

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)

A. V. Rocha¹, G. S. Cabanne², A. Aleixo^{3,5}, L. F. Silveira⁴, P. Tubaro² and R. Caparroz¹

¹Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Brazil

²Museo Argentino de Ciencias Naturales MACN - CONICET, Buenos Aires, Argentina

³Coordenação de Zoologia, Museu Paraense Emílio Goeldi, Belém, São Paulo, Brazil

⁴Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil

⁵Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

Corresponding author: Renato Caparroz, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Brazil. E-mail: renatocz@yahoo.com.br

Decision date: 28-Jul-2020

This article has been accepted for publication and 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/jav.02537].

Abstract.

Based on a phylogeographic and niche model analyses of the Narrow-billed Woodcreeper (*Lepidocolaptes angustirostris*), we evaluated the predictions of two diversification hypotheses related to the dry diagonal of South America: (I) isolation by distance (IBD) and (II) landscape heterogeneity. We also investigated the influence of the Pleistocene climatic oscillations on the diversification and population dynamic of this species, and discussed the implications of our findings for the taxonomy of this woodcreeper. We sampled 63 individuals including all subspecies described for *L. angustirostris*, and compared them using a mitochondrial (ND2) and a nuclear (FIB5) fragments. We performed a Mantel test and spatial autocorrelation analysis, reconstructed the phylogenetic relationships among haplotypes, investigated changes in population size and estimated divergence time among the genetic lineages. We refuted the hypothesis that bird species associated with open environments have less geographic differentiation and showed that strong genetic structure observed in *L. angustirostris* can not be explained by IBD. Paleo-modeling showed strong association between climatic stable areas and the genetic lineages suggesting that Pleistocene climatic oscillations have driven primarily the intraspecific diversification of this species. Secondly, the association between genetic lineages and the dry diagonal biomes, supporting that landscape heterogeneity may be acting as a secondary barrier restricting gene flow among the lineages. The genetic lineages found in *L. angustirostris* do not correspond to the subspecies described, indicating that this species can be considered as a single species with strong genetic structure. In conclusion, our data corroborate other studies indicating that Pleistocene climatic oscillations may have had a strong influence in the intraspecific divergence of dry diagonal fauna and each biome of the dry diagonal should be considered as independent units in further biogeographic studies.

Keywords: Biogeography, Caatinga, Cerrado, Chaco, Isolation by distance, STDFs

Introduction

The Neotropical region is the most biodiverse in the World (Rull 2008). Many models and processes have been invoked to explain the evolution of biodiversity in this region, most of them mainly focused on tropical humid forests (e.g. the hypothesis of Pleistocene refugia; Haffer 1969, Carnaval 2009). Although seasonally deciduous forests and open vegetation formations in South America form the largest continuous portion of dry vegetation in the continent (hereafter dry diagonal, Fig. 1A), and occupy almost a third of the region, agents and processes promoting biological diversification in this diagonal are still poorly known (Werneck 2011, Turchetto-Zolet et al. 2013).

The dry diagonal is formed by a complex mosaic of vegetation that apparently evolved under strong pressure of dry periods and marked seasonality (Furley and Metcalfe 2007), including biomes such as the Seasonally Dry Tropical Forests (SDTFs), the Cerrado (tropical savanna of central Brazil) and the dry Chaco (xerophytic woodlands of northern Argentina, western Paraguay, and southeastern Bolivia) (Fig. 1A). Unlike other biomes of the dry diagonal, the SDTFs exhibit a highly disjunct pattern and is divided by four main nuclei: (I) Caatinga, (II) Misiones, (III) Subandean Piedmont, and (IV) Chiquitano (Fig. 1A) (for details, see Linares-Palomino et al. 2011).

Biological diversification in the dry diagonal has been associated to a combination of events that occurred during the Tertiary (e.g. marine transgressions and the uplift of the Central Brazil Plateau), as well as during the Quaternary (e.g. the Pleistocene climatic fluctuations; Rull 2008, Werneck 2011). In general, the geological events have been associated with diversification of high taxonomic levels (e.g. Squamata genera, Giugliano et al. 2007), while Quaternary climatic oscillations have been associated to intraspecific differentiation (e.g. in plants, Ramos et al. 2007, Caetano et al. 2008; and insects, Moraes et al. 2009). Even though birds are the most diverse group of land vertebrates in the Neotropical region (Stotz et al. 1996), few studies have focused on their intraspecific evolution in the dry diagonal (Leite et al.

2009, Lima-Rezende et al. 2019, Savit and Bates 2015, Moura et al. 2018), mostly because early study found very low infraspecific differentiation in open area birds (Bates et al. 2003).

Birds inhabiting open environments tend to have high dispersion capacity, which would lead to high gene flow among populations and, therefore, less geographic differentiation (e.g. Bates et al. 2003, Leite et al. 2009). However, species with large geographical distribution, such as many inhabitants of the dry diagonal, may also have limited gene flow across their ranges due to isolation by distance (Wright 1946, Holderegger and Wagner 2006, Van Strien et al. 2015). Isolation by distance may occur within a population with continuous distribution when dispersion capacity decreases with increasing geographical distance (Hardy and Vekemans 1999). Thus, a prediction of isolation by distance is to find a strong correlation between genetic differences between populations and geographical distance.

Even though bird species are supposed to have high dispersal capacity, some open area birds may have restricted gene flow due to landscape heterogeneity, contributing to genetic structuring of populations (Manel et al. 2003). Although the open biomes of South America form a long and continuous diagonal, these biomes differ in many biotic and abiotic aspects (see Werneck 2011), which could convert the dry diagonal in a heterogeneous landscape for many of their organisms. This heterogeneity could have partitioned the region and thus promoted genetic differentiation among biomes (e.g., between Cerrado and Caatinga). Some phylogeographic studies of lizards (Fonseca et al. 2018) and birds (Moura et al. 2018) of the dry diagonal have found different genetic lineages associated with each biome, supporting the aforementioned idea. Even when landscape heterogeneity does not act as a primary force for speciation, it can contribute by acting as a secondary barrier to gene flow (Manel et al. 2003).

Accepted Article

According to the Pleistocene refugia hypothesis, contraction and fragmentation of habitat during glacial cycles promoted the formation of climatic refugia, leading to allopatric speciation in these refugia (Whitmore and Prance 1987). The formation of Pleistocene climatic refugia has been invoked to explain the genetic structure of some Neotropical forest species, like mammals (e.g. Carnaval et al. 2009), frogs (e.g. Kosciński et al. 2008) and birds (Batalha-Filho et al. 2012, Silva et al. 2019, Thom et al. 2020). However, there is controversy on the impact of Quaternary glacial cycles on the population dynamics and evolution of open area Neotropical organisms (Rull 2008, Werneck 2011). For instance, models of South American biomes suggested that the STDFs have been retracted and fragmented during the last Glacial Maximum (i.e. 21 kyr ago, Costa et al. 2017), while some authors have hypothesized that these STDFs have been connected during the Last Glacial Maximum, being contracted and fragmented during the current interglacial period (Prado and Gibbs 1993, Pennington et al. 2000). If habitat fragmentation associated to Global climate cycles impacted on evolution of open area organisms, we would expect to find an association between intraspecific lineages of these organisms and regions with long term habitat stability.

In this paper, we studied the Narrow-billed Woodcreeper (*Lepidocolaptes angustirostris*) as a model to investigate evolutionary processes in the South American dry diagonal. This passerine is broadly distributed (Fig. 1A), having high abundance across most of the dry diagonal. There are eight subspecies described for this woodcreeper which are separated into two groups based on plumage pattern: (I) group *angustirostris*, with brown back and heavily streaked pattern in the chest and vent (*L. a. angustirostris*, *L. a. praedatus*, *L. a. certhiolus* and *L. a. hellmayri*), and (II) group *bivittatus*, with more rufescent in the back and unstreaked pattern below (*L. a. bivittatus*, *L. a. griseiceps*, *L. a. coronatus* and *L. a. bahiae*) (Marantz et al. 2003). Up to now there has not been any genetic analysis to address if these subspecies and plumage groups represent independent evolutionary lineages.

Accepted Article

Here, we describe the phylogeographic pattern of *L. angustirostris* and evaluate hypotheses on the evolution of the dry diagonal. First, we addressed whether *L. angustirostris* is formed by a single panmictic population (absence or low genetic structure), or whether there is evidence of isolation by distance along the dry diagonal. After having rejected these previous hypotheses, we evaluated whether landscape heterogeneity promoted by differences among the biomes (Caatinga, Cerrado and Chaco) can explain the phylogeographic structure of this woodcreeper. We also applied paleo-distribution modeling to investigate whether the distribution of the climatic stable areas is concordant with distribution of the *L. angustirostris* genetic lineages and whether Pleistocene climatic fluctuations influenced its demographic history. Finally, we discussed the implications of our results in light of the taxonomy of this woodcreeper.

