Supporting Information

Giometto et al. 10.1073/pnas.1301552110

SI Text

Meaning of Power Laws. Power laws are ubiquitous in nature. They have been shown to describe the similarity of the part and the whole of many objects in nature, ranging from coastlines, clouds, or mountain ranges (1) to natural or artificial networks like rivers (2) and up to complex social interactions (3–5). In general, a homogeneous power function [e.g., $f(x) = Cx^a$ with C, a constants] is intrinsically self-similar: If x is rescaled (multiplied by a constant), then $f(x)$ is still proportional to x^a, albeit with a different constant of proportionality. Such functions reproduce themselves on rescaling, and therefore lack natural scales, do not harbor a characteristic unit, and are said to be scale-free or true on all scales. Power-law probability distributions of size may imply an infinite mean, unless a finite range of sizes is assumed. In such a case and depending on the value of the scaling exponent, the mean, variance, and progressive moments of the distributions diverge in the infinite range. To be finite, they must depend on a finite interval of sizes sampled; thus, the mean makes no sense as a property of a population (6). This analog to the syndrome of infinite variance (1) (i.e., the progressive divergence of the variance of a self-similar or self-affine signal as the sample size is enlarged) is widely held as the typical signature of scale invariance.

Finite-Size Scaling Distributions. A general account of finite-size scaling (7) in ecology is provided by Banavar et al. (8). General properties of finite-size scaling distributions (as in Eq. **1**) are detailed in a study by Banaver et al. (9) and its supplementary information. A brief account of the derivation of the most relevant results, adapted to the case at hand, is given in the following sections.

Normalization. Here, we study the normalization conditions for size distributions of the form:

$$p_k(m) = \frac{1}{m^{\Delta}} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right), \qquad \textbf{[S1]}$$

where $\phi > 0$ and, for dimensional reasons:

$$F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right) = \frac{1}{m_0^{1-\Delta}} \hat{F}\left(\frac{1}{m_0^{1-\phi}} \frac{m}{\langle m_k\rangle^{\phi}}\right), \qquad \textbf{[S2]}$$

where m_0 is the minimum mass of an organism or a cutoff in the system. From this point onward, we will measure all masses in units of m_0; thus, F and $\hat{F}$ coincide and $\frac{\langle m\rangle_k}{m_0} \to \langle m\rangle_k$ is arbitrarily large.

In order for the distribution in Eq. **S1** to be normalized (i.e., $\int dm\, p_k(m) = 1$), one needs to make the following assumptions:

$F(x)$ approaches a constant when $x \ll 1$.
$F(x)$ goes to zero sufficiently fast when $x \gg 1$.

With these conditions, one has:

$$\begin{aligned} 1 &= \int_1^{\infty} dm\, p_k(m) = \int_1^{\infty} dm \frac{1}{m^{\Delta}} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right) = \\ &= \langle m\rangle_k^{\phi(1-\Delta)} \int_{\langle m\rangle_k^{-\phi}}^{\infty} dx \frac{1}{x^{\Delta}} F(x) = \\ &= \langle m\rangle_k^{(1-\Delta)\phi} \left[\int_{\langle m\rangle_k^{-\phi}}^{1} dx\, x^{-\Delta} + \int_1^{\infty} dx\, x^{-\Delta} F(x) \right] = \\ &= a + b\langle m\rangle_k^{(1-\Delta)\phi}, \end{aligned} \qquad \textbf{[S3]}$$

where a and b are constants. Now, except for corrections to the scaling [From Eq. **S3**: $p_k(m) = m^{-\Delta} F\left(\frac{m}{\langle m\rangle_k}{}^{\phi}\right)\left(a + b\langle m\rangle_k^{(1-\Delta)\phi}\right)^{-1} = m^{-\Delta}\left[\frac{1}{a} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right) - \frac{b}{a^2}\langle m\rangle_k^{(1-\Delta)\phi} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right) + \ldots\right]$, which adds an additional term to Eq. **S1**.], everything is consistent if $a = 1$, $(1 - \Delta)\phi < 0$, and $\Delta > 1$ ($\phi > 0$) or if $\Delta = 1$, in which case one has:

$$\begin{aligned} 1 &= \int_1^{\infty} dm\, p_k(m) = \int_1^{\infty} dm \frac{1}{m} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right) = \\ &= \int_{\langle m\rangle_k^{-\phi}}^{\infty} dx \frac{1}{x} F(x), \end{aligned} \qquad \textbf{[S4]}$$

and two possibilities arise: (i) $\int_0^{\infty} dx \frac{F(x)}{x} = 1$, with $F(x) \to 0$ sufficiently fast for $x \to 0$ (which is consistent with our data), and (ii) $F(x) \sim_{x\sim 0} (-\ln x)^{-\alpha}$, such that $\int_{\langle m\rangle_k}^{\infty} dx \frac{1}{x} F(x) \sim (\ln\langle m\rangle_k)^{\max(0,1-\alpha)}$. If $\alpha > 1$, one is back to case i, whereas if $\alpha < 1$, one has logarithmic corrections to the scaling. In fact, if $\alpha < 1$, one finds $p_k(m) = \frac{1}{m}(\ln\langle m\rangle_k)^{1-\alpha} F\left(\frac{m}{\langle m\rangle_k^{\phi}}\right)$, (i.e., a logarithmic correction to the scaling).

