



A Brazilian bioregionalization reappraisal based on Angiosperm distribution: A biogeographical and conservation overview

Luísa Lucrecia^{1,*} , Thomas R. Meagher² , Paulo Takeo Sano¹ 

¹ Universidade de São Paulo, Institute of Biosciences, Botany Department, São Paulo, SP, Brazil.

² University of St Andrews, Centre for Biological Diversity, School of Biology, St Andrews, Fife, UK, United Kingdom.

*Corresponding author: luisalucrecia@gmail.com

ABSTRACT

Bioregionalizations constitute an important tool to synthesize biodiversity knowledge, essential for conservation planning. Bioregions currently recognized in the Neotropics have been delineated using different methods and biological groups, frequently suggesting conflicting results. This study aimed to reevaluate the circumscription of bioregions using the distribution of angiosperm species. Angiosperm records were compiled from online herbaria and filtered for taxonomical and geographical accuracy according to *Flora e Funga do Brasil*. A network analysis was performed with InfoMap Bioregions and similarity amongst recognized bioregions was evaluated with an Unweighted Pair Group Method with Arithmetic Mean. Beta diversity patterns were modelled using Generalized Dissimilarity Modelling (GDM) and five climatic variables. A conservation overview was constructed comparing sampling, conservation coverage, deforestation and fire occurrence across bioregions. Seven bioregions were identified and hierarchically clustered: Brazilian subregion (subdivisions not recovered in our analysis); Chacoan subregion, comprised of the Chacoan dominion (four provinces) and the Paraná dominion (two provinces). GDM recognized other heterogeneous areas within these bioregions. The Paraná dominion presented the largest historical anthropization, the Chacoan dominion presented the highest density of fire spots, and the Brazilian subregion had the highest coverage of conservation units while encompassing unknown biodiversity.

Keywords: biodiversity, biogeography; bioregions; conservation; floristic database.

Introduction

Since Augustin de Candolle (1820) proposed the first world regionalization, multiple other bioregionalization proposals have been published with varying biological models, geographical scope and methods (*e.g.*, Morrone, 2014; Nelson, 1978).

The primary goal of these publications was to document nature and its distribution patterns across space, representing abstractions of how biodiversity is organized in response to history and ecology, but numerous other uses have been recognized since (Kreft & Jetz, 2010; Morrone, 2009). Regionalization is a potentially useful

Received January 10, 2025; Accepted May 12, 2025

Editor-in-Chief: Thaís Elias Almeida; Associate Editor: Thiago André

How to cite:

Lucrecia L, Meagher TR, Sano PT *et al.* 2025. A Brazilian bioregionalization reappraisal based on Angiosperm distribution: A biogeographical and conservation overview. *Acta Botanica Brasilica* 39: e20240279. doi: [10.1590/1677-941X-ABB-2024-0279](https://doi.org/10.1590/1677-941X-ABB-2024-0279)



framework for addressing questions regarding ecology, evolutionary biology and systematics, providing a useful tool for the study of specific lineages and assisting the construction of a “grand synthesis” to set priorities for conservation (Bergamin *et al.*, 2017; Kreft & Jetz, 2010; Morrone, 2017; Whittaker *et al.*, 2005). Such a grand synthesis can also be used to guide sampling efforts to provide more geographically comprehensive measures of biodiversity and thus improve future synthesis (Ahrends, Burgess, *et al.*, 2011; Ahrends, Rahbek, *et al.*, 2011).

At present, the identification of bioregions constitutes a challenge; different organisms, types of distributional data and grid sizes render different patterns (Hausdorf & Hennig, 2003; Morrone, 2018). While bioregions have been proposed using different methods and model systems, circumscription of bioregions is often based on methods that are not fully explicit, precise or repeatable (Cabrera & Willink, 1973; Olson *et al.*, 2001; Wallace, 1876). A major drawback of early regionalizations, for instance, is the lack of tightly defined and repeatable criteria to recognize and delineate biogeographical units (Kreft & Jetz, 2010; Vilorio, 2005). However, as applications of bioregions have grown, methods and available datasets to recognize bioregions have also expanded, as well as the emergence of a better nomenclature for naming bioregions (Ebach *et al.*, 2008; Morrone, 2018).

In the last few decades, multiple regionalization schemes have been proposed for the Neotropics with different scopes and methods (*e.g.*, Amaral *et al.*, 2017; DRYFLOR *et al.*, 2016; Olson *et al.*, 2001; Reginato & Michelangeli, 2020). The regionalization scheme compiled by Morrone (2014) is the most geographically comprehensive, recognizing and naming the main neotropical bioregions according to multi-taxa distribution patterns and is periodically updated to incorporate new evidence (Morrone, 2017; Morrone *et al.*, 2022). Although integrating patterns from different model groups is one of the ultimate goals of global regionalization efforts (Morrone, 2018; Procheş & Ramdhani, 2012), the bioregions

recognized by Morrone *et al.* (2022) are a compilation of patterns rendered by a multitude of different methods. Since different methods and model groups usually render different biogeographical patterns (Amaral *et al.*, 2017; Morrone, 2018; Procheş & Ramdhani, 2012) and the use of few taxonomic groups might not be sufficient to identify relevant areas for biodiversity (Oliveira *et al.*, 2019), a comprehensive analysis with an explicit methodology and a comprehensive biological model could shed light on which patterns are expressive in a broader geographical scale.

Parallel to work on regionalization, Brazilian plant taxonomists and herbaria have committed since 2015 through such programs as INCT and REFLORA to make information from herbarium collections more readily accessible and to summarize taxonomic information regarding the Brazilian flora (Flora e Funga do Brasil, 2020; INCT, 2024; REFLORA, 2024). As a result, most Brazilian collection data are readily available online and the majority of plant species occurring in Brazil have their taxonomy and known distribution accessible on a freely available online platform (Flora e Funga do Brasil, 2020), facilitating analysis with large amounts of data and putting Brazilian biodiversity in a prime spot to reevaluate proposed bioregions using a comprehensive taxonomical model. Brazil is one of the most biodiverse countries in the world, and a reevaluation would provide an updated resource for conservation throughout the country. Therefore, the main objective of the present work is to reappraise the bioregions currently recognized by Morrone *et al.* (2022) for Brazil, especially at the hierarchical level of dominions and provinces, using flowering plant distributions, and to present an overview of the spatial distribution of sampling, conservation, and deforestation in identified bioregions.

Material and Methods

Every angiosperm record collected in Brazil with valid geographical coordinates and taxonomic identification at a species level was obtained from the online platforms GBIF (<https://www.gbif>).

org/), Jabot (<http://jabot.jbrj.gov.br/>) and Splink (<https://specieslink.net>) and compiled using R. Synonyms were transferred to their valid names according to Flora e Funga do Brasil (2020). The initial dataset was then filtered for geographic accuracy through the removal of centroids and invalid coordinates using the R package *CoordinateCleaner* (Zizka *et al.*, 2019). Records with geographical coordinates placed outside the state in which the specimen was collected, according to the collector's description, were removed as a second geographic filter. Finally, to reduce taxonomical bias, non-native species were removed from the dataset and specimens were then filtered according to their species distribution in different Brazilian states, according to species descriptions in Flora e Funga do Brasil (2020), following Reginato & Michelangeli (2020).

The filtered dataset was then used to delineate bioregions with Infomap Bioregions 2 (Edler *et al.*, 2017). Infomap Bioregions is a tool that uses network analysis to extract biotic associations from species occurrences and recognize highly interconnected groups of localities, incorporating presence-absence data instead of similarity measures and minimizing the problems of species turnover (Edler *et al.*, 2017; Vilhena & Antonelli, 2015). Bioregions were delineated using 0.25° grid cells and a minimum of 100 data points per cell. This approach, while more conservative to avoid clustering subsampled localities, rendered

multiple areas without information. To incorporate less sampled localities in the regionalization, a secondary analysis was performed with the same grid size and a minimum of 10 data points per cell. This second analysis, although more spatially comprehensive, is more susceptible to sampling artifacts, and, for that reason, its results were used solely to attribute less-sampled locations to circumscribed bioregions, instead of guiding the circumscription itself. Results were exported and compiled in QGIS software, and bioregions were delineated manually by drawing polygons connecting clustered grid cells (Fig. 1). The resultant boundaries depicted in Fig. 2 therefore represent abstractions of the Infomap Bioregions results (Fig. S1-S2) and subsequent assessments using this scheme. A presence-absence matrix was compiled for each delineated bioregion and their similarities were investigated using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA).

Recognized units and their similarity patterns were then compared to previously recognized bioregions, including Morrone *et al.* (2022) (a compilation of multi-taxa distributional patterns) and IBGE (2019) (classification patterns based on climate and vegetation physiognomy). These classifications are somewhat congruent and are widely used for conservation assessments, delineation of conservation policies within Brazil and reconstruction of the biogeographical history

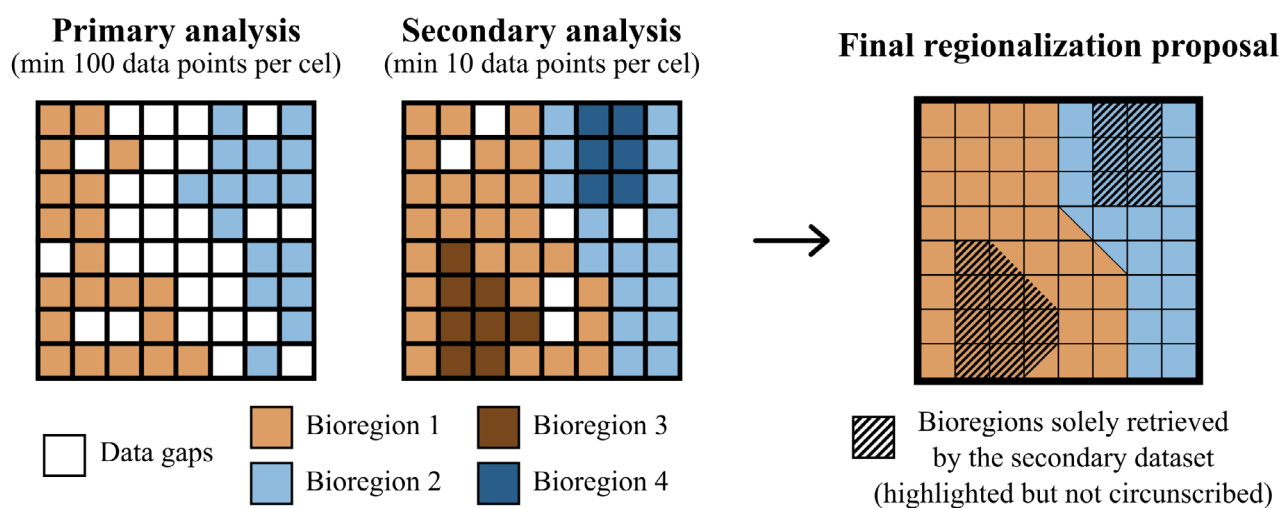


Figure 1. Delineation of bioregions according to Infomap Bioregions outputs of the primary and secondary analysis.



