Supporting Information
Covariations in ecological scaling laws fostered by community dynamics

Silvia Zaoli[a,1], **Andrea Giometto**[a,b,1,2], **Amos Maritan**[c,d,2], **and Andrea Rinaldo**[a,e,2]

[a]Laboratory of Ecohydrology, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; [b]Department of Physics, Harvard University, Cambridge, MA 02138; [c]Dipartimento di Fisica ed Astronomia, Università di Padova, I-35131 Padova, Italy; [d]Istituto Nazionale di Fisica Nucleare, I-35131 Padova, Italy; [e]Dipartimento di Ingegneria Civile, Edile ed Ambientale, Università di Padova, I-35131 Padova, Italy

The SI is organized as follows. Section 1 gives details concerning several theoretical results which underlie the derivation of the macroecological scaling linkages and generalizations of our scaling framework. Section 2 addresses the compatibility of our framework with empirical evidence of scaling macroecological laws. Section 3 provides details and results of community dynamics models which corroborate the validity and general applicability of our scaling framework.

1 Mathematical scaling framework

The choice of variables suitable to describe an ecological system is not obvious. In our case, the choice of n, m and A is the one that better fits the scope of accounting for resource constraints. Individual metabolic rate is assumed to depend on body mass as $b = cm^{\alpha}$, with $\alpha \leq 1$. Thus, m and n determine the metabolic rate of a species, and by summing over all species one obtains the total ecosystem metabolic rate B, i.e. its resource consumption rate. By assuming that the resources supply rate $\mathcal{R}$ is proportional to the ecosystem area A and that the total resource consumption rate (total metabolism B) is constrained by $\mathcal{R}$, we have $B \propto A$ (SI section 1.1 and 1.2). The three variables n, m and A thus allow us to enforce the constraint of finite resources supply by imposing $B \propto \mathcal{R} \propto A$.

1.1 Proportionality of resources supply rate and ecosystem area. In our scaling framework and community dynamics models, we fix the resource-supply per unit area to unity, that is $\mathcal{R}/A = r = 1$. The constant r determines the type of ecosystem described by the model (endowed with abundant or scarce resources, which could be interpreted e.g. as a tropical forest or a desert). $\mathcal{R}$ will also affect the total number of species S via the proportionality of the total consumption rate to the resources supply rate $B = cSI_{1,\alpha} = rA$, where c is the proportionality constant that appears in Kleiber's law. Thus, the number of species S varies linearly with r: $S = r/c \cdot A/I_{1,\alpha}$. Similarly, the total biomass and total abundance vary linearly with r given that $M = SI_{1,1}$ and $N = SI_{1,0}$. The power-law scaling behavior for these quantities is valid at fixed values of r, that is, for ecosystems of similar type, for example a set of islands, lakes or forests sharing the same climate and environmental conditions.

We expect $\mathcal{R} \propto A$ to hold strictly for a community composed of a single trophic level, such as plants competing for light in a forest. In a generalized case, say describing multitrophic levels, resources (and thus the total metabolism) might scale non-linearly with the area $\mathcal{R} \propto A^{\kappa}$ for certain trophic levels. For example, in a two-trophic levels community, the predator resource is the prey biomass, which might scale non-linearly with A (this case is discussed in section 1.8.2). However, if the prey biomass scales linearly with the area, which is most often the case (see section 1.8.2 and Eq. 14 of the main text), then $\kappa = 1$.

1.2 Proportionality of total metabolic rate and resources supply rate. In the derivation of macrecological laws, we assume that the total metabolism B of a community (i.e. the total resource consumption rate) is proportional to the resource supply rate $\mathcal{R}$. Such assumption is motivated by the following reasoning. Assume $B \propto \mathcal{R}^{\iota}$. The total metabolism per unit resources is thus $B/\mathcal{R} \propto \mathcal{R}^{\iota-1}$. If $\iota = 1$, $B/\mathcal{R}$ is constant in the limit of large A. If $\iota > 1$, the total metabolism per unit resources would diverge in the limit of large A, which would not be sustainable. If $\iota < 1$, conversely, the total metabolism per unit resources would tend to zero in the limit of large A, so that resources would be completely unexploited. We expect that, at stationarity, the total metabolism of the community will be the maximum sustainable one, which corresponds to a linear dependence of the total metabolic rate on the resources supply rate $\mathcal{R}$.

1.3 Ansatz for the joint probability distribution of mass and abundance. Consider an ecosystem of area A, and let $p(n, m|A)dm$ be the joint probability of finding a species of abundance $n \in \mathbb{N}$ (including $n = 0$) and typical mass $m \in [m, m + dm]$. The joint probability distribution $p(n, m|A)$ differs from $P(n, m|A)$ presented in the main text, as described in this section. We assume that the minimum viable mass for an organism is $m_0 > 0$ which is independent of A, so that $p(n, m|A)$ is null for $m < m_0$. We measure mass and area in units of m_0 and a reference area a_0, so that m and A are dimensionless.

For a fixed ecosystem area A and in the limit of infinite mass, the probability that a species has abundance $n > 0$ must go to zero, i.e. $\lim_{m\to\infty} p(n > 0|m, A) = 0$. Because $p(n > 0|m, A) = 1 - p(n = 0|m, A)$, one has $\lim_{m\to\infty} p(n = 0|m, A) = 1$, which implies that for large m there is a finite probability that $n = 0$, i.e. $p(n = 0|m, A) > 0$. Therefore, when approximating $p(n, m|A)$ with a continuous (in both n and m) probability density*, care must be taken to separate the value $n = 0$ from $n > 0$. Such procedure is analogous to the

*i.e., for $n > 0$, $p(n, m|A)dndm$ is the probability that a species has abundance in $(n, n+dn)$ and mass in $(m, m + dm)$.

S.Z., A.G., A.M. and A.R. designed research; S.Z. and A.G. performed research and analyzed data with assistance from A.M. and A.R.; and S.Z., A.G., A.M. and A.R. wrote the manuscript.

[1]S.Z. and A.G. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: andrea.rinaldo@epfl.ch, giometto@fas.harvard.edu or maritan@pd.infn.it

separation of the lowest energy state from the excited ones in the computation of the energy distribution of a Bose-Einstein condensate (1). For example, the normalization condition for $p(n, m|A)$ reads:

$$1 = \int_1^\infty p(0, m|A)dm + \int_1^\infty \int_0^\infty p(n, m|A)dndm, \quad \text{[S1]}$$

where we evaluated the contribution of $n = 0$ separately. We indicate with $p(m|A)$ and $p(n|m, A)$ the distributions derived from $p(n, m|A)$ via the identity $p(n, m|A) = p(m|A)p(n|m, A)$. When estimating the joint probability distribution of mass and abundance in real datasets or in simulations of community dynamics models, one only observes species that exist, i.e. one evaluates probability distributions conditional on $n > 0$. Thus, we indicate with $P(n, m|A) = p(n, m|A, n > 0)$ the joint probability distribution that one can estimate in real datasets or in community dynamics models simulations. Correspondingly, we indicate with $P(m|A) = p(m|A, n > 0)$ and $P(n|m, A) = p(n|m, A, n > 0)$ the distributions derived from $P(n, m|A)$. The relationship between $P(m|A)$ and $p(m|A)$ can be derived using Bayes' theorem as follows:

$$\begin{aligned} P(m|A) = p(m|A, n > 0) &= \frac{p(m|A)p(n > 0|m, A)}{p(n > 0|A)} \\ &= p(m|A)\frac{1 - p(n = 0|m, A)}{1 - p(n = 0|A)}, \end{aligned} \quad \text{[S2]}$$

which can be recast as:

$$p(m|A) = p(n = 0, m|A) + [1 - p(n = 0|A)]\, P(m|A), \quad \text{[S3]}$$

where we have used the fact that $p(n = 0|m, A) = p(n = 0, m|A)/p(m|A)$ to express $p(n = 0|m, A)$ in terms of $p(n = 0, m|A)$. $P(n|m, A)$, in turn, is $p(n|m, A)$ renormalized for $n > 0$, that is:

$$P(n|m, A) = \frac{p(n|m, A)}{\sum_{n>0} p(n|m, A)} = \frac{p(n|m, A)}{1 - p(n = 0|m, A)}. \quad \text{[S4]}$$

Eqs. S2 and S4 allow us to express $p(n, m|A)$ in terms of the experimentally or numerically observable distributions $P(m|A)$ and $P(n|m, A)$ as:

$$\begin{aligned} p(n, m|A) =& p(m|A)p(n|m, A) \\ =& [1 - p(n = 0|A)]\, P(m|A)P(n|m, A) \\ =& [1 - p(n = 0|A)]\, P(n, m|A), \end{aligned} \quad \text{[S5]}$$

which holds for all $n > 0$. Note that $P(n, m|A)$ is simply $p(n, m|A)$ renormalized for $n > 0$. Moments of $P(n, m|A)$ are computed as:

$$I_{j,k} = \int_1^\infty \int_0^\infty n^j m^k P(n, m|A)dndm, \quad \text{[S6]}$$

whereas moments of $p(n, m|A)$ are computed as:

$$\begin{aligned} i_{j,k} =& 0^j \int_1^\infty m^k p(0, m|A)dm + \int_1^\infty \int_0^\infty n^j m^k p(n, m|A)dndm \\ =& 0^j \int_1^\infty m^k p(0, m|A)dm + [1 - p(n = 0|A)] \\ & \times \int_1^\infty \int_0^\infty n^j m^k P(n, m|A)dndm \\ =& 0^j \int_1^\infty m^k p(0, m|A)dm + [1 - p(n = 0|A)]\, I_{j,k}, \end{aligned} \quad \text{[S7]}$$

and therefore for $j > 0$ one has: $i_{j,k}/I_{j,k} = [1 - p(n = 0|A)]$. Because $1 - p(n = 0|A)$ is limited, $i_{j,k}$ and $I_{j,k}$ have the same scaling with A.

Supported by the characterization of the stationary state of our stochastic community dynamics models (Section 3) and by the empirical scaling behavior of ecological patterns (Section 2), we put forward an *ansatz* for the analytical form of the joint probability distribution of species abundances and masses $P(n, m|A)$. Thus, for $P(m|A)$, $P(n|m, A)$ and $\langle n|m, A\rangle$ we assume:

$$P(m|A) = (\delta - 1)m^{-\delta} \quad \text{[S8]}$$

$$P(n|m, A) = \hat{G}\left(\frac{n}{\langle n|m, A\rangle}\right) g(m, A), \quad \text{[S9]}$$

where:

$$\langle n|m, A\rangle = m^{-\gamma} A^\Phi h\left(\frac{m}{A^\lambda}\right). \quad \text{[S10]}$$

The term $g(m, A)$ allows further dependencies on A and m required for normalization and is characterized in the next section. The functions $\hat{G}(x)$ and $h(x)$ have the properties:

$$\int_0^\infty x^j \hat{G}(x)dx < \infty \qquad j = 0, 1, 2 \quad \text{[S11a]}$$

$$h(x) = o\left(x^{-2+\delta+\gamma}\right) \qquad \text{as } x \to \infty \quad \text{[S11b]}$$

$$\lim_{x\to 0} h(x) = h_0, \quad \text{[S11c]}$$

where h_0 is a positive constant. The rate of the decay of h and $\hat{G}$ for large arguments is such as to allow convergence of the (j, k)-th moment of $P(n, m|A)$ for $j = 0, 1, 2$ and $k \in [0, 1]$. These are all the moments needed to derive ecological scaling laws (see Methods). An example of functions satisfying the above requests is $h(x) = \hat{G}(x) = e^{-x}$.

1.4 Normalization of $P(n, m|A)$. Here we derive the normalization condition for $P(n, m|A)$. The marginal distribution $P(m|A)$ given in Eq. S8 is already normalized, as $\int_1^\infty dm(\delta - 1)m^{-\delta} = 1$. It remains to impose normalization on $P(n|m, A)$:

$$\begin{aligned} 1 &= \int_0^\infty P(n|m, A)dn = g(m, A) \int_0^\infty \hat{G}\left(\frac{n}{\langle n|m, A\rangle}\right) dn \\ &= g(m, A)\langle n|m, A\rangle \int_0^\infty \hat{G}(x)dx. \end{aligned} \quad \text{[S12]}$$

Thus, Eq. S12 reads:

$$1 = g(m, A)\langle n|m, A\rangle, \quad \text{[S13]}$$

where we have imposed $\int_0^\infty \hat{G}(x)dx = 1$ without loss of generality. Therefore:

$$P(n|m, A) = \frac{1}{\langle n|m, A\rangle}\hat{G}\left(\frac{n}{\langle n|m, A\rangle}\right) = n^{-1} G\left(\frac{n}{\langle n|m, A\rangle}\right), \quad \text{[S14]}$$

where we defined $G(x) = x\hat{G}(x)$. In conclusion, the normalized joint probability density distribution $P(n, m|A)$ reads:

$$P(n, m|A) = (\delta - 1)n^{-1}m^{-\delta} G\left[\frac{nm^\gamma}{A^\Phi}\frac{1}{h(m/A^\lambda)}\right]. \quad \text{[S15]}$$

This result motivates the choice of the factor n^{-1} in Eq. 2 of the main text. Note that Eq. S15 can be recast as:

$$P(n,m|A) = (\delta-1)n^{-1}m^{-\delta}G\left[\frac{n}{A^{\Psi}}\frac{(m/A^{\lambda})^{\gamma}}{h(m/A^{\lambda})}\right], \qquad \text{[S16]}$$

with $\Psi = \phi - \gamma\lambda$. For scaling to hold, Eq. S16 must be valid for $x = n/A^{\Psi}$, $y = m/A^{\lambda}$ fixed and $n, m, A \to \infty$. Therefore, one has $\Psi = \phi - \gamma\lambda > 0$.

1.5 Average abundance conditional on mass and area $\langle n|m, A\rangle$. We show here that the distributions:

$$\begin{aligned} P(m|A) &= (\delta-1)m^{-\delta} \\ P(n|m,A) &= n^{-1}G\left[\frac{nm^{\gamma}}{A^{\Phi}}\frac{1}{h(m/A^{\lambda})}\right] \end{aligned} \qquad \text{[S17]}$$

lead to the average abundance conditional on mass and area $\langle n|m, A\rangle = A^{\Phi}m^{-\gamma}h(m/A^{\lambda})$. In fact:

$$\begin{aligned} \langle n|m,A\rangle &= \int_0^{\infty} nP(n|m,A)dn = \int_0^{\infty} G\left[\frac{nm^{\gamma}}{A^{\Phi}}\frac{1}{h(m/A^{\lambda})}\right]dn \\ &= A^{\Phi}m^{-\gamma}h\left(\frac{m}{A^{\lambda}}\right)\int_0^{\infty}G(x)dx = A^{\Phi}m^{-\gamma}h\left(\frac{m}{A^{\lambda}}\right), \end{aligned} \qquad \text{[S18]}$$

where we have redefined without loss of generality $G(x)$ as $G'(x) = G(Cx)$, where $C = \int_0^{\infty}G(x)dx$, so that $\int_0^{\infty}G'(x)dx = \int_0^{\infty}x^{-1}G'(x)dx = 1$. We will drop the prime symbol in the remainder of the text to simplify the notation.

1.6 Moments of $P(n, m|A)$. One can compute the scaling of the (j,k)th-moment ($j = 1, 2$, $k \in [0, 1]$) with A, for large A, as:

$$\begin{aligned} I_{j,k} &= \int_1^{\infty}\int_0^{\infty} n^j m^k P(n,m|A)\; dn\; dm \\ &= (\delta-1)\int_1^{\infty} m^{k-\delta}\int_0^{\infty} n^{j-1}G\left[\frac{nm^{\gamma}}{A^{\Phi}}\frac{1}{h\,(m/A^{\lambda})}\right]dn\; dm \\ &= (\delta-1)A^{j\Phi}\int_1^{\infty} m^{k-\delta-j\gamma}h^j\left(\frac{m}{A^{\lambda}}\right)dm\int_0^{\infty}x^{j-1}G\,(x)\,dx \\ &\propto A^{j\Phi}\int_1^{\infty} m^{k-\delta-j\gamma}h^j\left(\frac{m}{A^{\lambda}}\right)dm \\ &\propto A^{j\Phi+\lambda(1+k-\delta-j\gamma)}\int_{1/A^{\lambda}}^{\infty} y^{k-\delta-j\gamma}h^j(y)dy \\ &\propto A^{j\Phi+\lambda(1+k-\delta-j\gamma)}\left[h_0^j\int_{1/A^{\lambda}}^{\epsilon} y^{k-\delta-j\gamma}dy\right. \\ &\qquad\left. +\int_{\epsilon}^{\infty} y^{k-\delta-j\gamma}h^j(y)dy\right] \\ &\propto A^{j\Phi+\lambda(1+k-\delta-j\gamma)}\left[c_1A^{-\lambda(1+k-\delta-j\gamma)}+c_2\right] \\ &\propto A^{j\Phi+\max\{0,\lambda(1+k-\delta-j\gamma)\}}, \end{aligned} \qquad \text{[S19]}$$

with $\epsilon \ll 1$, c_1 and c_2 constants. We used the property Eq. S11a to ensure that the integral $\int_0^{\infty}x^{j-1}G\,(x)\,dx$ converges for $j = 1, 2$, and properties (S11b-S11c) to evaluate the integral in m (in particular, $h(x) \simeq h_0$ constant for $x \in (0, \epsilon]$). Eq. S19 can be used to derive several macroecological scaling laws, as outlined in the Methods section.

