Positive associations among rare species and their persistence in ecological assemblages

Joaquín Calatayud[1,2]*, Enrique Andivia[3,4], Adrián Escudero[5], Carlos J. Melián[6], Rubén Bernardo-Madrid[7], Markus Stoffel[8,9,10], Cristina Aponte[11], Nagore G. Medina[12,13], Rafael Molina-Venegas[14], Xavier Arnan[15], Martin Rosvall[1], Magnus Neuman[1], Jorge Ari Noriega[2], Fernanda Alves-Martins[2], Isabel Draper[13], Arantzazu Luzuriaga[5], Juan Antonio Ballesteros-Cánovas[8,10], César Morales-Molino[16,17], Pablo Ferrandis[18], Asier Herrero[3,19], Luciano Pataro[13], Leandro Juen[20], Alex Cea[21], Jaime Madrigal-González[10].

[1] Integrated Science Lab, Department of Physics, Umeå University, 901 87 Umeå, Sweden.

[2]Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), C/ José Gutiérrez Abascal 2, 28006, Madrid, Spain.

[3]Departamento de Ciencias de la Vida, Edificio de Ciencias, Universidad de Alcalá, ctra. Madrid-Barcelona km 33.6, 28805 Alcalá de Henares, Madrid, Spain.

[4]Departamento de Biodiversidad, Ecología y Evolución, Universidad Complutense de Madrid, C/ José Antonio Novais 12, 28040, Madrid, Spain.

[5]Departamento de Biología, Geología, Física y Química inorgánica. ESCET., Universidad Rey Juan Carlos, C/ Tulipán s/n, Móstoles, C.P. 28933, Madrid, Spain.

[6]Department of Fish Ecology and Evolution, Eawag, Seestrasse 79, 6047 Kastanienbaum, Switzerland.

[7]Departamento de Biología de la Conservación, Estación Biológica de Doñana CSIC, C/ Americo Vespucio, 26, 41092, Isla de la Cartuja, Sevilla, Spain.

[8]Dendrolab, Department of Earth Sciences, University of Geneva, rue des Maraîchers 13, CH-1205 Geneva, Switzerland.

[9]Department F.-A. Forel for Environmental and Aquatic Sciences, University of Geneva, Boulevard Carl-Vogt 66, 1205, Geneva, Switzerland.

[10]Climate Change Impacts and Risks in the Anthropocene (C-CIA), Institute for Environmental Sciences (ISE), University of Geneva, 66 Boulevard Carl Vogt, 1205, Geneva, Switzerland.

[11]Plant Science, The University of Melbourne, Burnley, 500 Yarra Boulevard, Richmond, Victoria, 3121, Australia.

[12]Department of Botany, Faculty of Sciences, University of South Bohemia, České Budějovice, Czech Republic.

[13]Departamento de Biología (Botánica), Facultad de Ciencias,

Universidad Autónoma de Madrid, calle Darwin 2, Madrid, Spain.

[14]Plant Ecology, Institute of Plant Sciences (IPS), University of Bern, Altenbergrain 21, 3013 Bern, Switzerland.

[15]CREAF (Centre de Recerca Ecològica i Aplicacions Forestals) Campus de Bellaterra (UAB), Edifici C, 08193, Cerdanyola del Vallès Catalunya, Spain.

[16]UMR CNRS 5805 EPOC – OASU, Université de Bordeaux Site de Talence - Bâtiment B18N Allée Geoffroy Saint-Hilaire, CS 50023, 33615 Pessac Cedex, France.

[17]Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern, Altenbergrain 21, 3013, Bern, Switzerland.

[18]Higher Technical School of Agricultural and Forestry Engineering and Botanical Institute, University of Castilla-La Mancha, 02071 Albacete, Spain.

[19]Department of Plant Biology and Ecology, Pharmacy Faculty, University of Basque Country, 01006 Vitoria-Gasteiz, Basque Country.

[20]Laboratótio do Ecologia e Zoologia de Invertebrados, Instituto de Ciências Biológicas, Universidade Federal do Pará, Rua Augusto Correia, 1, Barrio Guama, CEP 66075-110, Belém, Pará, Brazil.

[21]Departamento de Biología, Facultad de Ciencias, Universidad de La Serena, Avenida Raúl Bitrán Nachary, 1305, Casilla 554, La Serena, Chile.

* Correspondence to: j.calatayud.ortega@gmail.com

According to the competitive exclusion principle, species with low competitive abilities should be excluded by more efficient competitors, and yet they generally remain as rare species. Here, we describe the positive and negative spatial association networks of 326 disparate assemblages, showing a general organization pattern that simultaneously supports the primacy of competition and the persistence of rare species. Abundant species monopolize negative associations in about 90% of the assemblages. Contrarily, rare species are mostly involved in positive associations, forming small network modules. Simulations suggest that positive interactions among rare species and microhabitat preferences are the most likely mechanisms underpinning this pattern and rare species persistence. The

consistent results across taxa and geography suggest a general explanation for the maintenance of biodiversity in competitive environments.

Rare species, in terms of low abundance, are the main component of the diversity of ecological assemblages[1]. However, despite decades of intense investigation, general mechanisms behind the persistence of these species remain unclear. In theory, the widely assumed effects of competition between pairs of species should preclude the persistence of weak competitors and the high diversity observed in natural assemblages[2,3]. Explanations for this diversity paradox include the differential roles of niche partition, intraspecific competition, facilitation, indirect and neutral interactions[4-9], among others. Yet, there is no consensus to explain rare species persistence across taxa and environmental conditions so far.

The spatial arrangement of individuals plays a crucial role for unveiling the mechanisms underpinning species assembly and coexistence[10-21]. Because individuals within assemblages are not homogeneously distributed, their spatial organization may both reflect important assembly processes[10,11] and induce species coexistence *per se*[12]. For example, the patchy distribution of a dominant species might prevent the monopolization of resources and allow the existence of its subordinate species[12,13]. Hence, considering spatial aspects of coexistence appears to be an important step in elucidating assembly mechanisms[12]. The spatial sorting of species can be the outcome of divergent habitat preferences, dispersal abilities, and biotic interactions, although the role of interactions is thought to

prevail under rather homogeneous environmental conditions, and especially at very fine spatial scales[11,14,15]. The organization of species within assemblages can be translated into association networks of species that are spatially aggregated (positive networks) or segregated (negative networks). Association networks of disparate biological assemblages can provide valuable empirical evidence of the main forces driving the assembly of species[16-20], helping to reveal general mechanisms underlying species coexistence.

Here we describe a general pattern of positive and negative species associations that is consistent with the competitive exclusion paradigm but, at the same time, can explain the persistence of rare species in natural assemblages. We base our results on a dataset of 326 assemblages that meet the following criteria: (i) each assemblage comprises taxa from only one trophic guild, thereby excluding the possibility that species associations result from direct predation or parasitism; (ii) each assemblage shows reduced spatial extent and low environmental heterogeneity, to increase the likelihood that species associations are mainly due to biotic interactions; (iii) the abundance of at least ten species is recorded in a minimum of ten samples, to improve statistical power (Appendix S1); (iv) the dataset represents a wide variety of biomes (*e.g.*, tropical forests, deserts, temperate steppes and polar climates), thus avoiding biome-specific results; and (v) it encompasses a diversity of taxa (such as bryophytes, vascular plants, and insects among others), to ensure the generalization of our results across taxonomic and

functional groups. We generated positive and negative association networks for each assemblage by comparing the observed spatial association patterns among species to a null model[21]. Species pairs that significantly deviate from random expectations receive positive or negative links in their respective association networks (Fig. 1).

We first analysed whether the structure of positive and negative association networks can reflect predictions from the competitive exclusion principle. Given that competition is heavily emphasized in the literature[2], one would expect species to be more segregated than aggregated in natural systems. If so, negative networks should be more densely connected (*i.e.*, more links per species) than their positive counterparts. In accordance, negative networks were significantly more connected than their positive pairs in a notable 93.2% of all assemblages ($t = 17.006$, $P < 0.001$, Fig. 2a). Differences in connectivity remained similar after accounting for differences in network size ($t = -16.815$, $P < 0.001$, for 78.8% of the assemblages) or when calculating differences in the average number of links (i.e. average species degree; t = -14.689, $P < 0.001$, for 69.0% of the assemblages). Furthermore, if abundance is considered to be an expression of the species' competitive abilities[22], the number of segregations should be monopolized by the most abundant species. Accordingly, results indicated a strong positive correlation between abundance and species degree in negative networks (mean Spearman's ρ = 0.65, SD = 0.23), but a weak or even negative correlation in positive networks (mean ρ = 0.02, SD = 0.38), with

differences between networks being statistically significant (t= -23.881, $P < 0.001$, Fig. 2b). Moreover, we found evidence showing that a particular species is more often involved in negative associations when it becomes abundant (Appendix S2). Both the greater density of links and the relationship between species degree and abundance in negative networks support current knowledge about the prevailing role of competitive interactions in sustaining the dominance of abundant species.

Yet, if the competitive exclusion principle is supported across several assemblages, how can rare species persist? To search for potential mechanisms answering this question we looked at the role played by rare species in association networks. Curiously, we found that rare species are mostly involved in positive associations in 91.7% of the assemblages studied, where positive networks showed a higher incidence of less abundant species than their negative pairs (t= 22.425, $P < 0.001$, Fig. 2c). Such spatial aggregations, however, do not occur among every rare species in the assemblage. In fact, we found that 91.1% of positive networks were more modular than their negative counterparts (t= 39.676, P<0.001, Fig. 2d). This result remained ~~similar~~ after accounting for network size and connectivity (t= 11.306, P<0.001, for 67.3% of assemblages, Methods). Moreover, while 60.7% of positive networks were significantly modular, only 13.8% of negative ones showed this pattern (Methods). Taken together, these findings show that rare species tend to generate modular networks of positive spatial associations.

The patterns of negative and positive associations networks remain largely invariant regardless of different probability thresholds to detect significant associations, the use of quantitative links and assumptions of disparate null models (Appendix S1). This robust and conspicuous spatial organization ~~point to~~suggest that the underlying mechanisms can also be responsible for the persistence of rare species. On the one hand, dissimilar habitat preferences between dominant species and groups of weak competitors[14] may generate this pattern, enhancing also rare species persistence. Indeed, numerical simulations (Methods and Appendix S3) show that this possibility increased the probabilities of reproducing realized association network patterns, regardless of different interaction networks reflecting hypothesized assembly mechanisms (Fig. 3a and Appendix S3). This, however, mainly occurs when habitat preferences are strong, a situation that should arise under marked environmental gradients most likely far from the reality of the fine-scaled assemblages studied here (Methods and Table S1). Complementarily, positive interactions within groups of rare species may also contribute and/or generate these modular positive networks. This may moreover increase the persistence of weak competitors since, just as in harsh abiotic environments[23], the biotic harshness produced by superior competitors could be counterbalanced by positive interactions among rare species. Accordingly, simulations show that the inclusion of positive interactions within groups of weak competitors increases the chance of species persistence by 58.2% compared to assemblages

ruled by competition alone (Appendix S3). Our simulations also reveal that this hypothesis most likely reproduces the observed patterns in association networks compared to other stabilizing mechanisms such as neutral colonization-extinction dynamics[9], intransitive competition[24,25], differential density-dependent effects[26,27] or facilitation by nurse species[7] (Figs. 3b, S1 and S2; and Appendix S3). Interestingly, the combination of habitats and positive interactions yields the highest probability of reproducing the observed network patterns (Fig. 3a and Fig. S2). This further suggests that, even under strong differences in habitat preferences, stabilizing forces, such as facilitation or complementarity, ~~would be necessary to~~ enhance the coexistence of groups of rare species in reduced microhabitats[15,17]. Besides habitat selection, it seems that modular positive interactions among rare species can contribute to the pattern we found and the persistence of these species, which agrees with recent experimental evidence[28].

Overall, our results show that ecological assemblages are consistently organized in positive and negative association networks across the main biological groups (*i.e.*, animals and plants) and geography (Fig. 4 and Table S2). This ubiquity sheds light on the long-standing diversity paradox as the potential mechanisms leading to this organizational pattern can also enhance the persistence of rare species. Modular positive interactions among weak competitors emerged as a plausible mechanism even when assessed in conjunction with different microhabitat preferences. Questions remain

about the relative contribution and feedbacks of these positive interactions and microhabitats. Nevertheless, the generality of the findings presented here bring us closer to understanding the assemblage of the vast biodiversity on Earth.

Materials and Methods

Data acquisition

Assemblage data were collected from published studies in peer-reviewed journals and our own surveys (Appendix S4 and Table S1). Each assemblage consists of at least ten samples where the abundance of at least ten species of the same trophic guild was recorded. In order to minimize the effects of environmental heterogeneity and dispersion on spatial patterns, we only included datasets that showed (i) low environmental variability across samples (excluding surveys where any kind of environmental gradient was reported or no clear information about it was provided), (ii) a very reduced spatial extent (median = 0.1 ha; ranging from 0.01 to 25.6 ha), (iii) a very small grain size to increase the probability of physical and/or chemical contact among all species in the samples (median = 100 m^2; ranging from 0.002 to 400 m^2, respectively), and (iv) standardization among samples to avoid sampling biases (e.g. effects of area). Following these criteria, we gathered a total of 385 datasets distributed worldwide and representing a wide taxonomic spectrum, including bryophytes (n= 71), tracheophytes (n=279), anthozoans (n=7), and insects (n=28). Abundance was estimated as the number of individuals per sample in most of the assemblages, but a small number of assemblages included abundance data estimated as the percentage cover of the sampled surface (especially in bryophytes and plants). Since some null models only accept integer data (see below), we rounded percentages when necessary. Finally, we only used those assemblages where both positive and negative networks showed at least

two links (n=326).

Generation of association networks

For each assemblage, we calculated similarity in abundance distribution across samples for each species pair *i* and *j* using the Schoener's index[31],

$$S(i,j)=1-\sum_{k=1}^{N}\left|p_{ik}-p_{jk}\right|/2$$

where N is the number of samples and p_{ik} is the proportion of the total abundance of species *i* present in sample *k* ($p_{ik}=x_{ik}/\sum_{k=1}^{N} x_{ik}$). We compared observed similarities to 999 null values obtained through randomization of species abundances using a fixed-fixed algorithm (i.e., row and column totals are kept constant). For each observed similarity value, two one-tailed p-values were calculated as the proportion of null values (plus the observation) that were higher than or equal to and lower than or equal to the observed value for positive and negative associations, respectively. We considered an aggregation or segregation significant in those cases where associated p-values in any of the two tests were lower than or equal to 0.05. Alternative probability thresholds and null models provided quantitatively and qualitatively similar results (Appendix S1; Figs. S3 and S4). Significantly aggregated and segregated species pairs were used to generate unweighted links in the positive and negative association networks of each assemblage, respectively. It is important to note that the frequency of spurious associations (i.e., Type I errors) may be thought to be relatively high in species rich assemblages due to multiple comparisons (but see ref[32]). However, species pairwise similarities were compared against null values generated using a fixed-fixed assemblage-wise null model (i.e., a strict null model making null hypotheses among comparisons to be different but intrinsically interdependent). This partially alleviates the detection of false positives

while preventing the use of powerful false discovery rate methods[33]. Nevertheless, we used the same nominal error (i.e., $\alpha = 5\%$) to detect both positive and negative associations, making the rate of false discoveries equal in both types of networks, and allowing unbiased comparisons of their structures. Indeed, results remained largely constant when using different nominal errors (Appendix S1 and Fig. S3).

Network structure comparison

To explore whether positive and negative association networks reflected competitive processes we compared their connectivity and their relationships between abundance (calculated as the sum of the abundances across samples) and species degree (i.e., species' number of links) for each pair of network types. Connectivity is defined as the number of realised links relative to the number of potential links. This measure of connectivity may be negatively correlated with network size. Hence, we also used residual connectivity obtained from the residuals of a linear regression between the number of observed and potential links, both log-transformed[34]. On the other hand, the relationship between abundance and species degree was assessed using the Spearman's ρ correlation coefficient. Finally, to search for differences between network types we used a paired Student's t-test, where the alternative hypothesis was that negative networks present higher means than their positive pairs.

To determine if rare species have a larger participation in positive association networks, we compared the average relative abundance, weighted by the number of links of each species in the network, between the species involved in positive and negative networks. We also explored if positive networks were more modular than their negative pairs calculating modularity with the index proposed by Newman[35] (Q) along with the optimisation algorithm of Louvain[35]. The algorithm was run 100 times, and we

selected the partition that showed the highest modularity value. Since modularity can be related to network size and connectivity, we compared observed and null modularity values from random networks generated using a null model that maintains the number of links and nodes, as well as the degree sequence (implemented in the RandNetGen software[37]). Then, we computed relative modularity values as $Q_r = -2(P - 0.5)$, where P represents the proportion of null cases showing modularity higher than or equal to the observation. A paired Student's t-test was used to explore the differences between network types in all cases.

Finally, we explored whether the probability of finding the above-explained differences of positive and negative networks was related to the number of samples per assemblage (as indicative of sampling effort), an approximation of null model degrees of freedom (Appendix S1), latitude, longitude, taxonomic group (i.e., animals or plants) and species richness. To do so, we firstly generated four binomial dependent variables, based on whether i) the negative networks of each assemblage were more densely connected than their positive pairs; ii) the negative networks present higher positive abundance-degree relationships; iii) the positive networks tend to be composed of less abundant species; and iv) the positive networks were more modular. Then, we fitted logistic models with a logit link function.

Numerical simulations

We ran simulations to explore whether different interaction matrices and/or habitat preferences can generate the patterns observed in association networks. We designed a simulation model composed of 20 samples and ten species, whose individuals were randomly distributed at the outset. Individuals reproduce, colonise a randomly chosen sample or die, with probabilities dependent on the density of individuals and the sample

carrying capacity (K = 100). We subsequently incorporated the effects of both competition and positive interactions by modifying these probabilities depending on the species identities of co-occurring individuals (Appendix S3). That is, individuals of dominant species reduce the probability of reproduction and colonisation, while increasing mortality probability, of co-occurring individuals of subordinate species. Benefactor individuals have the opposite effects on beneficiary individuals' probabilities (Fig. S5a and b).

We further incorporated the effects of dissimilarities in habitat preferences by setting four habitats preferred by different groups of species (Fig. S5c). Specifically, the probabilities of reproduction, survival (i.e., one minus the mortality probability) and colonisation in non-preferred habitats were multiplied by a habitat-tolerance coefficient, β, ranging between 0 (null tolerance) and 1 (total tolerance; Appendix S3). Hence, when β = 0, individuals are highly specialist and only allowed to reproduce, survive or colonize in preferred habitat, whereas β = 1 corresponds to a neutral habitat scenario.

