Recent past connections between Amazonian and Atlantic forests by comparative phylogeography and paleodistribution models for didelphid mammals

Arielli Fabrício Machado

ariellifm@gmail.com

Universidade Federal do Rio Grande do Sul (UFRGS)

Maria Nazareth Ferreira da Silva

Instituto Nacional de Pesquisas da Amazônia (INPA)

Izeni Pires Farias

Universidade Federal do Amazonas (UFAM)

Marina Anciães

Instituto Nacional de Pesquisas da Amazônia (INPA)

Mario Silva Nunes

Universidade Federal do Amazonas (UFAM)

Cleuton Lima Miranda

Universidade Federal do Amazonas (UFAM)

Willian Thomaz Peçanha

Universidade Federal do Rio Grande do Sul (UFRGS)

Leandro Duarte

Universidade Federal do Rio Grande do Sul (UFRGS)

Additional Declarations: No competing interests reported.

Abstract

Despite the disjunct distribution of Amazonian and Atlantic forests, evidence suggests historical connections. Here we investigated the historical connections between three didelphid mammal species from Amazonian and Atlantic forests (*Caluromys philander*, *Marmosa murina* and *Marmosa demerarae*) to uncover these connections using a comparative phylogeography approach and paleodistribution models. We generated species distribution models from the present and past (Holocene through Pliocene) to evaluate changes in distributions over time focusing on the previously suggested connection regions. We inferred divergence times between populations of each species via the mitochondrial cytochrome b gene, estimated the phylogeographic relationships with haplotype networks, and calculated the genetic distances. All species showed Amazonian and Atlantic Forest divergences between 1 and 2 million years ago, while some populations diverged in more recent times. Paleodistribution was wider for periods before the estimated divergence times, becoming narrower after. We suggest that the climatic oscillations during the onset of the Pleistocene largely influenced the phylogenetic structuring of these forest-dwelling species, leading to the separation of populations of taxa currently distributed in Amazonian and Atlantic forests. Our findings also point to repeated connections over time, with both North- and South-eastern routes compatible with older connections, whereas younger connections were supported only for the Southeastern route. This represents the first simultaneous evaluation of past connections between Amazonian and Atlantic forests combining phylogeographic inferences with paleodistribution models. Our results highlight the relevance of environmental tolerance of species to their responses to climate change, such as more frequent increases in the paleodistribution of more resilient species.

Introduction

The Amazonian and Atlantic forests represent the largest tropical forests in the Americas, being among the most diverse regions in the world (Mittermeier et al 1998; Jenkins et al 2013). Despite the apparent disjunct distribution between them, the mosaic of forests in the interior of the Cerrado, Caatinga and Chaco arid biomes, the so-called 'dry diagonal' of vegetation of the South America, is considered a bridge of historical connection between these forests (Por 1992; Oliveira-Filho and Ratter 1995).

Several tree species from the Cerrado have a common geographical distribution between the Amazonian and Atlantic forests, and the same is true for some Caatinga species (Oliveira-Filho and Ratter 1995; Melo Santos et al 2007). Recently, some biogeographical studies have shown connections between these forests via reconstruction of phylogenetic relationships and ancestral areas for plants (Terra-Araujo et al 2015; Thode et al 2019). Accordingly, connections may also be evidenced for intra and interspecific relationships between the fauna of the Amazonian and Atlantic forests, such as found for birds (Batalha-Filho et al 2013; Batalha-Filho et al 2014; Carvalho et al 2017; da Silva et al 2020; Moura et al 2020); mammals (Cortés-Ortiz et al 2003; Costa 2003; Nascimento et al 2008; Pavan et al 2011); reptiles (Pellegrino et al 2011; Prates et al 2016); amphibians (Fouquet et al 2012; de Sá et al 2018) and insects (Peres et al 2017).

Evidence are for connections occurring in different periods over time, including both older, as prior to the Miocene (Pellegrino et al 2011; Fouquet et al 2012; Batalha-Filho et al 2013; Machado et al 2021) and recent connections, as those occurring during the Pleistocene (Costa 2003; Nascimento et al 2008; Batalha-Filho et al 2013; Prates et al 2016). For instance, phylogeographic studies for most mammal species tend to show more recent connections, associated with forest expansions and retractions caused by interglacial and glacial cycles during the Quaternary (Costa 2003; Nascimento et al 2008; Pavan et al 2011).

Three historical connection routes have been proposed to explain the phylogeographic associations observed for the fauna and flora from the Amazonian and Atlantic forests: one through the forests of the Paraná Basin passing to the southwest of the Brazilian Cerrado, Pantanal and Moist Chaco biomes named Southeastern-Northwestern route (SE-NW); a second through gallery forests of the Cerrado; and a third through Northeast Brazil via Caatinga and Cerrado/Caatinga ecotone forests named Northeastern route (NE) (Por 1992; Oliveira-Filho and Ratter 1995; Silva et al 1996; Costa 2003). These last two corridors would be related to the expansion of the gallery forests and wetlands during the Pleistocene climate oscillations (Batalha-Filho et al 2013).

Palynological and speleological data reveal short but recurring periods of rainfall intensification for the current northeastern semi-arid region of Caatinga during the Pleistocene and Holocene (De Oliveira et al 1999; Behling et al 2000; Auler et al 2004; Wang et al 2004). Some of this evidence suggests that during periods of strong rainfall in the Caatinga, two connecting routes between the Amazonian and the Atlantic forests would have been established: one near to the coast of northeastern Brazil, probably through the highland swamps and another in a transition area of the Cerrado biome through the seasonal forests of Chapada Diamantina (Oliveira et al 1999).

Paleogeographic studies indicate an increase in savannas during the late Miocene and early Pleistocene in the Cerrado due to the marked reduction in rainfall (Hoorn et al 2011; Ledru 2002). There is also evidence of range expansion for forest plant species from the Chaco/Pantanal in at least two periods: during the Last Interglacial (ca. 130–150 kya), for plants adapted to warm climate and during the Last Glacial Maximum (ca. 18–21 kya), for plants adapted to cold climates (Ledo and Colli 2017).

Studies focusing on biogeographic connections between the Amazonian and Atlantic forests have used either floristic similarities between these biomes (Oliveira-Filho and Ratter 1995; Méio et al 2003); phylogeographic patterns and phylogenetic relationships of animal species (Costa 2003; Batalha-Filho et al 2014); or species distribution models (SDMs) (Sobral-Souza et al 2015; Ledo and Colli 2017). Several studies have shown that integrating ecology and evolution helps the understanding of biogeographic processes leading to observed patterns of species distributions (Graham et al 2004; Anciães and Peterson 2009; Carnaval et al 2009; Carnaval et al 2014; Gutiérrez et al 2014; Machado et al 2019). With the advance of geographic information systems and increased availability of molecular data, it became possible to access numerous environmental variables at many temporal and spatial scales to investigate the geography of species distributions and estimating their ecological niches, inserting the phylogenetic

context in comparisons of species ranges and ecologies (Kozak et al 2008; Chan et al 2011; Alvarado-Serrano and Knowles 2014; Hoban et al 2019). As such, these tools applied to studies addressing historical connections between biomes can help the understanding of temporal dynamics affecting tropical forests, by predicting biogeographic routes that had been established between these biomes over time.

Thereby, this study aimed to investigate recent past connections between the Amazonian and Atlantic rainforests by reconstructing historical connections between populations of didelphid mammal species currently inhabiting these biomes. For this, we used comparative phylogeography coupled with paleodistribution modeling to predict changes in geographical distributions over time. Studies about historical connection routes between these biomes used either molecular data, geographic distribution data or paleodistributions of mammals, but never integrated these approaches into the same framework (but see Batalha-Filho et al 2013 for a focus on these connections at a macroevolutionary scale; and see da Silva et al (2020) for an integrative framework without a direct focus on such connections).

Here, we propose to test historical connecting routes between Amazonian and Atlantic forests in a comparative approach integrating molecular data and SDMs for mammals. Previous studies suggested recent connection routes through the Caatinga forests and Cerrado/Caatinga transition forests (Northeast) and older routes to the Chaco/Pantanal region (Southwest) (Batalha-Filho et al 2013; Sobral-Souza et al 2015; Ledo and Colli 2017), thus revealing conflicting results with expectations of either forest retraction or expansion during the Last Glacial Maximum (LGM) in South America, predicting expansion of paleodistributions of forest species (Haffer 1969; Leite et al 2016; Costa et al 2017).

To test the suggested hypotheses, we evaluated for scenarios of both forest retraction and expansion during the LGM, the Holocene (ca. 6 kya) and previous periods in the history of forest mammal species. Specifically, we address the following questions: 1) When did connections and disruptions between Amazonian and Atlantic forests occur for the populations of the species investigated here? 2) In which past periods did expansions or retractions in the distribution of specialist forest species occurred, revealing connections and disruptions between these forests? 3) Are the answers to these questions congruent when evaluated by different approaches, such as phylogenetic and ecological niche modeling? We are also interested in predicting in which regions these connections may have occurred in view of the alternative previously suggested connection routes between Amazonian and Atlantic forests.

Because previous studies indicated that the Northeast route (through the forests of the Caatinga Brazilian biome) would be more recent than the South route (through the Paraná Basin forests) (Por 1992; Batalha-Filho et al 2013; Ledo and Colli 2017), we expect to find more recent connections to the Northeast. Finally, we expect to find congruence between the results from the phylogeographic and palaeoecological approaches, such that paleodistributions should predict wider ranges for periods prior to the phylogenetic divergences followed by retractions, as evidence of past connections and separation between Amazonian and Atlantic forests, respectively.

Materials And Methods

Study area

The study area covers the region of distribution of Neotropical forests, encompassing the humid tropical forests of this region, focusing on the domains of Amazonia and the Atlantic Forest, as well as on Seasonal Dry Forests and riparian forests that occur between the Amazonia and the Atlantic Forest inland these Ecoregions: Cerrado, Caatinga, Cerrado/Caatinga ecotone and Chaco/Pantanal in South America. The Amazonia and the Atlantic Forest represent the main moist tropical forests of South America distributed mainly in Brazil. Currently, these forests have disjointed distribution in the region, isolated by the 'dry diagonal' which is composed of the open area of the Central Brazilian Cerrado, Semi-arid Caatinga of Northeastern Brazil and Chaco arid Ecoregions. Although the dry diagonal restricts the current spread between the Amazonia and the Atlantic Forest for much of the typical biodiversity of these rainforests, the Seasonally dry forests located inland (such as Deciduous and Semideciduous Forests; and riparian forests or gallery forests) are considered bridges of historical connection between the Amazonia and the Atlantic Forest (Por 1992; Costa 2003; Batalha-Filho et al 2013).

Sampled Species

We sampled three forest species of didelphid marsupials occurring between the Amazonia and the Atlantic Forest to investigate recent past connections between these forests through phylogeography and paleodistribution. These species are *Caluromys philander*, *Marmosa demerarae* and *Marmosa murina*, which are mammals of the order Didelphimorphia, family Didelphidae, representing the marsupials of the new world known as opossums. They occur in the moist forests of the Amazonia and the Atlantic Forest, as well as in seasonal deciduous, semi-deciduous forests and forest of galleries in the interior of the Caatinga, Cerrado and Chaco in South America (Emmons and Feer 1997; Gardner 2008).

