Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest

Running title (40 characters): Correlates of diversity in the Atlantic Forest

Andrea Paz[1,2], Jason L. Brown[1], Carlos L.O. Cordeiro[3], Julian Aguirre-Santoro[2,4], Claydson Assis[5], Renata Cecilia Amaro[5], Fabio Raposo do Amaral[6], Thuane Bochorny[7], Lucas F. Bacci[7], Mayara K. Caddah[8], Fernando d'Horta[5], Miriam Kaehler[5], Mariana Lyra[9], Carlos Henrique Grohmann[10], Marcelo Reginato[11], Karina Lucas Silva-Brandão[12], André Victor Lucci Freitas[13], Renato Goldenberg[14], Lúcia G. Lohmann[5], Fabián A. Michelangeli[2,15], Cristina Miyaki[16], Miguel T. Rodrigues[5], Thiago S. Silva[3,17], Ana C. Carnaval[1,2]

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

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

[3] Instituto de Geociências e Ciências Exatas, UNESP-Univ Estadual Paulista, Departamento de Geografia, Ecosystem Dynamics Observatory. Rio Claro, SP, Brazil

[4] Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá, Colombia

[5] Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil

[6] Departamento de Ecologia e Biologia Evolutiva, Universidade Federal de São Paulo, São Paulo, SP, Brazil

[7] Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP, Brazil

[8] Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

[9] Departamento de Zoologia and Centro de Aquicultura (CAUNESP), Instituto de Biociências, Universidade Estadual Paulista, Rio Claro, SP, Brazil

[10] Institute of Energy and Environment, University of São Paulo, São Paulo, SP 05586-060, Brazil

[11] Departamento de Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[12] Universidade Federal do ABC, UFABC, SP, Brazil

[13] Departamento de Biologia Animal and Museu de Zoologia, Instituto de Biologia, Unicamp, Campinas, São Paulo, Brazil

[14] Universidade Federal do Paraná, Curitiba, PR, Brazil

This is the peer reviewed version of the following article: Paz, A., Brown, J.L., Cordeiro, C.L.O., Aguirre-Santoro, J., Assis, C., Amaro, R.C., Raposo do Amaral, F., Bochorny, T., Bacci, L.F., Caddah, M.K., d'Horta, F., Kaehler, M., Lyra, M., Grohmann, C.H., Reginato, M., Silva-Brandão, K.L., Freitas, A.V.L., Goldenberg, R., Lohmann, L.G., Michelangeli, F.A., Miyaki, C., Rodrigues, M.T., Silva, T.S. and Carnaval, A.C. (2021), Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest. *Journal of Biogeography*, 48: 1377-1391, which has been published in final form at https://doi.org/10.1111/jbi.14083.

[15] Institute of Systematic Botany, The New York Botanical Garden, Bronx, New York, NY, USA
[16]Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil
[17]Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK

Acknowledgments

This work 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; AVLF 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); and, LGL though a CNPq Pq-1B grant (310871/2017-4), MK through CNPq grants (151281 / 2013-0 and 232621 / 2014-4), CHG through CNPq (304413/2018-6), CYM through CNPq (303713/2015-1, 306204/2019-3) and, FRA by FAPESP and the BIOTA program (2011/50143-7, 2011/23155-4, 2018/03428-5) and CNPq Pq-2 grant (312697/2018-0). We also thank the #CarnavalLab at the City College of New York (CCNY), 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.

Accepted for publication in Journal of Biogeography published by Wiley:
https://doi.org/10.1111/jbi.14083

Abstract

Aim

We compare patterns of diversity and their environmental correlates across nine clades of ecologically distinct groups of animals and plants co-existing in a single rainforest domain. We ask whether there are common correlates of diversity patterns, despite ecological differences across clades, to enable a unified platform to predict changes in the distribution of biodiversity in these groups. We focus on predictions of species richness, phylogenetic diversity, and phylogenetic endemism.

Location

Brazilian Atlantic Forest

Methods

Using carefully curated occurrence localities and phylogenetic data, we generated maps of (i) species richness, (ii) phylogenetic diversity, (iii) residuals of phylogenetic diversity regressed on species richness, and (iv) phylogenetic endemism for nine groups of plants and animals in the Atlantic forest. We also compiled a set of 30 potential environmental descriptors including records of current temperature and precipitation, climatic stability over time, and topography. Through a machine learning framework, we then explored the environmental correlates of each of these diversity measures for each group.

Results

The environmental variables used in this study were strong predictors of diversity for all studied groups. However, models for phylogenetic endemism had a lower predictive power. Although patterns of diversity are different among groups, correlates of diversity are very much consistent across taxa. For both species richness and phylogenetic diversity, current precipitation and precipitation stability over time were constantly ranked among the variables that most strongly correlate with diversity patterns. Differently from species richness and phylogenetic diversity, the correlates of phylogenetic endemism were less homogenous across groups. The results also suggest that the inclusion of climate stability over time, along with current climatic descriptors, is important when predicting diversity measures that reflect historical components, such as phylogenetic diversity and endemism.

Main conclusions

Investigating environmental correlates of diversity for multiple co-existing clades and diversity measures in a single geographic area allows for a better understanding of common patterns across taxa. In this study, we identified common environmental correlates of the patterns of species richness and phylogenetic diversity, but not of phylogenetic endemism, across different Atlantic Forest groups. This information can now be used to improve predictions of biodiversity changes at broad taxonomic and geographical scales in the Atlantic Forest.

Keywords

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

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 (Scholes et al., 2012; Pereira et al., 2013). With the increasing amount of satellite data being now freely available to the public, direct near-real time monitoring of some components of diversity has become a reality (Turner, 2014). Examples include global estimates of forest cover change (Hansen et al., 2013), plant functional diversity (Jetz et al., 2016), and penguin population locations and sizes in Antarctica (Fretwell & Trathan, 2009). However, many of the world's diverse groups of organisms cannot be directly observed through satellites, including most animals and non-canopy plants (Turner, 2014). Indirect estimates of diversity would be valuable for more inaccessible organisms. For instance, environmental variables that can be obtained from remote sensing sources (e.g., temperature, precipitation) and correlate with diversity patterns, can be used as proxies to predict the patterns of diversity themselves (Paz et al., 2020). Despite that, the utility of remote sensing tools for indirect biodiversity monitoring remains underexplored.