1.7 Relative species abundance (RSA). The RSA is the distribution of species abundances $P(n|A)$. In our framework it can be obtained by marginalizing $P(n, m|A)$ over m:

$$\begin{aligned} P(n|A) &= \int_1^{\infty} P(n,m|A)dm \\ &= (\delta-1)n^{-1}\int_1^{\infty} m^{-\delta}G\left(\frac{n}{\langle n|m,A\rangle}\right)dm. \end{aligned} \qquad \text{[S20]}$$

This integral cannot be computed in the general case, that is, without specifying G and h. We compute it here for the particular choice $G(x) = \frac{1}{\sqrt{\pi\sigma}}\,e^{-\frac{1}{\sigma}(\log x+\sigma/4)^2}$, with $\sigma > 0$ constant. Note that $\int_0^{\infty}G(x)dx = \int_0^{\infty}G(x)/xdx = 1$ as prescribed in sections 1.4 and 1.5. For $h(x)$, it is sufficient to know that $h(x)$ is monotonically decreasing to carry out the calculations, but to simplify the expressions we take $h(x) = h_0e^{-x}$ with $h_0 > 0$ constant. With these assumptions, $P(n|A)$ reads:

$$P(n|A) = \frac{1}{n}\frac{\delta-1}{\sqrt{\pi\sigma}}\int_1^{\infty} m^{-\delta}e^{f(m)}dm, \qquad \text{[S21]}$$

where we defined

$$f(m) = -\frac{1}{\sigma}\left[\log\left(\frac{nm^{\gamma}}{h_0A^{\Phi}}\right)+\frac{m}{A^{\lambda}}+\frac{\sigma}{4}\right]^2. \qquad \text{[S22]}$$

The integral in Eq. S21 cannot be computed analytically. However, noticing that the contribution to the integral is maximum when $m = m^*$ where m^* maximizes $f(m)$, we can approximate the integral for certain values of n. The approximation is akin to the Laplace method, but it is not possible to give an upper bound on the error made by the approximation. Nonetheless, the approximation can be compared to the numerical computation of $P(n|A)$ (see Fig. S1). The derivative of $f(m)$ reads:

$$f'(m) = -\frac{2}{\sigma}\left[\log\left(\frac{nm^{\gamma}}{h_0A^{\Phi}}\right)+\frac{m}{A^{\lambda}}+\frac{\sigma}{4}\right]\left(\frac{\gamma}{m}+\frac{1}{A^{\lambda}}\right). \qquad \text{[S23]}$$

Note that the derivative $f'(m)$ is negative for any $m \in [1, \infty]$ if $n > A^{\Phi}h_0e^{-1/A^{\lambda}}e^{-\sigma/4} \simeq A^{\Phi}h_0e^{-\sigma/4}$. Thus, for $n \gg A^{\Phi}h_0e^{-\sigma/4}$ (i.e. in the tail of the distribution) $f(m)$ is maximum at $m^*_{\text{tail}} = 1$ and the approximation gives:

$$\begin{aligned} P(n|A)_{\text{tail}} &= \frac{1}{n}\frac{\delta-1}{\sqrt{\pi\sigma}}\frac{e^{-\frac{1}{\sigma}\left[\log\left(\frac{n}{h_0A^{\Phi}}\right)+\frac{1}{A^{\lambda}}+\frac{\sigma}{4}\right]^2}}{\frac{2}{\sigma}\left[\log\left(\frac{n}{h_0A^{\Phi}}\right)+\frac{1}{A^{\lambda}}+\frac{\sigma}{4}\right]\left(\gamma+\frac{1}{A^{\lambda}}\right)} \\ &\simeq \frac{1}{n}\frac{\delta-1}{\sqrt{\pi\sigma}}\frac{e^{-\frac{1}{\sigma}\left[\log\left(\frac{n}{h_0A^{\Phi}}\right)+\frac{\sigma}{4}\right]^2}}{\frac{2\gamma}{\sigma}\left[\log\left(\frac{n}{h_0A^{\Phi}}\right)+\frac{\sigma}{4}\right]}. \end{aligned} \qquad \text{[S24]}$$

Note that the tail of the RSA resembles that of a lognormal distribution, which is typically found empirically (2), plus a correction of the form $C_1 + C_2 \log n$ at the denominator, where C_1 depends on A. For comparison, we plotted in Fig. S1 the (rescaled) lognormal tail $\frac{1}{n}\frac{\delta-1}{\sqrt{\pi\sigma}}\frac{\sigma}{2\gamma}e^{-\frac{1}{\sigma}\left[\log\left(\frac{n}{h_0A^{\Phi}}\right)+\frac{\sigma}{4}\right]^2}$.

If $n < A^{\Phi}h_0e^{-1/A^{\lambda}}e^{-\sigma/4} \simeq A^{\Phi}h_0e^{-\sigma/4}$, the maximum of $f(m)$ occurs at a value $\hat{m} > 1$. However, one cannot solve $f'(m) = 0$ analytically to determine $\hat{m}$. We can approximate the RSA at small and intermediate values of n as follows. The

behavior of $P(n|A)$ can be characterized for $n \simeq A^{\Phi} h_0 e^{-\sigma/4}$ by recognizing that, at such values of n, the value m^*_{body} maximizing $f(m)$ is close to 1. Therefore, we approximate $\log\left(\frac{nm^{\gamma}}{h_0 A^{\Phi}} e^{m/A^{\lambda}}\right) \simeq \log\left(\frac{nm^{\gamma}}{h_0 A^{\Phi}} e^{1/A^{\lambda}}\right)$ in $f'(m) = 0$ (see Eq. S23) and solve for m, yielding $m^*_{\text{body}} \simeq \left[\frac{A^{\Phi} h_0}{n} e^{-1/A^{\lambda}} e^{-\sigma/4}\right]^{\frac{1}{\gamma}}$. By applying the approximation method, one finds that the approximation for the RSA for $n \simeq A^{\Phi} h_0 e^{-\sigma/4}$ is:

$$P(n|A)_{\text{body}} = \frac{\delta - 1}{\gamma} \left[A^{\Phi} h_0 e^{-\frac{\sigma}{4}}\right]^{\frac{1-\delta}{\gamma}} n^{-1-\frac{1-\delta}{\gamma}}, \qquad \text{[S25]}$$

which is a power law with exponent $-1 - \frac{1-\delta}{\gamma}$. The extents of the tail and body of the RSA distribution depend on the values of h_0 and σ. Fig. S1 shows the RSA computed numerically via Eq. S21 (black curve) and its approximations computed via equations Eq. S24 (blue curve) and Eq. S25 (green curve).

1.8 Generalizations of the scaling framework.

1.8.1 Consequences of curvature in Kleiber's law

Various studies have claimed that Kleiber's law, the relationship linking metabolic rates to body size, displays curvature in a log-log plot (3–6), implying departures from a power-law behavior. For example, respiration rates of mammals have been claimed to increase with body size as $m^{2/3}$ until $m_1 \simeq 400$ g and as $m^{3/4}$ above (6). Conversely, respiration rates of trees have been claimed to increase linearly with tree biomass m until $m_1 \simeq 40$ g and as $m^{3/4}$ above (4, 6). Our framework can be used to infer the implications of such curvature on other macroecological patterns, as we show here.

In order to perform analytical calculations, we will make the simplifying assumption that Kleiber's law is described by the piecewise power-law:

$$b(m) = \begin{cases} c_0 m^{\alpha_1} \text{ if } m < m_1 \\ c m^{\alpha} \text{ if } m \geq m_1 \end{cases} \qquad \text{[S26]}$$

where $c_0 = c\, m_1^{\alpha - \alpha_1}$ is such that $b(m)$ is continuous in $m = m_1$. Using Eq. S26 we can compute the total community consumption rate per species:

$$\begin{aligned}
B/S &= \int_1^{\infty} \int_0^{\infty} n b(m) p(n, m|A) dn dm \\
&= (\delta - 1) \int_1^{\infty} \int_0^{\infty} b(m) m^{-\delta} G\left[\frac{nm^{\gamma}}{A^{\Phi}} \frac{1}{h\left(m/A^{\lambda}\right)}\right] dn dm \\
&\propto A^{\Phi} \int_1^{\infty} b(m) m^{-\gamma-\delta} h\left(\frac{m}{A^{\lambda}}\right) dm \int_0^{\infty} G(x) dx \\
&\propto A^{\Phi} \int_1^{m_1} c_0 m^{\alpha_1-\gamma-\delta} h\left(\frac{m}{A^{\lambda}}\right) dm \\
&\quad + A^{\Phi} \int_{m_1}^{\infty} c m^{\alpha-\gamma-\delta} h\left(\frac{m}{A^{\lambda}}\right) dm \\
&\propto A^{\Phi+\lambda(1+\alpha_1-\gamma-\delta)} \int_{A^{-\lambda}}^{m_1 A^{-\lambda}} c_0 x^{\alpha_1-\gamma-\delta} h(x) dx \\
&\quad + A^{\Phi+\lambda(1+\alpha-\gamma-\delta)} \int_{m_1 A^{-\lambda}}^{\infty} c x^{\alpha-\gamma-\delta} h(x) dx.
\end{aligned} \qquad \text{[S27]}$$

We now consider separately two possible scenarios:

a) $A \ll m_1^{1/\lambda}$:

$$\begin{aligned}
B/S &\propto c_0 A^{\Phi+\lambda(1+\alpha_1-\gamma-\delta)} \left[\int_{A^{-\lambda}}^{\epsilon} x^{\alpha_1-\gamma-\delta} h(x) dx \right. \\
&\qquad \left. + \int_{\epsilon}^{m_1 A^{-\lambda}} x^{\alpha_1-\gamma-\delta} h(x) dx\right] \\
&\quad + c A^{\Phi+\lambda(1+\alpha-\gamma-\delta)} \int_{m_1 A^{-\lambda}}^{\infty} x^{\alpha-\gamma-\delta} h(x) dx \\
&\propto c_0 A^{\Phi+\lambda(1+\alpha_1-\gamma-\delta)} \left[c_3 + c_4 A^{-\lambda(1+\alpha_1-\gamma-\delta)} \right. \\
&\qquad \left. + \int_{\epsilon}^{m_1 A^{-\lambda}} x^{\alpha_1-\gamma-\delta} h(x) dx\right] \\
&\quad + c A^{\Phi+\lambda(1+\alpha-\gamma-\delta)} \int_{m_1 A^{-\lambda}}^{\infty} x^{\alpha-\gamma-\delta} h(x) dx \\
&\propto A^{\Phi+\max\{0,\lambda(1+\alpha_1-\delta-\gamma)\}},
\end{aligned} \qquad \text{[S28]}$$

where $\epsilon \ll 1$, c_3 and c_4 are constants. In the first line we have used the limiting behavior $\lim_{x\to 0} h(x) = h_0$ constant and in the last line we have taken the limit $m_1 \to \infty$ before evaluating the integral.

b) $A \gg m_1^{1/\lambda}$:

$$\begin{aligned}
B/S &\propto c_0 A^{\Phi+\lambda(1+\alpha_1-\gamma-\delta)} \int_{A^{-\lambda}}^{m_1 A^{-\lambda}} x^{\alpha_1-\gamma-\delta} h(x) dx \\
&\quad + c A^{\Phi+\lambda(1+\alpha-\gamma-\delta)} \left(\int_{m_1 A^{-\lambda}}^{\epsilon} x^{\alpha-\gamma-\delta} h(x) dx \right. \\
&\qquad \left. + \int_{\epsilon}^{\infty} x^{\alpha-\gamma-\delta} h(x) dx \right) \\
&\propto c_0 c_4 A^{\Phi} \left(1 - m_1^{1+\alpha_1-\gamma-\delta}\right) \\
&\quad + c A^{\Phi+\lambda(1+\alpha-\gamma-\delta)} \left[c_5 + c_6 m_1^{1+\alpha-\gamma-\delta} A^{-\lambda(\gamma+\delta-1-\alpha)}\right] \\
&\propto A^{\Phi+\max\{0,\lambda(1+\alpha-\gamma-\delta)\}},
\end{aligned} \qquad \text{[S29]}$$

where $\epsilon \ll 1$, c_5 and c_6 are constants and we have used the properties of h.

Thus, in the limit of large area (i.e. above the crossover value $m_1^{1/\lambda}$) the scaling of B/S with A is independent on α_1, whereas below such crossover value B/S scales as $A^{\Phi+\max\{0,\lambda(1+\alpha_1-\gamma-\delta)\}}$. Thus, the curvature in the relationship between individual metabolic rates and body mass translates into a curvature in the scaling of the specific community consumption rate B/S with A. The scaling of the total number of species S with A is in turn determined by the proportionality of the total consumption rate to the ecosystem area, $B \propto A$. Imposing such proportionality we find the following scaling of S with A:

a) $A \ll m_1^{1/\lambda}$: $S \propto A^{1-\Phi-\max\{0,\lambda(1+\alpha_1-\delta-\gamma)\}}$.

b) $A \gg m_1^{1/\lambda}$: $S \propto A^{1-\Phi-\max\{0,\lambda(1+\alpha-\delta-\gamma)\}}$.

These equations are generalizations of the linking relationship Eq. 6 of the main text and show that the curvature in the relationship between individual metabolic rates and

body mass causes a curvature in the scaling of S with A. The difference in scaling exponents between the two regimes at small and large areas (i.e. $A \ll m_1^{1/\lambda}$ and $A \gg m_1^{1/\lambda}$) is $\Delta z = \max\{0, \lambda\,(1+\alpha_1-\delta-\gamma)\} - \max\{0, \lambda(1+\alpha-\gamma-\delta)\}$. Recalling the values of α_1 and α for mammals reported at the beginning of this section one finds $\Delta z = \max\{0, \lambda\,(5/3-\delta-\gamma)\} - \max\{0, \lambda(7/4-\gamma-\delta)\}$, which if $\delta+\gamma < 5/3$ and $\lambda > 0$ is equal to: $\Delta z = -\lambda/12$, i.e. a concave curvature. In general, if $\lambda(1+\alpha_1-\delta-\gamma) > 0$ and $\lambda(1+\alpha-\delta-\gamma) > 0$, a convex curvature in Kleiber's law implies a concave curvature in the species-area relationship, and viceversa. Note that for tropical forests, where we find $\lambda = 0$ (see Methods), a curvature in Kleiber's law does not imply any curvature in the species area relationship.

Of course, the curvature in Kleiber's law is likely to be described by a smoother functional relationship than the piecewise-power law assumed in Eq. S26. However, such assumption should not affect the asymptotic estimates for small and large areas, that is, the scaling of S with A for $A \ll m_1^{1/\lambda}$ and $A \gg m_1^{1/\lambda}$.

We note here that a recent investigation by Chisholm *et al.* (7) has highlighted curvatures in species-area relationships that are qualitatively similar to the one we described here, arguing that the origin of such curvature relies in a transition from a niche-structured regime on small islands to a colonization-extinction balance regime on large ones (7). Despite the qualitative similarities between the predicted curvatures in the two studies, we believe that further research is required to understand whether the empirical curvatures reported in Chisholm *et al.* (7) can be ascribed to a curvature in Kleiber's law as reported in this section. Because our results hold strictly in the limit of large areas, however, we believe that the 'small-island effect' described in Chisholm *et al.* (7) is a fine detail that cannot be reconciled with our framework in its present form. We do not believe that this observation should call the general framework into question, but rather it highlights how our results should be applied to large ecosystems where scaling arguments can be informative. May we note that our framework deals with several macroecological patterns at once and not exclusively with the species-area relationship, and one must trade the level of detail with which each pattern is reproduced with the capability to derive the covariations of several macroecological patterns analytically. Finer details of some of these patterns may be addressed by including additional ecological processes in our approach, although we deem that such path is beyond the scopes of this investigation, which aims at establishing a general null model framework.

1.8.2 Multiple trophic levels

So far we have assumed that all individuals in the ecosystem share the same resources. Thus, our results hold strictly for organisms within the same trophic level. Power law distributions and relationships, however, are empirically observed even on multi-trophic level contexts. For example, power-law size spectra across trophic levels are routinely observed in marine microbial ecosystems over several orders of magnitude (8). Here, we show that the conceptual design of our framework can be applied to multi-trophic level systems by means of a simple example. We assume that species' interactions are described by a simple foodweb made of two trophic levels, with one level feeding on abiotic factors like light (i.e. the producers) and the latter feeding on the lower trophic level (i.e. the consumers). In such an ecosystem, producers would be described by our framework with resources limited by ecosystem area $\mathcal{R} \propto A$, while consumers would be described by the same framework where the limiting resource is now the total producers' biomass, which scales as A^{μ_p} (the subscript p identifies the producers). Thus, one can describe the two trophic levels separately, with the producers' level being described by our current framework and the consumers level being described by a similar scaling framework where resources scale as $\mathcal{R} \propto A^{\mu_p}$. The effect of such modification on scaling relationship is easy to compute and affects the species-area relationship via a modification of Eq. S48b:

$$z_c = \mu_p - \Phi_c - \max\{0, \lambda_c(1+\alpha_c-\eta_c)\}, \qquad \text{[S30]}$$

which differs from Eq. S48b only for the exchange of 1 with μ_p (the subscript c identifies the consumers). Such modification propagates to equations Eq. S48c and Eq. S48d (recall that $N = SI_{1,0}$ and $M = SI_{1,1}$) which become:

$$\begin{aligned} \mu_c &= \mu_p + \max\{0, \lambda_c(2-\eta_c)\} - \max\{0, \lambda_c(1+\alpha_c-\eta_c)\}, \\ \nu_c &= \mu_p - \max\{0, \lambda_c(1+\alpha_c-\eta_c)\}. \end{aligned} \qquad \text{[S31]}$$

We note, however, that μ_p is most likely equal to one in most ecosystems, as found for example in our empirical analysis of tropical forests datasets. In such a case, the consumers level is described exactly by the framework that we presented in the main text. Obviously, more complicated schemes of multi-level foodwebs may be envisioned in the proposed framework. On this, research is forthcoming.

1.8.3 Area-independent limitation on maximum size

Eq. 5 of the main text predicts that, if $z > 0$, the maximum species' mass increases with A as $m_{\max} = A^{\xi}$. When A is very large, area-independent constraints could settle in to limit the maximum body size, either due to physiological limits or due to ecological dynamics making larger body sizes unfavorable. Because $\xi = z/(\delta-1)$, the critical value of A above which the maximum body size is independent of A is equal to $A_c = M_0^{z/(\delta-1)}$. This observation can be reconciled with our framework by generalizing Eq. 7 as:

$$P(m|A) = m^{-\delta} H\left(\frac{m}{M_0}\right), \qquad \text{[S32]}$$

where the cutoff function H is such that $\lim_{x\to 0} H(x) = \text{const}$, $\lim_{x\to\infty} H(x)/x^{\max\{2-\delta-\gamma, 1-\delta\}} = 0$ and is such that $\int_1^\infty P(m|A)dm = 1$. This generalization of Eq. 7 does not affect our results for $A < A_c$. In fact, the joint probability distribution in such a generalized setting reads:

$$\begin{aligned} P(n,m|A) &= n^{-1} m^{-\delta} H\left(\frac{m}{M_0}\right) G\left(\frac{n}{\langle n|m,A\rangle}\right) \\ &= n^{-1} m^{-\delta} H\left(\frac{m}{M_0}\right) G\left[\frac{n m^{\gamma}}{A^{\Phi} h(m/A^{\lambda})}\right] \end{aligned} \qquad \text{[S33]}$$

and integrals of this distribution (e.g. marginals and moments) depend on which of the two finite-size cutoffs (h and H) sets in at the lowest value of m. We show this by calculating the

moment $I_{j,k}$ with $j,k>0$:

$$I_{j,k} = \int_1^\infty \int_0^\infty n^{j-1} m^{k-\delta} G\left[\frac{nm^\gamma}{A^\Phi h(m/A^\lambda)}\right] H\left(\frac{m}{M_0}\right) dn dm$$
$$= A^{j\Phi} \int_1^\infty m^{k-\delta-j\gamma} h^j\left(\frac{m}{A^\lambda}\right) H\left(\frac{m}{M_0}\right) dm \int_0^\infty x^{j-1} G(x) dx, \quad [S34]$$

where the properties of G ensure the convergence of the integral. The scaling of $I_{j,k}$ with A is thus determined by which of the two functions $h^j(m/A^\lambda)$ and $H(m/M_0)$ decays earlier in m. For intermediate values of A (i.e. $A^\lambda < M_0$) one can replace $H(m/M_0)$ with $(\delta-1)$ in equations Eq. S33 and Eq. S34, and is left with the framework described in the main text. For larger values of A (i.e. $A > M_0^{1/\lambda}$), instead, one replaces the term $h(m/A^\lambda)$ with the constant h_0^j in equations Eq. S33 and Eq. S34. In the limit of very large area, therefore, the maximum mass M_0 ceases to increase with A, due to attained physiological or ecological constraints. Furthermore, Damuth's law displays no cutoff in m, i.e. $\langle n|m,A\rangle \simeq h_0 m^{-\gamma} A^\Phi$ (more precisely, the cutoff in Damuth's law would be unobservable due to the extremely low probability to observe a species with $m > M_0$), and the ecosystem is effectively described by the joint probability distribution:

$$P(n,m|A) = n^{-1} m^{-\delta} H\left(\frac{m}{M_0}\right) G\left(\frac{n}{\langle n|m,A\rangle}\right)$$
$$= n^{-1} m^{-\delta} H\left(\frac{m}{M_0}\right) G\left(\frac{nm^\gamma}{h_0 A^\Phi}\right), \quad [S35]$$

and thus by the modified $P(n|m,A) = n^{-1} G[nm^\gamma/(h_0 A^\Phi)]$. One can show that in this limit of very large area the linking relationship (1) goes unchanged, whereas equations (6–9) are replaced by: $z = 1-\Phi$, $\mu = 1$, $\nu = 1$ (i.e. the maxima in equations 6–8 disappear) and $\xi = 0$. This could be the situation for the two forest datasets analyzed in section 2.2.2, where one should have $\lambda = 0$ to verify the reasonable assumption $\mu = \nu = 1$ (see section 2.2.2 for details).