We ran simulations following six hypotheses explaining species assembly and coexistence (Fig. 3 and Appendix S3). i) A neutral interaction model, where all species were ecologically equivalent[9]. ii) A hierarchical competition model with one strong competitor. iii) An intraspecific density-dependent model, where superior competitors suffer more from intraspecific competition[5]. iv) A model incorporating intransitive competition[25], where the superior competitor is outcompeted by three species, which, in turn, outcompete all species except specific pairs (i.e., theoretically promoting the generation of empirical association patterns; see Appendix S3). v) A nurse model[7] with four superior competitors, three of which facilitate different pairs of subordinate species. vi) A model reflecting positive interactions within three groups of three rare species. Fourteen additional matrices with different settings to these six general models

were also explored (see Appendix S3; and Figs. S1 and S6).

Simulations were run using a wide range of combinations (n=216) where demographic rates (i.e., reproduction, mortality and dispersal) had different relative importance (Appendix 3.4). In addition, we also used five values of the habitat-tolerance parameter (β). For each interaction matrix and parameter combination, we ran 25 replicates of 5,000 iterations each. We quantified the probability of simulated association networks showing empirical patterns (i.e., differences between positive and negative networks in connectivity, abundance-degree relationship, abundance and modularity), as well as the probability of persistence of all species (i.e., non-extinction), as the proportion of all our replicates showing these patterns. Finally, these probabilities were averaged across the parameter space defined by demographic rates where the *Competition* model, under neutral habitat preferences, fulfilled expectations from the competitive exclusion principle (i.e., weak competitors went extinct; *P* (non-extinction) = 0; see Appendix S3 and Fig. S2).

Acknowledgments. We are very grateful to Joaquín Hortal and Stefano Allesina for their critical comments on an early version of the manuscript. The simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N. J.C. is supported by the Carl Tryggers Foundation for Scientific Research (CTS 16:384). E.A. is supported by a postdoctoral grant founded by the Universidad Complutense de Madrid. C.J.M is supported by the Swiss National Science Foundation (SNSF-31003A-144162). R.B.M. is supported by the Spanish Ministry of Science and Innovation Predoctoral Fellowship BES-2013-065753. M.S., J.A.B.C. and J.M.G acknowledge support from the University of Geneva. X.A. was supported by a Ramón y Cajal research contract by the Spanish Ministry of Economy and Competitiveness (RYC-2015-18448). M.R. is supported by the Swedish Research

Council grant 2016-00796. JAN was supported by a Colombian COLCIENCIAS PhD scholarship. F.A.M. is grateful to CAPES for a PhD scholarship (120147/2016-01). A.L., P.F. and J.M.G. were funded by AGORA Project (ref. CGL2016-77417-P; MINECO, Spain). C.M.M. was supported by an IdEx Bordeaux Postdoctoral Fellowship (VECLIMED project). A.H. was supported by the University of Alcalá Own Research Programme's 2018 Postdoctoral Grant and Basque Country Government funding support to FisioClimaCO2 (IT1022-16) research group. L.J. received productivity grants from of CNPq (process: 307597/2016-4).

Author contributions. J.C. and J.M.G. conceived the ideas; J.C. and J.M.G. designed the analyses with contributions from E.A., A.E., C.J.M. and R.B.M.; J.C., E.A., R.B.M., M.S., C.A., X.A., N.G.M., J.A.N., F.A.M., I.D., A.L., J.A.B.C., C.M.M., P.F., A.H., L.P., L.J., A.C., and J.M.G. collected the data; J.C. analyzed the data with assistance from C.J.M., M.R. and M.N.; J.C., E.A. and J.M.G. lead the writing in close collaboration with A.E., C.J.M., R.B.M., M.S., C.A., and R.M.V.; all authors contributed to the development and writing of the paper.

Code availability. We will make all R scripts used in this study freely available upon acceptance of the manuscript or if required by the referees during revision.

Data availability. We will make the dataset used in this study freely available upon acceptance of the manuscript or if required by the referees during revision.

References

1. Gaston, K.J. *Rarity*. (Chapman & Hall, London, 1994).
2. Hardin, G. The competitive exclusion principle. *Science* **131**, 1292-1297(1960).
3. Hutchinson, G. E. The paradox of the plankton. *Am. Nat.* **95**, 137-

145 (1961).

4. Schoener, T. W. Resource partitioning in ecological communities. *Science*, **185**, 27-39 (1974).
5. Yenni, G., Adler, P. B. & Ernest, S. M. Strong self-limitation promotes the persistence of rare species. *Ecology* **93**, 456-461 (2012).
6. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Evol. Syst.* **31**, 343-366 (2000).
7. Soliveres, S. et al. A missing link between facilitation and plant species coexistence: nurses benefit generally rare species more than common ones. *J. Ecol.* **103**, 1183-1189 (2015).
8. Grilli, J. et al. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210 (2017).
9. Hubbell, S. P. *The unified neutral theory of biodiversity and biogeography*, vol. 32 (Princeton University Press, 2001).
10. Durrett, R. & Levin, S. Spatial aspects of interspecific competition. *Theor. Popul. Biol.* **53**, 30-43 (1998).
11. McIntire, E. J. & Fajardo, A. Beyond description: the active and effective way to infer processes from spatial patterns. *Ecoloy* **90**, 46-56 (2009).
12. Arnan, X., Gaucherel C. & Andersen A.N. Dominance and species co-occurrence in highly diverse ant communities: a test of the interstitial hypothesis and discovery of a competition cascade. *Oecologia* **166**, 783-794 (2011).

13. Atkinson, W. D. & Shorrocks, B. Competition on a divided and ephemeral resource: a simulation model. *J. Anim. Ecol.* **50**, 461-471 (1981).

14. Hart, S. P., Usinowicz, J. & Levine, J. M. The spatial scales of species coexistence. *Nat. Ecol. Evol.* **1**, 1066 (2017).

15. Chacón-Labella, J. et al. Evidence for a stochastic geometry of biodiversity: the effects of species abundance, richness and intraspecific clustering. *J. Ecol.* **105**, 382-390 (2017).

16. Saiz, H. et al. Evidence of structural balance in spatial ecological networks. Ecography **40**, 733-741 (2017).

17. Freilich, M. A., Wieters, E., Broitman, B. R., Marquet, P. A. & Navarrete, S. A. Species co-occurrence networks: can they reveal trophic and non-trophic interactions in ecological communities?. *Ecology* **99**, 690-699 (2018).

18. A. Faisal, et al. Inferring species interaction networks from species abundance data: A comparative evaluation of various statistical and machine learning methods. *Ecol. Inform.* **5**, 451-464 (2010).

19. Barberán, A., Bates, S. T., Casamayor, E. O. & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME journal* **6**, 343 (2012).

20. Borthagaray, A. I., Arim, M. & Marquet, P. A. Inferring species roles in metacommunity structure from species co-occurrence networks. *Proc. R. Soc. Lond. B*, **281**, 20141425 (2014)

21. Ulrich, W. & Gotelli, N. J. Null model analysis of species

associations using abundance data. *Ecology* **91**, 3384-3397 (2010).

22. Tilman, D. Resource competition between plankton algae: an experimental and theoretical approach. *Ecology* **58**, 338-348 (1977).

23. Callaway, R.M. et al. Positive interactions among alpine plants increase with stress. *Nature* **417**, 844-848 (2002).

24. Allesina, S. & Levine, J. M. A competitive network theory of species diversity. *Proc. Nat. Acad. Sci.* **108**, 5638-5642(2011).

25. Gallien, L. et al. The effects of intransitive competition on coexistence. *Ecol.Lett.* **20**, 791-800 (2017).

26. Comita, L. S. et al. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**, 330-332 (2010).

27. Cody, M. L. & Diamond, J. M. *Ecology and evolution of communities*. (Belknap Press, Cambridge, 1975).

28. Bimler, M. D., Stouffer, D. B., Lai, H. R. & Mayfield, M. M. Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency. *J. Ecol.* **106**, 1839-1852 (2018).

29. Fruchterman, T. M. & Reingold, E. M. Graph drawing by force-directed placement. *Softw. Pract. Exp.* **21**, 1129-1164 (1991).

30. Bailey, R. G. Explanatory supplement to ecoregions map of the continents. *Environ. Conser.* **16**, 307-309, (1989).

31. Schoener, T. W. The Anolis lizards of Bimini: resource

474 partitioning in a complex fauna. *Ecology* **49**, 704-726 (1968).
475 32. Rothman, K. J. No adjustments are needed for multiple
476 comparisons. *Epidemiology* **1,** 43-46 (1990).
477 33. Benjamini, Y. & Yekutieli, D. The control of the false discovery
478 rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165-
479 1188 (2001).
480 34. Araújo, W.S. et al. Contrasting Effects of Land Use Intensity and
481 Exotic Host Plants on the Specialization of Interactions in Plant-
482 Herbivore Networks. *PLoS ONE* **10**, e0115606 (2015).
483 35. Newman, M. E. Modularity and community structure in
484 networks. *Proc. Nat. Acad. Sci.* U.S.A. **103**, 8577-8582 (2006).
485 36. Blondel, V. D. et al. Fast unfolding of communities in large
486 networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
487 37. Colomer-de-Simón, P., Serrano, M. A., Beiró, M. G., Alvarez-
488 Hamelin, J. I. & Boguná, M. Deciphering the global organization of
489 clustering in real complex networks. *Sci. Rep.* **3**, 2517 (2013).
490

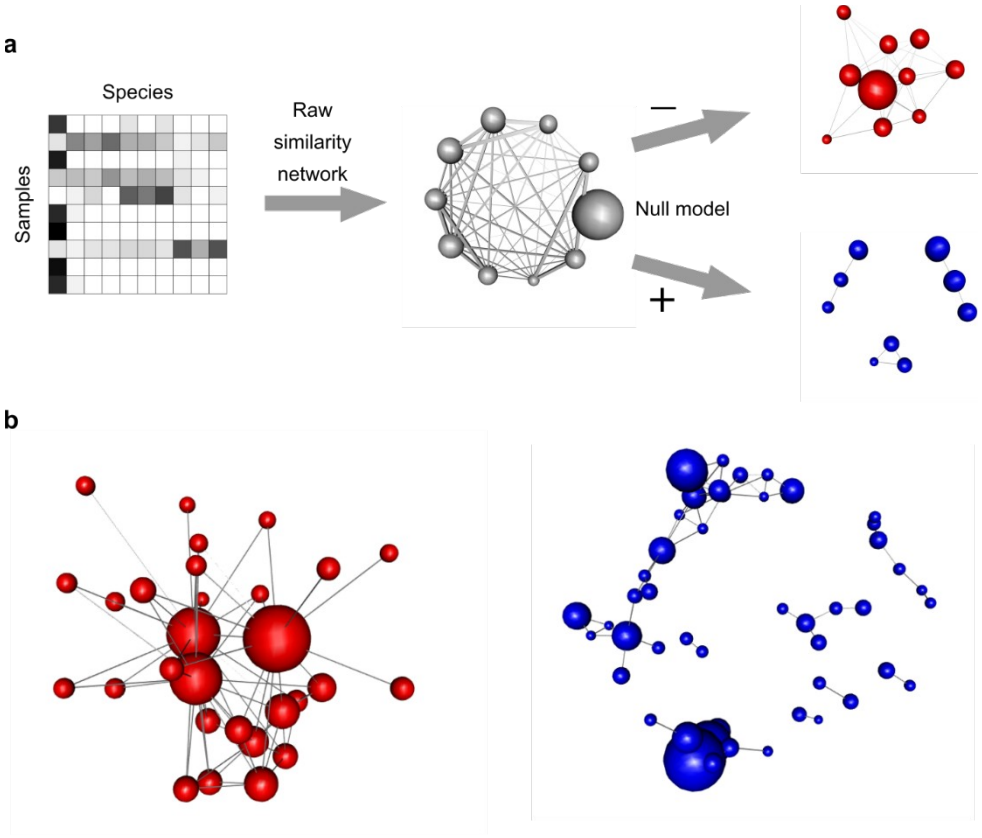
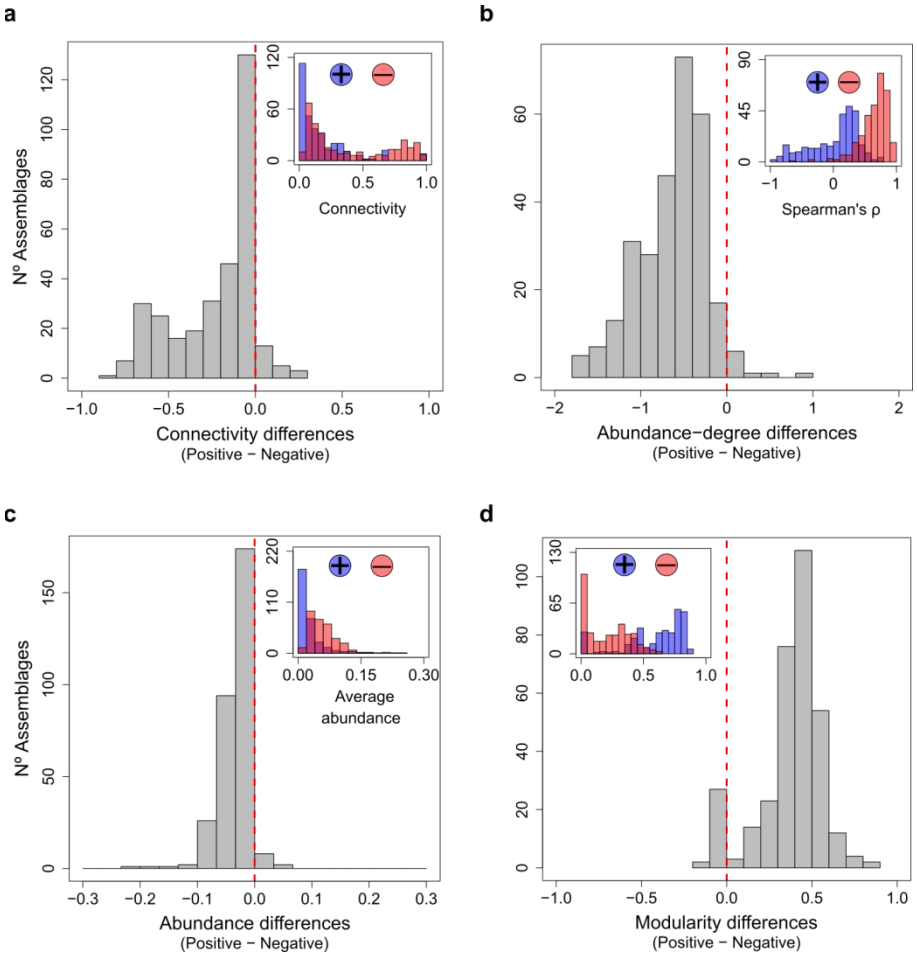


Fig. 1. Approaching assembly mechanisms through the lens of positive and negative association networks. a, Species segregations and aggregations can inform on the main mechanisms underlying ecological assemblages. These spatial patterns are measured between species pairs using the similarity in the spatial distribution of their individuals. Observed similarities are compared with those obtained by a null model to distinguish actual associations from those generated by chance. Species pairs whose individuals are more aggregated in samples than expected by chance receive a positive link in association networks (blue nodes). Species pairs whose individuals are more segregated than randomly expected receive a negative link in association networks (red nodes). **b**, Positive (blue) and negative (red) networks of a tropical rainforest tree

assemblage (see “Barra_Paraguacu” in Table S1). The size of the nodes is proportional to the species’ abundances at the assemblage level. Networks were plotted using the Fruchterman-Reingold force-directed layout algorithm[29].



509

510 **Fig. 2. The contrasting patterns of positive and negative**
511 **association networks**. **a**, The higher connectivity of negative
512 networks indicates that species segregation dominates over species
513 aggregation. **b,** These segregations are monopolized mostly by
514 dominant species, as shown by the strong relationship between
515 abundance and species degree (i.e., number of links of a species) in
516 negative networks. **c,** In contrast, less abundant species are more
517 prone to generate positive associations, although, **d**, these
518 associations only occur among specific groups of rare species, as
519 indicated by the higher modularity of positive networks. Main
520 histograms show the differences in network features between pairs of
521 positive and negative networks. Insets show the raw values for both

types of networks.

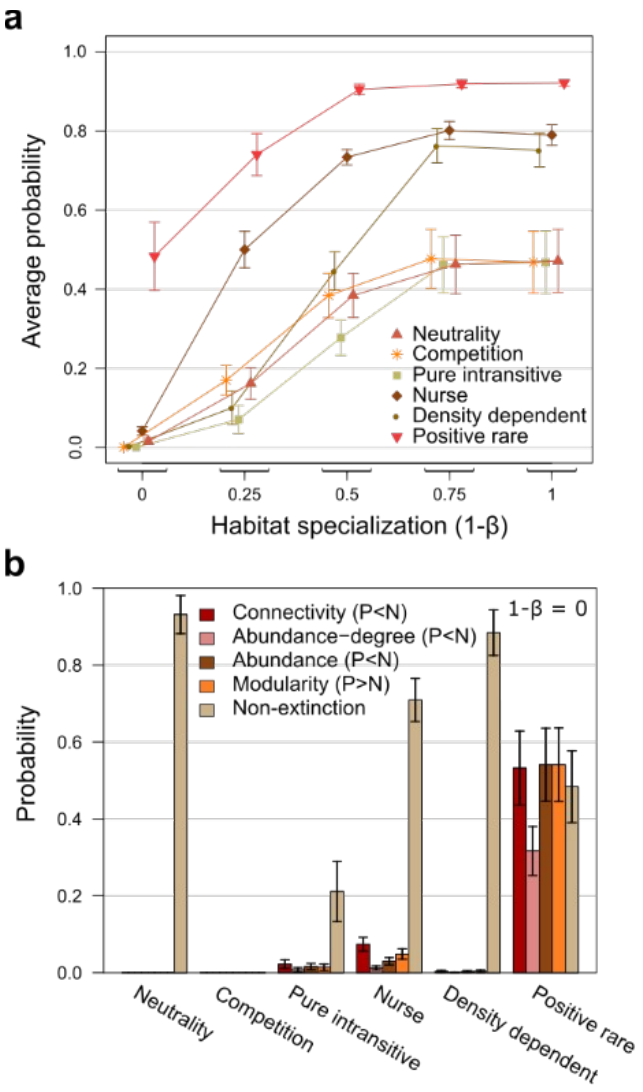


Fig. 3. Positive interactions among weak competitors alone or together with habitat preferences reproduce realized association patterns. **a**, Dissimilarities in habitat preferences between dominants and groups of rare species may generate the empirical patterns of association networks, regardless of different assembly mechanisms. However, this only occurs when habitat specialization is strong. Moreover, the combination of habitats and positive interactions among weak competitors (*Positive rare*) yields the highest probabilities. The y-axis represents the average probabilities of finding the four empirical patterns, and the x-axis

depicts a gradient of habitat specialisation (see Methods and Appendix S3). **b**, All theoretical models explaining species coexistence increase the chance of species persistence (Non-extinction) relative to simulated assemblages only driven by hierarchical competition. However, positive interactions among groups of rare species is the most likely model to generate simulated assemblages showing the same association networks as empirical assemblages (Connectivity, Fig. 2a; Abundance-degree, Fig. 2b; Abundance, Fig. 2c; and Modularity, Fig. 2d). The y-axis represents the probability of simulated association networks showing empirical differences between positive and negative networks across different combinations of reproduction, mortality and dispersal rates where interactions are expressed (see Methods and Appendix S3). Error bars depict confidence intervals at $\alpha = 0.05$. P: positive networks. N: negative networks.