One potential caveat of indirect sensing of biodiversity is that little consensus exists regarding the environmental variables that represent good predictors of the many different dimensions of biodiversity, especially in megadiverse and threatened tropical ecosystems. Furthermore, the selection of input variables is important and often different between studies (Williams et al., 2012). While species richness (SR) and endemism have been widely used as biodiversity metrics, and broadly employed to describe its spatial patterns, phylogenetic diversity (PD) and phylogenetic endemism (PE) are increasingly used to explicitly quantify the amount of evolutionary uniqueness of a region (Vane-Wright et al., 1991; Faith, 1992; Rosauer et al., 2009). Both PD and PE reflect how combinations of more distantly related species will encompass higher percentages of the overall evolutionary history than combinations of closely related species (Forest et al., 2007; Devictor et al., 2010). Yet, measures of phylogenetic diversity are often positively correlated with species richness. In other words, areas where these two measures of diversity (taxonomic and phylogenetic) are decoupled, have been shown to include more or less evolutionary history than expected given their species richness (Forest et al., 2007; Devictor et al., 2010; Safi et al., 2011; Fritz & Rahbek, 2012; Tucker & Cadotte, 2013). Importantly, however, both species richness and evolutionarily informed measures of diversity appear to be highly correlated to environmental variation - particularly temperature and precipitation. Still, the specific contribution of the individual climatic or landscape descriptors appears idiosyncratic when different taxa or different measures of diversity are compared across regions (Rompré et al., 2007; Laurencio & Fitzgerald, 2010; Peters et al., 2016; 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 groups of plants and animals along elevational gradients on Mount Kilimanjaro (Peters et al., 2016), a similar transect study in Switzerland identified precipitation, temperature, and topography as better predictors of the 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 several groups when exploring environmental correlates of diversity.

Studies of evolutionarily-informed measures of diversity also result in contrasting inferences. Humidity and precipitation are 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 might 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). On the other hand, phylogenetic endemism of African frogs appears to be related to Quaternary climatic stability (Barratt et al., 2017). Consensus is even harder to achieve when including different dimensions of diversity such as richness and phylogenetic diversity in the equation.

To provide a controlled comparison in the same area and inform the ability to indirectly monitor tropical biodiversity as a function of climatic conditions in a biological hotspot, 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, harboring one of the highest levels of diversity and endemism in the world (Ribeiro et al., 2009). It spans a region of complex topography and environments, making it an excellent location to study the potential drivers of diversity in dissimilar groups. 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, which could be useful in community-level prediction and indirect biodiversity monitoring.

We combine geo-referenced locality data with phylogenetic information for nine target clades: five groups of plants and four groups of animals. We used 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 investigated how well these patterns are predicted by each one of 30 abiotic correlates obtained from weather-station data and remote sensing sources (Vermote et al., 2015; Karger et al., 2017; Title & Bemmels, 2018), which describe spatial shifts in temperature, precipitation, humidity, and topography, as well as climatic stability over the last 120,000 years.

Methods

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 nine biological groups (five plant clades and four animal clades). Two of the plant datasets and all animal data were downloaded from Brown et al. (2020), including vetted occurrence points and maximum likelihood phylogenies based on mitochondrial DNA for animals and chloroplast DNA for plants. This dataset included a clade with 18 species (3,774 occurrence points; ca. 67% of the AF species included in this lineage) of tank-forming plants belonging to the Bromelioideae subfamily (Aguirre-Santoro et al., 2016; Aguirre-Santoro, 2017), a clade with 177 species (25,645 sampling points; ca. 70% of the AF species included in this lineage) of shrubs and small trees from the Miconieae tribe in the Melastomataceae (Goldenberg et al., 2008; Michelangeli et al., 2008; Caddah, 2013; Reginato & Michelangeli, 2016), a clade with 55 species (3,269 occurrence points, ca. 100% of the AF species included in this lineage) of clearwing butterflies of tribe Ithomiini in the subfamily Danainae (Nymphalidae), a clade of 19 species (227 occurrence points; ca. 76% of the AF species included in this lineage) of treefrogs from the genus *Boana* from the family Hylidae, a clade of 19 species (2,065 occurrence points; ~ 67% of the AF species included in this lineage) of the horned frogs of the *Proceratophrys* genus in the Odontophrynidae family (Brown et al., in press; Vasconcelos et al., 2014), and a clade of 22 species of birds (8,501 occurrence points; ca. 100% of the AF species included in this lineage) from tanagers belonging to the subfamily Thraupinae (Burns et al., 2014). In addition, we gathered three plant datasets (including species presence data and maximum likelihood phylogenies produced with alternative markers): the *Fridericia* and allies group of the tribe Bignonieae, in the plant family Bignoniaceae (hereon referred to as "bignones," Kaehler et al., 2019), and two clades of melastomes, the *Bertolonia* genus (Bacci et al., 2020), and the Cambessedesieae tribe (Bochorny et al., 2019). The bignones dataset includes 65 species (ca. 72% of the AF species included in this lineage), 5,115 presence points, and a phylogeny built from one plastid marker (*ndh*F) and one nuclear marker (*Pep*C). The *Bertolonia* dataset contains 31 species (ca. 88% of the AF species included in this lineage), 744 points, and a phylogeny built from nine nuclear, ribosomal and plastid markers (*nr*ITS and *nr*ET, *atp*F-*atp*H, *ndh*F, *psb*K-*psb*L, *rbc*L, *rpl*16, and *trn*S-*trn*G, *ADH*, and *PCRF1*). The Cambessedesieae dataset contains 54 species (ca. 81% of the AF species included in this lineage), 1,167 presence points, and a phylogeny built from six ribosomal, plastid, and nuclear markers (nrITS, nrETS, *atp*F-*atp*H, *psb*K-*psb*L, *trn*S-*trn*G, and *waxy*).