Although these species are typically from forest habitats, they show differences in the use of the forest strata. *C. philander* is the most arboreal among these species using rather the canopy, *M. demerarae* is the second most arboreal species, both use more frequently the upper to the middle levels of the forest (canopy and sub-canopy), while *M. murina* is usually seen on the ground to the understory (Vieira and Camargo 2012). Regarding the type of diet of these species, all of them are omnivorous, but *C. philander* is the most fruitful while *M. murina* consumes more invertebrates and *M. demerarae* more arthropods and fruits (Santori et al 2012). *M. murina* is the most tolerant species in terms of forest disturbances (Pardini 2004; Passamani et al 2005), while the other species are mostly associated with richer forests (Emmons and Feer 1997). Even so, all species can vary in the use of vertical space, and consequently in the diet during the dry season depending on the scarcity of resources (Santori et al 2012; Vieira and Camargo 2012).

Sampling data

Genetic data

We used sequences of mitochondrial DNA of the Cytochrome b (*Cytb*) region for the three didelphid mammal species to evaluate genetic relationships between Amazonian and Atlantic forests. The majority *Cytb* sequences analyzed for *Caluromys philander* (27) were generated for this work. Hence, the minority sampled sequences for *C. philander* (13) were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/GenBank/) and incorporated in the analyses (Supporting Information, Table S1). Samples from *Marmosa murina* and *Marmosa demerarae* were gathered from the GenBank (77 and 47, respectively).

The total genomic DNA was extracted following the phenol–chloroform extraction (Longmire et al 1997) and amplified the partial *Cytb* gene sequences using the polymerase chain reaction (PCR) using the set of primers OryzCytbF1 and OryzCytbR1 following the instructions of the manufacturer (Smith and Patton 1993). The PCR was performed with a final volume of 13 μl, containing 5.2 μl of H_2O; 1.5 μl of $MgCl_2$ (25 mM); 1.25 μl of dNTP (2.5 mM); 1.25 μl of 10x buffer (100mM Tris-HCl, 500mM KCl); 1 μl of each primer (0.2 μM); 0.8 μl of Taq DNA Polymerase (5 U/μl) and 1μl of the extracted DNA (concentration varied between 10 ng and 30 ng). PCR consisted of an initial denaturation at 5 minutes for 94°C; followed by 30 cycles at 1 minute for 94°C, 1 minute and 30 seconds for 57°C and 1 minute for 72°C; and a final extension at 10 minutes for 72°C. The PCR products were checked in 1% agarose gel stained with GelRed® (Biotium) and purified using ExoSAP following the protocol described by Werle et al. (1994) as suggested by the manufacturer. The sequencing reaction followed the protocol recommended by the manufacturer of the BigDye Sequencing Kit (Life Technologies), while the products were sequenced on an ABI 3130XL (Life Technologies) automatic sequencer. Sequences were deposited in the GenBank database (Supporting Information, Table S1).

Occurrence records

Georeferenced occurrence records for the three didelphid mammalian species distributed between Amazonian and Atlantic forests (*Caluromys philander*, *Marmosa demerarae* and *Marmosa murina*) were collected from the online databases of SpeciesLink (http://splink.cria.org.br/), GBIF - Global Biodiversity Information Facility (www.gbif.org/) and VertNet (http://vertnet.org/). Furthermore, geographical coordinates of sampled locations with available genetic data (from GenBank, see below), as well as records in others literature was employed to complete the entire known distribution area for each didelphid species (Supporting information, Table S2). All occurrence records were checked using the geographic information system program QGIS 3.6 (QGIS 2020). A total of 293 occurrence records were compiled for *C. philander*, 568 for *M. murina* and 199 for *M. demerarae*.

Environmental data

To generate the Species Distribution Model (SDMs) for each species, we used digital layers of selected continuous bioclimatic variables extracted from the PaleoClim database (Brown et al 2018; paleoclim.org), in ca. 5 km of resolution, which were cut considering the limits of the Neotropical region. To project SDMs to the past digital layers of continuous bioclimatic variables simulated for different past periods (Holocene to Pliocene) were also extracted from the PaleoClim database (Brown et al 2018) also

cutting considering the limits of the Neotropical region. This database provides current bioclimatic variables using the original data from CHELSA model simulations (Karger et al 2017) and the most paleoclimatic variables from PaleoClim were simulated by the CCSM model (Fordham et al 2017), or by the HadCM3 model for the Pliocene (Brown et al 2018; Hill 2015; Dolan et al 2015). The paleoclimatic variables made available by this database are simulated using the CCSM model for most periods and the HadCM3 model for the Pliocene.

Phylogeography

The *Cytb* sequences were aligned using the ClustalW method and edited in Geneious software Trial Version (http://www.geneious.com, Kearse et al 2012). In general, partial sequences of the mitochondrial *Cytb* gene were obtained for *Marmosa murina* and *Marmosa demerarae* with 801 bp (with 135 and 153 informative sites, respectively), and a smaller fragment with 500 bp for *Caluromys philander* (with 68 informative sites). These sequences represent 39, 29 and 15 localities, respectively (Table S1 in Supporting Information).

We estimated gene trees using Bayesian inference (BI) based on datasets originally generated from samples housed in scientific collections, with additional sequences retrieved from GenBank (Supporting information, Table S1). Sequences from four species of didelphidae were used as outgroup: *Didelphis marsupialis* [JF280998.1], *Monodelphis emiliae* [KM071602.1], *Marmosa paraguayanus* [HM106372.1] and *Caluromys derbianus* [MK817302.1] (Supporting information, Table S1).

The substitution models used were the HKY$_+$G for *C. philander*, HKY$_+$I$_+$G for *M. murina* and *M. demerarae*, selected as the best models for each species based on the Akaike Information Criterion (AIC) as implemented in jModelTest v2.1 (Darriba et al 2012). We estimated BI trees using BEAST v2.6.0 (Bouckaert et al 2019), assuming a Birth-Death speciation model, with estimated base frequencies, four gamma rate categories, and an uncorrelated lognormal relaxed-clock model distribution, as tree priors. Runs were carried out under an uncorrelated lognormal relaxed-clock model with calibration.

The phylogenetic tree was dated using calibration points based on a fossil record, incorporating a substitution rate of 0.04 that represents the average of didelphid substitution rates simulated by Leite et al (2016). We incorporated information from two fossils to calibrate the nodes. The calibrations used were 15.9 to 16.3 million years ago (mya) for the separation time between Caluromyinae and Didelphidae (Marshall 1976), and between 12.1 to 12.6 mya for the separation between *Monodelphis* and *Marmosa* following Jansa et al (2014). We used three independent Markov Chain Monte Carlo (MCMC) runs, each one with four streams per 15 million steps of the MCMC for *C. philander*, 10 million steps for *M. demerarae*, and 30 million steps of the MCMC for *M. murina*, sampling every 1000 generations and discarding the initial about 20% generations as burn-in, starting the initial trees with randomness, without restriction. MCMC convergence also was checked in Tracer v.1.6 (Rambaut et al 2014), where we confirmed if effective sample sizes (ESS) were above 200. After discarding the initial 20% generations, remaining trees were compiled using TreeAnnotator (Bouckaert et al 2019) into a maximum clade

credibility (MCC) tree obtaining posterior probability (PP) values for all nodes. The tree consensus generated was visualized and edited in Figtree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

In addition, summary diversity index, such as the number of variable sites (S), number of haplotypes (H), haplotype diversity (*h*), nucleotide diversity (π), and average number of nucleotide differences (k) were identified computationally using the software DnaSP v5.10 (Librado and Rozas 2009) and Arlequin v3.5 (Excoffier and Lischer 2010). The evolutionary relationships between the haplotypes for each dataset (i.e., *C. philander*, *M. murina* and *M. demerarae*) were estimated using the median-joining method implemented in the Network 10.0 (http://www.fluxus-engineering.com). Finally, we calculate the genetic distance between groups within each species using the software MEGA (Kumar et al 2018), with the Tamura 3-parameter model and Bootstrap method.

Species Distribution Models

Data treatment

To avoid the effect of sampling bias on the final Species Distribution Models (SDMs), a polarization file was used, which generates a grid of sampling probabilities, where cell values reflect the variation in sampling effort between pixels to weight background points models used for modeling (Elith et al 2010). The bias file was generated for species occurrence records using the kernel density estimation method with the *kernelUD* function of the adehabitatHR 0.4.16 R package (Calenge 2006). The digital layers of the environmental variables used to calibrate and test the SDMs were crop based on the known geographical distribution area of each species using the International Union for Conservation of Nature - IUCN polygons (IUCN 2020) including a buffer of five pixels, considering that the distribution area may be larger than the currently known. To project the results of the final SDMs were cropped to the same geographical extent covering the study area representing the Neotropics (x minimum: -95; x maximum: -25; y minimum: -40; y maximum: 24). All treatment of environmental variables was performed in the R program (R Core Team 2021) using the raster R package (Hijmans et al 2017).

The level of correlation between all cropped variables was assessed by Pearson's correlation test using the '*cor*' function of the stats 3.6.2 R package (R Core Team 2021). The selection of uncorrelated variables followed the same procedure used by Rissler and Apodaca (2007), considering the correlation threshold of 0.75. Only the variables with greater biological relevance for each species were used to generate the final models, which was assessed in the results of the Jackknife graphs generated in preliminary models in which each variable is tested in isolation and by excluding each other variable, revealing the gain and loss of models containing or missing each variable (Phillips 2006).

Thus, the variables selected for *Caluromys philander* were Temperature Seasonality [standard deviation*100] (bio4), Annual Precipitation [mm/year] (bio12), Precipitation of Driest Month [mm/month] (bio14), Precipitation Seasonality [coefficient of variation] (bio15), Precipitation of Warmest Quarter [mm/quarter] (bio18) and Precipitation of Coldest Quarter [mm/quarter] (bio19); for *Marmosa murina* were bio4, Mean Temperature of Wettest Quarter [°C*10] (bio8), bio12, bio14, bio15, bio18 and bio19; and

for *Marmosa demerarae* were the variables bio4, Mean Temperature of Driest Quarter [°C*10] (bio9), bio12, bio15, bio18 and bio19.

Models' construction

The final Species Distribution Models (SDMs) were built for each species using all the localities (occurrence records non-duplicated) recorded for each species, the bias file and the previously uncorrelated environmental variables selected based on the relevance and biological contribution of each species (as detailed above). The SDMs were built using the MaxEnt 3.3.3 (Phillips et al 2007), which is a prediction algorithm for incomplete data (presence only) based on the principle of maximum entropy, assuming that the best approximation for an unknown probability distribution is one that satisfies any constraint on its distribution (Phillips et al 2006; Elith et al 2011). Due to the scarcity of long-term studies that collect reliable absence data for the species, MaxEnt becomes a suitable and widely used algorithm as it has also been more successful in generating models from a small number of samples (Elith et al 2011). A total of 114 not duplicated occurrence records were used to construct the SDM for *Caluromys philander*, 137 for *Marmosa murina* and 70 for *Marmosa demerarae*.

The final models were generated using 10 independent replicates, the MaxEnt default parameters and the cross-validate method to evaluate models. We chose the logistic output for the visualization of SDMs in geographic space (potential distributions), where pixel values represent a range of environmental suitability from 0 (representing inadequate conditions) to 1 (maximum suitability) (Phillips 2006). The performance of the models was evaluated using the AUC (Area Under the Curve) method, with the threshold value of AUC > 0.7 to accept models (Phillips 2006 suggest values greater than 0.5). The SDMs were projected for 11 past times representing: Late-Holocene ~ 4 thousand years ago (kya), Mid-Holocene ~ 8 kya, Early-Holocene ~ 11 kya, Pleistocene Younger Dryas Stadial ~ 13 kya, Pleistocene Bølling-Allerød ~ 15 kya, Pleistocene Heinrich Stadial 1 ~ 17 kya, Pleistocene Last Glacial Maximum (LGM) ~ 21 kya, Pleistocene Last Interglacial (LIG) ~ 130 kya, Pleistocene MIS19 ~ 787 kya, Mid-Pliocene warm ~ 3 mya and Pliocene M2 ~ 3.3 mya. These paleodistribution were used to evaluate the historic connecting routes between the Amazonia and the Atlantic Forest through the past.

