

## Disentangling environmental and spatial processes of community assembly in tropical forests from local to regional scales

**Gabriel Arellano, J. Sebastián Tello, Peter M. Jørgensen, Alfredo F. Fuentes, M. Isabel Loza, Vania Torrez and Manuel J. Macía**

*G. Arellano (orcid.org/0000-0003-3990-5344)(gabriel.arellano.torres@gmail.com) and M. J. Macía, Depto de Biología, Univ. Autónoma de Madrid, Calle Darwin 2, ES-28049 Madrid, Spain. – J. S. Tello, Center for Conservation and Sustainable Development, Missouri Botanical Garden, PO Box 299, St. Louis, MO 63166-0299, USA. JST also at: Escuela de Biología, Pontificia Univ. Católica del Ecuador, Av. 12 de Octubre 1076 y Roca, Apdo. 17-01-2184, Quito, Ecuador. – P. M. Jørgensen, Missouri Botanical Garden, PO Box 299, St. Louis, MO 63166-0299, USA. – Alfredo F. Fuentes and M. I. Loza, Herbario Nacional de Bolivia, Campus Universitario Cota-Cota, calle 27, Correo Central Cajón Postal 10077, La Paz, Bolivia. MIL also at: Dept of Biology, Univ. of Missouri, St. Louis, MO 63121, USA. – V. Torrez, Dept of Biology, Univ. of Leuven, BE-3001 Leuven, Belgium.*

Understanding patterns and mechanisms of variation in the compositional structure of communities across spatial scales is one of the fundamental challenges in ecology and biogeography. In this study, we evaluated the effects of spatial extent (i.e. size of study region) on: 1) whether community composition can be better explained by environmental (i.e. niche-based) or spatial (e.g. dispersal-based) processes; and 2) how climate and soils contribute to the influence of environment on plant community composition. We surveyed community composition across a network of 398 forest plots spanning a ~4000 m elevational gradient in the Madidi region in northwestern Bolivia. Using redundancy analyses and hierarchical variation partitioning, we disentangled the effects of environmental and spatial predictors on species composition, further decomposing the environmental effect between its climatic and soil components. We repeated analyses for 200 sub-regions ranging in spatial extent from ~250 to ~17 500 km<sup>2</sup>. Our analyses show a high degree of idiosyncrasy in results that come from different sub-regions. Despite this variability, we were able to identify various important patterns in the structure of tropical plant communities in our study system. First, even though sub-regions varied in size by nearly two orders of magnitude, the total amount of explained variation in community composition was scale independent; at all spatial scales, environment and space accounted for about 25% of the differences in community composition among plots. Second, the measured environmental effect was higher than the spatial effect on average and in the vast majority of sub-regions. This was true regardless of the spatial extent of analysis. Finally, we found that both climatic and soil variables accounted for significant fractions of variation, but climate was always more important than soils.

Understanding patterns and mechanisms of variation in the compositional structure of communities across spatial scales is one of the fundamental challenges in ecology and biogeography (Sandel 2015, Tello et al. 2015). Most explanations for the structure of biological communities depend on either niche-based processes, dispersal-based processes or a combination of the two (Chave 2009, Tuomisto et al. 2012). Niche assembly theory proposes that the presence and abundance of species are determined by their deterministic interactions with the abiotic and biotic environment (Chase and Leibold 2003). As a consequence, species are sorted along environmental gradients based on their requirements, and environmental heterogeneity places a fundamental role in structuring communities. In contrast, dispersal assembly theory argues that niche preferences of species are not required to explain species distributions and community-level patterns (Hubbell 2001). Variation in community composition through space and time is proposed to be the result of dispersal limita-

tion, demographic stochasticity (i.e. random drift of species abundances), and random speciation (Hubbell 2001). Although these two views of community assembly seem conflicting, research in the last decade has led to the conclusion that most plant communities are distributed along a continuum from purely deterministic to entirely neutral, and both types of processes are believed to contribute simultaneously to the structure of biological communities (Gravel et al. 2006, Adler et al. 2007). The factors that determine the position of a community along this 'niche to dispersal assembly' gradient, however, remain poorly understood.

Spatial scale can have very important effects on patterns and mechanisms of assembly in biological communities (Levin 1992, Münkemüller et al. 2014, Sandel 2015, Tello et al. 2015). There are multiple aspects that define the spatial scale at which biological communities can be studied (Levin 1992, Sandel 2015), but one of its most important elements is extent, which encompasses the distribution of all local

communities or sampling units and defines the total size of the region under study (Dungan et al. 2002). Changes in spatial extent can have significant effects on how important different assembly processes might become at structuring biological communities (Tello et al. 2015). For example, an increase in spatial extent is typically associated with increases in environmental heterogeneity, which in turn could lead to stronger species sorting. This could lead to variation in community structure at large spatial extents to be mostly driven by niche-based processes, where community composition is highly predictable and strongly associated with environmental gradients (Normand et al. 2006, Kristiansen et al. 2012, Pansonato et al. 2013). In contrast, small spatial extents can be associated with less environmental heterogeneity compared with larger extents in the same landscape, and it is at these scales where dispersal-based assembly processes might become most important (Chave 2009). Indeed, some of the few studies that have been conducted at different spatial extents in tropical forests suggest that there might be an increase in the importance of environment relative to spatial variables as a predictor of community composition as the size of regions increases (Kristiansen et al. 2012, López-Martínez et al. 2013, but see Normand et al. 2006 for a counterexample).

Spatial extent has also been proposed to have an important influence on the identity of the environmental factors that are important to the structure of communities. For plant communities, Siefert et al. (2012) proposed that soil properties are the most important determinants of community composition at small extents (<2000 km<sup>2</sup>), but that climate becomes the overriding environmental factor as spatial scale increases. Several independent studies seem to support this idea in tropical forests, showing that climatic predictors become more important at larger (regional to continental) than at smaller (local) extents (Bjorholm et al. 2008, Slik et al. 2009, Blach-Overgaard et al. 2010). Conversely, other authors have suggested that soils are the most important factors determining community composition at any spatial scale in these ecosystems (Vormisto et al. 2004, John et al. 2007, Jones et al. 2008a, Guèze et al. 2013).

Previous studies on how spatial extent influences patterns and mechanisms of community assembly have been few in tropical forests, and have been conducted using small sample sizes or covering a small range of spatial extents (Normand et al. 2006, Kristiansen et al. 2012). This has prevented a clear and rigorous evaluation of the effects of spatial extent on the relative importance of various community assembly mechanisms. In this study, we use data from an extensive network of forest plots along a ~4000 m tropical elevational gradient to understand the effects of spatial extent on 1) the relative importance of environmental (i.e. niche-based) and spatial (e.g. dispersal-based) processes on the structure of tropical plant communities, and 2) the relative contributions of climate and soils to the variation in community composition along environmental gradients. We do this by repeating analyses for multiple sub-regions of varying size, and comparing empirical patterns of community structure to null model expectations. Quantitative results differed considerably across sub-regions, but our analyses allowed us to identify several important community structural patterns that rise above this variability.

