



OPEN

Consequences of spatial patterns for coexistence in species-rich plant communities

Thorsten Wiegand^{1,2}, Xugao Wang³, Kristina J. Anderson-Teixeira^{4,5}, Norman A. Bourg⁴, Min Cao⁶, Xiuqin Ci^{7,8}, Stuart J. Davies⁵, Zhanqing Hao^{3,9}, Robert W. Howe¹⁰, W. John Kress¹¹, Juyu Lian¹², Jie Li^{7,8}, Luxiang Lin⁶, Yiching Lin¹³, Keping Ma¹⁴, William McShea⁴, Xiangcheng Mi¹⁴, Sheng-Hsin Su¹⁵, I-Fang Sun¹⁶, Amy Wolf¹⁰, Wanhai Ye¹² and Andreas Huth^{1,2,17}

Ecology cannot yet fully explain why so many tree species coexist in natural communities such as tropical forests. A major difficulty is linking individual-level processes to community dynamics. We propose a combination of tree spatial data, spatial statistics and dynamical theory to reveal the relationship between spatial patterns and population-level interaction coefficients and their consequences for multispecies dynamics and coexistence. Here we show that the emerging population-level interaction coefficients have, for a broad range of circumstances, a simpler structure than their individual-level counterparts, which allows for an analytical treatment of equilibrium and stability conditions. Mechanisms such as animal seed dispersal, which result in clustering of recruits that is decoupled from parent locations, lead to a rare-species advantage and coexistence of otherwise neutral competitors. Linking spatial statistics with theories of community dynamics offers new avenues for explaining species coexistence and calls for rethinking community ecology through a spatial lens.

Understanding the mechanisms that maintain high species diversity in plant communities such as tropical forests has long challenged ecologists¹ and has stimulated major efforts in field and theoretical ecology^{2–5}. However, despite a multitude of coexistence mechanisms that have been proposed⁶ and recent advances in coexistence theories^{7–18}, this fundamental question has not been fully resolved^{8,9,14}. For example, theoretical models indicate that stable coexistence is difficult to reach in large communities^{11,13}. We argue that consideration of spatial patterns of plant individuals, such as intraspecific clustering and interspecific segregation, may allow for a better understanding of mechanisms of coexistence in species-rich communities¹⁷.

Although many studies suggest that spatial patterns and neighbourhood effects may play an important role in diversity maintenance^{17–21}, the integration of spatial patterns into coexistence theories of species-rich communities is difficult. A major difficulty is linking spatial processes at the individual level to community dynamics. One reason for this is a scale mismatch. The analytical models that form the basis of most coexistence theories^{7,8,11–16,22} have state variables that operate at the macroscale (that is, the population or community-level abundances), use parameters that describe average ‘mesoscale’ properties of the individuals (such as population-level interaction coefficients and demographic rates) and often rely on ‘mean-field’ approximations^{18,23} where spatial patterns are neglected.

However, spatial patterns and population-level interaction coefficients emerge at the mesoscale from the microscale behaviour of individuals and their interactions with other individuals and the environment. Therefore, studying the impact of spatial patterns on species coexistence requires multiscale approaches such as spatial moment equations^{18,23,24} that incorporate pattern-forming processes operating at the level of individuals and translate these into population and community dynamics.

We propose here such a multiscale approach. To this end, we first derive population-level interaction coefficients α_{fi} from individual-level interaction coefficients β_{fi} and neighbourhood crowding indices^{19,21} that are commonly used to describe interactions among tree individuals at the microscale, and then incorporate the emerging coefficients α_{fi} into analytical macroscale models. Our approach is based on separation of timescales (adiabatic approximation²⁵), given that mesoscale spatial patterns usually build up quickly and approach a quasi-steady state whereas the macroscale state variables (for example, abundances) change slowly²³. Therefore, we do not need to describe the dynamics of the quick mesoscale patterns explicitly (as, for example, is done in approaches based on moment equations^{18,23,24}) but concentrate instead on spatial patterns that transport the critical information from the microscale into macroscale models. This approach requires information on mesoscale spatial patterns that can be obtained from fully mapped forest plots

¹Department of Ecological Modelling, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany. ²German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. ³CAS Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences. ⁴Conservation Ecology Center, Smithsonian Conservation Biology Institute, Front Royal, VA, USA. ⁵Forest Global Earth Observatory (ForestGEO), Smithsonian Tropical Research Institute, Washington, DC, USA. ⁶CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences. ⁷Center of Conservation Biology, Core Botanical Gardens, Chinese Academy of Sciences. ⁸Centre for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences. ⁹School of Ecology and Environment, Northwestern Polytechnical University. ¹⁰Department of Natural and Applied Sciences, University of Wisconsin-Green Bay, Green Bay, WI, USA. ¹¹Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. ¹²Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences. ¹³Department of Life Science, Tunghai University. ¹⁴State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences. ¹⁵Taiwan Forestry Research Institute. ¹⁶Center for Interdisciplinary Research on Ecology and Sustainability, National Dong Hwa University. ¹⁷Institute of Environmental Systems Research, University of Osnabrück, Osnabrück, Germany. [✉]e-mail: thorsten.wiegand@ufz.de; wangxg@iae.ac.cn

such as those of the Forest Global Earth Observatory (ForestGEO) network⁴.

More specifically, we (1) derive species-level interaction coefficients from individual-level interactions using empirical information on spatial patterns in nine ForestGEO megaplots⁴, (2) integrate the resulting species-level interaction coefficients into analytical macroscale multispecies models and (3) study their consequences for multispecies dynamics and coexistence.

Results and discussion

Species-level interaction coefficients. We first derive species-level interaction coefficients from individual-level neighbourhood crowding indices^{19,21,26} that quantify how the performance of a focal individual depends on interactions with its neighbours. To this end, we describe the survival rate of a focal individual k of species f in dependence on the local number of neighbours as

$$s_{kf} = s_f \exp \left(-\beta_{ff} \left(n_{kff} + \underbrace{\sum_{i \neq f} (\beta_{fi}/\beta_{ff}) n_{kfi}}_{n_{kfp}} \right) \right), \quad (1a)$$

where s_f is a density-independent background survival rate of species f , the crowding indices n_{kfp} , n_{kfi} and n_{kfh} are the number of conspecifics, neighbours of species i and heterospecific neighbours within distance R of a focal individual k , respectively; the subscript ‘h’ indicates all heterospecifics together (that is, $n_{kfh} = \sum_{i \neq f} n_{kfi}$) and the crowding index n_{kfp} weights each heterospecific neighbour by its relative competition strength β_{fi}/β_{ff} (equation (1a)), with β_{fi} being the individual-level interaction coefficients between species f and i (Fig. 1a–c). The corresponding population-level survival rate is given by

$$\text{surv}_f = s_f \exp \left(-\alpha_{ff} \left(N_f(t) + \sum_{i \neq f} (a_{fi}/\alpha_{ff}) N_i(t) \right) \right), \quad (1b)$$

where $N_i(t)$ is the abundance of species i at time step t and α_{fi} is the population-level interaction coefficient between species f and i . To estimate surv_f we average the survival rates s_{kf} (equation (1a)) of all individuals k of species f :

$$\text{surv}_f = \frac{1}{N_f(t)} \sum_{k=1}^{N_f(t)} s_{kf} = s_f \int_x \int_y \exp(-\beta_{ff}(x+y)) p_x p_y dx dy \quad (2)$$

where p_x and p_y are the distributions of the crowding indices $x = n_{kff}$ and $y = n_{kfp}$ for individuals of species f , respectively.

To determine the distributions p_x and p_y , we analysed forest inventory data from nine 20–50 ha forest dynamics plots (Supplementary Table 1) in the ForestGEO network⁴. We used phylogenetic similarity between tree species as a surrogate for the relative competition strength β_{fi}/β_{ff} because it is available for the species in the nine plots (Methods). This is an established approach in species-rich communities^{19,26,27} to approximate niche differences in the absence of other data.

The number of con- and heterospecific neighbours and the heterospecific interaction index n_{kfp} vary widely among conspecifics and can be described by gamma distributions (Fig. 1 and Extended Data Figs. 1 and 2). Detailed analysis of the empirical crowding indices reveals additional relationships that are relevant for our subsequent analysis. First, we find that the crowding indices n_{kff} and n_{kfp} are not, or are only weakly, correlated for a given species f (Extended Data Fig. 3a). Second, we find for trees of a given species f high correlations between the two crowding indices n_{kff} and n_{kfp} (Extended

Data Fig. 3b) with a common factor B_f (that is, $n_{kfp} \approx B_f n_{kff}$). This result suggests operation of diffuse neighbourhood competition, in which the competition strength of heterospecifics is on average a factor B_f lower than that of conspecifics.

The integral of equation (2) can be solved analytically for independent gamma distributions p_x and p_y and yields

$$\text{surv}_f = s_f \exp \left(-\beta_{ff} \left(\gamma_{ff} \bar{n}_{ff} + \gamma_{fp} \bar{n}_{fp} \right) \right) \quad (3)$$

where \bar{n}_{ff} and \bar{n}_{fp} are the average values of the crowding indices n_{kff} and n_{kfp} , respectively, and γ_{ff} and γ_{fp} contain the variance-to-mean ratios of the gamma distributions p_x and p_y , respectively, but in our case have values close to one (Methods).

The last step in deriving pairwise population-level interaction coefficients is to relate the averages of the different crowding indices to the macroscale population abundances $N_f(t)$. We accomplish this by taking advantage of connections between crowding indices and the summary functions of spatial point process theory²¹. The mean of the crowding index n_{kff} (that is, the mean number of further conspecific neighbours within distance R) is proportional to Ripley’s K , a well-known quantity in point process theory^{28,29}:

$$\bar{n}_{ff} = K_{ff}(R) N_f(t)/A \quad (4a)$$

where $K_{ff}(R)$ is the univariate K function for species f and A is the area of the observation window. The K function describes the spatial pattern of conspecifics within a neighbourhood distance R , indicating clustering if $K_{ff}(R) > \pi R^2$, a random pattern if $K_{ff}(R) = \pi R^2$ and regularity if $K_{ff}(R) < \pi R^2$. In the following, we use the normalized K function $k_{ff}(R) = K_{ff}(R)/\pi R^2$ to quantify the spatial neighbourhood patterns, and therefore $\bar{n}_{ff} = k_{ff}(R) \frac{\pi R^2}{A} N_f(t)$.

Analogously, the mean number of heterospecific neighbours is given by

$$\bar{n}_{fh} = k_{fh}(R) \frac{\pi R^2}{A} \sum_{i \neq f} N_i(t), \quad (4b)$$

where the bivariate normalized K function $k_{fh}(R)$ indicates segregation to heterospecifics (subscript ‘h’) within distance R if $k_{fh}(R) < 1$. Independent placement occurs if $k_{fh}(R) = 1$, and attraction if $k_{fh}(R) > 1$.

Motivated by the finding $n_{kfp} \approx B_f n_{kff}$ (Extended Data Fig. 3b) we rewrite the mean crowding index \bar{n}_{fp} as

$$\bar{n}_{fp} = \underbrace{B_f}_{\frac{\bar{n}_{fp}}{\bar{n}_{fh}}} \bar{n}_{fh}, \quad (5)$$

where the point process summary function B_f indicates how much the competition strength of one heterospecific neighbour differs on average from that of one conspecific neighbour. The values of B_f depend mainly on the individual-level interaction coefficients β_{fi} but also on the spatial pattern of the different species and their relative abundances (equation (12)).

