



UNIVERSIDADE FEDERAL DE SERGIPE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E
CONSERVAÇÃO



**PADRÕES ESPACIAIS E DETERMINANTES DA DIVERSIDADE TAXONÔMICA,
FILOGENÉTICA E FUNCIONAL DE PLANTAS NA RESTINGA BRASILEIRA**

Tese de Doutorado

Eduardo Vinícius da Silva Oliveira

SÃO CRISTÓVÃO-SE
JULHO/2021

EDUARDO VINÍCIUS DA SILVA OLIVEIRA

**PADRÕES ESPACIAIS E DETERMINANTES DA DIVERSIDADE
TAXONÔMICA, FILOGENÉTICA E FUNCIONAL DE PLANTAS NA
RESTINGA BRASILEIRA**

Tese apresentada ao Programa de Pós-Graduação em Ecologia e Conservação da Universidade Federal de Sergipe, como requisito parcial para obtenção do título de Doutor em Ecologia e Conservação.

Orientador (a): Prof. Dr. Sidney Feitosa Gouveia

Coorientador (a): Prof.^a Dra. Myrna Friederichs Landim

SÃO CRISTÓVÃO-SE
JULHO/2021

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

O48p Oliveira, Eduardo Vinícius da Silva.
 / Eduardo Vinícius da Silva Oliveira; orientador Sidney Feitosa
 Gouveia. – São Cristóvão, SE, 2021.
 130 f. : il.

Tese (doutorado em Ecologia e Conservação) – Universidade
Federal de Sergipe, 2021.

1. Diversidade biológica. 2. Ecossistemas. 3. Biologia -
Classificação. 4. Biologia costeira. 5. Restingas. I. , orient. II. Título.

CDU 551.435.32

TERMO DE APROVAÇÃO

PADRÕES ESPACIAIS E DETERMINANTES DA DIVERSIDADE TAXONÔMICA, FILOGENÉTICA E FUNCIONAL DE PLANTAS NAS RESTINGA BRASILEIRA

por

EDUARDO VINICIUS DA SILVA OLIVEIRA

Tese apresentada ao Programa de Pós-Graduação em Ecologia e Conservação da Universidade Federal de Sergipe, como parte dos requisitos exigidos para a obtenção do título de Doutor em Ecologia e Conservação.

APROVADA pela banca examinadora composta por



PROF. DR. SIDNEY FEITOSA GOUVEIA
Programa de Pós-Graduação em Ecologia e Conservação da
Universidade Federal de Sergipe



PROF.ª DR.ª LEILA MEYER
Universidade Federal de Minas Gerais



PROF. DR. CRISÓFORO FABRÍCIO VILLALOBOS CAMACHO
Instituto de Ecología, A.C./México

FRANCISCO SOARES SANTOS
FILHO:30218152353

Digitally signed by FRANCISCO SOARES SANTOS FILHO:30218152353
DN: c=BR, o=ICP-Brasil, ou=Secretaria da Receita Federal do Brasil -
RFB, ou=RFB e-CPF A1, ou=(EM BRANCO), ou=26553520000115,
cn=FRANCISCO SOARES SANTOS FILHO:30218152353
Date: 2021.07.08 10:18:11 -03'00'

PROF. DR. FRANCISCO SOARES SANTOS FILHO
Universidade Estadual do Piauí



PROF. DR. PABLO ARIEL MARTINEZ
Programa de Pós-Graduação em Ecologia e Conservação da
Universidade Federal de Sergipe

São Cristóvão/SE, 07 de julho de 2021

AGRADECIMENTOS

A todas as pessoas que contribuíram de forma direta ou indireta para a conclusão deste trabalho, os meus mais sinceros agradecimentos. Agradeço especialmente:

À COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES) e a FUNDAÇÃO DE APOIO À PESQUISA E À INOVAÇÃO TECNOLÓGICA DO ESTADO DE SERGIPE (FAPITEC) pela concessão de bolsa de doutorado.

Ao meu orientador Prof. Dr. SIDNEY FEITOSA GOUVEIA, pelos seus valiosos ensinamentos, paciência, atenção e incentivo ao aprimoramento deste trabalho.

À minha coorientadora Profa. Dra. MYRNA FRIEDERICHS LANDIM, pelos seus valiosos ensinamentos e incentivo, desde a graduação. Obrigado por ter estado disponível quando eu precisei.

Ao PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E CONSERVAÇÃO (PPEC) pela oportunidade e aos seus professores pelos diversos ensinamentos.

À secretária do PPEC, JULIANA CORDEIRO pelo suporte nos mais variados assuntos.

À minha esposa JAMILLE FERREIRA, pelo carinho e atenção e por me ajudar a levantar nos momentos difíceis.

A toda a minha família, especialmente aos meus pais, MÁRCIA VARJÃO e ANTONIO MENDES pelo apoio, carinho e torcida.

Aos colegas do PPEC e do PIBILab pela troca de experiências e momentos de descontração, especialmente a CAROLINE SANTOS, DAVI ALVES, ERIVELTON ROSÁRIO, HAMILTON BARRETO, HELON OLIVEIRA, ISADORA CORREIA, MARCOS VINICIOS, RAFAELLA SANTANA e SAULO VALENÇA.

E por fim a VOCÊ que me dá a honra de ler esta Tese.

RESUMO GERAL

Dentre as consequências das interações dos organismos com o meio estão diversos padrões de coexistência de múltiplos organismos distintos no espaço ou padrões de comunidades. Desses padrões, aqueles relacionados a quantos e quais organismos compõem essas comunidades permanecem sendo debatidos. Essas duas questões são expressas em termos de diversidade e composição de espécies. Historicamente, esses estudos têm enfatizado a divisão por escalas espaciais, sobretudo localmente dentro da comunidade – ou diversidade alfa – e entre comunidades – ou diversidade beta. Recentemente, estudos sobre esses padrões têm contemplado diferentes aspectos da biodiversidade, além da abordagem tradicional baseada na taxonomia, mas incluindo suas relações evolutivas e suas características, ou seja, aspectos filogenéticos e funcionais. Contudo, ainda não há consenso sobre suas causas, especificamente quando envolvem grupos biológicos específicos em contextos ambientais específicos, seja pela sua complexidade inerente, seja pela escassez sistemática de informações sobre as espécies. Esse é o caso de plantas em ambientes costeiros tropicais, como a vegetação de Restinga. Neste estudo, desafiando diferentes tipos de lacunas de informação sobre as espécies de plantas litorâneas do Brasil, investigamos os padrões espaciais e os processos envolvidos na diversidade e composição dessas comunidades. No primeiro capítulo, após considerar o nível de conhecimento relativo à amostragem de plantas, investigamos os padrões geográficos e testamos diferentes hipóteses ecológicas para a variação na riqueza de espécies. Nossos resultados revelaram que, apesar das lacunas de conhecimento em mais de 3/4 da área, apontamos que a riqueza de espécies parece ser melhor explicada pela heterogeneidade de habitat, propriedades do solo e restrições de água. No segundo capítulo, investigamos os padrões de composição, avaliando regras de montagem dos padrões de estrutura funcional e filogenética das assembleias e o papel de diferentes preditores ambientais. Mostramos que, sob maior restrição hídrica, predominaram assembleias com espécies funcionalmente similares e pré-adaptadas as condições da Restinga. Em áreas associadas a ecossistemas megadiversos de florestas tropicais pluviais, a composição das assembleias tende a arranjos mais estocásticos ou com espécies mais distintas quanto as suas similaridades fenotípica e filogenética. Restrições de água foram os principais responsáveis em causar variação nesses padrões de estrutura filogenética e funcional. Por fim, no terceiro capítulo, avaliamos a contribuição relativa dos componentes da diversidade beta, i.e., substituição e aninhamento, se esses processos resultam de dispersão estocástica e que preditores ambientais respondem pelos padrões de dissimilaridade entre assembleias. Os padrões de diversidade beta taxonômica não correspondem a uma dinâmica estocástica, ao contrário dos padrões de dissimilaridade filogenética e funcional. A disponibilidade de água parece ser o fator determinante dos padrões espaciais de dissimilaridade dos três aspectos da diversidade beta. As três perspectivas avaliadas revelaram um padrão complexo de interação entre padrões regionais de diversidade – em que ecossistemas adjacentes doadores de espécies contribuíram de forma distinta – e diferentes graus de restrição à colonização de espécies devido a fatores ambientais, especificamente a disponibilidade de água. Nosso estudo foi o primeiro a descrever os padrões e determinantes das assembleias de plantas de Restinga, contribuindo para o entendimento da origem e dinâmica desse ecossistema e padrões em sistemas e contextos ambientais específicos. Apesar das lacunas de informações sobre a biodiversidade, essas limitações não devem impedir nossa capacidade de investigar esses padrões em escala geográfica.

Palavras-chave: déficits de conhecimento, montagem de assembleias, restrições hídricas, dispersão histórica, *pools* de espécies.

GENERAL ABSTRACT

Among the consequences of the interactions of organisms with the environment are several patterns of coexistence of multiple distinct organisms in space or patterns of communities. Of these patterns, those related to how many and which organisms compose these communities remain under debate. These two issues are expressed in terms of species diversity and composition. Historically, these studies have emphasized division by spatial scales, particularly locally within the community – or alpha diversity – and between communities – or beta diversity. Recently, studies on these patterns have contemplated different aspects of biodiversity, beyond to the traditional approach based on taxonomy, but including their evolutionary relationships and their characteristics, that is, phylogenetic and functional aspects. However, there is still no consensus on their causes, specifically when they involve specific biological groups in specific environmental contexts, either because of their inherent complexity or because of the systematic scarcity of information about the species. This is the case of plants in tropical coastal environments, such as Restinga vegetation. In this study, challenging different types of information gaps about coastal plant species in Brazil, we investigate the spatial patterns and processes involved in the diversity and composition of these communities. In the first chapter, after considering the level of knowledge regarding plant sampling efforts, we investigated geographic patterns and tested different ecological hypotheses for variation in species richness. Our results revealed that despite knowledge gaps in more than 3/4 of the area, we found that species richness appears to be best explained by habitat heterogeneity, soil properties and water constraints. In the second chapter, we investigate the compositional patterns, evaluating assembly rules of the functional and phylogenetic structure patterns of the assemblages and the role of different environmental predictors. We show that, under greater water constraint, assemblages with species functionally similar and pre-adapted to Restinga conditions predominated. In sites associated with megadiverse rainforest ecosystems, the composition of the assemblages tends to more stochastic arrangements or with more distinct species in terms of their phenotypic and phylogenetic similarities. Water constraints were the main responsible for causing variation in these patterns of phylogenetic and functional structure. Finally, in the third chapter, we evaluate the relative contribution of the beta diversity components, i.e., turnover and nestedness, if these processes result from stochastic dispersion and which environmental predictors account for the dissimilarity patterns between assemblages. Patterns of taxonomic beta diversity do not correspond to stochastic dynamics, unlike phylogenetic and functional dissimilarity patterns. Water availability appears to be the determining factor in the spatial patterns of dissimilarity of the three aspects of beta diversity. The three perspectives evaluated revealed a complex pattern of interaction between regional patterns of diversity – where adjacent ecosystems that are sources of species contributed in different ways – and different degrees of restriction on species colonization due to environmental factors, specifically water availability. Our study was the first to describe the patterns and drivers of Restinga plant assemblages, contributing to the understanding of the origin and dynamics of this ecosystem and patterns in contexts and specific environmental systems. Despite gaps in biodiversity information, these limitations should not impede our ability to investigate these patterns at a geographic scale.

Keywords: knowledge shortfalls, community assembly, water constraints, historical dispersion, species pools.

SUMÁRIO

Lista de Figuras	9
Lista de Tabelas	11
Introdução geral.....	14
CAPÍTULO 1 - Sampling effort and the drivers of plant species richness in the Brazilian coastal regions.....	22
Abstract.....	23
Introduction	23
Materials and Methods	26
Results	31
Discussion.....	34
CAPÍTULO 2 - Patterns and drivers of functional and phylogenetic structures of coastal woody vegetation of Eastern South America	50
Abstract.....	51
Introduction	52
Material and Methods	55
Results	60
Discussion.....	66
CAPÍTULO 3 - Diversidade-beta de múltiplos aspectos da diversidade de árvores em ecossistemas costeiros da América do Sul	84
Resumo	85
Introdução	86
Material e Métodos	89
Resultados.....	98
Discussão.....	105
Conclusões gerais	128

Lista de Figuras

CAPÍTULO 1 - Sampling effort and the drivers of plant species richness in the Brazilian coastal regions

Figure 1. Maps of a) distribution of the number of records, b) observed richness, c) expected richness, and d) inventory completeness for vascular plants in the Brazilian coastal vegetation. Cells with the bolder border have higher knowledge (completeness $\geq$ 50% and records number $\geq$ mean). Brazilian phytogeographic domains are delimited in gray. Grid cells have a 0.5° spatial resolution..... 32

Figure 2. Comparisons of the effects, grouped by hypotheses, of the environmental predictors on plant richness from the Brazilian coastal region based on partial regression coefficients. Results include portions uniquely accounted by ecological hypothesis, shared contributions and unexplained variance (residuals). The sum of partial coefficients exceeds 1.0 because of some degree of collinearity among predictors, which results in negative variances in the shared fraction of variance explained (omitted)..... 34

CAPÍTULO 2 - Patterns and drivers of functional and phylogenetic structures of coastal woody vegetation of Eastern South America

Figure 1. Functional and phylogenetic structure of assemblages of flowering woody plants in the Brazilian Restingas, Eastern coast of South America. Values of observed Phylogenetic Species Variability (PSV) and assemblages with clustering and overdispersed structure significantly different from chance are shown. The phylogenetic structure was obtained by averaging the 100 randomizations using different phylogenies. Those PSV values that differ from a null model are represented by the mean of the significance value (p-value)..... 62

Figure 2. The individual and shared proportion of variance explained (%) by the environmental factors as drivers of the phylogenetic and functional structure of woody plants in the Brazilian Restingas, grouped according to the corresponding hypothesis. Values are representative of the better models for each response variable..... 66

CAPÍTULO 3 - Diversidade-beta de múltiplos aspectos da diversidade de árvores em ecossistemas costeiros da América do Sul

Figura 1. Mapas de (A) riqueza de espécies de árvores por assembleia (0,5° de resolução espacial) obtida pela sobreposição da área de distribuição geográfica das espécies preditas pelos modelos de distribuição de espécies (SDMs) ao longo da zona costeira do Brasil, (B) delimitação da vegetação de Restinga considerada para a modelagem de distribuição das espécies de árvores e (C) localização do Brasil na América do Sul..... 100

Figura 2. Comparações entre os valores proporcionais de contribuição de substituição (cinza) e aninhamento (branco) para a diversidade beta taxonômica, filogenética e funcional dos conjuntos de assembleias pertencentes aos domínios fitogeográficos presentes na região costeira do Brasil..... 103

Figura 3. Distribuição de frequência de densidade dos atributos funcionais contínuos (área foliar e altura da planta) obtidos para as 360 espécies de árvores utilizadas para as análises de diversidade beta na região costeira do Brasil..... 107

Lista de Tabelas

CAPÍTULO 1 - Sampling effort and the drivers of plant species richness in the Brazilian coastal regions

Table 1. Hypotheses used for the gradients in plant species richness across the coastal zone of Brazil..... 28

Table 2. Summary of statistics comparing hypothesis for explaining plant species richness patterns from the coastal region of Brazil based on an ordinary least-squares (OLS) regression of the full model selected by forward-selection method and involving non-collinear predictors grouped by hypotheses. * = $p < 0.05$ 33

CAPÍTULO 2 - Patterns and drivers of functional and phylogenetic structures of coastal woody vegetation of Eastern South America

Table 1. Six best models ($\Delta AICc < 2$) composed of environmental predictors for the patterns of the functional structure of plants according to the Akaike Information Criterion corrected for sample size ($AICc$). CN: Condition number; CS: Climatic stability; ED: Edaphic; EH: Environmental heterogeneity. PRE-VAR: Historical precipitation change; TEMP-VAR: Historical temperature change; ALT: Altitude; SWC: Soil Water Capacity; NUT: soil nutrient availability. * = $p < 0.01$ 63

Table 2. Three best models ($\Delta AICc < 2$) composed of environmental predictors for the patterns of phylogenetic structure of plants according to the Akaike Information Criterion corrected for sample size ($AICc$). CN: Condition number; WD: Water deficit; CS: Climatic stability; ED: Edaphic; EH: Environmental heterogeneity. RAD: Solar radiation; WBD: Water balance deficit ($PET - AET$); PRE-VAR: Historical precipitation variability; TEMP-VAR: Historical temperature variability; ALT: Altitude; SWC: Soil Water Capacity; NUT: nutrient availability in the soil. * = $p < 0.01$ 64

CAPÍTULO 3 - Diversidade-beta de múltiplos aspectos da diversidade de árvores em ecossistemas costeiros da América do Sul

Tabela 1. Valores de particionamento da diversidade beta taxonômica, filogenética e funcional de assembleias de árvores ao longo da região costeira do Brasil. JTU: valor de substituição medido pelo índice de Jaccard, JNE: valor do aninhamento medido pelo índice de Jaccard, JAC: valor da diversidade beta geral medida pelo índice de Jaccard, SIM: valor de substituição medido pelo índice de Sørensen, SNE: valor de aninhamento medido pelo índice de Sørensen, SOR: valor da diversidade beta geral medida pelo índice de Sørensen..... 102

Tabela 2. Comparações entre modelos lineares múltiplos e seus componentes para os padrões da diversidade beta taxonômica, filogenética e funcional de árvores ao longo da região costeira do Brasil. Valores de slope e R² são padronizados. * = p < 0,05..... 104

A black and white photograph of a coastal landscape. In the foreground, there is a sandy path or clearing surrounded by dense, low-lying vegetation, including various shrubs and grasses. The path leads towards the middle ground, where a dense line of trees, possibly a mangrove or coastal forest, is visible. Beyond the trees, the ocean stretches to the horizon under a sky with scattered clouds. The overall scene is a natural, undisturbed coastal environment.

Introdução Geral

Restinga na Reserva Biológica de Santa Isabel,
litoral norte do Estado de Sergipe (Foto:
Eduardo V. S. Oliveira).

Introdução geral

As questões sobre a organização, dinâmica e causas por trás dos conjuntos de espécies que coexistem no espaço em um dado intervalo de tempo ocupam o campo da ecologia de comunidades. Dentre os problemas fundamentais deste campo estão os mecanismos por trás da quantidade e variedade de organismos, i.e., a diversidade, e de quem são esses organismos, i.e., a composição das comunidades (Cody & Diamond 1972; Morin 1999). O estudo da diversidade envolve sobretudo explicações geométricas (i.e., área), evolutivas (e.g., taxas de diversificação e tempo evolutivo) e ecológicas (e.g., energia, água e heterogeneidade ambiental) (Rosenzweig 1995). Por sua vez, o estudo da composição enfatiza os padrões de similaridade entre espécies dentro das comunidades (ou assembleias), i.e., a estrutura de assembleia, e os padrões de similaridade entre comunidades (Cody & Diamond 1972; Whittaker 1972). Recentemente, essas questões têm sido exploradas – com bastante sucesso e poder de generalização – em grandes escalas espaciais, através da abordagem de múltiplas comunidades ao longo do espaço geográfico ou ‘metacomunidades’ (Gaston 2000; Hillebrand 2004; Leibold et al. 2004).

A distinção entre similaridade dentro de e entre assembleias decorre da percepção do efeito das escalas espaciais envolvidas nos padrões observados. Tradicionalmente, o estudo desses padrões adota diferentes escalas espaciais, referindo-se à diversidade de espécies dentro e entre as assembleias, i.e., componentes alfa (α) e beta (β), respectivamente (Whittaker 1972). A diversidade α diz respeito à quantidade de entidades em uma comunidade ecológica local. A diversidade β , por sua vez, descreve a dissimilaridade entre essas entidades de locais diferentes (Whittaker 1972; Magurran 2004). Ainda, a diversidade β pode ser particionada em dois componentes distintos, ‘turnover’ ou substituição entre espécies e perda de espécies (‘nestedness’ ou aninhamento) entre assembleias (Baselga 2010). Este último aspecto é essencial para

fornecer entendimento de como a composição de espécies varia entre assembleias, em resposta a fatores ecológicos e geográficos (Alahuhta et al. 2017; Soininen et al. 2018).

Evidentemente, as questões sobre a diversidade e composição de comunidades não são independentes entre si. A diversidade pode emergir das regras de montagem de comunidades, i.e., uma limitação na composição pode restringir a diversidade (Weiher & Keddy 2004). Nesse sentido, a diversidade dependerá de fatores que limitam ou favorecem o estabelecimento das espécies. Esses fatores envolvem processos históricos relacionado à origem e distribuição de diferentes linhagens evolutivas, além de fatores ambientais locais como condições abióticas, disponibilidade de recursos e a presença de outras espécies (Ricklefs & Schluter 1993). Dessa forma, o entendimento dos padrões de diversidade e composição de espécies em comunidades tem se beneficiado de abordagens recentes que buscam integrar informações sobre os antecedentes evolutivos das espécies, e suas características ecológicas, as quais estariam relacionadas às adaptações ao contexto abiótico e biótico local (Davies & Buckley 2011; Tilman 2013; Mittelbach & McGill 2019). Perspectivas centradas no uso das relações evolutivas e nas características ecológicas das espécies, em contraste a uma perspectiva centrada na identidade das espécies, permitem uma compreensão mais profunda de como processos evolutivos e ecológicos influenciaram os padrões da diversidade (Purvis & Hector 2000; Pavoine & Bonsall 2011), estabelecendo, portanto, uma relação fundamental entre a composição de espécies resultada de mecanismos de montagem de comunidades e da quantidade de espécies, que reflete a própria diversidade (Tucker et al. 2017).

A investigação dos processos e mecanismos subjacentes aos padrões de diversidade em suas diferentes perspectivas (i.e., taxonômica, filogenética e funcional) é dificultada pela existência de diferentes déficits de conhecimento (Meyer et al. 2016; Freitas et al. 2020). Nesse caso, ao menos quatro tipos de déficits podem ser observados, i) Linneano, relacionado a espécies que ainda não foram descritas, ii)

Wallaceano, relacionado ao desconhecimento da distribuição geográfica das espécies, iii) Darwiniano, relacionado a ausência de filogenias completamente resolvidas ou com comprimentos de ramo desconhecidos e iv) Raunkiaeriano, relacionado a ausência de atributos funcionais das espécies (Diniz-Filho et al. 2009; Hortal et al. 2015). Usualmente, dados de ocorrência são afetados pelos déficits Linneano e Wallaceano (Stropp et al. 2016). Buscando não comprometer a generalidade e a validade das análises, uma alternativa seria incluir apenas as assembleias bem conhecidas, definidas a partir de um critério mínimo de completude de amostragem e de número de registros (e.g., Ballesteros-Mejia et al. 2013; Sousa-Baena et al. 2014; Stropp et al. 2016). Em se tratando dos déficits Darwiniano e Raunkiaeriano, é notória a necessidade de superá-los para melhorar nosso entendimento dos processos que explicam os padrões da diversidade com foco nos contextos evolutivo e ecológico (Diniz-Filho et al. 2013; Assis et al. 2018). Um ponto crítico é a avaliação da consequência desses déficits nas análises, avaliando as tendências associadas e seu poder em afetar a generalidade da inferência (Hortal et al. 2015). A redução do viés gerado pelos déficits Darwiniano e Raunkiaeriano pode envolver o uso de filogenias consensuais e abordagens de imputação baseadas em algoritmos de aprendizado (Diniz-Filho et al. 2013; Penone et al. 2014; Assis et al. 2018).

Diferentes aspectos relacionados aos padrões espaciais da diversidade permanecem pouco explorados em ecossistemas costeiros tropicais, como por exemplo, na Restinga, uma vegetação que ocupa os solos arenosos da Planície Litorânea do Brasil (Rizzini 1997). A Restinga compreende aproximadamente 79% da vegetação costeira do Brasil (Lacerda et al. 1993), abrangendo diferentes fitofisionomias, que variam de acordo com a distância do oceano e do tipo de substrato (Oliveira & Landim 2014). Logo após a zona de preamar, predominam espécies herbáceas psamófilas-reptantes, variando para indivíduos lenhosos de porte mais desenvolvido nas porções mais internas

e protegidas do continente (Araujo & Pereira 2009). Solos arenosos distróficos de habitats costeiros tropicais, como é o caso da Restinga no Brasil, são limitantes para as plantas (Hesp 1991; Cordazzo et al. 2006; Marques et al. 2015). Nestes solos, por causa dos grandes espaços intersticiais entre as partículas do substrato, existe uma baixa capacidade de retenção de água e nutrientes (Cordazzo et al. 2006; Maun 2009). Somado a isso, a perda de umidade evaporativa do solo e o aumento da evapotranspiração pelos ventos, podem condicionar plantas à seca (Cordazzo et al. 2006; Maun 2009; Silva & Souza 2018). Ambos, a escassez de água e a deficiência de nutrientes limitam o estabelecimento e o crescimento de plantas em habitats costeiros, exigindo uma série de mecanismos morfológicos e fisiológicos (Hesp 1991; Cordazzo et al. 2006).

Evidências têm mostrado que a flora de Restinga se originou dos ecossistemas adjacentes de Florestas Tropicais Pluviais (i.e., Floresta Atlântica e Amazônia) e de Florestas Tropicais Sazonalmente Secas (i.e., Caatinga) (Rizzini 1997; Fernandes e Queiroz 2015; Silva et al. 2015), sendo composta principalmente por espécies pertencentes a poucas famílias dominantes, como Fabaceae e Myrtaceae (Scarano 2002). Contudo, uma abordagem abrangente que explique a dinâmica ecológica de formação das comunidades vegetais desses ecossistemas permanece necessária. Devido a sua localização em uma das regiões mais urbanizadas do Brasil, a Restinga está entre os ecossistemas mais ameaçados do país (Marques et al. 2015). A continuidade da ocupação e descaracterização das áreas costeiras pode dificultar, se não impedir, análises biogeográficas futuras dentro dessa formação, o que torna urgente a realização de estudos sobre essa formação, no presente.

Diante do apresentado, neste estudo, buscamos investigar os padrões geográficos de assembleias de Angiospermas presentes na Restinga, considerando diferentes lacunas de informação sobre as espécies. Abordamos diferentes hipóteses sobre os padrões

observados e os processos potencialmente envolvidos na formação dessas assembleias. Avaliamos padrões de estrutura e diversidade, envolvendo aspectos taxonômicos, filogenéticos e funcionais, dentro e entre assembleias ao longo da região costeira brasileira. Especificamente, no **capítulo 1**, avaliamos o nível de conhecimento relativo à amostragem botânica na região e testamos o efeito de preditores macroecológicos sobre a riqueza de espécies. No **capítulo 2** investigamos qual mecanismo de montagem, se filtragem ambiental ou similaridade limitante (i.e., conduzindo a um padrão de agrupamento ou sobredispersão, respectivamente), possui um papel mais importante para os padrões de estrutura funcional e filogenética e avaliamos quais preditores ambientais influenciam os padrões observados. Por fim, no **capítulo 3** avaliamos a contribuição de substituição e do aninhamento, quais processos (i.e., se determinismo ou estocasticidade) e quais preditores ambientais são responsáveis pelos padrões de diversidade β nas três perspectivas.

Referências

- Alahuhta J, Kosten S, Akasaka M, Auderset D, Azzella MM & Bolpagni R et al. 2017. Global variation in the beta diversity of lake macrophytes is driven by environmental heterogeneity rather than latitude. *J Biogeogr* 44: 1758-1769.
- Araujo DSD & Pereira MCA. 2009. Sandy coastal vegetation. In: Del Claro K, Oliveira PS & Rico-Gray V (Eds), *Encyclopedia of Life Support Systems (EOLSS): tropical biology and conservation management – vol. IV*, Oxford: EOLSS-UNESCO, pp. 173-189.
- Assis LCS. 2018. Revisiting the Darwinian shortfall in biodiversity conservation. *Biodivers Conserv* 27: 2859-2875.
- Ballesteros-Mejia L, Kitching IJ, Jetz W, Nagel P & Beck J. 2013. Mapping the biodiversity of tropical insects: species richness and inventory completeness of African sphingid moths. *Glob Ecol Biogeogr* 22: 586-595.
- Baselga A. 2010. Partitioning the turnover and nestedness components of betadiversity. *Glob Ecol Biogeogr* 19: 134-143.

