



Defining centers of biodiversity, spatial patterns, and drivers of species distributions in *Coccoloba* (Polygonaceae)

Joilson Viana Alves^{1,*} , Daniel Mark Koenemann² , Rafael Batista Louzada¹ 

¹Universidade Federal de Pernambuco, Recife, PE, Brazil.

²Catholic International University, Charles Town, WV, United States.

*Corresponding author: j.vianaalves@hotmail.com

ABSTRACT

Coccoloba is the most diverse woody neotropical genus of the Polygonaceae, with four widely recognized biodiversity centers: the Atlantic Forest, the Amazon Rainforest, Central America, and the Caribbean Islands. However, the boundaries of these centers have never been tested. To address this, we compiled 6,989 records of 146 taxa of *Coccoloba*. The distribution pattern was analyzed using the *K-means* clustering; Centers of endemism were identified by PAE and Endemicity Analysis; and centers of richness and diversity were analyzed in Diva-Gis. The GDM model was used to understand the drivers of species distributions. Seven distribution patterns and eight areas of endemism were recovered. Areas of high richness and diversity are found in Central America, the Caribbean Islands and the Atlantic Forest. The Amazon Rainforest has no area of high species richness, although it has a small area with high diversity. The dissimilarity in *Coccoloba* species composition is strongly related to geographic distance, solar radiation in April, precipitation of wettest quarter, and mean temperature of warmest quarter drivers. We conclude that the geographic diversity centers previously suggested in the literature are too simplistic to accurately describe centers of biodiversity for *Coccoloba*, rather, they correspond to the main distribution patterns of *Coccoloba* species.

Keywords: Amazon Rainforest, Atlantic Forest, Biogeography, Central America, Caribbean Islands, Centers of Endemism, Eriogonoideae.

Introduction

Centers of biodiversity are geographical areas with elevated species diversity, richness, and endemism (Smith-Ramírez, 2004; Carpenter & Springer, 2005; Servonnat *et al.*, 2019; Lima *et al.*, 2020). The neotropical region is recognized as a major center of biodiversity (Antonelli & Sanmartín, 2011; Antonelli, 2022) and extends from

central Mexico to southern Brazil (Morrone *et al.*, 2022). It comprises heterogeneous areas noted for their great plant diversity, including moist forests (Oliveira-Filho *et al.*, 2021) and seasonally dry forests (Fernandes *et al.*, 2020; Maciel & Martins, 2021). Based on endemic species and ecoregions composed of climatic, geological, and biotic criteria, Morrone *et al.* (2022) propose 57 biogeographic provinces in the neotropical region. Studies focused on specific biomes,

Received April 30, 2024; Accepted October 12, 2024

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

How to cite:

Alves JV, Koenemann DM, Louzada RB. 2024. Defining centers of biodiversity, spatial patterns, and drivers of species distributions in *Coccoloba* (Polygonaceae). *Acta Botanica Brasilica* 38: e20240121. doi: [10.1590/1677-941X-ABB-2024-0121](https://doi.org/10.1590/1677-941X-ABB-2024-0121)



have revealed that many of these biogeographic provinces are considered centers of biodiversity for several groups of plants, for example the Chapada Diamantina, and the Southern of Espinhaço Range provinces for Asclepiadoideae (Bitencourt & Rapini, 2013) and Convolvulaceae (Alves & Buril, 2022), and the Atlantic province for ferns and lycophytes (Souza *et al.*, 2021). Recently, Rodrigues & Duarte (2024) used the genus *Myrcia* DC. to describe the spatial patterns of diversity and biogeographical history in the neotropical region, and highlight that the high species richness for the genus is located in the Atlantic Forest and Cerrado. However, native genera and widely distributed taxa are still scarcely used to investigate the centers of diversity throughout the neotropical region.

Antonelli (2022) highlights that centers of biodiversity in the Neotropics have suffered a recent and rapid deterioration due to anthropogenic factors. The delimitation of centers of biodiversity is necessary to identify priority areas for conservation and ensure the preservation of biodiversity (Dryflor *et al.*, 2016; Alves & Buril, 2022). Identifying priority areas for conservation is essential for protecting biodiversity, using resources effectively, maintaining ecosystem services, preventing extinctions, and promoting ecological sustainability (Pulido-Chadid *et al.*, 2023).

Coccoloba P.Browne is a monophyletic group (Burke *et al.*, 2010; Koenemann & Burke 2020), and the most diverse genus of the woody neotropical Polygonaceae (Burke *et al.*, 2010), with about 150 species distributed mainly in moist forests (Howard 1960; 1961; 1992; Brandbyge 1990; Melo 2004; Melo *et al.*, 2019). Koenemann & Burke (2020) suggested that the major *Coccoloba* clades are organized biogeographically, and correspond to major diversity centers in the genus (i.e. Atlantic Forest, Amazon Rainforest, Central America, and the Caribbean Islands). Despite many authors recognizing these four centers of diversity and endemism for *Coccoloba* (Brandbyge, 1990; Noa, 2013; Koenemann & Burke, 2020; Alves *et al.*, 2021; Alves & Buril, 2022; Ancona *et al.*, 2023), its geographic boundaries have never been defined or tested in a formal biogeographical analysis. To do so, we wanted to use a dataset to test these centers using objective methods. Alves *et al.* (2021) pointed out that the distribution of *Coccoloba* species in South America is influenced by the climatic conditions of neotropical moist forests, which offer suitable conditions for the occurrence of the species. These observations are supported by the large number of species present in the Amazon Forest and the Atlantic Forest. However, it is still unknown which drivers promote suitable conditions and the species composition of *Coccoloba*. On the other hand, compared to moist forests, seasonally dry forests (Cerrado and Caatinga biomes), have few *Coccoloba* species occurrences, generally with wide distribution (Melo, 2004; Alves *et al.*, 2021), suggesting that

these biomes act as ecological barriers for South America *Coccoloba* species.

Because of its distribution, *Coccoloba* is a good model to generally study larger patterns of biodiversity in neotropical forests. Here we address the following questions: (1) How can we best describe the distribution of *Coccoloba* species in neotropical forests? (2) What are the geographic boundaries of the centers of endemism, richness, and diversity in *Coccoloba*? (3) What are the main drivers and their effect on the distribution of *Coccoloba* species? We hope that the results indicate distribution patterns related to the four centers of diversity articulated in the *Coccoloba* in the literature. However, we expect that the geographic boundaries of the centers of endemism, richness, and diversity of the genus will correspond to smaller areas within the previously described centers. We expect this because the species composition of *Coccoloba* is not homogeneously distributed over the entire neotropical region (Howard, 1960; 1961; Brandbyge, 1990; Melo, 2004; Alves *et al.*, 2021).

Material and methods

Species occurrence dataset

Geographical coordinates for species occurrences were compiled from the Reflora Virtual Herbarium (<http://floradobrasil.jbrj.gov.br/>), SpeciesLink (<http://www.splink.org.br/>) and Global Biodiversity Information Facility - GBIF (<https://www.gbif.org/>). Vouchers without collector numbers, duplicate records, and species with one record were discarded. Non-georeferenced samples, when possible, were adjusted to the location indicated on the specimens or to the respective centroid of the municipality. When no such adjustment was possible, these samples were discarded. When necessary and possible, we confirmed determinations and made identifications for unidentified specimens. Dubious identifications and determinations not confirmed by experts in *Coccoloba* or Polygonaceae were discarded. Using the parameters described above we assembled a dataset with 6,989 records and 146 taxa (Table S1).

The *K-means* function available in the Stats package from R v.4.3.1 (R Core Team, 2023), was used to build an agglomerative cluster based on a '.csv' file containing the geographic coordinates of the species. We used the *fviz_nbclust* function, available in the Factoextra package (Kassambara & Mundt, 2020) with the wss method, to determine the best number of clusters. This function returned four clusters, each one of them, at first, representing one distribution pattern. The *K-means* function has limitations

in recognizing groups that represent species with wide continuous and disjunct distribution, so we checked any species present in more than one cluster and designated a new pattern (wide continuous or disjunct pattern), when necessary. Species with a known distribution in a geographic area, which had a greater number of samples in the cluster corresponding to that area (e.g., Central America), but with few samples emerging in another cluster corresponding to another area where there is no record of this species (e.g., Caribbean Islands), were considered in the cluster with the largest sampling (see Table S2). The analysis described above was performed in the RStudio (Posit team, 2024).