Material and methods

Sampling and genetic data

We sampled 63 individuals including all subspecies described for *L. angustirostris* in 37 localities (one to four individuals per locality) distributed along of the dry diagonal (Fig. 1A, Supplementary material Appendix 1 Table S1). Some tissue samples were collected in the field by AVR, RC, LFS and colleagues, while others were kindly provided as loans from Museo Argentino de Ciencias Naturales (MACN), Museu Paraense Emilio Goeldi (MPEG) and Burke Museum of Natural History and Culture (UWBM).

Total DNA was extracted from muscle and blood tissues using the commercial Quiagen DNAeasy kit following manufacturer's protocols. We amplified a mitochondrial fragment of the NADH dehydrogenase gene subunit 2 (ND2) using LMET (Groth, J. *apud* Ribas et al. 2005) and H6313 (Sorenson et al. 1999) primers and a nuclear fragment of the intron 5 of the beta-fibrinogen gene (FIB5) using FIB5 e FIB6 primers (Marini and Hackett 2002). PCR products were enzymatically purified using

Exo/SAP. The sequences were obtained either by Macrogen Inc., Unidad de Genómica at Argentina or Centro de Genômica de Alto Desempenho at Universidade Católica de Brasília facilities. We evaluated the sequences for quality and ambiguities using Geneious 6.1.6 (Biomatters). The alignment was made by ClustalW algorithm (Thompson et al. 1997). To infer the haplotype phase of nuclear locus, we use the Bayesian algorithm Phase (Stephens et al. 2001) implemented in DnaSP (Rozas et al. 2003). We choose the best evolution model for each locus based on BIC scores using JModelTest 2 (Darriba et al. 2012).

Panmixia and isolation by distance

The presence of panmixia among *L. angustirostris* individuals was verified based on the global Φ_{ST} value for each loci using Arlequin 3.0 (Excoffier et al. 2005). For the latter, each locality with ≥ 2 samples was considered as a subpopulation. To test isolation by distance as a driven of *L. angustirostris* diversification, we carried out a matrixes correlation analysis between the genetic (Φ_{st}) and geographic (Km) distances by Mantel test with 10,000 iterations implemented in Arlequin. As we identified some genetic populations (see results), we performed this analysis for all individuals as well as within each population. We also performed a spatial autocorrelation analysis to identify the distance classes where the correlation of Mantel is significant (Sokal et al. 1997). We conduct this analysis using Alleles in Space (AIS; Miller 2005) based on average genetic distance between pairs of individuals that fall within distance class y , considering 10 classes and 1,000 permutations. To avoid the bias of the low variability of the FIB5, we conducted this analysis only for the ND2.

Phylogenetic analysis and genetic structure

The relationship among the haplotypes were investigated by constructing a median-joining network (Bandelt et al. 1999) using PopART 1.7.2 (French et al. 2013). We also

Accepted Article

evaluated the phylogenetic relationship among the haplotypes based on a phylogenetic tree built with Bayesian inference (BI) analysis using a concatenated matrix of the ND2 (HKY+G substitution model) and FIB5 (JC69 substitution model). The BI was performed with coalescent model employed on BEAST 1.8 (Drummond and Rambaut 2007). We also performed a species tree analysis on STARBeast (Heled and Drummond 2010) using coalescent model and constant population size. For both analyses, we used the IUPAC systems and considered the ambiguities as information. The MCMC analyses were performed with 100 million generations sampled every 1,000 generations. We evaluated the run quality using Tracer 1.6 (Rambaut et al. 2014) by the ESS values (> 200). Trees were summarized in TreeAnnotator considering a burn-in of 10%. The topologies were visualized and edited in Figtree v. 1.4.2 (Rambaut and Drummond 2012). Then, all genetic lineages were characterized based on the number of haplotypes, haplotype diversity, nucleotide diversity and nucleotide differences for both molecular markers using DnaSP. We also estimated a pairwise genetic distance (Φ_{ST}) among populations to identify the most divergent populations using Arlequin.

Historical demography

To evaluate temporal variation in effective population size, we performed the following statistical tests: R_2 statistic (Ramos-Onsins and Rozas 2002), Tajima's D (Tajima 1989) and Fu's F_S (Fu 1997) for each genetic lineage in DnaSP. We also evaluated the dynamic of effective population size through time by extended Bayesian skyline plot (EBSP) model for each genetic lineage (Heled and Drummond 2008). We used 0.009 substitution/site/million years (Pacheco et al. 2011) and 0.0036 substitution/site/million years (Axelsson et al. 2004) as the substitution rate for ND2 and FIB5, respectively. We set 100 million runs and the parameters were saved each 1,000 runs. We evaluated the run quality using Tracer 1.6 by the ESS values (> 200). The graphic result is a

Accepted Article
correlation of median effective population size through the time in years and lower and upper 95% HPD values as confidence intervals for each population done in R software.

Divergence time

In order to estimate divergence time among genetic lineages, we performed an isolation with migration analysis assuming the population relationship (((E_Cerrado/Chaco) W_Cerrado) NE_Cerrado) Caatinga) obtained by species tree analysis (see Fig. 3), where each ancestral population were identical in size to their largest descendant population, and only one single migration parameter for all pairs of population (to simplify the model) as implemented in IMA2 (<https://bio.cst.temple.edu/~hey/software>). We used a floating-point run duration (the analysis ran until 22 million generations), one million generations as a burn-in, 15 as maximum time of population split, 60 as migration prior value, and 50 as maximum for population size. We used a geometric heating model, five chains and 0.9 for first and second heating parameters. We assumed one year as generation time and mutation rates' geometric mean to estimate values in demographic units.

Niche modeling and stable areas

To evaluate the influence of the climatic niche dynamic on *L. angustirostris* populations and predict areas of historical persistence of the species, we performed a paleo-distribution climatic niche analysis using 106 occurrence points (Fig. 4A, Supplementary material Appendix 1 Table S2) obtained from our group fieldworks, museums occurrence data, literature revision and Global Biodiversity Information Facility – GIBF (available on www.gbif.org). We created a grid with the same resolution of climatic data ($2.5^{\circ} \times 2.5^{\circ}$) and, to reduce the bias of occurrence matrix on the analysis, we choose only one occurrence point for each grid cell.

Accepted Article

Firstly, we evaluated niche models for the present period considering a) all climate variables, b) excluding the most correlated variables, and c) by selecting the most important variables according to the criterion of percent contribution of each variable (keeping variables with > 5%). These three approaches yielded similar models, but only using the full set of variables (AUC = 0.879) modeled the savannas of Amapá and other northern isolated populations as a climatically suitable areas. Thus, the discussion of the results was based on the scenarios using the full set of variables. Then, we projected the distribution at three time periods: Last Inter-Glacial (LIG, ~150 to 120 kyr BP), Last Glacial Maximum (LGM, 21 kyr BP) and Mid-Holocene (6 kyr BP) using the best scenario. For LGM and Mid-Holocene, two atmosphere-ocean circular models (CCSM4 and MIROC) were used and both projections were combined using the Maximum entropy algorithm in the MaxEnt (Phillips et al. 2010). We choose the 10-percentile training presence as threshold. The 19 bioclimatic variables (BIO1 – BIO19) were obtained from WorldClim database (Hijmans et al. 2005). Climatic stable areas were estimated overlapping the current potential distribution and the other paleo-projections (LIG, LGM and Mid-Holocene) using raster calculator tool in ArcMap 10.1. Areas that were suitable in all projections were defined as climatic stable ones for *L. angustirostris*.

Results

We collected a genetic data set that consisted in a final alignment of 1,449 bp (ND2, 978 bp; FIB5, 471 bp). The genetic variability of ND2 was higher than of FIB5 (Table 1). We found low (FIB5 $\Phi_{ST} = 0.09$, $p < 0.05$) to moderate genetic structure (ND2 $\Phi_{ST} = 0.46$, $p < 0.05$) considering all individuals as a single population, indicating that there is no random mating among them.

We found correlation between genetic and geographic distance considering all individuals ($r = 0.22$, $p = 0.000$), while no evidence of correlation was observed within

each genetic lineage, except for E_Cerrado ($r = 0.32$, $p = 0.009$) (Table 2). However, the spatial autocorrelation analysis for ND2 shows that IBD may occurred only in distances fewer than 1,200 km ($p = 0.003$, Supplementary material Appendix 1 Fig. S1).

Genetic lineages and genetic structure

We identified five haplogroups in the ND2 network: one group in Caatinga, three in Cerrado (NE_Cerrado, E_Cerrado and W_Cerrado) and one in Chaco (Fig. 1B).