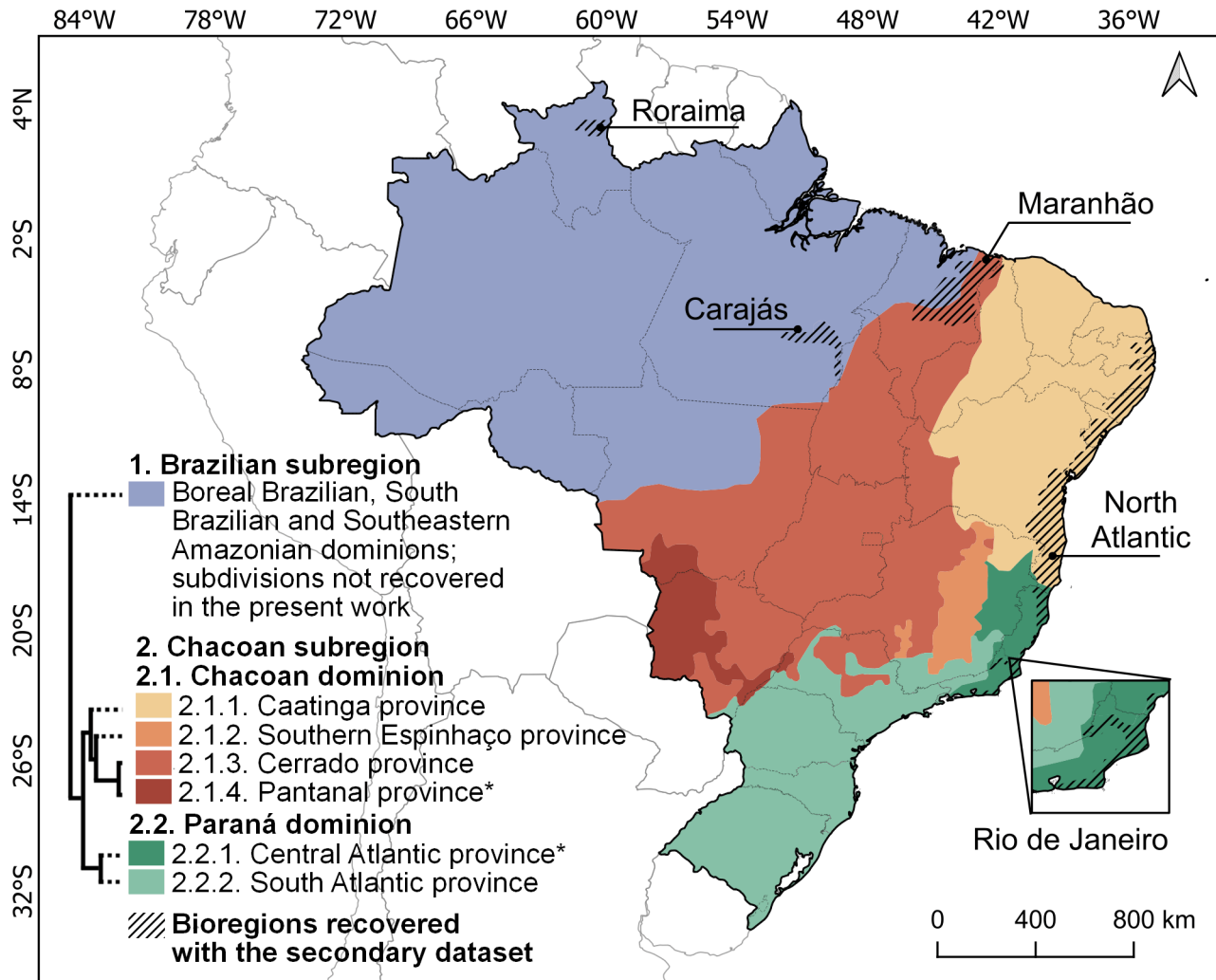


Figure 2. Bioregions recognized from InfoMap Biorregions and their similarity patterns according to UPGMA results. Provinces not recognized by Morrone *et al.* (2022) are highlighted with asterisks (*).

and biodiversity patterns of specific lineages in the Neotropics (e.g. Fortes *et al.*, 2025; Lopes *et al.*, 2024; Magri *et al.*, 2025). Therefore, while we also compare our results with other published classifications (e.g. Cabrera & Willink, 1973; Reginato & Michelangeli, 2020), we primarily paralleled our results with these two classifications.

To investigate beta diversity patterns, we used the primary dataset in a Generalized Dissimilarity Modelling (GDM) (Ferrier, 2002; Ferrier *et al.*, 2007). The method uses maximum likelihood and modelled I-splines to predict the best-supported function between climatic variation, geographical distance and floristic dissimilarities between pairs of sites (Ferrier *et al.*, 2007). Nineteen climatic

variables were compiled from CHELSA (Karger *et al.*, 2017) and elevation was retrieved from WorldClim (Fick & Hijmans, 2017), the median of each was measured for the grid cells analyzed, and these metrics were compiled in a matrix of explanatory variables using QGIS software (version 3.6.2). Colinearity in GDM is still underexplored, but it could potentially be problematic, as the analysis combines elements of generalized linear modelling and matrix regression (Leathwick *et al.*, 2006). For that reason, environmental variables were dropped based on biological knowledge (e.g., removal of variables not previously related to turnover in literature; Zuur *et al.*, 2010) and on the Variance Inflation Factor (VIF). The VIF was

measured in R with the package *usdm* (Naimi, 2023) and variables with a VIF higher than 10 were progressively excluded following Montgomery & Peck (1992). The final GDM was performed with five non-collinear variables (Annual Precipitation, Elevation, Precipitation Seasonality, Temperature Annual Range and Temperature Seasonality). In addition to environmental predictors, geographical distance between grid cells was included as an additional predictor. Dissimilarities between locations according to the fitted model were then plotted using the RGB color spectrum. All GDM procedures were performed in R using the package *gdm* (Fitzpatrick, 2022).

To map conservation efforts in each of the recognized bioregions, the spatial distribution of conservation units (UC) was obtained from TerraBrasilis (Assis *et al.*, 2019). Conservation units in Brazil are differentiated into two groups according to their objectives and permitted activities: sustainable use UCs restrict human activities with the main goal of regulating the sustainable use of natural resources; while the natural resources in full protection UCs are only available through indirect use, to preserve nature (Brasil, 2000). Since these fundamental differences might alter the conservation of natural landscapes, conservation units were differentiated into sustainable use UCs (UC-US) and full protection UCs (UC-PI). Also, since indigenous territories (TI) are positively associated with natural landscapes, the distribution of these was obtained from TerraBrasilis as a type of indirect conservation unit (Assis *et al.*, 2019; Garnett *et al.*, 2018).

The MapBiomass landcover classification of 2022 was used to delineate anthropogenic landscapes, such as crops, pastures, urban perimeters, and mining (MapBiomass, 2022; Souza *et al.*, 2020). The percentage of the territory attributed to natural vegetation or anthropogenic landscapes was determined using Google Earth Engine (Gorelick *et al.*, 2017). Although fire might be a natural occurrence in some Brazilian vegetation (Pivello, 2011), the occurrence of fire spots is highly correlated with economic objectives and

the rising demand for land clearing and land use associated with agriculture and cattle grazing (Arima *et al.*, 2007). For that reason, we used the occurrence of fire spots in 2023-2024 (obtained from the TerraBrasilis dataset; INPE, 2024) as a proxy for current deforestation. Fire spot density in each of the recognized bioregions and the percentage of territory within conservation units were calculated in QGIS software.

Results

The compiled dataset encompassed 1,166,285 angiosperm records and 23,741 species. The remaining 12,667 species of angiosperms occurring in Brazil (Flora e Funga do Brasil, 2020) did not have records with valid geographic data and therefore were not included in the network analysis. The analysis recovered 38 units based on species distribution, of which seven presented over 10 grid cells and were circumscribed as bioregions. The UPGMA clustered these bioregions in broader units, here putatively associated with dominions and subregions (Fig. 2). The complete list of species and endemic species for each recognized bioregion is available in the supplementary material (Table S1).

The GDM explained half the variation in our dataset (50.2%) and recovered some of the bioregions suggested by the network analysis (Fig. 3). Amongst the subdivisions not recovered is the main structure in three major dominions (Brazilian subregion, Chacoan and Paraná), suggesting a smooth beta diversity transition across these areas. Still, the analysis highlighted localities that might be climatically distinct within the circumscribed bioregions and thus might also contain distinct floristic composition, like the northernmost portion of the Brazilian subregion in the Roraima state (Fig. 3). In addition to geographic distance, the climatic predictors most strongly associated with turnover according to the model were annual precipitation, elevation, temperature seasonality, precipitation seasonality and temperature annual range (Fig. S3; Table S2).



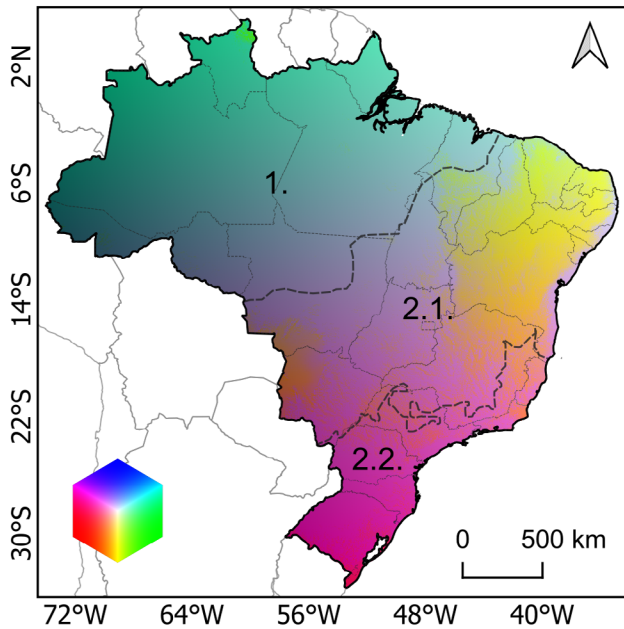


Figure 3. GDM results, indicating modelled beta diversity turnover according to climatic predictors. Colors on different sides of the spectrum represent higher predicted dissimilarities between locations.

Brazilian subregion (Amazon Biome)

The first circumscribed bioregion is largely congruent with IBGE delimitation of the Amazon Biome and Morrone *et al.* (2022) Brazilian subregion, including its Boreal Brazilian and South Brazilian dominions, and the South-eastern Amazonian dominion – attributed to the Chacoan subregion (Fig. 4). The circumscription of these dominions was not recovered, and the bioregion was highly distinct when compared to other circumscribed units. In the secondary analysis (with a minimum of 10 records per grid) the Carajás and Roraima regions were highlighted as the only distinct areas within the Brazilian subregion (Fig. 2). The GDM showed yet another pattern, highlighting a north/south structure, represented by the green/purple transition in Fig. 3, and a west/east structure, depicted by the darker/lighter green transition. The analysis also subtly highlighted the northernmost portion of the Brazilian subregion (north Roraima state) as a climatically distinct locality, depicted in bright green.

