#### **ORIGINAL PAPER**



# The role of environmental filtering, geographic distance and dispersal barriers in shaping the turnover of plant and animal species in Amazonia

Cristian Dambros, et al. [full author details at the end of the article]

Received: 18 January 2020 / Revised: 6 August 2020 / Accepted: 14 August 2020 © The Author(s) 2020

#### Abstract

To determine the effect of rivers, environmental conditions, and isolation by distance on the distribution of species in Amazonia. Location: Brazilian Amazonia. Time period: Current, Major taxa studied: Birds, fishes, bats, ants, termites, butterflies, ferns + lycophytes, gingers and palms. We compiled a unique dataset of biotic and abiotic information from 822 plots spread over the Brazilian Amazon. We evaluated the effects of environment, geographic distance and dispersal barriers (rivers) on assemblage composition of animal and plant taxa using multivariate techniques and distance- and raw-data-based regression approaches. Environmental variables (soil/water), geographic distance, and rivers were associated with the distribution of most taxa. The wide and relatively old Amazon River tended to determine differences in community composition for most biological groups. Despite this association, environment and geographic distance were generally more important than rivers in explaining the changes in species composition. The results from multi-taxa comparisons suggest that variation in community composition in Amazonia reflects both dispersal limitation (isolation by distance or by large rivers) and the adaptation of species to local environmental conditions. Larger and older river barriers influenced the distribution of species. However, in general this effect is weaker than the effects of environmental gradients or geographical distance at broad scales in Amazonia, but the relative importance of each of these processes varies among biological groups.

**Keywords** Tropical forest  $\cdot$  Species distribution  $\cdot$  Community composition  $\cdot$  Endemism areas  $\cdot$  Environmental filtering  $\cdot$  Ferns  $\cdot$  Termites  $\cdot$  Birds  $\cdot$  Bats  $\cdot$  Gingers  $\cdot$  Fish  $\cdot$  Palms  $\cdot$  Ants  $\cdot$  Butterflies

Communicated by Pedro V. Eisenlohr.

Published online: 02 September 2020

Cristian Dambros and Gabriela Zuquim have equally contributed to the study.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s10531-020-02040-3) contains supplementary material, which is available to authorized users.



#### Introduction

Identifying and understanding patterns in species distributions is essential for conservation planning and has long been recognized as crucial for defining conservation strategies in Amazonia (Guisan and Zimmermann 2000; Guisan and Thuiller 2005). Amazonian forests exhibit considerable internal heterogeneity (Emilio et al. 2010), but general knowledge of the distribution of the Amazonian biota is still limited: collection density is low, and taxonomically and geographically biased (Nelson et al. 1990; Hopkins 2007). Moreover, there is no consensus on the role of environmental and historical factors in predicting species composition at different spatial scales. Therefore, biogeographical studies can identify areas with unique sets of species and help to achieve the goal of preserving a representative mosaic of Amazonian habitats and the species they harbor.

How and why species composition varies among sites are some of the most frequent questions in ecology and biogeography. At broad scales, the distribution of organisms in space results from synergistic effects of species adaptations to the environment (Nekola and White 1999; Tuomisto et al. 2003) and diversification due to dispersal limitation (Hubbell 2001; Warren et al. 2014). While deterministic species responses to environmental conditions can give rise to patchy species distributions (Tuomisto et al. 2003), dispersal limitation and allopatric speciation can lead to differences in species composition across barriers and distant areas (Hubbell 2001; Warren et al. 2014).

In Amazonia, the most obvious potential dispersal barriers for terrestrial organisms are large rivers and associated floodplains. Accordingly, the Amazon River and its main tributaries have been recognized as important boundaries for the distribution of vertebrates for more than a century (Wallace 1852; Haffer 1974; Cracraft 1985; Moritz et al. 2000; Ribas et al. 2012; Boubli et al. 2015). The hypothesis that the development of the drainage system was a driver and maintainer of this pattern through allopatric speciation and/or preventing secondary contact between distinct populations (Wallace 1852; Ribas et al. 2012; Naka and Brumfield 2018) has been supported by occurrence (Cracraft 1985; Pomara et al. 2014) and phylogenetic data (Aleixo 2006; Ribas et al. 2012; Fernandes 2013; Naka and Brumfield 2018; Silva et al. 2019) for understory upland forest birds. The unique composition of understory bird and primate communities in different interfluves has led to the division of Amazonia into bird endemism areas delimited by large rivers, such as the Madeira, Tapajós, Rio Negro and Amazonas (Cracraft 1985; da Silva et al. 2005a, b). These divisions are widely used in conservation planning and are among the criteria for the definition of Amazonian ecoregions (Dinerstein et al. 2017).

Although the position of large Amazonian rivers matches the limits of the distributions of many understory birds and primates (Wallace 1852; Boubli et al. 2015; Silva et al. 2019; Maximiano et al. 2020), there has been controversy about to what degree species distribution patterns are related to rivers, especially when extrapolating for a wide range of organisms (Oliveira et al. 2017; Santorelli et al. 2018). For example, Santorelli et al. (2018) found that only 4 species with detectability above 50% out of almost 2000 species of the 14 taxonomic groups studied had their distributions delimited by the Madeira River. Among these 4 species were 2 birds and 2 primates, but no plant, invertebrate, or herpetofaunal species were found to have limits associated with the river at the studied localities. Intraspecific genetic structure associated with the position of Rio Negro river was reported for one tree species (Nazareno et al. 2017, 2019). Other studies have documented little to no river-barrier effect on ants (Souza et al. 2016; but see Winston et al. 2017), lizards (Souza et al. 2013), plants (Pomara et al. 2014; Tuomisto et al. 2016) and termites



(Dambros et al. 2017) at the community level. For these taxa, the main causes of differences in species distributions were associated with geographic distance (isolation by distance) or environmental differences.

Even in the absence of dispersal barriers, different parts of Amazonia may harbor different floras and faunas simply due to isolation by distance; natural populations are never panmictic because individuals typically disperse only a limited distance from where they are born (Hubbell 2001). In addition, spatially structured environmental heterogeneity related to environmental factors, such as soil properties and climate, can lead to differences in species-assemblage composition because distinct sets of species are favored in different environmental settings (Leibold and Mikkelson 2002; Tuomisto et al. 2003; Zuquim et al. 2014). Although some studies have controlled for both riverine barrier position and environmental heterogeneity in specific taxa (Pomara et al. 2014; Tuomisto et al. 2016; Maximiano et al. 2020), few studies have tried so far to disentangle the combined influence of dispersal barriers, geographic distance and environmental heterogeneity for a broad range of distinct taxonomic groups (but see Gascon et al 2000). Nevertheless, to conclude that a pattern observed is due to a single cause, it is important to consider the alternatives. Thus, we here advance in the still opened question on to what degree rivers have constrained species movements through time; and to what degree environmental filtering triggered by environmental differences among areas sampled in different sides of rivers drive species distributions for different taxonomic groups (Tuomisto and Ruokolainen 1997; Colwell 2000).

Species responses to the presence of barriers and environmental conditions are influenced by their dispersal capacity and width of tolerance to abiotic gradients (Pomara et al. 2014). However, biogeographic studies in Amazonia have generally tackled only one or a few taxa at a time, limiting their conclusions to the taxonomic group studied. Given that differences in sampling region, sampling design, length of environmental gradients and spatial extent among studies for different taxa can influence the results (Gilbert and Bennett 2010; Tuomisto et al. 2012), comparisons among these studies are questionable. Contrasting findings among different taxonomic groups or species may reflect different responses of taxonomic groups to the environment or to the presence of dispersal barriers. However, they may also be a consequence of differences in sampling schemes or in statistical methods employed (Fortin and Dale 2009). So far, no wide-scale comprehensive multi-taxa standardized assessment of the role of geographical distance, environment, and rivers to Amazonian biodiversity has been carried out. To draw general biogeographic conclusions, data collected using standardized protocols over large areas are necessary.

To understand how riverine barriers, contrasting environments (Tuomisto and Ruokolainen 1997; Tuomisto 2007), and spatial distance relate to the patterns of distribution of species belonging to a broad range of taxonomic groups, we integrate occurrence and abundance data collected using the same spatial grain. The analyses include data on three plant groups, three invertebrate groups, and three vertebrate groups sampled in forests and streams across all the major biogeographic regions of Amazonia. All surveys were based on a standardized sampling design, which allowed comparison of most taxa across the same river boundaries and along the same climatic and geographic gradients.



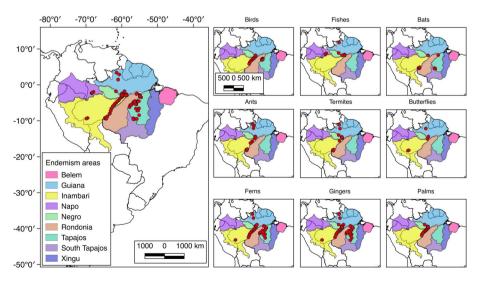
#### Methods

#### Data

## Sampling design

We compiled data generated by researchers from the Brazilian Biodiversity Research Program (PPBio), which adopts standardized protocols to create a comparable multi-taxa dataset. We sampled nine Amazonian lowland taxa: birds, fishes, bats, ants, termites, butterflies, ferns + lycophytes (hereafter called only ferns, for simplicity), gingers (Zingiberales), and palms. A total of 822 plots were sampled, and these were placed in 32 sampling grids of 5 to 72 plots each (Fig. 1). Within each grid, plots were regularly spaced, and the nearest neighboring plots were separated by a distance of 1 km. The distance between grids varied from dozens of kilometers to 1850 km, and grids were spread over an area encompassing about 2 million km<sup>2</sup>. Most plots represented tall, dense, lowland terrafirme tropical forest, but a few were established in white-sand vegetation that has a simpler structure (locally known as *Campinas* and *Campinaranas*).