Mapping species richness, phylogenetic diversity, and phylogenetic endemism

To assess which environmental variables (reflecting both past and present conditions) and landscape descriptors best explain diversity patterns in the target clades, we first superimposed distribution maps for 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) using all occurrence data available for each individual species with more than three locality points (we added all others as individual 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). Once done, those distribution maps were rasterized to match the spatial resolution of the predictor variables (~10 km). We acknowledge that this method might overestimate individual species distributions (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 then built for all nine groups, following the methods described in Brown et al. (in press). In brief, input data consisted of community composition matrices based on the superimposed maps of species ranges (alpha hulls and points when <= 3 points). For each group, the species maps were stacked and converted to a composition matrix in R. For phylogenetic diversity and phylogenetic endemism, we compiled phylogenetic trees, including branch lengths. For each of the nine groups, we imported the community composition matrix and the phylogeny to Biodiverse (Laffan et al., 2010). 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 pixel in the AF (Figure 1). Species richness was estimated by summing all species present in each pixel. Phylogenetic diversity was computed by summing branch lengths leading to all species present in a given pixel. Phylogenetic endemism, which combines endemism (estimated from the range of the species and the fraction contained in a given cell), and phylogenetic diversity, to estimate its level of restriction (Rosauer et al., 2009).

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. In the residual maps, 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.

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 1). Twenty-one of those were directly obtained from public databases. The variables include a 90m 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, downscaled to a 30” (~ 1km) 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 Atlantic Forest specific layers, two describing terrain, one for cloud cover, and six reflecting climatic stability over the last 120,000 years. The 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). Through an interactive discussion with experts having wide experience in the field, known Atlantic Forest areas were selected and combined into a multi-polygon vector geometry. The map is referenced to the WGS84 datum with geographic coordinates and is available in the Dryad repository associated to this manuscript. 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 (MODIS 09GA, Vermote et al., 2015) based on satellite data collected from 2000 to 2017, using Google Earth Engine (Gorelick et al., 2017). To determine whether historical climates were important in predicting present-day patterns of diversity, we built six layers to reflect climatic stability over the last 120,000 years. 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, Singarayer & Valdes, 2010; Carnaval et al., 2014). 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. For downstream analyses, all variables were resampled to a 5’ resolution (~10km) using the resample function of the R package raster 3.0-7 (Hijmans, 2019).

The complete dataset of environmental descriptors (30 layers; Table 1) 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 including 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 bio 1, CV bio 10, CV bio 16 and CV bio 17), topography (Altitude, Rugosity and TWI), and cloud cover (Table 1).

Correlates of biodiversity

To determine which of the environmental descriptors are the best predictors of SR, PD, PE, and the residuals of the PD, 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 (each map of a given diversity metric 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 (folds). We used each fold in turns as an internal validation dataset, utilizing the others for training. We repeated this procedure three times (repeat crossvalidation). All models were built with the R package caret 6.0-84 (Kuhn, 2016). A final ensemble model, built from a linear combination of the four algorithms based on the RMSE values, was built with the caretEnsemble function in R. Finally, we used the withheld 30% of pixels for model testing. 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.

Results

Patterns of SR, PD, and PE are different among groups (Figure 2). 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 (Figures 1, 2). Both the Miconieae and tanagers have higher diversity (both SR and PD) in the Serra do Mar coastal forest (coastal mountains, Figures 1, 2). On the other hand, 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 north-east (Bahia coastal forest) of the high diversity areas for the butterflies. The bromeliads have their peak diversity for both PD and SR in the north-east of the Bahia coastal forest, with the *Proceratophrys* frogs showing peak diversity in the north of Bahia and in 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; with the highest PD mainly in the Serra do Mar coastal forest.

In the case of PE, we detected two different general patterns with either small concentrations of high PE or more widespread areas of high PE values (Figure 2). Two groups, the butterflies and the tanager birds, show a pattern of widespread high PE, inland to the north of Rio de Janeiro for the butterflies and in the Serra do Mar coastal forest for the tanager birds. Both the bromeliads and the *Bertolonia* have very small areas of high PE in the north of the Bahia state. The Miconieae, the *Proceratophrys* frogs, and the Cambessedesieae, all show high PE in small areas of Espírito Santo and the border between Rio de Janeiro and São Paulo states. The Cambessedesieae also shows 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 residuals (positive or negative, Figure 3). Most groups have areas where PD is higher than expected given species richness, with *Proceratophrys* frogs showing this pattern along the entire mapped distribution (Figure 3). Butterflies have higher PD than expected given the number of species in the coastal region that extends from Alagoas to Paraiba, and in the southern interior region of the forest (red and dark orange areas in

Figure 3). We found a spatial concentration of higher PD than expected in the southern Atlantic Forest for another three groups: (i) the tanagers in the Paraná, Santa Catarina, and part of São Paulo states, (ii) both the *Boana* frogs, and (iii) bignones from the Serra do Mar Coastal Forest and the Santa Catarina and Parana states. The bignones, Miconieae, and Cambessedesieae also showed higher PD than expected in the north, around the Bahia Interior Forests region. Negative residuals, differently from positive residuals, are more spread out in geographical space (Figure 3, green and blue respectively). The few exceptions, with a concentration of negative residuals, 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 in the southern portion of the forest. In the Miconieae, they are mostly found in the Serra do Mar Coastal Forests (Figure 3, blue areas).

Of all biodiversity metrics, SR PD (R^2 0.86-0.98) and the residuals of PD (R^2 0.86-0.98, Figure 4) are best predicted by the environmental models. This metric is followed by SR (R^2 0.79-0.98. Model predictions of PE were more heterogeneous and generally lower (R^2 0-0.96, Figure 4), with three main exceptions: the bignones, the tanager birds, and the butterflies (R^2 of 0.94,0.95 and 0.96 respectively). Two other groups, the *Proceratophrys* frogs and the Miconieae (R^2 of 0.36 and 0.49 respectively), showed some predictive power.

The ability of the models to predict SR, PD, and PE also varied across clades. For instance, they 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 *Bertolonia* genus (lowest R^2 for all but PE and second to last for PE; Figure 4), and *Proceratophrys* frogs (which had one of the lowest R^2 values for predictions for SR, PD diversity and residuals; Figure 4).

Climatic variables, reflecting both present-day and past conditions, contributed highly to predictions of SR, PD, and phylogenetic residuals. In particular, precipitation-related variables were consistently identified as those of higher importance to predict SR and PD (blue in Figure 5). In eight out of the nine clades, variables reflecting current precipitation are those of highest importance for predicting SR; in one clade (bromeliads), stability in past precipitation was ranked first, but closely followed by current precipitation (Figure 5). Current temperatures were the second or third predictors of SR. Conversely, historical stability in temperature contributed less to predictions of diversity, ranking fourth to last for SR (except from bignones, which ranked second) (Figure 5a). 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. For the bromeliads, current temperature was the most important variable, followed by both 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 5b).