Inserting the expressions for \bar{n}_{ff} and \bar{n}_{fp} into equation (3) and comparing with equation (1b) leads to our first main result, the analytical expressions of the population-level interaction coefficients:

$$\alpha_{ff} = c \gamma_{ff} k_{ff} \beta_{ff} \quad \text{intraspecific interactions of species } f \quad (6a)$$

$$\alpha_{fi} = c \gamma_{fp} k_{fh} \beta_{ff} B_f \quad \text{interspecific interactions with species } i \quad (6b)$$

with scaling constant $c = \pi R^2/A$. Notably, equation (6b) indicates that the emerging population-level interaction coefficients α_{fi} are

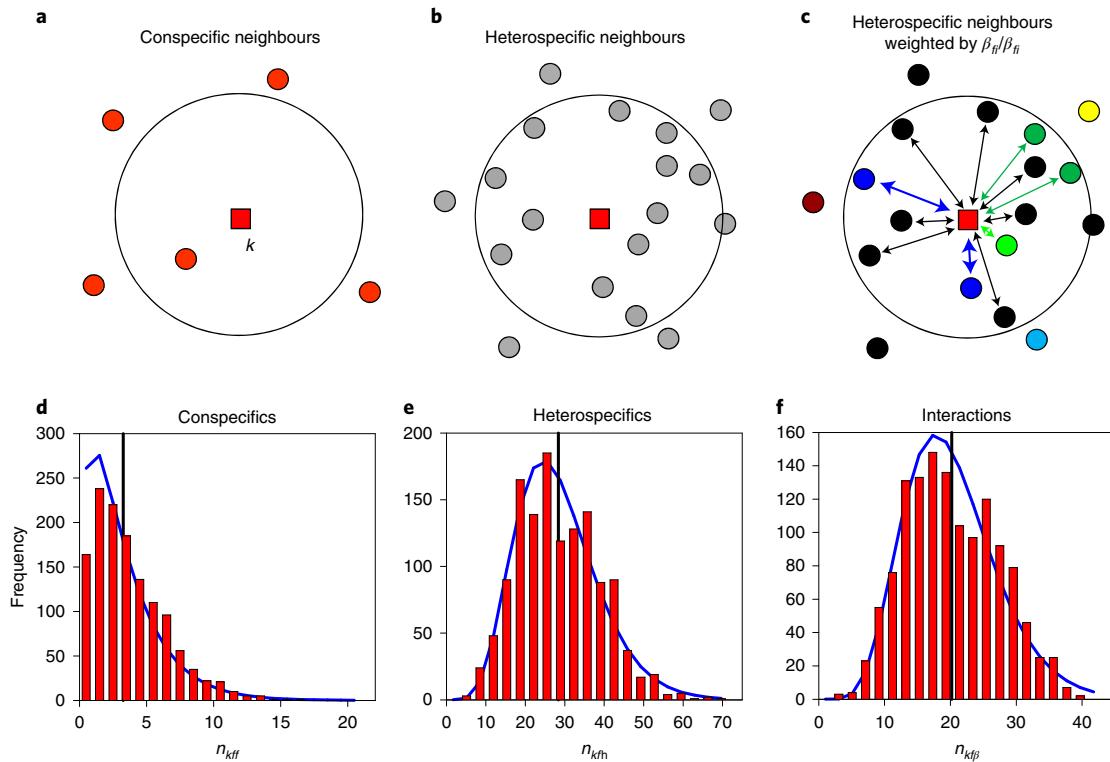


Fig. 1 | Neighbourhood crowding indices describe individual-level interactions and their intraspecific variability. **a**, The conspecific crowding index n_{kff} is the number of conspecific neighbours (filled red circles) within distance R (black circle) of the focal individual k (filled red square). **b**, The heterospecific crowding index n_{kfh} is the number of heterospecific neighbours (filled grey circles) within distance R of the focal individual k . **c**, The heterospecific interaction crowding index n_{kff}^* additionally weights heterospecifics by their relative competitive effect β_{fi}/β_{ff} , symbolized by the arrows. Different colours indicate different species. **d**, Distribution of the number n_{kff} of conspecific neighbours with diameter at breast height (dbh) ≥ 10 cm of the species *Castanopsis cuspidata* of the 25 ha Fushan plot. **e**, Corresponding distribution of heterospecific neighbours n_{kfh} . **f**, Corresponding distribution of the crowding index n_{kff}^* . In **d–f**, blue lines show gamma distributions with the same means and variance-to-mean ratios as the observed distributions, and the vertical black line indicates the mean value. See Supplementary Data Table 1 for additional examples.

the same for all heterospecifics (that is, $\alpha_{fi} = \alpha_{fh}$ for $i \neq f$). Thus, even if the individual-level interaction coefficients β_{fi} differ among species pairs, the emerging population-level interaction coefficients α_{fi} have a substantially simpler structure. This phenomenon is an example of simplicity emerging from complex species interactions³⁰ and is likely to occur only in species-rich communities^{9,12,31}.

The population-level interaction coefficients α_{fi} depend on several factors that can influence the macroscale balance between intra- and interspecific competition: (1) intraspecific clustering k_{ff} and interspecific segregation k_{fh} (Fig. 2a,b), (2) the relative competition strength B_f of one heterospecific neighbour (Fig. 2c) and (3) the shape of the response of survival to crowding (γ_f and γ_{ff} , which contain the variance-to-mean ratios of the distribution of the crowding indices). Note that absence of spatial patterns (that is, $k_{ff} = 1$ and $k_{fh} = 1$) and a linear approximation of equation (3) lead to $\gamma_f = \gamma_{ff} = 1$ and direct proportionality $\alpha_{fi} = c\beta_{fi}$ of the individual- and population-level interaction coefficients, as assumed by Lotka–Volterra models.

The information on spatial patterns extracted from the inventory data of our nine forests allows us to estimate the relative population-level interaction coefficients α_{fi}/α_{ff} for all pairs of species i and f (Fig. 3 and Extended Data Fig. 4). For example, at BCI, the values of α_{fi}/α_{ff} differ substantially from the corresponding individual-level coefficients β_{fi}/β_{ff} (Fig. 3a,c), and for 83% of all species pairs, we find $\alpha_{fi}/\alpha_{ff} < \beta_{fi}/\beta_{ff}$. Thus, the mesoscale spatial patterns can reduce, at the population level, the strength of heterospecific interactions relative to conspecific interactions by ‘diluting’ encounters with heterospecific neighbours relative to conspecific

neighbours. Spatial patterns therefore have a strong potential to alter the outcome of deterministic individual-level interactions.

Conditions for coexistence in the multiscale model. To study the consequences of the emerging spatial patterns for community dynamics and coexistence we insert the population-level interaction coefficients α_{fi} (equation (6)) into a simple macroscale model

$$\frac{N_f(t+\Delta t) - N_f(t)}{\Delta t} = N_f(t) \left[(r_f - 1) + s_f \exp \left(-\alpha_{ff} N_f(t) - \alpha_{fh} \sum_{i \neq f} N_i(t) \right) \right] \quad (7)$$

In this model we assume that survival is governed by neighbourhood competition with $\alpha_{fi} = \alpha_{fh}$, and the number of recruits of species f during a time step Δt is given by $r_f N_f(t)$, where r_f is the per capita reproduction rate of species f .

The carrying capacity of species f (that is, the equilibrium of equation (7) with $N_i(t) = 0$ for all $i \neq f$) is given by $K_f = -\ln \left(\frac{1-r_f}{s_f} \right) \alpha_{ff}^{-1}$. Note that our theory also applies, after redefinition of the carrying capacity, to alternative macroscale models (Supplementary Table 3). From equation (7) we find $K_f = N_f^* + \frac{\alpha_{fh}}{\alpha_{ff}} \sum_{i \neq f} N_i^* = N_f^* \left(1 - \frac{\alpha_{fh}}{\alpha_{ff}} \right) + \frac{\alpha_{fh}}{\alpha_{ff}} J^*$, where N_f^* is the abundance of species f in equilibrium and J^* the equilibrium community size (that is, $J^* = \sum_i N_i^*$; see also equation (14)). This leads, under the assumption that the population-level interaction coeffi-

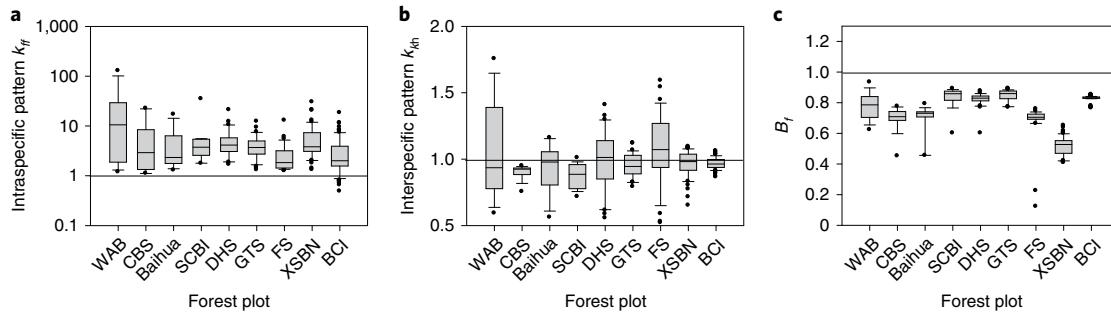


Fig. 2 | Emergent spatial patterns in the nine forest dynamics plots. a–c, Distribution of the values of the different measures of spatial patterns, taken over the focal species of the different forest plots. Boxplots show the 10th, 25th, 50th, 75th and 90th percentiles; outliers are indicated by filled black points. Intraspecific clustering is indicated by $k_{ff} > 1$ and interspecific segregation by $k_{fh} < 1$, and the less a heterospecific neighbour competes on average relative to a conspecific neighbour, the more B_f decreases. The neighbourhood radius used was $R = 10$ m. For the analysis, we used all individuals with dbh ≥ 10 cm and included focal species with more than 50 individuals. For forest plot names, see Supplementary Data Table 1.

cients α_{fh} are constant (Supplementary text), to a single equilibrium of the macroscale model for species f

$$N_f^* = \frac{-\ln\left(\frac{1-r_f}{s_f}\right)}{\frac{\alpha_{ff}K_f - \alpha_{fh}J^*}{\alpha_{ff} - \alpha_{fh}}} \quad (8)$$

that is positive if denominator and numerator are both positive or both negative. However, the invasion criterion (equation (18)) is only fulfilled if both are positive. In this case, equation (8) suggests two different ways a species can go extinct. First, the denominator indicates that a species with strong clustering k_{ff} will show a small equilibrium abundance since in this case $\alpha_{ff} \gg \alpha_{fh}$ (equation (6)). Large values of k_{ff} can be expected for species of low abundance under dispersal limitation, where recruitment happens close to conspecific adults.

Second, the numerator of equation (8) indicates a positive abundance of species f if $\alpha_{fh}K_f > \alpha_{ff}J^*$ and $\alpha_{fh}/\alpha_{ff} < 1$. Therefore, we introduce a new feasibility index

$$\mu_f = \frac{\alpha_{fh}J^*}{\alpha_{ff}K_f} \quad (9)$$

that indicates a positive abundance if $\mu_f < 1$ given that heterospecific interactions at the population level are weaker than conspecific interactions (that is, $\alpha_{fh}/\alpha_{ff} < 1$). The invasion criterion^{7,8} that tests whether a species with low abundance can invade the equilibrium community of all other species turned out to be basically the same as the feasibility condition (equation (9)) if the invading species does not show strong clustering (Methods and equation (18)). Note that we did not assume Allee effects⁸.

Further analysis that considers the dependency of J^* on the values of K_f and α_{fh}/α_{ff} shows that the values of μ_f must be similar for all species f to fulfil the condition $\mu_f < 1$, and that μ_f can show larger interspecific variability if the species richness S is smaller and/or if the mean of $\frac{\alpha_{fh}}{\alpha_{ff}} \left(1 - \frac{\alpha_{fh}}{\alpha_{ff}}\right)^{-1}$ is smaller (equations (14) and (15)). The feasibility index μ_f therefore governs species assembly by determining the subset of species of a larger species pool that can persist¹³, but any addition of a new species changes μ_f and may lead to reassembly of the community.

Using the observed abundances in the forest plots (and assuming equilibrium) allows us to test our theory. We can estimate from the observed abundances the carrying capacities K_f and therefore also the indices μ_f for all focal species of our nine plots (Fig. 4). For

282 of our 289 focal species, we found $\mu_f < 1$ and $\alpha_{fh}/\alpha_{ff} < 1$, which means that the two conditions for stable coexistence are indeed satisfied for nearly all species. However, this is not a given, as shown by the seven species from BCI with $\mu_f > 1$ and $\alpha_{fh}/\alpha_{ff} > 1$. Thus, our theory is compatible with the observed coexistence of most species at our nine forest plots if the assumption of approximately constant population-level interaction coefficients holds.

In addition, we get information on the typical values of μ_f and α_{fh}/α_{ff} that allows for insight into the stability of the communities. In agreement with the predictions of our theory, we find that μ_f tends to be smaller if α_{fh}/α_{ff} is smaller (Fig. 4c). Furthermore, the values of μ_f were, for most species, larger than the expectation of μ_f for the corresponding communities without interspecific variability in μ_f (equation (10a)) but with the same number of species and the same mean values of α_{fh}/α_{ff} (Fig. 4c), but all of the values were relatively close to the critical value of 1 (the median of all 289 species was 0.938; Fig. 4a).

Consequences of spatial patterns for coexistence. Our theory predicts that coexistence requires, in the limit of high species richness, that species approach functional identity with respect to the feasibility index μ_f (Methods). This resembles neutral theory^{2,32}, but our theory allows for trade-offs among demographic parameters and emerging spatial patterns to reach this equivalence (equation (9)). To study the consequences of spatial patterns for coexistence, we analysed a symmetric³³ version of our model where all species have the same parameters and follow the same stochastic rules and where all individuals compete identically (leading to $B_f = 1$). Thus, we eliminate any potential coexistence mechanism other than that resulting from spatial patterns.

