Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S1 Appendix.

Population growth rate in the case of stochastic developmental programs. Consider a population in which each group emerges as one of S initial types. These types could be the newborn groups of different size and/or composition. With time passing, a group grows from its initial size to maturity and subsequent fragmentation. The set of growth events (cells divisions, mutations, etc) may vary from group to group. We call such an event chain "developmental trajectory" and designate it as τ. Any two groups of the same initial type may adopt different developmental trajectories for a number of reasons, such as mutations, stochastic developmental programs, or different environmental conditions. We use the following parameters of the developmental trajectory: $i(\tau)$ – the initial state of the group leading to the given developmental trajectory, $p_k(\tau)$ – the probability that a group that emerged as initial type k will follow the trajectory τ, so $p_k(\tau) = 0$, if $k \neq i(\tau)$, $T(\tau)$ – the time necessary to the newborn group to complete the trajectory τ and $\mathbf{N}(\tau) = (N_1, N_2, \cdots, N_S)$ – the vector of numbers of each offspring type produced during the fragmentation at the end of the trajectory τ.

The population features an explicit maturation component: a newborn group does not reproduce until time $T(\tau)$ has passed. Thus, to describe the population dynamics and find the population growth rate λ, it is necessary to consider the population demography. To do so, we characterize each group at each moment of time by the age parameter η. We define the age in a way that the newborn group has $\eta = 0$, while the group that reached the end of the developmental trajectory and is about to fragment has $\eta = 1$. Along the trajectory, the age increases at a constant rate equal to $\frac{1}{T(\tau)}$, i.e. the rate of ageing differs between different trajectories.

From the perspective of the population dynamics, any two groups sharing the same developmental trajectory τ and age η are identical. Thus, the state of the whole population can be described by the density function $\zeta(\tau, \eta, t)$, which shows how many groups on the developmental trajectory τ have age η at the given time t. In the stationary regime, where the fraction of groups of each type stays constant, the density function grows exponentially,

$$\zeta(\tau, \eta, t) = \rho(\tau, \eta)e^{\lambda t}, \tag{8}$$

where $\rho(\tau, \eta)$ is the stationary density distribution of groups in a population.

Within a given developmental trajectory, ageing occurs at the same rate for all groups. Therefore, the dynamics of the density function at a given age η is determined by the balance between influx of maturing younger groups and the outflux of groups becoming too old. Both processes occur with the same rate $\frac{1}{T(\tau)}$, thus the density function must satisfy the transport equation

$$\frac{\partial \zeta}{\partial t} = -\frac{1}{T(\tau)}\frac{\partial \zeta}{\partial \eta}. \tag{9}$$

Combining Eqs. (8) and (9) we get

$$\lambda\rho = -\frac{1}{T(\tau)}\frac{\partial\rho}{\partial\eta}$$

The solution of this equation is

$$\rho(\tau,\eta) = \rho_0(\tau)e^{-\lambda T(\tau)\eta}, \tag{10}$$

where $\rho_0(\tau)$ is the stationary density distribution of newborn groups with $\eta = 0$.

To find $\rho_0(\tau)$, we use the fact that each newborn organism is produced as a result of the fragmentation of some mature organism. Thus, the rate of emergence of newborn organisms in the population (j_0) is the same as the rate of production of offspring in the course of reproduction of mature organisms (j_1).

For any developmental trajectory τ, the rate of entering into the newborn state per time unit is equal to

$$j_0(\tau) = \frac{\zeta(\tau,0,t)}{T(\tau)}, \tag{11}$$

where the right hand side of the equation is the product of the number of newborn groups and the rate of ageing. The number of offspring with developmental trajectory τ is equal to the product of the total number of offspring of type $i(\tau)$ produced by all mature organisms and the probability of the offspring to adopt this developmental trajectory ($p_{i(\tau)}(\tau)$)

$$j_1(\tau) = p_{i(\tau)}(\tau)\sum_{\tau'}\frac{N_{i(\tau)}(\tau')}{T(\tau')}\zeta(\tau',1,t), \tag{12}$$

where summation is performed over all possible developmental trajectories of parent groups.

Since each produced propagule is a newborn organism, $j_0(\tau) = j_1(\tau)$. Therefore,

$$\frac{\rho_0(\tau)}{T(\tau)} = p_{i(\tau)}(\tau)\sum_{\tau'}\frac{N_{i(\tau)}(\tau')}{T(\tau')}\rho_0(\tau')e^{-\lambda T(\tau')}. \tag{13}$$

To obtain the expression connecting the population growth rate λ with parameters of developmental trajectories τ, we multiply both parts by $N_j(\tau)e^{-\lambda T(\tau)}$ (note that in general $j \neq i(\tau)$) and sum over all possible developmental trajectories

$$\sum_{\tau}\frac{N_j(\tau)}{T(\tau)}\rho_0(\tau)e^{-\lambda T(\tau)} = \sum_{\tau}p_{i(\tau)}(\tau)N_j(\tau)e^{-\lambda T(\tau)}\left(\sum_{\tau'}\frac{N_{i(\tau)}(\tau')}{T(\tau')}\rho_0(\tau')e^{-\lambda T(\tau')}\right). \tag{14}$$

We define

$$X_i = \sum_{\tau}\frac{N_i(\tau)}{T(\tau)}\rho_0(\tau)e^{-\lambda T(\tau)} \tag{15}$$

$$Q_{i,j} = \sum_{\tau}p_i(\tau)N_j(\tau)e^{-\lambda T(\tau)}, \tag{16}$$

Note that $p_j(\tau) = 0$ if $j \neq i(\tau)$.

Taking into account that $p_j(\tau) = 0$ if $j \neq i(\tau)$, Eq (14) becomes

$$X_j = \sum_i Q_{i,j}X_i. \tag{17}$$

Also in the definition of $Q_{i,j}$, the result of summation over all trajectories τ is the same as over only developmental trajectories starting from the initial state of type j, since $p_j(\tau) = 0$, if

$j \neq i(\tau)$, because an organism emerged as one type has no access to developmental trajectories originated from other types.

Eq (17) can be satisfied only if

$$\det(Q - I) = 0, \tag{18}$$

where elements of matrix Q are defined by Eq (16) and I is identity matrix. This equation allows to infer the population growth rate λ if the parameters of each trajectory are known ($i(\tau)$, $p_i(\tau)$, $\mathbf{N}(\tau)$ and $T(\tau)$). In most interesting cases, this has to be done numerically.

Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S2 Appendix.

Existence of the neutral fitness landscape in the case of homogeneous groups. Consider the situation, where $w = 0$ and, therefore, the group properties depend only on the group size. A group of size i grows in size to $i + 1$ within time T_i. Here we show that if $T_i = \ln\left(\frac{i+1}{i}\right)$, all life cycles have the same growth rate $\lambda = 1$. We prove this by induction:

- The base of induction is given by Eq (4), which states that if $T_1 = \ln\left(\frac{2}{1}\right)$ and $T_2 = \ln\left(\frac{3}{2}\right)$, then $\lambda = 1$ for any life cycles fragmenting at size 3 or smaller.
- The step of induction must show that if the assumption of induction holds true for maximal size M, then under adding $T_M = \ln\left(\frac{M+1}{M}\right)$, the assumption also holds true for maximal size $M + 1$. To prove the step of induction, we only need to consider life cycles fragmenting exactly at the size $M + 1$ because life cycles fragmenting at sizes smaller than $M + 1$ have $\lambda = 1$ according to the assumption of induction.

