Supplementary Information: The possible modes of microbial reproduction are fundamentally restricted by distribution of mass between parent and offspring

Yuriy Pichugin[a,b] and Arne Traulsen[a]

[a]Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany

[b]Current address: Department of Ecology and Evolutionary Biology, Princeton University, 08544 Princeton, New Jersey, USA

Contents

1 Population dynamics of the model

We consider a population of organisms where the mass m of each organism grows at rate $g(m)$,

$$\frac{dm}{dt} = g(m). \tag{1}$$

A population of organisms can be described by the organism density $n(m,t)$ of organisms of mass m at time t. The dynamics of $n(m,t)$ is described by the von Foerster equation [1]

$$\frac{\partial}{\partial t}n(m,t) = -\frac{\partial}{\partial m}\left(g(m)n(m,t)\right), \tag{2}$$

which is also known as transport equation. This is a non-homogeneous quasilinear differential equation of first order, which can be solved by the method of characteristics. Applying the product rule and rewriting this in a standard form [2], we get

$$\frac{\partial}{\partial t}n(m,t) + g(m)\frac{\partial}{\partial m}n(m,t) = -\frac{dg(m)}{dm}n(m,t). \tag{3}$$

The nature of the method of characteristics is to find curves in the space (m,t,n), along which Eq. (3) becomes an ordinary differential equation. Since Eq. (3) contains differentiation with respect to two variables m and t, there are two characteristic curves: C_1 and C_2. Once they are found, the solution of differential equation is given by a union of characteristics $\Phi(C_1, C_2) = 0$, or in other terms $C_2 = F(C_1)$, such that the function F satisfies the initial and boundary conditions of the Cauchy problem.

The characteristics equations for Eq. (3) are

$$\frac{dt}{d\alpha} = 1, \tag{4}$$

$$\frac{dm}{d\alpha} = g(m), \tag{5}$$

$$\frac{dn}{d\alpha} = -\frac{dg(m)}{dm}n(m,t). \tag{6}$$

Combining Eqs. (4) with (5) and (5) with (6), we get a system of ordinary differential equations

$$\frac{dm}{dt} = g(m), \tag{7}$$

$$\frac{dn}{dm} = -\frac{1}{g(m)}\frac{dg(m)}{dm}n(m,t), \tag{8}$$

where Eq. (7) is identical to Eq. (1) because the growth dynamics of a single organism depends only on a single parameter t, i.e. is already a characteristic of Eq. (3).

The solutions of the characteristic equations are

$$t - T(m) = C_1, \tag{9}$$

$$g(m)n(m,t) = C_2, \tag{10}$$

where C_1 and C_2 are constants and we introduced a variable

$$T(m) = \int_1^m \frac{dx}{g(x)}, \tag{11}$$

such that the time it takes for an organism to grow from mass m_0 to m is $T(m_0, m) = T(m) - T(m_0)$. Note that if $g(1) = 0$ the integral may diverge for any m. However, as we show later, all observable results depend only on differences in a form $T(m) - T(m_0)$, which do not depend on $g(1)$. In other words, even if an organism of minimal viable mass is unable to grow, it does not affect the life cycles in which no organism ever experiences that minimal mass.

The general solution of Eq. (3) is given by $C_2 = F(C_1)$, so the organism density can be explicitly found as

$$n(m,t) = \frac{1}{g(m)} F(t - T(m)), \tag{12}$$

where $F(x)$ is a function determined by initial and boundary conditions.

After a sufficiently long time, the demographic distribution of the population becomes stationary. There, the organism density $n(m,t)$ can be factorised into a stationary demographic distribution $\rho(m)$ and growing population size $N(t)$, $n(m,t) = \rho(m)N(t)$. The demographic and time variables can be separated in Eq. (12), only if the function $F(x)$ is exponential, $F(x) = Ae^{\lambda x}$. Thus, for the stationary demography regime we find

$$n(m,t) = \frac{A}{g(m)} e^{\lambda(t - T(m))}, \tag{13}$$

and consequently

$$\begin{aligned} \rho(m) &= \frac{A}{g(m)} e^{-\lambda T(m)}, \\ N(t) &= e^{\lambda t}. \end{aligned} \tag{14}$$

The obtain the dynamics of population, we need to combine the general solution in Eq. 14 with the boundary conditions. These conditions are given by the fragmentation process. The fragmentation process removes organisms that reach mass m^* from the population and gives rise to organisms of masses $m_1, m_2, \ldots, m_z$. Therefore, the density distribution is not continuous at these points. At $m = m^*$, all organisms fragment, so $\rho(m > m^*) = 0$. At $m = m_1$, the smallest possible organisms are born, so $\rho(m < m_1) = 0$. At $m = m_i$, $i = \{1, \ldots, z\}$ offspring organisms are born, so the distribution $\rho(m)$ has a step, with a magnitude we define as $\Delta\rho_{m_i} = \rho(m_i + \epsilon) - \rho(m_i - \epsilon)$, see Fig SI 1. The values of steps can be inferred from the flow continuity: the number of organisms resulting from a fragmentation at a given moment is equal to the number of offspring organisms produced

$$\rho(m^*) g(m^*) = k_i \Delta\rho_{m_i} g(m_i), \tag{15}$$

where k_i is the number of organisms of mass m_i produced in the result of fragmentation.

At the same time, the value of $\rho(m^*)$ follows Eq. (14) taking into account all steps:

$$\rho(m^*) = \Delta\rho_{m_1} \frac{g(m_1)}{g(m^*)} e^{-\lambda T(m_1, m^*)} + \ldots + \Delta\rho_{m_z} \frac{g(m_z)}{g(m^*)} e^{-\lambda T(m_z, m^*)}. \tag{16}$$

Therefore, combining Eqs. (15) and (16), we get

$$\begin{aligned} 1 &= \sum_{k=1}^{z} e^{-\lambda T(m_k, m^*)}, \text{ where} \\ T(m_k, m^*) &= \int_{m_k}^{m^*} \frac{dx}{g(x)} \end{aligned} \tag{17}$$

and the equilibrium distribution of organism masses is

$$\rho(m) = \sum_i \rho(m_i) \frac{g(m_i)}{g(m)} e^{-\lambda T(m_i, m)}. \tag{18}$$

As a final remark, we would like to discuss the special case, where organisms may achieve a state with zero growth during the life cycle: $g(m_{\text{stop}}) = 0$ and $\min(m_i) \leq m_{\text{stop}} \leq m^*$. In such a scenario, any organism reached m_{stop} cannot grow beyond this size and its life cycle takes infinite time to complete – the integral $T(m_k, m^*)$ diverges. Hence the corresponding terms in the sum in Eq. 17 become zero. If at least one offspring is born to be larger than m_{stop}, then Eq. 17 has at least one term not equal to zero and can be solved normally, while the offspring with an arrested growth are discarded. If all offspring are born smaller than m_{stop} and the number of offspring is finite, then the population cannot grow and the population growth rate $\lambda = 0$. If all offspring are born smaller than m_{stop} and the number of offspring is infinite, mathematically, the model results are ambiguous (sum of infinite zero terms). However, the system itself has no biological meaning at that point.