1.8.4 Intra-specific size distributions

One assumption of our scaling framework is that all individuals of the same species have the same mass, although in reality individuals masses within the same species are distributed according to intra-specific size distributions. The scaling properties of intra-specific size distributions have been studied by Giometto *et al.* (9), where it was shown that protist species belonging to four different phyla and covering five orders of magnitude in mass are characterized by a universal size distribution:

$$p(m|\bar{m}) = \frac{1}{m} F\left(\frac{m}{\bar{m}}\right), \quad [S36]$$

where m is the mass of an individual, $\bar{m}$ is the characteristic mass of the species and $F(x) \to 0$ suitably fast for $x \to 0$ and $x \to \infty$, as detailed in Giometto *et al.* (9). When intra-specific size distributions are described by Eq. S36, the results of our scaling frameworks hold exactly. The moment (j,k) is computed as:

$$I_{j,k} = \int_0^\infty \int_1^\infty \int_1^\infty n^j m^k P(n,\bar{m}|A) p(m|\bar{m})\, dm\, d\bar{m}\, dn$$
$$= \int_0^\infty \int_1^\infty n^j P(n,\bar{m}|A) \langle m^k|\bar{m}\rangle\, d\bar{m}\, dn \quad [S37]$$

where :

$$\langle m^k|\bar{m}\rangle := \int_1^\infty m^k p(m|\bar{m}) dm. \quad [S38]$$

Usinq Eq. S36, we obtain:

$$\langle m^k|\bar{m}\rangle = \int_1^\infty m^{k-1} F\left(\frac{m}{\bar{m}}\right) dm = \bar{m}^k \int_{1/\bar{m}}^\infty x^{k-1} F(x) dx$$
$$\simeq \bar{m}^k \int_0^\infty x^{k-1} F(x) dx \propto \bar{m}^k \quad [S39]$$

where $x = m/\bar{m}$ and $\int_0^\infty x^{k-1} F(x) dx$ is a constant. Substituting this result in Eq. S37, we have $I_{j,k} = \int_0^\infty \int_1^\infty n^j \bar{m}^k P(n,\bar{m}|A) d\bar{m}\, dn$, which corresponds to Eq. S19. Thus, the moments $I_{j,k}$ computed in this generalized framework have the same scaling with the area as the ones computed by assuming that all individuals within a species have the same mass. Therefore, the linking relationships (2–5), whose derivation relies on the scaling of $I_{j,k}$ with A, are unchanged. Furthermore, we show that the linking relationship in Eq. 1 is also unchanged. In fact, assuming that intra-specific size distributions are given by Eq. S36, the size spectrum is given by:

$$s(m|A) = \frac{S}{N} \int_0^\infty n \int_1^\infty P(n,\bar{m}|A) m^{-1} F\left(\frac{m}{\bar{m}}\right) d\bar{m} dn$$
$$= \frac{S}{N} \int_0^\infty \int_1^\infty \bar{m}^{-\delta} G\left[\frac{n\bar{m}^\gamma}{A^\Phi h(\bar{m}/A^\lambda)}\right] m^{-1} F\left(\frac{m}{\bar{m}}\right) d\bar{m} dn$$
$$= m^{-\delta} \frac{S}{N} \int_0^\infty \int_{1/m}^\infty x^{-\delta} G\left[\frac{n(xm)^\gamma}{A^\Phi h(xm/A^\lambda)}\right] F\left(\frac{1}{x}\right) dx dn$$
$$= m^{-\delta-\gamma} A^\Phi \frac{S}{N} \int_{1/m}^\infty x^{-\delta-\gamma} h\left(\frac{xm}{A^\lambda}\right) F\left(\frac{1}{x}\right) dx \int_0^\infty G(y) dy, \quad [S40]$$

where $x = \bar{m}/m$, $y = [n(xm)^\gamma] / \left[A^\Phi h(xm/A^\lambda)\right]$ and $\int_0^\infty G(y) dy$ is a constant. We note that we cannot compute analytically the scaling of $s(m|A)$ with m from Eq. S40, because m appears both at the lower limit of the integral in x and in the argument of h. However, for large m, the lower limit of the integral in x tends to 0 and thus $s(m|A) \propto m^{-\delta-\gamma} \tilde{h}(m/A^\lambda)$ in the limit of large m, where $\tilde{h}(y) = \int_0^\infty x^{-\delta-\gamma} F(1/x) h(xy) dx$ has the same limiting behavior of $h(y)$ at $y \to 0$ and $y \to \infty$[†], and thus the linking relationship $\eta = \delta + \gamma$ (Eq. 1) still holds. Thus, introducing intra-specific variability in mass according to Eq. S36 does not alter the linking relationships (1, 6–9).

1.8.5 Tree intraspecific size distributions

As we discuss in section 2.2.1, trees are an exception to Eq. S36, given that a single species can cover several orders of magnitude in mass. We show in section 2.2.1 that the intra-specific size distributions of the most abundant tree species in tropical forests are characterized by the finite-size scaling form:

$$p(m|\bar{m}) = m^{-\Delta} \mathcal{F}\left(\frac{m}{\bar{m}^\Omega}\right), \quad [S41]$$

[†] Because $\lim_{x\to\infty} h(x) = 0$, $\lim_{x\to 0} h(x) = h_0$, $h(x) \leq h_0$ and $\int_0^\infty x^{-\delta-\gamma} F(1/x) dx < \infty$ if $\delta + \gamma > 1$, then $\lim_{y\to\infty} \tilde{h}(y) = 0$ and $\lim_{y\to 0} \tilde{h}(y) = \text{const}$ follows from the Lebesgue's dominated convergence theorem.

where $\Omega = 1/(2-\Delta)$ ensures that $\int mp(m|\bar{m})dm = \bar{m}$, $\Delta = 1.12 \pm 0.06$ and $\mathcal{F}(x)$ is a scaling function with limiting behaviors $\mathcal{F}(x) \to$ const for $x \to 0$ and $\mathcal{F}(x) \to 0$ more rapidly than any power of x for $x \to \infty$. The (j,k)th moment $I_{j,k}$ can still be computed exactly and some small corrections to the scaling exponents arise, compared to Eq. S19. In fact, using Eq. S41 we obtain:

$$\begin{aligned}\langle m^k|\bar{m}\rangle &= \int_1^\infty m^{k-\Delta}\mathcal{F}\left(\frac{m}{\bar{m}^\Omega}\right)dm \\ &= \bar{m}^{\Omega(1+k-\Delta)}\int_{1/\bar{m}^\Omega}^\infty x^{k-\Delta}\mathcal{F}(x)dx \\ &\propto \bar{m}^{\max\{0,\Omega(1+k-\Delta)\}}.\end{aligned} \quad [S42]$$

Note that, among the consequences of this result, we obtain that the variance of the intra-specific size distribution increases with the average body size: $\langle m^2|\bar{m}\rangle \propto \bar{m}^{\Omega(3-\Delta)}$. Substituting Eq. S42 in Eq. S37 gives:

$$\begin{aligned}I_{j,k} &= \int_0^\infty\int_1^\infty n^j\bar{m}^{\max\{0,\Omega(1+k-\Delta)\}}P(n,\bar{m}|A)d\bar{m}\ dn \\ &\propto A^{j\Phi+\max\{0,\lambda(1-\delta-j\gamma+\max\{0,\Omega(1+k-\Delta)\})\}},\end{aligned} \quad [S43]$$

where the calculations have been performed as in section 1.6. Eq. S43 differs only slightly from Eq. S19. As a result, a small correction applies to the linking relationship (2): the term α in Eq. 2 is substituted by $\frac{1+\alpha-\Delta}{2-\Delta}$, which for the empirical value $\Delta = 1.12 \pm 0.06$ results in a very small correction $C = \alpha - \frac{1+\alpha-\Delta}{2-\Delta} = (\alpha-1)\left(\frac{1-\Delta}{2-\Delta}\right) \simeq 0$ which is compatible with zero, given that Δ is compatible with one. Equations (3–5) are unchanged. The size spectrum is given by:

$$\begin{aligned}s(m|A) &= \frac{S}{N}\int_0^\infty n\int_1^\infty P(n,\bar{m}|A)m^{-\Delta}\mathcal{F}\left(\frac{m}{\bar{m}^\Omega}\right)d\bar{m}dn \\ &= \frac{S}{N}\frac{1}{m^\Delta}\int_1^\infty\frac{1}{\bar{m}^\delta}\mathcal{F}\left(\frac{m}{\bar{m}^\Omega}\right)\int_0^\infty G\left[\frac{n\bar{m}^\gamma}{A^\Phi h(\bar{m}/A^\lambda)}\right]dn\ d\bar{m} \\ &= \frac{S}{N}m^{-\Delta+(1-\delta)/\Omega}\int_0^m x^{(\delta-1)/\Omega-1}\mathcal{F}(x) \\ &\quad\times\int_0^\infty G\left[\frac{n(m/x)^{\gamma/\Omega}}{A^\Phi h\left[(m/x)^{1/\Omega}/A^\lambda\right]}\right]dn\ dx \\ &= \frac{S}{N}A^\Phi m^{-\Delta+(1-\delta-\gamma)(2-\Delta)}\int_0^\infty G(y)dy \\ &\quad\times\int_0^m x^{(2-\Delta)(\delta+\gamma-1)-1}\mathcal{F}(x)h\left[\frac{(m/x)^{1/\Omega}}{A^\lambda}\right]dx\end{aligned} \quad [S44]$$

where $x = m/\bar{m}^\Omega$, $\Omega = 1/(2-\Delta)$, $y = \left[n(m/x)^{1/\Omega}\right]/\left\{A^\Phi h\left[(m/x)^{1/\Omega}/A^\lambda\right]\right\}$ and $\int_0^\infty G(y)dy$ is a constant. Note that we cannot compute the scaling exponent of $s(m|A)$ with m exactly, because x and m are found in the arguments of both $\mathcal{F}$ and h. We can, however, derive an approximation that holds for large m. This is most easily seen if we assume that no finite-size effect is found in Damuth's law, i.e. $h =$ const, for which we have:

$$s(m|A) \propto \frac{S}{N}A^\Phi m^{(2-\Delta)(1-\delta-\gamma)-\Delta}\int_0^m x^{-(2-\Delta)(1-\delta-\gamma)-1}\mathcal{F}(x)dx, \quad [S45]$$

which, for large m, is a power-law ($s(m|A) \propto m^{-\eta}$) with:

$$\eta = \Delta - (2-\Delta)(1-\delta-\gamma), \quad [S46]$$

which generalizes Eq. 1 to the case of intra-specific size distributions as in Eq. S41. We note, in analogy to the calculations performed in the previous section, that for large m the upper limit of the integral in x in Eq. S44 tends to $+\infty$ and thus $s(m) \propto m^{(2-\Delta)(1-\delta-\gamma)-\Delta}\tilde{h}\left(m/A^{\lambda\Omega}\right)$ in the limit of large m, where the function $\tilde{h}(y) = \int_0^\infty x^{-(-2-\Delta)(1-\delta-\gamma)}\mathcal{F}(x)h\left[(y/x)^{1/\Omega}\right]dx$ has the same limiting behavior of $h(y)$ for $y \to 0$ and $y \to \infty$[‡]. Therefore, the linking relationship in Eq. S46 holds also when h is not identically constant. The correction to the linking relationship (1) is small because the difference between $\eta_1 = \delta + \gamma$ computed via Eq. 1 and $\eta_2 = \Delta - (2-\Delta)(1-\delta-\gamma)$ computed via Eq. S46 is $\eta_1 - \eta_2 = (\Delta-1)(\delta+\gamma-2)$ and $\Delta = 1.12 \pm 0.06$ is compatible with one, in which case equations (1) and Eq. S41 coincide. Using the values reported in Table S3, we find $\eta_1 - \eta_2 \simeq -0.07 \pm 0.03$, which is compatible with zero. Because the differences between the linking relationships (1) and (6) reported in the main text and the generalizations reported in this section are negligible and compatible with zero, we will neglect such generalizations in the rest of this study.

1.8.6 Alternative forms of $P(m|A)$

Our *ansatz* for $P(m|A)$ is a pure power function, Eq. S8. There are two possible relaxation of this hypothesis that do not compromise analytical tractability, in addition to the one already explored in section 1.8.3. The first is the addition of a cut-off at large masses, $P(m|A) = m^{-\delta}F_1\left(\frac{m}{A^\lambda}\right)$ with $\delta > 1$ (for normalization purposes), where the exponent λ is the same as the one of the cut-off in Damuth's law (Eq. S10) and where $F_1(x)$ is such that $\lim_{x\to 0}F_1(x) =$ const and $lim_{x\to\infty}F_1(x) = 0$. In this case, the computation of the moments $I_{j,k}$ is performed similarly to Eq. S34, where the role of $h^j\left(\frac{m}{A^\lambda}\right)$ is now played by the product $F_1\left(\frac{m}{A^\lambda}\right)h^j\left(\frac{m}{A^\lambda}\right)$, behaving similarly. The final result is the same as in Eq. S34. In the computation of the community size-spectrum, Eq. 16 of the main text, the cut-off is now given by the product of the two functions F_1 and h. In conclusion, this generalization does not change the linking relationships in Eq. 1, 6–8 of the main text. The generalization of Eq. 9 of main text to this case cannot be computed analytically. Some empirical studies (10, 11) found a $P(m|A)$ similar to a log-normal. We can describe this case as $P(m|A) = m^{-\delta}F_2\left(\frac{m}{A^\lambda}\right)$ where $F_2(x)$ is such that $\lim_{x\to 0,\infty}x^jF_1(x) = 0\ \forall j$. Normalization requires $\delta = 1$. Moments can be computed, again, as in Eq. S34 but noticing that in this case $\int_{1/A^\lambda}^\infty x^{k-\delta-\gamma j}h^j(x)F(x) \approx const$ in the limit of large A, as the expression inside the integral tends to 0 faster than any power of x when $x \to 0$. Therefore, we obtain $I_{j,k} \propto A^{\Phi j+\lambda(k-\lambda j)}$. Eq. 6–8 of the main text are then replaced by: $z = 1-\Phi-\lambda(\alpha-\gamma)$, $\nu = 1-\lambda\alpha$ and $\mu = 1+\lambda(1-\alpha)$. The community size-spectrum is computed

[‡] Because $\lim_{x\to\infty}h(x) = 0$, $\lim_{x\to 0}h(x) = h_0$, $h(x) \leq h_0$ and $\int_0^\infty x^{-\delta-\gamma}F(1/x)dx < \infty$ if $\delta+\gamma > 1$, then $\lim_{y\to\infty}\tilde{h}(y) = 0$ and $\lim_{y\to 0}\tilde{h}(y) =$ const follows from the Lebesgue's dominated convergence theorem.

as:

$$\begin{aligned} s(m|A) &= \frac{S}{N}\int_0^\infty nP(n,m|A)dn \\ &= \frac{S}{N}m^{-1}F_2\Big(\frac{m}{A^\lambda}\Big)\int_0^\infty dnG\Big(\frac{nm^\gamma}{A^\Phi h(m/A^\lambda)}\Big) \\ &\propto m^{-1}\frac{A^{\lambda\,\gamma}}{m}\,F_2\Big(\frac{m}{A^\lambda}\Big)h\Big(\frac{m}{A^\lambda}\Big) \\ &= m^{-1}H\Big(\frac{m}{A^\lambda}\Big) \end{aligned} \qquad [S47]$$

where $H(x) := x^{-\gamma}F_2(x)h(x)$ is such that $\lim_{x\to 0,\infty} x^j H(x) = 0\ \forall j$. Therefore, $s(m|A)$ is not a power-law, but has an internal mode, similarly to $P(m|A)$. The generalization of Eq. 9 of main text to this case cannot be computed analytically.

1.9 Compatibility with previous works. Southwood *et al.* (12) derived a linking relationship that is equivalent to our Eq. 5. Here, we briefly review their result and illustrate how it must be corrected due to a few miscalculations.

The authors start from the observation (13) that the total number of species of length class L scale as $S_L(L) \propto SL^{-\Delta}$, where $\Delta = 3/2$ and S is the total number of species in the ecosystem. The estimate $\Delta = 3/2$ made in that study (12) is incorrect, because the correction needed to account for logarithmic binning (14, 15) is missing. The correct estimate that accounts for logarithmic binning (14) is $\Delta = 5/2$. First, the authors of reference 9 say that the maximum length is obtained by imposing $S_L(L_{\max}) = SL_{\max}^{-\Delta} = 1$, i.e. $S \propto L_{\max}^{\Delta}$. This calculation is also incorrect, because it makes an improper use of probability distributions. In fact, $S_L(L)$ is the fraction of species with length in $[L, L+dL]$ and the maximum species' length is found by imposing that the probability of finding a species with length larger than $L_{\max}$ be equal to $1/S$ (this is a widely employed argument to estimate the largest random number drawn from a specified distribution (16, 17)) which results in $S \propto L_{\max}^{\Delta-1}$. Incidentally, these two miscalculations compensate each other and together lead to the estimate $S \propto L_{\max}^{3/2}$. However, and this is crucial to derive the correct linking relationship, the correct equation is $S \propto L_{\max}^{\Delta-1}$ and not the one suggested in reference 9, i.e. $S \propto L_{\max}^{\Delta}$. The equation relating S and L converts into a species-mass relationship via the scaling $L \propto m^{1/3}$, where we have assumed that body density does not scale with body mass. Specifically, one finds $S \propto m_{\max}^{\delta-1}$ with $\delta = (\Delta+2)/3$. Then, one can use the finding by Burness *et al.* (18), $m_{\max} \propto A^{\xi}$, to derive the linking relationship between z, ξ and δ. In fact, by comparing $S \propto m_{\max}^{\delta-1} \propto A^{\xi(\delta-1)}$ with the SAR $S \propto A^z$ one obtains the linking relationship $z = \xi(\delta-1)$, which coincides with our Eq. 5. However, we note that, in reference 9, z was meant as the exponent of the nested species-area relationship, differently from here. Nonetheless, the linking relationship is the same.

Our results thus agree with the earlier result by Southwood *et al.* (12), where one of the linking relationship was discovered. Our investigation reveals that such linking relationship is only one component of the broader set of linking relationships, Eqs. 1,6–9 of the main text.