Cerrado's haplogroups occur on northeastern (NE), eastern (E) and western (W) of biome, respectively (Fig. 1). A star-like pattern suggesting recent population expansion is clear on Chaco population. In contrast, FIB5 network shows one frequently and widely distributed haplotype (H_01), and few less frequently tip ones (Fig. 1B).

While IB (Fig. 2) supported the five haplogroups described above with high posterior probabilities values, excepted for some Cerrado's populations, species tree analysis (Fig. 3) showed low supported values for almost all nodes. Based on species tree analysis, the Caatinga populations were the first to diverge, followed by divergence of NE_Cerrado, W_Cerrado, Chaco and E_Cerrado. Basically, we found each genetic lineage restricted to one of the dry diagonal biomes, but not found unique lineages for the individuals from STDFs *nuclei* (Fig. 2), except for Caatinga. Instead, some samples from Misiones and Piedmont were recovered in Chaco lineage and samples from Chiquitano were recovered in W_Cerrado lineage.

Considering the five genetic lineages described above, the population genetic structure of *L. angustirostris* is strong (ND2: $\Phi_{st} = 0.76$, $p < 0.05$; FIB5: $\Phi_{st} = 0.62$, $p < 0.05$). The highest Φ_{st} values were observed between Caatinga and Chaco, while the lowest values were found when comparing E_Cerrado and W_Cerrado, and E_Cerrado and Chaco (Supplementary material Appendix 1 Table S3).

Demographic processes and divergence time

'This article is protected by copyright. All rights reserved.'

The demographic history recovered by Fu's F_S , Tajima's D and R_2 tests showed clear signals of recently range expansion only in Chaco lineage, while W_Cerrado showed range expansion signature only in R_2 and Fu's F_S tests based on ND2 (Table 2). A similar pattern was also observed by the EBSP analysis where Chaco lineage showed signal of demographic size shifts through time (Supplementary material Appendix 1 Fig. S3), while in W_Cerrado the sign of demographic size shifts was less evident.

All intraspecific diversification events of *L. angustirostris* likely occurred during the Pleistocene (Fig. 3). However, probably due to the low variability of genetic markers, the error associated with these estimates is very wide and does not rule out that some diversification events may have occurred in the Pliocene. The first diversification, separating Caatinga from the other lineages, occurred around 309 (180,580 - 3,881,490) kya, followed by divergence of NE_Cerrado around 254 kya (87,337 - 2,926,166), and W_Cerrado around 138 kya (33,009 - 1,776,670). Finally, the more recent diversification occurred between Chaco and E_Cerrado dated around 76 (13,592 - 739,794) kya.

Potential distribution model and stable areas

The projection for the LIG (~120 kya) suggests a period with lowest and fragmented suitable areas for *L. angustirostris* (Fig. 4D). Chaco was the only suitable patch closer to current distribution area at the LIG. It's possible to identify an expansion of suitable areas, and signals of patch connections at the LGM (21 kya; Fig. 4C). This expansion is more evident on Cerrado. Suitable area is slightly more extensive than the current geographical distribution of the species at the Middle Holocene (6 kya; Fig. 4B).

Summarizing, the suitable areas for *L. angustirostris* were smaller and more fragmented in the LIG, with the exception of Chaco, which is similar to its current distribution.

Since LIG, the amount of suitable habitat expanded, particularly in Cerrado. This expansion led to the reconnection of the fragments within the biome and establishing

wide contact with the other two biomes. In contrast, the suitable areas in the northeast did not increase as much and remained more isolated.

The major climatic stable area identified was within the current Chaco/Misiones/Piedmont region (Fig. 4E). We also identified four main patches located on western, middle, eastern and northeastern Cerrado and one fragmented patch on Caatinga. In general, we identify a geographical correspondence among the stable areas and the genetic lineages identified for *L. angustirostris* (Fig. 4E).

Discussion

Our phylogeographic study of *L. angustirostris* did not find evidence for the hypothesis stating that open environment species are able to disperse over long distances, and thus to have high gene flow among populations and less geographic genetic differentiation.

In contrast, our results indicated that *L. angustirostris* has strong genetic structure along the dry diagonal ($\Phi_{st} > 0.62$), having distinct lineages associated to each biome.

Lineage divergences were likely influenced by events occurring during the Pliocene/Pleistocene, perhaps related to climatic oscillations and landscape heterogeneity promoted by differences among the dry diagonal biomes.

Isolation by distance

We found evidence of isolation by distance in *L. angustirostris* only considering all individuals as a single population or within E_Cerrado lineage. However, in both cases, the Mantel test correlations may to be biased by the presence of samples with more than 1,200 km apart from each other (see Fig. 1), the threshold observed in the spatial autocorrelation analysis (Supplementary material Appendix 1 Fig. S1). Furthermore, significant Mantel correlation can also be observed due to the presence of genetic subpopulations (see Diniz-Filho et al. 2013), which was observed in this study, as

discussed below. This result agrees with that obtained for another open area bird species, the Rufous-collared Sparrow (*Zonotrichia capensis*, Campagna et al. 2014).

Lack of isolation by distance within the *L. angustirostris* lineages could have the following non-mutually exclusive explanations, such as: (I) the species capability of dispersion could be higher than expected; (II) the timescale for equilibrium between gene flow and genetic drift could have not been achieved due to recent population expansion (Castric and Bernatchez 2003, Pinsky et al. 2010); and (III) the heterogeneity landscape can restrict the dispersal of individuals irrespective of their dispersal capacity. Although these hypotheses have to be further investigated, our data support that the third one can be a good explanation.

Climatic oscillations and stable areas

The strong association between stable areas and genetic lineages (see Fig. 4E) suggests that climatic oscillations have been a primary force driving intraspecific diversification of *L. angustirostris*. Stable areas may have maintained isolated populations during the warm and wet interglacial periods, which expanded as the climate moved towards the colder and drier glacial periods, reestablishing contact and eventual gene flow between previous isolated lineages. However, establishing gene flow seems to have been more intense between Chaco and E_Cerrado, than between Cerrado and Caatinga lineages, as the latter was the first to diverge (see Figs. 2 and 3) and showed greater genetic divergence in relation to the other lineages (Supplementary material Appendix 1 Table S3), it seems have remained more isolated along the climatic oscillations. Niche modeling supported this historical isolation (see Fig. 4).

Our results are in accordance with other studies that found evidence of retraction of South American open biomes during warmer interglacial periods, and expansion during cold and dry glacial periods (Bueno et al. 2017, Lima-Rezende et al. 2019). As consequence of this spatio-temporal dynamic of suitable areas, it can be expected that

individuals evolved in isolation under different climatic and ecological pressures, and action of the genetic drift, accumulated differences that can limit gene flow between these lineages at the limits of respectively suitable areas after niche and population expansions (Hewitt 2004).

Interestingly, the Chaco lineage was the only one with signs of demographic expansion in the last 120,000 years (Supplementary material Appendix 1 Fig. S3), while distribution modeling indicated habitat stability in this region. Perhaps the stable population size observed in the other lineages may be due to either a non-proportional effect between the expansion of the niche and populations (for details, see Crick 2004), or the low geographic sampling coverage, in the case of NE_Cerrado lineage, or the low sampling size for the other lineages. However, these hypotheses should be further investigated.

The paleodistribution models were not in accordance to the idea that STDFs expanded during the Last Glacial Maximum, since the potential habitat of *L. angustirostris* populations inhabiting the STDFs were more disjunct during the LIG than the present-day (Fig. 4D). Our results also indicated that Chiquitano, Piedmont and Misiones *nuclei* were likely recently colonized by *L. angustirostris* coming from the neighboring biomes, likely favored by the habitat connectivity established among these different open formations. For instance, individuals sampled in the Chiquitano (see LOC 25 in Fig. 2) were included in the W_Cerrado lineage, while some individuals sampled in the Misiones and Piedmont (see LOC 24 and 26 in Fig. 2) were included in the Chaco lineage. Evidence supporting a recent connection between Chiquitano and Cerrado regions are the pollen record data indicated that the final establishment of Cerrado vegetation distribution took place in the Mid-Holocene (Ledru 2002), and the potential expansion of STDFs regions during this period indicated by paleodistribution modelling (Werneck et al. 2011). The genetic link between the woodcreeper populations

from Chaco and Misiones are also supported by our paleodistribution models, suggesting that these regions have been connected along the last 120.000 years.