The Brazilian subregion included 132,518 angiosperm records (11% of the dataset) and 7,625 species, of which 4,479 (59%) are endemic (Table 1; Fig. 5). This was the largest circumscribed bioregion, encompassing 47.4% of Brazil, although grid cells were sparse across its geographical range (Fig. 6). Sampling density in the area was the lowest across circumscribed bioregions, with 0.03 records/km² (Table 1). This subregion presents the highest percentage of territory within protected areas, with 17.9% of its area in Sustainable use UCs, 10.3% in full protection UCs, and 23.4% in indigenous territories (Fig. 5; Fig. 6). Fire spot density was 0.025 units/km², with fire spots concentrated in its south and east borders and historical anthropization was one of the lowest recorded in Brazil, with only 15% of its vegetation modified in anthropogenic landscapes (Fig. 5; Fig. 6).

Chacoan subregion

The remaining areas investigated were clustered in a larger floristic bioregion, putatively associated with the Chacoan subregion (Morrone *et al.*, 2022). This subregion had the highest sampling (1,033,767 records, 89% of the dataset) and Brazilian diversity (19,262 spp., of which 16,116 are endemic) and was further subdivided into two areas putatively associated with the Chacoan and Paraná dominions, being those further subdivided into four and two provinces, respectively.

Chacoan dominion

The Chacoan dominion contained 614,186 records (53% of the dataset) and 14,129 species, of which 6,940 (49%) are endemic (Fig. 5). The Chacoan dominion encompasses 40.1% of Brazil and it was further subdivided into four provinces according to the main dataset: Caatinga, Southern Espinhaço, Cerrado and Pantanal (Fig. 2). The secondary analysis (with minimum 10 records per grid) suggested a further division of the Caatinga province, segmenting its coastal area from the

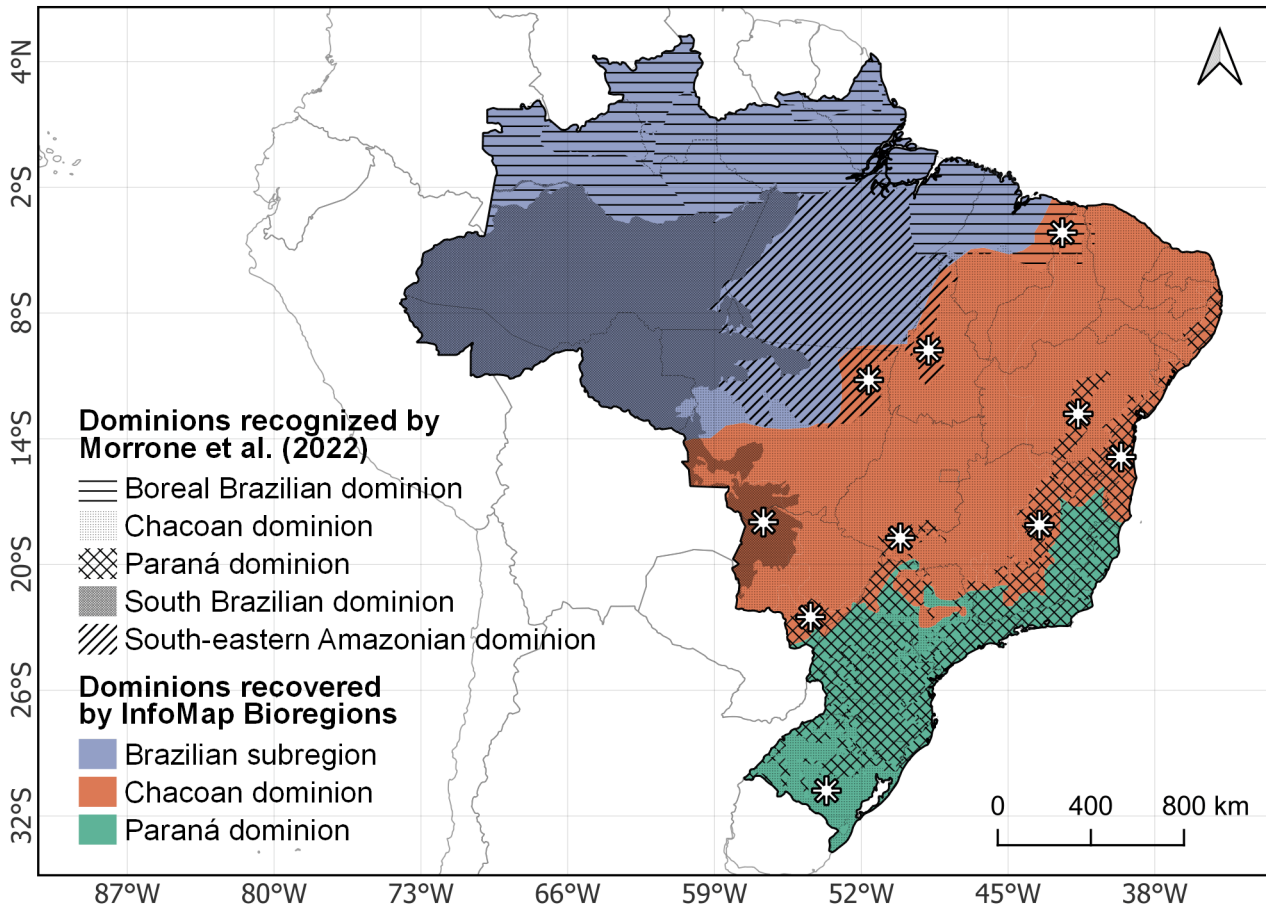


Figure 4. Dominions recognized from InfoMap Bioregions in contrast with the dominions recognized by Morrone *et al.* (2022). Extension incongruences are highlighted with asterisks (*).

interior, and a separation of the northmost area of the Cerrado province (a portion within Maranhão state somewhat congruent with the Pará province from Morrone *et al.*, 2022). The GDM somewhat recovered the circumscribed provinces, with Caatinga highlighted in yellow and Pantanal in orange in a matrix of pink/purple/green transition within the main section of the Chacoan (associated with the Cerrado province). The Southern Espinhaço was the only province not highlighted (Fig. 3).

Sampling in this dominion was heterogeneous (Fig. 6), more intensely concentrated in its central and eastern portions, coinciding with the state of Goiás, coastal states (Pernambuco, Rio Grande do Norte and Paraíba) and Chapada Diamantina (Bahia). Sampling density in the whole dominion was 0.18 records/km², although this parameter varied within recognized provinces (Fig. 5). The

Southern Espinhaço had the highest sampling effort with 0.57 records/km², followed by Caatinga (0.24 records/km²) and Pantanal (0.18 records/km²), while the Cerrado presented the smallest sampling density (0.12 records/km²).

Known diversity also varied within provinces (Fig. 5), with more species allocated in Caatinga (7,673 spp.) and Cerrado (7,594 spp.), while the Southern Espinhaço included 5,554 species and Pantanal had the lowest species richness, with 2,817 species. Endemism also varied, proportionally highest in Caatinga (31%; 2,375 spp.), Southern Espinhaço (22%; 1,204 spp.) and Cerrado (20%; 1,494 spp.). The Pantanal province included 276 endemic species, accounting for 10% of its flora. Despite the low endemism, it holds a remarkable concentration of biodiversity – including large populations of threatened species (Harris *et al.*, 2005).



Table 1. Sampling, diversity and conservation overview across recognized bioregions.

Bioregions	Diversity and Sampling				Conservation Units				Fire spots (unit/km ²)	Anthropogenic landscapes (%)
	Records	Species	Endemic Spp.	Sampling density (records/km ²)	UC-US (%)	UC-PI (%)	TI (%)	Total		
1. Brazilian subregion	132518	7625	4479	0,032	17,93%	10,28%	23,43%	51,64%	0,025	15,04%
2. Chacoan subregion	1033767	19262	16116	0,238						
2.1. Chacoan dominion	614186	14129	6940	0,181	4,90%	2,37%	2,61%	9,88%	0,025	25,99%
2.1.1. Caatinga province	258520	7673	2375	0,240	6,98%	2,16%	0,52%	9,66%	0,025	13,87%
2.1.2. Southern Espinhaço province	76216	5554	1204	0,567	3,74%	3,99%	0,03%	7,75%	0,012	88,89%
2.1.3. Cerrado province	238797	7594	1494	0,121	4,04%	2,32%	3,88%	10,25%	0,025	17,08%
2.1.4. Pantanal province	40653	2817	276	0,183	3,07%	2,86%	3,12%	9,05%	0,034	36,36%
2.2. Paraná dominion	419581	10844	5091	0,441	5,90%	2,48%	0,39%	8,78%	0,008	64,51%
2.2.1. Central Atlantic province	137113	6430	1494	0,891	5,85%	2,34%	0,19%	8,38%	0,010	71,89%
2.2.2. South Atlantic province	282468	8088	2695	0,354	5,92%	2,51%	0,43%	8,86%	0,008	63,14%
Anthropogenic landscapes (%)					11,79%	2,55%	2,16%			



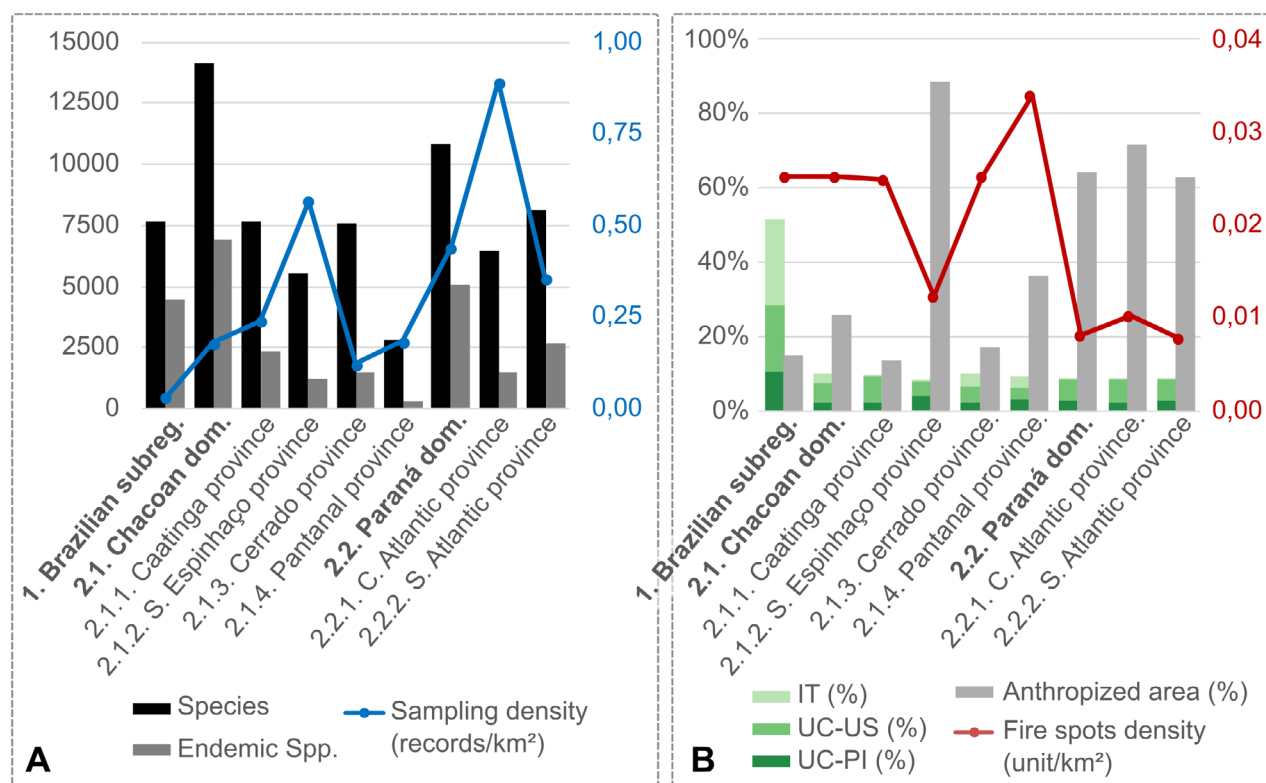


Figure 5. (a) Sampling density, species richness and endemism; and (b) conservation overview of the recognized bioregions.