Areas of endemism

Areas of endemism were identified using Parsimony Analysis of Endemism (PAE) and Endemicity Analysis (EA). PAE analyses consider only those grid cells with two or more records of endemic species (Morrone, 1994; Morrone, 2014). Species with a single record or records beyond the geographic boundaries of the neotropical region were discarded. A presence and absence matrix was built in grid cells of $3^{\circ} \times 3^{\circ}$ latitude-longitude degrees and converted to a '.tnt' file. In this analysis, species were used as characters and areas as species, additionally, an ancestor area, with all species absent was inserted (Morrone, 2014). Analyses were conducted in WinClada/Nona (Goloboff, 1997) with the following parameters: heuristic search; 10,000 maximum trees to keep (hold); 1,000 number of replication (mult*N); 10 starting trees per replicate (hold/) and search strategy as multiple TBR + TBR (mult*max*). This analysis returns a cladogram, where the unsupported nodes were collapsed and the final tree was obtained using a strict consensus. In this final tree, the clades supported by two or more endemic species were considered endemic areas (Morrone, 2014). The consistency index (Ci) was used to indicate which tree was more parsimonious, while the retention index (Ri) measured the fraction of potential synapomorphies retained as synapomorphies on the tree (Mickevich & Lipscomb, 1991).

Endemicity analysis (EA) was performed in the NDM/VNDM programs. Endemicity indices (EI) were assigned based on an optimality criterion to evaluate each area of endemism obtained (Szumik *et al.*, 2002; Szumik & Goloboff, 2004). Distinct from the PAE, the endemicity analysis provides a score of endemicity for each area, which indicates the contributions to endemism in the area (Szumik & Goloboff, 2004). Szumik *et al.* (2002) indicated that those areas with the highest scores of endemicity should be chosen as areas of endemism. For the endemicity analysis, we used the same taxa analyzed in the PAE, in grid cells of $3^{\circ} \times 3^{\circ}$ latitude-longitude degrees. The PAE matrix was converted to the '.xyd' format following Santos & Fuhlendorf (2018). A

heuristic search was carried out, keeping overlapping subsets with 50% unique species. A strict consensus considering a 50% minimum similarity of endemic species was applied.

A consensus map between the areas of endemism recovered by PAE and EA was constructed using the QGIS program. This same program was also used to verify and identify the overlap of areas of endemism between the analyses through the *union* function. We employed the following criteria to establish the consensus map: (1) overlap of the same areas recovered by both analyses; and (2) overlapping of a larger area over two or more smaller areas. In this last situation, we considered the composition of endemic species, and the geographical, and geological features of the area under consideration to establish an area of consensus (Escalante, 2011; Morrone, 2014). The final consensus of the areas of endemism was adjusted to the neotropical biogeographic provinces proposed by Morrone *et al.* (2022).

Areas of species richness and diversity

We used Jackknife2, a non-parametric estimator, for richness in the entire dataset, with 100 randomizations in 100% of the sample. The Shannon-Wiener index, which considers the number of categories (number of different species) and evenness of spatial distribution of individual categories (Dusek & Popelková, 2017), was used to obtain the diversity of *Coccoloba* species in the Neotropics. We used 146 categories to perform the diversity analyses (the same used in the richness analyses). Both richness and diversity analyses were conducted in grid cells of $3^{\circ} \times 3^{\circ}$ in the DIVA-GIS program (Hijmans *et al.*, 2001). The results obtained were then converted to shapefile format and exported to the QGIS program for editing and visualization.

Assessing the drivers of species distribution

To investigate the main drivers that influence the distribution patterns of *Coccoloba* species in the Neotropics, we obtained a set of 39 explanatory variables in high resolution (30 arc-second, around 1 km²), including elevation, seven soil quality variables, 19 climatic variables (representing the values of precipitation and temperature for the years 1970 to 2000), and 12 solar radiation variables (one for each month of the year). Elevation, climate variables, and solar radiation were obtained from WorldClim (Fick & Hijmans, 2017); soil quality variables were obtained from SoilGrids — global gridded soil information (<https://soilgrids.org/>).

Due to the different origins and resolutions between the soil layers and the other explanatory layers, we applied the *resample* function of the Raster package (Hijmans, 2023), to transfer the values between non-matching GeoTiff layers.



This action allowed us to use the *stack* and *extract* functions of the Raster package to access the values and prepare the data for the analysis described below.

To reduce the multicollinearity, we used the *cor* function of the Stats package (R Core Team, 2023), to calculate the Pearson correlation coefficient, and the *findCorrelation* function of the Caret package (Kuhn, 2008) to find highly correlated variables ($\geq 80\%$ of correlation) and eliminate them. Fifteen explanatory variables remained after removing highly correlated variables (Tab. 1). A non-parametric Kruskal-Wallis test was used to compare the median of the explanatory variables between the distribution patterns (significance level of 0.05). Violin plots were used to visualize and compare the distribution of data between the distribution patterns for each variable using the *ggplot* function of the ggplot2 package (Wickham, 2016).

We evaluated the drivers of *Coccoloba* species distribution through generalized dissimilarity modeling (GDM). The GDM analysis accounts for the non-linearity of community dissimilarity across environmental gradients and calculates three I-spline coefficients for each explanatory variable, where higher coefficients indicate higher rates of change of the response variable along the gradient of the explanatory variable (Ferrier et al., 2007; Mokany et al., 2022). All analyses were performed in R 4.3.1, using the packages gdm, ggplot2, ggrepel, and dplyr.

Results

Overall species distribution patterns

Seven categories were recovered representing seven spatial distribution patterns of the studied taxa. Four of these categories were recovered by the *K-means* function: the Caribbean (Fig. 1A), Central America (Fig. 1B), the Amazon (Fig. 1C), and Eastern South America (Fig. 1D). After examination, there were several species whose distributions did not fit cleanly into the groups identified by *K-means*. Instead, they fit into three basic patterns representing combinations of *K-means* clusters: a Caribbean-Central America disjunction distribution (Fig. 1E), a South American wide distribution (Fig. 1F) and a continuous Central American-Amazon distribution (Fig. 1G). According to the occurrence records used in this study, *Coccoloba* species are located mainly in the Amazon Rainforest region, where about 35% of all *Coccoloba* species occur, followed by the Caribbean Islands (32%), Central American (26%), and Eastern South America with about 22%.

Forty species follow the Caribbean pattern (Tab. 2). However, they are not distributed across all archipelagos. In this pattern, most species are concentrated mainly in Cuba, Hispaniola, and Jamaica. Examples of Caribbean pattern

Table 1. List of explanatory drivers considered after the selection of non-correlated variables using the Pearson coefficient, classified into four main categories: temperature, precipitation, soil quality, and solar radiation drivers.

Abbreviation	Distribution drivers (predictor)
Temperature	
bio02	Mean Diurnal Range (Mean of monthly qy (max temp - min temp))
bio04	Temperature Seasonality (standard deviation $\times 100$)
bio10	Mean Temperature of Warmest Quarter
Precipitation	
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio18	Precipitation of Warmest Quarter
Bio19	Precipitation of Coldest Quarter
Soil quality	
sq02	Nutrient retention capacity
sq04	Oxygen availability to roots
sq06	Calcium carbonate and gypsum
sq07	Soil texture, effective soil depth/volume, and soil phases constraining soil management (soil depth, rock outcrop, stoniness, gravel/concretions and hardpans)
Solar radiation	
srad02	Solar Radiation ($\text{kJ m}^{-2} \text{ day}^{-1}$) in February
srad04	Solar Radiation ($\text{kJ m}^{-2} \text{ day}^{-1}$) in April
srad11	Solar Radiation ($\text{kJ m}^{-2} \text{ day}^{-1}$) in November