- Cody ML & Diamond JM (Eds.). 1975. Ecology and evolution of communities. Cambridge: Harvard Univ Press, 560 p.
- Cordazzo CV, Paiva JB & Seeliger U. 2006. Guia ilustrado: plantas das dunas da costa sudoeste Atlântica. Pelotas: USEB, 107 p.
- Davies TJ & Buckley LB. 2011. Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Phil Trans R Soc* 366: 2414-2425
- Diniz-Filho JAF, Bini LM, Oliveira G, Barreto BS, Silva MMFP & Terribile LC et al. 2009. Macroecologia, biogeografia e áreas prioritárias para conservação no cerrado. *Oecologia Brasiliensis* 13: 470-497.
- Diniz-Filho JAF, Loyola RD, Raia P, Mooers AO & Bini LM. 2013. Darwinian shortfalls in biodiversity conservation. *Trends Ecol Evol* 28: 689-695.
- Fernandes MF & Queiroz LP. 2015. Floristic surveys of Restinga Forests in southern Bahia, Brazil, reveal the effects of geography on community composition. *Rodriguésia* 66: 51-73.
- Freitas TMDS, Stropp J, Calegari BB, Calatayud J, De Marco Jr P, Montag LFDA & Hortal J. 2021. Quantifying shortfalls in the knowledge on Neotropical Auchenipteridae fishes. *Fish Fish* 22: 87-104.
- Gaston KJ. 2000. Global patterns in biodiversity. *Nature* 405: 220-227.
- Hesp PA. 1991. Ecological processes and plant adaptations on coastal dunes. *J Arid Environ* 21: 165-191.
- Hillebrand H. 2004. On the generality of the latitudinal diversity gradient. *Am Nat* 163: 192-211.
- Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM & Ladle RJ. 2015. Seven shortfalls that beset large-scale knowledge on biodiversity. *Annu Rev Ecol Evol Syst* 46: 523-549.
- Lacerda LD, Araújo DSD & Maciel NC. 1993. Drycoastal ecosystems of the tropical Brazilian coast. In: van der Maarel E (Ed), *Dry Coastal-ecosystems: Africa, Asia, Oceania*, Amsterdam: Elsevier, pp. 477-493.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM & Hoopes MF et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 7: 601-613.
- Magurran AE. 2004. Measuring biological diversity. Oxford: Blackwell Science, 256 p.

Marques MCM, Silva SM & Liebsch D. 2015. Coastal plain forests in southern and southeastern Brazil: ecological drivers, floristic patterns and conservation status. *Braz J Bot* 38: 1-18.

Maun MA. 2009. The biology of coastal sand dunes. New York: Oxford University, 288 p.

Meyer C, Weigelt P & Kreft H. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecol Lett* 19: 992-1006.

Mittelbach GG & McGill BJ. 2019. Community Ecology. Oxford: Oxford University Press, 430 p.

Morin PJ. 1999. Community Ecology. Oxford: Blackwell Science, 424 p.

Oliveira EVS & Landim MF. 2014. Caracterização fitofisionômica das restingas da Reserva Biológica de Santa Isabel, litoral norte de Sergipe. *Scientia Plena* 10:1-10.

Pavoine S & Bonsall MB. 2011. Measuring biodiversity to explain community assembly: A unified approach. *Biol Rev* 86:792-812.

Penone C, Davidson AD, Shoemaker KT, Di Marco M, Rondinini C & Brooks TM et al. 2014. Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods Ecol Evol* 5:961-970.

Purvis A & Hector A. 2000. Getting the measure of biodiversity. *Nature* 405: 212-219.

Ricklefs RE & Schluter D (Eds.). 1993. Species diversity in ecological communities. Chicago: Univ. of Chicago Press, 416 p.

Rizzini AT. 1997. Tratado de Fitogeografia do Brasil: aspectos ecológicos, sociológicos, e florísticos. Rio de Janeiro: Âmbito Cultural, 746 p.

Rosenzweig ML. 1995. Species Diversity in Space and Time. New York: Cambridge University Press, 460 p.

Scarano FR. 2002. Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian Atlantic Rainforest. *Annals of Botany* 90: 517-524.

Silva JLA, Souza AF, Jardim JG & Goto BT. 2015. Community assembly in harsh environments: the prevalence of ecological drift in the heath vegetation of South America. *Ecosphere* 6: 1-18.

Silva KJP & Souza AF. 2018. Common species distribution and environmental determinants in South American coastal plains. *Ecosphere* 9: 1-15.

Soininen J, Heino J & Wang J. 2018. A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. *Glob Ecol Biogeogr* 27: 96-109.

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.

Stropp J, Ladle RJ, Malhado ACM, Hortal J, Gaffuri J, Temperley WH, Skøien JO. 2016. Mapping ignorance: 300 years of collecting flowering plants in Africa. *Glob Ecol Biogeogr* 25: 1085-1096.

Tilman D. 2013. Functional diversity. In: Levi SA (Ed), *Encyclopedia of Biodiversity*, San Diego: Academic Press, pp. 109–120.

Tucker CM, Cadotte MW, Carvalho SB, Davies TJ, Ferrier S & Fritz SA et al. 2017. A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biol Rev* 92: 698-715.

Weiher E & Keddy P. 2004. Assembly rules as general constraints on community. In: Weiher E & Keddy P (Eds), *Ecological assembly rules: perspectives, advances, retreats*, Cambridge: Cambridge University Press, pp. 251-271.

Whittaker RH. 1972. Evolution and measurement of species diversity. *Taxon* 21: 213-251.

CAPÍTULO 1

Sampling effort and the drivers of plant species richness in the Brazilian coastal regions ^[1]

[1] Artigo publicado no periódico Oecologia. Citação: Oliveira EVS, Alves DMC, Landim MF, Gouveia SF. 2021. Sampling effort and the drivers of plant species richness in the Brazilian coastal regions. Oecologia 195: 163-171. *This is a post-peer-review, pre-copy edit version of an article published in Oecologia. The final authenticated version is available online at: <https://doi.org/10.1007/s00442-020-04805-7>.*

Abstract

The causes of the gradients in species richness remain contentious because of multiple competing hypotheses, significant knowledge gaps, and regional effects of environmental and historical factors on species pools. Coastal zones are subject to particular sets of environmental constraints, thus identifying the drivers of species richness therein should shed light on the regional gradients of species diversity. Here, we investigate the geographic patterns and drivers of plant diversity across coastal regions while allowing for pervasive sampling deficiencies. Based on 142,708 records of flowering plant occurrences, we mapped species richness and estimated the level of knowledge across the coastal zone of Brazil. Based on inventory completeness, we used linear regression models to test the predictive power of environmental variables that represent different environmental hypotheses. Few cells (25%) were well-surveyed, reflecting little knowledge about the distribution and diversity of flowering plants on the highly-populated Brazilian coast. Still, we found support for the habitat heterogeneity hypothesis as the best explanation of the variation in species richness of flowering plants in this region. Soil properties and water constraints are also important factors. Although our work emphasises the paucity of information on plant diversity in tropical and human-dominated areas, we show that knowledge limitations should not curb our capability of addressing hypotheses about species diversity.

Keywords: Plant diversity; inventory completeness; macroecology; Wallacean shortfall; coastal plain.

Introduction

At large spatial scales, both ecological and historical processes drive the variation in species diversity (Ricklefs 2004; Condamine et al. 2012). However, after innumerable

hypotheses, no single factor appears to explain the diversity patterns for all biological groups, geographic regions, and scales of analyses (Willig et al. 2003). Recently, studies that evidence the context-dependency of the relationships between species diversity and environmental conditions have increased (Gouveia et al. 2013; Alves et al. 2018). In addition to improving our understanding about diversity gradients, these efforts also can reveal some particularities that result from the specific systems' features (i.e., environmental context, taxon, and scale) that deviates from the general trend and enable predictions at the regional scale (Willig et al. 2003; Gouveia et al. 2013).

For plants, different ecological hypotheses for the geographical gradients of species richness have been offered, especially to account for the latitudinal gradients of diversity (e.g., Currie 1991; Rohde 1992; Kreft and Jetz 2007). Most explanations are related to the environmental role of water and energy (e.g., Francis and Currie 2003; Hawkins et al. 2003; Kreft and Jetz 2007; Li et al. 2013), but also include environmental productivity (e.g., Currie 1991) and habitat heterogeneity (e.g., Yang et al. 2013). Nevertheless, the explanatory power of these hypotheses also varies considerably over geographic space (Eiserhardt et al. 2011), such as the case of water-related hypotheses that prevail in low latitudes, where energy availability is high (Hawkins et al. 2003).

The longstanding debate over the geographical gradients of diversity is plagued with persistent gaps of knowledge, even for conspicuous organisms such as plants (Meyer et al. 2016). The critical knowledge deficits regard to what species an organism belongs, and to where its species occur, which are known as Linnaean and Wallacean shortfalls, respectively (Hortal et al. 2015). Obtaining a complete knowledge of species diversity in a given locality is challenging because few inventories can account for all species, and studies often vary in sampling efforts (Magurran 2004). In this sense, data on regional-to-global patterns biodiversity are generally incomplete, compromising our

understanding of patterns and its causes (Hortal et al. 2015; Yang et al. 2013), which is further exacerbated by the existence of spatial biases in sampling effort (Oliveira et al. 2016).

An alternative to investigate biodiversity patterns based on limited information on the occurrence of species is to use the best-sampled and representative datasets possible. Techniques based on rarefaction curves that indicate the expected level of completeness of species richness (e.g., Sousa-Baena et al. 2014; Stropp et al. 2016) can provide a criterion of the minimum level of knowledge necessary to evaluate the quality of the inventories. Thereby, better inferences about the environmental drivers of targeted regional diversity are possible (e.g., Yang et al. 2013; Oliveira et al. 2016).

Tropical coastal zones harbour unique ecosystems that are affected by both continental and marine factors, including salt spray, low nutrients, intense winds, solar radiation, substrate mobility, and drought (Barbour 1992). Consequently, many species in these ecosystems possess physiological and morphological adaptations to these factors, such as the development of specialised stems, succulence, and salt accumulation in the leaves (Lüttge 2008). On the Eastern coast of South America, the Restingas comprises a plant formation that develops on sandy deposits parallel to coastlines (Rizzini 1997), and under the influence of soil characteristics (Bonilha et al. 2012). The Restingas comprise the vast majority of species and areas across the Brazilian coastal zone (Lacerda et al. 1993) and occur between the sea and tropical forest ecosystems (e.g., Atlantic and Amazon rainforest), savanna (e.g., the Cerrado), or the Caatinga seasonally dry forest (Rizzini 1997).

In Brazil, the coastal zone harbour ca. 70% of the country's human population and most of its largest cities (Moraes 2007). This high human density also means higher accessibility and shorter distances from research institutes, which in turn is expected to make these regions better known regarding their biodiversity, despite the higher impacts

(Hortal et al. 2015). Therefore, a high degree of knowledge about the flora in the Brazilian coastal zone is expected. In this study, we assembled the most extensive, curated, and standardised dataset of angiosperms from the Brazilian coastal vegetation to date. First, we evaluated the level of knowledge of its floristic diversity and assessed the spatial patterns of species richness. Then, we investigated the environmental drivers of the species diversity of herbs and trees (hereafter referred to as plants) across the entire region. We tested the influence of predictors related to water, energy, habitat heterogeneity, human influence, and soil properties. Due to the particularities of tropical coastal environments outlined above, together with the characteristics of the plants from these environments, we expected factors related to water and soil to be better predictors of the variation in the species richness across the region. Furthermore, because of the history of human presence across the Brazilian coastal zone, we expect a negative effect of human influence. However, we also expect an association between human influence and topography, with flatter areas linked to higher human activity (Sandel and Svenning 2013; Odgaard et al. 2014).

Materials and methods

Study area and occurrence data

To define our study area, we overlapped the limits of the Brazilian Coastal Plain with a vegetation map that included the Restingas (IBGE 2017; MMA 2017). This overlap was necessary because the contour map of the coastal plain was incomplete in many locations, while the map of the coastal vegetation did not cover the entire coastal plain. The coastal plains also included the mangroves. However, these ecosystems have very different environmental characteristics, contain species with particular adaptations, which result in a disparate species composition and richness and that are not expected to be explained by the hypotheses raised in this study (Kjerfve and Lacerda 1993).

Therefore, we removed mangroves from our study (ICMBio 2018). We then built a cells' grid with 55×55 km of resolution ($0.5^\circ \times 0.5^\circ$ of latitude-longitude at the Equator) onto our contour map, using SAM software (Spatial Analysis in Macroecology) (Fig. S1) (Rangel et al. 2010).

We obtained the records of vascular plants of the study area (i.e., contour map) from the BIEN 4.1 database (Botanical Information and Ecology Network, access on 04-Oct-2019). This database has standardised information on botanical records, with the correction of taxonomic, spatial, and temporal biases (Enquist et al. 2016). We retrieved these records with the help of the BIEN package (Enquist et al. 2016; Maitner et al. 2017) in R software (R Development Core Team 2018). We removed records of cultivated and non-native plants (BFG 2018), retaining only records with valid binomial names of native species (for a similar approach, see Ballesteros-Mejia et al. 2013). Then, we selected only the flowering plants and classified the species with relation the life forms, retaining herbs and trees as representative of the vegetation of the Restingas. Others life forms, such as bamboos, epiphytes, lianas, and vines, were excluded. This decision was taken because these life forms, as well as nonflowering plants, are expected to differ in their response to environmental factors (Mao et al. 2013; Lü et al. 2018). The steps described above were performed using the 'flora' package in R (R Development Core Team 2018; Carvalho 2020).

Data preparation

To investigate the effect of environmental predictors on the pattern of plant species richness, we first selected 25 candidate environmental predictors (e.g., Currie 1991; Francis and Currie 2003; Kreft and Jetz 2007; Li et al. 2013; Yang et al. 2013). We grouped these predictors into different models that represented general ecological hypotheses (Table 1). *Water hypothesis* was characterised by annual precipitation,

precipitation seasonality, precipitation of driest month, precipitation of wettest month (Fick and Hijmans 2017), air relative humidity (New et al. 1999), actual evapotranspiration, potential evapotranspiration, Priestley-Taylor Alpha Coefficient (i.e., water availability for vegetation; Trabucco and Zomer 2019), and wind speed (CEPEL/UFRJ 2017). *Energy hypothesis* was described by annual mean temperature, mean diurnal range, temperature seasonality, the maximum temperature of the warmest month, minimum temperature of the coldest month and solar radiation (Fick and Hijmans 2017). *Habitat heterogeneity hypothesis* was expressed by the altitude and terrain slope (Fischer et al. 2008). *Human impact hypothesis* was described by the Human Footprint index (WCS/CIESIN 2002). Finally, the *edaphic hypothesis* was described by carbon density, nitrogen density, available water capacity (Global Soil Data Task Group 2000), nutrient availability, rooting conditions, toxicity, and excess salts of the soil (Fischer et al. 2008). Nutrient availability is related to different characteristics as soil texture, soil organic carbon, soil pH, and total exchangeable bases. In turn, rooting conditions considers the soil textures, bulk density, coarse fragments, and soil depth and soil volume (Fischer et al. 2008).

Table 1 Hypotheses used for the gradients in plant species richness across the coastal zone of Brazil.

Hypothesis	Rationale	Reference
Water	Higher levels of environmental water permit the coexistence of more species.	Francis and Currie (2003) and Hawkins et al. (2003)
Energy	Higher energy supplies permit more resource, permitting the coexistence of	Currie (1991) and Tilman (1994)

Hypothesis	Rationale	Reference
	more species.	
Habitat Heterogeneity	Spatially complex habitats provide more niches, permitting the coexistence of more species.	Currie (1991)
Human impact	Human footprint promotes species loss, mainly due to habitat destruction.	McKinney (2008)
Edaphic	In nutrient-poor regions, higher nutrient availability favours the presence of more plant species.	Huston (1980) and Medinski et al. (2010)

Within each predictor group (i.e., hypothesis), we assessed multicollinearity among the containing variables. For that, we ran a bivariate correlation matrix based on Pearson's r coefficient, removing collinear factors (considering an $r \geq 0.7$ with each other). We kept those factors more correlated with species richness (for a similar approach, see Irl et al. 2015). After removing multicollinearity, we ended up with 14 predictors, being four representatives of the water hypothesis, four of the edaphic hypothesis, three of the energy hypothesis, two of the habitat heterogeneity hypothesis, and the human footprint, which stood for the human impact hypothesis.

For each grid cell, we calculated the number of records, the observed and expected richness, and the inventory completeness. To estimate the expected richness and the inventory completeness, we employed the approach proposed by Lobo et al. (2018), which adapts to individual grid cells the method of Ugland et al. (2003), tailored for entire regions. In the Lobo et al.' (2018) procedure, each collection record is treated as a sample unit of each cell (for a similar approach, see Pelayo-Villamil et al. 2018).

The expected species richness for each grid cell is given by the rational adjustment function (Ratkowski 1990). Next, we calculated the completeness, given as a percentage, based on the ratio between observed and predicted richness of each grid cell (Sousa-Baena et al. 2014). All steps described were performed with the package KnowBr (Lobo et al. 2018) in the R software (R Development Core Team 2018). Then, we obtained a geographic distribution of the well-surveyed cells. To be retained as well-surveyed cells, the cells had to fulfil two criteria: a completeness $\geq 50\%$ and a number of records $\geq$ mean of the number of records among all cells. The first criterion is considered sufficient in the definition of well-known cells (Sousa-Baena et al. 2014; Stropp et al. 2016; Lobo et al. 2018). The second one was employed because high values of completeness may result from a low sample size. Thus, the definition of a minimum sampling effort (given here as the number of records) can minimise this effect (Sousa-Baena et al. 2014; for alternative approaches, see Ballesteros-Mejia et al. 2013; Stropp et al. 2016).

Data analysis

To test the effect of environmental predictors on the geographical pattern of species richness, we included in the analyses only the well-surveyed cells. We first performed an ordinary least-squares (OLS) regression with all non-collinear predictors in a full model employing a forward selection procedure (Borcard et al. 2011). We then retained the relevant predictors from the full model (i.e., with significant effect), grouped them according to the hypotheses, and performed partial regression analyses to partition the variance of their effects on plant richness (Legendre and Legendre 2012). We tested the regression residuals for data normality, using the Shapiro-Wilk test, and for the presence of the spatial autocorrelation through spatial correlograms of Moran's *I* index

(see Legendre and Legendre 2012). We performed spatial analyses with the help of the SAM software (Rangel et al. 2010).

Results

Richness pattern and knowledge gaps

We obtained 142708 valid records of flowering plants along the Brazilian coastal zone, which were distributed in 180 of 202 grid cells. This dataset encompassed 7256 species (average = 240.3 species/cell). From the 194 families recorded in the database, Fabaceae was the more speciose one, with 708 species (Fig. S2). *Tapirira guianensis* Aubl. (Anacardiaceae) was the species more prevalent, being present in 95 cells. Only 57 species (about 1.0%) occurred in more than 50 cells, whereas 2200 species (30%) were registered at a single cell.

Regarding the sampling effort across the study region, we found a substantial asymmetry in the number of records among cells, with cells ranging from zero to 6474 records (average = 668.4 records/cell $\pm$ 1092.6 *sd*; Fig. 1a). Most cells (55%) had less than 250 records. Cells with a higher number of records (> 3000) were at the border of the Atlantic rainforest. In contrast, cells with none or fewer records were adjacent to the Cerrado (Central Brazilian Savanna), the Caatinga (Seasonal Dry Forest), and the Pampas (Lowland grasslands) ecoregions (Fig. 1a). Regarding the expected pattern of species richness, we estimated an average richness of 801.4 species/cell. For the regression analyses, we retained only 51 (25%) of the 202 grid cells. These were defined as the well-surveyed cells regarding botanical sampling, as given by the inventory completeness and the records number. We found a similar pattern between well-surveyed cells and those with a higher number of records (Fig. 1).

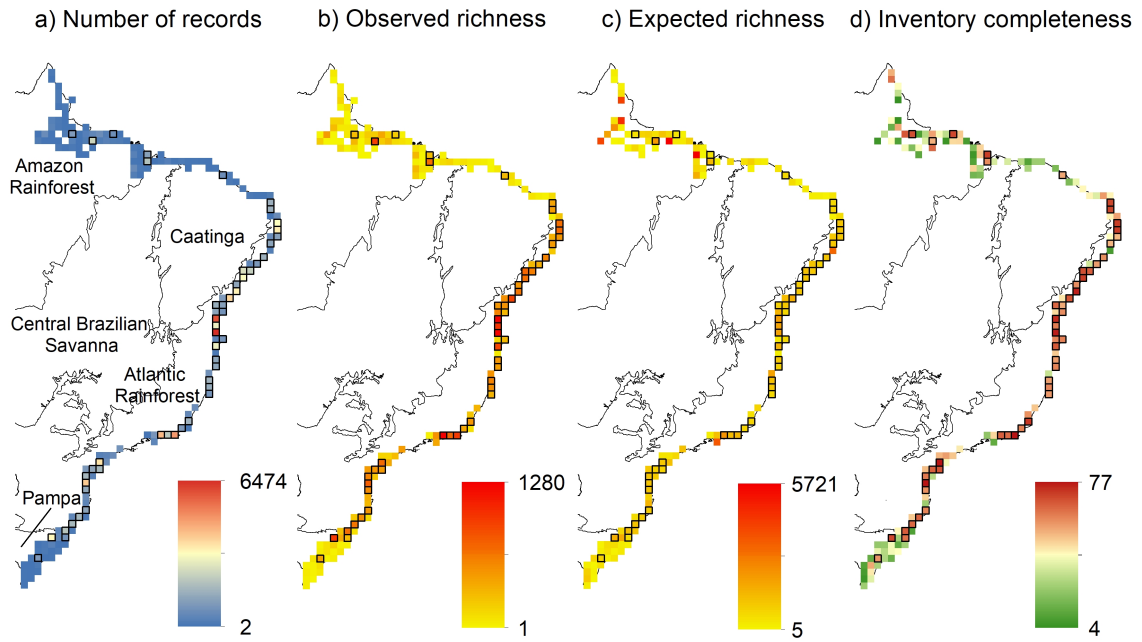


Fig. 1 Maps of a) distribution of the number of records, b) observed richness, c) expected richness, and d) inventory completeness for vascular plants in the Brazilian coastal vegetation. Cells with the bolder border have higher knowledge (completeness $\geq$ 50% and records number $\geq$ mean). Brazilian phytogeographic domains are delimited in gray. Grid cells have a 0.5° spatial resolution.

Ecological drivers of the plant richness

We found no spatial autocorrelation in the residuals of the regression analyses, thus exempting us from the need to employ spatial regression analyses (Fig. S3). The full linear model retained five explanatory variables for the pattern of species richness ($R^2 = 0.51$; Table 2), which were representative of three hypotheses (habitat heterogeneity, edaphic and water). In this model, excepting wind speed, all factors related positively to species richness, with terrain slope being the most important predictor, followed by altitude, nutrient availability and available water capacity of the soil, and wind speed (Table 2). In the partial analysis that partitioned the variance of the effect of these predictors arranged by individual hypothesis, habitat heterogeneity had the most

substantial effect (partial $R^2 = 0.389$), followed by edaphic factors (partial $R^2 = 0.182$), and water (partial $R^2 = 0.032$; Fig. 2). We found a low overlap of contribution between the predictors (Fig. 2). Explanatory variables representative of the hypotheses of energy and human impact had no significant effect on plant richness.

Table 2 Summary of statistics comparing hypothesis for explaining plant species richness patterns from the coastal region of Brazil based on an ordinary least-squares (OLS) regression of the full model selected by forward-selection method and involving non-collinear predictors grouped by hypotheses. * = $p < 0.05$.

Hypothesis	Predictor	Slope	R^2
Habitat heterogeneity	Terrain slope	0.51*	0.51*
	Altitude	0.39*	
Edaphic	Nutrient availability	0.36*	
	Available water capacity	0.31*	
Water	Wind speed	-0.19*	

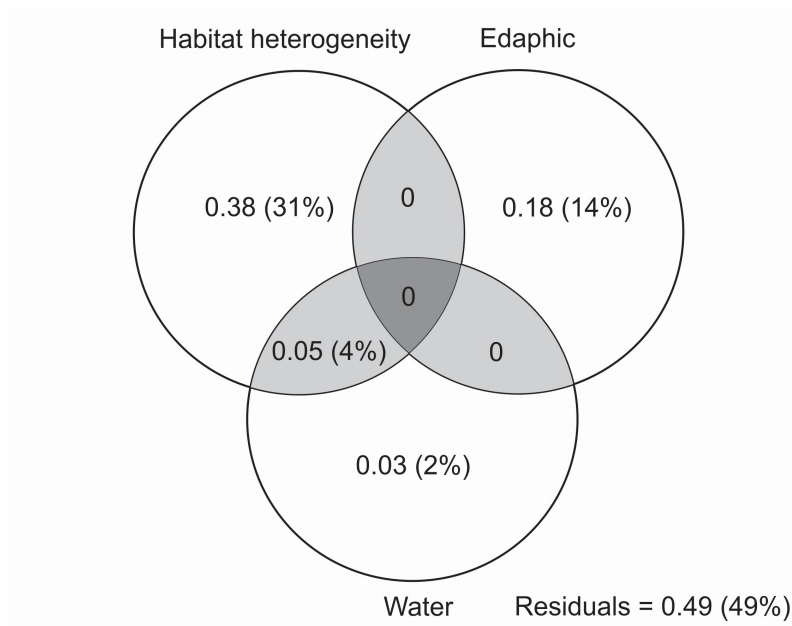


Fig. 2 Comparisons of the effects, grouped by hypotheses, of the environmental predictors on plant richness from the Brazilian coastal region based on partial regression coefficients. Results include portions uniquely accounted by ecological hypothesis, shared contributions and unexplained variance (residuals). The sum of partial coefficients exceeds 1.0 because of some degree of collinearity among predictors, which results in negative variances in the shared fraction of variance explained (omitted).

Discussion

Although the coastal region of Brazil harbours most of the Brazilian population and research centres, pervasive knowledge shortfalls persist for a biological group as conspicuous as flowering plants. No single cell had a knowledge level of species richness of 80% or above. Most of the well-surveyed cells had 50 to 70% of completeness. Variations in the levels of biological knowledge between regions are usually related to factors such as accessibility, landscape attractiveness, colonisation history (e.g., past samplings), and availability of research infrastructure (Ballesteros-Mejia et al. 2013; Oliveira et al. 2016; Stropp et al. 2016). This finding does not mean,

however, that the coastal region is less studied than other regions of the Tropics. Instead, this finding suggests that knowledge shortfalls of biodiversity appear to be a norm throughout the Tropics. If we accept an association between accessibility and knowledge of biodiversity, we are led to conclude that this region is actually better known than others. This appears to be the case if we look at the coastal region itself. The northern portion (i.e., associated with the Caatinga and Amazon regions), which is less accessible and has fewer research institutions, are less surveyed than the coastal region related to the Atlantic Rainforest.

Although the observed low level of knowledge about species richness may hinder a full account of its drivers, novel analytical tools allow drawing inference from the currently-known pattern, as long as a minimum set of criteria of inventory intensity is met (Sousa-Baena et al. 2014). By applying these techniques, and contrary to our predictions, we found support for the habitat heterogeneity hypothesis as the best explanation of the species richness from the Brazilian coastal region. The topography is known to be an essential driver of plant richness by increasing the number of habitat types (Lobo et al. 2001; Yang et al. 2013; Irl et al. 2015), which allows greater specialisation, reduced competition, hence enabling the establishment of different and more species (Tilman 1994; Cramer and Willig 2005). In the coastal plains, a single grid cell may contain dune fields, interdunes, flat terrain, beach ridges, water courses, lagoons and ecotones, all of which can contribute to increased habitat heterogeneity, hence different species combinations and higher species richness (Araujo and Pereira 2009).

The topography is also linked to the spatial pattern of human occupation, with more topographically complex and less disturbed areas acting as a refuge for plants (Sandel and Svenning 2013; Odgaard et al. 2014; Irl et al. 2015). As such, the richness–topography association could be thought of as indicative of human influence. The

Brazilian coastal plains do encompass large tracts of urban areas, agriculture, and livestock concentrated in the flatter and species-depleted Restingas (Lacerda et al. 1993; Araujo e Pereira 2009; Marques et al. 2015). However, the non-significant effect of the human footprint on plant richness lends no support to this claim, suggesting that the habitat diversity associated with the terrain heterogeneity is the fundamental factor driving the variation in species richness in the region.