Paleodistributions

Paleodistribution models were used to evaluate the connection routes between the Amazonia and the Atlantic Forest through the recent past times towards the present, measuring the expansion or retraction of the potential distribution over time compared to the current species distribution. For this, the SDMs generated with continuous environmental suitability (present and past) were transformed into binary models using a threshold value that maximizes the sensitivity and specificity in the model tests (Liu et al 2013). The threshold value that maximizes the sensitivity and specificity for the *C. philander* model was 0.32, 0.17 for the *M. murina* model and 0.30 for the *M. demerarae* model. These analyses were performed using the R package 'raster' (Hijmans et al 2017). Then, we summed the number of pixels for the region of each route (i.e., Northeastern route, NE route; Southeastern-Southeastern route, SE-SE route; and Southeastern-Northwestern route, SE-NW route) in the models of each period.

We delimited the areas of each connection route corresponding to three regions in the 'dry diagonal' according to the literature (Por 1992; Oliveira-Filho and Ratter 1995; Silva et al 1996; Costa 2003; Machado et al 2021) using the ecoregion polygons (Dinerstein et al 2017; https://ecoregions2017.appspot.com/). For this, we considered the boundaries of the Caatinga, Babaçu forests and adjacent semidecidual forests for NE route (which represent transition areas between Caatinga and the Amazonia), the Cerrado ecoregion boundaries (excluding areas from Northern Cerrado (which correspond to the NE route and areas from Southern Cerrado which correspond to the SE-NW route) and the limits of the ecoregions of the Pantanal, wet Chaco and adjacent semi deciduous forests for the SE-NW route.

Results

Genetic diversity

A total of 40 *Cytb* sequences were obtained for *Caluromys philander* (479 bp), 77 for *Marmosa murina* (801 bp) and 47 for *Marmosa demerarae* (801 bp), sampled in 15, 39 and 29 sites, respectively (Table S1 in Supporting Information). The number of segregating sites (S) between species ranged from 68 for *C. philander* and 153 for *M. demerarae* (Table 1), and the number of haplotypes (H) varied between 25 in *C. philander* and 54 in *M. murina*. Haplotype diversity (*h*) and nucleotide diversity (π) were much higher in *M. demerarae* samples than in the other species (Table 1), while the average number of nucleotide differences (k) was also higher, despite *M. murina* presenting many samples.

Table 1
Molecular diversity indices of *Cytb* (mtDNA) of three didelphid mammalian species in South America. Species: *C. philander* = *Caluromys philander*, *M. murina* = *Marmosa murina*, *M. demerarae* = *Marmosa demerarae*; N = Number of sequences; H = Number of haplotypes; S = number of variable sites; *h* ± SD = haplotype diversity ± standard deviation; π ± SD = nucleotide diversity ± Standard deviation; k = average number of nucleotide differences.

Species	N	H	S	$h \pm$ SD	$\pi \pm$ SD	k
C. philander	40	25	68	0.9026 ± 0.0331	0.01757 ± 0.00333	7.87051
M. murina	77	54	135	0.9836 ± 0.0063	0.02895 ± 0.00106	22.66576
M. demerarae	47	42	153	0.9944 ± 0.0060	0.03313 ± 0.00224	26.24144

Phylogeography and divergence time

Divergence times estimated from sequences indicate that populations of *Caluromys philander* from the Atlantic Forest diverged from those distributed in other regions (Cerrado, Southeastern Amazonia, Northeastern Amazonia, Northwestern Amazonia and Antilles) about 1.84 million years ago (mya), with 97% of support, although the divergence between populations from Cerrado and Southeastern Amazonia was recovered with low support (Fig. 1). In turn, divergence times for populations of *Marmosa murina* from the Northern Atlantic Forest were all well supported and dated for about 1.33 mya, although for one of the locations the sequence was grouped into the Cerrado and Southeastern Amazonia clade diverging

within this about 0.04 mya. In addition, populations from the Southwestern Atlantic Forest diverged from populations from Cerrado about 0.02 mya (Fig. 2). Lastly, divergence times between populations of *Marmosa demerarae* from the clade distributed in the Caatinga and Cerrado/Caatinga and the Northern Atlantic Forest and populations from the Cerrado, Northeastern Atlantic Forest and Southeastern Amazonia were estimated for about 2 mya, with high support. Populations from the Northeastern Atlantic Forest diverged from those from the Cerrado and some locations in the Southeastern Amazonia about 1.4 mya (Fig. 3).

The median-joining haplotype network based on the mtDNA showed agreement with the Bayesian consensus trees for all three species studied. However, in *C. philander* despite the several mutational steps between biogeographic populations, the presence of several median vectors suggests non-sampled or extinct ancestral sequences (Fig. 1). No widely distributed biogeographic population was found, although the Amazonia (Northwestern, Northeastern and Southeastern Amazonia) shelter most of the haplotypes found. Populations of Atlantic Forest and Pantanal appear as the most distant from each other with many mutational steps, when compared to other biogeographic populations. Although samples from the Pantanal appear to be more related to the Cerrado (and these with Antilles), in general, all populations presented restricted distributions for this species, and did not share haplotypes.

Considering *M. murina*, we found a structure similarity revealed by *C. philander* haplotype network, with several mutational steps and median vectors between biogeographic populations. Nonetheless, samples from the Atlantic Forest appear closer related with Cerrado and Southeastern Amazonia, even presenting many mutational steps, than other populations towards the north (i.e., Northeastern Amazonia and Antilles). The pattern for the *M. murina* network shows that Cerrado samples emerged as intermediaries between the Atlantic Forest and the Southeast of the Amazonia. Further, the samples from the Northeast of the Amazonia (the most diverse) are related to some samples from the Southeast of the Amazonia, with huge mutational steps. Thus, acting like a direct link of Northeast-Southeast Amazonia populations and other populations towards the south. While samples from the Antilles are only related to the Northeastern Amazonia (Fig. 2).

The haplotype network for *M. demerarae*, despite following a similar haplotype network structure, showed the most complex relationship when compared to the other species studied. Although revealed low frequency haplotypic variants, the network presents a high number of median vectors (i.e., non-sampled or extinct ancestral sequences) and mutational steps (Fig. 3). The network was divided into a reticulated portion consisting of Northern Atlantic Forest, Northeastern Atlantic Forest, Southeastern Amazonia and Caatinga haplotypes and a complex structure composed of haplotypes from all other locations. In addition, Southeastern Amazonia populations were the most diversified and widespread cluster, with three main connections: Northern Atlantic Forest and Caatinga, Northeastern Atlantic Forest, and Cerrado populations. The other biogeographic populations were verified haplotypes isolated and covered smaller areas when compared to Northern Atlantic Forest or Southeastern Amazonia biogeographic populations, except for Southwestern Amazonia with the majority haplotype occurrence grouped.

Genetic distance

For *C. philander*, the smallest genetic distance between Amazonian and Atlantic forests was for Northern Atlantic Forest and Southeastern Amazonia (0.039), followed by the distance between Northern Atlantic Forest and Northeastern Amazonia (0.041) and the distance between Northern Atlantic Forest and Cerrado was larger (0.048), revealing that these connections occurred through the Cerrado, but that the current populations of the Cerrado did not occupy this region at that time (Table 2). For *M. murina*, the smallest genetic distance between Amazonian and Atlantic forests was for Southwestern Atlantic Forest and Southeastern Amazonia (0.015), but the distance between Southwestern Atlantic Forest and Cerrado was even smaller (0.012), revealing that the Southwestern Atlantic Forest populations were derived from those of the Cerrado and not the Atlantic Forest. Then, considering the other populations of the Atlantic Forest for this species, the smallest genetic distance between Amazonian and Atlantic forests was for Northern Atlantic Forest and Southeastern Amazonia (0.038). However, the genetic distance between the Northern Atlantic Forest and the Cerrado was greater (0.040), revealing that these connections occurred through the Cerrado, but that the current populations of the Cerrado did not occupy this region at that time. For *M. demerarae*, the smallest genetic distance between Amazonian and Atlantic forests was for Northeastern Atlantic Forest and Southeastern Amazonia (0.029). However, the distance between Northeastern Atlantic Forest and Cerrado was even smallest (0.028), revealing that these connections occurred through the Caatinga (near the Northeastern Atlantic Forest is located) and through the Northern Cerrado.

Table 2
Genetic distance between populations from Amazonian and Atlantic forests of three species of didelphid mammals investigated in this study. Regions: CE = Cerrado; NA = Northern Atlantic Forest; PA = Pantanal; A = Antilles; NEA = Northeastern Amazonia; SEA = Southeastern Amazonia; SA = Southern Atlantic Forest; NWA = Northwestern Amazonia; NEAF = Northeastern Atlantic Forest; SWA = Southwestern Atlantic Forest.

C. philander	CE	NA	PA	A	NEA	SEA	SA	NWA
CE	-	-	-	-	-	-	-	-
NA	0.05	-	-	-	-	-	-	-
PA	0.11	0.13	-	-	-	-	-	-
A	0.03	0.05	0.12	-	-	-	-	-
NEA	0.02	0.04	0.12	0.03	-	-	-	-
SEA	0.02	0.04	0.12	0.03	0.01	-	-	-
SA	0.05	0.01	0.13	0.05	0.04	0.04	-	-
NWA	0.03	0.05	0.11	0.03	0.02	0.02	0.05	-
M. murina	NA	CE	SEA	SWA	A	NEA		
NA	-	-	-	-	-	-		
CE	0.04	-	-	-	-	-		
SEA	0.04	0.02	-	-	-	-		
SWA	0.04	0.01	0.01	-	-	-		
A	0.06	0.06	0.06	0.05	-	-		
NEA	0.05	0.04	0.04	0.04	0.04	-		
M. demerarae	SEA	NEA	SWA	NWA	CE	NEAF	NAF	CA
SEA	-	-	-	-	-	-	-	-
NEA	0.05	-	-	-	-	-	-	-
SWA	0.06	0.05	-	-	-	-	-	-
NWA	0.04	0.02	0.05	-	-	-	-	-
CE	0.03	0.05	0.06	0.04	-	-	-	-
NEAF	0.03	0.04	0.05	0.03	0.03	-	-	-
NAF	0.03	0.04	0.05	0.03	0.04	0.02	-	-
CA	0.03	0.04	0.05	0.03	0.03	0.02	0.01	-

Species Distribution Models

All SDMs showed high AUC value: 0.899 for the *C. philander* SDM, 0.956 for the *M. murina* SDM and 0.888 for the *M. demerarae* SDM. The variables of greatest contribution to the *C. philander* SDM were Temperature Seasonality [standard deviation*100] (bio4) with 22.4%, followed by Annual Precipitation [mm/year] (bio12) with 18.1% and Precipitation of Driest Month [mm/month] (bio14) with 17.7%. The variables of greatest contribution to the *M. murina* SDM were bio14 with 31.7%, followed by Mean Temperature of Wettest Quarter [°C*10] (bio8) with 18.7%, Precipitation of Coldest Quarter [mm/quarter (bio19) with 18% and bio4 with 17.7%. The variables of greatest contribution to the *M. demerarae* SDM were bio19 with 26.9%, followed by Mean Temperature of Driest Quarter [° C*10] (bio9) with 25.4% and bio4 with 25.3%.

