Supplemental Materials: Population dynamics of cancer cells with cell-state interconversions

Da Zhou[1], Dingming Wu[1], Zhe Li[2], Minping Qian[3] and Michael Q. Zhang[4,1,*]

1. MOE Key Laboratory of Bioinformatics; Bioinformatics Division/Center for Synthetic & Systems Biology, TNLIST; Department of Automation, Tsinghua University, Beijing 100084, China.
2. Computational Neuroscience Lab, School of Medicine, Tsinghua University, Beijing 100084, China.
3. School of Mathematical Sciences, Peking University, Beijing 100871, China.
4. Department of Molecular and Cell Biology, Center for Systems Biology, The University of Texas at Dallas, Richardson, TX 75080, USA

* Corresponding Author: Michael Q. Zhang, Department of Molecular and Cell Biology, Center for Systems Biology, The University of Texas at Dallas, Richardson, TX 75080, USA.
Email: michael.zhang@utdallas.edu

Running Title: Cancer stem cell model with cell-state conversions.

The models

Model of three cell states

We consider a population of three cell states: CSC, NSCC_1, NSCC_2. By integrating the conventional hierarchical CSC model with cell-state conversions, there are twelve cellular reactions included in our model:
1) CSC $\xrightarrow{\alpha_1}$ CSC+CSC;
2) CSC $\xrightarrow{\alpha_2}$ CSC+NSCC_1;
3) CSC $\xrightarrow{\alpha_3}$ CSC+NSCC_2;
4) CSC $\xrightarrow{\alpha_4}$ $\emptyset$.
5) NSCC_1 $\xrightarrow{\beta_1}$ NSCC_1+NSCC_1;
6) NSCC_1 $\xrightarrow{\beta_2}$ $\emptyset$;
7) NSCC_1 $\xrightarrow{\beta_3}$ CSC;
8) NSCC_1 $\xrightarrow{\beta_4}$ NSCC_2;
9) NSCC_2 $\xrightarrow{\gamma_1}$ NSCC_2+NSCC_2;
10) NSCC_2 $\xrightarrow{\gamma_2}$ $\emptyset$;
11) NSCC_2 $\xrightarrow{\gamma_3}$ CSC;
12) NSCC_2 $\xrightarrow{\gamma_4}$ NSCC_1.
Denote the numbers of CSCs, NSCC_1 and NSCC_2 as S_t $N_t^{(1)}$ and $N_t^{(2)}$ respectively, then the dynamics of this population can be described by a system of ordinary differential equations (ODEs) of (S_t, N_t^1, N_t^2)

$$\begin{cases} \dfrac{dS_t}{dt} = (\alpha_1 - \alpha_4)S_t + \beta_3 N_t^{(1)} + \gamma_3 N_t^{(2)} \\ \dfrac{dN_t^{(1)}}{dt} = \alpha_2 S_t + (\beta_1 - \beta_2 - \beta_3 - \beta_4)N_t^{(1)} + \gamma_4 N_t^{(2)} \\ \dfrac{dN_t^{(2)}}{dt} = \alpha_3 S_t + \beta_4 N_t^{(1)} + (\gamma_1 - \gamma_2 - \gamma_3 - \gamma_4)N_t^{(2)} \end{cases} \tag{1}$$

To fit the model to the data on the cell-state proportions, one often converts the population model to a proportion one. Let $Z_t = S_t + N_t^{(1)} + N_t^{(2)}$ to be the total population number, then $s_t = S_t/Z_t$, $n_t^{(1)} = N_t^{(1)}/Z_t$ and $n_t^{(2)} = N_t^{(2)}/Z_t$ are the proportions of CSCs, NSCC_1 and NSCC_2 respectively. The dynamics of cell-state proportions can be captured by the following nonlinear ODEs

$$\begin{cases} \dfrac{ds_t}{dt} = (\alpha_1 - \alpha_4)s_t + \beta_3 n_t^{(1)} + \gamma_3 n_t^{(2)} - s_t[A_1 s_t + A_2 n_t^{(1)} + A_3 n_t^{(2)}] \\ \dfrac{dn_t^{(1)}}{dt} = \alpha_2 s_t + (\beta_1 - \beta_2 - \beta_3 - \beta_4)n_t^{(1)} + \gamma_4 n_t^{(2)} - n_t^{(1)}[A_1 s_t + A_2 n_t^{(1)} + A_3 n_t^{(2)}] \\ \dfrac{dn_t^{(2)}}{dt} = \alpha_3 s_t + \beta_4 n_t^{(1)} + (\gamma_1 - \gamma_2 - \gamma_3 - \gamma_4)n_t^{(2)} - n_t^{(2)}[A_1 s_t + A_2 n_t^{(1)} + A_3 n_t^{(2)}] \end{cases} \tag{2}$$