If the mesoscale patterns k_{ff} and k_{fh} converge to a stochastic equilibrium, we find that the feasibility and invasion criteria are always fulfilled if $\alpha_{fh}/\alpha_{ff} < 1$ since

$$\mu_f = \frac{S \frac{\alpha_{fh}}{\alpha_{ff}}}{1 + \frac{\alpha_{fh}}{\alpha_{ff}} (S - 1)} < 1 \quad (10a)$$

$$\mu_f^i = \frac{(S - 1) \frac{\alpha_{fh}}{\alpha_{ff}}}{1 + \frac{\alpha_{fh}}{\alpha_{ff}} (S - 2)} < 1 \quad (10b)$$

Equations (10a) and (10b) follow from equations (16) and (18), respectively, if α_{fh}/α_{ff} is the same for all species f . Thus, the spatial patterns that emerge at the mesoscale from the individual-level interactions can stabilize if $\alpha_{fh}/\alpha_{ff} < 1$. The underlying mechanism is a positive fitness–density covariance³⁴ (Methods).

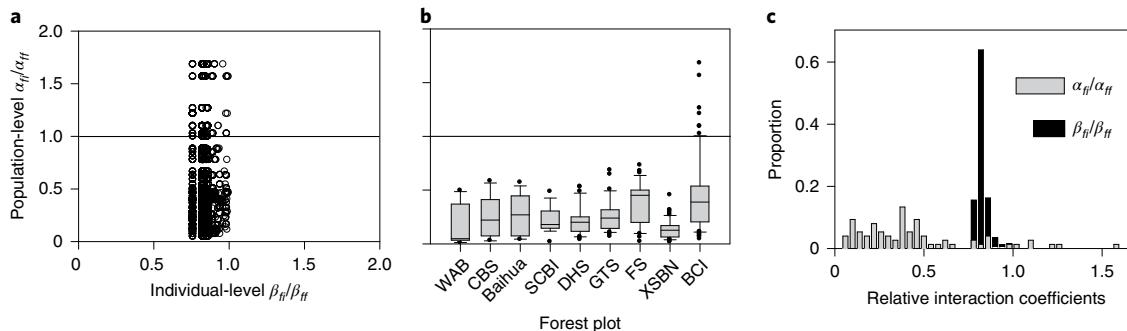


Fig. 3 | The emergent population-level interaction coefficients. **a**, Relationship between the relative population-level interaction coefficients $\alpha_{ii}/\alpha_{ii}^*$ and the corresponding relative individual-level interaction coefficients β_{ii}/β_{ii}^* (equation (6)) for the 75 focal species of the BCI plot. The β_{ii}/β_{ii}^* were based on phylogenetic dissimilarity. **b**, Distribution of the values of $\alpha_{ii}/\alpha_{ii}^*$ taken over the 289 focal species of the different forest plot. Boxplots show the 10th, 25th, 50th, 75th and 90th percentiles; outliers are indicated by filled black points. For full, separate distributions for tropical, subtropical and temperate forests, see Extended Data Fig. 4. **c**, Example for the distribution of the relative individual- and population-level interaction coefficients for the BCI plot. The neighbourhood radius used was $R=10$ m. For the analysis, we used all individuals with $dbh \geq 10$ cm and included focal species with more than 50 individuals. For forest plot names, see Supplementary Data Table 1.

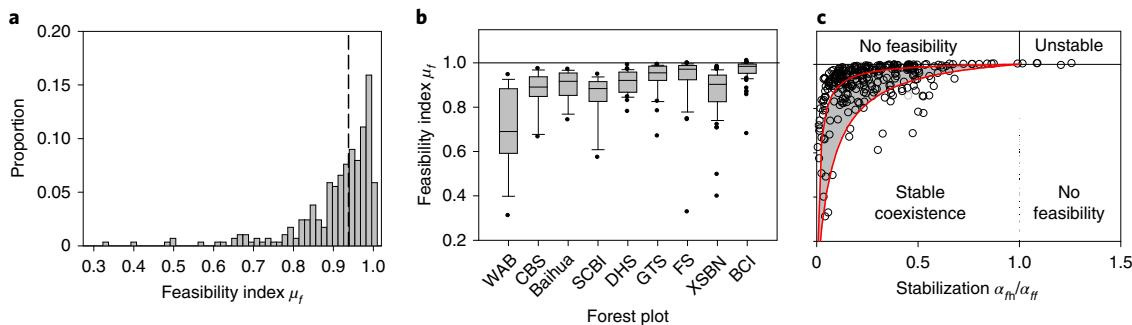


Fig. 4 | The feasibility index μ_i for the 289 focal species of different forest plots. **a**, The distribution of the index μ_i (equation (9)) for the 289 analysed tree species. The vertical dashed line indicates the median of the distribution. **b**, Distribution of μ_i , taken over the 289 analysed species of the different forest plot. Boxplots show the 10th, 25th, 50th, 75th and 90th percentiles; outliers are represented by black points. **c**, Relationship between the feasibility index μ_i and the ratio $\alpha_{ii}/\alpha_{ii}^*$ of heterospecific to conspecific population-level interaction coefficients. Values of $\mu_i < 1$ indicate a positive abundance. The red lines show the dependence of μ_i on $\alpha_{ii}/\alpha_{ii}^*$ expected for communities without interspecific variability in μ_i (equation (10a)) with 18 focal species (CBS plot, lower line) and 75 focal species (BCI plot, upper line). The neighbourhood radius used was $R=10$ m. For the analysis, we used all individuals with $dbh \geq 10$ cm and included tree species with more than 50 individuals. For plot names, see Supplementary Data Table 1.

To reveal the conditions that can lead to coexistence in a spatially explicit context, we use a spatially explicit and individual-based³⁵ implementation of the symmetric version of the multiscale model (equation (7) and Methods). Models of this type are able to produce realistic spatial patterns consistent with mapped species distributions of large forest plots^{17,35}. While our analytical approach in equation (7) only allows us to make simplified assumptions about the spatial component of the recruitment process, the simulation model allows us to explore the role of the spatial component of recruitment in more detail.

Indeed, the way recruits were placed was critical for coexistence. Randomly placed recruits produced unstable dynamics (Extended Data Fig. 5a) characterized by regularity (mean of $k_{ij}=0.92$) and segregation (mean of $k_{jh}=0.92$), both caused by competition¹⁸, and the instability was caused by con- and heterospecifics competing equally at the population level (that is, $\alpha_{jh}/\alpha_{jj} \approx 1$) (Extended Data Fig. 5d,g). When we followed the common approach of placing recruits with a kernel around conspecific adults^{17,18,35–39} to mimic dispersal limitation, we again found unstable dynamics (Extended Data Fig. 5b), despite intraspecific clustering and interspecific segregation (that is, $\alpha_{jh}/\alpha_{jj} < 1$; Extended Data Fig. 5n). The reason for

the instability was high clustering of rare species^{24,35} (Extended Data Fig. 6) that completely negated the potentially positive effects of $\alpha_{jh}/\alpha_{jj} < 1$.

In contrast, community dynamics can be stabilized if recruits are placed in small clusters but independent of the location of conspecific adults (Extended Data Fig. 5c). With this mechanism we mimic canopy gaps⁴⁰, animal seed dispersal⁴¹ or other mechanisms that can generate clustering independent of parent locations, as found at BCI⁴². Decoupling clustering from the parent locations does not lead to the negative relationship between clustering and abundance, and all measures of spatial patterns converged quickly into quasi-equilibrium (Extended Data Fig. 5f,i,o).

The simulation data reveal the spatial coexistence mechanism underlying the positive fitness–density covariance³⁴ (Extended Data Figs. 7 and 8). We find that the emerging spatial patterns lead to a situation where individuals of a common species are more likely to be near more neighbours and tend therefore to experience stronger competition (Extended Data Fig. 7). While the number of heterospecific neighbours remains approximately constant, the number of conspecific neighbours decreases with decreasing abundance if clustering does not change with abundance (equation (4a)). However,

if clustering increases with decreasing abundance, the rare-species advantage is weakened and the dynamics become unstable²⁴.

The data of several ForestGEO forest plots were compatible with a positive fitness–density covariance (Extended Data Fig. 8f–i) as they show that, when a species becomes rare, areas of higher conspecific crowding tend to have fewer total competitors. Comparison of the results of the stable versus unstable simulation showed that even relatively weak tendencies in this relationship are sufficient to stabilize the dynamics (Extended Data Figs. 8a,b). However, this was not the case for three temperate forest plots where the power-law clustering–abundance relationship showed exponents of $b < -0.5$ (that is, species tended to have high clustering at low abundances), but the other plots showed $b > -0.5$ (Extended Data Fig. 8n–p).

The apparent contradiction with previous theoretical studies^{18,24,35,37} where intraspecific clustering and interspecific segregation could not stabilize community dynamics thus arises as a consequence of the assumption of placing recruits close to their parents. This finding has important consequences for ecological theory because it shows, in contrast to the prevalent view^{36,43,44}, that spatial patterns alone can lead to coexistence of multiple species. This is even more important since the specific spatial patterns required for this coexistence mechanism also exist in real forests.

Conclusions

Understanding the mechanisms that maintain high species diversity in communities such as tropical forests is at the core of ecological theory, but these mechanisms are not yet fully resolved. Here, we argue that spatial patterns may play an important role in species coexistence of high diversity plant communities¹⁷. To test this hypothesis, we introduced a multiscale framework that reveals how pattern-forming processes operating at the level of individuals translate into mesoscale spatial patterns and how those patterns influence macroscale community dynamics.

We showed that the population-level interaction coefficients α_{fi} can have, for a broad range of common circumstances, a simpler structure than the underlying individual-level interaction coefficients β_{fi} . This simplicity, which emerged from spatially explicit species interactions³⁰, allowed for an analytical treatment of equilibrium, feasibility and invasion conditions of the corresponding macroscale models (equations (8–10)). Inserting the emerging α_{fi} coefficients into macroscale community models (for example, equation (7); Supplementary Table 3) should, in principle, allow us to take advantage of macroscale theory^{9,11–13,45}. However, our results also indicate that the population-level interaction coefficients may not be temporally constant as commonly assumed but depend on spatial patterns that may change with abundance. This is especially likely if recruitment is mainly located close to the parents.

It is also possible to expand our framework to take into account more detailed neighbourhood crowding indices that consider not only the number of neighboured trees of a given species but also their distance and size^{19,21,26}. This requires redefinition of the quantities k_{ff} and k_{fh} that describe intraspecific clustering and interspecific segregation, respectively, but does not change the overall structure of our equations. A special strength of our approach is that the population-level interaction coefficients contain measures of spatial neighbourhood patterns that can be directly estimated from fully mapped forest plots⁴. Together with additional information, this may allow for estimating network structures as well as stability of the whole community.

Our analysis revealed that communities of competing species can show a stable mode where the mesoscale patterns converge quickly into quasi-equilibrium and an unstable mode where negative relationships between species clustering and abundance emerge (Extended Data Figs. 5 and 6). The two modes are governed by the way species clustering is generated: the well-known unstable mode is related to clustering of recruits around their parents^{17,18,35–39}, whereas the stable mode is related to clustering in locations that are

independent from the parent locations, due, for example, to animal seed dispersal⁴¹ or canopy gaps⁴⁰. This result calls for a closer examination of the spatial relationship between the recruits and adults. Indeed, independent placement of recruits from conspecific large trees may not be unusual. For example, Getzin et al.⁴² found in detailed analyses of the BCI forest that recruits were for most species spatially independent of large conspecific trees. For the stable mode we could identify conditions for coexistence, and forthcoming work may extend to quantifying the ability of additional mechanisms such as niche differences^{7,8}, habitat associations⁴⁶, spatial and temporal relative nonlinearity^{7,8} and storage effects^{7,8} to alleviate the destabilizing increase of clustering if species become rare.

This study explicitly incorporates spatial patterns in theoretical models of plant communities and combines analytical theory with spatial simulations and field data analysis. Our finding that species with similar attributes may show stable coexistence has profound implications for ecological theory. Furthermore, the multiscale framework we propose here opens exciting new avenues to explain species coexistence through a spatial lens.

