Li, Y., Shipley, B., Price, J.N., de L. Dantas, V., Tamme, R., Westoby, M., Siefert, A., Schamp, B.S., Spasojevic, M.J., Jung, V., Laughlin, D.C., Richardson, S.J., Le Bagousse-Pinguet, Y., Schöb, C., Gazol, A., Prentice, H.C., Gross, N., Overton, J., Cianciaruso, M.V., Louault, F., Kamiyama, C., Nakashizuka, T., Hikosaka, K., Sasaki, T., Katabuchi, M., Frenette Dussault, C., Gaucherand, S., Chen, N., **Vandewalle, M.**, Batalha, M.A. (2018):
Habitat filtering determines the functional niche occupancy of plant communities worldwide
J. Ecol. **106** (3), 1001 – 1009

http://dx.doi.org/10.1111/1365-2745.12802

DR. YUANZHI LI (Orcid ID : 0000-0002-5258-877X)
DR. DANIEL C LAUGHLIN (Orcid ID : 0000-0002-9651-5732)
DR. CHRISTIAN SCHÖB (Orcid ID : 0000-0003-4472-2286)
MR. ANTONIO GAZOL (Orcid ID : 0000-0001-5902-9543)

Article type : Standard Paper

`Handling Editor: Peter Vesk`

Running title: Niche occupancy and community assembly

Title: Habitat filtering determines the functional niche occupancy of plant communities worldwide

Yuanzhi Li[1,2], Bill Shipley[1,*], Jodi N. Price[3,4], Vinícius de L. Dantas[5], Riin Tamme[6,7], Mark Westoby[8], Andrew Siefert[9], Brandon S. Schamp[10], Marko J. Spasojevic[11], Vincent Jung[12], Daniel C. Laughlin[13], Sarah J. Richardson[14], Yoann Le Bagousse-Pinguet[15,16], Christian Schöb[17], Antonio Gazol[18], Honor C. Prentice[19], Nicolas Gross[16,20,21], Jake Overton[22], Marcus V. Cianciaruso[23], Frédérique Louault[24], Chiho Kamiyama[25], Tohru Nakashizuka[26], Kouki Hikosaka[26], Takehiro Sasaki[27], Masatoshi Katabuchi[28], Cédric Frenette Dussault[1], Stephanie Gaucherand[29], Ning Chen[2], Marie Vandewalle[30], Marco Antônio Batalha[31]

[1]Département de Biologie, Université de Sherbrooke, Sherbrooke, QC, J1K 2R1, Canada.

[2]State Key Laboratory of Grassland Agro-Ecosystem, School of Life Sciences, Lanzhou University, Lanzhou, Gansu 730000, China.

[3]School of Plant Biology, University of Western Australia, Perth, Australia.

[4]Institute for Land, Water and Society, Charles Sturt University, Albury, NSW, Australia, 2640.

[5]Institute of Geography, Federal University of Uberlândia – UFU, 38400-902, Uberlândia, Brazil.

[6]Institute of Ecology and Earth Sciences, University of Tartu, Lai 40, Tartu, Estonia, 51005.

[7]Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia, 2052.

[8]Department of Biological Sciences, Macquarie University, Sydney NSW 2109 Australia.

[9]Department of Evolution and Ecology, University of California, Davis, Ca 95616.

[10]Department of Biology, Algoma University, 1520 Queen St. East, Sault Ste. Marie, Ontario, Canada P6A 2G4.

[11]Department of Biology, University of California Riverside, Riverside, CA, 92521, USA.

[12]CNRS UMR 6553, ECOBIO, Université de Rennes 1, Rennes, France.

[13]Department of Botany, University of Wyoming, 1000 East University Ave, Laramie, Wyoming, 82071, USA.

[14]Landcare Research, PO Box 69040, Lincoln 7640, New Zealand.

[15]Department of Botany, University of South Bohemia, Na Zlate stoce 1, 370 05 Ceske Budejovice, Czech Republic.

[16]Área de Biodiversidad y Conservación, Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933 Móstoles, Spain.

[17]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

[18]Instituto Pirenaico de Ecología (IPE-CSIC), Av. Montañana 1005, Zaragoza 50059, Spain.

[19]Department of Biology, Lund University, Sölvegatan 37, SE-223 62 Lund, Sweden.

[20]INRA, USC 1339 (Centre d'étude biologique de Chizé –CNRS), F-79360, Villiers en Bois, France.

[21]Centre d'étude biologique de Chizé, UMR 7372 CNRS – Université de La Rochelle, F-79360, Villiers en Bois, France.

[22]Landcare Research, Private Bag 3127, Hamilton, New Zealand.

[23]Departamento de Ecologia, Universidade Federal de Goiás, 74690-900, Goiânia, Goiás, Brazil.

[24]INRA UR874, UREP, Unité de Recherche sur l'Ecosystème Prairial, 5 Chemin de Beaulieu, F-63039 Clermont-Ferrand, France.

[25]Institute for the Advanced Study of Sustainability, United Nations University, 5-53-70 Jingumae, Shibuya, 150-8925, Japan.

[26]Graduate School of Life Sciences, Tohoku University, Aoba, Sendai 980-8578, Japan.

[27]Department of Biology, Faculty of Science, Chiba University, 1-33 Yayoicho, Inage-ku, Chiba 263-8522, Japan.

[28]Department of Biology, University of Florida, Gainesville, Florida, USA 32611.

[29]RSTEA, Unité de Recherche sur les Ecosystèmes Montagnards, BP 76, 38402 St-Martin d'Hères, cedex, France.

[30]Department of Conservation Biology, Helmholtz Centre for Environmental Research - UFZ Permoserstr. 1504318 Leipzig, Germany.