## Methods

### Study region and floristic data

For our study, we used a network of 398 0.1-ha (20 × 50 m) plots across the Madidi region on the eastern slopes of the Bolivian Andes (Fig. 1). The plot network included multiple types of mature tropical forests along a steep elevational gradient from Amazonia to high-elevation forests at around 4000 m in elevation (Fuentes 2005). Plots were located at least 250 m apart, avoiding large gaps or recent human disturbances. In each plot, we inventoried all woody plant individuals with a diameter at breast height (i.e. 130 cm from rooting point) equal to or greater than 2.5 cm. All individuals were identified to species or morphospecies, and 15 879 voucher specimens were collected to document each species at each site. Taxonomic identifications were standardized across the plot network. Individuals that could not be identified (1.69%) were excluded from analyses. The dataset contains information on the distribution of 118 895 individuals belonging to 2507 species. These data belong to the Madidi Project (<[www.mobot.org/madidi](http://www.mobot.org/madidi)>) and are available for consultation in Tropicos (<<http://tropicos.org/PlotSearch.aspx?projectid=20>>), which functions as a permanent data repository maintained by the Missouri Botanical Garden. Data availability and terms of use are described in the Data Sharing and Publication Policy document of the Madidi Project. Individuals interested in having access to the data need to submit a written request to the Project director (Peter M. Jørgensen, [peter.jorgensen@mobot.org](mailto:peter.jorgensen@mobot.org)).

### Environmental data: climate and soils

To explain the variation in species composition among plots, we used multiple environmental predictors that reflect both climatic and soil properties of each site. We estimated climatic conditions at each plot using the 19 bioclimatic variables in the WorldClim database (rasters of 30 arc-second resolution; Hijmans et al. 2005). These variables represent mean annual trends, annual temporal changes, as well as extreme or limiting conditions in temperature and precipitation derived from monthly data collected by weather stations.

Additionally, we measured soil properties at each plot using 10 variables representing edaphic conditions and resources. We collected superficial soil samples at each plot (0–30 cm below the litter layer), which were air-dried and sieved through a 2-mm sieve. Percentages of sand, silt and clay were measured with the hydrometer method (Reeuwijk 2002). Soil pH was measured in a 1:2.5 soil:H<sub>2</sub>O suspension. Organic carbon (C) was determined with the Walkley and Black method and total nitrogen (N) with the semi-micro Kjeldahl method. The C:N ratio was included as an additional soil variable. We also measured calcium (Ca), magnesium (Mg) and potassium (K) using two extractive methods for different sets of samples: 1) the 1M ammonium acetate solution method and 2) the Mellich-3 extraction method. These two methods are often linearly correlated in many types of soils and we used this fact to standardize all soil measurements into comparable values (details in Supplementary material Appendix 1 Methods A1 Fig. A1).

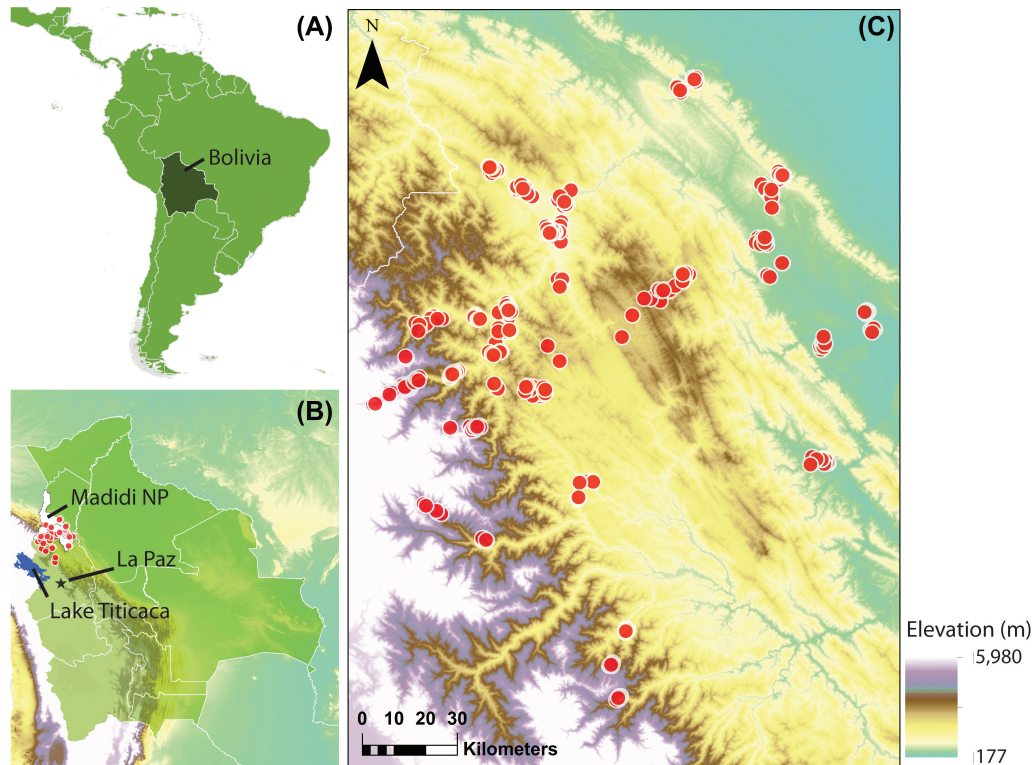


Figure 1. Distribution of 398 0.1-ha plots across the Madidi region in Bolivia. (A) Study region is in north-western Bolivia. (B) Forest plots are distributed along the eastern slopes of the Andes between latitudes  $-12.43^{\circ}$  and  $-15.72^{\circ}$ , and longitudes  $-69.48^{\circ}$  and  $-66.66^{\circ}$ , (C) ranging in elevation approximately from 250 to 4000 m.