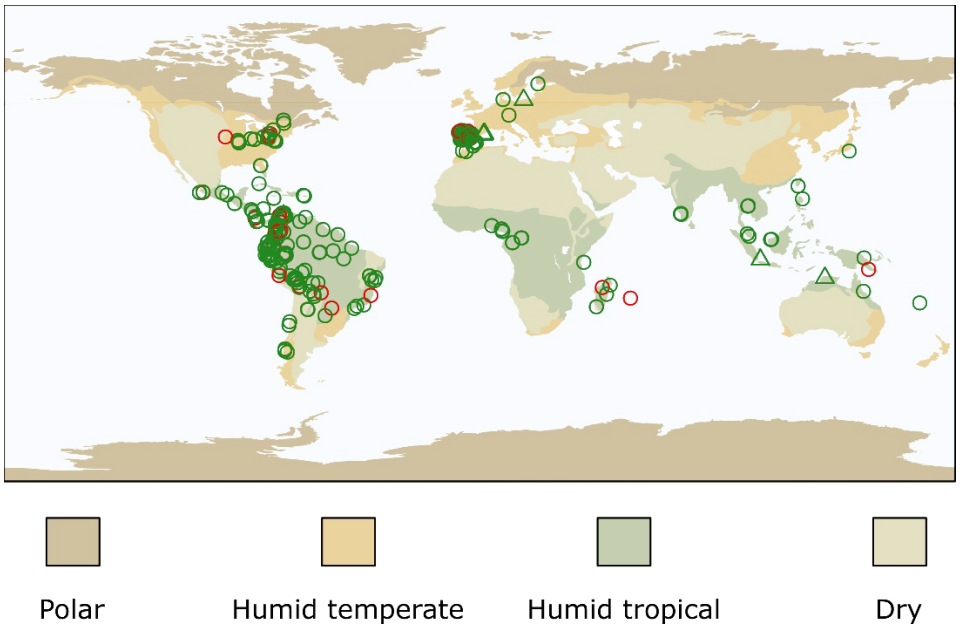


Fig. 4. Organization of association networks remains invariant across the globe and regardless of taxa. Circles and triangles represent plant and animal assemblages, respectively. Green colour depicts assemblages where positive networks were both composed of less abundant species and more modular than negative counterparts, whereas red colour shows assemblages where these patterns were not found. Map colours represent the Earth climatic zones proposed by Bailey[30].

Positive associations among rare species and their persistence in ecological assemblages

Joaquín Calatayud, Enrique Andivia, Adrián Escudero, Carlos J. Melián, Rubén Bernardo-Madrid, Markus Stoffel, Cristina Aponte, Nagore G. Medina, Rafael Molina-Venegas, Xavier Arnan, Martin Rosvall, Magnus Neuman, Jorge Ari Noriega, Fernanda Alves-Martins, Isabel Draper, Arantzazu Luzuriaga, Juan Antonio Ballesteros-Cánovas, César Morales-Molino, Pablo Ferrandis, Asier Herrero, Luciano Pataro, Leandro Juen, Alex Cea, Jaime Madrigal-González

Correspondence to: j.calatayud.ortega@gmail.com

Content

Appendix S1. Sensitivity analyses

S1.1. Sensitivity analysis to probability thresholding and weighted network analysis.

The establishment of positive and negative links in association networks relies on the comparison of observed similarities between pairs of species against those generated using a null model. Here, spatial associations that depart from null expectations more than a given probability threshold are classified as either positive or negative links. We used a 0.05 probability threshold for both negative and positive links. Although this constitutes a commonly used nominal error, the question remains as to whether the observed patterns of negative and positive networks are robust to different thresholds. Additionally, during the probability thresholding we only used information on link presence-absence yet link weights may provide richer information. Here, we first tested whether the observed patterns in association networks are robust to different probability thresholds. Then, we also explored whether the structure of weighted association networks is similar to the one of non-weighted networks.

To explore the robustness of association network patterns to different probability thresholds, we established positive or negative links based on 20 different nominal errors ranging from 0.010 to 0.499. Notice that we used a 0.499 probability, instead of 0.500, since positive and negative links were established based on the same null distribution and an association with a 0.500 probability could show the same probability of being negative or positive. Then, for each assemblage and network type, we computed connectivity, abundance-degree relationship, mean species abundance weighted by the number of links of the species in each network, and modularity as described in Methods. Finally, for each probability threshold we calculated the proportions of assemblages where: i) negative networks were more densely connected than

positive ones; ii) negative networks showed a stronger degree-abundance relationship than positive pairs; iii) negative networks showed higher mean abundance, weighted by the number of links, than positive counterparts; and iv) positive networks were more modular than negative ones.

The generation and analysis of quantitative association networks entail some methodological challenges. Regarding link weights, it may be possible to assign weights based on a standard effect sizes, $SES_{ij}=(x_{obs}-\bar{x}_{null})/sd(x_{null})$, where x_{obs} represents the observed pairwise similarity and x_{null} the null similarity values. This approach classifies positive and negative links while conferring a weight, but it has some important drawbacks, being among the most important the assumption that the null values follow a normal distribution, and the dependence of this measure on the number of null model realisations (see, for instance ref[1]). To avoid this issue, we calculated a non-standard effect size as, $ES_{ij}=-2(P_{ij}-0.5)$, where P_{ij} represents the probability for the association between species i and j of being higher than and equal to, $P_{ij}(x_{obs}\geq x_{null})$, or lower than and equal to null model expectations, $P_{ij}(x_{obs}\leq x_{null})$, respectively for positive and negative links. This is a non-parametric measure of magnitude ranging between -1 and 1. However, its estimation has a known limitation since the probability used for detecting positive associations is not complementary to the probability used for detecting negative ones (i.e. $P_{ij}(x_{obs}\geq x_{null})\neq 1-P_{ij}(x_{obs}\leq x_{null})$). This implies that for each pair of species we have two non-complementary probabilities that in an extreme case can be equal (i.e. if all null values are equal to the observation, both positive and negative acceptance probabilities are equal to 1). We overcame this limitation by computing the ES only for those associations (either positive or negative) showing probabilities of acceptance lower than 0.5 and assigning a zero otherwise,

$$ES_{ij} = \begin{cases} -2\left(P_{ij} - 0.5\right), if\ P_{ij} < 0.5 \\ 0, otherwise \end{cases}$$

This measure ranges between 0 and ≈1, and was calculated independently for positive and negative links. We used the 0.5 threshold since it ensures that each species pair has a maximum of one ES value higher than zero (i.e. if $P_{ij}\left(x_{obs} \geq x_{null}\right) < 0.5$, then $P_{ij}\left(x_{obs} \leq x_{null}\right) > 0.5$ always, and *vice versa*). The corrected ES provides two important pieces of information since it provides an estimate of the effect size and the sign of the spatial aggregation, either positive (if $ES_{positive} > 0$ and $ES_{negative} = 0$) or negative (if $ES_{positive} = 0$ and $ES_{negative} > 0$). Notice also that it should be possible that an association cannot be classified either as positive or negative (when $ES_{positive} = 0$ and $ES_{negative} = 0$). This, however, would only occur when a reduced null model space results in more than half of the null values being equal to the observation, which, as shown below (see Appendix S1.3), is very unlikely in our dataset.

We measured different properties of weighted positive and negative association networks using analogous indices to non-weighted networks. To explore whether negative associations were more prevalent than positive ones, we first developed a quantitative measure of connectivity for our particular networks. Non-weighted network link density (or connectivity) is measured as the number of observed links divided by the number of potential ones, $C = \frac{N}{0.5 \times n \times (n-1)}$, where N denotes the number of observed links and n the number of nodes. Analogously, weighted network density (herein, weighted connectivity) can be calculated as the sum of link weights relative to the maximum expected. In our case, the maximum weight expected for a single link is equal to 0.998. Here notice that probabilities are calculated as the proportion of 999 null values plus the observed value ($n = 1000$) being equal or different to the

observed value. Thus, when all null values are different from the observed one we have, $P_{ij}=1/1000$ and then, $ES_{ij}=-2\left(\frac{1}{1000}-0.5\right)=0.998$. The potential maximum weighted connectivity can be therefore measured as 0.998 times the number of potential links, $C_{w_{max}}=0.998\times 0.5\times n\times(n-1)$; where n represents the number of nodes. Accordingly, we measured weighted connectivity as,

$$C_w=\frac{\sum_{i=1}^{L} l_i}{0.998\times 0.5\times n\times(n-1)}$$

where L represents the total number of links and l_i the weight of link i. This index ranges from 0 (when there is no links) to 1 (for a fully connected graph with all links weights equal to 0.998).

We also studied the role of dominant and rare species in generating positive or negative associations. To do so, we explored the relationship between species abundance and the average weight of positive and negative links jointly. That is, we first multiplied negative links by -1, and then calculated the average link weight for each species, including both positive and negative links. When a given species has more and/or stronger negative links this index is negative, and vice versa. Then, we used the Spearman's rank correlation coefficient to measure the relationship between abundance and this index. Negative relationships between this mean link weight and species abundance indicate that abundant species have more and/or stronger negative links, whereas rare species have more and/or stronger positive links. Finally, to explore whether positive weighted networks were more modular than negative pairs, we used the Louvain algorithm[2] to optimize the quantitative version of the modularity index proposed by Newman[3].

Results from the sensibility analysis revealed that the configuration of positive and negative association networks was, in general, highly robust to the probability threshold used for defining significant associations (Fig. S3). The only exception was for the differences in connectivity between positive and negative networks. For this property, we found a decrease in the proportion of assemblages where negative networks were more connected than positive pairs as the probability threshold increases (Fig. S3). In any case, the main result holds even with less robust thresholds since the proportion of assemblages where negative networks were more densely connected was always higher than 0.60.

The above results also agree with the analysis of weighted association networks. That is, negative networks were more densely connected that positive pairs in 0.61 of the assemblages ($t = -14.244$, $P < 0.001$), and the mean link weight was negatively correlated with abundance in 0.95 of assemblages (average Spearman's $\rho = -0.528$, $sd = 0.285$). Similarly, 0.92 assemblages showed positive networks being more modular than negative pairs ($t = 31.169$, $P < 0.001$). Overall, these results showed that our findings are robust to different probability thresholds and the use of quantitative information.

S1.2. Sensitivity analysis to null model performance

Association networks inference relies on the accurate detection of significant spatial association patterns among species pairs (see Methods). The significance of pairwise associations was assessed by using a fixed-fixed null model (hereafter, null model I), which randomizes species abundances while keeps constant the marginal totals of rows and columns (see "r2dtable" algorithm in ref[4]). Thus, this model assumes that: (i) species populations and

carry capacities of samples are constant, and (ii) species can occur across all samples. This null model has been recommended due to its low type I and II error rates[5]. Moreover, its assumptions fit rather well with the reduced spatial scale and the putatively homogeneous conditions of the studied assemblages. Nevertheless, we explored whether alternative null models might yield different results. To do so, we used two other alternative null models (hereafter, null models II and III) to generate the association networks. The null model II is similar to null model I, but it also maintains all zero entries in the species per sample matrix (see "quasiswap count" algorithm in ref[4]). Thus, the null model II includes the assumption of the existence of forbidden occurrences, more typical of systems showing some degree of environmental heterogeneity. This further assumption should reduce the number of significant positive and negative associations. On the other hand, the null model III maintains the frequency of occurrence and abundance of species, as well as the species richness of samples (see "independent swap" in ref[5]). This null model assumes that species are divided in subpopulations distributed across samples and that each subpopulation has specific growth restrictions. Additionally, by constraining species frequency of occurrence and sample richness this null model assumes restrictions in species ranges more typical from systems with marked dispersal or environmental barriers. Although these assumptions are rather far from the reality of our assemblages (since we only used assemblages where samples were close enough to allow dispersion among them and with apparently low environmental variability, see Methods), we used this null model to show that our results are robust to assumptions associated with disparate null models. In all cases we used 999 null replicates and calculated significance as explained in Methods.

Once association networks were generated based on each null model approach, we assessed differences in structure between positive and negative networks. In particular, we studied

differences in: (i) connectivity, (ii) the relationship between abundance and species degree, (iii) species abundances, and (iv) modularity (see Methods). Overall, our results remained qualitatively and quantitatively similar irrespective of the null model approach, thus corroborating the robustness of the patterns observed in association networks (Table S3 and Fig. S4).

S1.3. Effects of null model space

The null model space —or "degrees of freedom"— defines the degree of differences among null model realizations. This critically determines the capacity to detect significant associations and, in turn, the association network patterns. For instance, if the null model space is small enough, observed and null values can be equal, precluding the detection of significant positive or negative associations. The null model space is only controlled by the constraints assumed and the related matrix characteristics. However, determining a large enough null model space *a priori* is a complicated task, even more when considering the different constraints of the null models we used. In an attempt to handle this, we selected assemblage surveys with at least ten species and ten sites, which constitute a proxy to ensure a large enough null model space in general. However, given the particular constraints of each null model, using matrix dimension as a proxy may not work in all situations. For instance, null models II and III keep non-zero elements (i.e. matrix fill), and this should then also contribute to the space of these null models. We, therefore, explored whether our results are affected by the null model space.

We estimated the null model space for each assemblage and null model as the average (and minimum) dissimilarity among null model realizations. To do so, we firstly calculated the proportion of non-overlapping individuals between each pair of realizations (e.g. realizations A and B) as,

$$Ov = \frac{\sum_{i=j=1}^{N} ¿ r_{ij}^{A} - r_{ij}^{B} \vee ¿}{2T} ¿$$

where r_{ij}^{A} represents the abundance of species i in site j of the realization A; N the number of entries in the matrix of species per sites and, T the total number of individuals. This dissimilarity measure is the matrix level equivalent of 1 minus the Schoener's index previously used and ranges between 0 (when both realizations are identical) and 1 (when there is no overlap between realizations). We computed this index for each pair of the 999 realizations of each null model and assemblage. Then, we measure the extent of the null model space using both the mean and minimum dissimilarity among all possible pairs. When the null model space is large we should find a large proportion of non-overlapping individuals between realizations. Finally, we explored the relationship between null model space and matrix dimension, as well as its effects on the probability of obtaining the patterns of positive and negative association networks.

On average across all assemblages the mean proportion of non-overlapping individuals was 0.34 ± 0.02; 0.58 ± 0.01; and 0.58 ± 0.01; respectively for null models I, II and III (mean and confidence interval at α = 0.05). Hence, even in the cases of the null model producing fewer individual switches, 34% of individuals were switched on average across different null model realisations. Similarly, the minimum proportion of non-overlapping individuals was 0.27 ± 0.02; 0.42 ± 0.01; and 0.38 ± 0.02; respectively for each null model. Hence, on average, at least 27% of individuals were switched. Moreover, we found that no null model realizations were equal, regardless of assemblage or null model (i.e. the minimum dissimilarity was always higher than zero). These results show that in all cases the null model space was large enough as to not generate two identical null matrices after 999 realizations, which strongly suggests that our

results are not affected by restrictions in the degrees of freedom of the null models. Indeed, we found non-significant effects of the minimum dissimilarity in the regression models explaining the probabilities of finding positive and negative network differences ($P > 0.01$ in all cases, see Table S2). Notice that we used the minimum dissimilarity in these models, instead of the mean dissimilarity, since it represents the most restrictive measure of null model space and it was highly correlated with the mean dissimilarity ($r = 0.99$). Finally, we found a strong correlation between the minimum dissimilarity and matrix size ($r = 0.75$, 0.76 and 0.77; respectively for null models I, II and III), which validates the use of matrix size as a proxy for null model space size.

Appendix S2. Association patterns across assemblages

Whether a particular species is more often involved in negative associations when common or vice versa is an interesting question.

We explored this idea using those species that were present in at least 5 assemblages (n=384).

We first standardised the abundance of species i in each assemblage as follows:

$$Std.A_i = \frac{\log(A_i) - \mu_{\log(A)}}{\sigma_{\log(A)}},$$

where $\mu_{\log(A)}$ is the mean log-transformed abundance of the species in a given assemblage and $\sigma_{\log(A)}$ the standard deviation. Then, we calculated the balance of positive, N_i^{pos}, and negative associations, N_i^{neg}, normalized by the number of species in the assemblage, N, also for species i in a given assemblage,

$$B_i = \frac{N_i^{neg} - N_i^{pos}}{N}.$$

This balance is negative if a given species has more positive than negative associations. Finally, we used the Spearman's correlation coefficient to explore the relationship between both variables. For most species (64.1 %) we obtained a positive relationship between both variables (mean Spearman's $\rho = 0.162 \pm 0.041$ at $\alpha = 0.05$), indicating that, in general, species tend to be involved in more negative associations when common. Generalisations of this result should be taken with caution since the species occurring in five or more assemblages only represent less than 5% of the total number of species used in this study and were exclusively plants.

Appendix S3. Extended numerical simulations

S3.1 Model

We ran simulations accounting for biotic and habitat factors to explore whether proposed mechanisms enhancing rare species persistence do indeed promote species coexistence while leading to the observed structural patterns of positive and negative association networks. Here, we explain the general features of the model and the following sections contain the biotic scenarios (S2.2), the habitat ones (S2.3), the simulation settings (S2.4), the analytical protocol (S2.5) and results and discussion (S2.6).

