

Positive associations among rare species and their persistence in ecological assemblages

Joaquín Calatayud^{1,2*}, Enrique Andivia^{3,4}, Adrián Escudero⁵, Carlos J. Melián⁶, Rubén Bernardo-Madrid⁷, Markus Stoffel^{8,9,10}, Cristina Aponte¹¹, Nagore G. Medina^{12,13}, Rafael Molina-Venegas¹⁴, Xavier Arnan¹⁵, Martin Rosvall¹⁶, Magnus Neuman¹, Jorge Ari Noriega¹⁷, Fernanda Alves-Martins², Isabel Draper¹³, Arantzazu Luzuriaga⁵, Juan Antonio Ballesteros-Cánovas^{8,10}, César Morales-Molino^{16,17}, Pablo Ferrandis¹⁸, Asier Herrero^{13,19}, Luciano Pataro¹³, Leandro Juen²⁰, Alex Cea²¹ and Jaime Madrigal-González¹⁰

According to the competitive exclusion principle, species with low competitive abilities should be excluded by more efficient competitors; 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. On the other hand, 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 probable 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 assemblages¹. However, despite decades of intense investigation, the general mechanisms behind the persistence of these species are 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 partitioning, intraspecific competition, facilitation, indirect and neutral interactions^{4–9}, among others. Yet, thus far, there is no consensus to explain rare species persistence across taxa and environmental conditions.

The spatial arrangement of individuals plays a crucial role in 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 mechanisms¹². 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, 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.

In the present study, 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 assemblages²¹ that meet the following criteria: (1) each assemblage comprises taxa from only one trophic guild, thereby excluding the possibility that species associations result from direct predation or parasitism; (2) each assemblage shows reduced

¹Integrated Science Laboratory, Department of Physics, Umeå University, Umeå, Sweden. ²Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales, Madrid, Spain. ³Departamento de Ciencias de la Vida, Edificio de Ciencias, Universidad de Alcalá, Madrid, Spain. ⁴Departamento de Biodiversidad, Ecología y Evolución, Universidad Complutense de Madrid, Madrid, Spain. ⁵Departamento de Biología, Geología, Física y Química Inorgánica, Universidad Rey Juan Carlos, Madrid, Spain. ⁶Department of Fish Ecology and Evolution, Eawag, Kastanienbaum, Switzerland. ⁷Departamento de Biología de la Conservación, Estación Biológica de Doñana-CSIC, Seville, Spain. ⁸Dendrolab, Department of Earth Sciences, University of Geneva, Geneva, Switzerland. ⁹Department F.-A. Forel for Environmental and Aquatic Sciences, University of Geneva, Geneva, Switzerland. ¹⁰Climate Change Impacts and Risks in the Anthropocene, Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland. ¹¹Plant Science, University of Melbourne, Burnley Campus, Richmond, Victoria, Australia. ¹²Department of Botany, Faculty of Sciences, University of South Bohemia, České Budějovice, Czech Republic. ¹³Departamento de Biología (Botánica), Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid, Spain. ¹⁴Plant Ecology, Institute of Plant Sciences, University of Bern, Bern, Switzerland. ¹⁵Centre de Recerca Ecològica i Aplicacions Forestals Campus de Bellaterra (UAB), Cerdanyola del Vallès, Spain. ¹⁶UMR CNRS 5805 EPOC, OASU, Université de Bordeaux Site de Talence-Pessac-Gradignan, Pessac, France. ¹⁷Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland. ¹⁸Higher Technical School of Agricultural and Forestry Engineering and Botanical Institute, University of Castilla-La Mancha, Albacete, Spain. ¹⁹Department of Plant Biology and Ecology, Pharmacy Faculty, University of Basque Country, Vitoria-Gasteiz, Spain. ²⁰Laboratório do Ecologia e Zoologia de Invertebrados, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém do Pará, Brazil. ²¹Departamento de Biología, Facultad de Ciencias, Universidad de La Serena, La Serena, Chile. *e-mail: j.calatayud.ortega@gmail.com