### Defining sub-regions and measuring spatial extent

To study the effects of spatial extent on the relative contribution of different predictors to plant community composition, analyses were repeated for multiple sub-regions that varied broadly in size. A sub-region was defined by a unique subset of 60 plots from the pool of 398 plots in the dataset. Spatial extent was then measured as the area of the minimum convex polygon containing all 60 plots in a sub-region. Among the many possible sub-regions ( $> 1.1 \times 10^{72}$  potential combinations of 60 plots), we chose 200 so that: 1) spatial extent varied by nearly two orders of magnitude from local (256 km<sup>2</sup>) to regional (17 649 km<sup>2</sup>) scales, 2) all extents were represented in similar proportions, 3) spatial extent and average elevation were not correlated, thus guaranteeing that elevation was not a confounding variable, and 4) each of the selected sub-regions shared less than 70% of its plots with any other sub-region (Supplementary material Appendix 1 Methods A1 Fig. A2). Further details on how sub-regions were constructed and their characteristics (e.g. species richness, environmental heterogeneity) are presented in the Supplementary Material Appendix 1 Methods A2 and A3 Fig. A2 and A3. All subsequent analyses were repeated independently for each of the 200 sub-regions.

### Response and predictor matrices

For each sub-region, we decomposed the variation in community composition among multiple spatial and environmental predictor matrices. For these analyses, the response

matrix of species composition was created by performing a principal coordinate analysis (PCoA; Legendre and Legendre 1998) on a pair-wise matrix of compositional distances among plots within a sub-region. As a measure of compositional distance, we used Chao's dissimilarity index, which has been recommended for datasets with great compositional heterogeneity (Chao et al. 2005). All axes of the PCoA were retained after applying Cailliez's correction to avoid negative eigenvalues (Legendre and Legendre 1998). Chao's dissimilarities and the PCoA were calculated with functions `vegdist` and `cmdscale` in the package 'vegan' (Oksanen et al. 2013) of the R statistical software (ver. 3.1.2).

The full matrix of candidate environmental predictors was composed of 19 climatic predictors and 10 soil predictors (pH, C, N, sand, silt, clay, Ca, Mg, K and C:N). The full matrix of candidate spatial predictors was composed of longitude, latitude, and the principal coordinates of a neighborhood matrix (PCNMs). PCNMs represent a spectral decomposition of the spatial relationships among plots within a sub-region, and can account for complex spatial structures (Borcard et al. 2004). Only PCNMs with positive eigenvalues were used. PCNMs were calculated with function `pcnm` in the R package 'vegan' (Oksanen et al. 2013).

For analyses in each sub-region, we needed to reduce the number of predictors in each predictor matrix so that a full model (including spatial, climatic and soil predictors) could be fitted to the data. To do this, we first ran three independent principal component analyses (PCAs) for climatic, soil, and spatial predictors. Then, we used the approach described by Blanchet et al. (2008) to conduct variable

selection separately on each predictor matrix containing all spatial, climatic or soil principal components (PCs) (R function `ordiR2step` in 'vegan'; Oksanen et al. 2013). The mean numbers of selected PCs per predictor matrix were 11.5, 6.25 and 12 for climate, soil and space, respectively. The selected climatic PCs represented on average 99.2% of the climatic variation across plots in a sub-region, the selected edaphic PCs represented about 88.9% of the soil variation, and the selected spatial PCs represented on average 46% of the variation in spatial variables. A matrix of environmental predictors (matrix **E**) was constructed as the combination of the selected climatic and soil variables (matrices **C** and **So** respectively), while a full matrix of all predictors was constructed combining the selected environmental (**E**) and spatial (**Sp**) variables. These reduced matrices were then used in the redundancy and variation partitioning analyses described next.

### Environmental effects, spatial effects and hierarchical variation partitioning

To understand how much variation in community composition can be explained in each sub-region, we conducted a distance-based redundancy analysis (dbRDA; Legendre and Anderson 1999) where the axes of the PCoA represented the response matrix (community composition) and the selected environmental and spatial PCs together were used as predictors. Then, we repeated similar analyses separately for the environmental and spatial predictor matrices. Finally, we conducted a hierarchical variation partitioning analysis (Cushman and McGarigal 2002) to decompose the compositional variation among plots into fractions accounted for by multiple predictor matrices. In particular, we used variation partitioning to disentangle the effects on species composition of environmental versus spatial predictors, and of climatic versus soil predictors (Fig. 2).

At the first level of the variation partitioning analysis, we decomposed variation in species composition between environmental and spatial predictors, leading to three fractions of explained variation:

- Explained exclusively by environment:  $E | Sp$ .* This is the amount of compositional variation that is correlated with environment, but that is not spatially structured. Given that measuring all relevant environmental predictors is difficult or impossible, this is potentially a minimum measure of the effect of environmentally mediated niche-based processes on community structure.
- Explained exclusively by space:  $Sp | E$ .* This represents the amount of variation that is spatially structured, but that does not correlate with environmental characteristics. Because it is likely that we have missed some important environmental determinants of species composition, this fraction of variation is best interpreted as a liberal overestimate of the importance of dispersal-based processes on community structure.
- Explained simultaneously by environment and space:  $E \cap Sp$ .* This is the amount of variation that correlates with environmental characteristics and that is also spatially structured. Previous analyses show that this fraction of variation is frequently large (Ruokolainen et al. 2007,

Jones et al. 2011, Chain-Guadarrama et al. 2012), but it is difficult to interpret. On the one hand, it can represent the effect of niche-based processes mediated by environmental variables that are spatially structured; while, on the other hand, it can represent dispersal-based spatial processes that are confounded with measured but irrelevant or unimportant environmental gradients.

At the second level of the variation partitioning analysis, we decomposed the variation explained only by environment ( $E | Sp$ ) into the contributions of climatic and soil predictors. Note that at this second level, the effects of space have already been removed (Fig. 2). This leads to another three fractions of variation:

- Explained exclusively by climate:  $C | (So \cup Sp)$*  This is the variation that can be predicted by climatic variables, even after any effects of soils and space have been factored out. This fraction of variation can be considered a minimum measure of the effects of climatically mediated niche-based processes on community structure.
- Explained exclusively by soils:  $So | (C \cup Sp)$ .* This represents the variation that can be predicted by soil variables, even after any effects of climate and space have been factored out. This fraction of variation can represent a minimum measure of the effects of niche-based processes mediated by soil properties.
- Explained simultaneously by climate and soils:  $(C \cap So) | Sp$ .* This fraction represents the amount of variation in community composition predicted simultaneously by either climatic or soil variables, after removing the effects of spatial predictors. Like for c, this fraction of variation can be difficult to interpret as it can correspond to either climate or soils mediated niche-based processes.