We finally note the differences of the current work with a previous scaling framework (19), namely: 1) the enforcement and the implications of resource limitation, 2) the validation based on a broad class of community dynamics models, 3) the extensive empirical verification, and 4) the richer set of macroecological laws that the scaling framework accounts for, most importantly Damuth's law, which allows us to reconcile the predicted linkages with empirical data and community dynamics models (see, e.g. Fig. S2). Such framework (19) took as starting point an equation compatible with our assumption on $P(n,m|A)$ (Eq. 2 of the main text). Specifically, Eq. 1 in Banavar *et al.* (19) is more general than our Eq. 6, but one needs to specify further properties of $P(n,m|A)$ in order to recover Damuth's law, as we did here. Furthermore, the scaling properties hypothesized for the function F that appears in Eq. 1 of Banavar *et al.* (19) (hypothesis $\mathcal{H}_1$ therein) is incompatible with Damuth's law and leads to different predictions for the pattern covariations (e.g. the linking relationship $\eta = \delta$) that are falsified by empirical data and by our model simulations (Fig. S2). Note that the introduction of a constraint on the total community consumption rate in the framework of Banavar *et al.* would not affect the relationship $\eta = \delta$, which is instead a byproduct of the assumptions on $P(n,m|A)$.

2 Compatibility with macroecological data

2.1 Empirical evidence of scaling ecological laws. Abundant empirical evidence exists of scaling ecological patterns in diverse types of ecosystems: forests, terrestrial (including mammals in particular) and aquatic ecosystems. A comparison of these empirical results highlights the non-universality of the values of scaling exponents. Fig. 1 in the main text shows evidences of Kleiber's law (panel a), Damuth's law (panel b), the Species-Area Relationship (SAR) (panel c) and the community size-spectrum $s(m)$ (panel d) for the three types of ecosystems. Regression lines in Fig. 1 are fits provided in the original papers (see legends), except for the patterns for forests and terrestrial ecosystems in panel b, which were fitted by linear least-squares fits on log-transformed data, and for the community size-spectra from BCI and Niwot Ridge datasets in panel d, which were fitted with maximum likelihood (20). Table S1 reports the estimates for the scaling exponents. Table S2 contains a compilation of references to empirical measures of the ecological patterns referred to in the main text.

2.2 Compatibility of linking relationships and data. Our framework predicts five relationships linking the scaling exponents of ecological laws (Eqs. 1, 6–9 of the main text):

$$\eta = \gamma + \delta, \qquad [S48a]$$

$$z = 1 - \Phi - \max\{0, \lambda(1+\alpha-\eta)\}, \qquad [S48b]$$

$$\mu = 1 + \max\{0, \lambda(2-\eta)\} - \max\{0, \lambda(1+\alpha-\eta)\}, \qquad [S48c]$$

$$\nu = 1 - \max\{0, \lambda(1+\alpha-\eta)\}, \qquad [S48d]$$

$$\xi = \frac{z}{\delta - 1}. \qquad [S48e]$$

Eqs. S48b and S48d also imply

$$z + \Phi = \nu. \qquad [S49]$$

Notice that there are only 5 independent exponents (e.g. γ, δ, Φ, α and λ), whereas the observable laws amount to 10: Kleiber's law and Eqs. 3, 12-18 of the main text. Note that Eq. 13 contains three laws because it describes the scaling of the average abundance of a species with body mass and with the area of the ecosystem and the scaling of its cut-off with the area of the ecosystem. Figure S3 summarizes the predictions on the values or bounds of scaling exponents

based on the linking relationships (S48a-e), for different possible values of the independent exponents. As each of the exponents appearing in Eqs. (S48a-e) could have different values in different ecosystems and in different environmental conditions, in order to verify the validity of Eqs. (S48a-e) each relationship must be verified on a single dataset, where all the exponents have been measured simultaneously.

2.2.1 Equation S48a (Eq. 1 of the main text)

We verified Eq. S48a on censuses of Barro Colorado Island (BCI) (21) (Fig. 2) and the Luquillo forest (22) (Fig. S4). These datasets report the trunk diameter and the species' identity of every tree having a diameter at breast height (dbh) >10 mm contained in a plot of 50 ha within the BCI forest (Panama), and a plot of 16 ha within the Luquillo forest (Puerto Rico). Diameters were converted into mass using an established allometric relationship between mass and diameter (23, 24), $m = 0.124d^{8/3}$ kg, with d expressed in cm. To compute $P(\bar{m}|A)$, where $\bar{m}$ is the typical species' mass, we used the mean mass of the species' individuals as our estimate of $\bar{m}$. To obtain estimates of δ and η the probability distributions $P(\bar{m}|A)$ and $s(m|A)$ were fitted via maximum likelihood to the functional forms $P(\bar{m}|A) = a_1\bar{m}^{-\delta}e^{-b_1\bar{m}}$ and $s(m|A) = a_2 m^{-\eta}e^{-b_2 m}$, where a_1 and a_2 are the normalization constants (which can be expressed in terms of b_1 and b_2), b_1 and b_2 are constants that accounts for possible finite-size effects, and δ and η are the power-law exponents[§]. Section 1.8.3 justifies the possible presence of an upper cutoff in $P(\bar{m}|A)$. To account for deviations from the power-law behavior at low values of m or $\bar{m}$ (these may arise for various reasons, like e.g. sampling protocols affecting the estimates of mean masses and mean abundances at small masses, as described in the next paragraph) we performed the maximum-likelihood estimation of δ and η by considering only the data with $m > m_k$ at various values of $m_k = a^k$ 0.124 kg in the range $(0.124 - 10^2)$ kg (with $1 < a < 2$ and k integer). Note that 0.124 kg is the mass of a tree with dbh=10 mm, i.e. the lower limit of the sampling protocol. If the data were distributed according to a pure power-law with no finite-size effects, such procedure would return approximately the same value of the exponent for any m_k. If the data were distributed according to a power-law with finite-size effects at small and large values of m, instead, one would observe an approximately constant estimate of the exponent at intermediate m_k and deviations from such estimate at small and large values of m_k (see e.g. Fig. S5). For each fit, we identified the extent of the power-law regime and our estimate of the exponent and the associated error are, respectively, the mean and standard deviation of the maximum-likelihood exponent at different values of $\bar{m}_k$ in the power-law regime.

The estimate of Damuth's law exponent γ, describing the decay of mean abundances with species' typical masses $\bar{m}$, requires a correction for a bias introduced by the sampling protocol on the estimates of mean abundances and mean masses at small values of $\bar{m}$. In fact, the sampling protocol in tropical forests censuses instructs to sample only the trees with dbh larger than 10 mm. The measured abundances of small species (i.e. those with typical diameter close to 10 mm) are therefore lower than the true ones because individuals with diameter $d \leq 10$ mm were not censused. As a result, the average abundance as a function of a typical species' mass initially increases with $\bar{m}$ and is followed by the decreasing power-law regime where the effect of the sampling protocol becomes unimportant (fig. 2A). The initial increase is a sampling artifact. In fact, we verified that this is the case by creating an artificial forest dataset where species' mean abundances follow Damuth's law exactly. Within such artificial forest, we distributed species mean masses $\bar{m}$ according to the power-law $p(\bar{m}) = (\delta - 1)\bar{m}^{-\delta}$. We drew the abundance of each species from a Poisson distribution with mean $\bar{m}^{-\gamma}$. Finally, we needed to assign a mass to each individual of each species. To do so, we characterized the intra-specific mass distributions of tropical trees $p(m|\bar{m})$, i.e. the probability that an individual has mass in $(m, m+dm)$ given that it belongs to a species with mean mass $\bar{m}$. We computed intra-specific mass distributions in the BCI and Luquillo forests (Fig. S6a) and found, looking at the species with more than 1500 (BCI) and 400 (Luquillo) individuals, that most species have intraspecific size distributions characterized by the finite-size scaling form:

$$p(m|\bar{m}) = m^{-\Delta}\mathcal{F}\left(\frac{m}{\bar{m}^{\Omega}}\right), \qquad [S50]$$

where $\Omega = 1/(2-\Delta)$ ensures that $\int mp(m|\bar{m})dm = \bar{m}$, $\Delta = 1.12 \pm 0.06$[¶] and $\mathcal{F}(x)$ is a scaling function with limiting behaviors $\mathcal{F}(x) \to$ const for $x \to 0$ and $\mathcal{F}(x) \to 0$ more rapidly than any power of x for $x \to \infty$. A similar result was found for unicellular protists (9). Furthermore, via data collapse (i.e. plotting $m^{\Delta}p(m|\bar{m})$ vs $m/\bar{m}^{\psi}$) we found that $\mathcal{F}(x) = q_0 e^{-q_1 x}$ provides a good fit to the data, where $q_0 = 0.17$ and $q_1 = 0.21$ are constants. Note that q_0 is not a parameter of the fit, as it is fixed by normalization. Having characterized the scaling form of intraspecific distributions, we could then randomly sample from such distributions the masses of individuals belonging to each species in our artificial forest. We then mimicked the sampling protocol by eliminating all individuals with mass lower than 0.124 kg (corresponding to a dbh of 10 mm) and computed Damuth's law in such a filtered dataset. Fig. S7c shows that, despite the fact that mean species abundances in the artificial forest follow Damuth's law exactly, the sampling protocol causes the emergence of a new regime at small $\bar{m}$ where the relationship between $\langle n|\bar{m}\rangle$ and $\bar{m}$ is monotonically increasing. This demonstrates that the sampling protocol introduces an artificial deviation from the power-law regime which has to be considered with care while interpreting empirical data. The sampling artifact can be corrected as follows. To derive our estimate for γ and the associated error, we binned the typical species masses logarithmically and computed the mean abundance of all species within each bin. Then, we varied the number of bins n_{bin} and computed the Damuth's law exponent $\gamma_{n_{bin}}$ via least-squares fitting of log-transformed data, weighted by the standard deviation of abundances within each bin. Our estimate for γ is the mean $\gamma = \langle\gamma_{n_{bin}}\rangle$ across several values of n_{bin}. To correct for the bias caused by the sampling protocol, we repeated such computation by considering only the species with mean mass $\bar{m} > \bar{m}_k$, with $\bar{m}_k = a^k$ 0.124 kg

[§] Specifically, one has: $a_1 = b_1^{1-\delta}/\Gamma(1-\delta, \bar{m}_0\delta)$, where Γ is the incomplete Gamma function. We maximized the log-likelihood: $\log L(\delta, b_1) = S\log\left[b_1^{1-\delta}/\Gamma(1-\delta, \bar{m}_0\delta)\right] - \delta\sum_{i=1}^{S}\log\bar{m}_i - b_1\sum_{i=1}^{S}\bar{m}_i$ with respect to δ and b_1, where $\bar{m}_0$ is the minimum species mass, S is the total number of species and i is an index that identifies the species. The maximum likelihood fitting of $s(m|A)$ was performed analogously.

[¶] The exponent estimate is computed using the method described in Bhattacharjee and Seno (25). The error is computed as the value of the exponent at which the error functional P_b defined in Bhattacharjee and Seno (25) is 1% larger than its value at the minimum.

in the range $(0.124 - 10^2)$ kg (with $1 < a < 2$ and k integer). If mean abundances followed Damuth's law exactly (in the absence of sampling bias and with sufficient statistics), such procedure would return the same value of γ for any $\bar{m}_k$. If finite-size effects were present at small and large values of $\bar{m}$, instead, one would observe an approximately constant estimate of the exponent at intermediate $\bar{m}_k$ and deviations from such estimate at small and large values of $\bar{m}_k$ (see e.g. Fig. S7b). At large values of $\bar{m}_k$ a finite-size effect may also be induced by low statistics. For each fit, we identified the extent of the power-law regime and our estimate of γ and the associated error are, respectively, the mean and standard deviation of the exponents estimated at the different values of $\bar{m}_k$ in the power-law regime.

In our analysis, we used the fifth, sixth and seventh censuses of BCI and the five censuses of the Luquillo forest available online in the Center for Tropical Forest Science (CTFS) dataset collection. All these censuses satisfy the linking relationship Eq. S48a within the errors. Whereas BCI censuses appear very similar to each other (and therefore also the exponent values estimated in different censuses, see Table S3), the Luquillo forest appears to be more dynamic (we note that the forest was hit by a major hurricane between censuses 2 and 3), with the values of γ decreasing in time after 1998 (census 2, see Table S4). Because the estimate of δ remains constant suggesting that climatic, ecological or anthropogenic dynamics affected only species' abundances in this forest, our framework would predict via eq. Eq. S48a that η would also decrease in time, and in fact this is also found in the data, with eq. Eq. S48a being verified in all censuses. We note that both the BCI and the Luquillo datasets reject the linking relationship $\eta = \delta$ predicted in a previous theoretical work (19).

Although we do not have an estimate of the exponents μ and ν for tropical forests, a reasonable assumption is $\mu = 1 = \nu$. Given that $\eta < 2$ for the BCI and the Luquillo datasets, such assumption is only verified if $\lambda = 0$, i.e. if the maximum body size does not scale with the ecosystem area. An analysis of this situation is provided in section 1.8.3.

2.2.2 Equations Eq. S48b and Eq. S48d in other datasets (Eqs. 6, 8 of the main text)

To test the validity of Eq. S48b we used a dataset gathering population densities of several species of lizards on 64 islands worldwide (LIZ) (26), with areas ranging from 10^{-1} to 10^5 km^2. In this dataset, we fitted the SAR with linear least-squares regression on log-transformed data and $P(\bar{m}|A)$ via maximum-likelihood (20) (Fig. S8a and b), where $\bar{m}$ are species' mean masses. The exponent $\Phi = 0.78$ (describing the dependence of $\langle n|\bar{m}, A\rangle$ on A) was obtained by maximizing the coefficient of determination R^2 of the linear least-squares regression of the pairs $\left(\log \bar{m}, \log \frac{n}{A^\Phi}\right)$ obtained by varying Φ in the interval [0,2] (Fig. S8c and d). The estimate of γ is obtained from the pairs $\left(\log \bar{m}, \log \frac{n}{A^\Phi}\right)$ computed with the optimal value of Φ with the same methods used for forests, with $\bar{m}_k = 2c^k 10^{-1}$ kg in the range $(10^{-1} - 10^2)$ kg (with $1 < c < 2$ and k integer). Note that this estimate of γ is different from the one given in Table S1, which was obtained by plotting densities versus typical masses (equivalent to taking $\Phi = 1$) in order to allow comparison with other data from the literature. The estimates for the scaling exponents in this dataset are reported in Table S5. Because in this dataset $\eta = \delta + \gamma$ is compatible with 2, Eqs. S48c and S48d imply $\mu = \nu$. Furthermore, because in general $\alpha \leq 1$, one has $\max\{0, \lambda(1+\alpha-\eta)\} = 0$ and therefore our framework predicts:

$$\nu = \mu = 1. \qquad [S51]$$

Eq. S48b (or, equivalently, Eq. S49) thus implies that in order to have $z > 0$, as found in the dataset analyzed here, Φ needs to be smaller than one. Since Φ describes the scaling of $\langle n|\bar{m}, A\rangle$ with A, our framework predicts that species' densities should decrease with increasing ecosystem area. This is indeed found in LIZ and the values of z and $1 - \Phi$ are compatible within the errors (see Table S5). Equivalently, using Eq. S49 and our estimate of z we find $\nu = 0.95 \pm 0.08$, which is compatible with the prediction (Eq. S51) $\nu = 1$.

2.2.3 Equation S48e (Eq. 9 of the main text)

To test the validity of Eq. S48e we used a dataset of mammals species presence/absence data on several islands in Sunda Shelf (SSI) (27), covering a wide range of island areas (10^1 to 10^6 km^2). The SAR and the scaling of the maximum body mass with the area were fitted with linear least-squares regression on log-transformed data, while $P(\bar{m}|A)$ was fitted with maximum-likelihood (20). Scaling exponents in this dataset are reported in Table S6. We find that eq. Eq. S48e is verified in the SSI dataset within one standard error, as $z = 0.23 \pm 0.02$ and $\xi(\delta - 1) = 0.29 \pm 0.1$.

3 Mathematical community dynamics models

3.1 Fixed number of species.

3.1.1 Basic model: exploration of parameters' space

Our basic model for the community dynamics of an ecosystem depends on a number of parameters (see Methods). As explained in the Methods section, a thorough exploration of the parameters' space is computationally unfeasible. Nonetheless, we verified that varying the values of the parameters that are most meaningful for the dynamics (i.e. Kleiber's law exponent α, the speciation rate w, the SAR exponent z and the exponent θ that describes the scaling of vital rates with body mass), the scaling characterization of the stationary state always holds and the linking relationships in Eqs. 1, 6–9 of the main text are always satisfied. Starting from the set of parameters $w = 10^{-3}$, $z = 1/4$, $\alpha = 3/4$ and $\theta = 1/4$ (parameters used to generate Fig. 3 of the main text) we varied one or two parameters at a time, keeping the other ones fixed. The parameters v_0 and c which appear in Eq. 20 of the main text were fixed to $v_0 = 1/2$ and $c = 10^{-5}$. Figs. (S9–S15) show the ecological patterns computed at stationarity for each set of parameters. Table S7 reports the estimates of the scaling exponents obtained for each set of parameters. All estimates were obtained as explained in the Methods section, unless otherwise stated. For each set of parameters, the relationships in Eqs. 1, 6–9 of the main text are satisfied within errors, the data collapses predicted by our scaling framework hold and the density scatter-plot of η versus $\delta + \gamma$ estimated at each time-step (Figs. S9–S15, panel e) is peaked along the 1:1 line, implying that the linking relationship (1) is satisfied, on average, at all times.

3.1.2 Variation on the speciation dynamics

In order to investigate the sensitivity of our results (i.e. the compatibility of the dynamic model with the scaling framework,

Eqs. 6–9 of the main text) to changes in the dynamic model assumptions, we investigated a variation of our basic model in which the species that goes extinct at each speciation event (to maintain S constant) is chosen randomly with a weight inversely proportional to its abundance, but independent of its mass. We ran this model with the same parameter values reported in the Methods section for the basic model and found that such modified model is compatible with our scaling framework, and thus with the predicted pattern covariations, which are verified within the errors. Table S7 reports the corresponding exponents values and Fig. S16 displays the macroecological patterns in this model.

3.2 Fluctuating number of species. In order to further investigate the sensitivity of our results to changes in the dynamic model assumptions, we relaxed the constraint of a fixed number of species S (as in the basic model) and we let $\langle S \rangle = A^z$ be an emergent property of the stochastic model. We achieved this by maintaining the ecological dynamics of births and deaths as in the basic model and by modifying the speciation dynamics in two different ways:

a. At each time step, the number of species that undergo speciation is drawn from a Poisson distribution with rate w. The species that undergo speciation are selected randomly. At speciation, a random number $n_i' < n_i$ of individuals of species i maintains the original mass m_i, whereas the remaining individuals are assigned to a new species j with mass $m_j = qm_i$, where q is drawn from a lognormal distribution with constant mean and variance. To avoid instability (i.e. extreme fluctuations that lead to the extinction of the community), we impose that the sum of the consumption rates of species i and j after speciation is equal to the consumption rate of species i before speciation. This is done by setting the abundance of species j such that $n_i' m_i^\alpha + n_j m_j^\alpha = n_i m_i^\alpha$, i.e. $n_j = (n_i - n_i')m_i^\alpha / m_j^\alpha$.

b. Same as in model a, but the species that undergo speciation are selected randomly with a weight proportional to their abundance, so that more abundant species are more likely to speciate.

