



RESEARCH ARTICLE

Cross-scale drivers of woody plant species commonness and rarity in the Brazilian drylands

Bruno X. Pinho^{1,2}  | Diego P. F. Trindade³ | Carlos A. Peres^{4,5} | Davi Jamelli¹ | Renato A. F. de Lima⁶  | Elâine M. S. Ribeiro⁷ | Felipe P. L. Melo¹ | Inara R. Leal¹ | Marcelo Tabarelli¹

¹Departamento de Botânica, Universidade Federal de Pernambuco, Recife, Brazil

²AMAP, Univ Montpellier, INRAe, CIRAD, CNRS, IRD, Montpellier, France

³Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia

⁴School of Environmental Sciences, University of East Anglia, Norwich, UK

⁵Instituto Juruá, Manaus, Brazil

⁶Departamento de Ecologia, Universidade de São Paulo, Brazil

⁷Laboratório de Biodiversidade e Genética Evolutiva, Universidade de Pernambuco – Campus Petrolina, Petrolina, Brazil

Correspondence

Bruno X. Pinho, AMAP, Univ Montpellier, Montpellier, France.
Email: bxpinho@hotmail.com

Funding information

Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 403770/2012-2; Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, Grant/Award Number: BFP-0164-2.05/19; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2013/08722-5; Newton Fund

Editor: Francesco Maria Sabatini

Abstract

Aim: Locally abundant species are typically widespread, while locally scarce species are geographically restricted—the so-called abundance-occupancy relationships (AORs). AORs help explain the drivers of species rarity and community assembly, but little is known about how variation around such relationship is driven by species traits and niche-based processes, particularly in tropical woody plants. We tested the hypothesis that AORs in tropical dryland woody plants are positive and mediated by niche and functional traits along environmental gradients.

Location: The Caatinga dry forest and Cerrado savannah, Brazil.

Methods: We aggregated abundance and occurrence data into grid-cells representing local (10-km) to landscape scales (50-km). We calculated species mean relative abundance at occupied grid-cells (local abundance) and the proportion of grid-cells occupied (occupancy), and estimated their niche breadth and marginality along multivariate environmental gradients.

Results: AORs were positive but weak at different scales in both regions due to some locally abundant but geographically restricted species, with most species being both locally and geographically rare. Cross-species variation in local abundance was largely unpredictable, but occupancy was strongly driven by niche and functional traits, with a prominent negative effect of niche marginality. Geographically restricted species were associated with rare habitats, such as wetter and less intensively used habitats. Large seeds and abiotic dispersal favoured occupancy in Caatinga at small and large spatial scales, respectively, whereas species with conservative leaves were more widespread across scales in Cerrado.

Main conclusions: Woody plants in dry tropical biotas exhibit weak AORs, a pattern likely related to low habitat availability and dispersal limitation. Caatinga and Cerrado biotas emerge as environmentally structured at multiple spatial scales, with several habitat-specialist rare species bearing specific regenerative and resource-use traits and relying on conditions threatened by climate change and land-use intensification. Examining AORs through the lens of niche, functional traits and spatial scales enables

mapping patterns and drivers of species commonness and rarity, enhancing understanding of species assembly and providing tools for biodiversity conservation.

KEYWORDS

abundance-occupancy relationships, Caatinga, Cerrado, functional traits, niche breadth, niche marginality, savannahs, scale-dependency, seasonally dry tropical forests

1 | INTRODUCTION

Understanding patterns and drivers of species commonness and rarity has long been a central topic in macroecology, with important implications for landscape management and biodiversity conservation (Enquist et al., 2019; Rabinowitz, 1981). Across several biotas, locally abundant species tend to be widespread geographically, whereas rare species tend to be both locally scarce and geographically restricted—the so-called abundance-occupancy relationships (hereafter AORs; Gaston et al., 2000). AORs provide a mechanistic basis for understanding the drivers of species commonness and rarity across spatial scales (Borregaard & Rahbek, 2010). In applied terms, local scarcity and geographic limitations may impose a “double jeopardy” for rare species by rendering them more extinction-prone due to demographic stochasticity and habitat loss (Gaston, 1998). This is particularly alarming in tropical ecosystems that support a myriad of rare plant species (ter Steege et al., 2013) exposed to land-use intensification, climate change and overexploitation of forest products (Laurance et al., 2014; Pinho et al., 2020). Therefore, understanding cross-scale determinants of species local abundance and distribution is urgent to prevent biodiversity loss.

Positive AORs have been detected for many taxa across terrestrial and aquatic biotas and represent one of the most pervasive patterns in ecology (e.g. Buckley & Freckleton, 2010; Caten et al., 2022; Craven et al., 2021; Heino & Tolonen, 2018; Webb et al., 2017). However, the shape of AOR curves varies strongly (Gaston, 1996), reflecting the many types of rarity already documented (Rabinowitz, 1981). For example, some plant species can be widespread but locally scarce such as most tropical rainforest trees (ter Steege et al., 2013), while others can be geographically restricted but locally abundant such as many tropical dry forest trees (Hubbell, 1979; Williams et al., 2010), yielding either neutral, negative or unimodal AORs (Borregaard & Rahbek, 2010; Sporbert et al., 2020). Exceptions to the general rule are expected to be particularly evident at regional scales, as dispersal limitation can preclude colonization of available sites, thereby limiting positive relationships between increased abundance and occupancy (Freckleton et al., 2005). Accordingly, the few studies reporting positive AORs in plants have been mainly deployed at small spatial scales (but see Craven et al., 2021, for a regional-scale example). Most importantly, the causal mechanisms of AOR patterns across scales remain elusive (Borregaard & Rahbek, 2010), especially for tropical forest tree species (but see Díaz et al., 2020, for a niche-based explanation).

A positive relationship between local abundance and the extent of species distribution can emerge in response to a combination of ecological, biological and artefactual mechanisms (Borregaard & Rahbek, 2010). Ecological mechanisms are mainly related to two non-mutually exclusive niche-based processes related to how species exploit available habitat. The “niche breadth hypothesis” suggests that species able to tolerate a wide range of abiotic conditions should be able to persist under contrasting environments and thus achieve high regional occupancy, but also better exploit local resources, thereby also maintaining large local populations (Brown, 1984). In fact, niche breadth consistently predicts geographic range size across biotas (Slatyer et al., 2013), but positive effects on local abundance are largely unclear (Sporbert et al., 2020) or driven by statistical artefacts (Díaz et al., 2020; Marino et al., 2020). On the other hand, according to the “niche marginality hypothesis,” the main driver of variation in species abundance and occupancy is not the degree of niche specialization, but how common or rare are the habitats they use. In other words, the niche marginality hypothesis states that species able to exploit the most common habitats tend to be more widespread and locally abundant than those adapted to marginal or rare habitat conditions (Hanski et al., 1993). Recent studies in aquatic communities (Heino & Tolonen, 2018; Marino et al., 2020) and tropical rainforests (Díaz et al., 2020) have strongly supported the niche marginality hypothesis, demonstrating that geographically restricted species are mostly associated with rare habitat conditions and thus geographically limited by habitat availability. Nonetheless, both positive AORs and the effects of niche marginality and breadth on species occupancy may arise from sampling artefacts, as occupancy of locally scarce species can be consistently underestimated, and other mechanisms (e.g. dispersal) may drive species-environment associations (Díaz et al., 2020; Hanski et al., 1993).

Beyond the above-mentioned mechanisms, plant functional traits may also be key drivers of AORs along environmental gradients, as they are related to species colonization ability and tolerance to environmental conditions (Grime & Pierce, 2012). However, this hypothesis is yet to be tested in tropical forest plants (e.g. Sporbert et al., 2021 for European vascular plants), and few studies have examined the combined effects of niche and traits in driving AORs in other biotas (e.g. Heino & Tolonen, 2018; Marino et al., 2020). Functional traits have been examined to understand plant species differences related to individual growth, survival and reproduction, but also differences in population performance and colonization ability along environmental gradients (Reich, 2014; Westoby et al., 2002). For instance, leaf and woody traits have been proposed

to reflect a continuum from species that can grow fast under high-resource availability by producing soft tissues (i.e. acquisitive strategies) to more conservative species that ensure survival under abiotic stress by slowly growing durable tissues (Reich, 2014). Also, seed traits may describe regenerative strategies, as large seeds may promote tolerance to stressful conditions and higher initial seedling performance under low resource availability, whereas small-seeded species tend to benefit from their high fecundity, enhancing seed shadows and potentially winning new sites by default (Muller-Landau, 2010; Westoby et al., 2002). By affecting basic population processes related to colonization, growth and mortality, functional traits likely mediate species abundance-occupancy patterns, particularly in biotas exposed to harsh environmental filters such as seasonally dry tropical forests and savannahs (Pennington et al., 2009).

Drylands occur as disjunct patches or “nuclei” across the Neotropics and they are particularly distinguished by their low and very seasonal precipitation regimes (Pennington et al., 2009). Most plant species are spatially restricted to a given dryland nucleus, leading to high floristic dissimilarity among nuclei (DRYFLOR, 2016). The Brazilian drylands comprise the Caatinga and Cerrado biogeographic regions, the former representing the largest tract of seasonally dry tropical forests in South America (Silva et al., 2017), while the latter represents the largest Neotropical savannah. Plant adaptations in Caatinga are particularly related to dealing with intense water stress, while many Cerrado plant lineages are associated with recent diversification of sister species from nearby wet forests via adaptation to dystrophic soils or fire (e.g. Simon et al., 2009). Like other drylands, Caatinga dry forests and Cerrado savannahs exhibit high plant species turnover mainly driven by soil properties and aridity levels (Bueno et al., 2018; Silva & Souza, 2018). As the most densely settled Neotropical semiarid region, the Brazilian drylands have experienced a trajectory of intense deforestation, forest degradation and increased aridity, leading to biodiversity loss, desertification and vulnerability of human populations, which rely on forest products and other ecosystem services to enhance their livelihood conditions (Silva et al., 2017; Strassburg et al., 2017). In this context, assessing how AORs are shaped by both functional traits and environmental conditions, including land-use and climate, should help elucidate cross-scale patterns of species rarity/commonness, and the prospects of biodiversity and associated ecosystem services under global change.

Here, we use a cross-scale approach (i.e. at different grain sizes) to answer the following questions: (1) what is the nature of AORs in woody plant species in the Caatinga and Cerrado biogeographic regions? (2) How do species niche attributes (marginality and breadth) and functional traits drive cross-species variation in local abundance and occupancy, and thus their deviation from AORs? In addition, as the detection of AORs and the effects of niche marginality and breadth can be affected by sampling artefacts (Díaz et al., 2020; Hanski et al., 1993), we used null models to examine the degree to which our findings can be given by chance. First, we expected strong positive AORs across scales in each region, since increases in local abundance should disproportionally affect colonization of new sites.