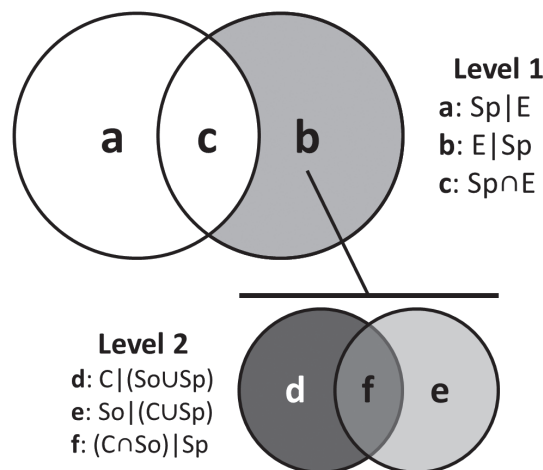


Figure 2. Conceptual representation of the hierarchical variation partitioning approach employed in this study. At the first level, variation is partitioned between environmental and spatial predictors; at the second level, the variation explained exclusively by environment (i.e., after removing the spatial effect) is further partitioned between climatic and soil predictors. E: environment; Sp: space; C: climate; So: soils. Fraction a: variation explained exclusively by environment ( $E | Sp$ ); b: exclusively by space ( $Sp | E$ ); c: simultaneously by environment and space ( $E \cap Sp$ ); d: exclusively by climate ( $C | (So \cup Sp)$ ); e: exclusively by soils ( $So | (C \cup Sp)$ ); f: simultaneously by climate and soils ( $(C \cap So) | Sp$ ).

Finally, to know if a given amount of explained variation in a particular sub-region was significant, we conducted a null model analysis (Tello and Stevens 2010). In this null model, we permuted the rows in the response matrix of community composition and repeated the dbRDA and variation partitioning analyses described above. This procedure provides estimates of the amount of explained variation that are expected when breaking any relationship between species composition and environmental/spatial predictors. The randomization was repeated 5000 times, leading to a frequency distribution of null amounts of variation for each fraction in each sub-region. The empirical fractions of variation were considered statistically significant if they were above the 95th percentile of the distribution of null values.

### Tests of scale dependency: relationships between amount of explained variation and spatial extent

To know if the size of sub-regions had an effect on the relative importance of different environmental or spatial effects, we fitted linear ordinary least-squares regression models to the relationships between observed fractions of variation and spatial extent. To determine whether these relationships were statistically significant, we compared the empirical regression coefficients to coefficients produced by fitting similar regression models using expected fractions of variation obtained from the null model described above. For each fraction of variation, this resulted in 5000 regression models relating null amounts of explained variation with spatial extent. The coefficients (intercept and slope) of these null regression models constitute expectations about the relationship between amounts of variation and spatial extent that should exist if environmental and spatial predictors have no effects on community composition at any spatial scale. We considered a regression coefficient statistically significant if it was below the 2.5th percentile or above the 97.5th percentile of the distribution of null coefficients. Slopes (i.e. coefficients of extent) significantly larger or smaller than expected suggest scale-dependency in how strongly community composition is associated with environmental or spatial predictors.

## Results

On average, about one fourth of the variation in community composition could be predicted by all environmental and spatial predictors combined (total explained variation ~26.1%; Fig. 3). This amount of explained variation was significantly higher than expected in all of the sub-regions analyzed (Fig. 4A), but it varied markedly in magnitude from 13.9% to 43.6% (Fig. 3). These marked differences among sub-regions, however, could not be accounted for by spatial extent ( $R^2_{adj} = 0.001$ ; Table 1, Fig. 4A). A similar pattern was found for the variation explained by the selected environmental predictors, which changed considerably among sub-regions, but had no relationship with spatial extent (Table 1, Fig. 4B). On the other hand, the variation explained by the selected spatial predictors clearly decreased as spatial extent increased (Table 1, Fig. 4C). Importantly, however, environment explained more variation than space on average (Student's  $t = 12.87$ ;

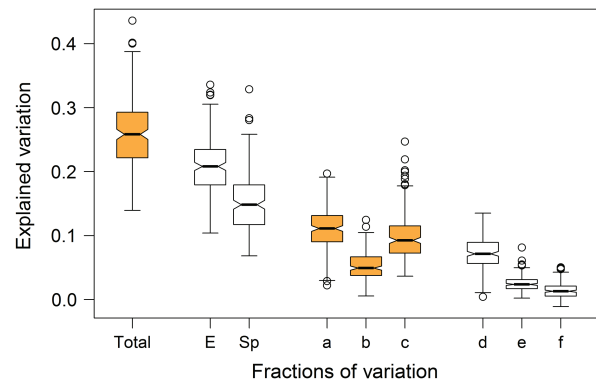


Figure 3. Differences in the amount of variation explained by different predictor matrices or different fractions in the variation partitioning analysis (Fig. 2). Total: variation explained by all environmental and spatial predictors combined; E: explained by environment; Sp: explained by space. Fraction a: variation explained exclusively by environment; b: exclusively by space; c: simultaneously by environment and space; d: exclusively by climate; e: exclusively by soils; f: simultaneously by climate and soils.

$p < 0.001$ ; Fig. 3), and in almost all (94.5%) of the individual sub-regions analyzed.

At the first level of the variation partitioning analysis, the variation explained exclusively by environment was on average ~11% (Fig. 3, 5A), and was significantly higher than expected by the null model in all of the sub-regions. This fraction of variation increased significantly but weakly with spatial extent from ~9% to ~13.5% (Table 1, Fig. 5A). Additionally, the variation explained exclusively by space was on average ~5.3% (Fig. 3, 5B). This fraction of variation was higher than null model expectations in 92.5% of sub-regions, but did not vary significantly with spatial extent (Table 1, Fig. 5B). These two fractions accounted for different amounts of explained variation: the variation explained exclusively by environment was higher than that explained exclusively by space on average (Student's  $t = 19.71$ ;  $p < 0.001$ ; Fig. 3), and in 94.5% of the individual sub-regions analyzed. Finally, the compositional variation simultaneously explained by environment and space was on average ~9.8% (Fig. 3), and was significantly higher than expected by the null model in all of the sub-regions (Fig. 5C). This fraction decreased slightly as spatial extent increased, ranging from ~12.5% at local extents to ~7% at large regional extents (Table 1, Fig. 5C).