For terrestrial taxa, each plot had a 250-m-long centerline following the terrain contour, to minimize within-plot variation in edaphic conditions. Plot width was adjusted for each taxonomic group due to differences in species density, diversity, and detectability (Magnusson et al. 2005). Aquatic plots were established in forest streams that are not subject to seasonal floods. Fishes were sampled in 50-m-long stretches along each stream (Mendonça et al. 2005). Except for one location, streams surveyed for fishes were located within the same sampling grids where other taxa were sampled and, in most cases, other taxa have been sampled in plots adjacent to the streams. Although not all terrestrial taxa were sampled in every plot, all plots were surveyed for at least two biological groups, and each



**Fig. 1** Location of the 32 sampling grids containing 822 permanent plots (red hexagons) of the Brazilian Program for Biodiversity Research Brazilian in Amazonia. Endemism areas are based on Cracraft 1985 with the addition of the Negro endemism area (Borges and da Silva 2012) and Tapajos Sul endemism area to the south of the Teles-Pires River based on Bates et al. (2004). Maplets show the plots (red hexagons) sampled for each of the groups



taxon was surveyed in at least two Amazonian endemism areas (Cracraft 1985). Moreover, there was overlap in the geographic distribution of surveys for all taxa. The number of plots sampled for each biological group varied from 45 (butterflies) to 475 (gingers), and the number of sampling grids varied from 4 (birds) to 20 (gingers) (Table S1).

# **Biological surveys**

Birds, fishes, bats, ants, termites, butterflies, ferns, gingers, and palms were sampled along a 250-m central line. The database of each biological group and species identities were carefully reviewed and taxonomically harmonized by specialists.

Understory birds, bats, butterflies, and ants were sampled using traps installed along the 250 m central line. Birds and bats were sampled with mist nets of 32 mm for birds and 19mm mesh for bats. Butterflies were captured using Van Someren-Rydon traps with fermented-fruit baits installed every 50 m along the central line and checked daily during 5 consecutive days. Ants were captured using pitfall traps, leaf-litter samples (Winkler sacks), and sardine baits on plastic plates. These three sampling methods tend to collect distinct ant assemblages according to species foraging mobility, which also reflects dispersal abilities. Pitfalls tend to trap relatively larger mobile forager ants, leaf-litter samples collect small cryptic and specialist ant fauna (Bestelmeyer et al. 2000), and sardine baits capture a small set of dominant ants from both groups (Baccaro et al. 2010). While the information about ant dispersal abilities is scarce, there is evidence that larger winged individuals may fly longer distances (Helms 2018). Therefore, analyses were done separately for the ant data obtained with each sampling method, and hereafter ants are referred to as mobile-forager (pitfall), cryptic (leaf-litter), or dominant (bait) ants. Fishes were sampled in 50 m-long aquatic plots that were blocked at both ends with fine-mesh nets, and fishes were then collected during daylight hours using seine and hand nets (Mendonça et al. 2005). Fish species were classified into categories of low or high dispersal capacity that were analyzed separately. We classified fish species with small body size, restricted habitat use and poor swimming capacity as "low dispersal capacity" species, and all others as "high dispersal capacity" species (Radinger and Wolter 2014). The adopted widths of plots for termites, palms, gingers, and ferns were 2, 4, 4, and 5 m, respectively, and all were sampled along the entire 250-m centerline. Termites were sampled in 5 to 10 sections of  $5 \times 2$  m interspaced along the centerline, with active search for nests lower than 2 m above ground, in tree logs, branches, soil, and leaf-litter. The soil was dug for a maximum of 50 cm. For ferns + lycophytes and gingers, all individuals with a leaf longer than 10 cm were counted and identified. Palm individuals with a minimum of 1 m height above ground to the tip of the highest leaf were also identified and counted. More detailed description of the sampling methodology and species identification for each group can be found in Supporting Information S1 and in Menger et al. (2017; understory birds), Espírito-Santo et al. (2009; fishes), Capaverde et al. (2018; bats), Souza et al. (2016; ants), Dambros et al. (2017; termites); Graça et al. (2017; butterflies), Zuquim et al. (2012; ferns + lycophytes); Figueiredo et al. (2014; gingers), and Costa et al. (2009; palms).

# Measured environmental variables: soil and water properties

In every terrestrial plot, topsoil (max. 10 cm depth) samples were taken every 50 m. In each plot, the samples were either mixed and taken to the soil laboratory to be analyzed as a single sample or analyzed separately, in which case average values of the six samples



were used to represent the soil characteristics of the plot. The composite sample was analyzed for soil clay content and exchangeable base-cation concentration (Ca, Mg, and K; Na concentrations were below detection limit) in the Thematic Laboratory of Soils and Plants at the National Institute for Amazonian Research (INPA), Manaus. Soil samples were not taken for 18 plots (2% of all plots) and for these, we extracted base-cation concentrations from a digital map (Zuquim 2017) that uses both direct soil measures and estimations based on the occurrence of soil-indicator fern species (Zuquim et al. 2019). To avoid circularity, fern + lycophyte inventories for the 18 plots from which soil information was derived from a digital map were not included in the analysis. Aquatic plots were environmentally characterized by the following water characteristics: pH, electric conductivity, dissolved oxygen, and temperature. These water variables were measured in the center of the stream channel and in the middle of the water column. Electrical conductivity and pH were measured using a portable Aqua-CheckTM Water Analyzer Operator (O.I. Analytical, College Station, TX, U.S.A.). Dissolved oxygen and temperature were measured using a Yellow Springs Instruments® (Yellow Springs, OH, USA.) model 58 portable oxygen meter thermometer.

## Retrieved environmental variables: tree cover and climate

We obtained percentage tree cover as modeled using the Advanced Very High Resolution Radiometer data (AVHRR; USGS 2017) to account for differences in vegetation structure. Climatic data were extracted from the Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA—Karger et al. 2017; https://chelsa-climate.org/, accessed 14/May/2017). We chose maximum temperature of the warmest month (bioclim 5), minimum temperature of the coldest month (bioclim 6) and the precipitation of the driest quarter (bioclim 17) out of the 19 bio-climatic (bioclim) variables available in CHELSA (Karger et al. 2017). We selected only these climatic variables to avoid an excessive number of correlated variables in the model and because these are the climatic variables that have been found to be strongly associated with the distributions of several taxa (Janzen 1967; Šímová et al. 2011). Bioclim 6 and 17 were highly correlated (r = 0.65, p < 0.001), therefore only bioclim 5 and bioclim 17 were included as independent predictors in subsequent models (Dormann et al. 2013).

All non-climatic variables obtained for terrestrial (tree cover, soil clay content, and soil bases) and aquatic plots (pH, electric conductivity, dissolved oxygen and temperature) were only weakly correlated with each other (r < |0.36|) and were used as independent predictor variables in regression models.

#### Classification of areas based on Amazonian rivers

Amazonia has been subdivided into areas of endemism based on the distribution of understory upland forest birds (Haffer 1974; Cracraft 1985; da Silva et al. 2005a, b). To test the relevance of these areas for Amazonian biota in general, we assigned each sampling plot to the corresponding bird area of endemism and used area membership as a predictor variable in subsequent analyses. A few plots in southern Amazonia were positioned in areas that have not previously been classified into endemism areas. Because earlier studies suggest that the Teles-Pires River is a dispersal barrier for some bird species (Bates et al. 2004), we added a new region to the South of the Tapajós endemism area



using the Teles-Pires River as a boundary between the two areas and assigned the southernmost plots to this endemism area (Fig. 1).

## **Analyses**

Two analytical approaches have been developed to tease apart the relative roles of spatial and environmental variables on species composition. These can be divided into two conceptually different groups: distance-based methods and raw-data-based methods. In distance-based methods, the response variable is a pairwise dissimilarity matrix (n  $\times$  n matrix) in which each site is compared with all other sites in turn, and the cell values quantify the degree of compositional dissimilarity between the two corresponding sites. The explanatory variables in distance-based methods are distance matrices quantifying geographical distance or degree of environmental difference between pairs of sites (Lichstein 2007; Tuomisto and Ruokolainen 2006). In raw-data methods, the response variable is a matrix where the rows represent objects (sites) and columns represent descriptors (often species). The explanatory variables are spatial coordinates and measured or estimated values of environmental factors (Borcard et al. 1992; Borcard and Legendre 2002; Dray et al. 2006). Distance-based and raw-data methods answer conceptually different questions given that the first asks if the differences in the response variable varies in relation to differences in the other factor whereas the second asks directly if the response variable varies in relation to the explanatory variables. There has been controversy about the relative merits and interpretation of the results from these methods (Tuomisto and Ruokolainen 2006; Laliberté 2008). As each approach targets a different null hypothesis, we used both approaches to obtain a more complete view of the distribution patterns of Amazonian biodiversity.