As most of the genetic variation analyzed in this study comes from the mitochondrial genome, the strong genetic structure observed in *L. angustirostris* may have been achieved by male-biased gene flow, as observed in other bird species (e.g. Gibbs et al. 2000). Both datasets analyzed here could be related, but the lower resolution of FIB5 does not allow clearly identify a geographical pattern. Furthermore, although studies focusing on the dispersion pattern in woodcreepers are practically nonexistent, genetic studies combining uniparental and biparental molecular markers have found that the incongruence between these markers can be explained as a consequence of the difference in the effective population size between these markers in other woodcreeper (Cabanne et al., 2008).

Landscape heterogeneity

The genetic lineages identified in *L. angustirostris* are spatially associated with the dry diagonal biomes, which suggest that landscape heterogeneity may be an important element for the intraspecific variation in this woodcreeper. Our findings agree with recent studies of lizards (*Vanzosaura rubricauda*, Recorder et al. 2014; *Polychrus acutirostris*, Fonseca et al. 2018) and birds (*Polioptila plumbea*, Moura et al. 2018) in which taxa widely distributed along the dry diagonal are composed by distinct lineages (or cryptic species). Furthermore, all these studies found endemic lineages to Cerrado and Caatinga, with restriction of gene flow between these two biomes. Likewise, ecotone regions seem to be an important barrier to gene flow for *L. angustirostris*, mainly between Cerrado and Caatinga. The lineage of Caatinga and two of Cerrado (NE_Cerrado and E_Cerrado) have a very narrow contact area in the northeastern Brazil, coinciding with the ecotone between these two biomes (see Fig. 1A). A similar pattern was also observed for a spider species (*Sicarius cariri*), when populations highly

structured were found in the northwestern Brazil (Magalhães et al. 2014). Although the genetic break identified between the Cerrado and Chaco lineages covers the ecotone between these biomes, our sampling does not allow us to establish a clear limit for this barrier (see Fig. 1A).

Gene flow among *L. angustirostris* lineages is likely to occur within and between the Cerrado and Chaco, because mitochondrial haplotype sharing was observed between the E_Cerrado and W_Cerrado (LOC20) and between E_Cerrado and Chaco (LOC27) lineages (see Figs. 1 and 2). The gene flow between these lineages was geographically restricted at the border of the lineage ranges, suggesting that the barriers to gene flow between these biomes are efficient, but less effective than those between the lineages from Cerrado and Caatinga. It is not possible to identify an evident geographical barrier between the lineage of Cerrado; however, it may be associated with the landscape compartmentalization of Cerrado, as observed in a tanager (*Neothraupis fasciata*, Lima-Rezende et al. 2019) and a lizard (*Micrablepharus atticolus*, Santos et al. 2014) species, which occur in this biome. The gene flow between the E_Cerrado and Chaco lineages was observed in an STDF patch (Serra da Bodoquena, LOC 27, Figs. 1 and 2), suggesting that these forests may have an important role in the dispersal of birds among these biomes (see below).

E_Cerrado was the lineage that had the greatest geographical distribution, including samples from northeastern Argentina (Entre Rios and Buenos Aires Provinces). These samples formed a well-supported clade in the IB tree (LOC 33, 36 and 37, Fig. 2), suggesting that this region may harbor a differentiated genetic group within this lineage. A recent local study showed that *L. angustirostris* from the Province of Buenos Aires represent a different genetic population compared to other individuals from the Chaco (Kopuchian et al. 2020). In this study, the authors attribute the Paraná-Paraguay River palaeochannel and the limit between the Espinal-Chaco ecoregions as a potential barriers to gene flow between these two populations. The limited sampling of

eastern South America does not allow our data to help clarify this issue, however it indicates that barriers to gene flow are much more efficient in restricting contact in the ecotone between Espinal – Chaco ecoregions than between the south (Buenos Aires Province) and north (Cerrado) of the distribution of this lineage.

Finally, evidence of historical gene flow (contact) was also observed between W_Cerrado and the disjunct savannas of Amapá (LOC 01, Figs. 1 and 2). Several paleo-corridors of fauna (Andean, Atlantic coastal, central Amazonian and along the Madeira River) have been proposed to explain the connection between Cerrado and the savannas of northern South America (Silva and Bates 2002, Ribeiro et al. 2016). Although our data do not allow to identify which corridor is the most likely, the connection with the W_Cerrado lineage suggests that the corridor across the Atlantic Coast is the least likely to explain the connection between the Cerrado and the savannas of Amapá for *L. angustirostris*. However, this issue can be further investigated.

Taxonomic implications

Our results indicate that neither subspecies of *L. angustirostris* is monophyletic (Fig. 2). In addition, neither genetic lineages of this species have had morphological diagnosis, as they do not match subspecies, as demonstrated by Bolívar-Leguizamón and Silveira (2015). Even NE_Cerrado lineage, which is formed entirely by samples from subspecies *L. a. coronatus*, cannot be considered a distinctive taxonomic unit, since samples assigned to this subspecies were also found in Caatinga lineage (see Fig. 2).

Considering that each plumage group is associated to either Caatinga and northeastern Cerrado (group *bivittatus*) or to the Chaco, southwestern Cerrado and neighboring regions (group *angustirostris*) (Fig. 1), would be expected that each plumage group represent an independent evolutionary lineage. However, our results showed that plumage color groups described by Marantz et al. (2003) did not also form clear monophyletic clades (Fig. 2). Thus, our results support to consider *L. angustirostris* as a

single species, but with strong genetic structure and plumage polymorphism, which is in agreement with a recent review of the plumage variation in the species (Bolívar-Leguizamón and Silveira 2015).

Plumage color variation in this species may be due to niche heterogeneity among regions, which would promote disruptive selection and plumage divergence in the presence of gene flow (see Galleotti et al. 2003). It is worth calling attention to the fact that despite *L. a. griseiceps* not being monophyletic, the individuals of this subspecies are unstreaked and form a well-supported clade within the W_Cerrado lineage (see Fig. 2), which is formed mainly by streaked individuals.

Finally, a song analysis would be useful to investigate the divergence among the lineages described in this study, as well as to evaluate whether those lineages represent organismic divergence at levels higher than the neutral genetic structure. For instance, a molecular study with *Zimmerius* genus shows plumage as a conservative characteristic and song as a better divergent signal (Rheindt et al. 2008). Thus, this pattern could also be true for *L. angustirostris*.

Concluding remarks

We refuted the hypothesis that bird species associated with open environments have less geographic differentiation. In contrast, all genetic lineages identified here are historically associated with one stable climatic area suggesting that diversification within *L. angustirostris* has been primarily influenced by Pleistocene climatic oscillation, which promoted allopatric diversification during interglacial periods. Secondly, the heterogeneous landscape along the dry diagonal may have been limiting gene flow among the genetic lineages after contact was re-established among them. Thus, our results are in agreement with other studies (Lima-Rezende et al. 2019, Silva et al. 2019) indicating that Pleistocene climatic oscillations have had a strong influence in the intraspecific divergence of dry diagonal avifauna. In addition, our

Accepted Article

results reinforce the initial proposition of Pennington et al. (2000), then stressed by others (e.g. Werneck et al. 2012, Fouquet et al. 2014), that the different biomes of the dry diagonal should be considered as independent units of the biogeographical point of view. Based on that, we recommend that next studies focusing on species with wide distributions in the dry diagonal to consider the intrinsic processes of each biome.

Acknowledgments – We are grateful to Laboratório de Genética e Biodiversidade/UnB team, especially Lilian G. Giugliano for contribution to research design and Cássia A. Lima-Rezende, for help during field and laboratory work. We also thank the owners and staff of all sampled locations for assistance during fieldwork. We thank the curators and staff of the institutions who kindly received us and allowed us to access the tissue samples and specimens under their care. We also thank to Drs. Wilson Lemos de Moraes Neto (Fazenda Fartura) and Bruno Ehlers (UPS) for support in laboratory and field activities. We thank MPEG, MZUSP, MACN and UBMW curators for providing tissue samples. We also thank the editor and two anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

Data availability

DNA sequences available from GenBank: MG459021 - MG459085 (ND2) and MG459086 - MG459150 (FIB5)

Funding – This study was supported by CNPq-Universal (445025/2014-0), CONICET (Arg), MINCyT (Arg), and was also part of the GENPAC (564036/2010-2) and PPBio (457444/2012-6) projects, supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Ministério da Ciência, Tecnologia e Inovação (MCTI) / Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). A.V.R

received scholarship from CAPES and CNPq. L.F.S. received grants from CNPq (302291/2015-6 and 308337/2019-0).

Conflict of interest – The authors declare that they have no competing interests.

Author contributions – All authors conceived the study. AVR, GSC and RC collected and analysed the data. All authors contributed to writing the paper.

Permits – All work was conducted under the Comitê de Ética da Universidade de Brasília approval (UnBDoc n° 75111/2013) and research permits were issued by the Brazilian Environmental Agency Instituto Chico Mendes de Conservação (ICMBio) (SISBIO n° 27682-1).