We designed a simulation model composed of 20 samples and ten different species – similar to the mean sampling size of the empirical assemblages, whose individuals were randomly distributed at the outset. Individuals reproduce, release propagules to colonize a randomly chosen sample (i.e. dispersion), or die[7]. These processes occur with given probabilities, affecting reproduction (α_R), mortality (α_M) and dispersal rates (α_D). Each sampling site had a constant carrying capacity, which reflects the amount of resources available at the sample, and resources proportionally decrease as the number of individuals increases. As resource decreases in the

sample the probability of reproduction (P_R) and successful colonization (P_c) decrease, whereas the probability of mortality (P_d) increases. Hence, in a neutral scenario –where individuals of all species are ecologically equivalent[8], the probabilities of reproduction, P_R, and mortality, P_D, of an individual i in a given sample will be determined by both, intrinsic factors and resource availability in the sample, such that

$$P_R(i) = \alpha_R \left(1 - \frac{\sum_{j=1} N_j}{K}\right) \qquad \text{(Eq. 1)},$$

$$P_D(i) = \alpha_M \frac{\sum_{j=1} N_j}{K} \qquad \text{(Eq. 2)},$$

where N_j is the number of individuals of species j and K represents the carrying capacity. Similarly, the probability of an individual i to colonise a given sample will be

$$P_C(i) = \frac{\alpha_D}{N_s}\left(1 - \frac{\sum_{j=1} N_j}{K}\right) \qquad \text{(Eq. 3)},$$

where N_s represents the number of samples. Note that when the number of individuals reaches the carrying capacity $\left(\sum_{j=1} N_j = K\right)$ the probabilities of reproduction and colonization are equal to 0, while the probability of mortality is equal to the intrinsic mortality ($P_D = \alpha_M$).

To incorporate the effects of different competitive abilities we assumed that, when co-occurring, individuals of dominant species acquire resources more efficiently than individuals of

less competitive species, decreasing their probabilities of reproduction and colonization and increasing the probability of mortality. Analogously, the effect of positive interactions can be incorporated assuming that benefactor individuals will increase the resource uptake efficiency of beneficiary ones. This, however, seems rather difficult to occur within assemblages of species belonging to the same trophic level, since some degree of resource interference can be always expected. We therefore took a conservative step and simulated complementarity based positive interactions[9]. That is, rather than simulating increases in resources uptake, we consider a positive interaction by setting the negative impact of competition under the expectations of a neutral interaction. By doing so "beneficiary" individuals will have increased probabilities of reproduction and colonization while reduced probability of mortality as compared to a neutral interaction (Fig. S5a). We modify equations 1, 2 and 3 to include the strength of competition and complementarity as follows,

$$P_R(i) = \alpha_R \left(1 - \frac{\sum_{j=1} N_j \theta_i^j}{K}\right) \quad \text{(Eq. 4)},$$

$$P_D(i) = \alpha_M \frac{\sum_{j=1} N_j \theta_i^j}{K} \quad \text{(Eq. 5)},$$

$$P_C(i) = \frac{\alpha_D}{N_s} \left(1 - \frac{\sum_{j=1} N_j \theta_i^j}{K}\right) \quad \text{(Eq. 6)},$$

where θ_i^j represents the effects of species j on species i. Notice that $\theta_i^j > 1$ comparatively reduce P_R and P_C while increasing P_D, representing dominance of species j on species i. On the other hand, $\theta_i^j < 1$ produce the opposite effect (i.e., a relative increase in P_R and P_C while decreasing P_D), meaning that species j complements species i. These equations also accommodate intraspecific interaction effects (i.e., when $i = j$), which allow us to explore intraspecific density-dependent processes (see below).

At the assemblage level, the interaction effects can be described by the *competition-complementarity assemblage matrix*, where the columns represent the effects of other species to a given species, the rows depict the effects of a given species to other species and the diagonal represent the intraspecific effects (Fig. S5b). Hence, the column marginal totals represent the total effect of all species to a given species. Large differences in these values between species would induce system-wide extinctions of weaker competitors. On the other hand, marginal total of rows represents the total effect exerted by a given species to the others, which will be related to system-wide species abundances at steady state (Fig. S5b). This simulation scheme allows exploring the patterns of spatial association networks generated by interaction matrices reflecting different assembly mechanisms. Notice, nevertheless, that the inference of the interaction effects in the competition-complementarity assemblage matrix (i.e. θ_i^j values) is challenging because the large number of parameters and the lack of empirical knowledge. To overcome this issue, we established constraints to the row and column marginal totals. Arbitrary θ_i^j values will not produce aberrant results as long as marginal totals fulfil these conditions. Specifically, the system-wide persistence of all species should rely on very similar column marginal totals between species, regardless of specific θ_i^j values. In the same way, row marginal totals can be

constrained to guarantee the generation of species abundance distributions like empirical ones (i.e. few abundant and many rare species).

S3.2 Biotic interactions scenarios

We studied 20 interaction matrices summarizing the proposed mechanisms on species assemblages and rare species persistence. Six of them are presented in the main text and here are identified by an asterisk. A neutral interaction matrix (***Neutrality* matrix***) was generated considering that all individuals were ecologically equivalent[8] with all competition-complementarity coefficients equal to 1 (Fig. S6). To strictly test that the empirical patterns of association networks cannot be generated by only negative interactions we generated 13 interaction matrices with only negative interactions but reflecting different stabilizing hypotheses. We first generated three matrices assuming hierarchical (or transitive) competition with one superior competitor (***Competition 1* matrix***); three superior competitors (***Competition 2* matrix**) and four superior competitors (***Competition 3* matrix**). Intransitive competition has been proposed as the underlying mechanism explaining species coexistence[10,11,12]. This hypothesis assumes that intransitive competition networks stabilize species abundances. Because of intransitive competition can take any form[9], we studied seven different interaction matrices. The first two ones represent a case where competitive effects were reciprocal between superior competitors and there were two superior competitors (***Intransitive 1* matrix**, see Fig. S6); and three superior competitors (***Intransitive 2* matrix**). In both cases, superior competitors outcompeted different species in the community theoretically enhancing the achievement of associations among rare species due to the same responses to dominant ones. The next two matrices had the same structure than the former ones, but in this case weak competitors also

exerted a negative effect on stronger ones (***Intransitive 3* and *4* matrices**, see Fig. S6). Finally, we generated three intransitive matrices with perfect intransitive competition[12]. In the first one, we assumed that a single superior competitor dominates over all the other species but one, which in turn, dominates over the (former) superior competitor (***Pure intransitive 1* matrix**, see Fig. S6). In the second intransitive matrix, we assumed the longest intransitive loop possible, where all species outcompete one species but are outcompeted by another one (***Pure intransitive 2* matrix**, see Fig. S6). In the last intransitive matrix explored, we strictly test if intransitive competition is not enough to generate the association network patterns we found in this study. For this, we generated an interaction matrix forcing the appearance of these patterns. This matrix was characterized by one superior competitor outcompeting the rest of the species but three species, which, in turn, outcompete all other species but different pairs of weak competitors (***Pure intransitive 3* matrix***, see Fig. S6). This should result in the appearance of a dominant species with a larger number of negative associations while positive association networks should tend to be modular. The final three matrices with only negative interactions reflect a stabilizing hypothesis based on asymmetrical density dependent effects[13,14]. This hypothesis assumes that rare species are not poor competitors, but they experience a high negative impact of intraspecific competition[14]. The interaction matrices for this hypothesis assumed the existence of a larger number of superior competitors that also show high negative effect exerted by conspecifics (see Fig. S6, ***Density dependence 1*, *2* and *3* matrices**). Notice that we used a large number of superior competitors since these species represent rare species and the number of rare species is, in general, large in empirical assemblages. Moreover, this scenario should produce virtually same results than assuming that other species in the community were interfering in resources-based interactions (e.g. when strong competitor suffer a higher predation pressure).

Facilitation from nurse species has also been hypothesized as a mechanism enhancing rare species persistence[15]. As in the case of intransitive competition, the interaction matrix representing this mechanism can take many forms. Thus, to test if facilitation from nurse species is not enough to generate the observed association patterns, we set three interaction matrices forcing the appearance of them. In the first one, we assumed that there are three superior competitors negatively affecting all species but two different species that are positively affected (***Nurse 1*** **matrix**, see Fig. S6). These three species can therefore be considered as nurse species facilitating different pairs of species. The second scenario was similar but also including a superior competitor that negatively affects the rest of the species (***Nurse 2*** **matrix***). In the last one, we assumed there are two superior competitors facilitating different sets of the community (***Nurse 3*** **matrix**). Perhaps one of the most parsimonious hypotheses explaining species coexistence would assume that competitive exclusion does not exist. Hence, we constructed an interaction matrix where this is the case (i.e., the positive matrix with all θ_i^jvalues lower < 1, Fig. S6). Modular positive interactions were assumed to explore if this mechanism would be enough to generate the observed empirical patterns (***Positive*** **matrix**). Generalized complementarity among rare species may also stabilize assemblages. Although this hypothesis should be not very plausible in real assemblages, we also generate an interaction matrix reflecting this possibility to explore if it can generate observed patterns in association networks (***Positive negative*** **matrix**). Finally, we generated a matrix reflecting modular positive interactions among poor competitors. This matrix was characterized by the presence of a dominant species and three groups of weak competitors interacting positively between them (***Positive rare****, see Fig. S6).

S3.3 Habitat preference scenarios

Besides biotic interactions, differences in microhabitat preferences may also lead to spatial aggregations and segregations[16] and, in turn, to the observed patterns of association networks. Indeed, if groups of rare species would persist in diverse microhabitats where dominant species performed worse, we could also expect negative networks to be centralized around dominant species and modular positive networks composed of rare species. To account for this, we strictly selected assemblages at fine scale and with apparently reduce environmental heterogeneity (Appendix S3) – where biotic interactions should prevail, partially solving this issue. However, unappreciated heterogeneity might exist in some assemblages, generating microhabitat conditions and reducing the role of biotic interactions. Moreover, even in the absence of abiotic heterogeneity, some microhabitats can be generated by the presence and interactions with other species in the community (e.g. soil bacteria for plants or predators for insects). These biotic habitats are more difficult to control for and might have similar effects as abiotic ones on the distribution of species' abundances.

The existence of microhabitats might question the link between spatial association patterns and biotic interactions within species of the same trophic guild. However, this would only apply when considering the theoretical extremes of a gradient between the effects of habitat specificity and competition/facilitation alike interactions. In real world, both processes more likely interact to determine the coexistence of species and the spatial distribution of their individuals, with their relative importance depending on the intensity of both the habitat specificity and competitive/facilitative interactions. This interaction between biotic and habitat factors hampers the feasibility of *a priori* predictions on the consequences of microhabitats on the structure of association networks, especially when occurring in concert with different competition/facilitation

interaction matrices. For instance, assuming a perfect hierarchical competition matrix, the existence of different microhabitats only preferred by groups of weak competitors should not be enough to produce the patterns observed in association networks. Indeed, within each microhabitat a weak competitor would turn into the strongest one, excluding remaining species that also prefer this microhabitat and preventing the existence of modular positive association networks. Here we have thus conducted numerical simulations to explore whether, and in combination with which interaction matrix, the existence of microhabitats and different habitat preferences can generate the observed association network patterns.

We distributed a number of habitats within the simulated assemblages and determined preferred and non-preferred habitat for each of the species. Following the rationale of the general model in section S2.1, we assumed that the probabilities of reproduction, mortality and successful colonization in preferred habitats is only determined by species interactions according to equations 4, 5 and 6; respectively. To simulate habitat specificity, we first assumed that the probabilities of reproduction and colonization decrease in non-preferred habitat according to a parameter (β) representing the habitat tolerance. To incorporate the effects of habitat specificity and biotic interactions on the probabilities of reproduction and colonization we modified equations 4 and 6 as follows,

$$P_{R_{NP}}(i) = \beta\, P_R(i) = \beta\alpha_R\left(1 - \frac{\sum_{j=1} N_j \theta_i^j}{K}\right) \qquad \text{(Eq. 7)},$$

$$P_{C_{NP}}(i) = \beta P_C(i) = \frac{\beta\alpha_D}{N_s}\left(1 - \frac{\sum_{j=1} N_j \theta_i^j}{K}\right) \qquad \text{(Eq. 8)}.$$

The tolerance parameter (β) ranged between 0 (i.e., when individuals are completely excluded in non-preferred habitat), and 1 (i.e., when habitat specificity does not exist and individuals' performance will be only driven by interactions). Thus, when $\beta = 1$, $P_{R_{NP}}$ and $P_{C_{NP}}$ are equal to P_R and P_c, respectively.

We assumed that habitat specificity influences the survival probability $P_{S_{NP}}$ according to the β parameter, thus modifying equation 5 as follows,

$$P_{S_{NP}}(i) = \beta\left(1 - P_D(i)\right) = \beta\left(1 - \alpha_M \frac{\sum_{j=1} N_j \theta_i^j}{K}\right) \qquad \text{(Eq. 9).}$$

By doing so when $\beta = 1$ (i.e. there is no habitat specificity), the probability of survival will have equal consequences to the probability of mortality in preferred habitats.

There exists a plethora of combinations to distribute habitats and species preference for them. Moreover, the scenarios grow exponentially when trying to combine habitat specificity and interaction matrices. In an attempt to reduce this wide space of combinations while strictly testing our hypothesis, we distributed habitat and habitat preferences in a way that theoretically should generate the association network patterns of surveyed assemblages. That is, assuming neutral interactions, the *a priori* most likely habitat preference scenario leading to the observed patterns would be characterized by a large habitat preferred by few dominant species and many microhabitats preferred by different groups of rare species. We therefore set four habitats (Fig. S5c) across the 20 samples: one occupying 11 samples and only preferred by a single dominant species; and the remaining three habitats distributed in three samples each and preferred by three weak competitors each. We chose this configuration since it is directly comparable and compatible to the interaction matrices presented in the main text (i.e. producing one abundant

species and three groups of three rare species, see Fig. S6). In that way and for instance, when combining the Competition 1 matrix, we set the single dominant species to be alone in its preferred and largest habitat, so that it cannot exclude weak competitors. In other words, combining this habitat scenario and these interaction matrices allow us to place the single superior competitor in the largest habitat and the remaining groups of weak competitors in their corresponding small habitats.

S3.4 Simulation settings

We first studied the patterns generated by the 20 interaction matrices when there is no habitat specificity. Similarly to the interaction effects, the parametrization of the intrinsic factors controlling reproduction (α_R), mortality (α_M) and dispersion (α_D) based on empirical values is challenging[17] and even more when considering the wide taxonomic spectrum covered in this study. To circumvent this fact, we explored a wide parameter space by using all possible combinations for these parameters based on six different values ranging from 0 to 1 (i.e. 0; 0.2; 0.4; 0.6; 0.8 and 1). Thus, for each interaction matrix we generated 216 ($6\alpha_R$ x $6\alpha_M$ x $6\alpha_D$) possible combinations. In addition, to account for the intrinsic stochasticity of these models we ran 25 replicates for each parameter combination and interaction matrix, making a total of 108k simulations (25 replicates x 216 parameter combinations x 20 interaction matrices).

We then explored the effects of combing habitat preferences and biotic interactions. We set a gradient of habitat specificity using five values for the β parameter evenly distributed between 0 and 1. Hence, habitat specificity ranges from complete exclusion in non-preferred habitats (β = 0, i.e. the probabilities of reproduction, colonization and survival are equal to zero) to not exits (β = 1; i.e. these probabilities are equal in preferred and non-preferred habitats). For the sake of

simplicity and interpretability (see above) we only combined the habitat scenario with the six interaction matrices presented in the main text (i.e. the Neutrality, Competition 1, Pure intransitive 1, Density dependence 1, Nurse 2 and Positive rare matrices, according to Fig. S6). For each interaction matrix and β value we replicated the simulations explained above, making a total of approx. 160K simulations (25 replicates x 216 parameter combinations x 6 interaction matrices x 5 habitat specificity values).

Each simulation was run during 5000 iterations since comparisons with 1000 and 3000 iterations showed that, on average, a steady state was reached (see Fig. S1). Initial conditions in all cases consisted in a carrying capacity of 100 individuals per sample. Only twenty individuals were allowed to disperse in each run. Species abundances were equal for all species and were set so that the sum of all individuals was equal to 5% of the system-wide carrying capacity (i.e. 10 individuals per species). Then, individuals were randomly placed among samples.

S3.5 Analytical Protocol

We explored the structure of the positive and negative association networks resulting from the scenarios following the same protocol as for the surveyed assemblages. That is, at the end of each simulation we measured species pairwise associations using the Schoener's index[18]. Then, we explored which pairs of species significantly deviated from a fixed-fixed null model (at $\alpha = 0.05$; see Methods). By doing so, we detected significant positive and negative species associations and derived the corresponding association networks. Once the association networks were generated, we explored i) whether negative networks were more densely connected than their positive pairs; ii) whether there was a stronger positive relationship between species degree and abundance in negative networks than positive ones; iii) whether positive networks were

composed of less abundant species and iv) if positive networks were more modular than positive ones. Moreover, to evaluate which interaction matrices promoted species coexistence, v) we also examined if all species survived in the system at the end of each simulation. We used the Louvain algorithm[2] to optimize the modularity index proposed by Newman[3] and selected the partition with the highest modularity value across 100 runs.

Finally, we identified which of the simulated interaction matrices alone or in conjunction with habitat preference more likely reproduce the observed differences in positive and negative association networks found on the surveyed assemblages. To do so, we firstly calculated the proportion of simulated assemblages, with the same interaction matrix and α_R, α_M, α_D and β values (n=25), showing the four above mentioned patterns (i.e. the proportion where: (i) negative networks were more densely connected, (ii) the abundance-degree relationship was stronger and positive in negative networks, (iii) positive networks tend to be composed of less abundant species and (iv) positive networks were more modular). Additionally, we also calculated the proportion of assemblages where all species survived at the end of the simulations. These proportions indicated the probability of finding empirical differences between positive and negative networks, as well as the probability of non-extinctions in the system. We, then, computed the average of these probabilities for each interaction matrix and investigated whether these proportions varied across the parameter space defined by α_R, α_M, α_D and β for each interaction matrix. Here notice that in the main text we present average probabilities over a parameter space where demographic rates (α_R, α_M, α_D) allows the expression of interactions. This space was defined using the *Competition 1* model under neutral habitat preferences (β =1) as baseline and selecting those combinations of rates where species went extinct (i.e. $P_{\text{Competition 1}}$ (non-extinction) = 0). By doing so, we focus on results of different demographic rates that fulfil

theoretical and empirical expectations when only hierarchical competition is operating (i.e. weak competitors go extinct). Results obtained using all possible combinations of rates were similar and are discussed below (Appendix S2.6.1, see also Fig. S2).

We conducted all analyses in the R software[19], using different functions of the "igraph"[20] and "vegan"[2] packages. R scripts for the simulation protocol are available upon request.