Variables related to precipitation also had higher importance to explain the residuals of PD in eight out of the nine target groups (Figure 5a, b), with a slightly higher contribution of climate stability as a correlate of the residuals relative to the other metrics (Figure 5a, b, d). 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 5d). Although generally low, cloud distribution was relevant to predicting the residuals for tanagers and *Boana* frogs, and slightly relevant for the bignones and the Cambessedesiae (Figure 5).

Unlike the other biodiversity metrics, PE does not seem to be better predicted by one specific type of environmental variable. Variables with highest contributions to the model are related to current precipitation (five out of the nine groups), current temperature (one out of the nine goups), and terrain (three out of nine). Of the groups with some predictive power for PE, four out of five have 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 is the highest contributor for patterns observed in Cambessedesiae, bromeliads, and butterflies, but ranks second and third for *Bertolonia* and Miconieae, respectively (Figure 5c).

Discussion

Models based on environmental variables describing temperature and precipitation represented good predictors of different dimensions of diversity in the Atlantic Forest, based on the nine focal clades (Figure 4). Nevertheless, predictions of PE were poorer than those of SR or PD, a result likely associated to the spatial restriction of this biodiversity measure in relation to the broader environmental predictors used, other studies have indeed suggested endemism might be explained by variation within broader study regions (Rosauer et al., 2009; Crisp et al., 2011). However, predictions were still good in half of the cases, with R^2 values ranging from 0.36 to 0.96. In other words, to a certain extent, we can still predict PE based on our set of predictor variables for some groups, although not as reliably. Our sampling for the AF was fairly complete, however, some narrow endemics are missing in the datasets and thus PE can be underestimated in certain cells. We acknowledge our threshold of >3 points for creating alpha shapes might be considered small. However, we created the distribution maps with alternative thresholds of 10 and 30 and have changes in only few groups of restricted distributions (Appendix A). We also ran analyses with an alternative method for map building, through minimum convex polygons with very similar predictive power and environmental predictors selected (Appendix A). Sampling multiple clades within a single geographic space allowed us to identify variables that are consistently important (or not) predictors of diversity in the Atlantic Forest, highlighting the importance of precipitation (both past and present), but limited contribution of topography.

Traditionally, temperature has been considered as the most important driver of diversity patterns, given the importance of this variable in the temperate zones (Rohde, 1992; Erwin, 2009; Peters et al., 2016). However, our results point to precipitation as a main predictor of species richness and phylogenetic diversity, while contributing to the mismatch between those variables. This result is in line with other recent tropical clade-based studies 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). Here we show that studying several clades in the same area does improve our ability to find general patterns in the potential environmental drivers of diversity, highlighting the importance of precipitation.

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 (i.e., PE and 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 may also be driven by the latter. The residuals, however, show areas of mismatch between the two, highlighting areas where evolutionary history is providing different information (Forest et al., 2007; Devictor et al., 2010), and might therefore give us better insights into predictors of evolutionary history. This relationship has been studied at a global scale for birds, where abiotic correlates of these residuals are spatially heterogeneous, suggesting that elevation is an important predictor of PD residuals in the tropics, along with contact among biomes (Voskamp et al., 2017). Here, we found relatively high importance of climatic stability as a predictor of PD residuals in the AF, a tropical realm that may point to further heterogeneity in the relative importance of environmental variables within realms.

We found some discrepancies among study groups in the strength of predictions and the importance of variables. 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 temperature stability (and sometimes precipitation stability) over the Quaternary 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 (Pabijan et al., 2012; Paz et al., 2015; Zamudio et al., 2016; Carstens et al., 2018).

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 the importance of precipitation in determining diversity patterns. In the face of global climatic changes, a similar framework may be useful as a biodiversity monitoring tool, particularly if tied to periodically retrieved remote sensing data (e.g., Vermote et al., 2015). Akin to near-real time snapshots of habitat change (Diniz et al., 2015), they can provide fairly accurate predictions of expected changes in diversity patterns driven by climatic shifts. Our analyses suggest that these near-time models, at least in the Atlantic Forest, and potentially in the tropics, will profit from the inclusion of climatic data describing current and past precipitation, replacing the need to include topography as an independent variable. Models may be refined for specific groups (or life-history characteristics) and diversity measures including, for example, stability measures to predict evolutionary history, or data on landscape configuration as an additional proxy to explain diversity distribution (Santos et al., 2020).

Table 1. Thirty environmental variables compiled for this study. Grey shading indicates the thirteen variables used for Machine Learning analyses, after eliminating variable with high collinearity given the original 30-variable dataset. All variables were resampled at a ~10km resolution. The table depicts the source and the category of each variable. Acronyms: SRTM: Shuttle Radar Topography Mission (Farr et al., 2007); TWI: Topographic wetness index; Envirem: Environmental rasters for ecological modeling (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 et al., 2015); Hadley Center: Hadley Center model (HadCM3, Singarayer & Valdes, 2010; Carnaval et al., 2014).

Layer	**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	
CV Bio 16	CV of Bio 16 (120,000 years, every 4,000 years)	Based on Hadley Center data	Precipitation Stability
CV Bio 17	CV of Bio 17 (120,000 years, every 4,000 years)	Based on Hadley Center data	

FIGURES

Figure 1. The study area, the Brazilian Atlantic Forest, including the Brazilian states encompassed by the Atlantic Forest domain. Lighter shades of gray indicate lower elevations, darker shades indicate higher elevations.

