Supplementary material for "Evolution of irreversible differentiation under stage-dependent cell differentiation"

Yuanxiao Gao[1,2*], Yuriy Pichugin[2,4], Arne Traulsen[2], and Román Zapién-Campos[2,3]

[1]School of Mathematics and Data Science, Shaanxi University of Science and Technology, 710021 Xi'an, China

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

[3]Centre for Life's Origins and Evolution, University College London, WC1 6BT, London, UK

[4]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

[*]Corresponding author: Yuanxiao Gao, yxgao@sust.edu.cn

February 1, 2025

Contents

1 S1: Population growth rate.

In our model, we treat both the number of cells and growth time as continuous, distilling the stochastic process down to two quantities for calculating the population growth rate: the expected offspring number of germ-like cells N and the amount of growth time for an organism to grow t. The expected population growth rate λ can be calculated by the following equation approximately

$$\lambda = \frac{\ln N}{t}. \tag{1}$$

The robustness of the approximation is tested in S3: Robustness of the population growth rates of stochastic differentiation strategies. Here, we use $f_g^{(i)}$ and $f_s^{(i)}$ to denote the fractions of germ-like cells and soma-like cells after the ith cell division. Since each organism starts with a single germ-like cell, $f_g^{(0)} = 1$ and $f_s^{(0)} = 0$. We use $p_{x\to y}^{(i)}$ to denote the transition probability from cell type x to y in the ith cell division, where x and y are either germ-like cells or soma-like cells. Based on Eq (1) in the main text, we have

$$\begin{aligned} g_{g\to g}^{(i)} &= g_{gg}^{(i)} + \frac{g_{gs}^{(i)}}{2} \\ g_{g\to s}^{(i)} &= g_{ss}^{(i)} + \frac{g_{gs}^{(i)}}{2} \\ s_{s\to g}^{(i)} &= s_{gg}^{(i)} + \frac{s_{gs}^{(i)}}{2} \\ s_{s\to s}^{(i)} &= s_{ss}^{(i)} + \frac{s_{gs}^{(i)}}{2}. \end{aligned} \tag{2}$$

After the ith cell division, the expected fraction of germ-like cells is $f_g^{(i)} = g_{g\to g}^{(i)} f_g^{(i-1)} + s_{s\to g}^{(i)} f_s^{(i-1)}$ and the expected fraction for soma-like cells is $f_s^{(i)} = g_{g\to s}^{(i)} f_g^{(i-1)} + s_{s\to s}^{(i)} f_s^{(i-1)}$, which can be expressed in

$$\begin{pmatrix} f_g^{(i)} \\ f_s^{(i)} \end{pmatrix} = \begin{pmatrix} g_{g\to g}^{(i)} & s_{s\to g}^{(i)} \\ g_{g\to s}^{(i)} & s_{s\to s}^{(i)} \end{pmatrix} \begin{pmatrix} f_g^{(i-1)} \\ f_s^{(i-1)} \end{pmatrix}. \tag{3}$$

The expected $f_g^{(n)}$ and $f_s^{(n)}$ can be calculated recursively by Eq (3)

$$\begin{pmatrix} f_g^{(n)} \\ f_s^{(n)} \end{pmatrix} = \begin{pmatrix} g_{g\to g}^{(n)} & s_{s\to g}^{(n)} \\ g_{g\to s}^{(n)} & s_{s\to s}^{(n)} \end{pmatrix} \cdots \begin{pmatrix} g_{g\to g}^{(i)} & s_{s\to g}^{(i)} \\ g_{g\to s}^{(i)} & s_{s\to s}^{(i)} \end{pmatrix} \cdots \begin{pmatrix} g_{g\to g}^{(1)} & s_{s\to g}^{(1)} \\ g_{g\to s}^{(1)} & s_{s\to s}^{(1)} \end{pmatrix} \begin{pmatrix} f_g^{(0)} \\ f_s^{(0)} \end{pmatrix}. \tag{4}$$

Since cells divide synchronously and no cell dies during growth, the expected number of germ-like cells $N_g^{(n)}$ and soma-like cells $N_s^{(n)}$ after the nth cell division are

$$\begin{pmatrix} N_g^{(n)} \\ N_s^{(n)} \end{pmatrix} = 2^n \begin{pmatrix} f_g^{(n)} \\ f_s^{(n)} \end{pmatrix}, \tag{5}$$

where $0 \leq f_g^{(n)}, f_s^{(n)} \leq 1$.

The cell division rate determines the growth duration of organisms. Since cells divide with a rate $r^{(i)} = \frac{1+b[f_s^{(i-1)}]^\alpha}{1+c[f_{g\to s}^{(i)}+\beta f_{s\to g}^{(i)}]}$ during the ith cell division, the waiting time for a cell division $t^{(i)}$ follows the exponential distribution $f(t^{(i)}) = r^{(i)} e^{-r^{(i)} t^{(i)}}$ ([1]), where $f_{g\to s}^{(i)} = f_g^{(i-1)} s_{s\to g}^{(i)}$ and $f_{s\to g}^{(i)} = f_s^{(i-1)} s_{s\to g}^{(i)}$, see Eq (6) in the main text. Thus the expected waiting time from the ith cell division to the $(i+1)$th cell division is $t^{(i)} = \frac{1}{r^{(i)}}$. The expected growth time for organisms with total n cell divisions is

$$t = \sum_{i=1}^{n} t^{(i)} = \sum_{i=1}^{n} \frac{1}{r^{(i)}} = \sum_{i=1}^{n} \frac{1 + c[f_g^{(i-1)} g_{g\to s}^{(i)} + \beta f_s^{(i-1)} s_{s\to g}^{(i)}]}{1 + b[f_s^{(i-1)}]^\alpha}. \tag{6}$$