Nevertheless, we expected the effect of local abundance on occupancy (i.e. positive deviations in AORs) to be mediated by specific functional traits, being particularly strong for opportunistic strategies (i.e. acquisitive resource-use, small and abiotically-dispersed seeds), due to potentially greater colonization and growth ability that favour ubiquitous occupancy across sites. Finally, we expected locally scarce and narrowly distributed (i.e. rare) species to occupy marginal niches and exhibit limited niche breadth, being associated to specific rare habitat conditions, such as sandy soils in Caatinga, relatively fertile soils in Cerrado, and more mesic and less disturbed sites in both regions. Overall, we expected similar patterns at both study regions, although water-related drivers should be particularly important in the Caatinga dry forest, while soil fertility and topographic factors are expected to be prominent drivers of species abundance and distribution across the Cerrado savannahs. Patterns documented here are discussed in light of the scale- and context-dependency of AOR drivers, plant community macro-organization, and the vulnerability of Brazilian drylands to global changes.

2 | METHODS

2.1 | Study regions

The Brazilian drylands comprise the Caatinga and Cerrado regions (Figure 1). The Caatinga dry forest region covers 912,529 km² or 11% of Brazil's territory (Silva et al., 2017), spreading over immense plateaus (mostly around 400 masl) with few montane ranges exceeding 1000 masl (Silva & Souza, 2018). Soils are mostly of Pre-Cambrian origin, nutrient-rich, stony and shallow, while sandy soils of sedimentary origin are deep and nutrient-poor, but they cover a relatively small fraction (~30%) of the Caatinga region (Moro et al., 2016). The Caatinga dry forest encompasses a wide variation in precipitation regimes, mostly from 400 to 1200 mm/year, with up to 70% of the annual precipitation concentrated into a single month (see Andrade et al., 2017, for further detail on precipitation regimes). Average annual temperature ranges from 19°C in the southern Caatinga to 28°C in the northern and eastern portions (Silva & Souza, 2018). The vegetation ranges from shrub-dominated to small-statured forest stands, where shrubs and small Fabaceae and Euphorbiaceae trees dominate (Fernandes et al., 2020; Queiroz et al., 2017). Floristically, the Caatinga is one of the most diverse dry forests anywhere, with many unique and old plant-lineages, including endemic species and genera (Queiroz et al., 2017). Recent evidence has indicated a pivotal role of varying aridity levels in driving plant community assembly and floristic affinities across the Caatinga dry forest (Silva & Souza, 2018).

The Cerrado region covers ~2 million km² or 23% of Brazil's territory plus small areas in Paraguay and Bolivia (Figure 1; Pennington et al., 2006). The region encompasses a highly heterogeneous environmental and vegetation macro-mosaic, from open-canopy savannahs (the predominant vegetation type) to closed-canopy seasonally dry forests (Bueno et al., 2018). The relief is mostly comprised of

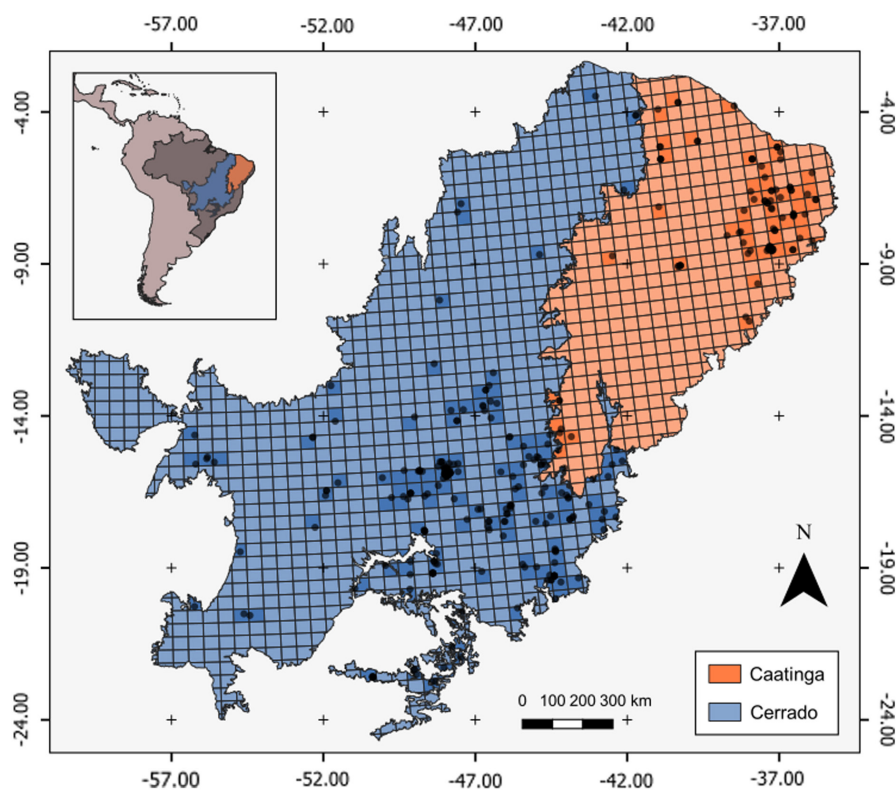


FIGURE 1 Location of study plots (solid dots) across 50-km spatial-grid-cells throughout the largest south American dryland regions, the Caatinga dry forest ($N = 121$ study plots) and Cerrado savannahs ($N = 209$ plots)

extensive plateaus and valleys, extending from lowlands (~100 m) to highlands of up to 1500 m. The climate is also seasonal, but generally wetter compared to the Caatinga, with higher annual precipitation (800–2000 mm) and less prolonged droughts (Bueno et al., 2018; see Figure S1). Average annual temperatures range from 18 to 28°C following a south–north gradient. Savannahs account for ~70% the region and occupy poor dystrophic soils under frequent dry-season fires, while deciduous and semi-deciduous forests occur as small patches in relatively fertile, less acidic soils (Bueno et al., 2018). The Cerrado holds one of the world's most diverse savannah floras, with many endemic vascular plants and ~1000 woody plant species, 10% of which is highly widespread across the region (Bridgewater et al., 2004).

2.2 | Abundance and occurrence data: From plots to spatial grid-cells

We first compiled a dataset of woody plant (i.e. sub-shrubs, shrubs and trees) abundance within 330 plots, 121 across Caatinga dry forests and 209 across Cerrado savannahs (solid dots in Figure 1). Plots differed in size and stem diameter cut-offs (see Table S1, for further sampling details in each region), covering 122,326 plant individuals in Caatinga and 385,884 in Cerrado. Most of the plot data was extracted from the Neotropical Tree Communities database (TreeCo; Lima et al., 2020) plus our own data as described elsewhere (Ribeiro et al., 2015; Rito et al., 2017).

To assess macroecological patterns at different spatial scales, while also reducing pseudo-replication issues caused by nearby

plots, we aggregated abundance data into spatial grid-cells of different sizes (10, 20, 30, 40 and 50 km) over the Caatinga and Cerrado regions (see example of 50-km grid scale in Figure 1). We were therefore able to examine cross-scale changes in species abundance/occupancy and their potential drivers, i.e. scale-dependent patterns. For clarity, however, only findings at the smallest and largest scales are reported in the main document, while further details on cross-scale patterns are provided in the Appendix S1. To avoid species with poorly known distributions/niches, only species occurring in at least three grid-cells were considered in further analyses. We also excluded palms, as they are frequently planted by locals in our focal regions. Finally, we filtered species for which functional trait data were available, which resulted in 135 and 417 woody plant species in the Caatinga and Cerrado regions, respectively, with 66 species occurring in both regions. The selected species accounted for ~85% of the total woody plant abundance across all plots, corresponding to 423,725 individual records. However, all individuals recorded across plots (i.e. including palms and poorly sampled species) were considered in estimates of species mean relative local abundances (see below). Although those few species lacking trait data were mostly rare, the selected species covered the entire spectrum of variation in both local abundance and occupancy across our two focal biotas (Figure S2).

To accurately estimate species occupancy patterns, we additionally compiled georeferenced occurrence data for each selected plant species from the plot dataset (see above) using the *Botanical Information and Ecology Network*—“BIEN” R package (Maitner et al., 2018). This botanical data acquisition includes several sources, such as GBIF and SpeciesLink, which comprise records

from hundreds of herbaria (Maitner et al., 2018). Valid records were checked using the “CoordinateCleaner” R package (Zizka et al., 2019). Records at the intraspecific level (e.g. subspecies) were merged at the species level. Synonyms and minor misspellings on specific epithets were double-checked using the “flora” R package (Carvalho, 2016), which follows the Brazilian Flora 2020 nomenclature (Brazil Flora Group, 2019).

We acknowledge that differences in sampling effort among plots may affect our estimations of species local abundance. However, (1) our focus is on *mean relative* abundance per species within plots where they occur, which is less affected; (2) abundance data were aggregated into spatial grid-cells, thereby reducing sample size discrepancies; and (3) we conducted a simple sensitivity analysis by removing spatial grid-cells within the lowest or highest 10% percentile in terms of incidence of plant records, which resulted in similar patterns as described below.

2.3 | Species local abundance and regional occupancy

After the data compilation, cleaning, spatializing and filtering steps, we calculated the local abundance and regional occupancy of each selected species, at each grid-cell scale in each biogeographic region. Local abundance was calculated as the mean relative abundance of a species across grid-cells at which it occurs (hereafter, local abundance), and thus measure how dominant (or subordinate) a given species tends to be where it is present. Occupancy is defined as the proportion of grid-cells occupied by any given species by collapsing plot-based data and individual occurrences. These are considered the best measures to examine the links between local- and regional-scale population processes (see Webb et al., 2012 for a debate on the use of different measures for testing AORs). In addition to occupancy across the grid-cells where the sample plots were distributed (i.e. the focus of this study, see Figure 1), we measured species occupancy over the entire study regions based on occurrence data (i.e. specimen collections unrelated to plot and occurrence data as described earlier) to examine whether species rarity patterns over entire regions are consistent with that observed across our focal grid-cells.

2.4 | Environmental and disturbance gradients

We overlaid the spatial grid-cells on global maps of key climate, soil, topographic and land-use variables that have been recognized as drivers of plant assembly across drylands (Aguirre-Gutiérrez et al., 2020; Pinho et al., 2019; Ribeiro et al., 2015, 2019; Sfai et al., 2018; Silva & Souza, 2018). Environmental variables were then averaged within grid-cells to measure species' niche attributes (see below). For peripheral grid-cells straddling the boundaries between biogeographic regions, averaged environmental variables only considered the area representing each region.