where $A_1 = \alpha_1 + \alpha_2 + \alpha_3 - \alpha_4$, $A_2 = \beta_1 - \beta_2$, $A_3 = \gamma_1 - \gamma_2$. Note that $s_t + n_t^{(1)} + n_t^{(2)} = 1$, so one of the equations is redundant. Let $n_t^{(2)} = 1 - s_t - n_t^{(1)}$, then the ODEs in Eq. **2** reduce to

$$\begin{cases} \dfrac{ds_t}{dt} = -(A_1 - A_3)s_t^2 - (A_2 - A_3)s_t n_t^{(1)} + A_4 s_t + A_5 n_t^{(1)} + \gamma_3 \\ \dfrac{dn_t^{(1)}}{dt} = -(A_2 - A_3)[n_t^{(1)}]^2 - (A_1 - A_3)s_t n_t^{(1)} + A_6 s_t + A_7 n_t^{(1)} + \gamma_4 \end{cases} \tag{3}$$

where $A_4 = (\alpha_1 - \alpha_4) - (\gamma_1 - \gamma_2) - \gamma_3$, $A_5 = \beta_3 - \gamma_3$, $A_6 = \alpha_2 - \gamma_4$, and $A_7 = (\beta_1 - \beta_2 - \beta_3 - \beta_4) - (\gamma_1 - \gamma_2) - \gamma_4$.

Model of two cell states

When focusing on the general cell-lineage relation between CSCs and NSCCs, we simplify the model to a population of only two cell states: CSC and NSCC. Then we consider the following reaction schemes:

CSC $\xrightarrow{\alpha_1}$ CSC+CSC;
CSC $\xrightarrow{\alpha_2}$ CSC+NSCC;
CSC $\xrightarrow{\alpha_4}$ $\emptyset$.
NSCC $\xrightarrow{\beta_1}$ NSCC+NSCC;
NSCC $\xrightarrow{\beta_2}$ $\emptyset$;
NSCC $\xrightarrow{\beta_3}$ CSC;
Denote the number of CSCs at time t as S_t , and the number of NSCCs as N_t. We call the model **bidirectional model** when $\beta_3 > 0$:

$$\begin{cases} \dfrac{dS_t}{dt} = (\alpha_1 - \alpha_4)S_t + \beta_3 N_t \\ \dfrac{dN_t}{dt} = \alpha_2 S_t + (\beta_1 - \beta_2 - \beta_3)N_t \end{cases} \tag{4}$$

When $\beta_3 = 0$, we have **unidirectional model**:

$$\begin{cases} \dfrac{dS_t}{dt} = (\alpha_1 - \alpha_4)S_t \\ \dfrac{dN_t}{dt} = \alpha_2 S_t + (\beta_1 - \beta_2)N_t \end{cases} \tag{5}$$

Note that $Z_t = S_t + N_t$ is the total number of the population, then $s_t = S_t/Z_t$ is the proportion of CSCs, and $n_t = 1 - s_t$ is the proportion of NSCCs. From Eq. **4** we have

$$\begin{cases} \dfrac{ds_t}{dt} = -As_t^2 + Bs_t + \beta_3, \\ \dfrac{dZ_t}{dt} = Z_t[As_t + (\beta_1 - \beta_2)]. \end{cases} \tag{6}$$

where $A = (\alpha_1 + \alpha_2 - \alpha_4) - (\beta_1 - \beta_2)$ and $B = (\alpha_1 - \alpha_4) - (\beta_1 + \beta_3 - \beta_2)$. Note that

$$A - B = \alpha_2 + \beta_3 > 0,$$

A is always larger than B. We term the equation

$$\frac{ds_t}{dt} = -As_t^2 + Bs_t + \beta_3, \tag{7}$$

the **proportion model**. When $\beta_3 > 0$, it corresponds to the bidirectional model; when $\beta_3 = 0$, it is for the unidirectional model.

The relation between our model and Markov chain model

To explain the phenotypic equilibrium in cell-state mixture in breast cancer cell lines, Gupta *et al.* introduced a Markov chain model of stochastic transitions between different phenotypic states of cancer cells (1). Now we discuss the relation between our model and theirs. We will show as follows that **their model can be a specific case of our framework only when the nonlinear terms in our model vanish**.

If the cell-state dynamics can be described as a continuous-time Markov chain, let q_{ij} denote the transition rate from cell state i to state j, then the Kolmogorov forward equation can be given by:

$$\begin{cases} \dfrac{dP_1(t)}{dt} = q_{11}P_1(t) + q_{21}P_2(t) + q_{31}P_3(t) \\ \dfrac{dP_2(t)}{dt} = q_{12}P_1(t) + q_{22}P_2(t) + q_{32}P_3(t) \\ \dfrac{dP_3(t)}{dt} = q_{13}P_1(t) + q_{23}P_2(t) + q_{33}P_3(t). \end{cases} \tag{8}$$

When the number of the cell population is large enough, by the Law of Large Numbers (2), $P_i(t)$ can be seen as the proportion of cell state i at time t. Namely, if we term CSCs, NSCC_1 and NSCC_2 as cell state

1, 2 and 3 respectively, then $P_1(t)$, $P_2(t)$ and $P_3(t)$ are respectively equivalent to s_t, $n_t^{(1)}$ and $n_t^{(2)}$ in Eq. **2**. Note that $P_3(t) = 1 - P_1(t) - P_2(t)$, similar to Eq. **3**, we have

$$\begin{cases} \dfrac{dP_1(t)}{dt} = (q_{11} - q_{31})P_1(t) + (q_{21} - q_{31})P_2(t) + q_{31} \\ \dfrac{dP_2(t)}{dt} = (q_{12} - q_{32})P_1(t) + (q_{22} - q_{32})P_2(t) + q_{32} \end{cases} \tag{9}$$

Note that in Eq. **3**, if we let $A_1 - A_3 = 0$ and $A_2 - A_3 = 0$, i.e. $A_1 = A_2 = A_3$, the equations become linear as follows:

$$\begin{cases} \dfrac{ds_t}{dt} = A_4 s_t + A_5 n_t^{(1)} + \gamma_3 \\ \dfrac{dn_t^{(1)}}{dt} = A_6 s_t + A_7 n_t^{(1)} + \gamma_4 \end{cases} \tag{10}$$

By comparing the mathematical forms of Eqs. **9** and **10**, we can find that they are equivalent to each other. That is, the Markov model can be equivalent to our model only when the nonlinear terms vanish. Note that A_1, A_2 and A_3 are the net growth rates of the population contributed by CSCs, NSCC_1 and NSCC_2 respectively, our result indicates that, **if different cell states equally contribute to the whole population growth, the changes of cell-state proportions can be equivalently described as a Markov chain with cell-state transitions**.

Theoretical analysis of the proportion model

Stability analysis

Based on the proportion model

$$\frac{ds_t}{dt} = -As_t^2 + Bs_t + \beta_3,$$

we will show that both the bidirectional and unidirectional models can show unique stable fixed points in the region of $(0, 1)$. That is, both the two models can display phenotypic equilibria in cell-state mixture.

Bidirectional model. In this case, $\beta_3 > 0$: $s_{eq} = -\beta_3/B$, $\frac{B+\sqrt{B^2+4A\beta_3}}{2A}$ and $\frac{B-\sqrt{B^2+4A\beta_3}}{2A}$ are the unique stable fixed points in $(0, 1)$ for $A = 0$, $A > 0$ and $A < 0$ respectively.

Unidirectional model. In this case, $\beta_3 = 0$:

1) When $A = 0$, $s_{eq} = 0$ is the unique stable state. That is, the proportion of CSCs will tend to zero, while NSCCs will take up the whole population.

2) When $A \neq 0$,

$$\frac{ds_t}{dt} = -As_t^2 + Bs_t = s_t(-As_t + B), \tag{11}$$

There are two fixed points in this case: $s_{eq} = 0$ and B/A. Only when $B > 0$, $s_{eq} = B/A$ is the unique stable fixed point in $(0, 1)$. Otherwise, $s_{eq} = 0$ is the unique stable fixed point.