S3.6 Results

We found qualitatively and quantitatively similar results using different numbers of iterations (n= 1000, 3000, and 5000; Fig. S1). Therefore, we will only refer to the results derived from simulations using 5000 iterations. Regarding the simulations where habitat preferences were neutral, the interaction matrix including a superior competitor and modular positive interactions among weak competitor species (*Positive rare* matrix) showed the highest averaged probabilities of finding the four empirical patterns (Fig. 3b and Fig. S1). Thus, in the absence of different habitat preferences, positive interactions among weak competitors are more likely to yield similar positive and negative association networks to those found in the empirical assemblages. Notice that the matrices only including modular positive interactions (*Positive* matrix) also produced to some extent high probabilities of finding the empirical networks patterns, especially for the differences in connectivity and modularity between negative and positive association networks. While obtaining modular positive networks agrees with *a priori* expectations for this matrix, higher connectivity in negative ones might be counterintuitive. This can, nonetheless, be explained attending to the indirect effects produced by the modular structure of complementarity based positive interactions. That is, in our simulations two co-occurring individuals complementing each other will obtain available resources faster, thus reducing the

performance of other “non-complemented” individuals and generating an indirect negative interaction. Given the number of species within each module, the number of these indirect negative interactions is larger than the number of direct positives ones, leading, therefore, to a higher connectivity in negative association networks.

Regarding species persistence, the simulations conducted with the *Positive rare* interaction matrix resulted in higher probabilities of survival of all species in the system, compared to the competition matrices (Fig. 3b; Fig. S1 and S2a). This supports the idea that the persistence of rare species can rely on the establishment of positive interactions among them, which would mitigate the negative impact of superior competitors. Survival probabilities were, nonetheless, also high for the rest of interaction matrices (Fig. 3b; Fig. S1 and S2a), which shows that these hypotheses might theoretically enhance species coexistence. However, the network structures obtained following these hypotheses were far from those empirically observed. Therefore, although neutrality, intraspecific density-dependent processes, intransitive competition and facilitation from nurse species might allow rare species persistence, our results suggest that these mechanisms are unlike to drive the empirical patterns of species assemblages.

It is worth mentioning that the averaged probabilities of finding the structural pattern observed on surveyed association network were lower when all possible combinations of demographic parameters were used (i.e. reproduction, mortality and dispersal rates; compare Fig. 3b with Fig. S2a). In fact, these values were never higher than 0.5. On the contrary, survival probabilities were considerably higher, even in the case of the competition matrices. These results show that different combinations of the demographic parameters can mask the effects of biotic interactions. The interplay between these factors has been previously proposed as a mechanism allowing species coexistence by buffering the negative effects of competition[21,22]. For

instance, low reproduction rates ($\alpha_R \cong 0$) will lead to under-saturated samples with increased resource availability and hence allowing species coexistence. On the other hand, relatively high dispersal abilities will allow weak competitors to more easily escape from samples dominated by strong competitors. Our results confirm that this interplay can certainly promote species coexistence in the absence of other mechanisms. Yet, they also show that differences in these demographic factors are not enough to generate the observed features of association networks (i.e. despite the *Competition 1* model yield a considerably high probability of non-extinction, the empirical patterns are not likely reproduced, Fig. S2a). Interestingly, the increment in non-extinction probabilities resulting from the *Positive rare matrix* when compared to the *Competition 1* matrix, was highly related to the average probability of finding the network features across the parameter space (R=0.77, $P < 0.001$). In other words, for those combinations of demographic parameters (α_R,α_R, and α_E) where the *Positive rare* matrix showed higher species survival probabilities than the *Competition 1* matrix, we also found a high probability of obtaining the observed structural patterns of surveyed association networks. Thus, when interactions become important (e.g. when non-extinction probability equals zero in the *Competition 1* matrix), positive interactions among weak competitors yields simulated association networks similar to empirical ones with higher probability (compare Fig. 3b with Fig. S2a).

Taken together, these results confirm that the structure of empirical association networks can reflect mechanisms allowing rare species persistence and that these mechanisms are most likely based on positive interactions among weak competitors, at least in the absence of habitat heterogeneity.

Habitat preference simulations revealed an increase in the probabilities of finding empirical association networks patterns for all interaction matrices. As expected, these probabilities increased as habitat specificity become more intense (Fig. 3b). These results suggest that the existence of diverse microhabitats may act as "refugia" for rare species, allowing their persistence and producing the empirical patterns, independently of interaction matrices. Yet, this uniquely occurs when habitat specificity is notably marked ($\beta < 0.5$; meaning that reproduction, survival and colonization is reduced at least to the half in non-preferred habitat), a situation that should only occur when microhabitat heterogeneity is very sharp and patent, which most surely is not case in our assemblages. Moreover, the combination of habitat preferences and the positive interaction among rare species yield the highest probabilities to find empirical patterns of association networks, regardless of habitat specificity levels (Fig. 3a and Fig. S2b). This shows that the effects of habitat preferences are enhanced by positive association among rare species (or *vice versa*). Hence, our results suggest that, even when habitat preferences are acting, modular positive interactions among rare species would still more likely promote the empirical patterns of association networks. Explanations for this are based on the stabilizing effects of complementarity or facilitation among rare species in preferred microhabitats. That is, in our simulations the persistence of different rare species into small microhabitats rely on large enough and evenly distributed (sub)populations that are able to buffer the effects of ecological drift[23]. In other words, even assuming that rare species interact neutrally, the reduced dimension of microhabitats would promote the "fixation" of one species per microhabitat due to stochastic processes in small population sizes[23].

Interestingly, the *Nurse* and the *Density dependence* matrices also produced empirical patterns at high probability, especially when habitat specificity was intense (Fig. 3a and Fig.

S2b). This can be explained by the interactions that are expressed in each microhabitat. That is, in the case of the *Nurse* matrix there are two positive and one neutral interactions between the species preferring same small habitats. In the case of the *Density dependent* matrix there were three species suffering more from intraspecific competition in each small habitat, which according to our model framework represents a kind of complementarity (i.e. an individual will perform better when co-occurring with an individual from other species than with an individual of its own species). Hence, in both cases some stabilizing forces are acting to enhance rare species coexistence in microhabitats. This further support the idea that, even when habitat specificity is intense, cooperation among rare species would more likely generate the empirical patterns of association networks while promoting the persistence of rare species.

Appendix S3. Abundance surveys and data preparation

We compared the structure of negative and positive association networks using a comprehensive databased of 427 assemblages worldwide distributed and representing a wide array of taxa including bryophytes, tracheophytes, vertebrates and arthropods. More information about these assemblages, including geographic coordinates, extent and gain, and the reference of the primary study in which they were included can be found in Table S1. We established a criterion of minimum number of samples and species on the assemblage datasets to be included in the study. Thus, all assemblage datasets consist in at least ten species of the same guild and a minimum of ten samples where the abundance of each species was recorded. The raw data used in this study is provided upon request.

S4.1 Abundance surveys in single trophic level assemblages

In this section, we provide information on sampling methodology for those assemblage datasets (n=87) that have not been extracted from published studies.

Tree species were sampled in two montane temperate mixed forests of Switzerland, Central Europe (see “alps_tree” and “jura_tree” in Table S1 for geographic coordinates). In each site, all the tree individuals (saplings included) were identified and quantified in ten non-overlapping plots of 100 m^2 with similar ecological conditions (i.e. assemblage samples). Each plot was at least 10 m apart from the closest one.

Abundance (i.e. number of aboveground individuals) of herbaceous plant species was sampled in squared sampling plots of 50 cm x 50 cm within a 30 m x 20 m stand located in sand dunes in central Iberian Peninsula (Cuéllar, Spain; see “erodium” in Table S1). The stand was selected in an open area without influence of trees or taller shrubs species. A total of 75 sampling plots were randomly distributed maintaining a minimum distance of 1 m between adjacent plots. Sampling was conducted during the period of maximum productivity in June 2016.

Angiosperm trees in Brazil were sampled in five different geographic locations in the state of Bahia (“Barra_Paraguacu”, ”Bonito_Bahia”, ”Itubera_Bahia”, ”Maracas_Bahia”, ”Semidecidous_Bahia” in Table S1). In each locality, several plots (ranging from 13 to 49) were established representing at least 10% of each forest area (i.e. assemblage samples). Due to differences in tree density between the different forest types, sampling plots had different sizes (100 m^2 for tropical rainforest and 400 m^2 for Brazilian savanna). Plots were at least 50 m apart from the closest one. Species abundance data included all the trees with at least 15 cm of perimeter at breast height in tropical rainforest sites and 10 cm in Brazilian savanna.

Herbaceous species in central and southern Spain were sampled in 71 sites on soils derived from gypsum outcrops on south-oriented gentle slopes. At each site, a 30 x 30 m square plot was established within homogeneous vegetated areas. The composition and structure of perennial vascular plants was assessed using four 30 m long linear transects parallel to the slope and 8 m apart in each plot. Species-specific cover was estimated visually in 12 quadrats (1.5 m x 1.5 m) placed in each plot (i.e. assemblage samples), with three quadrats per transect line (see "herbs_ym" from 1 to 36 and "herbs_yal" from 1 to 35 in Table S1).

Dung beetles in central Spain were sampled in a grassland located in Guadarrama Mountains at an altitude of 1500 m a.s.l. during three consecutive days (from April 28th to 30th in 2015, see "guada_dung_d_1", "guada_dung_d_2", "guada_dung_d_3" in Table S1). Each day, 10 traps baited with cattle dung were distributed around a circumference of approximately 50 m of radius in each plot (with traps being at least 30 m apart from each other). Baits were introduced into a piece of nylon stocking to avoid the stagnancy of beetles along different sampling events. All traps were checked every 30 min from dawn to dusk (approx. from 7:00 am to 8:00 pm), and all individuals we collected and subsequently identified in the laboratory. The individuals from the 10 traps were pooled together obtaining an estimation of the species abundances each 30 min. For this dataset we explored temporary patterns of species pairwise aggregation and segregation.

Dung beetles were also sampled in mountain forests in the north-western slope of Sierra Nevada de Santa Marta, Colombia (see "colom_dung_480", "colom_dung_875", "colom_dung_1280" in Table S1). Sampling was conducted in three different locations between December 1998 and January 1999. At each location, three altitudes at intervals of 400 m were selected (i.e. 480, 875 and 1280 m). Dung beetles were trapped using a linear transect of 10

pitfall traps without any preservative solution and with aerial bait. Traps were separated at least by 30 m, baited with 25 ml of fresh human dung and collected each hour during a period of 24 hours. The individuals from the 10 traps of each plot were pooled together obtaining an estimation of the species abundances each hour. For this dataset we also explored temporary patterns of species pairwise aggregation and segregation.

Abundance of ephemeral plant species was estimated in two localities in the semiarid ecosystem of the Fray Jorge National Park (Coquimbo Region, Chile; see "anuals_fray_1" and "anuals_fray_2" in Table S1). In each locality, 10 sampling quadrats were randomly distributed in a 50 m x 50 m area where species abundance was estimated as relative cover. Sampling was conducted during the period of maximum productivity in July 2010. These data are part of a more extensive experiment including artificial shading treatments and shrub influence. The sampling plots included in this study were all distributed in open habitats outside the influence of taller shrubs of trees.

S4.2 Data preparation

In this section, we provide information about data preparation for those assemblage datasets where it was necessary to fulfil the criteria on reduce spatial extent and low environmental heterogeneity.

Some ant assemblage datasets were obtained from Arnan & Blüthgen[24]. This dataset consisted in a total of 24 plot surveys (distributed in three different habitat types within four locations and during two seasons) in Barcelona, north-east Spain. Each plot contained 60 bait stations separated each other at least 5 m, where the identity and abundance of each species was recorded (see ref[24] for details). We considered each plot survey as an independent assemblage

and selected only ten of them following our criteria of at least ten different species at each assemblage (see “PHFIG_Spring”, “PHGRF_Spring”, “PHGRF_Summer”, “PHVAC_Spring”, “PHVAC_Summer”, “SHRFIG_Spring”, “SHRFIG_Summer”, “SHRGRF_Spring”, “SHRVAC_Spring” and “SHRVAC_Summer” in Table S1).

Figures

Fig. S1.

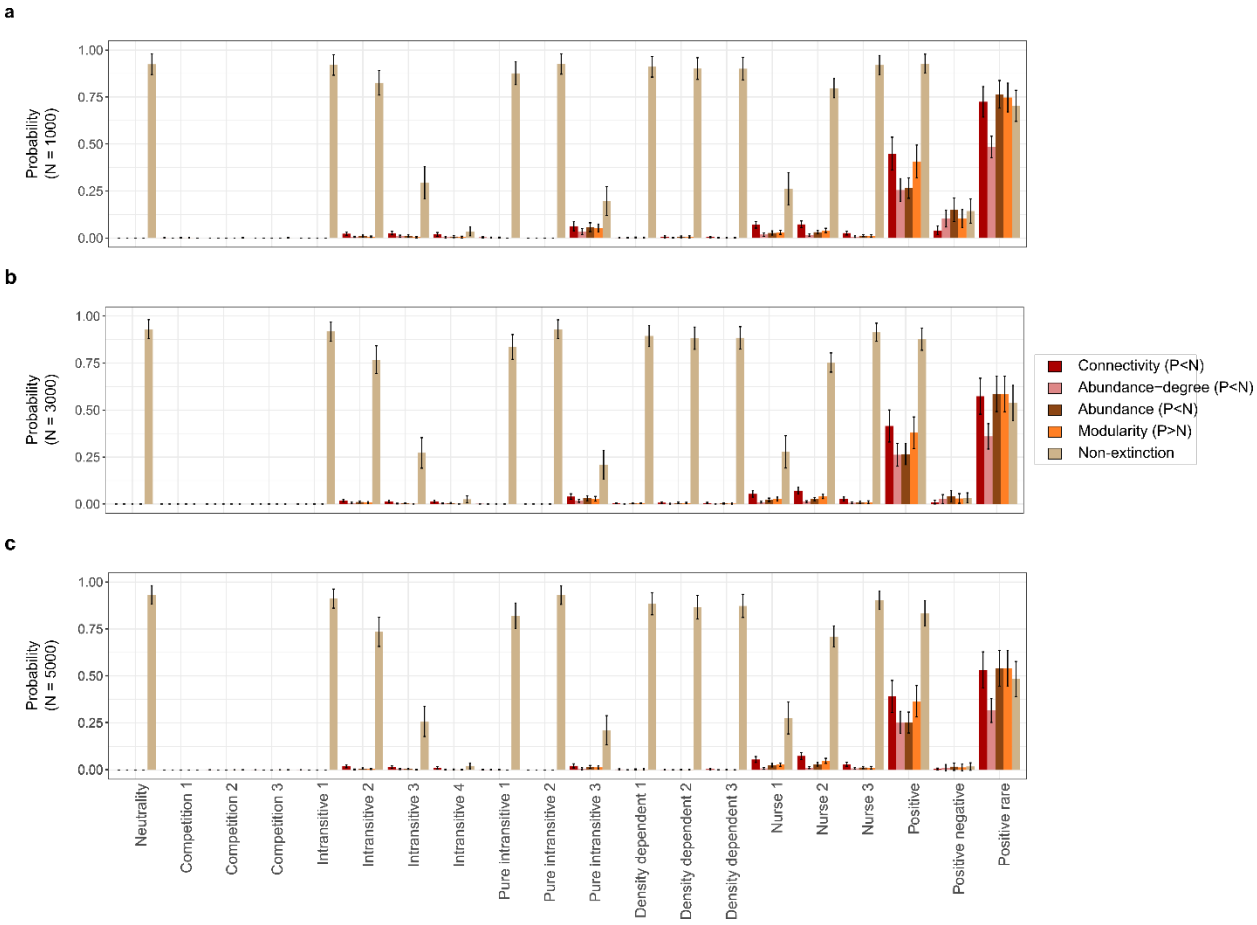


Fig. S1. Positive interactions among weak competitors yields the highest probability of reproducing empirical network patterns over 20 different models. Mean probability (y-axis) of finding simulated assemblages that show empirical network patterns under neutral habitats for each interaction matrix (x-axis) and for simulations with (**a**) 1000, (**b**) 3000 and (**c**) 5000 iterations (see also Fig. S6 and Appendix S2.6 for a detailed explanation of these results).

Fig. S2.

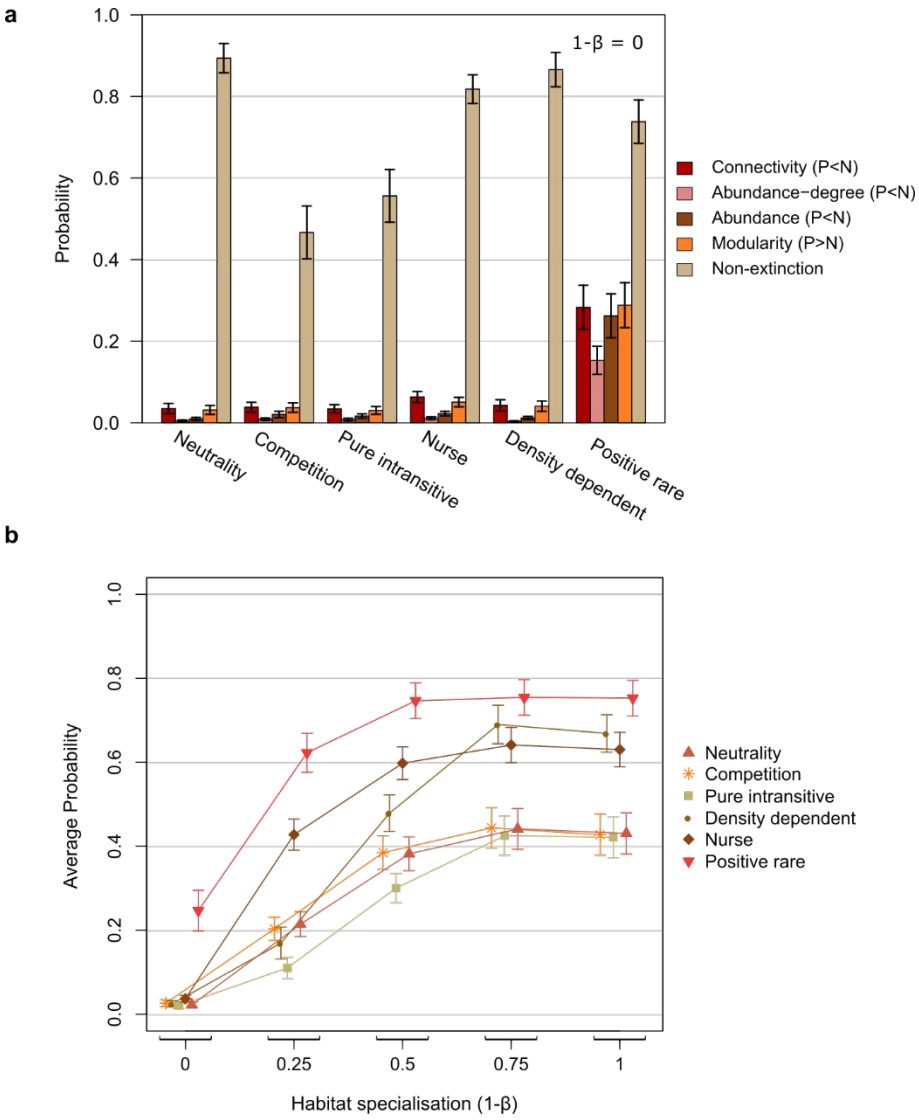


Fig. S2. Positive interactions among rare species is the model that best reproduces the empirical patterns for a broad range of demographic parameters. The effects of interactions are masked at some combinations of demographic parameters (i.e., reproduction, mortality and colonization) in simulated assemblages. **a**, This produces, for instance, that the *Competition* model do not always end up with the single dominant species (i.e., the non-extinction probability in this matrix is not always equal

to zero, as it would be expected based on the competitive exclusion principle; see also Appendix S2.6.1). Still, even when using all the possible combinations of demographic rates, positive interaction among rare species is the best model for reproducing the empirical patterns. **b**, This also holds for different levels of habitat specialization. **a**, The y-axis represents the probability of simulated association networks showing empirical differences between positive and negative networks across the entire parameter space defined by reproduction, mortality and dispersal rates (Appendix S2). **b**, The y-axis represents the average probabilities of finding the four empirical patterns across the completed parameter space, whereas the x-axis depicts a gradient of habitat specialization (Appendix S2). Error bars depict confidence intervals at $\alpha = 0.05$. P: positive networks. N: negative networks. See also Fig. 3.