Current SDMs for all species predicted disjunct areas of greater environmental suitability between the Amazonia and the Atlantic Forest (Fig. 4A, Fig. 5A, and Fig. 6A), mainly for *M. murina* (Fig. 5A). The SDM for *C. philander* predicted disjunct areas of greater environmental suitability between the Atlantic Forest, the Cerrado and the Amazonia (Fig. 4A). The result of this model also reveals areas of high environmental suitability for the occurrence of the species in regions where there is not record, such as regions of the Brejos de Altitude and Babaçu Forests in the Caatinga/Cerrado ecotone, as well as for the slope of the Andes (Fig. 4A). The SDM for *M. murina* predicted considerable disjunct areas of greater environmental suitability between the Amazonia and the Atlantic Forest (Fig. 5). The result of this model also reveals areas of high environmental suitability for the occurrence of the species in regions where there is not record, such as regions of the Southwestern Amazonia (Fig. 5A). The SDM for *M. demerarae* also predicted disjunct areas of greater environmental suitability between the Atlantic Forest, the Cerrado and the Amazonia (Fig. 6A). The result of this model also reveals areas of high environmental suitability for the occurrence of the species in regions where there is no record, such as regions of the Brejos de Altitude and Babaçu Forests in the Caatinga and Cerrado/Caatinga ecotone (Fig. 6A).

Paleodistribution changes

The results of the change in paleodistribution models over time revealed an increase in potential areas for all species in all connecting regions (Caatinga and Cerrado/Caatinga ecotone in Northeast Brazil, Cerrado and Chaco/Pantanal) for about 17 kya (Fig. 7). During the LGM (~ 21 kya), the results show a reduction in paleodistribution for all species for Cerrado regions, but for the other regions the paleodistributions remained stable, except for *M. demerarae* in the Caatinga and Cattinga/Cerrado ecotone, which reduced even less proportion (Fig. 7). During LIG (~ 130 kya), paleodistribution models show an increase in areas suitable for the occurrence of most species in most regions, except for *C. philander* which decreased in the Cerrado and remained reasonably stable for the other regions, for *M. murina* that did not show sudden changes in the Caatinga, Cerrado/Caatinga ecotone and Chaco/Pantanal regions and for *M. demerarae* in the Caatinga where it has reduced slightly (Fig. 7). For about 3 mya, the paleodistribution models reveal an increase in the areas suitable for the occurrence of *C. philander* and *M. demerarae* in the

Cerrado and Chaco-Pantanal but decrease in the Caatinga. For *M. murina*, there was a slight reduction for the Caatinga and the Cerrado route and stability for the Chaco-Pantanal (Fig. 7).

Comparative results

Comparison of all the results, genetic and ecological, revealed the most likely routes for each species and their timing. Combining these data, we found that the most likely route to *Caluromys philander* would be through the Caatinga, to *Marmosa murina* through the Cerrado and to *Marmosa demerarae* these connections would have occurred in two moments by two different routes, the Cerrado and the Caatinga (Table 3).

Table 3

Divergence time estimates between nodes that separate populations from the Amazonia and the Atlantic Forest for three species of didelphid mammals investigated in this study. The letters represent the nodes highlighted in the figures of the trees dated for each species. Age = estimated divergence time in million years ago, HPD% = 95% confidence intervals of Highest Posterior Density (HPD) and PP = Posterior probability values. Regions show the connection regions for each population of each species. CE = Cerrado and CA = Caatinga.

Species	Node	Age	HPD%	PP	Connection regions
Caluromys philander	A	1.84	1.04−2.71	0.97	Old CE
Marmosa murina	A	1.33	0.73−2.03	0.98	Old CE and CA
Marmosa demerarae	A	2.00	1.14−2.91	1.00	CE and CA
	B	1.37	0.71−2.08	0.98	CE and CA

Discussion

Many studies have investigated the connections between Amazonian and Atlantic forests through the phylogeography of species from these regions (Costa 2003; Nascimento et al 2008; Pavan et al 2011; Fouquet et al 2012; Terra-Araujo et al 2015; Prates et al 2016; Peres et al 2017; Capurucho et al 2018; and also see Ledo and Colli 2017 for a review), some also through paleodistribution models (Sobral-Souza et al 2015; Ledo and Colli 2017; Maciel et al 2017), but none through both approaches simultaneously focus on these connections. The research presented here is the first to investigate the connections between these forests through comparative phylogeography and paleodistribution simultaneously for small mammals. Thus, although those studies corroborated, through different approaches, the connections between Amazonian and Atlantic forests proposed by Por (1992), the present study brings a different proposal.

Among previous studies on molecular evidence of connections between the Amazonia and the Atlantic Forest for animals, few have been recent (for example, Costa 2003; Nascimento et al 2008; Pavan et al 2011; Batalha-Filho et al 2013; see a review in Ledo and Colli 2017), with only one covering didelphid mammals (Costa 2003). Although Costa (2003) has brought a lot of information and data for species of

didelphid mammals, many years have passed since its publication and, with this, there has been a considerable increase in the number of genetic data available. Although the results revealed by Costa (2003) for the time of divergence between the populations of the Amazonia and the Atlantic Forest are in accordance with those presented here, we bring additional results for *Marmosa demerarae*. In addition, after recent reviews of phylogenetic relationships between didelphid mammal species, some groups that were previously considered to be the same taxonomic group have been separated, for example, for *Marmosa murina* (Gutiérrez et al 2010; Faria et al 2013). Therefore, the data used here for the species investigated are up to date and the inferences we made considered these revisions.

Our study brings new evidence about the historical connections between Amazonian and Atlantic forests, revealing two times of connections and isolation between populations from these forests, one during the Pleistocene from about 2 and 1.3 million years ago (mya), and other during the Holocene (from about 0.02 to 0.1 mya), the most recent being for the Northeast region of Brazil for *M. demerarae*. Evaluating these molecular results together with changes in paleodistribution, we found increase of paleodistribution for two of the investigated species in periods prior to those of genetic divergence between populations of the Amazonia and the Atlantic Forest (as for about 3 mya), becoming narrower in younger periods (as for about 787 thousand years ago [kya]), except *M. murina* and for the Northeast route (NE). The beginning of the Pleistocene was a period of declining temperature in relation to previous periods and of greater climatic instability (Hansen et al 2013). Therefore, we suggest that climatic changes during the onset of the Pleistocene had great influence on the phylogenetic structure of these species, leading to the separation of the populations of Amazonian and Atlantic forests. Furthermore, we demonstrated for the first time a congruence between molecular and ecological evidence for the connections between these regions.

According to our results, there are at least two connection routes between Amazonian and Atlantic forests for the recent past, considering these species, which had also been previously suggested, the Northeastern and Southeastern routes (Por 1992; Oliveira-Filho and Ratter 1995; Silva et al 1996; Costa 2003). Apparently, *C. philander* reveals connections between these biomes across the Northern Atlantic Forest and Southeastern Amazonia. *M. murina* shows connections between Amazonian and Atlantic forests through the Cerrado and Southeastern Amazonia with the Northern Atlantic Forest representing a connection via the Southeastern route. Interestingly, we found connections between Amazonian and Atlantic forests in two different times for *M. demerarae*, the oldest dispersal event through Southeastern Amazonia and Cerrado to the Northern Atlantic Forest and the younger also through Southeastern Amazonia and Cerrado to the Northeastern Atlantic Forest. The first event probably represents connections through the SE-SE route, while the second and most recent event probably represents connections through the Northeastern route. Batalha-Filho et al (2013) suggests connections across the Northeast to more recent periods compared to the Southwest (Chaco and Pantanal regions). However, in general our results also revealed more recent connections to the Southeast route in agreement with other previous studies (Por 1992; Ledo and Colli 2017), although within a limited time scale. Our results corroborate Costa's (2003) proposal that the forests of central Brazil have played an important role as a present and past habitat for tropical forest species.

Unlike what was found in previous studies on the connections between the Amazonia and the Atlantic Forest through paleodistribution (Sobral-Souza et al 2015; Leite et al 2016; Costa et al 2017; Ledo and Colli 2017), we have not found a considerable expansion in the paleodistribution of the species investigated here during the Last Glacial Maximum (LGM). For two species there was a significant decrease in paleodistribution during LGM compared to the present, except for *M. murina* where there was a slight expansion mainly for Caatinga. This species is the most resilient species among those investigated here, occupying disturbed forest areas (Pardini 2004; Passamani et al 2005). Given this peculiarity of *M. murina*, it becomes evident that species respond to climate change according to its characteristics, which will pose great influence on the outcome of these idiosyncratic responses (Graham et al 1996).

As such, species-specific traits, such as environmental tolerance, will play different roles on the biogeographic history of each group (Papadopoulou and Knowles 2016; Machado et al 2021). For example, through paleodistribution models Ledo and Colli (2017) found expansion of species adapted to the cold in the LGM, while in the Last Inter-Glacial (LIG), the expansion involved species adapted to warmer climates. Neotropical marsupials, that is didelphid mammals, being better adapted to warmer climates (Hunsaker 2012), which explains why they do not follow this expansion observed for other groups during the LGM, but during the LIG. In addition, we emphasize that each species can respond in a different way to the climatic changes of each past time, according to their biological and/or ecological characteristics, in addition to their biogeographic history. These findings can be particularly important for the conservation of species through studies on the responses of species to future climate changes (Da Silva et al 2020). Another important point to highlight is that the genetic data used came mainly from online databases, showing that the increase in the availability of molecular data in recent years favors biogeographic studies. Thus, we encourage the molecular data available in the databases to be used to answer questions of biogeographic and evolutionary ecology.

The present study shows a congruence between the phylogenetic and paleodistribution models, something never evaluated simultaneously before the present study for the historical connections between the Amazonia and the Atlantic Forest. We corroborate two of the three previously proposed connection routes, namely through the Northeast and Southeast. We show that the connections between these forests have occurred more than once over recent past time through phylogeography and paleodistribution models. We found more recent connections for both the Northeastern and Southwestern routes. The present study also shows that the different biological and ecological characteristics of each group, as well as biogeographic history, are important factors to be considered when assessing responses to past climate changes and, consequently, for future climate changes. We also highlight that the increase in the availability of molecular data in recent years favors biogeographic studies and we encourage that these data be used to answer biogeographic and ecological-evolutionary questions.

Declarations

We are thankful to Renan Maestri, Fabricio Villalobos and Thales R. O. de Freitas for critical revision on manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Authors' contributions

AFM contributed to conceptualization, data curation, data analysis, writing – proofreading, and editing. MNFS and IPF contributed to obtaining data, writing – proofreading, and editing the manuscript. MA contributed to the conceptualization, writing – proofreading, and editing of the manuscript. MSN contributed to obtaining data and analyses. CLM contributed to writing – proofreading and editing the manuscript. WTP contributed to data analysis, writing – proofreading, and editing of the manuscript. LD contributed to the conceptualization, writing – proofreading, and editing of the manuscript.