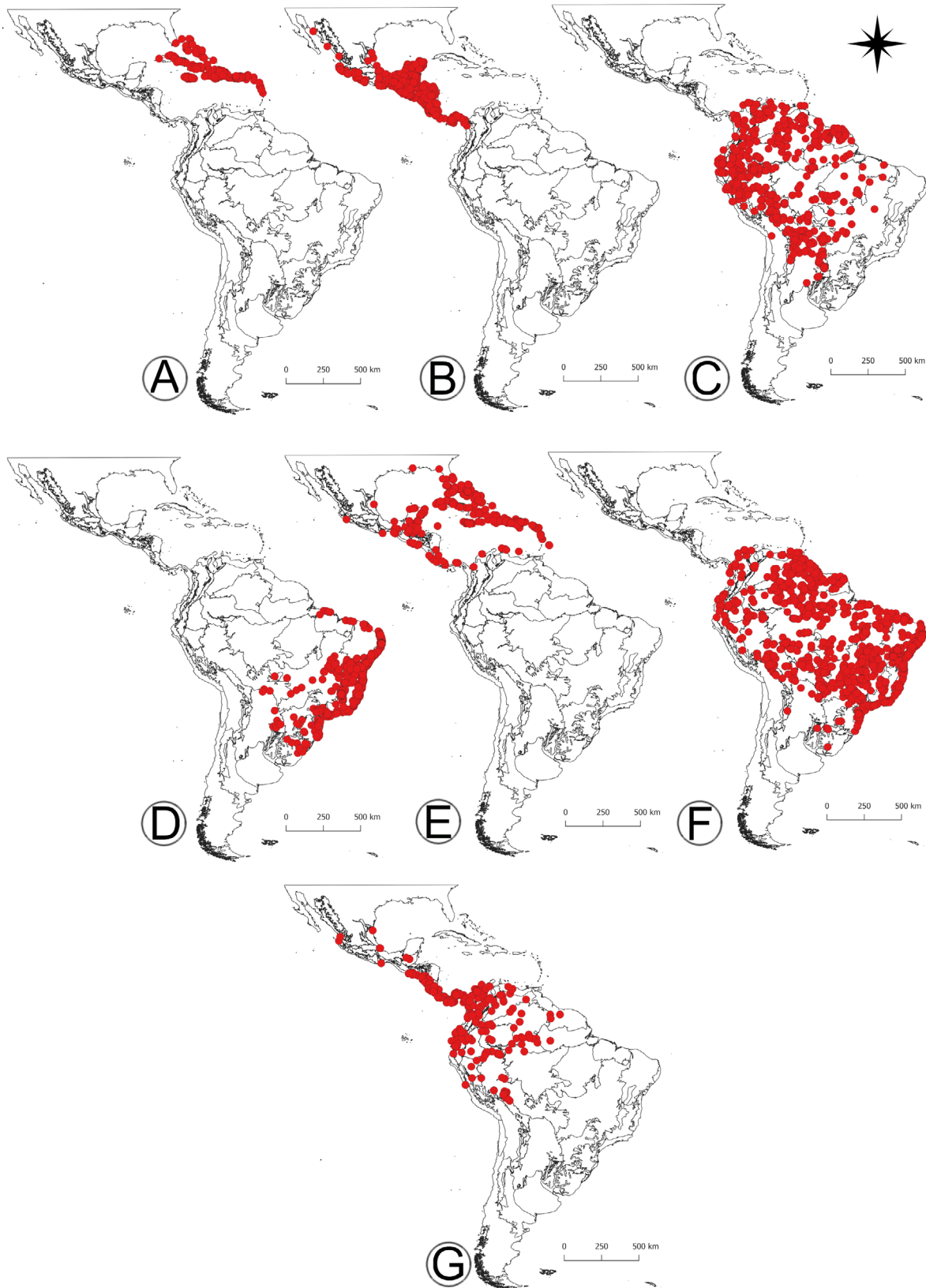


Figure 1. Distribution patterns of *Cocoloba* species as revealed by *K-means* clustering (A-D) and combinations of clusters after examining the *K-means* groups (E-G). (A) Caribbean Islands pattern. (B) Central American pattern. (C) Amazonian pattern. (D) Eastern South American pattern. (E) Caribbean-Central American disjunction pattern. (F) South American-wide distribution. (G) Central American-Amazonian pattern.



Table 2. Number of species and records of *Coccoloba* for each distribution pattern.

Patterns	Species	Records
Amazon	30	949
Caribbean Islands	40	569
Caribbean Islands-Central America	7	555
Central America	31	1,120
Central America-Amazon	5	506
East of South America	17	1,423
South America	16	1,867
Total	146	6,989

species include: *Coccoloba armata* C.Wright ex Griseb., *C. krugii* Lindau, and *C. pubescens* L. (see Table S1). Thirty-one species are categorized under the Central American pattern (Tab. 2), where we have placed species that occur only in North (e.g. *C. barbadosensis* Jacq.) or only in South Central America (e.g. *C. lasserii* Lundell), or are continuously distributed across this region (e.g. *C. acapulcensis* Standl.). Seven species showed a disjunction between Central America and the Caribbean Islands generated by the Caribbean Sea (e.g. *C. venosa* L and *C. diversifolia* Jacq.), and five species have continuous distributions spanning from South Central America to the Northern Amazon (e.g. *C. lehmannii* Lindau) (Tab. 2).

Thirty species follow the Amazonian pattern (e.g. *C. acuminata* Kunth) and 17 species follow the Eastern South America pattern (e.g. *C. laevis* Casar.). Sixteen species have a continuous distribution in South America (e.g. *C. mollis* Casar.). Species following the Amazonian pattern are concentrated in the Amazon Rainforest, including endemic and widely distributed species in this region. Species following the Eastern South America pattern are concentrated in the Atlantic Forest, including species endemic to this region, as well as some species occurring in adjacent regions (Espinhaço range and Pantanal).

Parsimony analysis of endemism (PAE)

The PAE analysis returned 40 trees. The strict consensus tree, with 495 steps, consistency index (Ci) = 26 and retention index (Ri) = 60, showed seven clades as areas of endemism (Fig. 2): Area 1, in Mesoamerica, is the largest area of endemism with 22 grid cells, included nearly all provinces in this area, except the Sierra Madre del Sur, Occidental and Oriental; Area 2, Cuba, supported by two grid cells; Area 3, Jamaica, supported by two grid cells; Area 4, with four grid cells, including the Lesser Antilles and Puerto Rico; Area 5, in the extreme north of the Amazon Rainforest, covering six grid cells and parts of Madeira, Rondônia and Yungas provinces; Area 6, the second largest area of endemism with 12 grid cells, including Chapada Diamantina and Atlantic provinces; and Area 7, supporting by two grids cells in the Espinhaço Ranges.

Endemicity analysis (EA)

Endemicity analysis revealed eight consensus areas of endemism (Fig. 3), some of which closely resemble the seven areas found in the PAE. Area 1 (score = 6.43506), was supported for 10 endemic species and corresponds to the north of Mesoamerica (Fig. 3A). Area 2, in southern Mesoamerica, (score= 4.30000) has six endemic species (Fig. 3B). Area 3 (score= 6.02555), was supported for 11 species, covers Cuba and Jamaica (Fig. 3C). Area 4 (score= 8.62179), in Hispaniola and Puerto Rico, corresponds to 12 species (Fig. 3D). Area 5 (score= 9.35972), supported by 14 species, corresponds to Jamaica, Cuba, Hispaniola, and Puerto Rico, however, this area overlaps Areas 3 and 4 (five species are shared with Area 3), and two species are shared with Area 4 (Fig. 3E, see also Tab. 3). Area 6 (score= 2.72753, supported by three species) covers the northern Amazon Rainforest and the Chaco (Fig. 3F). An overlap was observed in Area 7 (score= 2.74536, supported for four species) and Area 8 (score= 4.20600, supported for six species), which correspond to the northern part of the Atlantic Forest (Fig. 3G-H).

Consensus between PAE and EA

Seeing that PAE and EA revealed similarities in areas of endemism, we performed a consensus analysis between both analyses to delimit the endemic areas of *Coccoloba* (Fig. 4). This consensus analysis recovered eight areas of endemism: (1) Northern and (2) Southern Central America, (3) Cuba, (4) Jamaica, (5) Hispaniola, including Puerto Rico, (6) the Chaco province, including parts of Yungas and Rondônia provinces, (7) Atlantic Forest, including the Chapada Diamantina province, and (8) Southern Espinhaço province.

Areas of richness and diversity

Central America, the Caribbean Islands and the Atlantic Forest revealed grid cells with the highest species richness, each of the grid cells is estimated to contain twenty-three to twenty-nine *Coccoloba* species, based on the Jackknife2 analysis (Fig. 5A). In Central America, these areas are located in the Maya Biosphere Reserve, the National Parks