The Chacoan dominion included 26% of anthropogenic landscapes and a fire spot density of 0.025 units/km² (Table 1). The provinces with a higher proportion of anthropogenic landscapes were Southern Espinhaço (88.9%) and Pantanal (36.3%), while the Cerrado and Caatinga had 17% and 13.8%, respectively. Fire spot density, on the other hand, was higher in Pantanal (0.03 units/km²), followed by Cerrado and Caatinga (0.025 and 0.024 units/km², respectively) (Fig. 5; Fig. 6). The Southern Espinhaço, although presenting the largest proportion of anthropogenic landscapes, had one of the smallest fire spot densities, with 0.01 units/km².

Conservation units make up 9.8% of the Chacoan dominion (Fig. 5; Fig. 6), with 4.9% in sustainable use UCs, 2.3% in full protection UCs and 2.6% in indigenous territories. This coverage also varied amongst provinces: Cerrado had the highest coverage of conservation units (10.2%), followed by Caatinga (9.6%) and Pantanal (9%), while southern Espinhaço had the lowest

coverage (7.7%), although presenting the highest proportion of its territory under full protection conservation units (4%).

Paraná dominion

The Paraná dominion included 419,581 records (36% of the dataset) and 10,844 species, of which 5,091 (47%) are endemic (Fig. 5). The Paraná dominion was the smallest dominion identified, encompassing 11.2% of Brazil. It was further segmented into two provinces according to the main analysis: Central and South Atlantic. The secondary analysis (with minimum 10 records per grid) also suggests a subdivision of the Central Atlantic province, segmenting a small portion in the coast of Rio de Janeiro state from a larger area within Espírito Santo and Minas Gerais states (Fig. 2). The GDM did not suggest a strong dissimilarity structure within this dominion, with a small variation north/south – represented by the transition from light pink/yellow (north) to dark pink/red (south) (Fig. 3).



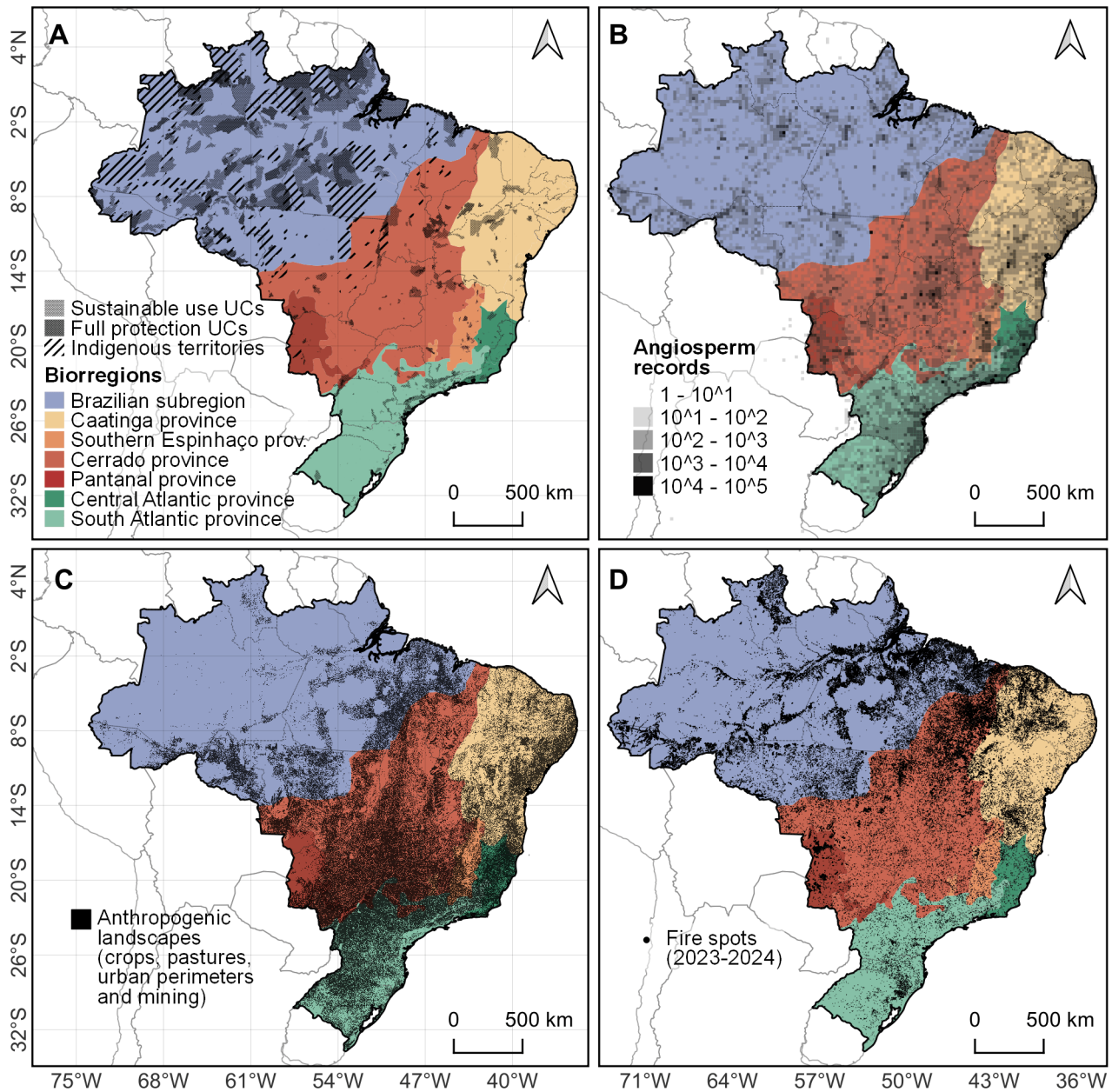


Figure 6. Spatial distribution of (a) conservation units, (b) sampling, (c) anthropogenic landscapes, and (d) fire spots in the recognized bioregions.

Sampling was heavily concentrated in the coastal portion, with large sampling gaps in areas closest to the Chacoan dominion (Fig. 6). Average sampling density was 0.44 records/km², although the Central Atlantic province presented a record density of 0.89 records/km² - the highest across all the recognized bioregions, while the South Atlantic province presented an average of 0.35 records/km². Diversity also varied within provinces (Fig. 5), with more species richness and

endemism occurring in the South Atlantic, with 8,088 spp., of which 2,695 are endemic (33%), compared to the Central Atlantic, with 6,430 species, of which 1,494 (23%) are endemic.

Historical anthropization in the Paraná dominion was the highest across recognized bioregions, with an average of 64.5% of anthropogenic landscapes (Fig. 5; Fig. 6). The Central Atlantic province included a proportionally higher amount of these, with 71.9%, while the

South Atlantic province had 63.1% of its territory anthropized. Fire spot density was low across the entire dominion (0.008 units/km²), but far more concentrated in the Central Atlantic (0.01 units/km²) than in the South Atlantic (0.007 units/km²). Conservation units comprise 8.8% of the dominion (Table 1; Fig. 5), with 5.9% in sustainable use UCs, 2.48% in full protection UCs and only 0.39% in Indigenous territories. This coverage varied little amongst recognized provinces: South Atlantic presented 8.36% of its territory in UCs while Central Atlantic had 8.38%.

Discussion

Congruence among recognized bioregions

The subregions, dominions and provinces recovered were overall congruent with the ones currently recognized by Morrone *et al.* (2022) and IBGE (2019) and reflect some patterns previously documented by other authors. The segmentation of Brazil into three main sectors (Brazilian subregion/Amazon, Chacoan dominion/Dry diagonal and Paraná dominion/Atlantic Forest), for instance, has been previously documented (Cupertino-Eisenlohr *et al.*, 2021; Morrone *et al.*, 2022). The floristic similarity patterns supported by our results have also been suggested by these authors, with the Brazilian subregion (Amazon) presenting greater affinity with Mesoamerica and the Andes while the Chacoan and Paraná dominions are more floristically related to each other, despite the Amazon and the Atlantic Forest presenting the same dominant biome (Olson *et al.* 2001). Still, a few assessments highlight past or current connections between these two massifs (Ledo & Colli, 2017) and illustrate such connections through the number of shared species between them – so far 757 spp. reported by Cupertino-Eisenlohr *et al.* (2021), now updated to 1.668 spp., from which most also occur in the dry diagonal (97.5%; 1.626 spp.) (Table S1).

The combination of dominions related to the Brazilian subregion in one cohesive bioregion, on the other hand, is the largest incongruence when

comparing our results with the circumscription proposed by Morrone *et al.* (2022) and other authors that have found internal structure within the Amazon (Luize *et al.*, 2024; Silva-Souza & Souza, 2020). This result can be a consequence of the scale used in the present analysis, as the internal structure within the Amazon has been recovered when its diversity is assessed individually using data from local assemblages (Luize *et al.*, 2024; Silva-Souza & Souza, 2020). Still, the high cohesiveness of the Brazilian subregion has been previously documented and hypothesized to be linked to the absence of major ecological and physical barriers to dispersal and establishment of plants (Cupertino-Eisenlohr *et al.*, 2021; Dexter *et al.*, 2017). The model group used can also be another source of disagreement, as those authors have worked exclusively with woody species patterns. Analysis that includes all angiosperm species mix ecological groups with distinct functions in the ecosystem (*e.g.*, terrestrial herbaceous, epiphytes and woody) and different biological models are known to render different biogeographical patterns (Amaral *et al.*, 2017).

Nonetheless, our findings could also be a reflection of low sampling effort and data scarcity, a known limitation of InfoMap Bioregions and biodiversity assessments in general (Edler *et al.*, 2017; Yang *et al.*, 2013). The Amazon currently represents the biggest sampling gap in Brazil (Hopkins, 2007; Oliveira *et al.*, 2016; Oliveira *et al.*, 2019; Sousa-Baena *et al.*, 2014), which is reinforced by our dataset (Fig. 6). This knowledge gap includes not only distributional data of known species (Wallacean shortfall) but also many uncatalogued species (Linnean shortfall): two of the more severe biodiversity shortfalls, as they represent a lack of basic knowledge that affects other aspects of biodiversity information (Hortal *et al.*, 2015). This Wallacean shortfall is illustrated through the comparison of the total number of angiosperm species we recorded for the Brazilian subregion (7,625 spp.) and the number found by combining floristic lists (10,655 spp.; Cardoso *et al.*, 2017). The Linnean shortfall, on the other hand, could be illustrated through the comparison



of the total number of arboreal species found by combining floristic lists (7,005 spp.; Cardoso *et al.*, 2017) with tree species richness estimates for the Amazon (over 15,000 spp.; Ter Steege *et al.*, 2020).

