs (*Matas Secas*), separated by at least 300 meters of a sudden altitudinal barrier. These species have recently diverged within the Cerrado biome, and both taxa appear to be highly adapted and limited by their specific environments (CR and MS).

Our results indicate that Pleistocenic climatic cycles and open areas dynamics likely favored the occupation by sabrewing lineages that were successful to establish in their species-specific environments. Climate fluctuations enabled the connections of such distant environments and shaped the structure of populations over time by constraining the migration between cycles. The subsequent specialization (local adaptation) to *Matas Secas* and *Campo Rupestre* allowed the maintenance of populations in these places, and lack of recent connections with the Amazon Forest favored lineage differentiation. Between MS and CR, however, the altitudinal gradient restricts gene flow between these populations. All results were congruent with the recent taxonomic revision. Here we could reconstruct successfully the historic dispersal of a generalist hummingbird species leading to the colonization of distinct environments of two major Neotropical biomes.

Acknowledgements

This research was funded by the Brazilian agencies FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). FRS is supported by a CNPq research fellowship. We thank Leonardo E. Lopes, Anderson V. Chaves, and Marcelo F. Vasconcelos for earlier discussions and fieldwork studies of high-altitude birds, including some samples of *Campylopterus spp.* deposited at CCT-UFMG. We thank Dr. Camila Ribas of INPA/Manaus collection, Dr. Luiz F. Silveira of MZUSP/São Paulo, and Dr. Alexandre Aleixo of MPEG/Belém for the tissue samples of *Campylopterus spp.* We thank Eloisa Helena Reis Sari, Ana Cecília Holler del Prette, and Thiago Araújo Quintão for suggestions.

Conflict of Interest

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Author Contributions

JCPO conceived the study, conducted the experiments, analyzed the data, and wrote the manuscript; GSC made substantial contributions to data analysis, supervised and revised the manuscript; FRS conceived and coordinated the study, obtained funds, discussed the findings, supervised and provided important feedback regarding the scientific significance of data, wrote and revised the manuscript. All authors read and approved the final version.

References

- Anhuf D, Ledru MP, Behling H, Da Cruz Jr FW, Cordeiro RC, Van der Hammen T, Karmann I, Marengo JA, De Oliveira PE, Pessenda L *et al.* (2006) Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeogr Palaeoclimatol Palaeoecol* 239:510–527.
- Antonelli A, Zizka A, Carvalho FA, Scharn R, Bacon CD, Silvestro D and Condamine FL (2018) Amazonia is the primary source of Neotropical biodiversity. *Proc Natl Acad Sci U S A* 115:6034–6039.
- Baker AJ, Haddrath O, McPherson JD and Cloutier A (2014) Genomic support for a moa–tinamou clade and adaptive morphological convergence in flightless ratites. *Mol Biol Evol* 31:1686–1696.
- Batalha-Filho H, Fjeldså J, Fabre P-H and Miyaki CY (2013) Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *J Ornithol* 154:41–50.
- Berv JS, Campagna L, Feo TJ, Castro-Astor I, Ribas CC, Prum RO and Lovette IJ (2021) Genomic phylogeography of the White-crowned Manakin *Pseudopipra pipra* (Aves: Pipridae) illuminates a continental-scale radiation out of the Andes. *Mol Phylogenet Evol* 164:107205.
- Braga PLM, Borges SH, Peres CA, Loisele BA, Blake JG, Menger J, Bueno AS, Anciães M, Abreu FHT, Maximiano MFA *et al.* (2022) Connecting Amazonian historical biogeography and local assemblages of understory birds: Recurrent guild proportionality within areas of endemism. *J Biogeogr* 49:324–338.
- Buzatti RSO, Pfeilsticker TR, de Magalhães RF, Bueno ML, Lemos-Filho JP and Lovato MB (2018) Genetic and historical colonization analyses of an endemic savanna tree, *Qualea grandiflora*, reveal ancient connections between Amazonian savannas and Cerrado core. *Front Plant Sci* 9:981
- Cabanne GS, Calderón L, Trujillo Arias N, Flores P, Pessoa R, d’Horta FM and Miyaki CY (2016) Effects of Pleistocene climate changes on species ranges and evolutionary processes in the Neotropical Atlantic Forest. *Biol J Linn Soc* 119:856–872.
- Cabanne GS, Campagna L, Trujillo-Arias N, Naoki K, Gómez I, Miyaki CY, Santos FR, Dantas GP, Aleixo A, Claramunt S *et al.* (2019) Phylogeographic variation within the Buff-breasted foliage-gleaner (Aves: Furnariidae: *Syndactyla rufosuperciliata*) supports an Andean-Atlantic forests connection via the Cerrado. *Mol Phylogenet Evol* 133:198–213.
- Capurro JM, Ashley MV, Ribas CC and Bates JM (2018) Connecting Amazonian, Cerrado, and Atlantic forest histories: Paraphyly, old divergences, and modern population dynamics in tyrant-manakins (*Neopelma/Tyrannetes*, Aves: Pipridae). *Mol Phylogenet Evol* 127:696–705.
- Cheng H, Sinha A, Cruz FW, Wang X, Edwards RL, d’Horta FM, Ribas CC, Vuille M, Stott LD and Auler AS (2013) Climate change patterns in Amazonia and biodiversity. *Nat Commun* 4:1411.
- Colli GR, Vieira CR and Dianese JC (2020) Biodiversity and conservation of the Cerrado: Recent advances and old challenges. *Biodivers Conserv* 29:1465–1475.
- Cumer T, Pouchon C, Boyer F, Yannic G, Rioux D, Bonin A and Capblancq T (2021) Double-digest RAD-sequencing: Do pre- and post-sequencing protocol parameters impact biological results? *Mol Genet Genom* 296:457–471.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al.* (2011) The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.

