Article

Latitudinal scaling of aggregation with abundance and coexistence in forests

https://doi.org/10.1038/s41586-025-08604-z

Received: 8 July 2023

Accepted: 7 January 2025

Published online: 26 February 2025

Open access

Check for updates

Thorsten Wiegand[1,2], Xugao Wang[3]✉, Samuel M. Fischer[1], Nathan J. B. Kraft[4], Norman A. Bourg[5], Warren Y. Brockelman[6], Guanghong Cao[7], Min Cao[8], Wirong Chanthorn[9], Chengjin Chu[10], Stuart Davies[11], Sisira Ediriweera[12], C. V. Savitri Gunatilleke[13], I. A. U. Nimal Gunatilleke[13], Zhanqing Hao[14], Robert Howe[15], Mingxi Jiang[16], Guangze Jin[17], W. John Kress[18], Buhang Li[10], Juyu Lian[19], Luxiang Lin[8], Feng Liu[20], Keping Ma[21], William McShea[5], Xiangcheng Mi[21], Jonathan A. Myers[22], Anuttara Nathalang[6], David A. Orwig[23], Guochun Shen[24], Sheng-Hsin Su[25], I-Fang Sun[26], Xihua Wang[24], Amy Wolf[15], Enrong Yan[24], Wanhui Ye[19], Yan Zhu[21] & Andreas Huth[1,2,27]

The search for simple principles that underlie the spatial structure and dynamics of plant communities is a long-standing challenge in ecology[1–6]. In particular, the relationship between species coexistence and the spatial distribution of plants is challenging to resolve in species-rich communities[7–9]. Here we present a comprehensive analysis of the spatial patterns of 720 tree species in 21 large forest plots and their consequences for species coexistence. We show that species with low abundance tend to be more spatially aggregated than more abundant species. Moreover, there is a latitudinal gradient in the strength of this negative aggregation–abundance relationship that increases from tropical to temperate forests. We suggest, in line with recent work[10], that latitudinal gradients in animal seed dispersal[11] and mycorrhizal associations[12–14] may jointly generate this pattern. By integrating the observed spatial patterns into population models[8], we derive the conditions under which species can invade from low abundance in terms of spatial patterns, demography, niche overlap and immigration. Evaluation of the spatial-invasion condition for the 720 tree species analysed suggests that temperate and tropical forests both meet the invasion criterion to a similar extent but through contrasting strategies conditioned by their spatial patterns. Our approach opens up new avenues for the integration of observed spatial patterns into ecological theory and underscores the need to understand the interaction among spatial patterns at the neighbourhood scale and multiple ecological processes in greater detail.

Species-rich plant communities such as tropical forests have been investigated by ecologists for decades, but explaining their high species richness remains a challenge for ecological theory[1–6,15,16]. Although numerous studies have been devoted to this issue, mechanistic connections among features of plant communities and species coexistence are incompletely understood[7–9]. For example, a key feature in forests is the spatial aggregation of tree species, which has long been used to infer mechanisms that contribute to coexistence[17–21]. This is because aggregation is related to ecological processes such as negative conspecific density dependence[17,18,22], dispersal limitation[2,23], mycorrhizal associations[12–14] and habitat association[24,25]. Conspecific spatial aggregation Ω is usually defined as the average density D of conspecific trees in neighbourhoods around individual trees of the same species divided by the mean tree density λ of the species in the forest plot[20,25]. Hence, Ω describes the extent to which trees of the same species tend to occur in spatial clusters. Several theoretical and observational studies suggest that conspecific aggregation is related to species abundance, whereby species with lower abundance show higher levels of aggregation[8,19–21,26–28]. However, other studies have reported only weak relationships between aggregation and species abundance[29] (Supplementary Text). Nevertheless, a relationship between aggregation and abundance could have important consequences for the maintenance of high species richness. This is because recent theoretical work[7–9] has indicated a connection between conspecific aggregation and the rare species advantage required for coexistence.

Different aggregation–abundance relationships are possible. At one extreme, when conspecific clusters form mostly near adults (for example, owing to short-distance dispersal[7,8,26] and/or mycorrhizal associations[12–14]), the density D of conspecific trees in the neighbourhood of individual trees will be similar for species with low and high abundance, but species with lower abundance will have fewer clusters (compare Fig. 1a and Fig. 1b). Thus, aggregation (that is, $\Omega = D/\lambda$)[20,25] increases if abundance (and therefore the mean tree density λ) decreases. At the other extreme, when local clusters are created away from conspecific adults[8] (for example, owing to clumped animal seed dispersal (zoochory) or canopy gaps[24,30–33]), fewer seeds will reach the

A list of affiliations appears at the end of the paper.

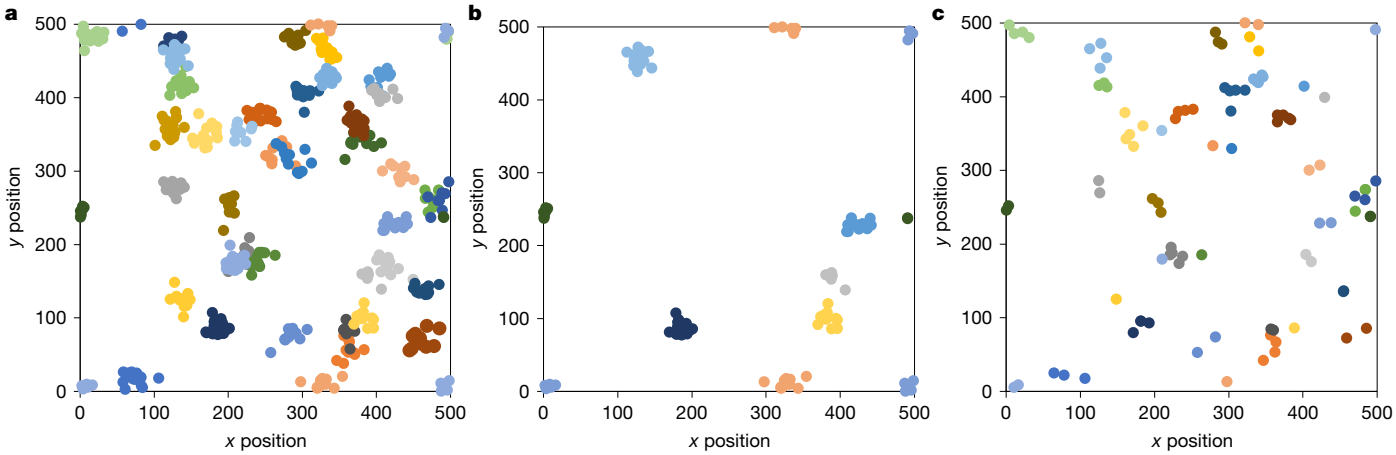


Fig. 1 | Different responses of conspecific spatial aggregation to changes in abundance. a, Illustration of a simulated pattern of a species with abundance $N = 500$ individuals in a 25-ha area with mean neighbourhood density $D = 0.0116$ trees per m^2, mean tree density $\lambda = 0.002$ trees per m^2 and aggregation $\Omega = D/\lambda = 5.8$ (colours represent different clusters of individual trees of the same species). **b**, Entire clusters of the pattern in **a** were removed ($N = 100$). Through this step, the neighbourhood density was approximately maintained ($D = 0.0107$), but because λ was reduced by a factor of 1/5, aggregation increased approximately 5 times ($\Omega = 26.7$). **c**, Individuals of the pattern in **a** were randomly removed ($N = 100$). Through this step, D and λ were reduced by a factor of 4.93 and 5.0, respectively, which approximately maintained aggregation ($\Omega = 5.9$). We estimated D and Ω for 10-m neighbourhoods around the focal individuals.

cluster locations if the species has lower abundance. Consequently, the neighbourhood density D will become proportional to abundance and therefore aggregation Ω becomes independent of abundance (compare Fig. 1a with Fig. 1c).

Given the links between conspecific aggregation, negative density dependence and coexistence[7–9,26], the aggregation–abundance relationship may be related to the latitudinal diversity gradient, which is proposed to be driven by ecological, evolutionary, regional or historical effects[34]. Here we conduct a comprehensive analysis of both how spatial neighbourhood patterns of trees derived from large forest inventories[6] and the relationship between aggregation and abundance change with latitude. We propose underlying ecological mechanisms and integrate our results into mathematical theory to investigate how the aggregation–abundance relationship may affect the rare species advantage and thereby species coexistence (Box 1).

Although aggregation can be defined in various ways[19,20,25,29,35] (Extended Data Fig. 1), we derive measures of spatial patterns from established approaches[8,36–38] that model the effect of neighbours on the performance of individual plants (Box 1). This links the competition of individual trees with the dynamics of species at the community scale. We illustrate our new theory using the example of neighbourhood competition, in which tree survival is reduced in areas of high tree density by competition for space, light or nutrients[39], or through predators or pathogens[17,18].

A latitudinal trend in aggregation

Using data on 720 focal species in 21 temperate, subtropical and tropical forest plots with sizes of 20–50 ha from a global network of forest research plots (CTFS-ForestGEO)[6] (Extended Data Table 1), we found that species with lower abundance tended to be more aggregated than species with higher abundance (Fig. 2). Notably, when describing the relationship between observed aggregation k_{ff}^* and abundance N_f^* for each forest plot by a power law $k_{ff}^* = a\,N_f^{*e}$ (refs. 21,27,28) (Fig. 2a,b), the exponent e followed a marked latitudinal gradient (Fig. 2c and Extended Data Fig. 2). Tropical forests showed a weak negative relationship between aggregation and abundance (that is, exponent values close to zero; Fig. 2a and Extended Data Fig. 3a–f). By contrast, in temperate forests, species with low abundance showed generally high aggregation, and aggregation strongly decreased with increasing abundance (that is, exponent values mostly below −0.58; Fig. 2b and Extended Data Fig. 3n–u). In contrast to conspecifics, heterospecific associations were not related to abundance (except for some weak correlations in temperate forests; Extended Data Fig. 4).

The latitudinal gradient in the relationship between conspecific aggregation and abundance (Fig. 2c) and the absence of such a relationship for heterospecifics (Extended Data Fig. 4) suggested that simple principles may drive the complex spatial structure and dynamics of plant communities across latitudinal gradients. We also observed similar latitudinal gradients in the proportion of species that show mostly animal seed dispersal[11] and in the proportion of species with an arbuscular mycorrhizal (AM) association[14] (Extended Data Table 1). Temperate forests are usually dominated by ectomycorrhizal (EM) tree species, whereas tropical forests are dominated by AM tree species[14].

For the combination of the two traits (that is, zoochory and AM association), there was an even stronger latitudinal gradient (Fig. 3a), which suggested an explanation of the observed latitudinal gradient in the aggregation–abundance relationship (Fig. 3b). Species-specific EM fungi facilitate conspecific recruitment by forming a physical sheath around young feeder roots[14,40] and counteracting negative competitor-driven or pathogen-driven effects[12–14], thereby leading to increased aggregation. Therefore, seed dispersal close to conspecific adults should be advantageous in temperate forests, where many species show an EM association. By contrast, seed dispersal farther away from conspecific adults should be more advantageous for AM trees, given that an AM association provides less protection against competitors or pathogens that accumulate near conspecifics than an EM association[13,40]. In summary, we propose that the key mechanism that leads to the different responses of aggregation to abundance is the way that local clusters emerge with respect to conspecific adults. That is, in tropical forests, mechanisms such as animal seed dispersal lead to the emergence of clusters away from adults, whereas in temperate forests, clusters form close to adults[8] (Fig. 1).

Regardless of the ultimate mechanisms that generate it, the systematic change in the relationship between aggregation and abundance with latitude (Fig. 2c) has important implications for coexistence dynamics and theory. Stable coexistence requires that the abundance of a newly invading (or an almost extinct) species increases[1,5,8,9] (that is, a rare species advantage). Common non-spatial models that feature this invasion criterion ignore the possibility of a negative aggregation–abundance relationship by assuming that the invading species does not suffer from conspecific competition[9]. This assumption was

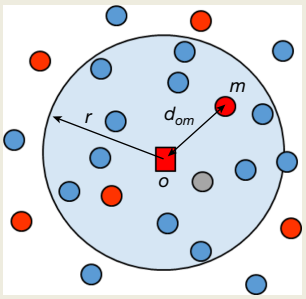
Box 1

Crowding indices, spatial patterns and their link with macroscale dynamics

Following earlier work on neighbourhood crowding indices[36,37], we assumed that the survival of an individual (red square) depends on its conspecific (red) and heterospecific (blue) neighbours within a distance r (blue shaded area). For an individual o of the focal species f, the crowding index C_{of} counts all conspecifics (index m) within distances r, but weights them by their distance d_{om} (equation (3)), assuming that distant neighbours compete less (n_f is the number of neighbours of individual o of species f within distance r). The crowding index H_{of} does the same with all heterospecifics (equation (4)), but the crowding index I_{of} weights the n_i neighbours of species i additionally by their relative competition strength β_{fi}/β_{ff} (equation (5a)), where β_{fi} is the neighbourhood-scale competition coefficient between species f and i, which gives the negative impact of one neighbour of species i on the survival of individuals of the focal species f. Total crowding $C_{of} + I_{of} = C_{of} + B_f H_{of}$ (equation (5a)) then determines the survival s_{of} of the focal individual[36] (equation (6)).

We found that the population averages $\overline{C_f}$ and $\overline{I_f}$ of the individual crowding indices determine the average survival rate $\overline{s_f}$ of species f (equations (6) and (9)), weighted by coefficients γ_{fC} and γ_{fI}, respectively, which arise through the nonlinear averaging of the survival s_{of} over all individuals o of the focal species[8]. To incorporate the average survival rate into a model of community dynamics (equation (11)), we decompose $\overline{C_f}$, $\overline{H_f}$ and $\overline{I_f}$ into species abundance N_f and measures of spatial patterns (equations (3), (4), (5b) and (10)), where $J = \sum_i N_i$ is the total community size. Conspecific aggregation k_{ff} describes the extent to which trees of the focal species (subscript ff) tend to occur in spatial clusters and heterospecific association k_{fh} describes the extent to which heterospecifics are spatially associated with trees of the focal species (subscript fh). We define k_{ff} and k_{fh} by dividing the average crowding indices $\overline{C_f}$ and $\overline{H_f}$ by their expectation in the absence of spatial patterns[20,25] (square brackets in equations (3) and (4)), where c is a scaling factor ($c = 2\pi r/A$; Methods) and A is the area of the plot.

The quantity B_f (defined in equation (5a)) is the average competitive strength of one heterospecific neighbour relative to that of one conspecific. We compared two scenarios for the neighbourhood-scale competition coefficient β_{fi}: one in which conspecifics and heterospecifics compete equally ($\beta_{fi}/\beta_{ff}=1$), and one in which phylogenetic similarity is a proxy for β_{fi}/β_{ff} because it is difficult in practice to estimate β_{fi}/β_{ff} for species-rich forests[37].



	Neighbourhood scale	Population level	Equation
Conspecific	$C_{of} = \sum_{m=1}^{n_f} \frac{1}{d_{om}}$	$\overline{C_f} = k_{ff}\,[c\,N_f]$	(3)
Heterospecific	$H_{of} = \sum_{i\neq f} \sum_{m=1}^{n_i} \frac{1}{d_{om}}$	$\overline{H_f} = k_{fh}\,[c\,(J - N_f)]$	(4)
Niche differences	$I_{of} = \sum_{i\neq f} \sum_{m=1}^{n_i} \frac{\beta_{fi}}{\beta_{ff}} \frac{1}{d_{om}}$	$B_f := \overline{I_f}\,/\overline{H_f}$	(5a)
		$\overline{I_f} = k_{fh}\,B_f\,[c\,(J - N_f)]$	(5b)
Survival	$s_{of} = s_f e^{-\beta_{ff}(C_{of}+I_{of})}$	$\overline{s_f} = s_f e^{-\beta_{ff}(\gamma_{fC}\overline{C_f}+\gamma_{fI}\overline{I_f})}$	(6)

met in our data for tropical forests, where aggregation was weakly related to abundance (Fig. 2a,c). In this case, mean conspecific neighbourhood densities $\overline{C_f}$ were almost linearly related to species abundance N_f (as constant aggregation k_{ff}^* led to $\overline{C_f} = ck_{ff}^* N_f^*$; Box 1 equation (3); Fig. 1c). However, an increase in aggregation with decreasing abundance, as observed for temperate forests (that is, higher latitudes; Fig. 2b,c), challenges the common assumption of invasion analysis. This is because local conspecific neighbourhood densities were almost independent of abundance, given that the exponent of the aggregation–abundance relationship approached values of −1 (from $\overline{C_f} = ck_{ff}^* N_f^*$ and $k_{ff}^* = a/N_f^*$ follows $\overline{C_f} = ca$; Box 1 equation (3); Fig. 1b). Thus, individuals of species with low and high abundance experience similar degrees of conspecific competition. Consequently, existing invasion analysis can break down[9]. Spatial aggregation of trees is therefore closely linked to species coexistence, and we need new theories to determine whether and under which circumstances species with low abundance are likely to increase[9].

