Eur. Phys. J. E (2013) **36**: 65
DOI 10.1140/epje/i2013-13065-4

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

Model on cell movement, growth, differentiation and de-differentiation: Reaction-diffusion equation and wave propagation

Mao-Xiang Wang[1,3], Yu-Jung Li[1], Pik-Yin Lai[1,a], and C.K. Chan[1,2,b]

[1] Department of Physics, Graduate Institute of Biophysics, and Center for Complex Systems, National Central University, Chungli, Taiwan 320, R.O.C.
[2] Institute of Physics, Academia Sinica, Nankang, Taipei, Taiwan 115, R.O.C.
[3] School of Science, Nanjing University of Science and Technology, Nanjing 210094, China

Received 11 September 2012 and Received in final form 17 April 2013
Published online: 27 June 2013 –

Abstract. We construct a model for cell proliferation with differentiation into different cell types, allowing backward de-differentiation and cell movement. With different cell types labeled by state variables, the model can be formulated in terms of the associated transition probabilities between various states. The cell population densities can be described by coupled reaction-diffusion partial differential equations, allowing steady wavefront propagation solutions. The wavefront profile is calculated analytically for the simple pure growth case (2-states), and analytic expressions for the steady wavefront propagating speeds and population growth rates are obtained for the simpler cases of 2-, 3- and 4-states systems. These analytic results are verified by direct numerical solutions of the reaction-diffusion PDEs. Furthermore, in the absence of de-differentiation, it is found that, as the mobility and/or self-proliferation rate of the down-lineage descendant cells become sufficiently large, the propagation dynamics can switch from a steady propagating wavefront to the interesting situation of propagation of a faster wavefront with a slower waveback. For the case of a non-vanishing de-differentiation probability, the cell growth rate and wavefront propagation speed are both enhanced, and the wavefront speeds can be obtained analytically and confirmed by numerical solution of the reaction-diffusion equations.

1 Introduction

The ability of cells or bacteria to divide and migrate often determines the properties of the spatial patterns in tissue structures and bacterial colonies. The process of wound healing [1,2] is a well-studied example in which the interplay of cell division and cell movement governs the healing pattern and speed. In another perspective, the ability of cells to differentiate to a different cell type is of utmost importance for carrying out specific biological functions in multi-cellular organisms. For example, stem cells can often differentiate into a certain lineage of cells under suitable environment and totipotent stem cells can even differentiate into several different lineages. For a system composed of single-cell organisms such as bacteria, the system can grow only by simple proliferation resulting in an increase in the population size. However for higher-level multi-cellular organisms, the growth can proceed within a single organism that is governed by the interplay of cell proliferation, differentiation, and regulated by cell apoptosis. The difference in growth mode between single cells and multi-cellular organisms can be appreciated in the following consideration. If an *Escherichia coli* is put in a suitable environment, there will be plenty of bacteria after some period of time. However, one can only get an adult frog if a frog zygote is incubated. The key point of the above example is that each *E. coli* has almost unlimited potential of replication, while the zygote's ability to proliferate is suppressed by the degree of differentiation status. Cell differentiation, along with proliferation are the most important events that promote the process for multi-cellular organ formation and tissue stratification [3]. In addition, they also play important roles in pathogenesis, such as tumorigenesis including benign and malignant tumors [4–6].

On the other hand, it is also worth noting that in developmental stages, cell-cell communications and interactions become much more frequent and important in the gastrulation phase than in the cleavage stage. This results in the aggregation of cells with similar properties with cell migration now playing a significant role in the developmental stage. Developmental biology experiments

[a] e-mail: pylai@phy.ncu.edu.tw
[b] e-mail: ckchan@gate.sinica.edu.tw

revealed that certain groups of cells undergo cell migration, resulting in pattern formation during early embryogenesis. Cells with similar properties will tend to aggregate and deposit, and then they may migrate in groups to the specific locations for proliferation by either cell-cell interactions or chemotaxis. For example, concentration gradients of BMP and Wnt along two orthogonal axes are responsible for both dorsal-ventral and anterior-posterior axis formation. In this phase, a small number of the embryonic stem cells retain their properties while others will differentiate into three major lineages, known as the ectoderm, mesoderm and endoderm. Cell-cell communications can be achieved by receptors, concentration differences, and/or cell-cell junctions.

Cell proliferation and migration can be even more pronounced in tumor cells. Benign tumors often proliferate themselves locally and they aggregate into a cluster forming a tumor mass. Their growth and proliferation will lead to pressure on other nearby tissues and cause signs of swelling, redness and painful sensations. On the other hand, malignant tumor cells can transform into the primitive cell type, similar to cell de-differentiation, thereby enhancing cell growth rate, mobility and invasion to other tissues [4]. The idea that de-differentiation of committed cells is a cause for cancer is a decade old, but it was challenged by the cancer stem cell hypothesis that cancer arises from specific progenitor populations [4]. Meanwhile, recent interesting findings of the p53 tumor suppressor can prevent de-differentiation [5,6] suggested that tumorigenesis and cell reprogramming might share the common mechanism of de-differentiation, and hence revived the close connection between carcinogenesis and de-differentiation. In addition, cell de-differentiation has been observed in various cell types such as pancreatic cells [7], neurons [8], epithelial cells [9], retinal cells [10], germ cell [11], myoblasts [12] and epidermal keratinocytes [13] in response to appropriate stimuli or factors. Studies have been focused on the invasiveness, motility and spreading of cancerous cells [14–19]. It has also been reported that de-differentiation indeed occurs and plays a significant role in the context of skin cell proliferation [20] both in *vivo* [21] and *in vitro* [22, 23] under various conditions [24]. These de-differentiated cells were observed to exhibit high proliferative ability [22–24], suggesting that de-differentiation can be an important factor for skin cell proliferation. Recently it has been demonstrated that in a simple autoregulated single progenitor cell model, a small but finite de-differentiation probability can lead to uncontrolled growth resembling carcinogenesis [25].

In typical discrete models, cells are represented on the lattice sites in the cellular Potts model with the cells randomly chosen and their states to be updated by discrete time steps obeying certain dynamical rules [26–28]. Computer simulations on cellular systems with the Potts model have been studied by various research groups [26–28], where the focus was mainly on the spatial arrangement and pattern formation as governed by cell-cell interactions. The role of cell proliferation can complicate the cell motion and patterns, which are much less studied. In principle, using the information from the microscopic cellular dynamics or interactions, macroscopic mathematical models can be derived from the underlying description in the cellular scale [29, 30]. Cell differentiation has been modeled using the cellular Potts model in the context of developmental biology and morphogenesis [31]. On the other hand, the effect of de-differentiation is much less addressed theoretically [25]. Meanwhile, cell movement and proliferation have been the major focus in modeling tumor spreading or cancer cell growth [14–18]. Various theoretical studies have been focused on modeling cancer growth and invasion, ranging from microscopic cellular dynamics via the computational approach [32] to mathematical modeling in terms of partial differential equations [29, 30, 33, 34].

In this work, we consider the even more complex situation of having the possibility of changing cell types as a result of differentiation or de-differentiation. We focus on the spatial and temporal evolution of the cell population taking into the account the effects of cell proliferation, motility, differentiation and de-differentiation. Cell differentiation can often be depicted conceptually in terms of some epigenetic energy landscape [35] in which totipotent stem cells are on the top, passing through the pluripotent state and then eventually roll down and commit to a certain lineage. The recent discovery of induced pluripotent stem cells (iPSC) [36], in which mature cells can be reprogrammed to become pluripotent and possibly capable of differentiating into different lineages, opens new avenues in cell fate control and stem cell therapy. Albeit with a low efficiency, iPSC from mouse embryonic cells/fibroblasts [36] and human fibroblasts [37] revealed that the stochastic nature of cell differentiation/de-differentiation [38] could be important in cell development. Using the concept of transition probabilities between different/same cell types, the processes of proliferation, differentiation, and de-differentiation can be described within the same framework allowing for a better quantitative understanding. Furthermore, it has been suggested that stochastic theory based on cell stabilization might be important in differentiation therapy for cancer [39]. In our model, cells can proliferate, differentiate/de-differentiate and migrate by diffusive motion, hence the population of different cell types can vary in space and time, which is in turn formulated in terms of reaction-diffusion type nonlinear coupled partial differential equations (PDEs). Interesting cell density wave propagation can be obtained and the characteristic wave speeds in various situations can be calculated analytically and verified by direct numerical solution of the PDEs. Our results indicate the existence of a wavefront propagating with a constant terminal speed resembling collective cell motion. In the simplest case of pure proliferation and diffusive motility, the wavefront profile can be obtained analytically by a singular perturbation method. For the case of proliferation with differentiation and motility, as the mobility of the down-lineage cells increases, its wave propagation dynamics can "switch gears" to propagate with a faster wavefront speed, while retaining a slower waveback that propagates with its upper-lineage ancestor cells. Furthermore, in the presence of cell de-differentiation, both the wavefront speeds and the cell growth rate will be significantly enhanced.