Surprisingly, edaphic factors, namely the nutrient availability and water capacity of the soil, had only a secondary influence on the spatial pattern of plant richness; exhibiting a positive effect, as expected. Nutrient concentration is purported as an essential driver of plant diversity in nutrient-depleted regions (Huston 1980; Medinski et al. 2010), which evidences the harshness of these sandy coastal plains for species establishment. In this formation, the lower availability of soil nutrients requires from plants to possess adaptations for resource allocation, nutrients redistribution, symbiosis with endomycorrhizal fungi and cyanobacteria for assimilating atmospheric nitrogen (Hesp 1991; Marques et al. 2015). Besides, the Restingas are recognised as an edaphic climax, i.e., vegetation in which development, distribution and richness are strictly related to soil properties (Magnago et al. 2012; Marques et al. 2015; Silva and Souza 2018a).

The minor, yet significant, effect of factors linked to water was also surprising. The only factor involved in this group of predictors was wind speed. Nonetheless, this wind effect together with that of the soil's water capacity suggests that the importance of water constraints on coastal plant richness may not be negligible. Wind accelerates evapotranspiration from plants; and the Restingas occurs on wind-exposed, sandy soils with low capacity for water retention (Marques et al. 2015; Silva and Souza 2018a). Given the critical role of water to plant metabolism, its influence on species establishment and richness should thus be expressed not through the environmental

water input, such as precipitation or air humidity (Francis and Currie 2003; Hawkins et al. 2003; Kreft and Jetz 2007), but rather through alternative pathways, such as the water removal through soil runoff and wind-driven desiccation. Ultimately, it is the water balance (input minus output) that will determine the water available for plants. Those water-removing processes require plants to possess adaptations in morphology, physiology, and reproductive and life cycles to colonise such environments (Maun 2009). These adaptations include small leaves to control transpiration, fewer stomata or higher stomatal concentration on the abaxial leaf surface, leaf folding, abundant epidermal hair, high cutinisation, succulence, deep roots system and C4 or CAM photosynthetic system (Hesp 1991; Rizzini 1997). Similar strategies exist in plants from other water-depleted tropical environments, such as the Caatinga dry forest from northeastern South America (e.g., Silva e Souza 2018b).

Still, the predictors included in the analyses do not explain all variation in species richness. The unexplained proportion of the plant richness may be related to other factors, such as missing environmental and historical factors. Recent studies have evidenced the influence of neighbour formations on the composition of species from coastal regions of Brazil through dispersal (Rizzini 1997; Fernandes and Queiroz 2015; Silva and Souza 2018a). This reasoning led the authors to conclude that the Restingas do not constitute a floristic group distinct from the adjacent phytogeographic domains, such as Caatinga and Atlantic Rainforest, but instead to be a derived formation from these adjacent domains (Silva and Souza 2018b; Cantidio and Souza 2019). This role played by dispersion into the coastal plant communities finds support in the low number of endemic species (Lacerda et al. 1993; Fernandes and Queiroz 2015). The Restingas occurs at geologically new substrates. Therefore, recent colonisation and short time for speciation account for this low level of endemism, and thus to the overall pattern of species richness (Scarano 2002). Based on our findings, however, we argue that this

biogeographical effect alone cannot explain the pattern of species richness. Since parental vegetal formations also respond to most of the environmental factors investigated here, the coastal region appears to subsume these environmental effects, together with its biogeographic history and ecological singularity. Further investigation regarding the interplay between biogeographical and ecological factors influencing the coastal vegetation is desirable, yet the carrying out of extensive surveys in this sensitive ecosystem is nonetheless mandatory.

In summary, we have found an overall low level of botanical knowledge for the Brazilian coastal vegetation, despite the high concentration of research centres in this region. Our results reinforce the need for improving sampling efforts and reducing knowledge gaps, including in the less sampled regions of the northern portion of the coast. Still, through using techniques that allow us to circumvent such inventory deficiencies, we identified a leading effect of the habitat heterogeneity, followed by soil properties and water balance as primary drivers of the geographic variation of plant richness along the coastal zone of Brazil. Despite intensely populated, the influence of the human footprint exerts no significant influence on the broad-scale pattern of species richness analysed, likely a consequence of the broad scale considered. Overall, these findings highlight the predominant role of structural factors of this region (i.e., topography and soil properties), also mediated by water constraints, as likely determinants of the local assembling of species that derive from adjacent ecosystems. Finally, this study underscores the importance of investigating specific systems, with their particular environmental context, as a means to gain an understanding of the drivers and their effects on plant species diversity; and that such effort should not be discouraged by existing data limitations.

References

- Alves DMCC, Diniz-Filho JAF, Souza KS, Gouveia FS, Villalobos F (2018) Geographic variation in the relationship between large-scale environmental determinants and bat species richness. *Basic Appl Ecol* 27:1-8. <https://doi.org/10.1016/j.baae.2017.12.002>
- Araujo DSD, Pereira MCA (2009) Sandy coastal vegetation. In: Del Claro K, Oliveira PS, Rico-Gray V (eds) *Encyclopedia of Life Support Systems (EOLSS): tropical biology and conservation management – vol. IV*. EOLSS-UNESCO PP, Oxford, pp 173-189.
- Ballesteros-Mejia L, Kitching IJ, Jetz W, Nagel P, Beck J (2013) Mapping the biodiversity of tropical insects: species richness and inventory completeness of African Sphingid moths. *Glob Ecol Biogeogr* 22:586-595. <https://doi.org/10.1111/geb.12039>
- Barbour MG (1992) Life at the leading edge: the beach plant syndrome. In: Seeliger U (ed) *Coastal plant communities of Latin America*. Academic Press, San Diego, pp 291-308.
- BFG - The Brazil Flora Group (2018) Brazilian Flora 2020: Innovation and collaboration to meet Target 1 of the Global Strategy for Plant Conservation (GSPC). *Rodriguésia* 69:1513-1527.
- Bonilha RM, Casagrande JC, Soares MR, Reis-Duarte RM (2012) Characterization of the soil fertility and root system of restinga forests. *Rev Bras Ciênc Solo* 36:1804-1813. <https://doi.org/10.1590/s0100-06832012000600014>
- Borcard D, Gillet F, Legendre P (2011) *Numerical ecology with R, Use R*. Springer Science, New York. <https://doi.org/10.1007/978-1-4419-7976-6>

- 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. <https://doi.org/10.1111/ecog.04564>
- Carvalho G (2020) Tools for Interacting with the Brazilian Flora 2020. Available: <http://www.github.com/gustavobio/flora>. Accessed 11 Mai 2020.
- CEPEL/UFRJ - Electric Energy Research Center/Federal University of Rio de Janeiro (2017) Wind Energy in Brazil. Available: <http://www.cepel.br>. Accessed 21 April 2017.
- Condamine FL, Sperling FAH, Wahlberg N, Rasplus J-Y, Kergoat GJ (2012) What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol Lett* 15:267-277. <http://dx.doi.org/10.1111/j.1461-0248.2011.01737>.
- Cramer MJ, Willig MR (2005) Habitat Heterogeneity, Species diversity and null models. *Oikos* 108:209-218. <https://doi.org/10.1111/j.0030-1299.2005.12944.x>
- Currie, D (1991) Energy and large-scale patterns of animal and plant-species richness. *Am Nat* 137: 27-49. <https://doi.org/10.1086/285144>
- Eiserhardt WL, Bjorholm S, Svenning JC, Rangel TF, Balslev H (2011) Testing the water-energy theory on American Palms (Arecaceae) using geo-graphically weighted regression. *Plos One* 6:1-8. <https://doi.org/10.1371/journal.pone.0027027>
- Enquist BJ, Condit R, Peet RK, Schildhauer M, Thiers BM (2016) Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. *PeerJ Preprints* 12:1-33. <https://doi.org/10.7287/peerj.preprints.2615v2>

- Fernandes MF, Queiroz LP (2015) Floristic surveys of Restinga Forests in southern Bahia, Brazil, reveal the effects of geography on community composition. *Rodriguésia* 66:51-73. <http://dx.doi.org/10.1590/2175-7860201566104>
- Fick SE, Hijmans RJ (2017) Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302-4315. <https://doi.org/10.1002/joc.5086>
- Fischer G, Nachtergaele F, Prieler S, van Velthuisen HT, Verelst L, Wiberg D (2008) Global agro-ecological zones assessment for agriculture. Available: <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/>. Accessed 21 November 2018.
- Francis AP, Currie DJ (2003) A globally consistent richness-climate relationship for angiosperms. *Am Nat* 161:523-536. <https://doi.org/10.1086/368223>
- Global Soil Data Task Group (2000) Global Gridded Surfaces of Selected Soil Characteristics. Available: <https://webmap.ornl.gov/ogcdown/index.jsp>. Accessed 21 November 2017.
- Gouveia SF, Hortal J, Cassemiro FAZ, Rangel TF, Diniz-Filho JAF (2013) Nonstationary effects of productivity, seasonality, and historical climate changes on global amphibian diversity. *Ecography* 36:104-113. <https://doi.org/10.1111/j.1600-0587.2012.07553.x>
- Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, Kaufman DM, Kerr JT, Mittelbach GG, Oberdorff T, O'Brien EM, Porter EE, Turner JRG (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84:3105-3117. <https://doi.org/10.1890/03-8006>
- Hesp PA (1991) Ecological processes and plant adaptations on coastal dunes. *J Arid Environ* 21:165-191. [https://doi.org/10.1016/S0140-1963\(18\)30681-5](https://doi.org/10.1016/S0140-1963(18)30681-5)

- Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ (2015) Seven shortfalls that beset large-scale knowledge on biodiversity. *Annu Rev Ecol Evol Syst* 46:523-549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
- Huston M (1980) Soil nutrients and tree species richness in Costa Rican forests. *J Biogeogr* 7:147-157. <https://doi.org/10.2307/2844707>
- IBGE - Instituto Brasileiro de Geografia e Estatística (2017) Bases de dados. Available: <ftp://geofp.ibge.gov.br>. Accessed 17 April 2017.
- ICMBio - Instituto Chico Mendes de Conservação da Biodiversidade (2018) Atlas dos Manguezais do Brasil. ICMBio, Brasília.
- Irl SDH, Harter DEV, Steinbauer MJ, Puyol DG, Fernández-Palacios JM, Jentsch A, Beierkuhnlein C (2015) Climate vs. topography – spatial patterns of plant species diversity and endemism on a high-elevation island. *J Ecol* 103:1621-1633. <https://doi.org/10.1111/1365-2745.12463>
- Kjerfve B, Lacerda LD (1993) Mangroves of Brazil. In: Lacerda LD, Diop S (eds) *Conservation and Sustainable Utilization of Mangroves in Latin America and Africa Regions*. International Society for Mangrove Ecosystems, Okinawa, pp 245-272.
- Kreft H, Jetz W (2007) Global patterns and determinants of vascular plant diversity. *PNAS* 104:5925-5930. <https://doi.org/10.1073/pnas.0608361104>
- Lacerda LD, Araújo DSD, Maciel NC (1993) Drycoastal ecosystems of the tropical Brazilian coast. In: van der Maarel E (ed) *Dry Coastal-ecosystems: Africa, Asia, Oceania*. Elsevier, Amsterdam, pp 477-493.
- Legendre P, Legendre L (2012) *Numerical ecology*. Elsevier, Amsterdam.
- Li L, Wang Z, Zerbe S, Abdusalih N, Tang Z, Ma M, Yin L, Mohammad A, Han W, Fang J (2013) Species Richness patterns and water-energy dynamics in the

drylands of Northwest China. Plos One 8:1-10.

<https://doi.org/10.1371/journal.pone.0066450>

Lobo JM, Castro I, Moreno JC (2001) Spatial and environmental determinants of vascular plant species richness distribution in the Iberian Peninsula and Balearic Islands. *Biol J Linn Soc* 73:233-253. <https://doi.org/10.1006/bijl.2001.0543>

Lobo JM, Hortal J, Yela JL, Millán A, Sánchez-Fernández D, García-Roselló E, González-Dacosta J, Heine J, González-Vilas L, Guisande C (2018) KnowBR: An application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecol Indic* 91:241-248. <https://doi.org/10.1016/j.ecolind.2018.03.077>

Lü L, Cai H, Yang Y, Wang Z, Zeng H (2018) Geographic patterns and environmental determinants of gymnosperm species diversity in China. *Biodivers Sci* 26:1133-1146. <https://doi.org/10.17520/biods.2018098>

Lüttge U (2008) *Physiological ecology of tropical plants*, 2nd edn. Springer-Verlag, Berlin-Heidelberg.

Magnago LFS, Martins SV, Schaefer CEGR, Neri AV (2012) Restinga forests of the Brazilian coast: richness and abundance of tree species on different soils. *An Acad Bras Ciênc* 84:807-822. <https://doi.org/10.1590/S0001-37652012000300023>

Magurran AE (2004) *Measuring biological diversity*, 1st edn. Blackwell Science, Oxford.

Maitner BS, Boyle B, Casler N, Condit R, Donoghue J, Durán SM, Guaderrama D, Hinchliff CE, Jørgensen PM, Kraft NJB, McGill B, Merow C, Morueta-Holme N, Peet RK, Sandel B, Schildhauer M, Smith SA, Svenning J-C, Thiers B, Violle C, Wiser S, Enquist BJ (2017) The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol Evol* 9:1-7. <https://doi.org/10.1111/2041-210X.12861>

- Mao LF, Chen S Bin, Zhang JL, Hou Y-H, Zhou G-S, Zhang X-S (2013) Vascular plant diversity on the roof of the world: Spatial patterns and environmental determinants. *J Syst Evol* 51:371-381. <https://doi.org/10.1111/j.1759-6831.2012.00240.x>
- Marques MCM, Silva SM, Liebsch D (2015) Coastal plain forests in southern and southeastern Brazil: ecological drivers, floristic patterns and conservation status. *Braz J Bot* 38:1-18. <https://doi.org/10.1007/s40415-015-0132-3>
- Maun MA (2009) The biology of coastal sand dunes, 1st edn. Oxford University, New York.
- McKinney ML (2008) Effects of urbanization on species richness: a review of plants and animals. *Urban Ecosyst* 11:161-176. <https://doi.org/10.1007/s11252-007-0045-4>
- Medinski TV, Mills AJ, Esler KJ, Schmiedel U, Jürgens N (2010) Do soil properties constrain species richness? Insights from boundary line analysis across several biomes in south western Africa. *J Arid Environ* 74:1052-1060. <http://dx.doi.org/10.1016/j.jaridenv.2010.03.004>
- Meyer C, Weigelt P, Kreft H (2016) Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecol Lett* 19:992-1006. <https://doi.org/10.1111/ele.12624>
- MMA - Ministério do Meio Ambiente (2017) i3Geo. Available: <http://mapas.mma.gov.br/i3geo/datadownload.htm>. Accessed 17 April 2017.
- Moraes CAR (2007) Contribuição para a gestão da zona costeira do Brasil: elementos para uma geografia do litoral brasileiro, 1st edn. Annablume, São Paulo.
- New M, Hulme M, Jones PD (1999) Representing 20th century space-time climate variability. Part I: Development of a 1961-1990 mean monthly terrestrial

- climatology. J Clim 12:829-856. [https://doi.org/10.1175/1520-0442\(1999\)012<0829:RTCSTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2)
- Odgaard M V., Bøcher PK, Dalgaard T, Moeslund JE, Svenning JC (2014) Human-driven topographic effects on the distribution of forest in a flat, lowland agricultural region. J Geogr Sci 24:76-92. <https://doi.org/10.1007/s11442-014-1074-6>
- Oliveira U, Paglia AP, Brescovit AD, Carvalho CJB, Silva DP, Rezende DT, Leite FSF, Batista JAN, Barbosa JPPP, Stehmann JR, Ascher JS, Vasconcelos MF, De Marco Jr P, Löwenberg-Neto P, Dias PG, Ferro VG, Santos AJ (2016) The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. Divers Distrib 22:1-13. <https://doi.org/10.1111/ddi.12489>
- Pelayo-Villamil P, Guisande C, Manjarrés-Hernández A, Jiménez LF, Granado-Lorencio C, García-Roselló E, González-Dacosta J, Heine J, González-Vilas L, Lobo JM (2018) Completeness of national freshwater fish species inventories around the world. Biodivers Conserv 27:3807-3817. <https://doi.org/10.1007/s10531-018-1630-y>
- R Development Core Team (2018) R: A language and environment for statistical computing. Available: <http://www.R-project.org>. Accessed 06 September 2018.
- Rangel TFLVB, Diniz-Filho JAF, Bini LM (2010) SAM: a comprehensive application for Spatial Analysis in Macroecology. Ecography 33:46-50. <https://doi.org/10.1111/j.1600-0587.2009.06299.x>
- Ratkowski DA (1990) Handbook of nonlinear regression models, 1st edn. Marcel Dekker, New York.
- Ricklefs RE (2004) A comprehensive framework for global patterns in biodiversity. Ecol Lett 7:1-15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>

- Rizzini AT (1997) Tratado de fitogeografia do Brasil: aspectos ecológicos, sociológicos e florísticos, 2nd edn. Âmbito Cultural, Rio de Janeiro.
- Rohde K (1992) Latitudinal Gradients in Species Diversity: The Search for the Primary Cause. *Oikos* 65:514-527.
- Sandel B, Svenning JC (2013) Human impacts drive a global topographic signature in tree cover. *Nat Commun* 4:1-4. <https://doi.org/10.1038/ncomms3474>
- Scarano FR (2002) Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian Atlantic Rainforest. *Ann Bot* 90:517-524. <http://dx.doi.org/10.1093/aob/mcf189>
- Silva KJP, Souza AF (2018a) Common species distribution and environmental determinants in South American coastal plains. *Ecosphere* 9:1-15. <https://doi.org/10.1002/ecs2.2224>
- Silva AC, Souza AF (2018b) Aridity drives plant biogeographical sub regions in the Caatinga, the largest tropical dry forest and woodland block in South America. *Plos One* 13:1-22. <https://doi.org/10.1371/journal.pone.0196130>
- Sousa-Baena MS, Garcia LC, Peterson AT (2014) Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Divers Distrib* 20:369-381. <https://doi.org/10.1111/ddi.12136>
- Stropp J, Ladle RJ, Malhado ACM, Hortal J, Gaffuri J, Temperley WH, Skøien JO (2016) Mapping ignorance: 300 years of collecting flowering plants in Africa. *Global Ecol Biogeogr* 25:1-12. <https://doi.org/10.1111/geb.12468>
- Tilman D (1994) Competition and Biodiversity in Spatially Structured Habitats. *Ecology* 75:2-16. <https://doi.org/10.2307/1939377>
- Trabucco A, Zomer RJ (2019). Global High-Resolution Soil-Water Balance: CGIAR Consortium for Spatial Information. Available: <http://www.cgiar-csi.org>. Accessed 11 October 2019.

- Ugland KI, Gray JS, Ellingsen KE (2003) The species-accumulation curve and estimation of species richness. *J Anim Ecol* 72:888-897.
<https://doi.org/10.1046/j.1365-2656.2003.00748.x>
- WCS/CIESIN - Wildlife Conservation Society/Center for International Earth Science Information Network (2002) Global Human Footprint Dataset. Available: <https://sedac.ciesin.columbia.edu/>. Accessed 09 June 2020.
- Willig MR, Kaufmann DM, Stevens RD (2003) Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. *Annu Rev Ecol Evol S* 34:273-309.
<https://doi.org/10.1146/annurev.ecolsys.34.012103.144032>
- Yang W, Ma K, Kreft H (2013) Geographical sampling bias in a large distributional database and its effects on species richness–environment models. *J Biogeogr* 40:1415-1426. <https://doi.org/10.1111/jbi.12108>

Supplemental Material

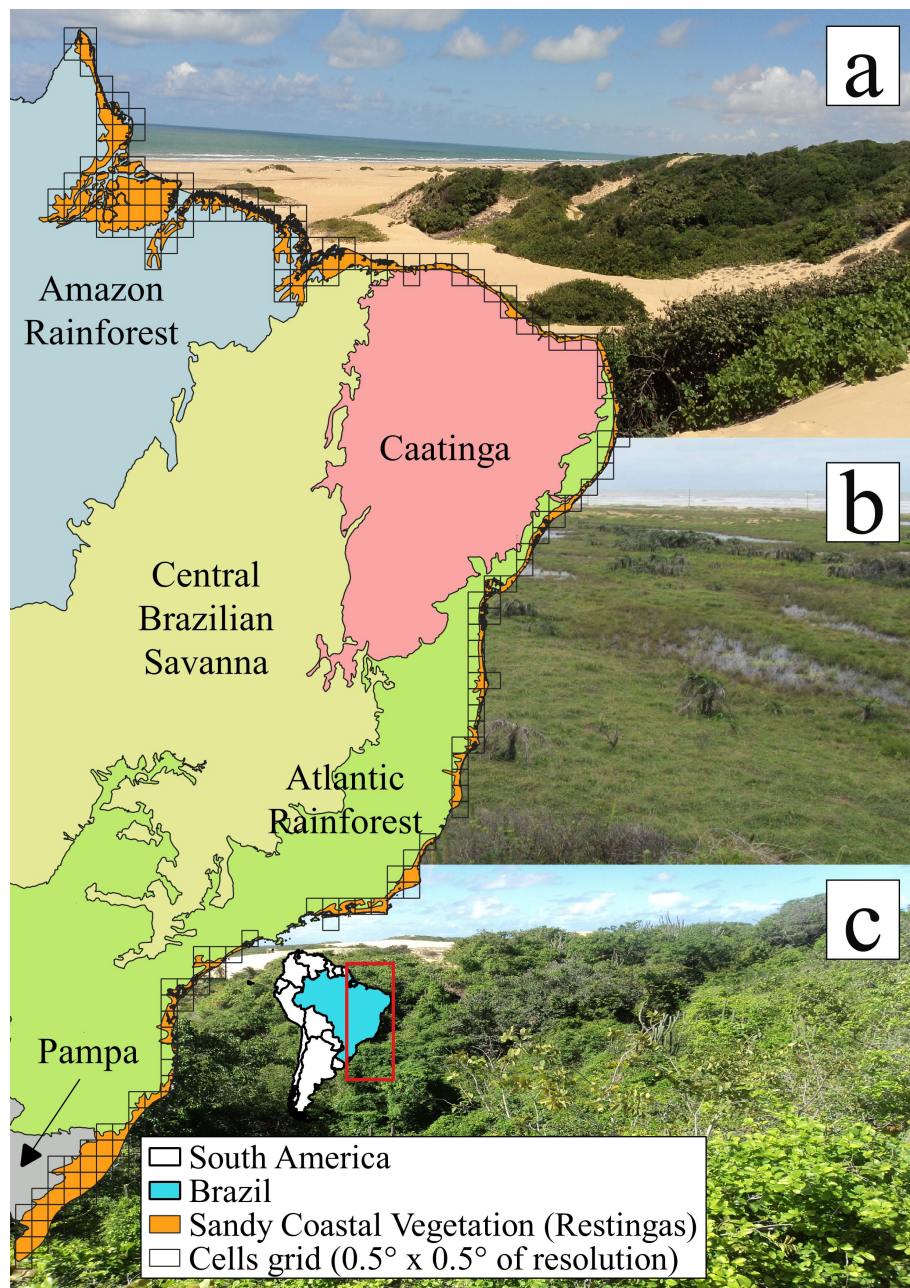


Figure S1 Delimitation of sandy coastal vegetation (Restingas) in the Eastern South America with adjacent phytogeographic domains. Sided images are typical formations of this vegetation, including a) Restingas scrubs (shrubby vegetation) on coastal dunes, b) Herbaceous Restinga (grasslands), and c) Restinga forest at the leeward of coastal dunes (Photos: Eduardo V. S. Oliveira).

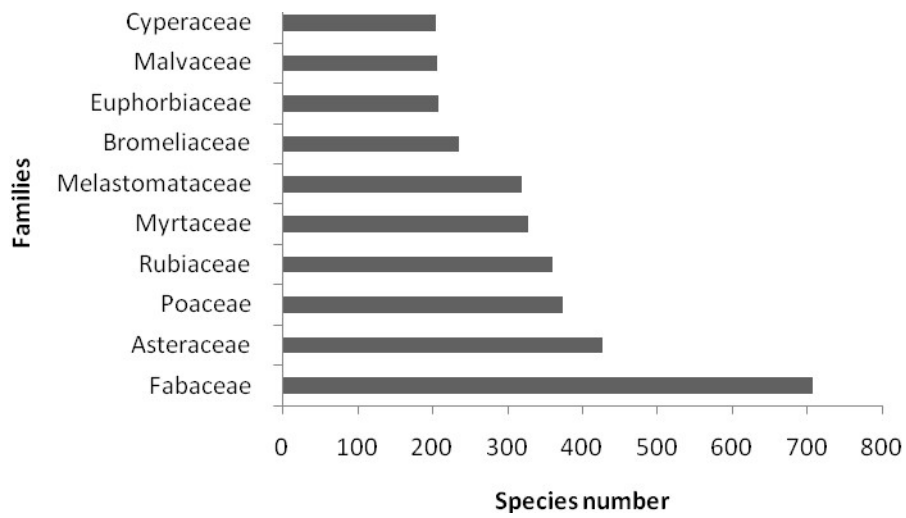


Figure S2 Most speciose families of plants at the sandy coastal vegetation (Restingas) in the Eastern South America.

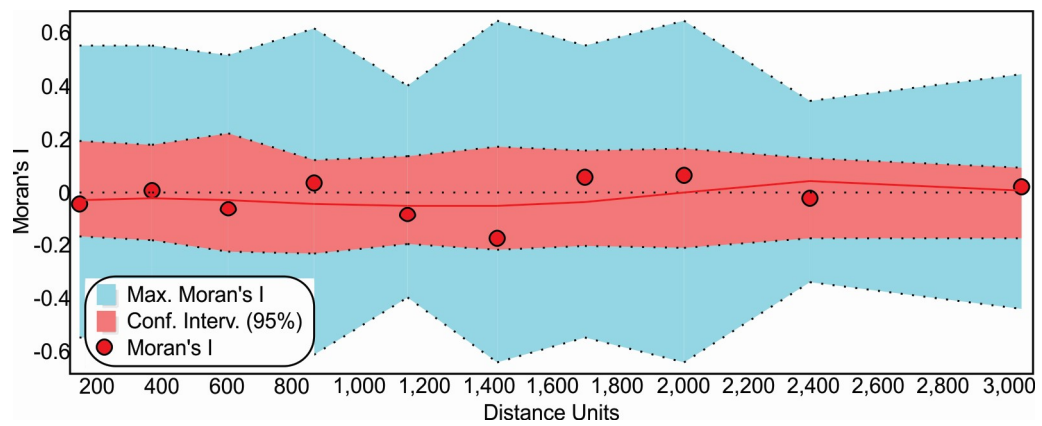


Figure S3 Spatial correlograms of Moran's *I* index of the regression residuals of the full model tested.

CAPÍTULO 2

Patterns and drivers of functional and phylogenetic structures of coastal woody vegetation of Eastern South America ^[1]

[1] Artigo submetido no periódico Journal of Plant Research. Citação: Oliveira EVS, Landim MF, Gouveia SF. Patterns and drivers of functional and phylogenetic structures of coastal woody vegetation of Eastern South America. Journal of Plant Research, submetido.

Abstract

Despite it remains challenging, understanding the formation of ecological communities has gained new insights from studies in tropical environments. Recent studies have hypothesized that hydric constraints determine plant assembly structure in coastal communities by favoring preadapted lineages from neighbor ecosystems, consistent with a functional clustering. However, this hypothesis, together with the evolutionary context involved in this pattern, remains elusive. Here, we approach these questions across the coastal zone of Brazil while dealing with multiple knowledge deficiencies in plant geographic distribution, traits, and phylogeny. After overcoming spatial, functional, and phylogenetic knowledge gaps, we measured the functional and phylogenetic structures through phylogenetic species variability (PSV). We then evaluated the presence of deterministic assembling processes, i.e., clustering or overdispersion. Using linear models in a model selection framework, we further contrasted different environmental hypotheses deemed to explain the patterns, namely hydric stress, historical climatic stability, edaphic constraints, and habitat heterogeneity. Plant coastal assemblages show varied patterns, but with a predominance of stochastic-to-clustering in both functional and phylogenetic structures, although overdispersion also occurred. Soil properties (namely available water capacity) and historical variability of temperature and precipitation best explained the gradient in the functional structure. Solar radiation, water deficit, and historical variability of precipitation were the best predictors of the variation in phylogenetic structure, with clustering being more frequent at drier regions. Our results support the water-stress conservatism hypothesis according to which water-depleted conditions in tropical coastal environments favor functionally similar species. This pattern appears to result from the occupation of closely related species, likely originated from the neighbor ecosystems. However, as we hypothesize, the species contribution by hyperdiverse rainforests seems to

counterbalance this effect, highlighting the influence of the richness of regional species pools on the observed structure.