To construct the matrix Q and find the growth rate of considered life cycles, we need to characterize the set of offspring and developmental trajectories. In an arbitrary life cycle, the fragmentation of a homogeneous group of size M results in production of offspring groups of sizes ranging from 1 to M. In total, M different types of offspring can be produced, so the size of the matrix Q is M by M. Each of the offspring will grow up to size $M + 1$ and then fragment, thus there is only one developmental trajectory for each type of offspring with $p_i(\tau) = 1$. The developmental time of the trajectory $\widetilde{T}(\tau)$ is given as the sum of incremental growth time

$$\widetilde{T}_k(\tau) = \sum_{j=k}^{M} T_j = \ln\left(\frac{M+1}{k}\right), \tag{19}$$

where k denotes the size of the newborn offspring.

An arbitrary life cycle can be characterized by the distribution of offspring sizes produced upon fragmentation N_i, where i denotes the size of offspring. By the conservation of cell number during reproduction $\sum_{i=1}^{M} iN_i = M + 1$. Therefore, according to Eq (16), for an arbitrary life cycle, the elements of matrix Q_{ij} are given by

$$Q_{ij} = N_i e^{-\lambda \ln\left(\frac{M+1}{j}\right)} \tag{20}$$

To prove the step of induction, we verify whether $\lambda = 1$ is the solution of Eq (18), with matrix Q given by Eq (20). Plugging $\lambda = 1$ into Eq (20), we have $Q_{ij} = N_i \frac{j}{M+1}$, so the Eq (18) becomes

$$\begin{vmatrix} N_1\frac{1}{M+1}-1 & N_1\frac{2}{M+1} & \cdots & N_1\frac{M}{M+1} \\ N_2\frac{1}{M+1} & N_2\frac{2}{M+1}-1 & \cdots & N_2\frac{M}{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ N_M\frac{1}{M+1} & N_M\frac{2}{M+1} & \cdots & N_M\frac{M}{M+1}-1 \end{vmatrix} = 0. \tag{21}$$

Based on the properties of determinant, we can take out the coefficients of each row and each column, then the left hand side of Eq (21) becomes

$$\frac{\prod_{i=1}^{M} iN_i}{(M+1)^M} \cdot \begin{vmatrix} 1-\frac{M+1}{1N_1} & 1 & \cdots & 1 \\ 1 & 1-\frac{M+1}{2N_2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-\frac{M+1}{MN_M} \end{vmatrix}. \tag{22}$$

For convenience, we neglect the coefficient and denote $\frac{M+1}{iN_i}$ as K_i. Thus, the determinant is

$$\begin{vmatrix} 1-K_1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-K_M \end{vmatrix}. \tag{23}$$

Next we calculate the determinant by splitting the first row,

$$\begin{aligned} \begin{vmatrix} 1-K_1 & 1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} &= \begin{vmatrix} -K_1 & 0 & 0 & \cdots & 0 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} \\ &+ \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix}. \end{aligned} \tag{24}$$

For the second part, splitting the second row, we can get

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} \\ &+ \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix}, \end{aligned} \tag{25}$$

The second term in Eq (25) is zero because the determinant has two identical columns, therefore only the first term remains. Continuing splitting the remaining rows of the first term of Eq (25), we finally obtain

$$
\begin{aligned}
\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} \\
&= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 0 & 0 & -K_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} \\
&= \cdots \\
&= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 0 & 0 & -K_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -K_M \end{vmatrix} \\
&= (-1)^{K-1} \prod_{i \neq 1}^{M} K_i.
\end{aligned}
\tag{26}
$$

Now, we look back at the first term in Eq (24), we split the second row

$$
\begin{aligned}
\begin{vmatrix} -K_1 & 0 & 0 & \cdots & 0 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} &= \begin{vmatrix} -K_1 & 0 & 0 & \cdots & 0 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} \\
&\quad + \begin{vmatrix} -K_1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix}.
\end{aligned}
\tag{27}
$$

For the second term at the right hand side of Eq (27), similar to Eq (25) in the last step, we can work out that it equals $(-1)^{M-1} \prod_{i \neq 2}^{M} K_i$. That means we can get $(-1)^{M-1} \prod_{i \neq j}^{M} K_i$ when split the j-th row. So we keep the same procedure to split the remaining rows of the first term in

Eq (27). After that, the initial determinant changes to

$$
\begin{aligned}
\begin{vmatrix} 1-K_1 & 1 & 1 & \cdots & 1 \\ 1 & 1-K_2 & 1 & \cdots & 1 \\ 1 & 1 & 1-K_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-K_M \end{vmatrix} &= \begin{vmatrix} -K_1 & 0 & 0 & \cdots & 0 \\ 0 & -K_2 & 0 & \cdots & 0 \\ 0 & 0 & -K_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -K_M \end{vmatrix} \\
&\quad + (-1)^{M-1} \sum_{j=1}^{M} \prod_{i \neq j}^{M} K_i \\
&= (-1)^M \prod_{i=1}^{M} K_i + (-1)^{M-1} \sum_{j=1}^{M} \prod_{i \neq j}^{M} K_i \\
&= (-1)^M \left(\prod_{i=1}^{M} K_i - \sum_{j=1}^{M} \prod_{i \neq j}^{M} K_i \right) \\
&= (-1)^M \left(\frac{(M+1)^M}{\prod_{i=1}^{M} iN_i} - \frac{(M+1)^{M-1} \sum_{i=1}^{M} iN_i}{\prod_{i=1}^{M} iN_i} \right) \\
&= 0,
\end{aligned}
\tag{28}
$$

where we used $K_i = \frac{M+1}{iN_i}$ and $\sum_{i=1}^{M} iN_i = M+1$ in the last two steps.

This proves that an arbitrary life cycle fragmenting at size $M+1$ has the growth rate $\lambda = 1$, if $T_i = \ln\left(\frac{i+1}{i}\right)$ for any $i \leq M$. This means that $T_i = \ln\left(\frac{i+1}{i}\right)$ is a neutral fitness landscape for the scenario of homogeneous groups.

Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S3 Appendix.

Life cycles of homogeneous groups.

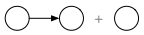
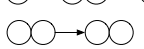
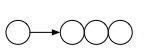
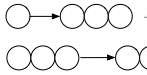
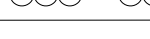
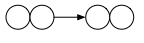
LC	τ	$N(\tau)$ $\circ\ \infty\ \infty\infty$	$p(\tau)$	$T(\tau)$
$1+1$		$(2,0,0)$	1	$t_{[1,0]}$
$2+1$		$(1,1,0)$	1	$t_{[1,0]}+t_{[2,0]}$
		$(1,1,0)$	1	$t_{[2,0]}$
$1+1+1$		$(3,0,0)$	1	$t_{[1,0]}+t_{[2,0]}$
$3+1$		$(1,0,1)$	1	$t_{[1,0]}+t_{[2,0]}+t_{[3,0]}$
		$(1,0,1)$	1	$t_{[3,0]}$
$2+2$		$(0,2,0)$	1	$t_{[2,0]}+t_{[3,0]}$
$2+1+1$		$(2,1,0)$	1	$t_{[1,0]}+t_{[2,0]}+t_{[3,0]}$
		$(2,1,0)$	1	$t_{[2,0]}+t_{[3,0]}$
$1+1+1+1$		$(4,0,0)$	1	$t_{[1,0]}+t_{[2,0]}+t_{[3,0]}$

Fig 1. Homogeneous groups have deterministic developmental trajectories for each type offspring group, i.e. $p(\tau)=1$**.**

In the absence of cells' interactions, all cells are identical i.e. the cell type has no influence on groups. Essentially, all groups can be treated as homogeneous groups, in which only group sizes affect growth rate. In this case, groups have fixed developmental trajectories. For instance, the life cycle 1+1+1 has to go through the unique developmental trajectory: two successive divisions and then producing three single cells (see Fig 1). In this unique developmental trajectory, only one initial type exist – independent cell, so $p(\tau)=1$ and $\mathbf{N}(\tau)=3$.