In both the distance-based and raw-data analyses presented below, we included the same predictor variables to test the effect of spatial isolation, environment, and barriers. In distance-based analysis, we used geographical distances calculated from geographical positioning to represent the effects of isolation by distance, the environmental distances to represent the effects of the environment, and differences in areas of endemism between plots to represent the effect of river barriers. In raw-data analyses, spatial coordinates, individual environmental variables, and endemism areas were directly used as predictor variables (see details below).

#### Distance-based analysis

In a first set of analyses, we asked what determines compositional dissimilarity between plots. We quantified compositional dissimilarity between plots using 1-Jaccard index, which is based on the proportion of unique species out of the total number of species observed in the two plots being compared. The Jaccard-based dissimilarity matrices were calculated for each taxonomic group separately. The explanatory distance matrices were also produced for each taxonomic group separately because different groups were surveyed in partly different sets of plots.

To test whether species dissimilarity of plots was related to geographic distance, environmental difference, or separation by a major river, we created distance matrices based on geographic location, environmental variables, and area of endemism. Each plot was georeferenced in the field using GPS and the coordinates were used to construct a matrix of geographic distances. Distance values were logarithmically transformed prior to



analysis to account for the tendency of distance decay to become slower at larger distances (Nekola and White 1999; Hubbell 2001; Tuomisto et al. 2003). Environmental-distance matrices were calculated as Euclidean distances based on environmental data. Soil bases were log-transformed prior to the calculation of the environmental distances because soil cations are usually more limiting at lower levels than at higher levels. We calculated the environmental distances independently for each variable as a simple difference between the values of each site. The endemism distance was defined as zero between plots in the same endemism area and one for plots in different endemism areas.

For each organismal group, we tested for the association between dissimilarity in species composition and the predictor distance matrices using distance-based multiple regressions and variance partitioning. Significance values were calculated by permutation using the MRM function in the ecodist package of R (Goslee and Urban 2007), which tests for the significance of each predictor variable after controlling for the effect of the other variables in the model. P-values calculated using permutation (i.e. a non-parametric test) do not assume normality in model residuals (Legendre et al. 1994); therefore, we did not test for normality.

Differences in climate variables (temperature and precipitation) between plots were correlated with geographical distance (r > 0.6, p < 0.05), so it was not possible to separate their effects on compositional dissimilarity for most groups (Fig. S3). Therefore, we will discuss the potential effects of both variables together. Moreover, the maximum temperature in the warmest month and precipitation in the driest quarter were strongly correlated with geographical distance and were not included in multiple regression models. Because differences in biogeographical regions were correlated to geographical distance, we avoid comparing the coefficients of these variables in a single model. We test the individual effects of these variables in simple regression models using distance matrices (single variable) and compare their shared effects in variance partitioning analyses. We therefore, discuss their individual and shared contributions to changes in species composition.

### Raw data analysis

In a second set of analyses, we took the raw-data approach and used the first axis of a Principal Coordinates Analysis (PCoA) based on the Jaccard dissimilarity matrix as a response variable for each taxon separately in multiple-regression models. In this case, the explanatory environmental variables for terrestrial organisms were tree cover, soil basecation concentration, and soil clay content. For fishes, we used water temperature, conductivity, pH, and dissolved oxygen. Endemism area was used as a categorical predictor variable representing the effect of rivers. Latitude and longitude were included as predictors to account for spatial gradients. Even without including high-order polynomials to represent the association of spatial gradients with the response variables, there was no spatial autocorrelation in model residuals for most groups (see results). However, spatial autocorrelation was present in the residuals for termites and gingers even after the inclusion of latitude and longitude as predictors. For these groups we used a Moran Eigenvector Map (MEM) analysis (Dray et al. 2012; Legendre and Gauthier 2014) in order to account for spatial autocorrelation in the residuals of the response variable. Spatial autocorrelation was calculated using Moran's I. Spatial autocorrelation using MEMs was only associated with fine-scale MEMs (MEMs with negative associated eigenvalue; Dray et al. 2012). Therefore, the removal of spatial autocorrelation did not change the broad-scale results shown in the simpler model using latitude and longitude as predictors. In order to simplify



analyses, use more easily interpretable spatial variables and make the results of all groups comparable, we only used the models with latitude and longitude as predictors. For each taxon, models containing all possible combinations of predictor variables were compared using the corrected Akaike Information Criterion (AICc). All environmental and spatial variables were standardized before analyses to zero mean and unit variance.

Biogeographic units are spatially structured, and it was virtually impossible to separate the effect of spatial positioning and area of endemism using raw-data approaches. Therefore, we conducted analyses using geographic positioning only or area of endemism only as predictor variables separately, each one along with environmental predictors. To test for differences in species composition between biogeographic units, we also ran a posteriori Tukey test comparing each pair of regions and correcting p-values for the use of multiple comparisons. Because the Amazon river is the largest and oldest river in the region (Hoorn et al. 2010) and has been hypothesized to have stronger effects on species composition than other rivers (e.g. Fluck et al. 2020), we expected differences between regions separated by the Amazon river to be stronger than between other regions. We present results from the Tukey test with this distinction (see "Results" section).

To investigate if differences in geographic extent of the data could affect the observed relationships between taxa and predictor variables, we re-ran all analyses described above with a subset of plots. Out of the whole dataset, we selected plots in order to restrict the geographic extent of the sampling of each taxonomic group to the same geographic extent of palms, which was the group with the smallest extent. Most results were qualitatively similar and are presented in Supporting Information II. We here present the results for the whole dataset analysis and focus on the results that are consistent regardless of the sampling extent.

To compare model coefficients between all predictors and taxonomic groups, all predictor and response variables were standardized (mean = 0; SD = 1) in both distance-based and raw-data-based approaches.

All analyses were carried out in the R environment (R Core Team 2019) using the packages vegan (Oksanen et al. 2018) and ecoDist (Goslee and Urban 2007) and functions created to automate the process of running multiple regression models on distance matrices for each different taxonomic group.

#### Results

We found a total of 1,889 species or morphospecies in the 822 sampled plots. Wideforaging ants were the group with the highest number of species per plot and in total (Table S1). Species-accumulation curves did not approach an asymptote for any of the taxa investigated (Fig. S1). The sampling of all taxa covered practically the same amplitude of climatic gradients, but the sampled gradient in soil-cation concentration was one to two orders of magnitude broader for butterflies, ferns + lycophytes, and gingers than for the other taxa (Table S1).

The strongest decay in compositional similarity with distance was observed for fish and palms (decay from 0 to 500 km: Fish<sub>low</sub> = 0.23, Fish<sub>high</sub> = 0.17, Palms = 0.20; Fig. 2; Fig. S2). These taxa also had the highest similarities between nearby sites (intercept for Jaccard index of 0.44 for palms and 0.35 for fishes; Fig. 2). A strong decay in similarity with geographic distance was also observed in ferns + lycophytes and gingers (Fig. 2), but the distance decay was weak or non-existent for bats (Fig. 2; Table 1; Fig. S2; Table S3).



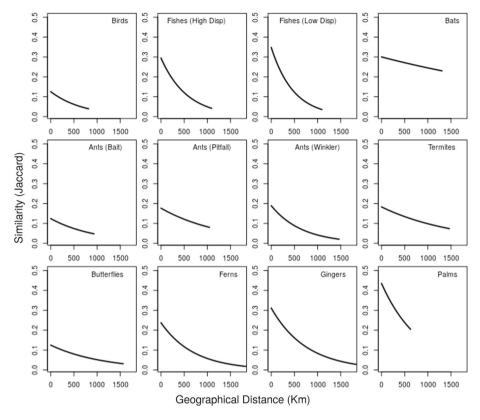


Fig. 2 Decay of similarity in species composition with geographical distance for birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms in Brazilian Amazonia. The similarity in species composition was quantified by subtracting dissimilarity values from unity (1—dissimilarity) using the Jaccard dissimilarity index calculated for each pair of plots

In regression models using distance matrices, geographical distance explained more than 17% of the variation in dissimilarity for all taxa except bats, for which geographic distance explained only 6%. Using this approach, geographical distance was associated with decay in species similarity for all taxa even after controlling for the effect of the measured environmental variables (Table 1) but it was much weaker for bats than for any other group. If only geographical distance and climate were included in the variation partitioning model, geographical distance alone explained slightly more variation in species composition than climate alone for all groups except bats, in which climate was more important (Fig. S3). However, geographical distance and climate variables were highly correlated (r > 0.6 for all taxa) and explained a similar portion of the variance in species composition for most groups, so it was difficult to disentangle the unique contribution of these variables to species turnover.