Successive Moments Ratios. A test for the validity of a scaling size distribution of the form $p_k(m) = \frac{1}{m} F\left(\frac{m}{\langle m\rangle_k}\right)$ is the proportionality of successive moments ratios $\langle m^j\rangle_k / \langle m^{j-1}\rangle_k$ ($j > 1$) to the first moment $\langle m\rangle_k$. In fact, if $p_k(m) = \frac{1}{m} F\left(\frac{m}{\langle m\rangle_k}\right)$, one has:

$$\begin{aligned} \frac{\langle m^j\rangle_k}{\langle m^{j-1}\rangle_k} &= \frac{\int dm\, m^j \frac{1}{m} F\left(\frac{m}{\langle m\rangle_k}\right)}{\int dm\, m^{j-1} \frac{1}{m} F\left(\frac{m}{\langle m\rangle_k}\right)} = \\ &= \frac{\langle m\rangle_k^j \int dx\, x^j \frac{1}{x} F(x)}{\langle m\rangle_k^{j-1} \int dx\, x^{j-1} \frac{1}{x} F(x)} \propto \langle m\rangle_k, \end{aligned} \qquad \textbf{[S5]}$$

where $x = \frac{m}{\langle m\rangle_k}$ and $j > 1$. In Fig. S1, we plot successive moments ratios calculated from our data [cultures in standard conditions (data are shown in Fig. 2; see also *Materials and Methods*)] and linear regressions on these data. The slopes of the linear regressions are compatible with the value of 1 (linear regressions on log-transformed data). Coefficient of determination R^2 values for the regressions are $R^2_{\langle m^2\rangle/\langle m\rangle} = 0.999$, $R^2_{\langle m^3\rangle/\langle m^2\rangle} = 0.996$, and $R^2_{\langle m^4\rangle/\langle m^3\rangle} = 0.988$. The same is shown for the data of *Chilomonas* sp., *Euglena gracilis*, and *Euplotes aediculatus* in different environmental conditions in Fig. S2 (data are shown in Fig. 3; see also *Materials and Methods*). The slopes of the linear regressions are compatible with the value of 1 (linear regressions on log-transformed data). Coefficient of determination R^2 values for the regressions are $R^2_{\langle m^2\rangle/\langle m\rangle} = 0.998$, $R^2_{\langle m^3\rangle/\langle m^2\rangle} = 0.987$, and $R^2_{\langle m^4\rangle/\langle m^3\rangle} = 0.962$.

Modeling Cell Growth and Division. To model growth and cell division, we studied the scaling properties of a simple model for these two processes (10), focusing our attention on unicellular organisms of an unspecified single species. Here, we show that the model produces stationary size distributions as in Eq. **1**.

Let $N(m, t)$ be the number of organisms of mass m at time t: A cell's mass grows exponentially in time with rate μ (i.e., $\dot{m} = \mu m$) and cell division occurs in time at a rate $b(m)$. Therefore, the fission rate depends on the mass of the cell [a mechanism known as sloppy size control (11)], and $b(m)dt$ is the probability that a cell of mass m divides in a time dt. We introduce a maximum possible size M for a cell [i.e., $N(m, t) = 0\ \forall m > M$], which requires (10):

$$\int_0^M dm\, b(m) = \infty. \qquad \textbf{[S6]}$$

Considering the balance of growth and division in an infinitesimal time interval dt and in the size interval $[m_1, m_2]$, expanding at first order in dt, one has:

$$\int_{m_1}^{m_2} dm \left(\frac{\partial N}{\partial t}(m,t) + \frac{\partial[\mu m N(m,t)]}{\partial m} + \right. \\ \left. + b(m)N(m,t) - 4b(2m)N(2m,t) \right) = 0. \qquad \textbf{[S7]}$$

The equation governing the balance of growth and cell division is then:

$$\frac{\partial N}{\partial t}(m,t) + \frac{\partial[\mu m N(m,t)]}{\partial m} + b(m)N(m,t) - \\ -4b(2m)N(2m,t) = 0. \qquad \textbf{[S8]}$$

Let us assume that $N(m, t) = \lambda(m)e^{kt}$ at stationarity, where $\lambda(m)$ is proportional to the stationary cell size distribution. Introducing $N(m, t) = \lambda(m)e^{kt}$ in Eq. **S8**, one finds for $\lambda(m)$:

$$\mu \frac{d[m\lambda(m)]}{dm} = -[k + b(m)]\lambda(m) + 4b(2m)\lambda(2m). \qquad \textbf{[S9]}$$