Fig. S3.

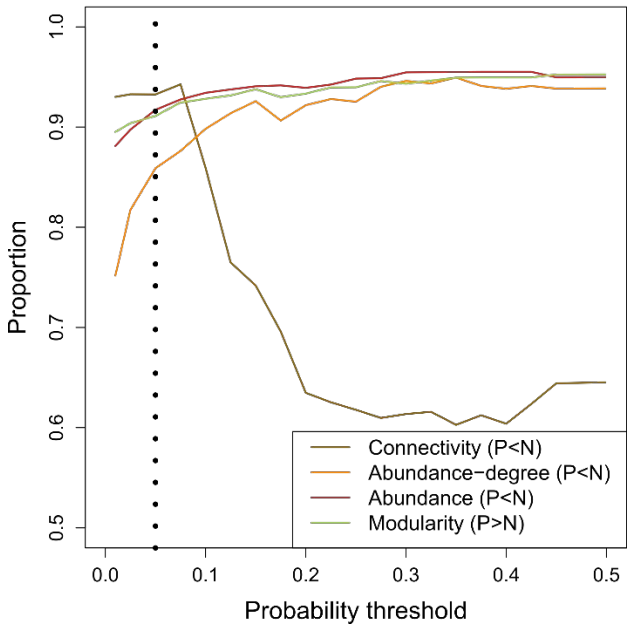


Fig. S3. The proportion of assemblages showing differences between positive and negative networks is insensitive to the probability threshold used to detected significant associations. The dotted line depicts the probability threshold used in the main text ($P = 0.05$, see Appendix S1.1).

Fig. S4.

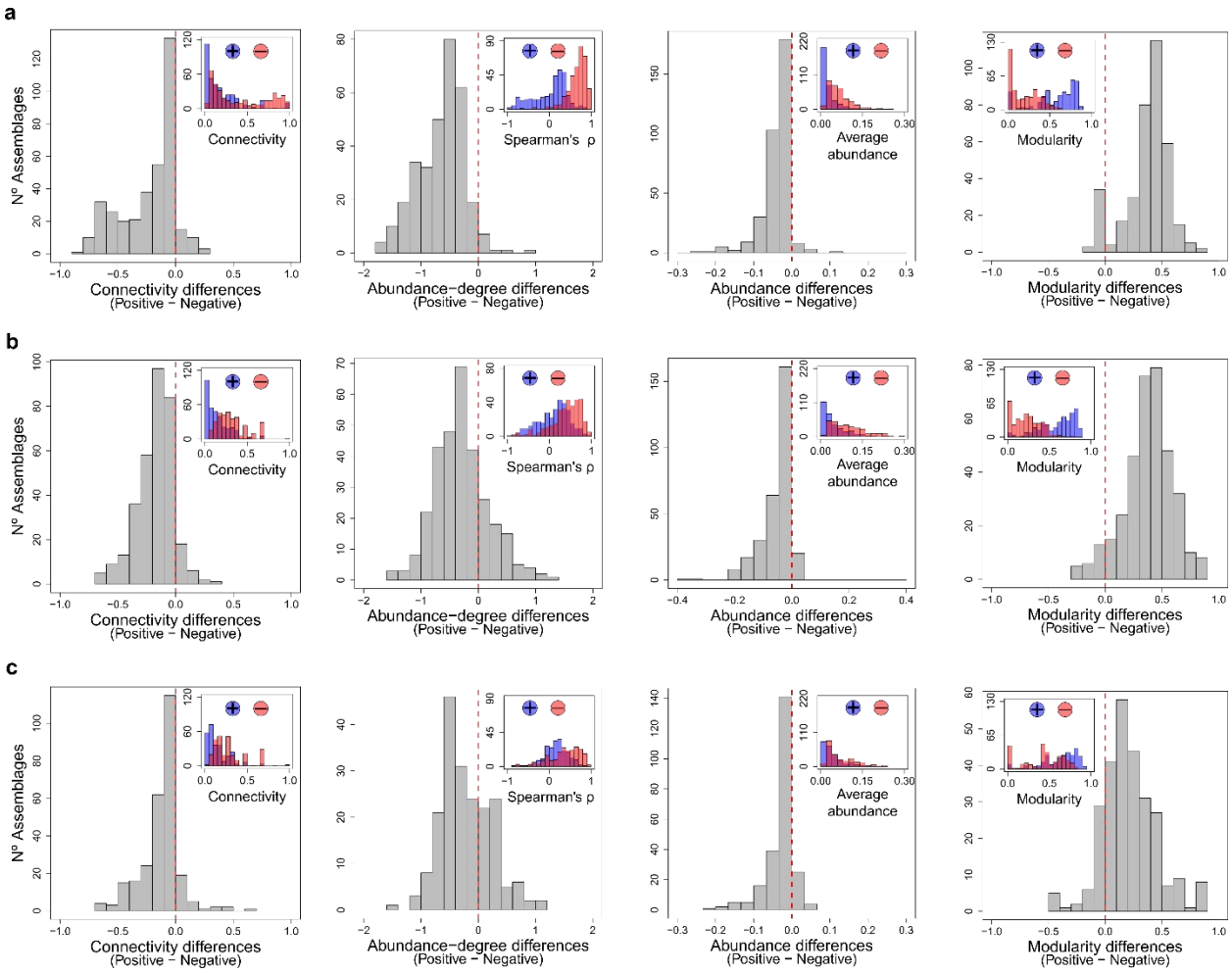


Fig. S4. Network patterns remain largely invariant when using different null models to detect significant associations. Results of (**a**) a fixed-fixed null model; (**b**) a fixed-fixed null model that also maintains matrix fill and; (**c**) a null model that keeps constant the species abundance and frequency of occurrence, as well as, sample species richness (null models I, II and III respectively in Appendix S1.2).

Fig. S5.

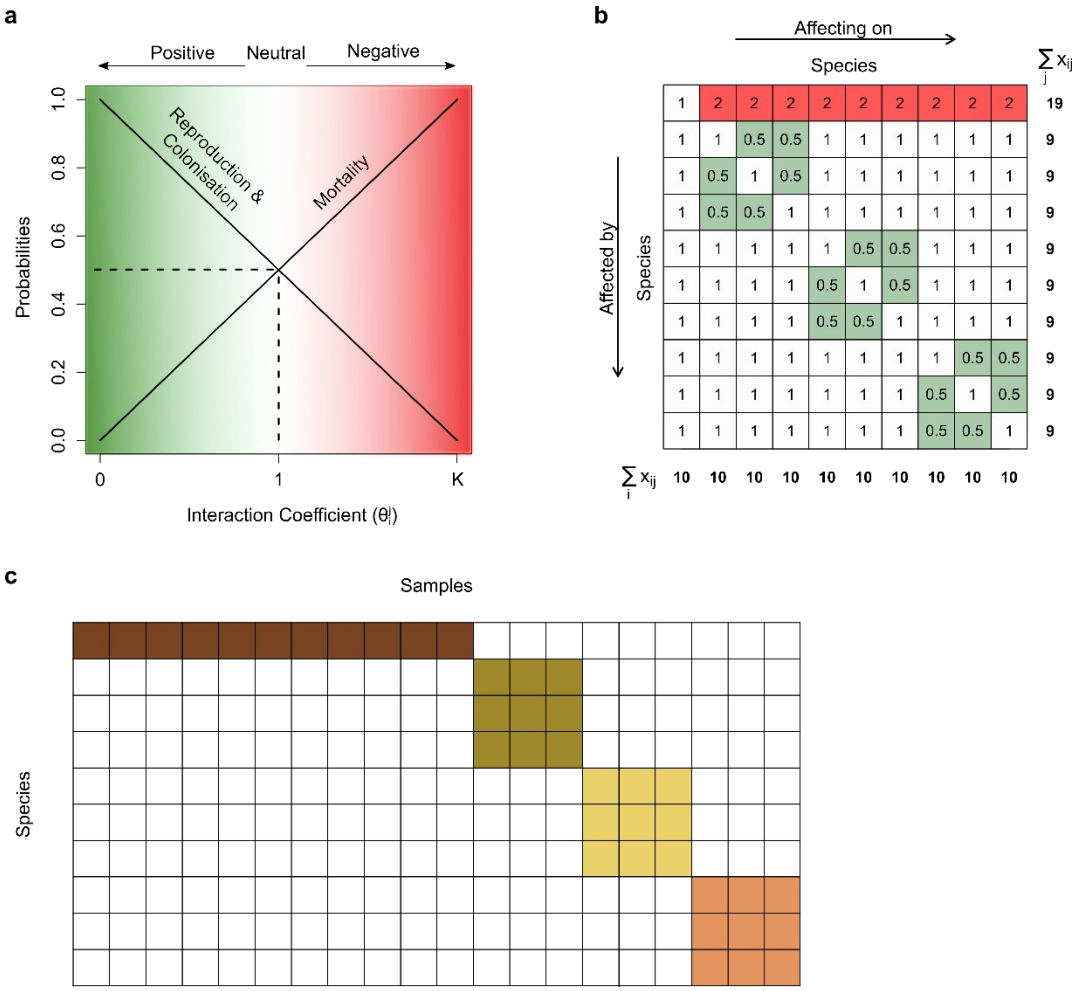


Fig. S5. a, Reproduction, mortality and colonization probabilities from an individual of species i vary according to the interaction effects exerted by an individual of species j (θ_i^j). Interaction coefficients lower than 1 lead to a decrease of mortality probabilities while reproduction and colonisation increase compared to a neutral interaction (where $\theta_i^j=\theta_i^i$). On the contrary, when interaction coefficients are higher than 1 ($\theta_i^j>1$), the probability of mortality increases, and the probabilities of reproduction and colonisation are lower than in a neutral scenario. In the x-right extreme, when the

interaction coefficient is equal to the carrying capacity ($\theta_i^j = K$), the probabilities of reproduction and colonization are equal to zero while the probability of mortality is equal to 1. **b**, An example of the interaction matrices used in the simulations. Values in column j represent the effects of other species on the focal species. Values in row i represent the effects of the focal species on the other species. The marginal total of column j ($\sum_i x_{ij}$) depicts the combined effects of all species on the focal species, whereas the marginal total of row i ($\sum_j x_{ij}$) represents the overall competitive ability of the focal species. **c**, Matrix representation of the habitat preferences. Each color presents one habitat and the species that prefer it.

Fig. S6.

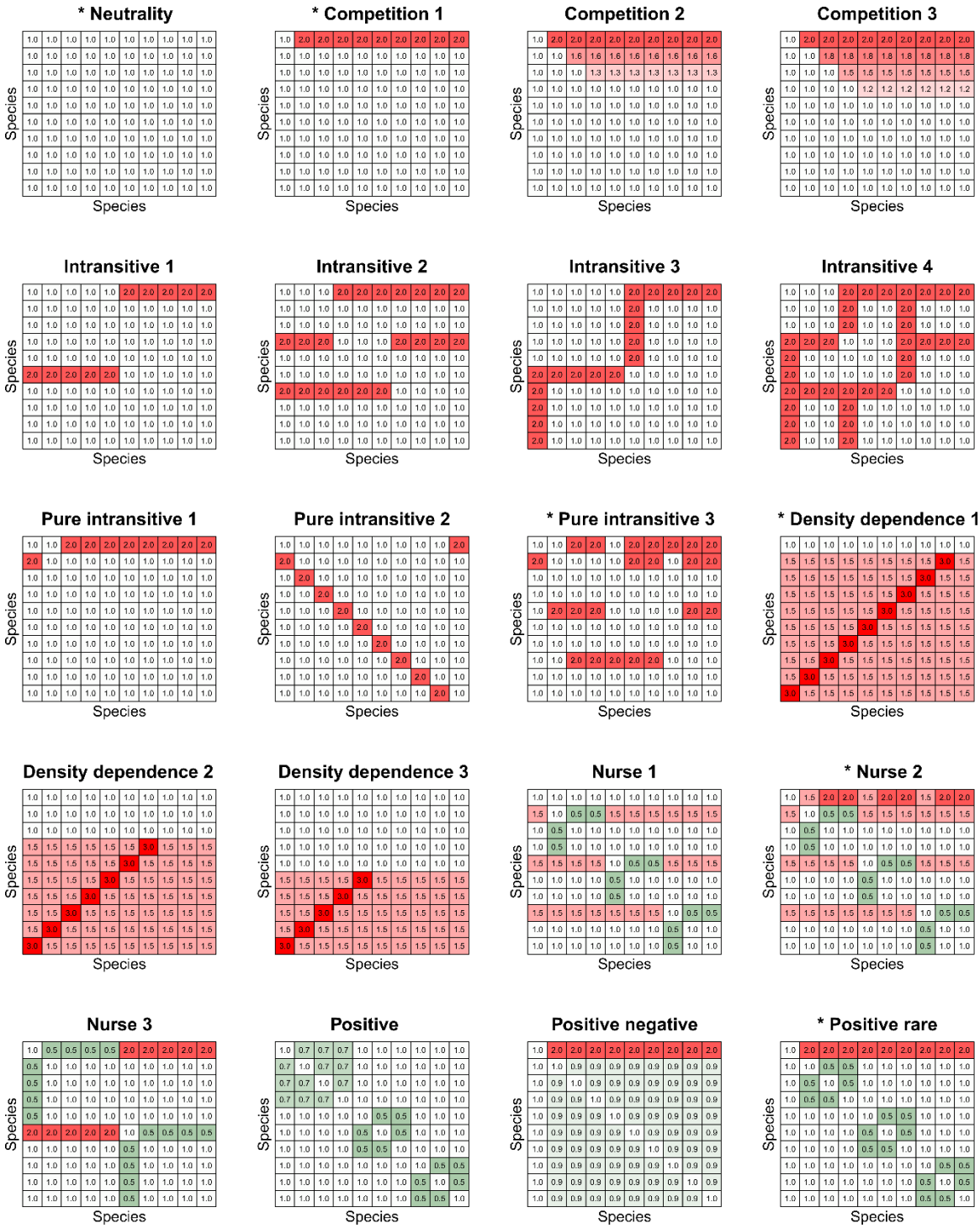


Fig. S6. **Interaction matrices reflecting different hypotheses about species coexistence and rare species survival**. Cell numbers represent

the interaction effects (θ). Red colour indicates competitive interactions (i.e., $\theta>1$), and green colour depicts positive ones (i.e., $\theta<1$). * denotes the matrices that were presented in the main text and combined with the habitat preference scenario (see Appendix S2.2).

Tables

Table S1.