At the second level of the variation partitioning analysis, the compositional variation explained exclusively by climate averaged 7.1% (Fig. 3), and was significantly higher than expected in 93% of sub-regions. This fraction of variation increased significantly with spatial extent (Table 1, Fig. 5D). On the other hand, the variation explained exclusively by soils was on average 2.5% (Fig. 3), and was significantly higher than expected in 57.5% of sub-regions. This fraction of variation showed a significant but extremely weak increase with spatial extent (Table 1, Fig. 5E). The variation explained exclusively by climate was higher than that explained exclusively by soils on average (Student's  $t = 23.18$ ;  $p < 0.001$ ; Fig. 3), and in 94.5% of the sub-regions analyzed. Finally, the amount of variation simultaneously explained by climate and soils was on average only 1.4% (Fig. 3). It was

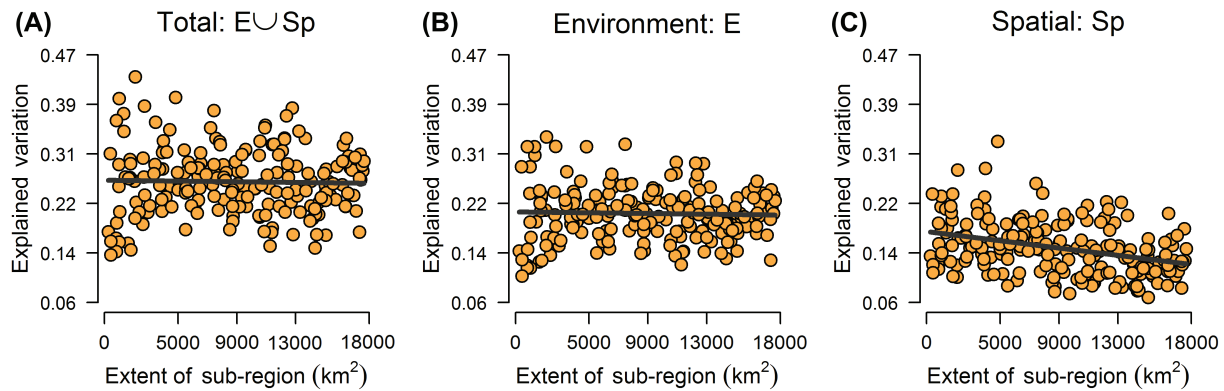


Figure 4. Changes across spatial extents in the total amount of variation in community composition that can be explained, and in the variation explained by all environmental or all spatial predictors. (A) Total variation explained, (B) all variation explained by environment, and (C) all variation explained by space. In each panel, orange circles represent sub-regions where the variation explained was significantly higher than expected by the null model, while white circles represent sub-regions for which it was not (see Methods for details). The black line shows the fit of a linear ordinary least-squares regression to the empirical relationship.

significantly higher than expected by the null model in 44% of the sub-regions, and showed a slight but significant increase with spatial extent (Table 1, Fig. 5F).

## Discussion

### The total amount of compositional variation that can be explained is highly variable but not scale dependent

In our study, size of sub-regions varied by nearly two orders of magnitude (roughly from 255 to 17 700 km<sup>2</sup>; Fig. A2).

Nevertheless, the spatial extent of analysis does not have an effect on the total amount of compositional variation that can be explained by environmental and spatial processes (Fig. 4A). Although no previous study has directly evaluated the effects of spatial extent on the predictability of composition in tropical communities, our results seem to disagree with the scale dependency implied in some previous studies conducted at different spatial extents in other study systems. For example, while Jones et al. (2008a, 2011) found that between 30–40% of the variation could be explained at small extents (~25 km<sup>2</sup>), other researchers found that only between 20–25% could be explained at intermediate extents (~10 000 km<sup>2</sup>; Svenning et al. 2004, Chain-Guadarrama

Table 1. Effects of scale (spatial extent) on the amounts of variation in community composition explained by environmental and spatial predictors (Fig. 2). Observed coefficients of linear regression models are compared against 5000 expected values given a null model where relationships between community composition and predictors have been eliminated. The mean null coefficients were zero in all cases. Intercepts significantly larger than expected by the null model suggest that at small scales community composition is significantly associated with environmental or spatial predictors. Extent coefficients significantly larger or smaller than expected (in bold) suggest scale-dependency on how strongly community composition is associated with environmental or spatial predictors. p-values represent the proportion of null coefficients that are ‘higher than or equal to’ (observed > expected) or ‘lower than or equal to’ (observed < expected) empirical estimates. ∩: intersection; ∪: union; |: ‘after removing the effects of’.

Fraction	Coefficient	R <sup>2</sup> <sub>adj.</sub>	Observed	Expected (95% confidence interval)	Observed > Expected (p)	Observed < Expected (p)
Total explained variation	Intercept		0.264	–0.006 to 0.006	<0.001	1.000
	Extent × 10 <sup>5</sup>	0.001	–0.028	–0.059 to 0.058	0.821	0.179
Environment	Intercept		0.211	–0.004 to 0.004	<0.001	1.000
	Extent × 10 <sup>5</sup>	0.001	–0.029	–0.038 to 0.039	0.929	0.071
Space	Intercept		0.178	–0.003 to 0.003	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.121</b>	– <b>0.304</b>	–0.030 to 0.028	1.000	< <b>0.001</b>
a Environment   Space	Intercept		0.085	–0.005 to 0.005	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.166</b>	<b>0.276</b>	–0.051 to 0.051	<0.001	1.000
b Space   Environment	Intercept		0.053	–0.005 to 0.005	< <b>0.001</b>	1.000
	Extent × 10 <sup>5</sup>	<0.001	0.001	–0.045 to 0.045	0.496	0.504
c Environment ∩ Space	Intercept		0.125	–0.004 to 0.004	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.175</b>	– <b>0.305</b>	–0.041 to 0.041	1.000	< <b>0.001</b>
d Climate   (Soils ∪ Space)	Intercept		0.054	–0.005 to 0.005	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.147</b>	<b>0.191</b>	–0.044 to 0.044	< <b>0.001</b>	1.000
e Soils   (Climate ∪ Space)	Intercept		0.021	–0.004 to 0.004	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.032</b>	<b>0.042</b>	–0.038 to 0.037	<b>0.014</b>	0.986
f (Climate ∩ Soils)   Space	Intercept		0.010	–0.003 to 0.003	<0.001	1.000
	Extent × 10 <sup>5</sup>	<b>0.036</b>	<b>0.044</b>	–0.028 to 0.029	<b>0.002</b>	0.998