Methods

Study areas. Nine large forest dynamics plots of areas between 20 and 50 ha were used in the present study (Supplementary Table 1). The forest plots are part of the ForestGEO network⁴ and are situated in Asia and the Americas at locations ranging in latitude from 9.15° N to 45.55° N. Tree species richness among the plots ranges from 36 to 468. All free-standing individuals with diameter at breast height (dbh) ≥ 1 cm were mapped, size measured and identified. We focused our analysis here on individuals with dbh ≥ 10 cm (resulting in a sample size of 131,582 individuals) and focal species with more than 50 individuals (resulting in 289 species). The 10 cm size threshold excludes most of the saplings and enables comparisons with previous spatial analyses^{20,35,47,48}. Shrub species were also excluded.

Some of our analyses require estimation of the ratio β_f/β_g that describes the relative individual-level competitive effect¹⁸ of individuals of species i on an individual of the focal species f . We used for this purpose phylogenetic distances based on molecular data, given in Myr, that assume that functional traits are phylogenetically conserved^{19,26,27}. In this case, close relatives are predicted to compete more strongly or to share more pests than distant relatives²⁶. To obtain consistent measures among forest plots, phylogenetic similarities were scaled between 0 and 1, with conspecifics set to 1, and a similarity of 0 was assumed for a phylogenetic distance of 1,200 Myr, which was somewhat larger than the maximal observed distance (1,059 Myr). This was necessary to avoid discounting crowding effects from the most distantly related neighbours²⁶.

Observed spatial patterns at species-rich forests. Figure 1 and Supplementary Data Table 1 show the intraspecific variation in our three crowding indices $n_{k_{ff}}$, $n_{k_{fh}}$ and $n_{k_{hf}}$ that can be approximated by gamma distributions. To assess how well the gamma distribution described the observed distribution, we used an error index defined as the sum of the absolute differences of the two cumulative distributions divided by the number of bins (spanning the two distributions). The maximal value of the error index is one, and a smaller value indicates a better fit.

Equations (6, 8 and 9) relate the measures of the emerging spatial patterns (that is, k_{ff} , k_{fh} and B_f) to macroscale properties and conditions for species coexistence. Even though our multiscale model (equation (7)) is simplified, it allows for a direct comparison with the emerging patterns in our nine fully stem-mapped forest plots. We estimate the key quantities of equations (8) and (9) directly from the forest plot data (Fig. 4), with the exception of the carrying capacities K_f , which were indirectly estimated from the observed species abundances (assuming approximate equilibrium; equation (8) and Supplementary Data Table 1). This allowed us to estimate the feasibility index μ_f (equation (9)). Because statistical analyses with individual-based neighbourhood models^{19,26} based on neighbourhood crowding indices have shown that the performance of trees depends on their neighbours for R between 10 and 15 m, we estimate all measures of spatial neighbourhood patterns with a neighbourhood radius of $R = 10$ m. Analyses with $R = 15$ or $R = 20$ gave similar results.

The spatial multispecies model and equilibrium. We use a general macroscale model to describe the dynamics of a community of S species:

$$\frac{N_f(t + \Delta t) - N_f(t)}{\Delta t} = N_f(t) \left[(r_f - 1) + s_f \exp \left(-\alpha_{ff} N_f(t) - \sum_{i \neq f} \alpha_{fi} N_i(t) \right) \right] \quad (11)$$

where r_f is the mean number of recruits per adult of species f within time step Δt , s_f is a density-independent background survival rate of species f and the α_{fi} are the

population-level interaction coefficients, yielding $\alpha_{ff} = c\gamma_{ff}k_{ff}\beta_{ff}$ and $\alpha_{fi} = c\gamma_{if}k_{fi}\beta_{if}$ (equation (6)). The β_{if} are the assumed individual-level interaction coefficients between individuals of species i and f ; $k_{ff} = K_f(R)/\pi R^2$ and $k_{fi} = K_f(R)/\pi R^2$ measure intraspecific clustering and interspecific segregation, respectively, with $K_f(R)$ being the univariate K function for species f and $K_{fh}(R)$ the bivariate K function describing the pattern of all heterospecifics 'h' around individuals of species f . A is the area of the observation window.

Following equation (5), B_f can be estimated as

$$B_f = \frac{\bar{n}_{f\beta}}{\bar{n}_{fh}} = \frac{\sum_{i \neq f} [ck_{fi}N_i(t)] \frac{\beta_{if}}{\beta_{ff}}}{\sum_{i \neq f} [ck_{fi}N_i(t)]} = \frac{\sum_{i \neq f} k_{fi}N_i(t) \frac{\beta_{if}}{\beta_{ff}}}{k_{fh} \sum_{i \neq f} N_i(t)}, \quad (12)$$

and is the weighted average of the relative individual-level interaction coefficients β_{if}/β_{ff} between species i and the focal species f , weighted by the mean number of individuals of species i in the neighbourhoods of the individuals of the focal species (that is, $c k_{fi} N_i(t)$). For competitive interactions, B_f ranges between zero and one; $B_f = 1$ indicates that heterospecific and conspecific neighbours compete equally, and smaller values of B_f indicate reduced competition with heterospecific neighbours. The denominator can be rewritten in terms of segregation k_{fh} to all heterospecifics and the total number of heterospecifics $\sum_{i \neq f} N_i(t)$.

The analytical expression of the equilibrium (equation (8)) relies on the assumption that the values of B_f are approximately constant in time. This assumption may not apply in our model during the initial burn-in phase of the simulations if the β_{if}/β_{ff} show large intraspecific variability (Supplementary Text and Figs. 1–5). The underlying mechanism is the central niche effect introduced by Stump¹⁵ where a species has reduced average fitness if it has high niche overlap with competitors.

Finally, the factors $\gamma_{ff} = \ln(1 + b_{ff}\beta_{ff}) / (b_{ff}\beta_{ff})^{-1}$ and $\gamma_{fp} = \ln(1 + b_{fp}\beta_{fp}) / (b_{fp}\beta_{fp})^{-1}$ describe the influence of the variance-to-mean ratios b_{ff} and b_{fp} of the gamma distribution of the crowding indices n_{kff} and n_{kfp} , respectively. For high survival rates during one time step (for example, >85%), the values of γ_{ff} and γ_{fp} are close to one; in this case the exponential function in equation (1a) can be approximated by its linear expansion and $\gamma_{ff} = \gamma_{fp} = 1$.

In equilibrium we have $(N_f(t + \Delta t) - N_f(t))/\Delta t = 0$, which leads, with equation (7), to:

$$N_f^* = \left(K_f - \frac{\alpha_{fh}}{\alpha_{ff}} J^* \right) \left(1 - \frac{\alpha_{fh}}{\alpha_{ff}} \right)^{-1} \quad (13)$$

with $K_f = -\ln\left(\frac{1 - r_f}{s_f}\right)(\alpha_{ff})^{-1}$ and the total number of individuals being $J^* = \sum_i N_i^*$. Rewriting equation (13) yields $\frac{K_f}{J^*} = \left(\frac{N_f^*}{J^*} \right) \left(1 - \frac{\alpha_{fh}}{\alpha_{ff}} \right) + \frac{\alpha_{fh}}{\alpha_{ff}}$. For $\alpha_{fh}/\alpha_{ff} < 1$ we therefore find $K_f < J^*$, which indicates that a multispecies forest would host more individuals than a monoculture. To estimate J^* we sum equation (13) over all species i and find $J^* = \sum_i \frac{K_i}{1 - \alpha_{ih}/\alpha_{ii}} - J^* \sum_i \frac{\alpha_{ih}/\alpha_{ii}}{1 - \alpha_{ih}/\alpha_{ii}}$. Therefore, we obtain

$$J^* = \sum_{i=1}^S \frac{K_i}{1 - \alpha_{ih}/\alpha_{ii}} \left(1 + \sum_{i=1}^S \frac{\alpha_{ih}/\alpha_{ii}}{1 - \alpha_{ih}/\alpha_{ii}} \right)^{-1} = \frac{Sm_K}{1 + Sm_a} \quad (14)$$

with $m_K = \frac{1}{S} \sum_i \frac{K_i}{1 - \alpha_{ih}/\alpha_{ii}}$ and $m_a = \frac{1}{S} \sum_i \frac{\alpha_{ih}/\alpha_{ii}}{1 - \alpha_{ih}/\alpha_{ii}}$ being averages over the S species of the community.

All species have positive abundances at equilibrium if the two conditions $\mu_f = \alpha_{fh}J^*/\alpha_{ff}K_f < 1$ and $\alpha_{fh}/\alpha_{ff} < 1$ are met (see equation (9)). We now show that the chance that these conditions are satisfied for all species is larger if the values of μ_f show little intraspecific variability. To understand this, we assume that the μ_f can be approximated by their mean $\bar{\mu}$. In this case $J^*/\bar{\mu}$ is also approximately constant and we can replace K_i in equation (14) by $(J^*/\bar{\mu})(\alpha_{ih}/\alpha_{ii})^{-1}$ and obtain

$$J^* = \sum_{i=1}^S \frac{\left(\frac{J^*}{\bar{\mu}}\right) \frac{\alpha_{ih}}{\alpha_{ii}}}{1 - \frac{\alpha_{ih}}{\alpha_{ii}}} \left(1 + \sum_{i=1}^S \frac{\frac{\alpha_{ih}}{\alpha_{ii}}}{1 - \frac{\alpha_{ih}}{\alpha_{ii}}} \right)^{-1} = \frac{J^*}{\bar{\mu}} \frac{Sm_a}{1 + Sm_a} \quad (15)$$

where S is the number of species in the community, and therefore

$$\bar{\mu} = \frac{Sm_a}{1 + Sm_a} < 1 \quad (16)$$

Thus, in the case of a perfect interspecific balance in μ_f we always have a feasible equilibrium if $\alpha_{fh}/\alpha_{ff} < 1$, and species can go extinct only if the intraspecific variability in μ_f becomes too large. The smaller the mean value of μ_f , the more variability in μ_f is allowed. Equation (16) shows that $\bar{\mu}$ is smaller if the number S of species in the community is smaller and/or if the mean value of m_a is smaller.

Equation (16) also suggests that communities with more species need to show stronger species equivalence in μ_f , because the term $Sm_a(1 + Sm_a)^{-1}$ approaches a value of one for a large number of species S . This finding mirrors the results of analyses of Lotka–Volterra models with random interaction matrices¹⁴ that showed that the larger the number of species S , the more difficult it becomes to generate a feasible community.

Invasion criterion. Using the population-level interaction coefficients (equation (6)) in the macroscale model, we now derive conditions for coexistence based on the invasion criterion^{7,8} for a species m . The growth rate of an invading species m with low density $M(t)$ into the equilibrium community of all other $S-1$ species $N_i(t)$ should be positive; thus, with equation (7), we have

$$(r_m - 1) + s_m \exp \left(-\alpha_{mm}M(t) - \alpha_{mh} \sum_{i=1}^{S-1} N_i^* \right) > 0. \quad (17)$$

Considering that $J_m^* = \sum_{i=1}^{S-1} N_i^*$ and $\alpha_{mm}M(t) \ll \alpha_{fh}J_m^*$ (that is, the invading species m is at low abundance and does not show strong clustering) we find $-\ln\left(\frac{1 - r_m}{s_m}\right) > \alpha_{mh}J_m^*$, and by dividing by α_{mm} we obtain the invasion condition

$$\mu_m^i = \frac{\alpha_{mh}J_m^*}{\alpha_{mm}K_m} < 1 \quad (18)$$

which is basically identical to the condition for feasibility (equation (9)), but here the community size J_m^* of the reduced community appears instead of the equilibrium community size J of all species, including species m . Thus, a new species m is more likely to invade if it has a high value of the carrying capacity K_m and it more strongly reduces heterospecific interactions relative to conspecific interaction (that is, α_{mh}/α_{mm} is smaller). However, if the species is too efficient (that is, has too large a capacity K_m and/or too low an α_{mh}/α_{mm}) it may increase the value of J too much (equation (14)), thereby causing the extinction of the weakest species with the highest values of μ_f (that is, a too-low value of K_m and a too-high value of α_{fh}/α_{ff}). Equation (18) also suggest that an equilibrium with $\mu_f > 1$ and $\alpha_{mh}/\alpha_{mm} > 1$ will be unstable.

Fitness–density covariance. To place our new spatial coexistence mechanism in the context of existing coexistence theory, we apply scale transition theory³⁴ to our model version where spatial effects are the only potential coexistence mechanism (that is, all species have the same parameters and all individuals compete equally; $\beta_{if}/\beta_{ff} = 1$, $B_f = 1$).