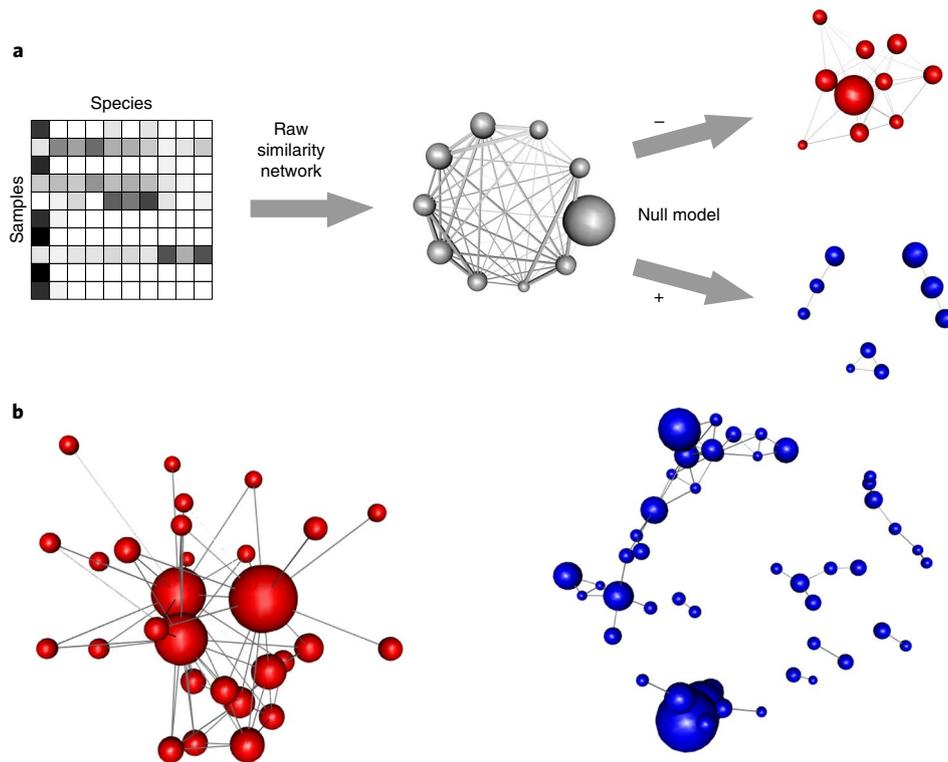


Fig. 1 | Approaching assembly mechanisms through the lens of positive and negative association networks. **a**, Species segregations and aggregations can inform on the main mechanisms underlying ecological assemblages. These spatial patterns are measured between species pairs using the similarity in 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 the association networks (blue nodes). Species pairs whose individuals are more segregated than randomly expected receive a negative link in the association networks (red nodes). **b**, Positive (blue) and negative (red) networks of a tropical rainforest tree assemblage (see 'Barra_Paraguacu' in Supplementary Table 1). 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 algorithm³⁷.

spatial extent and low environmental heterogeneity, to increase the probability that species associations are mainly due to biotic interactions; (3) the abundance of at least ten species is recorded in a minimum of ten samples, to improve statistical power (Supplementary Appendix 1); (4) the dataset represents a wide variety of biomes (for example, tropical forests, deserts, temperate steppes and polar climates), thus avoiding biome-specific results; and (5) 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 model²². Species pairs that significantly deviated from random expectations received 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 literature², one would expect species to be more segregated than aggregated in natural systems. If so, negative networks should be more densely connected (that is, more links per species) than their positive counterparts. In accordance, negative networks were more connected than their positive pairs in a notable 93.2% of all assemblages ($t = 17.01$, $P < 0.001$; Fig. 2a). Differences in connectivity remained similar after accounting for differences in network size ($t = -16.81$ and $P < 0.001$ for 78.8% of the assemblages) or when calculating differences in the average number of links (that is, average species degree; $t = -14.69$ and $P < 0.001$ for 69.0% of

the assemblages). Furthermore, if abundance is considered to be an expression of the species' competitive abilities²³, 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 $\rho = 0.65$, s.d. = 0.23), but a weak or even negative correlation in positive networks (mean $\rho = 0.02$, s.d. = 0.38), with differences between networks being statistically significant ($t = -23.88$, $P < 0.001$; Fig. 2b). Moreover, we found evidence showing that a particular species is more often involved in negative associations when it becomes abundant (Supplementary Appendix 2). 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 ($t = 22.42$, $P < 0.001$; Fig. 2c). However, such spatial aggregations 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 ($t = 39.68$, $P < 0.001$; Fig. 2d). This result remained similar after accounting for network size and connectivity ($t = 11.31$ and $P < 0.001$ for 67.3% of assemblages). Moreover, while

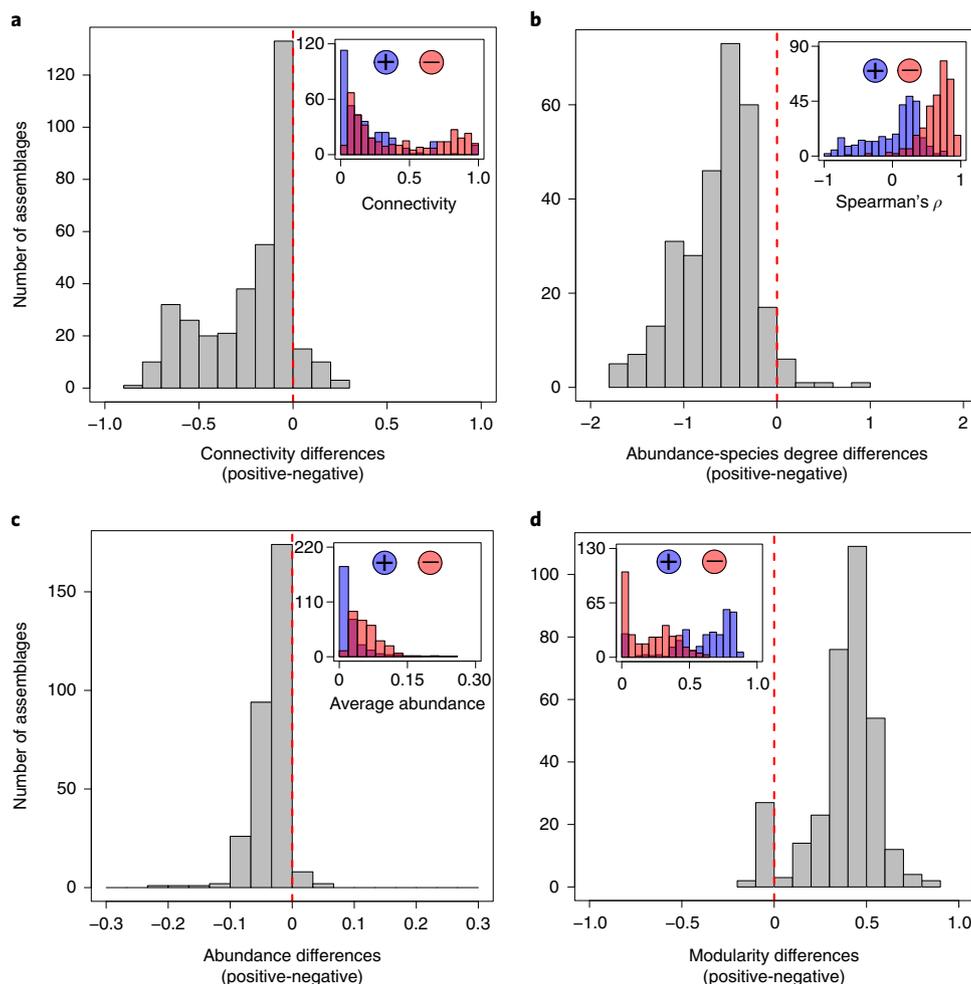


Fig. 2 | The contrasting patterns of positive and negative association networks. **a**, The higher connectivity of negative networks indicates that species segregation dominates over species aggregation. **b**, These segregations are monopolized by dominant species, as shown by the strong relationship between abundance and species degree (that is, number of links of a species) in negative networks. **c,d**, In contrast, less abundant species are more prone to generate positive associations (**c**), although these associations only occur among specific groups of rare species, as indicated by the higher modularity of positive networks (**d**). The main histograms show the differences in network features between pairs of positive and negative networks. The insets show the raw values for both types of networks, where the purple colour represents the overlap between both distributions.