- de Freitas EL, Campagna L, Butcher B, Lovette I and Caparroz R (2022) Ecological traits drive genetic structuring in two open-habitat birds from the morphologically cryptic genus *Elaenia* (Aves: Tyrannidae). *J Avian Biol* 2022:e02931.
- Dornas T, Dantas SM, Araújo-Silva LE, Morais F and Aleixo A (2022) Comparative phylogeography of birds across the Tocantins–Araguaia interfluvium reveals a new biogeographic suture in the Amazon Far East. *Front Ecol Evol* 10:826394.
- Eaton DAR and Overcast I (2020) Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36:2592–2594.
- Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620.
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* 5:184–186.
- Gronau I, Hubisz MJ, Gulko B, Danko CG and Siepel A (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet* 43:1031–1034.
- Haffer J (1969) Speciation in Amazonian forest birds: Most species probably originated in forest refuges during dry climatic periods. *Science* 165:131–137.
- Hoon C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartin I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927–931.
- Hopper SD, Lambers H, Silveira FAO and Fiedler PL (2021) OCBIL theory examined: Reassessing evolution, ecology and conservation in the world's ancient, climatically buffered and infertile landscapes. *Biol J Linn Soc* 133:266–296.
- Hughes PD and Gibbard PL (2018) Global glacier dynamics during 100 Ka Pleistocene glacial cycles. *Quat Res* 90:222–243.
- Jombart T and Ahmed I (2011) ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071.
- Kern AK, Akabane TK, Ferreira JQ, Chiessi CM, Willard DA, Ferreira F, Sanders AO, Silva CG, Rigsby C, Cruz FW *et al.* (2023) A 1.8 million year history of Amazon vegetation. *Quat Sci Rev* 299:107867.
- Kozlov AM, Darriba D, Flouri T, Morel B and Stamatakis A (2019) RAXML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453–4455.
- Ledo RM, Colli GR (2017) The historical connections between the Amazon and the Atlantic Forest revisited. *J Biogeogr* 44:2551–2563.
- Lim MCW, Bi K, Witt CC, Graham CH and Dávalos LM (2021) Pervasive genomic signatures of kocal adaptation to altitude across highland specialist Andean hummingbird populations. *J Hered* 112:229–240.
- Lima-Rezende CA, Rocha AV, Couto Júnior AF, Martins ÉS, Vasconcelos V and Caparroz R (2019) Late Pleistocene climatic changes promoted demographic expansion and population reconnection of a Neotropical savanna-adapted bird, *Neothraupis fasciata* (Aves: Thraupidae). *PLoS One* 14:e0212876.
- Lopes LE, Vasconcelos MFD and Gonzaga LP (2017) A cryptic new species of hummingbird of the *Campylopterus largipennis* complex (Aves: Trochilidae). *Zootaxa* 4268:1–33.
- Malinsky M, Trucchi E, Lawson DJ and Falush D (2018) RADpainter and fineRADstructure: Population Inference from RADseq Data. *Mol Biol Evol* 35:1284–1290.
- Marques EQ, Marimon-Junior BH, Marimon BS, Matricardi EAT, Mews HA and Colli GR (2020) Redefining the Cerrado–Amazonia transition: Implications for conservation. *Biodivers Conserv* 29:1501–1517.
- Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D and Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol Ecol Resour* 15:28–41.
- McGuire JA, Witt CC, Remsen JV, Corl A, Rabosky DL, Altshuler DL and Dudley R (2014) Molecular phylogenetics and the diversification of Hummingbirds. *Curr Biol* 24:910–916.
- Medeiros BAS and Farrell BD (2018) Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ* 6:e5089.
- Musher LJ, Ferreira M, Auerbach AL, McKay J and Cracraft J (2019) Why is Amazonia a 'source' of biodiversity? Climate-mediated dispersal and synchronous speciation across the Andes in an avian group (Tityrinae). *Proc R Soc B Biol Sci* 286:20182343.
- Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
- Neves DM, Dexter KG, Pennington RT, Bueno ML and Oliveira Filho AT (2015) Environmental and historical controls of floristic composition across the South American Dry Diagonal. *J Biogeogr* 42:1566–1576.
- Norambuena HV and Van Els P (2021) A general scenario to evaluate evolution of grassland birds in the Neotropics. *Ibis* 163:722–727.
- O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Mol Ecol* 27:1–14.
- Oliveira-Filho AT, Jarenkow J and Rodal MJN (2006) Floristic relationships of Seasonally Dry Forests of eastern South America based on tree species distribution patterns. In: Pennington RT and Ratter JA (eds) *Neotropical Savannas Seasonally dry forests*. CRC Press, Boca Raton, pp 159–192.
- Pacheco JF, Silveira LF, Aleixo A, Agne CE, Bencke GA, Bravo GA, Brito GRR, Cohn-Haft M, Maurício GN, Naka LN *et al.* (2021) Annotated checklist of the birds of Brazil by the Brazilian Ornithological Records Committee – second edition. *Ornithol Res* 29:94–105.
- Pennington RT, Prado DE and Pendry CA (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *J Biogeogr* 27:261–273.
- Peterson BK, Weber JN, Kay EH, Fisher HS and Hoekstra HE (2012) Double Digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in nodel and non-model species. *PLoS ONE* 7:e37135.
- Petkova D, Novembre J and Stephens M (2016) Visualizing spatial population structure with estimated effective migration surfaces. *Nat Genet* 48:94–100.
- Prado DE and Gibbs PE (1993) Patterns of species distributions in the dry seasonal forests of South America. *Ann Mo Bot Gard* 80:902–927.
- Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Rambaut A, Drummond AJ, Xie D, Baele G and Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904.
- Reis LS, Guimarães JT, Souza-Filho PW, Sahoo PK, Figueiredo MM, Souza EB, Giannini TC (2017) Environmental and vegetation changes in southeastern Amazonia during the late Pleistocene and Holocene. *Quat Int* 449:83–105.
- Ribas CC, Aleixo A, Nogueira ACR, Miyaki CY and Cracraft J (2012) A palaeobiogeographic model for biotic diversification