Integrating Eq. **S9** in $[0, M]$, one then has the following:

$$k\int_0^M dm\,\lambda(m) + \mu \int_0^M dm \frac{d[m\lambda(m)]}{dm} = \int_0^M dm\, b(m)\lambda(m). \qquad \textbf{[S10]}$$

Eq. **S10** imposes $\int_0^M dm\, b(m)\lambda(m) < \infty$; therefore, in Eq. **S6**, one has $\lim_{m\uparrow M}\lambda(m) = 0$, which implies $\int_0^M dm \frac{d[m\lambda(m)]}{dm} = 0$, and as a result:

$$k = \frac{\int_0^M dm\, b(m)\lambda(m)}{\int_0^M dm\,\lambda(m)}. \qquad \textbf{[S11]}$$

The scale invariance of $\lambda(m)$ can be deduced directly from Eq. **S9** as follows. The value of λ depends on m and M [i.e., $\lambda = \lambda(m, M)$]. We assume that the total mass present at $t = 0$ is equal to 1 (Eq. **S9** is linear in λ; therefore, if λ is a solution, so is $C\lambda$ with C as an arbitrary constant). Because M is the only scale in the problem, we assume that $b(m, M) = \hat{b}\left(\frac{m}{M}\right)$ and rewrite Eq. **S9** with $x = m/M$ as follows:

$$\mu \frac{d[x\lambda(Mx,M)]}{dx} = -\left[k + \hat{b}(x)\right]\lambda(Mx,M) + 4\hat{b}(2x)\lambda(2Mx,M), \qquad \textbf{[S12]}$$

where $x \in [0, 1]$. Therefore, one has the solution $\lambda(Mx, M) = \hat{\lambda}(x)/M$ [i.e., $\lambda(m, M) = \frac{1}{M}\hat{\lambda}\left(\frac{m}{M}\right)$], which satisfies:

$$\mu \frac{d\left[x\hat{\lambda}(x)\right]}{dx} = -\left[k + \hat{b}(x)\right]\hat{\lambda}(x) + 4\hat{b}(2x)\hat{\lambda}(2x) \qquad \textbf{[S13]}$$

in $x \in [0, 1]$, with $\int_0^1 dx\,\hat{\lambda}(x) = 1$ and $\hat{\lambda}(x) = 0\ \forall x > 1$. In particular, the size distribution $p(m)$ can be written as:

$$p(m) = \frac{1}{m}\frac{m}{M}\hat{\lambda}\left(\frac{m}{M}\right) = \frac{1}{m}G\left(\frac{m}{M}\right). \qquad \textbf{[S14]}$$

Computing the average mass $\langle m\rangle$, one finds that it is proportional to M:

$$\langle m\rangle = \int_0^M dm\, m\, p(m) = M\int_0^1 dy\, G(y) = cM, \qquad \textbf{[S15]}$$

such that the stationary size distribution is of the form:

$$p(m) = \frac{1}{m}F\left(\frac{m}{\langle m\rangle}\right), \qquad \textbf{[S16]}$$

which is precisely the scaling ansatz proposed and observed in the data.

The solution to Eq. **S13** can be written as $\lambda(x) = \lambda_n(x)$ with $2^{-n} \le x \le 2^{-n+1}$ for $n = 1, 2, \ldots$ and

$$\lambda_n(x) = e^{-\int_{2^{-n}}^{x} dy\, h(y)}\left[C_n + \frac{4}{\mu}\int_{2^{-n+1}}^{2x} dy\, \frac{\hat{b}(y)}{y}\lambda_{n-1}(y)\right], \qquad \textbf{[S17]}$$

where $\lambda_0(y) = 1$, $x \in [2^{-n}, 2^{-n+1}]$, $h(y) = [k + \hat{b}(y) + \mu]/y$, and C_0 depends on the normalization condition [i.e., $\int_0^1 dy\hat{\lambda}(y) = 1$]. The C_ns ($n \ge 1$) are determined recursively imposing the continuity $\hat{\lambda}_n(2^{-n+1}) = \hat{\lambda}_{n-1}(2^{-n+1})$ for $n \ge 2$. For instance, if $n = 2$:

$$\hat{\lambda}_1(x) = C_0 e^{-\int_{1/2}^{x} dy h(y)} \Rightarrow \hat{\lambda}_1\left(\frac{1}{2}\right) = C_0 \qquad \textbf{[S18]}$$

(note that $\hat{\lambda}_1(1) = 0$ due to the singularity in Eq. **S6**):

$$\hat{\lambda}_2(x) = e^{-\int_{1/2}^{x} dy\, h(y)} \left[C_1 + + \frac{4}{\mu} C_0 \int_{1/2}^{2x} dy \frac{\hat{b}(y)}{y} e^{-\int_{1/2}^{y} dz\, h(z)} \right], \qquad \textbf{[S19]}$$

and $\hat{\lambda}_2(1/2) = \hat{\lambda}_1(1/2) = C_0$ allows one to compute C_1/C_0. Iterating, one can compute $C_n/C_0\ \forall n \geq 1$.