References

- Axelsson, E., Smith, N. G. C., Sundström, H., Berlin, S. and Ellegren, H. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* 21: 1538–1547.
- Bandelt, H. J., Forster, P. and Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16: 37–48.
- Batalha-Filho, H., Fjeldså, J., Fabre, P. H. and Miyaki, C. Y. 2012. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *J. Ornithol.* 154: 41-50.
- Bates, J. M., Tello, J. G. and Silva, J. M. C. 2003. Initial assessment of genetic diversity in ten bird species of South American Cerrado. *Stud. Neotrop. Fauna Environ.* 38: 87-94.

BirdLife International. 2020. Species factsheet: *Lepidocolaptes angustirostris*.

Downloaded from <http://www.birdlife.org> on 23/03/2015.

- Bolívar-Leguizamón, S. and Silveira, L. F. 2015. Morphological variation and taxonomy of *Lepidocolaptes angustirostris* (Vieillot, 1818) (Passeriformes: Dendrocolaptidae). *Papéis avulsos de Zoologia* 55: 281–316.
- Cabanne, G. S., d'Horta, F. M., Sari, E. H. R., Santos, F. R. and Miyaki, C. Y. 2008. Nuclear and mitochondrial phylogeography of the Atlantic forest endemic *Xiphorhynchus fuscus* (Aves: Dendrocolaptidae): biogeography and systematics implications. *Mol. Phylogenet. Evol.* 49: 760-773.
- Caetano, S., Prado, D., Pennington, R. T., Beck, S., Oliveira-Filho, A., Spichiger, R. and Naciri, Y. 2008. The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Mol. Ecol.* 17: 3147–3159.
- Campagna, L., Kopuchian, C., Tubaro, P. L. and Lougheed, S. C. 2014. Secondary contact followed by gene flow between divergent mitochondrial lineages of a widespread Neotropical songbird (*Zonotrichia capensis*). *Biol. J. Linn. Soc.* 111: 863–868.
- Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T. and Moritz, C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science* 323: 785–789.
- Castric, V. and Bernatchez, L. 2003. The rise and fall of isolation by distance in the anadromous brook charr (*Salvelinus fontinalis* Mitchill). *Genetics* 163: 983–996.
- Costa, G. C., Hampe, A., Ledru, M-P., Martinez, P. A., Mazzochini, G. G., Shepard, D. B., et al. 2017. Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. *Glob. Ecol. Biogeogr.* 27: 285–297.
- Crick, H. Q. P. 2004. The impact of climate change on birds. *Ibis* 146: 48-56.

- Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Meth.* 9(8): 772.
- Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., Telles, M. P. C., Rangel, T. F. and Bini, L. M. 2013. Mantel test in population genetics. *Genet. Mol. Biol.* 36: 475–485.
- Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolut. Biol.* 7: 214.
- Excoffier, L., Laval, G. and Schneider, S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47–50.
- French, N., Yu, S., Biggs, P., Holland, B., Fearnhead, P., Binney, B., Fox, A., Grove-White, D., Leigh, J., Miller, W., Muellner, P. and Carter, P. 2013. Evolution of *Campylobacter* species in New Zealand. - In: Sheppard, S. K. and Meric, G. (eds), *Campylobacter Ecology and Evolution* - Caister Academic Press.
- Fonseca, E. M., Gehara, M., Werneck, F. P., Lanna, F. M., Colli, G. R., Sites, J. W., Rodrigues, M. T., Garda, A. A. 2018. Diversification with gene flow and niche divergence in a lizard species along the South American “diagonal of open formations”. *J Biogeogr.* 45: 1688–1700.
- Fouquet, A., Cassini, C. S., Haddad, C. F. B., Pech, N., Rodrigues, M. T. 2014. Species delimitation, patterns of diversification and historical biogeography of the Neotropical frog genus *Adenomera*. *J. Biogeog.* 41: 855–870.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
- Furley, P. A. and Metcalfe, S. E. 2007. Dynamic changes in savanna and seasonally dry vegetation through time. *Prog. Phys. Geog.* 31: 633–642.

- Galeotti, P., Rubolini, D., Dunn, P. O. and Fasola, M. 2003. Colour polymorphism in birds: Causes and functions. *J. Evolut. Biol.* 16: 635–646.
- Gibbs, H. L., Dawson, R. J. G. and Hobson, K. A. 2000. Limited differentiation in microsatellite DNA variation among northern populations of the yellow warbler: evidence for male-biased gene flow? *Mol. Ecol.* 9: 2137–2147.
- Giugliano, L. G., Collevatti, R. G. and Colli, G. R. 2007. Molecular dating and phylogenetic relationships among Teiidae (Squamata) inferred by molecular and morphological data. *Mol. Phylogenet. Evol.* 45: 168–179.
- Haffer, J. 1969. Speciation in Amazonian forest birds. *Science* 165: 131–137.
- Hardy, O. J. and Vekemans, X. 1999. Isolation by distance in a continuous population: reconciliation between spatial auto-correlation analysis and population genetics models. *Genetics* 83:145-154.
- Heled, J. and Drummond, A. J. 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27: 570–580.
- Hijmans, R.J., Cameron, S. E., Parr, J. L., Jones, P. G. and Jarvis, A. 2005. Very high-resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965–1978.
- Holderegger R, and Wagner, H. H. 2006. A brief guide to landscape genetics. *Landscape Ecol.* 21: 793–796.
- Koscinski, D., Handford, P., Tubaro, P. L., Sharp, S. and Lougheed, S. C. 2008. Pleistocene climatic cycling and diversification of the Andean treefrog, *Hypsiboas andinus*. *Mol. Ecol.* 17: 2012–2025.
- Kopuchian, C., Campagna, L., Lijtmaer, D. A., Cabanne, G. S., García, N. C., Lavinia, P. D., Tubaro, P. L., Lovette, I., Di Giacomo, A. S. 2020. A test of the riverine barrier hypothesis in the largest subtropical river basin in the Neotropics. *Mol. Ecol.* 00:1–13.

- Ledru, M.P. 2002. Late quaternary history and evolution of the cerrados revealed by palynological records. - In: Oliveira, P. S. and Marquis, R. J. (eds), The cerrados of Brazil. - Columbia University Press.
- Leite, K. C., Seixas, G. H., Berkunsky, I., Collevatti, R. G. and Caparroz, R. 2008. Population genetic structure of the blue-fronted Amazon (*Amazona aestiva*, Psittacidae: Aves) based on nuclear microsatellite loci: implications for conservation. *Genet. Mol. Res.* 7: 819–829.
- Lima-Rezende, C. A., Rocha, A. V., Júnior, A. F. C., Martins, É. S., 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.
- Linares-Palomino, R., Oliveira-Filho, A. T. and Pennington, R. T. 2011. Neotropical Seasonally Dry Forest: diversity, endemism, and biogeography of woody plants. - In: Dirzo, R., Young, H. S., Mooney, H. A. and Ceballos, G. (eds), Seasonally Dry Tropical Forest: Ecology and Conservation. Island Press, pp. 3–21.
- Magalhaes, I. L. F., Oliveira, U., Santos, F. R., Vidigal, T. H. D. A., Brescovit, A. D., Santos, A. J. 2014. Strong spatial structure, Pliocene diversification and cryptic diversity in the Neotropical dry forest spider *Sicarius cariri*. *Mol. Ecol.* 23: 5323–5336.
- Manel, S., Schwartz, M. K., Luikart, G. and Taberlet, P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* 18: 189–197.
- Marantz, C., Aleixo, A., Bevier, L. and Patten, M. 2003. Available at: <http://www.hbw.com/species/narrow-billed-woodcreeper-lepidocolaptes-angustirostris>.
- Marini, M. A. and Hackett, S. 2002. A multifaceted approach to the characterization of an intergeneric hybrid manakin (Pipridae) from Brazil. *The Auk* 119: 1114–1120.