[31]Department of Botany, Federal University of São Carlos, São Carlos, Brazil.

* Correspondence: Bill Shipley, Département de Biologie, Université de Sherbrooke, Sherbrooke, QC, J1K 2R1, Canada.

E-mail: bill.shipley@usherbrooke.ca

Summary

1. How the patterns of niche occupancy vary from species-poor to species-rich communities is a fundamental question in ecology that has a central bearing on the processes that drive patterns of biodiversity. As species richness increases, habitat filtering should constrain the expansion of total niche volume, while limiting similarity should restrict the degree of niche overlap between species. Here, by explicitly incorporating intraspecific trait variability, we investigate the relationship between functional niche occupancy and species richness at the global scale.

2. We assembled 21 datasets worldwide, spanning tropical to temperate biomes and consisting of 313 plant communities representing different growth forms. We quantified three key niche occupancy components (the total functional volume, the functional overlap between species and the average functional volume per species) for each community, related each component to species richness, and compared each component to the null expectations.

3. As species richness increased, communities were more functionally diverse (an increase in total functional volume), and species overlapped more within the community (an increase in functional overlap) but did not more finely divide the functional space (no decline in average functional volume). Null model analyses provided evidence for habitat filtering (smaller total functional volume than expectation), but not for limiting similarity (larger functional overlap and larger average functional volume than expectation) as a process driving the pattern of functional niche occupancy.

4. *Synthesis.* Habitat filtering is a widespread process driving the pattern of functional niche occupancy across plant communities and coexisting species tend to be more functionally similar rather than more functionally specialized. Our results indicate that including intraspecific trait variability will contribute to a better understanding of the processes driving patterns of functional niche occupancy.

Keywords: community assembly, habitat filtering, limiting similarity, niche occupancy, species richness, intraspecific trait variability, determinants of plant community diversity and structure

Introduction

Understanding the processes that drive the assembly of local communities from a regional species pool has been a fundamental goal in ecology for decades (Diamond 1975; Ricklefs & Travis 1980; Keddy 1992; Weiher & Keddy 1999; Hubbell 2001; Kraft *et al.* 2008; Jung *et al.* 2010). Recent advances suggest that variation in the patterns of biodiversity results from multiple assembly processes varying in relative importance (Maire *et al.* 2012; Spasojevic & Suding 2012; Takahashi & Tanaka 2016). For instance, niche-based processes such as habitat filtering (Keddy 1992; Diaz *et al.* 1998) and limiting similarity (MacArthur & Levins 1967) may be the main determinants of biodiversity in some communities, while stochastic processes (Hubbell 2001) may dominate in others. One classical approach for assessing the relative importance of different assembly processes involves mapping all species of a community onto an *n*-dimensional niche space (Hutchinson 1957), quantifying the niche occupancy structure, and investigating how it varies from species-poor to species-rich communities (Hutchinson 1978; Litvak & Hansell 1990). However, this approach has been hindered by the practical limitations of measuring resource niche axes for a large number of species in natural communities.

Trait-based ecology offers an alternative approach for investigating niche occupancy structure in functional space (Mouillot *et al.* 2005; Mason *et al.* 2011). An *n*-dimensional functional space is defined analogously to Hutchinson's (1957) *n*-dimensional niche space, except that the axes represent functional traits rather than resources (Rosenfeld 2002). The vector of *n* trait values possessed by a single individual defines its position in this functional space and the projection of all individuals of a species represents the functional volume occupied by this species. When all species of a community are mapped onto the functional space, the functional niche occupancy structure of the community can be described by three metrics (Fig. 1a): the total functional volume of the community (*T*), the functional overlap between species within the community (*O*) and the average functional volume per species (*A*). There is an underlying relationship (equation 1; see deviation in Appendix S1 in Supporting Information), given here for the first time, between species richness (*S*) and the three niche occupancy metrics:

$$S = \frac{T + O}{A} \qquad \text{(eqn 1)}$$

Equation 1 is a mathematical identity (is true by definition), not an empirical hypothesis that requires testing. The importance of equation 1 is that it not only specifies the logically possible responses of the niche occupancy metrics as species richness increases, but also makes explicit the fact that all three components of niche occupancy (*T*, *O* and *A*) must be quantified in order to completely explain variation in species richness (Fig. 1).

Even though Equation 1 applies to communities driven by any assembly process, different responses of the three metrics to species richness may reveal the relative importance of different community assembly mechanisms (Fig. 2): (a) neutrality, (b) habitat filtering, (c) limiting similarity, and (d) a combination of habitat filtering and limiting similarity. Neutral theory assumes functional equivalence of species (Hubbell 2001), and thus the pattern of functional niche occupancy in observed communities should be similar to that obtained by randomly assigning species from the regional species pool to local communities (Fig. 2a). As species are randomly sampled, some species would overlap with the pre-existing ones while others would fill empty space, resulting in increases T and O, and without changes in A (Fig. 2a). Habitat filtering excludes species with inappropriate trait combinations for given abiotic and biotic conditions, leading to trait convergence (Keddy 1992; Diaz *et al.* 1998). Here, we used a broad definition of habitat filtering, including both abiotic and biotic (e.g. competitive displacement) filters, because these processes often lead to similar patterns of functional niche occupancy (Kunstler *et al.* 2012). Limiting similarity reduces the likelihood of co-occurrence of species that overlap too much in their niche occupancy, leading to trait divergence (MacArthur & Levins 1967; Stubbs & Bastow 2004; Jung *et al.* 2010). If habitat filtering prevails in communities (Fig. 2b), T should be smaller than the null (neutral) expectation, while O and A could be larger or smaller than the null expectations, as they together satisfy Equation 1. If limiting similarity prevails in communities (Fig. 2c), then O should be smaller than the null expectation, while T and A could be larger or smaller than the null expectations, as long as they together satisfy Equation 1. Finally, if habitat filtering and limiting similarity jointly drive the functional niche occupancy (smaller T and O than the null expectations), then species should on average occupy smaller functional volume to satisfy Equation 1 (Fig. 2d).