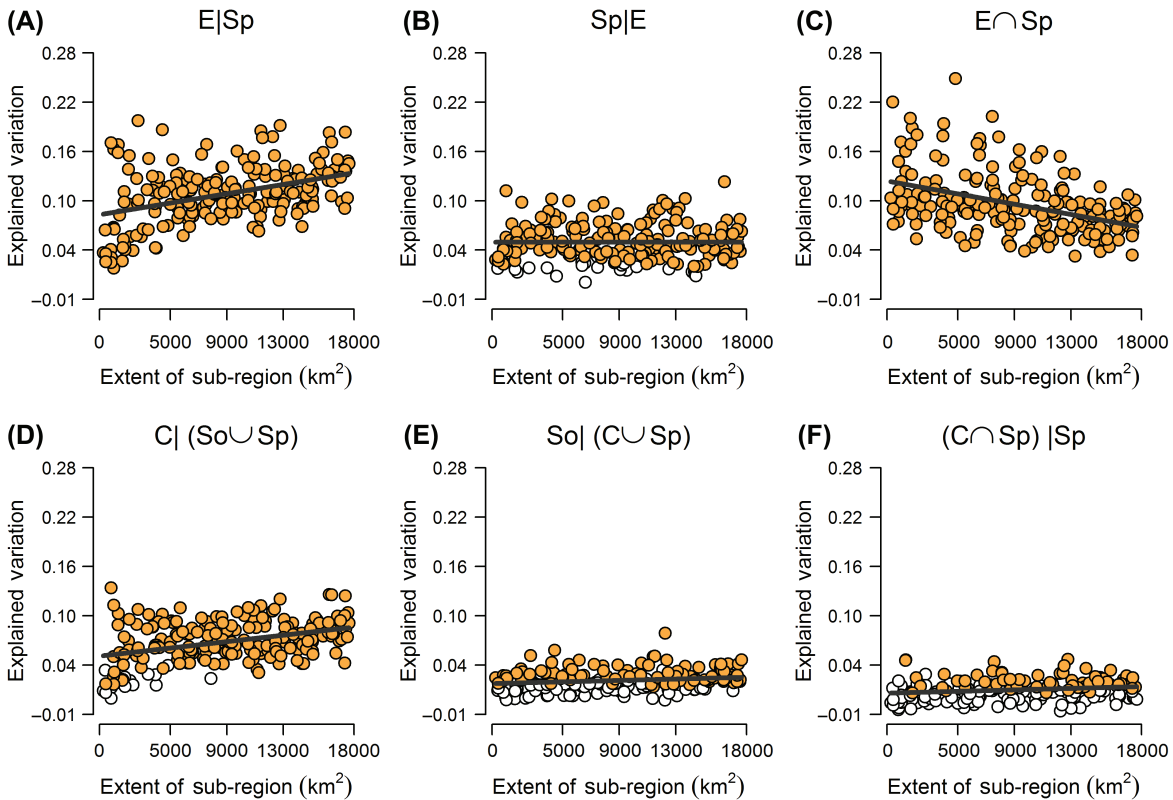


Figure 5. Changes across spatial extent in the fractions of variation in community composition explained by environmental or spatial predictors (Fig. 2). (A) Variation explained exclusively by environment (fraction a), (B) variation explained exclusively by space (fraction b), (C) variation explained simultaneously by environment and space (fraction c), (D) Variation explained exclusively by climate (fraction d), (E) variation explained exclusively by soils (fraction e), and (F) variation explained simultaneously by climate and soils (fraction f). Orange circles: significantly higher than expected by the null model; white circles: not significantly higher than the null model; black line: empirical relationship.

et al. 2012, Myers et al. 2013). Our results challenge the idea that increases in environmental heterogeneity or isolation (associated with increases in extent; Fig. A3) would lead to community composition to be more strongly related to environmental and spatial gradients at large spatial scales (Kristiansen et al. 2012).

Despite the lack of scale dependency, the total amount of explained variation fluctuated substantially from about 14% to 44% (Fig. 3). Additional analyses (not shown) indicated that neither mean elevation nor elevational range nor environmental heterogeneity of the sub-region account for significant amounts of additional variation (<7% of additional variation explained). Together, these results suggest that idiosyncrasies of the study system or the sampling scheme can lead to significant differences in how well community composition correlates with environmental and spatial predictors. These idiosyncrasies can make it challenging to know whether results from a particular study are dependent on details of the analyses, and could help account for some of the differences in results observed across previous studies (Chave 2009). Our study, however, embraces this potential variability by repeating analyses for many subsets of the data at various spatial scales. This approach allows us to identify some consistent patterns that rise above the variability, and which we discuss next.

### Environmental niche-based processes are more important than spatial processes from local to regional scales

The amount of variation explained exclusively by environment is more than double of that explained exclusively by space, and this is so regardless of the extent of analysis (Fig. 3, 4). This seems to be true for the average trend, as well as for the majority of sub-regions. Our results are in agreement with previous studies in tropical forests that have found similar patterns in regions spanning a broad range of spatial extents (Phillips et al. 2003, Macía et al. 2007, Ruokolainen et al. 2007, Jones et al. 2008b, Myers et al. 2013). While the amount of compositional variation explained exclusively by environment increases with spatial extent in our study system, the variation explained solely by spatial predictors remains constant (Fig. 5). This contradicts previous suggestions that the role of space is greater than that of environment at smaller scales (Normand et al. 2006, Laliberté et al. 2009, Hu et al. 2012). Instead, our results support the idea that niche-based processes are more important than spatial dispersal-based processes in structuring tropical plant communities from local to regional scales.

This conclusion relies on the interpretation of fractions of variation in partitioning analyses, an approach that has

been used broadly among previous studies. However, this approach has also been criticized (Smith and Lundholm 2010, Vellend et al. 2014), particularly regarding the appropriate mechanistic interpretation of spatial structure in community composition. Because all environmental gradients driving community structure are difficult or impossible to measure, the environmental effect (fraction a) is likely underestimated, while the spatial effect (fraction b) is likely over-estimated. This can cause erroneous conclusions about whether niche-based or dispersal-based assembly is more important in a particular system. However, our results suggest a pattern that is opposite to the inherent bias in the analyses, showing that environment is more important than space. Another important concern with variation partitioning analysis is that, given the nature of ecological data, environmental and spatial predictors are likely to be correlated, making it difficult to distinguish their effects. Indeed, analyses of empirical data show that the amount of compositional variation explained simultaneously by environmental and spatial predictors is typically large in tropical forests (Ruokolainen et al. 2007, Jones et al. 2011, Chain-Guadarrama et al. 2012). In our analyses, this fraction of variation was typically ~9.8%; larger than the variation explained exclusively by space, but typically smaller than that accounted for exclusively by environment (Fig. 3). On the one hand, this fraction of variation can represent niche-based processes mediated by environmental variables that are spatially structured; while, on the other hand, it can represent dispersal-based spatial processes confounded with spatially structured environmental predictors. It is our opinion that the first interpretation is the most likely, but the exact identity of the mechanisms behind this fraction of variation is unknown. Experimental approaches, although challenging in species-rich tropical regions, might be required to fully disentangle niche- from dispersal-based processes hidden behind this difficult-to-interpret fraction of variation.