2 Cell differentiation model

As we know, different cell types exist in an individual, which can be thought of corresponding to different states as described in the cellular Potts model [26]. In many biological events, cells have complicated interactions among them either over short distances such as adhesion with junctions, or via hormone over long distances. Here we focus on local short-range interactions, long-range interaction via signaling molecules can also be incorporated (see appendix F), but will be considered in our future studies. The local interactions between cells of different types are taken into account in the Potts model [27,28], which we will employ. Consider a system consisting of a collection of different cell types, each cell i can take q different states (cell types) labeled by σ_i. Similar to cellular Potts models [26], σ can take $0, 1, 2, \ldots, q-1$ possible states, with the state $\sigma = 0$ denoting a vacant site.

Similar to the idea of describing cell progression and mutations by suitable probability distributions in the kinetic theory approach of modeling multicellular systems [33,40], differentiation to different cell types is described by transition probabilities in our model. In our system, each cell can undergo cell division or differentiation into another cell type with a different state as long as there is a nearby vacant site. The differentiation of cells into different states of the cell is described by the probabilities $\Pi_{\alpha\beta}$, which is defined as the contributing probability that a vacant site will become state β if one of its neighbors is in state α. The $q \times q$ asymmetric matrix, $\Pi_{\alpha\beta}$ fully describes the differentiation pathways of the cells, and $\sum_\beta \Pi_{\alpha\beta} = 1$. Assuming that each neighbor has equal effect on the vacant site, then the (normalized) probability that the vacant site (state 0) will flip to state β is given by the weighted sum

$$p_\beta = \frac{1}{c}\sum_{i=1}^{c}\sum_{\alpha}\Pi_{\alpha\beta} \times \text{prob}(i\text{-th neighbor is an } \alpha\text{-cell}), \tag{1}$$

where c is the coordination number (number of nearest neighbors) of the lattice. It follows that

$$\begin{aligned}\sum_\beta p_\beta &= \frac{1}{c}\sum_{i=1}^{c}\sum_{\alpha}\sum_{\beta}\Pi_{\alpha\beta} \\ &\quad \times \text{prob}(i\text{-th neighbor is an } \alpha\text{-cell}) \\ &= \frac{1}{c}\sum_{i=1}^{c}\sum_{\alpha}\text{prob}(i\text{-th neighbor is an } \alpha\text{-cell}) = 1,\end{aligned} \tag{2}$$

verifying the expected normalization of the p_β's. The mobility of the cells are taken to diffuse randomly and their mobility is governed by the corresponding diffusion coefficients, D_α. The cell death is described by the probabilities $(g_\alpha, \alpha \neq 0)$ that each cell (non-vacant site) state becomes a vacant state. Here it is implicitly assumed that an apoptotic cell quickly dissociates and does not take up space.

Figure 1a shows schematically that a cell with high potential to differentiate, such as a zygote or stem cell

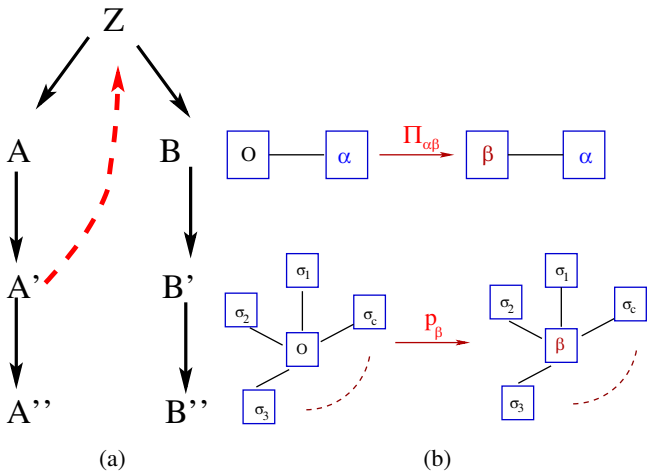


Fig. 1. (a) Schematic picture showing cell differentiation from the zygote (Z) to different lineages $A, A', A'', \ldots$ and $B, B', B'', \ldots$, etc. The solid arrows denote normal differentiation down the lineage, while the dashed arrow indicates backward de-differentiation (or possible carcinogenesis pathway). (b) Schematics showing the differentiation/transformation between different cell types and the associated probabilities. A vacant site is denoted by O, $\Pi_{\alpha\beta}$ denotes the probability contribution that a cell in state α will proliferate to its neighbor a new cell of state β. p_β denotes the overall probability of a vacant site to be transformed to a cell of state β.

(Z), can self-replicate and proliferate into different lineages $A, B, \ldots$, etc. The cells in state Z have high proliferation rates. Each lineage can further differentiate into higher level of differentiated cells down the lineage. The direction of differentiation is irreversible and the number of Z state cells is very little for a well-developed organism under a normal (healthy) situation. Carcinogenesis is modeled by the possibility that a fraction of differentiated cells can proliferate backward to the Z state, causing uncontrolled rapid growth in a cell population resulting in tumor [25]. Under special conditions, highly differentiated cells can be induced backward by proliferating to stem cells, as in the recent discovery of induced pluripotent stem cells [36].

3 Reaction diffusion equation

The lattice description of the present model can be cast in the continuum limit into partial differential equations allowing further mathematical analysis. In the continuum description, the lattice spacing is interpreted as the cell size h. Our goal here is to derive equations to describe the spatio-temporal evolution of the cell volume fractions $n_\alpha(\vec{r}, t)$ for the cell at states $\alpha = 0, 1, 2, \ldots, q-1$, with $\sum_\alpha n_\alpha = 1$. Unlike the cellular Potts model on a discrete lattice, here the n_α fields can overlap in space. Denoting the cell proliferation rate or "flipping rate" by τ_f, the evolution equations for n_α are

$$\frac{\partial n_\alpha}{\partial t} = D_\alpha \nabla^2 n_\alpha - g_\alpha n_\alpha + \frac{n_0 p_\alpha}{\tau_f}, \quad \text{for} \quad \alpha \neq 0, \tag{3}$$

where p_α is the probability of a vacant site to become an α-type cell due to proliferation from its neighbors as given by eq. (1). Hereafter, we consider the two-dimensional case. Since $n_\sigma(x,y)$ gives the probability of being a σ-cell at (x,y), using eq. (1), p_α can be expressed as (consider a square lattice for convenience)

$$p_\alpha = \frac{1}{4}\sum_\sigma \Pi_{\sigma\alpha}\Big(n_\sigma(x+h,y)+n_\sigma(x-h,y) +n_\sigma(x,y+h)+n_\sigma(x,y-h)\Big). \quad (4)$$

For continuum limit, the contributions from the four neighbors are expanded for small h to $\mathcal{O}(h^2)$, and finally one obtains

$$p_\alpha = \sum_\sigma \Pi_{\sigma\alpha} n_\sigma + \frac{h^2}{4}\sum_\sigma \Pi_{\sigma\alpha}\nabla^2 n_\sigma. \quad (5)$$