Keywords: Brazil, completeness, clustering, *Restinga* vegetation, water constraints

Introduction

The explanation of the process that underlies the species composition in ecological communities has challenged ecologists ever since the onset of the field (Clements 1905; Gleason 1939) and resisted time (MacArthur and Levins 1967; Diamond 1975; Pavoine and Bonsall 2011). Nevertheless, some general principles – or assembly rules – have been established to explain how and why only a subset of species available in the regional pool occurs in a local community (Belyea and Lancaster 1999; Götzenberger et al. 2012; Mittelbach and McGill 2019). Accordingly, the species' sorting in a locality depends on the interaction of three processes over temporal and spatial scales: (i) stochastic dispersal, (ii) abiotic filtering (i.e., conditions), and (iii) biotic filtering (e.g., resource use and partitioning). In other words, communities comprise species from the regional pool that can reach the site and possess features that enable them to cope with conditions and assimilate resources despite other species' presence (HilleRisLambers et al. 2012; Mittelbach and McGill 2019).

The detection of assembly rules has been a critical issue. It should be drawn from the imprint that those processes have left in the species in terms of features they possess, which express how they relate to abiotic and biotic conditions. This detection often involves similarity measures among the observed species regarding these 'functional traits' through metrics of assemblages' functional structure (Webb et al. 2002; Petchey and Gaston 2006; Emerson and Gillespie 2008). Because evolutionarily related species tend to share similar phenotypes (Wiens and Graham 2005), metrics based on phylogenetic distance first described the patterns of trait similarity of the

assemblages (Webb et al. 2002; Cavender-Bares et al. 2009). It is now clear that the functional structure may depart from the phylogenetic structure due to different processes, including phenotypic divergence and convergence between species (Cadotte et al. 2013; Gerhold et al. 2015; Davies 2021), or to subtle differences in their responses at different temporal scales (Nascimento et al. 2018).

Assemblages under more pronounced abiotic constraints, for instance, might have more similar species, i.e., functionally clustered (Pavoine et al. 2011). However, these species can be phylogenetically similar when trait similarities reflect phylogenetic relatedness or phylogenetically dissimilar when their trait similarity is due to evolutionary convergence (Davies 2021). Conversely, resource-depleted environments tend to harbor species that have endured more intense competition so that the resulting assemblage comprises different, less competing species. In this case, the functional structure tends to be more overdispersed, which may or may not be accompanied by an overdispersed phylogenetic structure, depending on the evolutionary modes of the attributes involved (Pausas and Verdú 2010; Pavoine et al. 2011). Therefore, understanding the patterns of functional and phylogenetic structures separately for different organisms at different environmental contexts is essential to elucidate how different evolutionary and ecological contexts conflate to form the existing diversity of ecological communities.

For angiosperms in dry regions, the water shortage seems to favor functionally clustered plant assemblages through filtering out water-sensitive species (Cornwell and Ackerly 2009; Qian and Sandel 2017; Kubota et al. 2018; Qian et al. 2019). In this regard, the premise of niche conservatism posits that community assembly under more significant environmental stress, such as cold, dry, or pronounced seasonality, should be mainly structured by environment filtering than by biotic interactions (Qian et al. 2019). In this line, Qian et al. (2016) proposed that sites under more intense hydric stress will

contain more closely related species, referring to this explanation as to the ‘water-stress conservatism hypothesis’. This hypothesis finds support in plant assemblages from multiple habitats, including coastal regions, where species sorting seems to be determined by their tolerance to abiotic conditions (Cornwell and Ackerly 2009; Brunbjerg et al. 2013; Jiménez-Alfaro et al. 2015; Sperandii et al. 2018).

On the Eastern coast of South America, sandy coastal plains harbour a plant formation known as the Restingas, which is influenced by the physical features of coastal environments, such as geologically recent sandy substrate, maritime wind blow, and high insolation (Villwock et al. 2005; DaSilva and Pinto-da-Rocha 2011). Evidence suggests that the Restingas’ flora originated from adjacent ecosystems related to the phytogeographic domains of the Atlantic rainforest, Caatinga seasonally dry forests, and Amazon rainforest (Rizzini 1997; Scarano 2002; Fernandes and Queiroz 2015). Recently, Oliveira et al. (2021) showed that environmental heterogeneity plus soil properties and water constraints are the main drivers of plant species richness in Restingas by limiting species accumulation in topographically flatter, water-depleted, and nutrient-poor soil conditions.

Based on these drivers of species richness (Oliveira et al. 2021), the water-stress conservatism hypothesis suggests that these assemblages should comprise species phenotypically more similar, thus functionally clustered (Melo-Jr and Boeger 2017; Oliveira et al. 2021). However, according to the biogeographic hypothesis of an origin of species composition at different neighbor formations (Rizzini 1997; Scarano 2002; Fernandes and Queiroz 2015), it remains the question of the species’ evolutionary associations linked to the functional structure. That is, did these assemblages result from the colonization of a few preadapted lineages (i.e., resulting in phylogenetic clustering) or by evolutionarily independent species (i.e., resulting in a phylogenetic random or overdispersed structure) originated from the regional pool? Here, we tested the

hypothesis of functional clustering of flowering woody plant assemblages in the Restingas of Brazil and addressed the associated pattern of phylogenetic structure, if clustered or not. We also tested whether the observed pattern is consistent with the water-stress conservatism hypothesis or could be better explained by alternative environmental hypotheses, namely the climatic stability hypothesis, the environmental heterogeneity hypothesis, or the edaphic hypothesis. We addressed these questions while handling knowledge shortfalls in the distribution, evolutionary relationships, and species functional traits.

Material and Methods

Study area and data handling

We defined the study area by creating a polygon from the overlapping contours of the coastal plains and plant formations associated with the Brazilian Coastal Zone (Oliveira et al. 2021). Then, onto this polygon, we created a cells grid with a resolution of 55 x 55 km (0.5° x 0.5° latitude-longitude at the Equator), using the SAM software (Spatial Analysis in Macroecology, Rangel et al. 2010).

Regarding plant occurrence data, we extracted occurrence records of angiosperms present in the study area from the BIEN database (Botanical Information and Ecology Network; Enquist et al. 2016, accessed on July 21, 2018). In this step, we only included unique records of native species with valid names. From this data set, we selected trees and shrubs only, following the classification of species by life-forms of BFG (2018). Then, we removed the woody plants of the Brazilian Mangroves (Kjerfve and Lacerda 1993) to obtain a set of species typical of the Restingas vegetation. We used the `BIEN` and `Flora` packages in R software (R Development Core Team 2018; Maitner et al. 2017; Carvalho 2020).

Environmental data

We selected environmental predictors related to water deficit, climatic stability, environmental heterogeneity, and edaphic factors. The water deficit hypothesis was described by solar radiation (Fick and Hijmans 2017) and water balance deficit (WBD). We included solar radiation because of its critical role in moisture removal from soil and plants through solar irradiance (Heck et al. 2020) although we acknowledged that this predictor has other direct effects on vegetation (Moeslund et al. 2013). WBD is the difference between potential evapotranspiration (PET) and actual evapotranspiration (AET) (Trabucco and Zomer 2019). When PET is higher than AET, WBD is positive, meaning that vegetation is constrained by water (Littell and Gwozdz 2011). The climatic stability hypothesis was described by average historical variability of temperature and precipitation, obtained from the difference between the present and the last glacial maximum (CCSM4; Fick and Hijmans 2017). The environmental heterogeneity hypothesis was described by altitude, and the edaphic hypothesis was described by the soil water capacity, nutrient availability, and salinity (Fischer et al. 2008). We removed the collinear variables, considering Pearson's $r \geq 0.7$.

Definition of well-surveyed cells

From the occurrence records, we calculated the inventory completeness (given as a percentage) for each grid cell, defined as the ratio between the observed and the predicted number of species (Lobo et al. 2018). We calculated this latter using the Rational curve for species saturation (Ratkowski 1990), using cells with more than 100 records and completeness higher than 50% as well-surveyed cells. This criterion is similar to that adopted by other authors (cf. Sousa-Baena et al. 2014; Stropp et al. 2016; Oliveira et al. 2021). Well-surveyed cells were selected to compose the grid used in the analysis. The procedures described were performed using the QGIS software (QGIS

Development Team 2020) and with the `Raster` (Hijmans et al. 2019) and `KnowBr` (Lobo et al. 2018) packages in R environment (R Development Core Team 2018).

Functional and phylogenetic structure

To investigate the functional structure, we selected three plant functional traits, leaf area (mm^2), whole plant height (m), and dispersal syndrome (biotic and abiotic). We selected these traits considering their consistency in the expression of plant responses to abiotic (i.e., environmental factors) and biotic filters (i.e., competitive interactions). For instance, height and leaf area relate to the plant responses to climate, soil resources (e.g., water and nutrient availability), and competitive interactions, whereas dispersal syndromes are related to plant colonization potential and disturbances (Lavorel and Garnier 2002; Cornelissen et al. 2003). We also choose these attributes based on their availability for the species from the study region. For each species, we obtained values for these attributes using the following databases and research sources: BIEN (Enquist et al. 2016), TRY - *Plant Trait Database* (Kattge et al. 2020), NeoTropTree (only height, Oliveira-Filho et al. 2017) and scientific publications (leaf area and dispersal syndrome). In this step, we used the `BIEN` package (Maitner et al. 2017) in R (R Development Core Team 2018).

We addressed data gaps in functional traits for some species through an imputation approach. We used the Random Forest method, based on machine learning algorithms, which adds missing values of categorical or continuous traits using other known variables (Pantanowitz and Marwala 2009). Good accuracy of the procedure depends on a proportion of at least 70% of data availability (Penone et al. 2014). We assessed the accuracy of the imputation method using two error measures: the Normalized Root Mean Squared Error (NRMSE) for continuous traits and the Proportion of Falsely Classified (PFC) for categorical traits. In both cases, the error

ranges from zero (i.e., accurate) to unity (i.e., estimate is not different from random) (Li et al. 2015; Lenz et al. 2020). We performed the imputation with 92% of the known traits, therefore with high reliability (see Results). We used the `missForest` (Stekhoven and Bühlmann 2012) package in R (R Development Core Team 2018) for this procedure.

To investigate the phylogenetic structure, we use a dated phylogeny for spermatophytes (ALLOTB; Smith and Brown 2018). This megaphylogeny was generated from molecular data (GenBank) and the Open Tree of Life Project (Smith and Brown 2018), being the most complete for this group of plants and with the highest proportion of well-resolved species present in Restingas so far (cf. Zanne et al. 2014; Gastauer and Meira-Neto 2017; Jin and Qian 2019). For species missing in the phylogeny, we randomly inserted them in the most derived clades, usually genera or families (e.g., Ma et al. 2016; Qian et al. 2016). Because this procedure adds uncertainty in the analyses, we ran the analyses for 100 possible phylogenies. All steps described were performed in R, using the `ape` (Paradis et al. 2004), `phytools` (Revell 2012) and `V.PhyloMaker` (Jin and Qian 2019) packages in R (R Development Core Team 2018).

We used the Phylogenetic Species Variability (PSV; Helmus et al. 2007) to describe both the functional (FSV hereafter) and phylogenetic (PSV) structures. The PSV expresses the hypothetical variation of a neutral attribute evolving under a Brownian movement between the species of an assemblage (Helmus et al. 2007) but can also be extended to describe functional structure based on the functional distance among species (Pausas and Verdú 2010). We generate the FSV using a square matrix, expressing the functional distances based on the Gower index (Gower 1971). We obtained the average PSV for each cell from the estimates obtained from all 100 phylogenetic hypotheses. The functional structure can be coupled to phylogenetic

structure because of phylogenetic conservatism in traits. To check upon this effect, we performed a linear regression analysis between FSV and PSV (Nascimento et al. 2018). As they were unrelated to each other (see Results), we used the obtained values of each measure without needing a correction of the phylogenetic effect. To evaluate whether communities tend to overdispersion or clustering, we generated standardized values of FSV and PSV. We compared the observed values for each community against a null model with 999 randomizations (Independent swap algorithm; Gotelli and Entsminger 2003). Functional and phylogenetic clustering occurs when the observed value of FSV or PSV, respectively, is lower than expected by chance, and overdispersion when it is higher than expected by chance (Helmus et al. 2007; Pausas and Verdú 2010). We performed these steps following Nascimento et al. (2018) and using the *vegan* (Oksanen et al. 2015), *FD* (Laliberté et al. 2015), and *picante* (Kembel et al. 2010) packages in R (R Development Core Team 2018).

Drivers of functional and phylogenetic structure

We used a model selection approach to test the hypotheses for the patterns of functional and phylogenetic structure. We used linear regression models and variance partitioning to evaluate the individual contribution of each of the variables present in the best models for the variation in the functional and phylogenetic structure (Legendre and Legendre 2012). We adopted Akaike's Information Criterion corrected for small sample size (AICc) to identify the most parsimonious model (Burnham and Anderson 2004). We also check the residuals of linear regressions for spatial autocorrelation using spatial correlograms generated by Moran's *I* index (Legendre and Legendre 2012). For these analyses, we used the SAM software (Rangel et al. 2010).

Results

We gathered 107,946 records of flowering woody plants for the Restingas, with cells ranging from two to 4,841 records. The inventory completeness varied between 4% and 82%, with only 59 (30% of the total) well-sampled cells (i.e., completeness > 50%) being thus retained for analysis. Considering only these well-sampled cells (hereafter referred to as assemblages), we found 3,623 species of woody plants (i.e., trees and shrubs angiosperms), of which 334 (~ 9%) were inserted in the phylogeny following the procedure described above. Regarding the imputation of functional traits through the Random Forest approach, error measures were NRMSE = 0.31 and PFC = 0.27, which indicates a percentual inaccuracy of traits estimation lower than 10%, thus assumed here as being acceptable for our purposes. Complete information about the species list, the spatial distribution of records, observed richness, estimated richness, and inventory completeness are summarized in the supplementary material (Fig. S1).

The relationship between functional and phylogenetic structure was statistically negligible ($R^2 = 0.01$), so the metrics were independent of each other. Overall, FSV values showed a weakly structured pattern in space, with most lower than 0.50 (~ 90%), ranging from 0.36 to 0.61 (mean = 0.45), indicating a generally stochastic, with a slight tendency towards functional similarity among species (Fig. 1). We found 13 assemblages (~ 22%) with values of FSV significantly lower than expected by chance, and eight assemblages (~ 13%) with FSV values significantly higher than expected by chance, i.e., overdispersed (Fig. 1). The highest values of FSV occurred in assemblages located in the Atlantic Rainforest and Amazon Rainforest domains (Fig. 1).

Regarding the phylogenetic structure, we asked if the observed pattern would parallel the functional structure or show phylogenetic overdispersion, indicating occupation by distinct lineages with convergent adaptations. We found a tendency towards phylogenetic clustering (Fig. 1), with low PSV values (min. = 0.34; max. =

0.39; mean = 0.36), thus analogous to the functional structure. The PSV values generated for each assemblage using different phylogenies remained nearly identical (standard deviation ± 0.01). Still, considering the inclusion of phylogenetic uncertainty in the analyzes, we found an average of 18 assemblages ($\sim 30\%$; min. = 17; max. = 19) with PSV significantly lower than expected by chance. In contrast, an average of two assemblages ($\sim 3\%$; min. = 1; max. = 3) had PSV values significantly higher than expected by chance (Fig. 1). Most higher values of PSV occurred in the Atlantic rainforest (at higher latitudes) and the Amazon rainforest. The lowest values of PSV of the entire study region were also observed in assemblages in the Atlantic rainforest (at lower latitudes) and Caatinga domains (Fig. 1).

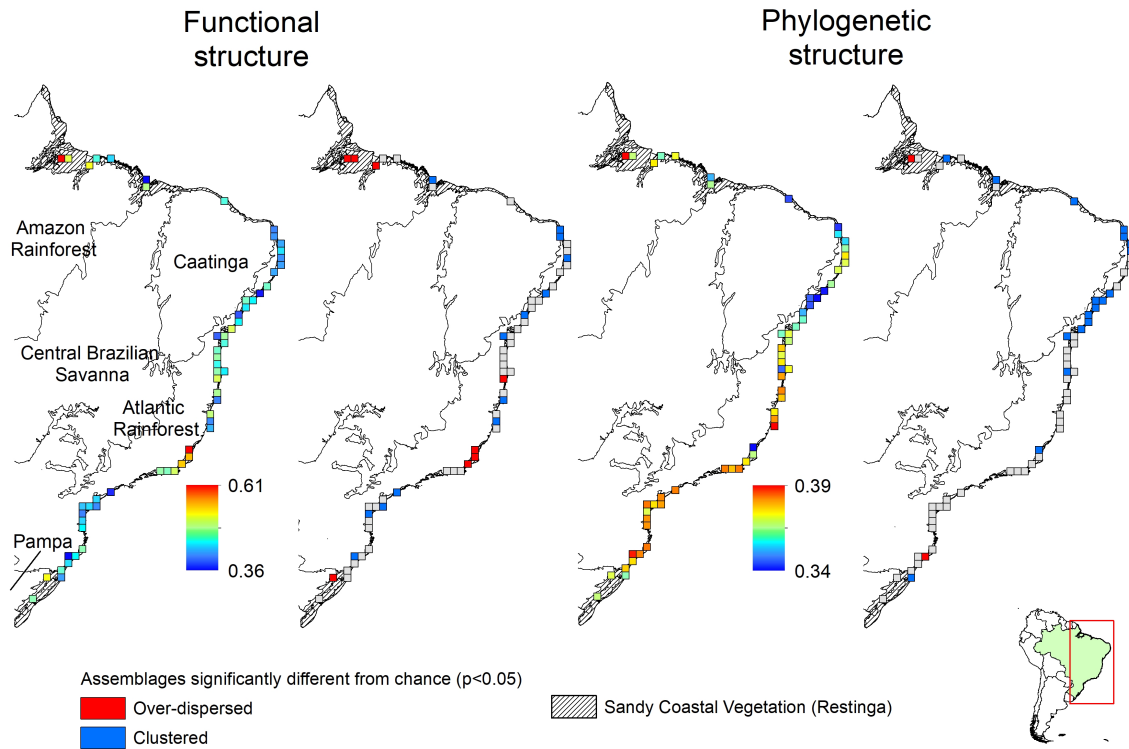


Fig. 1 Functional and phylogenetic structure of assemblages of flowering woody plants in the Brazilian Restingas, Eastern coast of South America. Values of observed Phylogenetic Species Variability (PSV) and assemblages with clustering and overdispersed structure significantly different from chance are shown. The phylogenetic structure was obtained by averaging the 100 randomizations using different phylogenies. Those PSV values that differ from a null model are represented by the mean of the significance value (p-value).

We found no spatial autocorrelation in the best model residuals of the relationships involving functional or phylogenetic structure, so we did not use spatial regression analyses (Fig. S2). From the model selection approach using the Akaike Information Criterion (AICc), soil water capacity (positive effect), water capacity (positive effect), and historical variability of both temperature and precipitation (negative effect) were in the best model for functional structure (Model 1 in Table 1).

Furthermore, available water capacity and historical temperature variability were retained in all six best models based on the AICc (Table 1).

Table 1 Six best models ($\Delta \text{AICc} < 2$) composed of environmental predictors for the patterns of the functional structure of plants according to the Akaike Information Criterion corrected for sample size (AICc). CN: Condition number; CS: Climatic stability; ED: Edaphic; EH: Environmental heterogeneity. PRE-VAR: Historical precipitation change; TEMP-VAR: Historical temperature change; ALT: Altitude; SWC: Soil Water Capacity; NUT: soil nutrient availability. * = $p < 0.01$.

Model	Hypothesis	Predictors	Slope (β)	CN	R^2	AICc	ΔAICc
1	ED	SWC	0.401*	1.85	0.307*	-213.11	0
	CS	TEMP-VAR	-0.270*				
		PRE-VAR	0.211*				
2	ED	SWC	0.510*	1.32	0.272*	-212.71	0.40
	CS	TEMP-VAR	-0.321*				
3	ED	SWC	0.372*	2.26	0.300*	-212.47	0.64
		NUT	0.212				
	CS	TEMP-VAR	-0.323*				
4	ED	SWC	0.418*	1.88	0.326*	-211.91	1.20
	CS	TEMP-VAR	-0.225*				
		PRE-VAR	0.199				
	EH	ALT	0.120				
5	ED	SWC	0.522*	1.35	0.291	-211.78	1.33
	CS	TEMP-VAR	-0.312*				
	EH	ALT	0.133				

Model	Hypothesis	Predictors	Slope (β)	CN	R^2	AICc	$\Delta AICc$
6	ED	SWC		2.41	0.318	-211.59	1.52
		NUT					
	CS	TEMP-VAR					
		PRE-VAR					

The best model for the spatial pattern in phylogenetic structure included solar radiation, water balance deficit, and historical variability of precipitation (Table 2). PSV was negatively related to solar radiation and water balance deficit and positively with historical precipitation variability. These three environmental predictors were also present in the three main models based on the AICc (Table 2).

Table 2 Three best models ($\Delta AICc < 2$) composed of environmental predictors for the patterns of phylogenetic structure of plants according to the Akaike Information Criterion corrected for sample size (AICc). CN: Condition number; WD: Water deficit; CS: Climatic stability; ED: Edaphic; EH: Environmental heterogeneity. RAD: Solar radiation; WBD: Water balance deficit (PET - AET); PRE-VAR: Historical precipitation variability; TEMP-VAR: Historical temperature variability; ALT: Altitude; SWC: Soil Water Capacity; NUT: nutrient availability in the soil. * = $p < 0.01$.

Models	Hypothesis	Predictors	Slope (β)	CN	R^2	AICc	$\Delta AICc$
1	WD	RAD	-0.364*	1.95	0.450*	-528.69	0
		WBD	-0.363*				
	CS	PRE-VAR	0.248*				

Models	Hypothesis	Predictors	Slope (β)	CN	R^2	AICc	$\Delta AICc$
2	WD	RAD	-0.408*	2.07	0.466*	-528.46	0.23
		WBD	-0.356*				
	CS	PRE-VAR	0.307*				
	ED	SWC	0.159				
3	WD	RAD	-0.404*	2.12	0.460*	-527.82	0.86
		WBD	-0.343*				
	CS	PRE-VAR	0.313*				
	ED	NUT	-0.140				

Regarding the variance partition, we found that available water capacity, from the edaphic hypothesis, had the highest proportion of explanation of the functional structure of woody plants (i.e., exclusive and shared portion), followed by historical variability of temperature and precipitation, both from the hypothesis of climatic stability (Fig. 2). In turn, water deficit, representative of the water deficit hypothesis, presented the highest proportion of explanation of the phylogenetic structure of woody plants, followed by solar radiation (water deficit hypothesis) and historical precipitation variability (climate stability hypothesis) (Fig. 2).

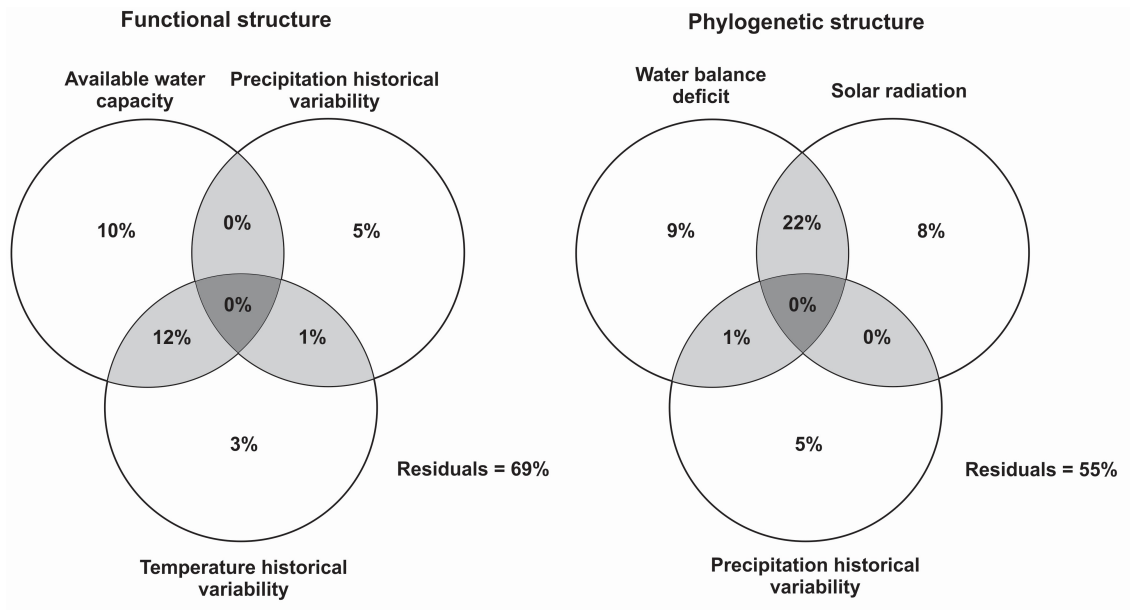


Fig. 2 The individual and shared proportion of variance explained (%) by the environmental factors as drivers of the phylogenetic and functional structure of woody plants in the Brazilian Restingas, grouped according to the corresponding hypothesis. Values are representative of the better models for each response variable.

Discussion

We found a variety of patterns of functional structure in the Restingas of Brazil, with a predominance of stochastic arrangements, a slight tendency towards clustering, and some overdispersion of species relative to their traits and phylogenetic relatedness. Overall, this result indicates that trait convergency is more widespread among the species in these assemblages. The lack of relationship between FSV and PSV was surprising given the general expectation of traits phylogenetic conservatism. This pattern suggests a complex pattern of species assembly likely influenced by regional processes. As comparisons involve assemblages associated with different formations (i.e., Amazon forest, Cerrado, Caatinga, Atlantic forest, and Pampas), it seems that different portions of the Restingas share different lineages with their neighbor formations, leading to these complex patterns. However, there is an overall tendency to

reduce both functional and phylogenetic variability in assemblages from Northeastern Brazil, which seems to be associated with the Caatinga and the drier portion of the Atlantic forest.

This occurrence of eco-evolutionary clustering, especially in the drier regions (i.e., Northeast and South), indicates that the patterns of species co-occurrence in the habitats of the Restingas of Brazil is, to some extent, influenced by environmental filtering (e.g., Yan et al. 2013), as in plant assemblages from other regions with pronounced environmental constraints (Swenson et al. 2011; Qian et al. 2014; Bruelheide et al. 2018; Kubota et al. 2018; Qian et al. 2019). According to this rationale, environmental filtering favors species with specific phenotypes in the communities usually derived from a few lineages (Pausas and Verdú 2010). In our data, the ten more speciose families (8% of all families) comprise about 60% of all species found in the Restingas. The same pattern holds at the order level, in which three of the 34 orders (Myrtales, Malpighiales, and Fabales) comprise half of all species (Scarano 2002; Massante and Gerhold 2020). In the Restingas, this moderate environmental filtering favors closely related and preadapted species or species from distinct lineages with either convergent phenotypes or prone to exhibit phenotypic plasticity to cope with the shortage of water and soil nutrients (Scarano 2002; Amorim and Melo-Jr 2017; Melo-Jr and Boeger 2017). In fact, phenotypic plasticity in this formation is well known, being more common in the leaves (e.g., leaf area and thickness), such as in *Myrsine venosa* A.DC., *Ocotea pulchella* (Nees & Mart.) Mez and *Pera glabrata* (Schott) Baill.), or in the anatomy of the stem (e.g., height and wood density), as is the case of *Andira fraxinifolia* Benth. and *Tibouchina clavata* (Pers.) Wurdack (Silva et al. 2016; Amorim and Melo-Jr 2017; Silva et al. 2017).

Nevertheless, the high share of assemblages with moderate phylogenetic variability and functional overdispersion, mainly in the Amazon and Atlantic rainforest,

seems not to be caused by chance. These areas have higher environmental humidity, mild temperatures and harbor more diverse species pools. Therefore, a converse tendency seems to be the case in these regions, with within-clade competition and phenotypic divergence (Gerhold et al. 2015; Davies 2021). This result may reflect the biogeographic legacy of these hyperdiverse domains. Biogeographical processes on large spatial scales determine the regional species pool from which ecological communities are assembled (Pausas and Verdu 2010). The Amazon and Atlantic rainforests possess some of the most remarkable species diversities and endemisms on the planet (Haffer 2008; Sobral-Souza and Lima-Ribeiro 2017). As such, these large species pools, together with their long histories of climatic cycles and geological heterogeneity, should account for their contribution of multiple species from a higher diversity of lineages and traits to the adjacent Restingas (Scarano 2002; Silva et al. 2015; Massante and Gerhold 2020).