No study, to our knowledge, has empirically measured all three components of niche occupancy by incorporating intraspecific trait variability in natural communities. Early studies investigated the relationship between niche occupancy in morphological space and species richness for animal communities containing a few species (reviewed by Ricklefs & Miles 1994). Most of these studies found a positive relationship between the morphological volume (similar to *T*) and species richness, and a non-significant relationship between the degree of niche packing and species richness (reviewed in Ricklefs & Miles 1994; but see Ricklefs 2009). Recently, Lamanna *et al.* (2014) and Swenson and Weiser (2014) investigated functional niche occupancy for tree communities and found that the total functional volume (*T*) increased with species richness but was always lower than the null expectation, which was consistent with habitat filtering (Fig. 2b). Swenson and Weiser (2014) further pointed out that the degree of niche packing increased with increasing species richness, but species were less packed in functional space than the null expectation, suggesting limiting similarity. Despite the progress made by these studies (Ricklefs & Miles 1994; Lamanna *et al.* 2014; Swenson & Weiser 2014), no study has included intraspecific trait variability (necessary for the estimation of *O* and *A*) and explicitly measured all three niche occupancy components. Instead, each species was positioned, by its mean trait values, as single point in functional space (Fig.1b; Ricklefs & Miles 1994; Lamanna *et al.* 2014; Swenson & Weiser 2014). The functional volume was only calculated for an entire community (Fig.1b) but not for individual species within a community (Fig. 1a). In this approach, it was not possible to tell whether species overlapped more within the community and/or more finely divided the functional space as species richness increased (Ricklefs & Miles 1994; Swenson & Weiser 2014). Moreover, these vegetation studies (Lamanna *et al.* 2014; Swenson & Weiser 2014) have investigated niche occupancy only for tree communities. Therefore, more comprehensive studies, involving herbaceous and shrub communities, are necessary in order to understand the generality of global mechanisms of community assembly.

In the present study, we assembled 21 datasets from across the world consisting of 313 plant communities for which information on intraspecific trait variability was available. These plant communities spanned tropical to temperate terrestrial biomes and covered a variety of growth forms (e.g. trees, shrubs, herbs). We quantified the functional spaces using pair-wise combinations of nine of the most common traits across all datasets. We then used the hypervolume method (Blonder et al. 2014) to calculate the niche occupancy metrics for all communities in the nine functional spaces. Finally, we related each metric to species richness, and compared them to the null expectation to test hypotheses of community assembly (Fig.2).

Materials and methods

Data collection

To estimate the functional niche occupancy metrics, we assembled datasets from published and unpublished studies on the basis of the following requirements (Table S1): (1) multiple traits were measured on the same plant individual, so that each individual could be placed in a functional space with its trait values as coordinates; (2) traits were measured for several individuals of a species (median of 10 individuals per species, Table S2) to estimate the functional volume occupied by each species; (3) traits were measured for all the dominant species of a community (more than 60% of total species richness was sampled, or more than 80% of total community abundance was represented by the sampled species, Table S2) to quantify the community niche occupancy. Therefore, species richness here refers to the number of species with available trait values included in the analyses (the effective species richness, Lamanna *et al.* 2014). Overall, we collected 21 datasets that met the above criteria (Table S1). Each dataset was located within a given geographical region and contained more than one community within the same region. A community was defined as the sampling unit (e.g. a quadrat or a plot) used within each dataset. However, we combined several neighbouring sampling units (e.g. within the same stratum, same transect or same site) as a combined community when the traits were not measured per sampling unit but at larger scale, or when a dataset included many rare species per sampling unit (Table S2). In total, our analyses were based on 313 communities spanning temperate to tropical biomes (Fig. S1) nested within the 21 datasets. Only the nine most common (out of 32) traits were included in the analyses (vegetative height, VH; specific leaf area, SLA; leaf dry matter content, LDMC; leaf area, LA; leaf carbon concentration, LCC; and leaf nitrogen concentration, LNC; specific stem density, SSD; bark thickness, BT; diameter at ground/breast level, DGH). All trait values were log-transformed to better approximate normality, and then standardized to have zero mean and unit standard deviation (SD) to make the functional volume metrics comparable across analyses (Lamanna *et al.* 2014).

Quantifying the functional niche occupancy metrics within a community

The functional volume of a single species can be estimated from a set of points (individuals of that species within a community are positioned in the functional space on the basis of their trait values) using a non-parametric method based on kernel density estimation (Stine & Heyse 2001). This approach has been applied to quantify niche breadth and niche overlap in one-dimensional space (Mouillot *et al.* 2005; Mason *et al.* 2011), and was extended to multi-dimensional space by Blonder *et al.* (2014). In the present study, each hypervolume (functional volume of a species) was constructed using a quantile threshold of 0.05, 1000 Monte Carlo samples per data point, and a fixed kernel bandwidth of 0.5SD (Blonder *et al.* 2014; Lamanna *et al.* 2014). Details about the parameter settings of the hypervolume method is given in Appendix S2. We also tested another fixed kernel bandwidth (the median intraspecific trait variation): both bandwidths gave similar results; we only report the results using bandwidth of 0.5SD.