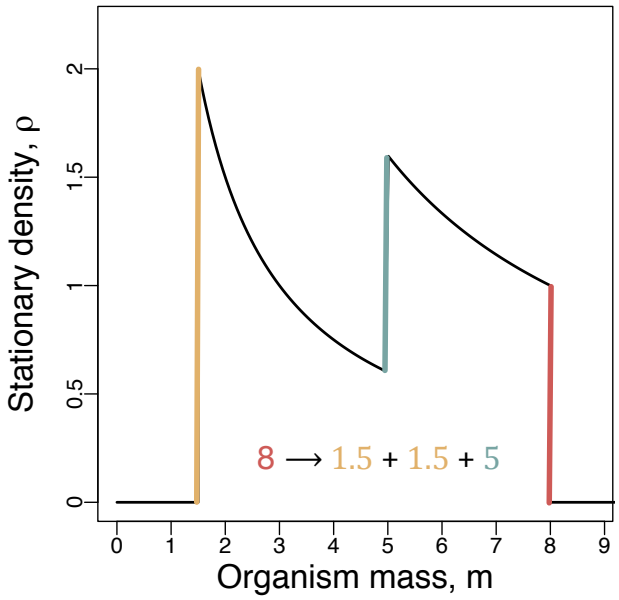

Figure SI 1: **Example of the organism density distribution**
The life cycle of population is given by $\mathcal{M} = \{1.5, 1.5, 5\}$. Hence, the distribution is piecewise continuous, interrupted at points $m = 1.5, 5, 8$. In this example $g(m) = 1$ and $T(m_0, m) = m - m_0$, so intervals of continuity demonstrate exponential decay. Since two organisms of mass 1.5 are produced at each fragmentation, the step there is twice as large as the step at mass 5, which describes a single offspring organism.

2 If mass is conserved, the evolutionarily optimal life cycle is always a binary fragmentation

In this section, we show that when the mass of the parent is equal to the combined mass of offspring, then for any life cycle with more than two offspring, it is possible to find another life cycle with a smaller number of offspring and a larger or equal growth rate. For that, we consider three life cycles

1. $\mathcal{M}_1 = \{m_1, m_2, \ldots, m_z\}$ with z offspring of the combined mass $m^* = \sum_{i=1}^{z} m_i$.
2. $\mathcal{M}_2 = \{m_1 + m_2, m_3, \ldots, m_z\}$ with $z - 1$ offspring of the same combined mass m^*.
3. $\mathcal{M}_3 = \{m_1, m_2\}$ with 2 offspring of the combined mass $m_1 + m_2$.

The growth rates of these life cycles are given by the solution of

$$\begin{aligned} f_1(\lambda) &= 1 - e^{-\lambda T(m_1, m^*)} - e^{-\lambda T(m_2, m^*)} - \sum_{i=3}^{z} e^{-\lambda T(m_i, m^*)} = 0, \\ f_2(\lambda) &= 1 - e^{-\lambda T(m_1+m_2, m^*)} - \sum_{i=3}^{z} e^{-\lambda T(m_i, m^*)} = 0, \\ f_3(\lambda) &= 1 - e^{-\lambda T(m_1, m_1+m_2)} - e^{-\lambda T(m_2, m_1+m_2)} = 0. \end{aligned} \tag{19}$$

The expressions $f_1(\lambda), f_2(\lambda), f_3(\lambda)$ are monotonically increasing functions of λ. At $\lambda = 0$, their values are all negative due to $z > 2$: $f_1(0) = 1 - z$, $f_2(0) = 2 - z$, and $f_3(0) = -1$. At $\lambda \to \infty$, these expressions approach 1. Therefore, each of the Eqs. (19) has a single root. We need to show that λ_1 – the root of $f_1(\lambda)$ – is located between roots of $f_2(\lambda)$ and $f_3(\lambda)$. In that case, one of these two life cycles has a faster growth rate.

Since $e^{-\lambda(T(x,y)+T(y,z))} = e^{-\lambda T(x,z)}$ and the multiplication of the expression for $f_3(\lambda)$ by $e^{-\lambda T(m_1+m_2,m^*)}$ does not change its root, the last equation in (19) is equivalent to

$$\tilde{f}_3(\lambda) = f_3(\lambda)e^{-\lambda T(m_1+m_2,m^*)} = e^{-\lambda T(m_1+m_2,m^*)} - e^{-\lambda T(m_1,m^*)} - e^{-\lambda T(m_2,m^*)} = 0. \tag{20}$$

Then, $f_1(\lambda) = f_2(\lambda)+\tilde{f}_3(\lambda)$, and therefore $f_1(\lambda_1) = 0 = f_2(\lambda_1)+\tilde{f}_3(\lambda_1)$. Hence, either $f_2(\lambda_1) < 0$, or $\tilde{f}_3(\lambda_1) < 0$, or $f_2(\lambda_1) = \tilde{f}_3(\lambda_1) = 0$.

- If $f_2(\lambda_1) < 0$, then by Bolzano's theorem, the root of $f_2(\lambda)$ is larger than λ_1. In other words, the life cycle $\mathcal{M}_2$ provides larger growth rate than $\mathcal{M}_1$.
- If $\tilde{f}_3(\lambda_1) < 0$, then by Bolzano's theorem, the root of $\tilde{f}_3(\lambda)$ and thus the root of $f_3(\lambda)$ is larger than λ_1. Consequently, the life cycle $\mathcal{M}_3$ has larger growth rate than $\mathcal{M}_1$.
- If $f_2(\lambda_1) = \tilde{f}_3(\lambda_1) = 0$, then all three roots coincide and all three considered life cycles have the same growth rate.

This proves our statement.

3 Evolutionarily neutral environments

Here, we show that at $g(m) = m$, all possible life cycles share the same population growth rate if the mass is conserved in the act of fragmentation $\sum_{k=1}^{z} m_i = m^*$. For an arbitrary life cycle $\mathcal{M} = \{m_1, m_2, \ldots, m_z\}$, its population growth rate λ is given by the solution of

$$1 - \sum_{k=1}^{z} e^{-\lambda T(m_k,m^*)} = 0, \tag{21}$$

where

$$T(m_i, m^*) = \int_{m_i}^{m^*} \frac{dy}{g(y)} = \int_{m_i}^{m^*} \frac{dy}{y} = \ln\left(\frac{m^*}{m_i}\right). \tag{22}$$

The expression on the left-hand side of Eq. (21) is continuous and monotonic in λ, so only a single solution of this equation exists. Plugging Eq. (22) into (21), we get

$$1 - \sum_{k=1}^{z} \left(\frac{m_i}{m^*}\right)^{\lambda} = 0. \tag{23}$$

The value $\lambda = 1$ solves this equation, as it makes it equivalent to the mass conservation $\sum_{k=1}^{z} m_i = m^*$, which must be satisfied.