We found that these models give rise to the empirically observed set of macroecological laws described in the main text, and of course the exponents of such laws depend both on the model specifications and on the model parameters. Most importantly, despite the differences in the speciation dynamics, we found that these models are also compatible with our scaling framework (Eqs. 6–9 of the main text), which specifies the scaling properties of the joint distribution $P(n, m|A)$. Thereby, macroecological patterns in these models comply with our predicted pattern covariations. Tables S9 and S10 report the exponents values measured in these models and Figs. S17–S20 display the corresponding macroecological patterns. Parameter values used to run the models are reported in the figures captions.

3.3 Value of η in the community dynamic models. The size spectrum exponent η in natural ecosystems typically assumes values $\eta \in (1, 2]$ (see Tables S1, S3 and S4), although values of $\eta > 2$ can also be found in marine environments (Table S1). All our community dynamics models yield values of η that are on average larger than 2 (the average is performed over time, see Tables S7, S8 and S9), although panels e in Figs. S9–S20 show that η can assume values smaller than 2 at any fixed time point (i.e. in snapshots of the ecosystem). Unfortunately, a suitably broad exploration of the parameters space in our models is computationally unfeasible, as the estimation of scaling exponents requires several hours of computation in a high-performance computer in order to properly estimate the tails of the distribution $P(n, m|A)$. However, based on our exploration of parameters' space, $\langle \eta \rangle = 2$ does seem to be a lower limit in our community dynamics models. In the attempt to find parameter sets that may allow for $\langle \eta \rangle < 2$, we found that increasing the mean $\bar{q}$ of the multiplicative factor q that specifies the mass of the descendant species at a speciation event (this may be seen as an implementation of Cope's rule (28), which postulates that descendant lineages tend to increase in body size) causes a reduction of the mean size spectrum exponent $\langle \eta \rangle$ (Figs. S19 and S20, and Table S10). Nonetheless, parameter sets that yield $\langle \eta \rangle < 2$ lead to communities that are very unstable and that rapidly go towards extinction. It thus appears that our community dynamics models are missing processes that would allow multiple species to coexist at a stable equilibrium with $\langle \eta \rangle < 2$. We speculate that one reason for this behavior may be the fact that our models assume a well-mixed system, unlike terrestrial ecosystems such as forests. In this sense, it may not be coincidental that values of $\eta > 2$ are typically found in aquatic ecosystems (Table S1) rather than terrestrial ones (Tables S1, S3 and S4). Further research will be dedicated to the investigation of macroecological linkages in metacommunities (29–32), with explicit focus on the implications of spatial structure and connectivity on scaling exponents values and linkages.

3.4 Specificity and universality. Our investigation of dynamic birth, death and speciation models corroborates the generality of our scaling framework and the predicted pattern covariations. In fact, we found that all the models investigated that are compatible with the empirically observed macroecological patterns described in the main text are all characterized by the same scaling properties of $P(n, m|A)$, which are encapsulated in our scaling framework (Eqs. 2–5 of the main text) and univocally specify the pattern covariations in Eqs. 1, 6–9 of the main text. Therefore, Eqs. 2–5 of the main text do not rely on specific assumptions about the population and speciation dynamics of a community, but rather specify the universal scaling properties that possibly any dynamic model compatible with the empirically observed macroecological laws must satisfy. Furthermore, the pattern covariations predicted by our scaling framework agree with empirical evidence (section 2) and with heuristic arguments, i.e. the many back-of-the-envelope calculations reported in the main text. It must be understood, however, that the scaling framework does not predict the values of scaling exponents, but rather their covariations.

The various community dynamic models studied here, instead, do predict scaling exponents values, which emerge from the rates and assumptions concerning birth, death and speciation events. However, there may exist several dynamic models capable of reproducing quantitatively one specific set of exponents' values, as it is often the case that several processes lead to the same pattern (33). Furthermore, we do not claim that our dynamic models describe any real ecosystem in all its

complexity, as such models are of course overly simplified to encompass the broad range of processes that may set the values of macroecological scaling exponents (e.g. species' interactions, landscape structure and ecological disturbances, to name a few). When modelling any natural process, the first step is that of abstraction: unnecessary details are removed, until one reaches the simplest model that is still compatible with the observed patterns, which for the purpose of this investigation are the functional forms of the various scaling relationships and distributions. In this sense, we believe that our models of birth, death and speciation capture the essential ingredients that produce the empirically-observed functional forms of macroecological laws and that set the scaling properties of the joint distribution of mass and abundance, and thus the pattern covariations. The exact values of the macroecological scaling exponents, instead, are most likely determined by several processes that are not included in our community dynamics models, but can be properly described by our scaling framework.

1. Huang K (1987) Statistical Mechanics, Wiley.
2. Azaele S, et al. (2016) Statistical mechanics of ecological systems: neutral theory and beyond. *Rev. Mod. Phys.* **88**, 035003.
3. Kolokotrones T, Savage V, Deeds EJ, Fontana W (2010) Curvature in metabolic scaling. *Nature* **464**, 753–756.
4. Mori S, et al. (2010) Mixed-power scaling of whole-plant respiration from seedlings to giant trees. *Proc. Natl Acad. Sci. USA* **107**, 1447–1451.
5. Marañón E, et al. (2013) Unimodal size scaling of phytoplankton growth and the size dependence of nutrient uptake and use. *Ecol. Lett.* **16**, 371–379.
6. Banavar JR, Cooke TJ, Rinaldo A, Maritan A (2014) Form, function, and evolution of living organisms. *Proc. Natl Acad. Sci. USA* **111**, 3332–3337.
7. Chisholm RA, Fung T, Chimalakonda D, O'Dwyer JP (2016) Maintenance of biodiversity on islands. *Proceedings of the Royal Society of London B: Biological Sciences* **283**.
8. Sheldon RW, Prakash A,Sutcliffe H (1972) The size distribution of particles in the ocean. *Limnol. Oceanogr.* **XVII**, 327–340.
9. Giometto A, Altermatt F, Carrara F, Maritan A, Rinaldo A (2013) Scaling body size fluctuations. *Proc. Natl Acad. Sci. USA* **110**, 4646–4650.
10. Siemann E, Tilman D, Haarstad J (1996) Insect species diversity, abundance and body size relationships. *Nature* **6576**, 704–706.
11. Labra FA, Hernández-Miranda E, Quiñones RA (2015) Dynamic relationships between body size, species richness, abundance, and energy use in a shallow marine epibenthic faunal community. *Ecology and Evolution* **5(2)**, 391–408.
12. Southwood TRE, May RM, Sugihara G (2006) Observations on related ecological exponents. *Proc. Natl Acad. Sci. USA* **103**, 6931–6933.
13. May RM (1988) How many species are there on Earth? *Science* **241**, 1441–1449.
14. Stegen JC, White EP (2008) On the relationship between mass and diameter distributions in tree communities. *Ecol. Lett.* **11**. 1287–1293.
15. White EP, Enquist BJ, Green, J. L. (2008) On estimating the exponents of power-law frequency distributions (Appendix A). *Ecology* **89**, 905–912.
16. Redner S (1990) Random multiplicative processes: An elementary tutorial. *Am. J. Phys.* **58**, 267–273.
17. Giometto A, Formentin M, Rinaldo A, Cohen JE, Maritan A (2015) Sample and population exponents of generalized Taylor's law. *Proc. Natl. Acad. Sci. USA* **112**, 7755–7760.
18. Burness GP, Diamond J, Flannery T. (2001) Dinosaurs, dragons, and dwarfs: the evolution of maximal body size. *Proc. Natl Acad. Sci. USA* **98**, 14518–14523.
19. Banavar JR, Damuth J, Maritan A, Rinaldo A (2007) Scaling in ecosystems and the linkage of macroecological laws. *Phys. Rev. Lett.* **98**, 068104.
20. Clauset A, Shalizi CR, Newman MEJ (2009) Power-Law Distributions in Empirical Data. *SIAM Review* **51**, 661.
21. Condit R, et al. (2012) Barro Colorado Forest Census Plot Data (Version 2012).
22. Zimmerman JK, Comita LS, Thompson J, Uriarte M, Brokaw N (2010) Patch dynamics and community metastability of a subtropical forest: compound effects of natural disturbance and human land use. *Landscape Ecology* **25**, 1099–1111.
23. Enquist BJ, Niklas KJ (2001) Invariant scaling relations across tree-dominated communities. *Nature* **410**, 655–660.
24. Simini F, Anfodillo T, Carrer M, Banavar JR, Maritan A (2010) Self-similarity and scaling in forest communities. *Proc. Natl Acad. Sci. USA* **107**, 7658–62.
25. Bhattacharjee SM, Seno F (2001) A measure of data collapse for scaling. *J. Phys. A-Mathematical Gen.* **34**, 6375–6380.
26. Novosolov M, et al. (2015) Power in numbers. The evolutionary drivers of high population density in insular lizards. *Global Ecol. Biogeogr.* **25**, 87–95.
27. Okie JG, Brown JH (2009) Niches, body sizes, and the disassembly of mammal communities on the Sunda Shelf islands. *Proc. Natl Acad. Sci. USA* **106 Suppl. 2**, 19679–19684.
28. Rensch B (1948) Histological Changes Correlated with Evolutionary Changes of Body Size. *Evolution (N. Y).* **2**, 218–230.
29. Wilson DS (1992) Complex Interactions in Metacommunities, with Implications for Biodiversity and Higher Levels of Selection. *Ecology* **73**, 1984–2000.
30. Leibold MA, et al. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* **7**, 601–613.
31. Muneepeerakul R, et al. (2008) Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin. *Nature* **453**, 220–222.
32. Bertuzzo E, et al. (2011) Spatial effects on species persistence and implications for biodiversity. *Proc. Natl Acad. Sci. USA* **108**, 4346–4351.
33. Newman MEJ (2007) Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46**, 323–351.
34. Kleiber M (1932) Body size and metabolism. *Hilgardia* **6**, 315–353.
35. Nielsen SL, Sand-Jensen K (1990) Allometric scaling of maximal photosynthetic growth rate to surface / volume ratio. *Limnol. Oceanogr.* **35**, 177–181.
36. Dodds PS, Rothman DH, Weitz JS (2001) Re-examination of the "3/4-law" of metabolism. *Journal of theoretical biology* **209**, 9–27.
37. Finkel ZV, Irwin AJ, Schofield O (2004) Resource limitation alters the 3/4 size scaling of metabolic rates in phytoplankton. *Marine Ecology Progress Series* **273**, 269–279.
38. Marañón E, Cermeno P, Rodriguez J, Zubkov M, Harris R (2007) Scaling of phytoplankton photosynthesis and cell size in the ocean. *Limnol. Ocean.* **52**, 2190–2198.
39. Lomolino MV (1982) Species-area and species-distance relationships of terrestrial mammals in the Thousand Island Region. *Oecologia* **54**, 72–75.
40. MacArthur RH, Wilson EO (1963) An Equilibrium Theory of Insular Zoogeography. *Evolution* **17**, 373–387.
41. Dodson S (1992) Predicting crustacean zooplankton species richness. *Limnology and Oceanography* **37**, 848–856.
42. Newmark WD (1986) Species-area relationship and its determinants for mammals in western North American national parks. *Biol. J. Linn. Soc.* **28**, 83–98.
43. Lonsdale WM (1999) Global Patterns of Plants Invasions and the Concept of Invasibility. *Ecology* **80**, 1522–1536.
44. Smith VH, et al. (2005) Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. *Proc. Natl Acad. Sci. USA* **102**, 4393–4396.
45. Preston FW (1962) The Canonical Distribution of Commonness and Rarity: Part I. *Ecology* **43**, 185–215.
46. Cohen JE, Xu M, Schuster WSF (2012) Allometric scaling of population variance with mean body size is predicted from Taylor's law and density-mass allometry. *Proc. Natl Acad. Sci. USA* **109**, 15829–15834.
47. Damuth J (1981) Population density and body size in mammals. *Nature* **290**, 699–700.
48. Cyr H, Peters RH, Downing JA (1997) Population density and community size structure: comparison of aquatic and terrestrial systems. *Oikos* **80**, 139–149.
49. Nee S, Read AF, Greenwood JD, Harvey PH (1991) The relationship between abundance and body size in British birds. *Nature* **351**, 312–313.
50. Cohen JE, Jonsson T, Carpenter SR (2003) Ecological community description using the food web, species abundance, and body size. *Proc. Natl Acad. Sci. USA* **100**, 1781–1786.
51. Muller-Landau HC, et al. (2006) Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecol. Lett.* **9**, 589–602.
52. White EP, Ernest SKM, Kerkhoff AJ, Enquist BJ (2007) Relationships between body size and abundance in ecology. *Trends Ecol. Evol.* **22**, 323–330.
53. Halfpenny J Small mammal disturbance data for Niwot Ridge from 1981/6/30 - 1990/8/23, yearly provided by the NSF supported Niwot Ridge Long-Term Ecological Research project and the University of Colorado Mountain Research Station.
54. Cavender-Bares KK, Rinaldo A, Chisholm SW (2001) Microbial size spectra from natural and nutrient enriched ecosystems. *Limnol. Oceanogr.* **46**, 778–789.
55. Rinaldo A, Maritan A, Cavender-Bares KK, Chisholm SW (2002) Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. *P. Roy. Soc. B: Bio* **269**, 2051–2059.
56. Marañón E (2015) Cell Size as a Key Determinant of Phytoplankton Metabolism and Community Structure. *Ann. Rev. Marine Science* **7**, 241–264.
57. Marquet PA, Taper ML (1998) On size and area: Patterns of mammalian body size extremes across landmasses. *Evol. Ecol.* **12**, 127–139.
58. Smith FA, et al. (2003) Body Mass of Late Quaternary Mammals. *Ecology* **84**, 3403.
59. Marquet PA, et al. (2005) Scaling and power-laws in ecological systems. *J. Exp. Biol.* **208**, 1749–1769.
60. Preston FW (1948) The Commonness, And Rarity, of Species. *Ecology* **29**, 254–283.
61. Taylor LR (1961) Aggregation, variance and the mean. *Nature* **189**, 732–735.
62. Taylor, L, Woiwod, I, & Perry, J. (1980) Variance and the large scale spatial stability of aphids, moths and birds. *J. Anim. Ecol.* **49**, 831–854.
63. Anderson, R. M, Gordon, D. M, Crawley, M. J, & Hassell, M. P. (1982) Variability in the abundance of animal and plant species. *Nature* **296**, 245–248.

Table S1. Estimates for the scaling exponents of the ecological patterns depicted in Fig. 1 (main text). Errors are SEM, CI stands for confidence interval. If no error is reported, none was given in the original paper.

Law	Forests	Terrestrial	Aquatic
Kleiber's law (panel a)	0.80 ± 0.01	0.67 ± 0.2	1.10, CI 95% : $[0.94, 1.21]$
Damuth's law (panel b)	0.26 ± 0.05	0.57 ± 0.08	0.73, CI 95% : $[0.73, 0.92]$
SAR (panel c)	0.27 ± 0.01	0.23 ± 0.02	0.094
$s(m)$	1.59, CI 95% : $[1.57, 1.63]$	1.5 ± 0.2	2.11

Table S2. References to empirical measurements of the ecological patterns discussed in the main text. RSA stands for Relative Species Abundance and Max. body mass stands for the scaling of the maximum body mass with the area of the ecosystem.

Law	Forests	Terrestrial	Aquatic
Kleiber's law	Mori *et al.* (2010) (4)	Kleiber (1932) (34) Dodds *et al.* (2001) (36)	Nielsen and Sand-Jensen (1990) (35) Finkel *et al.* (2004) (37) Marañón *et al.* (2007) (38)
SAR	Lomolino (1982) (39)	MacArthur and Wilson (1963) (40) Newmark (1986) (42) Okie *et al.* (2009) (27) Preston (1962) (45)	Dodson (1992) (41) Lonsdale (1999) (43) Smith *et al.* (2005) (44)
Damuth's law	Cohen *et al.* (2012) (46)	Damuth (1981) (47) Nee *et al.* (1991) (49) Novosolov *et al.* (2015) (26)	Cyr *et al.* (1997) (48) Cohen *et al.* (2003) (50)
$s(m)$	Muller-Landau *et al.* (2006) (51) Stegen and White (2008) (14) Condit *et al.* (2012) (21)	White *et al.* (2007) (52) Halfpenny (2016) (53)	Sheldon (1972) (8) Cavender-Bares *et al.* (2001) (54) Rinaldo *et al.* (2002) (55) Marañón *et al.* (2015) (56)
$P(m)$		Marquet and Taper (1998) (57) Smith *et al.* (2003) (58) Marquet *et al.* (2005) (59) Southwood *et al.* (2006) (12)	
Max. body mass		Burness *et al.* (2001) (18) Okie *et al.* (2009) (27)	
RSA		Preston (1948) (60)	
Taylor's law	Giometto *et al.* (2015) (17)	Taylor *et al.* (1961) (61) Taylor *et al.* (1980) (62) Anderson *et al.* (1982) (63)	

Table S3. Estimates of scaling exponents η, δ and γ in the BCI forest. Errors are computed as reported in the text.

BCI forest	Census 5	Census 6	Census 7
$s(m)$	$\eta = 1.43 \pm 0.04$	$\eta = 1.43 \pm 0.03$	$\eta = 1.44 \pm 0.04$
$P(m)$	$\delta = 1.03 \pm 0.03$	$\delta = 1.05 \pm 0.03$	$\delta = 1.07 \pm 0.05$
Damuth's law	$\gamma = 0.41 \pm 0.07$	$\gamma = 0.40 \pm 0.06$	$\gamma = 0.38 \pm 0.06$

Table S4. Estimates of scaling exponents η, δ and γ in the Luquillo forest (22). Errors are computed as reported in the text.

Luquillo forest	Census 1	Census 2	Census 3	Census 4	Census 5
$s(m)$	$\eta = 1.27 \pm 0.04$	$\eta = 1.18 \pm 0.03$	$\eta = 1.09 \pm 0.03$	$\eta = 1.09 \pm 0.04$	$\eta = 0.95 \pm 0.07$
$P(m)$	$\delta = 1.02 \pm 0.02$	$\delta = 1.01 \pm 0.03$	$\delta = 1.02 \pm 0.05$	$\delta = 1.02 \pm 0.03$	$\delta = 1.02 \pm 0.01$
Damuth's law	$\gamma = 0.21 \pm 0.08$	$\gamma = 0.23 \pm 0.04$	$\gamma = 0.16 \pm 0.04$	$\gamma = 0.09 \pm 0.06$	$\gamma = 0.03 \pm 0.04$

Table S5. Estimates of scaling exponents z, δ, Φ and γ for the LIZ dataset (26). Errors on z and δ are SEM, the error on γ is the SD, the error on Φ was obtained by bootstrapping. $P(\tilde{m})$ is computed gathering together species from all the islands in the dataset.

SAR	$z = 0.17 \pm 0.01$	R^2=0.46
Damuth's law	$\Phi = 0.78 \pm 0.08$	
Damuth's law	$\gamma = 0.53 \pm 0.03$	R^2=0.89
$P(m)$	$\delta = 1.45 \pm 0.06$	

Table S6. Estimates of scaling exponents z, ξ and δ for the SSI dataset (27). Errors on z, δ and ξ are SEM. R^2 is the coefficient of determination. $P(\tilde{m})$ is computed gathering together species from all the islands in the dataset.

SAR	$z = 0.23 \pm 0.02$	$R^2 = 0.93$
$M_{\max}$	$\xi = 0.49 \pm 0.09$	$R^2 = 0.76$
$P(m)$	$\delta = 1.6 \pm 0.2$	

Table S7. Scaling exponents measured in the basic model for different sets of parameters' values. Each row (delimited by horizontal lines) refers to a set, indicated by the parameter which value has been changed with respect to the parameters set described in the main text (indicated as "Main text set", specified in section 3.1). Under each value, the lower and upper ends of the confidence intervals are reported. The estimates and the confidence intervals were obtained as described in the Methods section.

	δ	η	γ	Φ	λ	χ	ω	μ	ν
Main text set	2.23	2.54	0.26	0.74	0.38	0.99	0.18	1.0016	0.9982
Fig. S9	2.05	2.28	0.20	0.72	0.32	0.87	0.17	1.0016	0.9980
	2.41	2.79	0.34	0.76	0.42	1.03	0.21	1.0017	0.9983
$w = 10^{-4}$	2.27	2.64	0.33	0.76	0.52	0.98	0.19	1.0030	0.9968
Fig. S10	2.01	2.22	0.28	0.74	0.50	0.92	0.17	1.0027	0.9967
	2.54	2.94	0.38	0.78	0.57	1.04	0.20	1.0034	0.9969
$w = 10^{-5}$	2.28	2.63	0.30	0.75	0.46	0.94	0.18	1.000	0.998
Fig. S11	2.00	2.21	0.25	0.74	0.41	0.73	0.16	0.996	0.995
	2.54	2.94	0.36	0.79	0.52	1.05	0.22	1.005	1.001
Fig. S12	2.00	2.21	0.26	0.74	0.73	0.93	0.18	1.0029	0.9989
	2.54	2.93	0.31	0.78	0.87	1.03	0.21	1.0031	0.9989
$\alpha = 1/4$	2.38	2.64	0.29	0.76	0.63	0.98	0.19	1.000	1.000
Fig. S13	2.01	2.22	0.25	0.73	0.60	0.94	0.19	0.998	0.998
	2.54	2.94	0.35	0.80	0.69	1.03	0.20	1.002	1.000
$\theta = 1/2$	2.13	2.62	0.49	0.78	0.47	0.95	0.22	1.002	0.997
Fig. S14	1.88	2.26	0.26	0.75	0.44	0.87	0.21	1.002	0.996
	2.36	2.94	0.56	0.86	0.53	1.03	0.26	1.003	0.998
$z = 1/2$	2.23	2.41	0.32	0.50	0.93	0.98	0.38	1.0	0.99
Fig. S15	2.05	2.28	0.31	0.49	0.92	0.96	0.36	1.00	0.99
	2.54	2.79	0.34	0.51	0.96	1.05	0.39	1.01	1.00

Table S8. Scaling exponents measured in the variation of the basic model described in section 3.1.2, Fig. S16. Under each value, the lower and upper ends of the confidence intervals are reported. The estimates and the confidence intervals were obtained as described in the Methods section.

δ	η	γ	Φ	λ	χ	ω	μ	ν
2.54	2.78	0.13	0.76	0.36	0.94	0.14	1.0021	0.998
2.22	2.29	0.04	0.74	0.35	0.7	0.14	1.0021	0.998
2.83	3.11	0.28	0.78	0.40	1.1	0.17	1.0022	0.998

Table S9. Scaling exponents measured in the community dynamics models with fluctuating numbers of species. Each row (delimited by horizontal lines) refers to a different model (see section 3.2). Model parameters are reported in Figs. S17 and S18. Under each value, the lower and upper ends of the confidence intervals are reported. The estimates and the confidence intervals were obtained as described in the Methods section.

Model	δ	η	γ	Φ	λ	χ	ω	μ	ν	z
a	2.52	2.83	0.28	0.51	0.92	1.00	0.32	1.01	1.00	0.50
Fig. S17	2.26	2.50	0.24	0.48	0.82	0.91	0.29	1.00	0.99	0.49
	2.77	3.15	0.34	0.56	0.99	1.08	0.33	1.01	1.00	0.50
b	2.52	2.69	0.13	0.50	0.60	0.98	0.32	1.00	1.00	0.50
Fig. S18	2.27	2.41	0.10	0.48	0.56	0.92	0.30	0.98	0.99	0.49
	2.76	2.96	0.18	0.52	0.66	1.11	0.34	1.02	1.01	0.50

Table S10. Scaling exponents measured in the models with fluctuating numbers of species and $\bar{q} > 1$ (see section 3.3). Each row (delimited by horizontal lines) refers to a model and parameter set. Model parameters are reported in Figs. S19 and S20. Under each value, the lower and upper ends of the confidence intervals are reported. The estimates and the confidence intervals were obtained as described in the Methods section. The parameter $\bar{q}$ is the mean of the multiplicative factor q that defines the descendant species' mass at each speciation event (cfr. section 3.2).

Model	δ	η	γ	Φ	λ	χ	ω	μ	ν	z
a with	1.96	2.22	0.31	0.52	0.80	0.94	0.50	1.06	0.98	0.49
$\bar{q} = 1.2$	1.75	1.98	0.26	0.50	0.75	0.75	0.45	0.91	0.95	0.48
Fig. S19	2.16	2.44	0.34	0.55	0.90	1.18	0.58	1.20	1.00	0.50
b with	1.93	2.06	0.25	0.53	0.93	0.91	0.64	1.13	0.93	0.47
$\bar{q} = 1.2$	1.71	1.83	0.19	0.50	0.89	1.65	0.57	0.97	0.87	0.43
Fig. S20	2.14	2.27	0.28	0.55	1.10	0.46	0.70	1.30	0.99	0.52

Table S11. Summary of the empirical tests performed. References are to equations in the main text.

Dataset	Measured exponents	Relationship that was verified
BCI	η, γ, δ (Table S3)	Eq.(1): $\eta = \delta + \gamma$
Luquillo	η, γ, δ (Table S4)	Eq.(1): $\eta = \delta + \gamma$
LIZ	z, Φ,γ, δ (Table S5)	Eq.(6): $z = 1 - \Phi - \max\{0, \lambda(1 + \alpha - \gamma - \delta)\}$
SSI	z, ξ, δ (Table S6)	Eq.(9): $\xi = \frac{z}{\delta - 1}$

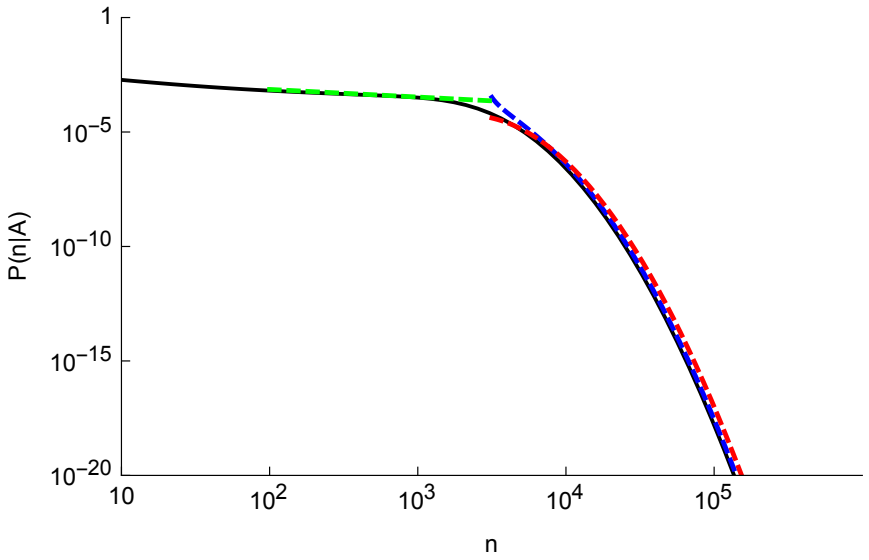

Fig. S1. Relative species abundance $P(n|A)$ computed numerically via Eq. S21 (solid black curve), with $A = 100$, $h_0 = 100$, $\sigma = 1/10$, $\delta = 3/2$, $\gamma = 3/4$, $\lambda = 3/4$ and $\Phi = 3/4$. Shown are the approximations to the RSA computed via Eq. S24 (dashed blue curve) and Eq. S25 (dashed green curve). The blue and red curves are plotted for $n > A^{\Phi} h_0 e^{-\sigma/4}$. The green curve is plotted for $100 \leq n \leq A^{\Phi} h_0 e^{-\sigma/4}$. A lognormal tail is plotted for comparison (dashed red curve).

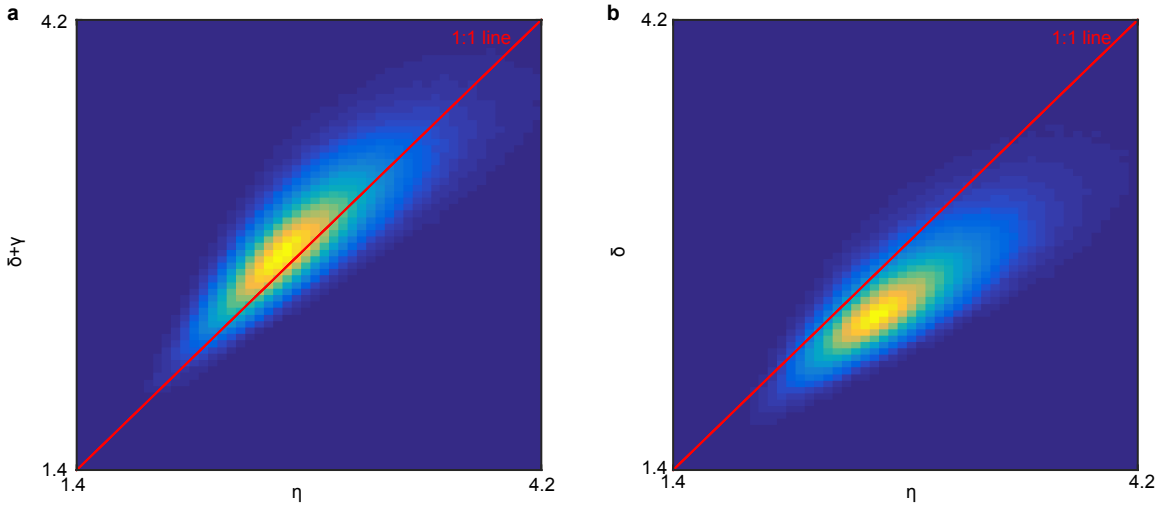

Fig. S2. Density-scatter plot of $\delta + \gamma$ (panel a) and δ (panel b) versus η in simulations of the stochastic community dynamics model (model *a*, see section 3.2), with the exponents estimated at each sampling time-point. The parameters of the stochastic community dynamics model are reported in Fig. S17; shown are simulation data for the largest simulated area $A = 10^3$. Density histograms are normalized to one, with blue representing the value zero and yellow the value one. At each time-step, the exponent γ was extimated by linear least-squares fit of ($\log(m_i)$,$\log(n_i)$) where the index i runs on all the species present in the ecosystem at that time-step. Note that the small deviations from the 1:1 line are also due to statistical errors in the estimation of the exponents at each time step. Panel b) shows that the prediction (19) that $\eta = \delta$ is not supported by simulation data.

General case — SSI

$$\eta = \delta + \gamma$$

$$\left.\begin{aligned} z &= 1 - \Phi - \max\{0, \lambda(1+\alpha-\eta)\} \\ \nu &= 1 - \max\{0, \lambda(1+\alpha-\eta)\} \end{aligned}\right\} \quad z + \Phi = \nu$$

$$\mu = 1 + \max\{0, \lambda(2-\eta)\} - \max\{0, \lambda(1+\alpha-\eta)\}$$

$$z = \xi(\delta - 1)$$

If η>2 — LIZ	If η≤2			
	1+α≤ η≤2		1≤ η≤1+α	
	λ>0	λ=0	λ>0	λ=0 — Forests
$z = 1-\Phi$ $\mu = 1$ $\nu = 1$	$z = 1-\Phi$ $\mu = 1+\lambda(2-\eta)$ $\nu = 1$ $\gamma < 1, \delta < 2$ $z < \xi$	$z = 1-\Phi$ $\mu = 1$ $\nu = 1$ $\gamma < 1, \delta < 2$ $z < \xi$	$z = 1-\Phi-\lambda(1+\alpha-\eta)$ $\mu = 1+\lambda(1-\alpha)$ $\nu = 1-\lambda(1+\alpha-\eta)$ $\gamma < \alpha,\ \delta < 1+\alpha$ $z < \xi\alpha,\ \Phi < 1-z$	$z = 1-\Phi$ $\mu = 1$ $\nu = 1$ $\gamma < \alpha,\ \delta < 1+\alpha$ $z < \xi\alpha$
	If ξ=1/2, α=3/4		If ξ=1/2, α=3/4	
	$z < \frac{1}{2}$		$z < \frac{3}{8}$	

Fig. S3. Scheme of predictions on the values or bounds of scaling exponents based on the linking relationships (S48a-e). Datasets names are located in different columns according to the available information on their exponents' values. Forests include Barro Colorado Island (21), Luquillo (22), SSI stands for Sunda Shelf Islands (27), LIZ for the dataset of lizard densities on islands worldwide (26). Note that the relationship $z < \xi\alpha$ is valid for forests only before the physiological constraint has been attained (see section 1.8.3).

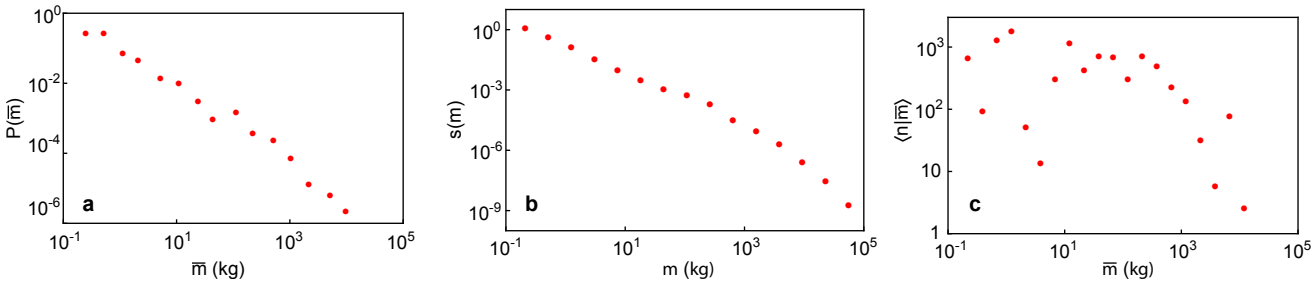

Fig. S4. Scaling patterns in the Luquillo forest, first census. a) $P(\bar{m})$ vs $\bar{m}$, b) $s(m)$ vs m, c) $\langle n|\bar{m}\rangle$ vs $\bar{m}$. Scaling exponents estimates are reported in the text. Note that finite-size effects may be present both at small and large values of m and $\bar{m}$, for example due to the sampling protocol (see text).

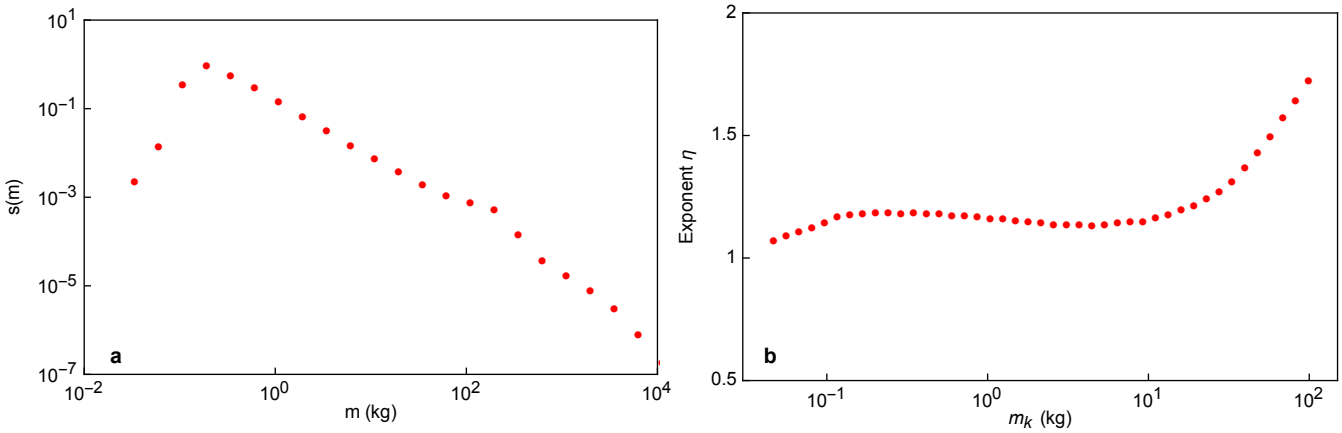

Fig. S5. a) Size spectrum in the Luquillo forest, second census. Finite-size effect are present both at small (i.e. $m < 0.2$ kg) and large (i.e. $m > 10^2$ kg) values of m. b) Size-spectrum exponent η estimated using only data with mass $m > m_k$. The estimated exponent initially increases (until $m_k \simeq 0.2$ kg) due to a finite-size effect, then is rather stable until the statistics is not sufficient to properly estimate it ($m_k > 20$ kg).

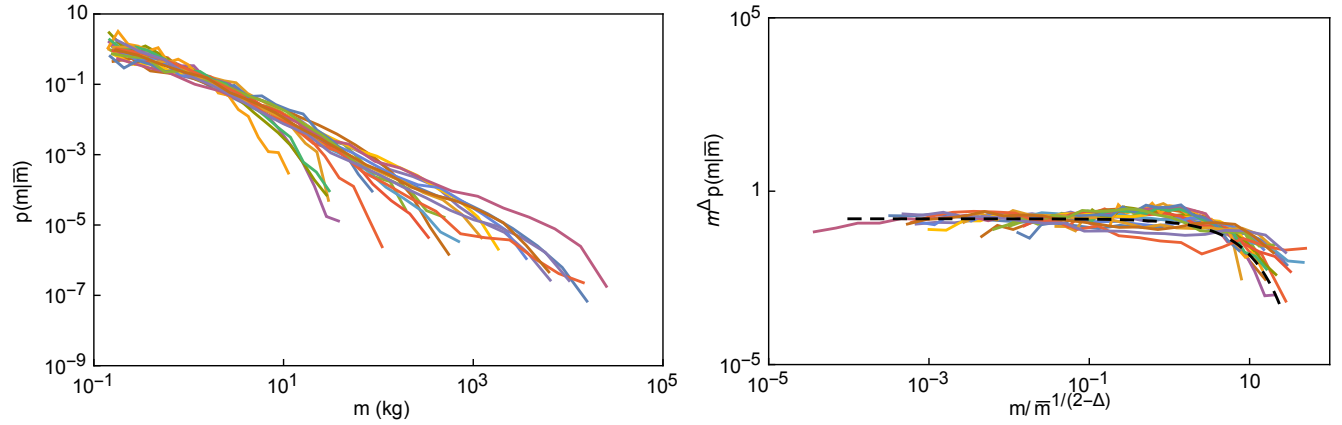

Fig. S6. a) Intraspecific tree size distributions in BCI (most abundant species). Each color corresponds to a different species. b) Intraspecific tree size distributions collapse according to Eq. S50 onto the same universal curve. The dashed black line is the best fit of $\mathcal{F}(x) = q_0 e^{-q_1 x}$ (see text).