References

1. Alvarado-Serrano DF, Knowles LL (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources* 14(2):233-248. doi.org/10.1111/1755-0998.12184.
2. Anciães M, Peterson AT (2009) Ecological niches and their evolution among Neotropical manakins (Aves: Pipridae). *Journal of Avian Biology* 40(6):591-604. doi.org/10.1111/j.1600-048X.2009.04597.x.
3. Auler AS, Wang X, Edwards RL, Cheng H, Cristalli PS, Smart PL, Richards DA (2004) Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *Journal of Quaternary Science 19*(7):693-701. https://doi.org/10.1002/jqs.876.
4. Batalha-Filho H, Pessoa RO, Fabre PH, Fjeldså J, Irestedt M, Ericson PG, ... and Miyaki CY (2014) Phylogeny and historical biogeography of gnateaters (Passeriformes, Conopophagidae) in the South America forests. *Molecular Phylogenetics and Evolution* 79:422-432. doi.org/10.1016/j.ympev.2014.06.025.

5. Batalha-Filho H, Fjeldså J, Fabre PH, Miyaki CY (2013) Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithology* 154(1):41-50. 10.1007/s10336-012-0866-7.

6. Behling H, Arz HW, Pätzold J, Wefer G. (2000) Late Quaternary vegetational and climate dynamics in northeastern Brazil, inferences from marine core GeoB 3104-1. *Quaternary Science Reviews* 19(10):981-994. https://doi.org/10.1016/S0277-3791(99)00046-3.

7. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology* 15(4):e1006650. 10.1371/journal.pcbi.1006650.

8. Brown JL, Hill DJ, Dolan AM, Carnaval AC, Haywood AM (2018) PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Nature – Scientific Data* 5:180254. doi.org/10.1038/sdata.2018.254.

9. Calenge C (2006) The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197:1035. doi.org/10.1016/j.ecolmodel.2006.03.017.

10. Capurucho JMG, Ashley MV, Ribas CC, Bates JM (2018) Connecting Amazonian, Cerrado, and Atlantic forest histories: Paraphyly, old divergences, and modern population dynamics in tyrant-manakins (*Neopelma/ Tyranneutes*, Aves: Pipridae). *Molecular phylogenetics and evolution* 127:696-705. doi.org/10.1016/j.ympev.2018.06.015.

11. Carnaval AC, Hickerson MJ, Haddad CF, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323(5915):785-789. 10.1126/science.1166955.

12. Carnaval AC, Waltari E, Rodrigues MT, Rosauer D, VanDerWal J, Damasceno R, Prates I, Strangas M, Spanos Z, Rivera D, Pie MR (2014) Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences* 281(1792):20141461. doi.org/10.1098/rspb.2014.1461.

13. Carvalho CDS, do Nascimento NFF, De Araujo HF (2017) Bird distributional patterns support biogeographical histories and are associated with bioclimatic units in the Atlantic Forest, Brazil. *Zootaxa* 4337(2):223-242. http://dx.doi.org/10.11646/zootaxa.4337.2.3.

14. Chan LM, Brown JL, Yoder AD (2011) Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution* 59:523–537. doi.org/10.1016/j.ympev.2011.01.020.

15. Cortés-Ortiz L, *et al* (2003) Molecular systematics and biogeography of the Neotropical monkey genus, *Alouatta*. *Molecular phylogenetics and evolution* 26(1):64-81. doi.org/10.1016/S1055-7903(02)00308-1.

16. Costa LP (2003) The historical bridge between the Amazonia and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography* 30:71–86. doi.org/10.1046/j.1365-2699.2003.00792.x.

17. Costa GC, Hampe A, Ledru MP, Martinez PA, Mazzochini GG, Shepard DB, ... and Carnaval AC (2017) Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. *Global Ecology and Biogeography 27*(3):285-297. https://doi.org/10.1111/geb.12694.

18. Da Silva EA, De Araujo HFP, Aleixo A, Antonelli A, Fernandes AM (2020) The effects of climate change on the distribution of South American antbirds (*Thamnophilus punctatus* complex) as affected by niche divergences and contact zone interactions between species. *Journal of Ornithology* 161(1):229-241. https://doi.org/10.1007/s10336-019-01721-3.

19. De Sá RO, Tonini JFR, van Huss H, Long A, Cuddy T, Forlani MC, ... and Haddad CF (2019) Multiple connections between Amazonia and Atlantic Forest shaped the phylogenetic and morphological diversity of Chiasmocleis Mehely, 1904 (Anura: Microhylidae: Gastrophryninae). *Molecular phylogenetics and evolution 130*:198-210. https://doi.org/10.1016/j.ympev.2018.10.021.

20. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8):772. doi.org/10.1038/nmeth.2109.

21. De Oliveira PE, Barreto AMF, Suguio K (1999) Late Pleistocene/Holocene climatic and vegetational history of the Brazilian caatinga: the fossil dunes of the middle São Francisco River. *Palaeogeography, palaeoclimatology, palaeoecology 152*(3-4):319-337. https://doi.org/10.1016/S0031-0182(99)00061-9.

22. Dinerstein E, Olson D, Joshi A, Vynne C, Burgess ND, Wikramanayake E, Hahn N, Palminteri S, Hedao P, Noss R, Hansen M (2017) An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* 67(6):534-545. doi.org/10.1093/biosci/bix014.

23. Dolan AM, Haywood AM, Hunter SJ, Tindall JC, Dowsett HJ, Hill DJ, Pickering SJ (2015) Modelling the enigmatic Late Pliocene Glacial Event—Marine Isotope Stage M2. *Global and Planetary Change* 128:47-60. doi.org/10.1016/j.gloplacha.2015.02.001.

24. Eisenberg JF, Redford KH, Reid FA (2000) Mammals of the Neotropics: Ecuador, Bolivia.

25. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17:43-57. https://doi.org/10.1111/j.1472-4642.2010.00725.x.

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

27. Faria MB, Oliveira JAD, Bonvicino CR (2013) Filogeografia de populações brasileiras de *Marmosa (Marmosa) murina* (Didelphimorphia, Didelphidae). *Revista Nordestina de Biologia* 21(2):27-52. arca.fiocruz.br/handle/icict/11605.

28. Fordham DA, Saltré F, Haythorne S, Wigley TM, Otto-Bliesner BL, Chan KC, Brook BW (2017) PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography* 40(11):1348-1358. doi.org/10.1111/ecog.03031.

29. Fouquet A, Loebmann D, Castroviejo-Fisher S, Padial JM, Orrico VG, Lyra ML, Roberto IJ, Kok PJ, Haddad CF, Rodrigues MT (2012) From Amazonia to the Atlantic forest: Molecular phylogeny of Phyzelaphryninae frogs reveals unexpected diversity and a striking biogeographic pattern emphasizing conservation challenges. *Molecular phylogenetics and evolution* 65(2):547-561. doi.org/10.1016/j.ympev.2012.07.012.

30. Graham CH, Ron SR, Santos JC, Schneider, C.J., Moritz, C. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, 58(8): 1781-1793. doi.org/10.1554/03-274.

31. Graham, R.W., Lundelius, E.L., Graham, M.A., Schroeder, E.K., Toomey, R.S., Anderson, E., ... & Guthrie, R.D. 1996. Spatial response of mammals to late Quaternary environmental fluctuations. *Science*, 272(5268): 1601-1606.

32. Gutiérrez, E.E., Jansa, S.A., Voss, R.S. 2010. Molecular systematics of mouse opossums (Didelphidae: Marmosa): assessing species limits using mitochondrial DNA sequences, with comments on phylogenetic relationships and biogeography. *American Museum Novitates*, (3692): 1-22. doi.org/10.1206/708.1.

33. Gutiérrez, E.E., Boria, R.A., Anderson, R.P. 2014. Can biotic interactions cause allopatry? Niche models, competition, and distributions of South American mouse opossums. *Ecography*, 37(8): 741-753. doi.org/10.1111/ecog.00620.

34. Haffer, J. 1969. Speciation in Amazonian forest birds. *Science*, 165: 131-137.

35. Hansen, J., Sato, M., Russell, G., Kharecha, P. 2013. Climate sensitivity, sea level and atmospheric carbon dioxide. *Phil. Trans. R. Soc. A.*, 371(2001): 20120294. DOI: 10.1098/rsta.2012.0294.

36. Hijmans RJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ (2017) Package 'dismo'. Circles 9(1):1-68. http://rspatial.org/sdm/.

37. Hill, D.J. 2015. The non-analogue nature of Pliocene temperature gradients. *Earth and Planetary Science Letters*, 425: 232-241. dx.doi.org/10.1016/j.epsl.2015.05.044.

38. Hoban, S., Dawson, A., Robinson, J.D., Smith, A.B., Strand, A.E. 2019. Inference of biogeographic history by formally integrating distinct lines of evidence: genetic, environmental niche and fossil. *Ecography*, 42: 1-21. doi.org/10.1111/ecog.04327.

39. Hoorn, C., & Wesselingh, F. (Eds.). (2011). *Amazonia: landscape and species evolution: a look into the past*. John Wiley & Sons. 464 pg. ISBN 1444360256.

40. Hunsaker, D.I. (Ed.). 2012. *The biology of marsupials*. Elsevier. 556 pg. ISBN 0323146201, 9780323146203.

41. IUCN. 2020. *The IUCN Red List of Threatened Species. Version 2020.* Available from http://www.iucnredlist.org. Accessed on 09 January 2020.

42. Jansa, S.A., Barker, F.K., & Voss, R.S. 2014. The early diversification history of didelphid marsupials: a window into South America's "Splendid Isolation". *Evolution*, 68(3): 684-695. doi.org/10.1111/evo.12290.

43. Jenkins, C.N., Pimm, S.L., Joppa, L.N. 2013. Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences*, *USA*, 110(28): E2602–E2610. doi.org/10.1073/pnas.1302251110.

44. Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. & Kessler, M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4: 170122. https://doi.org/10.1038/sdata.2017.122.

45. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... Thierer, T. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28* (12): 1647-1649. doi.org/10.1093/bioinformatics/bts199.

46. Kozak, K.H., Graham, C.H., Wiens, J.J. 2008. Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, 23(3): 141-148. doi.org/10.1016/j.tree.2008.02.001.

47. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, *35*(6), 1547. doi.org/10.1093/molbev/msy096.

48. Ledo, R.M.D., Colli, G.R. 2017. The historical connections between the Amazonia and the Atlantic Forest revisited. *Journal of Biogeography*, 44(11): 2551-2563. doi.org/10.1111/jbi.13049.

49. Ledru, M. P. (2002). Late Quaternary history and evolution of the cerrados as revealed by palynological records. Oliveira P.S. & Marquis R.J. (eds) *The Cerrados of Brazil: ecology and natural history of a Neotropical savanna*, 33-50. Columbia University Press. ISBN 0231120427.

50. Leite, Y.L., Costa, L.P., Loss, A.C., Rocha, R.G., Batalha-Filho, H., Bastos, A.C., ... Pardini, R. 2016. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences*, *113*(4): 1008-1013. doi.org/10.1073/pnas.1513062113.

51. Librado, P., & Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452. doi.org/10.1093/bioinformatics/btp187.

52. Liu, C., White, M., Newell, G. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 40(4): 778-789.

53. Longmire, J.L., Maltbie, M., Baker, R.J. 1997. Use of "lysis buffer" in DNA isolation and its implication for museum collections. *Occasional papers/Museum of Texas Tech University*, 163. doi.org/10.5962/bhl.title.143318.

54. Machado, A. F., Ritter, C. D., Miranda, C. L., Bredin, Y. K., Ramos Pereira, M. J., & Duarte, L. (2021). Potential mammalian species for investigating the past connections between Amazonia and the Atlantic Forest. *PloS one*, *16*(4), e0250016. doi.org/10.1371/journal.pone.0250016.

