Interacting cells driving the evolution of multicellular life cycles

Yuanxiao Gao¹, Arne Traulsen¹, Yuriy Pichugin¹*

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 $\lambda$ for life cycles of heterogeneous groups. To show how our approach 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.

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\tau$</th>
<th>$N(\tau)$</th>
<th>$p(\tau)$</th>
<th>$T(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>(2, 0)</td>
<td>$(1 - m)^2$</td>
<td>$\ln(2)$</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>(1, 1)</td>
<td>$2m(1 - m)$</td>
<td>$\ln(2)$</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>(0, 2)</td>
<td>$m^2$</td>
<td>$\ln(2)$</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>(0, 2)</td>
<td>$(1 - m)^2$</td>
<td>$\ln(2)$</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>(1, 1)</td>
<td>$2m(1 - m)$</td>
<td>$\ln(2)$</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>(2, 0)</td>
<td>$m^2$</td>
<td>$\ln(2)$</td>
</tr>
</tbody>
</table>

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

$$Q_{11} = 2(1 - m)^2 e^{-\lambda m^2}$$
The elements of matrix $Q$ are given by

\[ Q_{11} = 3n^4 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_1 \]
\[ + 4m n^3 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_2 \]
\[ + m^2 n^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_3 \]
\[ + 6m^3 n P_{1+1}^B e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_4 \]
\[ + 4m n^2 (n P_{1+1}^A + 2m P_{1+1}^B) e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_5 \]
\[ + 2m n^2 (2m P_{1+1}^A + n P_{1+1}^B) e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_6 \]
\[ + 2m^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_7 \]
\[ + 4m^n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_8 \]
\[ + 2m^3 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_9 \]

\[ Q_{12} = 2m n^3 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_2 \]
\[ + 2m n^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_3 \]
\[ + 2m^2 n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_4 \]
\[ + 4m^3 n P_{1+1}^B e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_5 \]
\[ + 4m^2 n P_{1+1}^A e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_6 \]
\[ + 2m^3 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_7 \]
\[ + 4m^2 n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_8 \]
\[ + 3m^2 n^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_9 \]
\[ Q_{21} = 2m n^3 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_1 \]
\[ + 2m n^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_2 \]
\[ + 2m^2 n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_3 \]
\[ + 4m^3 n P_{1+1}^B e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_4 \]
\[ + 4m^2 n P_{1+1}^A e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_5 \]
\[ + 2m^3 n P_{1+1}^B e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_6 \]
\[ + 2m^3 n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_7 \]
\[ + 4m^2 n^2 e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_8 \]
\[ + 3m^2 n e^{-\lambda t_{1,0} + t_{2,0}} \quad \leftarrow \tau_9 \]
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 \( P_{[1,1]}^A = \frac{1}{2} + w \frac{b-c}{4} \) and \( P_{[1,1]}^B = \frac{1}{2} - w \frac{b-c}{4} \), see details in the model section of the main text.

\[
Q_{22} = 3n^4 e^{-\lambda(t_{[0,1]}+t_{[0,2]})} \leftarrow \tau_{20}
\]

\[
+ 3m^2 n^2 e^{-\lambda(t_{[0,1]}+t_{[2,0]})} \leftarrow \tau_{21}
\]

\[
+ 4m^3 n e^{-\lambda(t_{[0,1]}+t_{[0,2]})} \leftarrow \tau_{22}
\]

\[
+ m^2 n^2 e^{-\lambda(t_{[0,1]}+t_{[2,0]})} \leftarrow \tau_{23}
\]

\[
+ 6m^3 n P_{1+1}^A e^{-\lambda(t_{[0,1]}+t_{[1,1]})} \leftarrow \tau_{24}
\]

\[
+ 4mn^2(2mP_{1+1}^A + nP_{1+1}^B) e^{-\lambda(t_{[0,1]}+t_{[1,1]})} \leftarrow \tau_{25}
\]

\[
+ 2mn^2(nP_{1+1}^A + 2mP_{1+1}^B) e^{-\lambda(t_{[0,1]}+t_{[2,0]})} \leftarrow \tau_{26}
\]

\[
+ 2m^4 e^{-\lambda(t_{[0,1]}+t_{[2,0]})} \leftarrow \tau_{27}
\]

\[
+ 2m^3 n e^{-\lambda(t_{[0,1]}+t_{[2,0]})} \leftarrow \tau_{28}
\]

Here, \( P_{1+1}^A = \frac{1}{2} + w \frac{b-c}{4} \), \( P_{1+1}^B = \frac{1}{2} - w \frac{b-c}{4} \), \( t_{[1,0]} = t_{[1,0]} = t_{[0,1]} = t_{[0,2]} \) and \( t_{[2,0]} = \ln \frac{b+c}{2} (1 - w d) \). Arrows indicate the contributions of
each developmental trajectory to $Q_{ij}$.

The solution of Eq (18) for life cycle 1+1+1 yields
\[
\lambda_{1+1+1} = 1 + w \frac{\ln(3/2)}{\ln(9)} (a + d) \left( (1 - 2m + 2m^2) + 2m(1 - m) \frac{b + c}{a + d} \right).
\] (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 a 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
\[
\lambda_{2+1} = 1 + w \frac{3\ln(3/2)}{2(5 + 8m)\ln(27/4)} (a + d) \left( (5 - 6m + 10m^2) + 2m(7 - 5m) \frac{b + c}{a + d} \right).
\] (38)

For the life cycles 1+1+1+1 and 2+2, we just list the growth rate
\[
\lambda_{1+1+1+1} = 1 + w \frac{1}{12\ln(2)} (a + d) \left( 3\ln(2) + (1 - \frac{b + c}{a + d}) \right)
\]
\[
(-2m(5\ln(2) - \ln(3)) + 2m^2(11\ln(2) - 4\ln(3)) - 8m^3(2\ln(2) - \ln(3)))
\],
\] (39)

\[
\lambda_{2+2} = 1 + w \frac{1}{4(8m^2 - 9m - 2)\ln(2)} (a + d) \left( -4\ln(2) + m(\ln(3) - \ln(2)) - 4m^2\ln(3)
\right)
\]
\[
+ 4m^3\ln(3) + \frac{b + c}{a + d} (-m(17\ln(2) + \ln(3)) + 4m^2(4\ln(2) - \ln(3)) - 4m^3\ln(3))
\].
\] (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
\[
\lambda_{x_1+...+x_n} = 1 + w \frac{1}{P_1(m)} (a + d) \left( P_2(m) + \frac{b + c}{a + d} P_3(m) \right),
\] (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 $\lambda$. With this approach, we are able to compute numerical values of $\lambda$ with very high accuracy, even if traditional closed form solutions are unavailable.
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 $P_{[1,1]}^A = \frac{1}{2} + w^{\frac{-1}{2}}$ and $P_{[1,1]}^B = \frac{1}{2} - w^{\frac{1}{2}}$, 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 $P_{[1,1]}^A = \frac{1}{2} + w_1\frac{A}{4}$ and $P_{[1,1]}^B = \frac{1}{2} - w_1\frac{B}{4}$, see details in the model section of the main text.