Analytical Form of the Universal Size Distribution. A fitting procedure suggests the viability of an analytical log-normal form for the universal size distribution; that is,

$$p(m) = \frac{1}{m\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln \frac{m}{\langle m \rangle} - \mu\right)^2}{2\sigma^2}}, \qquad \textbf{[S20]}$$

where σ^2 and μ are constants, $\langle m \rangle$ depends on the species, and ln is the natural logarithm (i.e., the logarithm to the base e). In order for the distribution in Eq. **S20** to have the scaling form $p(m) = 1/mF(m/\langle m \rangle)$, one has to impose that $\langle m \rangle = \int_0^\infty dm\, mp(m)$, which implies $\mu = -\sigma^2/2$ (i.e., μ and σ are not independent). We thus propose the following analytical form for the universal size distribution, which depends on only one parameter, σ^2:

$$p(m) = \frac{1}{m\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln \frac{m}{\langle m \rangle} + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}}. \qquad \textbf{[S21]}$$

The scaling function $F(x)$ is therefore of the form:

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln x + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}}, \qquad \textbf{[S22]}$$

as suggested by the fact that a parabola fits the log-transformed data $mp(m)$ vs. $m/\langle m \rangle$ (least-squares fit on log-transformed data) well. To have a good estimate of the mean of F, we performed the fit to Eq. **S22** in the common support of at least half of the protist species, with only one parameter, σ^2. The best estimate for the parameter is $\sigma^2 = 0.222(3)$, and the coefficient of determination is $R^2 = 0.92$. The fit of the scaling function is shown in Fig. S3*A*, superimposed on the measured single-species size distributions. Fig. S3*B* compares the fit of a log-normal scaling function with the ensemble average of the experimental size distributions, showing a remarkable overlap.

We now turn to an illustration of how the postulated universal scaling form for $p(m)$ might arise from general dynamical considerations (9, 12, 13). Consider ecosystem dynamics over ecological time scales. Ecological processes governing the abundances and niche occupancy of species are expected to change their characteristic size. One would expect, however, that offspring would have a mass proportional to the mass of the parent organism (13). Thus, fluctuations in size within same species ought to be measured on the order of percent variations, and the natural variable is $x = \log(m/\overline{m})$ (12, 13), with $\overline{m}$ being the characteristic mass of the reference species k (e.g., proportional to the mean $\overline{m} = \alpha\langle m \rangle$) (9, 14). In this framework, the results of ecological processes can be represented by a random walk in the variable x, because a fixed percent increase (decrease) of the mass corresponds to a shift to the right (left) of the variable x by a constant amount. In the simplest model, the results of ecological processes could be represented by an Ornstein–Uhlenbeck process (15, 16). This process is a modification of a Wiener process, where the walk tends to move toward a central location, which, in the context of phenotypic evolution (12), has been identified as the optimum in the adaptive zone for the phenotype. A phenotypic character like body size, therefore, is expected to be distributed around a fitness optimum in this framework. The physical analogy of this process is a noisy relaxation process (e.g., a spring fluctuating around its rest length in the presence of disturbances). The fraction of organisms $q(x, t)$ with a (log-)body mass x at time x is governed by the dynamical equation:

$$\frac{\partial q(x,t)}{\partial t} = D \frac{\partial^2 q(x,t)}{\partial x^2} + \frac{\partial [kxq(x,t)]}{\partial x}, \qquad \textbf{[S23]}$$

where k is a constant, $k > 0$. The stationary solution is obtained by setting the rate of change of $q(x, t)$ to zero, and it is known to be Gaussian (15, 16):

$$q(x) = \sqrt{\frac{k}{\pi D}} e^{-\frac{kx^2}{2D}}. \qquad \textbf{[S24]}$$

For the mass distribution, one then obtains $p(m)$:

$$p(m) = q(x) \frac{dx}{dm} = \frac{1}{m} \sqrt{\frac{k}{2\pi D}} e^{-\frac{k}{2D}\left(\ln \frac{m}{\langle m \rangle} - \ln\alpha\right)^2} \qquad \textbf{[S25]}$$

(i.e., a log-normal distribution of mass). Imposing Eq. **S25** to have mean $\langle m \rangle$, one finds $\alpha = \exp[-D/(2k)]$, and the distribution of size is therefore:

$$p(m) = q(x) \frac{dx}{dm} = \frac{1}{m} \sqrt{\frac{k}{2\pi D}} e^{-\frac{k}{2D}\left(\ln \frac{m}{\langle m \rangle} + \frac{D}{2k}\right)^2}, \qquad \textbf{[S26]}$$

(i.e., a log-normal distribution of mass with mean $\langle m \rangle$ as in Eq. **S21**). Therefore, the scaling function F in Fig. 2 and Fig. S3 is $F(x) = \sqrt{\frac{k}{2\pi D}} \exp\left[-\frac{k}{2D}\left(\ln x + \frac{D}{2k}\right)^2\right]$.