55. Machado, A.F., Nunes, M.S., Silva, C.R., dos Santos, M.A., Farias, I.P., da Silva, M.N.F., Anciães, M. 2019. Integrating phylogeography and ecological niche modelling to test diversification hypotheses

using a Neotropical rodent. *Evolutionary Ecology*, 33(1): 111-148. doi.org/10.1007/s10682-019-09968-1.

56. Maciel, J.R., Sánchez-Tapia, A., Siqueira, M.F.D., Alves, M. 2017. Palaeodistribution of epiphytic bromeliads points to past connections between the Atlantic and Amazonia forests. *Botanical Journal of the Linnean Society*, 183(3): 348-359. doi.org/10.1093/botlinnean/bow020.

57. Marshall, L.G. 1976. New didelphine marsupials from the La Venta fauna (Miocene) of Colombia, South America. *Journal of Paleontology*, 50(3): 402-418. www.jstor.org/stable/1303521.

58. Méio, B.B., Freitas, C.V., Jatobá, L., Silva, M.E., Ribeiro, J.F., & Henriques, R.P. 2003. Influência da flora das florestas Amazônica e Atlântica na vegetação do cerrado sensu stricto. *Brazilian Journal of Botany*, *26*(4), 437-444.

59. Melo Santos, A.M., Cavalcanti, D.R., Silva, J.M.C.D., Tabarelli, M. 2007. Biogeographical relationships among tropical forests in north-eastern Brazil. *Journal of Biogeography,* 34(3): 437-446. doi.org/10.1111/j.1365-2699.2006.01604.x.

60. Mittermeier, R.A., Myers, N., Thomsen, J.B., Da Fonseca, G.A., Olivieri, S. 1998. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology*, 12(3): 516-520. doi.org/10.1046/j.1523-1739.1998.012003516.x.

61. Moura, C.C.D.M., Fernandes, A.M., Aleixo, A., Pereira de Araujo, H.F., Mariano, E.D.F., Wink, M. 2020. Evolutionary history of the Pectoral Sparrow *Arremon taciturnus*: evidence for diversification during the Late Pleistocene. IBIS, *International Journal of Avian Science*. https://doi.org/10.1111/ibi.12813.

62. Nascimento, F.F., Bonvicino, C.R., De Oliveira, M.M., Schneider, M.P.C., & Seuánez, H.N. 2008. Population genetic studies of *Alouatta belzebul* from the Amazonian and Atlantic Forests. American Journal of Primatology: Official Journal of the American *Society of Primatologists*, 70(5), 423-431. https://doi.org/10.1002/ajp.20507.

63. Oliveira-Filho, A.T. de, Ratter, J.A. 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. *Edinburgh Journal of Botany*, 52(02): 141-194. doi.org/10.1017/S0960428600000949.

64. Pardini, R. 2004. Effects of forest fragmentation on small mammals in an Atlantic Forest landscape. *Biodiversity & Conservation*, *13*(13), 2567-2586. https://doi.org/10.1023/B:BIOC.0000048452.18878.2d.

65. Papadopoulou, A., & Knowles, L.L. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences*, 113: 8018−8024. https://doi.org/10.1073/pnas.1601069113.

66. Passamani, M., Jenilson, D., & Lopes, S.A. 2005. Mamíferos não-voadores em áreas com predomínio de Mata Atlântica da Samarco Mineração SA, município de Anchieta, Espírito Santo. *Biotemas*, *18*(1), 135-149. https://doi.org/10.5007/%25x.

67. Pavan, A.C., Martins, F., Santos, F.R., Ditchfield, A., & Redondo, R.A. 2011. Patterns of diversification in two species of short-tailed bats (*Carollia* Gray, 1838): the effects of historical fragmentation of

Brazilian rainforests. *Biological journal of the Linnean Society*, 102(3), 527-539. https://doi.org/10.1111/j.1095-8312.2010.01601.x.

68. Pellegrino, K.C.M., Rodrigues, M.T., James Harris, D., Yonenaga-Yassuda, Y., Sites Jr, J.W. 2011. Molecular phylogeny, biogeography and insights into the origin of parthenogenesis in the Neotropical genus *Leposoma* (Squamata: Gymnophthalmidae): Ancient links between the Atlantic Forest and Amazonia. *Molecular Phylogenetics and Evolution,* 61: 446-459. doi.org/10.1016/j.ympev.2011.07.010.

69. Peres, E.A., Silva, M.J., & Solferini, V.N. 2017. Phylogeography of the spider *Araneus venatrix* (Araneidae) suggests past connections between Amazonia and Atlantic rainforests. *Biological Journal of the Linnean Society*, 121(4): 771-785. https://doi.org/10.1093/biolinnean/blx036.

70. Phillips, S. 2006. A Brief Tutorial on Maxent. *AT&T Research*, 107-135. Available on http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on 20 January 2020.

71. Phillips, S.J., Dudík, M., Schapire, R.E. 2007. *Maxent software for species habitat modeling*. Version 3.3.3k. Available in https://www.cs.princeton.edu/~schapire/maxent/. Accessed on 10 January 2020.

72. Por, F.D. 1992. *Sooretama: the Atlantic rain forest of Brazil*. SPB Academic. The Hague: SPB Academic Publishing.

73. Prates, I., Rivera, D., Rodrigues, M. T., Carnaval, A. C. 2016. A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Molecular ecology*, *25*(20), 5174-5186. https://doi.org/10.1111/mec.13821.

74. QGIS Development Team. 2020. *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. Available in http://qgis.osgeo.org. Accessed on 10 January 2020.

75. R Core Team. 2021. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available in http://www.R-project.org/. Accessed on 10 January 2021.

76. Rambaut, A., Suchard, M.A, Xie, D. & Drummond, A.J. 2014. Tracer v1.6. Available at: beast.bio.ed.ac.uk/Tracer. Accessed on 1 June 2019.

77. Rissler, L.J., Apodaca, J.J. 2007. Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, 56(6): 924-942. doi.org/10.1080/10635150701703063.

78. Santori, R.T., Lessa, L.G., & Astúa, D. 2012. Alimentação, nutrição e adaptações alimentares de marsupiais brasileiros. *Os marsupiais do Brasil: biologia, ecologia e conservação (NC Cáceres, ed.). Ed. UFMS. Campo Grande*, 385-406. ISBN: 978-85-7613-410-7.

79. Silva, J.D. 1996. Distribution of Amazonian and Atlantic birds in gallery forests of the Cerrado region, South America. *Ornitologia neotropical*, 7(1), 1-18. https://sora.unm.edu/node/119226.

80. Smith, M.F., Patton, J.L. 1993. The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe. *Biological Journal of the Linnean Society,* 50:149-177. doi.org/10.1111/j.1095-8312.1993.tb00924.x.

81. Sobral-Souza, T., Lima-Ribeiro, M.S., Solferini, V.N. 2015. Biogeography of Neotropical Rainforests: past connections between Amazonia and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, 29(5): 643-655. doi.org/10.1007/s10682-015-9780-9.

82. Terra-Araujo, M.H., de Faria, A.D., Vicentini, A., Nylinder, S., Swenson, U. 2015. Species tree phylogeny and biogeography of the Neotropical genus *Pradosia* (Sapotaceae, Chrysophylloideae). *Molecular phylogenetics and evolution,* 87: 1-13. doi.org/10.1016/j.ympev.2015.03.007.

83. Thode, V.A., Sanmartín, I., & Lohmann, L.G. 2019. Contrasting patterns of diversification between Amazonian and Atlantic forest clades of Neotropical lianas (Amphilophium, Bignonieae) inferred from plastid genomic data. *Molecular phylogenetics and evolution*, 133: 92-106. https://doi.org/10.1016/j.ympev.2018.12.021.

84. Vieira, E.M., Camargo, N.F., & Cáceres, N. 2012. Uso do espaço vertical por marsupiais brasileiros. *Os marsupiais do Brasil: biologia, ecologia e conservação (NC Cáceres, org.). Editora UFMS, Campo Grande, Brazil*, 345-362. ISBN: 978-85-7613-410-7.

85. Wang, X., Auler, A.S., Edwards, R.L., Cheng, H., Cristalli, P.S., Smart, P. L., ... & Shen, C.C. 2004. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature*, *432*(7018): 740-743. https://doi.org/10.1038/nature03067.

86. Warren, D.L., Glor, R.E., Turelli, M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, 33(3): 607-611. doi.org/10.1111/j.1600-0587.2009.06142.x.