Following equation (1a), the expected fitness of an individual k of a focal species f (that is, its expected contribution to the population after some defined interval of time Δt) in the macroscale model (equation (7)) is

$$\lambda_{kf} = r_f + s_f f(W_k) = r_f + s_f \exp \left(-\beta_{ff} \left(n_{kff} + \sum_{i \neq f} n_{kfi} \right) \right) \quad (19)$$

where $W_k = n_{kff} + \sum_{i \neq f} n_{kfi}$ is the fitness factor of individual k , $f(W_k) = \exp(-\beta_{ff} W_k)$ is the fitness function and n_{kff} and $\sum_{i \neq f} n_{kfi}$ are the number of conspecific and heterospecific neighbours, respectively, of individual k within distance R . The spatial average of the fitness factor over the entire plot is

$$\bar{W}_k = c N_f(t) + c \sum_{i \neq f} N_i(t) = c J(t) \quad (20)$$

where $c = \pi R^2/A$, and $J(t) = \sum_i N_i(t)$ is the total number of individuals in the plot. Given that $J(t)$ converges very quickly into equilibrium J (Extended Data Fig. 5 and Supplementary Figs. 1 and 2), we find for the spatial average fitness $\bar{\lambda}_f = 1$.

The average individual fitness $\bar{\lambda}_f(t)$ of a focal species f is the average of λ_{kf} over all individuals k of species f and can be estimated for the macroscale model (equations (6) and (7)) as $\bar{\lambda}_f(t) = N_f(t + \Delta t)/N_f(t)$. A key ingredient of scale transition theory³⁴ is that the fitness–density covariance is given by $\text{cov} = \bar{\lambda}_f(t) - \bar{\lambda}_f$. With equation (3) and $\gamma_{ff} \approx \gamma_{fp} \approx 1$ and $\bar{n}_{ff} = \bar{n}_{fh}$ we find

$$\bar{\lambda}_f(t) - \bar{\lambda}_f = (r_f - 1) + s_f \exp(-\beta_{ff}(\bar{n}_{ff} + \bar{n}_{fh})) \quad (21)$$

where the mean of the crowding indices is given by $\bar{n}_{ff}(N_f) = ck_{ff}(N_f)N_f$ and $\bar{n}_{fh}(N_f) = ck_{fh}(J^* - N_f)$ (equation (4)). Therefore, if clustering k_{ff} and segregation k_{fh} are independent from abundance N_f , more abundant species have more neighbours, since

$$\bar{n}_{ff} + \bar{n}_{fh} = c(k_{ff} - k_{fh})N_f + ck_{fh}J^* \quad (22)$$

Thus, a positive fitness–density covariance in our model means that individuals of a common species are more likely to be near more trees in total.

Extended Data Fig. 7 shows the quantities $\bar{n}_{ff} + \bar{n}_{fh}$, \bar{n}_{ff} , \bar{n}_{fh} and $\bar{\lambda}_f - \bar{\lambda}_f$ plotted over abundance N_f for data generated by our spatially explicit simulation model for the scenarios of stable and unstable dynamics (Extended Data Fig. 5b,c). Indeed, the stable simulations show a positive fitness–density covariance, however, there is no such trend for the dynamics of the unstable community (Extended Data Fig. 7g,h).

Spatial patterns will act as positive fitness–density covariance if, when a species becomes rare, areas of high conspecific crowding have fewer competitors. We tested this for the data generated by our simulation model and for the nine forest plots (Extended Data Fig. 8). We could estimate for each focal species f the covariance between the number of conspecific neighbours (that is, n_{kf}) and the

total number of neighbours (that is, $n_{kff} + n_{kfh}$) and demand that the covariance should be mostly positive and larger for more abundant species. However, since the quantity n_{kff} appears in this test on both sides, a positive covariance can be expected. To compensate for this artefact, we instead used the covariance between the local dominance of conspecifics in the neighbourhood of individuals k (that is, $d_{kff} = n_{kff}(n_{kff} + n_{kfh})^{-1}$) and total number of neighbours (that is, $n_{kff} + n_{kfh}$) (Extended Data Fig. 8).

Spatially explicit simulation model. The model is a spatially explicit and stochastic implementation of the spatial multispecies model (equation (7)), similar to that of May et al.^{35,37} and Dettlo and Müller-Landau¹⁷, and simulates the dynamics of a community of S tree species in a given plot of a homogeneous environment (for example, 50 ha) in 5 yr time steps adapted to the ForestGEO census interval (Extended Data Fig. 5 and Supplementary Figs. 1 and 2). Only reproductive (adult) trees are considered, but size differences between them are not considered. During a given time step the model first simulates stochastic recruitment of reproductive trees and placement of recruits, and second, stochastic survival of adults that depends on the neighbourhood crowding indices for conspecifics (n_{kff}) and heterospecifics (n_{kfp}) (but excluding recruits). In the next time step, the recruits count as reproductive adults and are subject to mortality. No immigration from a metacommunity is considered. To avoid edge effects, torus geometry is assumed.

The survival probability of an adult k of species f is given by $s_f \exp(-\beta_{ff}(n_{kff} + n_{kfp}))$ (equation (1a)). The two neighbourhood indices n_{kff} and n_{kfp} describe the competitive neighbourhood of the focal individual k and sum up all conspecific and heterospecific neighbours, respectively, within distance R , but weight them with the relative individual-level interaction coefficients β_{ff}/β_{fp} (refs. ^{19,21,26}).

Each individual produces on average r_f recruits, and their locations are determined by a type of Thomas process²⁸ to obtain clustering. To this end, the spatial position of the recruits is determined by two independent mechanisms. First, a proportion $1 - p_d$ of recruits is placed stochastically around randomly selected conspecific adults by using a two-dimensional kernel function (here a Gaussian with variance σ^2). This is the most common way to generate species clustering in spatially explicit models^{17,18,35–39}. Specifically, we first randomly select one parent for each of these recruits among the conspecific adults and then determine the position of the recruit by sampling from the kernel. Second, the remaining proportion p_d of recruits is distributed in the same way around randomly placed cluster centres that are located independently of conspecific adults. This mode mimics spatial clustering of recruits independent of the parent locations⁴² in a simple way, such as contagious seed dispersal by animals⁵⁰ or forest gaps that may imprint clumped distributions of recruits of pioneer species⁴⁰. For each species we assume a density λ_k of randomly distributed cluster centres, which have, at each time step, a probability p_{fp} of changing location. For each of these recruits, we first randomly select one cluster centre among the cluster centres of the corresponding species and then determine the position of the recruit by sampling from the kernel. For the simulation shown in Extended Data Fig. 5a, the recruits were located at random positions within the plot.

Parameterization of the simulation model. Extended Data Fig. 5 shows simulations of the individual-based model conducted in a 200 ha area containing approximately 83,000 trees with, initially, 80 species. There was no immigration. The model parameters were the same for all species, and all species followed exactly the same model rules. We selected $\beta_{ff} = \beta_{fp}$ to obtain no differences in con- and heterospecific interactions and $s_f = 1$ (no background mortality), and we adjusted the parameters $\beta_{ff} = 0.0075$ and $r_f = 0.1$ to yield tree densities (415 ha^{-1}) and an overall 5 yr mortality rate (10%) similar to those of trees with $\text{dbh} \geq 10 \text{ cm}$ in the BCI plot⁵¹.

The Gaussian kernel used to place recruits around conspecific adults or around random cluster centres had a parameter $\sigma = 10 \text{ m}$. There were 40 random cluster centres in total for each species that had a probability of $p_{fp} = 0.3$ of changing location within one census interval. The only difference between the simulation shown in Extended Data Fig. 5b and the one shown in Extended Data Fig. 5c is that in the former, we used a proportion $p_d = 0.05$ of recruits to be placed around randomly distributed cluster centres (that is, 95% of the recruits were placed close to their parents), but in the latter, we selected $p_d = 0.95$ (that is, 95% of the recruits were placed around randomly distributed cluster centres). In our simulations, on average, one of these cluster centres received four recruits per time step, which were scattered within a radius of approximately 30 m, and received approximately 13 recruits during its lifetime (at each time step it had a probability of 0.3 of changing location). In contrast, in Extended Data Fig. 5a recruits were placed at random locations within the plot.

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

Data availability

The data that support the findings in this manuscript (and the raw data for Figs. 2–4 and Extended Data Figs. 2–4 and 8) can be found in Supplementary Data Table 1. To generate this data, we used the raw census data of the ForestGEO

network that can only be shared on request because most PIs have not made them publicly available. For data requests see <https://forestgeo.si.edu/sites-all>.

Code availability

The source code of the simulation model is provided in the Supplementary Information.

Received: 17 March 2020; Accepted: 1 March 2021;

Published online: 3 May 2021

References

- Hutchinson, G. E. The paradox of plankton. *Am. Nat.* **95**, 137–147 (1961).
- Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).
- Usinowicz, J. et al. Temporal coexistence mechanisms contribute to the latitudinal gradient in forest diversity. *Nature* **550**, 105–108 (2017).
- Anderson-Treixeira, K. J. et al. CTFS-ForestGEO: a worldwide network monitoring forests in an era of global change. *Glob. Change Biol.* **21**, 528–549 (2015).
- Comita, L. S. et al. Testing predictions of the Janzen–Connell hypothesis: a meta-analysis of experimental evidence for distance and density-dependent seed and seedling survival. *J. Ecol.* **102**, 845–856 (2014).
- Wright, J. S. Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia* **130**, 1–14 (2002).
- Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
- Barabás, G., D’Andrea, R. & Stump, S. M. Chesson’s coexistence theory. *Ecol. Monogr.* **88**, 277–230 (2018).
- Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* **546**, 56–64 (2017).
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M. & Mayfield, M. M. Rethinking community assembly through the lens of coexistence theory. *Annu. Rev. Ecol. Syst.* **43**, 227–248 (2012).
- Stone, L. The feasibility and stability of large complex biological networks: a random matrix approach. *Sci. Rep.* **8**, 8246 (2018).
- Saavedra, S. et al. A structural approach for understanding multispecies coexistence. *Ecol. Monogr.* **87**, 470–486 (2017).
- Serván, C. A., Capitán, J. A., Grilli, J., Morrison, K. E. & Allesina, S. Coexistence of many species in random ecosystems. *Nat. Ecol. Evol.* **2**, 1237–1242 (2018).
- Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional nature of species coexistence. *Proc. Natl Acad. Sci. USA* **112**, 797–802 (2015).
- Allesina, S. & Tang, S. Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).
- Barbier, M., Arnoldi, J.-F., Bunin, G. & Loreau, M. Generic assembly patterns in complex ecological communities. *Proc. Natl Acad. Sci. USA* **115**, 2156–2161 (2018).
- Dettlo, M. & Müller-Landau, H. C. Stabilization of species coexistence in spatial models through the aggregation–segregation effect generated by local dispersal and nonspecific local interactions. *Theor. Pop. Biol.* **112**, 97–108 (2016).
- Bolker, B. & Pacala, S. W. Spatial moment equations for plant competition: understanding spatial strategies and the advantage of short dispersal. *Am. Nat.* **153**, 575–602 (1999).
- Uriarte, M. et al. Trait similarity, shared ancestry and the structure of neighbourhood interactions in a subtropical wet forest: implications for community assembly. *Ecol. Lett.* **13**, 1503–1514 (2010).
- Wang, X. et al. Stochastic dilution effects weaken deterministic effects of niche-based processes on the spatial distribution of large trees in species rich forests. *Ecology* **97**, 347–360 (2016).
- Wiegand, T. et al. Spatially explicit metrics of species diversity, functional diversity, and phylogenetic diversity: insights into plant community assembly processes. *Annu. Rev. Ecol. Syst.* **48**, 329–351 (2017).
- Godoy, O. & Levine, J. M. Phenology effects on invasion success: insights from coupling field experiments to coexistence theory. *Ecology* **75**, 726–736 (2014).
- Dieckmann, U. & Law, R. in *The Geometry of Ecological Interactions: Simplifying Spatial Complexity* (eds. Dieckmann, U. et al.) Ch. 21 (Cambridge Univ. Press, 2000).
- Klausmeier, C. A. & Tilman, D. in *Competition and Coexistence* (eds. Sommer, U. & Worm, B.) Ch. 3 (Springer, 2002).
- Fahse, L., Wissel, C. & Grimm, V. Reconciling classical and individual-based approaches in theoretical population ecology: a protocol for extracting population parameters from individual-based models. *Am. Nat.* **152**, 838–852 (1998).