The environmental predictors that best explained the observed patterns and the direction of their associations reinforce our conclusions about the abiotic constraint to species sorting, particularly in drier regions. Regarding the effect on the phylogenetic structure, we found a more substantial effect of solar radiation, water balance deficit, and historical variability of precipitation, with a lower phylogenetic variability at warmer and drier regions. For the functional structure, we found that the trait convergence (i.e., clustering) further increases towards areas of lower soil water capacity (i.e., a positive association of this factor and structure), changed more in temperature since the last glacial maximum, and changed less in precipitation loads since this geological period. Altogether, these findings lend support to the water-stress conservatism hypothesis (Qian et al. 2016). Although our model included variables assigned to different hypotheses, they can be linked to multiple constraints to water balance for plants in coastal regions, and the Restingas in particular, such as the

dehydrating effect of radiation, the water shortage in soil with low water retention capacity, and historical changes in climate that increase the hydric stress in this region (Vitória et al. 2017; Massante and Gerhold 2020). This effect of historical climate change is noteworthy. These areas that changed less in precipitation and increased temperature should have undergone a gradual increase in evaporative rates, posing a gradually higher selective constraint to the plants of this region (e.g., Amorim and Melo-Jr 2017; Melo-Jr and Boeger 2017; Silva e Souza 2018).

Our results bear four simultaneous shortfalls in biodiversity information regarding (i) taxonomic diversity of species (i.e., Linnean shortfalls), (ii) the distribution (i.e., Wallacean shortfall), (iii) the evolutionary relationship (i.e., Darwinian shortfall), and the species functional traits (i.e., the Raunkiaerian shortfall) (*sensu* Hortal et al. 2015). The two former shortfalls were recently described and dealt with for this study system (Oliveira et al. 2021). For the uncertainty involving the evolutionary information, our results show that the phylogenetic structure remained nearly identical between the assemblages, even after considering 100 phylogenetic hypotheses. This lack of influence of the phylogenetic uncertainty probably results from the low proportion (9.0%) of missing species in the phylogenetic tree, and because phylogenies resolved at the family or genus level can be robust enough for investigating the relationship between phylogenetic structure and environmental predictors in plants (Qian and Zhang 2016; Qian and Jim 2021). Regarding the data gaps about functional traits, particularly the quantitative ones, the information assembled also yielded an acceptable accuracy (average error in the estimate trait values < 10%), suggesting that the Raunkiaerian shortfall should have a low influence on our estimates of the functional structure. Therefore, overall, we could still draw meaningful information about this system despite the incompleteness in data.

In summary, our study has shown that, in woody plant assemblages from the Restingas of Brazil, different donor formations appear to have contributed differently to the adjacent assembly patterns at the coast. On the one hand, water limiting factors are the main drivers of the assemblage structure at drier portions, favoring closely related and preadapted lineages, particularly in the Northeast region. On the other hand, the high species diversity from wetter formations, such as areas in the Amazon and the Atlantic forests, contributes to more species and lineages, leading to either a stochastic arrangement or even overdispersion among species regarding their traits. The influence of historical changes in climatic factors linked to a gradual increase of water stress and plant functional clustering suggests that the continuing temperature rises from global climate change should further affect these plant assemblages by exacerbating water stress. In addition to the expected habitat loss due to sea level rises and the intensification of human occupation at this zone, this effect should result in species and functional loss in these communities. Considering the persistent knowledge gaps about multiple aspects of the species, improving the coverage of surveys is critical to understanding this fragile ecosystem's response to these impacts and informs tailored conservation strategies.

References

- Amorim MW, Melo-Jr JCF (2017) Functional diversity of restinga shrub species on the coastal plain of Southern Brazil. *Int J Dev Res* 7:13189-13202.
- Belyea LR, Lancaster J (1999) Assembly Rules within a Contingent Ecology. *Oikos* 86:402-416. <https://doi.org/10.2307/3546646>
- BFG - The Brazil Flora Group (2018) Brazilian Flora 2020: Innovation and collaboration to meet Target 1 of the Global Strategy for Plant Conservation

(GSPC). Rodriguésia 69:1513-1527. <https://doi.org/10.1590/2175-7860201869402>

Bruehlheide H, Dengler J, Purschke O, Lenoir J, Jiménez-Alfaro B, Hennekens SM et al (2018) Global trait–environment relationships of plant communities. *Nat Ecol Evol* 2:1906-1917. <https://doi.org/10.1038/s41559-018-0699-8>

Brunbjerg Ak, Ejrnæs R, Svenning J-C (2012) Species sorting dominates plant metacommunity structure in coastal dunes. *Acta Oecol* 39:33-42. <https://doi.org/10.1016/j.actao.2011.11.002>

Burnham KP, Anderson DR (2004) Model selection and multimodel inference. Springer-Verlag, New York.

Cadotte M, Albert CH, Walker SC (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecol Lett* 16:1234-1244. <https://doi.org/10.1111/ele.12161>

Carvalho G (2020) Tools for Interacting with the Brazilian Flora 2020. <http://www.github.com/gustavobio/flora>. Accessed 11 Mai 2020.

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW (2009) The merging of community ecology and phylogenetic biology. *Ecol Lett* 12:693-715. <https://doi.org/10.1111/j.1461-0248.2009.01314.x>

Clements FE (1905) Research methods in ecology. University of Nebraska Press, Lincoln.

Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE et al (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust J Bot* 51:335-380. <https://doi.org/10.1071/BT02124>

- Cornwell WK, Ackerly DD (2009) Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecol Monogr* 79:109-126. <https://doi.org/10.1890/07-1134.1>
- DaSilva MB, Pinto-da-Rocha R (2011) História biogeográfica da Mata Atlântica: opiliões (Arachnida) como modelo para sua inferência. In: Carvalho CJB, Almeida EAB (orgs) *Biogeografia da América do Sul - Padrões e Processos*. Roca, São Paulo, pp 221-238.
- Davies TJ (2021) Ecophylogenetics redux. *Ecol Lett* 24:1073-1088. <https://doi.org/10.1111/ele.13682>
- Diamond JM (1975) Assembly of species communities. In: Cody ML, Diamond JL (eds) *Ecology and evolution of communities*. Harvard University Press, Cambridge, pp 342-444.
- Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* 23:619-630. <https://doi.org/10.1016/j.tree.2008.07.005>
- Enquist BJ, Condit R, Peet RK, Schildhauer M, Thiers BM (2016) Cyber infrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. *PeerJ Preprints* 12:1-33. <https://doi.org/10.7287/peerj.preprints.2615v2>
- Fernandes MF, Queiroz LP (2015) Floristic surveys of Restinga Forests in southern Bahia, Brazil, reveal the effects of geography on community composition. *Rodriguésia* 66:51-73. <http://dx.doi.org/10.1590/2175-7860201566104>
- Fick SE, Hijmans RJ (2017) Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302-4315. <https://doi.org/10.1002/joc.5086>

- Fischer G, Nachtergaele F, Prieler S, van Velthuisen HT, Verelst L, Wiberg D (2008) Global agro-ecological zones assessment for agriculture. <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/>. Accessed November 21 2018.
- Gastauer M, Meira-Neto JAA (2017) Updated angiosperm family tree for analyzing phylogenetic diversity and community structure. *Acta Bot Brasilica* 31:191-198. <https://doi.org/10.1590/0102-33062016abb0306>
- Gerhold P, Cahill Jr JF, Winter M, Bartish IV, Prinzing A (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Funct Ecol* 29:600-614. <https://doi.org/10.1111/1365-2435.12425>
- Gleason HA (1939) The Individualistic Concept of the Plant Association. *Am Midl Nat* 21:92-110. <https://doi.org/10.2307/2479933>
- Gotelli N, Entsminger G (2003) Swap Algorithms in Null Model Analysis. *Ecology* 84:532-535.
- Götzenberger L, de Bello F, Bråthen KA, Davison J, Dubuis A, Guisan A et al (2012) Ecological assembly rules in plant communities-approaches, patterns and prospects. *Biol Rev* 87:111-127. <https://doi.org/10.1111/j.1469-185X.2011.00187.x>
- Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* 27:857-874. <https://doi.org/10.2307/2528823>
- Haffer J (2008) Hypothesis to explain the origin of species in Amazonia. *Braz J Biol* 68:917-947. <https://doi.org/10.1590/S1519-69842008000500003>
- Heck K, Coltman E, Schneider J, Helmig R (2020) Influence of radiation on evaporation rates: a numerical analysis. *Water Resour Res* 56:1-18. <https://doi.org/10.1029/2020WR027332>

- Helmus MR, Bland TJ, Williams CK, Ives AR (2007) Phylogenetic measures of biodiversity. *Am Nat* 169:68-83. <https://doi.org/10.1086/511334>
- Hijmans RJ, Etten JV, Sumner M, Cheng J, Baston D, Bevan A et al (2019) Package: Raster. Geographic Data Analysis and Modeling. <https://cran.r-project.org/web/packages/raster/index.html>. Accessed 06 Mai 2020.
- HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM (2012) Rethinking Community Assembly through the Lens of Coexistence Theory. *Annu Rev Ecol Evol Syst* 43:227-48. <https://doi.org/10.1146/annurev-ecolsys-110411-160411>
- Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ (2015) Seven shortfalls that beset large-scale knowledge on biodiversity. *Annu Rev Ecol* 46:523-549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
- Jiménez-Alfaro B, Marcenò C, Guarino R, Chytrý M (2015) Regional metacommunities in two coastal systems: spatial structure and drivers of plant assemblages. *J Biogeogr* 42:452-462. <https://doi.org/10.1111/jbi.12437>
- Jin Y, Qian H (2019) V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* 42:1353-1359. <https://doi.org/10.1111/ecog.04434>
- Kattge J, Boenisch G, Diaz S, Lavorel S, Prentice IC, Leadley P et al (2020) TRY plant trait database - enhanced coverage and open access. *Glob Change Bio* 26:119-188. <https://doi.org/10.1111/gcb.14904>
- Kembel SW, Ackerly DD, Blomberg SP, Cornwell WK, Cowan PD, Helmus MR et al (2010) Package: Picante. Integrating Phylogenies and Ecology. <https://cran.r-project.org/web/packages/picante/index.html>. Accessed 21 Mai 2019.
- Kjerfve B, Lacerda LD (1993) Mangroves of Brazil. In: Lacerda LD, Diop S (eds) *Conservation and Sustainable Utilization of Mangroves in Latin America and*

- Africa Regions. International Society for Mangrove Ecosystems, Okinawa, pp 245-272.
- Kubota Y, Kusumoto B, Shiono T, Ulrich W (2018) Environmental filters shaping angiosperm tree assembly along climatic and geographic gradients. *J Veg Sci* 29:607-618. <https://doi.org/10.1111/jvs.12648>
- Laliberté E, Legendre P, Shipley B (2015) FD: measuring functional diversity from multiple traits, and other tools for functional ecology. <https://cran.r-project.org/web/packages/FD/index.html> Accessed 16 July 2019.
- Lavorel S, Garnier E (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Funct Ecol* 16:545-556. <https://doi.org/10.1046/j.1365-2435.2002.00664.x>
- Legendre P, Legendre L (2012) Numerical ecology. Elsevier, Amsterdam.
- Lenz M, Schulz A, Koeck T, Rapp S, Nagler M, Sauer M et al. (2020) Missing value imputation in proximity extension assay-based targeted proteomics data. *Plos One* 15:1-19. <https://doi.org/10.1371/journal.pone.0243487>. eCollection 2020
- Li H, Zhao C, Shao F, Li GZ, Wang X (2015) A hybrid imputation approach for microarray missing value estimation. *BMC Genomics* 16:1-11. <https://doi.org/10.1186/1471-2164-16-S9-S1>
- Littell JS, Gwozdz RB (2011) Climatic Water Balance and Regional Fire Years in the Pacific Northwest, USA: Linking Regional Climate and Fire at Landscape Scales. In: McKenzie D, Miller C, Falk D (eds) *The Landscape Ecology of Fire. Ecological Studies (Analysis and Synthesis)*. Springer, Dordrecht, pp 117-139.
- Lobo JM, Hortal J, Yela JL, Millán A, Sánchez-Fernández D, García-Roselló E et al (2018) KnowBR: An application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecol Indic* 91:241-248. <https://doi.org/10.1016/j.ecolind.2018.03.077>

- Ma Z, Sandel B, Svenning JC (2016) Phylogenetic assemblage structure of North American trees is more strongly shaped by glacial-interglacial climate variability in gymnosperms than in angiosperms. *Ecol Evol* 6:3092-3106. <https://doi.org/10.1002/ece3.2100>
- MacArthur R, Levins R (1967) The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *Am Nat* 101: 377-385.
- Maitner BS, Boyle B, Casler N, Condit R, Donoghue J, Durán SM et al (2017) The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol Evol* 9:1-7. <https://doi.org/10.1111/2041-210X.12861>
- Massante JC, Gerhold P (2020) Environment and evolutionary history depict phylogenetic alpha and beta diversity in the Atlantic coastal white-sand woodlands. *J Veg Sci* 31:634-645. <https://doi.org/10.1111/jvs.12900>
- Melo Júnior JCF, Boeger MRT (2017) Functional traits of dominant plant species of the Brazilian sandy coastal plain. *Int J Curr Res* 9:45585–45593.
- Mittelbach GG, McGill BJ (2019) *Community Ecology*. Oxford University Press, Oxford.
- Moeslund JE, Arge L, Bøcher PK, Dalgaard T, Svenning J-C (2013) Topography as a driver of local terrestrial vascular plant diversity patterns. *Nord J Bot* 31:129-144. <https://doi.org/10.1111/j.1756-1051.2013.00082.x>
- Nascimento ER, Correia I, Ruiz-Esparza JM, Gouveia SF (2018) Disentangling phylogenetic from non-phylogenetic functional structure of bird assemblages in a tropical dry forest. *Oikos* 127:1177-1185. <https://doi.org/10.1111/oik.04910>
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D et al (2018) *Vegan: community ecology package*. R package version 2.5-1. <<https://CRAN.R-project.org/package=vegan>>. Accessed 21 June 2019

- Oliveira EVS, Alves DMC, Landim MF, Gouveia SF (2021) Sampling effort and the drivers of plant species richness in the Brazilian coastal regions. *Oecologia*, 195:163-171. <https://doi.org/10.1007/s00442-020-04805-7>
- Oliveira-Filho AT (2017) NeoTropTree, Flora arbórea da Região Neotropical: Um banco de dados envolvendo biogeografia, diversidade e conservação. Universidade Federal de Minas Gerais. <http://www.neotropree.info>. Accessed 12 December 2019
- Pantanowics A, Marwala T (2009) Missing data imputation through the use of the random forest algorithm. *Adv Intell Syst Comput* 116:53-62.
- Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289-290. <https://doi.org/10.1093/bioinformatics/btg412>
- Pausas JG, Verdú M (2010) The Jungle of Methods for Evaluating Phenotypic and Phylogenetic Structure of Communities. *BioScience* 60:614-625. <https://doi.org/10.1525/bio.2010.60.8.7>
- Pavoine S, Bonsall MB (2011) Measuring biodiversity to explain community assembly: A unified approach. *Biol Rev* 86:792-812. <https://doi.org/10.1111/j.1469-185X.2010.00171.x>
- Pavoine S, Vela E, Gachet S, De Bélair G, Bonsall MB (2011) Linking patterns in phylogeny, traits, abiotic variables and space: A novel approach to linking environmental filtering and plant community assembly. *J Ecol* 99:165-175. <https://doi.org/10.1111/j.1365-2745.2010.01743.x>
- Penone C, Davidson AD, Shoemaker KT, Di Marco M, Rondinini C, Brooks TM et al (2014) Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods Ecol Evol* 5:961-970. <https://doi.org/10.1111/2041-210X.12232>

- Petchey OL, Gaston KJ (2006) Functional diversity: back to basics and looking forward. *Ecol Lett* 9:741-758. <https://doi.org/10.1111/j.1461-0248.2006.00924.x>
- QGIS Development Team (2020) QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>. Accessed 08 August 2020
- Qian H, Deng T, Jin Y, Mao L, Zhao D, Ricklefs RE (2019) Phylogenetic dispersion and diversity in regional assemblages of seed plants in China. *Proc Natl Acad Sci* 116:23192–23201. <https://doi.org/10.1073/pnas.1822153116>
- Qian H, Field R, Zhang JL, Zhang J, Chen S (2016) Phylogenetic structure and ecological and evolutionary determinants of species richness for angiosperm trees in forest communities in China. *J Biogeogr* 43:603-615. <https://doi.org/10.1111/jbi.12639>
- Qian H, Hao Z, Zhang J (2014) Phylogenetic structure and phylogenetic diversity of angiosperm assemblages in forests along an elevational gradient in Changbaishan, China. *J Plant Ecol* 7:154-165. <https://doi.org/10.1093/jpe/rtt072>
- Qian H, Jin Y (2021) Are phylogenies resolved at the genus level appropriate for studies on phylogenetic structure of species assemblages? *Plant Diversity* in Press:1-9. <https://doi.org/10.1016/j.pld.2020.11.005>
- Qian H, Sandel B (2017) Phylogenetic structure of regional angiosperm assemblages across latitudinal and climatic gradients in North America. *Glob Ecol Biogeogr* 26:1258-1269. <https://doi.org/10.1111/geb.12634>
- Qian H, Zhang J (2016) Are phylogenies derived from family-level supertrees robust for studies on macroecological patterns along environmental gradients? *J Syst Evol* 54:29-36. <https://doi.org/10.1111/jse.12161>
- R Development Core Team (2018) R: A language and environment for statistical computing. <http://www.R-project.org>S. Accessed 06 September 2018.

- Rangel TFLVB, Diniz-Filho JAF, Bini LM (2010) SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography* 33:46-50. <https://doi.org/10.1111/j.1600-0587.2009.06299.x>
- Ratkowski DA (1990) Handbook of nonlinear regression models. Marcel Dekker, New York.
- Revell LJ (2012). Package: phytools. Phylogenetic Tools for Comparative Biology (and Other Things). <https://cran.r-project.org/web/packages/phytools/index.html>. Accessed 06 Mai 2020.
- Rizzini AT (1997) Tratado de fitogeografia do Brasil: aspectos ecológicos, sociológicos e florísticos. Âmbito Cultural, Rio de Janeiro.
- Scarano FR (2002) Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian Atlantic Rainforest. *Ann Bot* 90:517-524. <http://dx.doi.org/10.1093/aob/mcf189>
- Silva KJP, Souza AF (2018) Common species distribution and environmental determinants in South American coastal plains. *Ecosphere* 9:1-15. <https://doi.org/10.1002/ecs2.2224>
- Silva J LA, Souza AF, Jardim JG & Goto BT (2015) Community assembly in harsh environments: the prevalence of ecological drift in the heath vegetation of South America. *Ecosphere* 6:1-18. <https://doi.org/10.1890/ES14-00548.1>
- Silva KR, Melo JCF, Torres MR (2016) Variações fenotípicas em *Andira fraxinifolia* Benth. (Fabaceae) em duas fitofisionomias de Restinga. *Hoehnea* 43:237-245. <https://doi.org/10.1590/2236-8906-83/2015>
- Silva MM, Melo JCF (2017) Plasticidade da folha e lenho de cinco espécies lenhosas em duas áreas de restinga no Sul do Brasil. *Iheringia* 72:173-180.
- Smith SA, Brown JH (2018) Constructing a broadly inclusive seed plant phylogeny. *Am J Bot* 105:302-314. <https://doi.org/10.1002/ajb2.1019>

- Sobral-Souza T, Lima-Ribeiro MS (2017) De volta ao passado: Revisitando a história biogeográfica das florestas neotropicais úmidas. *Oecologia Aust* 21:93-107. <https://doi.org/10.4257/oeco.2017.2102.01>
- Sousa-Baena MS, Garcia LC, Peterson AT (2014) Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Divers Distrib* 20:369-381. <https://doi.org/10.1111/ddi.12136>
- Sperandii MG, Bazzichetto M, Acosta ATR, Barták V, Malavasi M (2019) Multiple drivers of plant diversity in coastal dunes: A Mediterranean experience, *Sci Total Environ* 652:1435-1444. <https://doi.org/10.1016/j.scitotenv.2018.10.299>.
- Stekhoven DJ, Bühlmann P (2012) Missforest — non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28:112-118. <https://doi.org/10.1093/bioinformatics/btr597>
- Stropp J, Ladle RJ, Malhado ACM, Hortal J, Gaffuri J, Temperley WH, Skøien JO (2016) Mapping ignorance: 300 years of collecting flowering plants in Africa. *Global Ecol Biogeogr* 25:1-12. <https://doi.org/10.1111/geb.12468>
- Swenson NG, Enquist BJ, Pither J, Kerkhoff AJ, Boyle B, Weiser MD et al (2011) The biogeography and filtering of woody plant functional diversity in North and South America. *Glob Ecol Biogeogr* 21:798–808. <https://doi.org/10.1111/j.1466-8238.2011.00727.x>
- Trabucco A, Zomer RJ (2019). Global High-Resolution Soil-Water Balance: CGIAR Consortium for Spatial Information. <http://www.cgiar-csi.org>. Accessed 11 October 2019.
- Villwock JE, Lessa GC, Suguio K, Angulo RJ, Dillenburg DR (2005) Geologia e geomorfologia de regiões costeiras. In: Souza CRG, Suguio K, Oliveira AMS, Oliveira PE (orgs) *Quaternário do Brasil*. Editora Holos, Ribeirão Preto, pp 94-107.

- Vitória AP, Alves LF, Santiago LS (2019) Atlantic forest and leaf traits: an overview. *Trees - Struct Funct* 33:1535-1547. <https://doi.org/10.1007/s00468-019-01864-z>
- Webb, CO, Ackerly, DD, McPeck, MA, Donoghue, MJ (2002) Phylogenies and community ecology. *Ann Rev Ecol Syst* 33:475-505. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150448>
- Wiens J, Graham C (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu Rev Ecol Syst* 36:519–539. <https://doi.org/10.1146/annurev.ecolsys.36.102803.095431>
- Yan Y, Yang X, Tang Z (2013) Patterns of species diversity and phylogenetic structure of vascular plants on the Qinghai-Tibetan Plateau. *Ecol Evol* 3:4584-4595. <https://doi.org/10.1002/ece3.847>
- Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG et al (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature* 506:89-92. <https://doi.org/10.1038/nature12872>

Supplemental Material

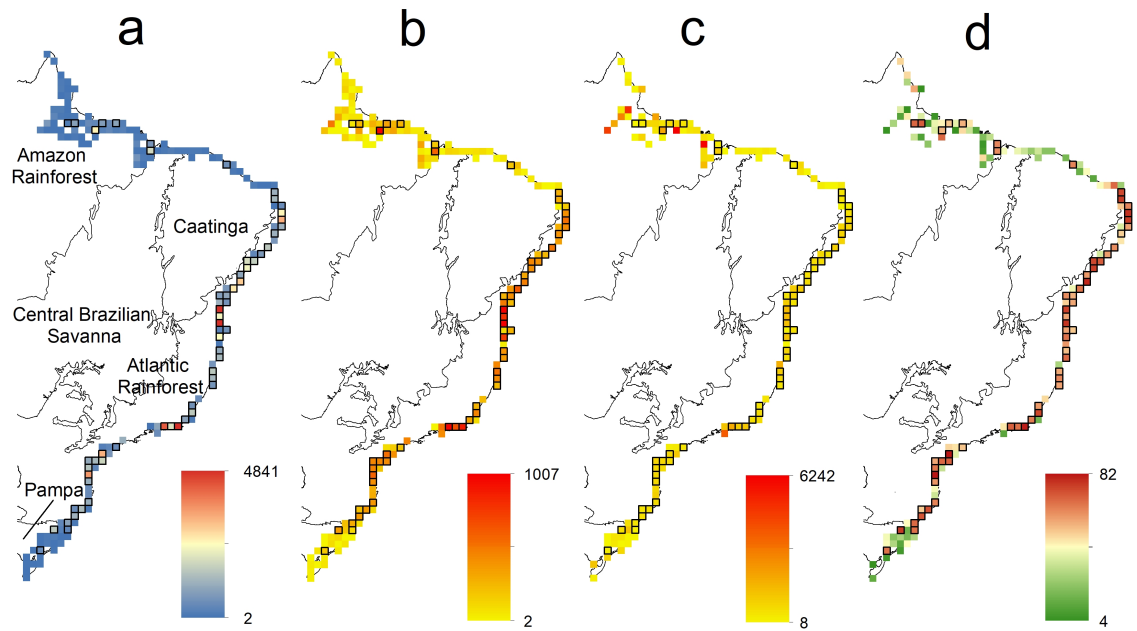


Fig. S1 Maps of **a** distribution of the number of records, **b** observed richness, **c** expected richness, and **d** inventory completeness for woody plants in the Brazilian coastal vegetation. Cells with the bolder border have higher knowledge (completeness $\geq 50\%$ and records number ≥ 100). Brazilian phytogeographic domains are delimited in gray. Grid cells have a 0.5° spatial resolution.

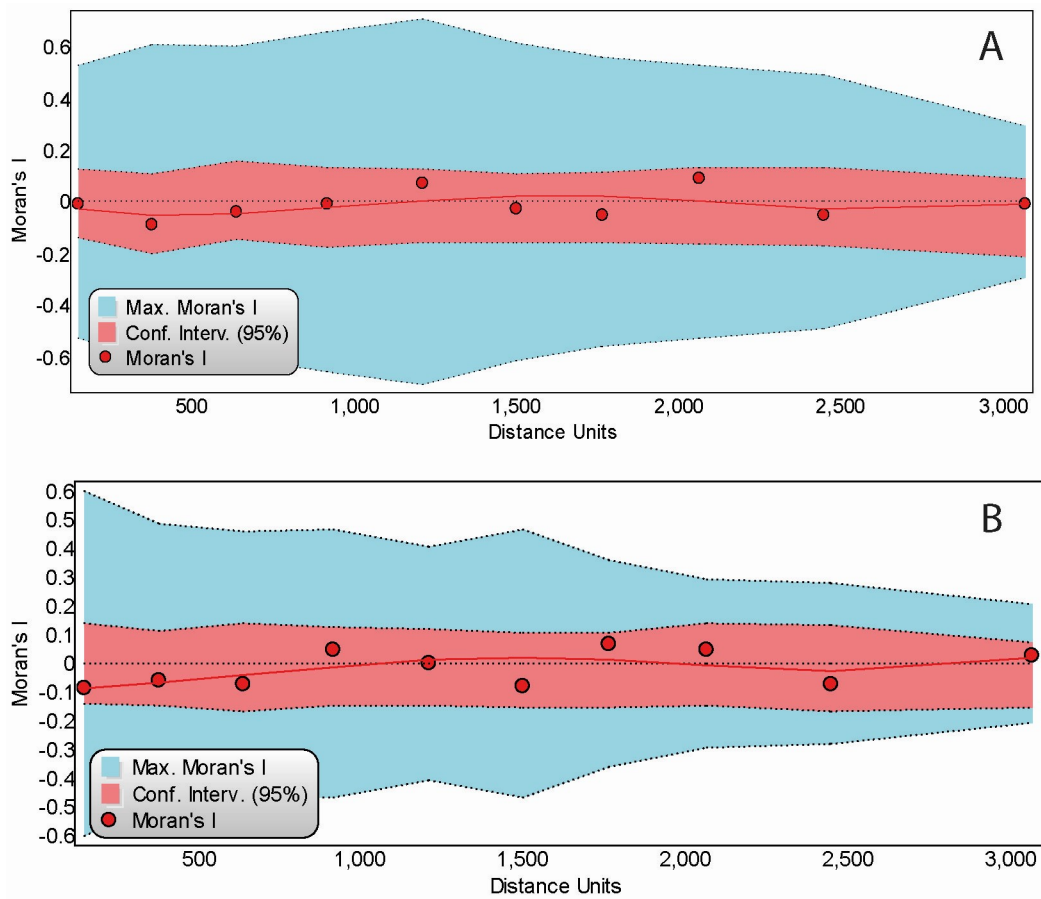


Fig. S2 Spatial correlograms of Moran's I index of the regression residuals of each better model tested: a) phylogenetic structure; and b) functional structure.