Including aggregation into the theory

A theory that describes how the response of conspecific aggregation to abundance influences species coexistence requires a dynamic population model that relates competition at the population level to spatial patterns in the neighbourhood of individual trees. Here we derived such an approach[8], which incorporated information on spatial patterns at the scale of individual trees provided by the ForestGEO datasets[6]. We exemplified our theory using a simple model, which uses the following assumptions: (1) reproduction is density-independent[7]; (2) survival of individual trees is reduced in areas of high local tree density, as described by neighbourhood crowding indices[36–38] (Box 1); and (3) immigration may occur at a constant rate:

$$\frac{N_{f,t+\Delta t} - N_{f,t}}{\Delta t} = \widetilde{\lambda}_f(N_{f,t})N_{f,t} = \underbrace{r_f N_{f,t}}_{\text{Reproduction}} - \left(1 - \underbrace{s_f \exp(-\beta_{ff} W_f(N_{f,t}))}_{\text{Survival}}\right) N_{f,t} + \underbrace{\upsilon_f r_f}_{\text{Immigration}} \quad (1a)$$

Here $N_{f,t}$ is the abundance of the focal species f at time t, the time interval Δt is the 5-year census interval, $\widetilde{\lambda}_f(N_{f,t})$ is the per capita population growth rate of species f as a function of species abundance $N_{f,t}$, s_f is a density-independent per capita background survival rate, r_f is the per capita recruitment rate, β_{ff} is the conspecific neighbourhood-scale competition coefficient and υ_f is a parameter that governs the magnitude of a constant immigration rate $r_f \upsilon_f$ (ref. 41). $W_f(N_{f,t})$, the fitness factor[42], is the average of the sum $C_{of} + I_{of}$ of conspecific and heterospecific neighbourhood crowding taken over all individuals o of a focal species f (Box 1, equation (6)) and incorporates information

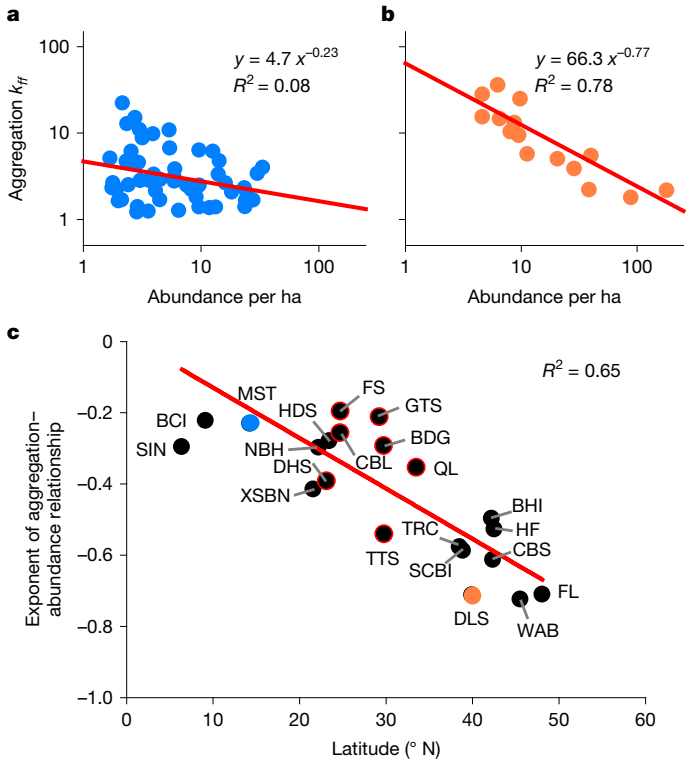


Fig. 2 | Latitudinal variation in the scaling of conspecific aggregation with abundance. a, Aggregation values for species in a tropical forest (Mo Singto (MST) plot) plotted over their respective abundances per ha (dots) and fitted linear regression between $\ln(k_{ff})$ and $\ln(N_f)$ (line). **b**, Same as **a**, but for a temperate forest (Donglingshan (DLS) plot). **c**, Latitudinal gradient in the exponent of the aggregation–abundance relationship for the 21 forest plots. To show the overall trend in the data, we fit a linear regression. Aggregation is defined in Box 1 equation (3) and was estimated for neighbourhoods of 15 m. We used in our analyses 720 species with at least 50 large trees[20] (diameter at breast height ≥ 10 cm). For plot characteristics, sample sizes and plot acronyms see Extended Data Table 1. Circles of subtropical plots are marked with a red edge.

on spatial patterns, abundance and the aggregation–abundance relationship into our model (equations (11b) and (11f)). To keep our example model simple, we did not consider tree size.

We used a spatially explicit and individual-based implementation of our model[8], in which spatial patterns (measured by k^*_{ff}, k^*_{fh}) emerged as a consequence of spatially explicit recruitment of offspring and negative density dependence (Methods). Changing only the way offspring are placed relative to conspecific adults led to the range of observed exponents of the aggregation–abundance relationship (Extended Data Fig. 5). When most offspring were placed close to their parents, locally high adult densities were maintained through the continuous placement of new individuals into these clumps, which was controlled through subsequent thinning due to density-dependent mortality. This mechanism led to an aggregation–abundance relationship similar to that of temperate forests[8], and we found that the community cannot be invaded by a species at low abundance (Extended Data Fig. 5a,f). However, if fewer recruits were placed close to their parents, then the dependence of aggregation on abundance was weaker (Extended Data Fig. 5b–e). Consequently, a species at low abundance can invade because it experiences reduced competition (that is, lower values of total crowding $\overline{C_f} + \overline{H_f}$) (Extended Data Fig. 5j,o). This rare species advantage emerges as a consequence of a positive fitness-density covariance[8,42] if a species at a lower abundance is surrounded by a lower number of conspecific neighbours (that is, lower values of $\overline{C_f}$; Extended Data Fig. 5o,t).

By taking a mean-field approach[43] (that is, diffuse competition at the community scale[8]) and assuming zero-sum dynamics[2] (that is, the total number of trees is fixed; Extended Data Fig. 5f–j), we decoupled the multispecies model (equation (1a)) by means of the fitness functions

$$W_f(N_{f,t}) = c\underbrace{\left[(k^*_{ff} - k^*_{fh})\left(\frac{N_{f,t}}{N^*_f}\right)^{b_f} + k^*_{fh}\right]N_{f,t}}_{\text{Conspecific crowding}} + \underbrace{cB_f k^*_{fh}(J^* - N_{f,t})}_{\text{Heterospecific crowding}} \quad (1b)$$

where k^*_{ff}, k^*_{fh}, N^*_f and J^* are the observed aggregation, heterospecific association, abundance and total community size, respectively, b_f is the exponent of a corrected power-law aggregation–abundance relationship (equation (12)), B_f is the average competitive strength of one heterospecific neighbour of species f relative to that of one conspecific and c is a scaling factor (Box 1). We parameterized our mathematical model (equations (1a) and (1b)) by using information from our large forest plots. All parameters of the model, and the measures of spatial patterns, were species-specific. However, given that we had only limited information, we used the same parameter value for all species for several parameters (that is, r_f, s_f and b_f). For the estimation of β_{ff}, we assumed that the observed abundance values were close to equilibrium (see Methods for details).

An expanded spatial-invasion criterion

The effect of aggregation on coexistence can be analysed by using an invasion criterion[1,9]. An invading (or almost extinct) species will be able to increase from a low abundance if its per capita population growth rate (equation (1a) and Fig. 4a) is sufficiently positive (that is, a rare species advantage). Our spatial-invasion criterion therefore requires that the (scaled) per capita population growth rate $\widetilde{\lambda}_f(N_s)/r_f$ should have at a low abundance N_s at least a value δ larger than zero (equation (17a)). For the case in which spatial patterns are the only mechanism to facilitate coexistence (that is, no niche overlap or immigration; $B_f = 1$, $\upsilon_f = 0$), we obtained the simple invasion criterion

$$\frac{1 - \left(\frac{N_s}{N^*_f}\right)^{b_f+1}}{1+\rho_f} > \delta \quad (2)$$

where N_s is the small invasion abundance (here $N_s = 5$ or 10) and N^*_f the observed abundance, b_f is the exponent of the corrected aggregation–abundance relationship (equation (12)), and the quantity $\rho_f = B_f(k^*_{fh}/(k^*_{ff} - k^*_{fh}))(J^*/N^*_f)$ (equation (17d)) is a risk factor, as larger values of ρ_f make it more difficult to meet the invasion criterion. The risk factor combines the main influence of niche differences (B_f), spatial pattern (k^*_{ff} and k^*_{fh}) and relative abundance (N^*_f/J) on the invasion criterion into a single compound index.

Equation (2) suggests that temperate and tropical forests use contrasting strategies to fulfil the invasion criterion (Fig. 4c). Tropical forests showed smaller negative values of the exponent b_f that allow them to have larger values of the risk factor, whereas temperate forests showed smaller values of the risk factor but larger negative values of the exponent b_f. Moreover, less abundant species (that is, larger values of N_s/N^*_f in equation (2)) must show, across all latitudes, larger aggregation k^*_{ff} than more abundant species to produce smaller values of the risk factor (the red versus black line in Fig. 4c). Moreover, larger niche differences (that is, B_f decreases) in combination with larger aggregation k^*_{ff} enhance coexistence (equations (18) and (19)). We anticipate that this effect would be particularly important for the persistence of low-density species.

To assess the order of magnitude of the effects of spatial structure, niche differences and immigration on the per capita population growth rate, we investigated four scenarios (Table 1). Coexistence was generally facilitated to a similar extent in the analysed forest plots by the observed spatial patterns (Table 1, Fig. 4b and Extended Data Fig. 6a,e). For example, model results for scenario 1 (without niche differences

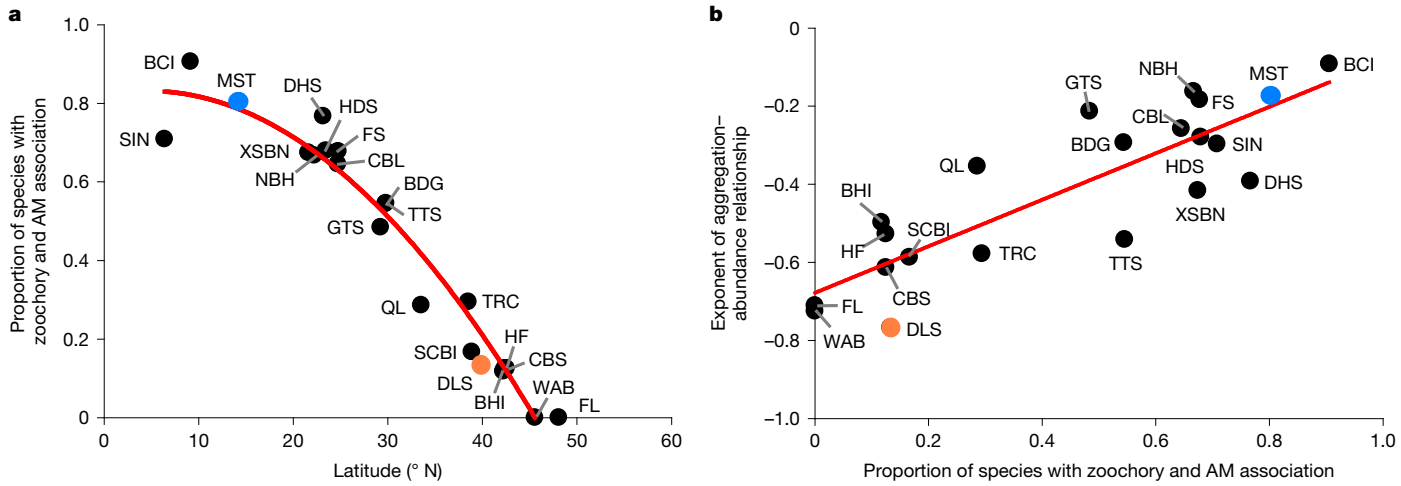


Fig. 3 | Latitudinal variation in the proportion of species showing mainly animal seed dispersal and AM association. a, Latitudinal gradient in the proportion of species per plot that show both mostly animal seed dispersal and AM association. **b**, Relationship between the exponent of the aggregation–abundance relationship and the proportion of species per plot that show mostly animal seed dispersal and AM association for the 21 forest plots. For plot characteristics, sample sizes, raw data, plot acronyms and relationships with animal seed dispersal and AM association, see Extended Data Table 1. To outline the overall trend in the data, we fit a polynomial regression of order 2 in **a** and we fit a linear regression in **b**. Most species showed either AM or EM associations. Coloured discs indicate the example plots presented in Fig. 2a,b.

and without immigration) produced at a small abundance of 10 individuals mean-scaled per capita growth rates of 0.036, 0.043 and 0.05 for tropical, subtropical and temperate forests, respectively (Table 1). Note that these values produced stable dynamics in our simulations (Extended Data Fig. 5i,j). Species in temperate forests tended to show slightly higher values of the scaled per capita growth rates on average compared with tropical forests (Table 1 and Fig. 4b).

Using expanded versions of the model, adding niche differences ($B_f < 1$) and/or a small constant immigration rate ($\upsilon_f = 0.1$) (Table 1) led to further increases in the scaled per capita growth rates of similar magnitude. Although the effect of a small immigration rate on the per capita population growth rate rapidly declined with increasing species abundance (Table 1 and Extended Data Fig. 6g,h), the effect of spatial patterns and niche differences persisted for larger abundance values (Extended Data Fig. 7).

Discussion

Developing an approach that integrates the observed spatial patterns of individual trees in forests with ecological processes into mathematical theory is a considerable challenge. Here we presented a unified framework that combines spatial point process theory with population models to derive expectations about how interactions between different spatial patterns and processes affect the ability of invading species to expand. The framework relies on spatial patterns, such as conspecific aggregation and heterospecific association, which link

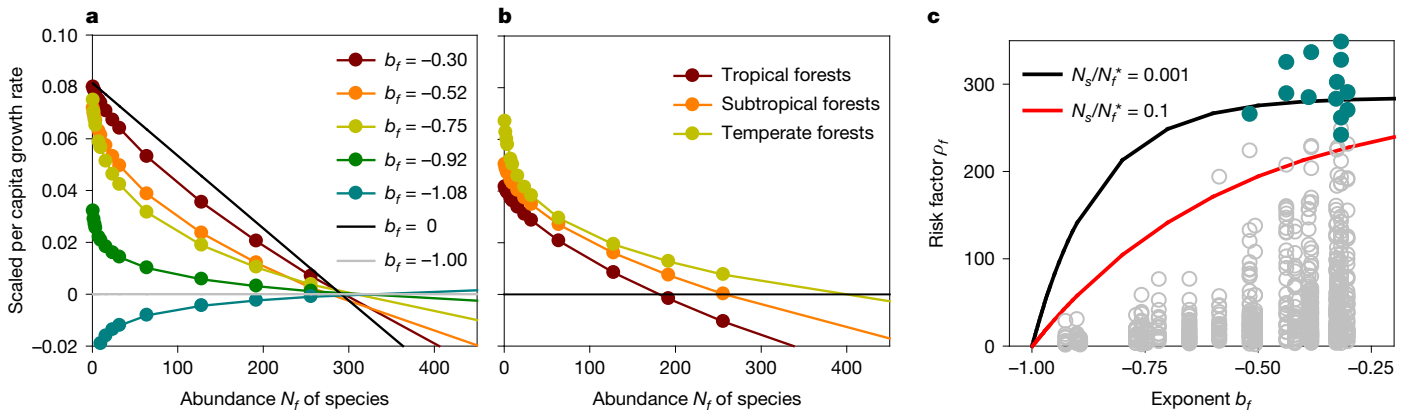


Fig. 4 | The per capita population growth rate and the spatial invasion criterion. a, The scaled per capita population growth rate $\widetilde{\lambda}_f(N_f)/r_f$ for five example species of different forests plotted over abundance N_f. We scaled $\widetilde{\lambda}_f(N_f)$ with the reproduction rate r_f to obtain a quantity that is comparable among forest plots. The species include *Castanopsis acuminatissima* of the MST plot ($b_f = -0.3$), *Ficus langkokensis* of the Xishuangbanna plot ($b_f = -0.52$), *Carya tomentosa* of the Tyson Research Center plot ($b_f = -0.75$), *Ostrya virginiana* of the Wabikon plot ($b_f = -0.92$) and *Maackia amurensis* of the Changbaishan plot ($b_f = -1.08$). We also show the theoretical values of $\widetilde{\lambda}_f(N_f)/r_f$ for $b_f = 0$ (black) and −1 (grey). **b**, Same as **a**, but for the averages over all focal species in tropical, subtropical and temperate forests. **c**, Trade-off between the exponent b_f of the aggregation–abundance relationship (x axis) and the maximal risk factor $\rho_{f,max}$ (equation (19)) that just satisfies the spatial invasion criterion (equations (2) and (18)): temperate forests (smaller b_f) show smaller risk factors ρ_f than tropical forests (larger value of b_f). The lines show $\rho_{f,max}$ for a small scaled growth rate $\delta = 0.0035$ (see equation (19)) and example ratios of $N_s/N_f^* = 5/50$ (red) and $N_s/N_f^* = 5/5{,}000$ (black), where N_s (=5) is the small invasion abundance and N_f^* (=50 and 5,000). The circles show for each species the risk factor ρ_f and the cyan discs mark the 33 out of 720 species that do not satisfy the criterion (19 species with $\rho_f > 350$ are not visible). The data are from scenario 1 (that is, no niche differences, no immigration and observed equilibrium abundances) (Extended Data Table 1).