The Laplacian term in p_α arises from the fact that the contribution of α-cell is due to the differentiation/proliferation from neighboring sites. Using the fact that $\sum_\alpha \Pi_{\sigma\alpha} = 1$ and $\sum_\sigma n_\sigma = 1$, one can verify explicitly that $\sum_\alpha p_\alpha = 1$. Using the expression of p_α in (5), the reaction-diffusion equation (3) becomes

$$\frac{\partial n_\alpha}{\partial t} = D_\alpha\nabla^2 n_\alpha - g_\alpha n_\alpha + \frac{n_0}{\tau_f}\sum_\sigma \Pi_{\sigma\alpha} n_\sigma +n_0 D_f \sum_\sigma \Pi_{\sigma\alpha}\nabla^2 n_\sigma, \quad \text{for} \quad \alpha\neq 0, \quad (6)$$

where $D_f \equiv \frac{h^2}{4\tau_f}$ is the "effective diffusion coefficient" arising from the proliferation/differentiation of the cells. The equation can be solved together with $n_0 = 1-\sum_{\mu\neq 0} n_\mu$. It would be convenient to rescale the space and time via $x \to x/\sqrt{D_f\tau_f}$ and $t \to t/\tau_f$, and defining $a_\alpha \equiv g_\alpha\tau_f/\Pi_{\alpha\alpha}$ and $b_\alpha \equiv D_\alpha/(D_f\Pi_{\alpha\alpha})$. Equation (6) can be written as

$$\frac{\partial n_\alpha}{\partial t} = \sum_\sigma \Pi_{\sigma\alpha}\big[(n_0+b_\alpha\delta_{\sigma\alpha})\nabla^2 n_\sigma +(n_0-a_\alpha\delta_{\sigma\alpha})n_\sigma\big], \quad \text{for} \quad \alpha\neq 0. \quad (7)$$

Notice that the effective diffusion coefficient in (7) increases with the local vacant site concentration, accounting for the excluded-volume interactions between the cells. The homogeneous steady-state solution is obtain by solving

$$n_0\sum_\sigma \Pi_{\sigma\alpha} n_\sigma = \Pi_{\alpha\alpha} a_\alpha n_\alpha, \quad \text{for} \quad \alpha = 1,2,\ldots,q-1. \quad (8)$$

Together with the equation $n_0 = 1-\sum_{\mu\neq 0} n_\mu$, one can eliminate variables and obtain an algebraic equation (polynomial equation of degree q) in n_0. It is easy to see that $n_\alpha = 0$ for $\alpha = 1,2,\ldots,q-1$ is always a homogeneous steady solution, and all other physical solutions have to satisfy the condition of $n_\alpha \geq 0$. Furthermore, the death rates of the cells should be small compared to the self-proliferation rates in order to sustain steady wavefront propagations, hence $g_\alpha < \Pi_{\alpha\alpha}/\tau_f$ or $a_\alpha < 1$ for $\alpha \neq 0$ is assumed throughout this paper. In this paper, we will focus on differentiation along a single lineage with $Z \to A \to A' \to \ldots$, with the cell fractions of $Z, A, A', \ldots$ corresponding to $n_1, n_2, n_3, \ldots$, respectively.

It should be noted that long-range interaction between cells via signaling molecules can affect cell differentiation fate as well as directing cell motion via chemotactic attractants. Chemical signaling is one of the most important processes for cell-cell communication that can lead to collective motion and quorum sensing in both bacteria and eukaryotic cells [41–44]. Although cell-cell interaction via signaling molecules is not considered in this work, it can be incorporated in our model by extending eq. (3) as described in appendix F.

3.1 Simple growth case: 2-states case (q = 2)

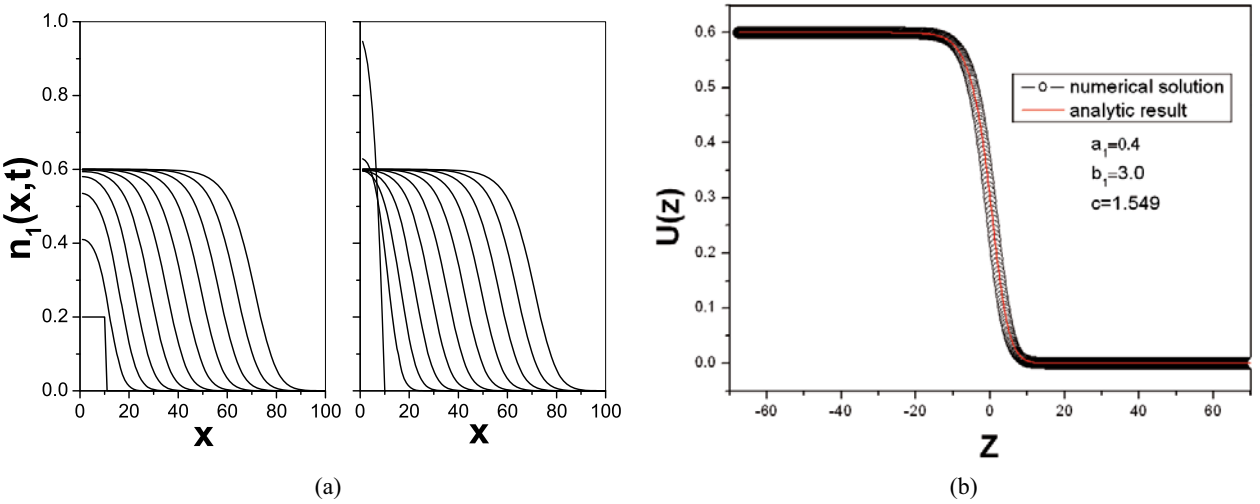
In this case, there are only vacancy and Z cells (corresponding to n_1), and eq. (7) reduces to

$$\frac{1}{\Pi_{11}}\frac{\partial n_1}{\partial t} = (1+b_1-n_1)\nabla^2 n_1 + n_1(1-a_1-n_1), \quad (9)$$

which resembles the well-studied Fisher-Kolmogorov equation [45] except for a concentration-dependent effective diffusion coefficient. The homogeneous steady-state solutions are simply $n_1 = 0$ and $1-a_1$. Much insight can be obtained by considering the one-dimensional case, which can be shown easily to admit a traveling wavefront solution $n(x,t) = U(x-ct)$ that satisfies the ODE $(1+b_1-U)U'' + cU + U(1-a_1-U) = 0$. The wavefront profile gives the boundary conditions $U(-\infty) = 1-a_1$ and $U(\infty) = 0$. Employing nonlinear dynamics analysis [46], one can show that a traveling wavefront solution exists for wave speed exceeding some minimal value, given by

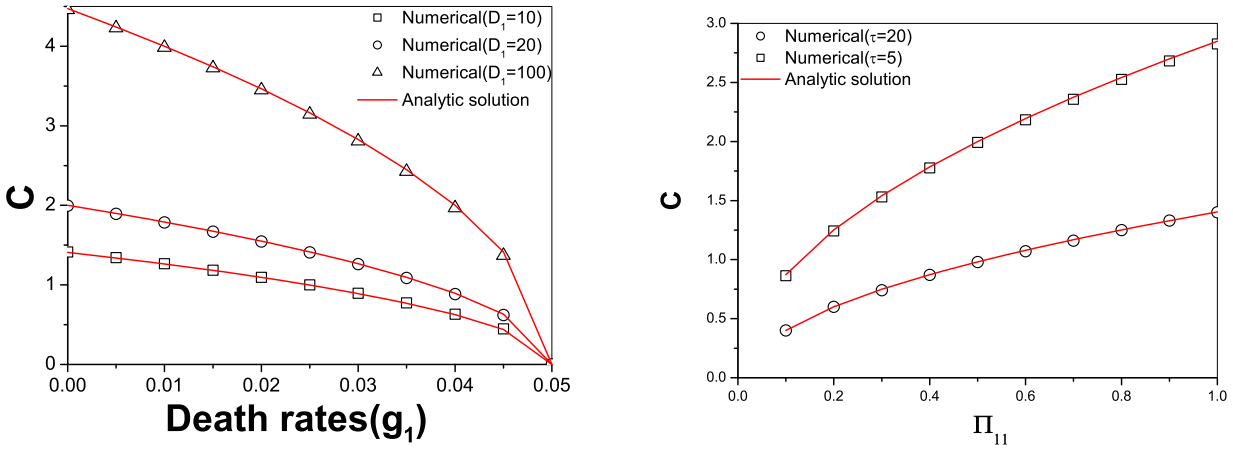
$$c \geq c_{\min} = 2\Pi_{11}\sqrt{(1+b_1)(1-a_1)}. \quad (10)$$