60.7% of positive networks were significantly modular, only 13.8% of negative ones showed this pattern. 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 are largely invariant regardless of different probability thresholds to detect significant associations, the use of quantitative links and assumptions of disparate null models (Supplementary Appendix 1). This robust and conspicuous spatial organization suggests 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 competitors¹⁴ may generate this pattern, also enhancing rare species persistence. Indeed, numerical simulations 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 Supplementary Appendix 3). However, this mainly occurs when habitat preferences are strong, a situation that should arise under marked environmental gradients most probably far from the reality of the fine-scaled assemblages studied in the present work. Complementarily, positive interactions

within groups of rare species may also contribute and/or generate these modular positive networks. Moreover, this may increase the persistence of weak competitors since, just as in harsh abiotic environments²⁴, 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. Our simulations also reveal that this hypothesis most probably reproduces the observed patterns in association networks compared to other stabilizing mechanisms, such as neutral colonization-extinction dynamics⁹, intransitive competition^{25,26}, differential density-dependent effects^{27,28} or facilitation by nurse species⁷ (Fig. 3b). Interestingly, the combination of habitats and positive interactions yields the highest probability of reproducing the observed network patterns (Fig. 3a). This further suggests that even under strong differences in habitat preferences, stabilizing forces, such as facilitation or complementarity, would 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