First, we investigate the simplest scenario, where the maximal size of the group was limited to two cells. There are three life cycles in total in this case: 1+1, 2+1 and 1+1+1. The matrices

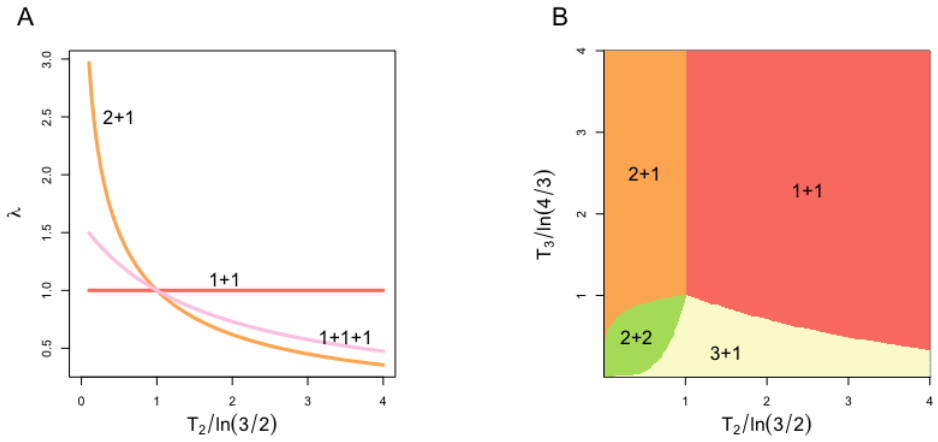


Fig 2. Growth rates and optimal life cycles in homogeneous groups on the condition of $n \leq 3$ and $n \leq 4$, respectively. A) describes the growth rates of life cycles when $n \leq 3$ i.e. 1+1, 1+1+1 and 2+1. **B**) shows the optimal life cycle when $n \leq 4$ with respect to T_2 and T_3. In both situations, T_i is the size increment time and we set $T_1 = \ln(2)$ for convenience.

Q corresponding to these life cycles are

$$\begin{aligned} Q_{1+1} &= \left(2e^{-\lambda T_1}\right), \\ Q_{2+1} &= \begin{pmatrix} e^{-\lambda(T_1+T_2)} & e^{-\lambda(T_1+T_2)} \\ e^{-\lambda T_2} & e^{-\lambda T_2} \end{pmatrix}, \\ Q_{1+1+1} &= \left(3e^{-\lambda(T_1+T_2)}\right). \end{aligned} \tag{29}$$

According to Eq 4, the growth rate of each life cycle are given by the solutions of

$$2e^{-\lambda_{1+1}T_1} - 1 = 0, \tag{30}$$

$$e^{-\lambda_{2+1}(T_1+T_2)} + e^{-\lambda_{2+1}T_2} - 1 = 0, \tag{31}$$

$$3e^{-\lambda_{1+1+1}(T_1+T_2)} - 1 = 0, \tag{32}$$

where λ_{1+1}, λ_{2+1} and λ_{1+1+1} are the growth rate of 1+1, 2+1 and 1+1+1, respectively, see Fig 2A. For small T_2, the largest growth rate is achieved by 2+1 life cycle. In this case, bi-cellular groups produce offspring cells faster than independent cells. Consequently, the life cycle 2+1, which allows production of bi-cellular groups (unlike unicellular life cycle 1+1) and preserving one offspring group in the most productive bi-cellular state (unlike 1+1+1) is most successful in growth competition. In the opposite limit of large T_2, the life cycle 1+1 leads to the largest population growth rate. In this case, independent cells are better off than bi-cellular groups. Thus, the best reproductive strategy is to avoid the growth to bi-cellular state, which can only be achieved with a single life cycle 1+1. In both situations of T_2, the growth rate of 1+1+1 is always between that of 1+1 and 2+1.

For the next scenario, we increase the maximal size of the group to three cells. This allows four new life cycles: 3+1, 2+2, 2+1+1 and 1+1+1+1. Their growth rates are given by the solutions of

$$e^{-\lambda_{3+1}(T_1+T_2+T_3)} + e^{-\lambda_{3+1}T_3} - 1 = 0 \tag{33}$$

$$2e^{-\lambda_{2+2}(T_2+T_3)} - 1 = 0 \tag{34}$$

$$2e^{-\lambda_{2+1+1}(T_1+T_2+T_3)} + e^{-\lambda_{2+1+1}(T_2+T_3)} - 1 = 0 \tag{35}$$

$$4e^{-\lambda_{1+1+1+1}(T_1+T_2+T_3)} - 1 = 0, \tag{36}$$

For large T_3, the life cycles which do not produce slow-growing three-cellular groups have the highest growth rates. Therefore, for large T_3, the optimal life cycles are the same as ones presented in the previous paragraph. For small T_3, the life cycles capable of producing three-cellular groups gain an evolutionary advantage. Specifically, 2+2 achieves the maximum growth rate when both T_2 and T_3 are comparatively small. In this case, an independent cell is the least productive state, whereas 2+2 is the only life cycle not producing independent cells. Life cycle 3+1 leads to the largest growth rate if T_3 is small but T_2 is large. There, the three-cellular group stands out as the most productive state, and 3+1 is the only life cycle keeping it as one of its offspring groups. Similarly to the previous scenario, life cycles with more than two offspring: 1+1+1, 2+1+1, 1+1+1+1, are never optimal. An important exception to this is the point $T_1 = \ln(2)$, $T_2 = \ln(\frac{3}{2})$ and $T_3 = \ln(\frac{4}{3})$, where all seven life cycles lead to the same growth rate ($\lambda = 1$).

Previously, we considered another model of life cycles evolution (1). There, the growth of groups from size i to size $i + 1$ occurs spontaneously with rate ib_i. Therefore, in that model, the time between cell divisions varies between groups of the same size, in contrast to the scenario considered here, where this time is always equal to T_i. Despite the differences between two models, they both share a number of findings: existence of the neutral point, only binary fragmentation is evolutionarily optimal, same optimal life cycles in the limit cases. Therefore, these features, are independent from the model design.

References

1. Pichugin Y, Peña J, Rainey P, Traulsen A. Fragmentation modes and the evolution of life cycles. PLoS Computational Biology. 2017;13(11):e1005860.

Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S4 Appendix.

Calculation of growth rates λ for life cycles of heterogeneous groups. To show how our approach can can be used in the case of heterogeneous groups, consider the simplest unicellular life cycle 1+1. There are two types of offspring possible: independent A and B cells, so the matrix Q has dimensions 2 by 2. When a cell divides into two, three outcomes are possible: no cell, one cell, or both daughter cells change the phenotype. Since the developmental trajectory ends after the first division, there are only six developmental trajectories possible for this life cycle, see Fig 3.