It is possible that, as a result of this knowledge gap, known heterogeneous areas within the Amazon, like the Pantepui region and its ferruginous outcrops, are homogenized within a matrix of ombrophilous forests despite known floristic and ecological differences and high volume of endemic species (Cabrera & Willink, 1973; Rull & Vegas-Vilarrúbia, 2020; Viana *et al.*, 2016). This heterogeneity could be indicated in the cluster performed with the secondary dataset through the circumscription of the Carajás and Roraima regions (Fig. 2), but other localities known to have somewhat similar floras to this area were not included in the regionalization (Barbosa-Silva *et al.*, 2022) – therefore indicating these circumscriptions might represent artifacts of punctual high sampling in few floristically distinct areas.

Within the Chacoan dominion, the circumscription of the Caatinga and Cerrado provinces agrees with every bioregionalization scheme proposed so far (Amaral *et al.*, 2017; Dryflor *et al.*, 2016; Morrone *et al.*, 2022; Reginato & Michelangeli, 2020). The circumscription of the Pantanal province within the Chacoan dominion and Chacoan subregion, on the other hand, has not been a pattern consistently reported. The area has been identified according to climate and vegetation physiognomy as a distinct unit (IBGE, 2019; Olson *et al.*, 2001); classified by Morrone *et al.* (2022) as a part of the Rondonia province (attributed to the South Brazilian dominion and the Brazilian subregion); and included within the Cerrado province according to floristic similarity patterns of woody species (Dryflor *et al.*, 2016; Ratter *et al.*, 2003; Silva-Souza & Souza, 2020). Although highlighted by Silva-Souza and Souza (2020) as a distinct subunit within Cerrado, the authors also highlight at least seven other subregions according to compositional data and in that assessment, the two Pantanal subregions

were not clustered as cohesive or distinct areas apart from Cerrado.

Alternatively, our results suggest the area is a province of the Chacoan dominion and most closely related to the Cerrado province. The high similarity between the Pantanal and Cerrado provinces concurs with previous assessments using woody species compositional data and ecological similarities between these localities, since they share a dominant cover of savanna vegetation (Cole, 1960; Silva & Bates, 2002). Still, we highlight that different biological models render wildly different biogeographical patterns and our results are more congruent with what has been retrieved for woody species (*e.g.*, Dryflor *et al.*, 2016; Ratter *et al.* 2003; Silva-Souza & Souza, 2020), while biogeographical assessments targeting herbs suggest the Pantanal to be most closely related to Southeast Cerrado and the Paraná dominion (Amaral *et al.*, 2017).

The circumscription of the Maranhão region by the secondary analysis, on the other hand, concurs with biogeographical assessments based on herbs, which suggest the area has a greater floristic similarity with others within the Brazilian subregion (Amaral *et al.*, 2017). The Maranhão region has also been circumscribed by Morrone *et al.* (2022) as a part of the Boreal Brazilian dominion (Brazilian subregion) and highlighted by Silva-Souza and Souza (2020) as the most dissimilar subunit within the Cerrado dominion. Still, primary analysis classified the region within the Chacoan subregion and Cerrado provinces. This incongruence might reflect the different biological models used in each dataset, but these results suggest further biogeographical structures that should be examined in the future. Another large incongruence when comparing our results with prior bioregionalizations (Moro *et al.*, 2024; Morrone *et al.*, 2022) is the circumscription of the Caatinga province, including both the Chapada Diamantina province and the Northern portion of the Paraná dominion. The Chapada Diamantina is a well-sampled location within the Caatinga province and, although described as a *campo rupestre* province within the Chacoan



dominion, was incorrectly attributed to the Paraná dominion in Morrone *et al.* (2022). In the original manuscript of this circumscription (Colli-Silva *et al.*, 2019), the area is easily distinguished when using rock outcrop endemics but harder to detect when analyzing all angiosperms occurring in the area, dissolving within Caatinga when improving analysis cluster cost. Therefore, our results reinforce Chapada Diamantina's floristic affinity with Caatinga. This floristic link might be related to the specific climatic conditions within the Semiarid and Caatinga unit, allowing characteristic species from Caatinga that are adapted to xeric conditions to penetrate Chapada Diamantina (Harley, 1995; Queiroz *et al.*, 2005); or it might be possible that the grid size used to assess floristic composition is too wide and it does not reflect accurately the heterogeneity of highly fragmented landscapes such as Chapada Diamantina (Harley, 1995; Lucrezia *et al.*, 2021).

We also highlight that our results suggest modifications in the geographical extent of the Southern Espinhaço province, another *Campo Rupestre* province (Colli-Silva *et al.*, 2019; Morrone *et al.*, 2022). Our results suggest the province should also include the Mantiqueira and Canastra Ranges, concurring with Reginato & Michelangeli (2020). The Southern Espinhaço province was recognized by Colli-Silva *et al.* (2019) with a *campo rupestre* dataset, therefore limited to records occurring in quartzite outcrops. Although Azevedo *et al.* (2024) have highlighted the role of edaphic predictors in assessing species composition in rock outcrops, our results suggest that the Southern Espinhaço province is not limited to quartzite outcrops, including neighboring montane environments. Nonetheless, these broadly defined floristic units do not refer to the biogeographical history of a single vegetation but instead represent the combined identity of the region. Therefore, this result may reflect the mixed biogeographical signal of rock outcrops and other types of vegetation. We suggest that these areas should be explored with other datasets, grid sizes and methods to understand how biodiversity is

structured, especially given their conservation relevance (Martinelli, 2007).

The division of the Paraná dominion into north/south sectors here recovered has been reported previously (*e.g.*, Brown *et al.*, 2020; Peres *et al.*, 2020; Reginato & Michelangeli, 2020) and putatively linked to climatic differences between these sectors (Peres *et al.*, 2020). Other assessments have further subdivided the area into smaller subunits (Brown *et al.*, 2020; Cantidio & Souza, 2019), although those vary according to the method and model group assessed. Our results are the first assessment that includes the northern and central sectors *sensu* Peres *et al.* (2020) within the Caatinga province (here referred to as North Atlantic; Fig. 2) and that further subdivides the southern portion of the Paraná dominion in Central and South Atlantic. This subdivision somewhat coincides with multi-taxa phylogenetic turnover patterns (Brown *et al.*, 2020) but has never been formally proposed at this scale. On a finer scale, on the other hand, Cantidio & Souza (2019) suggest the strongest secondary subdivision of the Atlantic Forest to be a coastal/mainland subdivision, which could be associated with the segmentation of the coast of Rio de Janeiro, suggested by our secondary analysis.

These circumscriptions might represent sampling artifacts rather than real patterns. We hypothesize the inclusion of the North Atlantic in the Caatinga might be a result of its high fragmentation, as most of the vegetation cover in this area has been lost (Rezende *et al.*, 2018). In that regard, it is possible that: (1) sampling within fragments of natural vegetation is very low; or (2) that the original biodiversity and biogeographical footprint of the area has already been lost at this scale, as it has been severely disturbed for years (Silva & Tabarelli, 2000). While the Atlantic Forest today represents the most well-sampled location in Brazil (Sousa-Baena *et al.*, 2014), its sampling might reflect known biases for biodiversity sampling, like proximity to urban centers and research institutes, therefore favoring sampling in its southern portion (Oliveira *et al.*, 2016).



The regionalization of the Rio de Janeiro coast by the secondary analysis might represent a similar artifact to the recognition of the Carajás region – an artifact of punctual high sampling in a floristically distinct, yet underexplored, bioregion. The north/south and east/west patterns suggested by the GDM for the Brazilian subregion somewhat reflect the tree beta diversity patterns recovered for the Amazon by Luize *et al.* (2024). The model estimates beta diversity turnover, and despite explaining a high proportion of variance (50%), it does not include other environmental variables known to be relevant for plant distribution, such as edaphic properties (Higgins *et al.*, 2011; Luize *et al.*, 2024). Still, the results highlight climatic patterns that might suggest interesting opportunities to improve sampling. Although others have highlighted sampling priorities based on sufficient knowledge of biodiversity and vegetation fragmentation (e.g., Oliveira *et al.* 2019), the GDM results might indicate areas that could be prioritized as expected to contain a distinct assemblage of species.

Conservation across recognized bioregions

The recognized bioregions are subjected to distinct threats and are heterogeneously anthropized and conserved. The Brazilian subregion, largely congruent with the Amazon, has the most recent history of anthropization. Deforestation in the area began in 1970 with the inauguration of the Transamazon Highway, reached its maximum by 2004 and was significantly reduced by 2012 in compliance with the Brazilian government's Action Plan for the Prevention and Control of Deforestation in the Legal Amazon. Deforestation in the area started to rise again from 2016 to 2020 (Fearnside, 2005; Silva Junior *et al.*, 2020), which is illustrated by the number of fire spots. The small amount of historical anthropogenic landscapes in the area (Fig. 5; Fig. 6) also reinforces that deforestation in the Brazilian subregion is recent, concurring with other global assessments suggesting its high conservation status (Dinerstein *et al.*, 2017). The main activities driving the last 50

years of anthropogenic expansion in the area are cattle grazing, agriculture, logging, and mining, especially in its southern and eastern portions, an area known as the 'arc of deforestation' (Fig. 6) (Fearnside, 2005; Ometto *et al.*, 2011). The spatial distribution of anthropogenic landscapes is correlated with the distribution of access routes, which are scarce (Fearnside, 2015; Garnelo *et al.*, 2020; Silva *et al.*, 2023).

The overall lack of access routes, while possibly correlated with the small historical deforestation in the area, also poses a challenge for human occupation (Garnelo *et al.*, 2020) and severely biases sampling of the area (Hughes *et al.*, 2021; Oliveira *et al.*, 2016). Therefore, while the area is historically preserved and encompasses a high coverage of conservation units, its biodiversity is severely miscomprehended, with a sampling density of 0.03 records/km². Our data reflects known patterns of Amazon subsampling (Oliveira *et al.*, 2016), whereas Oliveira *et al.* (2019) highlight a few well-known locations within the Brazilian subregion as high priority for conservation and most of its territory as gaps where evidence-based conservation decisions are not possible yet. The recognition of conservation priorities that led to the circumscription of current conservation units in the Amazon were mostly delineated using phytogeographic regions available at the time (e.g. Ducke & Black, 1953; Prance, 1973), vegetation types and the concept of Pleistocene refugia (Mittermeier *et al.*, 2005) – although most of these authors recognize data scarcity as a limitation of their conclusions.

Despite the high coverage of conservation units, fire spot density in the Amazon is high (Fig. 5). Recent anthropogenic activity and fire occurrences pose an irreversible threat to the vegetation in the area, as documented by Ducke & Black (1953) – forested areas within the Amazon, after burning, are converted into different phytophysognomies with smaller diversity that often never returns to its original state. The occurrence of periodic fires also reduces canopy cover, drastically changes vegetation structure along forest edges, sharply increases the density



of vines, lianas and ruderal species and enhances the likelihood of new fires, starting an ecological cascade that renders immense biodiversity loss (Tabarelli *et al.*, 2004).