Therefore, in the unidirectional model, there can still exist unique stable fixed point in $(0, 1)$ provided $B > 0$ (note that A is always larger than B, so $A>0$ also holds), and **CSCs proportion will tend to its final limit in** $(0, 1)$ **for any nonzero initial proportions**. Note that $B = (\alpha_1 - \alpha_4) - (\beta_1 - \beta_2)$, $B > 0$ implies that the net growth rate contributed by the symmetric cell division of CSCs is larger than that by NSCCs.

Transient analysis

We now compare the transient dynamics of bidirectional and unidirectional models. We will show that the two models differ in their transient dynamics especially starting from a very small proportion of CSCs.

Unidirectional model. According to Eq. **3**, the initial increase rate of CSCs proportion is $-As_0^2 + Bs_0$, where s_0 is the initial value of CSCs proportion. When s_0 is very small, $-As_0^2 + Bs_0$ approximates to zero.

That is to say, the model predicts a slow initial increase. As it increasing, the increase rate of the CSCs proportion is getting faster, and then gradually return to its final equilibrium.

Bidirectional model. The initial increase rate of CSCs now equals to $-As_0^2 + Bs_0 + \beta_3$. Compared to the slow initial increase in CSCs proportion predicted by the unidirectional model, the bidirectional model shows a transient rapid increase especially when β_3 is bounded away from zero. In this way, there forms a disparity in the initial growth rate between the two models.

This difference between the two models thus can be used to distinguish the bidirectional model from the unidirectional one. Suppose the unique stable points of the two models are both located at s^*, if the two models start from any nonzero state, then

$$\lim_{t\to\infty} s_t^{uni} = s^* = \lim_{t\to\infty} s_t^{bi},$$

where s_t^{uni} and s_t^{bi} are the solutions of the unidirectional and bidirectional models respectively. That is, the disparity between the two models will gradually shrink as time passes. Therefore, it is important to estimate the time before which the two dynamics can be distinguishable. For any ϵ, define the **characteristic time** as

$$t^*(\epsilon) := \max_t\{t : |s_t^{uni} - s_t^{bi}| \geq \epsilon\}.$$

When ϵ is small enough, $t^*(\epsilon)$ characterizes the time after which the two model would be essentially indistinguishable. Therefore, the difference between the two models before the characteristic time can be used to investigate the existence of cell-state conversion from NSCCs to CSCs.

Nonlinear term

It was shown that the nonlinear term plays a crucial role in determining if our proportion model can be reduced into a Markov chain. In particular, when the nonlinear term $A = (\alpha_1 + \alpha_2 - \alpha_4) - (\beta_1 - \beta_2)$ equals to zero, the proportion model in Eq. **7** can be equivalently described as a Markov chain of cell-state transitions between CSCs and NSCCs.

In fact, A being zero implies that the trajectories predicted by the proportion model are exponential-like curves (Fig. 1). The proportion model will reduce to a linear one

$$\frac{ds_t}{dt} = Bs_t + \beta_3 \tag{12}$$

by letting $A = 0$. Its solution can be given by

$$s_t = (s_0 + \frac{\beta_3}{B})\exp(Bt) - \frac{\beta_3}{B},$$

where s_0 is the initial value of s_t, and it is easy to see that s_t is an exponential function of time t. Furthermore, if we consider the dynamics of the total population $Z_t = S_t + N_t$ and from Eq. **6** we have

$$\frac{dZ_t}{dt} = Z_t[As_t + (\beta_1 - \beta_2)]. \tag{13}$$

When A tends to zero, this equation will also become a linear one

$$\frac{dZ_t}{dt} = Z_t(\beta_1 - \beta_2), \tag{14}$$

its solution will also be an exponential function

$$Z(t) = Z_0 \exp\{(\beta_1 - \beta_2)t\}.$$

This means that when $A = 0$, the total population grows exponentially with constant rate.

When $A \neq 0$, to investigate the geometry of the dynamics, we consider the second derivative of s_t