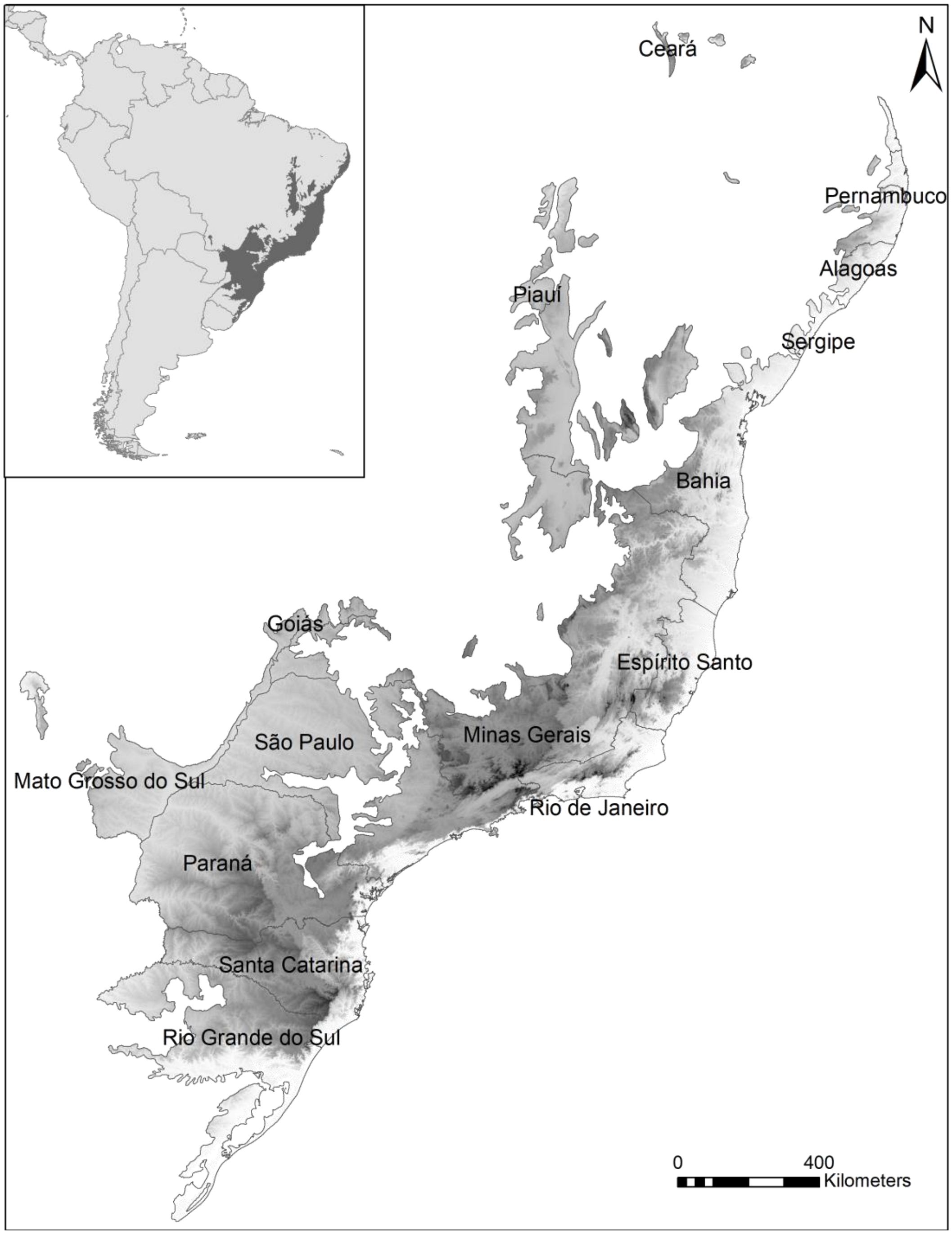


Figure 2. Maps of species richness (first column), phylogenetic diversity (middle column), and phylogenetic endemism (last column) for the nine study groups (rows), based on individual species minimum convex polygons around presence points, and molecular phylogenies. Warmer colors represent higher diversity, colder colors 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 (bignones), 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 (bignones; 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 (bignones; 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 (bignones; $2.31*10^{-6}$-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).