When the variance in compositional dissimilarities was partitioned among the predictor variables in multiple regression of distance matrices, endemism area was the factor with the highest unique contribution to explain changes in the dissimilarity of bird species (Fig. 3) and also had some explanatory power for palms, dominant ants, butterflies, and high-dispersal fish, but not for the other groups. Differences in soil exchangeable base-



Table 1 Distance-based regression coefficients relating indices of dissimilarity in species composition of birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms to geographical and environmental distance matrices in multiple regression models in Brazilian Amazonia. Dissimilarity in species composition was measured by the pairwise laccard dissimilarity index based on presence and absence data. All predictor and response variables were standardized (sd = 1; mean = 0) for each individual group before analyses. Maximum temperature in the warmest month and precipitation in the driest quarter were strongly correlated with geographical distance and were not included in multiple regression models (see Table S2 for simple regression coefficients). For all taxa except bats, geographical distance was a better predictor of species dissimilarity than climatic variables (Table S2). Significance of coefficients was calculated using permutation tests. Coefficients marked with \* and in bold correspond to values that were significant at p-value < 0.05. We defined the assemblages of as follow: Dominant ants were those sampled with bait; mobile-forager ants were those sampled with pitfall and cryptic ants were those sampled with Winkler

	Birds Fishes Disp)	Fishes (High Disp)	Fishes (Low Disp)	Bats	Dominant ants	Free-forager Cryptic ants		Termites	Termites Butterflies Ferns		Gingers	Palms
Distance	0.42* 0.48*	0.48*	0.317*	0.192*	0.411*	0.472*	0.563*	0.425*	0.403	0.357*	0.483*	0.489*
Tree cover	0.033			-0.06	0.198*	0.289*	-0.013	0.102*	-0.013	0.09*	0.035	0.136*
Clay	0.062*			0.023	0.023	0.138*	0.034	0.095*	-0.014	0.044	-0.028	0.226*
Soil bases	-0.001			0.171*	- 0.048	-0.008	*260.0	0.148*	0.044	0.231*	0.196*	0.158*
Oxygen		0.046	0.051									
Hd		0.11*	0.056									
Water		0.041	0.004									
temperature												
Conductivity		0.038	0.009									
$\mathbb{R}^2$	0.18	0.42	0.27	0.00	0.23	0.36	0.35	0.25	0.18	0.23	0.32	0.33



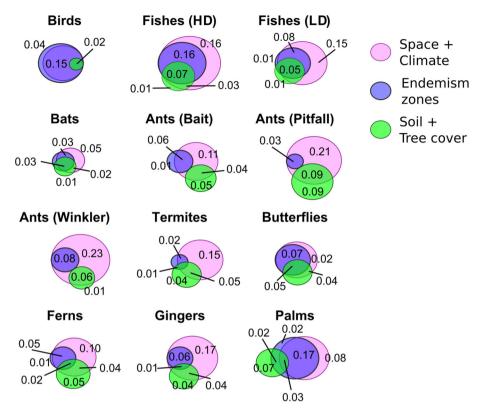


Fig. 3 Venn diagrams showing the relative contributions  $(r^2)$  of three groups of explanatory distance matrices to explaining the variation in community compositional dissimilarities (Jaccard index) of birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms in Brazilian Amazonia. Relative contributions were determined using multiple regressions on distance matrices. Overlapped areas represent the amount of variance in the response variable that was jointly explained by two or more groups of factors. Environmental distance matrices were calculated separately for each individual predictor variable: tree cover, soil clay concentration, and the sum of exchangeable cations (for terrestrial taxa), and for dissolved oxygen in the water, pH, temperature, and conductivity (for fishes)

cation concentration and tree cover had a unique contribution to compositional dissimilarities in all taxa, except birds and butterflies. The highest unique contribution of soil cations was observed in mobile-forager ants (pitfall), ferns and palms (Fig. 3, Table 1). Distance/climate explained a large and significant part of the variation in species dissimilarity even after controlling for the effects of the other variables for all groups except birds, bats, and low-dispersal-capacity fishes (Fig. 3). After controlling for spatial/climatic distances, part of the variation in compositional differences of all taxa could also be explained by differences in soil base-cation concentration (terrestrial species), water properties (fishes), tree cover and endemism area (Fig. 3).

The first PCoA axis representing the composition of species (Fig. 4) captured between 8% (mobile-forager ants) and 26% (palms) of the compositional dissimilarities between plots. Explanatory variables in raw-data-based regression models were able to capture more than 40% of the variation in community composition summarized in the first PCoA axis for all taxa (Table 2). Birds, palms, cryptic ants, and fishes with high dispersal capacity had more than 80% of the variance in their first PCoA axis explained by the



Table 2 Association of the first Principal Coordinates Axis (PCoA) of species composition of birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms with endemism regions, soil clay content, soil bases and tree cover (environmental variables), and latitude and longitude (spatial variables) in Brazilian Amazonia. Coefficients were obtained by using an Analysis of Covariance (ANCOVA). Model selection was based on AICc values and only variables included in the best model for each group are shown. The ranking of all models with delta AICc < 4 is presented as supporting material (Table S3). For Termites and Gingers, fine-scale spatial autocorrelation was found in model residuals. For these taxa, the use of MEMs instead of latitude and longitude values produced similar results (Table S3). All predictor and response variables were standardized (S = 1; mean = 0) for each group before analyses. Values in bold correspond to those with p-value < 0.05. R2 Exc.: (partial) R-squared explained exclusively by endemism zones, Environmental variables or Space. (-): Not sampled in Biogeographic region. We defined the assemblages of as follow: Dominant ants were those sampled with bait; mobile-forager ants were those sampled with pitfall and cryptic ants were those sampled with Winkler

	Birds	Fishes (high disp)	Fishes (low disp)	Bats	Dominant Mobile- ants forager	Mobile- forager ants	Cryptic ants	Termites	Termites Butterflies Ferns	Ferns	Gingers	Palms
Inambari (intercept)	- 0.988 0.833	0.833		- 0.472 - <b>1.761</b>	- 1.761	- 0.54	0.374	- 0.061	- 1.176	-0.681	0.162	- 0.747
Guiana (contrast)	2.033	ı		0.804	2.037	0.646	-0.559	0.044	1.961	0.494	0.575	1.393
Napo (contrast)	I	- 3.112		ı	ı	ı	I	ı	ı	ı	I	ı
Negro (contrast)	I	-1.654		ı	ı	ı	I	0.609	ı	ı	I	1.793
Rondonia (contrast)	0.28	1		- 0.099	I	I	- 0.378	- 0.381	I	1.42	- 1.08	I
Tapajos (contrast)	0.343	ı		0.124	1	1	ı	1	1	1.622	-1.514	1
Tapajos South (contrast)	1	ı		0.852	I	I	I	I	ı	ı	I	ı
Clay									-0.357	-0.157	0.243	0.312
Sum of Bases							-0.089	-0.279		0.455	-0.231	-0.083
Tree Cover	0.05				-0.303	0.361						
Oxygen pH												
Water temperature												
Conductivity		0.129										
Lat	0.013	<b>- 0.419</b>	-0.297	0.73	- 0.616	-0.935	-0.711	-0.707	-0.237	998.0	-0.341	-0.341 - 0.648
Long	0.256	0.059	0.702	-0.643	0.116	0.151	-0.019	0.545	0.179	-0.229	9.676	999.0
R2 Total	0.85	0.84	0.67	0.7	0.65	0.82	0.87	0.72	0.67	0.65	0.42	98.0



Table 2 continued												
	Birds	Birds Fishes (high disp)	Fishes (low Bats disp)	Bats	Dominant I ants f	Mobile- forager ants	Cryptic ants	Termites	Cryptic Termites Butterflies Ferns Gingers Palms ants	Ferns	Gingers	Palms
R2 Exc. Endemism	0.198 0.19	0.19	0 ~	0.097	0.154	0.018	0.033	0.033 0.032	0.085	0.056	0.108 0.118	0.118
R2 Exc. Environment	0.001	0 ~	0.01	0 ~	~ 0 0.067	0.123	0.002	0.039	0.084	0.148	0.072	0.074
R2 Exc. Space	0.002	0.17	0.67	0 ~	0.382	0.74	0.114	0.232	0.002	0.042	0.076	0.013



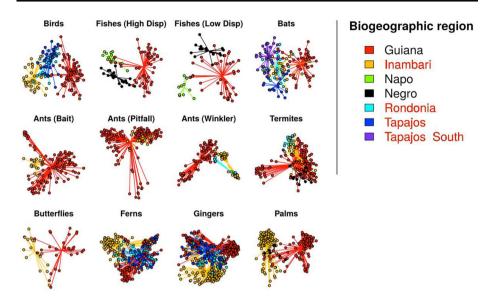
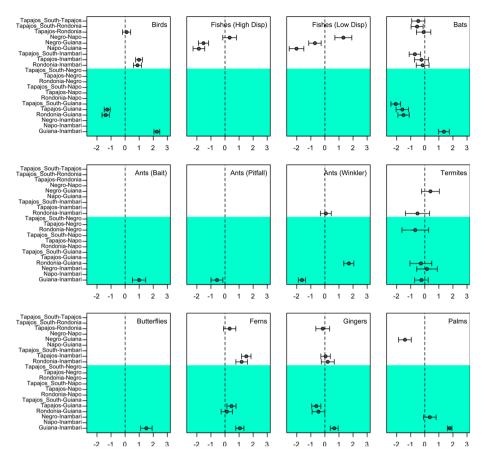