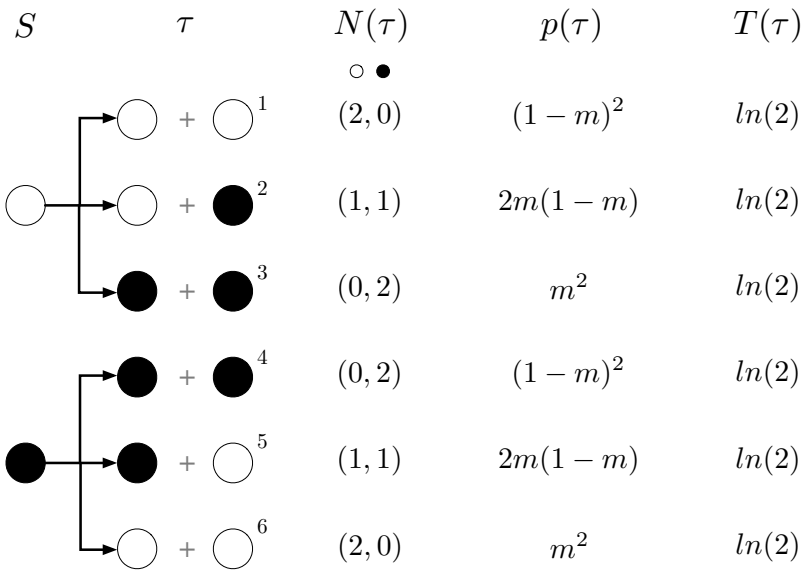


Fig 3. The full set of developmental trajectories in the life cycle 1+1. Here, the white and black circles denote A type cell and B type cell respectively.

To construct the matrix Q, we need to obtain the distribution of offspring (N_i), the probability of realization (p) and total developmental time (T) for each trajectory. Offspring distributions are apparent from Fig 3. The probability of each trajectory can be directly computed from the phenotype switch probability m. The developmental time is $T = T_1 = \ln(2)$ for each trajectory here. Therefore, the elements of Q are given by

$$Q_{11} = 2(1-m)^2 e^{-\lambda ln2} \qquad \longleftarrow \tau_1$$

$$
\begin{aligned}
&+ 2m(1-m)e^{-\lambda ln2} && \longleftarrow \tau_2,\\
Q_{12} =&2m(1-m)^2e^{-\lambda ln2} && \longleftarrow \tau_2\\
&+ 2me^{-\lambda ln2} && \longleftarrow \tau_3,\\
Q_{21} =&2(1-m)e^{-\lambda ln2} && \longleftarrow \tau_5\\
&+ 2m^2e^{-\lambda ln2} && \longleftarrow \tau_6,\\
Q_{22} =&2(1-m)^2e^{-\lambda ln2} && \longleftarrow \tau_4\\
&+ 2m(1-m)e^{-\lambda ln2} && \longleftarrow \tau_5,
\end{aligned}
$$

where arrows indicate the index of the developmental trajectory contributing a given term. Solution of the Eq (18) leads to $\lambda_{1+1} = 1$ in the life cycle 1+1.

Next, consider the life cycle 1+1+1. There are still two types of offspring possible: independent A and B cells, such that the matrix Q has dimensions 2 by 2. However, the life cycle requires two divisions to complete, so the number of possible developmental trajectories is increased to 20. Also, interactions play a role during the second division, so the probabilities p and developmental times T are more complicated, see Fig 4.

The elements of matrix Q are given by

$$
\begin{aligned}
Q_{11} =&3n^4e^{-\lambda(t_{[1,0]}+t_{[2,0]})} && \longleftarrow \tau_1\\
&+ 4mn^3e^{-\lambda(t_{[1,0]}+t_{[2,0]})} && \longleftarrow \tau_2\\
&+ m^2n^2e^{-\lambda(t_{[1,0]}+t_{[2,0]})} && \longleftarrow \tau_3\\
&+ 6m^3nP^B_{1+1}e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_4\\
&+ 4mn^2(nP^A_{1+1} + 2mP^B_{1+1})e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_5\\
&+ 2mn^2(2mP^A_{1+1} + nP^B_{1+1})e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_6\\
&+ 2m^4e^{-\lambda(t_{[1,0]}+t_{[0,2]})} && \longleftarrow \tau_8\\
&+ 2m^3ne^{-\lambda(t_{[1,0]}+t_{[0,2]})} && \longleftarrow \tau_9\\
Q_{12} =&2mn^3e^{-\lambda(t_{[1,0]}+t_{[2,0]})} && \longleftarrow \tau_2\\
&+ 2m^2n^2e^{-\lambda(t_{[1,0]}+t_{[2,0]})} && \longleftarrow \tau_3\\
&+ 2mn^2(nP^A_{1+1} + 2mP^B_{1+1})e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_5\\
&+ 4mn^2(2mP^A_{1+1} + nP^B_{1+1})e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_6\\
&+ 6m^3nP^A_{1+1}e^{-\lambda(t_{[1,0]}+t_{[1,1]})} && \longleftarrow \tau_7\\
&+ m^4e^{-\lambda(t_{[1,0]}+t_{[0,2]})} && \longleftarrow \tau_8\\
&+ 4m^3ne^{-\lambda(t_{[1,0]}+t_{[0,2]})} && \longleftarrow \tau_9\\
&+ 3m^2n^2e^{-\lambda(t_{[1,0]}+t_{[0,2]})} && \longleftarrow \tau_{10}\\
Q_{21} =&2mn^3e^{-\lambda(t_{[0,1]}+t_{[0,2]})} && \longleftarrow \tau_{12}\\
&+ 2m^2n^2e^{-\lambda(t_{[0,1]}+t_{[0,2]})} && \longleftarrow \tau_{13}\\
&+ 2mn^2(2mP^A_{1+1} + nP^B_{1+1})e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{15}\\
&+ 4mn^2(nP^A_{1+1} + 2mP^B_{1+1})e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{16}\\
&+ 6m^3nP^B_{1+1}e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{17}\\
&+ m^4e^{-\lambda(t_{[0,1]}+t_{[2,0]})} && \longleftarrow \tau_{18}\\
&+ 4m^3ne^{-\lambda(t_{[0,1]}+t_{[2,0]})} && \longleftarrow \tau_{19}
\end{aligned}
$$