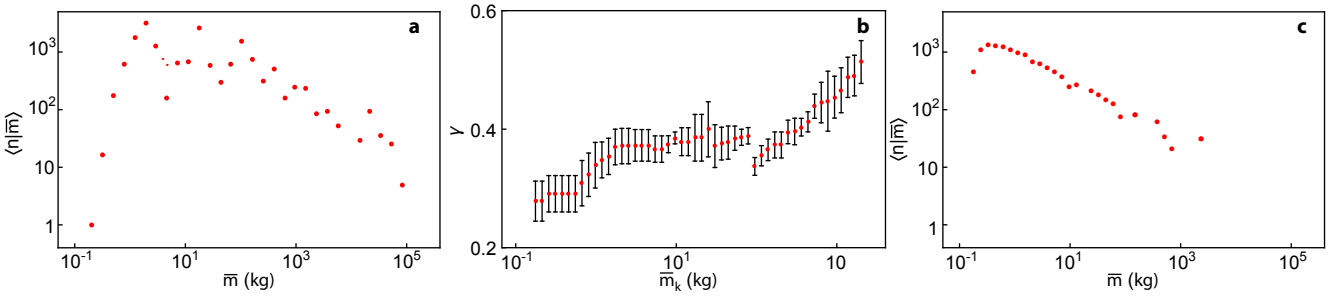

Fig. S7. a) Damuth's law ($\langle n|\bar{m}\rangle$ vs $\bar{m}$) in BCI, seventh census. Finite-size effect are present both at small (i.e. $\bar{m} < 10$ kg) and large (i.e. $\bar{m} > 10^3$ kg) values of $\bar{m}$. b) Damuth's law exponent γ estimated using only the species with typical mass $\bar{m} > \bar{m}_k$. The estimated exponent initially increases, then is stable until the statistics is not sufficient to properly estimate it (or finite-size effects become relevant). c) Damuth's law in an artificial forest after mimicking the sampling protocol (see text for details). The effect of the sampling bias is visible at small masses.

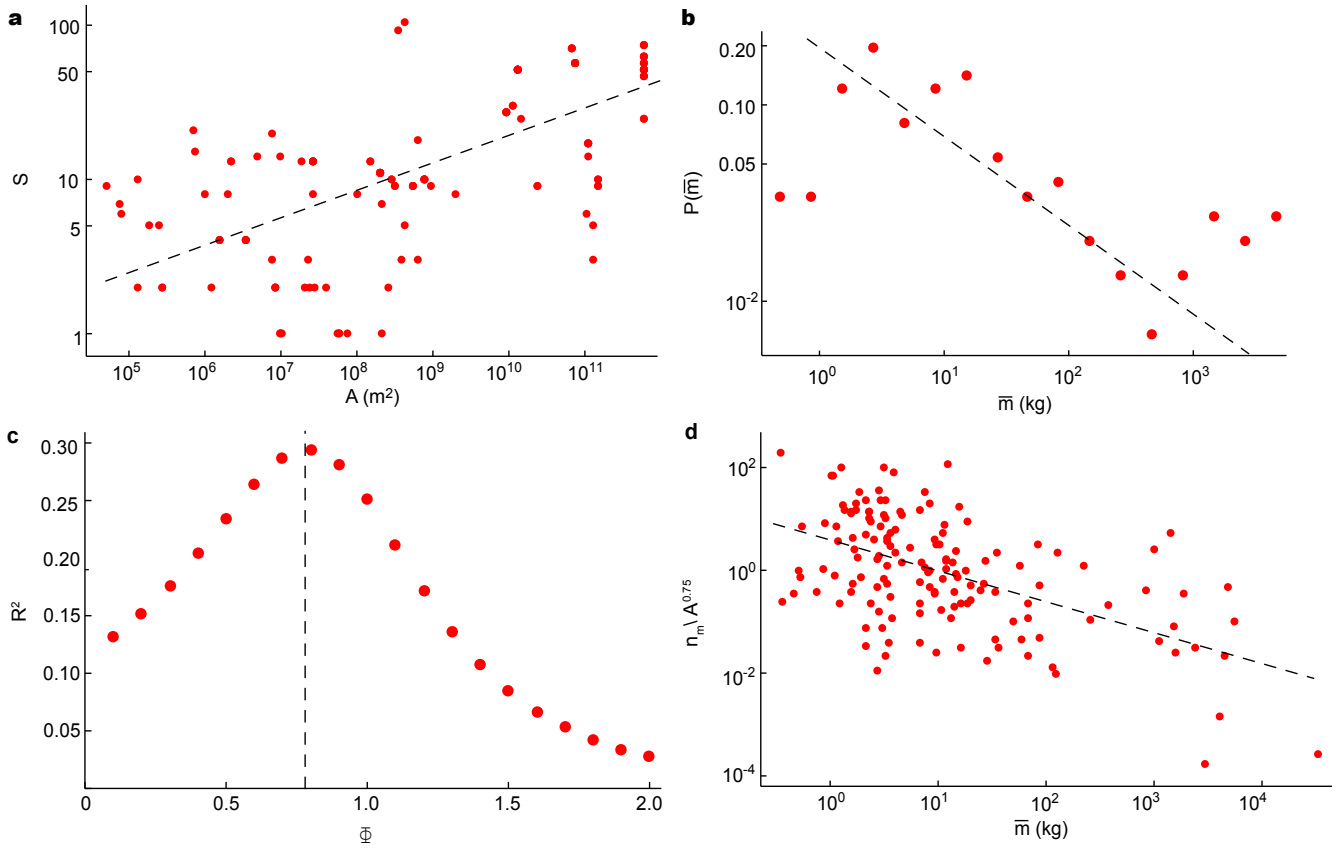

Fig. S8. Empirical evidence of scaling patterns in LIZ dataset (26): a) SAR; b) P($\bar{m}$); c) Coefficient of determination R^2 of the linear least-squares fit of $\left(\log \bar{m}, \log \frac{n(\bar{m},A)}{A^{\Phi}}\right)$ for $\Phi \in [0, 2]$. Dashed line in correspondence of the value $\Phi = 0.75$ giving the best fit; d) Non-binned Damuth's law plotted using the estimated value of Φ ($\bar{m}$ are species' mean masses). Best fit parameters are reported in Table S5, details on the fit in the SI text.

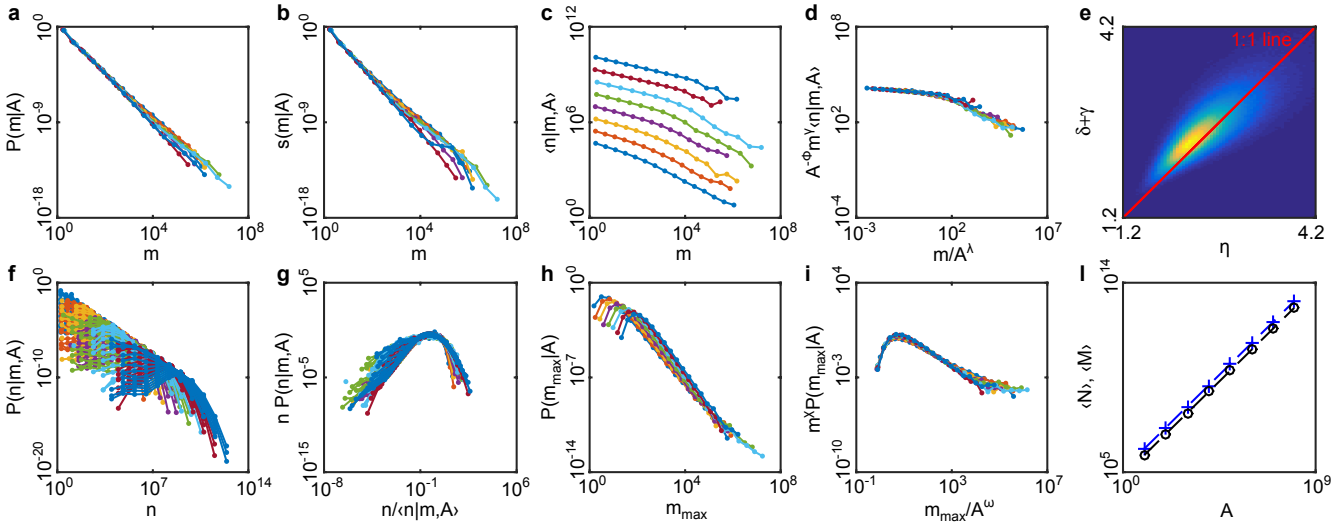

Fig. S9. Basic model statistics with parameters $z = 1/4$, $w = 10^{-3}$, $\alpha = 3/4$, $\theta = 1/4$ (simulation results shown in Fig. 3 of the main text). Different colors refer to different values of $A = 10^i$, from $i = 1$ (lower blue curve in panel c) to $i = 8$ (upper blue curve in panel c). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^8$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

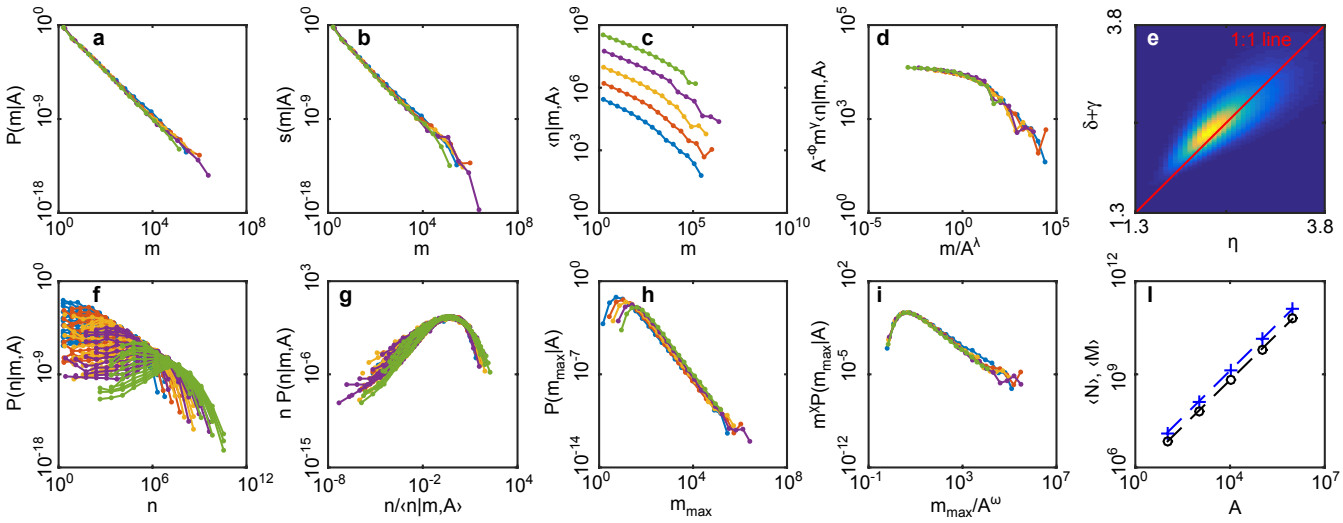

Fig. S10. Basic model statistics with parameters $z = 1/4$, $w = 10^{-4}$, $\alpha = 3/4$, $\theta = 1/4$. Different colors refer to different values of ecosystem area $A = 10^i$, with $i = 2$ (cyan); 3 (orange); 4 (yellow); 5 (purple); 6 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^6$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

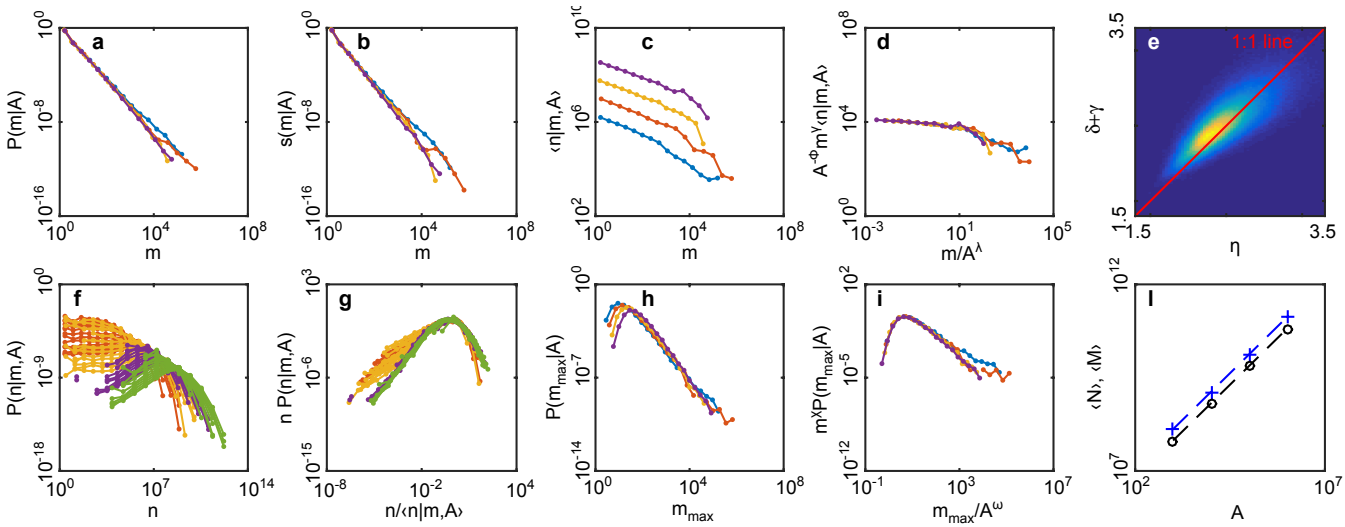

Fig. S11. Basic model statistics with parameters $z = 1/4$, $w = 10^{-5}$, $\alpha = 3/4$, $\theta = 1/4$. Different colors refer to different values of ecosystem area $A = 10^i$, with $i = 3$ (cyan); 4 (orange); 5 (yellow); 6 (purple). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^6$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

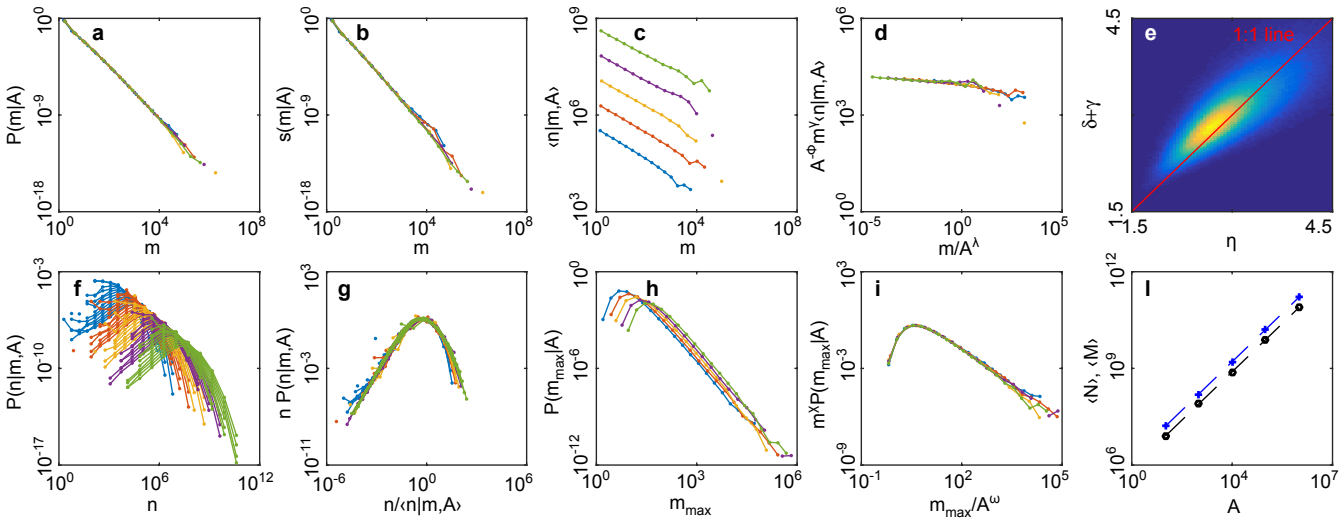

Fig. S12. Basic model statistics with parameters $z = 1/4$, $w = 10^{-3}$, $\alpha = 1/2$, $\theta = 1/4$. Different colors refer to different values of ecosystem area $A = 10^i$, with $i = 2$ (cyan); 3 (orange); 4 (yellow); 5 (purple); 6 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^6$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

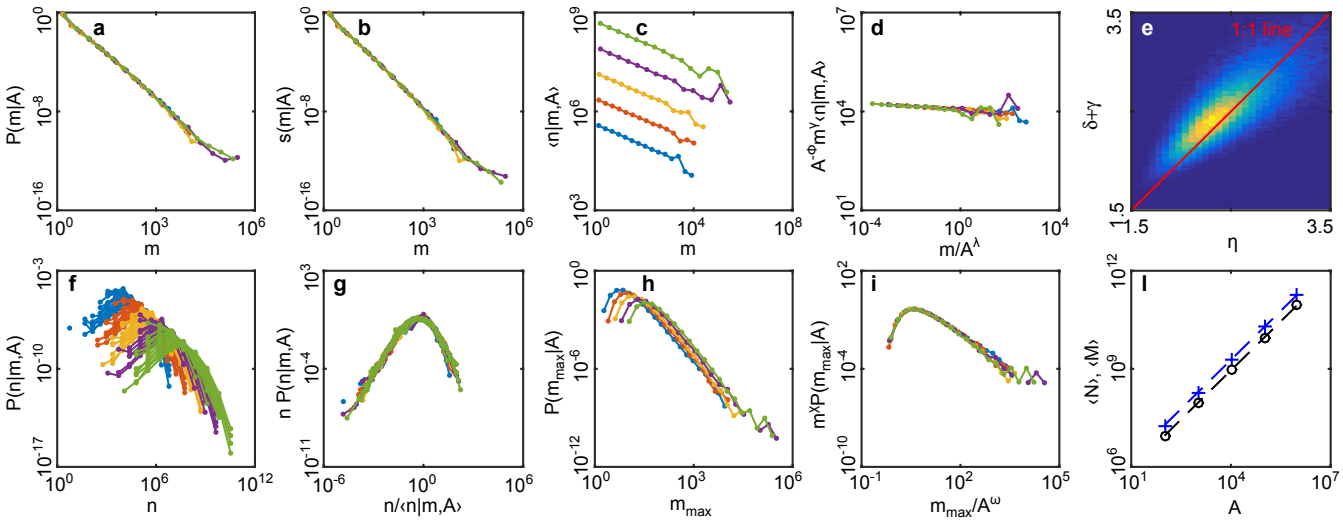

Fig. S13. Basic model statistics with parameters $z = 1/4$, $w = 10^{-3}$, $\alpha = 1/4$, $\theta = 1/4$. Different colors refer to different values of ecosystem area $A = 10^i$, with $i = 2$ (cyan); 3 (orange); 4 (yellow); 5 (purple); 6 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^6$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

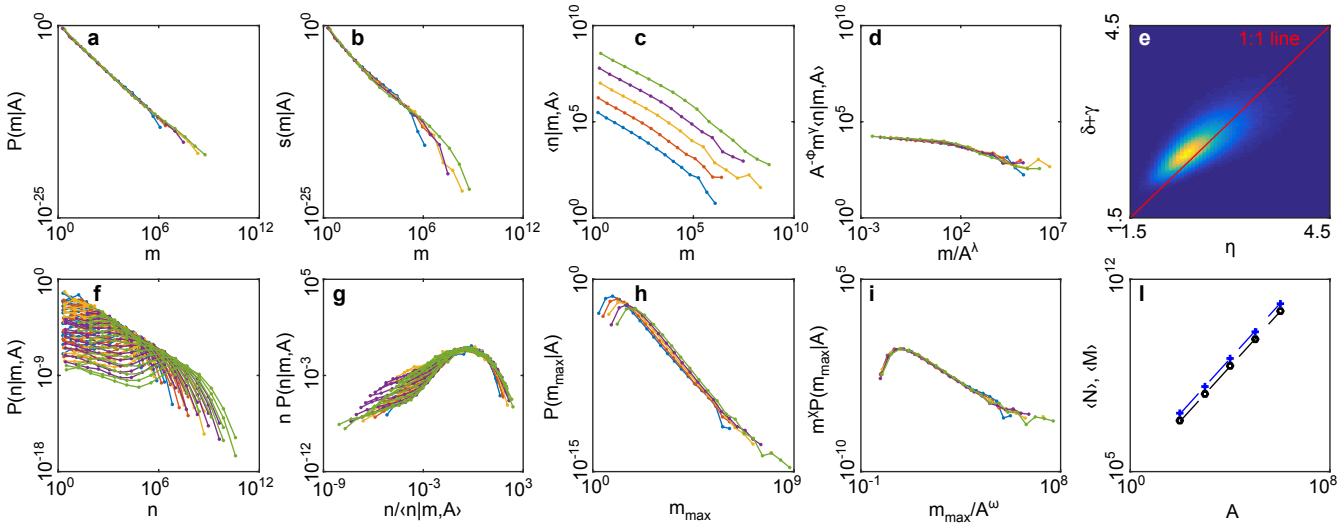

Fig. S14. Basic model statistics with parameters $z = 1/4$, $w = 10^{-3}$, $\alpha = 3/4$, $\theta = 1/2$. Different colors refer to different values of ecosystem area $A = 10^i$, with $i = 2$ (cyan); 3 (orange); 4 (yellow); 5 (purple); 6 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^6$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black dots and dashed lines) with A. See Table S7 for estimates of exponents' values.