Fig. 4 Ordination plots using the first two axes of Principal Coordinates Analyses (PCoA) to summarize similarities in assemblages of birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms in each plot in Brazilian Amazonia. Colors in the graphs represent the major bird endemism areas as defined in Cracraft (1985) with our addition of an endemism zone in the southern part of Tapajós region. Names of areas of endemism north of the Amazon River are written in black and those to the south of the Amazon are written in red. Explained variance in the first two PCoA axis—birds: 0.14; fishes (high disp): 0.30; fishes (low disp): 0.35; bats: 0.24; dominant ants (bait): 0.16; mobile-forager ants (pitfall): 0.15; cryptic ants (Winkler): 0.21; termites: 0.17; butterflies: 0.20; ferns + lycophytes: 0.21; gingers: 0.22; palms: 0.36. For several groups, Inambari and Guiana (red and yellow), regions separated by the Amazon River, are the most distinct regions in species composition

environmental, spatial and biogeographic variables included in the models. For low-dispersal-capacity fishes, bats, dominant and cryptic ants, termites, butterflies, and ferns, less than 10% of the variance in the first PCoA axis could be uniquely explained by endemism areas (Table 2). In contrast, for birds, areas of endemism uniquely explained nearly 20% of the variation in PCoA axis 1, while environmental variables and space per se played virtually no unique role (Table 2). Among animals, birds were also the group with the clearest compositional differences between Inambari and Guiana endemism areas (that are separated by the Amazon River) (Table 2). The variation in species composition between endemism areas limited by the Amazon river were strong for all animal groups except termites (Fig. 5; Tables 2 and S4: contrast between Inambari and Guiana). Among plants, Inambari and Guiana endemism areas had consistently different communities, but the differences between other endemism areas were not consistent among the plant groups (Fig. 5; Tables 2 and S4: contrast between Inambari and Guiana). When restricting the analysis to the geographical extent of palm samples, most of the results were similar and suggest the same relative role of rivers, environment and geographical distance as observed in the analysis with the full dataset (Supporting Information II). However, restricting the data to central Amazonia, where all taxa have been sampled, reduced the variance explained by soil cations and distance/climate (Supporting Information II, Fig. SII-3), probably due to the shorter length of the gradient sampled as a result of the lack of samples in nutrient-rich soils and in more seasonal areas.





**Fig. 5** Means and confidence intervals of the differences in community composition of birds, fishes, bats, ants, termites, butterflies, ferns, gingers and palms among endemism regions in Brazilian Amazonia. Coefficients shown in the x-axis were estimated as the mean difference between values of the PCoA first axis representing species composition in each sampling plot. Comparison pairs including one endemism area to the North and one to the South of the Amazon river are highlighted in green shadow. Non-highlighted pairs are separated by other Amazonian rivers. See Table S4 for confidence intervals and p-values. Variance in species composition captured by the first PCoA axis—birds: 0.09; fishes (high disp): 0.19; fishes (low disp): 0.18; bats: 0.14; dominant ants (bait): 0.09; mobile-forager ants (pitfall): 0.08; cryptic ants (Winkler): 0.14; termites: 0.10; butterflies: 0.12; ferns + lycophytes: 0.13; gingers: 0.13; palms: 0.26

## Discussion

Amazonian landscape evolution affects patterns in species distribution in several ways: by determining current environmental conditions (Tuomisto et al. 2003; Pomara et al. 2014), by imposing dispersal barriers (Ribas et al. 2012) and by constraining how far a species can establish from their center of origin (Dambros et al. 2017). When species movement is not limited by geographical barriers or distance, species can be found in all habitats for which they are adapted to survive and reproduce and local environmental conditions as well as biotic interactions determine species distributions (Hurtt and Pacala 1995; Hubbell 2005). In Amazonia, studies have associated species distributions to soil conditions (Tuomisto et al. 2003; Pomara et al. 2014), position of large rivers (Haffer 1974; Cracraft 1985; da



Silva et al. 2005a, b; Boubli et al. 2015; Silva et al. 2019; Maximiano et al. 2020), climate and geographic distance (Dambros et al. 2017; Fluck et al. 2020). We found that soil conditions and geographic distance were important predictors for all taxonomic groups, but their relative importance varied among taxonomic groups, and bellow we discuss the congruence and discrepancies among biological groups in their response to these factors.

#### Riverine barriers

Rivers have been reported to act as barriers that drive and maintain species diversity of many Amazonian birds, frogs and primate species (Burney and Brumfield 2009; Ribas et al. 2012; Boubli et al. 2015; Dias-Terceiro et al. 2015; Moraes et al. 2016; Godinho and da Silva 2018; Naka and Brumfield 2018). However, we found that riverine barriers had only a weak effect on species composition for most of the animal and plant groups studied here.

Rivers could explain changes in bird community composition, but the unique effect of rivers on birds (after controlling for other variables) was not very large (4%). The largest unique component of the explained variation in bird compositional changes was related to geographic distance and endemism areas combined (15%). This suggests that when species distribution boundaries coincide with rivers, the barrier effect may actually result from a combination of factors, especially at broad spatial scales. However, when only the plots in central Amazonia were considered, the relative contribution of the Amazon river to changes in species composition increased (Supporting Information SII), indicating that the effect of rivers is more evident over environmentally homogeneous areas. Similarly, when investigating changes in species composition at landscape scales, Maximiano et al. (2020) and Pomara et al. (2014) found strong evidence of bird species turnover across rivers that could not be attributed to differences in the environment or to geographical distance as such.

In most parts of Amazonia, and for most plant and animal groups, changes in species composition across rivers could equally well be explained by geographical distance or by differences in soil nutrient concentration. Understory birds and palms were the only taxa in which this was not the case, and the presence of rivers had some unique explanatory power. For understory birds, this conforms with results of earlier studies that have demonstrated that rivers are limits to species distribution, and act as primary or secondary barriers (Ribas et al. 2012; Naka and Brumfield 2018; Silva et al. 2019), even when environmental conditions are similar on both sides of the river (Pomara et al. 2014; Maximiano et al. 2020). This result also agrees with what is known about bird ecology. Many Amazonian understory birds avoid open habitats (Laurance et al. 2004), such as those found in floodplains, secondary forests (Antongiovanni and Metzger 2005), and low-canopy forest (Mokross et al. 2018). A more surprising result was the relatively strong effect of river barriers on palms. Birds are important dispersers of palm seeds (Zona and Henderson 1989), which could explain why these two taxa were congruent in their response to the Amazon River. However, data on the occurrence of palm species were restricted to a limited geographic extent, and further studies covering longer environmental gradients and larger geographic distances are needed to clarify the relative roles of rivers and other factors.

In spite of the overall weak effect of rivers on the distribution of most taxa, differences in species composition were consistently observed between the endemism areas separated by the Amazon River for most terrestrial organisms in raw-data-based analyses (Fig. 5).



This is consistent with the main west-east trans-Amazonian drainage having started to develop towards its current configuration few to several million years ago (Miocene-Hoorn et al. 2010, 2017, van Soelen et al. 2017, early Pliocene-Latrubesse et al. 2010, Neogene-Campbell et al. 2006), which allowed time for allopatric speciation or the accumulation of differences in species composition across the margins of the Amazon River (but see Rossetti et al. 2014 for an alternative hypothesis of recent origin in the Pleistocene).

For tributaries, the patterns were not as clear. For example, communities on the two sides of the Tapajós River (Tapajós and Rondonia endemism areas) were not significantly different for any of the groups investigated in this study (Fig. 5, but see Maximiano et al. 2020), which indicates that this river does not isolate communities as effectively as the Amazon river. Tributaries of the Amazon River are generally narrower and have been more dynamic over time (Ruokolainen et al. 2019; Rossetti et al. 2014; Hoorn et al. 2017), which results in a weaker effect compared to the Amazon River. Our results indicate that the Amazon River, which is the oldest, widest and with the largest discharge, has a stronger effect on species distribution and observed biogeographic patterns, whereas younger and more dynamic tributaries have weaker or no effect on most biological groups (but see e.g., Maximiano et al. 2020 and Silva et al. 2019).

# Geographic distance and environment

Our results are consistent with earlier findings of regular turnover of species along edaphic gradients in Amazonia in several plants (e.g. Tuomisto et al. 2003, 2016; Costa et al. 2009; Zuquim et al. 2012; Cámara-Leret et al. 2017) and animal groups (e.g. Menin et al. 2007; Dias-Terceiro et al. 2015; Dambros et al. 2017). The turnover is possibly a consequence of niche partitioning, in which different species are specialized to different parts of the environmental gradient (Leibold and Mikkelson 2002; Tuomisto et al. 2003; Zuquim et al. 2012). The relative importance of the environment in shaping biological assemblages may vary greatly among taxa depending on species ecological traits (Bie et al. 2012). The unique contribution of soils in explaining community turnover tended to be greater for plants than for animals. Plants obtain nutrients directly from the soil and evolve strategies that optimize the use of local resources, whereas animals obtain nutrients indirectly and have the ability to move in search of food.