CAPÍTULO 3

**Diversidade-beta de múltiplos aspectos da diversidade de árvores em
ecossistemas costeiros da América do Sul**

Resumo

A investigação da diversidade beta em comunidades ecológicas permite elucidar sua dinâmica de formação e suas causas através das dissimilaridades nas composições de espécies. Através da integração de informações geográficas, filogenéticas e fenotípicas das espécies, e usando modelos de distribuição de espécies, investigamos para a vegetação de Restinga do Brasil (1) se a dissimilaridade observada entre assembleias difere de uma dissimilaridade esperada pelo acaso (i.e., decaimento de similaridade com a distância); (2) quantificamos a contribuição relativa dos componentes da diversidade beta – substituição e aninhamento – para todos os aspectos da diversidade beta; e (3) testamos, através de modelos lineares, diferentes hipóteses ambientais para os padrões observados, incluindo o efeito da disponibilidade hídrica, energia térmica, heterogeneidade de habitat, restrições edáficas, estabilidade climática e influência humana. A dissimilaridade taxonômica entre assembleias, em geral, diferiu do esperado por uma dinâmica de dispersão estocástica. Já os padrões de dissimilaridade filogenética e funcional obedeceram a uma dinâmica típica de maior dissimilaridade com a distância (i.e., como esperado pelo acaso). Entretanto, esses padrões apresentaram diferentes contribuições de substituição e aninhamento, com uma predominância de substituição na dissimilaridade taxonômica e filogenética e de aninhamento na dissimilaridade funcional. A disponibilidade de água parece ser o fator determinante dos padrões espaciais de dissimilaridade taxonômica, filogenética e funcional. Os resultados descrevem um cenário de acesso de diferentes espécies e linhagens oriundas de distintas fontes de espécies para a Restinga (i.e., as formações adjacentes) proporcional à diversidade dessas fontes, mas favorecendo espécies que possuem um conjunto de adaptações específicas para tolerar as restrições hídricas desses ambientes.

Palavras-chave: Composição de espécies, dimensões da biodiversidade, modelagem de distribuição de espécies, Montagem de comunidades, Restinga.

Introdução

Como os padrões da diversidade são formados e mantidos ao longo de gradientes ambientais permanece uma das questões mais persistentes da ecologia (Gaston 2000; Ricklefs 2004). Essa diversidade biológica – que expressa a variedade de organismos vivos – é essencialmente multifacetada, necessitando ser abordada em diferentes escalas espaciais e descrita por diferentes aspectos dos organismos, incluindo suas relações evolutivas (i.e., diversidade filogenética), diferenças fenotípicas (i.e., diversidade funcional), além da descrição tradicional baseada em unidades taxonômicas (i.e., diversidade taxonômica) (Magurran 2004; Magurran & McGill 2011; Garnier et al. 2015). Em cada uma dessas perspectivas, diferentes combinações de fatores parecem interagir em diferentes intensidades para gerar os padrões observados, dificultando significativamente uma compreensão geral sobre o fenômeno (e.g., González-Maya et al. 2016; Li et al. 2019). Entretanto, investigar os padrões geográficos de diversidade em diferentes contextos (i.e., envolvendo diferentes grupos biológicos em diferentes ambientes) continua sendo essencial, seja para elucidar a origem, causas e susceptibilidades de comunidades específicas, seja para ganhar uma compreensão geral através das circunstâncias em que cada fator atua nos padrões encontrados (Purvis & Hector 2000; Magurran & McGill 2011).

Com relação a escalas espaciais, a diversidade em comunidades ecológicas tem sido tradicionalmente descrita dentro ou entre localidades, compreendendo os componentes espaciais alfa (α) e beta (β) da biodiversidade, respectivamente, além ainda da escala regional ou gama (γ), esta menos explorada (Magurran 2004). A diversidade β descreve a mudança da composição de espécies ao longo do espaço geográfico (Whittaker 1972) e envolve dois componentes distintos, a substituição entre espécies (i.e., ‘turnover’) e a perda de espécies (‘nestedness’ ou aninhamento) entre assembleias, respectivamente (Baselga 2010). A compreensão dos padrões de

diversidade beta permite acessar diferentes aspectos da origem, distribuição e manutenção da biodiversidade em comunidades ecológicas (Qian et al. 2021).

Algumas questões relevantes incluem, por exemplo, se os padrões de assembleias resultam de colonização estocástica ou de processos determinísticos na distribuição espacial das espécies (Mori et al. 2018) e se as diferenças de composição de espécies decorrem de conjuntos diferentes ou subconjuntos das mesmas espécies (i.e., substituição e aninhamento, respectivamente). Por exemplo, aninhamento pode refletir a influência da dinâmica seletiva de colonização e extinção determinadas pelas tolerâncias das espécies às condições ambientais (Whittaker & Fernández-Palacios 2007; Si et al. 2016). Por sua vez, substituição pode indicar uma série de mecanismos, como seleção de espécies (*sorting*) por gradientes ambientais, limitações de dispersão, dinâmica histórica de extinção e colonização (Svenning et al. 2011; Si et al. 2016). Um terceiro aspecto relevante envolve a relação entre diversidade beta e os fatores ambientais, já que estes podem indicar os processos responsáveis pelas diferenças na composição entre comunidades (Dobrovolski et al. 2012).

Apesar de ser informativa sobre esses processos, a abordagem tradicional baseada na diversidade beta taxonômica pode não ser suficiente para uma apreciação completa dessas questões. Tanto as diferenças relativas aos antecedentes evolutivos quanto ao conjunto de adaptações das espécies entre comunidades podem ajudar a elucidar o contexto evolutivo dos padrões de estabelecimento de comunidades ao longo do espaço geográfico (Zhang et al. 2013; Jiang et al. 2021). Além disso, o uso de similaridade evolutiva e de características fenotípicas das espécies ao longo de assembleias permite uma compreensão mais profunda dos processos gerando variação em composição, como por exemplo, isolamento geográfico, restrições filogenéticas, interações bióticas, distância geográfica e filtragem ambiental (Graham & Fine 2008; Magurran & McGill 2011; Siefert et al. 2013). Neste sentido, a diversidade beta funcional e filogenética

forneem informação adicional de como fatores evolutivos e ecológicos interagem para influenciar diferenças composicionais entre assembleias (Graham & Fine 2008; Swenson 2011; Yang et al. 2015).

Essas questões permanecem pouco exploradas em ecossistemas costeiros tropicais, como a Restinga. Essa formação vegetal ocorre em planícies costeiras arenosas da costa leste da América do Sul, e possui características físicas próprias, tais como áreas com topografia plana, insolação e atividade eólica elevadas, e solos pobres em água e nutrientes (Lacerda et al. 1993). Devido à formação recente da Planície Costeira e o baixo número de espécies endêmicas na Restinga, tem-se sugerido que sua flora se originou em ecorregiões vizinhas e geologicamente mais antigas (Rizzini 1997; Scarano 2002). Neste sentido, a composição de espécies presentes na Restinga seria uma mistura das formações contíguas, como às florestas tropicais pluviais (i.e., Mata Atlântica e Amazônia) e sazonalmente secas (i.e., Caatinga) (Rizzini 1997; Fernandes & Queiroz 2015; Silva et al. 2015). Estudos recentes mostraram que as restrições de água limitam o número de espécies localmente (Oliveira et al. 2021) ao tempo em que favorecem ocorrência de espécies funcionalmente similares (Oliveira et al. submetido), portanto fornecendo suporte à hipótese da origem dessas formações vegetais. Contudo, uma questão relacionada, mas complementar aos estudos anteriores, é se esses padrões de estrutura de assembleia com alta similaridade funcional entre espécies dentro de comunidades resultam das mesmas espécies, linhagens e características, ou se esses processos são comuns, mas envolvem outras combinações de espécies, seus antecedentes evolutivos e conjunto de características distintas ao longo do espaço. Para essas questões, a investigação dos padrões de diversidade beta envolvendo aspectos taxonômicos evolutivos e funcionais constitui uma abordagem apropriada.

Este estudo aborda esse tema. Considerando os aspectos taxonômico, filogenético e funcional, investigamos os padrões espaciais da diversidade beta ao longo de

assembleias de Angiospermas arbóreas da Restinga do Brasil. Especificamente, avaliamos 1) se a composição observada das assembleias difere de uma composição esperada por uma dinâmica estocástica; 2) qual a contribuição de substituição e do aninhamento para os padrões observados; e 3) quais preditores ambientais melhor explicam esses padrões. Estudos recentes sugerem que as assembleias de Restinga resultam dos diferentes domínios fitogeográficos adjacentes (Amazônia, Cerrado, Caatinga, Mata Atlântica e Pampas), que possuem riqueza mais baixa que essas formações (Scarano 2002; BFG 2018) e que poucas linhagens específicas são capazes de tolerar suas condições ambientais limitantes, refletindo a adaptação de espécies filogeneticamente próximas entre si e funcionalmente redundantes (Amorim & Melo-Jr 2017; Massante e Gerhold 2020). Portanto, esperamos encontrar descontinuidades nas composições de espécies ao longo da área de estudo, o que resultaria em variação não-estocástica (i.e., processos determinísticos) da dissimilaridade entre assembleias ao longo do espaço. Além disso, esperamos encontrar uma maior contribuição da substituição para a diversidade beta taxonômica e uma maior contribuição do aninhamento para a diversidade beta filogenética e funcional. Por fim, em estudo recente, observou-se que a os padrões de similaridade entre espécies dentro das assembleias da região resultam predominantemente do efeito da disponibilidade de água e da deficiência de nutrientes nos solos (Oliveira et al. submetido). Assim, esperamos que esses mesmos fatores estejam mais relacionados aos padrões de diversidade beta do que outros fatores, como energia, heterogeneidade topográfica, estabilidade climática e influência humana.

Material e Métodos

Região de estudo e dados das espécies

Nossa área de estudo compreende a Restinga, vegetação que ocupa os solos arenosos da Planície Litorânea do Brasil (Lacerda et al. 1993). Determinamos o contorno da Restinga a partir da sobreposição dos limites da Planície Litorânea do Brasil e das formações associadas a essa vegetação (IBGE 2017; MMA 2017; veja também Oliveira et al. 2021). Uma grade de células de resolução de 55 x 55 km (0,5° x 0,5° de latitude-longitude) foi construída sobre este contorno por meio do software *Spatial Analyses in Macroecology* (SAM; Rangel et al. 2010), mantendo apenas as células com ao menos 50% de área sobreposta ao polígono da área de estudo (e.g., Dobrovolski et al. 2012).

Extraímos registros de plantas com ocorrência para o contorno da Restinga através do pacote `BIEN` em R (R Development Core Team 2018; Maitner et al. 2017). Este pacote permite acesso ao banco de dados do BIEN (*Botanical Information and Ecology Network*), que compreende registros botânicos acurados e padronizados para o Novo Mundo (Maitner et al. 2017). Nós selecionamos apenas registros de árvores nativas de Angiospermas com nomes binomiais válidos utilizando o pacote `flora` em R (R Development Core Team 2018; Carvalho 2020).

Nós definimos três atributos funcionais de plantas para as espécies presentes na Restinga, sendo duas contínuas, (1) altura máxima (m) e (2) área foliar (mm²), e uma categórica, (3) síndrome de dispersão (biótica e abiótica). Estes atributos expressam características funcionais relevantes. A altura máxima está relacionada à habilidade competitiva e resistência a estresse abiótico. A área foliar reflete adaptações para aquisição de recursos e controle de evapotranspiração. A síndrome de dispersão, por sua vez, está ligada à própria capacidade de dispersão, mas também às interações possíveis (e.g., espécies zoocóricas) (Cornelissen et al. 2003). Valores de atributos funcionais foram obtidos com o pacote `BIEN` em R (R Development Core Team 2018; Maitner et al. 2017) ou extraídos da base de dados do TRY - *Plant Trait Database* (Kattge et al. 2020).

Preditores ambientais

Inicialmente selecionamos 23 variáveis ambientais e agrupamos em diferentes modelos lineares múltiplos, que por sua vez, representam hipóteses ecológicas para os padrões da diversidade em amplas escalas espaciais (e.g., Francis & Currie 2003; Hawkins et al. 2003; Kreft & Jetz 2007; Qian et al. 2013). Essas hipóteses incluem o efeito da disponibilidade hídrica, energia térmica, restrições edáficas, heterogeneidade de habitat, estabilidade climática e influência humana. A hipótese de disponibilidade de água foi descrita pela precipitação anual (PRE), precipitação do mês mais chuvoso (PRE-WET), precipitação do mês mais seco (PRE-DRY), sazonalidade da precipitação (PRE-SAZ) (Fick & Hijmans 2017), evapotranspiração potencial (PET) e evapotranspiração real (AET) (Trabucco & Zomer 2019). A hipótese de disponibilidade de energia foi expressa pela temperatura média anual (TEMP), variação diurna média da temperatura (TEMP-VAR), sazonalidade da temperatura (TEMP-SAZ), temperatura máxima do mês mais quente (TEMP-MAX) e temperatura mínima do mês mais frio (TEMP-MIN) (Fick & Hijmans 2017). A hipótese de fatores edáficos foi constituída por disponibilidade de nutrientes (NUT) e salinidade (SAL) (Fischer et al. 2008), além de porosidade (PORO), capacidade de retenção de água (CAP) e densidade de carbono (C) e nitrogênio (N) (Global Soil Data Task Group 2000). A hipótese da heterogeneidade de habitat foi descrita pela altitude (ALT) e inclinação da superfície (SLOPE) (Fischer et al. 2008). A hipótese de estabilidade climática compreendeu a mudança histórica na precipitação (PRE-HIST) e na temperatura (TEMP-HIST). Ambas as medidas representam a diferença entre a precipitação e a temperatura entre o último glacial máximo (~21.000 anos) e o presente (CCSM4; Fick & Hijmans 2017). Por fim, a hipótese de influência humana incluiu a densidade populacional (DEN) (CIESIN 2016) e infraestrutura (INFRA) (Fischer et al. 2008). Este último representa a porcentagem de terrenos urbanos ou construídos por célula da grade (Fischer et al. 2008). Em seguida, checamos

a correlação entre essas variáveis dentro de cada hipótese (i.e., cada modelo linear múltiplo) através de uma matriz de correlação de Pearson. Foram mantidas apenas as variáveis com coeficiente de correlação $r < 0,7$ com cada outra, sendo eliminadas três variáveis: TEMP e TEMP-MIN do modelo de disponibilidade de energia e PRE-DRY do modelo da água.

Modelos de distribuição de espécies

A investigação da diversidade beta requer a comparação de todo o conjunto de assembleias da área de estudo, incluindo assembleias contíguas no espaço geográfico. Isso por que se espera que assembleias mais próximas sejam mais semelhantes entre si do que assembleias mais distantes. A análise incluindo apenas as assembleias bem conhecidas envolveria assembleias geralmente mais distantes entre si, o que resultaria em uma predominância de assembleias mais dissimilares e com elevada substituição. Entretanto esse resultado poderia não refletir o padrão real de dissimilaridade entre assembleias. Devido a esse efeito, utilizamos uma abordagem baseada em composição de espécies obtida por modelos de distribuição de espécies (SDMs), que garantem uma caracterização de todas as assembleias ao longo da área de estudo (e.g., Zwiener et al. 2018; Allen et al. 2019).

Inicialmente, definimos 17 contornos de sensibilidade a partir da ocorrência das espécies nas células da grade e nos limites dos domínios fitogeográficos (Amazônia, Cerrado, Caatinga, Mata Atlântica e Pampa; IBGE 2017). Cada contorno representa a distribuição potencial máxima (área de *buffer*) de um grupo de espécies (i.e., área historicamente acessível), seja para um único domínio, ou uma combinação destes. Por exemplo, para as espécies com distribuição restrita a Mata Atlântica, um contorno de sensibilidade foi definido a partir dos limites deste domínio. De maneira similar, para as espécies com distribuição restrita à Amazônia e Cerrado, um contorno de sensibilidade

foi definido a partir dos limites de ambos. Este mesmo critério foi adotado para a definição dos demais contornos de sensibilidade (Fig. S1). Nesta etapa foram utilizados os programas SAM e QGIS (Rangel et al. 2010; QGIS Development Team 2020).

Para caracterizar o nicho climático das espécies, selecionamos camadas ambientais com base na sua proporção de importância para a explicação da distribuição das ocorrências de árvores. Nós selecionamos dez variáveis ambientais (TEMP, TEMP-VAR, TEMP-SAZ, TEMP-MAX, TEMP-MIN, PRE, PRE-WET, PRE-DRY, PRE-SAZ e ALT) e as recortamos de acordo com os limites dos contornos de sensibilidade. Para cada contorno nós selecionamos diferentes combinações de variáveis ambientais, considerando a colinearidade entre as variáveis, mantendo apenas as variáveis com $r < 0,7$ entre si (Tab. S1). A manipulação dos dados, detecção e a remoção de multicolinearidade foi realizada com os pacotes `raster` (Hijmans et al. 2017) e `virtualspecies` (Leroy et al. 2016) em ambiente R (R Development Core Team 2018).

Os modelos foram construídos com o algoritmo Maxent (Phillips et al. 2006). Este método utiliza apenas dados de presença e apresenta alta performance com baixo número de pontos de ocorrência (Elith et al. 2006; Phillips et al. 2006; Wisz et al. 2008). Para os modelos de distribuição, utilizamos apenas espécies com pelo menos 15 pontos de ocorrências. Nós adotamos este critério para garantir uma maior acurácia nos modelos construídos (e.g., Zwiener et al. 2018). Os modelos baseados em Maxent requerem pontos de *background* para caracterizar o nicho ecológico das espécies. Aqui definimos 100 pontos de *background* e realizamos um procedimento de validação cruzada (*K-Fold*) para cada uma das 50 repetições. Para converter os mapas contínuos em binários, definimos um limiar de corte (*threshold*) de adequabilidade dos modelos baseado na soma entre a especificidade e a sensibilidade do modelo. Os modelos preditivos foram obtidos a partir da média das repetições.

Avaliamos a acurácia dos modelos através da área sobre a curva (AUC, *Area Under the ROC Curve*) e o coeficiente Kappa. O AUC varia de 0 a 1, sendo, 1 a indicar um alto desempenho do modelo e 0,5 um baixo (Elith et al. 2006; Phillips et al. 2006). Por sua vez, o Kappa varia de -1 a +1, sendo que valores $> 0,75$ indicam uma excelente performance, entre 0,40 e 0,75 uma boa performance e $< 0,40$ uma performance fraca (Monserud & Leeemans 1992). Nesse sentido, modelos aceitáveis precisam ser maiores do que 0,40 (Weng & Zhou 2006). Apenas modelos com $AUC \geq 0,75$ e $Kappa > 0,40$ foram mantidos para as análises. A construção dos modelos preditivos foi realizada com o pacote SSDM em R (R Development Core Team 2018; Schmitt et al. 2017).

Uma consequência da utilização de SDMs é a dependência da qualidade dos modelos para a retenção das espécies nas assembleias. Ou seja, espécies cujos modelos foram inadequados não foram incluídas, o que poderia afetar a estimativa de diversidade beta. Para avaliar essa possível fonte de viés das análises de diversidade beta baseadas em SDMs, comparamos os resultados desta abordagem com as medidas de diversidade beta considerando apenas as células identificadas como bem amostradas (veja abaixo).

Diversidade beta taxonômica, filogenética e funcional

Inicialmente, nós geramos uma matriz de incidência utilizando a distribuição predita das espécies pelos modelos de distribuição para as células da grade. Então, foram preparados os conjuntos de dados necessários para calcular as matrizes de dissimilaridade taxonômica, filogenética e funcional, bem como realizar a partição da diversidade beta para cada medida de diversidade. Para as análises da diversidade beta filogenética, nós utilizamos a megafilogenia ‘ALLOTB’, elaborada por Smith e Brown (2018). Esta filogenia, disponibilizada em nível de espécies, foi construída a partir de dados moleculares e de relações taxonômicas (Smith e Brown 2018). Espécies presentes na matriz de incidência e ausentes nesta filogenia (26 espécies ou 7%) foram incluídas

aleatoriamente nos clados representando seus gêneros ou famílias (e.g., Qian e Zhang 2016). Nesta etapa, foram utilizados os pacotes `ape` (Paradis et al. 2004), `phytools` (Revell 2012) e `V. PhyloMaker` (Jin e Qian 2019) em R (R Development Core Team 2018).

Para as análises da diversidade beta funcional são necessários valores de atributos funcionais das espécies presentes na matriz de incidência. Entretanto, na busca realizada em bases de dados, foram observadas lacunas de valores para área foliar (15%) e síndromes de dispersão (11%). Estas lacunas foram resolvidas por meio de uma abordagem de imputação baseada em algoritmos de ‘*random forest*’ (Pantanowitz & Marwala 2009). Nós realizamos a imputação com 92% dos atributos conhecidos, superior ao mínimo recomendado para realizar o procedimento (70%) (Penone et al. 2014). Nós avaliamos o desempenho da imputação a partir de duas medidas, o NRMSE (*Normalized Root Mean Squared Error*) e o PFC (*Proportion of Falsely Classified*). Enquanto o primeiro avalia a imputação de atributos contínuos, o segundo avalia a imputação de atributos categóricos. Essas duas medidas se aproximam de zero quando a estimativa é precisa. Uma estimativa aleatória dos atributos funcionais é observada quando as medidas se aproximam de um (Li et al. 2015; Lenz et al. 2020). Ao final, nós obtivemos um valor de erro baixo (NRMSE = 0,28; PFC = 0,32), indicando uma boa acurácia da imputação. O procedimento de imputação foi executado em R (R Development Core Team 2018) com o pacote `missForest` (Stekhoven e Bühlmann 2012).

A partir dos atributos funcionais, geramos uma matriz expressando as distâncias funcionais pelo índice de Gower (Gower 1971). Nós computamos uma matriz de coordenadas com as distâncias funcionais (MCDF), utilizando os três primeiros eixos sintéticos sumarizando atributos funcionais obtidos de uma análise de coordenadas principais (PCoA) a partir da matriz de distância de Gower. A eficiência da MCDF foi

avaliada pelo desvio quadrado da média da distância de Gower (mSD). Neste procedimento, baixos valores representam alta qualidade no espaço funcional (Maire et al. 2015). Após avaliar a eficiência da MCDF, nós observamos um baixo valor de mSD (0,004), indicando uma alta qualidade no espaço funcional. Os procedimentos descritos foram executados em R (R Development Core Team 2018), com os pacotes `cluster` (Maechler et al. 2021) e `vegan` (Oksanen et al. 2018).

A partição da diversidade beta foi realizada para cada medida de diversidade (i.e., taxonômica, filogenética e funcional), sendo calculada a contribuição dos componentes de substituição e aninhamento (Baselga 2010). Este procedimento foi realizado utilizando a matriz de incidência para a diversidade beta taxonômica (índice de Jaccard); a matriz de incidência combinada com a filogenia para a diversidade beta filogenética (índice PhyloSor; Bryant et al. 2008); e a matriz de incidência combinada com a MCDF para a diversidade beta funcional (índice de Jaccard; Baselga 2010; Villéger et al. 2013). Neste mesmo procedimento, ainda foram geradas matrizes de dissimilaridade taxonômica, filogenética e funcional estimadas entre os pares de células da grade.

Além da partição realizada utilizando todas as células da grade, nós realizamos o particionamento utilizando diferentes combinações de células por domínios fitogeográficos adjacentes à distribuição da Restinga (i.e., Amazônia, Cerrado, Caatinga, Mata Atlântica e Pampa). Essa análise foi realizada com o objetivo de avaliar a contribuição dos componentes de substituição e aninhamento separadamente por domínio e checar se difere do padrão geral. Todos os passos descritos foram executados no software R (R Development Core Team 2018), utilizando os pacotes `vegan` (Oksanen et al. 2018), `betapart` (Baselga e Orme 2012), `picante` (Kembel et al. 2010) e a função ‘`multBetaDim`’ disponibilizada por Villéger et al. (2013).

Análises estatísticas

Para avaliar se a composição de espécies de plantas observada difere de uma composição determinada por processos estocásticos, nós geramos um modelo nulo com 999 repetições para testar a chance de descobrir valores de diversidade beta taxonômica, filogenética e funcional (i.e., matrizes de dissimilaridade) maiores do que o observado. Especificamente para a diversidade beta funcional, a abordagem multivariada adotada (Villéger et al. 2013) não permite uma exploração de um modelo nulo que seja análogo aos outros tipos de diversidade beta (i.e., taxonômico e filogenético). Dessa forma, usamos nesta etapa um procedimento alternativo que se baseia em uma matriz de dissimilaridade gerada através de uma matriz quadrada, expressando as distâncias funcionais pelo índice de Gower (Gower 1971).

Para testar as diferentes hipóteses ambientais para explicar os padrões de diversidade beta taxonômica, filogenética e funcional, agrupamos as variáveis em modelos representando cada hipótese e empregamos regressões múltiplas em matrizes de distância (MRMs) (Lichstein 2007). Para isso, as camadas ambientais foram transformadas em matrizes de dissimilaridade ambiental utilizando a distância Euclidiana. A dependência existente nos dados oriundos de uma abordagem pareada pode comprometer os valores de significância estatística das análises (Qian et al. 2013). Neste sentido, nós avaliamos a significância dos coeficientes de regressão e do R^2 através de um teste de permutação (= 999 repetições), procedimento similar ao adotado por outros autores (Qian et al. 2013; Zhang et al. 2013; Qian et al. 2020). Nós investigamos a relação entre a diversidade beta taxonômica, filogenética e funcional por testes de correlações de Mantel (Legendre & Legendre 2012). Todos os procedimentos descritos foram executados em R (R Development Core Team 2018), com os pacotes *vegan* (Oksanen et al. 2018), *ecodist* (Goslee e Urban 2007), *cluster* (Maechler

et al. 2021) e *picante* (Kembel et al. 2010) e os códigos presentes em Swenson (2014).

Avaliação da diversidade beta baseada em modelos de distribuições

A comparação dos resultados baseada em SDMs com a diversidade beta envolvendo todas as espécies conhecidas nas assembleias foi realizada utilizando apenas a diversidade beta taxonômica e filogenética medida entre pares de células da grade reconhecidas como bem amostradas. Utilizando os pontos de ocorrência das espécies, selecionamos as células que apresentaram uma completude de inventário $\geq 50\%$, combinado com um número de registros $\geq$ do que a sua média entre todas as células da grade. Esse critério permite selecionar apenas células bem conhecidas, reduzindo as lacunas de amostragem presentes nos dados (para detalhes, ver Oliveira et al. 2021). Então, nós obtivemos a composição observada de espécies para cada uma dessas células (= 41) e a árvore filogenética contendo essas espécies, de acordo com os mesmos procedimentos descritos anteriormente para a composição baseada nos SDMs (ver acima). Nós realizamos a partição da diversidade beta taxonômica e filogenética para as duas abordagens, i.e., baseada na composição por SDMs e observada. Finalmente, após eliminar colinearidade entre os preditores, foram realizadas as mesmas análises descritas para as assembleias obtidas por SDMs (como descrito acima). Nesta etapa, análises de completude de inventário foram realizadas em R, com o pacote *KnowBr* (Lobo et al. 2018). As demais análises utilizaram os mesmos pacotes descritos nas seções anteriores.

Resultados

Nós encontramos 40.178 registros de ocorrência, compreendendo 2.301 espécies de árvores ao longo da região costeira do Brasil. Nós realizamos a modelagem de nicho

ecológico com espécies de árvores presentes na região que apresentaram ao menos 15 pontos de ocorrência, o que restringiu o número de espécies analisadas a 594. Estas espécies foram aquelas que apresentaram uma maior distribuição na Restinga, ou seja, com o maior número de ocorrências nas células da grade (mínimo = 7 células; máximo = 100 células). Dessas 594 espécies, nós selecionamos para as nossas análises apenas as espécies que apresentaram modelos com uma acurácia suficiente, ou seja, com um $AUC \geq 0,75$ e $Kappa > 0,40$ (360 espécies ou 60%). Nós encontramos 18 modelos com acurácia excelente ($AUC > 0,90$ e $Kappa > 0,75$; 5% do total), 172 modelos com uma boa acurácia ($AUC = 0,80-0,89$ e $Kappa = 0,56-0,74$; 48%) e 170 modelos com acurácia aceitável ($AUC = 0,75-0,79$ e $Kappa = 0,41-0,55$; 47%) (Fig. S2). *Lacmellea bahiensis* J.F.Morales foi a espécie que obteve o maior valor de AUC (0,97) e $Kappa$ (0,84). A riqueza obtida pela sobreposição dos *ranges* das espécies preditos pelos SDMs variou de cinco a 233 espécies (média = 68 espécies) nas células da grade (Fig. 1). Apenas nove espécies ocorreram em menos de dez células (2% do total), estando à maioria das espécies presentes em mais de trinta células (208 ou 58%).