ID	Latitude	Longitude	Taxon or T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
allacher	48.70	11.50	Trees	10	21	100.000	0.100	25
allpahua	-3.95	-73.42	Trees	10	276	100.000	0.100	25
alps_tree	46.33	6.87	Trees	10	17	100.000	2.000	Appendix S3
alterdoc	-2.50	-54.97	Trees	10	34	100.000	0.100	25
altodemi	10.92	-73.83	Trees	10	87	100.000	0.100	25
altosapa	7.17	-75.90	Trees	10	64	100.000	0.100	25
amotape	-4.15	-80.62	Trees	10	54	100.000	0.100	25
anchicay	3.75	-76.83	Trees	10	143	100.000	0.100	25
ankarif	-16.32	46.82	Trees	10	96	100.000	0.100	25
antado	7.25	-75.92	Trees	10	162	100.000	0.100	25
ants_Darwin_A	-12.70	130.98	Ants	100	46	0.002	0.032	26
ants_Darwin_B	-12.70	130.98	Ants	100	38	0.002	0.032	26
ants_Darwin_C	-12.70	130.98	Ants	100	40	0.002	0.032	26
ants_Darwin_D	-12.70	130.98	Ants	100	37	0.002	0.032	26
anuals_fray_1	-30.63	-71.67	Herbs	10	18	0.090	0.250	Appendix S3
anuals_fray_2	-30.63	-71.67	Herbs	10	16	0.090	0.250	Appendix S3
araracua	-0.42	-72.33	Trees	11	265	100.000	0.100	25
avalanch	11.30	76.58	Trees	10	69	100.000	0.100	25
bablersp	38.53	-90.67	Trees	10	23	100.000	0.100	25
baitete	-5.17	145.80	Trees	10	147	100.000	0.100	25
bakosar	1.75	110.42	Trees	10	137	100.000	0.100	25
bankamp	40.50	-81.30	Trees	10	23	100.000	0.100	25
banyong	5.00	9.17	Trees	10	197	100.000	0.100	25
base_encina_a	31.06	-8.73	Bryophytes	14	11	0.040	0.023	27
base_encina_b	32.55	-4.92	Bryophyt	18	11	0.040	0.020	27

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
			es					
belem	-1.50	-47.98	Trees	10	129	100.000	0.100	25
benito	15.33	-92.25	Trees	10	30	100.000	0.100	25
berbicer	5.50	-58.80	Trees	10	101	100.000	0.100	25
beza1	-23.63	44.57	Trees	10	31	100.000	0.100	25
beza2	-23.63	44.57	Trees	10	40	100.000	0.100	25
bilsa	0.62	-79.85	Trees	10	121	100.000	0.100	25
blohmr	8.57	-67.58	Trees	10	65	100.000	0.100	25
boraceia	-23.38	-46.00	Trees	10	124	100.000	0.100	25
bosquede	19.50	-96.95	Trees	10	42	100.000	0.100	25
brio_fo_100	42.94	-6.51	Bryophytes	20	24	0.040	8.019	28
brio_fo_102	42.04	-3.61	Bryophytes	19	21	0.040	6.024	28
brio_fo_103	42.58	-3.95	Bryophytes	19	13	0.040	6.776	28
brio_fo_105	40.95	-4.38	Bryophytes	16	10	0.040	19.878	28
brio_fo_106	40.27	-2.83	Bryophytes	20	20	0.040	3.786	28
brio_fo_109	41.00	-2.71	Bryophytes	20	23	0.040	8.258	28
brio_fo_112	40.82	-3.35	Bryophytes	20	12	0.040	3.982	28
brio_fo_113	41.88	-4.10	Bryophytes	20	14	0.040	2.859	28
brio_fo_114	41.74	-2.63	Bryophytes	19	15	0.040	11.933	28
brio_fo_115	40.34	-4.15	Bryophytes	20	20	0.040	11.933	28
brio_fo_116	41.74	-5.12	Bryophytes	14	13	0.040	11.143	28
brio_fo_118	39.96	-6.66	Bryophytes	19	11	0.040	10.030	28
brio_fo_12	42.27	-6.07	Bryophytes	20	16	0.040	22.490	28
brio_fo_14	42.87	-5.63	Bryophytes	19	17	0.040	13.409	28
brio_fo_18	42.77	-4.13	Bryophytes	20	11	0.040	6.519	28
brio_fo_2	42.51	-5.32	Bryophytes	13	12	0.040	18.831	28

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
brio_fo_20	39.87	-5.31	Bryophytes	19	16	0.040	9.750	28
brio_fo_22	40.30	-3.37	Bryophytes	20	14	0.040	2.147	28
brio_fo_24	40.38	-3.44	Bryophytes	20	17	0.040	7.489	28
brio_fo_26	39.36	-5.71	Bryophytes	17	12	0.040	9.171	28
brio_fo_28	39.39	-6.43	Bryophytes	20	14	0.040	6.477	28
brio_fo_29	39.57	-7.00	Bryophytes	20	20	0.040	7.649	28
brio_fo_3	41.45	-5.81	Bryophytes	20	11	0.040	6.392	28
brio_fo_31	40.42	-3.87	Bryophytes	19	23	0.040	7.624	28
brio_fo_32	40.24	-3.92	Bryophytes	20	18	0.040	9.980	28
brio_fo_33	40.29	-3.25	Bryophytes	20	22	0.040	8.043	28
brio_fo_34	40.34	-4.26	Bryophytes	20	16	0.040	2.443	28
brio_fo_36	40.74	-3.85	Bryophytes	20	25	0.040	8.285	28
brio_fo_38	39.78	-4.55	Bryophytes	20	21	0.040	3.677	28
brio_fo_4	41.14	-6.73	Bryophytes	20	25	0.040	22.673	28
brio_fo_41	39.68	-4.63	Bryophytes	20	21	0.040	4.760	28
brio_fo_42	39.68	-4.91	Bryophytes	19	25	0.040	9.084	28
brio_fo_43	40.05	-5.13	Bryophytes	20	17	0.040	6.280	28
brio_fo_44	40.12	-4.64	Bryophytes	20	15	0.040	5.628	28
brio_fo_45	40.52	-5.72	Bryophytes	20	20	0.040	9.216	28
brio_fo_46	40.75	-5.53	Bryophytes	18	14	0.040	3.501	28
brio_fo_47	40.62	-6.60	Bryophytes	20	16	0.040	22.646	28
brio_fo_48	41.26	-6.21	Bryophytes	20	11	0.040	4.331	28
brio_fo_49	40.95	-5.98	Bryophytes	20	12	0.040	6.988	28
brio_fo_51	40.74	-4.55	Bryophytes	20	19	0.040	5.306	28
brio_fo_52	41.41	-3.21	Bryophytes	19	14	0.040	4.954	28
brio_fo_54	41.59	-6.17	Bryophyt	18	22	0.040	10.807	28

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
			es					
brio_fo_55	41.80	-6.21	Bryophytes	19	15	0.040	9.765	28
brio_fo_56	42.04	-6.10	Bryophytes	19	11	0.040	15.029	28
brio_fo_59	41.84	-4.82	Bryophytes	18	12	0.040	24.079	28
brio_fo_61	40.30	-2.25	Bryophytes	20	13	0.040	7.630	28
brio_fo_62	40.57	-4.15	Bryophytes	20	18	0.040	2.282	28
brio_fo_63	40.95	-3.76	Bryophytes	20	15	0.040	5.273	28
brio_fo_64	40.97	-3.80	Bryophytes	20	10	0.040	6.097	28
brio_fo_65	40.94	-3.77	Bryophytes	20	10	0.040	12.718	28
brio_fo_66	41.03	-3.94	Bryophytes	20	15	0.040	4.783	28
brio_fo_67	40.91	-3.99	Bryophytes	20	11	0.040	5.877	28
brio_fo_68	40.81	-3.78	Bryophytes	19	13	0.040	11.717	28
brio_fo_69	40.81	-3.61	Bryophytes	18	12	0.040	7.089	28
brio_fo_7	41.13	-5.47	Bryophytes	20	19	0.040	2.168	28
brio_fo_70	41.17	-3.88	Bryophytes	20	18	0.040	3.960	28
brio_fo_72	40.42	-3.25	Bryophytes	20	12	0.040	7.378	28
brio_fo_75	40.70	-2.41	Bryophytes	18	12	0.040	5.524	28
brio_fo_80	40.34	-5.18	Bryophytes	20	14	0.040	13.068	28
brio_fo_86	42.97	-4.90	Bryophytes	20	11	0.040	7.116	28
brio_fo_87	42.79	-4.71	Bryophytes	20	10	0.040	1.038	28
brio_fo_93	40.28	-5.84	Bryophytes	20	14	0.040	4.958	28
brio_fo_94	42.16	-7.60	Bryophytes	18	25	0.040	2.035	28
brio_fo_96	42.64	-7.14	Bryophytes	20	26	0.040	0.951	28
brisefer	-20.38	57.43	Trees	10	76	100.000	0.100	25
burling	38.95	-77.17	Trees	10	23	100.000	0.100	25
cabezade	-10.33	-75.30	Trees	10	153	100.000	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
calima	3.92	-77.00	Trees	10	263	100.000	0.100	25
camorin	-22.93	-43.37	Trees	10	157	100.000	0.100	25
campano	11.13	-74.20	Trees	10	106	100.000	0.100	25
candamo	-13.50	-69.83	Trees	10	236	100.000	0.100	25
capeira	-2.00	-79.97	Trees	10	60	100.000	0.100	25
carajas	-5.50	-51.00	Trees	10	175	100.000	0.100	25
carara	9.77	-84.53	Trees	10	153	100.000	0.100	25
carlosbo	-24.25	-46.93	Trees	10	149	100.000	0.100	25
carpanta	4.58	-73.67	Trees	10	47	100.000	0.100	25
cary	46.83	-73.75	Trees	10	20	100.000	0.100	25
cedarblu	39.77	-86.15	Trees	10	25	100.000	0.100	25
cedral	4.75	-75.55	Trees	10	141	100.000	0.100	25
cega_climb	41.36	-4.26	Trees	40	33	50.000	2.000	29
centinel	-0.58	-79.33	Trees	10	140	100.000	0.100	25
ceroneb1	0.83	-66.18	Trees	10	95	100.000	0.100	25
ceroneb2	0.83	-66.18	Trees	10	81	100.000	0.100	25
cerroayp	-4.58	-79.53	Trees	10	56	100.000	0.100	25
cerroelp	13.00	-85.92	Trees	10	68	100.000	0.100	25
cerroesp	10.47	-72.83	Trees	10	77	100.000	0.100	25
cerroolu	12.30	-85.40	Trees	10	96	100.000	0.100	25
chamela1	19.50	-105.50	Trees	10	89	100.000	0.100	25
chamela2	19.50	-105.50	Trees	10	87	100.000	0.100	25
chamela3	19.50	-105.50	Trees	10	102	100.000	0.100	25
chaquima	-14.57	-68.47	Trees	10	81	100.000	0.100	25
chiba	35.15	140.15	Trees	10	47	100.000	0.100	25
chirinos	-5.42	-78.88	Trees	10	101	100.000	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
choros_op	-29.26	-71.41	Herbs	50	14	2.250	0.450	30
choros_un	-29.26	-71.41	Herbs	50	15	2.250	0.450	30
cochacas	-11.85	-71.32	Trees	10	169	100.000	0.100	25
colom_dung_1280	10.87	-73.72	Beetles	18	13	1 hour	24 hours	Appendix S3
colom_dung_480	10.87	-73.72	Beetles	19	24	1 hour	24 hours	Appendix S3
colom_dung_875	10.87	-73.72	Beetles	16	12	1 hour	24 hours	Appendix S3
colorado	9.97	-75.17	Trees	10	121	100.000	0.100	25
colosoi	9.50	-75.80	Trees	10	111	100.000	0.100	25
constanc	-4.25	-72.75	Trees	10	233	100.000	0.100	25
cuangos	-3.48	-78.23	Trees	10	170	100.000	0.100	25
cueva	11.80	-73.47	Trees	11	92	100.000	0.100	25
cuevas	6.67	-76.50	Trees	10	126	100.000	0.100	25
cuivre	39.20	-91.00	Trees	10	29	100.000	0.100	25
curundu	8.98	-79.55	Trees	10	85	100.000	0.100	25
curuyuqr	-18.75	-62.30	Trees	10	51	100.000	0.100	25
cutervo	-6.17	-78.67	Trees	10	120	100.000	0.100	25
cuyas	-4.53	-79.73	Trees	10	54	100.000	0.100	25
cuzcoama	-12.58	-69.15	Trees	10	168	100.000	0.100	25
davies	-17.80	145.57	Trees	10	116	100.000	0.100	25
ducke	-3.00	-59.97	Trees	10	238	100.000	0.100	25
encanto	-14.63	-60.70	Trees	10	111	100.000	0.100	25
erodium	41.37	-4.25	Herbs	73	19	0.250	0.060	Appendix S3
esmerald	0.90	-79.62	Trees	10	101	100.000	0.100	25
farall	3.50	-76.58	Trees	10	133	100.000	0.100	25
fincam	2.27	-76.20	Trees	10	108	100.000	0.100	25
galeraz	10.80	-75.25	Trees	10	52	100.000	0.100	25

ID	Latitude	Longitude	Taxon or T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
genting	3.97	101.63	Trees	10	194	100.000	0.100	25
guada_dung_d_1	40.75	-4.02	Beetles	16	16	0.5 hours	8 hours	Appendix S3
guada_dung_d_2	40.75	-4.02	Beetles	21	18	0.5 hours	11 hours	Appendix S3
guada_dung_d_3	40.75	-4.02	Beetles	19	18	0.5 hours	10 hours	Appendix S3
herbs_yal_1	38.48	-0.85	Shrubs	12	18	2.250	0.090	Appendix S3
herbs_yal_10	38.48	-0.85	Shrubs	12	24	2.250	0.090	Appendix S3
herbs_yal_11	38.18	-1.24	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_yal_12	38.18	-1.24	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_yal_13	38.18	-1.24	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_yal_14	38.18	-1.24	Shrubs	12	11	2.250	0.090	Appendix S3
herbs_yal_15	38.18	-1.25	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_yal_16	38.48	-0.84	Shrubs	12	24	2.250	0.090	Appendix S3
herbs_yal_17	37.86	-1.87	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_yal_18	37.86	-1.87	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_yal_19	37.86	-1.87	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_yal_2	38.49	-1.62	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_yal_20	38.47	-0.83	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_yal_21	37.86	-1.87	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_yal_22	37.13	-2.05	Shrubs	12	17	2.250	0.090	Appendix S3
herbs_yal_23	37.13	-2.05	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_yal_24	37.13	-2.06	Shrubs	12	16	2.250	0.090	Appendix S3
herbs_yal_25	37.13	-2.06	Shrubs	12	11	2.250	0.090	Appendix S3
herbs_yal_26	37.13	-2.06	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_yal_27	37.09	-2.09	Shrubs	12	17	2.250	0.090	Appendix S3
herbs_yal_28	37.09	-2.09	Shrubs	12	17	2.250	0.090	Appendix S3
herbs_yal_29	37.09	-2.09	Shrubs	12	17	2.250	0.090	Appendix

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
								S3
herbs_yal_3	38.49	-1.62	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_yal_30	38.47	-0.83	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_yal_31	37.09	-2.09	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_yal_32	38.47	-0.83	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_yal_33	38.47	-0.83	Shrubs	12	11	2.250	0.090	Appendix S3
herbs_yal_34	38.45	-0.78	Shrubs	12	18	2.250	0.090	Appendix S3
herbs_yal_35	38.49	-1.62	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_yal_4	38.49	-1.62	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_yal_5	38.44	-0.78	Shrubs	12	22	2.250	0.090	Appendix S3
herbs_yal_6	38.44	-0.78	Shrubs	12	17	2.250	0.090	Appendix S3
herbs_yal_7	38.25	-0.91	Shrubs	12	21	2.250	0.090	Appendix S3
herbs_yal_8	38.26	-0.91	Shrubs	12	16	2.250	0.090	Appendix S3
herbs_yal_9	38.18	-1.24	Shrubs	12	11	2.250	0.090	Appendix S3
herbs_ym_1	40.14	-3.62	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_10	40.22	-3.58	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_11	40.19	-3.59	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_12	40.21	-3.59	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_13	40.21	-3.59	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_ym_14	39.82	-3.43	Shrubs	12	16	2.250	0.090	Appendix S3
herbs_ym_15	39.80	-3.43	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_ym_16	39.85	-3.33	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_17	40.05	-3.05	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_ym_18	40.06	-3.05	Shrubs	12	17	2.250	0.090	Appendix S3
herbs_ym_19	40.06	-3.06	Shrubs	12	18	2.250	0.090	Appendix S3
herbs_ym_2	40.15	-3.43	Shrubs	12	13	2.250	0.090	Appendix S3

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
herbs_ym_20	40.07	-3.07	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_ym_21	40.08	-3.07	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_22	40.08	-3.05	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_ym_23	40.09	-3.02	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_ym_24	40.13	-2.99	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_25	40.13	-2.99	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_ym_26	40.02	-3.13	Shrubs	12	10	2.250	0.090	Appendix S3
herbs_ym_27	40.01	-3.38	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_28	39.85	-3.63	Shrubs	12	11	2.250	0.090	Appendix S3
herbs_ym_29	39.87	-3.63	Shrubs	12	13	2.250	0.090	Appendix S3
herbs_ym_3	40.14	-3.62	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_30	40.24	-3.08	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_31	40.25	-3.05	Shrubs	12	16	2.250	0.090	Appendix S3
herbs_ym_32	40.20	-3.14	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_ym_33	40.02	-3.11	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_ym_34	40.17	-3.08	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_ym_35	40.33	-2.93	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_36	40.37	-2.98	Shrubs	12	14	2.250	0.090	Appendix S3
herbs_ym_4	40.13	-3.62	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_5	40.12	-3.62	Shrubs	12	12	2.250	0.090	Appendix S3
herbs_ym_6	40.15	-3.43	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_7	40.18	-3.42	Shrubs	12	20	2.250	0.090	Appendix S3
herbs_ym_8	40.27	-3.14	Shrubs	12	15	2.250	0.090	Appendix S3
herbs_ym_9	40.24	-3.06	Shrubs	12	15	2.250	0.090	Appendix S3
heustobm	39.55	-84.72	Trees	10	20	100.000	0.100	25
heustomf	39.55	-84.72	Trees	10	24	100.00	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
						0		
huamani	-0.67	-77.67	Trees	10	162	100.000	0.100	25
humboldt	-8.83	-75.00	Trees	10	158	100.000	0.100	25
incahuar	-15.92	-67.58	Trees	10	160	100.000	0.100	25
indiana	-3.52	-73.70	Trees	10	223	100.000	0.100	25
indianca	40.50	-95.72	Trees	10	25	100.000	0.100	25
jatunsac	-1.70	-77.60	Trees	10	244	100.000	0.100	25
jauneche	-1.10	-79.63	Trees	10	96	100.000	0.100	25
jejuimi	-24.13	-55.53	Trees	10	87	100.000	0.100	25
jenarohe	-4.92	-73.75	Trees	10	245	100.000	0.100	25
jonesmil	40.80	-79.35	Trees	10	15	100.000	0.100	25
jura_tree	47.48	7.77	Trees	10	17	100.000	8.000	Appendix S3
kanealle	41.67	-78.80	Trees	10	13	100.000	0.100	25
kennedy	11.80	-74.20	Trees	10	67	100.000	0.100	25
kenting	21.92	120.83	Trees	10	69	100.000	0.100	25
khaoyai	14.33	101.83	Trees	10	145	100.000	0.100	25
kitlope1	53.70	-127.83	Trees	10	10	100.000	0.100	25
korup	5.73	8.70	Trees	10	133	100.000	0.100	25
lagenoa	-11.80	-75.42	Trees	10	112	100.000	0.100	25
laplanad	1.13	-77.97	Trees	10	130	100.000	0.100	25
laraya	8.33	-74.92	Trees	10	160	100.000	0.100	25
laselva	10.43	-84.20	Trees	10	136	100.000	0.100	25
lasjoyas	19.58	-104.13	Trees	10	39	100.000	0.100	25
laurelri	39.95	-79.37	Trees	10	14	100.000	0.100	25
linhares	-19.30	-40.70	Trees	10	220	100.000	0.100	25
luquillo	18.18	-65.83	Trees	10	42	100.000	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
madden	9.10	-79.60	Trees	10	131	100.000	0.100	25
madidi	-13.58	-68.77	Trees	10	198	100.000	0.100	25
madidiri	-13.58	-68.77	Trees	10	176	100.000	0.100	25
magsasay	10.40	-84.50	Trees	10	148	100.000	0.100	25
makokou1	0.57	12.87	Trees	10	138	100.000	0.100	25
makokou2	0.57	12.87	Trees	10	118	100.000	0.100	25
manaus	-3.13	-60.20	Trees	10	102	100.000	0.100	25
maquipuc	0.12	-78.62	Trees	10	137	100.000	0.100	25
mariquit	5.25	-74.83	Trees	10	94	100.000	0.100	25
martin	-39.50	-73.17	Trees	10	19	100.000	0.100	25
miazi	-4.30	-78.67	Trees	10	186	100.000	0.100	25
mirador	-40.23	-73.30	Trees	10	16	100.000	0.100	25
mishnfl	-3.78	-73.50	Trees	10	266	100.000	0.100	25
mishws	-3.78	-73.50	Trees	10	195	100.000	0.100	25
mogote	18.42	-66.25	Trees	10	47	100.000	0.100	25
montgome	43.17	-77.62	Trees	10	18	100.000	0.100	25
motozint	15.33	-92.20	Trees	10	16	100.000	0.100	25
mtcam	4.80	9.00	Trees	10	124	100.000	0.100	25
mtsthila	45.62	-73.58	Trees	10	12	100.000	0.100	25
mudumal1	11.60	76.53	Trees	10	15	100.000	0.100	25
mudumal2	11.60	76.53	Trees	10	19	100.000	0.100	25
murri	6.58	-76.83	Trees	10	179	100.000	0.100	25
nadugani	11.45	76.38	Trees	10	106	100.000	0.100	25
nanjensh	22.00	120.83	Trees	10	85	100.000	0.100	25
ndakani	2.37	16.15	Trees	10	96	100.000	0.100	25
ndakanni	2.37	16.32	Trees	10	122	100.00	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
						0		
newcaldo	-22.17	166.83	Trees	11	172	100.000	0.100	25
nosymang	-15.50	49.77	Trees	10	210	100.000	0.100	25
nuevomun	-10.65	-66.77	Trees	10	149	100.000	0.100	25
nwbranch	39.30	-77.30	Trees	10	20	100.000	0.100	25
omofor	7.00	5.00	Trees	10	72	100.000	0.100	25
osasiren	8.50	-83.92	Trees	10	144	100.000	0.100	25
palanan	17.13	122.52	Trees	10	144	100.000	0.100	25
parqueer	-24.58	-64.70	Trees	10	41	100.000	0.100	25
pasoh30	3.00	102.33	Trees	10	210	100.000	0.100	25
pasoh40	3.00	102.33	Trees	10	221	100.000	0.100	25
perinet	-18.92	48.42	Trees	10	214	100.000	0.100	25
perromue	-1.60	-80.70	Trees	10	69	100.000	0.100	25
persever	-14.63	-62.62	Trees	10	116	100.000	0.100	25
PHFIG_Spring	41.70	2.27	Ants	60	10	0.003	0.113	24
PHGRF_Spring	41.27	1.92	Ants	53	14	0.003	0.113	24
PHGRF_Summer	41.27	1.92	Ants	56	15	0.003	0.113	24
PHVAC_Spring	41.58	1.93	Ants	57	10	0.003	0.113	24
PHVAC_Summer	41.58	1.93	Ants	58	11	0.003	0.113	24
pipeline	9.17	-79.75	Trees	10	161	100.000	0.100	25
Barra_Paraguacu	-12.63	-38.80	Trees	49	60	100.000	0.490	Appendix S3
Bonito_Bahia	-11.99	-41.10	Trees	13	28	400.000	0.520	Appendix S3
Itubera_Bahia	-13.72	-39.20	Trees	29	155	100.000	0.290	Appendix S3
Maracas_Bahia	-13.47	-40.44	Trees	42	35	400.000	1.680	Appendix S3
Semidecidous_Bahia	-11.91	-41.25	Trees	31	64	100.000	0.310	Appendix S3
potomac	38.80	-76.57	Trees	11	25	100.000	0.100	25
providen	13.35	-81.37	Trees	10	65	100.000	0.100	25
pugu	-6.83	39.80	Trees	10	88	100.000	0.100	25