- within Amazonia over the past three million years. *Proc R Soc B Biol Sci* 279:681–689.
- Rocha AV, Cabanne GS, Aleixo A, Silveira LF, Tubaro P and Caparroz R (2020) Pleistocene climatic oscillations associated with landscape heterogeneity of the South American dry diagonal explains the phylogeographic structure of the narrow-billed woodcreeper (*Lepidocolaptes angustirostris*, Dendrocolaptidae). *J Avian Biol* 51:jav.02537.
- Rodríguez-Flores CI, Ornelas JF, Wethington S and Arizmendi MC (2019) Are hummingbirds generalists or specialists? Using network analysis to explore the mechanisms influencing their interaction with nectar resources. *PLoS One* 14:e0211855.
- Sambrook J and Russell DW (2006) Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harb Protoc* 2006:pdb.prot4455.
- Schuchmann K-L (1999) Family Trochilidae (Hummingbirds). In: Hoyo J, Elliott A and Sargatal J (eds). *Handbook of the birds of the world*. Lynx editions, Barcelona, pp 468–680.
- Silva SM, Peterson AT, Carneiro L, Burlamaqui TCT, Ribas CC, Sousa-Neves T, Miranda LS, Fernandes AM, d’Horta FM, Araujo-Silva LE *et al.* (2019) A dynamic continental moisture gradient drove Amazonian bird diversification. *Sci Adv* 5:eaat5752.
- Smeds L, Qvarnström A and Ellegren H (2016) Direct estimate of the rate of germline mutation in a bird. *Genome Res* 26:1211–1218.
- Smith BT, McCormack JE, Cuervo AM, Hickerson Michael J, Aleixo A, Cadena CD, Pérez-Emán J, Burney CW, Xie X, Harvey MG *et al.* (2014) The drivers of tropical speciation. *Nature* 515:406–409.
- Thrasher DJ, Butcher BG, Campagna L, Webster MS and Lovette IJ (2018) Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. *Mol Ecol Resour* 18:953–965.
- Trujillo-Arias N, Dantas GPM, Arbeláez-Cortés E, Naoki K, Gómez MI, Santos FR, Miyaki CY, Aleixo A, Tubaro PL and Cabanne GS (2017) The niche and phylogeography of a passerine reveal the history of biological diversification between the Andean and the Atlantic forests. *Mol Phylogenet Evol* 112:107–121.
- Tziperman E and Gildor H (2003) On the mid-Pleistocene transition to 100-kyr glacial cycles and the asymmetry between glaciation and deglaciation times. *Paleoceanogr Paleoclimatol* 18:1001.
- Vasconcelos MF and D’Angelo S (2018) First avifaunal survey of a Cerrado dry forest enclave on the right bank of the São Francisco River, Minas Gerais, Brazil, with insights on geographic variation of some species. *Pap Avulsos Zool* 58:e20185815.
- Wang X, Edwards RL, Auler AS, Cheng H, Kong X, Wang Y, Cruz FW, Dorale JA and Chiang HW (2017) Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature* 541:204–207.
- Weir JT, Bermingham E and Schluter D (2009) The Great American Biotic Interchange in birds. *Proc Natl Acad Sci U S A* 106:21737–21742.
- Werneck FP (2011) The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. *Quat Sci Rev* 30:1630–1648.
- Werneck FP, Gamble T, Colli GR, Rodrigues MT and Sites Jr JW (2012) Deep diversification and long-term persistence in the South American ‘dry diagonal’: Integrating continent-wide phylogeography and distribution modeling of geckos: deep divergence of South American ‘dry diagonal’ biomes. *Evolution* 66:3014–3034.
- Willeit M, Ganopolski A, Calov R and Brovkin V (2019) Mid-Pleistocene transition in glacial cycles explained by declining CO₂ and regolith removal. *Sci Adv* 5:eaav7337.
- Willis EO (1992) Zoogeographical origins of eastern Brazilian birds. *Ornitol Neotrop* 3:1–15.
- Wright JD (2013) Global climate change in marine stable isotope records. In: Noller JS, Sowers JM and Lettis WR (eds). *American Geophysical Union, Washington, D.C.*, pp 427–433.
- Yamaguchi R (2022) Intermediate dispersal hypothesis of species diversity: New insights. *Ecol Res* 37:301–315.
- Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28:3326–3328.