Finally, unlike the variation explained exclusively by environment or exclusively by space, the variation explained simultaneously by both predictor matrices decreases significantly with spatial extent (Fig. 5). It seems that the observed lack of scale dependency in the total variation explained by environment (Fig. 4B) results from scale dependency in opposite directions of the variation explained exclusively by environment (Fig. 5A) and the variation confounded between environmental and spatial predictors (Fig. 5C). Thus, while the effects of the niche-based processes that we are able to disentangle seem to increase with increasing spatial scale, the processes hidden behind this fraction of variation clearly decrease in importance as the size of regions increases.

### **Climate is more important than soils, but both contribute to the structure of tropical plant communities**

Our results show that both climatic and soil predictors contribute complementarily to the effect of environment on the structure of plant communities, and this pattern seems to be true from local to regional spatial scales (Fig. 3, 5). These results are in agreement with previous research across long climatic (Bjorholm et al. 2008, Slik et al.

2009, Blach-Overgaard et al. 2010) or elevational gradients (Sesnie et al. 2009), which have shown that climate is often a strong predictor of community composition. Previous studies have also found that soils are important in structuring tropical plant communities at local (John et al. 2007, Jones et al. 2008a, b), intermediate (Costa et al. 2005, Bohlman et al. 2008, Andersen et al. 2010, Damasco et al. 2013), and regional spatial extents (Tuomisto et al. 2003, Vormisto et al. 2004, Macía et al. 2007, Guèze et al. 2013). These studies often report a greater importance of soils than what we found in our analyses, but they typically have not considered the variation that can also be explained by other alternative mechanisms, like climatic gradients or spatial dispersal-based processes. Our results robustly demonstrate that heterogeneity in both climate and soils are complementary mechanisms leading to community structure across a very broad range of spatial extents.

Additionally, our results also show that climate has an effect that is stronger than that of soils, a trend that remains unchanged by spatial extent (Fig. 3, 5). These results indicate that climatic conditions could be a strong determinant of plant species distributions even at small local scales (Jones et al. 2011). Thus, our results contradict Siefert et al. (2012) suggestion that soils are the most important factor controlling the distribution of plant species and community structure at small extents, and that this pattern is reversed as extent increases. One possible explanation for this disagreement in results is that our smallest sub-regions analyzed are around ~250 km<sup>2</sup>, while Siefert et al. study included also much smaller extents. So although we do not find evidence for a change in the relative importance of climatic and soil predictors at around 2000 km<sup>2</sup> as proposed by Siefert et al., it is possible that this change occurs at smaller scales than those we considered. Indeed, soils have been found to be important determinants of community structure in other systems at scales of only a few squared kilometers or even meters (John et al. 2007, Jones et al. 2008a, b).

### **Conclusions**

This study represents a robust evaluation of scale dependency on the relative importance of environmental and spatial processes structuring tropical plant communities. Despite the limitations of variation partitioning analyses, the best available evidence suggests that niche-based environmental processes are more important than spatial processes (potentially dispersal-based), and that this is true from local (~250 km<sup>2</sup>) to regional (~17 000 km<sup>2</sup>) spatial scales. On the other hand, an important amount of variation in community composition is explained simultaneously by environmental and spatial predictors, and this fraction of variation decreases modestly as spatial extent increases. Experimental approaches in plant community ecology are noticeably challenging in species rich tropical regions, but they might be required to fully disentangle niche- from dispersal-based processes hidden behind this difficult-to-interpret fraction of variation. We also found that both climatic and soil variables can account for significant and complementary proportions of the variation in community composition. However, in contrast to the results of a recent meta-analysis (Siefert et al. 2012), climate is a stronger predictor of community structure across



all spatial extents analyzed. Finally, we found that the total amount of variation in community composition that can be accounted for by either environmental or spatial predictors differs considerably across sub-regions. However, these differences in community predictability do not correspond to differences in spatial extent. Our analyses show a high degree of idiosyncrasy in the quantitative results that come from different sub-regions. Despite this variability, however, our approach allows us to identify various important patterns in the structure of communities in this hyper-diverse region of the planet.

*Acknowledgments* – G. Arellano and J. S. Tello contributed equally to the preparation of this manuscript. Author contributions: GA, JST and MJM conceived the hypotheses and designed the manuscript. PMJ and MJM conceived and designed the field data collection. GA, AFF, MIL, VT and MJM conducted fieldwork. AFF coordinated the vouchers determination. GA and JST analyzed the data and wrote the manuscript; other authors provided editorial suggestions. We thank the Bolivian Dirección General de Biodiversidad, the Servicio Nacional de Áreas Protegidas, and the local communities for permits, access, and collaboration during fieldwork. L. E. Cayola, A. Araújo-Murakami, J. M. Quisbert-Quispe, M. Cornejo, T. B. Miranda, R. Seidel, N. Y. Paniagua, C. Maldonado and V. Cala provided much of the data for the present study. I. Jiménez, H. Tuomisto and T. Wiegand have provided helpful comments on our manuscript. Many researchers, students and volunteers helped in the field, in the herbarium and with the taxonomic determination of specimens. This study received financial support from the following institutions: Consejería de Educación (Comunidad de Madrid), National Geographic Society (8047-06, 7754-04), US National Science Foundation (DEB#0101775, DEB#0743457), Missouri Botanical Garden, Universidad Autónoma de Madrid, Banco Santander and the Davidson and Taylor families.

## References

Adler, P. B. et al. 2007. A niche for neutrality. – *Ecol. Lett.* 10: 95–104.

Andersen, K. M. et al. 2010. Soil-based habitat partitioning in understorey palms in lower montane tropical forests. – *J. Biogeogr.* 37: 278–292.

Blanchet, F. et al. 2008. Forward selection of explanatory variables. – *Ecology* 89: 2623–2632.

Bjorholm, S. et al. 2008. To what extent does Tobler's 1st law of geography apply to macroecology? A case study using American palms (Arecaceae). – *BMC Ecol.* 8: 11.

Blach-Overgaard, A. et al. 2010. Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. – *Ecography* 33: 380–391.

Bohlman, S. A. et al. 2008. Importance of soils, topography and geographic distance in structuring central Amazonian tree communities. – *J. Veg. Sci.* 19: 863–874.

Borcard, D. et al. 2004. Dissecting the spatial structure of ecological data at multiple scales. – *Ecology* 85: 1826–1832.

Chain-Guadarrama, A. et al. 2012. Determinants of rain-forest floristic variation on an altitudinal gradient in southern Costa Rica. – *J. Trop. Ecol.* 28: 463–481.

Chao, A. et al. 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. – *Ecol. Lett.* 8: 148–159.

Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and contemporary approaches. – Univ. of Chicago Press.

Chave, J. 2009. Spatial variation in tree species composition across tropical forests: pattern and process. – In: Carson, W. P. and Schnitzer, S. (eds), *Tropical forest community ecology*. Wiley.

Costa, F. R. C. et al. 2005. Mesoscale distribution patterns of Amazonian understorey herbs in relation to topography, soil and watersheds. – *J. Ecol.* 93: 863–878.

Cushman, S. and McGarigal, K. 2002. Hierarchical, multi-scale decomposition of species-environment relationships. – *Landscape Ecol.* 17: 637–646.

Damasco, G. et al. 2013. Disentangling the role of edaphic variability, flooding regime and topography of Amazonian white-sand vegetation. – *J. Veg. Sci.* 24: 384–394.

Dungan, J. L. et al. 2002. A balanced view of scale in spatial statistical analysis. – *Ecography* 25: 626–640.

Fuentes, A. 2005. Una introducción a la vegetación de la región de Madidi. – *Ecol. Bolivia* 40: 1–31.

Gravel, G. et al. 2006. Reconciling niche and neutrality: the continuum hypothesis. – *Ecol. Lett.* 9: 399–409.

Guèze, M. et al. 2013. Determinants of tree species turnover in a southern Amazonian rain forest. – *J. Veg. Sci.* 24: 284–295.

Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – *Int. J. Climatol.* 25: 1965–1978.

Hu, Y. H. et al. 2012. Strong neutral spatial effects shape tree species distributions across life stages at multiple scales. – *PLoS ONE* 7: e38247.

Hubbell, S. 2001. *The unified neutral theory of biodiversity and biogeography*. – Princeton Univ. Press.

John, R. et al. 2007. Soil nutrients influence spatial distributions of tropical tree species. – *Proc. Natl Acad. Sci. USA* 104: 864–869.

Jones, M. M. et al. 2008a. Explaining variation in tropical plant community composition: influence of environmental and spatial data quality. – *Oecologia* 155: 593–604.

Jones, M. M. et al. 2008b. Differences in the degree of environmental control on large and small tropical plants: just a sampling effect? – *J. Ecol.* 96: 367–377.

Jones, M. M. et al. 2011. Microhabitat partitioning promotes plant diversity in a tropical montane forest. – *Global Ecol. Biogeogr.* 20: 558–569.

Kristiansen, T. et al. 2012. Environment versus dispersal in the assembly of western Amazonian palm communities. – *J. Biogeogr.* 39: 1318–1332.

Laliberté, E. et al. 2009. Assessing the scale-specific importance of niches and other spatial processes on beta diversity: a case study from a temperate forest. – *Oecologia* 159: 377–388.

Legendre, P. and Legendre, L. 1998. *Numerical ecology*. – Elsevier.

Legendre, P. and Anderson, M. J. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. – *Ecol. Monogr.* 69: 1–24.

Levin, S. 1992. The problem of pattern and scale in ecology. – *Ecology* 73: 1943–1967.

López-Martínez, J. O. et al. 2013. Partitioning the variation of woody plant  $\beta$ -diversity in a landscape of secondary tropical dry forests across spatial scales. – *J. Veg. Sci.* 24: 33–45.

Macía, M. J. et al. 2007. Congruence between floristic patterns of trees and lianas in a southwest Amazonian rain forest. – *Ecography* 30: 561–577.

Münkemüller, T. et al. 2014. Scale decisions can reverse conclusions on community assembly processes. – *Global Ecol. Biogeogr.* 23: 620–632.

Myers, J. A. et al. 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. – *Ecol. Lett.* 16: 151–157.

Normand, S. et al. 2006. Geographical and environmental controls of palm beta diversity in paleo-riverine terrace forests in Amazonian Peru. – *Plant Ecol.* 186: 161–176.

- Oksanen, J. et al. 2013. *vegan*: community ecology package. – R package ver. 2.0-10, <<http://CRAN.R-project.org/package=vegan>>.
- Pansonato, M. P. et al. 2013. Spatial scale or amplitude of predictors as determinants of the relative importance of environmental factors to plant community structure. – *Biotropica* 45: 299–307.
- Phillips, O. L. et al. 2003. Habitat association among Amazonian tree species: a landscape-scale approach. – *J. Ecol.* 91: 757–775.
- Reeuwijk, L. P. van. 2002. Procedures for soil analysis. – Int. Soil Ref. Inf. Centre, Tech. paper 9.
- Ruokolainen, K. et al. 2007. Are floristic and edaphic patterns in Amazonian rain forests congruent for trees, pteridophytes and Melastomataceae? – *J. Trop. Ecol.* 23: 13–25.
- Sandel, B. 2015. Towards a taxonomy of spatial scale-dependence. – *Ecography* 38: 358–369.
- Sesnie, S. E. et al. 2009. Landscape-scale environmental and floristic variation in Costa Rican old-growth rain forest remnants. – *Biotropica* 41: 16–26.
- Siefert, A. et al. 2012. Scale dependence of vegetation–environment relationships: a meta-analysis of multivariate data. – *J. Veg. Sci.* 23: 942–951.
- Slik, J. W. F. et al. 2009. Environmental correlates for tropical tree diversity and distribution patterns in Borneo. – *Divers. Distrib.* 15: 523–532.
- Smith, T. W. and Lundholm, J. T. 2010. Variation partitioning as a tool to distinguish between niche and neutral processes. – *Ecography* 33: 648–655.
- Svenning, J.-C. et al. 2004. Ecological determinism in plant community structure across a tropical forest landscape. – *Ecology* 85: 2526–2538.
- Tello, J. S. and Stevens, R. D. 2010. Multiple environmental determinants of regional species richness and effects of geographic range size. – *Ecography* 33: 796–808.
- Tello, J. S. et al. 2015. Elevational gradients in  $\beta$ -diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. – *PLoS ONE* 10: e0121458.
- Tuomisto, H. et al. 2003. Floristic patterns along a 43-km long transect in an Amazonian rain forest. – *J. Ecol.* 91: 743–756.
- Tuomisto, H. et al. 2012. Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning results. – *Ecography* 35: 961–971.
- Vellend, M. et al. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. – *Oikos* 123: 1420–1430.
- Vormisto, J. et al. 2004. Palm distribution patterns in Amazonian rainforests: what is the role of topographic variation? – *J. Veg. Sci.* 15: 485–494.

Supplementary material (available online as Appendix oik-02426 at <[www.oikosjournal.org/appendix/oik-02426](http://www.oikosjournal.org/appendix/oik-02426)>). Appendix 1.