

RESEARCH ARTICLE

Journal of
Biogeography

WILEY

Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest

Andrea Paz^{1,2} | Jason L. Brown^{1,3} | Carlos L. O. Cordeiro⁴ | Julian Aguirre-Santoro^{2,5} | Claydson Assis⁶ | Renata Cecilia Amaro⁶ | Fabio Raposo do Amaral⁷ | Thuane Bochorny⁸ | Lucas F. Bacci⁸ | Mayara K. Caddah⁹ | Fernando d'Horta⁶ | Miriam Kaehler⁶ | Mariana Lyra¹⁰ | Carlos Henrique Grohmann¹¹ | Marcelo Reginato¹² | Karina Lucas Silva-Brandão¹³ | André Victor Lucci Freitas¹⁴ | Renato Goldenberg¹⁵ | Lúcia G. Lohmann⁶ | Fabián A. Michelangeli^{2,16} | Cristina Miyaki¹⁷ | Miguel T. Rodrigues⁶ | Thiago S. Silva^{4,18} | Ana C. Carnaval^{1,2}

¹Department of Biology, City College of New York, New York, NY, USA

²Ph.D. Program in Biology, Graduate Center, City University of New York, New York, NY, USA

³Department of Zoology, School of Biological Sciences, Southern Illinois University, Carbondale, IL, USA

⁴Instituto de Geociências e Ciências Exatas, UNESP-Universidade Estadual Paulista, Departamento de Geografia, Observatório de Dinâmicas Ecológicas, Rio Claro, SP, Brazil

⁵Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá, Colombia

⁶Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil

⁷Departamento de Ecologia e Biologia Evolutiva, Universidade Federal de São Paulo, São Paulo, SP, Brazil

⁸Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP, Brazil

⁹Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

¹⁰Departamento de Biodiversidade & Centro de Aquicultura (CAUNESP), Instituto de Biociências, Universidade Estadual Paulista, Rio Claro, SP, Brazil

¹¹Institute of Energy and Environment, Universidade de São Paulo, São Paulo, SP, Brazil

¹²Departamento de Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

¹³Universidade Federal do ABC, Santo André, UFABC, SP, Brazil

¹⁴Departamento de Biologia Animal & Museu da Biodiversidade, Instituto de Biologia, Unicamp, Campinas, São Paulo, Brazil

¹⁵Universidade Federal do Paraná, Curitiba, PR, Brazil

¹⁶Institute of Systematic Botany, The New York Botanical Garden, New York, NY, USA

¹⁷Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil

¹⁸Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, UK

Correspondence

Andrea Paz, Department of Biology, City College of New York, New York, NY, USA.
Email: paz.andreita@gmail.com

Funding information

Fulbright Colombia; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2013/50297-0, 2011/50225-3, 2017/26162-8, 2003/10335-8, 2011/50146-6, 2011/50143-7, 2011/23155-4 and 2018/03428-5; National Science Foundation, Grant/Award Number: DEB 1343578, DEB-1343612 and DEB 1256742; NASA; Conselho Nacional de

Abstract

Aim: There is little consensus on which environmental variables are best at predicting multiple dimensions of diversity. We ask whether there are common environmental correlates of diversity, despite ecological differences, across nine clades of plants and animals distributed along a single rainforest domain. For that, we compare the environmental correlates of species richness, phylogenetic diversity, and phylogenetic endemism.

Location: Brazilian Atlantic Forest.

Desenvolvimento Científico e Tecnológico, Grant/Award Number: 303834/2015-3, 310871/2017-4, 151281/2013-0, 232621/2014-4, 304413/2018-6, 303713/2015-1, 306204/2019-3, 301778/2015-9, 301778/2015-9, 470939/2004-4 and 312697/2018-0; RedeLep-SISBIOTA-Brasil/CNPq, Grant/Award Number: 563332/2010-7

Editor: Fumin Lei

Taxon: Five clades of plants (Bromelioideae, Miconieae, *Bertolonia*, Cambessedesieae, and the *Fridericia* and allies) and four clades of animals (butterflies in the tribe Ithomiini, frogs in the genera *Boana* and *Proceratophrys*, and birds in the subfamily Thraupinae).

Methods: Using curated occurrence localities and phylogenetic data, we generated maps of (a) species richness, (b) phylogenetic diversity, (c) residuals of phylogenetic diversity regressed on species richness, and (d) phylogenetic endemism for all groups. We also compiled a set of 30 environmental descriptors, including records of current temperature and precipitation, climatic stability over time, and topography. Through a machine learning framework, we explored the environmental correlates of each of these diversity measures for each group.

Results: The environmental variables used here were strong predictors of diversity for all studied groups. However, models for phylogenetic endemism had lower predictive power. Although patterns of diversity are different among groups, correlates of diversity are consistent across taxa. For both species richness and phylogenetic diversity, current precipitation and precipitation stability over time were consistently ranked among the variables that strongly correlate with diversity patterns. The correlates of phylogenetic endemism were less homogeneous across groups. The results suggest that including climate stability over time is important when predicting diversity measures that reflect historical components.

Main conclusions: Investigating environmental correlates of diversity for multiple clades and diversity measures in a single geographic area allows for a better understanding of common patterns across taxa. This study shows that environmental conditions, particularly precipitation, are good predictors of the patterns of species richness and phylogenetic diversity - but not phylogenetic endemism-across different Atlantic Forest groups.

KEYWORDS

biodiversity correlates, climate stability, phylogenetic diversity, phylogenetic endemism, precipitation, species richness

1 | INTRODUCTION

Over the last decade, stakeholders from governmental, academic, and conservation organizations have shown a growing interest in the creation of systems that remotely monitor biodiversity over broad spatial scales (Scholes et al., 2008, 2012). A call for the establishment of a Global Observation Network (GEOBON) followed, along with a proposal to establish standardized measurements of essential biodiversity variables (Pereira et al., 2013; Scholes et al., 2012). With the increasing amount of satellite data being now freely available to the public, direct near-real time monitoring of biodiversity has become a reality (Turner, 2014). Examples include global estimates of forest cover change (Hansen et al., 2013) and of plant functional diversity (Jetz et al., 2016), and even penguin population monitoring, in Antarctica (Fretwell & Trathan, 2009). However, many of the world's diverse groups of organisms cannot be directly observed through satellites, as

the case of most animals and non-canopy plants. Indirect ways to monitor biodiversity would be valuable for these less accessible organisms (Turner, 2014).

It has been proposed that environmental conditions measured *in situ* or *ex situ* (via weather stations, or satellite observations of temperature and precipitation, for instance) can be used as proxies to predict diversity metrics whenever they correlate strongly with diversity patterns (Paz et al., 2020). Despite that, the utility of environmental monitoring for biodiversity prediction remains underexplored. One potential caveat of this indirect approach to monitoring biodiversity is that little consensus exists about the environmental variables that represent good predictors of the many different dimensions of biodiversity, especially in megadiverse and threatened tropical ecosystems. While the selection of biodiversity metrics and input variables is important, it often differs across studies (Williams et al., 2012). For instance, species richness (SR) and endemism have been widely



TABLE 1 Description of the biological data included in this study (five plant clades and four animal clades). Table includes the per-clade number of species included in the analysis, the total number of species occurrences utilized per clade, the level of completeness of the species sampling in the Atlantic Forest (AF), given the representation of each clade, the genes used in phylogenetic reconstruction (for new datasets), and data sources (both original and secondary). A link to Dryad is provided in the Data Availability Statement.

Clade	Taxonomic level	# of species	# of occurrences	Completeness of dataset in the AF (approximate %)	Phylogeny used	Data Sources
Bromelioideae	Subfamily	18	3,774	67%	Phylogenetic tree in Brown et al. (2020)	(Aguirre-Santoro, 2017; Aguirre-Santoro et al., 2016; Brown et al., 2020)
Miconieae	Tribe	177	25,645	70%	Phylogenetic tree in Brown et al. (2020)	(Brown et al., 2020; Caddah, 2013; Goldenberg et al., 2008; Michelangeli et al., 2008; Reginato & Michelangeli, 2016)
Ithomiini	Tribe	55	3,269	100%	Phylogenetic tree in Brown et al. (2020)	(Brown et al., 2020)
<i>Boana</i>	Genus	19	227	76%	Phylogenetic tree in Brown et al. (2020)	(Brown et al., 2020; Vasconcelos et al., 2014)
<i>Proceratophrys</i>	Genus	19	2,065	67%	Phylogenetic tree in Brown et al. (2020)	(Brown et al., 2020)
Thraupinae	Subfamily	22	8,501	100%	Phylogenetic tree in Brown et al. (2020)	(Burns et al., 2014)
Fridericia & allies	Group	65	5,115	72%	New phylogenetic tree based on <i>ndhF</i> and <i>PepC</i> markers, available in Dryad	(Kaehler et al., 2019, this study, Dryad)
<i>Bertolonia</i>	Genus	31	744	88%	New phylogenetic tree based on <i>nrITS</i> and <i>nrET</i> , <i>atpF-atpH</i> , <i>ndhF</i> , <i>psbK-psbL</i> , <i>rbcl</i> , <i>rpl16</i> , and <i>trnS-trnG</i> , <i>ADH</i> , and <i>PCRF1</i> , available in Dryad	(Bacci et al., 2020, this study, Dryad)
Cambessedesieae	Tribe	54	1,167	81%	New phylogenetic tree based on <i>nrITS</i> , <i>nrETS</i> , <i>atpF-atpH</i> , <i>psbK-psbL</i> , <i>trnS-trnG</i> , and <i>waxy</i> , available in Dryad	(Bochorny et al., 2019, this study, Dryad)

used as biodiversity metrics, and broadly employed to describe its spatial patterns, but phylogenetic diversity (PD) and phylogenetic endemism (PE) are increasingly used to explicitly quantify the amount of evolutionary uniqueness of a region (Faith, 1992; Rosauer et al., 2009; Vane-Wright et al., 1991). Both phylogenetic diversity and phylogenetic endemism reflect how combinations of more distantly related species will encompass higher percentages of the overall evolutionary history of the community relative to combinations of closely related species (Devictor et al., 2010; Forest et al., 2007). While measures of phylogenetic diversity are often positively correlated with species richness, sites where these variables are decoupled are of particular interest because they include unusually higher or lower amounts of evolutionary history than expected given the local species pool (Devictor et al., 2010; Forest et al., 2007; Fritz & Rahbek, 2012; Safi et al., 2011; Tucker & Cadotte, 2013).