We initially considered a set of 10 environmental variables: four climatic variables related to the average and seasonality of temperature and precipitation (mean annual precipitation—MAP, mean annual temperature—MAT, temperature seasonality—TS, cumulative water deficit—CWD); two edaphic variables related to soil structure (clay content—Clay) and fertility (cation exchange capacity—CEC); two topographic variables (elevation and slope); and two land-use variables (tree cover and landscape integrity). The first climate variables (MAP, MAT and TS) were extracted from the high-resolution (30 arc-sec or ~1 km at the equator) WorldClim geo-database, version 2.0 (Fick & Hijmans, 2017), which includes average monthly data from 1970 to 2000. CWD represents a measure of the annual deficit in water availability (lower values indicate drier climates with relatively prolonged droughts) obtained from Chave et al. (2014). Soil data were obtained from the World Soil Information Database (Hengl et al., 2017), a collection of world soil maps based on machine learning at 250-m resolution. Information on land-use patterns were obtained from Global Forest Watch (2014). “Tree cover” refers to data obtained from Landsat images in 2000 at 30-m resolution, representing the per cent of natural vegetation in a given grid (see Hansen et al., 2013 for further details on remote sensing procedures). “Landscape integrity” represents an index score based on the combination of four datasets related to forest extent and both acute and chronic disturbances, based on direct and indirect degradation gradients (Grantham et al., 2020). Correlations among environmental variables were generally low, except between elevation and MAT, and landscape integrity and tree cover, which were highly correlated ($r > .9$, Figure S3). We therefore removed elevation and landscape integrity from further analyses, as these were more correlated with other environmental variables. Importantly, however, we note that variation in MAT is associated with elevational gradients, while changes in tree cover should reflect variation in the intensity of land-use regimes. The distribution of environmental variables across the study regions is shown in Figure S1.

2.5 | Species niche attributes

To describe species niches, two key measures were adopted: niche marginality and niche breadth, based on the outlying mean index (OMI) analysis (Dolédec et al., 2000). OMI is a non-centred PCA of the differences between the mean used and the mean available conditions for each environmental variable and species. In other words, this analysis is explicitly designed to account for the niche of each species in the assemblage, by searching for linear combinations of environmental conditions that maximize the average species niche marginality. In this context, “niche marginality” is defined as the squared Euclidean distance between the centroid of species in multivariate environmental space (i.e. the average conditions of the grids in which they occur) and the average available environmental conditions across studied grids (Dolédec et al., 2000). Therefore, high niche marginality indicates species specialized to rare habitat conditions. On the other hand, low values may indicate either

widespread species or those specialized to common habitats, that is a hypothetical species uniformly distributed across all grids or alternatively those species restricted to the mean regional environmental condition, would both exhibit zero marginality as their centroids approach the region centroid. This ordination technique also ensures that species' niche breadth or tolerance to environmental conditions are measured, as the range of a given species distribution in the multivariate environmental space. After standardizing environmental variables to z-scores (i.e. mean = 0, $SD = 1$), we performed principal components analysis on the environmental data with the *dudi.pca* function, and retained the first three principal component axes to calculate the marginality and breadth of species niches using the *niche* function, both in the "ade4" R package (Chessel et al., 2012).

2.6 | Functional traits

To address the influence of species traits as drivers of abundance, occupancy and AORs, we compiled data describing four key functional traits that are widely recognized as drivers of species assembly along environmental gradients (including human-related disturbances) in seasonally dry tropical forests (Hulshof et al., 2013). We considered plant resource-use traits from both leaves (specific leaf area—SLA [cm^2/g]) and stems (wood density—WD [g/cm^3]) and regenerative traits (seed mass—SM [mg], primary seed dispersal syndrome—DS [0 = abiotic, 1 = biotic]). Plant trait data were partly available in Sfair et al. (2018), Pinho et al. (2019) and Ribeiro et al. (2019), and were additionally compiled from the BIEN (Maitner et al., 2018), TreeCo (Lima et al., 2020) and TRY databases (Kattge et al., 2020).

2.7 | Data analysis

First, local abundance, seed mass, SLA and niche attributes were log-transformed to reduce skewness in data distribution. We also standardized predictors before running each model at each grid size scale for each biogeographic region (Caatinga and Cerrado). To assess the drivers of occupancy (logit-transformed), we used beta-regression models, given the proportional nature of the response variable which is bounded between 0 and 1 (Ferrari & Cibrari-Neto, 2004), while similar linear regression models were used to assess the drivers of local abundance (log-transformed). This variable transformation procedure has been commonly adopted in the specialized AOR literature (e.g. Gaston et al., 2000; Heino & Tolonen, 2018) as it approximates variables and model residuals to a normal distribution. Local abundance and occupancy were also considered as predictors of each other (i.e. AOR terms), in addition to niche attributes (marginality and breadth) and functional traits. In beta-regression models for occupancy, interactive terms between local abundance and traits were considered to address trait-mediated mechanisms in AORs. Such interactive terms are informative in AOR patterns for different functional groups, indicating the signal and strength of their deviation from the overall AOR (Webb et al., 2017). We additionally

performed partial regressions to assess the unique effects of individual predictors. All variance inflation factor (VIF) values of regression models were <4 , ensuring that all predictors could be retained across models. Beta regression models were applied using the "betareg" R package (Cibrari-Neto & Zeileis, 2010). As each observation of our response variables (local abundance and occupancy) integrates multiple spatial coordinates, it was not possible to assess spatial autocorrelation in model residuals. Instead, we examined spatial autocorrelation in our abundance matrices by applying Mantel tests and plotting correlograms using the "mantel.correlog" function from the "vegan" package (Oksanen et al., 2020), which did not show any signs of spatial structure ($p > .1$; Figure S4).

To account for potential sampling/statistical artefacts in observed AORs and the effects of niche breadth and marginality on occupancy (see Craven et al., 2021; Díaz et al., 2020), we applied three separated null models, one for each focal relationship. In all null models, we randomized species distribution across grid-cells while retaining a constant species frequency and grid-cells richness, using the "independentswap" algorithm from the "picante" R package (Kembel et al., 2010). This randomization procedure removes any species-environment association, and adequately tests the significance of niche-based processes (e.g. Díaz et al., 2020) and AORs (e.g. Craven et al., 2021) against potentially biased sampling effects. In each randomization of the community matrix, we calculated the local abundance and occupancy of each species in the case of AOR null models, and niche marginality and breadth in their respective null models, while maintaining other variables included in beta-regression models constant. Finally, the slopes of focal relationships in each of the 999 null model iterations were calculated and compared to observed slopes. We considered significant deviations from random expectations if fewer than 5% of all null slopes were greater (i.e. more positive or negative) than the observed slopes. All analyses were performed in R 4.0.4 (R Core Team, 2018).

3 | RESULTS

3.1 | AORs across biogeographic regions and spatial scales

At all scales of analysis in both the Caatinga and Cerrado regions, most species were locally scarce and geographically restricted, while few species were locally abundant and/or widespread (Figure 2a,c). Over 70% of species in each region occupied fewer than 20% of occupied spatial grid-cells (i.e. cells with available data), averaging less than 2% of the total number of individuals at any given scale both in Caatinga and Cerrado (Figure S5). On the other hand, the most ubiquitous species occupied over 80% of all spatial grid-cells in both regions, and the most locally abundant species on average represented up to 40% and 10% of total abundance of occupied grid-cells in Caatinga and Cerrado, respectively (Figure S5). Species that were either widespread or geographically restricted at one scale tended to be also at other scales (Table S2) and across the entire regions