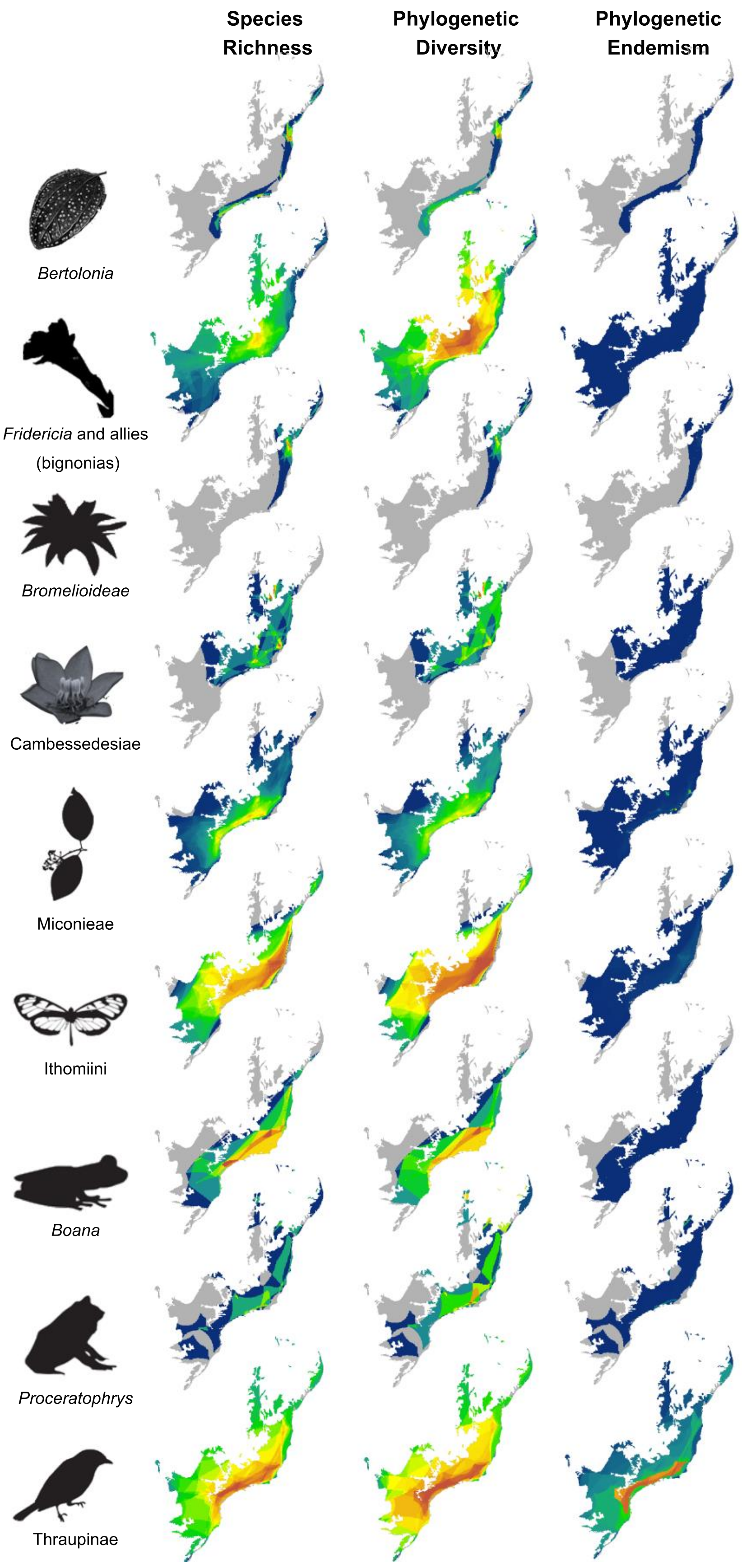

Species Richness
Phylogenetic Diversity
Phylogenetic Endemism
Bertolonia
Fridericia and allies (bignonias)
Bromelioideae
Cambessedesiae
Miconieae
Ithomiini
Boana
Proceratophrys
Thraupinae

Figure 3. 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 (bignones), Bromelioideae, Cambessedesiae, and Miconieae), and four groups of animals (the Ithomiini butterflies, the *Boana* and *Proceratophrys* frogs, and the tanagers of the Thraupinae subfamily).

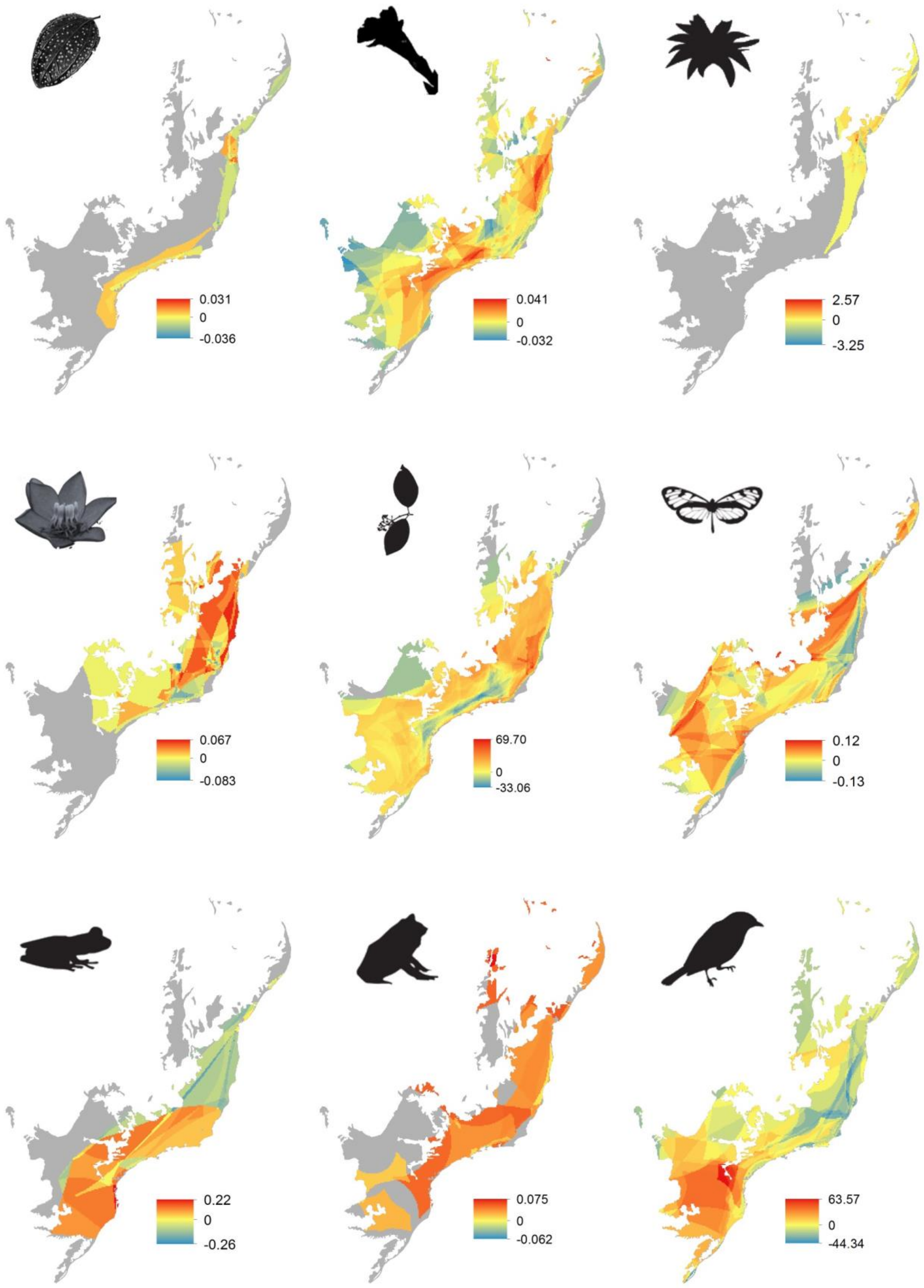

0.031
0
-0.036
0.041
0
-0.032
2.57
0
-3.25
0.067
0
-0.083
69.70
0
-33.06
0.12
0
-0.13
0.22
0
-0.26
0.075
0
-0.062
63.57
0
-44.34

Figure 4. 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*, *Proceratophrys* frogs, and the tanagers of the Thraupinae subfamily.