- Miller, M. P. 2005. Alleles In Space (AIS): Computer software for the joint analysis of inter-individual spatial and genetic information. *J. Heredity* 96: 722–724.
- Moraes, E. M., Yotoko, K. S. C., Manfrin, M. H., Solferini, V. N. and Sene, F. M. 2009. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. *J. Biogeogr.* 36: 2136–2147.
- Moura, C. M. M., de Araujo, H. F. P., Aleixo, A., Wink, M. and Fernandes, A. M. 2018. The role of landscape change and paleoclimatic events in shaping the evolutionary history of the *Polioptila* gnatcatchers (Passeriformes, Polioptilidae) with emphasis on species associated with open habitats. *J. Avian Biol.* 49: ejav-012409.
- Pacheco, M. A., Battistuzzi, F. U., Lentino, M., Aguilar, R. F., Kumar, S. and Escalante, A. A. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. *Mol. Biol. Evol.* 28: 1927–1942.
- Phillips, S., Dudik, M. and Schapire, R. 2010. Available at: <https://www.cs.princeton.edu/~schapire/maxent/>.
- Pinsky, M., Humberto, L., Montes Jr, R. and Palumbi, S. R. 2010. Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution* 64: 2688– 2700.
- Pennington, R. T., Prado, D. E. and Pendry, C. A. 2000. Neotropical seasonally dry forests and Quaternary vegetation change. *J. Biogeogr.* 27: 261–263.
- Prado, D. E. and Gibbs, P. E. 1993. Patterns of species distributions in the Dry Seasonal Forests of South America. *Ann. Missouri Botan. Gard.* 80: 902–927.
- Rambaut, A. and Drummond, A. J. 2012. FigTree: Tree figure drawing tool, v1.4.2. Institute of Evolutionary Biology, University of Edinburgh.
- Rambaut A, Drummond AJ, Xie D, Baele G and Suchard MA (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* syy032.

- Ramos, A. C. S., Lemos-Filho, J. P., Ribeiro, R. A., Santos, F. R. and Lovato, M. B. 2007. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of Quaternary climate changes in the Brazilian cerrado. *Ann. Bot.* 100: 1219-1228.
- Ramos-Onsins, S. E. and Rozas, J. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19: 2092–2100.
- Recoder, R.S., Werneck, F. P., Teixeira Jr., M., Colli, G. R., Sites Jr., J. W. and Rodrigues, M. T. 2014. Geographic variation and systematic review of the lizard genus *Vanzosaura* (Squamata, Gymnophthalmidae), with the description of a new species. *Zool. J. Linn. Soc.* 171: 206-225.
- Rheindt, F. E., Norman, J. A. and Christidis, L. 2008. DNA evidence shows vocalizations to be a better indicator of taxonomic limits than plumage patterns in *Zimmerius* tyrant-flycatchers. *Mol. Phylogenet. Evol.* 48: 150–156.
- Ribas, C. C., Gaban-Lima, R., Miyaki, C. Y. and Cracraft, J. 2005. Historical biogeography and diversification within the Neotropical parrot genus *Pionopsitta* (Aves: Psittacidae). *J. Biogeogr.* 32: 1409–1427.
- Ribeiro, V., Werneck, F. P., Machado, R. B. 2016. Distribution dynamics of South American savanna birds in response to Quaternary climate change. *Austral. Ecol.* 41: 768–777.
- Rozas, J., Sánchez-De, I., Barrio, J. C., Messeguer, X. and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
- Rull, V. 2008. Speciation timing and neotropical biodiversity: The Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. *Mol. Ecol.* 17: 2722–2729.
- Santos, M. G., Nogueira, C., Giugliano, L. G. and Colli, G. R. 2014. Landscape evolution and phylogeography of *Micrablepharus atticolus* (Squamata,

- Gymnophthalmidae), an endemic lizard of the Brazilian Cerrado. *J. Biogeog.* 41: 1506–1519.
- Savit, A. Z. and Bates, J. M. 2015. Right around the amazon: the origin of the circum-amazonian distribution in *Tangara cayana*. *Folia Zoolog.* 64: 273–283.
- Silva, J. M. C. and Bates, J. M. 2002. Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. *Bioscience* 52: 225–233.
- Silva, S. M., Peterson, A. T., Carneiro, L., Burlamaqui, T. C. T., Ribas, C. C., Sousa-Neves, T., Miranda, L. S., Fernandes, A. M., d'Horta, F. M., Araújo-Silva, L. E., Batista, R., Bandeira, Cinthia, H. M. M., Dantas, S. M., Ferreira, M., Martins, D. M., Oliveira, J., Rocha, T. C., Sardelli, C. H., Thom, G., Rêgo, P. S., Santos, M. P., Sequeira, F., Vallinoto, M. and Aleixo, A. 2019. A dynamic continental moisture gradient drove Amazonian bird diversification. *Sci. Adv.* 5(7): eaat5752.
- Sokal, R. R., Oden, N. L. and Thomson, B. A. 1997. A simulation study of microevolutionary inferences by spatial autocorrelation analysis. *Biol. J. Linn. Soc.* 60: 73–93.
- Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T. and Mindell, D. P. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol. Phylogenet. Evol.* 12: 105–114.
- Stephens, M., Smith, N. J. and Donnelly, P. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68: 978.
- Stotz, D. F., Fitzpatrick, J. W., Parker III, T. A. and Moskovits, D. K. 1996. Neotropical birds: ecology and conservation. - Chicago Univ. Press.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis, version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.

- Thom, G., Xue, A. T., Sawakuchi, A. O., Ribas, C. C., Hickerson, M. J., Aleixo, A. and Miyaki, C. Y. 2020. Quaternary climate changes as speciation drivers in the Amazon floodplains. *Sci. Adv.* 6: 1-11, eaax4718.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Ac. Res.* 24: 4876–4882.
- Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., Palma-Silva, C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. *Mol. Ecol.* 22: 1193-1213.
- Van Strien, M. J., Holderegger, R. and Van Heck, H. J. 2015. Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*, 114: 27– 37.
- Werneck, F. P. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quat. Sci. Rev.* 30: 1630–1648.
- Werneck, F. P., Gamble, T., Colli, G. R., Rodrigues, M. T. and Sites, J. W. 2012. Deep diversification and long-term persistence in the south american “dry diagonal”: integrating continent-wide phylogeography and distribution modelling of geckos. *Evolution* 66: 3014–3034.
- Werneck, F. P., Costa, G. C., Colli, G. R., Prado, D. E. and Sites Jr., J. W. 2011. Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecol. Biogeogr.* 20: 272-288.
- Whitmore, T. C. and Prance, G. T. 1987. *Biogeography and Quaternary History in Tropical America*. Clarendon Press, Oxford.
- Wright, S. S. 1946. Isolation by distance under diverse systems of mating. *Genetics* 31: 39-59.

BIOSKETCH

All authors are interested in Neotropical diversification patterns. Author contributions: A.V.R, R.C, G.S.C and L.F.S. concept the idea; A.V.R, R.C, G.S.C, and A.A designed the research; A.V.R, R.C., L.F.S. and G.S.C. performed the fieldwork; A.V.R. performed the laboratory work; A.V.R, R.C and G.S.C analyzed the data. All authors contributed to the discussion of the results and preparation of the manuscript. All authors read and approved the final manuscript.

Figure Legends

Figure 1. **A)** Geographic distribution of the *Lepidocolaptes angustirostris* tissue samples showing the subspecies analyzed in this study (symbols). The numbers correspond to localities described on Supplementary material Appendix 1 Table S1. Hachure area represents the *L. angustirostris* geographic distribution according BirdLife International (2015). The biomes composing the dry diagonal of South America are showed by colors: Cerrado - dark grey and Chaco - light grey. Green areas represent the SDTF nuclei: Caatinga (Ca), Misiones (Mi), Chiquitanos (Chi) and Piedmont (Pi) (following Linares-Palomino et al. 2011). Genetic lineages recovered in *L. angustirostris* are shown in colored symbols and phylogenetic relationships among lineages are shown on the right corner of the map (see also Fig. 3). Photos on left: unstreaked *L. a. angustirostris* from savannas of Macapá, Amapá, Brazil (top); streaked *L. a. angustirostris* from Chaco, Argentina (bottom). **B)** Median-joining networks showing the relationships between *L. angustirostris* haplotypes based on ND2 (top) and FIB5 (bottom) datasets. Network colors indicate the genetic lineages and the size of the circles is proportional to the number of individuals. Dashes represent the number of mutations separating the haplotypes. The numbers (LOC 1 to LOC 37) correspond to localities show in the map and described on Supplementary material Appendix 1 Table S1. Localities sharing haplotypes from different haplogroups (LOC 20 and 27) are shown in bold.