In the Chacoan dominion, exploration of natural resources started earlier, with the gold rush in the 18th century, especially in the Southern Espinhaço province (Costa & Rios, 2022; Derby, 1906). This area, while highlighted as the most anthropized of all provinces recognized, was circumscribed based on the species composition of highlands, while most of its anthropogenic landscapes are placed in the surrounding lowlands (Fig. 6). The area has been consistently highlighted for its immense biodiversity and endemism in rock outcrops (Colli-Silva *et al.*, 2019; Echternacht *et al.*, 2011), and its conservation units are placed within these well-sampled localities (Fig. 6). The greatest threat for the well-sampled highlands is mining and intensive farming (Castro Pena *et al.*, 2017) and the areas set as conservation priorities within this province were designated based on these phytophysiognomies (Monteiro *et al.*, 2018). The surrounding lowlands, on the other hand, do not have their biodiversity catalogued or threats mapped.

Similar to the Brazilian subregion, the Chacoan dominion had most of its destruction carried out in the past 50 years through the construction of highways and expansion of agriculture (Ratter *et al.*, 1997). Estimates of deforestation between 1990-2010 within the Cerrado and Caatinga provinces show that historically Cerrado had a more intensive history of deforestation (Beuchle *et al.*, 2015), which is reflected in the higher percentage of anthropogenic landscapes (Fig. 5). The area was largely devastated by agriculture and cattle grazing, with its aluminum-rich and acidic soils being modified by heavy applications of lime and fertilizer (Ratter *et al.*, 1997). Other global assessments found the Caatinga as the most threatened province within the Chacoan dominion, highlighting it as one of the Nature Imperiled global ecoregions (Dinerstein *et al.*, 2017). Both Caatinga and Cerrado are here reported with similar fire densities, demonstrating

they still face significant vegetation loss; although fire in these areas might also occur naturally in lower frequencies, especially in savanna phytophysiognomies (Martins *et al.*, 2024; Miranda *et al.*, 2009).

Despite the high anthropization in these provinces, sampling is still very low – indicating that biodiversity might be lost without ever having been catalogued. Conservation in these provinces is also insufficient: the highest gaps in conservation units in Brazil are in the Chacoan dominion, specifically in Caatinga, Cerrado and Pantanal (Fonseca & Venticinque, 2018). The Pantanal province, while highlighted as a conservation gap by Fonseca & Venticinque (2018), presented one of the highest percentages of anthropogenic landscapes (36.36%) and the highest concentration of fire spots amongst all the evaluated units (0,033 units/km²). The greatest conservation threats in the province include cattle grazing and the introduction of invasive grass species (Harris *et al.*, 2005). Additionally, the anthropization of neighboring Cerrado has led to severe erosion and the deposition of pesticides in the lowlands of the Pantanal, modifying the environment and threatening its aquatic biota (Harris *et al.*, 2005). Therefore, our results suggest the Pantanal province might be the most threatened environment under the Chacoan dominion.

Finally, the areas associated with the Atlantic Forest, mostly recognized here as the Paraná dominion, started its anthropization over 500 years ago, intensifying with increased urbanization on the coast of Brazil (Morellato & Haddad, 2000). The natural vegetation cover of the Paraná dominion was drastically reduced by 1990 (Morellato & Haddad, 2000; Rezende *et al.*, 2018), coinciding with our data, which suggests the highest percentage of anthropogenic landscapes in this area. Recent estimates of vegetation cover range from 10 to 16% of natural vegetation remaining, with urban expansion as one of its greatest threats and vegetation remnants extremely fragmented (Lembi *et al.*, 2020; Seto *et al.*, 2012). Additionally, ca. 83% of vegetation fragments are small (under



50 ha) and close to forest edges, indicating that matrix influences may have strong effects on many forests' ecological processes (Ribeiro *et al.*, 2009). Still, these fragments persist as the sole remnants of an area consistently highlighted for its astounding diversity and endemism (Myers *et al.*, 2000; Rezende *et al.*, 2018).

The historical exploitation of the Atlantic Forest is also reflected in the high density of access routes and urban centers, which facilitate sampling (Oliveira *et al.*, 2016). This is illustrated by the high sampling density observed (0.44), especially in the Central Atlantic province (0.89 records/km²). Still, this area is also highlighted as a dark spot of Linnean and Wallacean shortfalls and a collection priority in a few scenarios, especially when considering its low coverage of conservation units (Ondo *et al.*, 2024). Its high sampling is probably correlated with the high species richness observed in the area (10,844 spp.) despite the low geographical extent. Sampling and funding are notoriously correlated with perceived species richness, as well as proximity to urban centers and research institutions, all of which are high in the Paraná dominion (Ahrends, Burgess, *et al.*, 2011; Ahrends, Rahbek, *et al.*, 2011; Hughes *et al.*, 2021; Oliveira *et al.*, 2016). Since these variables play a significant role in documented biodiversity and species richness, we do not infer differential diversity in recognized bioregions. Although the number of species catalogued in each is brought as a result, we highlight these numbers to be heterogeneously underestimated.

The coverage of conservation units, on the other hand, is still overestimated when accounting for every type of conservation unit as a conservation effort. While the objective of sustainable use UCs is not to safeguard native vegetation (Brasil, 2000), these conservation units are accounted for when discussing conservation efforts in the country. While most anthropogenic landscapes in Brazil were recorded outside conservation units, UCs also presented anthropization in various degrees. The average deforestation percentage of Brazil was 10.9%, while deforestation in full protection UCs and indigenous territories was 2.55% and 2.16%, respectively. Alternatively,

the percentage of anthropogenic landscapes within sustainable-use UCs was higher than the national average, with 11.79% (Table 1). These values suggest that the sustainable use of UCs might not always be effective for the conservation of natural landscapes – in accordance with its objectives (Brasil, 2000). Extractive reserves, for instance, were created to balance the needs of local communities with the sustainable use of the resources on which they depend (Mittermeier *et al.*, 2005). Therefore, while altered landscapes might entail partial or full conservation of different components of biodiversity (Redford & Richter, 1999), their inadvertent inclusion as a conservation effort might severely mislead our estimates of preserved natural landscapes in the country – as this category is the most frequent amongst conservation units (Fig. 5; Table 1).

Finally, sampling in Brazil is still insufficient to delineate biodiversity patterns satisfactorily. According to Sobral & Stehmann (2009), the average sampling density in 2009 for Brazil was 0.59 records/km². Our dataset, after filtering for spatial and taxonomical incongruences, suggests an even smaller average sampling density of 0.13 records/km². Shepherd (2003) indicated that values around 1.0 would be sufficient to generally inform species richness in an area, although still not enough to comprehend the entire biological diversity of a determined area. The author also suggests values under 0.5 to be widely insufficient to comprehend biological diversity and estimate species richness.

Our work represents another contribution toward mapping areas that could be highlighted for future work and widen conservation and sampling, in addition to other valuable contributions published in the last few years (*e.g.*, Oliveira *et al.*, 2016; Oliveira *et al.*, 2019). Still, we emphasize that important and complex patterns exist at different scales, as has been highlighted in the literature mentioned above (*e.g.*, Amaral *et al.*, 2017; Brown *et al.*, 2020; Cantidio & Souza, 2019; Colli-Silva *et al.*, 2019; Luize *et al.*, 2024; Moro *et al.*, 2024; Ratter *et al.*, 2003; Silva-Souza & Souza, 2020).



Given the above, we conclude that recognized bioregions are heterogeneously sampled, conserved, and modified: While the Paraná dominion presents a long history of anthropization and the highest coverage of anthropogenic landscapes in the country, it also presents the highest sampling density and the lowest density of fire spots. The Chacoan dominion and the Brazilian subregion, on the other hand, have a smaller proportion of area anthropized but the greatest concentration of fire spots, indicating a current threat instead of historical destruction. Additionally, these areas hold more heterogeneity (climatically and regarding floristic regions with unique species compositions) while still encompassing smaller sampling.

Giulietti *et al.* (2005) recommended that future work focus on expanding herbarium collections and identifying and sampling hotspots in regional projects involving multiple institutions and specialists. Twenty years later, the same recommendation holds true: maintaining and growing herbarium collections is still necessary to comprehend biodiversity, as well as large projects involving diverse teams with different specialists. On the other hand, we could redirect our sampling priorities to not only cover areas already perceived as species-rich and endemic but also focus on unknown regions (like the Brazilian subregion) to avoid losing areas of underestimated or unknown value (Ahrends, Burgess, *et al.*, 2011).

Acknowledgments

The authors thank the Brazilian curators and taxonomists involved in the Brazilian Flora and Funga project and the data digitization of Brazilian specimens, both initiatives that made this project possible. LL thanks Silvio Nihei and Lucas Denadai for their valuable contributions to the preliminary manuscript and the two anonymous manuscript reviewers for their critical comments, which greatly improved the final manuscript. LL also thanks CAPES for the funding (Proc. 88887.910185/2023-00) and PTS thanks CNPq (Proc. PQ313151/2023-7) and FAPESP (Proc. 2023/12044-4) for research grants.

Author's Contributions

Luísa Lucresia was responsible for the conceptualization, data curation, formal analysis and writing the original draft. PTS and TRM were both responsible for supervision and reviewing and editing the manuscript.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Supplementary Material

The following online material is available for this article:

Figure S1. InfoMap Bioregions output of primary analysis (minimum of 100 data points per cell).

Figure S2. InfoMap Bioregions output of secondary analysis (minimum of 10 data points per cell).

Figure S3. GDM model fit and fitted I-splines for predictors associated with species turnover.

Table S1. Species occurrences in the delineated subregions and provinces.

Table S2. GDM Modelling Summary.

References

- Ahrends A, Burgess ND, Gereau RE. 2011. Funding begets biodiversity. *Diversity and Distributions* 17: 191-200. doi: 10.1111/j.1472-4642.2010.00737.x.
- Ahrends A, Rahbek C, Bulling MT. 2011. Conservation and the botanist effect. *Biological Conservation* 144: 131-140. doi: 10.1016/j.biocon.2010.08.008.
- Arima EY, Simmons CS, Walker RT, Cochrane MA. 2007. Fire in the Brazilian Amazon: A Spatially Explicit Model for Policy Impact Analysis*. *Journal of Regional Science* 47: 541-567. doi: 10.1111/j.1467-9787.2007.00519.x.
- Amaral AG, Munhoz CBR, Walter BMT, Aguirre-Gutiérrez J, Raes N. 2017. Richness pattern and phytogeography of the Cerrado herb-shrub flora and implications for conservation. *Journal of Vegetation Science* 28: 848-858. doi: 10.1111/jvs.12541.
- Assis LF, Ferreira KR, Vinhas L *et al.* 2019. TerraBrasilis: A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping. *ISPRS International Journal of Geo-Information* 8: 513. doi: 10.3390/ijgi8110513.
- Azevedo L, Zappi DC, De Oliveira DMG *et al.* 2024. On the rocks: Biogeography and floristic identity of rocky ecosystems in eastern South America. *Journal of Systematics and Evolution* 62: jse.13052. doi: 10.1111/jse.13052.
- Barbosa-Silva RG, Andrino CO, Azevedo L *et al.* 2022. A wide range of South American inselberg floras reveal cohesive biome patterns. *Frontiers in Plant Science* 13: 928577. doi: 10.3389/fpls.2022.928577.
- Bergamin RS, Bastazini VAG, Vélez-Martin E, Debastiani V, Zanini KJ, Loyola R, Müller SC. 2017. Linking beta diversity patterns to protected areas: Lessons from the Brazilian Atlantic Rainforest. *Biodiversity and Conservation* 26: 1557-1568. doi: 10.1007/s10531-017-1315-y.