Both species richness and evolutionarily informed metrics of diversity appear to be highly correlated to environmental variation, especially temperature and precipitation. Still, the specific contribution of individual climatic or landscape descriptors appears idiosyncratic when different taxa or different biodiversity metrics are compared across regions (Laurencio & Fitzgerald, 2010; Peters et al., 2016; Rompré et al., 2007; Zellweger et al., 2016). One possible reason for the observed mismatch across systems and taxa is that few studies have evaluated the correlates of species richness for multiple taxa occupying the same ecosystem. For example, while local temperature has been flagged as a main predictor of species richness in different plants and animals along elevational gradients on Mount Kilimanjaro (Peters et al., 2016), a similar transect in Switzerland identified precipitation, temperature, and topography as better predictors of the local bird, plant, and butterfly species diversity, respectively (Zellweger et al., 2016). In the tropics, herpetological surveys throughout Costa Rica (Laurencio & Fitzgerald, 2010) indicated that topography is an important predictor of richness, while a study of terrestrial vertebrates in Papua New Guinea recovered the same pattern for all organisms studied, except for reptiles (Tallowin et al., 2017). The latter exemplifies the importance of considering multiple groups when exploring environmental correlates of diversity.

Studies of evolutionarily informed measures of diversity can also provide conflicting results. Humidity and precipitation appear related to phylogenetic diversity in amphibian communities in Brazil (da Silva et al., 2012). However, patterns of phylogenetic diversity of northern Europe beetles are related to maximum temperature (Heino et al., 2015). Furthermore, evolutionarily informed measures may be influenced by historical climates and, in particular, by how much climatic variation an area has experienced. For example, long-term climatic stability has been flagged as an important predictor of avian phylogenetic diversity at a global scale (Voskamp et al., 2017). In African frogs, phylogenetic endemism appears to be related to Quaternary climatic stability (Barratt et al., 2017).

To provide a controlled comparison across organisms living in the same biogeographical region, and inform the ability to indirectly monitor tropical biodiversity as a function of climatic conditions, we compare patterns of diversity and their environmental correlates across nine clades of ecologically distinct groups of animals and plants in the Brazilian Atlantic Forest (AF). This domain is a known biodiversity hotspot, harbouring one of the highest levels of diversity and endemism in the world (Ribeiro et al., 2009). It spans a region of complex topography and environment, making it an excellent location to study the potential drivers of diversity in dissimilar groups. Here, we specifically ask if there is a set of environmental predictors that work sufficiently well across different dimensions of diversity and taxa in the Atlantic Forest, and which could be useful in community-level prediction and indirect biodiversity monitoring.

For that, we combine geo-referenced locality data with phylogenetic information from nine target clades: five groups of plants and four groups of animals. We use these data to map species richness, phylogenetic diversity, phylogenetic endemism, and the mismatch between phylogenetic diversity and species richness (i.e., the residuals of their regression analysis). Using a machine learning framework, we then investigate how well these patterns are predicted by each one of 30 abiotic correlates obtained from weather-station data and remote sensing sources that describe spatial shifts in temperature, precipitation, humidity, and topography, as well as climatic stability over the last 120,000 years (Karger et al., 2017; Title & Bemmels, 2018; Vermote & Wolfe, 2015).

2 | MATERIALS AND METHODS

2.1 | Phylogenetic information

To generate maps of species richness (SR), phylogenetic diversity (PD), and phylogenetic endemism (PE) for each target clade, we first obtained phylogenetic information and species distribution data for the nine biological groups (five plant clades and four animal clades; see Table 1 for details about species numbers, occurrence data, and data sources). Two of the plant datasets and all animal data were downloaded from Brown et al. (2020), which includes vetted occurrence points and maximum likelihood phylogenies based on mitochondrial DNA (for the animals) and chloroplast DNA (for the plants; Table 1). This dataset includes a clade of tank-forming plants belonging to the Bromelioideae subfamily and a clade of shrubs and small trees from the Miconieae tribe in the Melastomataceae family. It also includes clearwing butterflies of the tribe Ithomiini, treefrogs from the genus *Boana*, horned frogs of the genus *Proceratophrys*, and a clade of birds belonging to the subfamily Thraupinae. Adding to this dataset, we gathered literature data for three other plant clades, including species occurrence information and maximum likelihood phylogenies produced with alternative markers (Table 1). These clades include the *Fridericia* and allies group in the Bignoniaceae tribe of the plant family



Bignoniaceae (heron referred to as “bignones”), and two additional clades of Melastomataceae: the genus *Bertolonia*, and the tribe Cambessedesieae.

2.2 | Mapping species richness, phylogenetic diversity, and phylogenetic endemism

To assess which environmental variables and landscape descriptors best explain diversity patterns in each target clade, we first superimposed the distribution maps of every species in each clade. To avoid circularity in testing for environmental predictors of biodiversity, we did not use correlative species distribution models for this step. Instead, we created alpha hulls (Burgman & Fox, 2003), a generalization of convex hulls that allows for a non-convex union of points, resulting in lower overestimation of species distributions. To create the alpha hulls, we used all occurrence data available for each species that had more than three locality points. Species with three or fewer sampling sites were represented by their individual locality points. All occurrence data were vetted by the co-authors, who bring expertise in the systematics and natural history of each of the groups sampled here. The alpha hulls were built using the R package rangeBuilder (Rabosky et al., 2016) and a dynamic selection of alpha for each species, with alpha varying in steps of 1 (Meyer et al., 2017).

Because four occurrence points may be considered too small for creating alpha shapes, we also created maps based on alpha shapes built only for those species with 11 or more occurrence points, and another set of maps based on alpha shapes built only for those species with 31 or more points. This allowed us to verify that a larger threshold did not impact the results significantly: diversity patterns changed only in a few groups known to have small-ranged species (see Appendix S1 in Supporting Information). To ensure that the use of alpha shapes did not bias our study, we also ran our analyses with an alternative method for map building—minimum convex polygons—ensuring that the predictive power of the downstream analyses, and the selection of environmental predictors, are comparable across the two methodologies (see Appendix S1).

The distribution maps were then rasterized to match the spatial resolution of the predictor variables (~10 km). We acknowledge that this method might overestimate the distribution of individual species (or omit unsampled populations), but such problems are more limiting at finer spatial scales (Peterson, 2017; Peterson et al., 2018). Maps of species richness, phylogenetic diversity, and phylogenetic endemism were built for all nine groups, following the methods described in Brown et al. (2020). Species richness was estimated by summing all species present in each cell. Phylogenetic diversity denoted the sum of branch lengths leading to all species found in a given cell. Phylogenetic endemism estimated the spatial restriction of phylogenetic diversity, and was computed by weighing the branches by their ranges (Rosauer et al., 2009). To obtain these metrics, we superimposed the maps of species ranges (alpha hulls or occurrence points, the latter used when a species was represented by ≤ 3 occurrence sites). For each group, the species maps were stacked and converted

to a community composition matrix in R. To estimate phylogenetic diversity and phylogenetic endemism per clade, we imported the group's phylogenetic tree (including branch lengths) into Biodiverse (Laffan et al., 2010), along with its community composition matrix. We then used the spatial analysis tab to calculate species richness, Faith's PD index (Faith, 1992), and phylogenetic endemism (Rosauer et al., 2009) for every cell.

2.3 | Concordance between measures: mapping the residuals of phylogenetic diversity

Because measures of PD are highly correlated with SR, particularly Faith's PD (Forest et al., 2007), we regressed those two maps (PD onto SR) and mapped the residuals of the regression for each clade. The mapped residuals highlight areas where the information from these two diversity measures is different: values higher than 0 represent areas in which PD is higher than expected, given SR; negative values depict areas with less PD than expected, given SR. Hereon, we refer to this variable for each clade as the PD residuals.

2.4 | Environmental variables

To assess how much of the spatial patterns of SR, PD, and PE are explained by environmental descriptors, we compiled environmental data for the entire extension of the forest. We opted to use 30 variables, each one describing a climatic or landscape feature that has the potential to correlate with local biodiversity metrics (Table 2). Twenty-one of those were directly obtained from public databases. The variables include a 90 m Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM, Farr et al., 2007), the Topographic Wetness Index (TWI) from Envirem (Title & Bemmels, 2018), and 19 bioclimatic variables reflecting temperature and precipitation. The latter were downscaled to a 30" (~1 km) resolution, using climatologies at high resolution for the Earth's land surface algorithm (CHELSA; Karger et al., 2017), and the ANUCLIM method (Xu & Hutchinson, 2010).