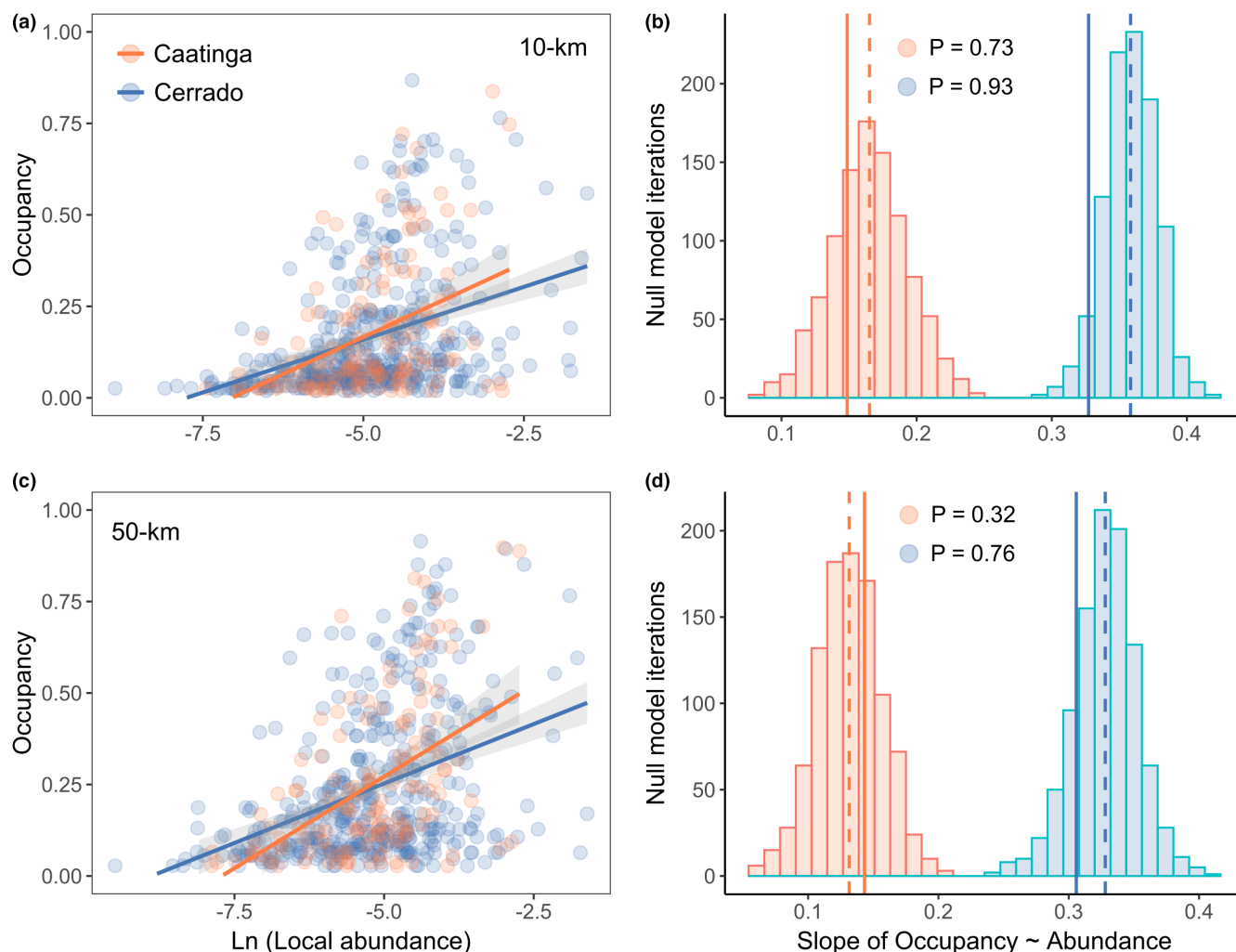


FIGURE 2 Relationships between species occupancy (i.e. proportion of spatial grid-cells occupied) and mean relative abundance at occupied grid-cells in the Caatinga dry forest ($N = 135$ species) and Cerrado savannahs ($N = 417$ species). Left panels: Observed relationships across (a) 10-km grids ($p < .001$; pseudo- $R^2 = .11$ and $.21$ for Caatinga and Cerrado, respectively) and (c) 50-km grids ($p < .05$ and $< .001$; pseudo- $R^2 = .05$ and $.24$). Right panels: Testing for non-random relationships in each study region at (b) 10-km and (d) 50-km scales. Histograms describe the distribution of expected slopes between local abundance and occupancy after accounting for the effects of other predictors, from 999 iterations of a null model that maintain total species abundance and grid-cell species richness, while randomizing the distribution of individuals across grid-cells. Dashed lines indicate the mean expected slope from null models, whereas solid lines indicate observed slopes in the beta regression model for occupancy. Observed relationships did not deviate from the random expectation in any region or spatial scale

(Table S3). Common species were represented by different taxonomic groups in each region, for which some families were highly abundant and ubiquitous, mainly Fabaceae and Euphorbiaceae in Caatinga dry forests, and also Vochysiaceae in Cerrado savannahs (see Table S4 for a checklist of the top-ranking 20 species in terms of local abundance and occupancy in each region).

Abundance-occupancy relationships were positive but weak at all spatial scales in both the Caatinga (Pseudo- $R^2 = 5\%$ – 11%) and Cerrado (Pseudo- $R^2 = 20\%$ – 25%), with very similar patterns evident across scales and biotas (Figure 2a,c; see Figure S6 for a more complete cross-scale AOR analysis). After accounting for the effects of species traits and niche attributes, local abundance explained at most 7% and 17% of variation in occupancy among plant species across the Caatinga and Cerrado, respectively.

3.2 | Environmental gradients, species distribution and niche patterns

According to the outlying mean index analysis, the most important environmental variables affecting plant species distribution across Caatinga dry forests were tree cover, soil fertility, soil clay content, climatic water deficit and annual precipitation (Figure 3a-b and Table S5). Across Cerrado savannahs, plant species distributions were mostly driven by mean annual temperature (which is strongly related to elevation, see above *environmental and disturbance gradients*), climatic water deficit and tree cover (Figure 3c-d, Table S5).

There was considerable variation in species niche marginality and breadth among Caatinga and Cerrado plant species, with no clear association among these two niche attributes at any scale

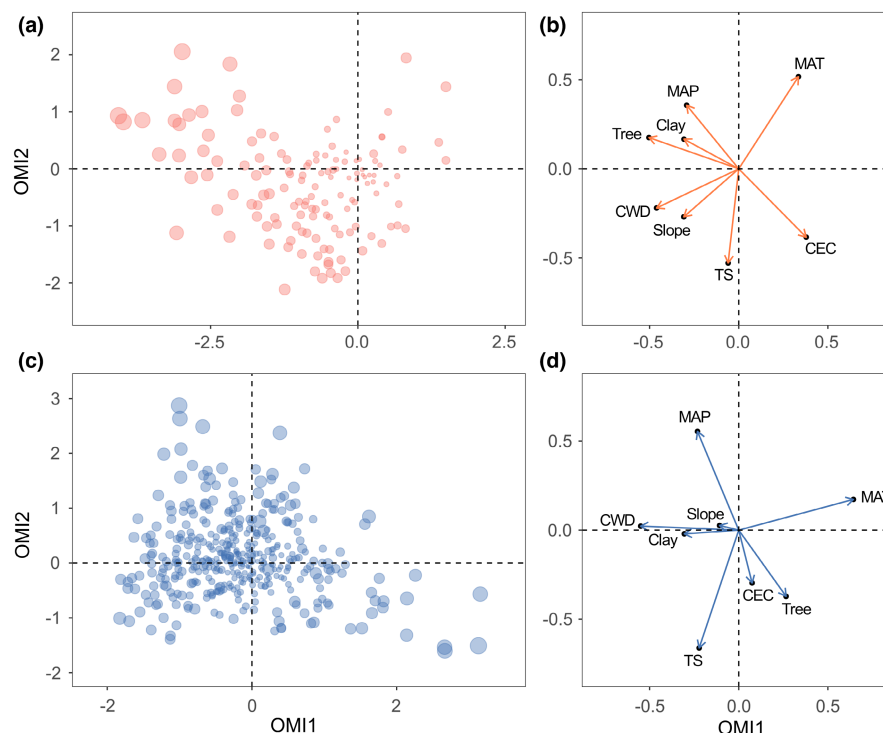


FIGURE 3 Distribution of plant species along the first two niche axes based on the outlying mean index (OMI) analysis and the related vectors of environmental drivers, in Caatinga (a-b; $N = 135$ species) and Cerrado (c-d; $N = 417$ species). Circle sizes in (a) and (c) are proportional to the variation in species niche marginality, which increase as they move away from zero. This represents the average available environmental conditions across the study grids in each region. Thus, larger circles depict high niche marginality values (i.e. species occurring in less common habitats). Species deviation from the regional average habitat conditions is driven by different environmental variables as indicated by vectors in the right panels. See Table S5 for correlation coefficients between species mean position along individual environmental gradients and their scores along the two main OMI axes. CEC, soil cation exchange capacity; Clay, soil clay content; CWD, cumulative water deficit; MAP, mean annual precipitation; MAT, mean annual temperature; Tree, tree cover; TS, temperature seasonality

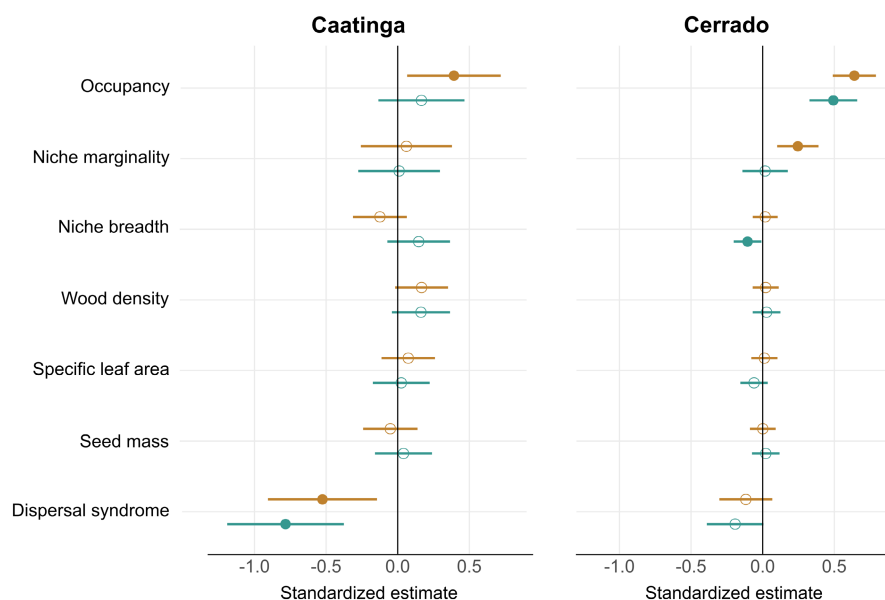


FIGURE 4 Drivers of cross-species variation in *local abundance* (i.e. mean relative abundances at occupied grid-cells). Points represent the standardized estimates from linear regression models examining the relative effects of occupancy, niche attributes, and functional traits across 10-km (brown) grids (adj. $R^2 = .16$ and $.22$ for Caatinga and Cerrado, respectively) and 50-km (green) resolution (adj. $R^2 = .14$ and $.21$). Filled circles denote significant relationships. See related statistics in Table S6

(Figure S7). Species characterized by high niche marginality in Caatinga dry forests were mostly associated with relatively wet climates, high tree cover and high soil clay content (Figure 3a-b), conditions typically found in transition zones between the Caatinga and neighbouring regions. Across the Cerrado, the

environmental signature of marginal niches was less clear, as both the driest and wettest conditions, or the most and least forested scenario represented marginal habitats (Figure 3c-d), while intermediate environmental conditions characterized common habitats (see Figure S1). Such opposite marginal habitat conditions

are likely to reflect the transition zones between the Cerrado and both drier and wetter adjacent biomes.

Finally, there was no clear functional signature associated with niche marginality (Figure S8) or niche breadth (Figure S9), as species with similar niche attributes were widely distributed in the functional space in both Caatinga and Cerrado.

3.3 | Drivers of local abundance

Species mean relative abundance at occupied grid-cells were significantly but weakly related to niche attributes and functional traits, which together with occupancy explained up to 19% and 25% of variance among species in the Caatinga and Cerrado, respectively (Figure 4 and Table S6). In addition to the positive weak effects of occupancy (maximum partial- $R^2 = 8\%$) across scales and regions (Figure 4), functional traits and niche attributes affected the local abundance of Caatinga and Cerrado plant species, respectively. Specifically, seed dispersal syndrome was the best predictor of local abundance in Caatinga plant species, in which abiotically dispersed species were more abundant, especially at large scales (Figure 4a). Local abundance in Cerrado plant species was negatively related to niche marginality at small scales and positively affected by niche breadth at large scales (Figure 4b).