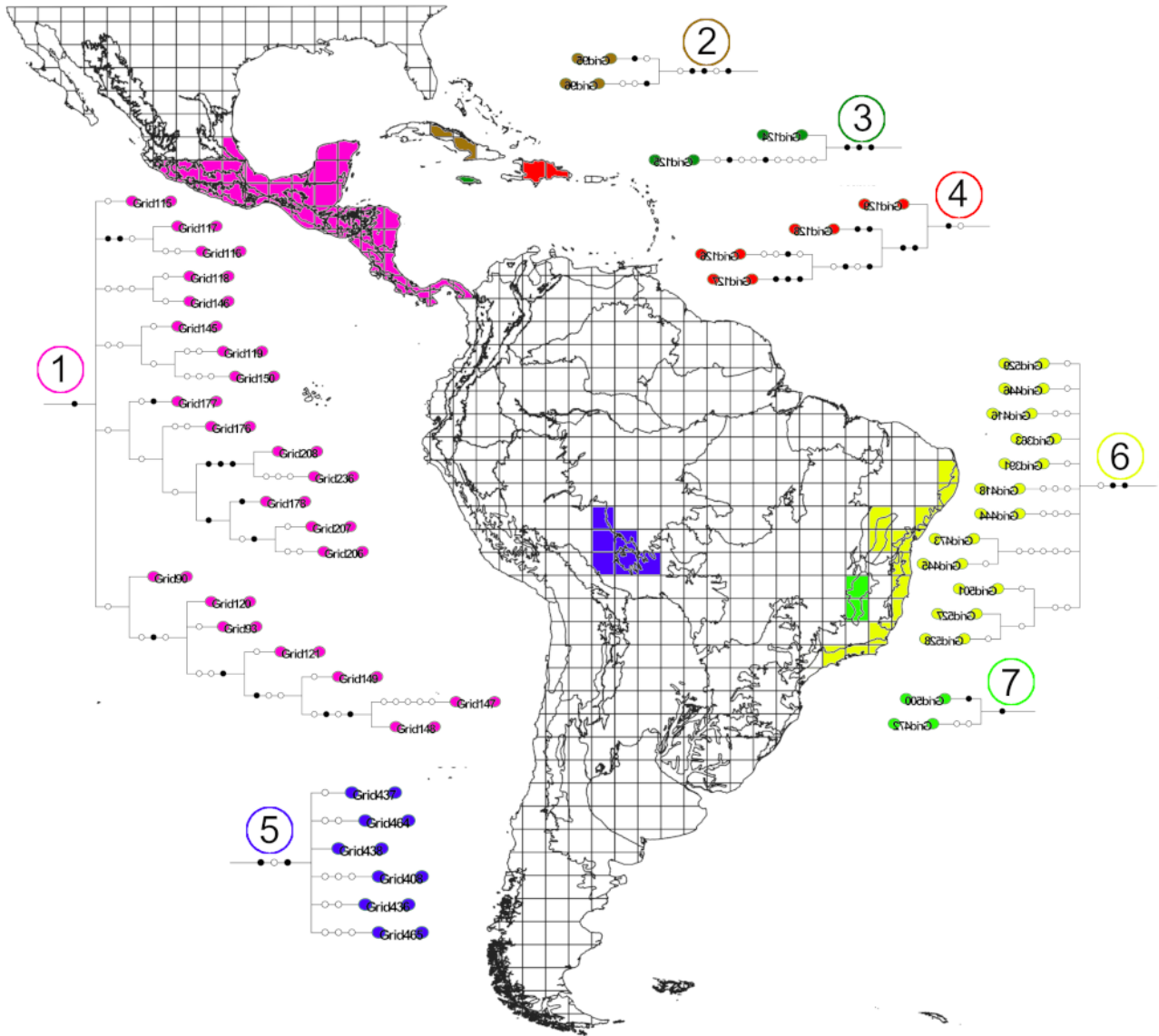


Figure 2. Results of the Parsimony Analysis of Endemism showing seven areas of endemism for *Cocoloba*: Central America (Area 1), Cuba (Area 2), Jamaica (Area 3), the Lesser Antilles (Area 4), Amazon Rainforest (Area 5), Chapada Diamantina (Area 6), and southern Espinhaço Range (Area 7).

Table 3. Summary of Endemism Analysis. AE = area of endemism. MS = maximum score of area. N = number of endemic species that support the area of endemism. Area 1 = northern Mesoamerica; Area 2 = southern Mesoamerica; Area 3 = Cuba and Jamaica; Area 4 = Hispaniola and Puerto Rico; Area 5 = Jamaica, Cuba, Hispaniola, and Puerto Rico; Area 6 = northern Amazon Rainforest and the Chaco; Areas 7-8 = Atlantic Forest, including Chapada Diamantina and southern Espinhaço Range. *species shared between Areas 3 and 5; **species shared between Areas 4 and 5; ***species shared between Areas 7 and 8.

AE	MS	N	Species and endemism index
Area 1	6.43506	10	<i>Cocoloba belizensis</i> (0.000-0.500); <i>C. chiapensis</i> (0.000-0.667); <i>C. cozumelensis</i> (0.364-1.000); <i>C. floresii</i> (0.000-0.773); <i>C. hirtella</i> (0.559-0.639); <i>C. lindeniana</i> (0.767-1.000); <i>C. lindaviana</i> (0.000-0.568); <i>C. lundellii</i> (0.567-0.786); <i>C. reflexiflora</i> (0.400-0.967); <i>C. tunii</i> (0.000-0.818)
Area 2	4.30000	6	<i>C. darienensis</i> (0.731-1.000); <i>C. johnstonii</i> (0.864-1.000); <i>C. lasseri</i> (0.727-0.865); <i>C. liportizii</i> (0.000-0.500); <i>C. manzinellensis</i> (0.894-1.000); <i>C. porphyrostachys</i> (0.000-0.405)
Area 3	6.02555	11	<i>C. armata</i> (0.545); <i>C. benitensis</i> (0.412)*; <i>C. caesia</i> (0.538)*; <i>C. cowellii</i> (0.500)*; <i>C. geniculata</i> (1.000); <i>C. microphylla</i> (0.667); <i>C. nipensis</i> (0.591)*; <i>C. northropiae</i> (0.385); <i>C. pallida</i> (0.467); <i>C. retusa</i> (0.304); <i>C. rufescens</i> (0.367)*



Table 3. Cont.

AE	MS	N	Species and endemism index
Area 4	8.62179	12	<i>C. buchii</i> (0.692-1.000); <i>C. ceibensis</i> (0.818-0.889); <i>C. fuertesii</i> (0.667-0.692)**; <i>C. incrassata</i> (0.909-1.000); <i>C. jimenezii</i> (0.773-0.833); <i>C. leoganensis</i> (0.633-0.654)**; <i>C. leonardii</i> (0.000-0.833); <i>C. microstachya</i> (0.000-0.733); <i>C. nodosa</i> (0.818-0.889); <i>C. picardae</i> (0.818-0.889); <i>C. pyrifolia</i> (0.000-0.567); <i>C. sintenisii</i> (0.500-0.773)
Area 5	9.35972	14	<i>C. benitensis</i> (0.500)*; <i>C. caesia</i> (0.450)*; <i>C. cowellii</i> (0.750)*; <i>C. fuertesii</i> (0.556)**; <i>C. leoganensis</i> (0.625)**; <i>C. longifolia</i> (0.750); <i>C. nipensis</i> (0.563)*; <i>C. oligantha</i> (0.688); <i>C. plumieri</i> (0.750); <i>C. rufescens</i> (0.833)*; <i>C. shaferi</i> (0.667); <i>C. troyana</i> (0.750); <i>C. wrightii</i> (0.479); <i>C. zebra</i> (0.750)
Area 6	2.72753	3	<i>C. guaranitica</i> (0.776); <i>C. paraguariensis</i> (0.788); <i>C. spinescens</i> (0.914)
Area 7	2.74536	4	<i>C. oblonga</i> (1.000); <i>C. peltada</i> (0.548)***; <i>C. rosea</i> (0.468)***; <i>C. schwackeana</i> (0.480)
Area 8	4.20600	6	<i>C. alnifolia</i> (0.512); <i>C. bullata</i> (0.795); <i>C. fastigiata</i> (0.578); <i>C. peltata</i> (0.750)***; <i>C. rosea</i> (0.667)***; <i>C. schwackeana</i> (0.654)***