Compositional differences of most taxa exhibited a strong association with geographic distance. Although these associations may represent the effect of differences in climate or in other unmeasured spatially-structured environmental variables (Tuomisto et al. 2003), the climatic variation among sites was relatively small (Table S1) and for most groups geographical distance was still important after controlling for differences in climate and soil (Fig. S3; but see bats). Therefore, it is reasonable to expect that the vast geographical distances separating areas within Amazonia are limiting the dispersal of many organisms. Geographic distance and environment may also interact, given that species adaptations to soil conditions that are patchily distributed in Amazonia create a mosaic of habitats that provides different establishment opportunities for propagules once they have reached the area.

Although we could not disentangle the effects of distance and climate, in general, the results were consistent with expectations based on the dispersal ability of the taxonomic groups: species with higher dispersal capacity were more strongly determined by local environmental conditions (Hubbell 2005). Ferns and lycophytes, the plants with the smallest propagules (and presumably the highest dispersal capacity), were more strongly



associated with soil gradients than the other plant groups (palms and gingers). Among fishes, those with high dispersal capacity were more strongly associated with water conditions than fishes with low dispersal capacity. Mobile-forager ants had stronger association with soil conditions than other ant groups, possibly because the other groups live in leaf litter rather than directly on the soil (Fig. 3). Birds were the group with the weakest association with local environmental variables (Fig. 3) but the largest joint effect of space and rivers, which suggests strong dispersal limitation. Interestingly, bats were only weakly associated with the distances in environmental variables measured. Moreover, geographical distance had almost no explanatory power for this group. In contrast to most understory birds, bats often fly long distances in densely-vegetated areas (Trevelin et al. 2013) and it is possible that unmeasured environmental variables, e.g. vegetation-clutter, terrain elevation and food (Marciente et al. 2015; Bobrowiec and Tavares 2017; Capaverde et al. 2018) could be better predictors then the ones included in our models.

## Limitations

Sampling multiple groups at common localities along the entire Amazonian region is extremely challenging. Although we have used data obtained over large areas and with high overlap for most groups, some differences in sampling between groups existed. Some groups were not sampled in the more seasonal northern Amazonia or in western Amazonia where abrupt changes in soil conditions occur (Higgins et al. 2011; Tuomisto et al. 2016), and the spatial distribution of sampled plots caused differences among taxa in the length of the gradients sampled among taxa. For example, birds and palms were mainly sampled on nutrient-poor soils in central Amazonia and thus, the effect of environment and space may be partially hidden as described in the veiled gradients concept (McCoy 2002). Besides, due to the natural spatial structure of climatic gradients, the strong correlation between climate and space prevented a better assessment of the unique effects of these factors. Finally, the degree of taxonomic knowledge varies between taxonomic groups. Hundreds of species are described in Amazonia each year especially in invertebrate and plant groups. Species complexes may hide effects of the environment when similar taxa partition their niche. In the case of birds, probably the historically most intensively studied group, analyses at the subspecific level detected a stronger river barrier effect than analyses at the species level (Maximiano et al. 2020). Consideration of these limitations should be included in planning the location of future biological surveys.

## Conclusion

So far, most studies of biodiversity-distribution patterns have addressed single taxa (Ribas et al. 2012; Zuquim et al. 2012; Dambros et al. 2017), and the rare attempts to integrate plant and animal groups have been spatially restricted to comparisons across distances of dozens to a few hundred kilometers (Landeiro et al. 2012; Pomara et al. 2014; Tuomisto et al. 2016). We used a comprehensive and standardized broad-scale dataset for several animal and plant taxa to explore how different biological groups perceive the environment and geographical barriers at a semi-continental scale. Our results are consistent with the idea that variation in community composition in Amazonia reflects both dispersal limitation (isolation by distance or large rivers) and the adaptation of species to local environmental conditions. However,the relative importance of each of these processes varies among biological groups. The wide and relatively old Amazon River tended to determine



differences in community composition for all biological groups. Soil gradients tended to be relatively good predictors for all plant groups and some animal groups such as ants and termites.

As most studies are undertaken in only a limited number of locations and for a limited number of taxa at a time, there is an urgent need for standardized surveys of biodiversity across the whole Amazonian landscape. There is great uncertainty about the distributions of most Amazonian species. Our results advance the understanding of spatial heterogeneity of Amazonian communities, providing basic information for conservation planning. A good representation of all endemism areas based on river barriers may be an important strategy for planning new conservation units for birds and primates. However, our results indicate that the importance of rivers and environmental heterogeneity in determining patterns of diversity distribution differ greatly among taxa, and that optimized conservation planning needs to be based on data from a variety of organisms with distinct life histories. For example, in addition to locating protected areas in different areas of endemism, regions of distinct habitat types within these areas should also be prioritized, uniting the interfluvial and ecological factors to maximize biodiversity conservation.

Acknowledgements We would like to dedicate this study to Maria Aparecida de Freitas whose relentless work as PPBio field manager made studies like this possible. Cida dedicated her professional life to installing and maintaining field infrastructure for ecological studies and, unfortunately, left us too early. Cida was a synonym of vivacious and spirited hard-work and was a great inspiration for us. We are grateful to the numerous people who participated in fieldwork, helped with the arrangements for field expeditions, or shared their expertise for species identification and soil analyses. We thank the Brazilian authorities for granting the permits to carry out fieldwork and collect voucher specimens. Funding was provided to the Brazilian Program for Biodiversity Research (PPBio) and the National Institute for Amazonian Biodiversity (INCT-CENBAM) through many grants from the Fundação de Amparo a Pesquisa do Estado do Amazonas (FAPEAM), the National Council for Scientific and Technological Development (CNPq), and the Coordination for the Improvement of Higher Education Personnel (CAPES). CGF (3848108/2010-1), FB (309600/2017-0), EMV (308040/2017-1), JLPS (302301/2019-4), RPL (436007/2018-5), FRCC (481297/ 2011-1; 473474/2008-5), CR (308927/2016-8) and WEM (301873/2016-0) were supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). PEDB (062.01173/2015) was supported by Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 to TE; 12401129 to JM and 1667237 to HMVES. We thank the FinCEAL program (grant to GZ), which funded a visit of CD to Finland that resulted in this publication. GZ was supported by the Finnish Cultural Foundation and the Academy of Finland (grant to HT).

Author contributions CD and GZ are interested in distribution patterns of Amazonian organisms and how these relate to the broader context of geological history and evolutionary processes, with focus on ferns and termites respectively. WM is an ecologist who planned the Brazilian Biodiversity Research Program (PPBio) and designed its sampling scheme. This study was conducted by researchers involved in Long Term Ecological Research projects within the PPBio. The RAPELD system used by the PPBio implements surveys of multiple taxonomic groups using standardized sampling protocols that allow comparisons of multiple taxa at multiple spatial scales, from local to continental. The team includes taxonomists, field- and theory-oriented students and researchers in biology and related areas that have published hundreds of papers about biodiversity in the Amazon region. More information is available at https://ppbio.inpa.gov.br/.

Funding Open access funding provided by University of Turku (UTU) including Turku University Central Hospital.

**Data availability** A tutorial, functions and R codes to replicate the analyses are available at <a href="https://github.com/csdambros/BioGeoAmazonia">https://github.com/csdambros/BioGeoAmazonia</a> and data for termites is provided as an example. Most of the PPBio data for the other taxonomic groups, spatial and environmental variables are available at <a href="https://ppbiodata.inpa.gov.br/metacatui/">https://ppbiodata.inpa.gov.br/metacatui/</a> (searches perform better in Portuguese).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

#### References

- Aleixo A (2006) Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus Xiphorhynchus (Aves: Dendrocolaptidae). Biol J Lin Soc 89:383–395
- Antongiovanni M, Metzger JP (2005) Influence of matrix habitats on the occurrence of insectivorous bird species in Amazonian forest fragments. Biol Conserv 122:441–451
- Baccaro FB, Ketelhut SM, De Morais JW (2010) Resource distribution and soil moisture content can regulate bait control in an ant assemblage in Central Amazonian forest. Austral Ecol 35:274–281
- Bates JM, Haffer J, Grismer E (2004) Avian mitochondrial DNA sequence divergence across a headwater stream of the Rio Tapajós, a major Amazonian river. J Ornithol 145:199–205
- Bestelmeyer BT, Agosti D, Leeanne F, Alonso T, Brandão CRF, Brown WL, Delabie JHC, Silvestre R (2000) Field techniques for the study of ground-living ants: an overview, description, and evaluation. In: Agosti D, Majer JD, Tennant A, Schultz T (eds) Ants: standart methods for measuring and monitoring biodiversity. Smithsonian Institution Press, Washington, DC, pp 122–144
- Bie T, Meester L, Brendonck L, Martens K, Goddeeris B, Ercken D, Hampel H, Denys L, Vanhecke L, Gucht K, Wichelen J, Vyverman W, Declerck SAJ (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. Ecol Lett 15:740–747
- Bobrowiec P, Tavares V (2017) Establishing baseline biodiversity data prior to hydroelectric dam construction to monitoring impacts to bats in the Brazilian Amazon. PLoS ONE 12(9):e0183036
- Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol Model 153:51–68
- Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. Ecology 73:1045–1055
- Borges SH, da Silva JMC (2012) A new area of endemism for Amazonian birds in the Rio Negro Basin. Wilson J Ornithol 124:15–23
- Boubli JP, Ribas C, Alfaro JWL, Alfaro ME, da Silva MNF, Pinho GM, Farias IP (2015) Spatial and temporal patterns of diversification on the Amazon: a test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. Mol Phylogenet Evol 82:400–412
- Cámara-Leret R, Tuomisto H, Ruokolainen K, Balslev H, Munch Kristiansen S (2017) Modelling responses of western Amazonian palms to soil nutrients. J Ecol 105:367–381
- Campbell KE, Frailey CD, Romero-Pittman L (2006) The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon River system. Palaeogeogr Palaeoclimatol Palaeoecol 239:166–219
- Capaverde UD Jr, Pereira LGDA, Tavares VDC, Magnusson WE, Baccaro FB, Bobrowiec PED (2018) Subtle changes in elevation shift bat-assemblage structure in Central Amazonia. Biotropica 50:674–683
- Colwell RK (2000) A barrier runs through it ... or maybe just a river. Proc Natl Acad Sci 97:13470–13472 Costa FR, Guillaumet J-L, Lima AP, Pereira OS (2009) Gradients within gradients: the mesoscale distribution patterns of palms in a central Amazonian forest. J Veg Sci 20:69–78
- Cracraft J (1985) Historical biogeography and patterns of differentiation within the south American Avifauna: areas of endemism. Ornithol Monogr 36:49–84
- Da Silva JMC, Rylands AB, da Fonseca GA (2005a) The fate of the Amazonian areas of endemism. Conserv Biol 19:689–694
- Da Silva JMC, Rylands AB, da Fonseca GAB (2005b) The fate of the Amazonian areas of endemism. Conserv Biol 19:689–694
- Dambros CS, Morais JW, Azevedo RA, Gotelli NJ (2017) Isolation by distance, not rivers, control the distribution of termite species in the Amazonian rain forest. Ecography 40:1242–1250