3.4 | Drivers of occupancy

In contrast to the low predictability of local abundance, occupancy patterns were strongly associated with niche attributes and functional traits, which together with local abundance explained up to 71% and 74% of variance in this population attribute across Caatinga and Cerrado woody plant species, respectively (Figure 5

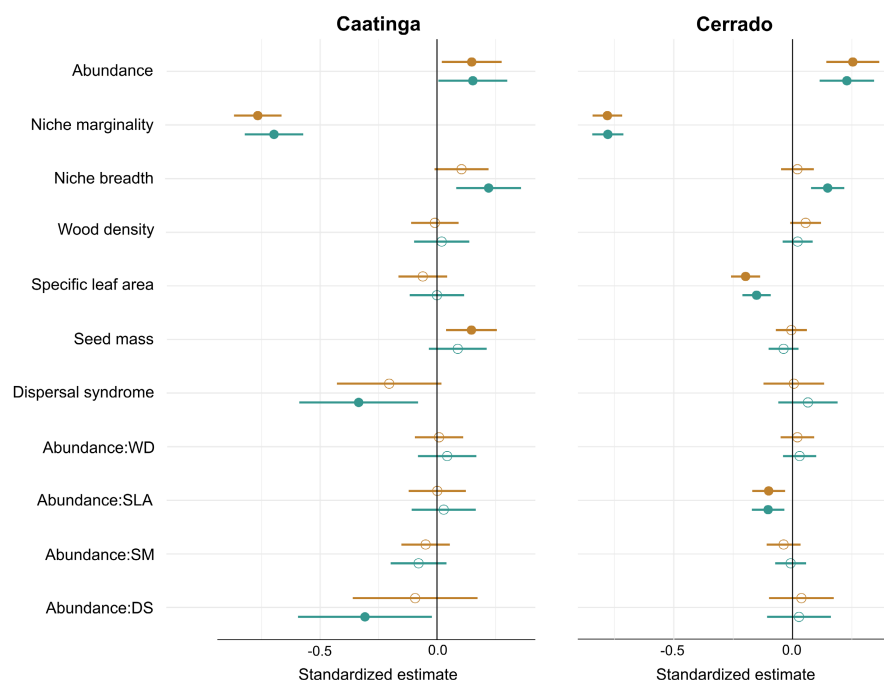
and Table S7). The effects of individual predictors were consistent in direction but variable in strength across spatial scales (Figure 5). Overall, niche marginality was the main predictor of occupancy (Figure 5), with consistent negative effects across scales and regions (Figure 6a,c; see Figure S10 for a more complete cross-scale analysis of such effects). In fact, niche marginality accounted for a large fraction of the explained variation in species occupancy across the Caatinga (partial- $R^2 = 46\%$ to 62% , depending on the scale) and Cerrado (partial- $R^2 = 59\%$ to 61%). In contrast, niche breadth was positively but weakly related to species occupancy (maximum partial- $R^2 = 8\%$), with significant effects only at relatively large scales (Figure 7a,c; see Figure S11 for a more complete cross-scale analysis).

Functional traits had additional direct effects on occupancy and also mediated AORs (Figure 5, Table S7). Specifically, large-seeded Caatinga woody plant species tended to be more widespread at small grid resolution regardless of their local abundance, while at larger scales abiotically dispersed species were most widespread and presented stronger AORs across Caatinga dry forests (Figure 5a). Occupancy and positive AORs in Cerrado species were consistently driven by low specific leaf area (Figure 5b). In fact, widespread Cerrado species had a clearer functional signature towards low specific leaf area and, to a lesser extent, high wood density and large seeds, while widespread and rare Caatinga species were widely distributed in the functional space (Figure S11).

3.5 | Disentangling artefacts from ecological effects

According to our null model analysis, observed positive AORs did not deviate from what would be expected from random dispersal of individuals across grid-cells, at any scale in any region (Figure 2b,d). In

FIGURE 5 Drivers of cross-species variation in occupancy (i.e. proportion of grid-cells occupied). Points represent the standardized estimates from beta regression models examining the relative effects of local abundance, niche attributes, functional traits and their interactions with local abundance, across 10-km (brown) grids (adj. Pseudo- $R^2 = .71$ and $.73$ for Caatinga and Cerrado, respectively) and 50-km (green) resolution (adj. Pseudo- $R^2 = .61$ and $.74$). Filled circles denote significant relationships. DS, dispersal syndrome; SLA, specific leaf area; SM, seed mass; WD, wood density. See related statistics in Table S7



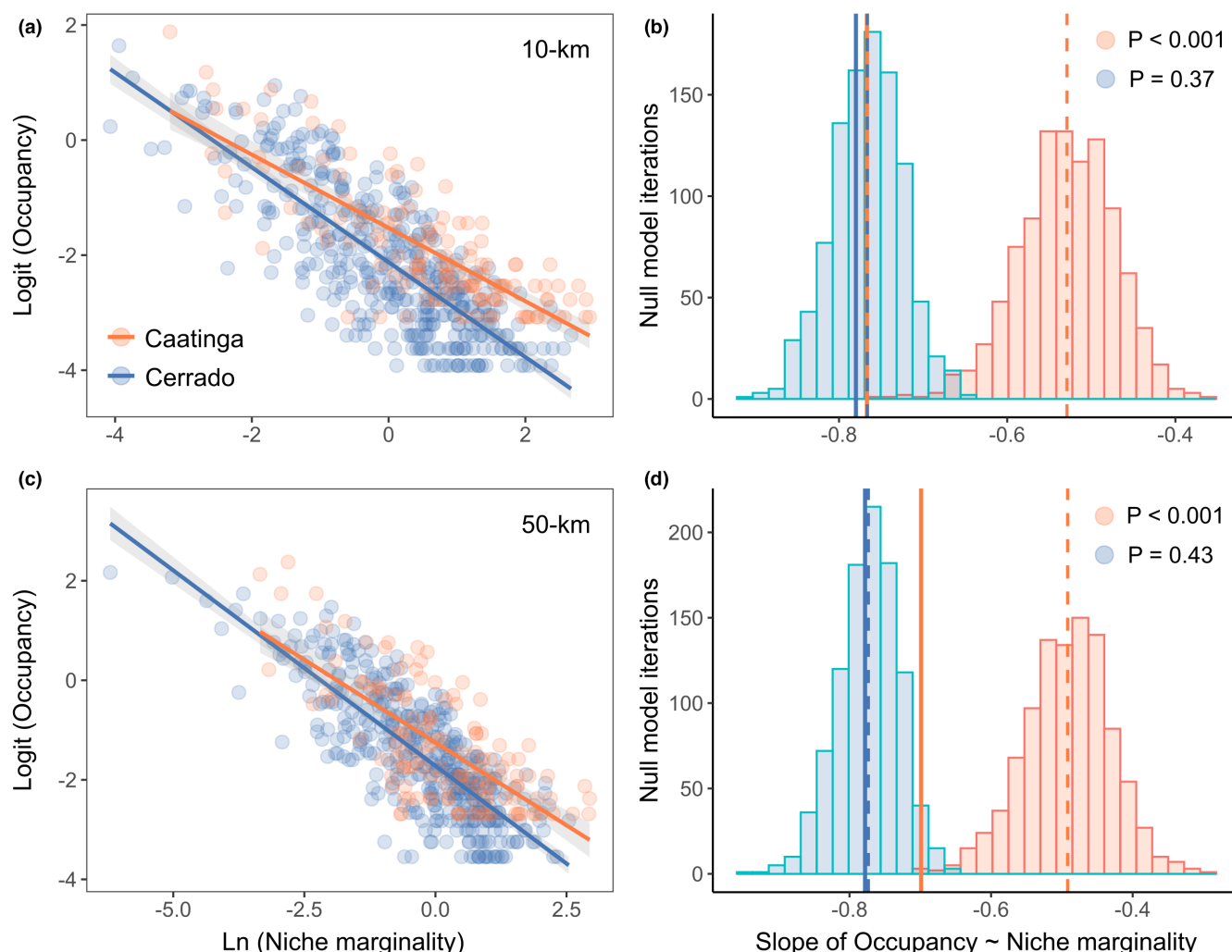


FIGURE 6 Relationships between species occupancy and niche marginality in the Caatinga dry forest and Cerrado savannah. Left panels: Observed relationships across (a) 10-km grids ($p < .001$; pseudo- $R^2 = .64$ and $.58$ for Caatinga and Cerrado, respectively) and (c) 50-km grids ($p < .001$; pseudo- $R^2 = .48$ and $.59$). Right panels: Testing for non-random relationships in each region at (b) 10-km and (d) 50-km scales. Histograms describe the distribution of expected slopes between occupancy and niche marginality (after accounting for the effects of other predictors), from 999 iterations of a null model that remove any associations between species distribution and abiotic variables, while preserving their observed abundance and occupancy. Dashed lines indicate the mean expected slope from the null models, whereas solid lines indicate observed slopes in the beta regression model (Table S7). Observed relationships were significantly more negative than expected by chance in Caatinga across all scales ($p < .05$), but did not deviate from random expectation in Cerrado at any of the scales

contrast, the strong relationships observed between occupancy and niche marginality were significantly more negative than expected by chance in the Caatinga dry forest (at all scales), but did not deviate from random expectations in the Cerrado at any scale (Figure 6b,d). Similarly, the observed positive effects of niche breadth on occupancy only deviated from random expectations in the Caatinga dry forest, at the largest scale (Figure 7b,d).

4 | DISCUSSION

Our results suggest that (1) AORs are positive but weak as some species are locally abundant but geographically restricted across scales and regions; (2) variation around AORs (commonness and rarity) in

dryland woody plant species is strongly driven by habitat availability (i.e. niche marginality) and, to a lesser extent, by environmental tolerance (i.e. niche breadth) and functional traits. Moreover, our null models suggest that positive AORs may arise from sampling artefacts in estimating occupancy of locally scarce species. Also, while species-environment associations leading to niche effects are distinct from a random-dispersal expectation across Caatinga dry forests, a more dispersal-driven species assembly likely occurs across Cerrado savannahs. In any case, many Caatinga and Cerrado plants could be described as habitat-specialist rare species, as they are both narrowly distributed and occupy rare habitats. In Caatinga, these important habitats are characterized by wetter climates, clay soils and less intensive land-use regimes, while in Cerrado, they comprise either extremely wet or dry conditions. These conditions