After calculating the functional volume of each species within a community, the three niche occupancy components were quantified (Litvak & Hansell 1990). Total functional volume (*T*) was quantified as the union of all individual functional volumes, functional overlap (*O*) was quantified as the sum of the intersections among the functional volumes of individual plants weighted by the level of the intersection (i.e. the number of species occupying the same functional space) and average functional volume per species (*A*) was calculated as the mean of the functional volume of all species (Fig.1a). The mathematical formulas for calculating the niche occupancy metrics are given in Appendix S1. The convex hull method has also been used to quantify functional volumes (Cornwell *et al.* 2006; Swenson & Weiser 2014). However, the convex hull method is relatively more sensitive to outliers than the hypervolume method (see an example in Appendix S2) and is unable to calculate the union of multiple (more than 10) convex hulls (Baselga & Orme 2012), and thus was not applicable for our study. A detailed comparison of different ways of estimating the niche volumes is given in Appendix S3.

A larger number of individuals per species are required to achieve a robust estimate of species' functional volume in high-dimensional functional space (Blonder *et al.* 2014). To better estimate species' functional volume and to allow the inclusion of more datasets, we restricted our analyses to two-dimensional functional spaces (LDMC ~ SLA, LA ~ SLA, SLA ~ VH, LDMC ~ VH, LA ~ LDMC, LA ~ VH, SSD ~ BT, SSD ~ DGH, LCC ~ LNC). The niche occupancy metrics were quantified for each community in each functional space.

Regressions and null model analyses

Linear mixed-effects models, using the "nlme" package in R (Pinheiro *et al.* 2017), were used to investigate the relationships between the niche occupancy metrics and species richness, while taking into account the variation of intercepts and slopes between datasets. In each case, we regressed a niche occupancy metric on species richness allowing for random variation in intercepts and allowing for random variation in slopes if it significantly improved the fit of the model.

We constructed null models to test whether the observed metrics (T, O and A) in each community were significantly different from the null expectations. The null models were built in five steps (see the R codes in Appendix 4). (i) For each community, we first defined its "regional" species pool as consisting of all the species occurring in the dataset to which the community belonged. We did not use a "global" species pool (all species occurring across datasets) because there is no doubt that trait filtering at large biogeographical scales occurs (e.g. vegetative height is lower in grasslands than in forests), and limiting similarity due to biological interactions is only expected to occur between species that can potentially coexist. (ii) Given a community containing *x* species and *t* traits, we randomly selected *x* species from the regional species pool for which the *t* traits were measured (not all species in the regional species pool are available for the *t* traits). (iii) For each selected species, we randomly selected a community from the regional species pool containing that species (most species were present in more than one community) and assigned the *t* traits of the individuals from the selected community to the species (the functional volume of a species in a null community was also estimated by the *t* traits of individuals from one community). (iv) For each null community (the *x* randomly selected species and randomly assigned *t* traits for each selected species), we calculated the functional volume of each selected species and the three niche occupancy metrics. (v) We repeated steps (ii)-(iv) 1000 times to generate a null distribution of communities for each metric. Finally, to determine whether the observed metrics for each community were greater or smaller than the null expectation, we calculated the standardized effect size (SES) as the observed metric minus the mean of the null distribution divided by the standard deviation of the null distribution. A Wilcoxon signed-rank test was used to test whether the SES value of each metric was significantly different from zero. A positive SES value indicates that the observed metric is larger than the null expectation while a negative SES value indicates that the observed metric is smaller than the null expectation. All analyses were conducted in R (R Core Team 2014).

Results

In the nine two-dimensional functional spaces, both the total functional volume (*T)* and the amount of overlap between coexisting species (*O)* increased with increasing species richness *(S)*, while the average functional volume per species (*A)* did not vary significantly or weakly increased with *S* (Table 1 and Fig. 3). The increasing rates (slopes) of *T*~*S* and *O*~*S* varied significantly between datasets in most cases (Table 1). *T* was generally smaller than the null expectation across different communities, although not significantly so in three out of nine functional spaces (Fig. 4a). Both *O* (Fig. 4b) and *A* (Fig. 4c) were larger than the null expectations, but were not significant in the two functional spaces quantified by stem traits (BT~SSD and DGH~SSD, Fig.4b and 4c).

Discussion

Our study addresses a longstanding and fundamental question in ecology: how the pattern of functional niche occupancy varies from species-poor to species-rich communities (Hutchinson 1978). Overall, we found that the total functional volume expanded and the functional overlap increased with increasing species richness, while the average functional volume did not change significantly. Variation in intercepts and slopes between datasets is likely to reflect differences in the community size and the trait-sampling effort between datasets, as well as differences in other variables such as vegetation type. However, these different sources of variation cannot actually be separated here and need to be investigated in future studies that are based on a more systematic sampling design.

The increase in total functional volume implied that communities were more functionally diverse in species-rich communities than in species-poor communities. However, the expansion of the total functional volume associated with increasing species richness was constrained and thus lower than the null expectation across communities (Fig. 4a), suggesting habitat filtering as found in previous studies (Lamanna *et al.* 2014; Swenson & Weiser 2014). In other words, although larger functional volumes were occupied as species richness increased, species were still tightly packed in functional space relative to the neutral expectation. Curiously, regressions between the standardized effect size (*SES*) of total functional volume and species richness indicated that, species were more tightly packed (habitat filtering was stronger) in species-rich communities than species-poor communities ($SES\_T < 0$ and decreased with S, Fig. 4 and Table S3). In our study, habitat filtering was not attributed to large-scale biogeographic factors such as climate differences, because the species pool used in our null model only included species that occurred within the same geographical region. To know what and how environmental factors (e.g. soil fertility, water availability or disturbance regimes) drive habitat filtering and the pattern of functional niche occupancy (Le Bagousse-Pinguet *et al.* 2017), requires future studies by investigating how the pattern of functional niche occupancy varied along the environmental gradients.