- Dias-Terceiro RG, Kaefer IL, de Fraga R, de Araújo MC, Simões PI, Lima AP (2015) A matter of scale: historical and environmental factors structure anuran assemblages from the Upper Madeira River, Amazonia. Biotropica 47:259–266
- Dinerstein E, Olson D, Joshi A, Vynne C, Burgess ND, Wikramanayake E, Hahn N, Palminteri S, Hedao P, Noss R, Hansen M, Locke H, Ellis EC, Jones B, Barber CV, Hayes R, Kormos C, Martin V, Crist E, Sechrest W, Price L, Baillie JEM, Weeden D, Suckling K, Davis C, Sizer N, Moore R, Thau D, Birch T, Potapov P, Turubanova S, Tyukavina A, de Souza N, Pintea L, Brito JC, Llewellyn OA, Miller AG, Patzelt A, Ghazanfar SA, Timberlake J, Klöser H, Shennan-Farpón Y, Kindt R, Lillesø JB, van Breugel P, Graudal L, Voge M, Al-Shammari KF, Saleem M (2017) An ecoregion-based approach to protecting half the terrestrial realm. Bioscience 67:534–545
- Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46
- Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model 196:483–493
- Dray S, Pélissier R, Couteron P, Fortin M-J, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, De Cáceres M et al (2012) Community ecology in the age of multivariate multiscale spatial analysis. Ecol Monogr 82:257–275
- Emilio T, Walker Nelson B, Schietti J, Desmoulière SJ-M, Espírito-Santo HMV, Costa FRC (2010) Assessing the relationship between forest types and canopy tree beta diversity in Amazonia. Ecography 33:738–747
- Espírito-Santo HMV, Magnusson WE, Zuanon J, Mendonça FP, Landeiro VL (2009) Seasonal variation in the composition of fish assemblages in small Amazonian forest streams: evidence for predictable changes. Freshw Biol 54:536–548
- Fernandes AM (2013) Fine-scale endemism of Amazonian birds in a threatened landscape. Biodivers Conserv 22:2683–2694
- Figueiredo FOG, Costa FRC, Nelson BW, Pimentel TP (2014) Validating forest types based on geological and land-form features in central Amazonia. J Veg Sci 25:198–212
- Fluck IE, Cáceres N, Hendges CD, Brum MDN, Dambros CS (2020) Climate and geographic distance are more influential than rivers on the beta diversity of passerine birds in Amazonia. Ecography 43:860–868
- Fortin M-J, Dale MRT (2009) Spatial autocorrelation in ecological studies: a legacy of solutions and myths. Geogr Anal 41:392–397
- Gascon C, Malcolm JR, Patton JL, da Silva MNF, Bogart JP, Lougheed SC, Peres CA, Neckel S, Boag PT (2000) Riverine barriers and the geographic distribution of Amazonian species. Proc Natl Acad Sci 97:13672–13677
- Gilbert B, Bennett JR (2010) Partitioning variation in ecological communities: do the numbers add up? J Appl Ecol 47:1071–1082
- Godinho MBDC, Da Silva FR (2018) The influence of riverine barriers, climate, and topography on the biogeographic regionalization of Amazonian anurans. Sci Rep 8:3427
- Goslee S, Urban D (2007) ecodist: dissimilarity-based functions for ecological analysis. R package version,
- Graça MB, Pequeno PACL, Franklin E, Souza JLP, Morais JW (2017) Taxonomic, functional and phylogenetic perspectives on butterfly spatial assembly in northern Amazonia. Ecol Entomol 42:816–826
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186
- Haffer J (1974) Avian speciation in tropical South America, with a systematic survey of the toucans (Ramphastidae) and jacamars (Galbulidae), 1st ed. Nuttall Ornithological Club, Cambridge
- Helms JA (2018) The flight ecology of ants (Hymenoptera: Formicidae). Myrmecol News 26:19-30
- Higgins MA, Ruokolainen K, Tuomisto H, Llerena N, Cardenas G, Phillips OL, Vásquez R, Räsänen M (2011) Geological control of floristic composition in Amazonian forests. J Biogeogr 38:2136–2149
- Hoorn C, Bogot-AGR, Romero-Baez M, Lammertsma EI, Flantua SGA, Dantas EL, Dino R, Carmo DA, Chemale Jr F (2017) The Amazon at sea: onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. Glob Planet Chang 153:51–65
- Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A (2010) Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927–931



- Hopkins MJG (2007) Modelling the known and unknown plant biodiversity of the Amazon Basin. J Biogeogr 34:1400–1411
- Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton
- Hubbell SP (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. Funct Ecol 19:166–172
- Hurtt GC, Pacala SW (1995) The consequences of recruitment limitation: reconciling chance, history and competitive differences between plants. J Theor Biol 176:1–12
- Janzen DH (1967) Why mountain passes are higher in the tropics. Am Nat 101:233-249
- Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M (2017) Climatologies at high resolution for the earth's land surface areas. Sci Data 4:170122
- Laliberté E (2008) Analyzing or explaining beta diversity? Comment. Ecology 89:3232-3237
- Landeiro VL, Bini LM, Costa FR, Franklin E, Nogueira A, de Souza JL, Moraes J, Magnusson WE (2012) How far can we go in simplifying biomonitoring assessments? An integrated analysis of taxonomic surrogacy, taxonomic sufficiency and numerical resolution in a megadiverse region. Ecol Ind 23:366–373
- Latrubesse EM, Cozzuol M, da Silva-Caminha SAF, Rigsby CA, Absy MA, Jaramillo C (2010) The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. Earth-Sci Rev 99:99–124
- Laurance SG, Stouffer PC, Laurance WF (2004) Effects of road clearings on movement patterns of understory rainforest birds in central Amazonia. Conserv Biol 18:1099–1109
- Legendre P, Gauthier O (2014) Statistical methods for temporal and space-time analysis of community composition data. Proc R Soc B 281:20132728–20132728
- Legendre P, Lapointe F-J, Casgrain P (1994) Modeling brain evolution from behavior: a permutational regression approach. Evolution 48:1487–1499
- Leibold MA, Mikkelson GM (2002) Coherence, species turnover, and boundary clumping: elements of meta-community structure. Oikos 97:237–250
- Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecol 188:117–131
- Magnusson WE, Lima AP, Luizão R, Luizão F, Costa FR, de Castilho CV, Kinupp VF (2005) RAPELD: a modification of the Gentry method for biodiversity surveys in long-term ecological research sites. Biota Neotrop 5:19–24
- Marciente R, Bobrowiec PED, Magnusson WE (2015) Ground-vegetation clutter affects phyllostomid bat assemblage structure in lowland Amazonian forest. PLoS ONE 10:e0129560
- Maximiano MFDA, d'Horta FM, Tuomisto H, Zuquim G, Van doninck J, Ribas CC (2020) The relative role of rivers, environmental heterogeneity and species traits in driving compositional changes in southeastern Amazonian bird assemblages. Biotropica. https://doi.org/10.1111/btp.12793
- McCoy ED (2002) The "veiled gradients" problem in ecology. Oikos 99:189-192
- Mendonça FP, Magnusson WE, Zuanon J (2005) Relationships between habitat characteristics and fish assemblages in small streams of Central Amazonia. Copeia 2005:751-764
- Menger J, Magnusson WE, Anderson MJ, Schlegel M, Pe'er G, Henle K (2017) Environmental characteristics drive variation in Amazonian understorey bird assemblages. PLoS ONE 12:e0171540
- Menin M, Lima AP, Magnusson WE, Waldez F (2007) Topographic and edaphic effects on the distribution of terrestrially reproducing anurans in Central Amazonia: mesoscale spatial patterns. J Trop Ecol 23:539
- Mokross K, Potts JR, Rutt CL, Stouffer PC (2018) What can mixed-species flock movement tell us about the value of Amazonian secondary forests? Insights from spatial behavior. Biotropica 50:664–673
- Moraes LJCL, Pavan D, Barros MC, Ribas CC (2016) The combined influence of riverine barriers and flooding gradients on biogeographical patterns for amphibians and squamates in south-eastern Amazonia. J Biogeogr 43:2113–2124
- Moritz C, Patton JL, Schneider CJ, Smith TB (2000) Diversification of rainforest faunas: an integrated molecular approach. Annu Rev Ecol Syst 31:533–563
- Naka LN, Brumfield RT (2018) The dual role of Amazonian rivers in the generation and maintenance of avian diversity. Sci Adv 4:eaar8575
- Nazareno AG, Dick CW, Lohmann LG (2017) Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. Mol Ecol 26:3636–3648
- Nazareno AG, Dick CW, Lohmann LG (2019) A Biogeographic barrier test reveals a strong genetic structure for a canopy-emergent amazon tree species. Sci Rep 9:18602
- Nekola JC, White PS (1999) The distance decay of similarity in biogeography and ecology. J Biogeogr 26:867–878