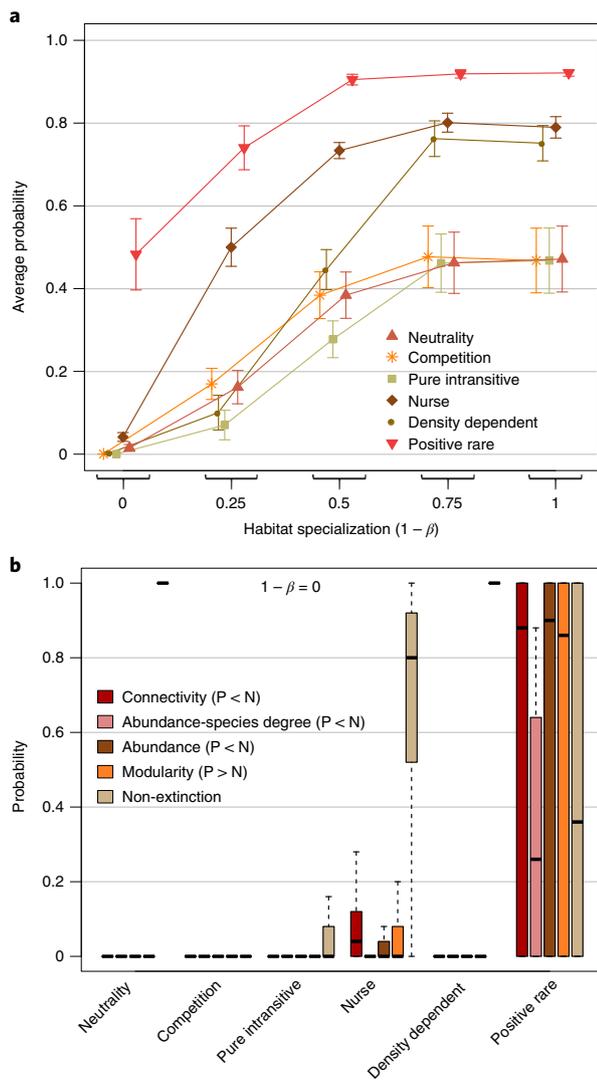


Fig. 3 | Positive interactions among weak competitors alone or together with habitat preferences reproduce realized association patterns. a, Dissimilarities in habitat preferences between dominant species and groups of rare species may generate the empirical patterns of association networks, regardless 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 specialization (see Methods and Supplementary Appendix 3). The error bars depict the confidence intervals at $\alpha = 0.05$. **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 probable model to generate simulated assemblages showing the same association networks as empirical assemblages (connectivity, Fig. 2a; abundance-species 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 Supplementary Appendix 3). In the boxplots, the centre line shows the median probabilities, being the box between the 25th and 75th percentiles and the whiskers at minimum and maximum probability once the outliers are discounted. Outliers are not shown. P, positive networks; N, negative networks.

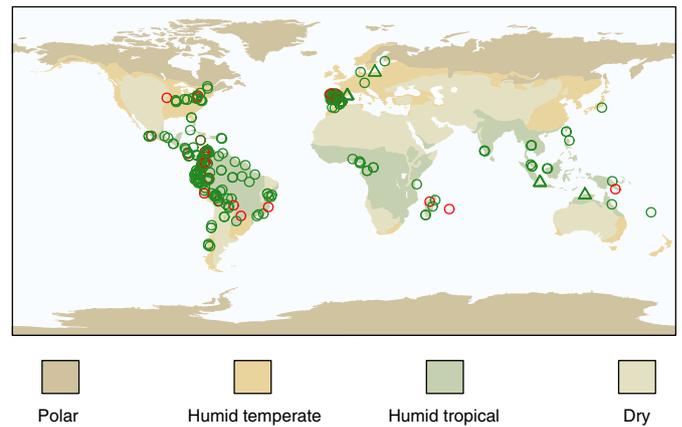


Fig. 4 | The organization of association networks remains invariant across the globe and regardless of taxa. The circles and triangles represent plant and animal assemblages, respectively. Green depicts assemblages where positive networks were both composed of less abundant species and more modular than their negative counterparts, whereas red shows assemblages where these patterns were not found. The map colours represent the Earth climatic zones as proposed by Bailey³⁸ (available at www.unep-wcmc.org; accessed October 2017).

the persistence of these species, which agrees with recent experimental evidence²⁹.

Overall, our results show that ecological assemblages are consistently organized in positive and negative association networks across the main biological groups (that is, animals and plants) and geography (Fig. 4 and Extended Data Fig 1). This ubiquity sheds light on the long-standing diversity paradox since 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 feedback of these positive interactions and microhabitats. Nevertheless, the generality of the findings presented in this study bring us closer to understanding the assemblage of the vast biodiversity on Earth.

Methods

Data acquisition. Assemblage data were collected from published studies in peer-reviewed journals and our own surveys²¹ (Supplementary Appendix 4 and Supplementary Table 1). Each assemblage consists of at least ten samples where the abundance of at least ten species of the same trophic guild was recorded. To minimize the effects of environmental heterogeneity and dispersion on spatial patterns, we only included datasets that showed (1) low environmental variability across samples (excluding surveys where any kind of environmental gradient was reported or no clear information about it was provided), (2) a very reduced spatial extent (median = 0.1 ha, ranging from 0.01 to 25.6 ha), (3) a very small grain size to increase the probability of physical and/or chemical contact among all species in the samples (median = 100 m², ranging from 0.002 to 400 m², respectively), and (4) standardization among samples to avoid sampling biases (for example, 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 later in the Methods), 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 index³⁰:

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

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 (that is, row and column totals were 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 (Supplementary Appendix 1 and Supplementary Figs. 1 and 2). 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 (that is, type I errors) may be thought to be relatively high in species-rich assemblages due to multiple comparisons (however, see Rothman³¹). However, species pairwise similarities were compared against null values generated using a fixed-fixed assemblage-wise null model (that is, 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 methods³². Nevertheless, we used the same nominal error (that is, $\alpha = 0.05$) 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, the results remained largely constant when using different nominal errors (Supplementary Appendix 1 and Supplementary Fig. 1).