Substituting Eq (5) and Eq (6) into Eq (1), we have

$$\lambda = \frac{\ln N}{t} = \frac{n \ln 2 + \ln f_g^{(n)}}{\sum_{i=1}^{n} \frac{1+c\left[f_g^{(i-1)} g_{g\to s}^{(i)} + \beta f_s^{(i-1)} s_{s\to g}^{(i)}\right]}{1+b[f_s^{(i-1)}]^\alpha}}, \tag{7}$$

where n is the number of total cell divisions of organisms, $f_g^{(i)}$ and $f_s^{(i)}$ are fractions of germ-like cell and soma-like cell after the ith cell division, $g_{g\to s}^{(i)}$ and $s_{s\to g}^{(i)}$ are the transition probabilities between germ-like cell and soma-like cell at the ith cell division ($1 \leq i \leq n$). We have $f_g^{(0)} = 1$ and $f_s^{(0)} = 0$. For the non-differentiation strategy ND^i, no soma-like cells are produced during growth, i.e. $g_{g\to g} = 1$ and $g_{g\to s} = s_{s\to g} = s_{s\to s} = 0$. Therefore, $f_g^{(i)} = 1$, $f_s^{(i)} = 0$. Thus from Eq (7) the population growth rate of ND^i which is denoted by λ_{ND^i} is $\ln 2$. Biologically, the population growth rate of ND^i describes the number of cells doubling per unit of time. As we defined strategies based on the series of cell differentiation probabilities, the population growth rate of a strategy should be a distribution rather than a fixed value. But for calculation convenience, we took Eq (7) as an approximation of a stochastic differentiation strategy. In S3: Robustness of the population growth rates of stochastic differentiation strategies, we show the approximation is reliable in finding the optimal differentiation strategy.

2 S2: Numerical calculation of the population growth rate of the stage-dependent differentiation strategy.

We introduce the method of calculating population growth rate numerically in stage-dependent cell differentiation. To find the optimal strategy, at a fixed benefit and cost condition, we use the Monte-Carlo methods randomly to sample differentiation strategies and then calculate and compare the population growth rates

of organisms under these strategies. Grid search is used to find the optimal strategy in different conditions of benefits and costs. We first look at the cell differentiation probabilities in the first division step $d^{(1)} = [g_{gg}^{(1)}, g_{gs}^{(1)}, g_{ss}^{(1)}, s_{gg}^{(1)}, s_{gs}^{(1)}, s_{ss}^{(1)}]$. Each probability can take the value 0 or other values by increasing 0.1 from 0 at each time until reaching the highest value 1, thus there are 11 possible values for each probability, i.e $0, 0.1, 0.2, \ldots, 1$. We first define the number of probability combinations. Since $g_{gg}^{(1)} + g_{gs}^{(1)} + g_{ss}^{(1)} = 1$ and $g_{ss}^{(1)} = 1 - g_{gg}^{(1)} + g_{gs}^{(1)}$, as long as we know the values of $g_{gg}^{(1)}$ and $g_{gs}^{(1)}$, we know g_{ss}. When $g_{gg} = 0$, g_{gs} can take the 11 values from 0 to 1, Thus, there are totally $\sum_{i=1}^{11} i = 66$ combinations for $g_{gg}^{(1)}, g_{gs}^{(1)}, g_{ss}^{(1)}$. The same number of combinations exist for soma-like cells. Thus, there are a total of $66 \times 66 = 4356$ combinations for $d^{(1)}$. As long as $d^{(1)}$ is chosen, we need to identify $d^{(2)}$, $d^{(3)}, \ldots, d^{(n)}$. $d^{(2)}$ deviates from $d^{(1)}$ by either a 0 or δ_2. That is $0 \leq |g_{gg}^{(1)} - g_{gg}^{(2)}| \leq \delta_2$. The same for the other probabilities $g_{gs}, g_{ss}, s_{gg}, s_{gs}, s_{ss}$. The choice of $d^{(i+1)}$ depends on number of neighbours of $d^{(i)}$, which further depends on the elements $d^{(i)}$. For $\delta^{(i+1)} = 0.1$, if $g_{gg} = s_{ss} = 1$ in $d^{(i)}$, then g_{gg} and s_{ss} can only be decreased or be constant. Thus, there are 5 choices for $d^{(i+1)}$. However, if the elements in $d^{(i)}$ are either 0.3 or 0.4, then each element can be increased, decreased, or unchanged. Therefore, it has 13 choices for choosing $d^{(i+1)}$. Let's take $d_1 = [0.3, 0.3, 0.4, 0.3, 0.3, 0.4]$ as an example, then it's neighbours are

$$[0.3, 0.3, 0.4, 0.3, 0.3, 0.4],$$
$$[0.4, 0.2, 0.4, 0.3, 0.3, 0.4],$$
$$[0.4, 0.3, 0.3, 0.3, 0.3, 0.4],$$
$$[0.2, 0.4, 0.4, 0.3, 0.3, 0.4],$$
$$[0.2, 0.3, 0.5, 0.3, 0.3, 0.4],$$
$$[0.3, 0.4, 0.3, 0.3, 0.3, 0.4],$$
$$[0.3, 0.2, 0.5, 0.3, 0.3, 0.4],$$
$$[0.3, 0.3, 0.4, 0.4, 0.2, 0.4],$$
$$[0.3, 0.3, 0.4, 0.4, 0.3, 0.3],$$
$$[0.3, 0.3, 0.4, 0.2, 0.4, 0.4],$$
$$[0.3, 0.3, 0.4, 0.2, 0.3, 0.5],$$
$$[0.3, 0.4, 0.3, 0.3, 0.3, 0.4],$$
$$[0.3, 0.3, 0.4, 0.3, 0.2, 0.5].$$

Note that $d^{(1)}$ is considered as one of its neighbours. To generate a stage-dependent differentiation strategy, we first chose $d^{(1)}$ from the combination pool and then chose $d^{(2)}$ from $d^{(1)}$'s neighbors and repeat the process until obtaining $d^{(n)}$. We choose each strategy randomly following a uniform distribution. As long as we have classified strategies, we will have a pool of each strategy and then we choose strategies from the pools.

Specificity, we first choose the last probabilities and then randomly choose other probabilities backward in rounds of cell division. For example, for choosing a RD strategy, we first randomly pick the probabilities at the nth round of cell division, which should satisfy $0 < g_{g\to s}^{(n)}, s_{s\to g}^{(n)} < 1$. Then we randomly choose the probabilities at the $(n-1)$th round of cell division and so on until the first one.

In stage-independent cell differentiation, we calculate the population growth rates of each strategy in the cell differentiation probabilities pool. We seek the optimal strategy which leads to the fastest growing among these 4356 strategies. To find the optimal strategy at a given parameter point, we first chose $M = 1000$ values for $d^{(1)}$ from the cell differentiation pool. Then for each chosen $d^{(1)}$, we randomly chose $R = 100$ stage-dependent strategies, all generated from this $d^{(1)}$. $R = 100$ is the sampling size of the stage-dependent strategies from the same initial differentiation probabilities $d^{(1)}$. Then we compute the population growth rate of the 10^5 strategies and choose the strategy leading to the largest population growth rate. Next, we optimize that strategy further. For the optimal strategy with the largest population growth rate, we compare its population growth rate with a slightly modified strategy. The modified strategies include the one removing $d^{(1)}$ but compensating with an $d^{(n+1)}$ or removing $d^{(n)}$ by compensating with an $d^{(0)}$. Specifically, for the focused $D = [d^{(1)}, d^{(2)}, \ldots, d^{(n)}]$, we check whether $D' = [d^{(0)}, d^{(1)}, d^{(2)}, \ldots, d^{(n-1)}]$ or $D' = [d^{(2)}, \ldots, d^{(n)}, d^{(n+1)}]$ leads to a higher population growth rate over D. Here the $d^{(0)}$ is one neighbour of $d^{(1)}$, and $d^{(n+1)}$ is one neighbour of $d^{(n)}$. If the D' leads to a higher population growth rate, we keep the process until we find the D' which makes the population growth rate stay at the maxima. We aim to find a local optimum close to the strategy that was identified in our grid search. Local optimization stops when the largest steady population growth rate in the local neighbourhood is identified. Overall, we first search the optimal D globally by randomly choosing $d^{(1)}$, represented by M and R. The values of M and R and the number of duplications used in the main text were chosen to ensure the optimal strategy converging to a unique strategy. Then, we used a local grid search by modifying $d^{(1)}$ or $d^{(n)}$ of a strategy until finding the optimal D. Besides, we have constructed initial sampling strategies from the middle of $d^{(i)}$ sequences. We first identified $d^{(\frac{n+1}{2})}$ if n is odd and $d^{(\frac{n}{2})}$ if n is even, and then constructed the rest $d^{(i)}$s. The results show that there is almost no differences in terms of searching for optimal strategies between the two methods.

3 S3: Robustness of the population growth rates of stochastic differentiation strategies.

In our model, we calculated the population growth rate of a stochastic strategy based on its expected growth time and expected number of germ-like cells. That is, we treat a stochastic differentiation strategy that may contain many potential developmental trajectories as a deterministic one. Theoretically, the population growth rate of a stochastic strategy should be a random variable. Next, we show that the method used in the model is a

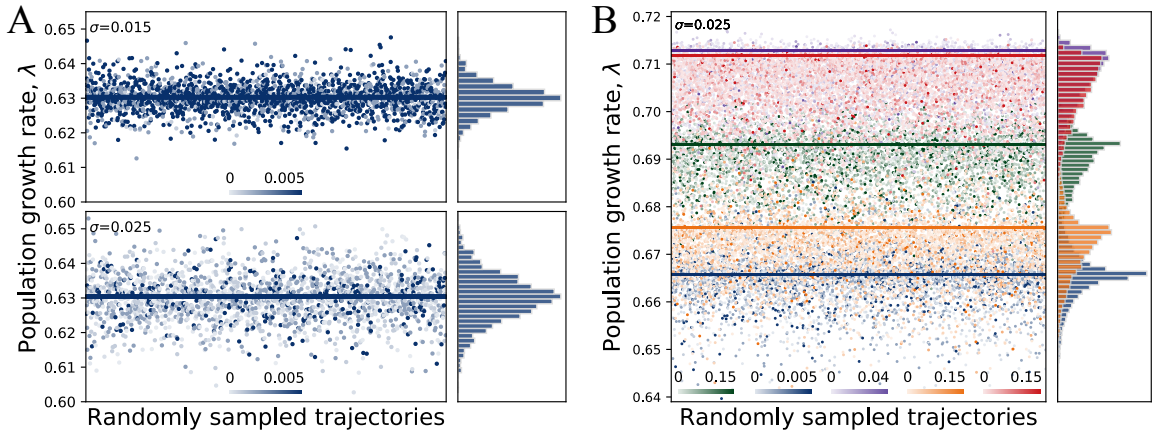

Figure 1: **Comparison of population growth rates by approximation and random sampling. A.** Population growth rate comparison of potential random trajectories (strategies) of a randomly chosen RD strategy under Gaussian distribution with variance 0.015 and 0.025 respectively. The small blue dots represent the potential trajectories. The lines represent the expected population growth rates calculated based on Eq (7). The color of the dots represents the probability of the randomly chosen RD choosing the dots. The histograms represent the distribution frequency of λ. **B.** Population growth rate of randomly sampled optimal strategies of each category (ND^i, RD, IGD, ISD and $IGSD$). The optimal strategy is obtained based on the calculation in S1: Population growth rate and the grid search method in S2: Numerical calculation of the population growth rate of the stage-dependent differentiation strategy. The color of the dots represents the probability of the optimal given strategy choosing the strategy. The histograms represent the distribution frequency of λ. The colors represent the same strategy as that in Fig 3 in the main text. Parameters for all panels $\delta = 0.05$, $n = 5$ and $b = c = 1$. For calculating the population growth rate of each strategy, see the appendix S2: Numerical calculation of the population growth rate of the stage-dependent differentiation strategy.