One might wonder whether the size distribution obtained in Eq. **S26** is in agreement with our model of cellular growth and division. The cellular growth and division model, as treated in the previous section, assumes the existence of a maximum mass M and allows one to study the scaling properties of the stationary size distribution. It is possible, however, to relax this hypothesis, allowing the cells to assume all masses in the range $[0, \infty]$ and to obtain an implicit relation for the stationary size distribution $p(m)$ (17), which allows one to compute the asymptotic behavior of the distribution for large mass (i.e., $m \to \infty$). In the notation of the previous section, the size distribution for large m satisfies the relation:

$$p(m) \to \frac{1}{m} \exp\left(- \int_{\alpha}^{m/\langle m \rangle} dy \frac{k + d(y)}{\mu y} \right), \qquad \textbf{[S27]}$$

which behaves as a log-normal distribution if we further assume a division rate $d(y)$ increasing logarithmically with size [i.e., $d(y) \propto \ln y$].

On the Goodness of Data Collapses. Data collapse is a tool widely used in statistical physics to establish scaling laws and extract information on their exponents (18). Traditionally, the procedure to produce a data collapse is to rely on the direct visualization of it and on "eyeballing" the exponent that gives the best collapse. A less subjective method has been proposed (19), which introduces a measure (error functional E; Figs. 2 and 3, *Insets*) to quantify the goodness of a collapse. Let Δ be the exponent that we tune to find the best collapse: $E(\Delta)$ is the cumulative area enclosed between all pairs of curves that we try to collapse, within their common support, for the value Δ of the exponent. The value Δ^* of the exponent, which minimizes $E(\Delta)$, is taken as the best estimate for the exponent (i.e., the smaller the area, the better is the collapse). Errors are associated with the determination of Δ^* and are obtained from the width of the minimum. Further details can be found in the study by Bhattacharjee and Seno (19).

Community Size Spectra. In this section, we show how a power law community-size spectrum arises as a sum of single-species size distributions of finite-size scaling form (14). Assume that the size distribution of species k is of the form:

$$p_k(m)=\frac{1}{m}F\left(\frac{m}{\langle m\rangle_k}\right). \tag{S28}$$

Let N_k be the stationary abundance of species k in an ecosystem [i.e., $N_k = N_k(t \to \infty)$] and S be the total number of species. The community size spectrum is defined as:

$$f(m)=\sum_{k=1}^{S} N_k p_k(m)\Big/\sum_{k=1}^{S} N_k. \tag{S29}$$

We assume, supported by a number of observations (20), that the population abundance of the kth species scales as:

$$N_k \propto \langle m\rangle_k^{\alpha}, \tag{S30}$$

where $\alpha < 0$ implies that the total number of organisms decreases with increasing typical size. From Eqs. **S28** and **S30**, one has that:

$$f(m)\propto\sum_{k=1}^{S} N_k\, p_k(m)\propto\sum_{k=1}^{S}\langle m\rangle_k^{\alpha} m^{-1} F\left(\frac{m}{\langle m\rangle_k}\right). \tag{S31}$$

Let $g(\overline{m})$ be the fraction of species of typical size $\overline{m}$. The above equation can be rewritten, treating $\langle m\rangle_k$ as a continuous variable for easiness of computation, as:

$$f(m)\propto\frac{1}{m}\int d\overline{m}g\left(\overline{m}\right)\overline{m}^{\alpha}F\left(\frac{m}{\overline{m}}\right)\propto\int dx\, g(xm)\, x^{\alpha}m^{\alpha}F\left(\frac{1}{x}\right). \tag{S32}$$

Theoretical predictions from a scaling macroecological framework (9, 14) and considerations on the total number of species on Earth (21, 22) suggest a pure power-law behavior for $g(\overline{m})$:

$$g\left(\overline{m}\right)\propto\frac{1}{\overline{m}^{\beta}}, \tag{S33}$$

which, because of normalization, is assumed to hold between an upper cutoff and a lower cutoff. One then has for the size spectrum $f(m)$:

$$f(m)\propto m^{\alpha-\beta}\int dx\, x^{\alpha-\beta}F\left(\frac{1}{x}\right)\propto m^{\alpha-\beta}, \tag{S34}$$

which has the form of a power law. Note, however, that the result still holds for log-normal species abundance distributions (22). In the case of a limited range of sizes, one might argue that the number of species S within the range of sizes investigated could be assumed as constant to first order. This, of course, is the particular case for which $\beta = 0$.