26. Fortunel, C., Valencia, R., Wright, S. J., Garwood, N. C. & Kraft, N. J. B. Functional trait differences influence neighbourhood interactions in a hyperdiverse Amazonian forest. *Ecol. Lett.* **19**, 1062–1070 (2016).
27. Cadotte, M. W., Davies, T. J. & Peres-Neto, P. R. Why phylogenies do not always predict ecological differences. *Ecol. Monogr.* **87**, 535–551 (2017).
28. Wiegand, T. & Moloney, K. A. *A Handbook of Spatial Point Pattern Analysis in Ecology* (CRC Press, 2014).
29. Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. *Statistical Analysis and Modelling of Spatial Point Patterns* (Wiley, 2008).
30. Wang, S. Simplicity from complex interactions. *Nat. Ecol. Evol.* **2**, 1201–1202 (2018).
31. Wilson, W. G. et al. Biodiversity and species interactions: extending Lotka–Volterra community theory. *Ecol. Lett.* **6**, 944–952 (2003).
32. Hubbell, S. P. Neutral theory and the evolution of functional equivalence. *Ecology* **87**, 1387–1398 (2006).
33. Rosindell, J., Hubbell, S. P. & Etienne, R. S. The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* **26**, 340–348 (2011).
34. Chesson, P. Scale transition theory: its aims, motivations and predictions. *Ecol. Complex.* **10**, 52–68 (2012).
35. May, F., Wiegand, T., Lehmann, S. & Huth, A. Do abundance distributions and species aggregation correctly predict macroecological biodiversity patterns in tropical forests? *Glob. Ecol. Biogeogr.* **25**, 575–585 (2016).
36. Murrell, D. When does local spatial structure hinder competitive coexistence and reverse competitive hierarchies? *Ecology* **91**, 1605–1616 (2010).
37. May, F., Wiegand, T., Huth, A. & Chase, J. M. Scale-dependent effects of conspecific negative density dependence and immigration on biodiversity maintenance. *Oikos* **129**, 1072–1083 (2020).
38. Rosindell, J. & Cornell, S. J. Species-area relationships from a spatially explicit neutral model in an infinite landscape. *Ecol. Lett.* **10**, 586–595 (2007).
39. Chave, J. & Leigh, G. L. A spatially explicit neutral model of betadiversity in tropical forests. *Theor. Pop. Biol.* **62**, 153–168 (2002).
40. Hubbell, S. P. et al. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* **283**, 554–557 (1999).
41. Chanthorn, W., Getzin, S., Wiegand, T., Brockelman, W. Y. & Nathalang, A. Spatial patterns of local species richness reveal importance of frugivores for tropical forest diversity. *J. Ecol.* **106**, 925–935 (2018).
42. Getzin, S., Wiegand, T. & Hubbell, S. P. Stochastically driven adult-recruit associations of tree species on Barro Colorado Island. *Proc. R. Soc. B* **281**, 20140922 (2014).
43. Chesson, P. & Neuhauser, C. Intraspecific aggregation and species coexistence. *Trends Ecol. Evol.* **17**, 210–211 (2002).
44. Ruokolainen, L. & Hanski, I. Stable coexistence of ecologically identical species: conspecific aggregation via reproductive interference. *J. Anim. Ecol.* **85**, 638–647 (2016).
45. Stump, S. Multispecies coexistence without diffuse competition; or, why phylogenetic signal and trait clustering weaken coexistence. *Am. Nat.* **190**, 213–228 (2017).
46. Harms, K. E., Condit, R., Hubbell, S. P. & Foster, R. B. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J. Ecol.* **89**, 947–959 (2001).
47. Wiegand, T. et al. Testing the independent species' arrangement assertion made by theories of stochastic geometry of biodiversity. *Proc. R. Soc. B* **279**, 3312–3320 (2012).
48. Wiegand, T., Gunatilleke, C. V. S., Gunatilleke, I. A. U. N. & Huth, A. How individual species structure diversity in tropical forests. *Proc. Natl Acad. Sci. USA* **104**, 19029–19033 (2007).
49. Erickson, D. L. et al. Comparative evolutionary diversity and phylogenetic structure across multiple forest dynamics plots: a mega-phylogeny approach. *Front. Genet.* **5**, 358 (2014).
50. Howe, H. F. Scatter- and clump-dispersal and seedling demography: hypothesis and implications. *Oecologia* **79**, 417–426 (1989).
51. Wiegand, T., May, F., Kazmierczak, M. & Huth, A. What drives the spatial distribution and dynamics of local species richness in tropical forests. *Proc. R. Soc. B* **284**, 20171503 (2017).

Acknowledgements

X.W. was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant XDB31030000), the National Natural Science Foundation of China (Grant 31961133027), the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (Grant ZDBS-LY-DQC019) and the K.C. Wong Education Foundation. T.W. and A.H. were supported by the ERC advanced grant 233066. M.C. and X. M. were supported by the National Natural Science Foundation (32061123003 and 31770478). L.L. was supported by the Joint Fund of the National Natural Science Foundation of China-Yunnan Province (U1902203). The FS plot project was supported by the Taiwan Forestry Bureau, the Taiwan Forestry Research Institute and the Ministry of Science and Technology. The BCI censuses have been made possible through support of the US National Science Foundation (awards 8206992, 8906869, 9405933, 9909947, 0948585 to S.P. Hubbell), the John D. and Catherine D. MacArthur Foundation and the Smithsonian Tropical Research Institute. We also thank the hundreds of people who contributed to the collection and management of the data from the plots. This work builds on discussion of the working group sNiche (Expanding neo-Chessonian coexistence theory towards a stochastic theory for species-rich communities) supported by sDiv, the Synthesis Centre of iDiv (DFG FZT 118). We thank S. Lehmann, D. Alonso and S. Harpole for discussion, and especially S. Stump for valuable feedback on earlier drafts.

Author contributions

T.W. and X.W. conceived and designed the project. T.W. implemented the simulation model, conducted the simulations, analysed the results and prepared figures and tables. T.W. and A.H. led the writing of the manuscript. X.W. assembled and analysed the plot data and conducted the spatial analyses. K.J.A.-T., N.B., M.C., X.C., S.J.D., Z.H., R.H., W.J.K., J. Lian, J. Li, L.L., Y.L., K.M., W.M., X.M., S.-H.S., I.-F.S., A.W., X.W. and W.Y. contributed to the acquisition of the data used in the paper and in revising the manuscript. All authors have given final approval to publish this manuscript and agree to be accountable for the aspects of the work that they conducted.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-021-01440-0>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41559-021-01440-0>.

Correspondence and requests for materials should be addressed to T.W. or X.W.

Peer review information *Nature Ecology & Evolution* thanks Simon Stump and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

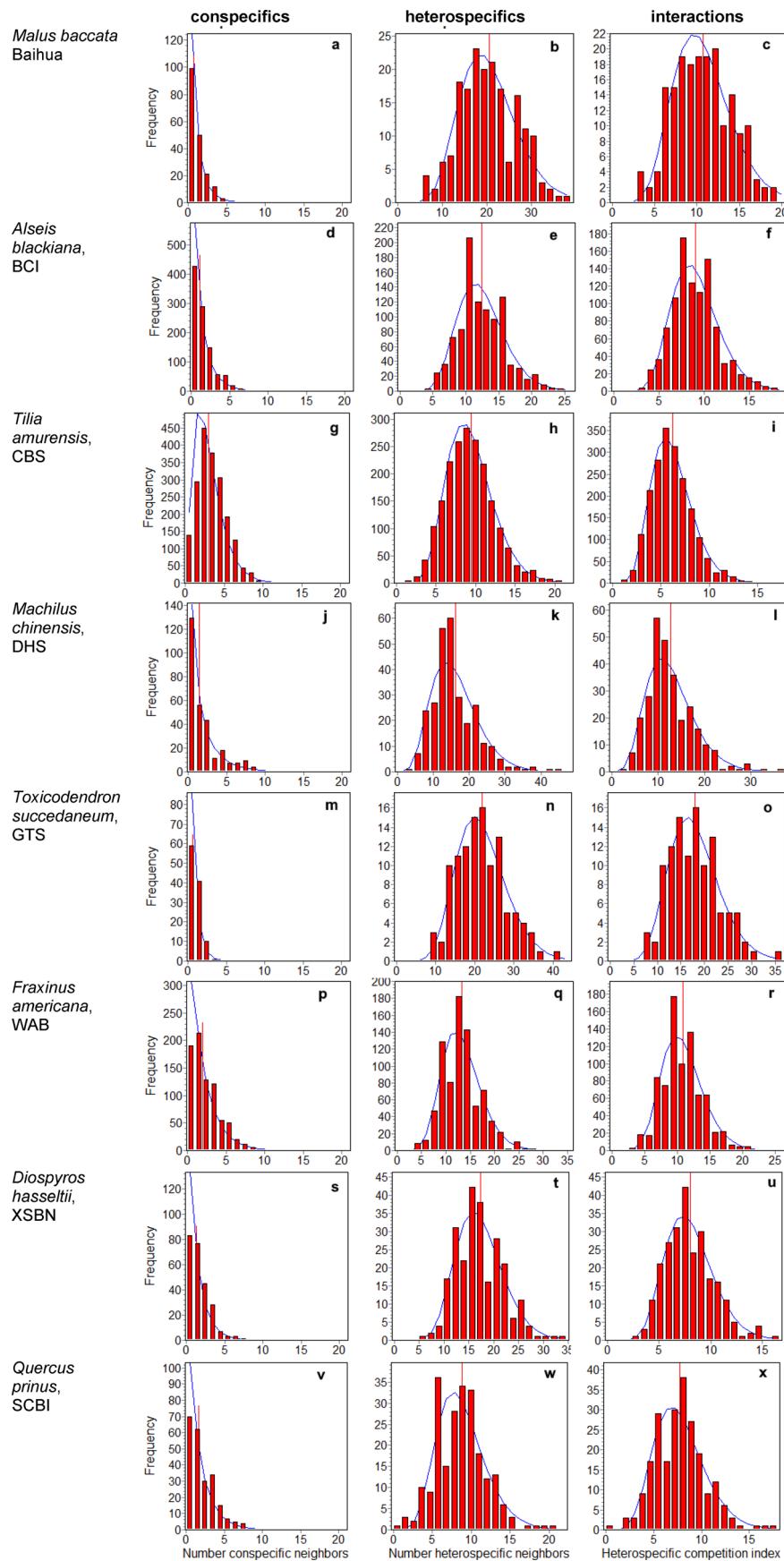
Reprints and permissions information is available at www.nature.com/reprints.

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



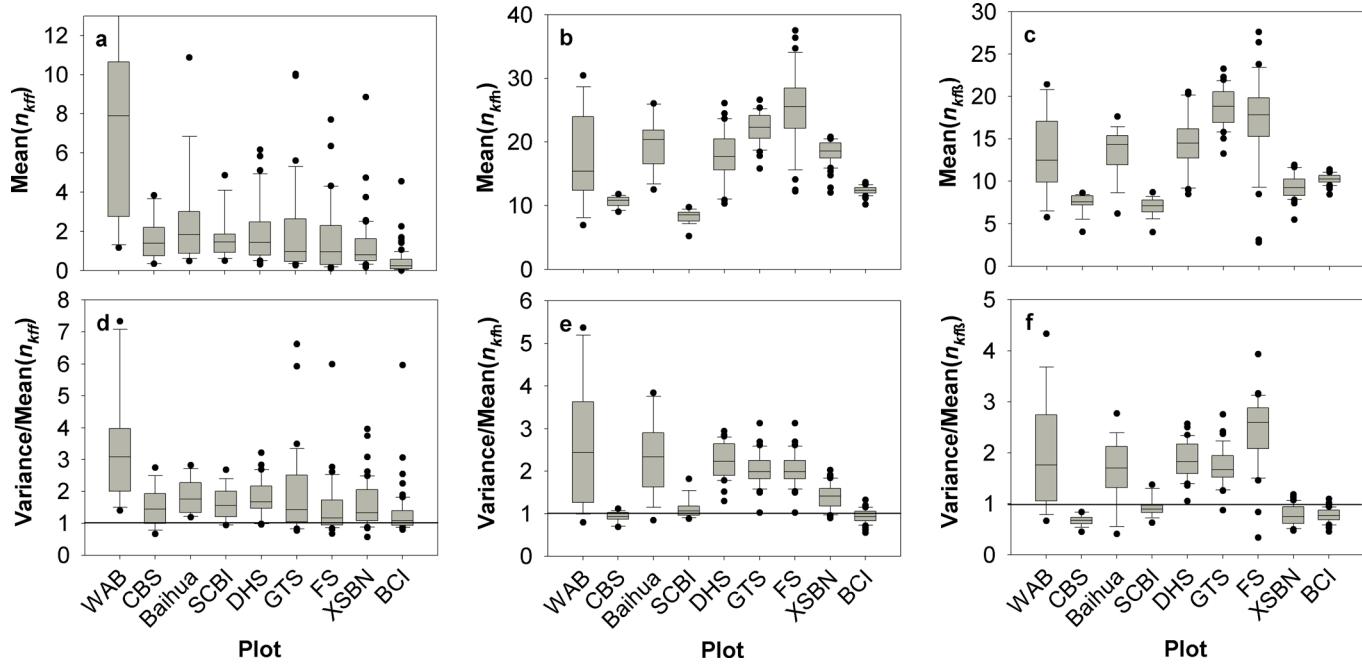
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

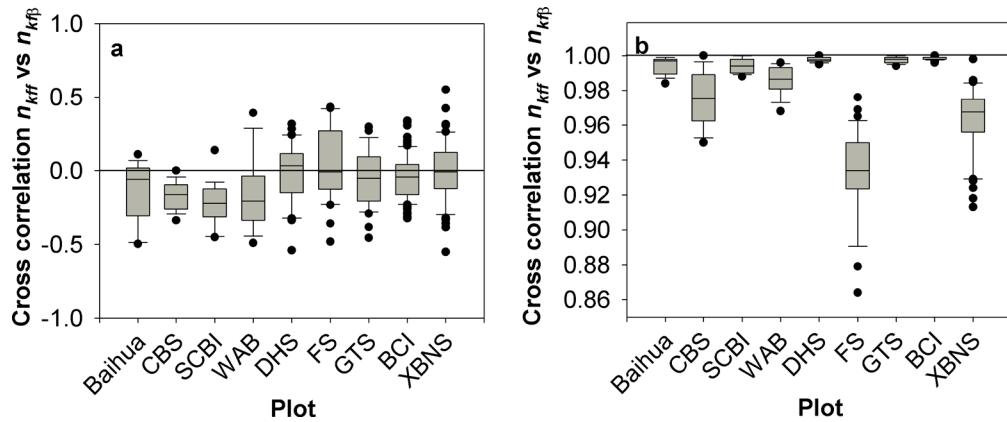


Extended Data Fig. 1 | See next page for caption.