We generated the remaining nine layers: two describe the Atlantic Forest terrain, one describes cloud cover, and six reflect climatic stability over the last 120,000 years. A map of Atlantic Forest Domain was created in a GIS environment based on the Vegetation Map of Brazil at a 1:5,000,000 scale (IBGE, 2004). Upon consultation with the field biologists in our group (MTR, FAM, MR, personal communication), we used that map to identify all areas known to be occupied by the Atlantic Forest and combined them into a multi-polygon vector geometry. The map is referenced to the WGS84 datum with geographic coordinates, and it is available in the Dryad repository associated with this manuscript (see Data Availability Statement). The first two variables, Slope and Rugosity (calculated as the standard deviation of Slope, Grohmann et al., 2011), were derived from the DEM (SRTM; Farr et al., 2007). Mean cloud coverage was derived from NASA's Moderate Resolution Imaging Spectroradiometer

TABLE 2 Thirty environmental variables compiled for this study. Asterisks denote the thirteen variables used in the Machine Learning analyses, after eliminating those with high collinearity. All variables were resampled at a ~10 km resolution. The table depicts the source and category of each variable. Acronyms: SRTM, Shuttle Radar Topography Mission (Farr et al., 2007); TWI, Topographic Wetness Index; Envirem, Environmental Rasters for Ecological Modelling (Title & Bemmels, 2018); CHELSA, Climatologies at High Resolution for the Earth's Land Surface Algorithm (Karger et al., 2017); MODIS, NASA's Moderate Resolution Imaging Spectroradiometer (Vermote & Wolfe, 2015); Hadley Center, Hadley Center Model (HadCM3, Carnaval et al., 2014; Singarayer & Valdes, 2010). CV, Coefficient of Variation

Variable	Description	Source	Category
Bio 1	Annual Mean Temperature	CHELSA	Current temperature
Bio 2*	Mean Diurnal Range	CHELSA	
Bio 3	Isothermality (BIO2/BIO7) (* 100)	CHELSA	
Bio 4	Temperature Seasonality (standard deviation *100)	CHELSA	
Bio 5	Max Temperature of Warmest Month	CHELSA	
Bio 6	Min Temperature of Coldest Month	CHELSA	
Bio 7	Temperature Annual Range (BIO5-BIO6)	CHELSA	
Bio 8*	Mean Temperature of Wettest Quarter	CHELSA	
Bio 9	Mean Temperature of Driest Quarter	CHELSA	
Bio 10	Mean Temperature of Warmest Quarter	CHELSA	
Bio 11	Mean Temperature of Coldest Quarter	CHELSA	
Bio 12	Annual Precipitation	CHELSA	Current precipitation
Bio 13*	Precipitation of Wettest Month	CHELSA	
Bio 14	Precipitation of Driest Month	CHELSA	
Bio 15	Precipitation Seasonality (Coefficient of Variation)	CHELSA	
Bio 16	Precipitation of Wettest Quarter	CHELSA	
Bio 17	Precipitation of Driest Quarter	CHELSA	
Bio 18*	Precipitation of Warmest Quarter	CHELSA	
Bio 19*	Precipitation of Coldest Quarter	CHELSA	
Altitude*	Digital Elevation Model (DEM)	SRTM	Topographic
Slope*	Slope calculated from the DEM	Based on SRTM	
Rugosity	Rugosity calculated from the DEM	Based on SRTM	
TWI*	TWI calculated from the DEM	Envirem	
Clouds*	Average Cloud Coverage calculated from MODIS imagery	Based on MODIS (2000-2017)	Clouds
CV Bio 1*	CV of Bio 1 (120,000 years, every 4,000 years)	Based on Hadley Center data	Temperature stability
CV Bio 10*	CV of Bio 10 (120,000 years, every 4,000 years)	Based on Hadley Center data	
CV Bio 11	CV of Bio 11 (120,000 years, every 4,000 years)	Based on Hadley Center data	
CV Bio 12	CV of Bio 12 (120,000 years, every 4,000 years)	Based on Hadley Center data	Precipitation stability
CV Bio 16*	CV of Bio 16 (120,000 years, every 4,000 years)	Based on Hadley Center data	
CV Bio 17*	CV of Bio 17 (120,000 years, every 4,000 years)	Based on Hadley Center data	

(MODIS 09GA, Vermote & Wolfe, 2015) based on satellite data collected from 2000 to 2017, using Google Earth Engine (Gorelick et al., 2017). Lastly, we built six layers to reflect climatic stability over the last 120,000 years, to determine whether historical climates were important in predicting present-day patterns of diversity. For that, we used existing bioclimatic descriptors available every 4,000 years for the past 120,000 years, and obtained through the Hadley Center model (HadCM3, Carnaval et al., 2014; Singarayer & Valdes, 2010). We summarized long-term variation in three temperature attributes (i.e., Annual Mean Temperature [Bio 1], Mean Temperature of the Warmest Quarter [Bio 10], and Mean Temperature of the Coldest Quarter [Bio 11]), and three precipitation attributes (i.e., Annual

Precipitation [Bio 12], Precipitation of the Wettest Quarter [Bio 16], and Precipitation of the Driest Quarter [Bio 17]). For each one of the six variables, we computed the coefficient of variation over the past 120,000 years. To ensure uniformity in the resolution of the data, all variables were resampled to a 5' (~10 km) resolution, using the resample function of the R package raster 3.0-7 (Hijmans, 2019). Only then they were incorporated in the downstream analyses. For downstream analyses, all variables were resampled to a 5' resolution (~10 km) using the resample function of the R package raster 3.0-7.

The complete dataset of environmental descriptors (30 layers; Table 2) reflected variables that may be highly correlated in the Atlantic Forest area. We thus ran a Variance Inflation Factor (VIF) analysis to



reduce collinearity, using the R package *usdm* 1.1–18 (Naimi et al., 2014), and keeping only those variables with $VIF < 5$. After eliminating the highly co-linear variables, we were left with a dataset that included 13 environmental descriptors, which were used in all machine learning analyses. Together, these variables represent present-day climate (Mean Diurnal Range [Bio 2], Mean Temperature of the Wettest Quarter [Bio 8], Precipitation of the Wettest Month [Bio 13], Precipitation of the Warmest Quarter [Bio 18], and Precipitation of the Coldest Quarter [Bio 19]), Climatic Stability over the past 120,000 years [CV of Bio 1, CV of Bio 10, CV of Bio 16, and CV of Bio 17], Topography [Altitude, Rugosity, and TWI], and Cloud Cover; Table 2).

2.5 | Correlates of biodiversity

To determine which of the environmental descriptors are the best predictors of SR, PD, PE, and PD residuals, we used four machine learning algorithms to generate correlative models of each biodiversity metric. We then combined the four resulting models in an ensemble prediction for each metric. The machine learning algorithms were Random Forests (*rf* from Liaw & Wiener, 2002), Neural Network (*nnet* from Venables & Ripley, 2002), Support Vector Machines (*svmRadial* from Karatzoglou et al., 2004), and Generalized Linear Models. While running each algorithm, we randomly split each dataset (here represented by each map of a given diversity metric and for a given group) into two sets: one containing 70% of the pixels (for model training), and one containing 30% of the pixels (completely withheld for model testing). For the training of each model, we randomly split the training data into 10 subgroups, or folds. We used each fold in turns as an internal validation dataset, utilizing the other folds for training. We repeated this procedure three times, for cross-validation.

All four types of models were built with the R package *caret* 6.0–84 (Kuhn, 2016). Within *caret*, models were specified by tag, using the methods “*rf*,” “*nnet*,” “*svmRadial*,” or “*glm*” to designate random forest, neural network, support vector machine, and generalized linear models, respectively. A *trainControl* object was used to specify the resampling method, which was implemented as described above. A final ensemble model, built from a linear combination of the four algorithms and based on root mean square error values, was built with the *caretEnsemble* function in R. To obtain an estimate of variable importance in the ensemble model, we computed a weighted average of the variable contributions estimated from the individual models, using the weight of the models in the ensemble. For that, we used the *varImp* function of the *caret* package for R. The script used for these analyses is available in the Dryad repository associated with this manuscript.

3 | RESULTS

Patterns of SR, PD, and PE are different among groups (Figure 1). Bignones and *Cambessedesieae* show a concentration of SR and

PD in the interior forests, from Bahia to Minas Gerais— a result even more striking in the phylogenetic diversity maps (Figure 1). Both the *Miconieae* and tanagers have higher diversity (both SR and PD) in the Serra do Mar coastal forest (coastal mountains, Figure 1). The butterflies show peaks for both PD and SR in the coastal forests from São Paulo to central Bahia. *Bertolonia* shows the opposite pattern, with higher SR and PD both south (from São Paulo state to Santa Catarina) and northeast (Bahia coastal forest) of the high-diversity areas for the butterflies. The bromeliads have their peak diversity for both PD and SR in the northeastern region of the Bahia coastal forest, with the *Proceratophrys* frogs showing peak diversity in northern Bahia and part of the Serra do Mar coastal forest. Frogs of the genus *Boana* have two peaks of higher SR, one in the Serra do Mar coastal forest, and a second one in the Bahia interior forests. For this group, highest PD levels are observed mainly in the Serra do Mar coastal forest.