Table 1 | Effects of spatial structure, niche differences and immigration on scaled per capita population growth rate

	Scenario 1, spatial patterns, no niches			Scenario 2, spatial patterns and niches		
Parameters	N_s=1, v_f=0	N_s=10, v_f=0	N_s=32, v_f=0	N_s=1, v_f=0.1	N_s=10, v_f=0.1	N_s=32, v_f=0.1
All forests	0.049	0.041	0.041	0.090 (0.042)	0.078 (0.037)	0.063 (0.030)
Tropical	0.042	0.036	0.036	0.091 (0.050)	0.080 (0.044)	0.064 (0.035)
Subtropical	0.050	0.043	0.043	0.078 (0.027)	0.068 (0.025)	0.056 (0.021)
Temperate	0.067	0.050	0.050	0.114 (0.047)	0.091 (0.041)	0.074 (0.036)
Immigration[a]	0.1000	0.010	0.003	0.1000	0.010	0.003

We show the average values of the scaled per capita population growth rate $\bar{\lambda}_f(N_s)/r_f$, taken separately for all species in tropical, subtropical and temperate forests and for small abundance values of N_s=1, 10 and 32 individuals. Scenario 1 assumes no niche differences ($\beta_{fi}=\beta_{ff}$), whereas scenario 2 includes niche differences ($\beta_{fi}<\beta_{ff}$). The numbers in parenthesis in scenario 2 show the increase in $\bar{\lambda}_f(N_s)/r_f$ due to niche differences, and the row 'Immigration' shows the increase due to immigration. Note that the corresponding model without spatial patterns, niches and immigration shows $\bar{\lambda}_f(N_s)=0$.

[a]The increase in $\bar{\lambda}_f(N_s)/r_f$ due to immigration with v_f=0.1.

neighbourhood-scale competition of individual trees with species dynamics at the community scale. Our theory led to a closed-form expression for the per capita population growth rate of species as a function of spatial patterns, demography, niche overlap and immigration (equations (1a) and (1b) and (13)), which facilitated a general understanding of spatial coexistence in forests (equations (2) and (18)).

Several important ecological insights resulted from this new theory. First, we showed that the spatial patterns found in forests have in general a stabilizing effect under neighbourhood competition[7,8], as they substantially increased the per capita population growth rate of an invading species compared with a value of zero for the non-spatial case (Table 1). This result challenges the prevalent perspective that spatial patterns alone cannot promote coexistence[44,45]. However, this assertion arises from the assumption of previous models[7,26,35,44,46] that place recruits close to their parents, which leads to destabilizing negative aggregation–abundance relationships (with exponent values close to −1) (Fig. 1 and Extended Data Fig. 5). Our results therefore highlight the crucial role that spatial patterns and the aggregation–abundance relationship have in shaping coexistence outcomes.

Second, our spatial theory suggests that temperate and tropical forests both satisfy the spatial invasion criterion required for coexistence to a similar extent (Table 1), but they do so in contrasting ways. Tropical forests are not subject to destabilizing negative aggregation–abundance relationships (that is, they show low negative values of the exponent b_f) owing to the combined effect of animal seed dispersal and an AM association, but show higher values of the risk factor ρ_f (Fig. 4c and Extended Data Fig. 8b,c). By contrast, temperate forests show a stronger negative dependency of aggregation on abundance, but this disadvantage is outweighed by lower values of the risk factor through higher relative species abundance and/or stronger aggregation facilitated by an EM association[12] (Fig. 4c and Extended Data Fig. 8a,c,d).

Third, the negative aggregation–abundance relationship found here for temperate forests challenges the implicit assumption of common non-spatial implementations of the invasion criterion[1,9] that an invading species does not suffer from conspecific competition[9] (such as shown in Extended Data Fig. 5t). In violation of this assumption, if aggregation is caused by short-distance dispersal, most individuals will be close to a conspecific competitor, even if the species abundance is low (Extended Data Fig. 5p). This mechanism leads to conspecific population-level competition coefficients that are not constant as commonly assumed[47], but increase with decreasing abundance (equation (11c)). We developed a new theory that overcame this issue and enabled us to determine whether and under what circumstances the abundance of invading species are likely to increase in a spatial context (equation (18)). We propose that similar criteria can also be derived for other population models (Supplementary Text).

Our spatial analysis of 21 large forest plots revealed a latitudinal gradient in the strength of the relationship between conspecific aggregation and abundance (Extended Data Fig. 2d). Our model simulations (Extended Data Fig. 5) provided evidence that this latitudinal gradient is related to seed dispersal, and on the basis of our correlation analysis (Fig. 3), we propose that it arises as an interaction between animal seed dispersal[11] and mycorrhizal associations[12–14]. Nonetheless, given that multiple ecological variables correlate with latitude, other mechanisms could contribute to this pattern. However, a recent global study[10] outlined the joint evolution of mycorrhizal symbiosis, seed dispersal and pollination in tree species because they each interact in ways that are mediated by the spatial structure of tree populations. Their results support our dispersal–mycorrhiza hypothesis, whereby most AM-associated trees have biotic seed dispersal and biotic pollination, with long dispersal distances, whereas most EM-associated trees have abiotic seed dispersal and wind (abiotic) pollination mode, with shorter dispersal distances[10]. The dispersal–mycorrhiza hypothesis therefore provides an extra dimension to the study of negative conspecific density dependence. An integrated understanding of the interacting effects of animal seed dispersal and mycorrhizal associations will be fundamental to our understanding of the forces that structure forest diversity and composition.

The simple spatial invasion criterion corroborated our hypothesis that simple principles may drive the complex spatial structure and dynamics of plant communities across latitudinal gradients, but this structural simplicity hides substantial biological complexity. Moving forward, we suggest that methods that model the forces that structure forest diversity and composition should incorporate mechanistic representations of the processes that can potentially drive the aggregation–abundance relationship. Here we showed that they may include seed dispersal[11,48], habitat association[24] and its interactions with mycorrhizal associations[10,12–14] as main processes. A priority would be to consider species-specific aggregation–abundance relationships and to study how the exponent b_f depends on the traits and the dispersal syndrome of the species. More detailed versions of our spatial theory may also consider variation in tree size, as included in classical neighbourhood crowding indices[36,37], to describe asymmetric competition. Preliminary analyses suggest that our key results hold when including the size of trees and their growth.

Overcoming limitations of non-spatial models, as often used in contemporary coexistence theory[1,4,15,16,47,49], requires approaches that explicitly consider the lower-level processes that generate the phenomenon of interest[50]. Our results demonstrated the utility of functions that describe the parameters of the population models, here the population-level competition coefficients α_{fi} (between species i and f; equation 11d), as a function of neighbourhood-scale competition coefficients β_{fi} and measures of spatial patterns (that is, k_{ff}, k_{fh} and B_f; equations (11c) and (11d))[8]. Our scaling approach had the advantage that the parameters of these functions can be determined from field measurements, such as the ForestGEO plots[6] in our case. Taking this approach, we demonstrated that spatial patterns that emerge from neighbourhood-scale processes have a key role in species coexistence,

which underscores the need to understand the mechanisms that underpin the spatial heterogeneity of forests in greater detail.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08604-z.

1. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
2. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).
3. Wright, J. S. Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia* **130**, 1–14 (2002).
4. 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).
5. Barabás, G., D'Andrea, R. & Stump, S. M. Chesson's coexistence theory. *Ecol. Monogr.* **88**, 277–230 (2018).
6. Davies, S. J. et al. ForestGEO: understanding forest diversity and dynamics through a global observatory network. *Biol. Conserv.* **253**, 108907 (2021).
7. Detto, M. & Muller-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).
8. Wiegand, T. et al. Consequences of spatial patterns for coexistence in species-rich plant communities. *Nat. Ecol. Evol.* **5**, 965–973 (2021).
9. Ellner, S. P., Snyder, R. E., Adler, P. B. & Hooker, G. Toward a "modern coexistence theory" for the discrete and spatial. *Ecol. Monogr.* **92**, e1548 (2022).
10. Yamawo, A. & Ohno, M. Joint evolution of mutualistic interactions, pollination, seed dispersal mutualism, and mycorrhizal symbiosis in trees. *New Phytol.* **243**, 1586–1599 (2024).
11. Rogers, H. S., Donoso, I., Traveset, A. & Fricke, E. C. Cascading impacts of seed disperser loss on plant communities and ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **52**, 641–666 (2021).
12. Tedersoo, L., Bahram, M. & Zobel, M. How mycorrhizal associations drive plant population and community biology. *Science* **367**, eaba1223 (2020).
13. Sasaki, T. et al. Role of mycorrhizal associations in tree spatial distribution patterns based on size class in an old-growth forest. *Oecologia* **189**, 971–980 (2019).
14. Delavaux, C. S. et al. Mycorrhizal feedbacks influence global forest structure and diversity. *Commun. Biol.* **6**, 1066 (2023).
15. Stone, L. The feasibility and stability of large complex biological networks: a random matrix approach. *Sci. Rep.* **8**, 8246 (2018).
16. Gibbs, T., Levin, S. A. & Levine, J. M. Coexistence in diverse communities with higher-order interactions. *Proc. Natl Acad. Sci. USA* **119**, e2205063119 (2022).
17. Janzen, D. H. Herbivores and the number of tree species in tropical forests. *Am. Nat.* **104**, 501–528 (1970).
18. Connell, J. H. Diversity in tropical rain forests and coral reefs. *Science* **199**, 1302–1310 (1978).
19. Hubbell, S. P. Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* **203**, 1299–1309 (1979).
20. Condit, R. et al. Spatial patterns in the distribution of tropical tree species. *Science* **288**, 1414–1418 (2000).
21. Morlon, H. et al. A general framework for the distance–decay of similarity in ecological communities. *Ecol. Lett.* **11**, 904–917 (2009).
22. Detto, M., Visser, M. D., Wright, S. J. & Pacala, S. W. Bias in the detection of negative density dependence in plant communities. *Ecol. Lett.* **22**, 1923–1939 (2019).
23. Seidler, T. G. & Plotkin, J. B. Seed dispersal and spatial pattern in tropical trees. *PLoS Biol.* **4**, e344 (2006).
24. McFadden, I. R. et al. Disentangling the functional trait correlates of spatial aggregation in tropical forest trees. *Ecology* **100**, e02591 (2019).
25. Wiegand, T. & Moloney, K. A. *A Handbook of Spatial Point Pattern Analysis in Ecology* (Chapman and Hall/CRC, 2014).
26. 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).
27. Davis, M. A., Curran, C., Tietmeyer, A. & Miller, A. Dynamic tree aggregation patterns in a species-poor temperate woodland disturbed by fire. *J. Veg. Sci.* **16**, 167–174 (2005).
28. Gilbert, S. et al. Beyond the tropics: forest structure in a temperate forest mapped plot. *J. Veg. Sci.* **21**, 388–405 (2010).
29. He, F., Legendre, P. & LaFrankie, J. V. Distribution patterns of tree species in a Malaysian tropical rain forest. *J. Veg. Sci.* **8**, 105–114 (1997).
30. Muller-Landau, H. C. et al. Interspecific variation in primary seed dispersal in a tropical forest. *J. Ecol.* **96**, 653–667 (2008).
31. Wright, J. S. et al. Interspecific associations in seed arrival and seedling recruitment in a neotropical forest. *Ecology* **97**, 2780–2790 (2016).
32. 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).
33. 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).
34. Pontarp, M. et al. The latitudinal diversity gradient: novel understanding through mechanistic eco-evolutionary models. *Trends Ecol. Evol.* **34**, 211–223 (2019).
35. Detto, M. & Muller-Landau, H. C. Rates of formation and dissipation of clumping reveal lagged responses in tropical tree populations. *Ecology* **97**, 1170–118 (2016).
36. 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).
37. Canham, C. D. et al. Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecol. Appl.* **16**, 540–554 (2006).
38. Hülsmann, L. et al. Latitudinal patterns in stabilizing density dependence of forest communities. *Nature* **627**, 564–571 (2024).
39. Lewis, S. L. & Tanner, E. V. J. Effects of above- and belowground competition on growth and survival of rain forest tree seedlings. *Ecology* **81**, 2525–2538 (2000).
40. Bennett, J. A. et al. Plant–soil feedbacks and mycorrhizal type influence temperate forest population dynamics. *Science* **355**, 181–184 (2017).
41. Azaele, S., Pigolotti, S., Banavar, J. R. & Maritan, A. Dynamical evolution of ecosystems. *Nature* **444**, 926–928 (2006).
42. Chesson, P. Scale transition theory: its aims, motivations and predictions. *Ecol. Complex.* **10**, 52–68 (2012).
43. O'Dwyer, J. & Chisholm, R. A mean field model for competition: from neutral ecology to the red queen. *Ecol. Let.* **17**, 961–969 (2014).
44. Murrell, D. When does local spatial structure hinder competitive coexistence and reverse competitive hierarchies? *Ecology* **91**, 1605–1616 (2010).
45. Chesson, P. & Neuhauser, C. Intraspecific aggregation and species coexistence. *Trends Ecol. Evol.* **17**, 210–211 (2002).
46. 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).
47. Broekman, M. J. E. et al. Signs of stabilisation and stable coexistence. *Ecol. Lett.* **22**, 1957–1975 (2019).
48. Beckman, N. G. & Sullivan, L. L. The causes and consequences of seed dispersal. *Annu. Rev. Ecol. Evol. Syst.* **54**, 403–427 (2023).
49. 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).
50. Loreau, M. *From Populations to Ecosystems* (Princeton Univ. Press, 2010).

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025

[1]Department of Ecological Modelling, Helmholtz Centre for Environmental Research–UFZ, Leipzig, Germany. [2]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [3]CAS Key Laboratory of Forest Ecology and Silviculture, Institute of Applied Ecology Chinese Academy of Sciences, http://www.iae.cas.cn. [4]Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA, USA. [5]Conservation Ecology Center, Smithsonian's National Zoo and Conservation Biology Institute, Front Royal, VA, USA. [6]National Biobank of Thailand (NBT), National Science and Technology Development Agency, Klong Luang, Thailand. [7]Administration Bureau of Naban River Watershed National Nature Reserve, https://www.xsbn.gov.cn/nbhbhq. [8]CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, http://www.xtbg.ac.cn. [9]Department of Environmental Technology and Management, Faculty of Environment, Kasetsart University, Bangkok, Thailand. [10]State Key Laboratory of Biocontrol, School of Ecology, Shenzhen Campus of Sun Yat-sen University, https://eco.sysu.edu.cn. [11]Forest Global Earth Observatory (ForestGEO), Smithsonian Tropical Research Institute, Washington, DC, USA. [12]Department of Science and Technology, Uva Wellassa University, Badulla, Sri Lanka. [13]Department of Botany, University of Peradeniya, Kandy, Sri Lanka. [14]School of Ecology and Environment, Northwestern Polytechnical University, https://see.nwpu.edu.cn. [15]Department of Natural and Applied Sciences, University of Wisconsin–Green Bay, Green Bay, WI, USA. [16]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, http://www.whiob.ac.cn. [17]School of Ecology, Northeast Forestry University, https://ecology.nefu.edu.cn. [18]Department of Botany, National Museum of Natural History Smithsonian Institution, Washington, DC, USA. [19]Key Laboratory of National Forestry and Grassland Administration on Plant Conservation and Utilization in Southern China, South China Botanical Garden, Chinese Academy of Sciences, https://www.scbg.ac.cn. [20]Yunnan Academy of Forestry and Grassland, https://www.yafg.ac.cn. [21]Zhejiang Qianjiangyuan Forest Biodiversity National Observation and Research Station, State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, http://www.ibcas.ac.cn. [22]Department of Biology, Washington University in St Louis, St Louis, MO, USA. [23]Harvard Forest, Harvard University, Petersham, MA, USA. [24]Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, https://sees.ecnu.edu.cn. [25]Taiwan Forestry Research Institute, https://www.tfri.gov.tw. [26]Center for Interdisciplinary Research on Ecology and Sustainability, National Dong Hwa University, https://www.ndhu.edu.tw. [27]Institute of Environmental Systems Research, University of Osnabrueck, Osnabrueck, Germany. ✉e-mail: wangxg@iae.ac.cn

Methods

Study areas

Twenty-one large forest dynamic plots of areas between 20 and 50 ha with similar numbers of tropical, subtropical and temperate forests were used in the current study (Extended Data Table 1 and Supplementary Table 1). The forest plots are part of the ForestGEO network[6] and are located in Asia and the Americas ranging in latitudes from 6° 40′ N to 48° 08′ N. Tree species richness among these plots ranges from 36 to 468. All free-standing individuals with diameter at breast height (d.b.h.) ≥ 1 cm were mapped, their size measured and identified. We focused our analysis here on individuals with d.b.h. ≥ 10 cm (resulting in 313,434 individuals) and tree species with more than 50 individuals (resulting in initially 737 species). The 10-cm size threshold excluded most of the saplings and enabled comparisons with previous spatial analyses. Shrub species were excluded. We also excluded 15 species with low aggregation (that is, $k_{ff}^{*} < k_{fm}^{*}$; Box 1), which would lead to negative growth rates at small abundance values: ten of them from BCI, two from MST, two from NBH and one from FS (for definitions of plot acronyms, see Extended Data Table 1). These (generally less abundant) species are probable relicts of an earlier successional episode when they were more abundant[19,51]. We also excluded the two species *Picea mariana* and *Thuja occidentalis* of the Wabikon forest that are restricted to a patch of successional forest that was logged approximately 40 years ago[52].