good approximation for seeking the optimal strategy in an average sense. To simulate the consecutive stochastic differentiation probabilities, at a given stage we need to know the differentiation probability distribution that the next consecutive probabilities follow. Without loss of generality, we assume that the differentiation probabilities follow Gaussian distribution. Then, the coming cell differentiation probability of a cell type is a variable with the last past differentiation probability as the mean. For an arbitrary strategy $D = [d^{(1)}, d^{(2)}, \ldots, d^{(n)}]$, we can get $g^{(i)}_{g\to s}$ and $s^{(i)}_{s\to g}$ for each i, $i = 0, 1, \ldots, n$. Then the variable $g^{(i+1)}_{g\to s}$ follows the Gaussian distribution $g^{(i+1)}_{g\to s} \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu = g^{(i)}_{g\to s}$. To capture the population growth rate of the stochastic differentiation strategy D, we randomly choose the new $g^{(i)}_{g\to s}$ and the new $s^{(i)}_{s\to g}$ from the Gaussian distribution with mean $g^{(i)}_{g\to s}$ and $s^{(i)}_{s\to g}$ respectively, $i = 1, 2, 3, \ldots, n$. Each sampling will generate a new strategy D^*, which is a potential developmental strategy based on D. For each D^*, we can calculate its population growth rate based on Eq (7). Then we adopt the Monte Carlo method to capture the potential population growth rate distribution by randomly choosing a large number of D^* and calculating their population growth rate. Based on our numerical calculation,

we found that our approximation is along well with the expected population growth rate of a randomly chosen strategy (Fig 1A). The value of variance σ is undefined. As here we focus on the mean behavior of a strategy, thus variance only impacts the range of population growth rate. Furthermore, we testified whether the conclusion under the approximation is consistent with the statistical results introduced above. We found that the optimal strategies are the same (Fig 3 in the main text and Fig 1B), indicating the robustness of the approximation method. However, we should note the expected population growth rate of a stochastic differentiation strategy may not be equal to our approximation. The former is $\sum^{k} \lambda_k p_k$, where k is the all possible trajectories of D^*, p_k is the corresponding probability of choosing trajectory k, and λ_k is the population growth rate under trajectory k. p_k is multiplication of p_k^i which is the probability of choosing a differentiation probability for either germ-like cell $g_{g\to s}^{(i)}$ or soma-like cell $s_{s\to g}^{(i)}$ in D^*, where $i = 1, 2, \ldots, n$. In the numerical calculation (Fig 1), we roughly classify 8 intervals i.e. 8 different probabilities for generating $g_{g\to s}^{(i)}$ or $s_{s\to g}^{(i)}$ for a given i. The 8 intervals are classified based on boundaries of $\mu + j * \sigma$, where $j = -3, -2, -1, 1, 2, 3$. As we seek the optimal strategy, which depends on the relative difference between different strategies i.e. the rank of the population growth rate of different strategies, we employ the approximation to seek the optimal strategy in the model.

4 S4: Robustness of the updated cell differentiation benefit function.

In the section, we prove that the updated power function for cell differentiation cost F_C does not significantly impact the results. We get the same conclusion as that in the previous model [2]. We found that stage-independent irreversible differentiation ID^i (ISD) emerges at the intermediate values of cell differentiation benefits and cell differentiation costs (Fig 2). Meanwhile, ISD is preferred in large organisms than in small organisms.

5 S5 Appendix. Optimality of non-differentiation strategy ND^i.

ND^i is optimal in the absence of cell differentiation benefits for any maximal cell division number n.

When the cell differentiation benefit is absent, i.e. $b = 0$ and $c > 0$, we find that ND^i is the optimal strategy based on Eq (7).

ND^i is optimal in the absence of costs for $n = 1$.

Next, when there is only one cell division ($n = 1$), we prove that ND^i leads to the largest population growth rate under $b > 0$ and $c = 0$. Since $f_g^{(0)} = 1$ and $f_s^{(0)} = 0$, based on Eq (6), the growth time is

$$t = \sum_{i=1}^{n=1} t^{(i)} = \frac{1}{r^{(1)}} = \frac{1}{1 + b[f_s^{(0)}]^{\alpha}} = 1. \tag{8}$$

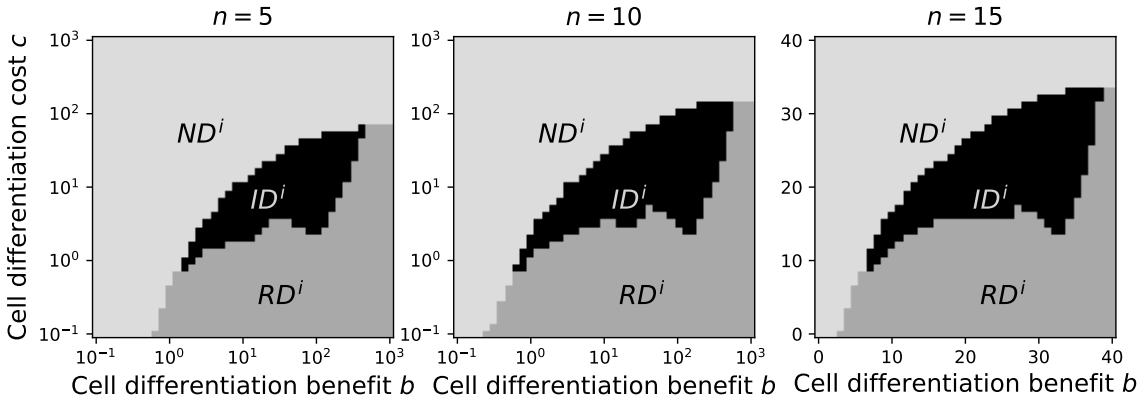

Figure 2: **The emergence parameter space for stage-independent irreversible differentiation strategy ISD under cell differentiation benefits b and cell differentiation costs c.** The emergence conditions for each stage-independent cell differentiation strategy are optimal under cell differentiation benefits and cell differentiation costs with varying maximum cell division $n = 5$, $n = 10$, and $n = 15$. Here, the ISD strategy is the only irreversible differentiation among stage-dependent strategies. Parameters for all panels $\delta = 0.1$, $\alpha = \beta = 1$. For calculating the population growth rate of each strategy, see the appendix 2.

Substituting Eq (8) into Eq (7) and using Eq (4), we have

$$\lambda = \frac{\ln N}{t} = \ln 2 + \ln f_g^{(1)} = \ln 2 + \ln g_{g\to g}^{(1)}. \tag{9}$$

Since $0 \leq g_{g\to g}^{(1)} \leq 1$, the optimal strategy is ND^i which has $g_{g\to g}^{(1)} = 1$. Thus, ND^i is the optimal strategy under $b > 0$, $c = 0$ and $n = 1$.