For phylogenetic endemism, we detected two general patterns: areas of high PE were either small, or widespread (Figure 1). Two groups, the butterflies and the tanager birds, show high levels of PE through relatively large portions of the forest (from inland areas to northern Rio de Janeiro in the butterflies, along the Serra do Mar coastal forest in the tanager birds). In contrast, both the bromeliads and *Bertolonia* have very small areas of high PE, both in northern Bahia. The *Miconieae*, the *Proceratophrys* frogs, and the *Cambessedesieae* show high levels of PE in small areas of Espírito Santo and at the border between the states of Rio de Janeiro and São Paulo. The *Cambessedesieae* also show a small area of high PE in the forests of inland Bahia. Finally, frogs of the *Boana* genus and bignones show no apparent areas of high PE.

All groups show spatial concentration of PD residuals (positive or negative, Figure 2). Most groups have areas where PD is higher than expected given species richness, with *Proceratophrys* frogs showing this pattern along the entire distribution (Figure 2). Butterflies have higher PD than expected given the number of species in the coastal region, extending from Alagoas to Paraíba, and in the southern interior region of the forest (red and dark orange areas in Figure 2). We found areas of higher PD than expected in the southern Atlantic Forest for another three groups: (a) the tanagers (in Paraná, Santa Catarina, and part of São Paulo), (b) the *Boana* frogs, and (c) bignones (the latter two in the Serra do Mar Coastal Forest, in the states of Santa Catarina and Paraná). The bignones, *Miconieae*, and *Cambessedesieae* showed higher PD than expected in the northern forests, around the Bahia Interior Forests region. Negative residuals, denoting areas of lower PD than expected given species richness, are more distributed in geographical space (Figure 2, green and blue, respectively). The few exceptions are observed in tanager birds, butterflies, and the plant tribe *Miconiae*. For the tanagers, areas holding less phylogenetic diversity than expected are concentrated in the north, mostly north of Minas Gerais state, including the states of Espírito Santo and Bahia. In the butterflies, these areas are found in the state of Bahia, but also in small clusters

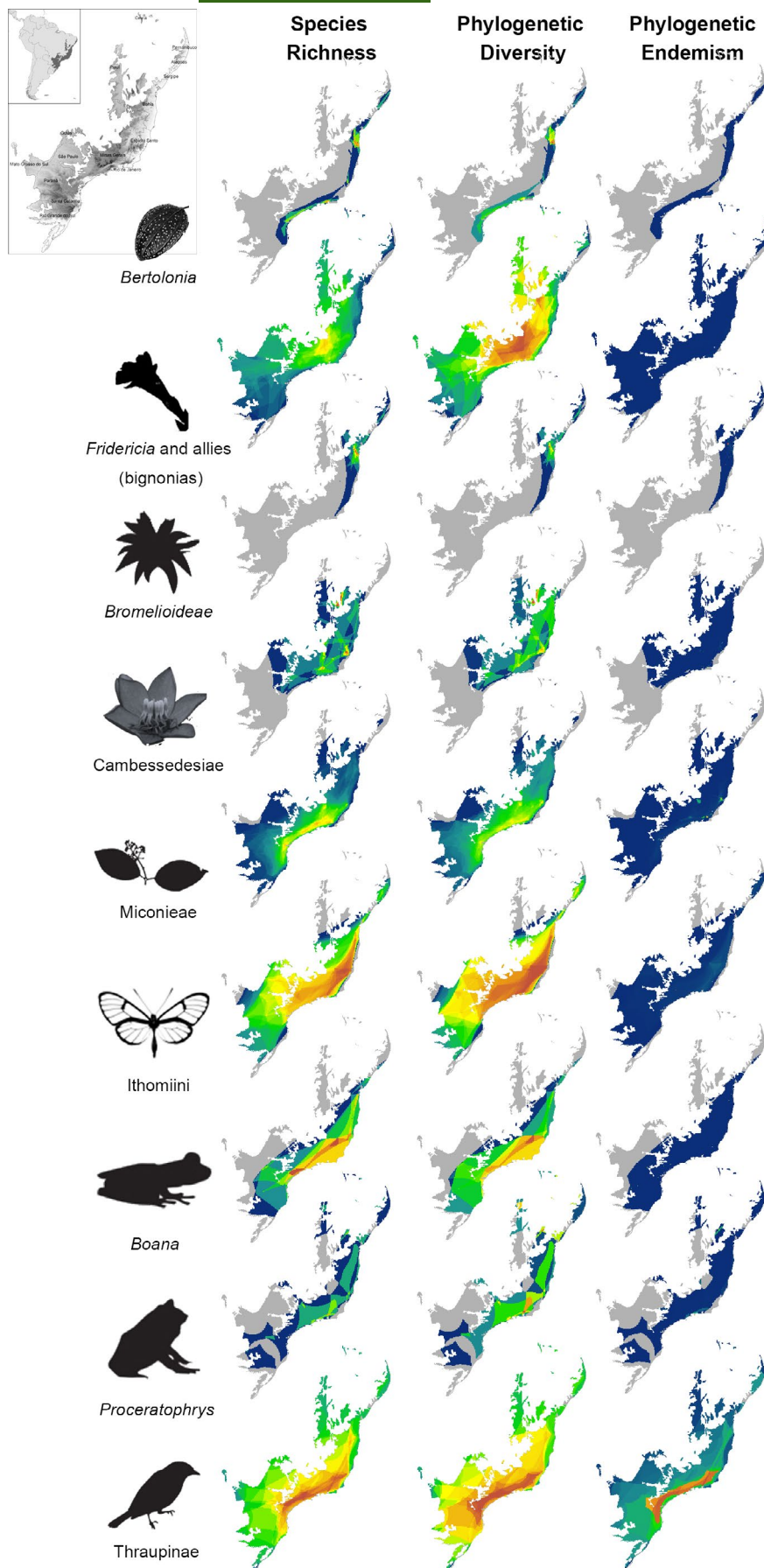


FIGURE 1 Maps of species richness (first column), phylogenetic diversity (middle column), and phylogenetic endemism (last column) for the nine study groups (rows) in the Brazilian Atlantic Forest (state boundaries in top left inset). Warmer colours represent higher diversity, colder colours depict lower diversity; for all maps, values are stretched to maximum-minimum. Taxa represented from top to bottom are: five clades of plants: *Bertolonia*, *Fridericia* and allies (bignonias), Bromelioideae, Cambessedesiae, Miconieae, and four groups of animals: the Ithomiini butterflies, the *Boana* and *Proceratophrys* frogs, and the tanager birds of the Thraupinae subfamily. In the species richness maps, the number of species for each group varies as follows: *Bertolonia* (1–5), *Fridericia* and allies (bignonias; 1–33), Bromelioideae (1–8), Cambessedesiae (1–8), Miconieae (1–81), Ithomiini (2–39), *Boana* (1–7), *Proceratophrys* (1–5), and Thraupinae (1–19). For the phylogenetic diversity maps, the value of PD for each group varies as follows: *Bertolonia* (0.11–0.21), *Fridericia* and allies (bignonias; 0.04–0.23), Bromelioideae (5.37–15.48), Cambessedesiae (0.17–0.52), Miconieae (25.73–424), Ithomiini (0.14–1.52), *Boana* (0.74–1.86), *Proceratophrys* (0.41–0.72), and Thraupinae (39.64–345.69). For the phylogenetic endemism maps, the value of PE for each group varies as follows: *Bertolonia* (4.35×10^{-5} –0.029), *Fridericia* and allies (bignonias; 2.31×10^{-6} –0.0013), Bromelioideae (0.003–0.54), Cambessedesiae (1.87×10^{-5} –0.041), Miconieae (0.001–3.64), Ithomiini (2.69×10^{-6} –0.0045), *Boana* (6.17×10^{-5} –0.1), *Proceratophrys* (1.9×10^{-5} –0.0074), and Thraupinae (0.00034–0.046) [Colour figure can be viewed at wileyonlinelibrary.com]