87. Werle, E., Schneider, C., Renner, M., Völker, M., Fiehn, W. 1994. Convenient singlestep, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res*, 22: 4354-4355. doi.org/10.1093/nar/22.20.4354.
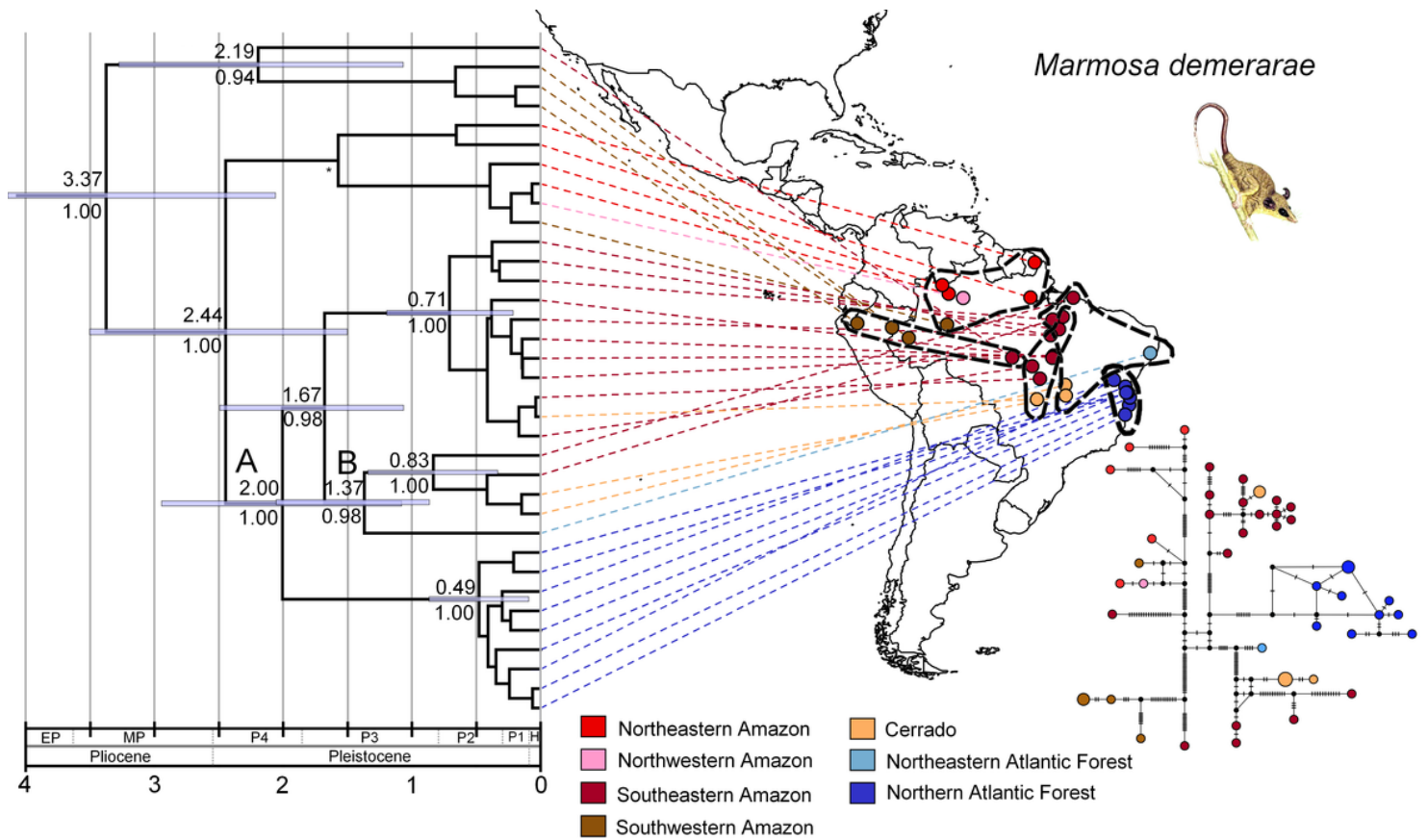
Figures
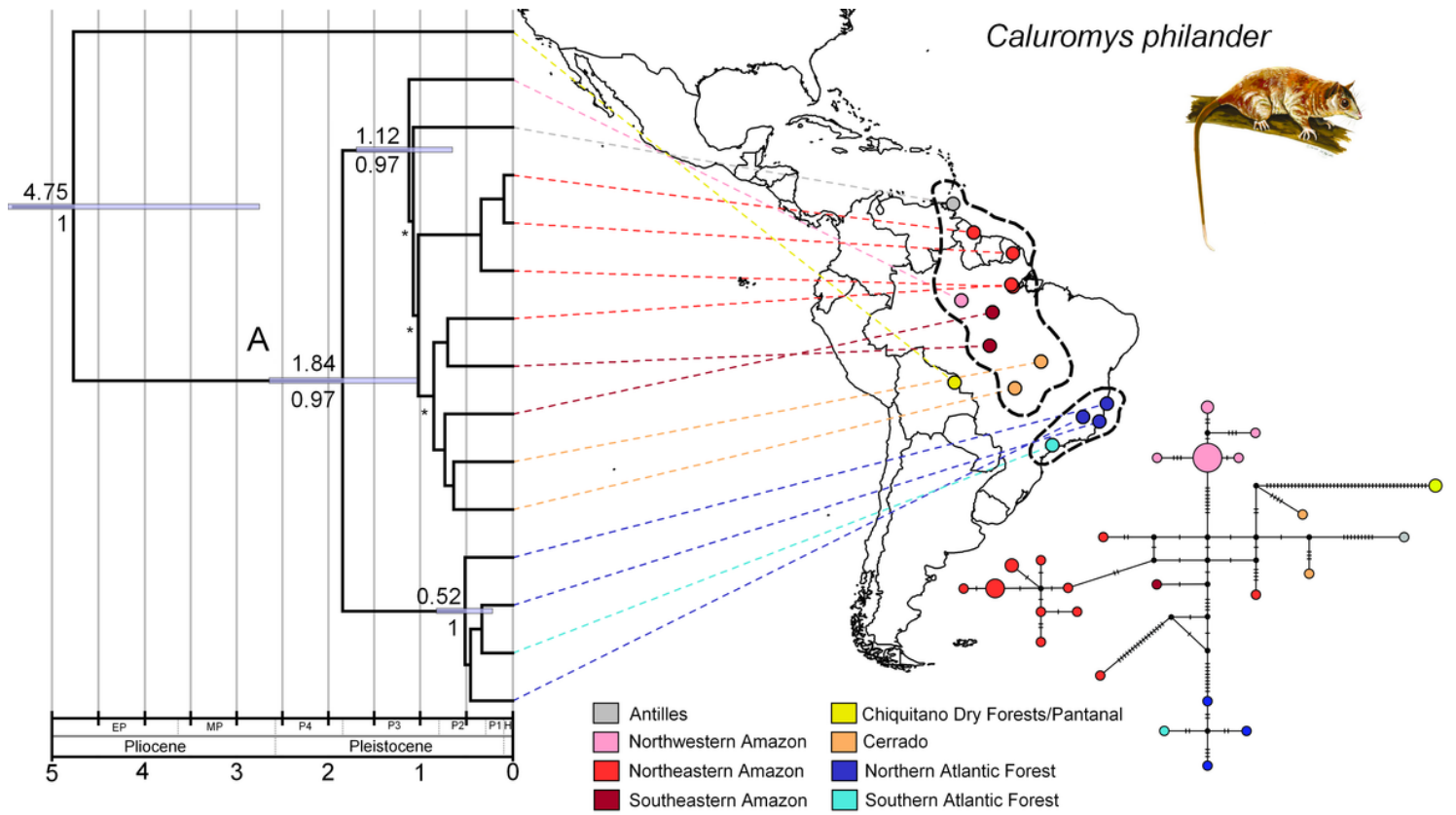
Figure 1

Time-calibrated phylogenetic tree for *Caluromys philander* based on cytochrome b gene (*Cytb*) plotted in geographical space. The axis represents the time scale in millions of years indicating the respective geological ages (EP = Early Pliocene, MP = Late Pliocene, P4 = First period of Early Pleistocene, P3 = Second period of Early Pleistocene, P2 = Middle Pleistocene, P2 = Late Pleistocene, H = Holocene). Node values represent the node supports and range of the HPD%, and asterisks nodes without support. Letters show divergence times between populations from the Amazonia and the Atlantic Forest detailed in Table 1. The colors of the occurrence records represent biogeographic regions: gray = Antilles, pink = Northwest Amazonia, light red = Northeast Amazonia, dark red = Southeast Amazonia, yellow = Chiquitano Dry Forests/Pantanal, orange = Cerrado, dark blue = North Atlantic Forest and light blue = South Atlantic Forest. Down right, is shown the haplotype networks of *Cytb*: the circles represent shared haplotypes; the size circle represents the number of sharing haplotypes and colors the biogeographical regions corresponding to the map. Species illustration by Eisenberg and Redford (2000).
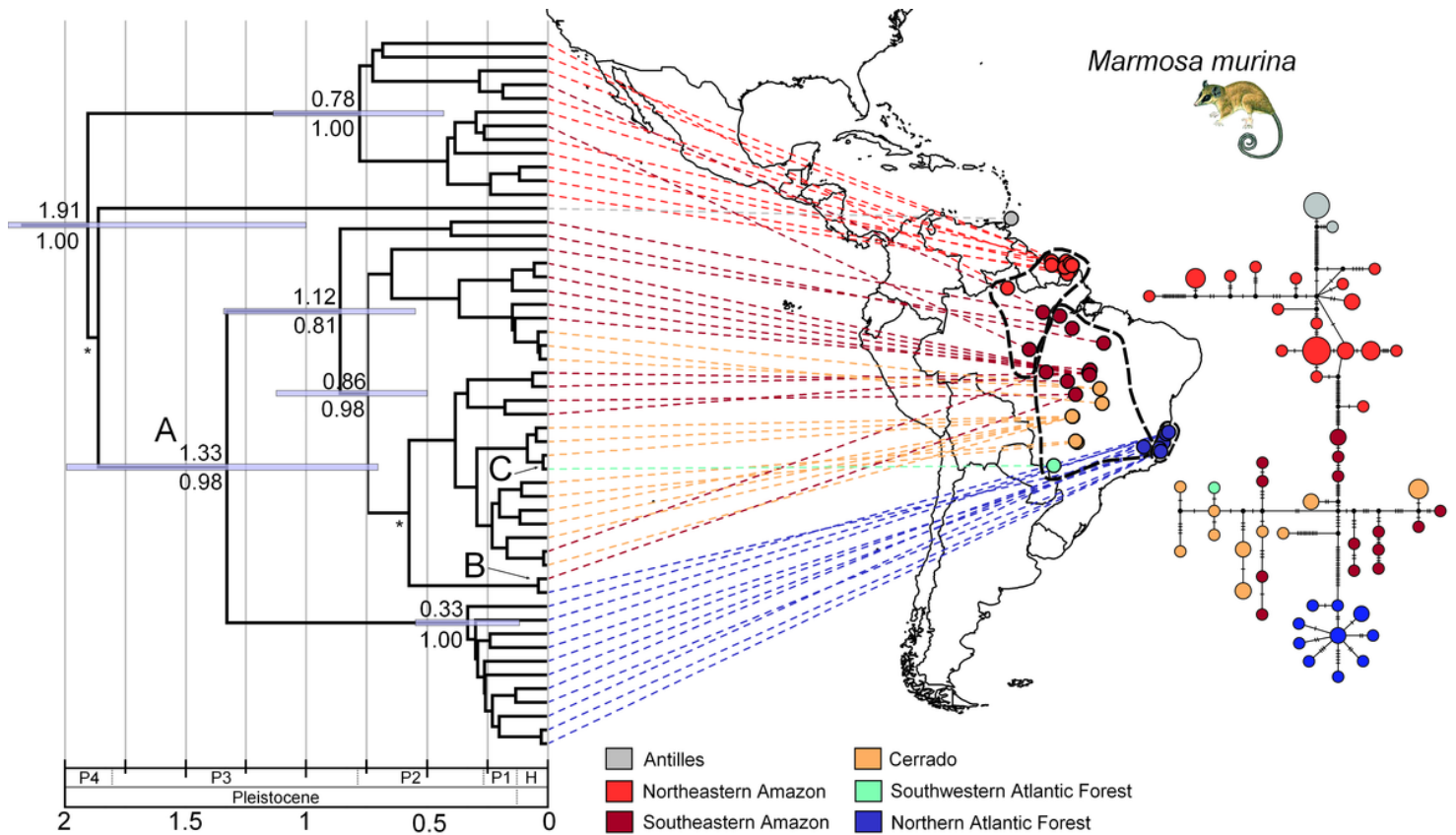
Figure 2

Time-calibrated phylogenetic tree for *Marmosa murina* based on cytochrome b gene (*Cytb*) plotted in the geographical space. The axis represents the time scale in millions of years indicating the respective geological ages (EP = Early Pliocene, MP = Late Pliocene, P4 = First period of Early Pleistocene, P3 = Second period of Early Pleistocene, P2 = Middle Pleistocene, P2 = Late Pleistocene, H = Holocene). Node values represent the node supports and HPD%, and asterisks nodes without support. Letters highlight divergences between Amazonia and Atlantic Forest. The colors of the occurrence records represent biogeographic regions: dark blue = Northern Atlantic Forest, orange = Cerrado, dark red = Southeastern Amazonia, light red = Northern Amazonia and gray = Antilles. Haplotype networks of *Cytb*: the circles represent shared haplotypes; the size circle represents the number of sharing haplotypes and colors the biogeographical regions corresponding to the map. Species illustration by Eisenberg and Redford (2000).