6 S6: Optimal strategies of $n = 5$ under a larger range of parameter space.

Optimal strategies of $n = 5$ under a larger range of parameter space. Here, we show that RD is optimal when benefits are far larger than costs, see Fig 3. ND^i is optimal when differentiation costs are far larger than benefits.

7 S7: Stage-independent strategies lead to higher population growth rates than stage-dependent strategies.

We found that stage-dependent strategies lead to higher population growth rates than stage-independent strategies. Stage-dependent differentiation leads to higher standard deviations in the fractions of germ-like cells, cell

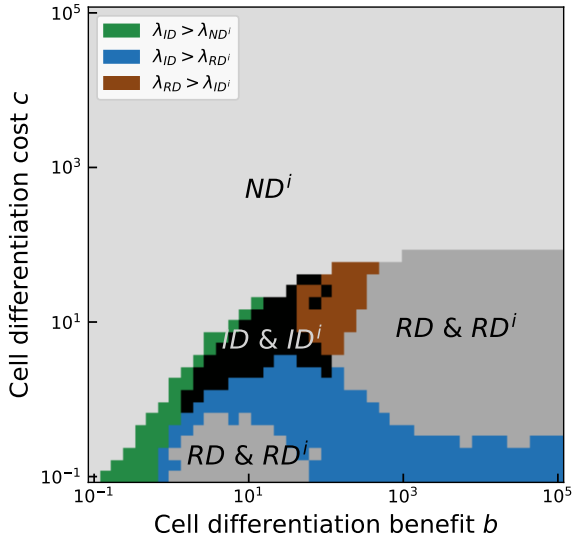

Figure 3: **Comparison of the optimal strategies between stage-independent differentiation and stage-dependent differentiation under the large scale of benefits and costs.** The colors show the same meaning as that in Fig 2 in the main text. Parameters: $0 \leq \delta_i \leq 0.1$, $\alpha = \beta = 1$, and $n = 5$. Parameters of calculating optimal strategy: the number of initial sampling $d^{(1)}$, $M = 1000$, the number of stage-dependent strategies starting with a given $d^{(1)}$, $R = 100$, for more detail, see S2: Numerical calculation of the population growth rate of the stage-dependent differentiation strategy. At each pixel, the frequency of each optimal strategy was calculated across 20 replicates.

division rates, and population growth rates than stage-independent differentiation, see Fig 4. Since cells with the same type divide with the same probabilities in stage-independent differentiation, the fractions of the germ-like cell of different categories approach stationary states over time, see Fig 4A. For example, the fractions of germ-like cells in a RD^i strategy approach a constant value, while the fractions in ID^i decrease monotonically to 0. However, the fractions of the germ-like cells of different categories under stage-dependent differentiation fluctuate irregularly and the patterns are more complex, see Fig 4B. For example, the fractions of germ-like cells have a minimum value in the middle of cell divisions in IGD on average. This pattern shows that IGD invests into soma-like cells first and later lets soma-like cells contribute to offspring i.e. produce germ-like cells. Compared with ID^i, ID leads to three very different developmental trajectories, where IGD keeps higher fractions of germ-like cells, ISD has the lowest fractions and $IGSD$ has intermediate fractions. Moreover, we found that stage-independent ID^i and ISD both lead to higher standard deviations in population growth rates. This is because ID^i and ISD contain many strategies without producing any offspring. Based on Eq (7), we know when $f_g^{(n)} < \frac{1}{2^n}$, then $\lambda < 0$.

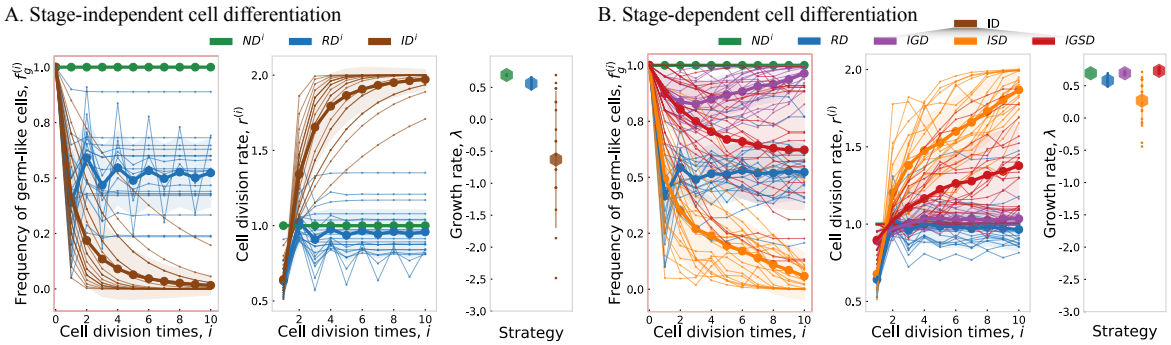

Figure 4: Frequencies of the germ-like cells, cell division rates, and population growth rates of randomly chosen stage-independent strategies: ND^i, RD^i, and ID^i (**A**) and stage-dependent strategies: ND^i, RD, and ID (IGD, ISD, and $IGSD$) (**B**). We should note that only the fraction of germ-like cells after the last division $f_g^{(n)}$ and the cell division rates contribute to the population growth rate. Thick dots and shaded areas (error bars in the last panels) are the averaged values and the standard deviation of the strategies in each category, respectively. Parameters in both panels: $n = 10$, $b = c = 1$, $0 \leq \delta \leq 0.1$. We chose 20 random strategies in each category and then recorded the frequency of germ-like cells, the cell division rates after each cell division, together with their final population growth rates (thin lines). For the details of choosing the random samples of strategies and the numerical calculation of the population growth rate, see S2: Numerical calculation of the population growth rate of the stage-dependent differentiation strategy.