Most forest plots (18 out of our 21 plots, those with more than 1 census) enabled the estimation of the average mortality risk of individuals with d.b.h. ≥ 10 cm within one census period. We estimated mortality across all species and obtained for each forest plot one average mortality rate for trees with d.b.h. ≥ 10 cm (Extended Data Table 1). We also determined for all species used in our analyses the mycorrhizal association types based on available global datasets[53–55] and website sources (https://www.mycorrhizas.info/). To determine whether a species is mainly dispersed by animals (zoochory), we used the Seed Information Database (https://ser-sid.org/) of the Society for Ecological Restoration and the Royal Botanic Gardens Kew and available literature[56]. Species without descriptions of mycorrhizal associations and dispersal modes were assigned according to their congeneric species. The proportion of focal species with zoochory, with AM association and with both are shown in Extended Data Table 1.

Proxy for pairwise competition strength between species

Some of our analyses required the ratio β_{fi}/β_{ff} (Box 1) that describes the relative competitive effect of individuals of species *i* on an individual of the focal species *f* at the neighbourhood scale[36,57]. In general, it is challenging to derive estimates for the pairwise competition coefficients[37] because this would require unfeasibly large datasets to obtain a sufficient number of neighboured *f*–*j* species pairs for less abundant species. We therefore compared two scenarios. In scenario 1, we assumed that conspecific and heterospecific individuals compete equally, thus $\beta_{fi}/\beta_{ff} = 1$. In scenario 2, we assumed that individuals that are close relatives compete more strongly or share more natural competitors or pathogens than distant relatives[58] (that is, $\beta_{fi}/\beta_{ff} < 1$). As proxy for this effect, we used phylogenetic distances[59], given in millions of years (Myr), as a surrogate for the relative competition strength because they are available for the species in our plots based on molecular data or the Phylomatic informatics tool[60]. To obtain consistent measurements for the ratio β_{fi}/β_{ff} among forest plots, phylogenetic similarities β_{fi}/β_{ff} were scaled between 0 and 1, with conspecifics set to 1 and a similarity of 0 assumed for a phylogenetic distance of 1,200 Myr, which was larger than the maximal observed distance (1,059 Myr). This was necessary to avoid discounting crowding effects from the most distantly related neighbours[58].

For plots without molecular data, we used the V.PhyloMaker2 package (v.0.1.0)[61] to generate a phylogenetic tree for each plot using GBOTB.extented.WP.tre updated from the dated megaphylogeny GBOTB[62] as a backbone. For the other eight plots with molecular data, we followed a previously reported method[63] to build the phylogenetic tree based on DNA barcode data. We then used the cophenetic function in the picante package (v.1.8.2)[64] to calculate phylogenetic distance for each plot. For this, we assumed that functional traits are phylogenetically conserved[36,58,65]. The analyses to generate phylogenetic trees and to calculate phylogenetic distances were performed using R (v.4.3.2)[66].

Crowding indices describing competition of individual trees at the neighbourhood scale

We assumed in our example model that survival of a focal tree *k* is reduced in areas of high local density of conspecifics and heterospecifics (that is, neighbourhood crowding), for example, through competition for space, light or nutrients, or predators or pathogens[17,18,39,67], whereas reproduction is density-independent with per capita rate r_f. However, our approach was also able to deal with alternative assumptions on the processes driven by neighbourhood crowding. For example, analogous models were derived for crowding effects on the reproductive rate and/or the establishment of offspring (Supplementary Text).

We describe the neighbourhood crowding around a given tree *o* of a focal species *f* by commonly used neighbourhood crowding indices[36,37,58,68–70] (Box 1), but used separate indices for conspecific and heterospecific trees. The conspecific crowding index C_{of} of a given individual *o* of a given focal species *f* counts the number n_f of conspecific neighbours *j* that have a distances d_{oj} smaller than a given neighbourhood radius *r*, but weights each neighbour *o* by its inverse distance $1/d_{oj}$, assuming that farther away neighbours compete less (equation (7a)). The heterospecific crowding index H_{of} does the same with all heterospecifics (equation (7b)), and the heterospecific interaction crowding index I_{of} weights heterospecifics additionally by their relative competitive strength β_{fi}/β_{ff} (equation (7c); see above 'Proxy for pairwise competition strength between species'). Thus, we estimated for each individual *o* three crowding indices:

conspecific crowding:

$$C_{of} = \sum_{j=1}^{n_f} \frac{1}{d_{oj}} \tag{7a}$$

heterospecific crowding:

$$H_{of} = \sum_{i \neq f} \sum_{j=1}^{n_i} \frac{1}{d_{oj}} \tag{7b}$$

with niche differences:

$$I_{of} = \sum_{i \neq f} \sum_{j=1}^{n_i} \frac{\beta_{fi}}{\beta_{ff}} \frac{1}{d_{oj}} \tag{7c}$$

where n_i is the number of neighbours of species *i* within distance *r* of the focal individual, d_{oj} is the distance between the focal individual *o* and its *j*th neighbour of species *i*, and β_{fi}/β_{ff} is the competitive effect of one individual of species *i* relative to that of the focal species *f* (refs. 36,37,68–70).

Survival probability of individual trees

To link the survival of an individual *o* to its crowding indices C_{of} and I_{of}, we followed earlier work on individual neighbourhood models[36,37,58,69,70] and assumed that the survival probability s_{of} of a tree *o* of species *f* is given by

$$s_{of} = s_f \exp(-\beta_{ff}(C_{of} + I_{of})), \tag{8}$$

where s_f is a density-independent background survival rate of species *f* and β_{ff} is the neighbourhood-scale conspecific competition coefficients of species *f* (ref. 36). Statistical analyses with neighbourhood crowding

indices have shown that the growth and survival of trees depend on their neighbours mostly within distances r of up to 10 or 15 m (ref. 70). We therefore estimated all measures of spatial neighbourhood patterns with a neighbourhood radius of $r = 15$ m.

Average survival rate of species

We used scale transition theory[42] and spatial point process theory[25] to transfer the individual-based microscale information on the number and distance of conspecific and heterospecific neighbours of focal individuals, which are provided by the ForestGEO census maps, into macroscale models of community dynamics[8]. To this end, we averaged the survival probabilities s_{of} of all individuals o of the focal species f (equation (8)), to obtain the average population-level survival rate $\bar{s}_f$, for which we derived a closed-form expression for gamma-distributed crowding indices[8]:

$$\bar{s}_f = s_f \exp(-\beta_{ff}(\gamma_{fC}\overline{C_f} + \gamma_{fI}\overline{I_f})) \tag{9}$$

where $\overline{C_f}$ and $\overline{I_f}$ are the average crowding indices, and $\gamma_{fC} = \ln(1 + D_{fC}\beta_{ff})/(D_{fC}\beta_{ff})$ and $\gamma_{fI} = \ln(1 + D_{fI}\beta_{ff})/(D_{fI}\beta_{ff})$ arise through the averaging step because of the nonlinearity in equation (8)[8], and are driven by the dispersion (that is, the variance-to-mean ratio) D_{fC} and D_{fI} of the distribution of the crowding indices $\overline{C_f}$ and $\overline{I_f}$, respectively. In our case of high survival, where $D_{fC}\beta_{ff}$ and $D_{fI}\beta_{ff}$ are both small, γ_{fC} and γ_{fI} are near one and can be neglected (that is, $\gamma_{fC} \approx 1$ and $\gamma_{fI} \approx 1$).

Link between average crowding indices and spatial patterns

To incorporate the population-level survival rate (equation (9)) into our population model, we decomposed the average crowding indices into species abundance and measures of spatial patterns (Box 1). In brief, we did this by expressing the crowding indices in terms of the pair correlation function, a basic summary function of spatial statistics[25], and the mean density $\lambda_f = N_f/A$ of the species f across the whole plot of area A (see equations (S1)–(S8) in the Supplementary Text). The resulting measures k_{ff} and k_{fh} of spatial patterns quantify the increase or decrease in average conspecific and heterospecific neighbourhood crowding, respectively, relative to the reference case without spatial patterns. For conspecifics, we expect under a random distribution of the focal species a mean crowding index of $\overline{C_f} = cN_f$ (where $c = 2\pi r/A$ is a scaling factor, with A being the area of the plot and r the radius of the neighbourhood; see equation (S7) in the Supplementary Text), and for heterospecifics, we expect under independent placement (of the focal species with respect to the heterospecifics) a mean crowding index of $\overline{H_f} = c\sum_{i\neq f} N_i$ (ref. 25). We therefore obtained

$$\overline{C_f} = k_{ff}(cN_f) \tag{10a}$$

$$\overline{H_f} = k_{fh}\left(c\sum_{i\neq f} N_i\right) \tag{10b}$$

$$\overline{I_f} = \underbrace{\frac{\overline{I_f}}{\overline{H_f}}}_{B_f}\overline{H_f} = B_f \underbrace{k_{fh}(c\textstyle\sum_{i\neq f} N_i)}_{\overline{H_f}}, \tag{10c}$$

where we define B_f in equation (10c) as $B_f = \overline{I_f}/\overline{H_f}$ to be the average competitive strength of one heterospecific neighbour relative to that of one conspecific. In a subsequent step, we assumed that B_f is approximately constant in time (that is, our mean-field approximation).

The quantity k_{ff} in equation (10a) measures spatial patterns in conspecific crowding of species f ($k_{ff} > 1$ indicates aggregation, and $k_{ff} < 1$ regularity), and the quantity k_{fh} in equation (10b) measures patterns of heterospecific association around the focal species f ($k_{fh} < 1$ indicates segregation, and $k_{fh} > 1$ attraction). Note that our measure of conspecific aggregation, which weights neighbours by distance, is highly correlated to Condit's omega measure of aggregation[20] that counts the number of neighbours without weighting by distance (Extended Data Fig. 1). We also found that the strength of the latitudinal gradient in the exponent of the aggregation–abundance relationship (expressed as the R^2) was for a radius of, for example, $r > 10$ m basically independent of the neighbourhood area over which conspecific aggregation was measured (Extended Data Fig. 2c,f). This was expected because of the distance-weighting (Box 1), whereby distant neighbours contribute little to total neighbourhood crowding.

Mean-field assumption

A crucial insight used in our approach[8] is that crowding competition of individual trees, as described by equation (8), leads to diffuse competition at the population level in species-rich communities. That is, when taking a mean-field approximation[43,71], the species-specific competition strengths of heterospecifics can be replaced in the macroscale model by a temporally constant average heterospecific competition strength B_f (ref. 8) (equation (10c)), which summarizes the emerging effects of the pairwise neighbourhood-scale competition coefficients β_{fi}/β_{ff} at the population level. For species-rich forests at or near a stationary state, B_f is a good approximation of a species-specific constant (see the supplementary text in ref. 8). As we will see, a constant B_f simplifies the matrix of the community-level competition coefficients (equation (11d)) and enables analytical expressions of the invasion condition and the equilibria N_f^* of our multispecies model (equations (11a) and (11b)) for the case that aggregation is independent of abundance.

Zero-sum assumption

Local density dependence on survival as assumed here (equation (8)) controls local tree densities and causes approximate zero-sum dynamics[2], in which the total number J of individuals remains approximately a constant J^*. For example, zero-sum dynamics emerged in our individual-based simulations of the extended multispecies model (blue lines in Extended Data Fig. 5f–j). The number of heterospecifics is therefore given in good approximation by $\sum_{i\neq f} N_i = J^* - N_{f,t}$. Using the zero-sum approximation together with the mean-field approximation (equation (10c)) in equation (11f) decouples the multispecies dynamics and enabled us to investigate the dynamics of individual species in good approximation (equation (13)).

The corrected aggregation–abundance relationship

In Fig. 2 and Extended Data Fig. 3, we fitted a phenomenological power-law for the species of a given forest plot, where the x value was the logarithm of abundance N_f and the y value the corresponding logarithm of k_{ff} (refs. 21,27,28). However, this led to values of k_{ff} close to zero for large abundance values, which would indicate strongly regular patterns[25] not found in the data. Instead, in the extreme case without an aggregation mechanism (that is, random placement of offspring), crowding competition will lead to the repulsion of conspecifics comparable to heterospecific association k_{fh} of heterospecifics. Thus, to avoid a bias, we used the quantity $k_{ff} - k_{fh}$ as the y value in our fit (Extended Data Fig. 9). This required the assumption that $k_{ff} > k_{fh}$, which was satisfied for 98% of our species (722 out of a total of 737 species; see the section 'Study areas'). We leave investigation of the specific cases where $k_{ff} < k_{fh}$ to future studies.

Basic multispecies model M1

Our basic multispecies model (M1) for the per capita growth rate of species f is given by

$$\frac{N_{f,t+1} - N_{f,t}}{\Delta t}\frac{1}{N_{f,t}} = \widetilde{\lambda}_f(N_{f,t}) = (r_f - 1) + s_f \exp(-\beta_{ff} W_f(N_{f,t})), \tag{11a}$$

where $N_{f,t}$ is the abundance of species f at time step t, s_f is a density-independent per capita background survival rate, r_f is the per capita

recruitment rate and β_{ff} is the neighbourhood-scale conspecific competition coefficients of species f. The biological information on neighbourhood crowding competition was incorporated into the fitness factor[42] W_f, which results from combining equations (9) and (10):

$$W_f(N_f) = c\left\{k_{ff}N_f + B_f k_{fh}\sum_{i\neq f} N_i\right\}. \quad (11b)$$

The community-level competition coefficients α_{fi} of this model are therefore given by

$$\alpha_{ff} = c\,\beta_{ff}\,k_{ff} \quad (11c)$$

$$\alpha_{fi} = c\,\beta_{ff}\,(B_f\,k_{fh}) \text{ for } i \neq f. \quad (11d)$$

Thus, even if the conspecific neighbourhood-scale competition coefficients β_{ff} were constant, the corresponding community-level coefficients α_{ff} were not necessarily constant. Instead, they depended on abundance if aggregation k_{ff} depended on abundance[8], as observed in many of our forest dynamics plots (Fig. 2). Not considering the effect that crowding can have on the competition coefficients α_{ff} and α_{fi} is a common (implicit) assumption of models used in coexistence theory[4,15,42,47,49]. Note that equation (11a) together with the fitness factor of equation (11b) enabled us to construct a reference model for the constants k_{ff} and k_{fh}, for which the equilibria N_f^* and the conditions for feasibility and invasiveness can be analytically derived[8,72], and the corresponding non-spatial model (that is, $k_{ff}^* = k_{fh}^* = 1$) without niche differences (that is, $B_f = 1$) led to a per capita population growth rate of zero for all abundances.

Extended multispecies model M2

However, we wanted to extend our model M1 (equations (11a) and (11b)) to include a dependence of aggregation on abundance as new aspect. To introduce this new model, we first defined k_{ff}^* and k_{fh}^* as the observed value of aggregation and heterospecific association, respectively, and k_{ff} as aggregation that depended on abundance. We then assumed that heterospecific association k_{fh}^* is independent of abundance (as suggested by Extended Data Fig. 4) and that the quantity $(k_{ff} - k_{fh})$ follows a power law with respect to abundance (Extended Data Fig. 9 and the section 'The corrected aggregation–abundance relationship'). To formulate our extended multispecies model (M2), we therefore rewrote the fitness factor of equation (11b) by adding and subtracting the term $k_{fh}^* N_{f,t}$:

$$W_f(N_{f,t}) = c\left\{(k_{ff}^* - k_{fh}^*)N_{f,t} + k_{fh}^* N_{f,t} + B_f k_{fh}^*\sum_{i\neq f} N_{i,t}\right\}. \quad (11e)$$

We obtained our extended fitness factor

$$W_f(N_{f,t}) = c\left\{(k_{ff}^* - k_{fh}^*)\left(\frac{N_{f,t}}{N_f^*}\right)^{b_f} N_{f,t} + k_{fh}^* N_{f,t} + B_f k_{fh}^*\sum_{i\neq f} N_{i,t}\right\} \quad (11f)$$

by replacing the quantity $(k_{ff}^* - k_{fh}^*)$ in equation (11e) with a new function $L(N_{f,t})$ depending on abundance:

$$L(N_{f,t}) = (k_{ff}^* - k_{fh}^*)\left(\frac{N_{f,t}}{N_f^*}\right)^{b_f}, \quad (12)$$

where N_f^* is the observed species abundance. Equation (12) simplifies for the observed abundance (that is, $N_{ft} = N_f^*$) to $L(N_f^*) = (k_{ff}^* - k_{fh}^*)$, but otherwise represents the desired power-law with respect to abundance.

Decoupled multispecies models M3

The overall objective of our extended model was to study the effect that a possible dependence of aggregation on abundance (Fig. 2) has on the ability of a newly invading (or an almost extinct) species to increase its abundance. Therefore, we decoupled our extended multispecies model M2 into multiple single-species models (M3) by assuming approximate zero-sum dynamics[2] (see the section 'Zero-sum assumption' above), in which the number of heterospecifics is given by $\sum_{i\neq f} N_i = J^* - N_{f,t}$. We obtained from equation (11f) the fitness function

$$W_f(N_{f,t}) = c\left\{(k_{ff}^* - k_{fh}^*)\left(\frac{N_{f,t}}{N_f^*}\right)^{b_f} N_{f,t} + (1 - B_f)k_{fh}^* N_{f,t} + B_f k_{fh}^* J^*\right\} \quad (13)$$

that results together with equation (11a) in a closed-form expression for the per capita population growth rate $\widetilde{\lambda}_f(N_f)$ of the focal species f at low abundance, which nevertheless includes the key information on crowding competition with heterospecifics through the parameters B_f and k_{fh}^*, and total community size J^*.