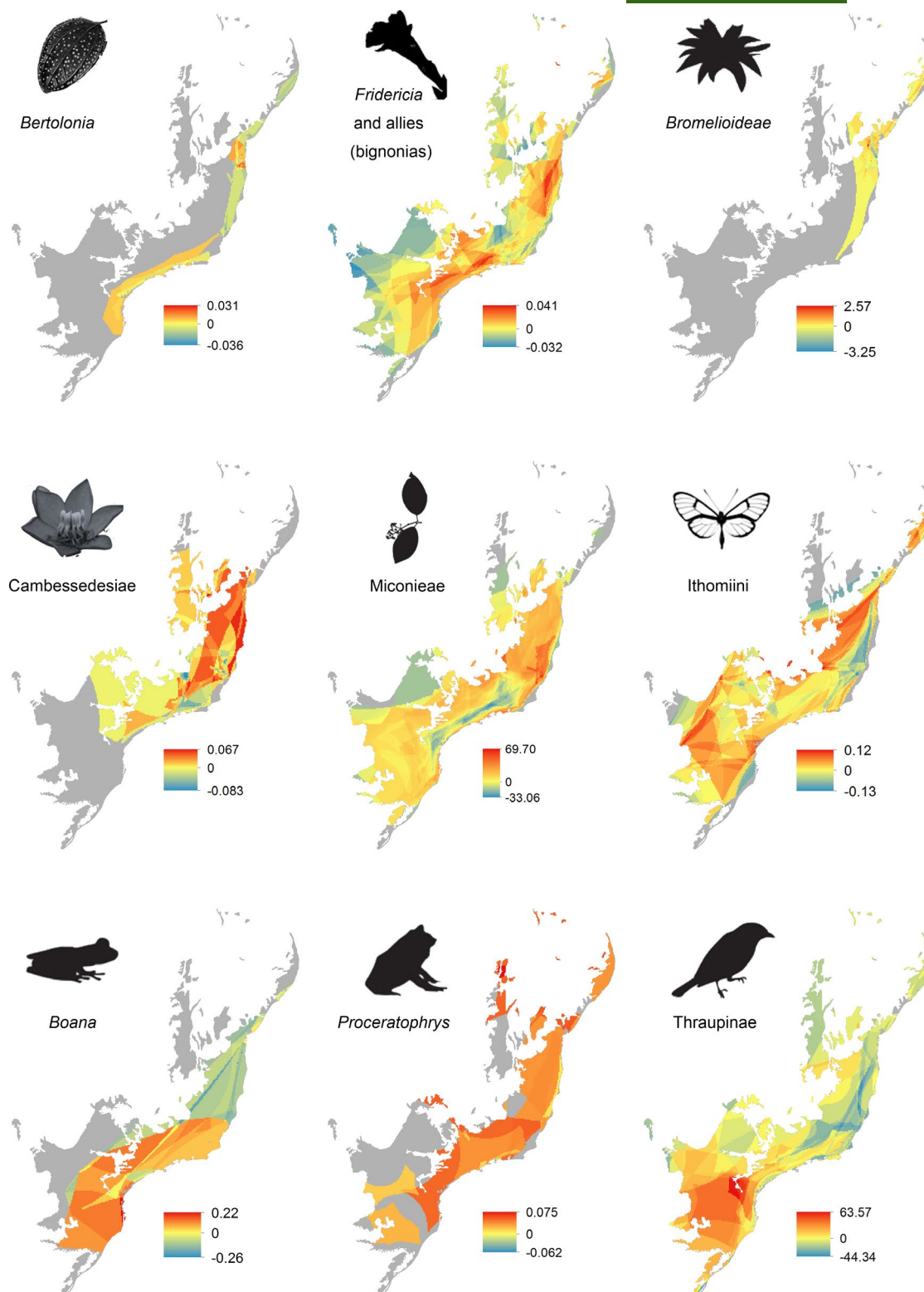


FIGURE 2 Maps of residuals of phylogenetic diversity regressed on species richness (PD residuals) for the nine study groups. Values of positive residuals are shown through a red scale; values of negative residuals are depicted through a blue scale. Taxa, represented from left to right and top to bottom, are: five groups of plants (*Bertolonia*, *Fridericia* and allies [bignonias], Bromelioideae, Cambessedesiae, and Miconieae), and four groups of animals (the Ithomiini butterflies, the *Boana* and *Proceratophrys* frogs, and the tanagers of the Thraupinae subfamily) [Colour figure can be viewed at wileyonlinelibrary.com]

in the southern portion of the forest. In the Miconieae, they are mostly found in the Serra do Mar Coastal Forests (Figure 2, blue areas).

Of all biodiversity metrics, PD (R^2 0.86–0.98) and PD residuals (R^2 0.86–0.98, Figure 3) were best predicted by the environmental models. These metrics were followed by SR (R^2 0.79–0.98). Model

predictions of PE were more heterogeneous and generally weaker (R^2 0–0.96, Figure 3), with three main exceptions: the bignones, the tanager birds, and the butterflies (R^2 of 0.94, 0.95 and 0.96 respectively). Predictions of PE in two other groups, the *Proceratophrys* frogs and the Miconieae, reflected more intermediate R^2 values (R^2 of 0.36 and 0.49 respectively).

The ability of the models to predict SR, PD, and PE also varied across clades. For instance, R^2 values were consistently high in butterflies ($R^2 > 0.96$), bignones ($R^2 > 0.93$), and tanager birds ($R^2 > 0.95$), but lower in the plants of the genus *Bertolonia* and *Proceratophrys* frogs. In fact, the latter had one of the lowest R^2 values for predictions of SR, PD diversity and residuals (Figure 3).

Climatic variables, reflecting both present-day and past conditions, contributed highly to predictions of SR, PD, and PD residuals. In particular, precipitation-related variables were consistently identified as those of higher importance to predict SR and PD (Figure 4, blue). In eight out of the nine clades, variables reflecting current precipitation were identified as those of highest importance in predictions of SR. In one clade (bromeliads), stability in past precipitation was ranked first, but closely followed by current precipitation (Figure 4). Current temperatures were ranked as the second or third most important predictors of SR. Conversely, historical stability in temperature contributed less to predictions of diversity, ranking fourth to last for SR (except for bignones, which ranked second; Figure 4a). For PD, precipitation was of higher importance in seven out of the nine clades. In one clade, the *Boana* frogs, past precipitation closely followed current precipitation. In models of the bromeliads, current temperature was the most important variable, followed by current and past precipitation. For the other groups, current temperature ranked second, third, or fourth, while temperature stability ranked third or fourth in importance (last for the bromeliads; Figure 4b).

Variables related to precipitation also had higher importance to explain PD residuals in eight out of the nine target groups (Figure 4a, b), with a slightly higher contribution of climate stability as a correlate of PD residuals relative to the other metrics (Figure 4a–c). The importance of current temperatures as predictors of PD residuals was mixed, being ranked first in the case of tanagers and butterflies, second in the case of *Bertolonia*, bromeliads and *Proceratophrys*, and third to sixth in all other clades (Figure 4c). Although generally low, cloud distribution was relevant to predicting the residuals for tanagers and *Boana* frogs, and slightly relevant for the bignones and Cambessedesiae (Figure 4).

Unlike the other biodiversity metrics, PE did not seem to be better predicted by one specific type of environmental variable. Variables with highest contributions to the model were related to current precipitation (five out of the nine groups), current temperature (one out of the nine groups), and terrain (three out of nine). Of the five groups showing PE models with some predictive power, four had current precipitation as the main predictor (bignones, birds, Miconieae, and *Proceratophrys*); in one, topography was closely followed by current precipitation. Only in this metric there is a more meaningful correlation with topography, which was recovered in at least four groups. Topography was the highest contributor for PE

patterns observed in Cambessedesiae, bromeliads, and butterflies, but ranked second and third for *Bertolonia* and Miconieae, respectively (Figure 4d).

4 | DISCUSSION

Models based on environmental variables that describe temperature and precipitation are able to predict the accumulation of different dimensions of diversity in the Atlantic Forest (Figure 3). Species richness and phylogenetic diversity are well described by climatic descriptors in all nine taxonomic groups sampled here. Phylogenetic endemism, however, is explained by environmental variables in only half of the taxonomic groups, with R^2 values ranging from 0.36 to 0.96. The fact that the predictions of PE were weaker than those of SR or PD is likely associated with the spatial restriction of this biodiversity measure in relation to the broader environmental predictors used here. In agreement with this finding, other studies have suggested that endemism might be explained by variation within broader study regions (Crisp et al., 2001; Rosauer et al., 2009). In addition, while our sampling of the AF was fairly complete, a few narrow endemic species are still missing in our datasets; as such, it is possible that PE may be underestimated in certain cells.

Sampling multiple clades within a single geographic space allowed us to identify variables that are consistently important (or not) as predictors of diversity in the Atlantic Forest. The results highlight the importance of precipitation, both in the past and present, and the small contribution of topography as an additional variable. Traditionally, temperature has been considered as the most important driver of diversity patterns globally – likely given its importance in temperate zones (Erwin, 2009; Peters et al., 2016; Rohde, 1992). However, our results point to precipitation as a main predictor of species richness and phylogenetic diversity, while also contributing to the mismatch between those variables. This result is in line with other recent clade-based studies of tropical species that highlight the importance of rainfall as a driver of species richness in small mammals (Mason-Romo et al., 2017), trees (Krishnadas et al., 2016), bats (Grimshaw & Higgins, 2017), fruit-feeding butterflies (Santos et al., 2020), and anurans (Vasconcelos et al., 2010). By analysing multiple clades in the same biogeographical domain, we flag commonalities regarding environmental correlates (and potential drivers) of diversity, highlighting the importance of precipitation at the community and ecosystem scales.

Contrary to our expectations of historical climatic stability and topography as strong predictors of evolutionary history, PD was not explained by any of those variables. However, climatic stability was an important predictor of the other two measures that reflect evolutionary history—PE and PD residuals—with topographic variables contributing to the explanation of PE. This result might be explained by the relative importance of evolutionary history in each measure. In this case, PD is highly correlated with SR, and thus the predictors