Because intraspecific trait variability was included in our study, we were able to address a previously unanswered question: do species overlap more within the community and/or more finely divide the functional space as species richness increases? We found an increase in functional overlap but no decline in average functional volume with increasing species richness, indicating that species overlapped more rather than more finely divided the functional space. Moreover, the greater functional overlap (Fig. 4b) and the larger average functional volume (Fig. 4c) than the null expectation (Fig. 4b) both suggest that limiting similarity (MacArthur & Levins 1967) is not a fundamental processes regulating the pattern of functional niche occupancy at the considered spatial scales used here. Instead, our results suggest that habitat filtering alone determines the functional niche occupancy of the studied plant communities worldwide (Fig.2b).

Our results are based on the largest and most representative collection of available datasets to date, but there are some limitations that should be addressed in future studies as more extensive datasets become available. First, some habitats (e.g. tundra, desert, and boreal sites) were not included or were underrepresented in our datasets. Second, although our analysis included traits that covered key plant-strategy axes (Wright *et al.* 2004; Díaz *et al.* 2016), several types of traits (e.g. secondary compounds, root, phenological and seed traits) reflecting other potentially important functional axes (Ricklefs & Marquis 2012) on which species might be divergent, were not considered. In addition, analyses were restricted to two-dimensional functional spaces (pairwise traits combination) because estimating species' functional volume in high-dimensional space requires larger sample size per species (Blonder *et al.* 2014). Third, trait sampling efforts (e.g. the percentage of species and number of individuals per species sampled, TableS2) were not consistent across studies. Hence the influence of rare species (Umana *et al.* 2015) and the influence of trait sample size (Appendix S2) on functional niche occupancy could not be determined in our study. Finally, the spatial scale of communities varied among datasets and, in some cases, might not be fine enough to detect neighbourhood scale patterns of niche differentiation resulting from limiting similarity (de Bello *et al.* 2013), especially for the combined communities.

A perspective on functional niche occupancy that incorporates intraspecific trait variability provides new insights into community assembly and extends the generality of previous findings to the global scale (Lamanna *et al.* 2014; Swenson & Weiser 2014). We propose that, after habitat filtering has excluded poorly adapted species from a plant community, the remaining species coexist because they are more functionally similar, rather than because they are more functionally specialized.

Author's Contributions

YL designed the project, analysed the data and wrote the manuscript with input from BS. All other authors collected the data and contributed to the revisions of the manuscript.

Acknowledgements

We thank Benjamin Blonder for the help in further developing his original method for n-dimensional hypervolume quantification. This research was funded by an NSERC Discovery grant to BS and by a CSC (China Scholarship Council) scholarship to YL. Data collection of "Hezuo" was funded by National Natural Science Foundation of China (No. 31270472). Data “Panama” was provided by Julie Messier. Data collection of "Saaremaa" was funded by the European Union through the European Social Fund (MOBILITAS post-doctoral grant MJD47).

VLD was supported by São Paulo Research foundation (processes: 2010/01835-0, 2013/50169-1 and 2014/06453-0). Data collection of "New Zealand" was supported by Australian Research Council. AS was supported by the National Science Foundation (DEB-03089). DCL was funded by a grant (UOW1201) from the Royal Society of New Zealand Marsden Fund. YLBP was supported was supported by the EU Education for Competitiveness Operational Programme (reg.no. CZ.1.07/2.3.00/30.0006), by the European Social Fund and Czech State Budget, and the Marie Sklodowska-Curie Actions Individual Fellowship (MSCA-IF) within the European Program Horizon 2020 (DRYFUN Project 656035). NG was support by the AgreenSkills+ fellowship programme, which has received funding from the EU's Seventh Framework Programme under grant agreement N° FP7-609398 (AgreenSkills+ contract). CS was supported by the Swiss National Science Foundation (PZ00P3_148261). Data collection of "Brazil" was funded by the research network GENPAC (Geographical Genetics and Regional Planning for natural resources in Brazilian Cerrado) supported by CNPq/MCT/CAPES Brazil (563621/2010-9) and PELD/CNPq (558187/2009-9, 403833/2012-4) and FAPEG (2012102677001109). MVC was supported by grant from CNPq (#306843/2012-9). MAB was supported by the São Paulo Research Foundation (grant 2008/57502-0) and the Brazilian National Council for Scientific and Technological Development (grant 470653/2010-8).

Data accessibility

Data deposited in the Dryad Digital Repository: doi:10.5061/dryad.cn642

References:

Baselga, A. & Orme, C.D.L. (2012) betapart: an R package for the study of beta diversity. *Methods in Ecology and Evolution,* **3,** 808-812.

Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hypervolume. *Global Ecology and Biogeography,* **23,** 595-609.

Cornwell, W.K., Schwilk, D.W. & Ackerly, D.D. (2006) A trait-based test for habitat filtering: convex hull volume. *Ecology,* **87,** 1465-1471.

de Bello, F., Vandewalle, M., Reitalu, T., Lepš, J., Prentice, H.C., Lavorel, S. & Sykes, M.T. (2013) Evidence for scale- and disturbance-dependent trait assembly patterns in dry semi-natural grasslands. *Journal of Ecology,* **101,** 1237-1244.

Diamond, J.M. (1975) Assembly of Species Communities. *Ecology and evolution of communities*. Harvard University Press.

Diaz, S., Cabido, M. & Casanoves, F. (1998) Plant functional traits and environmental filters at a regional scale. *Journal of Vegetation Science,* **9,** 113-122.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Joseph Wright, S., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D. & Gorné, L.D. (2016) The global spectrum of plant form and function. *Nature,* **529,** 167-171.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.

Hutchinson, G.E. (1957) Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology,* **22,** 415-427.