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 (that is, species' number of links) for each pair of network types. Connectivity is defined as the number of realized links relative to the number of potential links. This measure of connectivity may be negatively correlated with network size. Hence, we also used the residual connectivity obtained from the residuals of a linear regression between the number of observed and potential links, both log-transformed³³. 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 by calculating modularity with the index proposed by Newman³⁴ (Q) along with the optimization algorithm of Blondel et al.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 v.1 software³⁶). Then, we computed the 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 aforementioned 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 (Supplementary Appendix 1), latitude, longitude, taxonomic group (that is, animals or plants) and species richness. To do so, we first generated four binomial dependent variables, based on whether (1) the negative networks of each assemblage were more densely connected than their positive pairs; (2) the negative networks presented higher positive abundance-species degree relationships; (3) the positive networks tended to be composed of less abundant species; and (4) 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 10 species whose individuals were randomly distributed at the outset. Individuals reproduce, colonize 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 (Supplementary Appendix 3). That is, individuals of dominant species reduce the probability of reproduction and colonization, while increasing the death probability, of co-occurring individuals

of subordinate species. Benefactor individuals have the opposite effects on the probabilities of beneficiary individuals (Supplementary Fig. 3a,b).

We further incorporated the effects of dissimilarities in habitat preferences by setting four habitats preferred by different groups of species (Supplementary Fig. 3c). Specifically, the probabilities of reproduction, survival (that is, 1 minus the death probability) and colonization in non-preferred habitats were multiplied by a habitat tolerance coefficient, β , ranging between 0 (null tolerance) and 1 (total tolerance; Supplementary Appendix 3). Hence, when $\beta = 0$, individuals are highly specialist and only allowed to reproduce, survive or colonize in the preferred habitat, whereas $\beta = 1$ corresponds to a neutral habitat scenario.

We ran simulations following six hypotheses explaining species assembly and coexistence (Fig. 3 and Supplementary Appendix 3): (1) a neutral interaction model, where all species were ecologically equivalent⁶; (2) a hierarchical competition model with one strong competitor; (3) an intraspecific density-dependent model, where superior competitors suffer more from intraspecific competition⁵; (4) a model incorporating intransitive competition²⁶, where the superior competitor is outcompeted by three species, which, in turn, outcompete all species except specific pairs (that is, theoretically promoting the generation of empirical association patterns; see Supplementary Appendix 3); (5) a nurse model⁷ with four superior competitors, three of which facilitate different pairs of subordinate species; and (6) 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 Supplementary Appendix 3 and Supplementary Fig. 4).

Simulations were run using a wide range of combinations ($n = 216$) where demographic rates (that is, reproduction, mortality and dispersal) had different relative importance (Supplementary 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 (that is, differences between positive and negative networks in connectivity, abundance-species degree relationship, abundance and modularity), as well as the probability of persistence of all species (that is, 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 the expectations from the competitive exclusion principle (that is, weak competitors went extinct; $P_{\text{non-extinction}} = 0$; see Supplementary Appendix 3 and Supplementary Fig. 5).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The dataset used in this study is freely available at <https://doi.org/10.6084/m9.figshare.9906092>.

Code availability

The R scripts used in this study are freely available at <https://doi.org/10.6084/m9.figshare.9906092>.

Received: 22 February 2019; Accepted: 31 October 2019;