Figure 3

Time-calibrated phylogenetic tree for *Marmosa demerarae* based on cytochrome b gene (*Cytb*) plotted in geographical space. The axis represents the time scale in millions of years indicating the respective geological ages (EP = Early Pliocene, MP = Late Pliocene, P4 = First period of Early Pleistocene, P3 = Second period of Early Pleistocene, P2 = Middle Pleistocene, P2 = Late Pleistocene, H = Holocene). Node values represent the node supports and range of the HPD%, and asterisks nodes without support. Letters highlight divergences between Amazonia and Atlantic Forest. The colors of the occurrence records represent biogeographic regions: light red = Northeastern Amazonia, pink = Northwestern Amazonia, dark red = Southeastern Amazonia, brown = Southwestern Amazonia, orange = Cerrado, violet = Caatinga, dark blue = Northern Atlantic Forest and light blue = Northeastern Atlantic Forest. Haplotype networks of *Cytb*: the circles represent shared haplotypes; the size circle represents the number of sharing haplotypes and colors the biogeographical regions corresponding to the map. Species illustration by Eisenberg and Redford (2000).
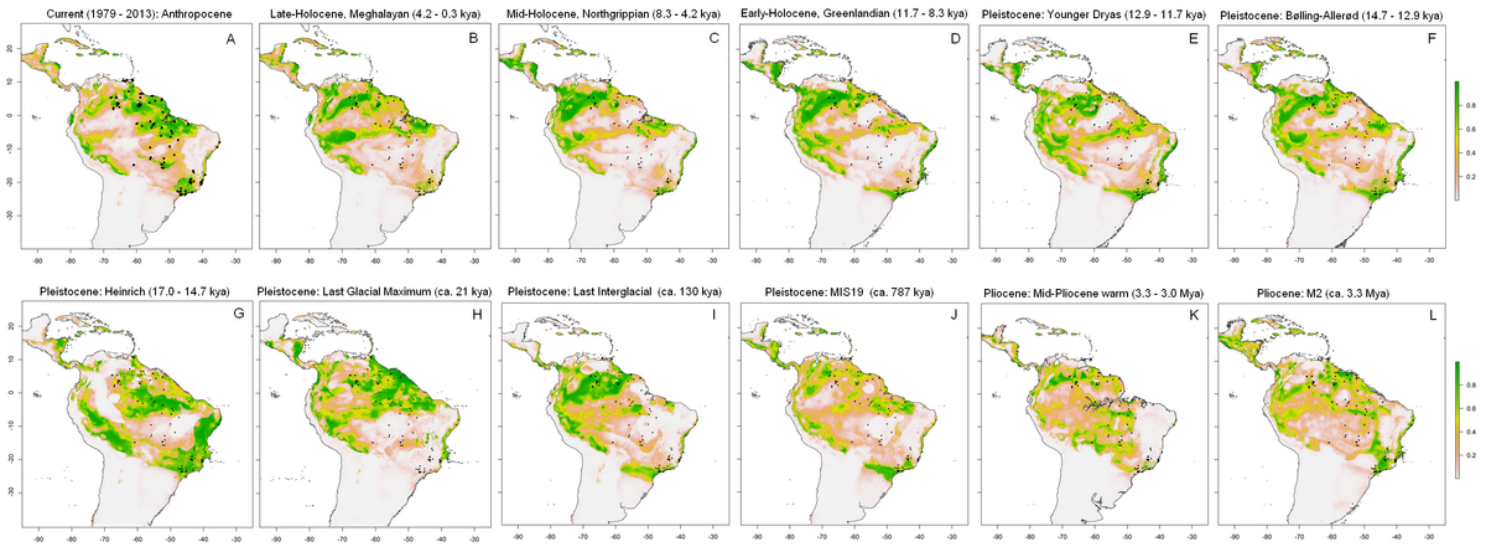
Figure 4

Species distribution models for *Caluromys philander* from top-left to bottom-right for the present and 11 past times. Color scales represent environmental suitability values ranging from 0 (in gray) to 1 (in green).
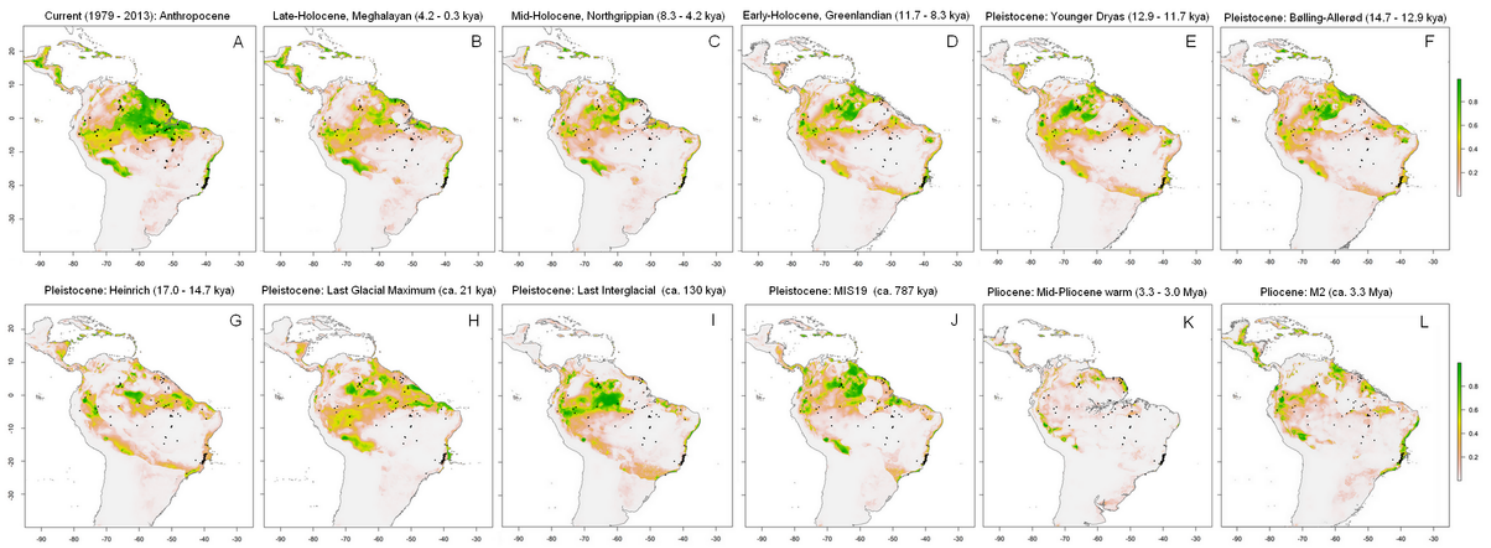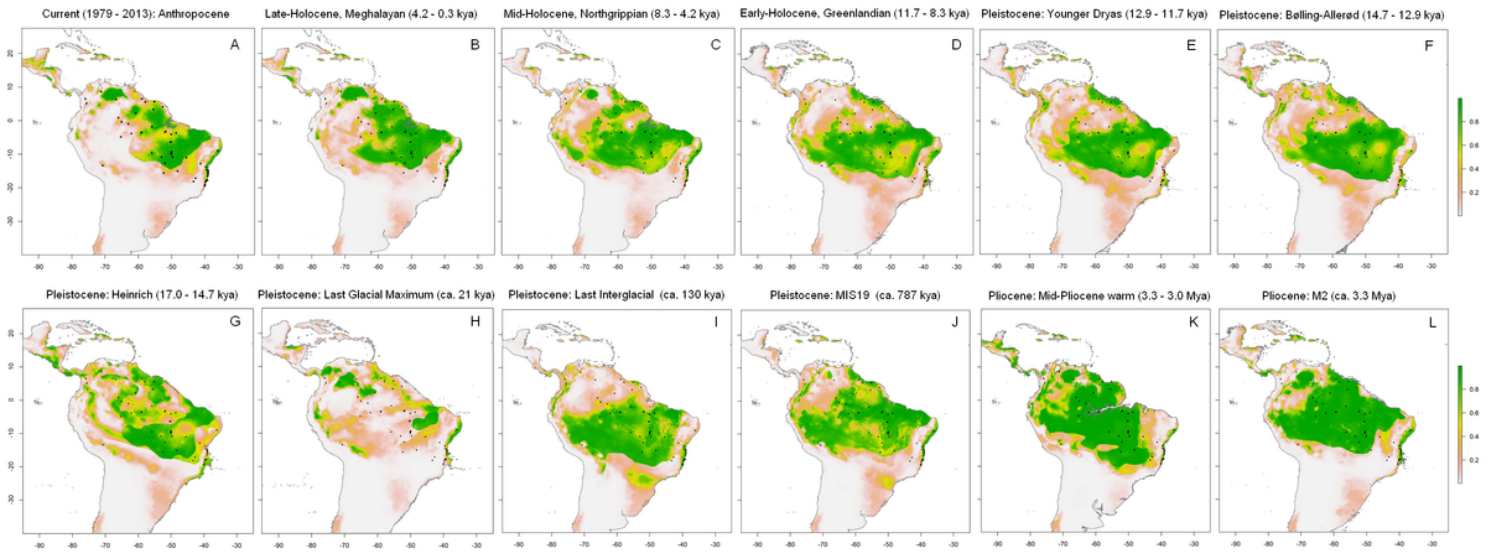

Figure 5

Species distribution models for *Marmosa murina* from top-left to bottom-right for the present and 11 past times. Color scales represent environmental suitability values ranging from 0 (in gray) to 1 (in green).
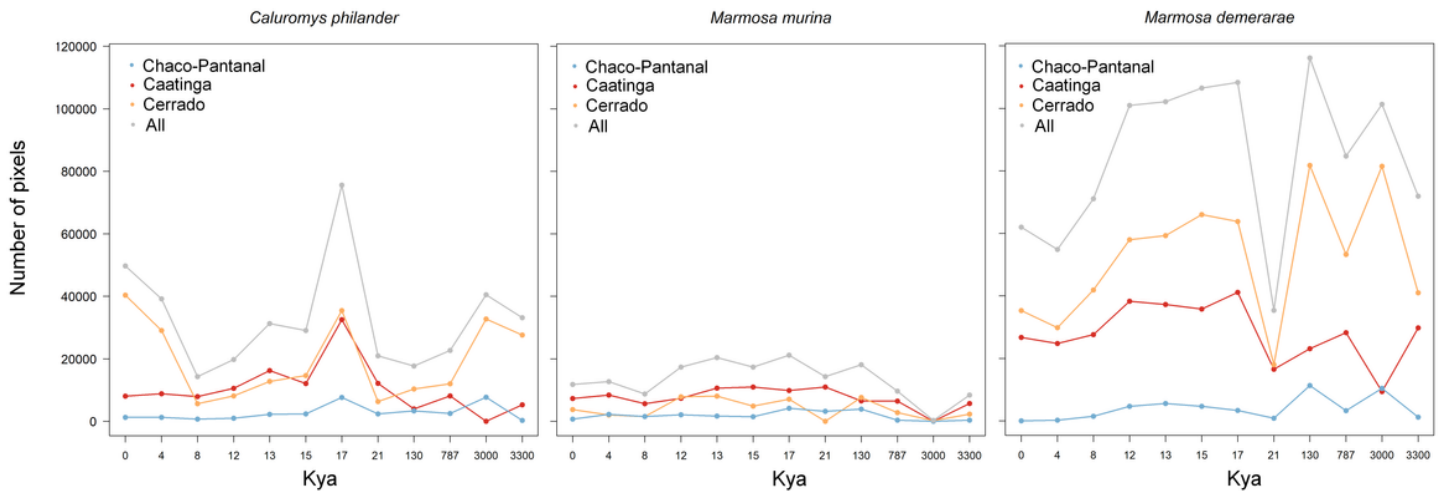
Figure 6

Species distribution models for *Marmosa demerarae* from top-left to bottom-right for the present and 11 past times. Color scales represent environmental suitability values ranging from 0 (in gray) to 1 (in green).

Figure 7

Change in paleodistribution models over time of each didelphid mammalian species for connection routes between Amazonia and Atlantic Forest. The y-axis represents the number of pixels of suitable areas for species occurrence in the binary models; The x-axis represents time on a scale of a thousand years ago (Kya). The Caatinga region is represented in red, the Cerrado in yellow, the Chaco-Pantanal in blue and all regions in gray.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx