LETTERS

Density dependence explains tree species abundance and diversity in tropical forests

Igor Volkov[1], Jayanth R. Banavar[1], Fangliang He[2], Stephen P. Hubbell[3,4] & Amos Maritan[5]

The recurrent patterns in the commonness and rarity of species in ecological communities—the relative species abundance—have puzzled ecologists for more than half a century[1,2]. Here we show that the framework of the current neutral theory in ecology[3–10] can easily be generalized to incorporate symmetric density dependence[11–14]. We can calculate precisely the strength of the rare-species advantage that is needed to explain a given RSA distribution. Previously, we demonstrated that a mechanism of dispersal limitation also fits RSA data well[3,4]. Here we compare fits of the dispersal and density-dependence mechanisms for empirical RSA data on tree species in six New and Old World tropical forests and show that both mechanisms offer sufficient and independent explanations. We suggest that RSA data cannot by themselves be used to discriminate among these explanations of RSA patterns[15]—empirical studies will be required to determine whether RSA patterns are due to one or the other mechanism, or to some combination of both.

Ecologists have long sought to explain the high levels of tree diversity that often occur in tropical forests. One aspect of this challenge is to understand the evolutionary origin and maintenance of this diversity on large spatial and temporal scales[16]. Another is to understand how such extraordinarily high alpha (local) tree diversity can be maintained on very local scales in particular tropical forests. For example, there are over a thousand tree species in a 52-hectare plot in Borneo (Lambir, Sarawak, Table 1). Numerous mechanisms have been proposed to explain tropical tree species coexistence on local scales; many of these hypotheses invoke density- and frequency-dependent mechanisms. Two of the most prominent of these hypotheses are the Janzen–Connell hypothesis[11,12] and the Chesson–Warner hypothesis[13].

The Janzen–Connell hypothesis is that seeds that disperse farther away from the maternal parent are more likely to escape mortality from host-specific predators or pathogens. This spatially structured mortality disfavours the population growth of locally abundant species relative to uncommon species by reducing the probability of species' self-replacement in the same location in the next generation.

The Chesson–Warner hypothesis is that a rare-species reproductive advantage arises when species have similar per capita rates of mortality but reproduce asynchronously, and there are overlapping generations. Processes that hold the abundance of a common species in check inevitably lead to rare-species advantage because the space or resources freed up by density-dependent deaths are then exploited by less-common species. Therefore, among-species frequency dependence is the community-level consequence of within-species density dependence, and thus they are two different manifestations of the same phenomenon. There is accumulating empirical evidence that such density- and frequency-dependent processes may play a large part in maintaining the diversity of tropical tree communities[17–22].

Density and frequency dependence are familiar mechanisms in population biology, but it is surprising how rarely their consequences for species diversity and RSA in communities have been discussed (see Ch. 3 of ref. 4). Here we show that these mechanisms are sufficient to explain precisely the species abundance patterns in six tropical forest communities on three continents.

The neutral theory of biodiversity provides a convenient theoretical framework for linking community diversity patterns to the fundamental mechanisms of population biology (such as birth, death and migration) and speciation[4]. The celebrated statistical distribution for RSA, Fisher's log-series[1], can be shown to arise directly from the stochastic equations of population growth under neutrality at the speciation-extinction equilibrium. More significantly, Fisher's log-series arises when the birth and death rates are density-independent[3].

According to the theory, the mean number of species with n individuals, $\langle \phi_n \rangle$, in a community at the stochastic speciation–extinction equilibrium takes the general form:

$$\langle \phi_n \rangle = SP_0 \left\langle \prod_{i=0}^{n-1} \frac{b_{i,k}}{d_{i+1,k}} \right\rangle_k$$

where $\langle \ldots \rangle_k$ represents the arithmetic average over all species, S is the average number of species present in the ecosystem, P_0 is a constant, and $b_{i,k}$ and $d_{i,k}$ are birth and death rates for the kth species with i individuals. Here we have subscripted the birth and the death rates for arbitrary species k to indicate that these rates could, in principle, be species-specific for an asymmetric community. In contrast, "symmetry occurs at the species level when no change in community dynamics or the fates of individuals occurs upon switching the species of any two given populations in the community. Any given population behaves as it would previously, despite its new species label, and its effects on other populations remain the same, regardless of their species labels." (P. Chesson, personal communication).

We note that what is important in determining the mean number of species, $\langle \phi_n \rangle$, are not the absolute rates of birth or death but their ratio, $r_{i,k} = \frac{b_{i,k}}{d_{i+1,k}}$. Indeed, $\langle \phi_n \rangle$ is proportional to $\langle r_{1,k} r_{2,k} \ldots r_{n-1,k} \rangle_k$. This formulation is sufficiently general to represent communities of either symmetric or asymmetric species. Such a situation could arise, for example, from niche differences or from differing immigration fluxes resulting from the different relative abundances of the species in the metacommunity. Hereafter, however, we consider only the symmetric case of a community of non-interacting species with identical vital demographic rates. For large community size, this formulation is equivalent to the case of zero-sum dynamics studied by Hubbell[4] (see Supplementary Information).

[1]Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA. [2]Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2H1, Canada. [3]Department of Plant Biology, The University of Georgia, Athens, Georgia 30602, USA. [4]The Smithsonian Tropical Research Institute, Box 2072, Balboa, Panama. [5]Dipartimento di Fisica 'G. Galilei', Universitá di Padova, via Marzolo 8, 35131 Padova, Italy.

Table 1 | Model parameters for the six data sets

Dispersal limitation model	S	J	θ_1	m	θ_2	c	x
BCI, Panama	225	21457	48.1	0.09	47.5	1.80	0.9978
Yasuni, Ecuador	821	17546	204.2	0.43	213.2	0.51	0.9883
Pasoh, Malaysia	678	26554	192.5	0.09	189.5	1.95	0.9932
Korup, Cameroon	308	24591	52.9	0.54	53.0	0.24	0.9979
Lambir, Malaysia	1004	33175	288.8	0.11	301.0	2.02	0.9915
Sinharaja, Sri Lanka	167	16936	27.3	0.55	28.3	0.38	0.9983

Model comparison	L_1	L_2	Deviance	P-value
BC1, Panama	−314.0	−315.0	2.0	0.16
Yasuni, Ecuador	−301.0	−303.6	5.2	0.02
Pasoh, Malaysia	−363.7	−365.3	3.2	0.07
Korup, Cameroon	−322.3	−323.1	1.6	0.21
Lambir, Malaysia	−390.5	−391.2	1.4	0.24
Sinharaja, Sri Lanka	−258.9	−258.5	0.8	0.37

Maximum-likelihood estimates of the dispersal limitation model[3] and the density-dependent symmetric model parameters (upper table) and comparison between the models (lower table) for the six data sets of tropical forests. In the six plots coordinated by Center for Tropical Forest Science of the Smithsonian (http://www.ctfs.si.edu), we considered trees with diameter at breast height ≥10 cm. S is the number of species, J is the total abundance and θ_1 and θ_2 are the biodiversity parameters in the dispersal limitation model[3] and equation (1) respectively (note that θ_2 is a function of c, x and S and that both models have the same number of fitting parameters). The comparison of the models was carried out with the likelihood ratio test[10,23,24]. The lower table presents deviance (twice the difference in the log-likelihoods L_1 and L_2 of the dispersal limitation model[3] and the density-dependent symmetric model respectively) between the two models and the corresponding P-value of the χ^2-distribution with one degree of freedom. The main result is that the dispersal-limitation model and the simple symmetric density-dependent model presented here are statistically comparable to each other in their ability to fit the tropical forest data.

We define $\hat{r}_n = \frac{b_n}{d_{n+1}}\frac{n+1}{n}$, where the factor $\frac{n+1}{n}$ is chosen to obtain $\hat{r}_n = x$ for the Fisher log-series $\langle\phi_n\rangle \propto x^n/n$. In Fisher's case, $\hat{r}_n$ does not change with population density and is an intraspecific parameter that measures the relative vital rates of birth and death of a population. To obtain intraspecific density dependence, $\hat{r}_n$ becomes a function of the population density n. Within our framework, $\frac{\langle\phi_{n+1}\rangle}{\langle\phi_n\rangle} = \frac{n}{n+1}\hat{r}_n$.

We now introduce the modified symmetric theory that captures density dependence (rare-species advantage or common-species disadvantage). In the modified theory, $\hat{r}_n$ will be a decreasing function of abundance, thereby incorporating density dependence. The equations of density dependence in the per capita birth and death rates for an arbitrary species of abundance n are:

$$\frac{b(n)}{n} = b \times \left[1 + \frac{b_1}{n} + o\left(\frac{1}{n^2}\right)\right]$$

and

$$\frac{d(n)}{n} = d \times \left[1 + \frac{d_1}{n} + o\left(\frac{1}{n^2}\right)\right]$$

for $n > 0$ as the leading terms of a power series in $(1/n)$, $\frac{b(n)}{n} = b \times \sum_{l=0}^{\infty} b_l n^{-l}$ and $\frac{d(n)}{n} = d \times \sum_{l=0}^{\infty} d_l n^{-l}$, where b_l and d_l are constants. This expansion captures the essence of density dependence by ensuring that the per capita birth rate to death rate ratios decrease and approach a constant value for large n. This happens because the higher-order terms are negligible. Note that the quantity that controls the RSA distribution is the ratio b_n/d_{n+1}. Thus the birth and death rates, b_n and d_n, are defined up to multiplicative factors $f(n+1)$ and $f(n)$ respectively, where f is any arbitrary well-behaved function.

We expect that the per capita birth rate or the fecundity will go down as the abundance increases, whereas the mortality ought to increase with abundance. Indeed, the per capita death rate can be arranged to be an increasing function of n, as observed in nature, by choosing an appropriate function f and adjusting the birth rate appropriately so that the ratio b_n/d_{n+1} remains the same. For example, the choice $f(n) = n/(n+c)$ yields a constant per capita death rate $d_n = dn$ and a fecundity that decreases with increasing abundance.

This mathematical formulation of density dependence may seem unusual to ecologists familiar with the logistic or Lotka–Volterra systems of equations, in which density dependence is typically described as a polynomial expansion of powers of n truncated at the quadratic level. However, this classical expansion is not valid in

our context because the range of n is from 1 to an arbitrarily large value, not to some fixed carrying capacity. Therefore an expansion in terms of powers of $(1/n)$ is more appropriate. For this symmetric model, noting that $\langle\phi_n\rangle = SP_0\prod_{i=0}^{n-1}\frac{b_n}{d_{n+1}}$, we readily arrive at the following relative species-abundance relationship:

$$\langle\phi_n\rangle = \theta\frac{x^n}{n+c} \quad (1)$$

where $x = b/d$, and for parsimony we have made the simple assumption that $b_1 = d_1 = c$. The biodiversity parameter θ is the normalization constant that ensures that the average number of species in the community is S and is given by $\theta = S\frac{1+c}{cx}F^{-1}(1+c, 2+c, x)$, where $F(1+c, 2+c, x)$ is the standard hypergeometric function. The parameter c measures the strength of the symmetric density dependence in the community, and it controls the shape of the RSA distribution. Note that when $c \to 0$ (the case of no density dependence), we obtain the Fisher log-series. In this case, as shown in ref. 3, θ captures the effects of speciation.
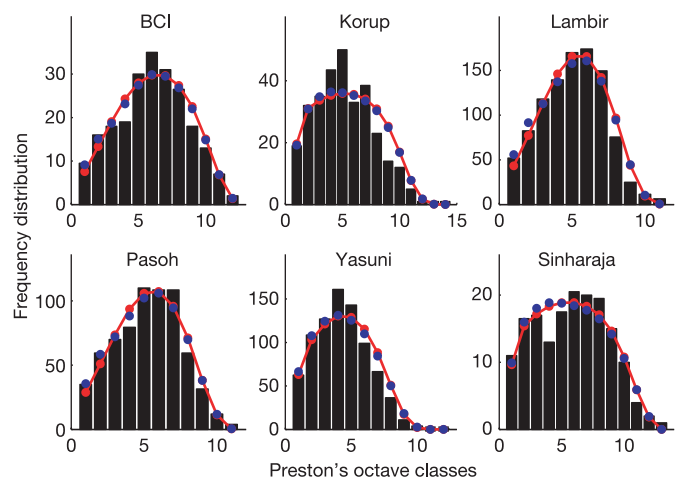

Figure 1 | Fits of density-dependent symmetric model (red line) and dispersal-limitation model[3] (blue circles) to the tree species abundance data from the BCI, Yasuni, Pasoh, Lambir, Korup and Sinharaja plots, for trees ≥10 cm in stem diameter at breast height (see Table 1). The frequency distributions are plotted using Preston's binning method as described in ref. 3. The numbers on the x axis represent Preston's octave classes.
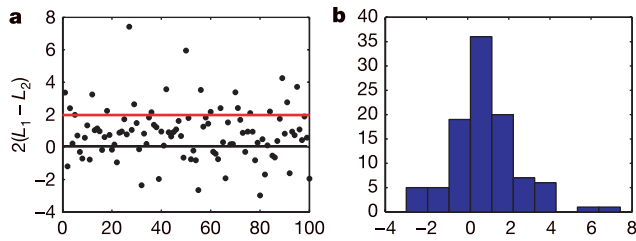
Figure 2 | Test of the equivalence of the dispersal limitation model and the density-dependent symmetric model. We have randomly generated 100 communities using the dispersal–limitation model, following the Etienne[23] algorithm with the values of m and θ_1 given in Table 1 for the BCI plot ($m = 0.09$ and $\theta_1 = 48.1$). For each of these data sets we calculated the log-likelihoods L_1 and L_2 of the dispersal-limitation model[3] and the density-dependent symmetric model respectively. **a,** The deviances for each of the data set; **b,** the corresponding histogram. The red line in **a** corresponds to the deviance estimated from the BCI data. The deviance is distributed almost symmetrically around zero (with a mean value of 0.85) and the BCI data is indistinguishable from the random pseudo-samples.