Overall, it is clear that to obtain a scaling community size spectrum (Eq. **S29**), a necessary condition is adaptive fine-tuning of the specific abundances. This is epitomized by the relation in Eq. **S30**, which, in turn, implies the thinning relations that are recurrent in the literature of macroecological empirical laws (9, 23).

Equal Biomass in Each Size Class. The case of $f(m) \propto m^{-2}$, which is routinely found in the literature (14, 24), is of special interest because it agrees with the assumption of constant biomass in each size class (25). In this case, the total mass in a range $(m, m + \Delta m)$ with $\Delta m/m \simeq 1$ is:

$$\int_{m}^{m+\Delta m} dx\, x f(x)\propto\log\left[1+\frac{\Delta m}{m}\right]\simeq 1, \tag{S35}$$

such that the total biomass in a range between m and $m + \Delta m$, where Δm is the typical variance at scale m, is independent of m.

***Bodo saltans* Size Distribution.** In Fig. 2, the size distribution of *Bodo saltans* (BOD) might seem not to collapse as well as that of the other species. This is due to an instrumentation limit and is not inherent to the BOD size distribution. BOD, in fact, has a mean volume of 17 μm^3 and a mean equivalent diameter of 3.1 μm^3. These values lie on the leftmost side of the size spectrum, where the debris peak is dominant and the instrument hardly resolves the peak of the protist culture (Fig. S4, *Inset*). As an effect of this, it is hard to separate the protist peak from the debris; consequently, the left side of the BOD size distribution is overestimated, causing a deviation from the other collapsing curves in Fig. 2. To show that this is indeed the case, we measured the size distribution of *Chlorogonium euchlorum* (CHO) with the 150-μm capillary of the Cell Counter and Analyzer System (CASY) model TTC (Roche Applied Science), whereas we used the 60-μm CASY capillary to obtain the data reported in the main text. Using the larger capillary, the protist and the debris peaks appear closer to each other (with there being fewer size channels in the [0, 20]-μm interval; Fig. S5, *Inset*), and as a result, the size of the CHO size distribution on the left is overestimated. If we plot the corresponding rescaled distribution, together with the other data of Fig. 2, we observe the same kind of deviation observed for BOD (Fig. S5). We argue, therefore, that if the instrumentation could resolve better the small-sized region of the BOD size distribution, it would collapse even better than it does in Fig. 2.

1. Mandelbrot B (1973) *The Fractal Geometry of Nature* (Freeman, San Francisco).
2. Rodriguez-Iturbe I, Rinaldo A (1997) *Fractal River Basins: Chance and Self-Organization* (Cambridge Univ Press, New York).
3. Bascompte J, Solé RV (1995) Rethinking complexity: Modelling spatiotemporal dynamics in ecology. *Trends Ecol Evol* 10(9):361–366.
4. Solé RV, Manrubia SC, Benton M, Kauffman S, Bak P (1999) Criticality and scaling in evolutionary ecology. *Trends Ecol Evol* 14(4):156–160.
5. Bak P (1997) *How Nature Works. The Science of Self-Organized Criticality* (Copernicus–Springer, Berlin).
6. Chisholm SW (1992) Phytoplankton size. *Primary Productivity and Biogeochemical Cycles in the Sea*, eds Falkowski PG, Woodhead AD (Plenum, New York), pp 213–237.
7. Fisher ME, Barber MN (1972) Scaling theory for finite-size effects in critical region. *Phys Rev Lett* 28(23):1516–1519.
8. Banavar JR, Green JL, Harte J, Maritan A (1999) Finite size scaling in ecology. *Phys Rev Lett* 83(20):4212–4214.
9. Banavar JR, Damuth J, Maritan A, Rinaldo A (2007) Scaling in ecosystems and the linkage of macroecological laws. *Phys Rev Lett* 98(6):068104.
10. Diekmann O, Lauwerier HA, Aldenberg T, Metz JAJ (1983) Growth, fission and the stable size distribution. *J Math Biol* 18(2):135–148.
11. Tyson J, Diekmann O (1986) Sloppy size control of the cell-division cycle. *J Theor Biol* 118(4):405–426.
12. Lande R (1976) Natural-selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314–334.
13. Clauset A, Erwin DH (2008) The evolution and distribution of species body size. *Science* 321(5887):399–401.
14. Rinaldo A, Maritan A, Cavender-Bares KK, Chisholm SW (2002) Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. *Proc Biol Sci* 269(1504):2051–2059.
15. van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam).