Published online: 16 December 2019

References

- Gaston, K. J. *Rarity* (Chapman and Hall, 1994).
- Hardin, G. The competitive exclusion principle. *Science* **131**, 1292–1297 (1960).
- Hutchinson, G. E. The paradox of the plankton. *Am. Nat.* **95**, 137–145 (1961).
- Schoener, T. W. Resource partitioning in ecological communities. *Science* **185**, 27–39 (1974).
- Yenni, G., Adler, P. B. & Ernest, S. K. M. Strong self-limitation promotes the persistence of rare species. *Ecology* **93**, 456–461 (2012).
- Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
- 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).
- Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213 (2017).
- Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).
- Durrett, R. & Levin, S. Spatial aspects of interspecific competition. *Theor. Popul. Biol.* **53**, 30–43 (1998).
- McIntire, E. J. & Fajardo, A. Beyond description: the active and effective way to infer processes from spatial patterns. *Ecology* **90**, 46–56 (2009).
- 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**, 1066–1073 (2017).
15. Chacón-Labela, J., de la Cruz, M., Escudero, A. & Gomez-Aparicio, L. 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. Faisal, A., Dondelinger, F., Husmeier, D. & Beale, C. M. 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. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* **6**, 343–351 (2012).
20. Borthagaray, A. I., Arim, M. & Marquet, P. A. Inferring species roles in metacommunity structure from species co-occurrence networks. *Proc. Biol. Sci.* **281**, 20141425 (2014).
21. Calatayud, J. et al. Positive associations among rare species and their persistence in ecological assemblages. *figshare* <https://doi.org/10.6084/m9.figshare.9906092> (2019).
22. Ulrich, W. & Gotelli, N. J. Null model analysis of species associations using abundance data. *Ecology* **91**, 3384–3397 (2010).
23. Tilman, D. Resource competition between plankton algae: an experimental and theoretical approach. *Ecology* **58**, 338–348 (1977).
24. Callaway, R. M. et al. Positive interactions among alpine plants increase with stress. *Nature* **417**, 844–848 (2002).
25. Allesina, S. & Levine, J. M. A competitive network theory of species diversity. *Proc. Natl Acad. Sci. USA* **108**, 5638–5642 (2011).
26. Gallien, L., Zimmermann, N. E., Levine, J. M. & Adler, P. B. The effects of intransitive competition on coexistence. *Ecol. Lett.* **20**, 791–800 (2017).
27. Comita, L. S., Muller-Landau, H. C., Aguilar, S. & Hubbell, S. P. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**, 330–332 (2010).
28. Cody, M. L. & Diamond, J. M. *Ecology and Evolution of Communities* (Harvard Univ. Press, 1975).
29. 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).
30. Schoener, T. W. The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49**, 704–726 (1968).
31. Rothman, K. J. No adjustments are needed for multiple comparisons. *Epidemiology* **1**, 43–46 (1990).
32. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188 (2001).
33. de Araújo, W. S., Vieira, M. C., Lewinsohn, T. M. & Almeida-Neto, M. Contrasting effects of land use intensity and exotic host plants on the specialization of interactions in plant-herbivore networks. *PLoS ONE* **10**, e0115606 (2015).
34. Newman, M. E. Modularity and community structure in networks. *Proc. Natl Acad. Sci. USA* **103**, 8577–8582 (2006).
35. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
36. Colomer-de-Simón, P., Serrano, M. Á., Beiró, M. G., Alvarez-Hamelin, J. I. & Boguñá, M. Deciphering the global organization of clustering in real complex networks. *Sci. Rep.* **3**, 2517 (2013).
37. Fruchterman, T. M. J. & Reingold, E. M. Graph drawing by force-directed placement. *Softw. Pract. Exp.* **21**, 1129–1164 (1991).
38. Bailey, R. G. Explanatory supplement to ecoregions map of the continents. *Environ. Conserv.* **16**, 307–309 (1989).