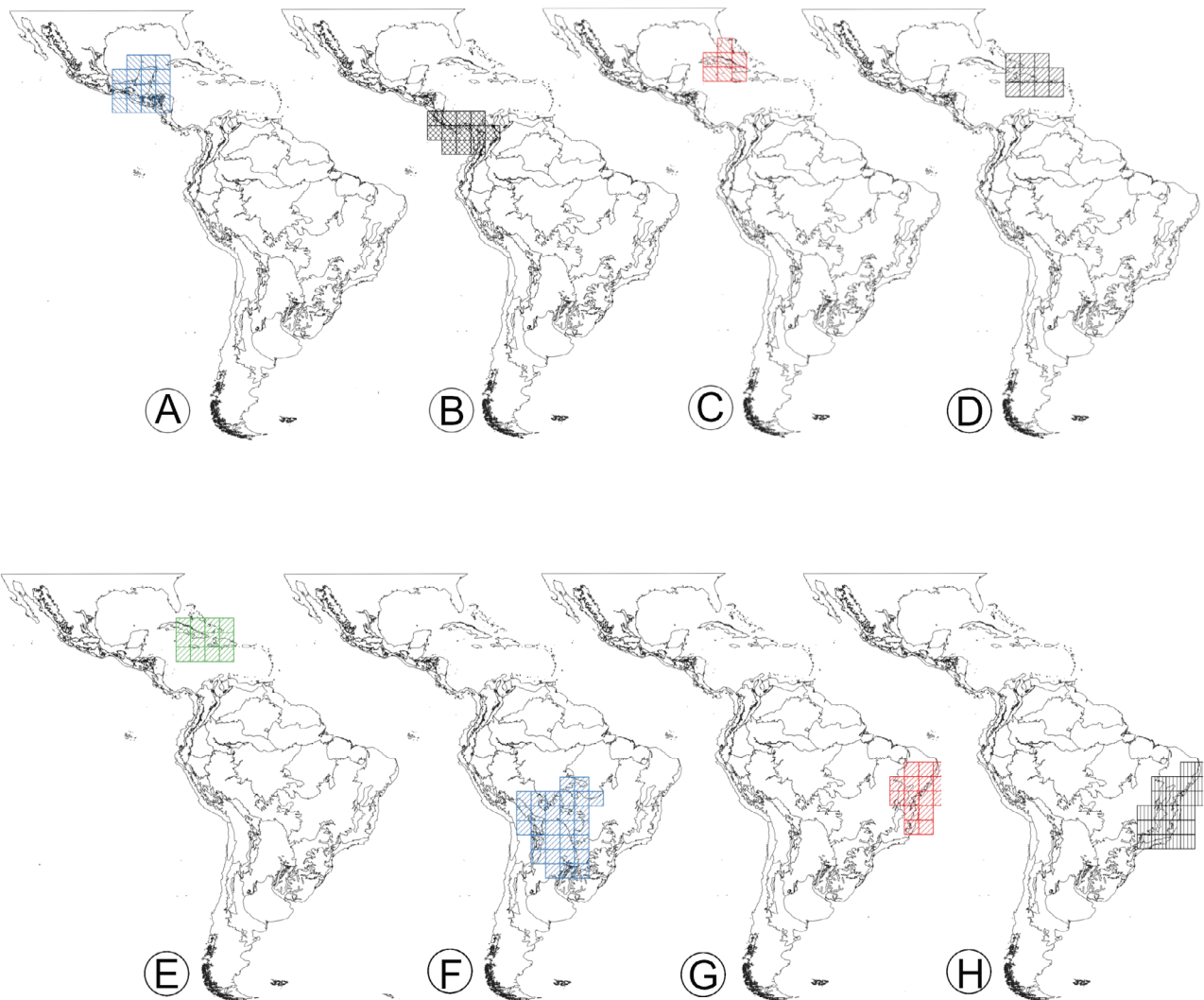


Figure 3. Results of the Endemism Analysis showing eight areas of endemism for *Coccoloba*: (A) northern Mesoamerica, (B) southern Mesoamerica, (C) Cuba and Jamaica, (D) Hispaniola and Puerto Rico, (E) Jamaica, Cuba, Hispaniola, and Puerto Rico, (F) northern Amazon Rainforest and the Chaco, (G-H) Atlantic Forest, including Chapada Diamantina and southern Espinhaço Range.

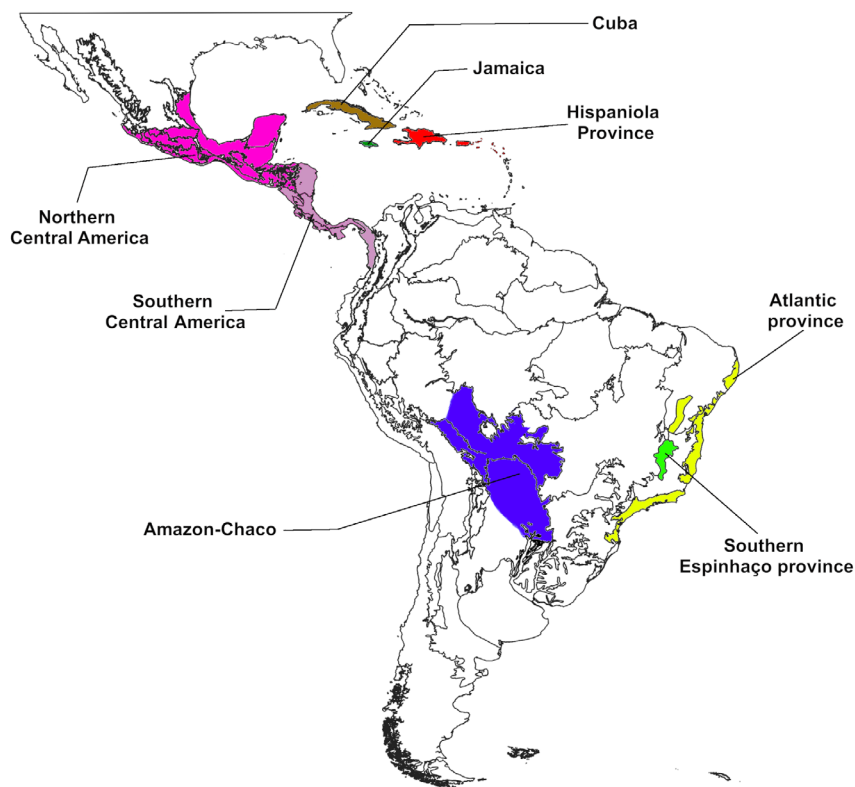


Figure 4. Consensus of *Cocoloba* endemism centers. The consensus between both PAE and EA results identified eight centers of endemism: Central America comprises two centers—northern and southern Central America; Cuba, Jamaica, and Hispaniola are the endemism centers of the Caribbean Islands. In South America, the Atlantic Forest, including Chapada Diamantina and southern Espinhaço Range, as well as southern Amazon and northern Chaco, were considered centers of endemism.

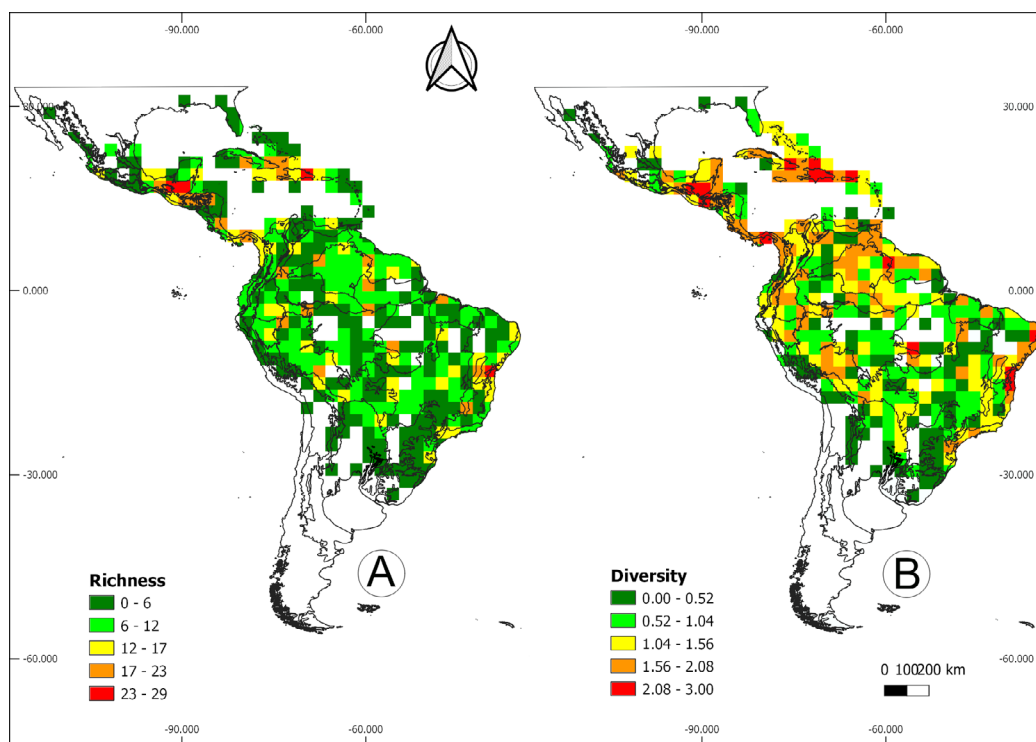


Figure 5. Areas of richness and diversity of *Cocoloba* in the Neotropics. (A) The richness of *Cocoloba* is high in Mesoamerica, the Caribbean Islands, and Atlantic Forest. (B) Areas with high diversity are located in Mesoamerica, the Caribbean Islands, Atlantic Forest, and Amazon Rainforest.



Sierra del Lacandón, Mirador and Tiger, in Guatemala, and the Biosphere Reserve Calakmul, in Mexico. In the Caribbean Islands, the highest areas of richness are found in Hispaniola, east of the Dominican Republic. The last area with high species richness is in Brazil, in the South of Bahia state in the Atlantic Forest area. The Amazon Rainforests did not show areas of high richness. These same areas of high species richness showed high species diversity (2.08 to 3.00 on the Shannon index). Species diversity was especially pronounced in Mesoamerica, the Caribbean Islands, and the Atlantic Forest. The Amazon Rainforest revealed two grid cells with high diversity located in east Madeira and Guiana provinces (Fig. 5B).

Drivers of species distribution

The Kruskal-Wallis test revealed significant statistical differences in the median for all explanatory variables presented ($p < 2.2e-16$). The violin plots comparing the distribution of data between the distribution patterns for each variable show that, in general, for each variable,

the median values of the distribution patterns located in South America are inversely proportional to those located in Central America and the Caribbean Islands (Fig. 6).