Model parameterization

Our extended multispecies model M2 (equations (11a) and (11f)) approximated an underlying individual-based model in the tradition of earlier spatially explicit work[7,35,44,46,73] (Extended Data Fig. 5), but used the empirically observed spatial patterns instead of explicitly modelling their dynamics[8]. In the most general case, the models M2 and M3 have therefore seven parameters per species: three demographic parameters (r_f, s_f and β_{ff}); three parameters that quantify the spatial patterns (that is, k_{ff}^*, k_{fh}^*, B_f); and the exponent b_f of the aggregation–abundance relationship. The parameterization of the models depended on the objective of the model application and the available data. In the example application of our theory, we wanted to feature the effects of spatial patterns (Fig. 2) on coexistence, and we derived a specific parameterization adapted to our data and objective.

We estimated the species-specific measures k_{ff}^* and k_{fh}^* of the observed spatial patterns directly from the ForestGEO plot data based on equations (7) and (10). The exponent b_f of the corrected (power law) aggregation–abundance relationship (equation (12)) was estimated by linear regression, where the x value was $\ln(N_f)$ and the y value the corresponding $\ln(k_{ff} - k_{fh})$ (Extended Data Fig. 9). The parameter b_f captured for a given forest plot the average species response of aggregation to abundance (Extended Data Fig. 9), and we used this value as a parameter for all focal species. This approach is based on a species-for-time substitution, which assumes that the power-law exponent b_f derived for multiple species of one census would be the same as the b_f derived for the same species but at multiple points in time. The results of our individual-based simulations (Extended Data Fig. 5a–f) supported this assumption. Note that effects of habitat association or details of dispersal will influence the values of k_{ff}^* (and k_{fh}^*) and contribute to the observed departures from the power law aggregation–abundance relationship, particularly for tropical forests (compare Extended Data Figs. 3 and 5).

We assumed that all mortality was driven by neighbourhood crowding and therefore set the background survival rate to $s_f = 1$. For estimation of the parameter B_f, we used the matrix of β_{fi}/β_{ff} (see the section 'Proxy for pairwise competition strength between species' above) and derived the crowding indices H_{of} and I_{of} for each individual o (equations (7b) and (7c)). The value of B_f was then given by $B_f = \overline{I_f}/\overline{H_f}$, where $\overline{I_f}$ and $\overline{H_f}$ are the population-level averages of the individual crowding indices H_{of} and I_{of}.

To determine the unknown value of β_{ff}, the neighbourhood-scale conspecific competition coefficients of species f that determines the strength of crowding competition, we assumed that the focal species

is close to equilibrium (that is, $\widetilde{\lambda}_f(N_f^*) = 0$) and obtained by rewriting equation (11a):

$$\beta_{ff} = -\ln((1-r_f)/s_f)/W_f(N_f^*) \tag{14}$$

At equilibrium, we found with equation (11b) that $W_f(N_f^*) = ck_{ff}N_f$ + constant , thus all else equal, β_{ff} is negatively related to the observed abundance: species with lower observed abundance experienced stronger negative impacts on conspecifics than more common species, a pattern frequently observed in plant communities[22,67,73–77]. The effect of departures from the equilibrium assumption on the invasion criterion can be assessed by using equation (S17) in the Supplementary Text.

Inserting equation (14) into equation (11a) led to our final equation for the per capita population growth rate, for which the fitness function is given by equation (13):

$$\frac{N_{f,t+1}-N_{f,t}}{\Delta t}\frac{1}{N_{f,t}} = \widetilde{\lambda}_f(N_{f,t}) = (r_f - 1) + s_f\left(\frac{1-r_f}{s_f}\right)^{\frac{W_f(N_{f,t})}{W_f(N_f^*)}} \tag{15a}$$

For small per capita recruitment rates r_f and large background survival ($s_f = 1$), we obtained

$$\frac{N_{f,t+1}-N_{f,t}}{\Delta t}\frac{1}{N_{f,t}} = \widetilde{\lambda}_f(N_{f,t}) \approx r_f\left(1-\frac{W_f(N_{f,t})}{W_f(N_f^*)}\right), \tag{15b}$$

which suggests the use of the scaled per capita population growth rate $\widetilde{\lambda}_f(N_{f,t})/r_f$ to remove the deterministic effects of the recruitment rate r_f. Note that the effect of individual variation in model parameters on the invasion criterion can be directly investigated by using equations (13) and (15).

Adding a small immigration rate

We reformulated our mathematical model (equation (15a)) in terms of the change in the number of individuals during one time step and added a small constant immigration rate $r_f v_f$ (ref. 41) (r_f is the reproduction rate), thus:

$$\frac{N_{f,t+1}-N_{f,t}}{\Delta t} = \widetilde{\lambda}_f(N_{f,t})N_{f,t} + v_f r_f, \tag{16}$$

which is equivalent to adding the term $v_f r_f/N_{f,t}$ to the per capita population growth rate (equation (15a)). This approach differs from the way immigration is usually modelled in neutral theory[2,26,78]. Note that the term $v_f r_f/N_{f,t}$ decreases rapidly with increasing abundance $N_{f,t}$ and does therefore influence the dynamics only for small abundance values, particularly if the values of v_f is small as assumed here (Extended Data Fig. 7a–d).

The invasion criterion

Stable coexistence requires that the abundance of a newly invading (or an almost extinct) species increases[1,5,8,9] (that is, a rare species advantage), and we wanted to investigate the effect of spatial patterns on the invasion criterion. On the basis of our closed-form expression for the per capita population growth rate $\widetilde{\lambda}_f(N_f)$ (equations (13) and (15a)), invasion analysis reduces to the task of identifying the conditions under which $\widetilde{\lambda}_f(N_s)$ maintains a sufficiently high value for low abundance N_s. We therefore demanded that $\widetilde{\lambda}_f(N_s)$ should be slightly larger than zero ($\widetilde{\lambda}_{\min}$) so that species can escape the effects of demographic stochasticity; that is, our invasion criterion is given by $\widetilde{\lambda}_f(N_s) > \widetilde{\lambda}_{\min}$.

From equation (15a) we found that the invasion condition $\widetilde{\lambda}_f(N_s) > \widetilde{\lambda}_{\min}$ translated into the condition

$$W_f(N_s)/W_f(N_f^*) < 1-\delta, \tag{17a}$$

with $\delta = 1 - \ln((\widetilde{\lambda}_{\min} - r_f + 1)/s_f)/\ln((1-r_s)/s_f)$. For our case of small reproduction rates r_f and high background survival (that is, $s_f = 1$), we obtained $\delta \approx \widetilde{\lambda}_{\min}/r_f$. We therefore investigated the ratio $W_f(N_s)/W_f(N_f^*)$ in more detail, which can be expressed as (equation (S14) in the Supplementary Text):

$$\frac{W_f(N_s)}{W_f(N_f^*)} = \frac{\left(\frac{N_s}{N_f^*}\right)^{b_f+1} + \kappa_f\left(\frac{N_s}{N_f^*}\right) + \rho_f}{1+\kappa_f+\rho_f} \tag{17b}$$

with

$$\kappa_f = (1-B_f)\frac{k_{fh}^*}{k_{ff}^* - k_{fh}^*} \tag{17c}$$

$$\rho_f = B_f\frac{k_{fh}^*}{k_{ff}^* - k_{fh}^*}\frac{J^*}{N_f^*} \tag{17d}$$

The condition in equation (17a) led together with equation (17b) to the invasion criterion

$$1-\frac{W_f(N_s)}{W_f(N_f^*)} = \frac{1-\left(\frac{N_s}{N_f^*}\right)^{b_f+1} + \kappa_f\left(1-\frac{N_s}{N_f^*}\right)}{1+\rho_f+\kappa_f} > \delta. \tag{18}$$

For the case without niche differences (that is, $\kappa_f = 0$), equation (18) suggests that two main mechanisms allow species to increase their rare species advantage: either a smaller negative value of the exponent b_f of the aggregation–abundance relationship or a smaller value of ρ_f (Fig. 4c). We therefore name ρ_f 'risk factor', because larger values of ρ_f lead to smaller values of the per capita growth rate. The maximal value of the risk factor that satisfies the invasion criterion is given by

$$\rho_{f,\max} = \frac{1}{\delta}\left\{(1-\delta) - \left(\frac{N_s}{N_f^*}\right)^{b_f+1} + \kappa_f\left[(1-\delta) - \left(\frac{N_s}{N_f^*}\right)\right]\right\} \tag{19}$$

Thus, for each combination of N_s/N_f^*, b_f, κ_f and δ, we obtained one critical value of $\rho_{f,\max}$ in which the species can invade if ρ_f is smaller than this critical value (Fig. 4c).

It is instructive to investigate the biological mechanism that determines the values of the three factors ρ_f, $(N_s/N_f^*)^{b_f+1}$ and κ_f that determine the invasion criterion (equations (18) and (19)). First, the risk factor ρ_f (equation (17d)) becomes smaller if the observed relative species abundance N_f^*/J^* of the focal species in the community becomes larger (this favours forests at higher latitudes; Extended Data Fig. 8a), if aggregation increases (that is, $k_{fh}^*/(k_{ff}^* - k_{fh}^*)$ decreases; Extended Data Fig. 8d), and if niche differences become larger (that is, B_f decreases). Because of the overwhelming effect of the relative abundance, the risk factor ρ_f decreases strongly with latitude for our dataset (Extended Data Fig. 8c) and for species with lower abundance, aggregation k_{ff}^* would increase and reduce ρ_f. Second, a smaller negative value of the exponent b_f of the aggregation–abundance relationship leads to a smaller value of $(N_s/N_f^*)^{b_f+1}$ (this favours forests at lower latitudes; Extended Data Fig. 8b).

Finally, as expected, it is more likely that the invasion criterion is fulfilled if niche differences become larger (that is, B_f becomes smaller). In this case, ρ_f becomes smaller (equation (17d)) and κ_f becomes larger (equation (17c)), and because $N_s/N_f^* << (1-\delta)$, the inequality of equation (19) is easier to satisfy. The invasion criterion (equation (19)) can be satisfied for niche differences even if $b_f < -1$. Indeed, some species of the CBS plot, which showed an exponent of $b_f = -1.077$, fulfil the

invasion criterion, mostly owing to weak aggregation that led to large values of $k_{fh}/(k_{ff} - k_{fh})$ and larger κ_f (Extended Data Fig. 7e,f). Thus, as expected, the stabilizing mechanism of niche differences ($B_f < 1$) has a key role if spatial patterns alone provide only weak stabilization.

Scenarios investigated

We considered four scenarios to investigate the effects of the spatial mechanism of neighbourhood crowding and immigration on the ability of the 720 study species to increase when having low abundance. Scenario 1 assumed that conspecifics and heterospecifics compete equally (that is, no niche differences; $\beta_{fi} = \beta_{ff}$, $B_f = 1$), and scenario 2 considers niche differences between species approximated by phylogenetic dissimilarity (see the section 'Proxy for pairwise competition strength between species' above). Scenarios 3 and 4 are the same as scenarios 1 and 2, but also assume a small constant immigration with parameter $\upsilon_f = 0.1$. For the mean reproduction rate of $r_f = 0.1$ per time step across all plots, this results in an immigration rate of $r_f \upsilon_f = 0.01$, or 1 immigrant every 100 time steps.

Spatially explicit simulation model

Model description. This model description was adapted from a previous publication[8]. We used the individual-based simulations to verify that the observed patterns (that is, the power law of the abundance–aggregation relationship) can emerge in principle from the minimal mechanisms included in the spatial multispecies model (equations (8) and (11a)).

The individual-based model considered only reproductive (adult) trees, but no size differences. During a given 5-year time step, the model simulated first stochastic recruitment of reproductive trees and placement of recruits, and second, stochastic survival of adults as given by equation (8), depending on their neighbourhood crowding indices (equations (7a)–(7c)) estimated from the community of adult trees. In the next time step, the recruits counted as reproductive adults and were subject to mortality. No immigration from a metacommunity was considered (that is, $\upsilon_f = 0$; equations (1a)). To avoid edge effects, torus geometry was assumed.

Each individual produced r_f recruits on average, and their locations were determined by a type of Thomas process[25] to obtain a clustered distribution of recruits. In our model, the spatial position of the recruits was determined using two independent mechanisms. First, a proportion $1 - p_d$ of recruits was placed stochastically around randomly selected conspecific adults (parents) by using a two-dimensional kernel function (here a Gaussian with variance σ^2). This is the most common way in most spatially explicit models to generate species clustering. Technically, we first selected for each of these recruits randomly one parent among the conspecific adults and then determined the position of the recruit by sampling from the kernel. Second, the remaining proportion p_d of recruits was distributed in the same way, but around randomly placed cluster centres that are located independently of conspecific adults.

Parameterization of the simulation model. The simulation model used here is described in detail in a previous study[8]. However, we used here distance-weighting for the estimation of the crowding indices. Thus, in the source code (the supplementary information in ref. 8) we used DistanceWeighting = 1 instead of DistanceWeighting = 0.

The simulations of the individual-based forest model were conducted in 5-year time steps over 25,000 years (equivalent to 5,000 census periods) in an area of $A = 200$ ha, and comprised approximately 86,000 trees with initially 80 species. The model parameters were the same for all species, and all species followed exactly the same model rules. We selected $\beta_{fi} = \beta_{ff}$ to obtain no differences in conspecific and heterospecific interactions and $s_f = 1$ (no background mortality), a standard deviation of $\sigma = 10$ m of the kernel function, and we adjusted the parameters $\beta_{ff} = 0.02$ and $r_f = 0.1$ to produce tree densities (430 per ha) and an overall 5-year mortality rate (10%) similar to that of trees with d.b.h. ≥ 10 cm of the BCI plot. The radius of the neighbourhood used to estimate the crowding indices was $r = 20$ m, and the number of randomly assigned cluster centres was 16.

Reporting summary

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

Data availability

The minimum datasets that enable the repetition, interpretation, verification and extension of the research is provided in Supplementary Table 1. This table provides for all species the indices of spatial patterns and all other quantities used in the model analysis and to generate Table 1, Figs. 2–4 and Extended Data Figs. 2–4 and 6–9. Additional raw data for Fig. 3 are shown in Extended Data Table 1. The output data of the individual-based forest simulation model used to generate Extended Data Fig. 5 is provided in Supplementary Table 2. To estimate the spatial pattern indices, we used the raw census data of the ForestGEO network. The raw census data of several sites are publicly available (https://forestgeo.si.edu/explore-data). For the other sites, they are available upon reasonable request and with permission of the principal investigators of the corresponding ForestGEO sites (https://forestgeo.si.edu/sites-all). Mycorrhizal association types were based on available global datasets[53–55] and website sources (https://www.mycorrhizas.info). To determine whether a species is mainly dispersed by animals (zoochory), we used the Seed Information Database (https://ser-sid.org) of the Society for Ecological Restoration and the Royal Botanic Gardens Kew. Source data are provided with this paper.

Code availability

Source code for the simulation model written in Delphi (Pascal), which contains the procedures to repeat the results shown in Extended Data Fig. 5 and to estimate the spatial pattern indices, can be found in the supplementary information of a previous study[8].

51. Flügge, A. J., Olhede, S. C. & Murrell, D. J. The memory of spatial patterns: changes in local abundance and aggregation in a tropical forest. *Ecology* **93**, 1540–1549 (2012).
52. Wang, X. et al. Phylogenetic and functional area relationships in two temperate forests. *Ecography* **36**, 883–893 (2013).
53. Soudzilovskaia, N. A. et al. FungalRoot v.2.0—an empirical database of plant mycorrhizal traits. *N. Phytol.* **235**, 1689–1691 (2022).
54. Akhmetzhanova, A. A. et al. A rediscovered treasure: mycorrhizal intensity database for 3000 vascular plant species across the former Soviet Union. *Ecology* **93**, 689–690 (2012).
55. Wang, B. & Qiu, Y.-L. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**, 299–363 (2006).
56. Jin, J. et al. Stronger latitudinal phylogenetic patterns in woody angiosperm assemblages with higher dispersal abilities in China. *J. Biogeogr.* **51**, 269–279 (2024).
57. Fortunel, C. et al. Topography and neighborhood crowding can interact to shape species growth and distribution in a diverse Amazonian forest. *Ecology* **99**, 2272–2283 (2018).
58. 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. Letters* **19**, 1062–1070 (2016).
59. 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).
60. Webb, C. & Donoghue, M. Phylomatic: tree assembly for applied phylogenetics. *Mol. Ecol. Notes* **5**, 181–183 (2005).
61. Jin, Y. & Qian, H. V. PhyloMaker2: an updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Divers.* **44**, 335–339 (2022).
62. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).
63. Kress, W. J. et al. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc. Natl Acad. Sci. USA* **106**, 18621–18626 (2009).
64. Kembel, S. W. et al. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).
65. Cadotte, M. W., Davies, T. J. & Peres-Neto, P. R. Why phylogenies do not always predict ecological differences. *Ecol. Monogr.* **87**, 535–551 (2017).
66. The R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022).
67. Mangan, S. A. et al. Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**, 752–755 (2010).