In terms of the proper units and original parameters, $c_{\min} = 2\sqrt{\Big(\frac{\Pi_{11}}{\tau_f}-g_1\Big)(D_1+\Pi_{11}D_f)}$. For the two-dimensional case, although (9) does not admit an exact radially symmetric wavefront solution, it can shown easily using similar analysis as in the standard Fisher-Kolmogorov equation [45], that the asymptotic solution admits a propagating wavefront with minimum speed $c_{\min}$. For the situation where cell motility is small, $D_1 \ll \Pi_{11}D_f$, one gets the growth speed $c_{\min} \sim h\sqrt{\frac{\Pi_{11}}{\tau_f}\Big(\frac{\Pi_{11}}{\tau_f}-g_1\Big)}$. And for small death rate $c_{\min} \simeq \Pi_{11}h/\tau_f$ agreeing with the simple consideration of compact radially growth which is also verified by our simulation results. On the other hand, if cell motility is large, *i.e.* $D_1 \gg \Pi_{11}D_f$, then one recovers the well-known result from the Fisher-Kolmogorov equation, $c_{\min} \simeq 2\sqrt{\Big(\frac{\Pi_{11}}{\tau_f}-g_1\Big)D_1}$.



(a) (b)

Fig. 2. Traveling wavefront for the simple growth case. (a) Evolution of two different initial wave packets into the same steady traveling wavefront with speed $c_{\min}$ that agree well with the analytical results given by eq. (10). $a_1 = 0.4$, $b_1 = 3$, $\Pi_{11} = 0.5$. (b) Wavefront profile: the singular perturbation solution (solid curve) given by (13) agrees well with the numerical solution of the PDE in (9) (symbols). The wavefront speed $c = c_{\min} = 1.549$ is obtained from the theoretical result in (10) and agrees well with the measurement of steady propagating profile from the numerical solution.



(a) (b)

Fig. 3. Traveling wavefront speed for the simple growth case. (a) Steady wavefront speeds of plane waves as a function of the death rate g_1 for various values of D_1. Symbols: speeds measured from numerical solutions. Curves: analytical results given by eq. (10). (b) Steady wavefront speeds of plane waves as a function of the cell flipping probability Π_{11} for various values of τ_f.

Figure 2a shows the numerical solution of eq. (9) in one spatial dimension from two arbitrary initial pulse shapes. Both of them eventually evolved into the same traveling wavefront with the minimal steady wave speed given by $c_{\min}$. A similar behavior has also been observed in the Fisher-Kolmogorov equation [46].

The dependence of the steady wavefront speed as a function of the death rate g_1 for different D_1 (fig. 3a), and as a function of the flipping probability Π_{11} (fig. 3b) are obtained from the numerical solutions. As shown in fig. 3, the waves travel with the analytic result of the minimal speed $c_{\min}$ given by (10) with the predicted dependence on the parameters. Furthermore, by considering the change in cell population in time Δt, one has

$$\begin{aligned}\Delta N_1 &= \int_{-\infty}^{\infty} \mathrm{d}x[U(x - ct - c\Delta t) - U(x - ct)] \\ &= -c\Delta t \int_{-\infty}^{\infty} \mathrm{d}x U'(x - ct) \\ &= -c\Delta t \int_{-\infty}^{\infty} \mathrm{d}z U'(z). \end{aligned} \tag{11}$$

Hence the steady cell growth rate of the Z-cell population, N_Z, is given by

$$\frac{\mathrm{d}N_Z}{\mathrm{d}t} = \frac{\Delta N_1}{\Delta t} = -c \int_{-\infty}^{\infty} \mathrm{d}U = c(1 - a_1). \tag{12}$$

For general situations of growth with differentiation and/or de-differentiation, the population growth rates for different cell types can be calculated in terms of the steady state cell populations and wavefront speeds as described in appendix C.

3.2 Wavefront profile

With $z = x - ct$, the propagating wavefront profile $U(z)$ can be calculated using singular perturbation theory [46]. Defining $\xi \equiv z/c = \sqrt{\epsilon}$, $U(z) \equiv u(\xi)$, $\epsilon \equiv 1/c^2$ as the small parameter for perturbation expansion, one gets the asymptotic wavefront profile (see appendix B),

$$\begin{aligned} U(z) = & \frac{1-a_1}{1+e^{\frac{1-a_1}{c}z}} - \left(\frac{1-a_1}{2c}\right)^2 \operatorname{sech}^2 \frac{(1-a_1)z}{2c} \\ & \times \left(\left[2 + \frac{1+a_1}{b_1}\right] \ln\cosh \frac{(1-a_1)z}{2c} \right. \\ & \left. + \frac{1-a_1}{b_1}\left[\frac{(1-a_1)z}{2c} + \tanh\frac{(1-a_1)z}{2c}\right]\right) \\ & + O\left(\frac{1}{c^4}\right). \end{aligned} \tag{13}$$

Figure 2b shows the analytic result given by (13) together with the steady wave profile obtained from the numerical solution, showing very good agreement. From the analytic result in (13), one easily gets

$$U'(0) = -\frac{(1-a_1)^2}{4c} - \frac{(1-a_1)^2}{4b_1c^3} + O\left(\frac{1}{c^4}\right), \tag{14}$$

i.e. the steeper the wavefront profile, the slower is the propagation speed.

4 Simple model for cell differentiation and de-differentiation: 3-states case

The system now consists of vacancy, undifferentiated Z type cells (n_1), and differentiated A type lineage cells (n_2). In this case eq. (7) reads

$$\begin{aligned} \frac{\partial n_1}{\partial t} = & \Pi_{11}(1 + b_1 - n_1 - n_2)\nabla^2 n_1 \\ & + \Pi_{21}(1 - n_1 - n_2)\nabla^2 n_2 \\ & + \Pi_{11}(1 - a_1 - n_1 - n_2)n_1 \\ & + \Pi_{21}(1 - n_1 - n_2)n_2, \end{aligned} \tag{15}$$

$$\begin{aligned} \frac{\partial n_2}{\partial t} = & \Pi_{22}(1 + b_2 - n_1 - n_2)\nabla^2 n_2 \\ & + \Pi_{12}(1 - n_1 - n_2)\nabla^2 n_1 \\ & + \Pi_{22}(1 - a_2 - n_1 - n_2)n_2 \\ & + \Pi_{12}(1 - n_1 - n_2)n_1. \end{aligned} \tag{16}$$

The dynamics is qualitatively different in the absence or presence of de-differentiation, and will be investigated in detail in the following.

4.1 Cell differentiation

In this case, $\Pi_{21} \equiv \Pi_{AZ} = 0$, the homogenous steady-state solutions are $(n_1, n_2) = (0, 0)$, $(0, 1 - a_2)$, and for $a_2 > a_1$, another the non-zero fixed point solution $(n_1, n_2) = (n_1^{(S)}, n_2^{(S)}) \equiv \left(\frac{(1-a_1)(a_2-a_1)\Pi_{22}}{a_1\Pi_{12}+(a_2-a_1)\Pi_{22}}, \frac{a_1(1-a_1)\Pi_{12}}{a_1\Pi_{12}+(a_2-a_1)\Pi_{22}}\right)$ emerges. Note that for both $n_1^{(S)}$ and $n_2^{(S)} > 0$, one must have $a_2 > a_1$ (*i.e.* $g_2/\Pi_{22} > g_1/\Pi_{11}$). Assuming plane wavefront solution of the form $n_1(x,t) = U_1(x - ct)$ and $n_2(x,t) = U_2(x - ct)$, one gets a pair of second-order coupled ODEs. Careful analysis on the nonlinear dynamics of the 4-dimensional dynamical system reveals that the fixed point $(U_1, U_2, U_1', U_2') = (0, 0, 0, 0)$ is stable suggesting that wave profiles $U_i(z)$ will end with $(n_1, n_2) = (0, 0)$ in the $z \to \infty$ limit end. For $a_1 \geq a_2$, there are only two fixed points and the only possible wavefront connects from $(0, 1 - a_2, 0, 0)$ to $(0, 0, 0, 0)$. Then using the physical requirement of non-negative U_1 and U_2 imposed the constraint that the fixed point $(0, 0, 0, 0)$ must be a stable node, not a stable focus and revealed that a plane wavefront can propagate for $n_1(x,t)$ and $n_2(x,t)$ with speed c above the minimal value (see appendix C)

$$\begin{aligned} & c \geq C_0 \equiv \\ & \max\left[2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}, 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}\right]. \end{aligned} \tag{17}$$

The above analytical results are checked against numerical solution of (15) and (16) in one dimension. Figure 4d and e show the asymptotic profiles of propagating wavefronts for Z and A having the same speeds that agree well with the minimal value given by (17). However, in the parameter regime of large b_2 (strong diffusion of A), a new local wavefront of A emerges, while the wavefront of Z continues to exist but propagates with a slower speed (fig. 4g). In this case the n_2 profile consists of a wavefront with a speed given by eq. (17) and also a waveback propagating with the same speed as the slow wavefront of Z. Furthermore, the profile of n_1 and n_2 indicate that the dynamics connects the $(n_1^{(S)}, n_2^{(S)}, 0, 0)$ fixed point to the $(0, 1-a_2, 0, 0)$, fixed point and then finally to the $(0, 0, 0, 0)$ fixed point. The stability of steady propagating wave solution is further studied by numerical integrating the PDEs. Figure 4 shows the phase diagrams for regimes with different values of a_1, a_2 (fig. 4a) and b_1, b_2 (fig. 4b). The phase boundaries in both phase diagram are straight lines given by $2\Pi_{11}\sqrt{(1-a_1)(1+b_1)} = 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}$ suggesting that the propagating wavefront solution (of n_2) switches to a new mode when C_0 in (17) takes another branch of maximal value governed by the parameter of the A-cell (*i.e.* a_2, b_2 and Π_{22}).

The wavefront speed is summarized in fig. 5. For small values of b_2, steady wavefront propagation of both n_1 and

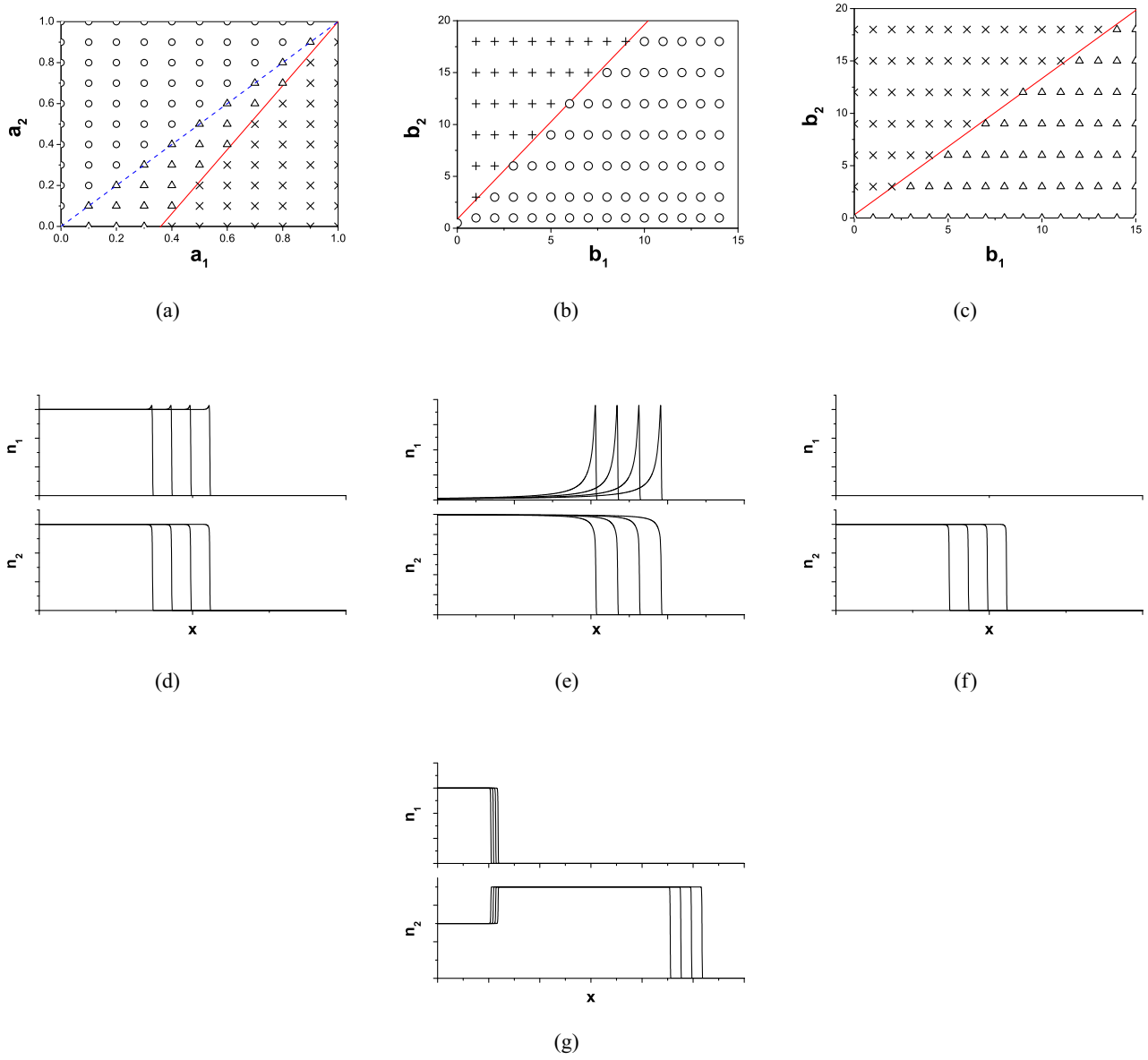


Fig. 4. Phase diagram for steady wave propagation for the case of $\Pi_{AZ} = 0$ obtained from numerical solution of the PDE system. $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$. The symbols denote different wave propagation modes as observed from the numerical solution of PDEs. (a) a_1-a_2 phase diagram, with other parameters fixed at $b_1 = 3$, $b_2 = 3$. $\circ$ denotes the regions $a_2 > a_1$ (above the dashed line) in which stable single wavefront profiles for n_1 and n_2 as depicted in (d). $\triangle$ denotes a stable wavefront profile in n_2 and a stable spiking profile in n_1 as shown in (e). $\times$ denotes a stable wavefront profile in n_2 but vanishing n_1 as shown in (f). The solid straight line is boundary given by $2\Pi_{11}\sqrt{(1-a_1)(1+b_1)} = 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}$, when the speed of the wavefront switches to another value as given in eq. (D.5). The dashed straight line is the $a_1 = a_2$ line. (b) b_1-b_2 phase diagram for $a_2 > a_1$, with other parameters fixed at $a_1 = 0.4$, $a_2 = 0.5$. $+$ denotes the propagation of a faster wavefront along with a slow waveback as shown in (g). $\circ$ and the solid straight line have the same meaning as in (a). (c) b_1-b_2 phase diagram for $a_1 \geq a_2$, with other parameters fixed at $a_1 = 0.5$, $a_2 = 0.4$.

n_2 with the same speed is stable, with values given by the minimal propagating speed in (D.5). As b_2 is increased beyond a critical value given by the switching of the $c_{\min}$ to the other maximal value in (D.5), the n_2 wavefront loses its stability giving rise to a new mode with a wave-back propagating with a slower speed which is the same as the wavefront of n_1 and another faster wavefront with speed given by (D.5). Our results indicate the existence of a wavefront propagating with a constant terminal speed, which is qualitatively similar to some experimental observation in cell migration [1,2]. Furthermore, the existence of a faster wavefront in our theory resembles the recent experimental observation of a moving front of leader cells with faster speed in cell migration patterns [1].