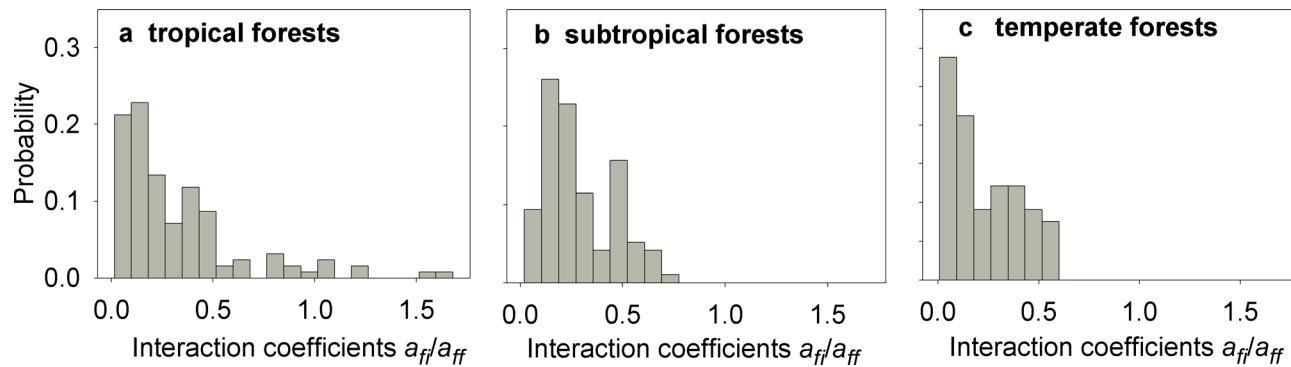
Extended Data Fig. 1 | Examples for intraspecific variability in the crowding indices. **a, b, c,** *Malus baccata*, Baihua plot. **d, e, f,** *Alseis blackiana*, BCI plot. **g, h, i,** *Tilia amurensis*, CBS plot. **j, k, l,** *Machilus chinensis*, DHS plot. **m, n, o,** *Toxicodendron succedaneum*, GTS plot. **p, q, r,** *Fraxinus americana* WAB plot. **s, t, u,** *Diospyros hasseltii* XSBN Plot. **v, w, x,** *Quercus prinus*, SCBI plot. Left: number of conspecific neighbours (n_{kff}), middle: number of heterospecific neighbours (n_{kfh}), and right: heterospecifics neighbours weighted by their relative competitive effect β_f/β_{ff} (n_{kfp}). We used a 10 m plant neighbourhood^{19,26} and phylogenetic similarity as surrogate for pairwise interaction strength^{19,26}. Solid blue lines show gamma distributions with the same mean and variance-to-mean as the observed distributions, and the vertical red lines indicates the mean values. The 95% percentiles for the error indices quantifying the departures from a Gamma distribution for all 289 focal species were 0.051, 0.027, and 0.022 for n_{kff} , n_{kfh} , n_{kfp} respectively, indicating a reasonable fit.



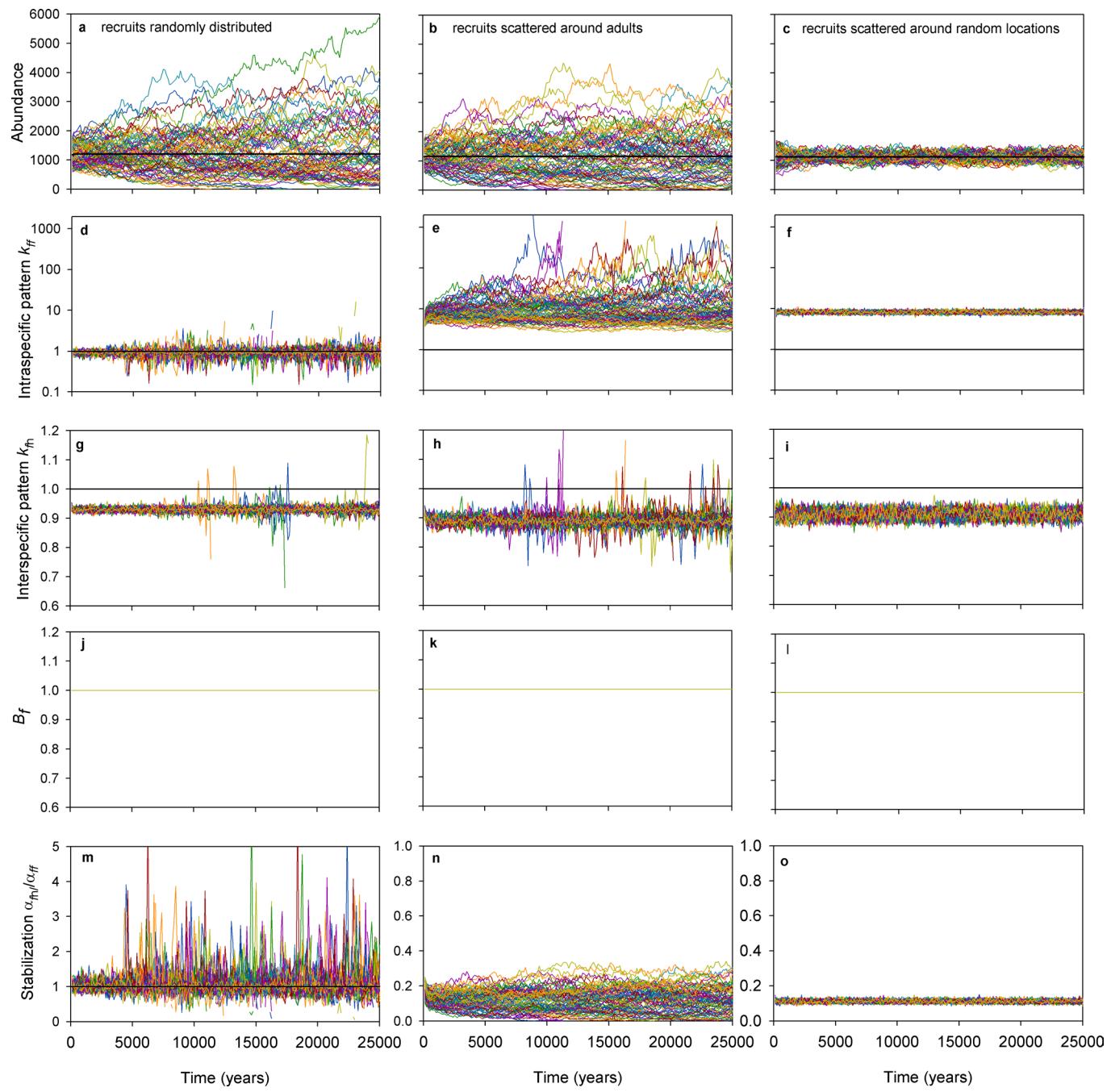
Extended Data Fig. 2 | Characteristics of the neighbourhood crowding indices. Distribution of the mean and the variance-to-mean ratio of the crowding indices for the different species at each forest plot with boxplots indicating 10th, 25th, 50th, 75th, 90th percentiles and outliers. **a.** The mean \bar{n}_{ff} of the conspecific neighbourhood crowding index n_{kff} over species. **b.** The mean \bar{n}_{fh} of the heterospecific neighbourhood crowding index n_{kfh} over species. **c.** The mean \bar{n}_{fb} of the interaction neighbourhood crowding index n_{kfb} over species. **d.** The variance-to-mean ratio b_{kf} of the conspecific neighbourhood crowding index n_{kff} over species. **e.** The variance-to-mean ratio b_{kh} of the heterospecific neighbourhood crowding index n_{kfh} over species. **f.** The variance-to-mean ratio b_{kb} of the interaction neighbourhood crowding index n_{kfb} over species. The neighbourhood radius was $R=10$ m. We used for the analysis all individuals with dbh ≥ 10 cm and included focal species with more than 50 individuals. For plot names see Supplementary Table 1.



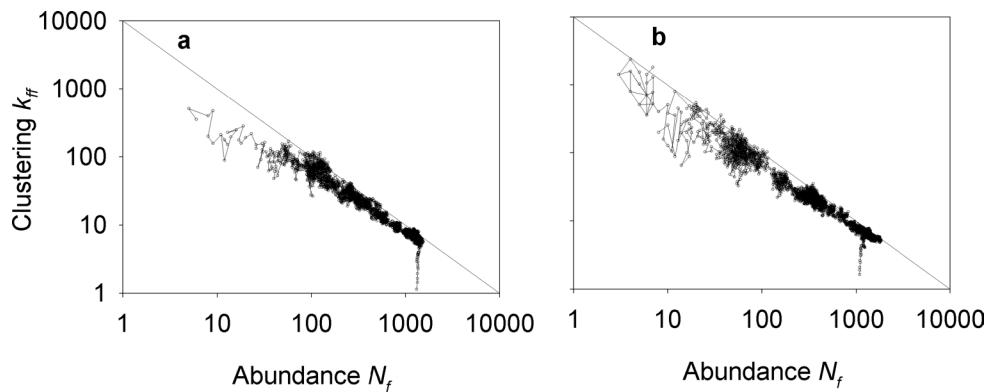
Extended Data Fig. 3 | Correlation between different crowding indices. We estimated for all individuals of a species f the correlation between their crowding indices n_{kff} and n_{kfp} (**a**) and n_{kff} and n_{kfp} (**b**). The crowding index n_{kff} counts the conspecific neighbours of individual k within distance $R=10\text{ m}$, n_{kfp} counts the corresponding number of heterospecific neighbours, and n_{kfp} weights each heterospecific neighbour by its relative competition strength β_p/β_{ff} . The boxplots show the distribution of the Pearson correlation coefficients for each focal species, separately for the nine forest plots, indicating 10th, 25th, 50th, 75th, 90th percentiles and outliers. We used for the analysis all individuals with $\text{dbh} \geq 10\text{ cm}$ and included focal species with more than 50 individuals. For plot names see Supplementary Table 1.