8 S8: Either IGD or RD is optimal in the absence of cell differentiation costs when maximal cell division $n > 1$.

To show the optimal strategy is either IGD or RD under $b > 0$ and $c = 0$. We first prove $\lambda_{IGD} > \lambda_{IGSD} > \lambda_{ND^i}$, and then prove $\lambda_{RD} > \lambda_{ISD} > \lambda_{ND^i}$.

The strategy IGD is optimal among IGD, $IGSD$ and ND^i.

To prove $\lambda_{IGD} > \lambda_{IGSD} > \lambda_{ND^i}$, we begin with the proof of $\lambda_{IGSD} > \lambda_{ND^i}$. Unlike the ND^i strategy, ISD, RD, IGD and $IGSD$ are categories which include many strategies. As long as one strategy in $IGSD$ has a greater population growth rate than ND^i, we say $IGSD$ is more optimal than ND^i. Think of an $IGSD$ strategy with only a non-zero cell differentiation probability from germ-like cells to soma-like cells and zero differentiation probabilities the other way around. Let's assume it is the ith cell division that makes $g_{g\to s}^{(i)} > 0$, thus we have $g_{g\to s}^{(j)} = 0$ for $j \neq i$ and $s_{s\to g}^{(i)} = 0$ for any i. Based on Eq (4), the cell frequencies after the nth

division are

$$\begin{aligned}\begin{pmatrix} f_g^{(n)} \\ f_s^{(n)} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} g_{g\to g}^{(i)} & 0 \\ g_{g\to s}^{(i)} & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f_g^{(0)} \\ f_s^{(0)} \end{pmatrix} \\ &= \begin{pmatrix} g_{g\to g}^{(i)} \\ g_{g\to s}^{(i)} \end{pmatrix},\end{aligned} \tag{10}$$

where $f_g^{(0)} = 1$ and $f_s^{(0)} = 0$. For convenience, we denote $g_{g\to g}^{(i)} = g^*$ and $g_{g\to s}^{(i)} = 1 - g^*$. From Eq (7), we have

$$\begin{aligned}\lambda_{IGSD} &= \frac{n\ln 2 + \ln g^*}{\frac{n-1}{1+b} + \frac{1}{1+b(1-g^*)^\alpha}} \\ &\geq \frac{n\ln 2 + \ln g^*}{\frac{n+b}{1+b}} \\ &= \ln 2 + \frac{(n-1)b}{n+b}\ln 2 + \frac{1+b}{n+b}\ln g^* \\ &= \lambda_{ND^i} + \frac{\ln 2^{(n-1)b}(g^*)^{1+b}}{n+b},\end{aligned} \tag{11}$$

where we use $(1 - g^*) \geq 0$ to obtain the inequality and $\lambda_{ND^i} = \ln 2$ (appendix 1). Since $b > 0$ and $n > 1$, as long as $2^{(n-1)b}(g^*)^{1+b} \geq 1$, we obtain $\lambda_{IGSD} > \lambda_{ND^i}$. That is $g^* \geq \frac{1}{2^{\frac{(n-1)b}{1+b}}}$. Therefore, when $b > 0$, we can always find an $IGSD$ strategy with a $g_{g\to s}^{(i)} \leq 1 - \frac{1}{2^{\frac{(n-1)b}{1+b}}}$ and all other $g_{g\to s}^{(i)} = 0$ and $s_{s\to g}^{(i)} = 0$, which leads to higher population growth rate than ND^i. Thus, $\lambda_{IGSD} > \lambda_{ND^i}$. The proof of $\lambda_{IGD} > \lambda_{IGSD}$ is in the appendix 9. Taken these together, we have $\lambda_{IGD} > \lambda_{IGSD} > \lambda_{ND^i}$.

RD is optimal among RD, ISD and ND^i.

Next, we prove $\lambda_{RD} > \lambda_{ISD} > \lambda_{ND^i}$. We first prove $\lambda_{ISD} > \lambda_{ND^i}$. We prove that there exists an ISD strategy leading to a higher λ than $\lambda_{ND^i} = \ln 2$ (appendix 1). Consider the ISD with $s_{s\to s}^{(i)} = 1$, but with at least one i which makes $g_{g\to s}^{(i)} > 0$ i.e. $g_{g\to g}^{(i)} = 1 - g_{g\to s}^{(i)} < 1$ for $1 \leq i \leq n$. The above constraint corresponds with the definition of the ISD strategy. Based on Eq (4), the cell frequencies after the nth division are

$$\begin{aligned}\begin{pmatrix} f_g^{(n)} \\ f_s^{(n)} \end{pmatrix} &= \begin{pmatrix} g_{g\to g}^{(n)} & 0 \\ g_{g\to s}^{(n)} & 1 \end{pmatrix} \cdots \begin{pmatrix} g_{g\to g}^{(i)} & 0 \\ g_{g\to s}^{(i)} & 1 \end{pmatrix} \cdots \begin{pmatrix} g_{g\to g}^{(1)} & 0 \\ g_{g\to s}^{(1)} & 1 \end{pmatrix} \begin{pmatrix} f_g^{(0)} \\ f_s^{(0)} \end{pmatrix} \\ &= \begin{pmatrix} \prod_{i=1}^n g_{g\to g}^{(i)} \\ 1 - \prod_{i=1}^n g_{g\to g}^{(i)} \end{pmatrix},\end{aligned} \tag{12}$$

where $n \geq 1$. Substituting Eq (12) into Eq (7) and together with $c = 0$, we find the population growth rate of the ISD strategy

$$\lambda_{ISD} = \frac{n \ln 2 + \ln \prod_{i=1}^{n} g_{g\to g}^{(i)}}{1 + \sum_{i=2}^{n} \frac{1}{1 + b[1 - \prod_{k=1}^{i-1} g_{g\to g}^{(k)}]^{\alpha}}}, \tag{13}$$

where the first item in the denominator represents the time for the first cell division $t_1 = 1$ because of $f_s^{(0)} = 0$. We define the second item of the denominator of Eq (13) as $F(n) = \sum_{i=2}^{n} f^{(i)}$, where $f^{(i)} = \frac{1}{1 + b[1 - \prod_{k=1}^{i-1} g_{g\to g}^{(k)}]^{\alpha}}$. Next, we prove $F(n)$ is a bounded function. Since $0 \leq g_{g\to g}^{(k)} \leq 1$, thus sequence $\{\prod_{k=1}^{i-1} g_{g\to g}^{(k)}\}_{i=2}^{\infty}$ decreases with increasing i. That is, $\{f^{(i)}\}_{i=2}^{\infty}$ is a positive but decreasing sequence. $f^{(2)}$ is the largest one in $\{f^{(i)}\}_{i=2}^{\infty}$. Therefore, the sequence $\{F(n) = \sum_{i=2}^{n} f^{(i)}\}$ is an accelerating discrete sequence with respect to n. We have

$$F(n) \leq (n-1) f^{(2)} = \frac{n-1}{1 + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}}. \tag{14}$$