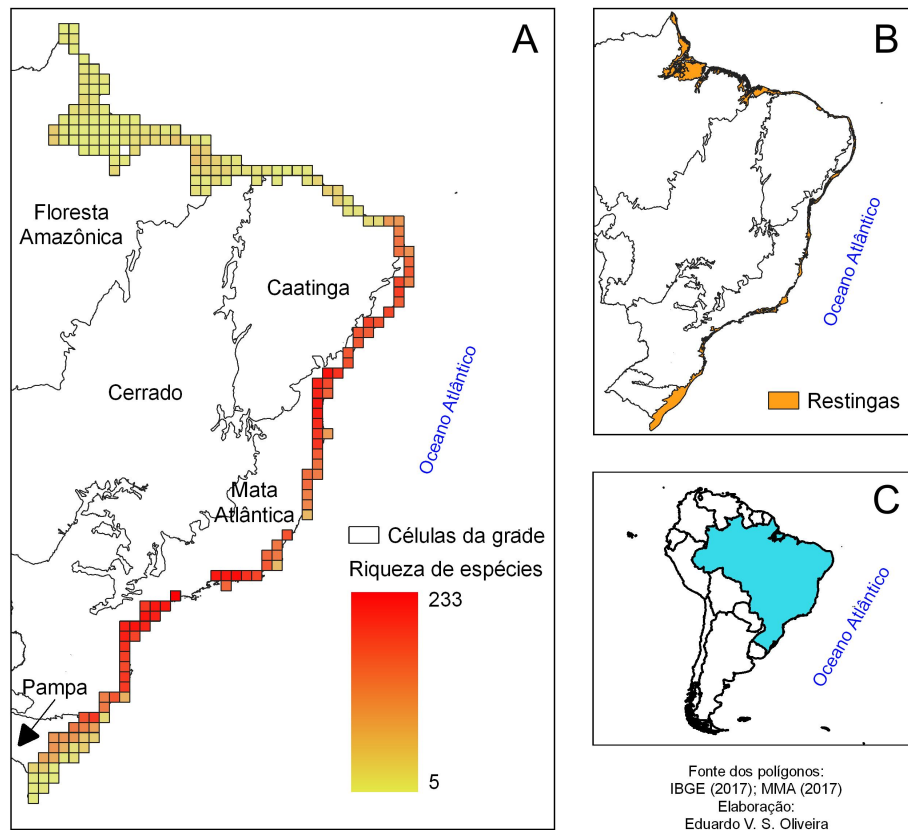


Figura 1. Mapas de (A) riqueza de espécies de árvores por assembleia ($0,5^\circ$ de resolução espacial) obtida pela sobreposição da área de distribuição geográfica das espécies preditas pelos modelos de distribuição de espécies (SDMs) ao longo da zona costeira do Brasil, (B) delimitação da vegetação de Restinga considerada para a modelagem de distribuição das espécies de árvores e (C) localização do Brasil na América do Sul.

Nós encontramos 17.485 (ou 86%) pares de assembleias com valores de dissimilaridade taxonômica diferindo de um modelo estocástico de montagem de assembleias. Desse valor, 7.877 (ca. de 39%) pares de assembleias apresentaram uma dissimilaridade menor do que o esperado pelo modelo nulo, enquanto 9.608 (ca. de 47%) pares de assembleias apresentaram uma dissimilaridade maior do que o esperado pelo modelo nulo. Mudanças na composição taxonômica entre pares de assembleias

foram explicadas pelo acaso em 2.715 (ca. de 14%) dos valores de dissimilaridade taxonômica. Em contraste, na maioria dos valores de dissimilaridade filogenética e funcional foi explicada pelo acaso (filogenética = 16.949 pares [ca. de 83%]; funcional = 17.465 [ca. de 85%]), enquanto menos pares de assembleias diferiram de um modelo estocástico de montagem de assembleias (filogenética = 3.352 pares [ca. de 17%]; funcional = 2.937 [ca. de 15%]). Desses valores, a maioria dos pares de assembleias apresentou uma dissimilaridade maior do que a esperada pelo modelo nulo (filogenética = 1.695 pares [ca. de 8,6%]; funcional = 1.730 [ca. de 9%]), em oposição aqueles pares que apresentaram uma dissimilaridade menor do que a esperado pelo modelo nulo (filogenética = 1.657 pares [ca. de 8,4%]; funcional = 1.207 [ca. de 6%]).

Nós encontramos uma maior contribuição da substituição na dissimilaridade taxonômica e filogenética ao longo da região costeira do Brasil (Tab. 1). Em contraste, encontramos uma maior contribuição do aninhamento para a dissimilaridade funcional (Tab. 1). Com relação à análise por domínio fitogeográfico, a substituição foi mais importante em todos os domínios, especialmente na Mata Atlântica e Floresta Amazônica, tanto para diversidade beta taxonômica como para filogenética (Fig. 2). A proporção da substituição foi menor para a diversidade beta filogenética em comparação a diversidade beta taxonômica. Além disso, nós observamos um único caso de aninhamento filogenético para o domínio Pampa (Fig. 2). No caso da diversidade beta funcional, o aninhamento foi o padrão mais comum em todos os domínios fitogeográficos (Fig. 2).

Tabela 1. Valores de particionamento da diversidade beta taxonômica, filogenética e funcional de assembleias de árvores ao longo da região costeira do Brasil. JTU: valor de substituição medido pelo índice de Jaccard, JNE: valor do aninhamento medido pelo índice de Jaccard, JAC: valor da diversidade beta geral medida pelo índice de Jaccard, SIM: valor de substituição medido pelo índice de Sørensen, SNE: valor de aninhamento medido pelo índice de Sørensen, SOR: valor da diversidade beta geral medida pelo índice de Sørensen.

<u>β-Taxonômica</u>			<u>β-Filogenética</u>			<u>β-Funcional</u>		
JTU	JNE	JAC	SIM	SNE	SOR	JTU (±)	JNE (±)	JAC
						0,247	0,539	
0,975	0,017	0,993	0,932	0,050	0,983	(0,198)	(0,307)	0,787
98%	2%	100%	95%	5%	100%	31%	69%	100%

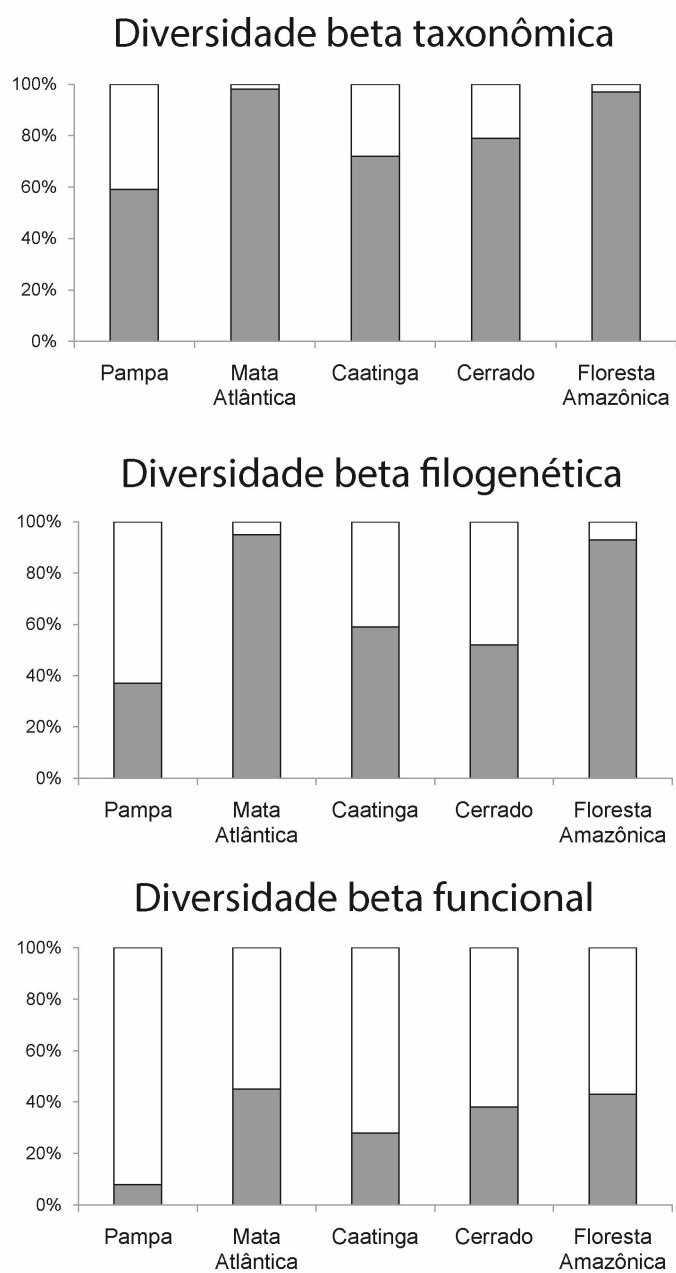


Figura 2. Comparações entre os valores proporcionais de contribuição de substituição (cinza) e aninhamento (branco) para a diversidade beta taxonômica, filogenética e funcional dos conjuntos de assembleias pertencentes aos domínios fitogeográficos presentes na região costeira do Brasil.

Os modelos de regressão lineares por matrizes de distância indicaram que a hipótese da disponibilidade hídrica possui o maior poder de explicação da dissimilaridade beta taxonômica, filogenética e funcional de espécies de árvores para a região costeira do Brasil (Tab. 2). A evapotranspiração potencial (PET) foi o preditor com maior contribuição (i.e., coeficiente de regressão padronizado) nas regressões lineares realizadas entre o modelo da água e as dissimilaridades beta taxonômica e filogenética. Por sua vez, precipitação média anual (PRE) foi o preditor com maior contribuição nas regressões lineares realizadas entre o modelo da água e a dissimilaridade funcional (Tab. 2). Nós observamos que os fatores associados ao modelo da energia possuem uma importância secundária para os padrões de diversidade beta taxonômica, filogenética e funcional. A temperatura máxima do mês mais quente foi à variável com maior contribuição nas regressões lineares realizadas entre o modelo da energia com as dissimilaridades taxonômica, filogenética e funcional. Os modelos da água e energia apresentaram o maior poder de explicação (R^2) para a diversidade beta filogenética, seguido da diversidade beta taxonômica e funcional (Tab. 2).

Tabela 2. Comparações entre modelos lineares múltiplos e seus componentes para os padrões da diversidade beta taxonômica, filogenética e funcional de árvores ao longo da região costeira do Brasil. Valores de *slope* e R^2 são padronizados. * = $p < 0,05$.

Modelo	Preditor	β -taxonômica		β -filogenética		β -funcional	
		<i>Slope</i>	R^2	<i>Slope</i>	R^2	<i>Slope</i>	R^2
Energia	TEMP-VAR	0,034*	0,255*	0,010	0,270*	-0,050	0,075*
	TEMP-MAX	0,287*		0,410*		0,369*	
	TEMP-SAZ	0,255*		0,130*		-0,160*	
Água	PRE	0,133*	0,362*	0,180*	0,373*	0,410*	0,216*
	PRE-WET	0,179*		0,170*		-0,030	
	PRE-SAZ	0,186*		0,085*		0,070	
	AET	0,112*		0,064*		0,022	
	PET	0,266*		0,355*		0,170*	
Heterogeneidade	ALT	-0,040	0,034*	0,250	0,063*	0,210*	0,054*

		β -taxonômica		β -filogenética		β -funcional	
de habitat							
Edáfico	SLOPE	0,180*		-0,040*		-0,120*	
	PORO	0,007	0,052*	-0,016	0,052*	-0,007	0,052*
	NUT	-0,028		-0,070*		0,056	
	SAL	0,156*		0,187*		0,170*	
	C	0,034		0,031		-0,070	
	N	0,101*		0,113*		0,130*	
	CAP	0,085*		0,083**		0,120*	
Estabilidade climática	TEMP-HIST	0,190*	0,049*	0,140*	0,046*	-0,100*	0,040*
	PRE-HIST	0,140*		0,170*		0,150*	
Influência Humana	DEN	0,080*	0,012*	0,100	0,012	0,160	0,034
	INFRA	-0,060*		-0,040		-0,070	

Nós encontramos uma correlação de Mantel maior entre as matrizes de dissimilaridade taxonômica e filogenética ($r = 0,90$; $p < 0,01$) do que entre dissimilaridade funcional e filogenética ($r = 0,74$; $p < 0,01$) e que entre a dissimilaridade taxonômica e funcional ($r = 0,64$; $p < 0,01$).

Nossos resultados obtidos com os SDMs não diferiram dos resultados considerando todas as espécies conhecidas das células bem amostradas. Os resultados foram congruentes em se tratando do particionamento (Tab. S2) e dos modelos de regressão lineares por matrizes de distância (Tabs. S3 e S4) em ambos, diversidade beta taxonômica e filogenética. São apresentadas as células de maior conhecimento e as estimativas utilizadas para o seu reconhecimento, i.e., número de registros, riqueza observada, riqueza estimada e completude de inventário, como apêndice (Fig. S3).

Discussão

Nossos resultados revelaram que a substituição taxonômica e filogenética apresentou uma maior influência na variação geográfica da composição de assembleias de árvores ao longo da região costeira do Brasil, corroborando e expandindo resultados recentes

para a Restinga nas regiões sul e sudeste do Brasil (Inague et al. 2021). É importante destacar que uma alta substituição filogenética implica necessariamente em uma alta substituição taxonômica, mas não o inverso. Ou seja, é possível haver uma substituição taxonômica sem um padrão filogenético correspondente, nos casos em que espécies são substituídas por outras do mesmo clado (e.g., gênero) (Vellend 2010). Portanto, nossos resultados refletem uma tendência à substituição de diferentes linhagens ao longo do espaço. Essa maior substituição taxonômica e filogenética é característica da existência de mais de um *pool* regional de espécies ao longo da região (Svenning et al. 2011), dando, portanto, respaldo à hipótese de um aporte de diferentes conjuntos de espécies a partir dos distintos domínios fitogeográficos adjacentes à Restinga (Silva et al. 2015). Nessa vegetação, a dispersão possui papel primário na condução da variação florística (Silva & Souza 2018a) e as distintas formações adjacentes (i.e., Florestas Tropicais Pluviais e Sazonalmente Secas) serviriam de fonte da flora taxonomicamente e filogeneticamente variada ao longo da Restinga.

Ao contrário da variação na identidade das espécies (i.e., dissimilaridade taxonômica e filogenética), nós observamos um predomínio de aninhamento funcional entre as assembleias, com uma similaridade nas espécies entre assembleias distintas quanto a seus atributos. Esse resultado indica uma tendência à seleção de espécies com adaptações específicas às condições ambientais típicas da Restinga, conforme encontramos em resultados recentes de uma tendência à maior similaridade funcional dentro das assembleias, especialmente no Nordeste do Brasil (Oliveira et al. submetido). As condições da Restinga favorecem espécies de diferentes linhagens, como descrito acima, acrescentando assim um elemento importante para a compreensão da variação da diversidade de plantas nas zonas costeiras neotropicais. Apesar da grande diversidade de linhagens, as espécies apresentam atributos funcionais similares. Por exemplo, das 360

espécies utilizadas para elaborar os modelos de distribuição, cerca de 92% delas possuem área foliar menor que 20 cm² e 80% possuem altura menor que 20 m (Fig. 3).

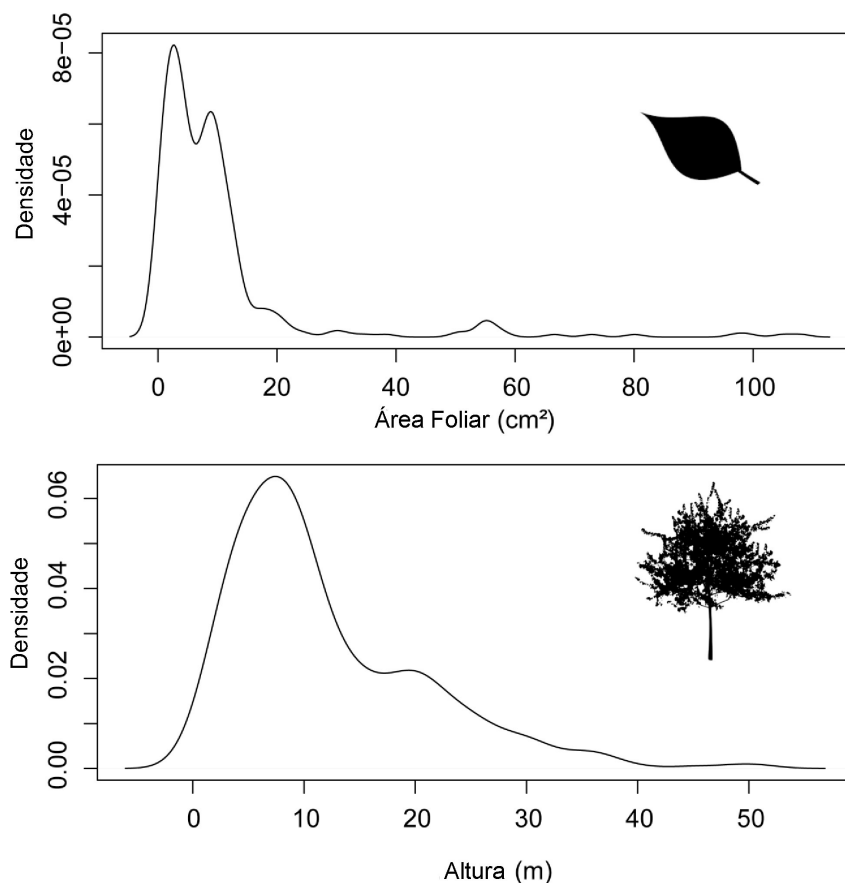


Figura 3. Distribuição de frequência de densidade dos atributos funcionais contínuos (área foliar e altura da planta) obtidos para as 360 espécies de árvores utilizadas para as análises de diversidade beta na região costeira do Brasil.

A predominância de dissimilaridade taxonômica diferente do esperado pelo acaso reforça a ideia de descontinuidades entre os diferentes *pools* de espécies ao longo da costa. Contudo, o padrão predominantemente estocástico observado na dissimilaridade filogenética reflete um padrão típico de decaimento de similaridade com a distância, neste caso com substituição de linhagens ao longo da região (e.g., Qian et al. 2013). Em conjunto, esses resultados indicam uma substituição gradual de linhagens ao longo da

área de estudo, com a ocorrência de espécies geograficamente restritas (Siefert et al. 2013; Yang et al. 2015). Embora haja muitas espécies oriundas de clados dominantes e amplamente distribuídos na Restinga (e.g., famílias Fabaceae e Myrtaceae), assembleias locais parecem resultar de diferentes subclados, resultado da separação evolutiva entre *pools* regionais (Daru et al. 2017).

Por sua vez, a diversidade beta funcional também foi influenciada por processos neutros, consistente com um padrão de decaimento de similaridade com a distância, porém, resultante de um maior aninhamento funcional. Este resultado acrescenta um elemento importante neste cenário de substituição gradual de linhagens e espécies majoritariamente distintas, em que a mudança nas espécies é acompanhada por uma redução no conjunto de tipos funcionais entre assembleias (como é característico do padrão de aninhamento) (Siefert et al. 2013). Esse resultado é também consistente com resultados recentes (Oliveira et al. submetido) que mostram uma redução na variabilidade funcional dentro das assembleias em regiões específicas, como o Nordeste do Brasil, onde os efeitos das condições abióticas são mais agudos.

Esses resultados discutidos acima, sobretudo relativo às diferenças regionais são corroborados com a análise de partição da diversidade beta taxonômica e filogenética por domínios fitogeográficos. Na Mata Atlântica e Amazônia, por exemplo, encontramos um predomínio da substituição. Além de estar refletindo o legado biogeográfico dos domínios adjacentes (e.g., Pinto-Ledezma et al. 2018), a maior contribuição da substituição para assembleias da Mata Atlântica e Amazônia pode estar relacionada à maior riqueza de espécies em relação aos demais (BFG 2018; Pinto-Ledezma et al. 2018). Ou seja, os maiores *pools* de espécies desses ambientes oferecem uma maior diversidade de espécies potenciais para colonizar os ambientes costeiros. Neste contexto, quais espécies disponíveis irão se estabelecer será função dos seus padrões de ocupação local e ordem de colonização (i.e., efeito prioritário; Drake 1991;

De Meester et al. 2016). O padrão encontrado na Restinga associadas aos Pampas reforça essa conclusão. Nesta região encontramos a menor contribuição da substituição para a diversidade beta taxonômica, uma maior contribuição do aninhamento para a diversidade beta filogenética e uma maior contribuição do aninhamento na diversidade beta funcional, o que, portanto, estaria relacionado à menor riqueza de espécies de árvores neste domínio em relação aos demais (BFG 2018). No caso da dissimilaridade funcional, a predominância do aninhamento em todos os domínios reforça a conclusão acima sobre o papel das condições abióticas na zona costeira na seleção de espécies com adaptações específicas.

Nossos resultados foram consistentes com relação à importância do modelo da água e importância secundária do modelo da energia sobre as dimensões da diversidade beta avaliadas (Oliveira et al. 2021; Massante e Gerhold 2020; Oliveira et al. submetido). Na região costeira do Brasil, são reconhecidos diferentes tipos climáticos, variando, marcadamente, quanto à sazonalidade da precipitação. Enquanto algumas partes da costa são menos influenciadas por secas no litoral (e.g., região sudeste), outras partes possuem até seis meses secos (e.g., costa setentrional ou semiárida) (IBGE 2002). Essa heterogeneidade no regime de precipitação, combinado com as variações de temperatura — que podem afetar a disponibilidade de água no ambiente (Moeslund et al. 2013) — contribuem por causar variação na diversidade beta taxonômica, filogenética e funcional de árvores na região. Em ambientes secos, variações florísticas ao longo do espaço geográfico respondem a variação na disponibilidade de água e temperatura, incluindo a Restinga (Silva & Souza 2018a) e florestas sazonalmente secas como a Caatinga (Silva & Souza 2018b).

A correspondência observada entre as três dimensões com relação ao efeito da água e secundariamente da energia pode ser explicada devido ao grau de correlação existente entre estas dimensões (Qian et al. 2013; Siefert et al. 2013; Yang et al. 2015;

Pinto-Ledezma et al. 2018). Entretanto, como mostramos, esses padrões guardam aspectos próprios quanto à sua variação espacial (i.e., se refletem estocasticidade ou não) e quanto à contribuição dos componentes de substituição e aninhamento. Essas diferenças, especialmente em função da dissimilaridade filogenética e funcional permitem um acessar o contexto biogeográfico (i.e., que e como linhagens evolutivas contribuíram) na formação das assembleias e como fatores ecológicos locais modulam as diferenças composicionais das assembleias (Wong et al. 2018). Por exemplo, a maior quantidade de explicação dos modelos para beta diversidade filogenética (i.e., maior coeficiente de determinação), sugere que a dissimilaridade filogenética possa estar captando um maior efeito dos determinantes ambientais, e assim, fornecendo mais informações sobre os fatores que estruturam as assembleias de Restinga (Swenson 2011).

Em síntese, neste estudo, apresentamos a primeira análise de diferentes aspectos de diversidade beta de plantas ao longo da Restinga brasileira. Fornecemos a evidência mais contundente até o presente da contribuição dos diferentes *pools* regionais de espécies ao longo da extensão da costa brasileira através dos padrões de dissimilaridade taxonômica, filogenética e funcional das espécies. Além disso, reforçamos resultados recentes da predominância de restrições abióticas – especificamente a disponibilidade de água – como o fator responsável pela variação geográfica na composição de plantas arbóreas na região. Mostramos como a variação na biodiversidade das formações da América do Sul parece está intimamente relacionada ao provimento de diferentes espécies e linhagens para as assembleias mais recentes das zonas costeiras, ao passo em que as condições locais limitam o estabelecimento das espécies com adaptações específicas para suportar as condições locais limitantes.

Referências

- Allen JM, Germain-Aubrey CC, Barve N, Neubig KM, Majure LC & Laffan SW et al. 2019. Spatial Phylogenetics of Florida Vascular Plants: The Effects of Calibration and Uncertainty on Diversity Estimates. *iScience* 11: 57-70.
- Amorim MW & Melo-Jr JCF. 2017. Functional diversity of restinga shrub species on the coastal plain of Southern Brazil. *Int J Dev Res* 7: 13189-13202.
- Baselga A. 2010. Partitioning the turnover and nestedness components of betadiversity. *Glob Ecol Biogeogr* 19: 134-143.
- Baselga A & Orme CDL. 2012. betapart: an R package for the study of beta diversity. *Methods Ecol. Evol.* 3: 808-812.
- BFG - Brazil Flora Group. 2018. Brazilian flora 2020: innovation and collaboration to meet target 1 of the global strategy for plant conservation (GSPC). *Rodriguésia* 69: 1513-1527.
- Bryant JB, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ & Green JL. 2008. Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *PNAS* 1: 11505-11511.
- Carvalho G. 2020. Tools for Interacting with the Brazilian Flora 2020. Disponível em: <<http://www.github.com/gustavobio/flora>>. Acesso em: maio/2020.
- CIESIN - Center for International Earth Science Information Network. 2016. Documentation for the Gridded Population of the World, Version 4 (GPWv4). Disponível em: <<http://www.ciesin.org/>>. Acesso em: novembro/2018.
- Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE et al. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust J Bot* 51: 335-380.

- Daru BH, Elliott TL, Park DS & Davies TJ. 2017. Understanding the Processes Underpinning Patterns of Phylogenetic Regionalization. *Trends Ecol Evol* 32: 845-860.
- De Meester L, Vanoverbeke J, Kilsdonk LJ & Urban MC. 2016. Evolving perspectives on monopolization and priority effects. *Trends Ecol Evol* 31: 136-146.
- Dobrovolski R, Melo AS, Cassemiro FAZ & Diniz-Filho JAF. 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Glob Ecol Biogeogr* 21: 191-197.
- Drake JA. 2011. Community-assembly mechanics and the structure of an experimental species ensemble. *Am Nat* 137: 1-26.
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S & Guisan A et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.
- Emerson BC & Gillespie RG. 2008. Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* 23: 619-630.
- Fernandes MF & Queiroz LP. 2015. Floristic surveys of Restinga Forests in southern Bahia, Brazil, reveal the effects of geography on community composition. *Rodriguésia* 66: 51-73.
- Fick SE & Hijmans RJ. 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37: 4302-4315.
- Fischer G, Nachtergaele F, Prieler S, van Velthuisen HT, Verelst L & Wiberg D. 2008. Global agro-ecological zones assessment for agriculture. Disponível em: <<http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases>>. Acesso em: novembro/2018.
- Francis AP & Currie DJ. 2003. A globally consistent richness-climate relationship for angiosperms. *Am Nat* 161: 523-536.

- Garnier E, Navas M-L & Grigulis K. 2015. Plant functional diversity: organism traits, community structure, and ecosystem properties. Oxford: Oxford University Press, 249 p.
- Gaston KJ. 2000. Global patterns in biodiversity. *Nature* 405: 220-227.
- Global Soil Data Task Group. 2000. Global gridded surfaces of selected soil characteristics. Disponível em: <<https://webmap.ornl.gov/ogcdown/index.jsp>>. Acesso em: novembro/2017.
- González-Maya JF, Viquez-R LR, Arias-Alzate A, Belant JL & Ceballos G. 2016, Spatial patterns of species richness and functional diversity in Costa Rican terrestrial mammals: implications for conservation. *Diversity Distrib.* 22: 43-56.
- Goslee SC & Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* 22: 1-19.
- Gower JC. 1971. A general coefficient of similarity and some of its properties, *Biometrics* 27: 857-874.
- Graham CH & Fine PV. 2008. Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecol Lett* 11: 1265-1277.
- Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF & Kaufman DM et al. 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84: 3105-3117.
- Hijmans RJ, van Etten J, Sumner M, Cheng J, Baston D & Bevan A et al. 2019. Package: Raster. Geographic Data Analysis and Modeling. Disponível em: <<https://cran.r-project.org/web/packages/raster/index.html>>. Acesso em: maio/2020.
- IBGE – Instituto Brasileiro de Geografia e Estatística. 2002. Mapa climático do Brasil. Rio de Janeiro: IBGE.