Acknowledgements

We thank J. Hortal and S. 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 at HPC2N. J.C. is supported by the Carl Tryggers Foundation for Scientific Research (no. CTS 16:384). E.A. is supported by a postdoctoral grant (no. CT39/17) funded by the Universidad Complutense de Madrid. C.J.M. is supported by the Swiss National Science Foundation (grant no. SNSF-31003A-144162). R.B.-M. is supported by the Spanish Ministry of Science and Innovation predoctoral fellowship no. BES-2013-065753. M.S., J.A.B.-C. and J.M.-G. acknowledge support from the University of Geneva (project: C-CIA; no. 309354). X.A. is supported by a Ramón y Cajal research contract by the Spanish Ministry of Economy and Competitiveness (MINECO, no. RYC-2015-18448). M.R. is supported by the Swedish Research Council grant no. 2016-00796. J.A.N. was supported by a Colombian COLCIENCIAS doctoral scholarship (no. 617-2013). F.A.-M. is grateful to CAPES for a doctoral scholarship (no. 120147/2016-01). A.L., P.F. and J.M.-G. were funded by the AGORA Project (MINECO, no. CGL2016-77417-P). 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 2018 postdoctoral grant and Basque Country Government funding support to FisioclimaCO2 (IT1022-16) research group. L.J. received productivity grants from of CNPq (process no. 307597/2016-4).

Author contributions

J.C. and J.M.-G. conceived the study. 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. analysed the data with assistance from C.J.M., M.R. and M.N. J.C., E.A. and J.M.-G. led 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.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-019-1053-5>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-019-1053-5>.

Correspondence and requests for materials should be addressed to J.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

<i>Dependent</i>	<i>Explanatory</i>	<i>Coeffs.</i>	<i>z-value</i>	<i>p-value</i>	<i>Dev. expl.</i>
<i>Connectivity (P<N)</i>	Plant vs Animal	-0.773	-0.599	0.549	0.06
	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	
	<i>Abundance-degree (P<N)</i>	Plant vs Animal	-19.518	-0.001	
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	
<i>Abundance (P<N)</i>		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	
	<i>Modularity (P>N)</i>	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	

Extended Data Fig. 1 | The differences between positive and negative network properties were in general unaffected by sampling effort, null model degrees of freedom, 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 (that is, 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.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Simulated data was generated using own scripts in R software (version 3.3.3).

Data analysis

All analyses were conducted in R software (version 3.3.3), using widely recognized packages and own scripts.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The dataset used in this study is freely available at <https://doi.org/10.6084/m9.figshare.9906092>

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We analysed 385 sites X species abundances matrices at a local assemblage level. We also conducted numerical simulations including most recent hypotheses.
Research sample	We collected data on assemblages worldwide distributed and representing a wide array of taxa including bryophytes, tracheophytes, and arthropods from published studies and own surveys.
Sampling strategy	During the analysis of empirical data, we only collected assemblages with at least ten samples and ten species. We established this minimum matrix dimension (n=100) to ensure statistical power in subsequent null model analysis. Moreover, we found that our results were insensitive to null model degrees of freedom. During simulations, we computed 25 replicates for each of the 216 parameter combinations used. We selected this number of replicates and combinations to explore a wide space of possible parameter effects while accounting for intrinsic model stochasticity.
Data collection	The data was gathered from literature surveys and using specific sampling technics of each studied taxa. 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.
Timing and spatial scale	Literature surveys were conducted from March to September 2017. Field work was conducted in different years, seasons and with different duration depending on the studied taxa.
Data exclusions	No data were excluded from the analyses.
Reproducibility	No experiment was conducted.
Randomization	We split animal and plant assemblages due to they form the major taxonomical groups.
Blinding	Blinding was not relevant to our study.
Did the study involve field work?	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

Field work, collection and transport

Field conditions	Field conditions were extremely variable among assemblages, but as homogenous as possible within assemblages.
Location	Worldwide.
Access and import/export	Habitat accession and data collection was done in accordance to local and national laws for each survey.
Disturbance	Disturbance was minimised as much as possible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	The study did not involve laboratory animals
Wild animals	Dung beetles in central Spain were captured using 10 traps baited with cattle dung. Dung beetles in Sierra Nevada de Santa

Wild animals

Marta, Colombia, were trapped using 10 pitfall traps without any preservative solution and with aerial bait of fresh human dung. In both cases, individuals were killed in the field using alcohol to subsequent identification in the laboratory.

Field-collected samples

The study did not involved samples collected from then field.

Ethics oversight

No ethical approval was required

Note that full information on the approval of the study protocol must also be provided in the manuscript.