ID	Latitude	Longitude	Taxonor T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
puyehue	-40.72	-72.30	Trees	10	16	100.000	0.100	25
quiapaca	-18.33	-59.50	Trees	10	83	100.000	0.100	25
quinceoc	19.73	-104.25	Trees	10	47	100.000	0.100	25
ranchoqu	8.70	-83.55	Trees	10	124	100.000	0.100	25
riachuel	-27.00	-58.00	Trees	10	47	100.000	0.100	25
rioheath	-12.83	-68.83	Trees	10	135	100.000	0.100	25
riomanso	7.50	-76.80	Trees	11	220	100.000	0.100	25
rionegro	-9.83	-65.67	Trees	10	169	100.000	0.100	25
riopal1	-0.57	-79.33	Trees	10	119	100.000	0.100	25
riopal2	-0.57	-79.33	Trees	10	122	100.000	0.100	25
riotavar	-13.35	-69.67	Trees	10	200	100.000	0.100	25
roundslo	17.33	-77.42	Trees	10	45	100.000	0.100	25
roundtop	17.33	-77.42	Trees	10	57	100.000	0.100	25
ruissalo	60.53	22.47	Trees	10	10	100.000	0.100	25
sabanaru	10.50	-72.92	Trees	10	52	100.000	0.100	25
sacram	-16.30	-67.80	Trees	10	100	100.000	0.100	25
sakaera2	14.50	102.00	Trees	10	37	100.000	0.100	25
sakaerat	14.50	102.00	Trees	10	81	100.000	0.100	25
salta	-24.67	-64.50	Trees	10	25	100.000	0.100	25
sanfelas	29.68	-82.43	Trees	10	19	100.000	0.100	25
sansebas	-1.60	-80.70	Trees	10	96	100.000	0.100	25
santacru	-17.77	-63.70	Trees	10	63	100.000	0.100	25
santotom	4.92	-74.83	Trees	10	76	100.000	0.100	25
saul	3.63	-53.20	Trees	10	149	100.000	0.100	25
semengoh	1.83	110.80	Trees	10	224	100.000	0.100	25
shiringa	-10.33	-75.17	Trees	10	191	100.00	0.100	25

ID	Latitude	Longitude	Taxon or T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
						0		
SHRFIG_Spring	41.70	2.27	Ants	58	16	0.003	0.113	24
SHRFIG_Summer	41.70	2.27	Ants	56	13	0.003	0.113	24
SHRGRF_Spring	41.27	1.92	Ants	51	11	0.003	0.113	24
SHRVAC_Spring	41.58	1.93	Ants	54	11	0.003	0.113	24
SHRVAC_Summer	41.58	1.93	Ants	53	10	0.003	0.113	24
Siemianice1	54.50	17.05	Beetles	52	27	300 cm3	1.000	5
Siemianice2	54.50	17.05	Beetles	49	26	300 cm3	1.000	5
Siemianice3	54.50	17.05	Beetles	27	26	300 cm3	1.000	5
Siemianice4	54.50	17.05	Beetles	27	26	300 cm3	1.000	5
Siemianice5	54.50	17.05	Beetles	63	27	300 cm3	1.000	5
Siemianice6	54.50	17.05	Beetles	55	26	300 cm3	1.000	5
Siemianice7	54.50	17.05	Beetles	33	27	300 cm3	1.000	5
Siemianice8	54.50	17.05	Beetles	47	26	300 cm3	1.000	5
sierraro	22.83	-83.00	Trees	11	47	100.000	0.100	25
sietecue	4.58	-73.67	Trees	10	77	100.000	0.100	25
suderhac	54.60	9.28	Trees	10	15	100.000	0.100	25
tahuampa	-3.78	-73.50	Trees	10	169	100.000	0.100	25
tamblat2	-12.78	-69.28	Trees	10	158	100.000	0.100	25
tambo	-12.78	-69.28	Trees	10	156	100.000	0.100	25
tamboall	-12.83	-69.28	Trees	10	187	100.000	0.100	25
tambupl	-12.82	-69.72	Trees	10	131	100.000	0.100	25
tarapoto	-6.58	-76.42	Trees	10	102	100.000	0.100	25
tayrona	11.33	-74.30	Trees	10	68	100.000	0.100	25
tidroute	41.70	-79.40	Trees	10	20	100.000	0.100	25
tikus_1	-5.86	106.58	Corals	10	54	30.000	0.040	31
tikus_2	-5.86	106.58	Corals	10	30	30.000	0.040	31
tikus_3	-5.86	106.58	Corals	10	29	30.000	0.040	31

ID	Latitude	Longitude	Taxon or T. level	Nº samples	Nº species	Grain (m^2)	Extent (ha)	Reference
tikus_4	-5.86	106.58	Corals	10	35	30.000	0.040	31
tikus_5	-5.86	106.58	Corals	10	26	30.000	0.040	31
tikus_6	-5.86	106.58	Corals	10	23	30.000	0.040	31
tikus_7	-5.86	106.58	Corals	10	25	30.000	0.040	31
trunk_alcornoque	35.27	-6.07	Bryophytes	13	14	0.040	0.013	32
trunk_encina_a	32.51	-6.03	Bryophytes	12	11	0.040	0.031	27
trunk_encina_b	33.74	-4.29	Bryophytes	12	11	0.040	0.031	27
trunk_loro	34.96	-4.67	Bryophytes	15	17	0.040	0.013	32
tutunend	5.77	-76.58	Trees	10	274	100.000	0.100	25
tuxtlas	18.58	-95.13	Trees	10	102	100.000	0.100	25
tysongla	38.50	-90.52	Trees	10	26	100.000	0.100	25
tysonwoo	38.50	-90.52	Trees	10	23	100.000	0.100	25
uchire	10.15	-65.43	Trees	10	67	100.000	0.100	25
ucumari	4.00	-75.50	Trees	10	110	100.000	0.100	25
ufhortic	29.67	-82.33	Trees	10	32	100.000	0.100	25
uppsala	59.85	17.63	Trees	10	13	100.000	0.100	25
valleyvi	38.25	-90.62	Trees	10	23	100.000	0.100	25
varirata	-9.50	147.50	Trees	10	210	100.000	0.100	25
vencer	-5.75	-77.67	Trees	10	164	100.000	0.100	25
yanaigua	-19.70	-62.10	Trees	10	32	100.000	0.100	25
yanam1	-3.43	-72.85	Trees	10	216	100.000	0.100	25
yanam2	-3.43	-72.85	Trees	10	228	100.000	0.100	25
yanamtah	-3.47	-72.83	Trees	10	164	100.000	0.100	25

Table S1. Geographic coordinates (Latitude and Longitude), taxon or trophic level, number of samples, number of species and grain (m^2) and extent (ha)

for the assemblage datasets (ID) analyzed in the study. Notice that some coordinates were approximated from locations described in original sources.

Table S2.

Dependent	Explanatory	Coeffs.	z-value	p-value	Dev. expl.
Connectivity (P<N)					0.06
	Plant vs Animal	-0.773	-0.599	0.549	
	Latitude	-0.749	-1.581	0.114	
	Longitude	0.365	0.918	0.359	
	Species richness	0.328	0.5	0.617	
	Sampling effort	0.029	0.083	0.934	
	Null model d.f.	-0.646	-0.922	0.357	
Abundance-degree (P<N)					0.24
	Plant vs Animal	-19.518	-0.001	0.999	
	Latitude	-0.681	-1.105	0.269	
	Longitude	-0.491	-0.715	0.475	
	Species richness	2.448	1.893	0.058	
	Sampling effort	183.519	0.01	0.992	
	Null model d.f.	-2.541	-1.544	0.123	
Abundance (P<N)					0.36
	Plant vs Animal	-14.376	-0.009	0.993	
	Latitude	-0.063	-0.062	0.95	
	Longitude	-0.391	-0.413	0.679	
	Species richness	10.984	1.404	0.16	
	Sampling effort	6.581	1.168	0.243	
	Null model d.f.	1.902	1.083	0.279	
Modularity (P>N)					0.37
	Plant vs Animal	1.659	1.579	0.114	
	Latitude	0.324	0.429	0.668	
	Longitude	0.541	0.827	0.408	
	Species richness	15.408	3.175	0.015	
	Sampling effort	0.195	0.481	0.631	
	Null model d.f.	0.755	0.684	0.494	

Table S2. The differences between positive and negative network properties were in general unaffected by sampling effort (estimated as the number of samples), null model degrees of freedom (see Appendix S1.3), species richness, latitude, longitude or taxa. Generalized linear model summary statistics including explained deviance (Dev. expl.) for each model. Connectivity (P<N): Probability of negative networks to be more densely connected than their positive pairs. Abundance-degree (P<N): Probability of dominant species to monopolizing negative links but not positive ones (i.e., a stronger positive abundance-degree relationship in negative networks). Abundance (P<N): probability of positive networks tending to be composed of less abundant species.

Modularity (P>N): probability of positive networks being more modular than their negative pairs.

Table S3.

Feature	H_1	Null model I (n=326)			Null model II (n=305)			Null model III (n=252)		
		t-value	p-value	Prop. cases	t-value	p-value	Prop. cases	t-value	p-value	Prop. cases
Connectivity	**P < N**	17.006	<0.001	0.932	18.686	<0.001	0.905	11.155	<0.001	0.865
Degree vs abundance	**P < N**	23.881	<0.001	0.859	10.938	<0.001	0.702	7.479	<0.001	0.512
Abundance	**P < N**	22.425	<0.001	0.971	16.538	<0.001	0.869	10.306	<0.001	0.778
Modularity	**P > N**	39.676	<0.001	0.911	32.343	<0.001	0.951	14.272	<0.001	0.849

Table S3. Summary of results from different null model approaches. Results are based on the comparison of (i) connectivity, (ii) abundance-degree relationship, (iii) species abundance distribution and (iv) modularity between positive association networks (P) and their negative pairs (N). H_1: Alternative hypothesis in paired Student's tests. Prop. Cases: the proportion of assemblages where the alternative hypothesis was observed. The number of assemblages analysed (i.e. those showing at least two links in both types of networks) is also provided for each null model. P: Positive association networks. N: Negative association networks.

References

1. C. Song, R. P. Rohr, S. Saavedra, Why are some plant–pollinator networks more nested than others?. *J. Anim. Ecol.* ***86***, 1417-1424 (2017).
2. V. D. Blondel, et al. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
3. M. E. Newman, Modularity and community structure in networks. *Proc. Nat. Acad. Sci.* U.S.A. **103**, 8577-8582 (2006).
4. J. Oksanen, et al. Vegan: Community Ecology Package. R package version 2.4-3 (2017). https://CRAN.R-project.org/package=vegan.
5. W. Ulrich, N. J. Gotelli, Null model analysis of species associations using abundance data. *Ecology* **91**, 3384-3397 (2010).
6. S. Kembel, et al. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463-1464 (2010).
7. P. Kareiva, Population Dynamics in Spatially Complex Environments: Theory and Data. *Philos. Trans. R. Soc. Lond. B Biol. Sci. B* **330**, 175-190 (1990).
8. S. P. Hubbell, *The unified neutral theory of biodiversity and biogeography* (Princeton University Press, 2001), vol. 32.
9. M. Cadotte, et al. The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecol. Lett.* **16**, 1234-1244 (2013).

10. B. Kerr, M. A. Riley, M. W. Feldman, B. J. Bohannan, Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature* **418**, 171-174 (2002).

11. S. Allesina, J. M. Levine, A competitive network theory of species diversity. *Proc. Nat. Acad. Sci.* **108**, 5638-5642 (2011).

12. L. Gallien, et al. The effects of intransitive competition on coexistence. *Ecol. Lett.* **20**, 791-800 (2017).

13. L. S. Comita, et al. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**, 330-332 (2010).

14. G. Yenni, P. B. Adler, S. M. Ernest, Strong self-limitation promotes the persistence of rare species. *Ecology* **93**, 456-461 (2012).

15. S. Soliveres, et al. A missing link between facilitation and plant species coexistence: nurses benefit generally rare species more than common ones. *J. Ecol.* **103**, 1183-1189 (2015).

16. S. P. Hart, J. Usinowicz, J. M. Levine, The spatial scales of species coexistence. *Nat. Ecol. Evol.* **1**, 1066 (2017).

17. N. J. Gotelli, & W. Ulrich, Statistical challenges in null model analysis. *Oikos*, **121**, 171-180 (2012).

18. T. W. Schoener, The Anolis lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49**, 704-726 (1968).

19. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2017).

20. G. Csardi, T. Nepusz, The igraph software package for complex network research, *InterJournal, Complex Systems* **1695**, 1-9 (2006).

21. R. Durrett, S. Levin, Spatial aspects of interspecific competition. *Theor. Popul. Biol.* **53**, 30-43 (1998).

22. P. Amarasekare, Competitive coexistence in spatially structured environments: a synthesis. *Ecol. Lett.* **6**, 1109-1122 (2003).

23. M. Vellend, Conceptual synthesis in community ecology. *Q. Rev. Biol.* **85**, 183-206 (2010).

24. X. Arnan, N. Blüthgen, Using ecophysiological traits to predict climatic and activity niches: lethal temperature and water loss in Mediterranean ants. *Glob. Ecol. Biogeogr.* **24**, 1454-1464 (2015).

25. O. Phillips, J. S. Miller, *Global patterns of plant diversity: Alwyn H. Gentry's forest transect data set* (Missouri Botanical Press, 2002), Vol. 89.

26. X. Arnan, C. Gaucherel, A .N. Andersen, Dominance and species co-occurrence in highly diverse ant communities: a test of the interstitial hypothesis and discovery of a competition cascade. *Oecologia* **166**, 783-794 (2011).

27. I. Draper, F. Lara, B. Albertos, R. Garilleti, V. Mazimpaka, Epiphytic bryoflora of the Atlas and Antiatlas mountains, including a synthesis on the distribution of the epiphytic bryophytes in Morocco. *J. Bryol.* **28**, 312-330 (2006).

28. N. G. Medina, et al. Epiphytic bryophytes of forests in Central and North inland Iberian Peninsula. *Front. Biogeogr.* **7**, 21-28 (2015).

29. J. Calatayud, J. Madrigal-González, E. Gianoli, J. Hortal, A. Herrero, Uneven abundances determine nestedness in climbing plant-host interaction networks. *Perspect. Plant. Ecol.* **26**, 53-59 (2017).

30. J. Madrigal-González, et al. Facilitation of the non-native annual plant *Mesembryanthemum crystallinum* (Aizoaceae) by the endemic cactus *Eulychnia acida* (Cactaceae) in the Atacama Desert*. Biol. Invasions* **15**, 1439-1447 (2013).

31. B. E. Brown, Damage and recovery of coral reefs affected by El Niño related seawater warming in the Thousand Islands, Indonesia. *Coral Reefs* **8**, 163-170 (1990).

32. I. Draper, V. Mazimpaka, B. Albertos, R. Garilleti, F. Lara, A survey of the epiphytic bryophyte flora of the Rif and Tazzeka mountains (northern Morocco). *J. Bryol.* **27**, 23-34 (2005).