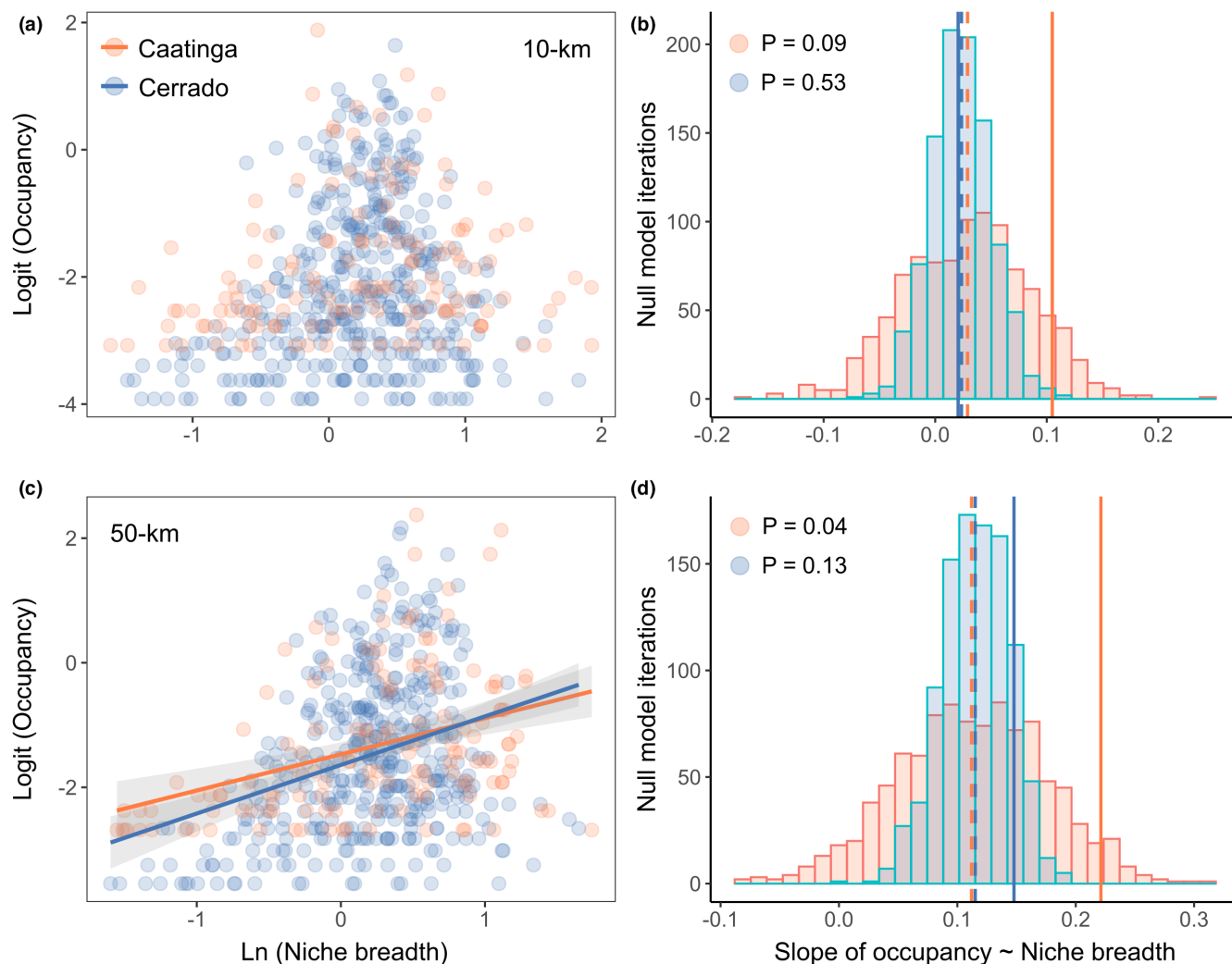


FIGURE 7 Relationships between species occupancy and niche breadth in the Caatinga dry forest and Cerrado savannahs. Left panels: Observed relationships across (a) 10-km grids ($p > .05$; pseudo- $R^2 = .04$ and $.02$ for caatinga and Cerrado, respectively) and (c) 50-km grids ($p < .001$; pseudo- $R^2 = .16$ and $.08$). Right panels: Testing for non-random relationships in each study region at (b) 10-km and (d) 50-km scales. Histograms describe the distribution of expected slopes between occupancy and niche breadth (after accounting for the effects of other predictors), from 999 iterations of a null model that remove any associations between species distribution and abiotic variables, while preserving their observed abundance and occupancy. Dashed lines indicate the mean expected slope from the null models, whereas solid lines indicate observed slopes in the beta regression model (Table S7). The observed relationships did not deviate from random expectations ($p > .05$), except for the 50-km scale in Caatinga, where it was significantly more positive than expected by chance

are consistent with those observed across transition zones between neighbouring biomes, revealing the importance of such marginal habitats for maintaining regional plant diversity and associated ecological functions. Trait-mediated mechanisms in AORs reflect specific environmental or ecological filters evident across regions and scales, with prominent positive effects of seed mass and abiotic seed dispersal on occupancy at relatively small and large spatial scales in Caatinga, respectively, while conservative leaves favour species occupancy and positive AORs across scales in Cerrado. These findings demonstrate context- and scale-dependency on the drivers of species rarity and AOR patterns, and reveal a deterministic taxonomic and functional macro-organization of Caatinga dry forests and Cerrado savannahs along environmental gradients from local to landscape scales.

Strong positive AORs are expected to emerge if similar processes underlie species local abundance and regional occupancy or in the case of strong positive synergy among these population attributes (Borregaard & Rahbek, 2010), which is not consistent with our data on Caatinga and Cerrado woody plant species. In fact, although fairly common, AORs are weak in many taxa and spatial scales (Caten et al., 2022), as these relationships are influenced by functional traits, niche-based processes and sampling artefacts (Craven et al., 2021; Díaz et al., 2020; Hanski et al., 1993). We acknowledge that sampling artefacts related to incomplete and uneven distribution of available data may affect estimation of occupancy of locally scarce species and thus lead to positive (albeit weak) AORs by chance. However, the main reason for the lack of strong positive AORs in our study was the presence of geographically restricted but locally abundant

species. This is in fact a type of rarity, which compensates for geographic rarity by high performance under suitable conditions (Williams et al., 2009). This pattern agrees with previous evidence of relatively aggregated distributions exhibited by tropical dry forest tree species (Hubbell, 1979; Williams et al., 2010). Also, the low predictability of local abundance likely results from biotic interactions and undetected environmental filters, including chronic human disturbances that lead to either declines or proliferation of plant local populations (Ribeiro et al., 2019; Sfair et al., 2018). In contrast, species occupancy appears to be highly predictable and mostly driven by similar niche-based processes across scales in both Caatinga and Cerrado. Geographically widespread woody plant species in these regions, from local to landscape scales, are mostly those adapted to common habitats, whereas rare species typically rely on rare habitat conditions. This is consistent with classic postulates in macroecology (Hanski et al., 1993) as well as recent evidence from tropical wet forest regions (Díaz et al., 2020) and other biotas (e.g. Heino & Tolonen, 2018; Marino et al., 2020), highlighting that habitat availability is a key occupancy driver across taxa, biotas and spatial scales. In contrast, the niche breadth hypothesis, which has been the cornerstone of niche-based theory (Brown, 1984), only exerted weak support at large scales, as species with wider niches were often geographically restricted. Such pattern is likely associated with dispersal limitation (Pagel et al., 2020), which has been previously invoked as a pervasive process explaining patterns of patchy distribution in woody plant species across drylands (Pennington et al., 2009).

In addition to niche-based processes, our findings suggest that functional traits related to plant resource-use, dispersal and regeneration strategy directly affect species occupancy patterns and mediate AORs from local to landscape scales. Interestingly, observed trait-mediated mechanisms agree with specific ecological filters evident at different scales in each region. For instance, local establishment filters are likely best pinpointed across small grids, over which large-seeded Caatinga plant species were more widespread, presumably because large seeds promote high initial seedling performance and survival across stressful (e.g. desiccated) local conditions (Muller-Landau, 2010; see Costa et al., 2020 and Trindade et al., 2020 for Caatinga examples). On the other hand, abiotic dispersal favours species occupancy across Caatinga dry forests particularly at larger spatial scales, where dispersal limitation is expected to be a more prominent driver of species distribution and community assembly (Asefa et al., 2020; Rito et al., 2021). The negative deviation for animal-dispersed species from AORs at larger scales may reflect a dispersal failure due to severe level of defaunation and widespread extirpations of large-bodied frugivore populations throughout the Caatinga (Bogoni et al., 2020). Finally, Cerrado species occupancy was apparently promoted by durable leaves at all spatial scales, which is consistent with high tolerance to dystrophic soils and wildfires, the typical filters for plant establishment in this context (Dantas et al., 2013). However, the effects of functional traits on occupancy across both biotas were relatively weak compared to the strong effects of niche attributes, which likely reflects the success of alternative functional profiles. For instance, both drought-avoidance and drought-tolerance strategies can overcome

common abiotic filters in tropical dry forests (Hulshof et al., 2013; see Pinho et al., 2019), as shown by the diverse functional profiles across both common and rare Caatinga plant species (Figure S12).

Collectively, our findings indicate the occurrence of spatially organized regional floras along environmental and disturbance gradients (Bueno et al., 2018; Silva & Souza, 2018). Nevertheless, species-environment associations observed in the Cerrado, to some extent, may have emerged from random processes, as indicated by our null model analysis. This result does not invalidate niche patterns observed across Cerrado species, but rather indicates that, given their abundance and the regional mosaic of environmental conditions, we cannot clearly distinguish those patterns from what would be expected from random dispersal. There are two possible explanations for this pattern. First, an oligarchy of few highly widespread species occupies virtually the entire Cerrado region (Bridgewater et al., 2004; Figure S5) and thus consistently exhibit low niche marginality as their niche position will always be close to the regional average environmental conditions. Alternatively, such pattern may reflect the patchy distribution of Cerrado habitats from local to regional scales (Bueno et al., 2018; Silva et al., 2006). In sum, we remain unable to determine whether the observed aggregated distribution of Cerrado species in particular habitats results from environmental selection or random dispersal. In any case, patchy gradients of physical conditions in both the Caatinga and Cerrado have apparently provided a constellation of opportunities for plant specialization as indicated by the large number of habitat-specialist rare species, which is consistent with the high level of species endemism in both floras (Pennington et al., 2006).

This general macroecological pattern poses critical implications for biodiversity conservation and restoration strategies, given that a considerable number of Caatinga and Cerrado plant species bear two key attributes associated with extinction risk: limited distribution and habitat specialization (Hubbell et al., 2008; Rabinowitz, 1981). The restricted regional distribution of most plant species in these drylands is a pattern shared by many tropical biotas (e.g. Hubbell et al., 2008), and a primary indicator of extinction risk (IUCN, 2020). In addition, rare Caatinga habitats experience relatively wet climates with less prolonged droughts and high forest cover, which does not bode well as these conditions are most threatened by climate change and land-use intensification (IPCC, 2019). On the other hand, habitat-specialist rare Cerrado species were associated with transitional zones into adjacent drier or wetter biomes, reinforcing the importance of these marginal habitats as strongholds of regional biodiversity (Françoso et al., 2016). Although some of the rare species we detected may straddle the boundaries between adjacent biomes and are restricted to extremes of their distribution, many are either Caatinga or Cerrado specialists. These rare habitats may therefore represent their last refugia or are irreplaceable for meta-population persistence (Freckleton et al., 2005).