Figure 3 | Plot of $\hat{r}_n$ derived from equation (1) versus n for the six data sets of tropical trees. For large values of n, $\hat{r}_n$ asymptotes to a value slightly less than 1. The BCI data (cyan circles) at small n is almost invisible because it coincides with the Pasoh data set (red circles).

We now show how equation (1) estimates the strength of symmetric density dependence that is consistent with the observed RSA distributions of tree species in six large tropical forest plots on three continents: Barro Colorado Island (BCI), Panama; the Yasuni National Park, Ecuador; the Pasoh Forest Reserve, peninsular Malaysia; the Korup National Park, Cameroon; the Lambir Hills National Park, Sarawak, Malaysia; and the Sinharaja World Heritage Site, Sri Lanka. These site plots are part of a global network of large plots managed by the Center for Tropical Forest Science of the Smithsonian Tropical Research Institute. These New and Old World tropical forests have had long separate ecological and evolutionary histories, but despite these different histories, the symmetric theory with density dependence fits each of the RSA distributions very well.

Figure 1 shows the fits of equation (1) and the dispersal limitation model[3] to the tree abundance data collected from the six permanent plots of tropical forest. These plots are 50 hectares except for Lambir (52 hectares), Yasuni (25 hectares) and Sinharaja (25 hectares). The results in Table 1 and Fig. 1 show that the RSA data of tree species in these plots are equally well described both by the density-dependent model and the dispersal-limitation model[3] (also see Fig. 2).

The rare-species advantage is illustrated in Fig. 3 and is of the same order of magnitude in the different forests. The key quantity that controls the RSA is the birth rate to death rate ratio $\hat{r}_n$ defined above. The curves in Fig. 3 were derived from the parameters in Table 1, which in turn were obtained from the empirical RSA data using the maximum-likelihood method. At stochastic steady state, community size (mass balance) is maintained by the slow rate of decline in common species (at large n in Fig. 3) exactly balanced by the growth of rare species, and by the very slow input of new species by speciation.

Several important ecological insights result from this new theory. First, we have shown that an assumption of asymmetric density dependence, for example, postulating different carrying capacities for each species, is not necessary to explain patterns of RSA at least in these six tropical forests: a much simpler symmetric hypothesis is sufficient. Second, we have shown that the population sizes that exhibit rare-species advantage consistent with the observed RSA data are all quite small. The transition to a 'Fisher log-series'-like value for $x = b/d$ that is slightly less than replacement occurs at what would be considered low population densities of tree species in these forests (<1 tree per hectare). Third, applied to spatially mapped ecological communities, the theory reveals the scales on which density dependence occurs and on what larger scales it gives way to density independence. In these six tropical forests, the theory is consistent with density dependence operating at very similar and small scales of abundance and spatial length, and weakening at larger scales.
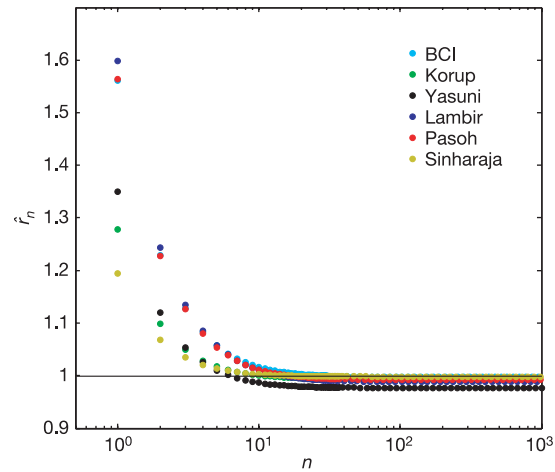
Finally, we have demonstrated that symmetric density dependence gives an equally sufficient mechanistic explanation for RSA patterns, in addition to and independent of dispersal limitation[3]. In Table 1, we show the fits of the two mechanisms to the RSA data from the six forests, from which it is clear that both mechanisms yield fits that cannot be distinguished statistically in quality. However, the ecological explanation that accompanies each of these mechanisms is very different. According to the dispersal mechanism, the explanation for the lower frequency of rare species compared to species of middling abundance is that rare species are more extinction-prone, and when they go extinct in a community, they take longer to re-immigrate than common species do. According to the density-dependence mechanism, on the other hand, the reduced steady-state frequency of rare species arises because populations of rare species grow differentially faster into higher abundance categories owing to a rare-species advantage. An important conclusion is that we cannot deduce the mechanisms causing a particular RSA pattern from RSA data alone (see Supplementary Information). Because these mechanisms are not mutually exclusive, it must be left to empirical research to uncover the relative contributions of each mechanism to observed RSA patterns. However, we do note one distinction between the two mechanisms. The dispersal-limitation mechanism generally implies that we are considering a local community into which immigration is possible. However, the density dependence mechanism can apply equally well to local communities or to the metacommunity.