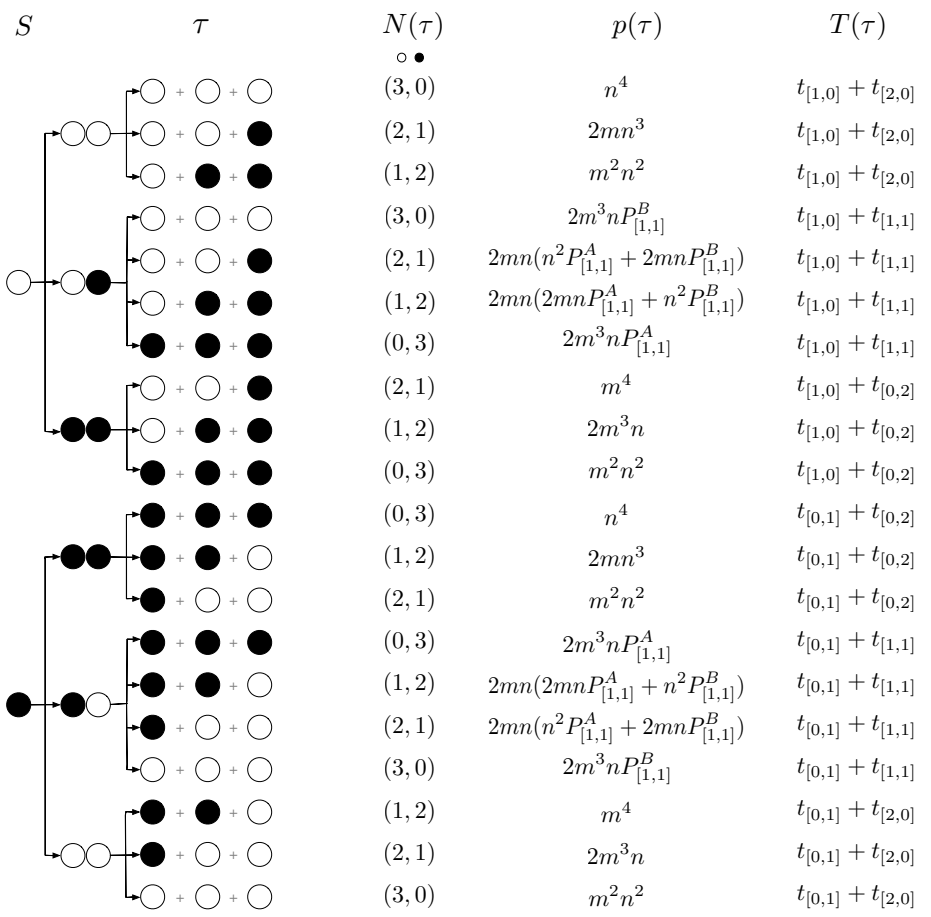


Fig 4. The full set of developmental trajectories in the life cycle 1+1+1. White circles represent A type cells and black circles represent B type cells. For simplicity of notation, we use $n = 1 - m$, $t_{[i,j]}$ is the time before the next cell division for a complex with i A type cells and j B type cells to divide; and $P^A_{[1,1]} = \frac{1}{2} + w\frac{b-c}{4}$ and $P^B_{[1,1]} = \frac{1}{2} - w\frac{b-c}{4}$, see details in the model section of the main text.

$$
\begin{aligned}
& + 3m^2n^2e^{-\lambda(t_{[0,1]}+t_{[2,0]})} && \longleftarrow \tau_{20} \\
Q_{22} = & 3n^4e^{-\lambda(t_{[0,1]}+t_{[0,2]})} && \longleftarrow \tau_{11} \\
& + 4mn^3e^{-\lambda(t_{[0,1]}+t_{[0,2]})} && \longleftarrow \tau_{12} \\
& + m^2n^2e^{-\lambda(t_{[0,1]}+t_{[0,2]})} && \longleftarrow \tau_{13} \\
& + 6m^3nP^A_{1+1}e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{14} \\
& + 4mn^2(2mP^A_{1+1} + nP^B_{1+1})e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{15} \\
& + 2mn^2(nP^A_{1+1} + 2mP^B_{1+1})e^{-\lambda(t_{[0,1]}+t_{[1,1]})} && \longleftarrow \tau_{16} \\
& + 2m^4e^{-\lambda(t_{[0,1]}+t_{[2,0]})} && \longleftarrow \tau_{18} \\
& + 2m^3ne^{-\lambda(t_{[0,1]}+t_{[2,0]})} && \longleftarrow \tau_{19}
\end{aligned}
$$

Here, $P^A_{1+1} = \frac{1}{2} + w\frac{b-c}{4}$, $P^B_{1+1} = \frac{1}{2} - w\frac{b-c}{4}$, $t_{[1,0]} = t_{[1,0]} = ln2$, $t_{[2,0]} = ln\frac{3}{2}(1 - wa)$, $t_{[1,1]} = ln2 + ln\frac{3}{2}(1 - w\frac{b+c}{2})$ and $t_{[2,0]} = ln\frac{3}{2}(1 - wd)$. Arrows indicate the contributions of

each developmental trajectory to Q_{ij}.

The solution of Eq (18) for life cycle 1+1+1 yields

$$\lambda_{1+1+1} = 1 + w\frac{\ln(3/2)}{\ln(9)}(a+d)\left((1-2m+2m^2) + 2m(1-m)\frac{b+c}{a+d}\right). \quad (37)$$

Our final example is the life cycle 2+1, where groups grow to three cells and fragment into a bi-cellular group and an independent cell. Here, five offspring types are possible: independent cells could be either A or B type and the bi-cellular group could have composition AA, AB, or BB, see Figs. 5, 6. Therefore, Q is 5×5 matrix. There are 48 developmental programs possible and we refrain from showing here how elements of matrix Q are constructed in this case. The solution of the Eq (18) for life cycle 2+1 yields

$$\lambda_{2+1} = 1+w\frac{3\ln(3/2)}{2(5+8m)\ln(27/4)}(a+d)\left((5-6m+10m^2) + 2m(7-5m)\frac{b+c}{a+d}\right). \quad (38)$$

For the life cycles 1+1+1+1 and 2+2, we just list the growth rate λ

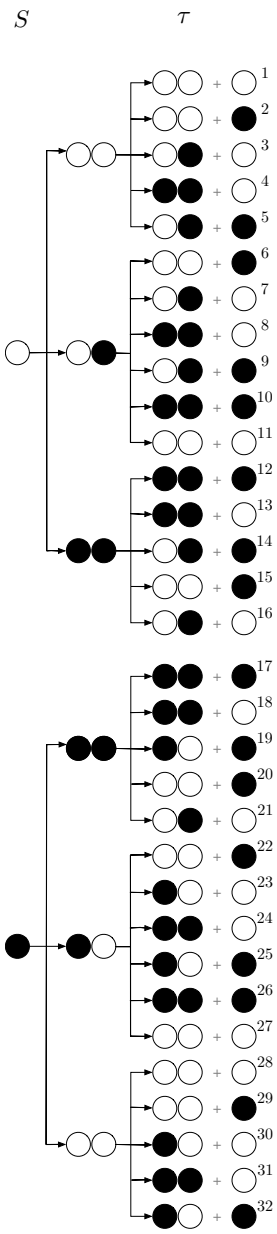
$$\begin{aligned}\lambda_{1+1+1+1} =& 1 + w\frac{1}{12\ln(2)}(a+d)\left(3\ln(2) + (1-\frac{b+c}{a+d})\right.\\ &\left.\left(-2m(5\ln(2)-\ln(3)) + 2m^2(11\ln(2)-4\ln(3)) - 8m^3(2\ln(2)-\ln(3))\right)\right),\end{aligned} \quad (39)$$

$$\begin{aligned}\lambda_{2+2} =& 1 + w\frac{1}{4(8m^2-9m-2)\ln(2)}(a+d)\left(-4\ln(2) + m(\ln(3)-\ln(2)) - 4m^2\ln(3)\right.\\ &\left.+4m^3\ln(3) + \frac{b+c}{a+d}(-m(17\ln(2)+\ln(3)) + 4m^2(4\ln(2)-\ln(3)) - 4m^3\ln(3))\right).\end{aligned} \quad (40)$$

For more complex life cycles, the analytical expressions are too large to be meaningful by naked eye analysis. Therefore, we used a combination of analytical and numerical approaches. After the linearisation with respect to w, the growth rate for any life cycle has the form

$$\lambda_{x_1+\ldots+x_n} = 1 + w\frac{1}{P_1(m)}(a+d)\left(P_2(m) + \frac{b+c}{a+d}P_3(m)\right), \quad (41)$$

where $P_1(m), P_2(m), P_3(m)$ are some polynomials of m of the finite power. We obtained exact expressions for these polynomials using the symbolic algebra software. However, the tracking of all developmental programs means that the computation load grows exponentially with the maximal group size M. For life cycles, such as 3+2+1+1, computation of the polynomials required an extraordinary amount of RAM (> 70 Gb) and the outcome is neither human-tractable nor even printable. This memory constraints is the factor, limiting the maximal group size considered to $M = 7$. Therefore, in our study, we only stored the numerical values of the coefficients of m in $P_1(m), P_2(m), P_3(m)$ and used them to compute λ. With this approach, we are able to compute numerical values of λ with very high accuracy, even if traditional closed form solutions are unavailable.