68. 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).
69. Lasky, J. R., Uriarte, M., Boukili, V. K. & Chazdon, R. L. Trait-mediated assembly processes predict successional changes in community diversity of tropical forests. *Proc. Natl Acad. Sci. USA* **111**, 5616–5621 (2014).
70. Uriarte, M., Condit, R., Canham, C. D. & Hubbell, S. P. A spatially explicit model of sapling growth in a tropical forest: does the identity of neighbours matter? *J. Ecol.* **92**, 348–360 (2004).
71. Fung, T., O'Dwyer, J. & Chisholm, R. Effects of temporal environmental stochasticity on species richness: a mechanistic unification spanning weak to strong temporal correlations. *Oikos* https://doi.org/10.1111/oik.08667 (2022).
72. Shigesada, N., Kawasaki, K. & Teramoto, E. The effects of interference competition on stability, structure and invasion of a multi-species system. *J. Math. Biol.* **21**, 97–113 (1984).
73. Bolker, B., Pacala, S. W. & Neuhauser, C. Spatial dynamics in model plant communities: what do we really know? *Am. Nat.* **162**, 135–148 (2003).
74. Comita, L., 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).
75. Chisholm, R. A. & Muller-Landau, H. C. A theoretical model linking interspecific variation in density dependence to species abundances. *Theor. Ecol.* **4**, 241–253 (2011).
76. Yenni, G., Adler, P. B. & Ernest, S. K. M. Strong self-limitation promotes the persistence of rare species. *Ecology* **93**, 456–461 (2012).
77. Mack, K. M. L. & Bever, J. D. Coexistence and relative abundance in plant communities are determined by feedbacks when the scale of feedback and dispersal is local. *J. Ecol.* **102**, 1195–1201 (2014).
78. Chisholm, R. A. & Lichstein, J. W. Linking dispersal, immigration and scale in the neutral theory of biodiversity. *Ecol. Lett.* **12**, 1385–1393 (2009).

Acknowledgements We thank numerous local field and laboratory staff, technicians, interns, volunteers and researchers for their contributions to the collection and management of the forest inventory data used in the analyses; and D. Alonso, J. Capitan, members of the Kraft laboratory and R. Condit for feedback on earlier drafts. A detailed list of funding sources, acknowledgements and references for each forest site is available in the Supplementary Notes. T.W. and A.H. were supported by an ERC Advanced Grant (SpatioCoexistence, project 101141989). Xugao Wang was supported by the National Key Research and Development Programme of China (2022YFF1300501), the CAS Project for Young Scientists in Basic Research (YSBR-108) and the Key Research Programme of Frontier Sciences, Chinese Academy of Sciences (ZDBS-LY-DQC019). S.M.F. and A.H. received funding from the Deutsche Forschungsgemeinschaft (DFG; grant number 431504473).

Author contributions T.W., Xugao Wang, S.M.F. and A.H. conceived and designed the project. T.W. implemented the models, conducted the simulations, analysed the results and prepared figures and tables. T.W., A.H., S.M.F., Xugao Wang and N.J.B.K. led the writing of the manuscript. Xugao Wang assembled and analysed the plot data and conducted the spatial analyses. The other authors contributed forest census data. All authors reviewed, approved and had the opportunity to comment on the manuscript.

Competing interests The authors declare no competing interests.

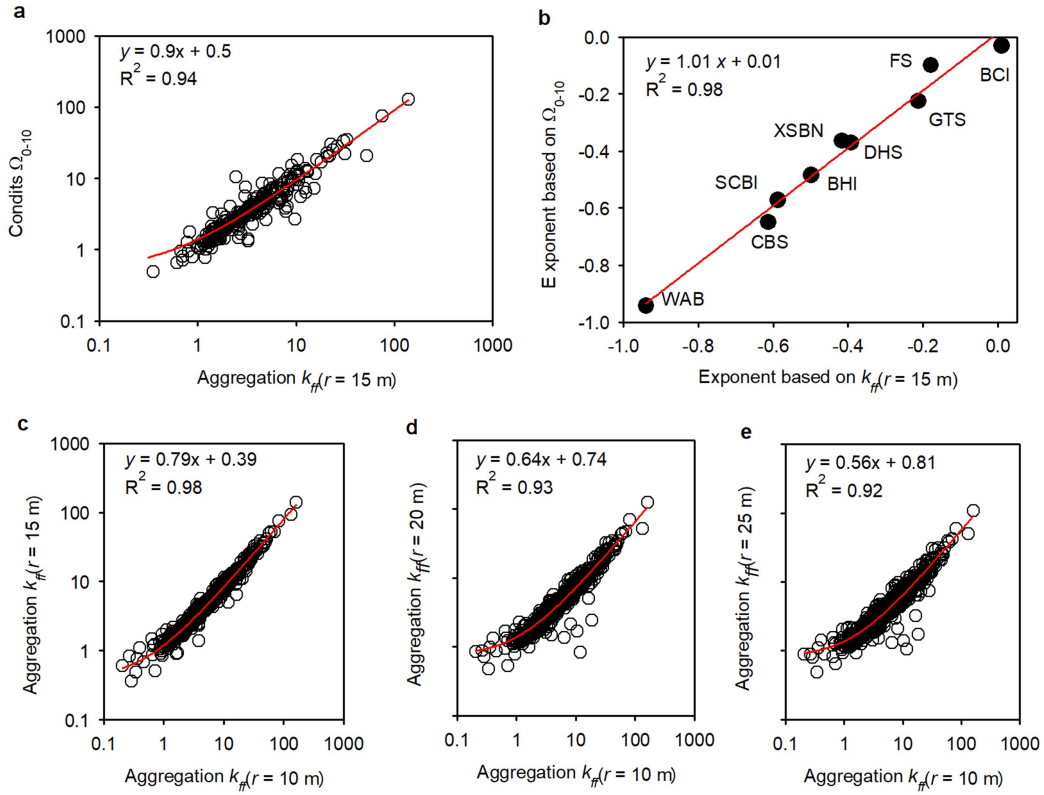
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08604-z.

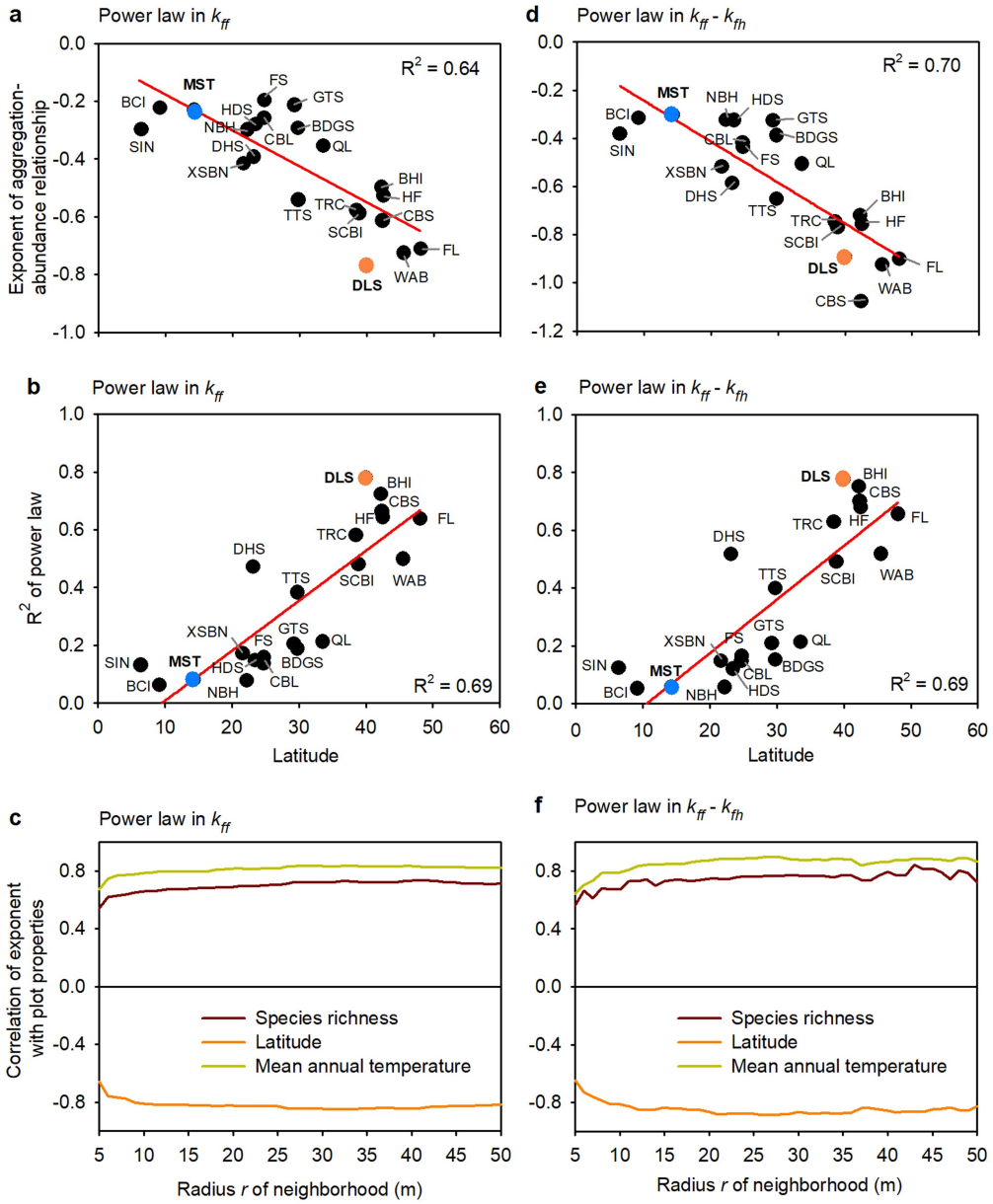
Correspondence and requests for materials should be addressed to Xugao Wang.

Peer review information *Nature* thanks Ryan Chrisholm, Fangliang He and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

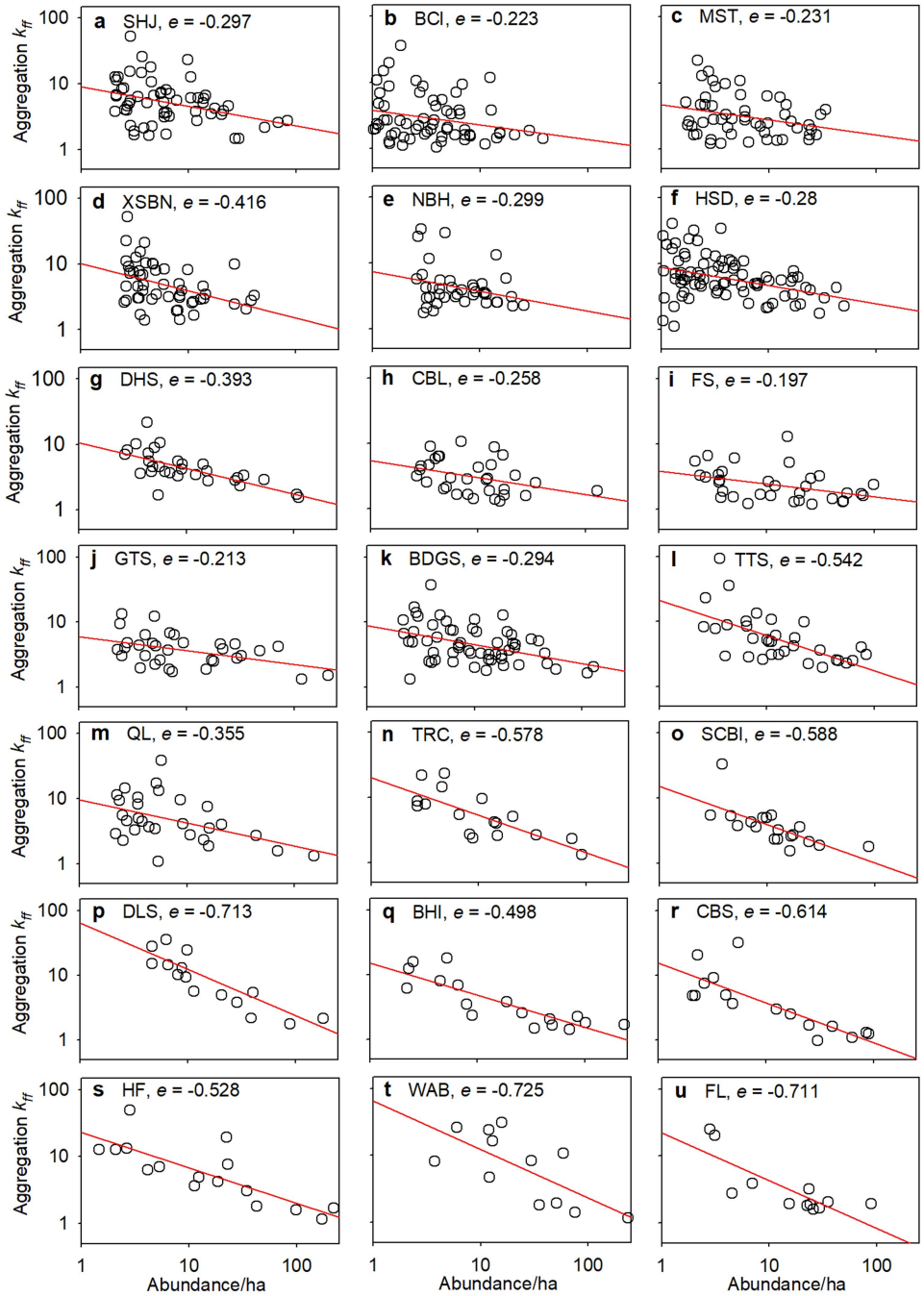


Extended Data Fig. 1 | Sensitivity of the aggregation-abundance relationships to the measure of aggregation. a, Condits's $\Omega_{1\text{-}10}$ aggregation measure for the species analysed in Wiegand et al. (ref. 8) plotted over the corresponding values of the distance-weighted aggregation measure $k_{ff}(r = 15\,\text{m})$ used in the present study. **b**, exponents of the aggregation-abundance power-law derived with Condits's $\Omega_{1\text{-}10}$ plotted over the corresponding values of the distance-weighted aggregation measure $k_{ff}(r = 15\,\text{m})$. **c**, comparison of the distance-weighted aggregation measure $k_{ff}(r)$ for the neighborhood radii 10 m vs. 15 m. **d**, same as c), but for the neighborhood radii 10 m vs. 20 m. **e**, same as c), but for the neighborhood radii 10 m vs. 25 m. To outline the overall tendency in the data we fitted linear regressions to the data. Note the double-logarithmic scale of panels a), c), d), and e).



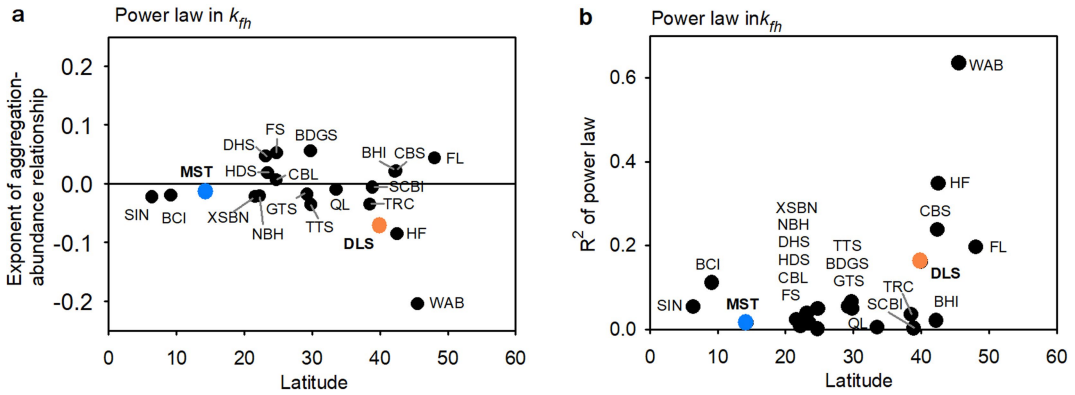
Extended Data Fig. 2 | Latitudinal gradients in the exponent of the abundance-aggregation relationship. a, the exponent e of the power law $k_{ff}(N_f) = a N_f^e$ of aggregation k_{ff} with abundance N_f over latitude. **b**, the R^2 of the linear regression between $\ln(k_{fh})$ and $\ln(N_f)$ for each species f. **c**, the Pearson correlation coefficient between the power law exponent e and the plot properties species richness, latitude and mean annual temperature for the scaling of k_{ff} with N_f. **d**, same as a) but for the exponent of the abundance scaling of the difference between aggregation k_{ff} and heterospecific association k_{fh}, **e**, same as b) but for the linear regression between $\ln(k_{ff} - k_{fh})$ and $\ln(N_f)$. **f**, same as c, but for the scaling of $k_{ff} - k_{fh}$ with N_f. Our measures of aggregation k_{ff} and heterospecific association k_{fh} are based on neighbourhood crowding indices that count the number of neighbours within distance r of the focal individual, but each neighbour is weighted by the inverse of its distance to the focal individual (Box 1). The neighbourhood distance was in all panels r = 15 m.