Internet Resources

- Andrews S (2010) FastQC A quality control tool for high throughput sequence data, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed 2 August 2022)
- BirdLife International (2021) Bird species distribution maps of the world. Version 2021.1, <https://datazone.birdlife.org/species/requestdis/> (accessed 28 November 2022).
- Gill F, Rasmussen P and Donsker D (eds) (2021) *IOC World Bird List 12.1*, <https://www.worldbirdnames.org/new/> (accessed 28 November 2022).
- WikiAves (2022) *Asa-de-sabre-da-mata-seca (Campylopterus calcirupicola)*, <https://www.wikiaves.com.br/asa-de-sabre-da-mata-seca/> (accessed 28 November 2022).

Supplementary material

- The following online material is available for this article:
- Figure S1 – Comparative results of ddRAD runs.
- Figure S2 – Map with Structure results represented as individual pie charts reflecting the geographical samples distribution.
- Figure S3 – Paired principal component analysis. Scatterplots showing the top four PCs.
- Figure S4 – Hierarchical clustering analysis based on identity-by-state (IBS) matrix from SNP data representing the genetic relationships.
- Table S1 – Voucher specimens and location details of the samples belonging to the *Campylopterus largipennis* complex.
- Table S2 – Pairwise FST values sampled localities.
- Table S3 – Population parameters estimated from three independent runs of G-Phocs analysis.
- Table S4 – Migration bands from G-Phocs analysis.

Associate Editor: Carlos F. M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.