The GDM model (Fig. 7) revealed that geographic distance and environmental drivers accounted for 20.9% of the variation in the dissimilarity in *Coccoloba* species composition. Both explanatory variables influenced the dissimilarity in species composition in the neotropical region. Geographical distance has a strong influence on dissimilarity in the *Coccoloba* species composition (sum of I-spline coefficient: 23.540), followed by solar radiation (kJ m⁻² day⁻¹) in April (srad04) (sum of I-spline coefficient: 3.409), precipitation of wettest quarter (bio16), and mean temperature of warmest quarter (bio10) (sum of I-spline coefficient: 2.310 and 1.903, respectively). The remaining variables had a low influence on dissimilarity (sum of I-spline coefficient < 1). The precipitation of warmest quarter (bio18), solar radiation (kJ m⁻² day⁻¹) in November (Srad11), and calcium carbonate and gypsum (sq06) variables did not have an influence on the dissimilarity in species composition in the Neotropics (Tab. 4).

Table 4. Percentage of deviance explained in generalized dissimilarity modeling (GDM), model and null deviance, intercept, and the sum of I-spline predictor variables of the GDM models for *Coccoloba* species dissimilarity. The I-spline coefficient values were arranged in ascending order.

GDM modelling summary	
Percent deviance explained	20.986
GDM model deviance	1780454.605
Null deviance	2253327.058
Intercept	0.764
Predictor variables and sum of I-spline coefficients	
Geographical distance	23.540
srad04	3.409
bio16	2.310
bio10	1.903
bio04	0.831
sq02	0.790
bio15	0.759
bio02	0.493
bio19	0.404
srad02	0.328
sq04	0.135
sq07	0.098
bio14	0.089
bio18	0
Srad11	0
sq06	0

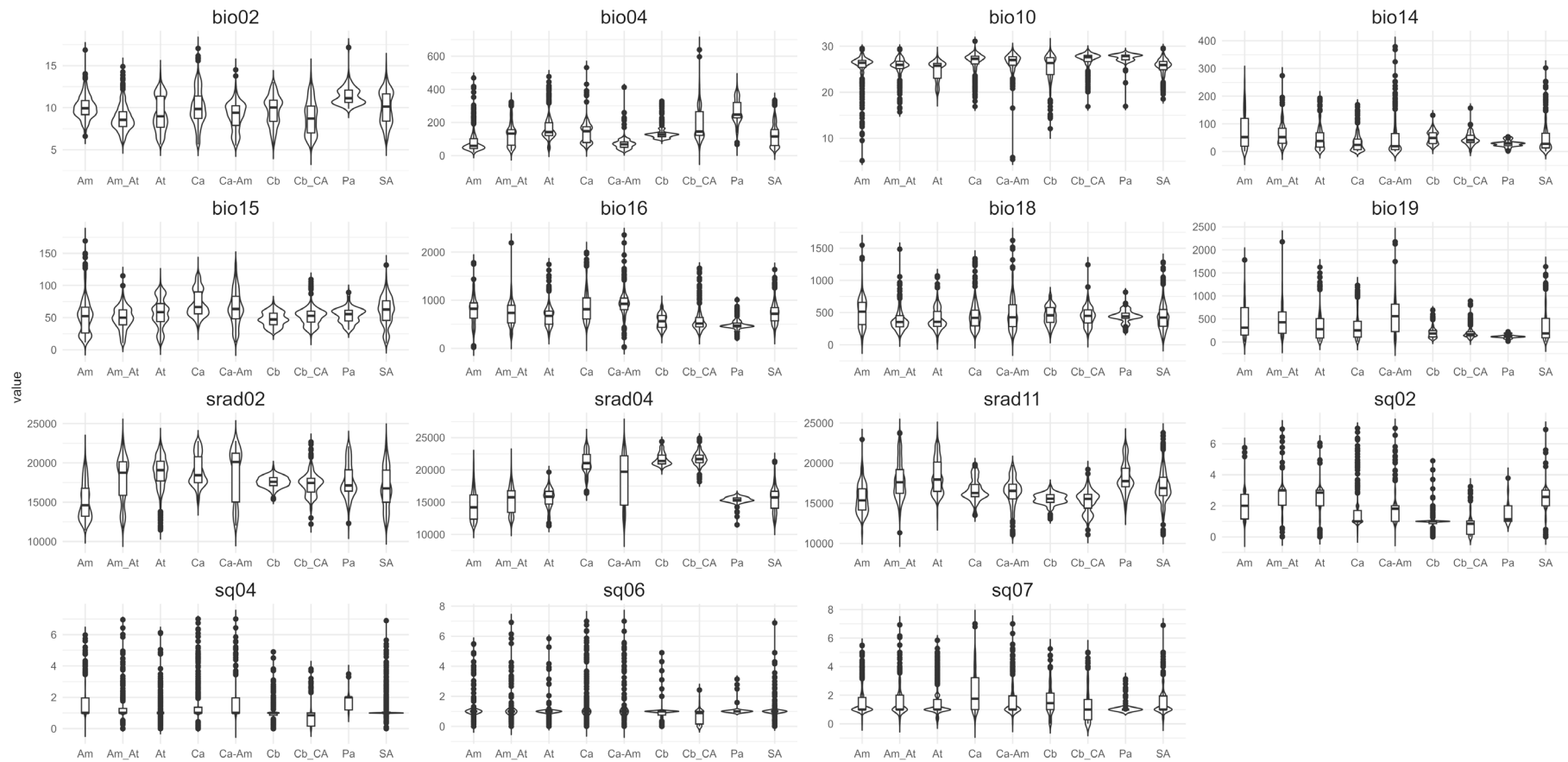


Figure 6. Violin plots showing the median values for each distribution pattern in relation to explanatory variables.



Generalized Dissimilarity Modelling

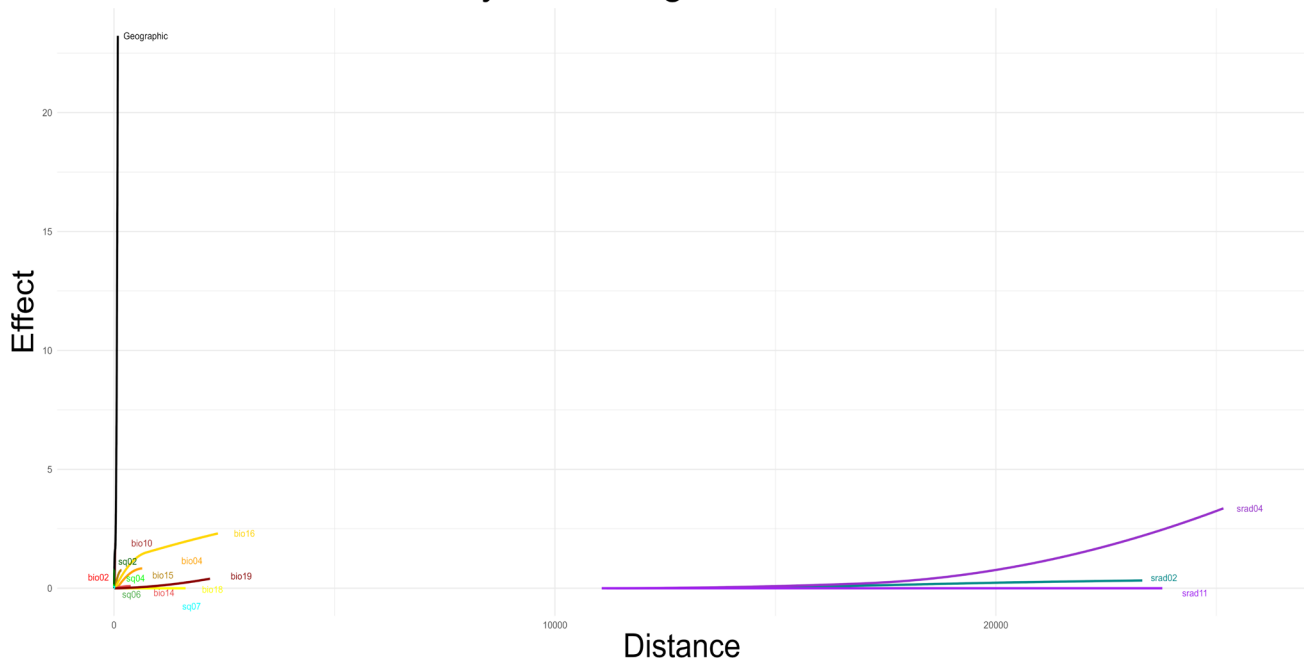


Figure 7. Relationship between predictor dissimilarities and the fitted values of the I-splines (ecological distance) calculated by generalized dissimilarity modeling.

Discussion

Our main results revealed that the four centers of biodiversity previously suggested in the literature correspond to the main distribution patterns of *Coccoloba* species (Fig. 1, Fig. 4–5). These distribution patterns are strongly influenced by solar radiation, precipitation, temperature, and nutrient retention capacity of the soil (Fig. 7, Tab. 4). On the other hand, areas of high endemism, diversity, and richness are located in smaller areas within these major centers (Fig. 4–5).