S	τ	$N(\tau)$	$p(\tau)$	$T(\tau)$
		○●○○○●●		
	1	$(1,0,1,0,0)$	n^4	$t_{[1,0]}+t_{[2,0]}$
	2	$(0,1,1,0,0)$	$2/3\ mn^3$	$t_{[1,0]}+t_{[2,0]}$
	3	$(1,0,0,1,0)$	$4/3\ mn^3$	$t_{[1,0]}+t_{[2,0]}$
	4	$(1,0,0,0,1)$	$1/3\ m^2n^2$	$t_{[1,0]}+t_{[2,0]}$
	5	$(0,1,0,1,0)$	$2/3\ m^2n^2$	$t_{[1,0]}+t_{[2,0]}$
	6	$(0,1,1,0,0)$	$2mn(1/3\ n^2P^A_{[1,1]}+2/3\ mnP^B_{[1,1]})$	$t_{[1,0]}+t_{[1,1]}$
	7	$(1,0,0,1,0)$	$2mn(2/3\ n^2P^A_{[1,1]}+4/3\ mnP^B_{[1,1]})$	$t_{[1,0]}+t_{[1,1]}$
	8	$(1,0,0,0,1)$	$2mn(2/3\ mnP^A_{[1,1]}+1/3\ n^2P^B_{[1,1]})$	$t_{[1,0]}+t_{[1,1]}$
	9	$(0,1,0,1,0)$	$2mn(4/3\ mnP^A_{[1,1]}+2/3\ n^2P^B_{[1,1]})$	$t_{[1,0]}+t_{[1,1]}$
	10	$(0,1,0,0,1)$	$2m^3nP^A_{[1,1]}$	$t_{[1,0]}+t_{[1,1]}$
	11	$(1,0,1,0,0)$	$2m^3nP^B_{[1,1]}$	$t_{[1,0]}+t_{[1,1]}$
	12	$(0,1,0,0,1)$	m^2n^2	$t_{[1,0]}+t_{[0,2]}$
	13	$(1,0,0,0,1)$	$2/3\ m^3n$	$t_{[1,0]}+t_{[0,2]}$
	14	$(0,1,0,1,0)$	$4/3\ m^3n$	$t_{[1,0]}+t_{[0,2]}$
	15	$(0,1,1,0,0)$	$1/3\ m^4$	$t_{[1,0]}+t_{[0,2]}$
	16	$(1,0,0,1,0)$	$2/3\ m^4$	$t_{[1,0]}+t_{[0,2]}$
	17	$(0,1,0,0,1)$	n^4	$t_{[0,1]}+t_{[0,2]}$
	18	$(1,0,0,0,1)$	$2/3\ mn^3$	$t_{[0,1]}+t_{[0,2]}$
	19	$(0,1,0,1,0)$	$4/3\ mn^3$	$t_{[0,1]}+t_{[0,2]}$
	20	$(0,1,1,0,0)$	$1/3\ m^2n^2$	$t_{[0,1]}+t_{[0,2]}$
	21	$(1,0,0,1,0)$	$2/3\ m^2n^2$	$t_{[0,1]}+t_{[0,2]}$
	22	$(0,1,1,0,0)$	$2mn(1/3\ n^2P^A_{[1,1]}+2/3\ mnP^B_{[1,1]})$	$t_{[0,1]}+t_{[1,1]}$
	23	$(1,0,0,1,0)$	$2mn(2/3\ n^2P^A_{[1,1]}+4/3\ mnP^B_{[1,1]})$	$t_{[0,1]}+t_{[1,1]}$
	24	$(1,0,0,0,1)$	$2mn(2/3\ mnP^A_{[1,1]}+1/3\ n^2P^B_{[1,1]})$	$t_{[0,1]}+t_{[1,1]}$
	25	$(0,1,0,1,0)$	$2mn(4/3\ mnP^A_{[1,1]}+2/3\ n^2P^B_{[1,1]})$	$t_{[0,1]}+t_{[1,1]}$
	26	$(0,1,0,0,1)$	$2m^3nP^A_{[1,1]}$	$t_{[0,1]}+t_{[1,1]}$
	27	$(1,0,1,0,0)$	$2m^3nP^B_{[1,1]}$	$t_{[0,1]}+t_{[1,1]}$
	28	$(1,0,1,0,0)$	m^2n^2	$t_{[0,1]}+t_{[2,0]}$
	29	$(0,1,1,0,0)$	$2/3\ m^3n$	$t_{[0,1]}+t_{[2,0]}$
	30	$(1,0,0,1,0)$	$4/3\ m^3n$	$t_{[0,1]}+t_{[2,0]}$
	31	$(1,0,0,0,1)$	$1/3\ m^4$	$t_{[0,1]}+t_{[2,0]}$
	32	$(0,1,0,1,0)$	$2/3\ m^4$	$t_{[0,1]}+t_{[2,0]}$

Fig 5. The developmental trajectories of the single A-cell and B-cell in the life cycle 2+1. White circles represent A type cells and black circles represent B type cells. For simplicity of notation, we use: $n = 1 - m$; $t_{[i,j]}$ is the time before the next cell division for a complex with i A type cells and j B type cells to divide; and $P^A_{[1,1]} = \frac{1}{2} + w\frac{b-c}{4}$ and $P^B_{[1,1]} = \frac{1}{2} - w\frac{b-c}{4}$, see details in the model section of the main text.