For details on the forest plots see Extended Data Table 1 and for data Supplementary Data Table 1. To show the overall tendency in the data we fitted in a), b), d) and e) linear regressions. The plot acronyms: BCI: Barro Colorado Island (tropical forest); BDGS: Badagongshan (subtropical forest); BHI: Baihua (temperate forest); CBL: Chebaling (subtropical forest); CBS: Changbaishan (temperate forest); DHS: Dinghushan (subtropical forest); DLS: Donglingshan (temperate forest); FL: Fenglin (temperate forest); FS: Fushan (subtropical forest); GTS: Gutianshan (subtropical forest); HF: Harvard Forest (temperate forest); HSD: Heishiding (tropical–subtropical forest); MST: Mo Singto (tropical forest); NBH: Nabanhe (tropical forest); QL: Qinling (subtropical–temperate forest); SCBI: Smithsonian Conservation Biology Institute (temperate forest); SHJ or SIN: Sinharaja (tropical forest); TRC: Tyson Research Center (temperate forest); TTS: Tiantong (subtropical forest); WAB: Wabikon (temperate forest); XSBN: Xishuangbanna (tropical forest).

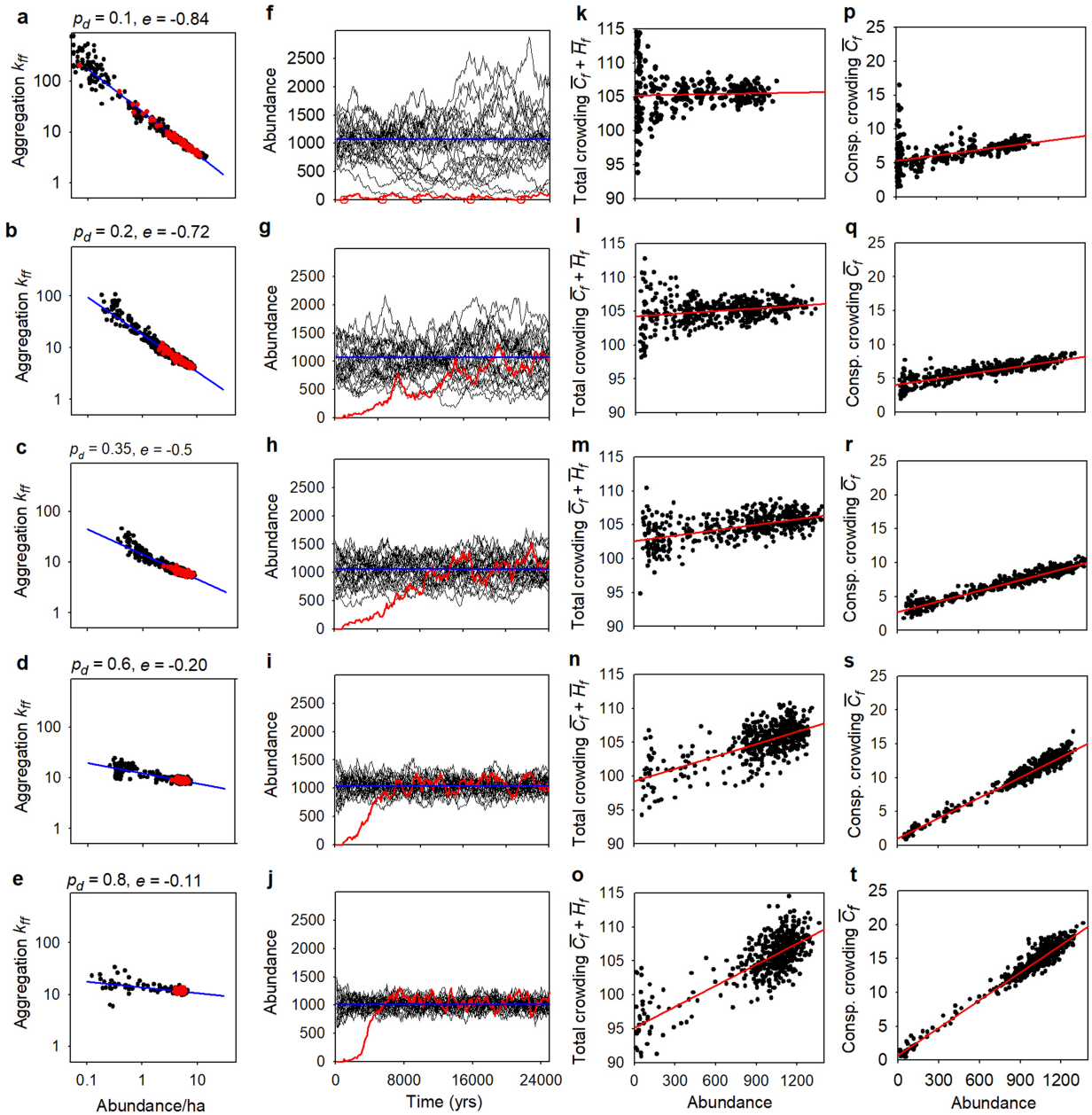


Acronym	Plot name
SHJ:	Sinharaja
BCI:	Barro Colorado Island
MST:	MoSingto
XSBN:	Xishuangbanna
NBH:	Nabanhe
HSD:	Heishiding
DHS:	Dinghushan
CBL:	Chebaling
FS:	Fushan
GTS:	Gutianshan
BDGS:	Badagongshan
TTS:	Tiantong
QL:	Qinling
TRC:	Tyson Research Center
SCBI:	Smithsonian Conservation Biology Institute
DLS:	Donglingshan
BHI:	Baihua
CBS:	Changbaishan
HF:	Harvard Forest
WAB:	Wabikon
FL:	Fenglin

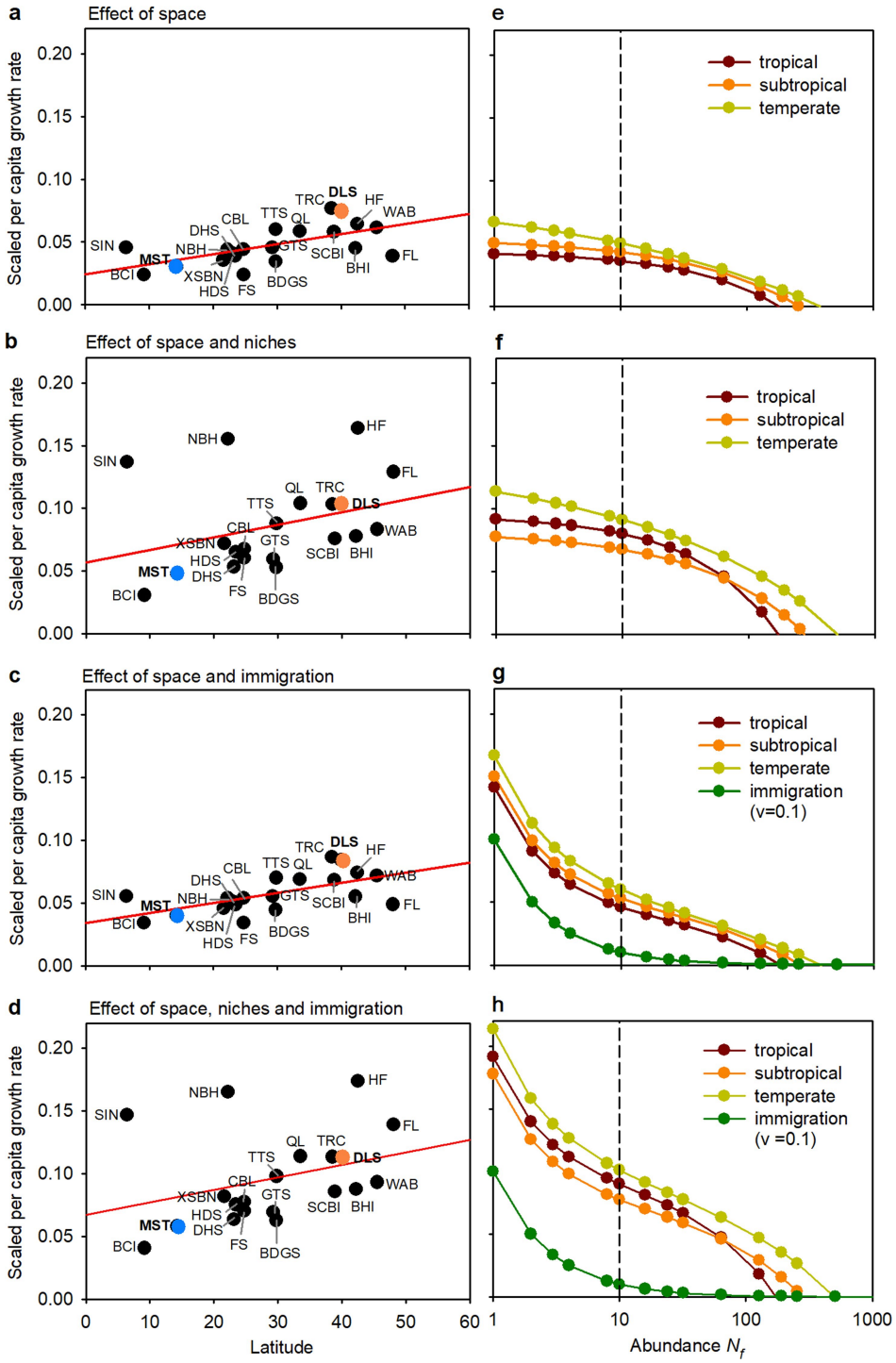
Extended Data Fig. 3 | Aggregation – abundance relationships for the 21 ForestGEO plots. The panels (a – u) show the relationship between aggregation k_{ff} (y-axis) and abundance per ha n_f (x-axis) for each plot. The value of e is the exponent of the power law $k_{ff}(n_f) = a\, n_f^e$ (red lines; see also Fig. 2c), estimated by linear regression analysis of $\log(k_{ff})$ over $\log(n_f)$. The neighbourhood distance was in all panels $r = 15$ m. We analysed focal species f with more than 50 large individuals (i.e., dbh ≥ 10 cm). (a – f): tropical, (f – m): subtropical, and (m – u) temperate forests.



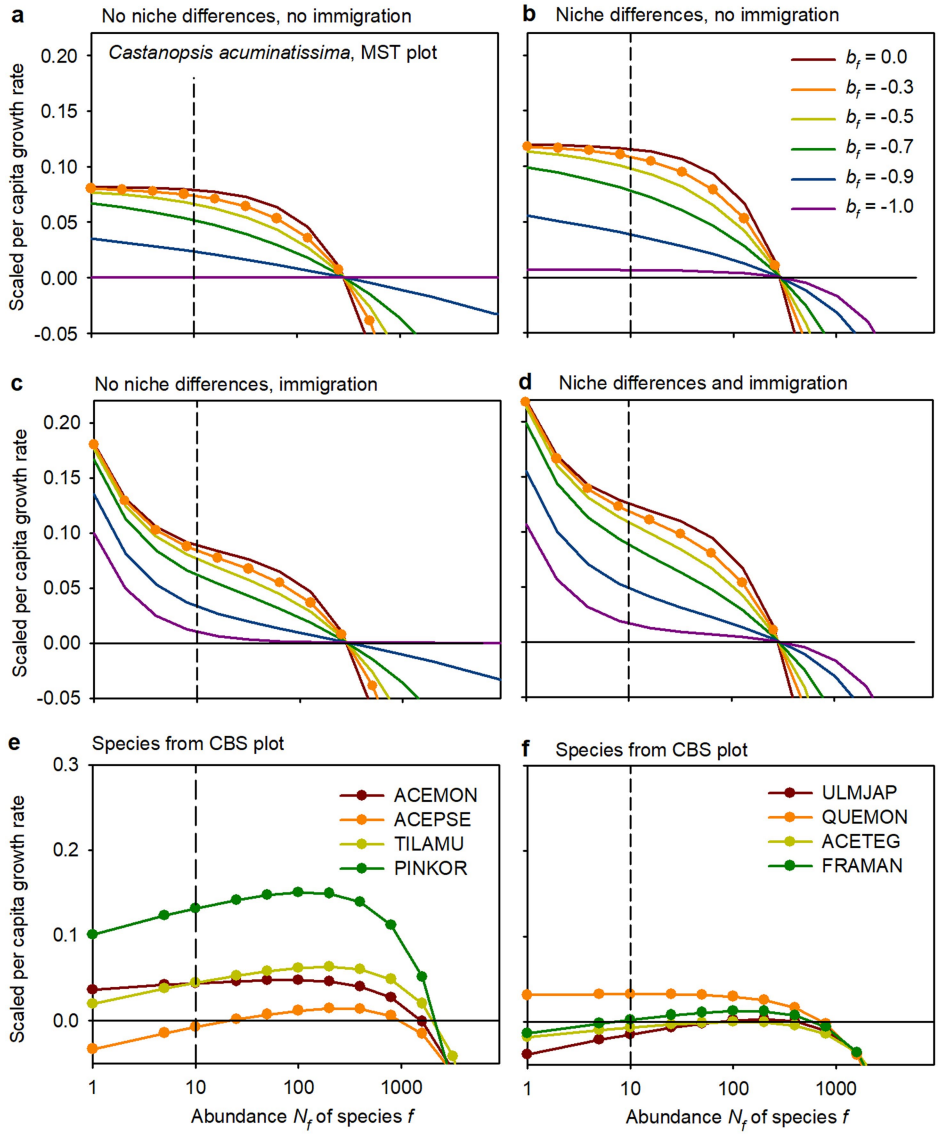
Extended Data Fig. 4 | Abundance-dependency of the association k_{fh} to heterospecifics. a, exponent of the power law of k_{fh} with respect to species abundance N_f for the 21 the ForestGEO data sets. **b**, the R^2 of the linear regression between $\ln(k_{fh})$ and $\ln(N_f)$. The neighbourhood distance was in all panels $r = 15$ m. For plot acronyms see Extended Data Fig. 2.



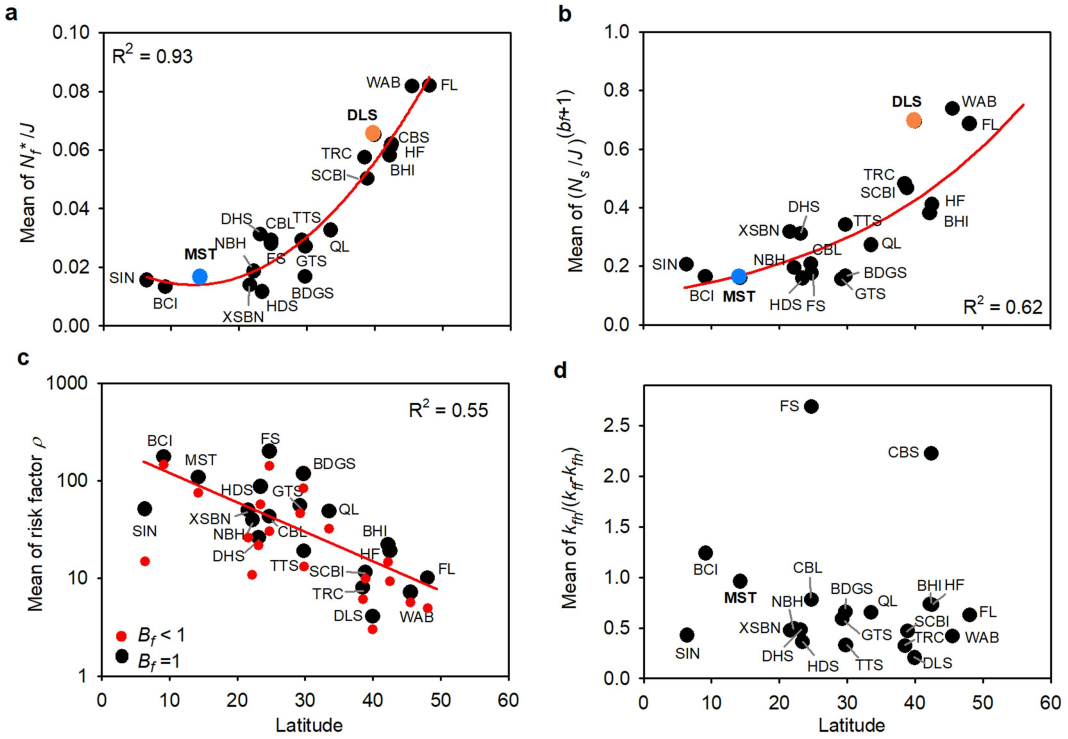
Extended Data Fig. 5 | The aggregation-abundance relationship and the rare species advantage in individual-based simulations. We used an individual-based implementation of the community model of eqs. 1 and 4 with initially 80 species on an area of 200 ha without immigration, simulated for 5000 time steps (25,000 years, $\Delta t = 5$ years). A proportion p_d of recruits was placed with a Gaussian kernel around randomly distributed cluster centres, and the rest with the same kernel around their parents (see methods). The different simulations differ only in the parameter $p_d = 0.1, 0.2, 0.35, 0.6, 0.8$ (from top to bottom). **a**–**e**: temporal aggregation-abundance relationship for 2 or 3 selected species, taken every 50 years (black disks), and for all 80 species at year 25,000 (red discs). Fit of the temporal relationship with a power law with exponent e (blue line), **f**–**j**: time series showing the abundances of the first 25 (out of 80) species, where one species (red) invades at year 1000 (and 2 timesteps after extinction) starting with 50 individuals. The average abundance is shown as blue line. **k**–**o**: Total crowding index $\overline{C_f} + \overline{H_f}$ (i.e., distance-weighted number of neighbours) of the invading species in dependence on abundance [in k) we used a species that went extinct]. The red line shows the fit with a linear regression. **p**–**t**: same as k)-o), but only for conspecific crowding $\overline{C_f}$. The model parameters were the same for all species: $r_f = 0.1/\Delta t$, $s_f = 0$, $B_f = 1$, $c = 0.000063$, $r = 20$ m, $\beta_{ff} = 0.02$, $\sigma = 10$ m (see methods), and $N_f^* = 1070$ and $J^* = 86{,}000$ emerged. For comparison with Table 1, the normalized per capita population growth rates $\tilde{\lambda}_f(N_s = 10)/r_f$ at the last timestep were 0.019, 0.012, 0.022, 0.031, and 0.038 for parameter values of $p_d = 0.1, 0.2, 0.35, 0.6$, and 0.8, respectively. The raw data with the output of the individual-based model can be found in the Supplementary Information as Supplementary Data Table 2.