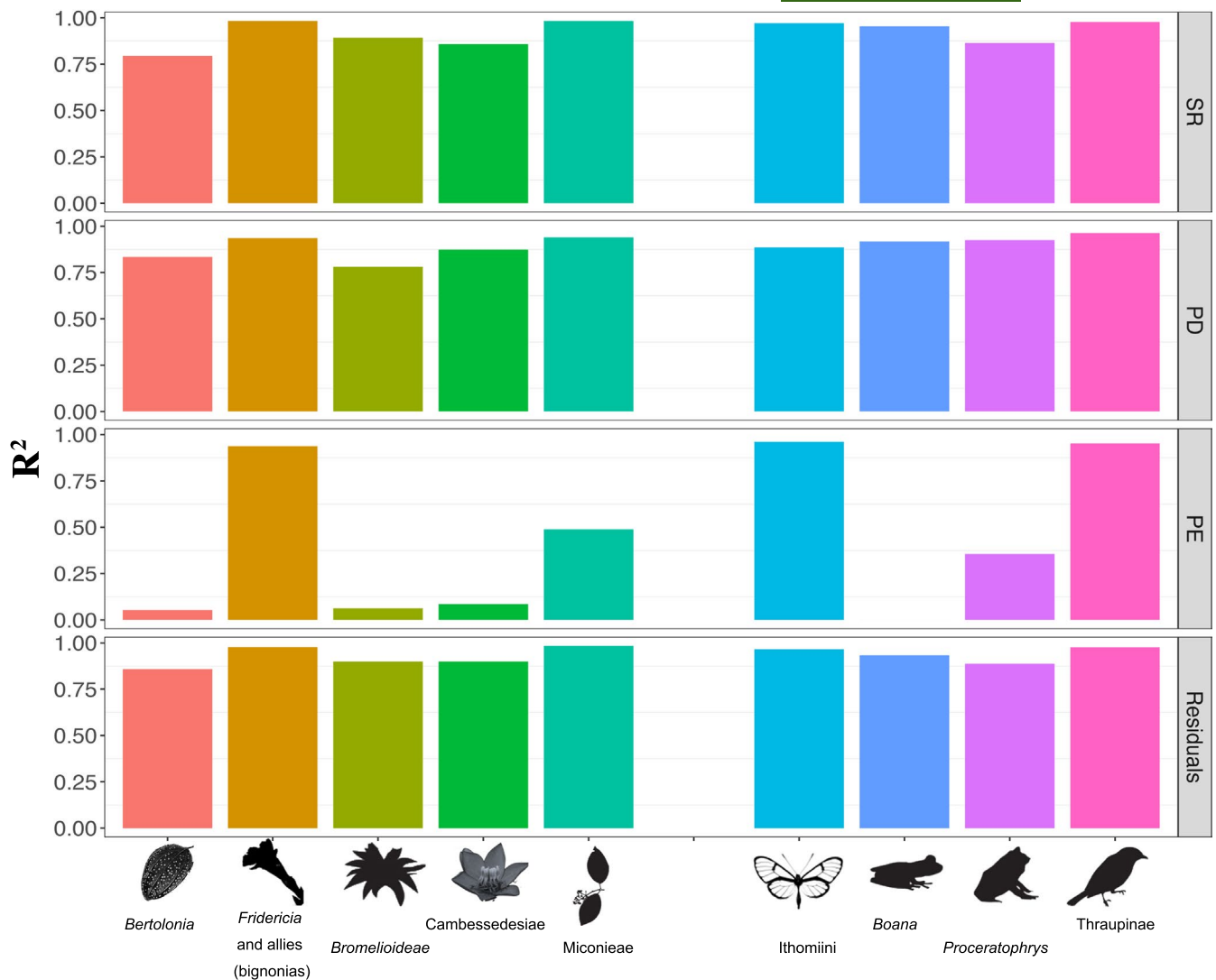


FIGURE 3 R^2 of ensemble machine learning models for each studied group and measure of diversity. From top to bottom, diversity measures are species richness (SR), phylogenetic diversity (PD), phylogenetic endemism (PE), and residuals of phylogenetic diversity regressed on species richness (residuals). The study groups, from left to right, are *Bertolonia*, *Fridericia* and allies (bignones), *Bromelioideae*, *Cambessedesiae*, *Miconieae*, *Ithomiini* butterflies, *Boana* and, *Proceratophrys* frogs, and the tanagers of the *Thraupinae* subfamily. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/jbi.14083)]

may also be driven by the latter. The residuals, however, show areas of mismatch between the two, highlighting regions where an analysis of evolutionary history provides different information (Devictor et al., 2010; Forest et al., 2007). As such, PD residuals may give us better insights into predictors of evolutionary history. At a global scale, this relationship has been studied with birds; abiotic correlates of PD residuals have been shown to be spatially heterogeneous, with elevation and ecotones as important predictors of PD residuals in the tropics (Voskamp et al., 2017). Here, we found relatively high importance of climatic stability as a predictor of PD residuals in the AF.

We found discrepancies in the strength of the predictions and the importance of variables among the nine taxonomic groups studied here. More specifically, *Proceratophrys* frogs showed the lowest predictive power in all models, which might be linked to the more restricted distributions of its species. Another example is the increased contribution of former temperature stability (and former

precipitation stability) to explain PE in bromeliads, tanagers, and frogs. This result is congruent with a previous study highlighting the importance of climatic stability for predicting PE in African frogs (Barratt et al., 2017). These discrepancies suggest that environmental correlates of diversity may be more similar among study groups with similar natural histories, rather than region-dependent. Indeed, life-history traits have been proposed to explain differences in how shared barriers lead to different levels of intraspecific isolation (or gene flow) in co-distributed species, as well as differences in the impact of climatic changes in population history and demography (Carstens et al., 2018; Pabijan et al., 2012; Paz et al., 2015; Zamudio et al., 2016).

By mapping diversity metrics of multiple taxonomic groups that co-occur in a single domain, our approach led not only to strong predictions of different dimensions of biodiversity based on past and present abiotic variables, but also highlighted which and how

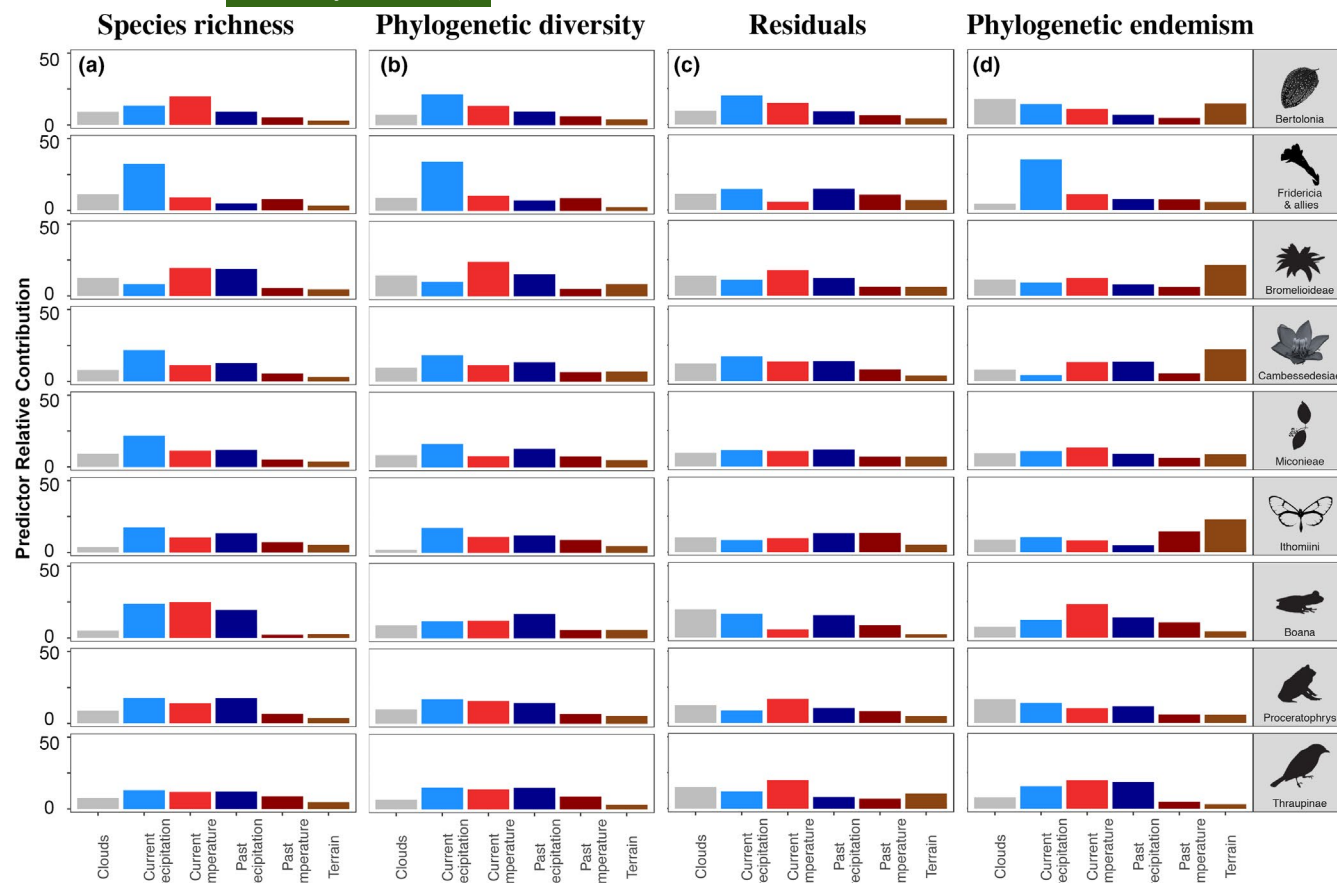


FIGURE 4 Relative importance of predictors of diversity for all studied groups and measures of diversity. Each panel corresponds to one measure of diversity: (a) Species richness (SR), (b) Phylogenetic diversity (PD), (c) Residuals of the PD/SR regression, and (d) Phylogenetic endemism. The 13 predictor variables are grouped in six categories, from left to right: clouds, current precipitation, current temperature, precipitation stability, temperature stability, and topography (more details in Table 2) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/jbi.14083)]

particular descriptors may be more or less effective in predicting regional diversity patterns. In the face of global climatic changes, an extension of this framework may be useful as an indirect biodiversity monitoring tool, particularly if these correlations are applied to periodically retrieved satellite data (e.g., Vermote & Wolfe, 2015). Akin to tools that allow for near-real time descriptions of habitat change (Diniz et al., 2015), they may act as a lens to help us estimate biodiversity trends under ongoing and future climatic shifts. Our analyses suggest that models of species richness and phylogenetic diversity in the Atlantic Forest will profit from the inclusion of climatic data describing current and past precipitation, potentially eliminating the need to include topography as an additional variable. These models may be refined for specific groups or life-history traits, and include other environmental descriptors such as climatic stability (to better account for evolutionary history) or landscape configuration (to incorporate processes acting on range distributions; Santos et al., 2020).