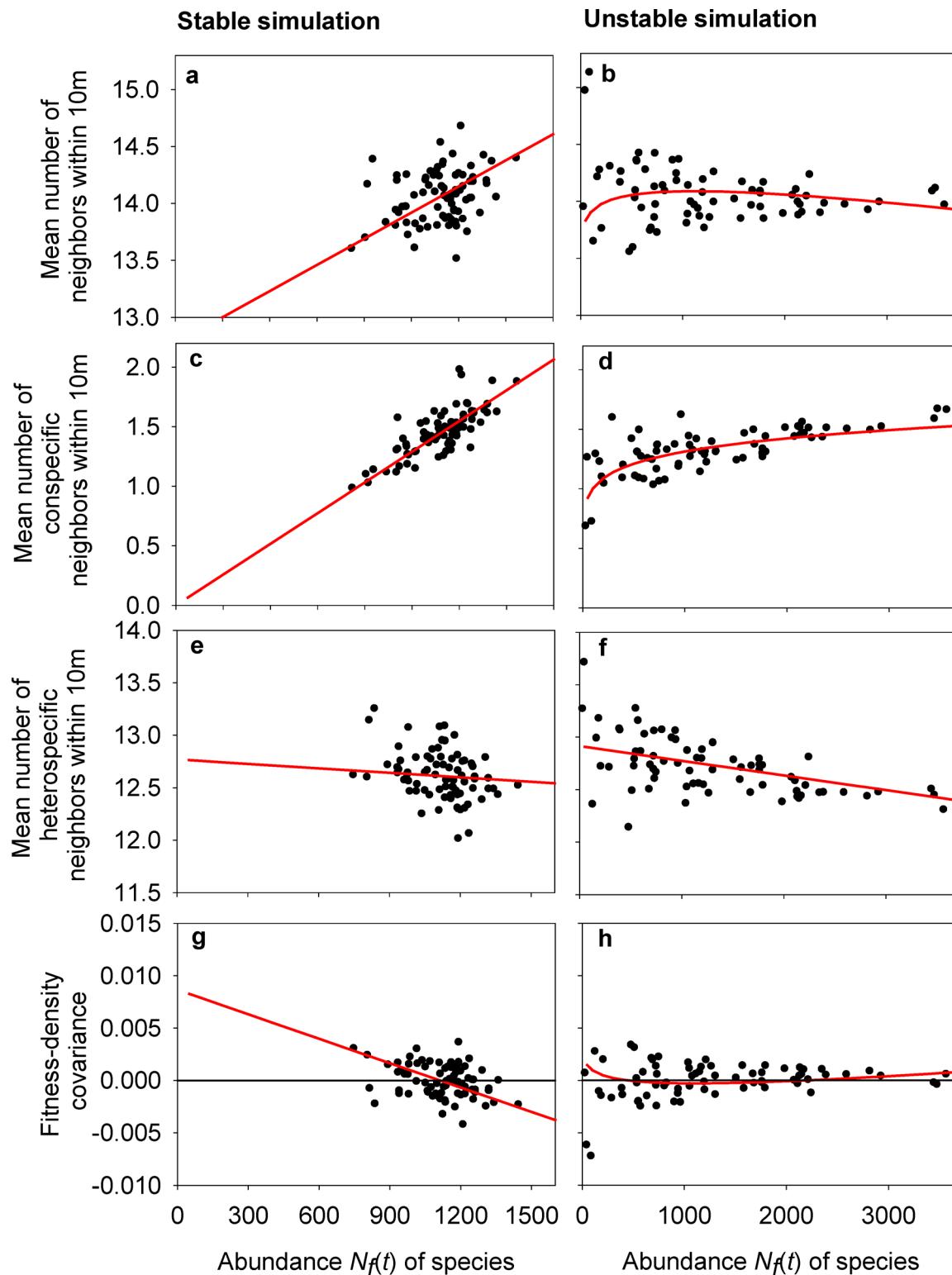
Extended Data Fig. 4 | The distribution of the relative population-level interaction coefficients. **a**, The distribution of the relative population level interaction coefficients α_{ff}/α_{ff} for the focal species of the tropical forests, resulting from equation 6. **b**, same as a, but for subtropical forests. **c**, same as a, but for temperate forests. Other conventions as in Fig. 3.



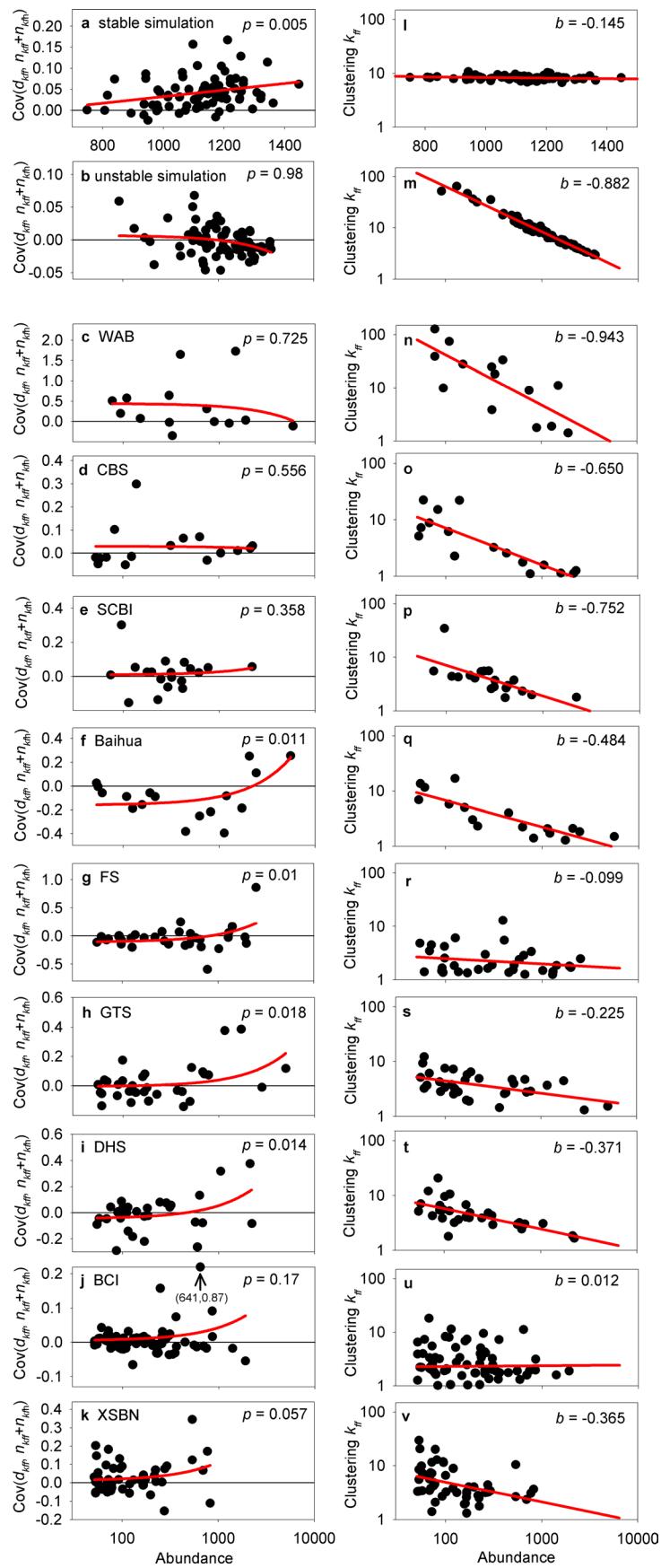
Extended Data Fig. 5 | Spatial patterns can stabilize community dynamics of symmetric species. Individual-based simulations of a symmetric model with initially 80 identical species, simulated on an area of 200 ha without immigration for 5000 time steps (25,000 years). Different colours correspond to different species. **left**, Recruits were randomly distributed, the dynamics is unstable with 2 extinctions. **middle**, Recruit were mostly scattered around conspecific adults, the dynamics is unstable with 7 extinctions. **right**, recruits were mostly scattered around random cluster centre, the dynamics is stable without extinctions. **a–c**, Species abundances with the bold black line indicating the expected mean abundance $J/80$. **d–f**, Intraspecific pattern k_{ff} . **g–i**, Interspecific pattern k_{fn} of heterospecifics with respect to the focal species f . **j–l**, the mean relative interaction strength B_f of a heterospecific neighbour of an individual of species f . **m–o**, Stabilization, being the population-level heterospecific interaction strength relative to the corresponding conspecific interaction strength (that is, α_{fn}/α_{ff}). Note the different scale of the y-axes. Spatial patterns were measured at a 10 m neighbourhood.



Extended Data Fig. 6 | The relationship between clustering and abundance. Shown are detailed results of the simulation of Extended Data Fig. 5b for two species that went extinct. **a**, Dynamics of the relationship between abundance $N_f(t)$ and clustering k_{ff} for species 1. **b**, same for species 3. The solid line indicates the power law $k_{ff}=10,000/N_f(t)$.



Extended Data Fig. 7 | Mechanism underlying the fitness-density covariance in our simulation model. Quantities involved in the estimation of the fitness-density covariance (equation 21) for model simulations of the stable scenario (**left**) and the unstable scenario (**right**). The data were taken from time step 5000 of the simulations shown in Extended Data Fig. 5c (stable dynamics) and of Extended Data Fig. 5b (unstable dynamics). Panels show the mean number $\bar{n}_{ff} + \bar{n}_{fh}$ of all neighbours (**a, b**), of conspecific neighbours \bar{n}_{ff} (**c, d**), of heterospecific neighbours \bar{n}_{fh} (**e, f**), and the resulting fitness-density covariance $\tilde{\lambda}_f - \lambda_f$ (**g, h**), of all species plotted over their abundance. The red lines show the expected relationship based on equations (21) and (22). We fitted power law relationships $k_{ff}(N_f) = a N_f^b$ to the data. For stable dynamics we find $k_{ff}(N_f) = 8.2$, $k_{fh}(N_f) = 0.905$ and for unstable dynamics $k_{ff}(N_f) = 3723/N_f^{0.883}$ and $k_{fh}(N_f) = 0.890$.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Rare species advantage and abundance –clustering relationships in model simulations and ForestGEO plots. The covariance between the local dominance d_{kff} of species f (that is, $d_{kff} = n_{kff}/(n_{kff} + n_{kfh})$) and the total number of neighbours (that is, $n_{kff} + n_{kfh}$). A positive relationship of the covariance with abundance indicates that, when a species becomes rare, areas of higher conspecific crowding have fewer competitors. Shown are results of model simulations (**a, b**) and the nine ForestGEO forest plots (**c–k**). The p-value is for the null hypothesis that the slope of the linear regression (red lines) is not positive. (**l–v**) The corresponding relationships between clustering k_{ff} and abundance N_f . The value of b is the slope of the power law $k_{ff}(N_f) = a N_f^b$ (red line). We used for the analysis focal species f with more than 50 individual. For plot names see Supplementary Table 1, and for raw data and details of the linear regressions see Supplementary Data Table 1.

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 our [Editorial Policies](#) 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 The study did not involve field data collection. We wrote Pascal (Dephi 5.0) code for the simulation model to conduct the example simulations (results reported in Extended Figures 5, 6, 7, and Supplementary Figures S1-S5). The code of the simulation model is provided as supplementary file.

Data analysis The code for estimating the crowding indices and other indices of spatial patterns from the ForestGeo data (results reported in Figs. 1, 2, Extended Data Figures 1, and 2) is included in the code of the simulation model. The same procedures can be used for simulated data and census data. The simple analyses for Figs. 2, 3, 4 and Extended Data Figures 2, 3, 4 and 8 are provided as Supplementary Data Table. This table contains also the raw data (including sample sizes) and estimation of the p-values shown in Extended Data Figure 8.

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 and 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 data that support the findings in this manuscript (and the raw data for figures 2, 3 and 4 and Extended Data Figures 2, 3, 4 and 8) can be found in the Supplementary Information as Data Table. To generate this data, we used the raw census data of the ForestGEO network that can only be shared on request because most PI's did not make them publicly available. For data request see <https://forestgeo.si.edu/sites-all>.

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

This study develops a theoretical multiscale framework to reveal how pattern-forming processes operating at the level of individual trees translate into mesoscale spatial patterns and how they influence macroscale population and community dynamics. The approach includes analysis of mathematical and simulation models and comparison of the simulation results with data of fully-mapped forest plots of the Forest Global Earth Observatory (ForestGEO).

Research sample

We used ForestGeo (<https://forestgeo.si.edu/explore-data>) data sets of nine large forest dynamics plots of areas between 20 and 50 ha.

Sampling strategy

The nine forest plots included in the study have been completely censused for trees at least 10 cm in diameter, so there is no sampling within each forest.

Data collection

The study did not involve data collection. The data from model simulations was collected as described in the methods.

Timing and spatial scale

The study did not involve data collection. We used for each forest dynamics plot (size ranging between 20 and 50ha) data of one census.

Data exclusions

We used only trees with sizes $\geq 10\text{cm dbh}$ (diameter at breast height), which is an established approach in the field. This size threshold excludes most of the saplings and enables comparisons with previous spatial analyses. Tree species with a minimum of 50 individuals ($\text{dbh} \geq 10\text{ cm}$) were used in the study to ensure that estimates of spatial patterns were reliable.

Reproducibility

The estimation of the measures of spatial patterns and the simulation model is specified exactly in the publication such that it could be reproduced. Additionally the simulation code (that includes estimation of measures of spatial patterns) is added as Supplementary File.

Randomization

Not applicable to this study.

Blinding

Blinding was not applicable to this study.

Did the study involve field work? Yes No

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 | Involved 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 and archaeology |
| <input checked="" type="checkbox"/> | <input 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 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | | |
|-------------------------------------|---|
| n/a | Involved 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 |