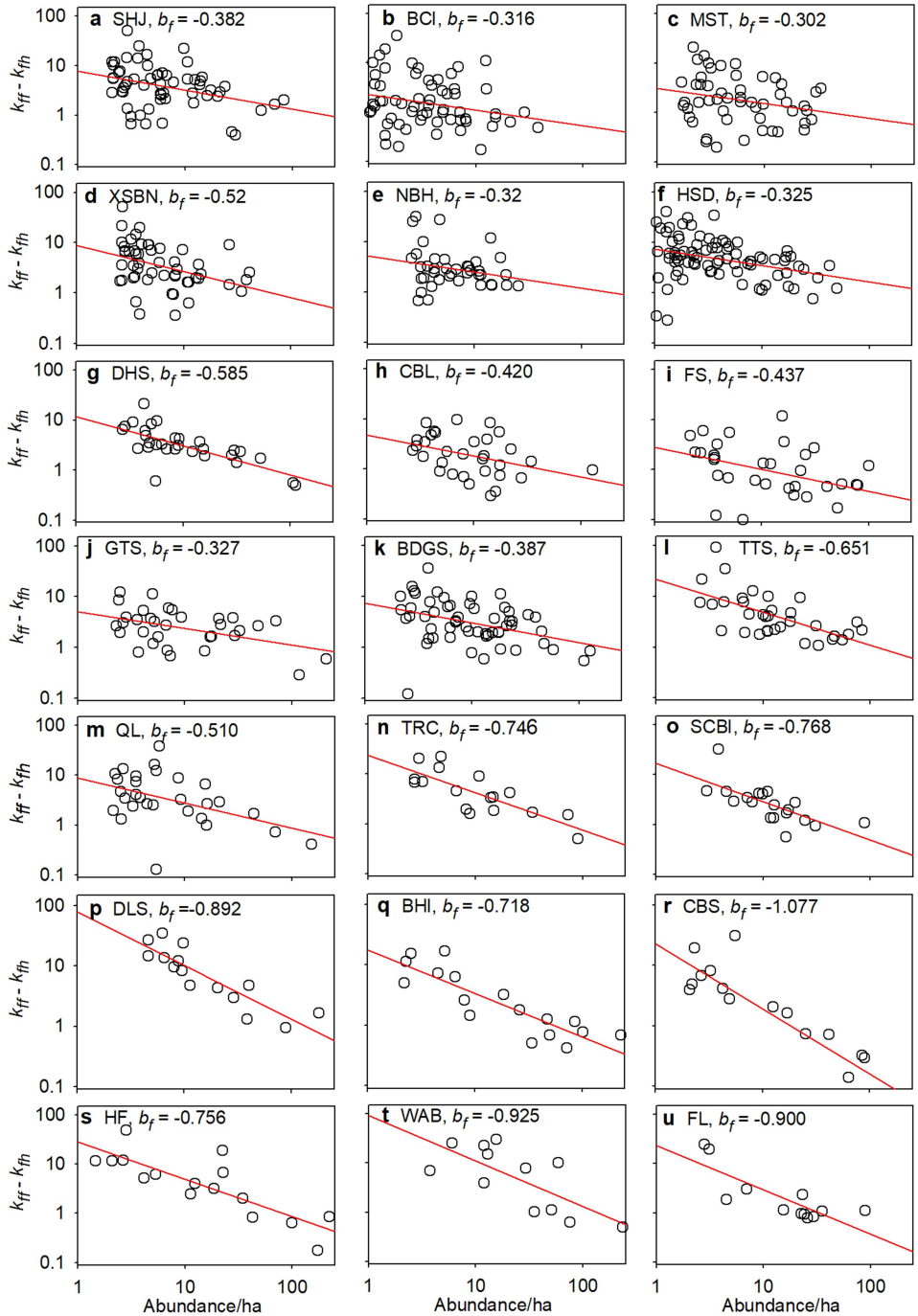
Extended Data Fig. 6 | The scaled per-capita population growth rate $\widetilde{\lambda}_f(N_s)/r_f$, as influenced by spatial structure, niche differences and immigration for 21 ForestGEO plots. a, the scaled per capita population growth rate $\widetilde{\lambda}_f(N_s)/r_f$ for a small abundance of $N_s = 10$ individuals, per forest plot for scenario 1 (no immigration: $\upsilon_f = 0$, no niche differences: $\beta_{fi} = \beta_{ff}$) that represents the "pure" effect of spatial structure. **b**, same as a), but for scenario 2 that adds niche differences ($\upsilon_f = 0, \beta_{fi} < \beta_{ff}$). **c**, same as a), but for scenario 3 adds immigration to scenario 1 ($\upsilon_f = 0.1, \beta_{fi} = \beta_{ff}$). **d**, scenario 4 adds niche differences and immigration ($\upsilon_f = 0.1, \beta_{fi} < \beta_{ff}$). **e**: $\widetilde{\lambda}_f(N_s)/r_f$ for scenario 1 in dependence on abundance N_f, averaged separately over all species of tropical, subtropical and temperate forests. The dashed vertical line indicates the small abundance of $N_s = 10$. **f**, same as e), but for scenario 2. **g**, same as e), but for scenario 3, and **h**, same as e), but for scenario 4. In scenarios 2 and 4 we assumed that more closely related species compete more strongly. The models were parameterized for 720 species of the 21 ForestGEO plots. We excluded the temperate forest at CBS with exponent $b_f < -1$. To outline the overall tendency in the data we fitted in panels a) to d) a linear regression to the data.



Extended Data Fig. 7 | Dependency of the scaled per capita population growth rate on the exponent of the aggregation-abundance relationship. **a**, Examples for the influence of the exponent b_f on the scaled per capita population growth rate of the species *Castanopsis acuminatissima* from the Mo Singto (MST) plot in Thailand with parameters $b_f = -0.3$, $N_f^* = 291$, $J = 15{,}665$, $r_f = 0.136/\Delta t$, $k_{ff} = 6.27$, $k_{fh} = 1.02$, $B_f = 1$ and $c = 0.000314$. The dashed vertical line indicates the small abundance of $N_s = 10$. **b**, same as a), but for niche differences ($B_f = 0.692$). **c**, and **d**, same as a) and b), but with immigration parameter $v_f = 0.1$. **e**, and **f**, examples for the scaled per capita population growth rate of species at the CBS plot, which showed a power law exponent of $b_f = -1.077$, for scenario 2 with niche differences (i.e., $B_f < 1$). We assumed that individuals compete at the individual scale more strongly if they are phylogenetically more similarity. The species acronyms: ACEMON (*Acer mono*), ACEPSE (*A. pseudosieboldianum*), ACETEG (*A. tegmentosum*), FRAMAN (*Fraxinus mandshurica*), PINKOR (*Pinus koraiensis*), QUEMON (*Quercus mongolica*), TILAMU (*Tilia amurensis*), ULMJAP (*Ulmus japonica*).



Extended Data Fig. 8 | Properties of species leading to a high per capita population growth rate at low abundances. The panels a) – d) show different quantities that determine the per capita population growth rate $\tilde{\lambda}_f(N_f)$ (eq. 11a, 13). **a**, the mean relative abundance N_f^*/J^*, which drives the risk factor ρ_f, averaged over all species of a given plot in dependence on latitude, **b**, same as a), but for the quantity $(N_s/N_f^*)^{b_f+1}$ as it appears in eq. 2. **c**, same as a), but for the mean of the risk factor ρ_f for niche differences (red) and no niche differences (black), **d**, same as a), but for the mean of the quantity $k_{fh}/(k_{ff}-k_{fh})$. To outline the overall tendency in the data we fitted in panels a) a polynomial regression of order 2, and in panels b) and c) an exponential regression. For plot acronyms see Extended Data Fig. 2.



Acronym	Plot name
SHJ:	Sinharaja
BCI:	Barro Colorado Island
MST:	MoSingto
XSBN:	Xishuangbanna
NBH:	Nabanhe
HSD:	Heishiding
DHS:	Dinghushan
CBL:	Chebaling
FS:	Fushan
GTS:	Gutianshan
BDGS:	Badagongshan
TTS:	Tiantong
QL:	Qinling
TRC:	Tyson Research Center
SCBI:	Smithsonian Conservation Biology Institute
DLS:	Donglingshan
BHI:	Baihua
CBS:	Changbaishan
HF:	Harvard Forest
WAB:	Wabikon
FL:	Fenglin

Extended Data Fig. 9 | Aggregation – abundance relationships for the 21 ForestGEO plots. The panels (a – u) show the relationships between corrected aggregation $L(N_f) = (k_{ff} - k_{fh})$ (y-axis) and abundance N_f (x-axis) (eq. 8). The value of b_f is the slope of the power law $L(N_f) = a_f N_f^{bf}$ (red lines), estimated by linear regression analysis of $\log(k_{ff} - k_{fh})$ over $\log(N_f)$. We used for the analysis focal species f with more than 50 individuals. (a – f): tropical, (f – m): subtropical, and (m – u) temperate forests.

Extended Data Table 1 | Characteristics of the selected forest plots

Plot	Acronym	Latitude	Elevation	No. species	Trees ≥10cm dbh	Focal species‡	Area (ha)	Mortality rate§	Exponent b_r‖	Prop. zoo¶	Prop. AM#	Prop. AM and zoo¶ #
a Tropical												
Sinharaja	SHJ	6.40	500	239	17070	59	25	0.087	-0.38	0.75	0.79	0.71
Barro Colorado Island	BCI	9.15	140	304	20647	63 (-10)	50	0.113	-0.32	0.89	1.00	0.91
Mo Singto	MST	14.26	770	262	15665	52 (-2)	30	0.136	-0.30	0.88	0.91	0.80
Xishuangbanna*	XSBN	21.61	789	468	12263	52	20	0.14	-0.52	0.81	0.85	0.67
Nabanhe	NBH	22.23	932	290	8512	40 (-2)	20	-	-0.32	0.78	0.86	0.67
Heishiding†	HSD	23.45	567	245	30992	82	50	0.1	-0.33	0.84	0.81	0.68
b Subtropical												
Dinghushan*	DHS	23.16	350	210	11906	30	20	0.18	-0.59	0.87	0.80	0.77
Chebaling	CBL	24.72	495	228	10159	32	20	-	-0.42	0.78	0.84	0.65
Fushan*	FS	24.76	667	110	19261	33 (-1)	25	0.071	-0.44	0.85	0.81	0.68
Gutianshan*	GTS	29.25	581	159	18215	33	24	0.09	-0.33	0.73	0.68	0.48
Badagongshan	BDGS	29.77	1406	232	25121	57	25	0.06	-0.39	0.67	0.77	0.54
Tiantong	TTS	29.81	454	154	14967	34	20	0.08	-0.65	0.68	0.79	0.55
Qinling*†	QL	33.54	1431	121	11698	29	25	-	-0.51	0.45	0.61	0.29
c Temperate												
Tyson Research Center	TRC	38.52	203	46	6520	17	20	0.085	-0.75	0.82	0.47	0.29
Smithsonian Conservation Biology Institute	SCBI	38.89	330	65	8165	19	25	0.08	-0.77	0.68	0.42	0.17
Donglingshan	DLS	39.96	1400	51	9558	15	20	0.04	-0.89	0.20	0.60	0.13
Baihua*	BHI	42.22	793	63	17197	17	24	0.07	-0.72	0.24	0.53	0.12
Changbaishan*	CBS	42.38	801	52	9995	16	25	0.05	-1.08	0.25	0.69	0.13
Harvard Forest	HF	42.54	354	55	23901	16	35	0.14	-0.76	0.50	0.25	0.13
Wabikon*	WAB	45.55	501	36	14021	12 (-2)	25.2	0.04	-0.93	0.00	0.42	0.00
Fenglin	FL	48.08	447	46	8601	12	30	0.11	-0.90	0.08	0.50	0.00

* plots with bar code phylogeny, genus level phylogenies were used for all other plots; † Heishiding is classified as tropical-subtropical forest, Qinling is classified as subtropical-temperate forest; ‡ The numbers give the number of focal species with at least 50 individuals with dbh ≥ 10 cm, we excluded species with $k_{ff} < k_{fh}$ (the numbers in parenthesis) and two species at the Wabikon forest located in a patch of successional forest that was logged approximately 40 yr ago; § The average mortality rate for trees with dbh ≥10 cm in the plot for a $\Delta t = 5$ year period. Assuming approximate equilibrium, we used the values of the mortality rate to parameterize the unknown per capita recruitment rates r_i. We used a recruitment rate of $0.1/\Delta t$ for the 3 plots without mortality data. The mortality rate of the TRC plot was estimated from a 13.4 ha section of the 20 ha plot; || The exponent b_i of the aggregation-abundance relationship based on corrected aggregation $k_{ff} - k_{fh}$; ¶ Proportion of species showing mostly animal seed dispersal; # Proportion of species with arbuscular mycorrhizal association.

nature portfolio

Corresponding author(s): Xugao Wang

Last updated by author(s): Jan 31, 2025

Reporting Summary

Nature Portfolio 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 Portfolio 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. Source code of the simulation model to generate Extended Data Fig. 5 and to estimate the spatial pattern indices can be found in the Supplementary Information of Wiegand et al. (2021).
Data analysis	Analyses to generate phylogenetic tree and calculate phylogenetic distance for each plot were performed in R 4.3.2. We used PhyloMaker 2 Package version number 0.1.0 and Picante package version number 1.8.2.

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

A minimum datasets that allows to repeat, interpret, verify and extend the research in the article can be found in the Supplementary Information as Supplementary

Data Table 1, it provides for all species the indices of spatial patterns and all other quantities used in the model analysis and to generate Table 1, Figures 2, 3, 4, and Extended Data Figs. 2, 3, 4, 6, 7, 8, and 9. Additional raw data for Figure 3 are shown in Extended Data Table 1. The output data of the individual-based forest simulation model used to generate Extended Data Fig. 5 can be found in in the Supplementary Information as Supplementary Data Table 2. To estimate the spatial pattern indices, we used the raw census data of the ForestGEO network. The raw census data of several sites are publicly available (https://forestgeo.si.edu/explore-data), for the other sites they are available upon reasonable request and with permission of the principal investigators of the corresponding ForestGEO sites (https://forestgeo.si.edu/sites-all). Mycorrhizal association types were based on available global datasets (ref. 53, 54, 55) and website sources (https://www.mycorrhizas.info). To determine if a species is mainly dispersed by animals (zoochory) we used the Seed Information Database (https://ser-sid.org) of the Society for Ecological Restoration and the Royal Botanic Gardens Kew.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

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 conducts a systematic latitudinal analysis of conspecific aggregation and heterospecific association among neighboured trees across 21 large 20-50 ha inventory plots of temperate to tropical forest. To derive theoretical expectations for how the observed spatial patterns may impact species coexistence, we incorporate them into a mathematical model featuring neighbourhood crowding competition and estimated the per capita population growth rate of each species for low abundances. The approach includes analysis of fully-mapped forest plots of the Forest Global Earth Observatory (ForestGEO) to derive indices of conspecific aggregation and heterospecific segregation of 720 species and analysis of mathematical and simulation models driven by these data.
Research sample	We used ForestGeo (https://forestgeo.si.edu/explore-data) data sets of 21 large forest dynamics plots of areas between 20 and 50 ha.
Sampling strategy	The 21 forest plots included in the study have been completely censused for trees at least 1 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 >= 10cm dbh (diameter at breast height), with shrub species being excluded, 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 (dbh >= 10 cm) were used in the study as focal species to ensure that estimates of spatial patterns were reliable. We also excluded 15 of 737 species because they showed lower levels of aggregation than heterospecific association (a special case not covered by our model), and we excluded the two species Picea mariana and Thuja occidentalis of the Wabikin forest that are located in a patch of successional forest that was logged approximately 40 years ago.
Reproducibility	The estimation of the measures of spatial patterns is specified exactly in the publication such that it could be reproduced (Box 1). The code of the simulation model was published as supplementary information in a previous publication (Wiegand et al, 2021, ref. 8).
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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks | N/A

Novel plant genotypes | N/A

Authentication | N/A