$$\frac{d^2 s_t}{dt^2} = (B - 2As_t)(-As_t^2 + Bs_t + \beta_3). \tag{15}$$

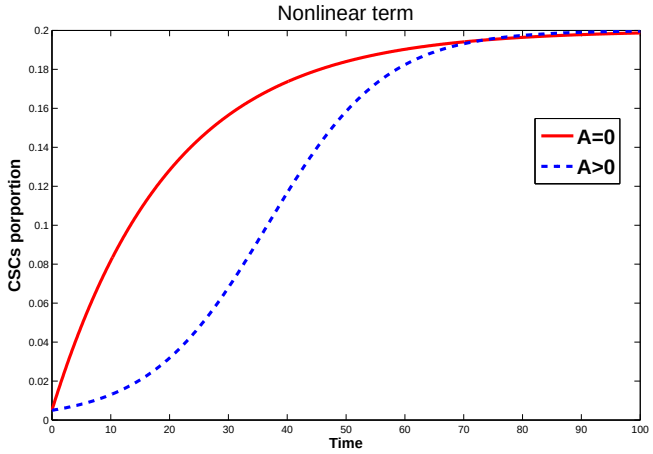

Figure 1: When $A = 0$, the trajectory is exponential-like, whereas the trajectory will be sigmoid-like when $A \neq 0$ (e.g. $A > 0$)

Suppose $A > B > 0$ (the analysis of the other cases are similar), there are two zero points $s = \frac{B}{2A}$ and $\frac{B+\sqrt{B^2+4A\beta_3}}{2A}$ in the region of $(0, 1)$. Starting from very small s_0, it is easy to see that s_t is a sigmoid-like curve, the critical point is $s = \frac{B}{2A}$, at which s_t turns from convex to concave (Fig. 1). Therefore the existence of nonlinear term also greatly affects the geometry of the trajectories, making **the populations grow as sigmoid-like curves**.

Simulations

To illustrate and intuitively show our theoretical results, we now consider the pseudo time-series data produced by Monte Carlo simulation method. Based on the six cellular reactions in the two cell-state model, we produced two cases of data: In case 1, we set $\alpha_1 = 0.2$, $\alpha_2 = 0.1$, $\alpha_4 = 0.1$, $\beta_1 = 0.25$, $\beta_2 = 0.05$, $\beta_3 = 0.07$ and $A = (\alpha_1 + \alpha_2 - \alpha_4) - (\beta_1 - \beta_2) = 0$; In case 2, we set $\alpha_1 = 0.6$, $\alpha_2 = 0.4$, $\alpha_4 = 0.1$, $\beta_1 = 0.25$, $\beta_2 = 0.05$, $\beta_3 = 0.005$, and $A = 0.74 > 0$. In each case, 20 stochastic trajectories were produced, and then we averaged them into one mean trajectory. Fig. 2 shows that in both cases, the trajectories tend to the stable equilibrium at about 0.4. However, they differs in their transient behavior. Note that $A = 0$ in case 1, corresponding to exponential-like curve; In case 2, $A > 0$, it is shown that the trajectory is just sigmoid-like as our theoretical result predicted.

References

[1] P.B. Gupta, C.M. Fillmore, G. Jiang, S.D. Shapira, K. Tao, C. Kuperwasser, and E.S. Lander. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146(4):633–644, 2011.

[2] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, Belmont, CA, USA, 2005.

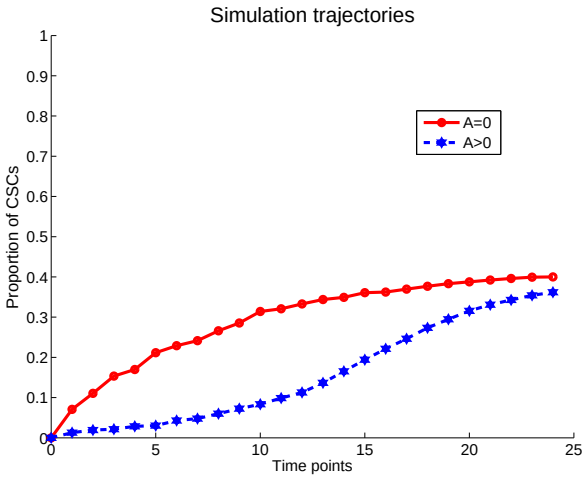

Figure 2: The red continuous line was produced by case 1, where $\alpha_1 = 0.2$, $\alpha_2 = 0.1$, $\alpha_4 = 0.1$, $\beta_1 = 0.25$, $\beta_2 = 0.05$, and $\beta_3 = 0.07$ (A=0); The blue dashed line was produced by case 2, where $\alpha_1 = 0.6$, $\alpha_2 = 0.4$, $\alpha_4 = 0.1$, $\beta_1 = 0.25$, $\beta_2 = 0.05$, and $\beta_3 = 0.005$ (A=0.74>0). In each case, 20 stochastic trajectories were produced by Monte Carlo simulations, and then we averaged them into one mean trajectory shown here.