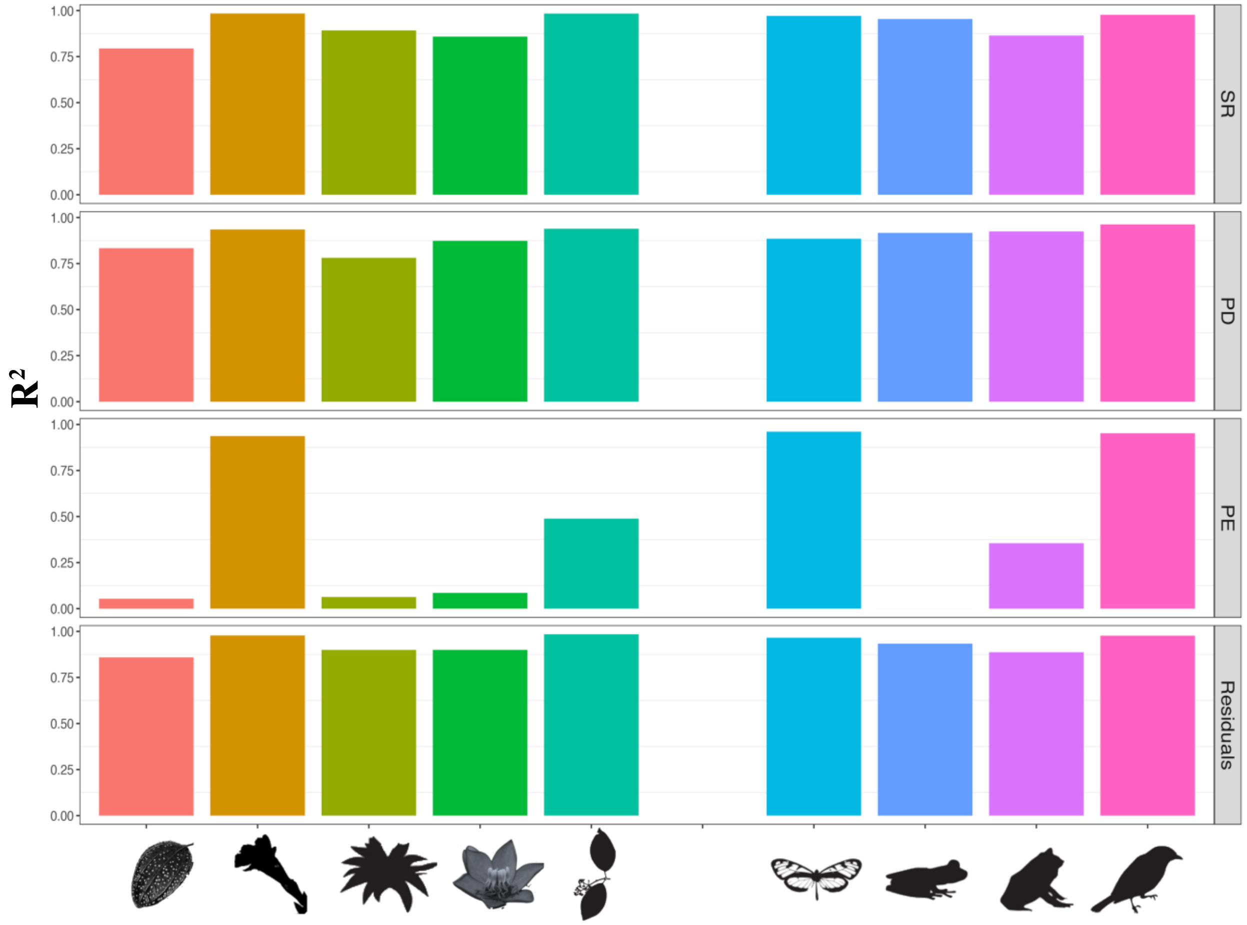
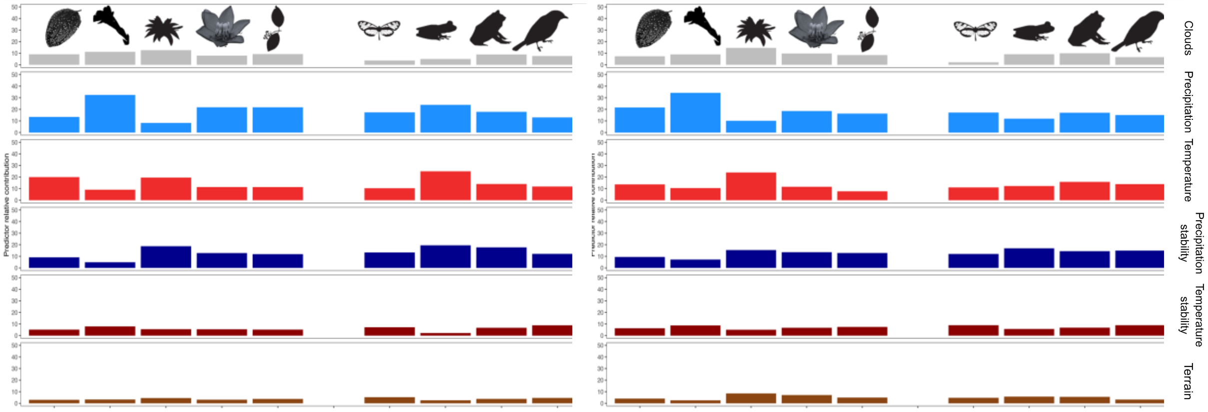
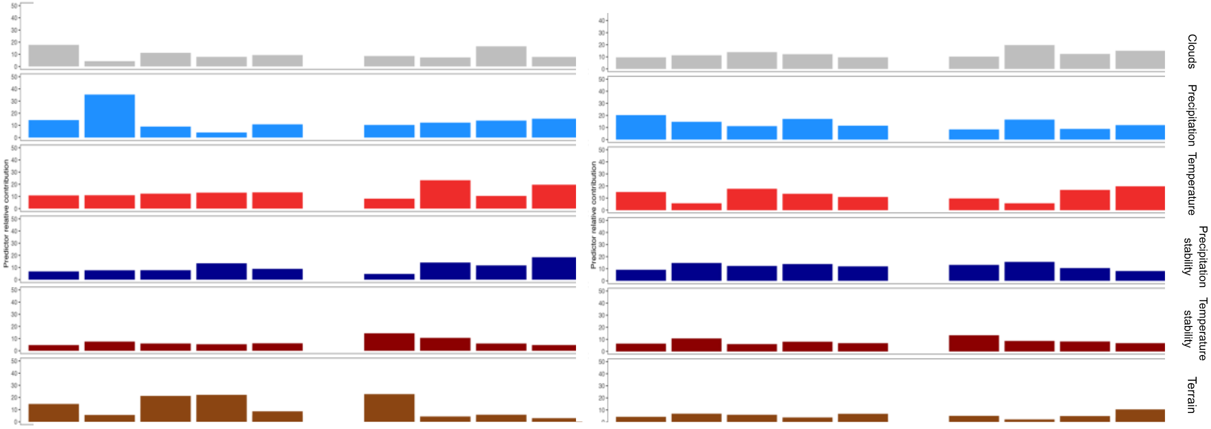