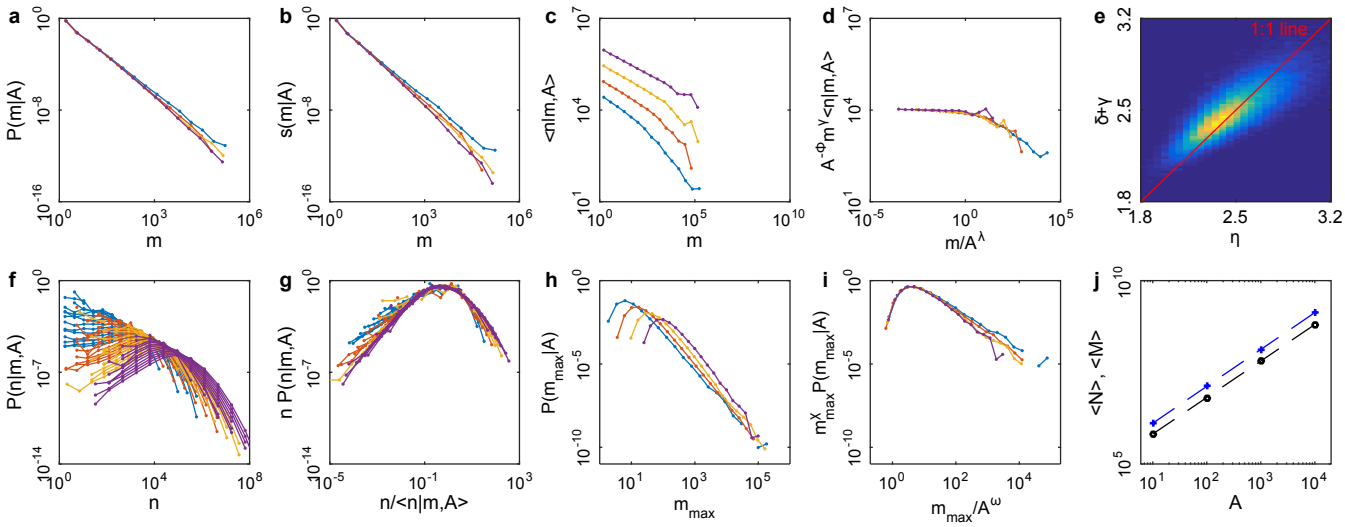

Fig. S15. Basic model statistics with parameters $z = 1/2$, $w = 10^{-3}$, $\alpha = 3/4$, $\theta = 1/4$. Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); 2 (orange); 3 (yellow); 4 (purple). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^4$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black crosses and dashed lines) with A. See Table S7 for estimates of exponents' values.

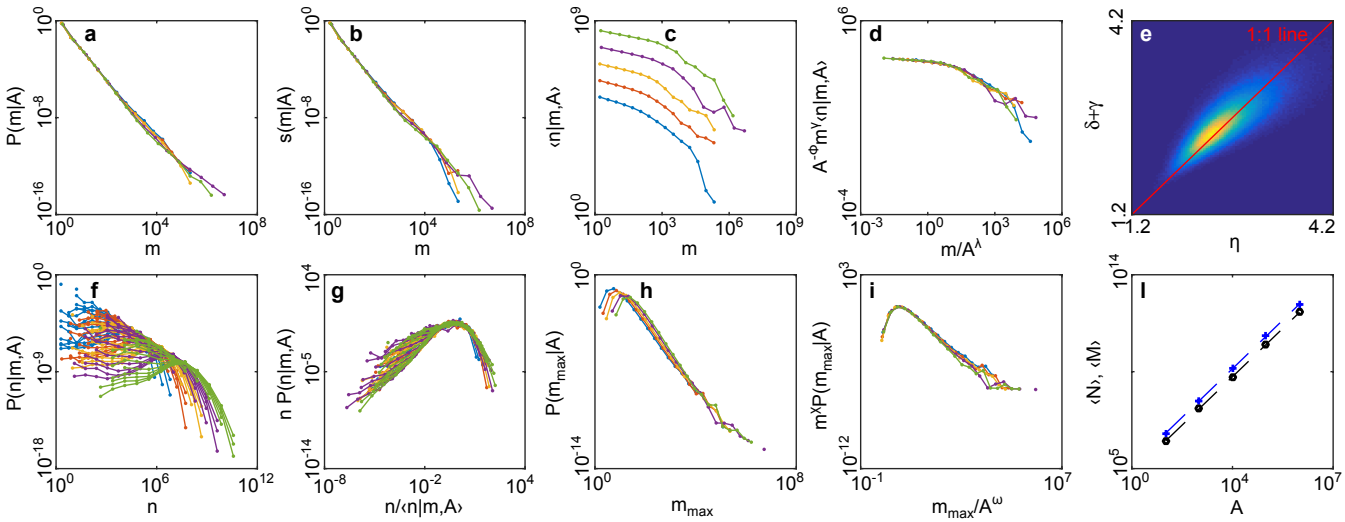

Fig. S16. Variation of the basic model, described in section 3.1.2. Statistics computed with parameters $z = 1/4$, $w = 10^{-6}$, $\alpha = 3/4$, $\theta = 1/2$. Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); $3/2$ (orange); 2 (yellow); $5/2$ (purple); 3 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^3$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines) and average total abundance $\langle N\rangle$ (black crosses and dashed lines) with A. See Table S8 for estimates of exponents' values.

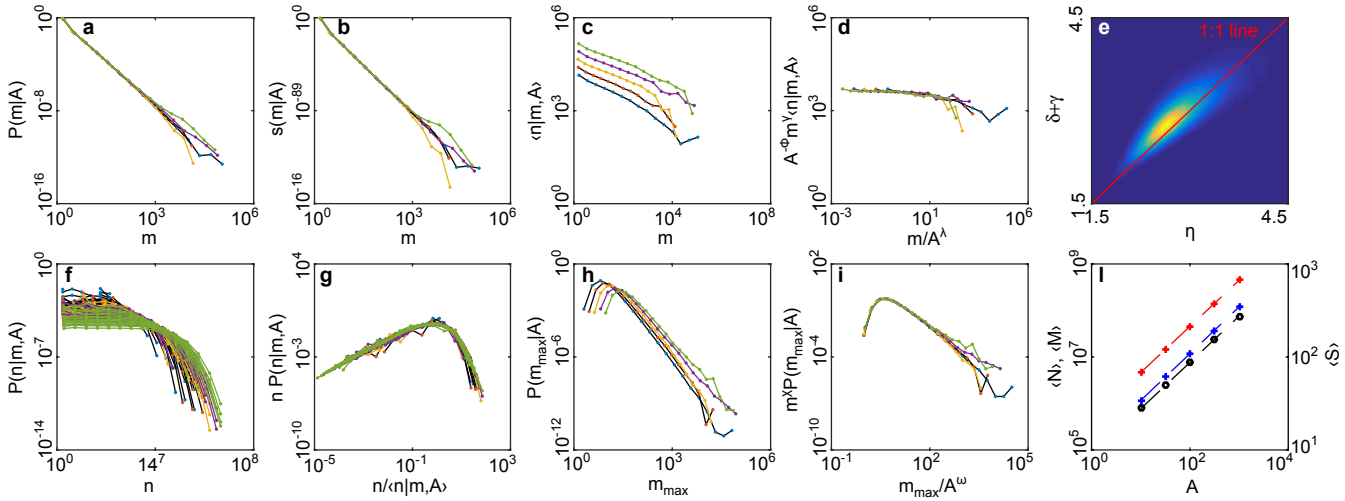

Fig. S17. Model *a* (section 3.2) statistics computed with the parameter set: $\alpha = 3/4$, $\theta = 1/4$, $w = 10^{-2}$, $v_0 = 1/2$ and $c = 10^{-5}$. Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); $3/2$ (orange); 2 (yellow); $5/2$ (purple); 3 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta+\gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^3$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines), average total abundance $\langle N\rangle$ (black crosses and dashed lines) and average number of species $\langle S\rangle$ (red crosses and dashed lines) with A. See Table S9 for estimates of exponents' values.

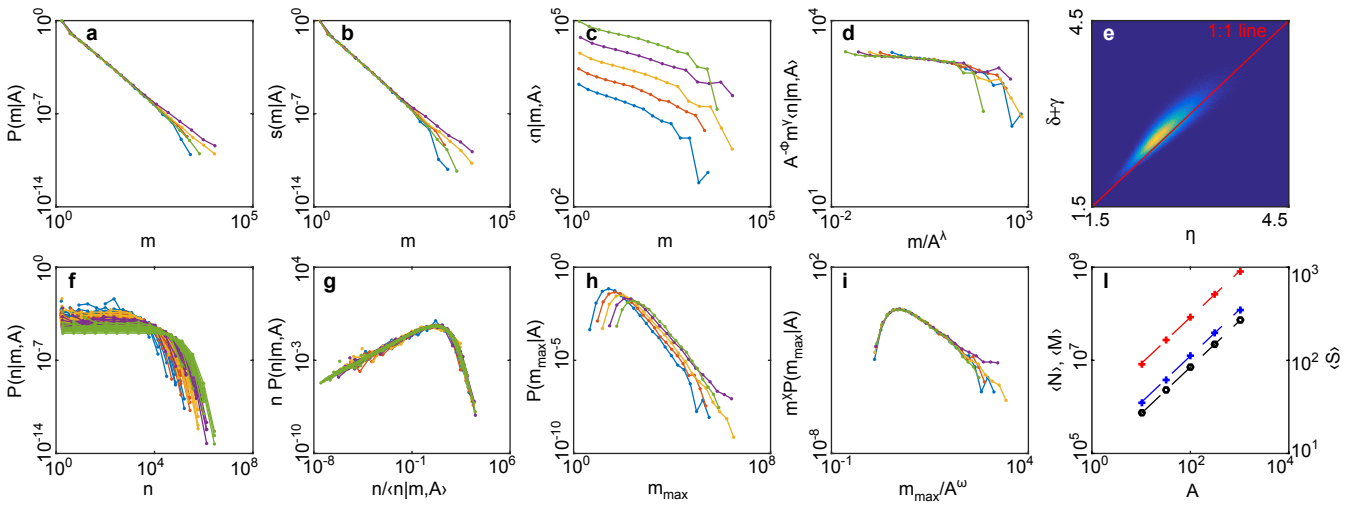

Fig. S18. Model *b* (section 3.2) statistics computed with the parameter set: $\alpha = 3/4$, $\theta = 1/4$, $w = 10^{-2}$, $v_0 = 1/2$ and $c = 10^{-5}$. Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); $3/2$ (orange); 2 (yellow); $5/2$ (purple); 3 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta+\gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^3$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines), average total abundance $\langle N\rangle$ (black crosses and dashed lines) and average number of species $\langle S\rangle$ (red crosses and dashed lines) with A. See Table S9 for estimates of exponents' values.

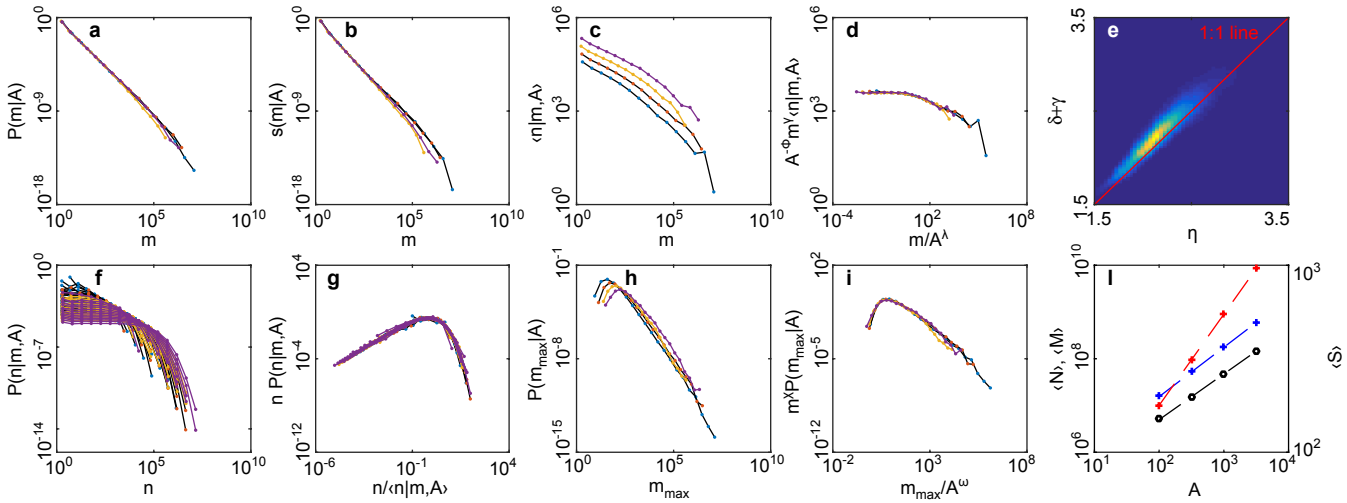

Fig. S19. Model *a* (section 3.2) statistics computed with the parameter set: $\alpha = 3/4$, $\theta = 1/4$, $w = 10^{-2}$, $v_0 = 1/2$, $c = 10^{-5}$ and $\bar{q} = 1.2$ ($\bar{q}$ is the mean of the multiplicative factor q that defines the descendant species' mass at each speciation event, cfr. section 3.2). Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); $3/2$ (orange); 2 (yellow); $5/2$ (purple); 3 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m, A\rangle$, $P(n|m, A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta + \gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^3$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines), average total abundance $\langle N\rangle$ (black crosses and dashed lines) and average number of species $\langle S\rangle$ (red crosses and dashed lines) with A. See Table S10 for estimates of exponents' values.

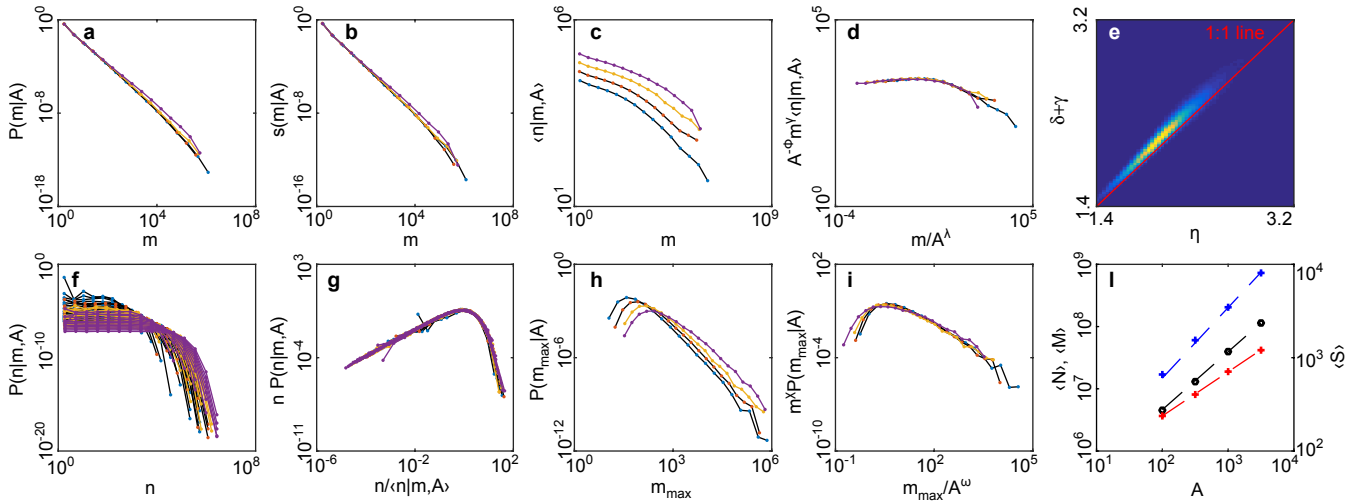

Fig. S20. Model *b* (section 3.2) statistics computed with the parameter set: $\alpha = 3/4$, $\theta = 1/4$, $w = 10^{-2}$, $v_0 = 1/2$, $c = 10^{-5}$ and $\bar{q} = 1.2$ ($\bar{q}$ is the mean of the multiplicative factor q that defines the descendant species' mass at each speciation event, cfr. section 3.2). Different colors refer to different values of ecosystem area $A = 10^i$, where $i = 1$ (cyan); $3/2$ (orange); 2 (yellow); $5/2$ (purple); 3 (green). Panels a-c, f and h show respectively $P(m|A)$, $s(m|A)$, $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$ estimated at stationarity. Panels d, g and i show collapses of simulation data for $\langle n|m,A\rangle$, $P(n|m,A)$ and $P(m_{\max}|A)$, respectively. Panel e shows the density scatter-plot of $\delta+\gamma$ versus η. The density histogram is normalized to one, with blue representing the value zero and yellow the value one. Shown are simulation data for the largest area $A = 10^3$. Panel j shows the scaling of the average total biomass $\langle M\rangle$ (blue crosses and dashed lines), average total abundance $\langle N\rangle$ (black crosses and dashed lines) and average number of species $\langle S\rangle$ (red crosses and dashed lines) with A. See Table S10 for estimates of exponents' values.