In synthesis, we provide a mechanistic basis for understanding cross-scale patterns of woody plant species commonness and rarity in Caatinga dry forests and Cerrado savannahs, which may apply to other tropical ecosystems. Our results highlight that considering

AOR patterns across scales, jointly with niche attributes and functional traits, enhances both theoretical and applied aspects of tropical dryland community organization. From a theoretical perspective, we show that AORs are weak in tropical dryland woody plants, a pattern related to low habitat availability, unpredictability in local abundances and likely dispersal limitation. In this context, climate, edaphic and disturbance gradients drive plant species distributions from local to landscape scales. Brazilian dryland plant assemblages therefore emerge as, at least to some extent, environmentally structured across multiple spatial scales, with a high proportion of habitat-specialist rare species associated with a wide spectrum of habitats. From an applied perspective, conservation actions should prioritize the site protection of rare plant species, their animal mutualists, and their rare habitats, since they will become increasingly vulnerable under forecasted scenarios of increasing aridity and land-use intensification.

ACKNOWLEDGEMENTS

This study was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq (grant 403770/2012-2), and a Newton Fund Institutional Partnership, awarded by the Royal Society to IRL and CAP, between the University of East Anglia (UK) and Universidade Federal de Pernambuco (Brazil). BXP was funded by a postdoc fellowship (BFP-0164-2.05/19) from Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco – FACEPE. The compilation of TreeCo plot and trait data was supported by the São Paulo Research Foundation—FAPESP (grant 2013/08722-5). FPLM, IRL and MT also thank CNPq for productivity grants. No specific permits were required to carry out this study. We thank Artur Wanderley for ideas in designing this study.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest related to this paper.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ddi.13587>.

DATA AVAILABILITY STATEMENT

All the data that support findings of this study are available in the Dryad repository: <https://doi.org/10.5061/dryad.zcrjdfnfc>.

ORCID

Bruno X. Pinho  <https://orcid.org/0000-0002-6588-3575>

Renato A. F. de Lima  <https://orcid.org/0000-0002-1048-0138>

REFERENCES

- Aguiar-Gutiérrez, J., Malhi, Y., Lewis, S. L., Fauset, S., Adu-Breda, S., Affum-Baffoe, K., Baker, T. R., Gvozdevaite, A., Hubau, W., Moore, S., Peprah, T., Ziemnińska, K., Phillips, O. L., & Oliveras, I. (2020). Long-term droughts may drive drier tropical forests towards increased functional, taxonomic and phylogenetic homogeneity. *Nature Communications*, 11, 3346.
- Andrade, E. M., Aquino, D., Chaves, L. C. G., & Lopes, F. B. (2017). Water as capital and its uses in the caatinga. In J. M. C. Silva, I. R. Leal, & M. Tabarelli (Eds.), *Caatinga: The largest tropical dry forest region in South America*. Springer International Publishing.
- Asefa, M., Wen, H.-D., Cao, M., & Hu, Y.-H. (2020). Key community assembly processes switch between scales in shaping beta diversity in two primary forests, Southwest China. *Forests*, 11, 1106.
- Bogoni, J. A., Peres, C. A., & Ferraz, K. M. (2020). Extent, intensity and drivers of mammal defaunation: A continental-scale analysis across the neotropics. *Scientific Reports*, 10, 1–16.
- Borregaard, M., & Rahbek, C. (2010). Causality of the relationship between geographic distribution and species abundance. *The Quarterly Review of Biology*, 85, 3–25.
- Brazil Flora Group. (2019). *Brazilian Flora 2020 project – Projeto Flora do Brasil 2020*, v393.179. Inst. de Pesquisas Jardim Botânico do Rio de Janeiro.
- Bridgewater, S., Ratter, J. A., & Ribeiro, J. F. (2004). Biogeographic patterns, β -diversity and dominance in the Cerrado biome of Brazil. *Biodiversity and Conservation*, 13, 2295–2317.
- Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, 124, 255–279.
- Buckley, H. L., & Freckleton, R. P. (2010). Understanding the role of species dynamics in abundance–occupancy relationships. *Journal of Ecology*, 98, 645–658.
- Bueno, M. L., Dexter, K. G., Pennington, R. T., Pontara, V., Neves, D. M., Ratter, J. A., & Oliveira-Filho, A. T. (2018). The environmental triangle of the Cerrado domain: Ecological factors driving shifts in tree species composition between forests and savannas. *Journal of Ecology*, 106, 2109–2120.
- Carvalho, G. (2016). *flora: tools for interacting with the Brazilian flora 2020*. R package ver. 0.2.7. <https://CRAN.R-project.org/package=flora>
- Caten, C. T., Holian, L., & Dallas, T. (2022). Weak but consistent abundance–occupancy relationships across taxa, space and time. *Global Ecology and Biogeography*, 31, 968–977.
- Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., ... Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, 20, 3177–3190.
- Chessel, D., Dufour, A.-B., & Dray, S. (2012). ADE-4: Analysis of ecological data. *Exploratory and euclidean methods in environmental sciences*. <http://CRAN.R-project.org/package=ADE4>
- Cibbari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, 34, 1–24.
- Costa, A., Stropp, J., de Carvalho, N. A., Alves-Martins, F., Ladle, R. J., & Malhado, A. C. M. (2020). Environmental correlates of seed weight of tropical semi-arid woody species. *Plant and Soil*, 446, 369–378.
- Craven, D., Weigelt, P., Wolkis, D., & Kreft, H. (2021). Niche properties constrain occupancy but not abundance patterns of native and alien woody species across Hawaiian forests. *Journal of Vegetation Science*, 32, e13025.
- Dantas, V., Batalha, M. A., & Pausas, J. G. (2013). Fire drives functional thresholds on the savanna–forest transition. *Ecology*, 94, 2454–2463.
- Díaz, D. M., Blundo, C., Cayola, L., Fuentes, A. F., Malizia, L. R., & Myers, J. A. (2020). Untangling the importance of niche breadth and niche position as drivers of tree species abundance and occupancy across biogeographic regions. *Global Ecology and Biogeography*, 29, 1542–1553.
- Dolédéc, S., Chessel, D., & Gimaret-Carpentier, C. (2000). Niche separation in community analysis: A new method. *Ecology*, 81, 2914–2927.
- DRYFLOR. (2016). Plant diversity patterns in neotropical dry forests and their conservation implications. *Science*, 353, 1383–1387.
- Enquist, B. J., Feng, X., Boyle, B., Maitner, B., Newman, E. A., Jørgensen, P. M., Roehrdanz, P. R., Thiers, B. M., Burger, J. R., Corlett, R. T., TLP, C., Dauby, G., Donoghue, J. C., Foden, W., Lovett, J. C.,

- Marquet, P. A., Merow, C., Midgley, G., Morueta-Holme, N., ... McGill, B. (2019). The commonness of rarity: Global and future distribution of rarity across land plants. *Science Advances*, 5(11), eaaz0414.
- Fernandes, M. F., Cardoso, D., & de Queiroz, L. P. (2020). An updated plant checklist of the Brazilian Caatinga seasonally dry forests and woodlands reveals high species richness and endemism. *Journal of Arid Environments*, 174, 104079.
- Ferrari, S. L. P., & Cibrari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31, 799–815.
- Fick, S. E., & Hijmans, R. J. (2017). Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315.
- Françoso, R. D., Haidar, R. F., & Machado, R. B. (2016). Tree species of South America central savanna: Endemism, marginal areas and the relationship with other biomes. *Acta Botanica Brasilica*, 30, 78–86.
- Freckleton, R. P., Gill, J. A., Noble, D., & Watkinson, A. R. (2005). Large-scale population dynamics, abundance–occupancy relationships and the scaling from local to regional population size. *Journal of Animal Ecology*, 74, 353–364.
- Gaston, K. J. (1996). The multiple forms of the interspecific abundance–distribution relations. *Oikos*, 76, 211–220.
- Gaston, K. J. (1998). Rarity as double jeopardy. *Nature*, 394, 229–230.
- Gaston, K. J., Blackburn, T. M., Greenwood, J. J. D., Gregory, R. D., Quinn, R. M., & Lawton, J. H. (2000). Abundance–occupancy relationships. *Journal of Applied Ecology*, 37, 39–59.
- Grantham, H. S., Duncan, A., Evans, T. D., Jones, K. R., Beyer, H. L., Schuster, R., ... Watson, J. E. M. (2020). Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nature Communications*, 11, 5978.
- Grime, J. P., & Pierce, S. (2012). *The evolutionary strategies that shape ecosystems*. Wiley-Blackwell.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342, 850–853.
- Hanski, I., Kouki, J., & Halkka, A. (1993). Three explanations of the positive relationship between distribution and abundance of species. In R. E. Ricklefs & D. Schluter (Eds.), *Species diversity in ecological communities* (pp. 108–116). University of Chicago Press.
- Heino, J., & Tolonen, K. T. (2018). Ecological niche features override biological traits and taxonomic relatedness as predictors of occupancy and abundance in lake littoral macroinvertebrates. *Ecography*, 41, 2092–2103.
- Hengl, T., de Jesus, M. J., Heuvelink, G. B. M., Gonzalez, R. M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., RA, M. M., Batjes, N. H., Leenaars, J. G., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, 12, e0169748.
- Hubbell, S. (1979). Tree dispersion, abundance, and diversity in a tropical dry Forest. *Science*, 203, 1299–1309.
- Hubbell, S. P., He, F., Condit, R., Borda-de-Água, L., Kellner, J., & ter Steege, H. (2008). How many tree species are there in the Amazon and how many of them will go extinct? *Proceedings of the National Academy of Sciences*, 105, 11498–11504.
- Hulshof, C., Martínez-Yrizar, A., Búrquez, A., Boyle, B., & Enquist, B. (2013). Plant functional trait variation in tropical dry forests: A review and synthesis. In A. Sanchez-Azofeifa & J. S. Powers (Eds.), *Tropical dry forests in the Americas: Ecology, conservation and management*. CRC Press.
- IPCC – Intergovernmental Panel on Climate Change. (2019). *Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. [Core Writing Team, A. Arneth, H. Barbosa, T. Benton et al. (Eds.)]. Intergovernmental Panel on Climate Change (IPCC).
- IUCN. (2020). *The IUCN Red List of Threatened Species*. Version 2019-3.
- Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar, C., Aleixo, I., ... Wirth, C. (2020). TRY plant trait database – Enhanced coverage and open access. *Global Change Biology*, 26, 119–188.
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464.
- Laurance, W. F., Sayer, J., & Cassman, K. G. (2014). Agricultural expansion and its impacts on tropical nature. *Trends in Ecology and Evolution*, 29, 107–116.
- Lima, R. A. F., Oliveira, A. A., Pitta, G. R., Gasper, A. L., Vibrans, A. C., Chave, J., ter Steege, H., & Prado, P. I. (2020). The erosion of biodiversity and biomass in the Atlantic Forest biodiversity hotspot. *Nature Communications*, 11, 6347.
- Maitner, B. S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S. M., Guaderrama, D., Hinchliff, C. E., Jørgensen, P. M., Kraft, N. J. B., McGill, B., Merow, C., Morueta-Holme, N., Peet, R. K., Sandel, B., Schildhauer, M., Smith, S. A., Svenning, J.-C., Thiers, B., ... Enquist, B. J. (2018). The BIEN R package: A tool to access the botanical information and ecology network (BIEN) database. *Methods in Ecology and Evolution*, 9, 373–379.
- Marino, N. A. C., Céréghino, R., Gilbert, B., Petermann, J. S., Srivastava, D. S., de Omena, P. M., Bautista, F. O., Guzman, L. M., Romero, G. Q., Trzcinski, M. K., Barberis, I. M., Corbara, B., Debastiani, V. J., Dézerald, O., Kratina, P., Leroy, C., MacDonald, A. A. M., Montero, G., Pillar, V. D., ... Farjalla, V. F. (2020). Species niches, not traits, determine abundance and occupancy patterns: A multi-site synthesis. *Global Ecology and Biogeography*, 29, 295–308.
- Moro, M. F., Lughadha, E. N., de Araújo, F. S., & Martins, F. R. (2016). A phytogeographical Metaanalysis of the semiarid caatinga domain in Brazil. *The Botanical Review*, 82, 91–148.
- Muller-Landau, H. C. (2010). The tolerance-fecundity trade-off and the maintenance of diversity in seed size. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 4242–4247.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szöcs, E., Wagner, H. (2020). *Community ecology package*. R package version 2.5-7. <https://CRAN.R-project.org/package=vegan>
- Pagel, J., Treurnicht, M., Bond, E. J., Kraaij, T., Nottebrock, H., Schutte-Vlok, A., Tonnabel, J., Esler, K. J., & Schurr, F. M. (2020). Mismatches between demographic niches and geographic distributions are strongest in poorly dispersed and highly persistent plant species. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 3663–3669.
- Pennington, R. T., Lavin, M., & Oliveira-Filho, A. (2009). Woody Plant diversity, evolution and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annual Review of Ecology Evolution and Systematics*, 40, 437–457.
- Pennington, R. T., Lewis, G. P., & Ratter, J. A. (2006). *Neotropical savannas and seasonally dry forests*. CRC Press.
- Pinho, B. X., Peres, C. A., Leal, I. R., & Tabarelli, M. (2020). Critical role and collapse of tropical mega-trees: A key global resource. *Advances in Ecological Research*, 62, 253–294.
- Pinho, B. X., Tabarelli, M., Engelbrecht, B. M. J., Sfair, J., & Melo, F. P. L. (2019). Plant functional assembly is mediated by rainfall and soil conditions in a seasonally dry tropical forest. *Basic and Applied Ecology*, 40, 1–11.
- Queiroz, L. P., Cardoso, D., Fernandes, M. F., & Moro, M. (2017). Diversity and evolution of flowering plants of the caatinga domain. In J. C. Silva, I. Leal, & M. Tabarelli (Eds.), *Caatinga: The largest tropical dry Forest region in South America*. Springer.

- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org>
- Rabinowitz, D. (1981). Seven forms of rarity. In H. Synge (Ed.), *The biological aspects of rare plant conservation* (pp. 205–217). John Wiley & Sons Ltd.
- Reich, P. B. (2014). The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *Journal of Ecology*, 102, 275–301.
- Ribeiro, E. M. S., Arroyo-Rodríguez, V., Santos, B. A., Tabarelli, M., & Leal, I. R. (2015). Chronic anthropogenic disturbance drives the biological impoverishment of the Brazilian caatinga vegetation. *Journal of Applied Ecology*, 52, 611–620.
- Ribeiro, E. M. S., Lohbeck, M., Santos, B. A., Arroyo-Rodríguez, V., Tabarelli, M., & Leal, I. R. (2019). Functional diversity and composition of caatinga woody flora are negatively impacted by chronic anthropogenic disturbance. *Journal of Ecology*, 107, 2291–2302.
- Rito, K. F., Arroyo-Rodríguez, V., Cavender-Bares, J., Santo-Silva, E. E., Souza, G., Leal, I. R., & Tabarelli, M. (2021). Unraveling the drivers of plant taxonomic and phylogenetic β -diversity in a human-modified tropical dry forest. *Biodiversity and Conservation*, 30, 1049–1065.
- Rito, K. F., Arroyo-Rodríguez, V., Queiroz, R. T., Leal, I. R., & Tabarelli, M. (2017). Precipitation mediates the effect of human disturbance on the Brazilian caatinga vegetation. *Journal of Ecology*, 105, 828–838.
- Sfair, J. C., De Bello, F., Franca, T. Q., Baldauf, C., & Tabarelli, M. (2018). Chronic human disturbance affects plant trait distribution in a seasonally dry tropical forest. *Environmental Research Letters*, 13, e025005.
- Silva, J. F., Fariñas, M. R., Felfili, J. M., & Klink, C. A. (2006). Spatial heterogeneity, land use and conservation in the cerrado region of Brazil. *Journal of Biogeography*, 33, 536–548.
- Silva, J. M., Leal, I. R., & Tabarelli, M. (2017). *Caatinga: The largest tropical dry forest region in South America*. Springer International Publishing.
- Silva, A. C., & Souza, A. F. (2018). Aridity drives plant biogeographical sub regions in the caatinga, the largest tropical dry forest and woodland block in South America. *PLoS One*, 13, e0196130.
- Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by *in situ* evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 20359–20364.
- Slatyer, R. A., Hirst, M., & Sexton, J. P. (2013). Niche breadth predicts geographical range size: A general ecological pattern. *Ecology Letters*, 16, 1104–1114.
- Sporbert, M., Keil, P., Seidler, G., Bruehlheide, H., Jandt, U., Aćić, S., Biurrun, I., Campos, J. A., Čarni, A., Chytrý, M., Čuštěrevska, R., Dengler, J., Golub, V., Jansen, F., Kuzemko, A., Lenoir, J., Marcenò, C., Moeslund, J. E., Pérez-Haase, A., ... Welk, E. (2020). Testing macroecological abundance patterns: The relationship between local abundance and range size, range position and climatic suitability among European vascular plants. *Journal of Biogeography*, 47, 2210–2222.
- Sporbert, M., Welk, E., Seidler, G., Jandt, U., Aćić, S., Biurrun, I., Campos, J. A., Čarni, A., Cerabolini, B. E. L., Chytrý, M., Čuštěrevska, R., Dengler, J., De Sanctis, M., Dziuba, T., Fagúndez, J., Field, R., Golub, V., He, T., Jansen, F., & Bruehlheide, H. (2021). Different sets of traits explain abundance and distribution patterns of European plants at different spatial scales. *Journal of Vegetation Science*, 32, e13016.
- Strassburg, B., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., Latawiec, A. E., Filho, F. J. B. O., Scaramuzza, C. A. M., Scarano, F. R., Soares-Filho, B., & Balmford, A. (2017). Moment of truth for the Cerrado hotspot. *Nature Ecology and Evolution*, 1, e0099.
- ter Steege, H., Pitman, N. C. A., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., Phillips, O. L., Castilho, C. V., Magnusson, W. E., Molino, J. F., Monteagudo, A., Núñez Vargas, P., Montero, J. C., Feldpausch, T. R., Coronado, E. N., Killeen, T. J., Mostacedo, B., Vasquez, R., ... Silman, M. R. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342(6156), 1243092.
- Trindade, D. P. F., Sfair, J. C., de Paula, A. S., Barros, M. F., & Tabarelli, M. (2020). Water availability mediates functional shifts across ontogenetic stages in a regenerating seasonally dry tropical forest. *Journal of Vegetation Science*, 31, 1090–1101.
- Webb, T. J., Barry, J. P., & McClain, C. R. (2017). Abundance–occupancy relationships in deep sea wood fall communities. *Ecography*, 40, 1339–1347.
- Webb, T. J., Freckleton, R. P., & Gaston, K. J. (2012). Characterizing abundance–occupancy relationships: There is no artefact. *Global Ecology and Biogeography*, 21, 952–957.
- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A., & Wright, I. J. (2002). Plant ecological strategies: Some leading dimensions of variation between species. *Annual Review of Ecology and Systematics*, 33, 125–159.
- Williams, J. N., Viers, J. H., & Schwartz, M. W. (2010). Tropical dry forest trees and the relationship between local abundance and geographic range. *Journal of Biogeography*, 37, 951–959.
- Williams, S. E., Williams, Y. M., VanDerWal, J., Isaac, J. L., Shoo, L. P., & Johnson, C. N. (2009). Ecological specialization and population size in a biodiversity hotspot: How rare species avoid extinction. *Proceedings of the National Academy of Sciences*, 106, 19737–19741.
- Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Ritter, C. D., Edler, D., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svanteson, S., Wengtrom, N., Zizka, V., & Antonelli, A. (2019). COORDINATECLEANER: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecology and Evolution*, 10, 744–751.

BIOSKETCH

Bruno X. Pinho is a postdoc fellow at AMAP, Université de Montpellier, France. He is broadly interested in plant ecology, biogeography and biodiversity conservation, with particular focus on the functional assembly of plants in response to environmental gradients and human-imposed disturbances across tropical forests.

Author contributions: B.X.P. conceived the ideas and designed the methodology, with significant contribution from D.P.F.T.; B.X.P. compiled the data and led the analyses, with the support of D.J. in spatial analysis; B.X.P. wrote a first draft and led the writing; E.M.S.R., I.R.L. and R.A.F.L. provided plant abundance and trait data; M.T., C.P. and F.P.L.M. acquired funds and supervised the work; and all authors contributed critically to improve the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Pinho, B. X., Trindade, D. P. F., Peres, C. A., Jamelli, D., de Lima, R. A. F., Ribeiro, E. M. S., Melo, F. P. L., Leal, I. R., & Tabarelli, M. (2022). Cross-scale drivers of woody plant species commonness and rarity in the Brazilian drylands. *Diversity and Distributions*, 28, 1497–1511. <https://doi.org/10.1111/ddi.13587>