Therefore, under mass conservation, at $g(m) = m$, the population growth rate of an arbitrary life cycle is $\lambda = 1$.

4 Michaelis-Menten-Monod kinetics of cell growth

Here we consider the model of the cell growth based on Michaelis-Menten-Monod kinetics. In this model, the substrate S comes by adsorbing the nutrients in environment U through the cell wall, and either reversibly binds with free enzymes E forming complexes C, or decays / escapes back to the environment. Substrate-enzyme complexes C either break apart into components or are processed into the product biomass P, releasing the free enzyme. These processes can be written in the form of reactions

$$\emptyset \underset{e}{\overset{F(m)}{\rightleftharpoons}} S \tag{24}$$

$$S + E \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C \xrightarrow{k_0} P + E, \tag{25}$$

where e is the substrate decay rate (either by degradation or by leaving the cell), $F(m)$ is the rate of substrate acquisition proportional to the cell surface area, k_+ and k_- are the rates of direct and reverse reactions of the substrate binding, and k_0 is the rate of substrate processing. The product of the reaction P is the gained biomass Δm.

Following the original analysis of Michaelis and Menten, we assume that the dynamics of the substrate is fast, so at any time, the substrate is in chemical equilibrium with the complexes, i.e., $k_0 \ll k_-$. Then, the dynamics of cell growth occurs at a much slower rate than the dynamics of compounds inside the cell. The fast dynamics of compounds in a cell of mass m is described by the set of differential equations

$$\begin{aligned}\frac{dS}{dt} =&\frac{F(m)}{m} - eS - k_+SE + k_-C,\\ \frac{dE}{dt} =&(k_0 + k_-)C - k_+SE,\\ \frac{dC}{dt} =&-(k_0 + k_-)C + k_+SE = -\frac{dE}{dt},\end{aligned} \tag{26}$$

where S, E, and C represent the concentrations of substrate, free enzymes, and substrate-enzyme complexes in the cell. The slow dynamics of the cell growth is given by the rate of biomass production. The productivity – the rate of biomass gain per mass unit – is equal to the rate of the product concentration increase

$$\frac{g(m)}{m} = \frac{dP}{dt} = k_0C. \tag{27}$$

The system of differential equations (26) converges to a stationary regime, where concentrations of S, E and C are constant. To find these concentrations, we first note that at the fast time scale, the combined amount of free and bounded enzymes is preserved, $\frac{d}{dt}(C + E) = 0$. Thus, we can set $C + E = \frac{E_0(m)}{m}$, where $E_0(m)$ is the total amount of enzymes present in a cell of mass m. The stationary regime is described by

$$\begin{aligned}&\frac{F(m)}{m} - eS - k_+S\left(\frac{E_0(m)}{m} - C\right) + k_-C = 0,\\ &(k_0 + k_-)C - k_+S\left(\frac{E_0(m)}{m} - C\right) = 0.\end{aligned} \tag{28}$$

The solution of these equations is

$$\begin{aligned}S =&\frac{1}{2e}\left(\frac{F(m)}{m} - k_e - k_0\frac{E_0(m)}{m} + \sqrt{\left(k_e + k_0\frac{E_0(m)}{m} - \frac{F(m)}{m}\right)^2 + 4k_e\frac{F(m)}{m}}\right),\\ C =&\frac{1}{2k_0}\left(\frac{F(m)}{m} + k_e + k_0\frac{E_0(m)}{m} - \sqrt{\left(k_e + k_0\frac{E_0(m)}{m} - \frac{F(m)}{m}\right)^2 + 4k_e\frac{F(m)}{m}}\right),\end{aligned} \tag{29}$$

where we introduced $k_e = \frac{k_- + k_0}{k_+}e$. The parameter k_e characterizes the effectiveness of substrate utilization (higher k_e means a lower amount of free substrate in the cell).

For $k_0 = 0$, we find for S

$$\begin{aligned}S|_{k_0=0} =&\frac{1}{2e}\left(\frac{F(m)}{m} - k_e + \sqrt{\left(k_e - \frac{F(m)}{m}\right)^2 + 4k_e\frac{F(m)}{m}}\right)\\ =&\frac{1}{2e}\left(\frac{F(m)}{m} - k_e + \left(k_e + \frac{F(m)}{m}\right)\right)\\ =&\frac{F(m)}{em}.\end{aligned} \tag{30}$$

For C, we need to take the limit of $k_0 \to 0$ more carefuly,

$$\begin{aligned}
C|_{k_0=0} &= \lim_{k_0\to 0} \frac{1}{2k_0}\left(\frac{F(m)}{m} + k_e + k_0\frac{E_0(m)}{m} - \sqrt{\left(k_e - \frac{F(m)}{m}\right)^2 + 2\left(k_e - \frac{F(m)}{m}\right)k_0\frac{E_0(m)}{m} + 4k_e\frac{F(m)}{m}}\right) \\
&= \lim_{k_0\to 0} \frac{1}{2k_0}\left(\frac{F(m)}{m} + k_e + k_0\frac{E_0(m)}{m} - \sqrt{\left(k_e + \frac{F(m)}{m}\right)^2 + 2\left(k_e - \frac{F(m)}{m}\right)k_0\frac{E_0(m)}{m}}\right) \\
&= \lim_{k_0\to 0} \frac{1}{2k_0}\left(\frac{F(m)}{m} + k_e + k_0\frac{E_0(m)}{m} - \left(k_e + \frac{F(m)}{m}\right)\sqrt{1 + 2\frac{k_e - \frac{F(m)}{m}}{\left(k_e + \frac{F(m)}{m}\right)^2}k_0\frac{E_0(m)}{m}}\right) \\
&\approx \lim_{k_0\to 0} \frac{1}{2k_0}\left(\frac{F(m)}{m} + k_e + k_0\frac{E_0(m)}{m} - \left(k_e + \frac{F(m)}{m}\right) - \frac{k_e - \frac{F(m)}{m}}{k_e + \frac{F(m)}{m}}k_0\frac{E_0(m)}{m}\right) \\
&= \frac{1}{2}\frac{E_0(m)}{m}\left(1 - \frac{k_e - \frac{F(m)}{m}}{k_e + \frac{F(m)}{m}}\right) \\
&= \frac{E_0(m)}{m}\frac{F(m)}{F(m) + mk_e}.
\end{aligned} \tag{31}$$

Thus, from Eqs. (27) and (31) at the slower time scale, the growth rate of cell is given by

$$g(m) = mk_0C \approx k_0\frac{E_0(m)F(m)}{F(m) + mk_e}. \tag{32}$$

To specify how the total amount of enzymes ($E_0(m)$) depends on the cell mass, we assume that the cell biomass (m) is divided between enzymes $E_0(m)$ and constant component M_0, representing non-processing biomass such as DNA. Therefore,

$$E_0(m) = m - M_0. \tag{33}$$

We define the rate of substrate absorption as being proportional to their concentration in the environment and the surface area of the cell. The surface area depends on the cell biomass and the cell shape. We consider two shapes: a sphere and a cylinder (rod shaped). Expressing the surface area of these figures in terms of their volume, we found the rates of substrate absorption as

$$\begin{aligned}
F_{\text{sph}}(m) &= k_aU\left(36\pi m^2\right)^{1/3}, \\
F_{\text{rod}}(m) &= 2k_aU\left(m/R + \pi R^2\right),
\end{aligned} \tag{34}$$

where U is the concentration of the nutrients in the environment, R is the constant radius of a rod-shaped cell, k_a is the absorption coefficient (which we set to one, $k_a = 1$).

5 Optimal equal split strategies of cell division at extreme conditions

Numerical simulations show that the optimal division strategy for our model of cell growth is always an equal split into two parts, $\mathcal{M} = \{m^*/2, m^*/2\}$. In this section, we derive analytical results for two limit cases: low and high concentration of nutrients (U).

The optimal mass of the size at fragmentation, m^*, is given by equation $g(m^*/2)/(m^*/2) = g(m^*)/m^*$, see SI 5.3. We found that the growth rate is given by Eq. (32). Hence, at the optimal life cycle,

$$\frac{1}{m^*/2}\frac{E_0(m^*/2)F(m^*/2)}{F(m^*/2) + \frac{1}{2}m^*k_e} = \frac{1}{m^*}\frac{E_0(m^*)F(m^*)}{F(m^*) + m^*k_e}. \tag{35}$$

5.1 Low nutrient concentration

For low nutrient concentration $U \ll 1$, the substrate acquisition rate is low ($F(m^*) \ll 1$) and in the limit $F(m^*) \to 0$, Eq. (35) can be reduced to

$$E_0(m^*/2)F(m^*/2) = \frac{1}{4}E_0(m^*)F(m^*) \tag{36}$$

For rod-shaped cells, the substrate acquisition rates and the amount of enzymes in a cell are $F_{\mathsf{rod}}(m) = 2k_aU\left(m/R + \pi R^2\right)$ and $E_0(m) = m - M_0$. Substituting this into Eq. (36), we obtain the optimal mass for a newborn rod-shaped cell

$$\frac{m^*_{\mathsf{rod}}}{2} = \frac{3}{2} \cdot \frac{M_0}{1 - \frac{M_0}{\pi R^3}}. \tag{37}$$

This makes the lower bound of the newborn rod-shaped cell as $\min(\frac{m^*_{\mathsf{rod}}}{2}) = \frac{3}{2}M_0$.

For spherical cells, the substrate acquisition rates and amount of enzymes in a cell are:

$$\begin{aligned} F_{\mathsf{sph}}(m) &= k_aU\left(36\pi m^2\right)^{1/3}, \\ E_0(m) &= m - M_0. \end{aligned} \tag{38}$$

In this case, we obtain

$$\frac{m^*_{\mathsf{sph}}}{2} = \frac{1 - \frac{2^{2/3}}{4}}{1 - \frac{2^{2/3}}{2}} M_0 \approx 2.92 M_0. \tag{39}$$

This results in approximately twice as large minimal cell mass as in the case of rod-shaped cells. Thus, we expect that when nutrient concentration is low, rod-shaped cells are smaller than spherical cells.

5.2 High nutrient concentration

Next, we consider the opposite regime of high nutrient concentrations $U \gg 1$. There, the substrate acquisition rate should be large as well $F(m) \gg 1$. The growth rate of a cell of mass m can be approximated as

$$\begin{aligned} g(m) &= k_0 \frac{E_0(m)F(m)}{F(m) + mk_e} \\ &= k_0 \frac{E_0(m)}{1 + mk_e/F(m)} \\ &\approx k_0 E_0(m)\left(1 - \frac{mk_e}{F(m)}\right) \\ &= k_0 E_0(m)\left(1 - \frac{mk_e}{Uf(m)}\right), \end{aligned} \tag{40}$$

where at the last step, we explicitly emphasized that the substrate acquisition rate $F(m)$ is proportional to the nutrient concentration $F(m) = Uf(m)$, see Eq. (34).

The optimal mass of equal split fragmentation is again (see SI 5.3) $g(m^*/2)/(m^*/2) = g(m^*)/m^*$. Inserting Eq. (40) into this, we can find at which nutrient concentration U, fragmentation at the mass m^* is evolutionary optimal,

$$\begin{aligned} U &= k_e m^* \frac{\frac{E_0(m^*)}{f(m^*)} - \frac{E_0(m^*/2)}{f(m^*/2)}}{E_0(m^*) - 2E_0(m^*/2)} \\ &= \frac{k_e m^*}{M_0}\left(\frac{m^* - M_0}{f(m^*)} - \frac{m^*/2 - M_0}{f(m^*/2)}\right). \end{aligned} \tag{41}$$

This solution provides us with the function $U(m^*)$, while we are interested in the inverse function $m^*(U)$. To find this, consider an effective substrate acquisition rate in a form

$$f(m) = \alpha m^\beta. \tag{42}$$

Then,

$$\begin{aligned} U &= \frac{k_e}{\alpha M_0}\left(1-2^{-(1-\beta)}-(1-2^{-\beta})\frac{M_0}{m^*}\right)m^{*2-\beta} \\ &\approx \frac{k_e}{\alpha M_0}\left(1-2^{-(1-\beta)}\right)m^{*2-\beta} \end{aligned} \tag{43}$$

For the two shapes considered in this study, we have $\alpha = (36\pi)^{1/3}k_a$, $\beta = 2/3$ for spherical cells and for rod-shaped cells $\alpha = 2k_a/R, \beta = 1$.

Hence, for spherical cells with $\beta < 1$, at $m^* \gg M_0$ the optimal mass of newborn cells changes with nutrients concentrations as

$$m^* \sim U^{\frac{1}{2-\beta}}. \tag{44}$$

For the linear case ($\beta = 1$), which corresponds to rod-shaped cells, we can consider a more detailed asymptotic

$$f(m) = \alpha m + \gamma. \tag{45}$$

Inserting this into Eq. (41), we get

$$U = \frac{k_e}{\alpha}\left(1+\frac{\gamma}{\alpha M_0}\right)\frac{1}{\left(1+\frac{\gamma}{\alpha m^*}\right)}\frac{1}{\left(1+\frac{2\gamma}{\alpha m^*}\right)} \tag{46}$$

In the limit $m^* \to \infty$, the function $U(m^*)$ saturates at $U^* = \frac{k_e}{\alpha}\left(1+\frac{\gamma}{\alpha M_0}\right)$. As a consequence, once the concentration of the nutrients U approaches U^*, the optimal mass of a newborn cell goes to infinity, and at $U > U^*$, the unconstrained growth becomes the limit of life cycle optimization. For the case of rod-shaped cells, $\alpha = 2k_a/R$ and $\gamma = 2k_a\pi R^2$, so $U^* = \frac{k_e R}{2k_a}\left(1+\frac{\pi R^3}{M_0}\right)$.

Finally, the scenario $\beta > 1$ corresponds to a case where the surface of a cell grows faster than its volume, which is not biologically meaningful. Yet, mathematically, this leads to productivity $g(m)/m$ being a monotonically increasing function, which makes unconstrained growth to be the evolutionarily limit.

5.3 At the internal ($m_i > 1$) equal split optimal life cycle, the fragmentation event does not change the productivity of organism

All evolutionary optimal reproduction modes found in numerical simulations of the model presented in 4 were equal splits with offspring sizes larger than the minimal viable mass ($m_i > 1$).

The binary equal split reproduction mode is characterized by the fragments set $\mathcal{M} = \{m^*/2, m^*/2\}$. Hence, the growth rate is provided by a solution of the equation

$$1 - 2e^{-\lambda T(m^*/2, m^*)} = 0, \tag{47}$$

which, in this case, can be explicitly solved

$$\lambda = \frac{\log(2)}{T(m^*/2, m^*)}. \tag{48}$$

The optimal mass of this type of fragmentation is achieved at m^* minimizing the growth time $T(m^*/2, m^*)$. For internal optima ($m^*/2 > 1$), this means that the first derivative of the growth rate $T(m^*/2, m^*)$ is equal to zero.

$$\frac{d}{dm^*}T(m^*/2, m^*) = \frac{d}{dm^*}\int_{m^*/2}^{m^*}\frac{dy}{g(y)} = \frac{1}{g(m^*)} - \frac{1}{2g(m^*/2)} = 0. \tag{49}$$

We can rewrite the last expression as

$$\frac{g(m^*)}{m^*} = \frac{g(m^*/2)}{m^*/2}. \tag{50}$$

This can be interpreted as the productivities of cells, just before and just after the fragmentation event, are equal in evolutionarily optimal non-boundary equal split life cycles.

6 Rosette colonies

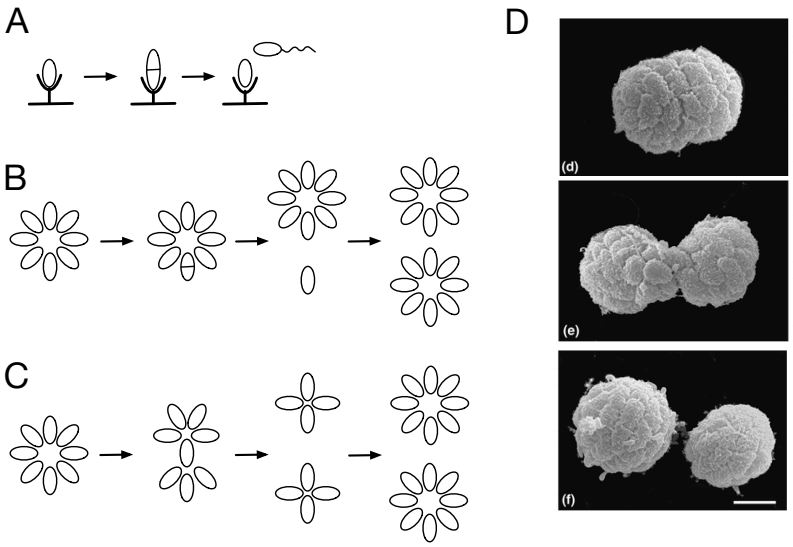

Figure SI 2: **Rosette colonies may reproduce by single-cell propagules or by an equal split, but the latter variant is predominant.**
A. Unicellular sedentary morphotypes of choanoflagellates are attached to the substrate by the anchor. Thus one of the offspring cells is born as a swimming type and can attach to substrate later. **B.** There is a possibility to use the same mechanism in the rosette colony, where a dividing cell releases the offspring, which later develops into another colony. Such a reproduction mode is mechanistically simpler than equal split, yet it is not dominant among rosette colonies. **C.** Rosette colonies reproduce by equal split into two equal-sized daughter colonies. Due to the spherical shape of the colony, such a splitting requires collective actions of multiple cells. **D.** Rosette colonies of magnetotactic bacteria reproduce by means of an equal split. Adopted from [3].

We do not have enough information to derive the productivity profiles of rosette colonies from first principles. However, the very structure of the colony implies some restrictions on the growth rate profiles. Due to the limited space around the center of the colony, the maximal cell count of a rosette colony must be limited. Rosettes with cell count exceeding this limit must experience high mechanical stress applied to member cells, which will harm their functional capacity.

Following this idea, we design productivity profiles to satisfy two conditions:

1. An increase in the colony size increases the productivity, while the colony size is below the critical value (representing a completely packed rosette), $\frac{d}{dm}\frac{g(m)}{m} > 0$ at $m < 20$.

2. A completed rosette ($m = 20$) is unable to incorporate any more cells, so the further growth leads to the fast decrease in productivity $\frac{d}{dm}\frac{g(m)}{m} < 0$ at $m > 20$, until it hits zero at $g(21) = 0$.

The maximal number of cells in a complete rosette was set to 20. We generated 10^5 random productivity profiles satisfying these criteria. Each such profile is a composition of two monotonic parts: For colony sizes between 1 and 20, the productivity grows monotonically with size, while at the sizes between 20 and 21, the productivity declines monotonically to zero.

Each of these monotonic parts was generated iteratively. At each step, the partial profile is a piece-wise linear function, connecting a series of joints on a $(m, g(m)/m)$ space. Each partial profile was initiated with only two joints given by reference points (e.g. $g(1) = 1$ and $g(20) = 10$ for the first part), hence it originated as a line. Then, we randomly sampled a single point within the rectangle defined by the reference points and used it as an intermediary joint. After the first iteration, the partial profile was composed of two linear pieces, see Fig. SI 3. At the next iteration, each of these pieces was broken in

two by the same procedure. For each partial profile, we continued the iteration 8 times, so the resulting profile contained $2 \times 256 = 512$ linear pieces.

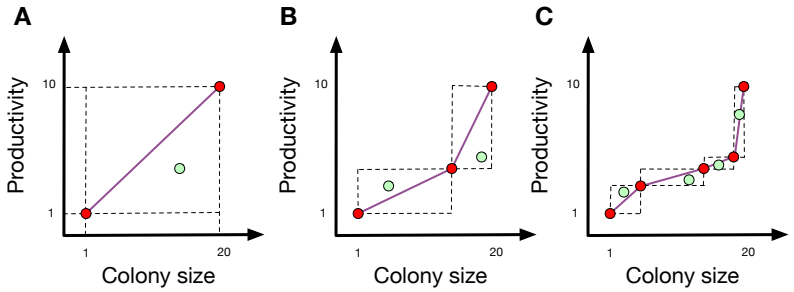

Figure SI 3: **Random profiles of productivity of rosette colonies were drawn iteratively.**
A. Each random profile was initiated as a line connecting two points. At the first step, the joint was randomly sampled within a rectangle defined by the two boundary points. **B.** After the first step, the profile connects three points and contains two linear pieces. Each pair of sequential joints defines a rectangle. Additional joints were randomly sampled within each, increasing the total number of internal joints to 3. **C.** The procedure was repeated to produce a random monotonic profile of productivity.

7 Model with constrained population size

Throughout the paper, we used a model that featured an unconstrained exponential growth of the population. However, no real population grows without constraints, and the population size is always bounded. In this section, we consider a model where the population size is constrained by a density dependent death rate, based on the Lotka-Volterra equation widely used in eco-evolutionary models [4]. There, in addition to the process of growth and fragmentation, any organism is subjected to death upon a collision with any other organism in the population with a rate $\frac{1}{K}$ determining the carrying capacity of the population. Then, the state of a population containing several competing sub-populations with different fragmentation strategies $\mathcal{M}_i$, $i = 1, 2, 3, \ldots$ is described by the set of population densities $n_i(m,t)$. The population dynamics is given by

$$\frac{\partial}{\partial t} n_i(m,t) = -\frac{\partial}{\partial m}\left(g(m) n_i(m,t)\right) - \frac{n_i(m,t)}{K} \sum_j \int_1^{m_j^*} n_j(m',t) dm', \tag{51}$$

where an additional term proportional to $\frac{1}{K}$ is the density dependent death rate.

Here, we show that the evolutionary optimality of life cycles in this model is the same as in the original model with unconstrained growth.

First, we note that after a sufficiently long time, the demographic distribution within each competing sub-population equilibrates. Hence, we will search for the solution of Eq. (51) in the form

$$n_i(m,t) = \rho_i(m) f_i(t) N(t), \tag{52}$$

$$\int_1^{m_i^*} \rho_i(m) dm = 1, \tag{53}$$

$$\sum_i f_i(t) = 1, \tag{54}$$

where $\rho_i(m)$ is the fraction of organisms of mass m within the sub-population i, $f_i(t)$ is the fraction of organisms belonging to the sub-population i in the whole population, and $N(t)$ is the number of organisms in the whole population. The dynamics of $f_i(t)$ determines the evolutionary success of the fragmentation strategy i in the given population: $f_i = 0$ means the extinction of strategy i, and

$f_i(t) = 1$ means that i is the only strategy in the population. Note, that with these definitions, $N(t) = \sum_j \int_1^{m_j^*} n_j(m', t)dm'$. Thus, the population dynamics in the considered system is given by

$$\rho_i(m)\frac{d}{dt}\left(f_i(t)N(t)\right) = -f_i(t)N(t)\frac{d}{dm}\left(g(m)\rho_i(m)\right) - \frac{\rho_i(m)f_i(t)N^2(t)}{K}. \tag{55}$$

Dividing both sides by $\rho_i(m)f_i(t)N(t)$, we can separate the variables m and t

$$\frac{1}{f_i(t)N(t)}\frac{d}{dt}\left(f_i(t)N(t)\right) + \frac{N(t)}{K} = -\frac{1}{\rho_i(m)}\frac{d}{dm}\left(g(m)\rho_i(m)\right). \tag{56}$$

The right-hand side of Eq. (56) is independent on K and on the abundance of other competing subpopulations, and therefore, must be identical to the model with unconstrained growth. Then we can state that

$$-\frac{1}{\rho_i(m)}\frac{d}{dm}\left(g(m)\rho_i(m)\right) = \lambda_i, \tag{57}$$

where the constant λ_i is the population growth rate of the fragmentation strategy i in the model of unconstrained growth. Then, Eq. (56) can be written as

$$\begin{aligned}
&\frac{1}{f_i(t)N(t)}\frac{d}{dt}\left(f_i(t)N(t)\right) + \frac{N(t)}{K} = \lambda_i,\\
&\frac{1}{N(t)}\frac{d}{dt}N(t) + \frac{1}{f_i(t)}\frac{d}{dt}f_i(t) + \frac{N(t)}{K} = \lambda_i,\\
&\frac{d}{dt}f_i(t) = f_i(t)\lambda_i - f_i(t)\frac{1}{N(t)}\frac{d}{dt}N(t) - \frac{f_i(t)N(t)}{K}
\end{aligned} \tag{58}$$

Next, we sum over i, to find $\frac{1}{N(t)}\frac{d}{dt}N(t)$. Since $\sum_i f_i(t) = 1$, then $\sum_i \frac{d}{dt}f_i(t) = 0$ and then

$$\frac{1}{N(t)}\frac{d}{dt}N(t) = \sum_i f_i(t)\lambda_i - \frac{N(t)}{K}, \tag{59}$$

Plugging Eq. (59) into Eq. (58), we get

$$\frac{d}{dt}f_i(t) = f_i(t)\left(\lambda_i - \sum_j f_j(t)\lambda_j\right) = f_i(t)(\lambda_i - \langle\lambda\rangle). \tag{60}$$

The resulting dynamics of f_i in Eq. (60) depends only on the population growth rates λ_i and is independent of the population carrying capacity K. Therefore, the winner of selection competition in the constrained system follows exactly the same rules as in the case of the model with unconstrained growth.

8 Size-independent productivity profile promotes multiple fission under reproduction with mass loss

To find the evolutionarily optimal life cycle under reproduction with a mass loss with size-independent productivity $g(m)/m = 1$, we prove two statements:

1. Among all reproduction modes with a fixed number of offspring and a fixed maturity size, fragmentation into equal pieces is a local optimum.
2. Among equal split reproduction modes with fixed maturity size, a fragmentation into more pieces is evolutionarily advantageous if the biomass loss does not decrease with offspring number.
3. Among equal split reproduction modes with the maximal number of offspring, a fragmentation into larger sizes is evolutionarily advantageous if the biomass loss grows slower than linear with maturity size.

As a result, the most optimal life cycle is a growth to the largest possible size and then split into as many equal offspring pieces as possible.

8.1 Among reproduction modes with a fixed number of offspring and fixed maturity size, fragmentation into equal pieces is a local optimum.