- Beuchle R, Grecchi RC, Shimabukuro YE *et al.* 2015. Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Applied Geography* 58: 116-127. doi: 10.1016/j.apgeog.2015.01.017.
- Brown JL, Paz A, Reginato M. 2020. Seeing the forest through many trees: Multi-taxon patterns of phylogenetic diversity in the Atlantic Forest hotspot. *Diversity and Distributions* 26: 1160-1176. doi: 10.1111/ddi.13116.
- Cabrera AL, Willink A. 1973. *Biogeografía de América Latina*. Washington, Secretaria General de la Organización de los Estados Americanos.
- Cantidio LS, Souza AF. 2019. Aridity, soil and biome stability influence plant ecoregions in the Atlantic Forest, a biodiversity hotspot in South America. *Ecography* 42: 1887-1898. doi: 10.1111/ecog.04564.
- Cardoso D, Särkinen T, Alexander S *et al.* 2017. Amazon plant diversity revealed by a taxonomically verified species list. *Proceedings of the National Academy of Sciences* 114: 10695-10700. doi: 10.1073/pnas.1706756114.
- Castro Pena JC, Goulart F, Wilson Fernandes G *et al.* 2017. Impacts of mining activities on the potential geographic distribution of eastern Brazil mountaintop endemic species. *Perspectives in Ecology and Conservation* 15: 172-178. doi: 10.1016/j.pecon.2017.07.005.
- Cole MM. 1960. Cerrado, Caatinga and Pantanal: The Distribution and Origin of the Savanna Vegetation of Brazil. *The Geographical Journal* 126: 168-179. doi: 10.2307/1793957.
- Colli-Silva M, Vasconcelos TNC, Pirani JR. 2019. Outstanding plant endemism levels strongly support the recognition of campo rupestre provinces in mountaintops of eastern South America. *Journal of Biogeography* 46: 1723-1733. doi: 10.1111/jbi.13585.
- Costa MAD, Rios FJ. 2022. The gold mining industry in Brazil: A historical overview. *Ore Geology Reviews* 148: 105005. doi: 10.1016/j.oregeorev.2022.105005.
- Cupertino-Eisenlohr MA, Oliveira-Filho AT, Simon MF. 2021. Patterns of variation in tree composition and richness in Neotropical Non-Flooded Evergreen Forests. *Applied Vegetation Science* 24: e12522. doi: 10.1111/avsc.12522.
- Derby OA. 1906. The Serra do Espinhaço, Brazil. *The Journal of Geology* 14: 374-401.
- De Candolle AP. 1820. *Essai élémentaire de géographie botanique*. In: *Dictionnaire des Sciences Naturelles*. Paris, F.G. Levrault. vol. XVIII, p. 359-422.
- Dexter KG, Lavin M, Torke BM *et al.* 2017. Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of the National Academy of Sciences* 114: 2645-2650. doi: 10.1073/pnas.1613655114.
- Dinerstein E, Olson D, Joshi A *et al.* 2017. An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* 67: 534-545. doi: 10.1093/biosci/bix014.
- Dryflor, Banda-R K, Delgado-Salinas A *et al.* 2016. Plant diversity patterns in neotropical dry forests and their conservation implications. *Science (New York, N.Y.)* 353: 1383-1387. doi: 10.1126/science.aaf5080.
- Ducke A, Black GA. 1953. Phytogeographical notes on the Brazilian Amazon. *Anais Da Academia Brasileira de Ciências* 25: 1-46.
- Ebach MC, Morrone JJ, Parenti LR, Vilorio ÁL. 2008. International Code of Area Nomenclature. *Journal of Biogeography* 35: 1153-1157. doi: 10.1111/j.1365-2699.2008.01920.x.
- Echternacht L, Trovó M, Oliveira CT, Pirani JR. 2011. Areas of endemism in the Espinhaço Range in Minas Gerais, Brazil. *Flora - Morphology, Distribution, Functional Ecology of Plants* 206: 782-791. doi: 10.1016/j.flora.2011.04.003.
- Eidler D, Guedes T, Zizka A, Rosval, M, Antonelli A. 2017. Infomap Bioregions: Interactive Mapping of Biogeographical Regions from Species Distributions. *Systematic Biology* 66: 197-204. doi: 10.1093/sysbio/syw087.
- Fearnside PM. 2005. Deforestation in Brazilian Amazonia: History, Rates, and Consequences. *Conservation Biology* 19: 680-688. doi: 10.1111/j.1523-1739.2005.00697.x.
- Fearnside PM. 2015. Highway Construction as a Force in the Destruction of the Amazon Forest. In: Van Der Ree R, Smith DJ, Grilo C (orgs.). *Handbook of Road Ecology*. New Jersey, Wiley. p. 414-424.
- Ferrier S. 2002. Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning: Where to from Here? *Systematic Biology* 51: 331-363. doi: 10.1080/10635150252899806.
- Ferrier S, Manion G, Elith J, Richardson K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13: 252-264. doi: 10.1111/j.1472-4642.2007.00341.x.
- Fick SE, Hijmans RJ. 2017. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302-4315. doi: 10.1002/joc.5086.
- Fitzpatrick M. 2022. gdm: Generalized Dissimilarity Modeling. R package version 1.5.0-9.1. <https://cran.r-project.org/web/packages/gdm/index.html>.
- Flora e Funga do Brasil. 2020. *Flora e Funga do Brasil*. <https://floradobrasil.jbrj.gov.br/>. 01 Dec. 2024.
- Fonseca CR, Ventocinque EM. 2018. Biodiversity conservation gaps in Brazil: A role for systematic conservation planning. *Perspectives in Ecology and Conservation* 16: 61-67. doi: 10.1016/j.pecon.2018.03.001.
- Fortes EA, Landis JB, Steege HT, Specht CD, Doyle JJ, Mansano VDF. 2025. Nuclear phylogenomics of Eperua (Leguminosae) highlights the role of habitat and morphological lability in dispersal and diversification across Amazonia and in the Caatinga-Cerrado ecotone. *Molecular Phylogenetics and Evolution* 202: 108236. doi: 10.1016/j.ympev.2024.108236.
- Garnelo L, Parente RCP, Puchiarelli MLR, Correia PC, Torres MV, Herkrath FJ. 2020. Barriers to access and organization of primary health care services for rural riverside populations in the Amazon. *International Journal for Equity in Health* 19: 54. doi: 10.1186/s12939-020-01171-x.
- Garnett ST, Burgess ND, Fa JE *et al.* 2018. A spatial overview of the global importance of Indigenous lands for conservation. *Nature Sustainability* 1: 369-374. doi: 10.1038/s41893-018-0100-6.
- Giulietti AM, Harley RM, Queiroz LP, Wanderley MDGL, Van Den Berg C. 2005. Biodiversity and Conservation of Plants in Brazil. *Conservation Biology* 19: 632-639. doi: 10.1111/j.1523-1739.2005.00704.x.
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202: 18-27. doi: 10.1016/j.rse.2017.06.031.
- Harley RM. 1995. *Introdução* (Zappi DC, Trad.). In: Stannard BL (ed.). *Flora of the Pico das Almas, Chapada Diamantina – Bahia, Brazil*. Royal Botanical Gardens. p. 43-77.
- Harris MB, Tomas W, Mourão G, Da Silva CJ, Guimarães E, Sonoda F, Fachim E. 2005. Safeguarding the Pantanal Wetlands: Threats and Conservation Initiatives. *Conservation Biology* 19: 714-720. doi: 10.1111/j.1523-1739.2005.00708.x.
- Hausdorf B, Hennig C. 2003. Biotic Element Analysis in Biogeography. *Systematic Biology* 52: 717-723. doi: 10.1080/10635150390235584.
- Higgins MA, Ruokolainen K, Tuomisto H *et al.* 2011. Geological control of floristic composition in Amazonian forests. *Journal of Biogeography* 38: 2136-2149. doi: 10.1111/j.1365-2699.2011.02585.x.
- Hopkins MJG. 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. *Journal of Biogeography* 34: 1400-1411. doi: 10.1111/j.1365-2699.2007.01737.x.
- Hortal J, De Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ. 2015. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 46: 523-549. doi: 10.1146/annurev-ecolsys-112414-054400.