Hutchinson, G.E. (1978) *An Introduction to Population Ecology*. Yale University Press.

Jung, V., Violle, C., Mondy, C., Hoffmann, L. & Muller, S. (2010) Intraspecific variability and trait-based community assembly. *Journal of Ecology,* **98,** 1134-1140.

Keddy, P.A. (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science,* **3,** 157-164.

Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. *Science,* **322,** 580.

Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmermann, N.E., Kattge, J. & Coomes, D.A. (2012) Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters,* **15,** 831-840.

Lamanna, C., Blonder, B., Violle, C., Kraft, N.J.B., Sandel, B., Šímová, I., Donoghue, J.C., Svenning, J.-C., McGill, B.J., Boyle, B., Buzzard, V., Dolins, S., Jørgensen, P.M., Marcuse-Kubitza, A., Morueta-Holme, N., Peet, R.K., Piel, W.H., Regetz, J., Schildhauer, M., Spencer, N., Thiers, B., Wiser, S.K. & Enquist, B.J. (2014) Functional trait space and the latitudinal diversity gradient. *Proceedings of the National Academy of Sciences,* **111,** 13745-13750.

Le Bagousse-Pinguet, Y., Gross, N., Maestre, F.T., Maire, V., de Bello, F., Fonseca, C.R., Kattge, J., Valencia, E., Leps, J. & Liancourt, P. (2017) Testing the environmental filtering concept in global drylands. *Journal of Ecology***,** n/a-n/a.

Litvak, M.K. & Hansell, R.I. (1990) A community perspective on the multidimensional niche. *The Journal of Animal Ecology***,** 931-940.

MacArthur, R. & Levins, R. (1967) The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist,* **101,** 377-385.

Maire, V., Gross, N., Börger, L., Proulx, R., Wirth, C., Pontes, L.d.S., Soussana, J.-F. & Louault, F. (2012) Habitat filtering and niche differentiation jointly explain species relative abundance within grassland communities along fertility and disturbance gradients. *New Phytologist,* **196,** 497-509.

Mason, N.W., de Bello, F., Doležal, J. & Lepš, J. (2011) Niche overlap reveals the effects of competition, disturbance and contrasting assembly processes in experimental grassland communities. *Journal of Ecology,* **99,** 788-796.

Mouillot, D., Stubbs, W., Faure, M., Dumay, O., Tomasini, J.A., Wilson, J.B. & Do Chi, T. (2005) Niche overlap estimates based on quantitative functional traits: a new family of non-parametric indices. *Oecologia,* **145,** 345-353.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R.C. (2017) Linear and Nonlinear Mixed

Effects Models.

Ricklefs, R.E. (2009) Aspect diversity in moths revisited. *The American Naturalist,* **173,** 411-416.

Ricklefs, R.E. & Marquis, R.J. (2012) Species richness and niche space for temperate and tropical folivores. *Oecologia,* **168,** 213-220.

Ricklefs, R.E. & Miles, D.B. (1994) Ecological and evolutionary inferences from morphology: an ecological perspective. *Ecological morphology: integrative organismal biology* (eds P.C. Wainwright & S.M. Reilly). University of Chicago Press.

Ricklefs, R.E. & Travis, J. (1980) A morphological approach to the study of avian community organization. *The Auk,* **97,** 321-338.

Rosenfeld, J.S. (2002) Functional redundancy in ecology and conservation. *Oikos,* **98,** 156-162.

Spasojevic, M.J. & Suding, K.N. (2012) Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *Journal of Ecology,* **100,** 652-661.

Stine, R.A. & Heyse, J.F. (2001) Non-parametric estimates of overlap. *Statistics in Medicine,* **20,** 215-236.

Stubbs, W.J. & Bastow, W.J. (2004) Evidence for limiting similarity in a sand dune community. *Journal of Ecology,* **92,** 557-567.

Swenson, N.G. & Weiser, M.D. (2014) On the packing and filling of functional space in eastern North American tree assemblages. *Ecography,* **37,** 1056-1062.

Takahashi, K. & Tanaka, S. (2016) Relative importance of habitat filtering and limiting similarity on species assemblages of alpine and subalpine plant communities. *Journal of Plant Research*, **129,** 1-9.

Team, R.C. (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Umana, M.N., Zhang, C., Cao, M., Lin, L. & Swenson, N.G. (2015) Commonness, rarity, and intraspecific variation in traits and performance in tropical tree seedlings. *Ecology Letters*, **18,** 1329-1337.

Weiher, E. & Keddy, P.A. (1999) *Ecological assembly rules: perspectives, advances, retreats*. Cambridge University Press, Cambridge.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature*, **428,** 821-827.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1 Mathematical formulae for niche occupancy metrics and derivation of equation 1.

Appendix S2 Parameter settings of the hypervolume method.

Appendix S3 Comparison of different ways of estimating a hypervolume.

Appendix S4 R codes for the metrics calculations and null model analyses

Figure S1 Geographic distribution of the 313 communities.

Table S1 List of datasets collected.

Table S2 Detailed information of (combined) communities and trait sampling efforts within each dataset.