ACKNOWLEDGMENTS

This study used data collected by a collaborative project co-funded by the Fundação de Amparo à Pesquisa do Estado de São

Paulo (FAPESP; 2013/50297-0), the National Science Foundation (NSF; DEB 1343578, DEB-1343612), and NASA. Additional funds were provided to AP through a Fulbright-Colciencias fellowship; AVL by FAPESP (2011/50225-3), the Conselho Nacional de Desenvolvimento Científico e Tecnológico Pq-1A grant (CNPq; 303834/2015-3), RedeLep-SISBIOTA-Brasil/CNPq (563332/2010-7), and NSF (DEB 1256742); LGL through a CNPq Pq-1B grant (310871/2017-4); MK by CNPq (151281/2013-0 and 232621/2014-4); MLL by FAPESP (2017/26162-8); CHG by CNPq (304413/2018-6); CYM by CNPq (303713/2015-1, 306204/2019-3); MTR by FAPESP (2003/10335-8 and 2011/50146-6) and CNPq (301778/2015-9, 301778/2015-9 and 470939/2004-4); and FRA by FAPESP (2011/50143-7, 2011/23155-4, and 2018/03428-5) and a CNPq Pq-2 grant (312697/2018-0). The authors also thank the #Carnavallab at the City College of New York (CCNY), Robert P. Anderson, Brian T. Smith, Youhua Chen, and one anonymous reviewer for helpful suggestions that greatly improved this manuscript, Keith S. Brown Jr. for providing a substantial portion of the Ithomiini dataset used, and Leila Meyer for assistance with the scripting to generate alpha hulls. Butterfly samples are under SISGEN ADF1F75.



DATA AVAILABILITY STATEMENT

All biological data used in this study were made available by the original publications cited in the Methods section (see Table 1 for a comprehensive list). Data from Brown et al., 2020 was used as provided in that publication [(https://doi.org/10.5061/dryad.qz612jmbm jmbm), data for three new groups, bignonies, *Bertolonia* and Cambessedesiae are available in the Dryad repository for this manuscript (occurrences and phylogenies). All environmental data are available from the original sources (see Table 2 for details). All maps in raster format, the shapefile of the Atlantic Forest boundary, and scripts used in all analyses can be found in Dryad [(https://doi.org/10.5061/dryad.qz612jmbm)].

ORCID

Andrea Paz <https://orcid.org/0000-0001-6484-1210>

Jason L. Brown <https://orcid.org/0000-0002-7153-2632>

Lucas F. Bacci <https://orcid.org/0000-0001-5825-9216>

Marcelo Reginato <https://orcid.org/0000-0002-3511-6586>

Ana C. Carnaval <https://orcid.org/0000-0002-4399-1313>

REFERENCES

- Aguirre-Santoro, J. (2017). Taxonomy of the Ronnbergia Alliance (Bromeliaceae: Bromelioideae): New combinations, synopsis, and new circumscriptions of Ronnbergia and the resurrected genus Wittmackia. *Plant Systematics and Evolution*, 303, 615–640.
- Aguirre-Santoro, J., Stevenson, D., & Michelangeli, F. (2016). Molecular phylogenetics of the Ronnbergia Alliance (Bromeliaceae, Bromelioideae) and insights about its morphological evolution. *Molecular Phylogenetics and Evolution*, 100, 1–20.
- Bacci, L. F., Amorim, A. M., Michelangeli, F. A., & Goldenberg, R. (2020). Flower morphology is correlated with distribution and phylogeny in Bertolonia (Melastomataceae), an herbaceous genus endemic to the Atlantic Forest. *Molecular Phylogenetics and Evolution*, 149, 106844.
- Barratt, C. D., Bwong, B. A., Onstein, R. E., Rosauer, D. F., Menegon, M., Doggart, N., Nagel, P., Kissling, W. D., & Loader, S. P. (2017). Environmental correlates of phylogenetic endemism in amphibians and the conservation of refugia in the Coastal Forests of Eastern Africa. *Diversity and Distributions*, 23, 875–887.
- Bochorny, T., Michelangeli, F. A., Almeda, F., & Goldenberg, R. (2019). Phylogenetics, morphology and circumscription of Cambessedesiae: A new Neotropical tribe of Melastomataceae. *Botanical Journal of the Linnean Society*, 190, 281–302.
- Brown, J. L., Paz, A., Reginato, M., Renata, C. A., Assis, C., Lyra, M., Caddah, M. K., Aguirre-Santoro, J., d'Horta, F., Raposo do Amaral, F., Goldenberg, R., Lucas Silva-Brandão, K., Freitas, A. V. L., Rodrigues, M. T., Michelangeli, F. A., Miyaki, C. Y., & Carnaval, A. C. (2020). Seeing the forest through many trees: Multi-taxon patterns of phylogenetic diversity in the Atlantic Forest hotspot. *Diversity and Distributions*, 26(9), 1160–1176. <https://doi.org/10.1111/ddi.13116>
- Burgman, M. A., & Fox, J. C. (2003). Bias in species range estimates from minimum convex polygons: Implications for conservation and options for improved planning. *Animal Conservation*, 6, 19–28.
- Burns, K. J., Shultz, A. J., Tittle, P. O., Mason, N. A., Barker, F. K., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2014). Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of Neotropical songbirds. *Molecular Phylogenetics and Evolution*, 75, 41–77.
- Caddah, M. (2013). Estudos taxonomicos e filogeneticos em Miconia sect. Discolor (Melastomataceae, Miconieae), PhD thesis. UNICAMP, Campinas.
- Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D. F., VanDerWal, J., Damasceno, R., Prates, I., Strangas, M., Spanos, Z., Rivera, D., Pie, M. R., Firkowski, C. R., Bornschein, M. R., Ribeiro, L. F., & Moritz, C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20141461.
- Carstens, B. C., Morales, A. E., Field, K., & Pelletier, T. A. (2018). A global analysis of bats using automated comparative phylogeography uncovers a surprising impact of Pleistocene glaciation. *Journal of Biogeography*, 45, 1795–1805.
- Crisp, M. D., Laffan, S., Linder, H. P., & Monro, A. (2001). Endemism in the Australian flora. *Journal of Biogeography*, 28, 183–198.
- Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W., & Mouquet, N. (2010). Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: The need for integrative conservation strategies in a changing world. *Ecology Letters*, 13, 1030–1040.
- Diniz, C. G., Souza, A. A. D. A., Santos, D. C., Dias, M. C., Luz, N. C., Da, M. D. R. V., De, M. J. S. A., Gomes, A. R., Narvaes, I. D. S., Valeriano, D. M., Maurano, L. E. P., & Adami, M. (2015). DETER-B: The new amazon near real-time deforestation detection system. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3619–3628.
- Erwin, D. H. (2009). Climate as a driver of evolutionary change. *Current Biology*, 19, R575–R583.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1–10.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45, 1–33.
- Forest, F., Grenyer, R., Rouget, M., Davies, T. J., Cowling, R. M., Faith, D. P., Balmford, A., Manning, J. C., Proches, S., van der Bank, M., Reeves, G., Hedderson, T. A. J., & Savolainen, V. (2007). Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature*, 445, 757–760.
- Fretwell, P. T., & Trathan, P. N. (2009). Penguins from space: Faecal stains reveal the location of emperor penguin colonies. *Global Ecology and Biogeography*, 18, 543–552.
- Fritz, S. A., & Rahbek, C. (2012). Global patterns of amphibian phylogenetic diversity. *Journal of Biogeography*, 39, 1373–1382.
- Goldenberg, R., Penneys, D. S., Almeda, F., Judd, W. S., & Michelangeli, F. (2008). Phylogeny of Miconia (Melastomataceae): Initial insights into broad patterns of diversification in a megadiverse neotropical genus. *International Journal of Plant Sciences*, 169, 963–979.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Remote sensing of environment Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27.
- Grimshaw, J. R., & Higgins, C. L. (2017). Environmental correlates of phylogenetic structure in Mexican bat communities. *Journal of Mammalogy*, 98, 1657–1666.
- Grohmann, C. H., Smith, M. J., & Riccomini, C. (2011). Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland. *IEEE Transactions on Geoscience and Remote Sensing*, 49, 1200–1213.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st century forest cover change. *Science*, 342, 850–854.
- Heino, J., Alahuhta, J., & Fattorini, S. (2015). Phylogenetic diversity of regional beetle faunas at high latitudes: patterns, drivers and chance along ecological gradients. *Biodiversity and Conservation*, 24, 2751–2767.