- Nelson B, Ferreira CAC, da Silva MF, Kawasaki, ML (1990) Endemism centres, refugia and botanical collection density in Brazilian Amazonia. Nature 345:714–716
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) vegan: community ecology package
- Oliveira U, Vasconcelos MF, Santos AJ (2017) Biogeography of Amazon birds: rivers limit species composition, but not areas of endemism. Sci Rep 7:2992
- Pomara LY, Ruokolainen K, Young KR (2014) Avian species composition across the Amazon River: the roles of dispersal limitation and environmental heterogeneity. J Biogeogr 41:784–796
- R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Radinger J, Wolter C (2014) Patterns and predictors of fish dispersal in rivers. Fish Fish 15:456-473
- Ribas CC, Aleixo A, Nogueira ACR, Miyaki CY, Cracraft J (2012) A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. Proc R Soc B 279:681–689
- Rossetti DF, Cohen MCL, Bertani TC, Hayakawa EH, Paz JDS, Castro DF, Friaes Y (2014) Late quaternary fluvial terrace evolution in the main southern Amazonian tributary. CATENA 116:19–37
- Ruokolainen K, Moulatlet GM, Zuquim G, Hoorn C, Tuomisto H (2019) Geologically recent rearrangements in central Amazonian river network and their importance for the riverine barrier hypothesis. Front Biogeogr 11:3.
- Santorelli S, Magnusson WE, Deus CP (2018) Most species are not limited by an Amazonian river postulated to be a border between endemism areas. Sci Rep 8:2294
- Silva SM, Peterson AT, Carneiro L, Burlamaqui TCT, Ribas CC, Sousa-Neves T et al (2019) A dynamic continental moisture gradient drove Amazonian bird diversification. Sci Adv 5:eaat5752. https://doi.org/10.1126/sciadv.aat5752
- Šímová I, Storch D, Keil P, Boyle B, Phillips OL, Enquist BJ (2011) Global species–energy relationship in forest plots: role of abundance, temperature and species climatic tolerances. Glob Ecol Biogeogr 20:842–856
- Souza SM, Rodrigues MT, Cohn-Haft M (2013) Are Amazonia rivers biogeographic barriers for lizards? A study on the geographic variation of the spectacled lizard *Leposoma osvaldoi* Avila-Pires (Squamata, Gymnophthalmidae). J Herpetol 47:511–519
- Souza JLP, Baccaro FB, Landeiro VL, Franklin E, Magnusson WE, Pequeno PACL, Fernandes IO (2016)
  Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants. Divers Distrib 22:111–122
- Trevelin LC, Silveira M, Port-Carvalho M, Homem DH, Cruz-Neto AP (2013) Use of space by frugivorous bats (Chiroptera: Phyllostomidae) in a restored Atlantic forest fragment in Brazil. For Ecol Manag 291:136–143
- Tuomisto H (2007) Interpreting the biogeography of South America: commentary. J Biogeogr 34:1294–1295
- Tuomisto H, Ruokolainen K (1997) The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. Biodivers Conserv 6:347–357
- Tuomisto H, Ruokolainen K (2006) Analyzing or explaining beta diversity? Understanding the targets of different methods of analysis. Ecology 87:2697–2708
- Tuomisto H, Ruokolainen K, Yli-Halla M (2003) Dispersal, environment, and floristic variation of western Amazonian forests. Science 299:241–244
- Tuomisto H, Ruokolainen L, Ruokolainen K (2012) Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning results. Ecography 35:961–971
- Tuomisto H, Moulatlet GM, Balslev H, Emilio T, Figueiredo FOG, Pedersen D, Ruokolainen K (2016) A compositional turnover zone of biogeographical magnitude within lowland Amazonia. J Biogeogr 43:2400–2411
- USGS (2017) https://www.landcover.org/data/treecover/data.shtml. Accessed 15 May 2017. The data is currently available at https://earthexplorer.usgs.gov/.
- van Soelen EE, Kim J-H, Santos RV, Dantas EL, de Almeida FV, Pires JP, Roddaz M, Damsté JSS (2017) A 30 Ma history of the Amazon River inferred from terrigenous sediments and organic matter on the Ceará Rise. Earth Planet Sci Lett 474:40–48
- Wallace AR (1852) On the monkeys of the Amazon. Proc Zool Soc Lond 20:107-110
- Warren DL, Cardillo M, Rosauer DF, Bolnick DI (2014) Mistaking geography for biology: inferring processes from species distributions. Trends Ecol Evol 29:572–580
- Winston ME, Kronauer DJ, Moreau CS (2017) Early and dynamic colonization of Central America drives speciation in Neotropical army ants. Mol Ecol 26:859–870
- Zona S, Henderson A (1989) A review of animal-mediated seed dispersal of palms. Selbyana 11:6–21 Zuquim G (2017) Soil exchangeable cation concentration map of the Amazonian. GeoTif



- Zuquim G, Tuomisto H, Costa FRC, Prado J, Magnusson WE, Pimentel T, Braga-Neto R, Figueiredo FOG (2012) Broad Scale Distribution of ferns and lycophytes along environmental gradients in central and Northern Amazonia, Brazil. Biotropica 44:752–762
- Zuquim G, Tuomisto H, Jones MM, Prado J, Figueiredo FOG, Moulatlet GM, Costa FRC, Quesada CA, Emilio T (2014) Predicting environmental gradients with fern species composition in Brazilian Amazonia. J Veg Sci 25:1195–1207
- Zuquim G, Stropp J, Moulatlet GM, Van-Doninck J, Quesada CA, Figueiredo FOG, Costa FRC, Ruokolainen K, Tuomisto H (2019) Making the most of scarce data: mapping soil gradients in data-poor areas using species occurrence records. Methods Ecol Evol 1:12. https://doi.org/10.1111/2041-210x.13178

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## **Affiliations**

- ☐ Gabriela Zuquim gabriela.zuquim@utu.fi
- Department of Ecology and Evolution, Federal University of Santa Maria, Santa Maria, Brazil
- <sup>2</sup> Department of Biology, Aarhus University, Aarhus, Denmark
- Department of Biology, University of Turku, Turku, Finland
- <sup>4</sup> Facultad de Ciencias de la Tierra y Agua, Universidad Regional Amazónica Ikiam, Tena, Ecuador
- <sup>5</sup> Biodiversity Coordination, National Institute for Amazonian Research, Manaus, Brazil
- Department of Biology, Federal University of Amazonas, Manaus, Brazil
- Biological Dynamics of Forest Fragments Project, National Institute for Amazonian Research, Manaus, Brazil
- Department of Ecology, University of Brasília, Brasília, Brazil
- Programa Nacional de Pós-Doutorado, Programa de Pós-Graduação em Ecologia, Institute of Biology, University of Campinas, Campinas, Brazil
- Nucleus of Research in Aquatic Science and Fisheries, Federal University of Pará, Belém, Brazil
- Graduate Program in Ecology, Federal University of Rio Grande do Norte, Natal, Brazil
- Graduate Program in Ecology, Biological Sciences Sector, Federal University of Paraná, Curitiba, Brazil
- Department of Genetics, Ecology and Evolution, Federal University of Minas Gerais, Belo Horizonte, Brazil
- Federal Institute of Education, Science and Technology of Amazonas, Campus Presidente Figueiredo, Brazil



- <sup>15</sup> National Institute of Atlantic Forest, Santa Teresa, Brazil
- Department of Systematics and Ecology, Federal University of Paraíba, João Pessoa, Brazil
- Department of Zoology, Federal University of Minas Gerais, Belo Horizonte, Brazil
- Pontifical Catholic University of Paraná, Campus Toledo, Brazil
- Department of Ecology, Federal University of Rio Grande do Norte, Natal, Brazil
- Federal Institute of Education, Science and Technology of Amazonas, Campus Coari, Brazil