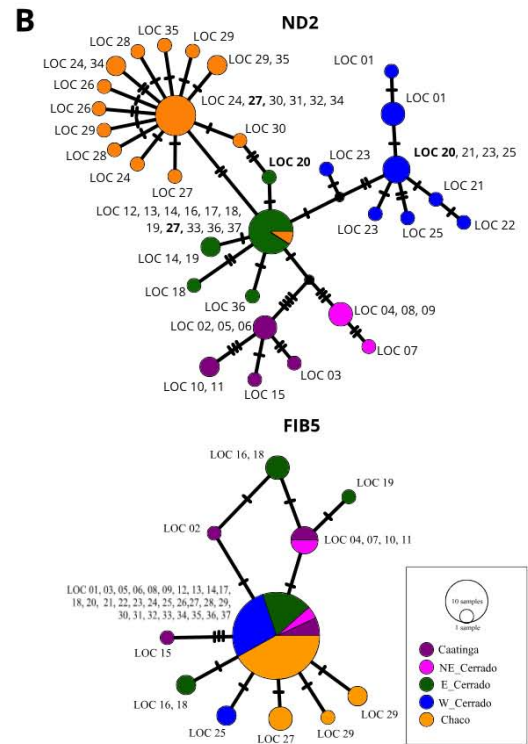
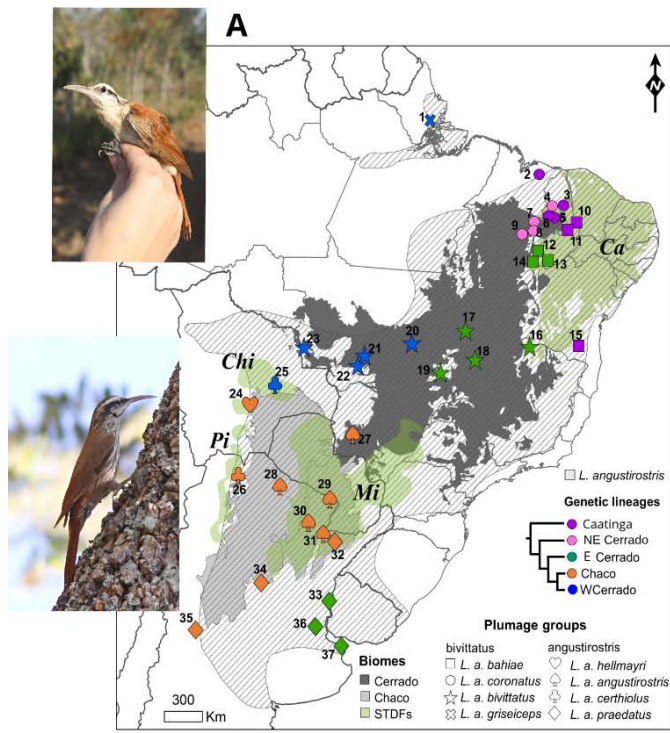


Figure 2. The majority rule consensus tree obtained for *Lepidocolaptes angustirostris* from the mixed-model Bayesian analysis of the concatenated dataset (ND2 + FIB5). The support values indicated at the node are the posterior probability (threshold > 0.8). The light brown bars indicate the throat and breast plumage color pattern. Branches were colored according to the genetic lineage as indicated in the colored bars on the right. Samples from the STDF nuclei (Caatinga, Missiones, Chiquitanos and Piedmont) were showed in bold. Samples collected in Caatinga/Cerrado ecotone are indicated with a hashtag. Localities sharing sequences form different lineages are shown in underline.

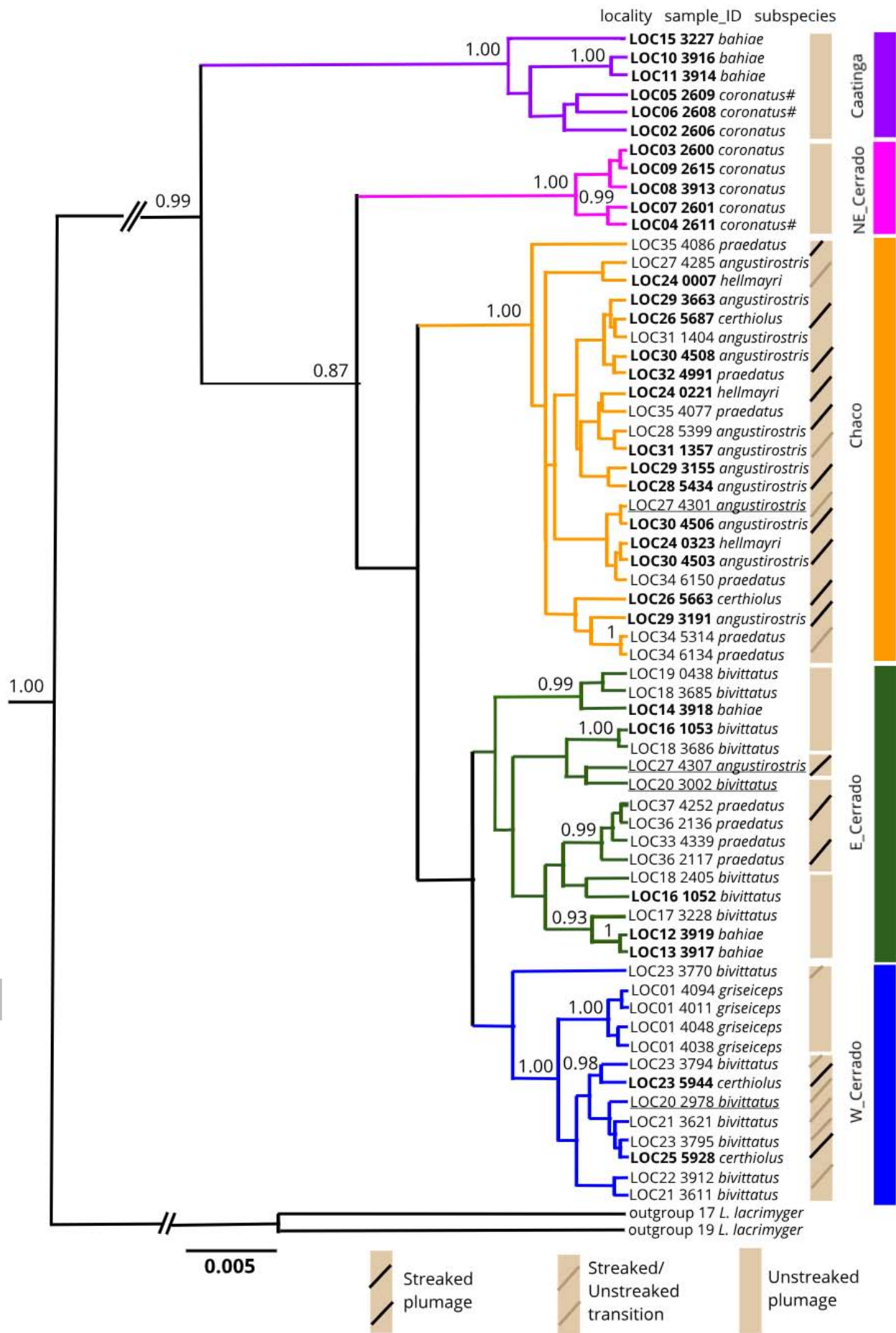


Figure 3. Multilocus species tree showing the relationships among the *Lepidocolaptes angustirostris* genetic lineages using coalescent model and constant population size for ND2 and FIB5 datasets. Support values (posterior probabilities) are shown for each node. Divergence time values (in years ago) and 95% CI (in parenthesis) between lineages estimated based on isolation with migration model using IMA2 are shown in each node.