- Hijmans, R. J. (2019). *raster: Geographic data analysis and modeling. R package*.
- IBGE (2004). Retrieved from ftp://geofpt.ibge.gov.br/informacoes_ambientes/vegetacao/vetores/brasil_5000_mil/Vegetacao_5000mil.zip
- Jetz, W., Cavender-bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., Guralnick, R., Kattge, J., Latimer, A. M., Moorcroft, P., Schaepman, M. E., Schildhauer, M. P., Schneider, F. D., Schrodt, F., Stahl, U., & Ustin, S. L. (2016). Monitoring plant functional diversity from space. *Nature Plants*, 2, 1–5.
- Kaehler, M., Michelangeli, F., & Lohmann, L. G. (2019). Fine tuning the circumscription of Fridericia (Bignoniaceae, Bignoniaceae). *Taxon*, 68, 751–770.
- Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab- An S4 package for kernel methods in R. *Journal of Statistical Software*, 11, 1–20.
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Data descriptor: Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 1–20.
- Krishnadas, M., Kumar, A., & Comita, L. S. (2016). Environmental gradients structure tropical tree assemblages at the regional scale. *Journal of Vegetation Science*, 27, 1117–1128.
- Kuhn, M. (2016). caret: Classification and regression training. R package version 6.0-71. Retrieved from <https://CRAN.R-project.org/package=caret>
- Laffan, S. W., Lubarsky, E., & Rosauer, D. F. (2010). Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography*, 33, 643–647.
- Laurencio, D., & Fitzgerald, L. A. (2010). Environmental correlates of herpetofaunal diversity in Costa Rica. *Journal of Tropical Ecology*, 26, 521–531.
- Liaw, A., & Wiener, M. (2002). Classification and regression by random Forest. *R News*, 2, 18–22.
- Mason-Romo, E. D., Farias, A. A., & Ceballos, G. (2017). Two decades of climate driving the dynamics of functional and taxonomic diversity of a tropical small mammal community in western. *PLoS One*, 12, e0189104.
- Meyer, L., Diniz-Filho, J. A. F., & Lohmann, L. G. (2017). A comparison of hull methods for estimating species ranges and richness maps. *Plant Ecology and Diversity*, 10, 389–401.
- Michelangeli, F., Judd, W. S., Penneys, D. S., Slean, J. Jr, Becquer, E., Goldenberg, R., & Martin, C. (2008). Multiple events of dispersal and radiation of the tribe Miconieae (Melastomataceae) in the Caribbean. *Botanical Review*, 74, 53–77.
- Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, 37, 191–203.
- Pabijan, M., Wollenberg, K. C., & Vences, M. (2012). Small body size increases the regional differentiation of populations of tropical mantellid frogs (Anura: Mantellidae). *Journal of Evolutionary Biology*, 25, 2310–2324.
- Paz, A., Ibañez, R., Lips, K. R., & Crawford, A. J. (2015). Testing the role of ecology and life history in structuring genetic variation across a landscape: A trait-based phylogeographic approach. *Molecular Ecology*, 24, 3723–3737.
- Paz, A., Reginato, M., Michelangeli, F. A., Goldenberg, R., Caddah, M., Aguirre-Santoro, J., Kaehler, M., Lohmann, L., & Carnaval, A. C. (2020). Predicting patterns of plant diversity and endemism in the tropics using remote sensing data: A study case from the Brazilian Atlantic rainforest. In J. Cavender-Bares, J. Gamon, & P. Townsend (Eds.), *Remote sensing of plant biodiversity* (pp. 255–266). Springer.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., ... Wegmann, M. (2013). Essential Biodiversity Variables. *Science*, 339, 277–278.
- Peters, M. K., Hemp, A., Appelhans, T., Behler, C., Classen, A., Detsch, F., Ensslin, A., Ferger, S. W., Frederiksen, S. B., Gebert, F., Haas, M., Helbig-Bonitz, M., Hemp, C., Kindeketa, W. J., Mwangomo, E., Ngeresa, C., Otte, I., Röder, J., Rutten, G., ... Rosseel, Y. (2016). Predictors of elevational biodiversity gradients change from single taxa to the multi-taxa community level. *Nature Communications*, 7, 13736.
- Peterson, A. T. Problems with reductive polygon-based methods for estimating species' ranges: reply to Pimm et al 2017. *Conservation Biology*, 31, 948–951.
- Peterson, A. T., Navarro-Sigüenza, A. G., & Gordillo, A. (2018). Assumption-versus data-based approaches to summarizing species' ranges. *Conservation Biology*, 32, 568–575.
- Rabosky, A. D., Cox, C. L., Rabosky, D. L., Title, P. O., Holmes, I. A., Feldman, A., & McGuire, J. A. (2016). Coral snakes predict the evolution of mimicry across New World snakes. *Nature Communications*, 7, 1–9.
- Reginato, M., & Michelangeli, F. (2016). Untangling the phylogeny of *Leandra* sensu str. (Melastomataceae, Miconieae). *Molecular Phylogenetics and Evolution*, 96, 17–32.
- Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. (2009). The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, 142, 1141–1153.
- Rohde, K. (1992). Latitudinal gradients in species diversity: The search for the primary cause. *Oikos*, 65, 514.
- Rompré, G., Robinson, W. D., Desrochers, A., & Angehr, G. (2007). Environmental correlates of avian diversity in lowland Panama rain forests. *Journal of Biogeography*, 34, 802–815.
- Rosauer, D., Laffan, S. W., Crisp, M. D., Donnellan, S. C., & Cook, L. G. (2009). Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. *Molecular Ecology*, 18, 4061–4072.
- Safi, K., Cianciaruso, M. V., Loyola, R. D., Brito, D., Armour-Marshall, K., & Diniz-Filho, J. A. F. (2011). Understanding global patterns of mammalian functional and phylogenetic diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 2536–2544.
- Santos, J. P., Sobral-Souza, T., Brown, K. S., Vancine, M. H., Ribeiro, M. C., & Freitas, A. V. L. (2020). Effects of landscape modification on species richness patterns of fruit-feeding butterflies in Brazilian Atlantic Forest. *Diversity and Distributions*, 26, 196–208.
- Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jürgens, N., Larigauderie, A., Muchoney, D., Walther, B. A., & Mooney, H. A. (2008). Toward a global biodiversity observing system. *Science*, 321, 1044–1045.
- Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H. R., Tuama, É. Ó., Faith, D. P., Mooney, H. A., Ferrier, S., Jongman, R. H. G., Harrison, I. J., Yahara, T., Pereira, H. M., Larigauderie, A., & Geller, G. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, 4, 1–8.
- Silva, F. R., Almeida-Neto, M., do Prado, V. H. M., Haddad, C. F. B., & de Cerqueira Rossa-Feres, D. (2012). Humidity levels drive reproductive modes and phylogenetic diversity of amphibians in the Brazilian Atlantic Forest. *Journal of Biogeography*, 39, 1720–1732.
- Singarayer, J. S., & Valdes, P. J. (2010). High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr. *Quaternary Science Reviews*, 29, 43–55.
- Tallowin, O., Allison, A., Algar, A. C., Kraus, F., & Meiri, S. (2017). Papua New Guinea terrestrial-vertebrate richness: Elevation matters most for all except reptiles. *Journal of Biogeography*, 44, 1734–1744.
- Title, P. O., & Bemmels, J. B. (2018). ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, 41, 291–307.
- Tucker, C. M., & Cadotte, M. W. (2013). Unifying measures of biodiversity: Understanding when richness and phylogenetic diversity should be congruent. *Diversity and Distributions*, 19, 845–854.



- Turner, W. (2014). Sensing biodiversity. *Science*, 346, 301–302.
- Vane-Wright, R. I., Humphries, C. J., & Williams, P. H. (1991). What to protect?—Systematics and the agony of choice. *Biological Conservation*, 55, 235–254.
- Vasconcelos, T. S., Prado, V. H. M., Da Silva, F. R., & Haddad, C. F. B. (2014). Biogeographic distribution patterns and their correlates in the diverse frog fauna of the Atlantic forest hotspot. *PLoS One*, 9, 1–9.
- Vasconcelos, T. D. S., Santos, T. G., Dos, H. C. F. B., & Rossa-Feres, D. D. C. (2010). Climatic variables and altitude as predictors of anuran species richness and number of reproductive modes in Brazil. *Journal of Tropical Ecology*, 26, 423–432.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer.
- Vermote, E., & Wolfe, R. (2015). MYD09GA MODIS/Aqua surface reflectance daily L2G global 1km and 500m SIN grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Retrieved from <https://doi.org/10.5067/MODIS/MYD09GA.006>
- Voskamp, A., Baker, D. J., Stephens, P. A., Valdes, P. J., & Willis, S. G. (2017). Global patterns in the divergence between phylogenetic diversity and species richness in terrestrial birds. *Journal of Biogeography*, 44, 709–721.
- Williams, K. J., Belbin, L., Austin, M. P., Stein, J. L., & Ferrier, S. (2012). Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*, 26, 2009–2047.
- Xu, T., & Hutchinson, M. (2010). *ANUCLIM user guide, version 6.1*. Australian National University.
- Zamudio, K. R., Bell, R. C., & Mason, N. A. (2016). Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. *Proceedings of the National Academy of Sciences*, 113, 8041–8048.
- Zellweger, F., Baltensweiler, A., Ginzler, C., Roth, T., Braunisch, V., Bugmann, H., & Bollmann, K. (2016). Environmental predictors of species richness in forest landscapes: Abiotic factors versus vegetation structure. *Journal of Biogeography*, 43, 1080–1090.

BIOSKETCH

Andrea Paz is a PhD candidate at the Carnavallab at the City College of New York. She is interested in the geographical patterns of biodiversity and the processes generating and maintaining those patterns. In particular, she seeks to understand how species change their distributions in response to environmental changes both in the past and present.

Author contributions: A.P., J.L.B., A.C.C., and T.S.S. conceived the ideas. A.P. and J.L.B. analysed the data with help from T.S.S. all authors contributed data and participated in discussions that led to data analyses and interpretations of results. A.P. and A.C.C. led the writing, and all authors read and approved the final version of the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Paz A, Brown JL, Cordeiro CL, et al. Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest. *J Biogeogr.* 2021;48:1377–1391. <https://doi.org/10.1111/jbi.14083>