- Hughes A., Orr MC, Ma K *et al.* 2021. Sampling biases shape our view of the natural world. *Ecography* 44: 1259-1269. doi: 10.1111/ecog.05926.
- IBGE. 2019. Biomas e sistema costeiro-marinho do Brasil [Mapa]. <https://www.ibge.gov.br/geociencias/informacoes-ambientais/vegetacao/15842-biomas.html>. 26 Nov. 2022.
- INCT. 2024. Sobre o INCT-HVFF. INCT-HVFF. <https://incthvff.wixsite.com/inct-hvff>. 01 Dec. 2024.
- INPE. 2024. Terrabrasilis. <https://terrabrasis.dpi.inpe.br/en/download-2/>. 01 Dec. 2024.
- Brasil. 2000. Lei No 9.985, de 18 de Julho de 2000. Institui o Sistema Nacional de Unidades de Conservação da Natureza e dá outras providências, No. 9.985. https://www.planalto.gov.br/ccivil_03/leis/19985.htm. 01 Dec. 2024.
- Karger DN, Conrad O, Böhner J *et al.* 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4: 1. doi: 10.1038/sdata.2017.122.
- Kreft H, Jetz W. 2010. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography* 37: 2029-2053. doi: 10.1111/j.1365-2699.2010.02375.x.
- Leathwick JR, Elith J, Hastie T. 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling* 199: 188-196. doi: 10.1016/j.ecolmodel.2006.05.022.
- Ledo RMD, Colli GR. 2017. The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of Biogeography* 44: 2551-2563. doi: 10.1111/jbi.13049.
- Lembi RC, Cronemberger C, Picharillo C. 2020. Urban expansion in the Atlantic Forest: Applying the Nature Futures Framework to develop a conceptual model and future scenarios. *Biota Neotropica* 20: e20190904. doi: 10.1590/1676-0611-bn-2019-0904.
- Lopes JC, Fonseca LHM, Johnson DM *et al.* 2024. Dispersal from Africa to the Neotropics was followed by multiple transitions across Neotropical biomes facilitated by frugivores. *Annals of Botany* 133: 659-676. doi: 10.1093/aob/mcad175.
- Lucresia L, Stadnik A, Campos L, Roque N. 2021. Myrtaceae floristic survey and vegetation distribution in a central portion of Chapada Diamantina, Brazil. *Phytotaxa* 498: 2. doi: 10.11646/phytotaxa.498.2.1.
- Luize BG, Tuomisto H, Ekelschot R *et al.* 2024. The biogeography of the Amazonian tree flora. *Communications Biology* 7: 1240. doi: 10.1038/s42003-024-06937-5.
- Magri RA, Luebert F, Cabral A. 2025. Historical biogeography of Vellozia (Velloziaceae) reveals range expansion in South American mountaintops after climatic cooling events and increased diversification rates after the occupation of Southern Espinhaço Province. *Botanical Journal of the Linnean Society* 207: 115-127. doi: 10.1093/botlinnean/boae072.
- MapBiomas. (2022). Coleção BETA da série anual de Mapas de Cobertura e Uso da Terra do Brasil. <https://arapyau.org.br/mapbiomas-lanca-nova-colecao-sobre-a-cobertura-e-uso-da-terra-no-brasil/>. 01 Dec. 2024.
- Martinelli G. 2007. Mountain biodiversity in Brazil. *Brazilian Journal of Botany* 30: 587-597. doi: 10.1590/S0100-84042007000400005.
- Martins SFS, Dos Santos AM, Silva CFAD, Rudke AP, Alvarado ST, Melo JLDS. 2024. The drivers of fire in the Caatinga Biome in Brazil. *Forest Ecology and Management* 572: 122260. doi: 10.1016/j.foreco.2024.122260.
- Miranda HS, Sato MN, Neto WN, Aires FS. 2009. Fires in the cerrado, the Brazilian savanna. In: Cochrane MA. *Tropical Fire Ecology*. Heidelberg, Springer Berlin Heidelberg. p. 427-450.
- Mittermeier RA, da Fonseca GAB, Rylands AB, Brandon K. 2005. A Brief History of Biodiversity Conservation in Brazil. *Conservation Biology* 19: 601-607.
- Monteiro L, Machado N, Martins E *et al.* 2018. Conservation priorities for the threatened flora of mountaintop grasslands in Brazil. *Flora* 238: 234-243. doi: 10.1016/j.flora.2017.03.007.
- Montgomery DC, Peck EA. 1992. *Introduction to Linear Regression Analysis*. New Jersey, Wiley.
- Morellato LPC, Haddad CFB. 2000. *Introduction: The Brazilian Atlantic Forest*. *Biotropica* 32: 786-792. doi: 10.1111/j.1744-7429.2000.tb00618.x.
- Moro MF, Amorim VO, Queiroz LP *et al.* 2024. Biogeographical Districts of the Caatinga Dominion: A Proposal Based on Geomorphology and Endemism. *The Botanical Review* 90: 376-429. doi: 10.1007/s12229-024-09304-5.
- Morrone JJ. 2009. *Evolutionary Biogeography: An Integrative Approach with Case Studies*. New York, Columbia University Press.
- Morrone JJ. 2014. Biogeographical regionalisation of the Neotropical region. *Zootaxa* 3782: 1-110. doi: 10.11646/zootaxa.3782.1.1.
- Morrone JJ. 2017. *Neotropical Biogeography: Regionalization and Evolution*. Boca Raton, CRC Press.
- Morrone JJ. 2018. The spectre of biogeographical regionalization. *Journal of Biogeography* 45: 282-288. doi: 10.1111/jbi.13135.
- Morrone JJ, Escalante T, Rodríguez-Tapia G, Carmona A, Arana M, Mercado-Gómez JD. 2022. Biogeographic regionalization of the Neotropical region: New map and shapefile. *Anais Da Academia Brasileira de Ciências* 94: e20211167. doi: 10.1590/0001-376520220211167.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 6772. doi: 10.1038/35002501.
- Naimi B. 2023. usdm: Uncertainty Analysis for Species Distribution Models. R package version 2.1-7. <https://CRAN.R-project.org/package=usdm>.
- Nelson G. 1978. From Candolle to croizat: Comments on the history of biogeography. *Journal of the History of Biology* 11: 269-305. doi: 10.1007/BF00389302.
- Oliveira U, Paglia, AP, Brescovit, AD *et al.* 2016. The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Diversity and Distributions* 22: 1232-1244. doi: 10.1111/ddi.12489.
- Oliveira U, Soares-Filho BS, Santos AJ *et al.* 2019. Modelling Highly Biodiverse Areas in Brazil. *Scientific Reports* 9: 6355. doi: 10.1038/s41598-019-42881-9.
- Olson DM, Dinerstein E, Wikramanayake ED *et al.* 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, 51: 933. doi: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.
- Ometto JP, Aguiar APD, Martinelli LA. 2011. Amazon deforestation in Brazil: Effects, drivers and challenges. *Carbon Management* 2: 575-585. doi: 10.4155/cmt.11.48.
- Ondo I, Dhanjal-Adams KL, Pironon S *et al.* 2024. Plant diversity darkspots for global collection priorities. *New Phytologist* 244: 719-733. doi: 10.1111/nph.20024.
- Peres EA, Pinto-da-Rocha R, Lohmann LG, Michelangeli FA, Miyaki CY, Carnaval AC. 2020. Patterns of Species and Lineage Diversity in the Atlantic Rainforest of Brazil. In: Rull V, Carnaval AC (orgs.). *Neotropical Diversification: Patterns and Processes*. Switzerland, Springer International Publishing. p. 415-447.
- Pivello VR. 2011. The Use of Fire in the Cerrado and Amazonian Rainforests of Brazil: Past and Present. *Fire Ecology* 7: 24-39. doi: 10.4996/fireecology.0701024.
- Procheş Ş, Ramdhani S. 2012. The World's Zoogeographical Regions Confirmed by Cross-Taxon Analyses. *BioScience* 62: 260-270. doi: 10.1525/bio.2012.62.3.7.



- Queiroz LP, França F, Giuliatti AM (2005). Caatinga. In: Juncá FA, Funch LS, Rocha W (orgs.). Biodiversidade e conservação da Chapada Diamantina. Brasília, Ministério do Meio Ambiente. p. 95-120
- Prance GT. 1973. Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon Basin, based on evidence from distribution patterns in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. *Acta Amazonica* 3: 5-26. doi: 10.1590/1809-43921973033005.
- Ratter JA, Bridgewater S, Ribeiro JF. 2003. Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation Iii: Comparison Of The Woody Vegetation Of 376 Areas. *Edinburgh Journal of Botany* 60: 57-109. doi: 10.1017/S0960428603000064.
- Ratter JA, Ribeiro JF, Bridgewater S. 1997. The Brazilian Cerrado Vegetation and Threats to its Biodiversity. *Annals of Botany* 80: 223-230. doi: 10.1006/anbo.1997.0469.
- Redford KH, Richter BD. 1999. Conservation of Biodiversity in a World of Use. *Conservation Biology* 13: 1246-1256. doi: 10.1046/j.1523-1739.1999.97463.x.
- REFLORA. 2024. *Herbário Virtual*. Herbário Virtual REFLORA. <https://reflora.jbrj.gov.br/reflora/herbarioVirtual/>. 20 Nov. 2024.
- Reginato M, Michelangeli FA. 2020. Bioregions of Eastern Brazil, Based on Vascular Plant Occurrence Data. In: V. Rull & A. C. Carnaval (Orgs.), *Neotropical Diversification: Patterns and Processes* (p. 475-494). Springer International Publishing. https://doi.org/10.1007/978-3-030-31167-4_18
- Rezende CL, Scarano FR, Assad ED *et al.* 2018. From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation* 16: 208-214. doi: 10.1016/j.pecon.2018.10.002.
- Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FJ, Hirota MM. 2009. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation* 142: 1141-1153. doi: 10.1016/j.biocon.2009.02.021.
- Rull V, Vegas-Vilarrúbia T. 2020. The Pantepui "Lost World": Towards a Biogeographical, Ecological and Evolutionary Synthesis of a Pristine Neotropical Sky-Island Archipelago. In: Rull V, Carnaval AC (orgs.). *Neotropical Diversification: Patterns and Processes*. Cham, Springer International Publishing. p. 369-413.
- Seto KC, Güneralp B, Hutrya LR. 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences* 109: 16083-16088. doi: 10.1073/pnas.1211658109.
- Shepherd GJ. 2003. Plantas Terrestres (Avaliação do Estado do Conhecimento da Diversidade Biológica do Brasil, p. 60). COBIO/MMA - GTB/CNPq - NEPAM/UNICAMP.
- Silva CFA, Andrade MO, Santos AM, Melo SN. 2023. Road network and deforestation of indigenous lands in the Brazilian Amazon. *Transportation Research Part D: Transport and Environment* 119: 103735. doi: 10.1016/j.trd.2023.103735.
- Silva JMC, Bates JM. 2002. Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot. *BioScience* 52: 225. doi: 10.1641/0006-3568.
- Silva JMC, Tabarelli M. 2000. Tree species impoverishment and the future flora of the Atlantic forest of northeast Brazil. *Nature* 404: 72-74. doi: 10.1038/35003563.
- Silva Junior CHL, Pessoa ACM, Carvalho NS, Reis JBC, Anderson LO, Aragão LEOC. 2020. The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. *Nature Ecology & Evolution* 5: 144-145. doi: 10.1038/s41559-020-01368-x.
- Silva-Souza KJP, Souza AF. 2020. Woody plant subregions of the Amazon forest. *Journal of Ecology* 108: 2321-2335. doi: 10.1111/1365-2745.13406.
- Sobral M, Stehmann JR. 2009. An analysis of new angiosperm species discoveries in Brazil (1990-2006). *TAXON* 58: 227-232. doi: 10.1002/tax.581021.
- Sousa-Baena MS, Garcia LC, Peterson AT. 2014. Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Diversity and Distributions* 20: 369-381. doi: 10.1111/ddi.12136.
- Souza CM, Shimbo JZ, Rosa MR *et al.* 2020. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing* 12: 2735. doi: 10.3390/rs12172735.
- Tabarelli M, Silva JMC, Gascon C. 2004. Forest fragmentation, synergisms and the impoverishment of neotropical forests. *Biodiversity and Conservation* 13:1419-1425. doi: 10.1023/B:BIOC.0000019398.36045.1b.
- Ter Steege H, Prado PI, Lima, RAFD *et al.* 2020. Biased-corrected richness estimates for the Amazonian tree flora. *Scientific Reports* 10: 10130. doi: 10.1038/s41598-020-66686-3.
- Viana PL, Mota NFDO, Gil ADSB *et al.* 2016. Flora das cangas da Serra dos Carajás, Pará, Brasil: História, área de estudos e metodologia. *Rodriguésia* 67: 1107-1124. doi: 10.1590/2175-7860201667501.
- Vilhena DA, Antonelli A. 2015. A network approach for identifying and delimiting biogeographical regions. *Nature Communications* 6: 6848. doi: 10.1038/ncomms7848.
- Viloria Á. 2005. Las mariposas (Lepidoptera: Papilionoidea) y la regionalización biogeográfica de Venezuela. In: Llorente-Bousquet J, Morrone JJ. *Regionalización biogeográfica en Iberoamérica y tópicos afines*. México, Las Prensas de Ciencias/UNAM. p. 441-459.
- Wallace AR. 1876. The geographical distribution of animals. With a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface. New York, Harper and Brothers.
- Whittaker RJ, Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ. 2005. Conservation Biogeography: Assessment and prospect. *Diversity and Distributions* 11: 3-23. doi: 10.1111/j.1366-9516.2005.00143.x.
- Yang W, Ma K, Kref, H. 2013. Geographical sampling bias in a large distributional database and its effects on species richness-environment models. *Journal of Biogeography* 40: 1415-1426. doi: 10.1111/jbi.12108.
- Zizka A, Silvestro D, Andermann T *et al.* 2019. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* 10: 744-751. doi: 10.1111/2041-210X.13152.
- Zuur AF, Ieno EN, Elphick, CS. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1: 3-14. doi: 10.1111/j.2041-210X.2009.00001.x.