For size-independent productivity $g(m)/m = 1$, the growth time is

$$T(m_i, m^*) = \int_{m_i}^{m^*} \frac{dy}{y} = \log\left[\frac{m^*}{m_i}\right]. \tag{61}$$

Then, the Euler-Lotka equation (17) for an arbitrary life cycle $\mathcal{M} = \{m_1, m_2, \ldots, m_z\}$ becomes

$$\sum_{i=1}^{z} \left(\frac{m_i}{m^*}\right)^{\lambda} = 1 \tag{62}$$

If the reproduction incurs a mass loss $\sum_i m_i < m^*$, then the growth rate λ of an arbitrary life cycle is confined within the interval $(0, 1)$.

Specifically, we consider an equal split life cycle $\mathcal{M}_e = \{m, m, \ldots, m\}$ with z offspring parts, in which mass is lost upon reproduction: $zm < m^*$.

To demonstrate that the life cycle $\mathcal{M}_e$ is a local optimum, we introduce a small perturbation of mass distribution across offspring: $\mathcal{M}'_e = \{m + \delta m_1, m + \delta m_2, \ldots, m + \delta m_z\}$, such that the total mass of offspring remains the same $\sum_i \delta m_i = 0$. We need to show that the growth rate of the original equal split life cycle is larger than of the disturbed life cycle $\lambda(\mathcal{M}_e) > \lambda(\mathcal{M}'_e)$.

First, we show that the equal split life cycle $\mathcal{M}_e$ is a stationary point, i.e., the first order derivative of the growth rate with respect to perturbation δm_i is equal to zero:

$$\begin{aligned}
&\sum_{i=1}^{z} \left(\frac{m + \delta m_i}{m^*}\right)^{\lambda+\delta\lambda} = 1, \\
&\sum_{i=1}^{z} \left(\frac{m}{m^*}\right)^{\lambda} \log\left[\frac{m}{m^*}\right] \delta\lambda + \sum_{i=1}^{z} \lambda \left(\frac{m}{m^*}\right)^{\lambda-1} \frac{\delta m_i}{m^*} = 0, \\
&z \left(\frac{m}{m^*}\right)^{\lambda} \log\left[\frac{m}{m^*}\right] \delta\lambda + \lambda \left(\frac{m}{m^*}\right)^{\lambda-1} \sum_{i=1}^{z} \frac{\delta m_i}{m^*} = 0, \\
&z \left(\frac{m}{m^*}\right)^{\lambda} \log\left[\frac{m}{m^*}\right] \delta\lambda = 0, \\
&\delta\lambda = 0.
\end{aligned} \tag{63}$$

Here, in the first line, we used

$$\left(f^h\right)' = h' \log[f] f^h + h f' f^{h-1} \tag{64}$$

were we plugged in $f \to m/m^*$ and $h \to \lambda$.

This result implies that $\mathcal{M}_e$ can be an evolutionary optimum. To prove that it is a maximum of the population growth rate, we need to show that the second derivative is negative.

To compute the second derivative of the growth rate, we use

$$\left(f^h\right)'' = f^h \left(h'' \log[f] + (h' log[f])^2\right) + f^{h-1} \left(h f'' + 2 f' h' + 2 f' h' h \log[f]\right) + f^{h-2} h(h-1)(f')^2 \tag{65}$$

This expression can be further simplified. First, since we use the substitution $h \to \lambda$, and we already find that the first derivative of growth rate is equal to zero, $h' = 0$. Second, since we use the substitution $f \to m/m^*$, the second derivative of mass with respect to disturbance δm_i is equal to zero as well $f'' = 0$. Hence, in our case:

$$\left(f^h\right)'' = f^h h'' \log[f] + f^{h-2} h(h-1)(f')^2 \tag{66}$$

For the second derivative of the growth rate $\delta^2\lambda$, we get

$$\begin{aligned}
&\sum_{i=1}^{z}\left(\frac{m+\delta m_i}{m^*}\right)^{(\lambda+\delta\lambda)}=1,\\
&\sum_{i=1}^{z}\left(\frac{m}{m^*}\right)^{\lambda}\log\left[\frac{m}{m^*}\right]\delta^2\lambda+\sum_{i=1}^{z}\left(\frac{m}{m^*}\right)^{\lambda-2}\lambda(\lambda-1)\left(\frac{\delta m_i}{m^*}\right)^2=0,\\
&z\left(\frac{m}{m^*}\right)^{\lambda}\log\left[\frac{m}{m^*}\right]\delta^2\lambda+\left(\frac{m}{m^*}\right)^{\lambda-2}\lambda(\lambda-1)\sum_{i=1}^{z}\left(\frac{\delta m_i}{m^*}\right)^2=0, \qquad (67)\\
&\delta^2\lambda=-\left(\frac{m}{m^*}\right)^{\lambda-2}\cdot\frac{\lambda(\lambda-1)\sum_{i=1}^{z}\left(\frac{\delta m_i}{m^*}\right)^2}{z\left(\frac{m}{m^*}\right)^{\lambda}\log\left[\frac{m}{m^*}\right]},\\
&\delta^2\lambda=-\frac{\lambda(\lambda-1)\sum_{i=1}^{z}\left(\delta m_i\right)^2}{zm^2\log\left[\frac{m}{m^*}\right]}.
\end{aligned}$$

The second derivative is negative because two negative signs cancel each other: The sum in the numerator is positive, $zm^2>0$. At the same time, $\log\left[\frac{m}{m^*}\right]<0$ in the denominator because the mass of offspring m is smaller than the parent mass m^*. Finally, we have $\lambda(\lambda-1)<0$ because under reproduction with a mass loss $0<\lambda<1$

Hence, the equal split is a locally optimal life cycle for size-independent productivity profile. From now on, we focus only on the equal split life cycles.

8.2 Among equal split reproduction modes with fixed maturity size, fragmentation into more pieces is evolutionarily advantageous.

For an equal split life cycle with $g(m)/m=1$, the growth rate can be found explicitly

$$\lambda=\frac{\log[z]}{\log[\frac{m^*}{m}]}=\frac{\log[z]}{\log[\frac{zm^*}{m^*-l(m^*,z)}]}, \qquad (68)$$

where $l(m^*,z)$ is an amount of biomass lost in the reproduction event.

To compute the evolutionarily optimal number of offspring in such a life cycle, we compute the derivative of the growth rate with respect to the number of offspring z

$$\frac{d}{dz}\lambda=\frac{(m^*-l)\log[\frac{m^*}{m^*-l}]+z\log[z]\frac{\partial}{\partial z}l}{z(m^*-l)\log^2[\frac{zm^*}{m^*-l}]}. \qquad (69)$$

The denominator is strictly positive because the biomass loss l is smaller than the parent size m^*. The first term in the nominator $(m^*-l)\log[m^*/(m^*-l)]$ is also strictly positive for the same reason. Therefore, producing more offspring is advantageous, unless $\frac{\partial}{\partial z}l<0$, i.e., the biomass loss decreases with the number of offspring produced (and this is a quite unrealistic assumption).