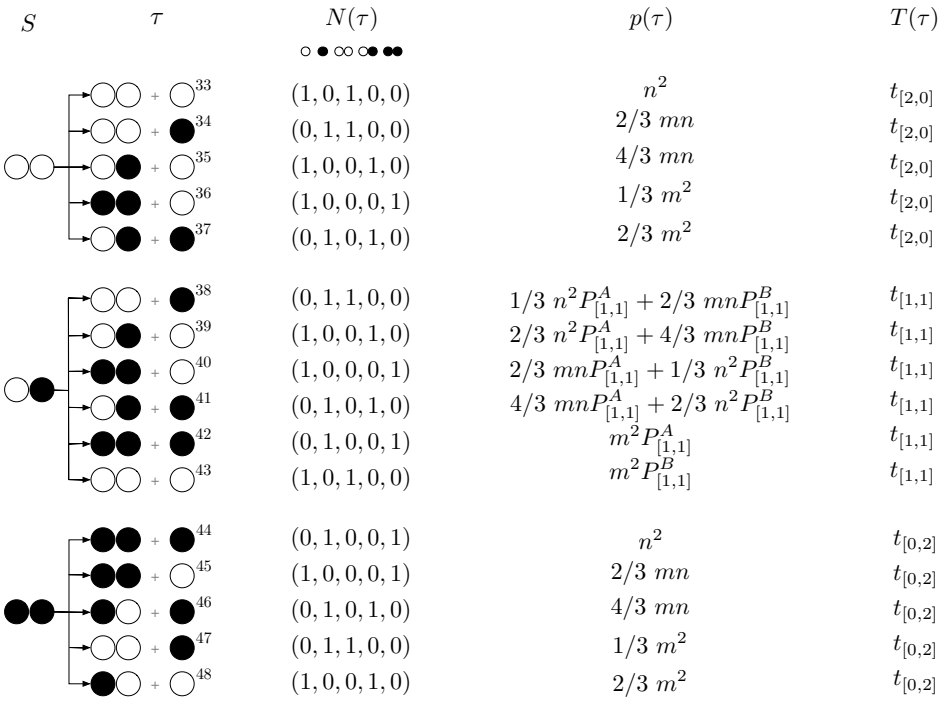


Fig 6. The developmental programs of bicellular newborn groups cells in the life cycle 2+1. White circles represent A type cells and black circles represent B type cells. For simplicity of notation, we use: $n = 1 - m$; $t_{[i,j]}$ is the time before the next cell division for a complex with i A type cells and j B type cells to divide; and $P^A_{[1,1]} = \frac{1}{2} + w\frac{b-c}{4}$ and $P^B_{[1,1]} = \frac{1}{2} - w\frac{b-c}{4}$, see details in the model section of the main text.

Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S5 Appendix.

Profiles of growth rates of the life cycles. In this appendix, we present profiles of growth rates at different conditions. The growth rate is determined by three parameters: ψ, ϕ and m. The greatest diversity of evolutionarily optimal life cycles is observed at $\psi > 0$ and small m, see Fig 7. In this case, we observed two clusters of life cycles, where life cycles behave quite similar. One cluster contain the multiple fission life cycles such as 1+1, 1+1+1. The second cluster is the group propagules life cycles such as 3+2, 4+3. The slope of multiple fission life cycles are increasing with colony size, see Fig 7C. More similar growth rate patterns are observed for the group propagules life cycles, which have identical growth rates at $\phi = 1$, see Fig 7B. Most other life cycles are between the area of multiple fission life cycles and group propagules life cycles, which can never be optimal. For $\psi > 0$ and large m, only one multiple fission life cycle, 1+1+1, is evolutionary optimal, see Fig 8A. Its area of optimality is located between unicellularity (1+1) at large negative ϕ and binary fragmentation with multicellular propagules (2+2 and 4+3) at large positive ϕ. Considering the dependence of growth rate from the phenotype switching probability m, we found that at $\phi \gg 1$, growth rate profiles are concave functions of m, see Fig 8B. Growth rates of most life cycles are generally bound between binary fragmentation with multicellular propagules (such as 2+2 and 4+3) and multiple fragmentation with unicellular propagules (such as 1+1+1 and 1+1+1+1). For $\phi \ll -1$, the pattern is very similar, with an exception, that growth rate profiles are convex, instead of concave, and the hierarchy of life cycles is reversed, see Fig 8C. This leads to the great diversity of evolutionary optimal life cycles at $\phi < 0$ and small m (including also transitional life cycles 2+1 and 2+1+1, as well as binary fragmentation 2+2), see Fig 8D.

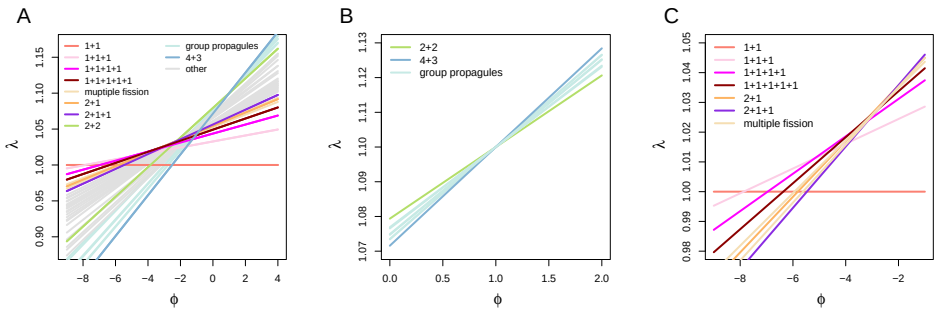


Fig 7. The growth rates of the considered life cycles as a function of ϕ for $\psi > 0$. Panel A: according to the weak selection approximation, growth rates λ are linear functions of ϕ. For all life cycles, the slope of the line is non-negative, thus, life cycles with smaller slope dominate at $\phi \ll -1$ (1+1 has slope zero) and life cycles with larger slope dominate at $\phi \gg 1$ (4+3 has the largest slope for $M \leq 7$). Panel B: all life cycles with multicellular offspring share the same growth rate at $\phi = 1$ ($\phi = -1$ under $\psi < 0$). Panel C: a sequence of multiple fission life cycles is optimal at the negative ϕ. At all panels $m = 0.06$. In all panels, multiple fission includes 1+1+1+1+1+1, 1+1+1+1+1+1+1; group fission includes 3+2, 3+3, 2+2+2, 4+2, 5+2.

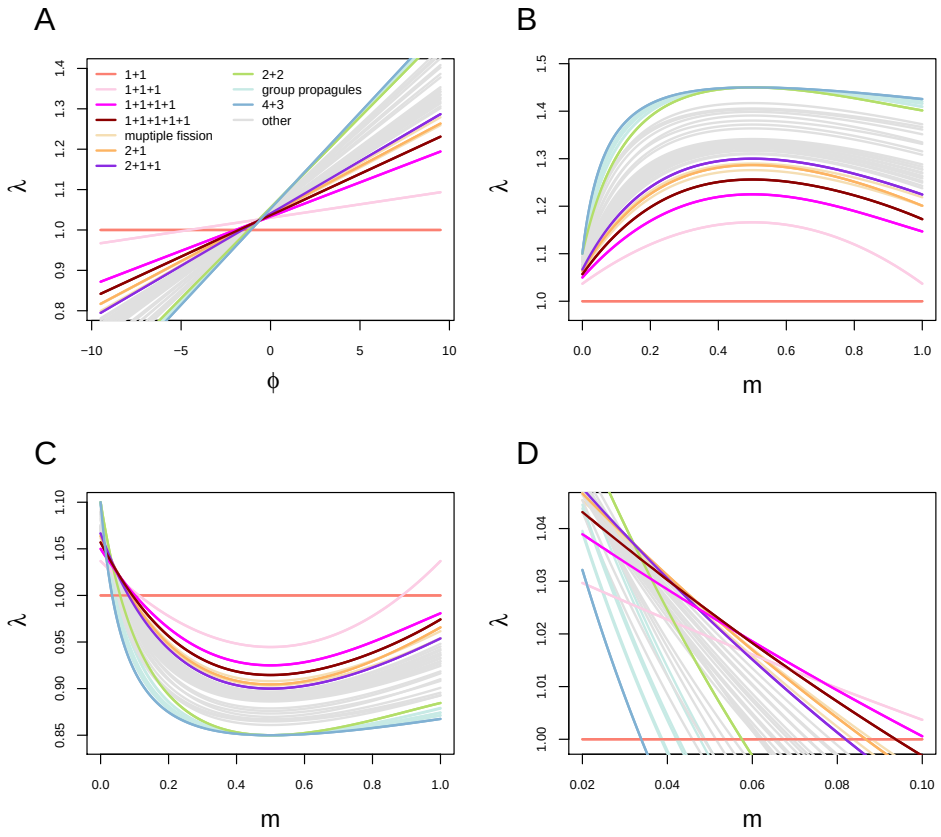


Fig 8. Multiple life cycles are optimal for $\psi > 0$. A Growth rates of all considered life cycles as function of ϕ at $m = 0.9$ (cf. Fig 4A for $m = 0.06$). **B** Growth rates of all considered life cycles as function of m at $\phi = 8$. **C** Growth rates of all considered life cycles as function of m at $\phi = -4$. **D** Detailed view of the panel C in the range of small m showing that large number of evolutionary optimal life cycles at different m.