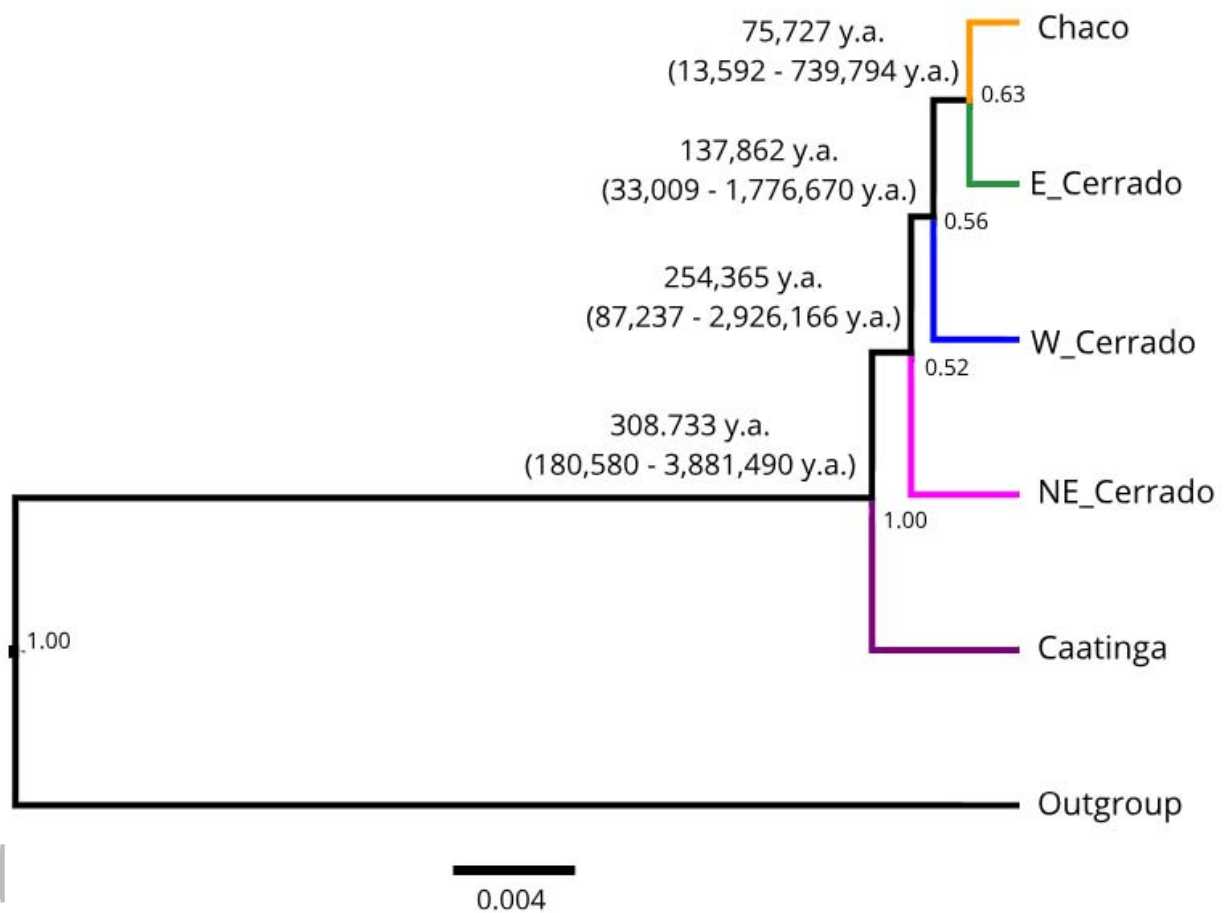


Figure 4. Potential niche modelling based on climatic variables available from WorldClim applying 10 percentile training presence as threshold. A: current projection using all climatic variables (see methods for details); green dots show the occurrence point used in modeling (for more information see Table S2); B: Projection for Mid-Holocene (6k); light grey represents suitable areas for just one database and dark grey the concordant suitable areas. C: Projection for Last Maximum Glacial (21k); light grey represents suitable areas for just one database and dark grey the concordant suitable areas. D: Projection for Last Interglacial (~120k); E: Stable areas defined as suitable ones for all periods are indicated by yellow areas. Genetic lineages are represented by colors: purple: Caatinga; pink: NE_Cerrado; blue: W_Cerrado; green: E_Cerrado and orange: Chaco.

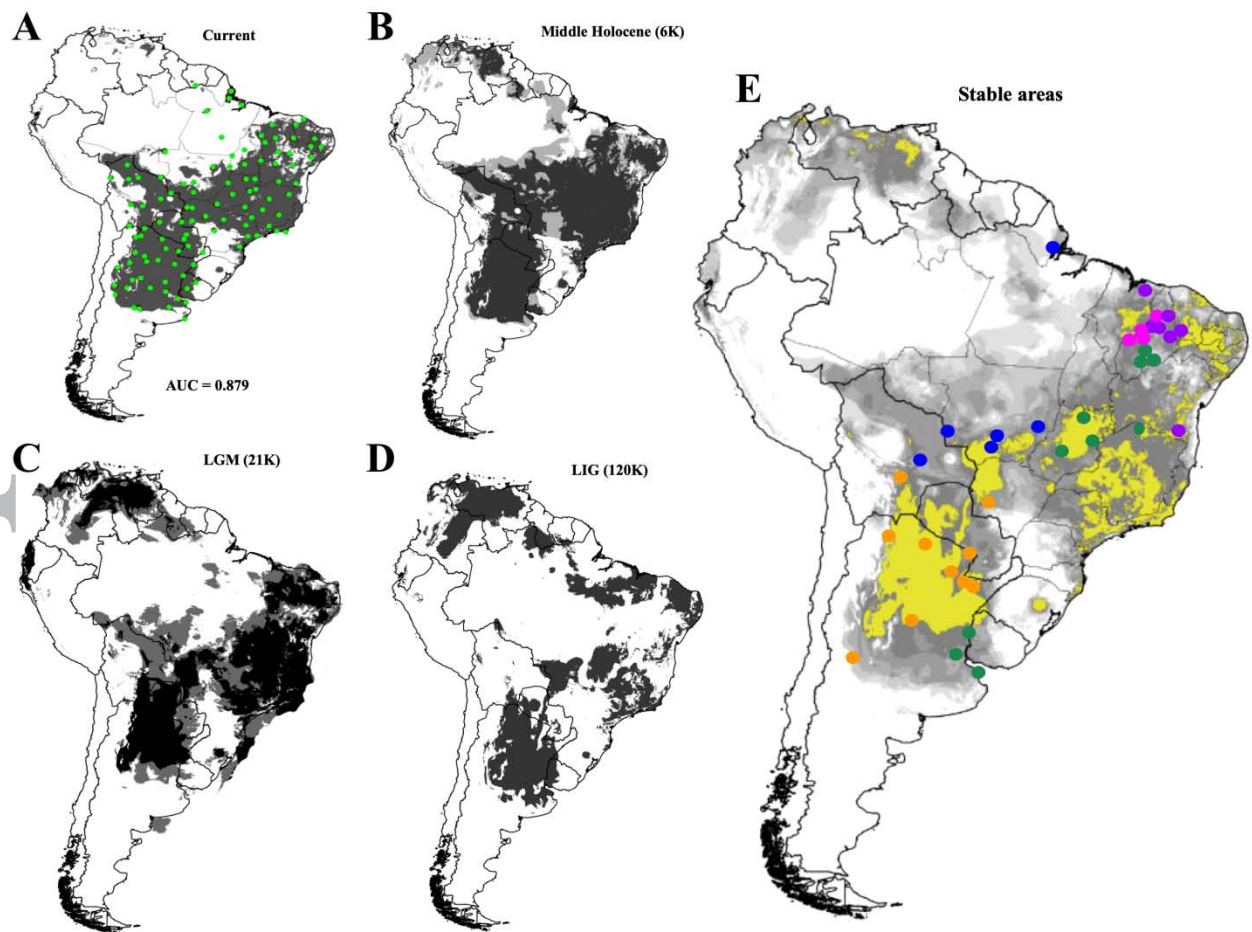


Table Legends

Table 1. Genetic diversity values for *Lepidocolaptes angustirostris* lineages based on two genetic markers (ND2 and FIB5). Numbers within parenthesis indicate sample size.

h = number of haplotypes; Hd = haplotype diversity; π = nucleotide diversity; k = nucleotide differences.

	Lineages											
	All samples (63)		Caatinga (6)		NE_Cerrado (5)		W_Cerrado (14)		E_Cerrado (14)		Chaco (24)	
	ND2	FIB5	ND2	FIB5	ND2	FIB5	ND2	FIB5	ND2	FIB5	ND2	FIB5
h	32	11	6	4	2	2	9	2	8	4	15	5
Hd	0.94	0.40	1.00	0.79	0.40	0.53	0.90	0.07	0.91	0.55	0.87	0.24
π	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01
k	4.71	0.61	3.60	1.70	0.80	0.53	3.10	0.07	2.87	1.05	1.90	0.25

Table 2. Mantel test results and demographic estimative of population expansion within each genetic lineage identified in *Lepidocolaptes angustirostris* using two different DNA alignment dataset (ND2 and FIB5). Statistically significant p-values ($p < 0.05$) are shown in bold.

Locus	Population	Mantel (r value)	p	Fu's F_S	p	Tajima'D	p	R^2	p
ND2	Caatinga	-0,428	1.000	0.121	0.473	-11.007	0.168	0.178	0.323
	NE_Cerrado	-0.539	1.000	10.986	0.739	-0.710	0.252	0.433	100.0
	W_Cerrado	0.820	0.107	-30.189	0.037	-13.987	0.082	0.085	0.004
	E_Cerrado	0.323	0.009	-20.800	0.097	-0.673	0.282	0.112	0.108
	Chaco	0.088	0.431	-12.551	0.000	-24.242	0.000	0.042	0.000
FIB5	Caatinga	0.440	0.005	0.480	0.626	-0.171	0.477	0.150	0.364
	NE_Cerrado	0.152	0.338	0.966	0.851	14.442	0.921	0.286	0.696
	W_Cerrado	0.290	0.068	-11.545	0.338	-11.514	0.115	0.186	0.973
	E_Cerrado	-0.131	0.739	0.453	0.631	0.050	0.571	0.131	0.136
	Chaco	0.102	0.316	-41.630	0.001	-16.634	0.002	0.054	0.083