Substituting the right-hand side of inequality (14) into Eq (13), we have

$$\begin{aligned}
\lambda_{ISD} &= \frac{n \ln 2 + \ln \prod_{i=1}^{n} g_{g\to g}^{(i)}}{\sum_{i=1}^{n} \frac{1}{1 + b(1 - \prod_{i=1}^{n} g_{g\to g}^{(i)})^{\alpha}}} \\
&\geq \frac{n \ln 2 + \ln \prod_{i=1}^{n} g_{g\to g}^{(i)}}{1 + \frac{n-1}{1 + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}}} \\
&= \frac{\left(n \ln 2 + \ln \prod_{i=1}^{n} g_{g\to g}^{(i)}\right)\left[1 + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}\right]}{n + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}} \\
&= \ln 2 + \frac{(n-1) b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha} \ln 2 + \left[1 + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}\right] \ln \prod_{i=1}^{n} g_{g\to g}^{(i)}}{n + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}}.
\end{aligned} \tag{15}$$

As long as there exist a strategy which makes the right side of Eq (15) greater than $\ln 2$, we have $\lambda_{ND^i} < \lambda_{ISD}$. Then we need to identify the conditions for

$$(n-1) b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha} \ln 2 + \left(1 + b\left(1 - g_{g\to g}^{(1)}\right)^{\alpha}\right) \ln \prod_{i=1}^{n} g_{g\to g}^{(i)} > 0 \tag{16}$$

to hold. As $b > 0$ and $\alpha > 0$, $(n-1)b\left(1-g_{g\to g}^{(1)}\right)^{\alpha}\ln 2 \geq 0$. Since $0 \leq g_{g\to g}^{(i)} \leq 1$, $\prod_{i=1}^{n} g_{g\to g}^{(i)} \leq 1$ and $\ln\prod_{i=1}^{n} g_{g\to g}^{(i)} \leq 0$. The second item of Eq (16) is negative. There exists a sequence $\{g_{g\to g}^{(i)}\}$, which makes $\prod_{i=1}^{n} g_{g\to g}^{(i)} \to 1^{-}$ and

$$\begin{aligned}\ln\prod_{i=1}^{n} g_{g\to g}^{(i)} &> -\frac{(n-1)b\left(1-g_{g\to g}^{(1)}\right)^{\alpha}\ln 2}{1+b\left(1-g_{g\to g}^{(1)}\right)^{\alpha}} \\ &= -\frac{(n-1)\ln 2}{1+\frac{1}{b\left(1-g_{g\to g}^{(1)}\right)^{\alpha}}},\end{aligned} \tag{17}$$

which makes the Eq (16) hold. With the above proof, we conclude that $\lambda_{ISD} > \lambda_{ND^i}$ with only cell differentiation benefit. From Eq (17), we found that more ISD strategies are better than ND^i under high benefits b. However, when b is small, only ISD with $g_{g\to g}^{(i)} \to 1$ leads higher population growth rate than ND^i. The proof of $\lambda_{RD} > \lambda_{ISD}$ can be found in the appendix 9. Thus, we have $\lambda_{RD} > \lambda_{ISD} > \lambda_{ND^i}$. The results show that when there is a benefit and no costs, differentiation strategies (ISD, $IGSD$, IGD and RD) are better over ND^i. Either IGD or RD is optimal under $b > 0$ and $c = 0$.

9 S9: $IGSD$ and ISD cannot be optimal in the absence of either cell differentiation benefit or cost.

In the appendix 5, we have proved ND^i is optimal in the absence of differentiation benefits, i.e. $b = 0$ and $c > 0$. Thus, we prove $IGSD$ and ISD can be optimal in the absence of differentiation costs, i.e. $c = 0$ and $b > 0$. Since we also have proved that ND^i is optimal under $n = 1$ when $c = 0$ and $b > 0$ in appendix 5. Therefore, we only need to prove that the optimal strategy can neither be $IGSD$ nor ISD when $b > 0$, $c = 0$ and $n \geq 2$.

We first prove $\lambda_{IGD} > \lambda_{IGSD}$. For a given $IGSD$ strategy, we can always modify it and obtain an IGD strategy, which leads to a higher λ than the given $IGSD$ strategy. For a given $IGSD$ strategy, we know its transition probabilities $s_{s\to g}^{(n)} = 0$. We modify the $IGSD$ strategy by setting $0 < s_{s\to g}^{(n)} = k \leq 1$ to get a IGD strategy. The constructed IGD strategy produces more offspring than the given $IGSD$ strategy as its final number of germ-like cells is $N = 2^n(f_g^{(n-1)}g_{g\to g}^{(n)} + f_s^{(n-1)}s_{s\to g}^{(n)})$, which is greater than that of the $IGSD$ as $s_{s\to g}^{(n)} = k > 0$ in the IGD strategy. Since there is no cell differentiation cost ($c = 0$), cell division rates are the same among all strategies. Thus, $\lambda_{IGD} > \lambda_{IGSD}$.