At the negative ψ, only two life cycles were found to be optimal, see Fig 9. The shape of individual growth rate profiles remain similar to the case of positive ϕ but the relative position changes significantly. Thus, the spectrum of observed life cycles is much less diverse.

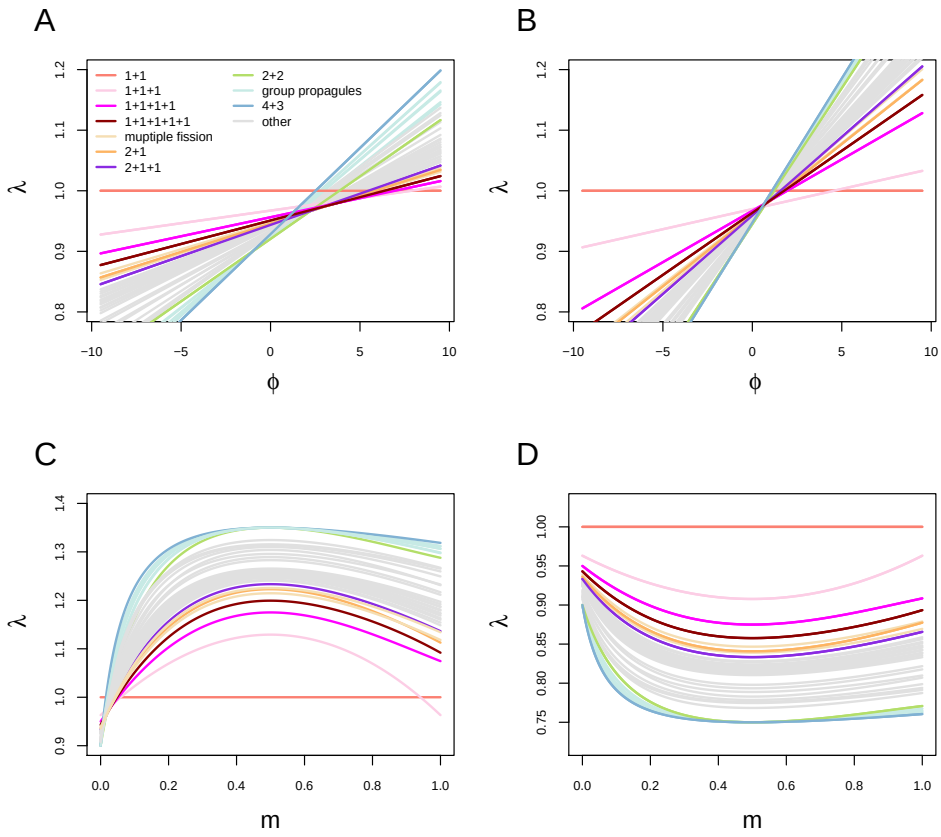


Fig 9. Only two life cycles are optimal for $\psi < 0$. A Growth rates of all considered life cycles as function of ϕ at $m = 0.06$. **B** Growth rates of all considered life cycles as function of ϕ at $m = 0.9$. **C** Growth rates of all considered life cycles as function of m at $\phi = 8$. **D** Growth rates of all considered life cycles as function of m at $\phi = -4$.

Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao[1], Arne Traulsen[1], Yuriy Pichugin[1*]

1 Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

* pichugin@evolbio.mpg.de

Supporting information

S6 Appendix.

Optimal life cycles landscape under the self-interaction game. In our model setting, we set the payoff of single cells to zero based on the assumption that no other cells can impact their strategies. While, theoretically single cells can also play self-interaction games to get payoff based on their cell types. Intuitively, the self-interaction game would produce the same results as the non self-interaction game, as in which only the final synergistic or antagonistic effects can really impact the outcome. To check this idea, we set the payoff of cells in a cluster to

$$\begin{aligned} \alpha_{[i,j]} &= \frac{ia + jb}{i + j}, \\ \beta_{[i,j]} &= \frac{ic + jd}{i + j}, \end{aligned} \tag{8}$$

where $\alpha_{[i,j]}$ and $\beta_{[i,j]}$ are the average payoff of A type cells and B type cells in a group of i A-cells and j B-cells, respectively. This payoff definition allows the single cells also have non zero payoff values i.e. payoff a for A cell type and d for B cell type. Meanwhile, all other settings in the model are unchanged. Then, we investigate the optimal life cycles for population with colony size M less than seven. The results are pretty similar between the non self-interaction game (see Fig 4) and the self-interaction game (see Fig 10).

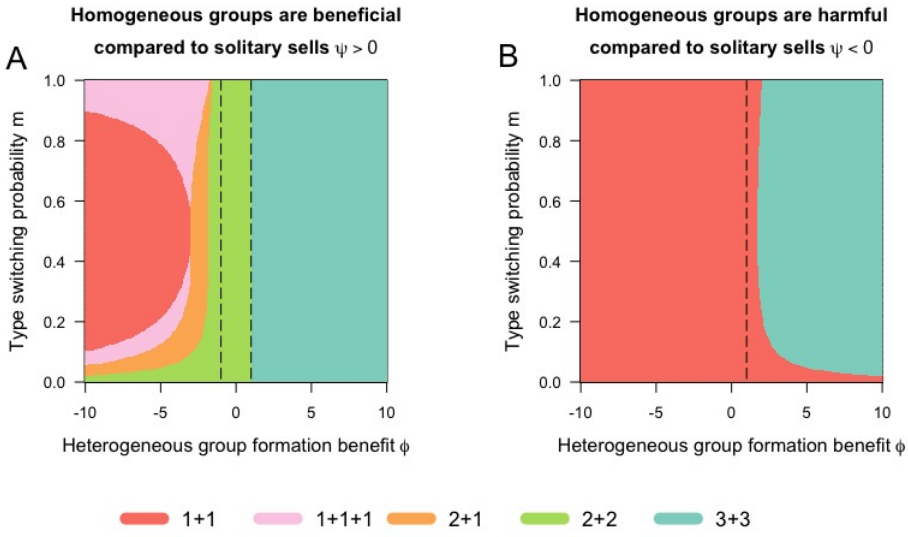


Fig 10. Similar optimal life cycles under the self-interaction game compared with the no self-interaction game. While we have only 5 optimal life cycles in this case, in general the results are very similar, with a large number of life cycles emerging only for $\psi > 0$ and $\phi < -1$. **A** Optimal life cycles for $\psi > 0$ under the self-interaction game. Dashed lines are $\phi = -1$ and $\phi = 1$, respectively. **B** Optimal life cycles for $\psi < 0$ under the self-interaction game. Dashed line is $\phi = 1$.