- IBGE – Instituto Brasileiro de Geografia e Estatística. 2017. Disponível em: <<ftp://geoftp.ibge.gov.br>>. Acesso em: abril/2017.
- Inague GM, Zwiener VP & Marques MCM. 2021. Climate change threatens the woody plant taxonomic and functional diversities of the Restinga vegetation in Brazil. *Perspect Ecol Conserv* 19: 53-60.
- Jiang L, Lv G, Gong Y, Li Y, Wang H & Wu D. 2021. Characteristics and driving mechanisms of species beta diversity in desert plant communities. *Plos One* 16: 1-17.
- Jin Y & Qian H. 2019. V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* 42: 1353-1359.
- Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P & Bönsch G et al. 2020. TRY plant trait database - enhanced coverage and open access. *Glob Change Biol.* 26: 119-188.
- Kembel SW, Ackerly DD, Blomberg SP, Cornwell WK, Cowan PD & Helmus MR et al. 2010. Package: Picante. Integrating Phylogenies and Ecology. Disponível em: <<https://cran.r-project.org/web/packages/picante/index.html>>. Acesso em: maio/2019.
- Kreft H & Jetz W. 2007. Global patterns and determinants of vascular plant diversity. *PNAS* 104: 5925-5930.
- Lacerda LD, Araújo DSD & Maciel NC. 1993. Drycoastal ecosystems of the tropical Brazilian coast. In: Maarel E van der (Ed.), *Dry Coastal-ecosystems: Africa, Asia, Oceania*, Amsterdam: Elsevier, p. 477-493.
- Legendre P & Legendre L. 2012. *Numerical ecology*. Amsterdam: Elsevier, 990 p.
- Leroy B, Meynard CN, Bellard C & Courchamp F. 2016. virtualspecies, an R package to generate virtual species distributions. *Ecography* 39: 599-607.

- Lenz M, Schulz A, Koeck T, Rapp S, Nagler M & Sauer M et al. 2020. Missing value imputation in proximity extension assay-based targeted proteomics data. *Plos One* 15: 1-19.
- Li H, Zhao C, Shao F, Li GZ, & Wang X. 2015. A hybrid imputation approach for microarray missing value estimation. *BMC Genomics* 16: 1-11.
- Li Z, Jiang X, Wang J, Meng X, Heino J & Xie Z. 2019. Multiple facets of stream macroinvertebrate alpha diversity are driven by different ecological factors across an extensive altitudinal gradient. *Ecol Evol* 9: 1306-1322.
- Lichstein JW. 2007. Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol* 188: 117-131.
- Lobo JM, Hortal J, Yela JL, Millán A, Sánchez-Fernández D & García-Roselló E et al. 2018. KnowBR: An application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecol Indic* 91: 241-248.
- Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K & Studer M et al. 2021. "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al. Disponível em: <<https://cran.r-project.org/web/packages/cluster/index.html>>. Acesso em: março/2021.
- Magurran AE. 2004. Measuring biological diversity. Oxford: Blackwell Science, 256 p.
- Magurran AE & McGill BJ (Eds.). 2011. Biological diversity: frontiers in measurement and assessment. Oxford: Oxford University Press, 359 p.
- Maire E, Grenouillet G, Brosse S & Villéger S. 2015. How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Glob Ecol Biogeogr* 24: 728-740.

- Maitner BS, Boyle B, Casler N, Condit R, Donoghue J & Durán SM et al. 2018. The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol Evol* 9: 373-379.
- Massante JC & Gerhold P. 2020. Environment and evolutionary history depict phylogenetic alpha and beta diversity in the Atlantic coastal white-sand woodlands. *J Veg Sci* 31: 634-645.
- MMA - Ministério do Meio Ambiente. 2017. Disponível em: <<http://mapas.mma.gov.br/i3geo/datadownload.htm>>. Acesso em: abril/2017.
- Moeslund JE, Arge L, Bøcher PK, Dalgaard T & Svenning J-C. 2013. Topography as a driver of local terrestrial vascular plant diversity patterns. *Nord J Bot* 31: 129-144.
- Monserud RA & Leemans R. 1992. Comparing global vegetation maps with the Kappa statistic. *Ecol Model* 62: 275-293.
- Mori AS, Isbell F & Seidl R. 2018. β -diversity, community assembly, and ecosystem functioning. *Trends Ecol Evol* 33: 549-564.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P & McGlinn D et al. 2018. Vegan: community ecology package. R package version 2.5-1. Disponível em: <<https://CRAN.R-project.org/package=vegan>>. Acesso em: junho/2018.
- Oliveira EVS, Alves DMC, Landim MF & Gouveia SF. 2021. Sampling effort and the drivers of plant species richness in the Brazilian coastal regions. *Oecologia* 195: 163-171.
- Oliveira EVS, Landim MF & Gouveia SF. Submetido. Patterns and drivers of phylogenetic and functional structures of coastal woody vegetation of Eastern South America. *Journal of Plant Research*.
- Pantanowics A & Marwala T. 2009. Missing data imputation through the use of the random forest algorithm. *Adv Intell Syst Comput* 116: 53-62.

- Paradis E, Claude J & Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.
- Penone C, Davidson AD, Shoemaker KT, Di Marco M, Rondinini C & Brooks TM et al. 2014. Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods Ecol Evol* 5: 961-970.
- Phillips SJ, Anderson RP & Schapired RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecol Model* 190: 231-259.
- Pinto-Ledezma JN, Larkin DJ & Cavender-Bares J. 2018. Patterns of beta diversity of vascular plants and their correspondence with biome boundaries across North America. *Front Ecol Evol* 6: 1-13.
- Purvis A & Hector A. 2000. Getting the measure of biodiversity. *Nature* 405: 212-219.
- QGIS Development Team. 2020. QGIS Geographic Information System: Open Source Geospatial Foundation Project. Disponível em: <<http://qgis.osgeo.org>>. Acesso em: abril/2020.
- Qian H, Swenson NG & Zhang J. 2013. Phylogenetic beta diversity of angiosperms in North America. *Glob Ecol Biogeogr* 22: 1152-1161.
- Qian H, Jin Y, Leprieur F, Wang X & Deng T. 2020. Geographic patterns and environmental correlates of taxonomic and phylogenetic beta diversity for large-scale angiosperm assemblages in China. *Ecography* 43: 1706-1716.
- Qian H, Jin Y, Leprieur F, Wang X & Deng T. 2021. Patterns of phylogenetic beta diversity measured at deep evolutionary histories across geographical and ecological spaces for angiosperms in China. *J Biogeogr* 48: 773-784.
- Qian H & Zhang J. 2016. Are phylogenies derived from family-level supertrees robust for studies on macroecological patterns along environmental gradients? *J Syst Evol* 54: 29-36.

- R Development Core Team. 2018. R: A language and environment for statistical computing. Disponível em: <<http://www.R-project.org>>. Acesso em: setembro/2018.
- Rangel TFLVB, Diniz-Filho JAF & Bini LM. 2010. SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography* 33: 46-50.
- Revell LJ. 2012. Package: phytools. Phylogenetic Tools for Comparative Biology (and Other Things). Disponível em: <<https://cran.r-project.org/web/packages/phytools/index.html>>. Acesso em: maio/2020.
- Ricklefs RE. 2004. A comprehensive framework for global patterns in biodiversity. *Ecol Lett* 7: 1-15.
- Rizinni AT. 1997. Tratado de Fitogeografia do Brasil: aspectos ecológicos, sociológicos, e florísticos. Rio de Janeiro: Âmbito Cultural, 746 p.
- Scarano FR. 2002. Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian Atlantic Rainforest. *Ann Bot* 90: 517-524.
- Schmitt S, Pouteau R, Justeau D, de Boissieu F & Birnbaum P. 2017. ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods Ecol Evol* 8: 1795-1803.
- Si X, Baselga A, Leprieur F, Song X & Ding P. 2016. Selective extinction drives taxonomic and functional alpha and beta diversities in island bird assemblages. *J Anim Ecol* 85: 409-418.
- Siefert A, Ravenscroft C, Weiser MD & Swenson NG. 2013. Functional beta diversity and tree community assembly. *Glob Ecol Biogeogr* 22: 682-691.
- Silva KJP & Souza AF. 2018a. Common species distribution and environmental determinants in South American coastal plains. *Ecosphere* 9:1-15.

- Silva AC & Souza AF. 2018b. Aridity drives plant biogeographical sub regions in the Caatinga, the largest tropical dry forest and woodland block in South America. *Plos One* 13: 1-22.
- Silva J LA, Souza AF, Jardim JG & Goto BT. 2015. Community assembly in harsh environments: the prevalence of ecological drift in the heath vegetation of South America. *Ecosphere* 6: 1-18.
- Smith SA & Brown JH. 2018. Constructing a broadly inclusive seed plant phylogeny. *Am J Bot* 105: 302-314.
- Stekhoven DJ & Bühlmann P. 2012. Missforest — non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28: 112-118.
- Svenning J-C, Fløjgaard C & Baselga A. 2011. Climate, history and neutrality as drivers of mammal beta diversity in Europe: insights from multiscale deconstruction. *J Anim Ecol* 80: 393-402.
- Swenson NG. 2011. Phylogenetic Beta Diversity Metrics, Trait Evolution and Inferring the Functional Beta Diversity of Communities. *Plos One* 6: 1-9.
- Swenson NG. 2014. Functional and Phylogenetic Ecology in R. New York: Springer, 217 p.
- Trabucco A & Zomer RJ. 2019. Global High-Resolution Soil-Water Balance: CGIAR Consortium for Spatial Information. Disponível em: <<http://www.cgiar-csi.org>>. Acesso em: outubro/2019.
- Vellend M. 2010. Conceptual synthesis in community ecology. *Q Rev Biol* 85:183-206.
- Villéger S, Grenouillet G & Brosse S. 2013. Decomposing functional b-diversity reveals that low functional b-diversity is driven by low functional turnover in European fish assemblages. *Glob Ecol Biogeogr* 22: 671-681.
- Weng ES & Zhou GS. 2006. Modeling distribution changes of vegetation in China under future climate change. *Environ Model Assess* 11: 45–58.

- Whittaker RH. 1972. Evolution and measurement of species diversity. *Taxon* 21: 213-251.
- Whittaker RJ & Fernández-Palacios. 2007. *Island Biogeography Ecology, evolution, and conservation*. New York: Oxford University Press, 416 p.
- Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A & NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. *Divers Distrib* 14: 763-773.
- Wong JSY, Chan YKS, Ng CSL, Tun KPP, Darling ES & Huang D. 2018. Comparing patterns of taxonomic, functional and phylogenetic diversity in reef coral communities. *Coral Reefs* 37: 737-750.
- Yang J, Swenson N, Zhang G, Ci X, Cao M & Sha L et al. 2015. Local-scale Partitioning of Functional and Phylogenetic Beta Diversity in a Tropical Tree Assemblage. *Sci Rep* 5: 1-10.
- Zhang J-L, Swenson NG, Chen S-B, Liu X-J, Li Z-S, Huang J-H, Mi X-C & Ma K-P. 2013. Phylogenetic beta diversity in tropical forests: Implications for the roles of geographical and environmental distance. *J Syst Evol* 51: 71-85.
- Zwiener VP, Lira-Noriega A, Grady CJ, Padial AA & Vitule JRS. 2018. Climate change as a driver of biotic homogenization of woody plants in the Atlantic Forest. *Glob Ecol Biogeogr* 27: 298-309.

Apêndice

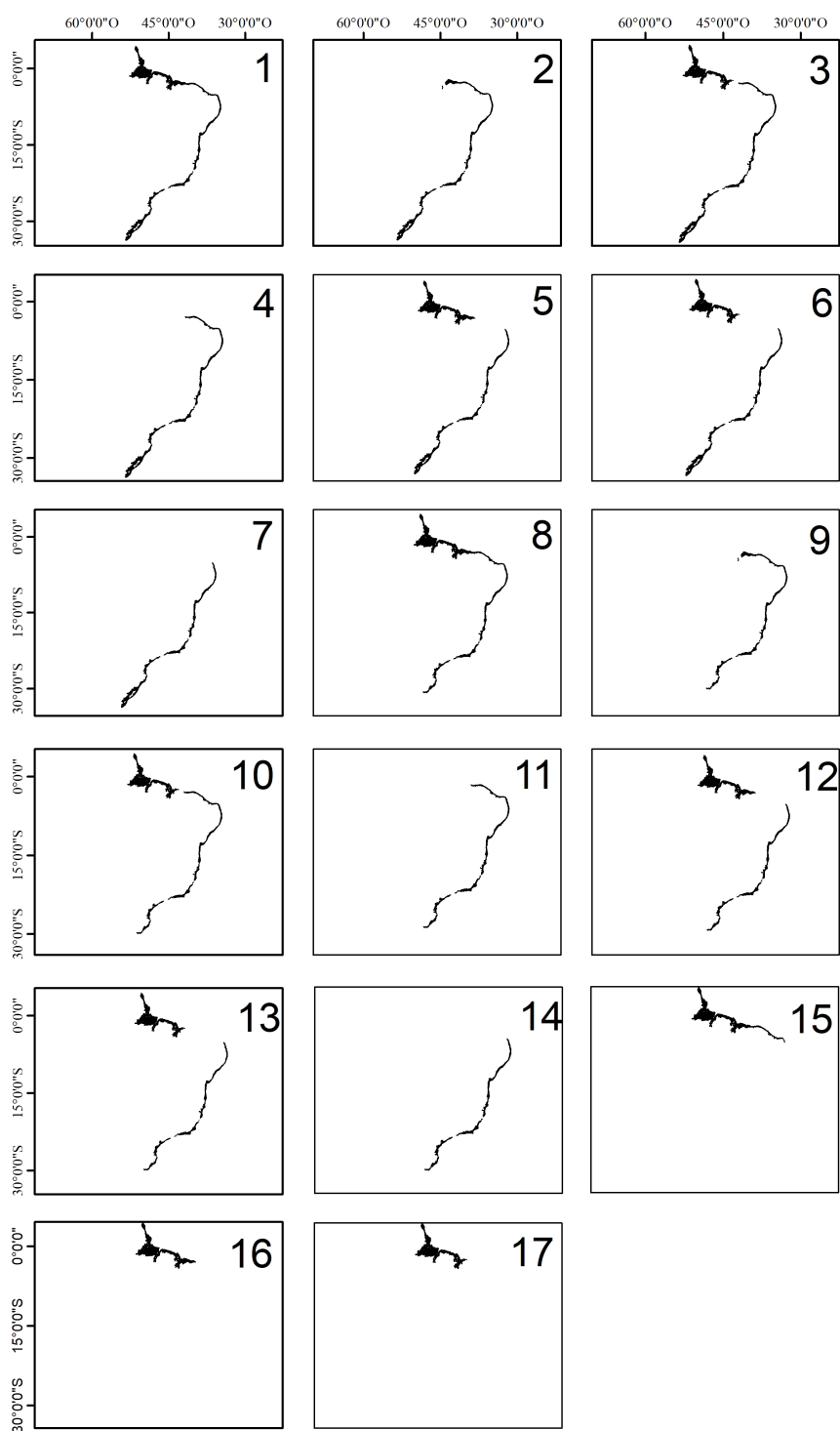


Figure S1. Diferentes contornos de sensibilidade (área de *buffer* das espécies) utilizados para a construção dos modelos de distribuição das espécies de árvores na região costeira do Brasil.

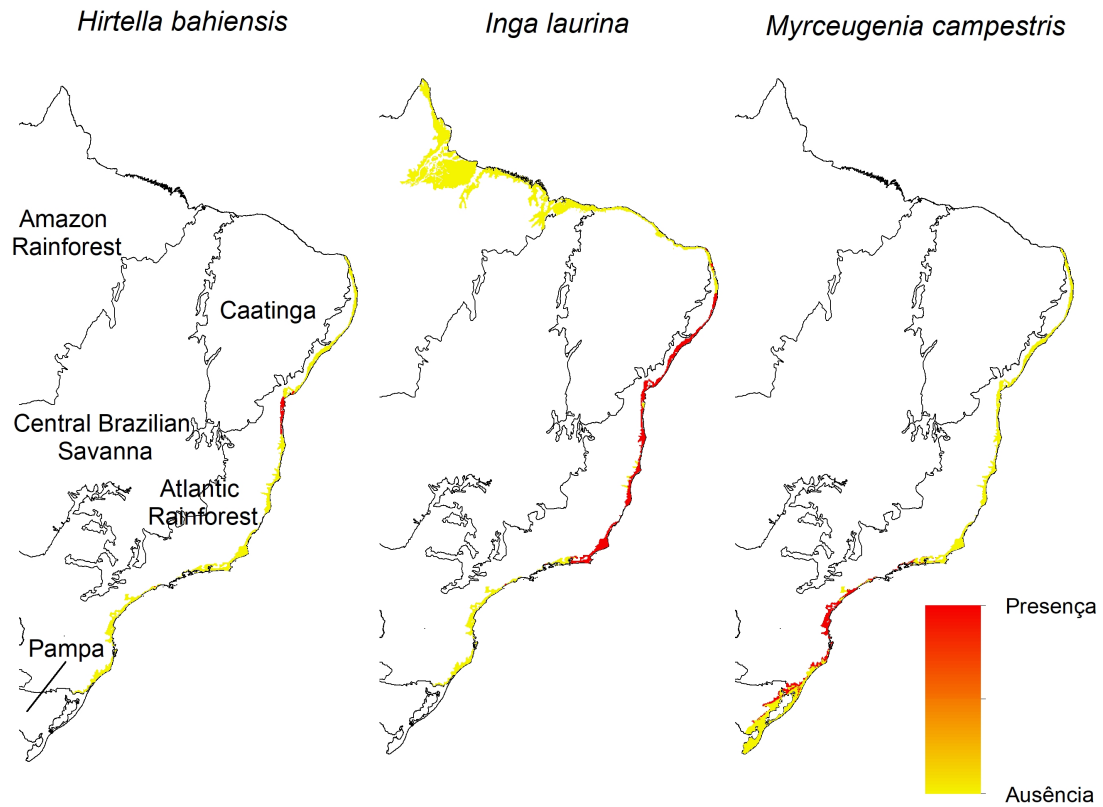


Figure S2. Mapas binários (*threshold* = especificidade + sensibilidade) gerados a partir de modelos de distribuição de espécies (SDMs) de três das 360 espécies de árvores utilizadas nas análises de diversidade beta. Estes mapas foram utilizados para estimar a presença ou ausência das espécies em cada célula da grade ao longo da região costeira do Brasil. São apresentadas espécies (da esquerda para a direita) com uma acurária excelente ($AUC > 0,90$ e $Kappa > 0,75$), boa ($AUC = 0,80-0,89$ $Kappa 0,56-0,74$) e aceitável ($AUC = 0,75-0,79$ e $Kappa 0,41-0,55$), respectivamente.

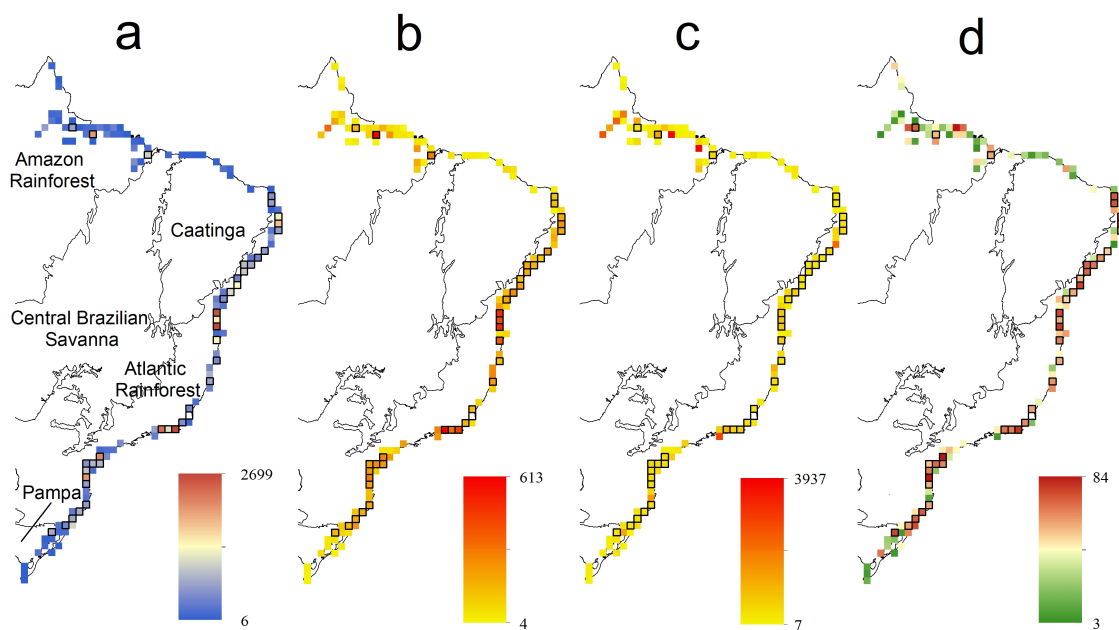


Figure S3. Mapas de distribuição de número de registros (a), riqueza observada (b), riqueza esperada (c) e completude de inventário (d) para espécies de árvores na região costeira do Brasil. Células com contorno na borda foram consideradas como bem amostradas (completude $\geq 50\%$ e número de registros ≥ 307).

Tabela S1. Grupo de variáveis ambientais não correlacionadas ($r < 0,70$) com respeito a cada um dos 17 contornos de sensibilidade adotados para a construção de modelos de distribuição de espécies (SDMs) de árvores para a região costeira do Brasil (detalhes sobre os contornos são mostrados na Fig. S1). Variáveis marcadas foram utilizadas para a construção dos SDMs utilizando cada respectivo contorno de sensibilidade. TEMP: temperatura média anual; TEMP-VAR: variação diurna média da temperatura; TEMP-SAZ: sazonalidade da temperatura; TEMP-MAX: temperatura máxima do mês mais quente; TEMP-MIN: temperatura mínima do mês mais frio; PRE: precipitação anual; PRE-WET: precipitação do mês mais chuvoso; PRE-DRY: precipitação do mês mais seco; PRE-SAZ: sazonalidade da precipitação; ALT: altitude.

Variável	Contornos de sensibilidade																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
TEMP						x	x		x				x			x	x
TEMP-VAR	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x
TEMP-SAZ	x							x						x	x	x	x
TEMP-MAX		x		x					x					x	x	x	x
TEMP-MIN			x		x					x	x	x				x	
PRE		x		x	x	x	x	x	x	x	x		x	x	x		x
PRE-WET	x	x	x	x					x		x	x				x	
PRE-DRY		x	x	x	x	x	x							x	x		x
PRE-SAZ	x							x		x	x	x	x			x	
ALT	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Tabela S2. Valores de particionamento da diversidade beta taxonômica e filogenética a partir de uma composição observada e obtida por modelos de distribuição de espécies (SDMs) para assembleias de árvores reconhecidas como bem amostradas ao longo da Restinga. JTU: valor de substituição medido pelo índice de Jaccard, JNE: valor do aninhamento medido pelo índice de Jaccard, JAC: valor da diversidade beta geral medida pelo índice de Jaccard, SIM: valor de substituição medido pelo índice de Sørensen, SNE: valor de aninhamento medido pelo índice de Sørensen, SOR: valor da diversidade beta geral medida pelo índice de Sørensen.

Composição	<u>β-Taxonômica</u>		<u>β-Filogenética</u>			
	JTU	JNE	JAC	SIM	SNE	SOR
Observada	0,957	0,015	0,972	0,856	0,050	0,907
	98%	2%	100%	94%	6%	100%
SDMs	0,926	0,028	0,954	0,780	0,093	0,874
	97%	3%	100%	89%	11%	100%

Tabela S3. Comparações entre modelos lineares múltiplos e seus componentes para os padrões da diversidade beta taxonômica referente a uma composição observada e obtida por modelos de distribuição de espécies (SDMs) para assembleias de árvores reconhecidas como bem amostradas ao longo da Restinga. Valores de *slope* e R^2 são padronizados. * = $p < 0,05$.

β -taxonômica		Composição			
		Observada		SDMs	
Modelo	Preditor	<i>Slope</i> (β)	R^2	<i>Slope</i> (β)	R^2
Água	PET	0,400*	0,280*	0,430*	0,340*
	AET	0,230*		0,140*	
	PRE-WET	0,110*		0,022*	
Energia	TEMP-MAX	0,404*	0,165*	0,509*	0,260*
	TEMP-VAR	0,090		0,041	
Estabilidade climática	TEMP-HIST	0,390*	0,150*	0,280*	0,080*
	PRE-HIST	0,160*		0,099	
Edáfico	CAP	0,194*	0,080*	0,230*	0,140*
	SAL	0,153*		0,220*	
	NUT	0,057		-0,075	
	PORO	0,051		0,160	
	C	-0,038		-0,075	
Heterogeneidade de habitat	SLOPE	0,160*	0,030*	0,016*	0,030*
	ALT	0,070		0,070	
Influência Humana	INFRA	0,020	0,005	-0,050	0,003

Tabela S4. Comparações entre modelos lineares múltiplos e seus componentes para os padrões da diversidade beta filogenética referente a uma composição observada e obtida por modelos de distribuição de espécies (SDMs) para assembleias de árvores reconhecidas como bem amostradas ao longo da Restinga. Valores de *slope* e R^2 são padronizados. * = $p < 0,05$.

β -filogenética		Composição			
		Observada		SDMs	
Modelo	Preditor	<i>Slope</i> (β)	R^2	<i>Slope</i> (β)	R^2
Água	PET	0,440*	0,340*	0,460*	0,470*
	AET	0,310*		0,130*	
	PRE-WET	0,070		0,350*	
Energia	TEMP-MAX	0,420*	0,180*	0,100*	0,340*
	TEMP-VAR	-0,090		-0,003	
Estabilidade climática	TEMP-HIST	0,380*	0,160*	0,190*	0,040*
	PRE-HIST	0,170*		0,130	
Edáfico	CAP	0,170*	0,100*	0,250*	0,140*
	SAL	0,200*		0,230	
	NUT	0,070		-0,090	
	PORO	0,070		0,150	
	C	-0,060		-0,047	
Heterogeneidade de habitat	SLOPE	0,150*	0,020*	0,200*	0,040*
	ALT	0,050		0,040	
Influência Humana	INFRA	-0,050	0,002	-0,070	0,005

Conclusões gerais

Conclusões gerais

Nós apresentamos a primeira análise de diferentes escalas de diversidade (α , β) em suas múltiplas perspectivas (i.e., taxonômica, filogenética e funcional) para espécies de plantas da Restinga brasileira. Nós encontramos um baixo nível de conhecimento botânico, apesar da presença da maioria dos centros de pesquisa nesta região. Este fato reforça a necessidade de aumentar os esforços de amostragem, notavelmente em regiões menos amostradas. Ainda assim, utilizando técnicas que nos permitiram contornar essas deficiências de inventário, nós fomos capazes de identificar os determinantes dos padrões espaciais da diversidade. Nós encontramos assembleias adjacentes a porção sul do domínio de Mata Atlântica (i.e., sul da Bahia e região Sudeste) com a maior riqueza de espécies e diversidade de linhagens, enquanto não identificamos um padrão claro para a estrutura funcional. Nós evidenciamos um efeito da heterogeneidade de habitat, além de propriedades de solo e restrições de água na variação da riqueza de espécies. Além disso, nossos resultados dão suporte ao efeito das restrições de água como determinante primário da estrutura funcional e filogenética, favorecendo linhagens pré-adaptadas e proximamente relacionadas nas porções mais áridas da costa. Por outro lado, a colonização por espécies de formações úmidas adjacentes, contribui para mais espécies e linhagens, conduzindo a arranjos estocásticos ou mesmo dispersão entre as espécies quanto as suas características fenotípicas e relação filogenética. As restrições de água também foram responsáveis pela variação geográfica na composição de plantas nas três perspectivas avaliadas. Nós observamos uma substituição taxonômica não aleatória e uma dissimilaridade filogenética e funcional associadas à distância geográfica. Em conjunto, nossos resultados apontam para um cenário em que as mudanças na composição e diversidade de plantas estão relacionadas a uma interação entre os distintos pools regionais de espécies, com seus diferentes padrões de diversidade, que por sua vez contribuem de forma diferente com espécies dispersoras, e

pelas condições locais encontradas ao longo da zona costeira. Nossos estudos permitem agregar informações importantes para uma compreensão abrangente sobre os fatores determinantes da diversidade atual de plantas da Restinga. Contudo, enfatizamos a necessidade de se contornar as múltiplas e significativas lacunas de informação sobre a biodiversidade das zonas costeiras tropicais. Por fim, ressaltamos a importância de investigar sistemas específicos, com seu contexto ambiental particular, seja como um meio de se entender a dinâmica e as causas dos padrões de comunidades ecológicas específicas, seja como meio para avançar na compreensão geral dos fenômenos ecológicos em nível de comunidades, através da dependência de contexto.