Next, we prove $\lambda_{RD} > \lambda_{ISD}$. Given an ISD strategy, we have $s_{s\to g}^{(n)} = 0$. Construct a RD strategy that has the same transition probability matrixes as the given ISD strategy for the first $(n-1)$ cell divisions. For the nth transition probability matrix, we keep the $g_{g\to g}^{(n)}$ and $g_{g\to s}^{(n)}$ the same as that in the given ISD strategy. However, we set $s_{s\to g}^{(n)} > 0$ rather than $s_{s\to g}^{(n)} = 0$ as that in ISD strategy. $s_{s\to g}^{(n)} > 0$ implies $s_{s\to s}^{(n)} = 1 - s_{s\to g}^{(n)} <$

1. Then, ISD and RD have the same germ-like cells during the first $n-1$ cell divisions. The fraction of germ-like cells for the ISD strategy after the nth cell divisions is $f_g^{(n-1)} g_{g\to g}^{(n)} + f_s^{(n-1)} s_{s\to g}^{(n)} = f_g^{(n-1)} g_{g\to g}^{(n)}$ as $s_{s\to g}^{(n)} = 0$. Whereas, the fraction of germ-like cells for the RD strategy after the nth cell divisions is $f_g^{(n-1)} g_{g\to g}^{(n)} + f_s^{(n-1)} s_{s\to g}^{(n)}$. Thus, the constructed RD has an extra $2^n f_s^{(n-1)} s_{s\to g}^{(n)}$ germ-like cells compared with the ISD strategy. Since the cell division rate at the nth cell division depends on the fraction of soma-like cells at the $(n-1)$th cell division, the cell division rates $r^{(i)}$ for the two strategies are the same, $1 \le i \le n$. Thus, from Eq (7), we have

$$\begin{aligned}
\lambda_{RD} &= \frac{\ln\{2^n[f_g^{(n-1)} g_{g\to g}^{(n)} + f_s^{(n-1)} s_{s\to g}^{(n)}]\}}{\sum_{i=1}^{n} \frac{1}{1+b[f_s^{(i-1)}]^\alpha}} \\
&> \frac{\ln\{2^n[f_g^{(n-1)} g_{g\to g}^{(n)}]\}}{\sum_{i=1}^{n} \frac{1}{1+b[f_s^{(i-1)}]^\alpha}} \\
&= \lambda_{ISD}.
\end{aligned} \tag{18}$$

Therefore, $\lambda_{RD} > \lambda_{ISD}$.

10 S10: Stage-dependent differentiation promotes irreversible cell differentiation under the effects of benefit function forms α and the ratio of differentiation costs between germ-like cells and soma-like cells β.

Under the effects of α and β, we found that stage-dependent differentiation favors irreversible cell differentiation over stage-independent cell differentiation. IGD replaces stage-independent RD when α and β are both small, see Fig 5. Under this scenario, the cell transition probability $s_{s\to g}$ has a smaller effect in decreasing the population growth rate than the transition probability $g_{g\to s}$. Thus, IGD produces a higher fraction of germ-like cells and bears less cell differentiation costs, leading to a higher population growth rate. When α is around 1, IGD leads to faster growth than ND^i. The reason is analogous to the one given in the main text.

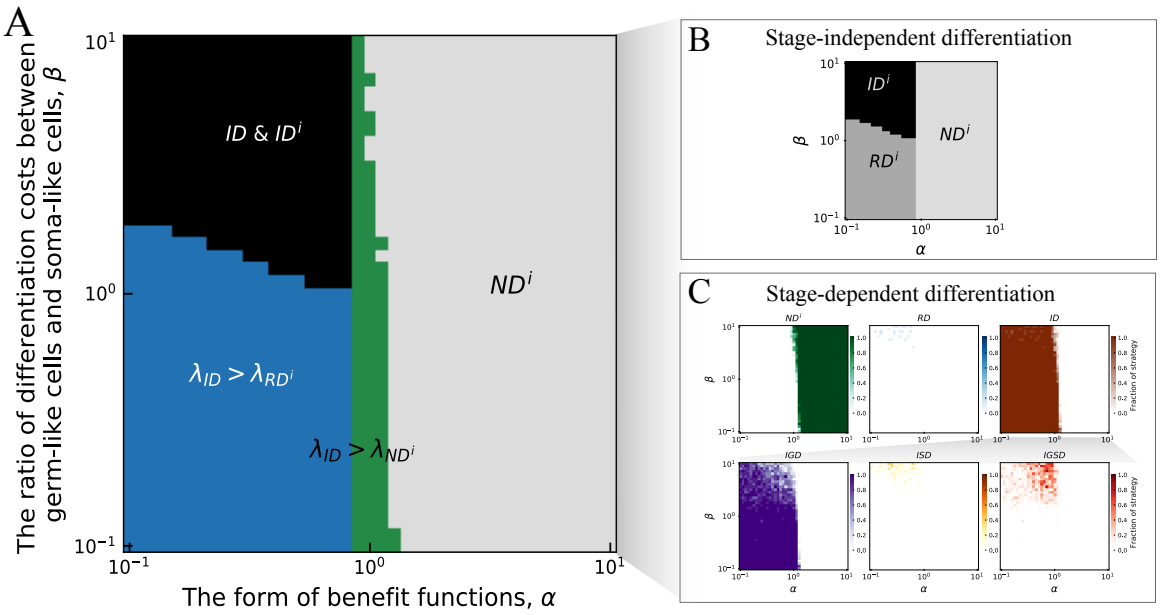

Figure 5: **The effects of α and β on the population growth rates of cell differentiation strategies. A.** Comparison of the optimal strategy evolved in stage-independent and stage-dependent differentiation strategies depending on α and β. The areas of grey and black represent the parameter space in which the same strategy are optimal both under stage-independent and stage-dependent cell differentiation. The green area represents stage-dependent ID leading to a larger population growth rate than stage-independent ND^i. The blue strip represents stage-dependent ID leading to a larger population growth rate than stage-independent RD^i. **B.** The parameter space of optimal stage-independent differentiation strategy at different values of α and β. **C.** The frequencies of each stage-dependent strategy depending on α and β. Parameters for all panels $\delta = 0.1$, $n = 5$ and $b = c = 1$. For calculating the population growth rate of each strategy, see the appendix 2.

References

[1] Linda JS Allen. *An introduction to stochastic processes with applications to biology*. CRC press, 2010.

[2] Yuanxiao Gao, Hye Jin Park, Arne Traulsen, and Yuriy Pichugin. Evolution of irreversible somatic differentiation. *eLife*, 10:e66711, 2021.