Figure 5. Relative importance of predictors of diversity for all studied groups and measures of diversity. Each panel corresponds to one measure of diversity a) Species richness (SR), b) Phylogenetic diversity (PD), c) Phylogenetic endemism (PE), and d) Residuals of the PD/SR regression. The 13 predictor variables are grouped in six categories, from top to bottom: clouds, current temperature, current precipitation, precipitation stability, temperature stability, and topography (more details in Table 1).



a Species richness

b Phylogenetic diversity



c Phylogenetic endemism

d Residuals

Data availability statement

All used raw data are available in the individual referenced publications. A table pointing to each individual data source, all maps in raster format, the shapefile of the Atlantic Forest boundaries, as well as scripts for analyses, can be found in Dryad [link to be added after acceptance].

Author contributions

AP, JLB, ACC and TSS conceived the ideas; AP and JLB analyzed the data with help from TSS; all authors contributed data and participated in discussions that led to data analyses and interpretations of results; AP and ACC led the writing, and all authors read and approved the final version of the manuscript.

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 Cambessedesieae: A new Neotropical tribe of Melastomataceae. *Botanical Journal of the Linnean Society*, **190**, 281–302.

Brown J., Paz A., Reginato M., Amaro C., Assis R., Lyra M., Caddah M., Aguirre-Santoro J., D'Horta F., Amaral F.R. do, Goldenberg R., Silva-Brandão K.L., Freitas A., Rodrigues M., Michelangeli F.A., Miyaki C., & Carnaval A. (2020) Seeing the forest through many trees:

multi-taxon patterns of phylogenetic diversity in the Atlantic Forest hotspot. *Diversity and Distributions*, In press.

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., Title 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 filogenéticos em Miconia sect. Discolor (Meslastomataceae, 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., Trewick S.A., & Cook L.G. (2011) Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, **26**, 66–72.

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**, 1030–40.

Diniz C.G., Souza A.A.D.A., Santos D.C., Dias M.C., Luz N.C. Da, Moraes D.R.V. De, Maia 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) Available at: ftp://geoftp.ibge.gov.br/informacoes_ambientais/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 (Bignonieae, 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. 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 randomForest. *R News*, **2**, 18–22.

Mason-Romo E.D., Farıas A.A., & G C. (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., Skean Jr J., 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. *Remote Sensing of Plant Biodiversity* (ed. by J. Cavender-Bares, J. Gamon, and P. Townsend), pp. 255–266. Springer, Cham,

Pereira H.M., Ferrier S., Walter M., Geller G.N., Jongman H.G., Scholes R.J., Bruford M.W., Brummitt N., Butchart S.H.M., Cardoso A.C., Coops N.C., & Dulloo E. (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., Ngereza C., Otte I., Röder J., Rutten G., Schellenberger Costa D., Tardanico J., Zancolli G., Deckert J., Eardley C.D., Peters R.S., Rödel M.-O., Schleuning M., Ssymank A., Kakengi V., Zhang J., Böhning-Gaese K., Brandl R., Kalko E.K.V., Kleyer M., Nauss T., Tschapka M., Fischer M., Steffan-Dewenter I., Gaston K.J., Brown J.H., Fine P.V.A., Rhode K., Mittelbach G.G., Allen A.P., Brown J.H., Gillooly J.F., Hawkins B.A., Kreft H., Jetz W., Hutchinson G.E., Brown J.H., Hurlbert A.H., Stegen J.C., Belmaker J., Jetz W., Romdal T.S., Grytnes J.-A., Colwell R.K., Rahbek C., Gotelli N.J., Brehm G., Colwell R.K., Kluge J., Novotny V., Basset Y., Wiens J.J., Crisp M.D., Kühsel S., Blüthgen N., Classen A., Chao A., Colwell R.K., Lin C.-W., Gotelli N.J., Chao A., Chao A., Burnham K.P., Anderson D.R., Burnham K.P., Anderson D.R., Huyvaert K.P., Grace J.B., Colwell R.K., Brehm G., Cardelus C.L., Gilman A.C., Longino J.T., Duan M., Kessler

M., Rahbek C., Rahbek C., Wiens J.J., Graham C.H., Jablonski D., Merckx V.S.F.T., Wu Y., Kozak K.H., Wiens J.J., Tittensor D.P., Baur B., Meier T., Baur A., Schmera D., McCain C.M., Colwell R.K., Forister M.L., Zhou J., Allen A.P., Gillooly J.F., Savage V.M., Brown J.H., Rolland J., Condamine F.L., Jiguet F., Morlon H., Bagchi R., Johnson D.J., Beaulieu W.T., Bever J.D., Clay K., Hemp A., Hemp A., Trautwein M.D., Wiegmann B.M., Beutel R., Kjer K.M., Yeates D.K., Misof B., Grafen A., Zancolli G., Steffan-Dewenter I., Rödel M.-O., Hemp C., Hemp C., Wang D., Atzberger C., Eilers P.H.C., Colwell R.K., Hurtt G.C., Johnson J.B., Omland K.S., Hsieh T.C., Ma K.H., Chao A., & 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. (2017) 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 apecies 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**, 4061–72.

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 of London. Series B, Biological sciences, **366**, 2536–44.

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., Faith D.P., Mooney H.A., Heip C.H.R., 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.

da 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.D.S., Santos T.G. Dos, Haddad 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.

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.

Venables W.N. & Ripley B.D. (2002) *Modern Applied Statistics with S.* Springer, New York.

Vermote E., Wolfe R., & NASA GSFC and MODAPS SIPS - NASA (2015) *MYD09GA MODIS/Aqua Surface Reflectance Daily L2G Global 1kmand 500m SIN Grid.*

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, Janet 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, Canberra.

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.