Drivers and *Coccoloba* species distribution

The GDM model pointed out that the dissimilarity of *Coccoloba* species composition is strongly related to geographic distance, environmental (solar radiation (kJ m⁻² day⁻¹) in April, precipitation of wettest quarter, mean temperature of warmest quarter, temperature seasonality), and edaphic conditions (nutrient retention capacity). The importance of these drivers differs significantly depending on the group and the area of study. For example, Figueiredo *et al.*, (2018) suggested that soil is more important than climatic drivers in predicting the distribution of Amazonian plant species, while Rago *et al.* (2021) highlighted that solar radiation is an important component in determining the plant community in the Patagonian steppe. Our results pointed out that these drivers act at different intensities in

the distribution patterns located in South America compared to those in Central America and the Caribbean Islands, which are closely related to habitat heterogeneity in the Neotropics (Antonelli & Sanmartín, 2011).

Koenemann & Burke (2020) highlighted that *Coccoloba* species distributed in both Caribbean Islands and Central America (e.g. *C. venosa*), or just in the Caribbean Islands (e.g. *C. pubescens* L.), do not occur in South America, and vice versa; and species distributed on the Coast of Brazil, are not in Amazonia, except for a few species with wide distributions in South America (Melo, 2004; Alves *et al.*, 2021; Alves & Buriel, 2022; Alves *et al.*, 2024). This difference in species composition across different areas of the Neotropics highlights the heterogeneity of this region. Also, one of the aspects highlighted by the GDM model is that only solar radiation drivers influence species distribution at macro scales, while precipitation, temperature, and soil drivers influence species distribution at local scales (Fig. 7).

Although the environmental drivers found here are a starting point for understanding the composition of *Coccoloba* species, it is still unclear which historical and evolutionary processes promoted the distribution of these lineages throughout neotropical forests, resulting in the current spatial distribution patterns. To understand the evolutionary history of the genus, it is necessary to expand phylogenetic studies in *Coccoloba*.

The Cerrado and Caatinga biomes (stricto sensu) appear to act as an ecological barrier in the distribution of most

South American *Coccoloba* species. This observation is supported by the low richness in grid cells in these regions and species distribution. In the *Cerrado* biome, *Coccoloba* species occurrences are generally associated with high humidity environments and water availability, like riparian forests along river banks and gallery forests (Melo, 2004); while in the *Caatinga* biome, *Coccoloba* species occur just in the *brejos de altitude*, which are enclaves of montane ombrophilous forest in areas of high altitude and humidity, forming vegetational islands in the *Caatinga* biome (Alves *et al.*, 2021).

The Caribbean-Central American distribution pattern is a disjunction promoted by the Caribbean Sea. This suggests that species following this pattern (and other wide patterns: South America and Amazon-Central America) have a long-distance dispersal mechanism. Birds have been identified as fruit and seed dispersers for *Coccoloba gigantifolia* Melo, Cid Ferreira & Gribel, and *C. uvifera* L. (Ferreira *et al.*, 2023). Our field observations of *Coccoloba alnifolia* Casar. and *C. warmingii* Meisn. along riparian forests, suggest that its fruits can be transported by water. The fruit of *Coccoloba* is an achene with a ruminant endosperm, surrounded by an accrescent and succulent hypanthium (Melo, 2004; Ferreira *et al.*, 2023). These features suggest that the hydrochory can be considered as a long-dispersal mechanism in *Coccoloba*.

Endemism centers of *Coccoloba*

Centers of endemism in the Caribbean Islands are associated with the geographic isolation of their archipelagos, promoted by the Atlantic Ocean and the Caribbean sea. These areas, recovered in this study, correspond with those suggested by Howard (1949; 1957), when he studied *Coccoloba* species in the Caribbean Islands and observed many endemic species in Cuba, Jamaica, and the Virgin and Bahama islands, such as *Coccoloba microphylla* Griseb., and *C. geniculata* Lindau. In Central America, Montane Forests appear to provide a biogeographic barrier determining the two great centers of endemism for *Coccoloba* in this region: northern and southern Central America.

Koenemann & Burke (2020) suggested that *Coccoloba* species from Central America and the Caribbean Islands are phylogenetically close, likely due to the intense migration of species, and therefore Central America and the Caribbean should be considered as a single biogeographic unit. Based on the distribution patterns, floristic composition, endemism patterns, richness, and diversity of *Coccoloba* species recovered in this study, we disagree with that assessment and here propose that these two regions should be considered as distinct biogeographic units.

The Atlantic Forest taken as a whole is one of the major centers of endemism for *Coccoloba*, supported by some species distributed along this region (e.g. *C. rosea* Meisn., *C. alnifolia*, and *C. warmingii*). These results corroborate findings in the literature, where the Atlantic Forest has

already been pointed out as one of the important centers of endemism for the genus (Brandbyge, 1990; Melo, 2004; Koenemann & Burke, 2020). Both Espinhaço Range sectors were identified by the PAE and EA. However, Chapada Diamantina is more closely related to the Atlantic Forest in terms of endemism and floristic composition, because both areas share species that occur exclusively in these two regions. This finding is consistent with the results discussed by Alves *et al.* (2007), who pointed out that some angiosperm groups show disjunction between the Campos Rupestres of the Espinhaço Range and the Restinga in the Atlantic Forest, supported here by *C. alnifolia*, *C. striata*, and *C. warmingii*.

Despite the large number of *Coccoloba* species occurring in the Amazon Rainforest, there is a low number of endemic species in this region. Melo (2004) suggested that the low number of endemic species in the region is the result of the majority of species having wide distributions. The consensus area of endemism recovered in this region (Fig. 4) is located in the extreme southern Amazon Rainforest, and extends to the northeast of the Chaco. Examples of species that follow this distribution pattern are *Coccoloba guaranitica* Hassler, *C. paraguariensis* Lindau, and *C. tiliacea* Lindau.

Richness and diversity of *Coccoloba*

Areas of great richness are typically located in areas with correspondingly high diversity. The Caribbean Islands, despite the small land area of the archipelago, show great diversity when compared with *Coccoloba* species diversity in the continental region (e.g. the Amazon Rainforest). Areas of richness and diversity in Central America and coastal Brazil (southern Bahia state) are associated with great environmental conservation units: the Biosphere Reserves of Montes Azules and Maya in Central America, and the Biological Reserve of Una in coastal Brazil. These biodiversity patterns are associated with solar radiation, precipitation, temperature, and nutrient retention capacity of the soil drivers. Additionally, Carnaval & Moritz (2008) suggest that topographic variables and the influence of ocean factors are considered important environmental predictors of plant diversity in coastal regions.

Coccoloba species are found in the highest numbers in South America (the Amazon and Atlantic Forests), followed by the Caribbean Islands and Central America. The distribution patterns in South America are influenced by the ecological boundaries of the Atlantic Forest and the Amazon Rainforest. The moist forests provide the environmental conditions to maintain the species composition, and the dry forest (*Caatinga* and *Cerrado*) acts as an ecological barrier for most *Coccoloba* species in South America. The drivers responsible for the dissimilarity of *Coccoloba* species distributions in the neotropical region are the solar radiation in April (srad04), precipitation of wettest quarter (bio16) and mean temperature of warmest quarter (bio10).



Nevertheless, our analyses also revealed that the four traditional centers of *Coccoloba* diversity, in their entirety, could not be considered centers of biodiversity. Rather, areas of heightened endemism, diversity, and richness for *Coccoloba* are located in smaller areas within these regions. The centers of diversity and endemism previously suggested in the literature correspond to the main distribution patterns of *Coccoloba* species.

Acknowledgements

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the research scholarship granted to the first author (process 88887.666901/2022-00).

Authors' Contributions

J.V.A. and D.M.K. wrote the main manuscript text. J.V.A. did the data collection, conducted the statistical analyses, and prepared figures. D.M.K. reviewed the grammar of the English text. R.B.L. supervised all the steps. All authors reviewed the manuscript.

Conflict of Interest

The authors declare no competing interests.

Supplementary material

The following online material is available for this article:

Table S1. The dataset with full information for each specimen analyzed. This dataset was used as the template for all analyses performed here.

Table S2. Clusters returned by *K-means* based on geographic coordinates of *Coccoloba* species. Cluster numbers 1, 2, 3, and 4 represent, respectively, eastern South America, the Amazon Rainforest, Central America, and the Caribbean Islands. The last column represents the distribution patterns after checking the species present in more than one cluster.