16. Uhlenbeck GE, Ornstein LS (1930) On the theory of the Brownian motion. *Phys Rev* 36(5):823–841.
17. Rading MM, Engel TA, Lipowsky R, Valleriani A (2011) Stationary size distributions of growing cells with binary and multiple cell division. *J Stat Phys* 145(1):1–22.
18. Barenblatt GI (1996) *Introduction to Phase Transitions and Critical Phenomena* (Cambridge Univ Press, Cambridge, UK).
19. Bhattacharjee SM, Seno F (2001) A measure of data collapse for scaling. *J Phys A Math Gen* 34(33):6375–6380.
20. Damuth J (1981) Population-density and body size in mammals. *Nature* 290(5808): 699–700.
21. May RM (1988) How many species are there on Earth? *Science* 241(4872):1441–1449.
22. Fenchel T (1983) There are more small than large species. *Oikos* 68(2):375–378.
23. Southwood TR, May RM, Sugihara G (2006) Observations on related ecological exponents. *Proc Natl Acad Sci USA* 103(18):6931–6933.
24. Camacho J, Sole RV (2001) Scaling in ecological size spectra. *Europhys Lett* 55(6): 774–780.
25. Rodriguez J, Mullin MM (1986) Relation between biomass and body-weight of plankton in a steady-state oceanic ecosystem. *Limnol Oceanogr* 31(2):361–370.

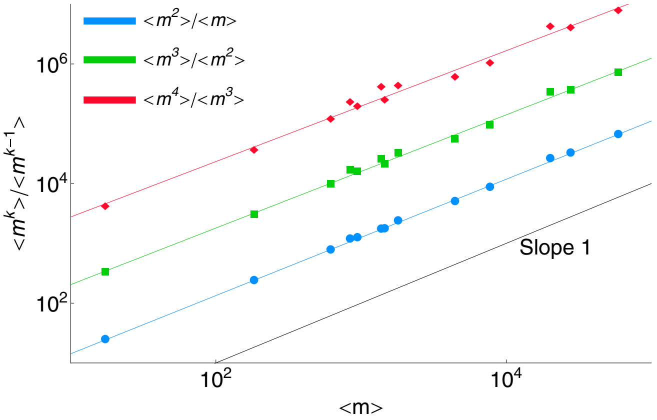

Fig. S1. Ratios of successive moments of m are proportional to $\langle m\rangle$. Moments were calculated from size distributions in standard conditions (i.e., calculated from the size distributions shown in Fig. 2 and *Materials and Methods*).

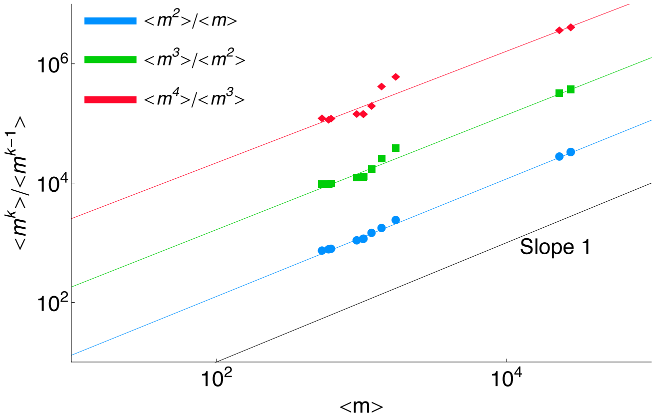

Fig. S2. Ratios of successive moments of m are proportional to $\langle m\rangle$. Moments were calculated from size distributions in nonstandard conditions (i.e., calculated from the size distributions shown in Fig. 3 and *Materials and Methods*).

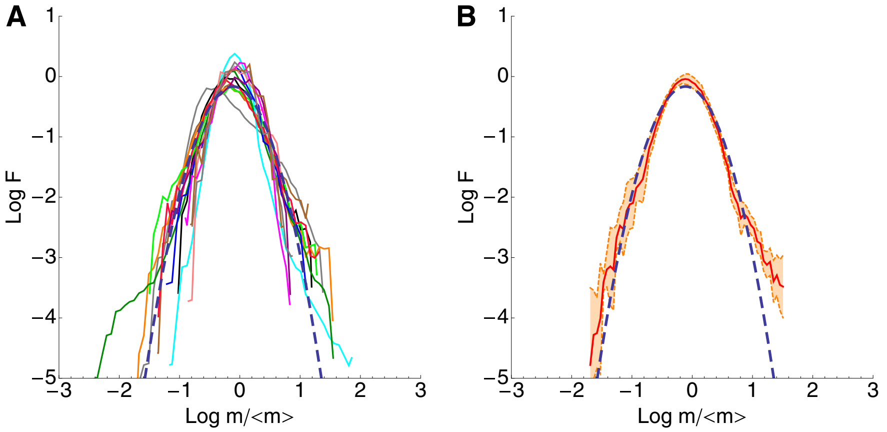

Fig. S3. (*A*) Same collapse of rescaled size distributions as in Fig. 2, with the best log-normal fit superimposed (dashed blue line). The scaling function F is a parabola in a log-log plot; thus, the universal size distribution *p(m)* is log-normal. Colors are as in Figs. 1 and 2. (*B*) Average of all curves in *A* (red curve) shows a remarkable resemblance of the best log-normal fit (dashed blue line). The orange region is the 99.7% confidence interval around the average.