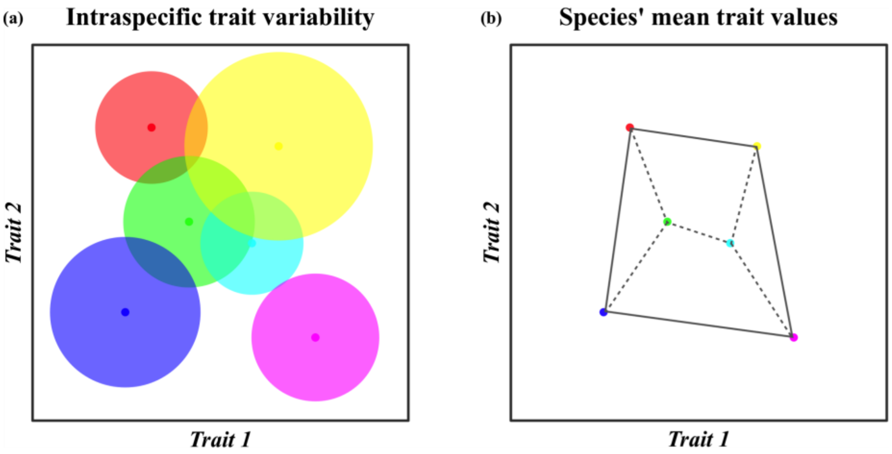
Table S3 Linear mixed-effects models for relationships between standard effect size of niche occupancy metrics and species richness.

Table 1. Linear mixed-effects models for relationships between niche occupancy metrics (*T*, *O* and *A*) and species richness (*S*). In each model, random variation in intercept between datasets was allowed and random variation in slope was included if it significantly improved the model ($p<0.05$). Numbers in brackets are the standard deviations of slopes between datasets if applicable. Pearson correlation coefficients were measured between the pair of traits used for the axes of each functional space. Columns "Datasets" and "Communities" give the number of datasets and communities available for each trait combination, respectively. *T* for total functional volume, *O* for functional overlap, *A* for average functional volume. Traits abbreviations: vegetative height (VH), specific leaf area (SLA), leaf dry matter content (LDMC), leaf area (LA), leaf carbon concentration (LCC), leaf nitrogen concentration (LNC), specific stem density (SSD), bark thickness (BT), diameter at ground/breast level (DGH).

Functional spaces	Correlations	Datasets	Communities	*T*		*O*		*A*	
				Slope	*P*-value	Slope	*P*-value	Slope	*P*-value
SLA~LDMC	-0.547	17	239	0.341 (0.262)	<0.001	2.331 (0.948)	<0.001	0.012	0.213
LA~SLA	0.033	14	198	0.263 (0.078)	<0.001	2.208 (0.656)	<0.001	0.016	0.01
VH~SLA	-0.277	14	160	0.228 (0.088)	<0.001	2.155 (0.494)	<0.001	0.004	0.06
VH~LDMC	0.51	12	157	0.265 (0.121)	<0.001	2.257 (0.672)	<0.001	0.006	0.032
LA~LDMC	0.389	10	130	0.342 (0.152)	<0.001	2.326 (0.822)	<0.001	0.016	0.093
VH~LA	0.752	11	127	0.105	<0.001	1.729 (0.284)	<0.001	0.002	0.21

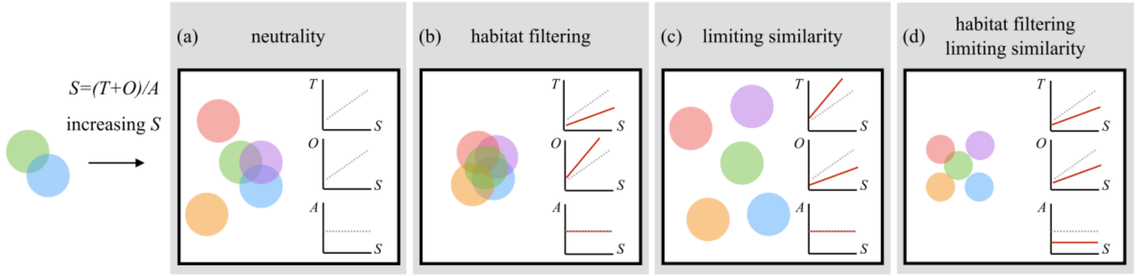
SSD~BT	-0.019	3	44	0.232	<0.001	3.056	<0.001	0.015	<0.001
SSD~DGH	0.024	3	44	0.273 (0.212)	0.04	3.152	<0.001	0.012	<0.001
LCC~LNC	-0.683	4	28	0.083	<0.001	1.47	<0.001	0.001	0.745

Figure 1



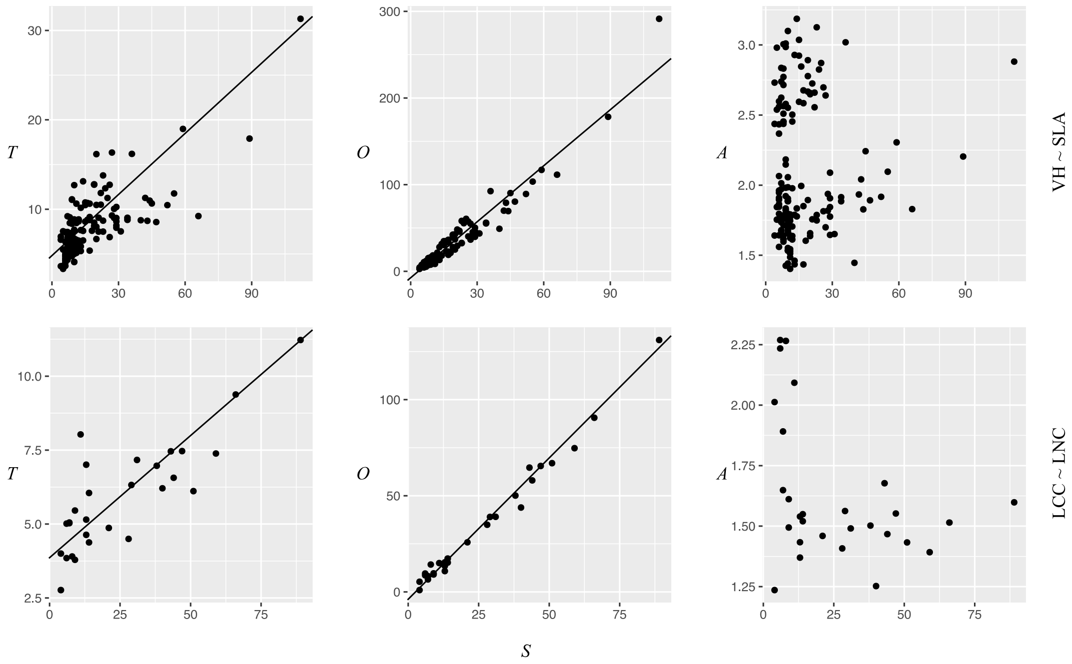
Functional niche occupancy with (a) and without (b) intraspecific trait variability. In case (a), each species occupies a specific functional volume (a coloured circle), then functional niche occupancy can be described by: the total functional volume (*T*, union of the circles), the functional overlap (*O*, sum of intersections among the circles) and the average functional volume per species (*A*, the average volumes of circles). In case (b), each species is represented by a single point based on the specie mean trait values. Total functional volume (e.g. convex hull of all species' points as solid lines) and degree of species packing (mean nearest neighbour distance between species' points as dotted lines) were usually used to describe functional niche occupancy (Ricklefs & Miles 1994; Swenson & Weiser 2014).