References

- Alves JV, Melo E, Buril MT. 2021. Polygonaceae Juss. de um fragmento de Mata Atlântica e padrão de distribuição das espécies no Nordeste do Brasil. *Arrudea* 7: 39-50.
- Alves JV, Buril MT. 2022. Flora da Estação Ecológica de Tapacurá, São Lourenço da Mata, Pernambuco - Brasil: Polygonaceae Juss. *Arrudea* 8: 37-45.
- Alves JV, Koenemann DM, Louzada R. 2024. How many sources of evidence are needed to delimit a taxon? Disentangling the identity of two *Coccoloba* species (Polygonaceae). *Plant Systematics and Evolution* 310: 1-15.
- Alves RJV, Cardin L, Kropf MS. 2007. Angiosperm disjunction "Campos rupestres -restingas": A re-evaluation. *Acta Botanica Brasilica* 21: 676-685.
- Ancona JJ, Ortiz-Diaz JJ, Tun-Garrido J. 2023. *Coccoloba efigeniana* and *C. ibarrae*, two new species of *Coccoloba* sect. *Paniculatae* Meisn. (Polygonaceae). *Phytotaxa* 599: 101-102.
- Antonelli A, Sanmartín I. 2011. Why are there so many plant species in the Neotropics? *Taxon* 60: 403-414.
- Antonelli A. 2022. The rise and fall of neotropical biodiversity. *Botanical Journal of the Linnean Society* 199: 8-24.
- Bitencourt C, Rapini A. 2013. Centres of Endemism in the Espinhaço Range: Identifying cradles and museums of *Asclepiadoideae* (Apocynaceae). *Systematics and Biodiversity* 11: 525-536.
- Brandbyge J. 1990. The diversity of micromorphological features in the genus *Coccoloba* (Polygonaceae). *Nordic Journal of Botany* 10: 25-44.
- Burke JM, Sanchez A, Kron K, Luckow M. 2010. Placing the woody tropical genera of the Polygonaceae: A hypothesis of character evolution and phylogeny. *American Journal of Botany* 97: 1377-1390.
- Carpenter KE, Springer VG. 2005. The center of marine shore fish biodiversity: The Philippine Islands. *Environmental Biology of Fishes* 72: 467-480.
- Carnaval AC, Moritz C. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic Forest. *Journal of Biogeography* 35: 1187-1201.
- Dryflor BRK, Delgado-Salinas A, Dexter KG, Linares-Palomino R, Oliveira-Filho A, Prado D *et al.* 2016. Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* 353: 1383-1387.
- Dusek R, Popelková R. 2017. Theoretical view of the Shannon index in the evaluation of landscape diversity. *Acta Universitatis Carolinae Geographica* 47: 5-13.
- Escalante T. 2011. De cómo el análisis de parsimonia de endemismos (PAE) tampoco explica la selección natural. *Revista Mexicana de Biodiversidad* 82: 1057-1059.
- Fernandes MF, Cardoso D, Queiroz LP. 2020. An updated plant checklist of the Brazilian Caatinga seasonally dry forests and woodlands reveals high species richness and endemism. *Journal of Arid Environments* 174: 104079.
- Ferreira SAN, Gentil DFO, Ferreira CAC. 2023. Emergence of *Coccoloba gigantifolia* Melo, Cid Ferreira & Gribel seedlings as a function of raceme and fruit maturation. *Journal of Seed Science* 45: e202345015.
- 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.
- Fick SE, Hijmans RJ. 2017. WorldClim 2: New 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302-4315.
- Figueiredo FOG, Zuquim G, Tuomisto H, Moulatlet GM, Balslev H, Costa FRC. 2018. Beyond climate control on species range: The importance of soil data to predict distribution of Amazonian plant species. *Journal of Biogeography* 45: 190-200.
- Goloboff P. 1997. Self-Weighted optimization: Tree searches and character state reconstructions under implied transformation costs. *Cladistics* 13: 225-245.
- Hijmans RJ, Guarino L, Cruz RE. 2001. Computer tools for spatial analyses of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources* 1127: 15-19.
- Hijmans RJ. 2023. raster: Geographic Data Analysis and Modeling. R package version 3.5-15. <https://raster.org/raster>. 01 Aug. 2024.
- Howard RA. 1949. The genus *Coccoloba* in Cuba. *Journal of the Arnold Arboretum* 30: 388-424.
- Howard RA. 1957. Studies in the genus *Coccoloba*, III. The Jamaican species. *Journal of the Arnold Arboretum* 38: 81-106.

- Howard RA. 1960. Studies in the genus *Coccoloba*, IX. A critique of the South American species. *Journal of the Arnold Arboretum* 41: 213-390.
- Howard RA. 1961. Studies in the genus *Coccoloba*, X. New species and a summary of distribution in South America. *Journal of the Arnold Arboretum* 42: 87-95.
- Howard RA. 1992. Collected notes on *Coccoloba* L. (Polygonaceae). *Brittonia* 44: 356-367.
- Kassambara A, Mundt F. 2020. Factoextra: Extract and visualize the results of multivariate data analyses: R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>. 6 Jan. 2024.
- Koenemann DM, Burke JM. 2020. A molecular phylogeny for the genus *Coccoloba* (Polygonaceae) with an assessment of biogeographic patterns. *Systematic Botany* 45: 567-575.
- Kuhn M. 2008. Building predictive models in R using the caret Package. *Journal of Statistical Software* 28: 1-26.
- Lima RAF, Souza VC, Siqueira MF, Steege H. 2020. Defining endemism levels for biodiversity conservation: Tree species in the Atlantic Forest hotspot. *Biological Conservation* 252: 1-9.
- Maciel EA, Martins FR. 2021. Rarity patterns and the conservation status of tree species in South American savannas. *Flora* 285: 151942.
- Melo E. 2004. As espécies de *Coccoloba* P.Browne (Polygonaceae) da Amazônia brasileira. *Acta Amazonica* 34: 525-551.
- Melo E, Ferreira CA, Gribel R. 2019. A new species of *Coccoloba* P. Browne (Polygonaceae) from the Brazilian Amazon with exceptionally large leaves. *Acta Amazonica* 49: 324-329.
- Mickevich MF, Lipscomb D. 1991. Parsimony and the choice between different transformations for the same character set. *Cladistics* 7: 111-139.
- Mokany KC, Ware C, Woolley SNC, Ferrier S, Fitzpatrick MC. 2022. A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. *Global Ecology and Biogeography* 31: 802-821.
- Morrone JJ. 1994. On the identification of areas of endemism. *Systematic Biology* 43: 438-441.
- Morrone JJ. 2014. Parsimony analyzes of endemism (PAE) revisited. *Journal of Biogeography* 41: 842-854.
- Morrone JJ, Escalante T, Rodríguez-Tapia G, Carmona A, Arana M, Mercado-Gomez JD. 2022. Biogeographic regionalization of the Neotropical region: New map and shapefile. *Anais da Academia Brasileira de Ciências* 94: e20211167.
- Noa IEC. 2013. *Coccoloba berazainae* and *C. cristalensis* (Polygonaceae) from Cuba, a new species and a new combination. *Willdenowia* 43: 319-323.
- Oliveira-Filho A, Dexter T, Pennington KG, Simon RT, Bueno ML, Neves DM. 2021. On the floristic identity of Amazonian vegetation types. *Biotropica* 53: 767-777.
- Posit Team. 2024. RStudio integrated development environment for R. <https://www.posit.co/>. 6 Jan. 2024.
- Pulido-Chadid K, Virtanen E, Geldmann J. 2023. How effective are protected areas for reducing threats to biodiversity? A systematic review protocol. *Environmental Evidence* 12: 18.
- R Core Team. 2023. R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>. 10 Apr. 2024.
- Rago RMM, Urretavizcaya MF, Defossé GE. 2021. Relationships among forest structure, solar radiation, and plant community in ponderosa pine plantations in the Patagonian steppe. *Forest Ecology and Management* 502: 119749.
- Rodrigues AV, Duarte L. 2024. Mapping species richness and evolutionary regions of the genus *Myrcia*. *Journal of Biogeography* 51: 827-839.
- Santos CMD, Fuhlendorf M. 2018. GeX: An automated tool for generating XYD files for analyzes of endemism using VNDM. *Cladistics* 35: 125-129.
- Servonnat M, Kaye R, Siringan, FP, Munar J, Yap HT. 2019. Imperatives for Conservation in a Threatened Center of Biodiversity. *Coastal Management* 47: 453-472.
- Smith-Ramírez C. 2004. The Chilean coastal range: A vanishing center of biodiversity and endemism in South American temperate rainforests. *Biodiversity and Conservation* 13: 373-393.
- Souza JF, Bueno ML, Salino A. 2021. Atlantic Forest: Centres of diversity and endemism for ferns and lycophytes and conservation status. *Biodiversity and Conservation* 30: 2207-2222.
- Szumik C, Cuzzo F, Goloboff PA, Chalup AE. 2002. An optimality criterion to determine areas of endemism. *Systematic Biology* 51: 806-816.
- Szumik C, Goloboff A. 2004. Areas of endemism: An improved optimality criterion. *Systematic Biology* 53: 968-977.
- Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York, Springer-Verlag.