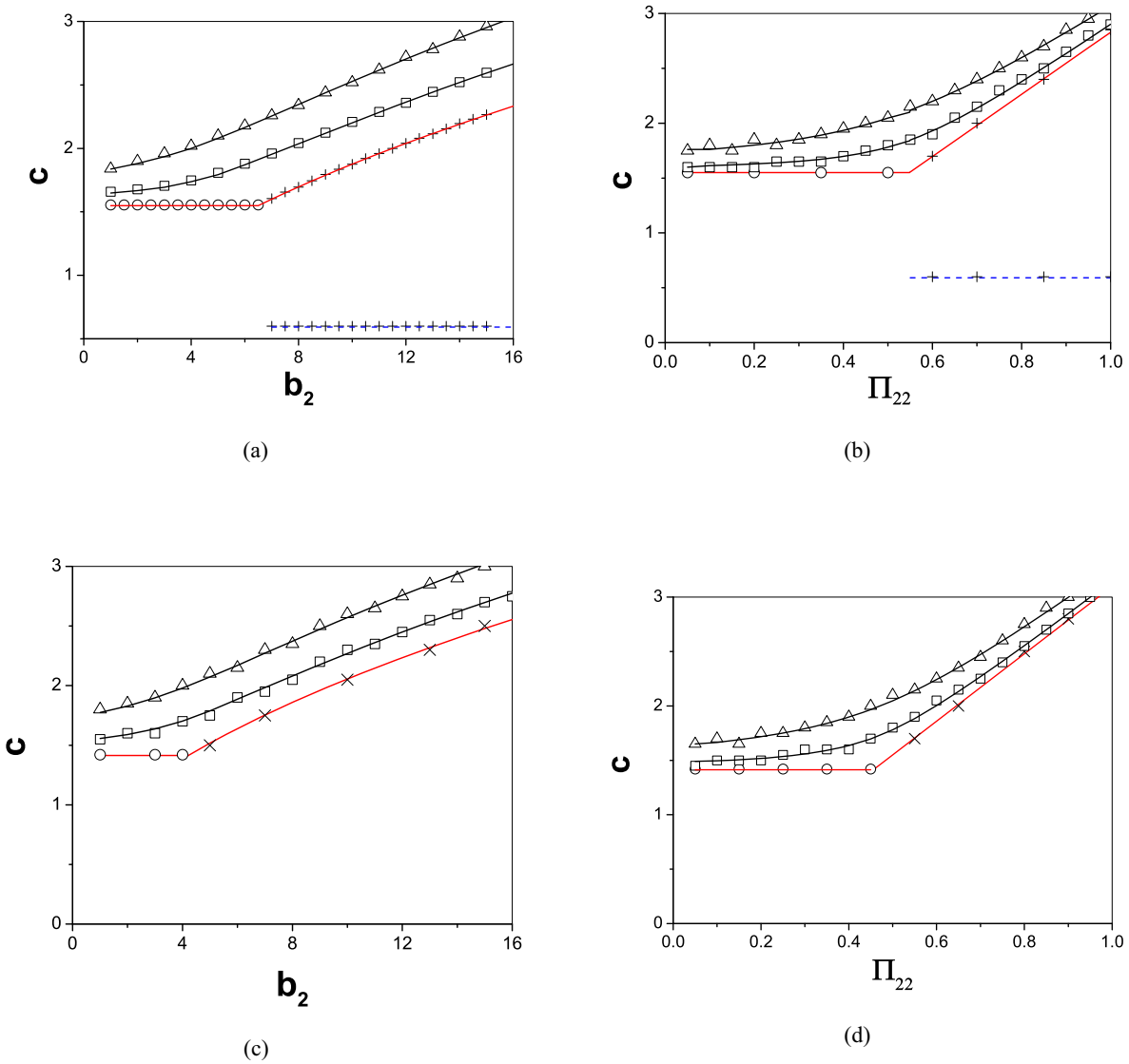


Fig. 5. Wavefront speeds, c, of plane waves for the case of 3-state model for cell differentiation ($\Pi_{AZ} = 0$) and de-differentiation ($\Pi_{AZ} > 0$). (a) c as a function of b_2 for $a_2 > a_1$, with $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$; $a_1 = 0.4$, $a_2 = 0.5$, $b_1 = 3$. Symbols: speeds measured from numerical solution of the PDEs. (○): wavefront speeds of n_2, which is the same as the wavefront speed of n_1 for $b_2 \lesssim 6.5$. (+): wavefront speeds of the n_2 is shown in the upper branch, the lower branch is the wavefront speed of n_1 which is the same as the waveback speed of n_2. Lower solid red curves: analytical results given by eq. (17). Dashed line: analytic result of lower limit of front speed of n_1 given by eq. (D.7). (△) and (□): measured wave speed for the $\Pi_{AZ} > 0$ case. Upper solid black curves: analytic results from fixed point analysis for the $\Pi_{AZ} = 0.1$ (□) and 0.4 (△) cases. (b) c as a function of Π_{22} with $b_2 = 3$, other parameters have the same values as in (a). The symbols and curves have the same meaning as in (a). (c) c as a function of b_2 for $a_1 > a_2$, with $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$; $a_1 = 0.5$, $a_2 = 0.4$, $b_1 = 3$. (×): measured wavefront speed of n_2 for $\Pi_{AZ} = 0$. (d) c as a function of Π_{22} for $a_1 > a_2$ with $b_2 = 3$, other parameters have the same values as in (c).

The existence of a faster wavefront and a slower waveback when the mobility and/or self-proliferation rate of the down-lineage cells is sufficiently large can be rationalized as follows. Since for steady wavefront propagation, the growth in population is simply given by the product of the wavefront speed and the difference in the steady-state homogeneous concentrations which the wavefront connects (see eq. (12)), therefore the growth rates of the cells are given by

$$\frac{\mathrm{d}N_Z}{\mathrm{d}t} = c_{\mathrm{slow}} n_1^{(S)}\,, \tag{18}$$

$$\frac{\mathrm{d}N_A}{\mathrm{d}t} = c_{\mathrm{slow}} n_2^{(S)} + (1 - a_2)(c_{\mathrm{fast}} - c_{\mathrm{slow}}), \tag{19}$$

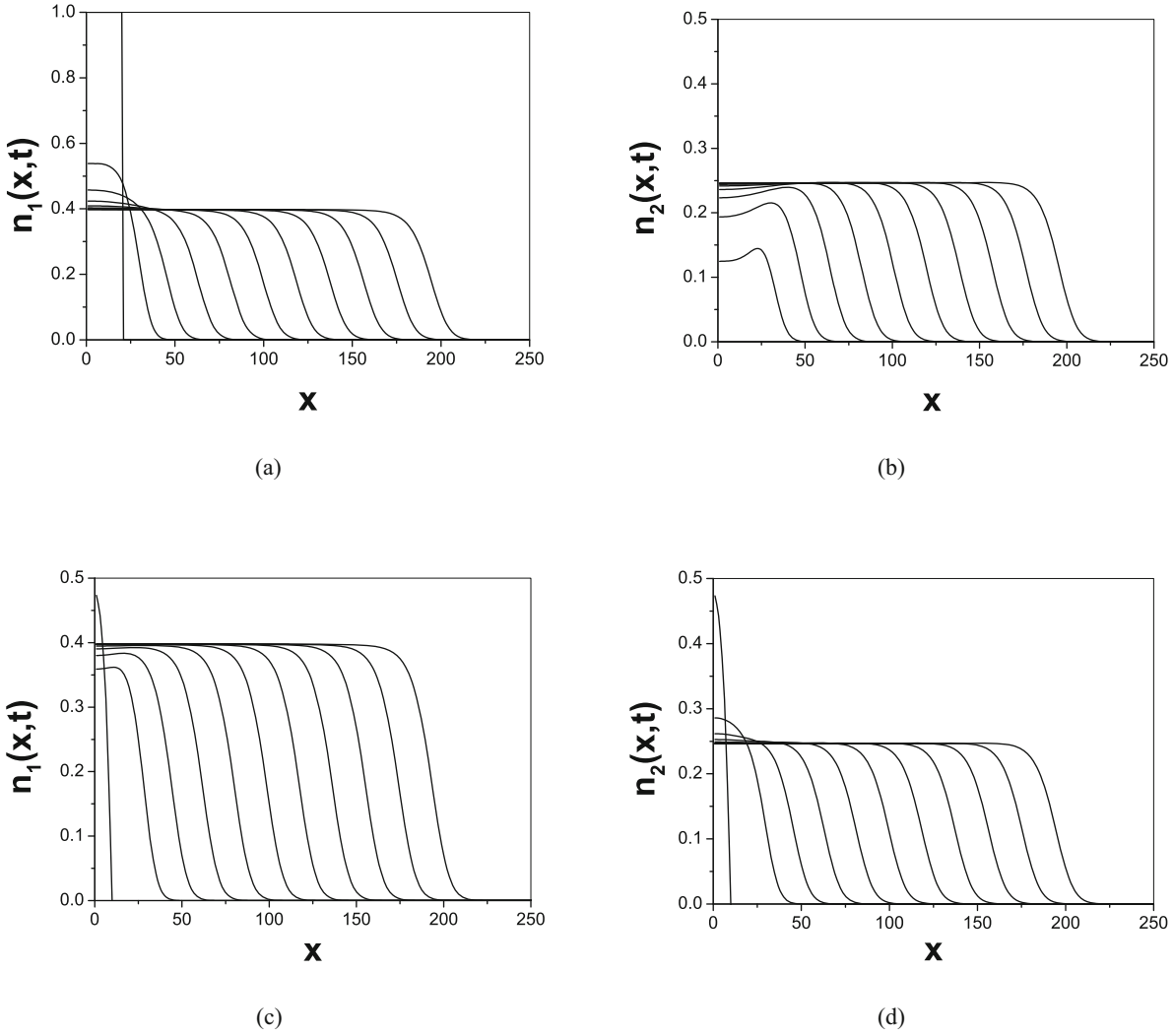


Fig. 6. Time evolution towards steady plane wavefront profiles obtained from numerical solution of (16) for the case of 3-state model with de-differentiation ($\Pi_{AZ} > 0$). $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$, $\Pi_{21} = 0.1$; $a_1 = 0.4$, $a_2 = 0.5$, $b_1 = 3$. Profiles of (a) $n_1(x,t)$, (b) $n_2(x,t)$ for $b_2 = 3$. Profiles of (c) $n_1(x,t)$, (d) $n_2(x,t)$ for $b_2 = 7$.

where $c_{\rm fast} = 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}$ and $c_{\rm slow} = 2\Pi_{11} \times \sqrt{(a_2-a_1)(a_2+b_1)}$ are the speeds (as given by (D.5) and (D.7)) of the faster wavefront and slower waveback respectively. Careful examination of the cell growth rates in (18) and (19) suggests that the growth of A has two contributions. The first part (first term in (19)) is originated from the differentiation from their upper-lineage Z and advancing with the slow speed $c_{\rm slow}$. Another part (second term in (19)) is derived from their self-proliferation giving rise to a faster front speed. The above results for the cell population growth follows from the general scenario as outlined in appendix C.

4.2 Cell de-differentiation

In this case, $\Pi_{21} \equiv \Pi_{AZ} > 0$, the cell can flip backwards from the differentiated state A to Z. Such a de-differentiation back to stem cells is an important route for carcinogenesis. When we carry out a nonlinear dynamics analysis, the homogenous steady-state solutions are $(n_1, n_2) = (0, 0)$, and the non-zero (n_1^*, n_2^*), whose values are given by

$$n_1^* = \frac{\gamma n_0^*(1-n_0^*)}{a_1 - (1-\gamma)n_0^*}, \qquad n_2^* = 1 - n_0^* - n_1^*, \tag{20}$$

with

$$n_0^* = \frac{a_1 + a_2 - \sqrt{(a_1+a_2)^2 - 4(1-\mu)a_1 a_2}}{2(1-\mu)}, \tag{21}$$

where $\gamma \equiv \frac{\Pi_{21}}{\Pi_{11}}$ and $\mu \equiv \frac{\Pi_{12}\Pi_{21}}{\Pi_{11}\Pi_{22}}$. Assuming the two cell species admit propagating wavefront solutions with the same speed, one can carry out similar analysis as in previous cases. Again one obtains a four-dimensional dynamical

system with two fixed points $(0,0,0,0)$ and $(n_1^*, n_2^*, 0, 0)$. The eigenvalues at the fixed points can be computed from the quartic characteristic polynomial equations. Similar to previous cases, the requirement of a stable node at the fixed point $(0,0,0,0)$ sets a lower limit of the propagating wavefront speed. The minimal wavefront speed as a function of b_2 obtained from such an analysis is shown in fig. 5 (upper black curve). Direct solution of the PDEs also indicates stable propagating wavefronts, with speeds faster than that of the $\Pi_{AZ} = 0$ case. Figure 6 shows the evolution of n_1 and n_2 into steady traveling wavefront for different parameters. In this case, the wavefronts for Z and A are always stable and propagate with the same speeds in all the parameter regimes we studied. The measured wavefront speeds are shown in fig. 5 (triangles) as a function of b_2 showing excellent agreement with the analytical results.

4.3 Cell growth rates

Our theory can predict the steady cell growth rates analytically in the absence of de-differentiation and making use of the observation or conjecture that the wavefront speed always assumes its minimal value, as described in appendix C. For $a_1 > a_2$, there are only two fixed points: $\vec{X}_0 = (0,0)$ and $\vec{X}_1 = (0, 1-a_2)$ and the cell population growth rates are given by eq. (12) in appendix C. Denoting $\vec{N} = (N_Z, N_A)$, one has

$$\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} = (\vec{X}_1 - \vec{X}_0)C_0 = C_0 \begin{pmatrix} 0 \\ 1-a_2 \end{pmatrix}. \qquad (22)$$

where

$$C_0 \equiv \max\left[2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}, 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}\right]$$

as defined in (D.5). For $a_2 > a_1$, another fixed point $\vec{X}_2 = (n_1^{(S)}, n_2^{(S)}) \equiv \left(\frac{(1-a_1)(a_2-a_1)\Pi_{22}}{a_1\Pi_{12}+(a_2-a_1)\Pi_{22}}, \frac{a_1(1-a_1)\Pi_{12}}{a_1\Pi_{12}+(a_2-a_1)\Pi_{22}}\right)$ emerges, and we have

$$\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} = 2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}\vec{X}_2,$$

$$\text{for} \quad 2\Pi_{11}\sqrt{(1-a_1)(1+b_1)} = C_0,$$

$$\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} = C_1\vec{X}_2 + (2\Pi_{22}\sqrt{(1-a_2)(1+b_2)} - C_1)\vec{X}_1,$$

$$\text{for} \quad 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)} = C_0. \qquad (23)$$

The growth in cell populations can also be measured directly from the numerical solutions. Figure 7 shows the variation of steady total cell growth rate, as a function of the parameter b_2 indicating a monotonic growth rate. The theoretical prediction, $\mathrm{d}N_Z/\mathrm{d}t + \mathrm{d}N_A/\mathrm{d}t$ given by (23), shows perfect agreement. For the case with non-vanishing de-differentiation probability Π_{AZ}, the cell growth rates can also be calculated precisely from our theory in a similar way as described in appendix C, although explicit expressions for the wave speeds, and hence the growth rates, cannot be obtained. The total growth rate of the cells obtained theoretically and also measured from direct numerical solution of the PDEs are shown in fig. 7 showing excellent agreement. As compared with the case of no de-differentiation, the cell growth rate is always significantly enhanced. Thus both the cell growth rates and wave propagation speeds are increased in the presence of finite de-differentiation, this provide a plausible understanding of the reason why cancerous cells can grow and spread faster than normal cells.