More offspring produced from the same maturity mass means smaller offspring. Therefore, we may only consider life cycles in which the parent organism splits into offspring of the minimal viable mass 1.

8.3 Among equal split reproduction modes with the maximal number of offspring, a fragmentation into larger sizes is evolutionarily advantageous, if the biomass loss grows slower than linear with maturity size.

For a life cycle equal split with offspring size equal to 1 and $g(m)/m=1$, the growth rate is

$$\lambda=\frac{\log[z]}{\log[m^*]}=\frac{\log[m^*-l(m^*)]}{\log[m^*]}, \qquad (70)$$

where $l(m^*)$ is the biomass loss in the reproduction act. Then, assuming a power law dependence $l(m^*) = Am^{*B}$, the first derivative of the population growth rate with respect to maturity size is

$$\begin{aligned}
\frac{d}{dm^*}\lambda =& \frac{d}{dm^*}\frac{\log[m^* - Am^{*B}]}{\log[m^*]} \\
=& \frac{1}{\log^2[m^*]}\left(\frac{(1 - ABm^{*B-1})\log[m^*]}{m^* - Am^{*B}} - \frac{\log[m^* - Am^{*B}]}{m^*}\right) \\
=& \frac{1}{\log^2[m^*]}\left(\frac{(1 - ABm^{*B-1})\log[m^*]}{m^* - Am^{*B}} - \frac{\log[m^*] + \log[1 - Am^{*B-1}]}{m^*}\right) \\
=& \frac{1}{m^*\log[m^*]}\left(\frac{1 - ABm^{*B-1}}{1 - Am^{*B-1}} - 1 - \frac{\log[1 - Am^{*B-1}]}{\log[m^*]}\right) \\
=& \frac{1}{m^*\log[m^*]}\left(\frac{(1 - B)Am^{*B-1}}{1 - Am^{*B-1}} - \frac{\log[1 - Am^{*B-1}]}{\log[m^*]}\right) \\
=& \frac{1}{m^*\log[m^*]}\left(\frac{(1 - B)l(m^*)/m^*}{1 - l(m^*)/m^*} - \frac{\log[1 - l(m^*)/m^*]}{\log[m^*]}\right).
\end{aligned} \tag{71}$$

The biomass loss cannot exceed the maturity size, $l(m^*) < m^*$, thus $0 < l(m^*)/m^* < 1$. Therefore, in the resulting expression, the first term is positive, when $B < 1$ and the second (subtrahend) term is strictly negative. As a result, the population growth rate is improved with the increase in the maturity size if the biomass loss grows slower than linear.

9 The biomass loss necessary for multiple fragmentation to evolve is on the order of magnitude of the mass of a minimal viable cell

Binary fragmentation is always optimal under reproduction without biomass loss. With an increase in biomass loss, it eventually concedes the evolutionary optimality to fragmentation into multiple parts. How large the critical biomass loss at which the transition from binary to multiple fragmentation occurs is, depends on the productivity function $g(m)/m$. To estimate the scale of the critical biomass loss, we used a unimodal productivity $g(m)/m = 1 + He^{-(m-m_{\text{opt}})^2/(2\sigma^2)}$ and screened through a range of parameter values, see Fig. SI 4. We found that for a wide range of parameter values, the critical loss is of the same scale as the mass of the minimal viable cell. Only when the maximal productivity bonus (H) is small, the transition from binary to multiple fragmentation occurs at small losses. This corresponds to a scenario where the organism size has only little impact on productivity. Otherwise, a significant loss of biomass is necessary for multiple fragmentation to evolve.

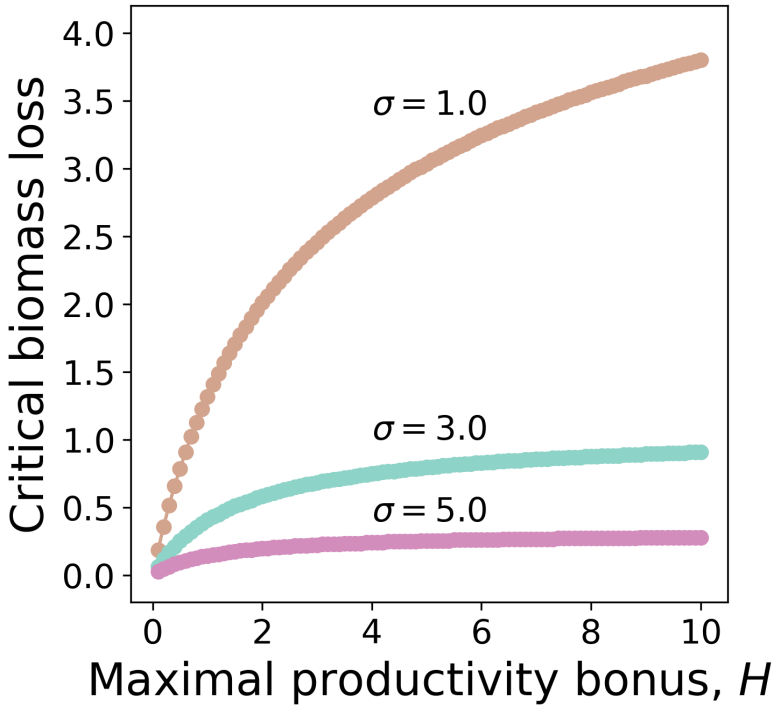

Figure SI 4: **Critical biomass loss increases with the maximal productivity bonus.** Multiple fragmentation evolves, if the biomass loss is larger than the critical value. For a unimodal productivity function $g(m)/m = 1 + He^{-(m-m_{\text{opt}})^2/(2\sigma^2)}$, the critical loss is on the order of a minimal viable cell. This loss grows with the maximal productivity bonus H and decreases with a widening of optimality peak (σ) (In these simulations, we assumed that the optimal grows occurs when the organism size is five times the minimum viable cell size, $m_{\text{opt}} = 5.0$).

References

[1] H. von Foerster. Some remarks on changing populations. In F. Stohlman, editor, *The Kinetics of Cellular Proliferation*, pages 382 – 407. Grune and Stratton, 1959.

[2] M. Renardy and R.C. Rogers. *An introduction to partial differential equations.*, volume 13. Springer Science and Business Media., 2006.

[3] C.N. Keim, J.L. Martins, F. Abreu, A.S. Rosado, H.L. de Barros, R. Borojevic, U. Lins, and M. Farina. Multicellular life cycle of magnetotactic prokaryotes. *FEMS Microbiology Letters*, 240(2):203 – 208, 2004.

[4] J. Murray. *Mathematical Biology*. Springer, Berlin, 1993.