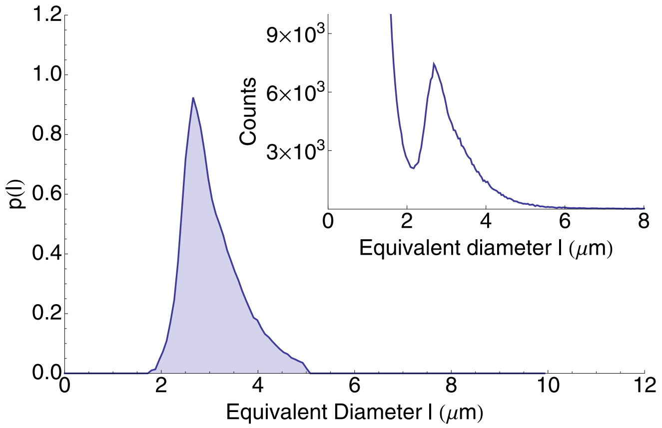

Fig. S4. Size distribution of BOD as a function of the equivalent diameter. (*Inset*) Output of the CASY instrument for a BOD culture. The protist peak is close to the debris peak on the leftmost side of the spectrum.

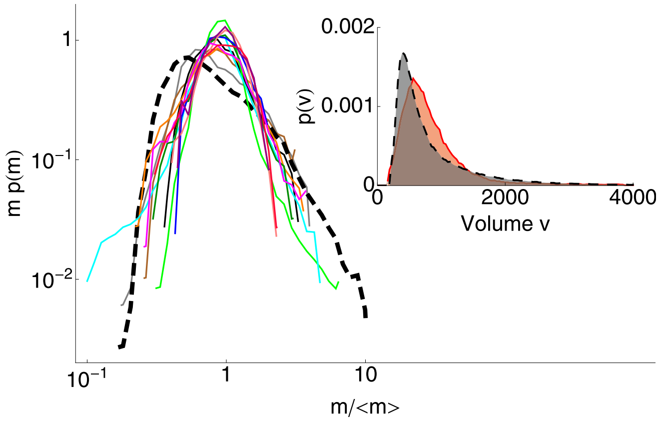

Fig. S5. Collapse of rescaled size distributions (as in Fig. 2), where the curve of CHO was measured with a large capillary (150 μm, black dashed line). The size of the distribution on the left is overestimated and causes the collapse to fail. Colors are as in Figs. 1 and 2. (*Inset*) CHO size distributions measured with the 60-μm capillary (red curve) and with the 150-μm capillary (black dashed line). The 60-μm capillary correctly resolves the whole spectrum, whereas the 150-μm capillary overestimates the left side of the distribution.

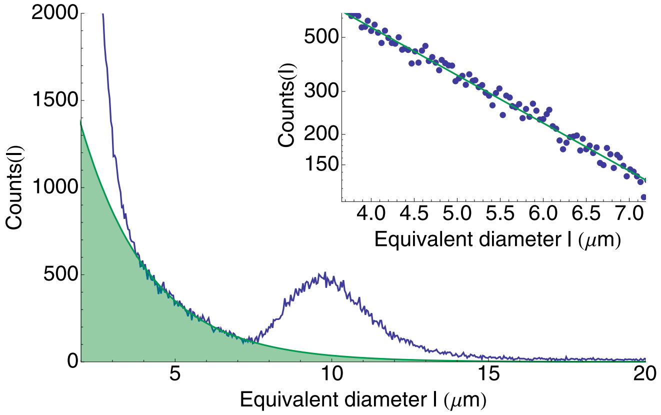

Fig. S6. Output of a measurement for *Chilomonas* sp. (CHI) culture. The superimposed exponential decay is the exponential fit of the debris in the region adjacent to the protist peak. (*Inset*) Decay of the debris to the left of the protist peak is exponential (straight line in a log-linear plot).

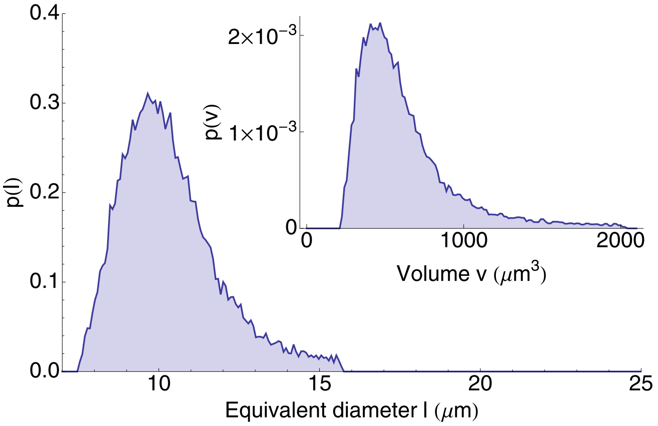

Fig. S7. Size distribution of *Chilomonas* sp. (CHI) as a function of the equivalent diameter. (*Inset*) Transformed volume size distribution of CHI.