1. Fisher, R. A., Corbet, A. S. & Williams, C. B. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12,** 42–58 (1943).
2. Preston, F. W. The commonness and rarity of species. *Ecology* **29,** 254–283 (1948).
3. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Neutral theory and relative species abundance in ecology. *Nature* **424,** 1035–1037 (2003).
4. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. New Jersey, 2001).
5. Bell, G. The distribution of abundance in neutral communities. *Am. Nat.* **155,** 606–617 (2000).
6. McKane, A., Alonso, D. & Solé, R. V. Mean-field stochastic theory for species-rich assembled communities. *Phys. Rev. E* **62,** 8466–8484 (2000).
7. Bell, G. Neutral macroecology. *Science* **293,** 2413–2418 (2001).
8. Vallade, M. & Houchmandzadeh, B. Analytical solution of a neutral model of biodiversity. *Phys. Rev. E* **68,** 061902 (2003).
9. Houchmandzadeh, B. & Vallade, M. Clustering in neutral ecology. *Phys. Rev. E* **68,** 061912 (2003).

10. Alonso, D. & McKane, A. J. Sampling Hubbell's neutral theory of biodiversity. *Ecol. Lett.* **7,** 911–914 (2004).
11. Janzen, D. H. Herbivores and the number of tree species in tropical forest. *Am. Nat.* **104,** 501–528 (1970).
12. Connell, J. H. in *Dynamics of Populations* (eds Den Boer, P. J. & Gradwell, G. R.) 298–312 (Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands, 1971).
13. Chesson, P. L. & Warner, R. R. Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* **117,** 923–943 (1981).
14. Chesson, P. A need for niches? *Trends Ecol. Evol.* **6,** 26–28 (1991).
15. Etienne, R. S. & Olff, H. A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* **7,** 170–175 (2004).
16. Morley, R. J. *Origin and Evolution of Tropical Rainforests* Ch. 5 (Wiley, New York, 2000).
17. Augspurger, C. K. Seedling survival of tropical tree species: interactions of dispersal distance, light gaps, and pathogens. *Ecology* **65,** 1705–1712 (1984).
18. Hubbell, S. P., Condit, R. & Foster, R. B. Presence and absence of density dependence in a neotropical tree community. *Trans. R. Soc. Lond.* **330,** 269–281 (1990).
19. Gilbert, G. S., Hubbell, S. P. & Foster, R. B. Density and distance-to-adult effects of a canker disease of trees in a moist tropical forest. *Oecologia* **98,** 100–108 (1994).
20. Condit, R., Hubbell, S. P. & Foster, R. B. Density dependence in two understory tree species in a neotropical forest. *Ecology* **75,** 671–680 (1994).
21. Harms, K. E., Wright, S. L., Calderón, O., Hernández, A. & Herre, E. A. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* **404,** 493–495 (2000).
22. Wright, S. J. Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia* **130,** 1–14 (2002).
23. Etienne, R. S. A new sampling formula for neutral biodiversity. *Ecol. Lett.* **8,** 253–260 (2005).
24. Hilborn, R. & Mangel, M. *The Ecological Detective: Confronting Models with Data* Ch. 7 (Princeton Univ. Press, Princeton, New Jersey, 1997).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.R.B. (banavar@psu.edu) or A.M. (maritan@pd.infn.it).