Figure 2



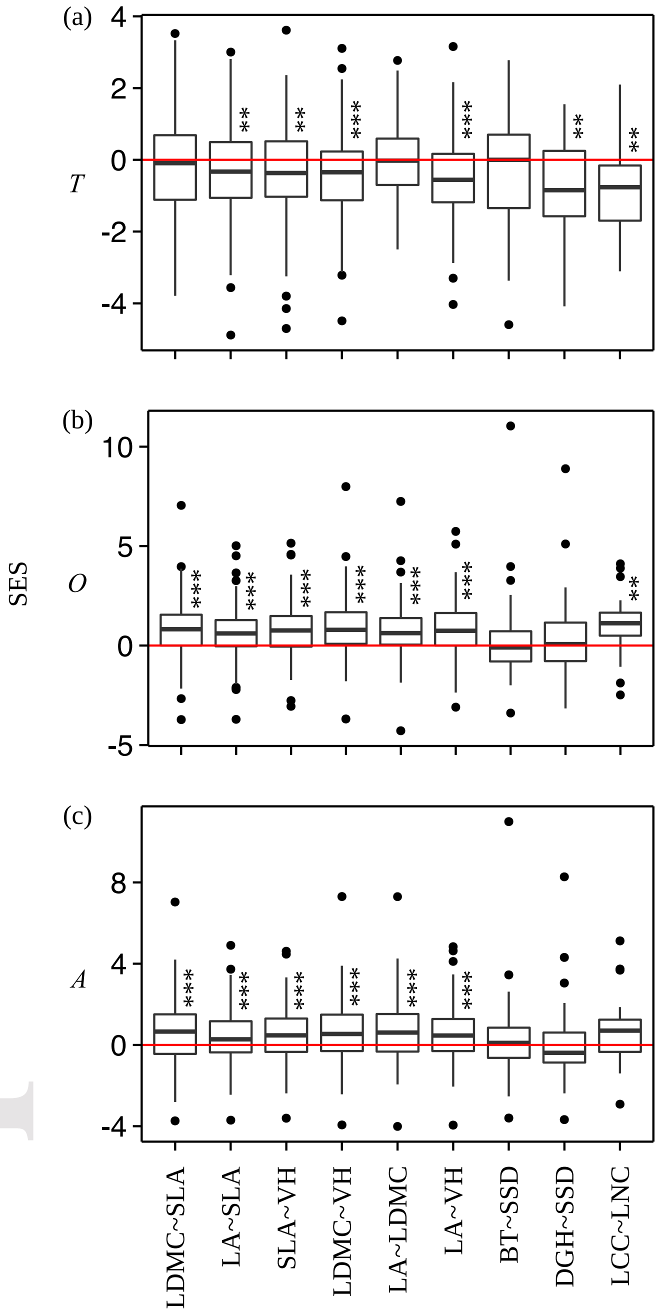
Predicted patterns of functional niche occupancy under different processes of community assembly: (a) pure neutral process, (b) only habitat filtering, (c) only limiting similarity, (d) both habitat filtering and limiting similarity. Different coloured circles represent the functional volumes occupied by different species within a community. Grey dashed and red solid lines represent relationships between niche occupancy metrics (*T*, *O* and *A*) and species richness (*S*) for null and observed communities, respectively. *T* for total functional volume, *O* for functional overlap, *A* for average functional volume.

Figure 3



Scatter plots showing the relationships between niche occupancy metrics (*T*, *O* and *A*) and species richness (*S*) in two representative functional spaces (VH~SLA and LCC~LNC). The regression lines are fitted with the average slope and intercept between datasets (Table 1). *T* for total functional volume, *O* for functional overlap, *A* for average functional volume. Trait abbreviations: vegetative height (VH); specific leaf area (SLA), leaf carbon concentration (LCC), and leaf nitrogen concentration (LNC).

Figure 4



Standard effect sizes (SES) of total functional volume of an entire community (*T*), the functional overlap between species (*O*), and the average functional volume per species (*A*) in the nine functional spaces. Abbreviations: vegetative height, VH; specific leaf area, SLA; leaf dry matter content, LDMC; leaf area, LA; leaf carbon concentration, LCC; and leaf nitrogen concentration, LNC; specific stem density, SSD; bark thickness, BT; diameter at ground/breast level, DGH. The red line shows the null expectation and the symbols indicate the significance level of Wilcoxon signed-rank tests (*** for $p<0.001$, ** for $p<0.01$, * for $p<0.05$).