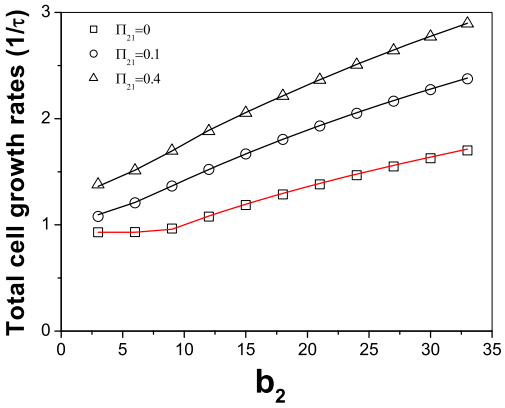


Fig. 7. Total cell growth rate for the 3-state model, defined as the increase in the number of A and Z cells in one time interval τ, as a function of b_2. $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$, $a_1 = 0.4$, $a_2 = 0.5$, $b_1 = 3$. Symbols: growth rate measured from the numerical solution of the PDEs. Curves: theoretical results from (23).

5 Z → A → A′: 4-states case

The system now consists of vacancy, undifferentiated Z type cells (n_1), and differentiated type lineage cells A (n_2) and A' (n_3). In the case of absence of de-differentiation, $\Pi_{AZ} = \Pi_{A'Z} = \Pi_{A'A} = 0$, the homogenous steady-state solutions (fixed points) are $\vec{X}_0 = (0,0,0)$, $\vec{X}_1 = (0,0,1-a_3)$, and for $a_3 > a_2$, another fixed point $\vec{X}_2 = (0, n_2^{**}, n_3^{**})$ emerges, where $n_2^{**} \equiv \frac{(1-a_2)(a_3-a_2)\Pi_{33}}{a_2\Pi_{23}+(a_3-a_2)\Pi_{33}}$, $n_3^{**} \equiv \frac{a_2(1-a_2)\Pi_{23}}{a_2\Pi_{23}+(a_3-a_2)\Pi_{33}}$. If furthermore $a_2 > a_1$, another non-zero fixed point $\vec{X}_3 = (n_1^{***}, n_2^{***}, n_3^{***})$ emerges, where

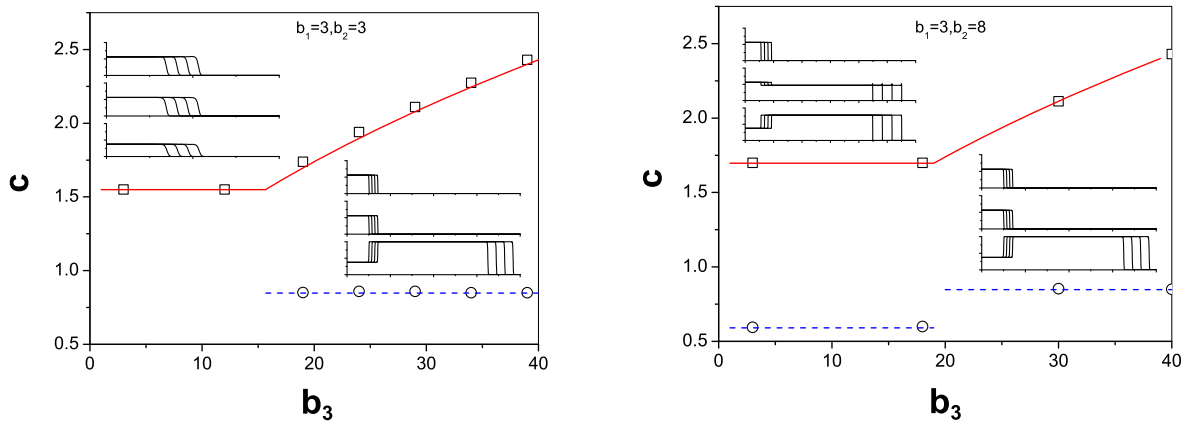


Fig. 8. Wave profiles and speeds as a function of b_3 of 4-state model of cell differentiation ($\Pi_{21} = \Pi_{32} = \Pi_{31} = 0$). $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{12} = 0.1$, $\Pi_{23} = 0.1$, $\Pi_{33} = 0.3$; $a_1 = 0.4$, $a_2 = 0.5$, $a_3 = 0.6$. The symbols show the wave profile speeds as measured from the numerical solution of the PDEs. The wave profiles of n_1, n_2 and n_3 (from top down respectively) in different regimes of b_3 are also shown. (a) $b_1 = b_2 = 3$. (□): wavefront speeds of all three profiles (which are the same) for smaller regime of b_3 and wavefront speed of n_3 for large b_3 regime. (○): wavefront speeds of n_1 and n_2 which is the same as the waveback speed of n_3. The red curve is $\mathrm{Max}[2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}, 2\Pi_{33}\sqrt{(1-a_3)(1+b_3)}]$ as given in eq. (27). The lower dashed line is $2\Pi_{11}\sqrt{(a_3-a_1)(a_3+b_1)}$ as given by eq. (28). Note that the change of speeds occurs at $b_3 = 47/3$ as predicted by eq. (27). The small and large regimes in b_3 corresponds to the flow $\mathbf{X}_3^* \to \mathbf{X}_0^*$ and $\mathbf{X}_3^* \to \mathbf{X}_1^* \to \mathbf{X}_0^*$, respectively. (b) $b_1 = 3$, $b_2 = 8$. The red curve is $\mathrm{Max}[2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}, 2\Pi_{33}\sqrt{1-a_3)(1+b_3)}]$ as given in eq. (27). The lower dashed lines are $2\Pi_{11}\sqrt{(a_2-a_1)(a_2+b_1)}$, $2\Pi_{11}\sqrt{(a_3-a_1)(a_3+b_1)}$ as given by eq. (29) and eq. (28) respectively. Note that the change of speeds occurs at $b_3 = 19$ as predicted by eq. (27). Symbols have similar meanings as in (a), except now in the smaller b_3 regime the n_2 and n_3 wavefront have the same faster speeds (□), the wavefront of n_1 and the wavebacks of n_2 and n_3 have the same slow speeds (○). The small and large regimes in b_3 corresponds to the flow $\mathbf{X}_3^* \to \mathbf{X}_2^* \to \mathbf{X}_0^*$ and $\mathbf{X}_3^* \to \mathbf{X}_1^* \to \mathbf{X}_0^*$, respectively.

$$n_1^{***} \equiv \frac{(1-a_1)(a_2-a_1)(a_3-a_1)\Pi_{22}\Pi_{33}}{a_1^2\Pi_{12}\Pi_{23}+a_1(a_3-a_1)\Pi_{12}\Pi_{33}+(a_3-a_1)(a_2-a_1)\Pi_{22}\Pi_{33}}, \tag{24}$$

$$n_2^{***} \equiv \frac{a_1(1-a_1)(a_3-a_1)\Pi_{12}\Pi_{33}}{a_1^2\Pi_{12}\Pi_{23}+a_1(a_3-a_1)\Pi_{12}\Pi_{33}+(a_3-a_1)(a_2-a_1)\Pi_{22}\Pi_{33}}, \tag{25}$$

$$n_3^{***} \equiv \frac{a_1^2(1-a_1)\Pi_{12}\Pi_{23}}{a_1^2\Pi_{12}\Pi_{23}+a_1(a_3-a_1)\Pi_{12}\Pi_{33}+(a_3-a_1)(a_2-a_1)\Pi_{22}\Pi_{33}}. \tag{26}$$

Assuming local plane wavefront solution of the form $n_i(x,t) = U_i(x-ct)$ for $i = 1, 2, 3$, and carrying out the nonlinear dynamics analysis of the 6-dimensional dynamical system as in the previous cases, one easily finds that two fixed points always exist: $\mathbf{X}_0^* \equiv (0,0,0,0,0,0)$, and $\mathbf{X}_1^* \equiv (0,0,1-a_3,0,0,0)$.

Notice that $\mathbf{X}_0^*$ is always stable, and a plane wavefront ending with $n_1 = n_2 = n_3 = 0$ can propagate with a speed (see appendix E)

$$C_0^{(4)} \equiv \max\left[2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}, 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}, 2\Pi_{33}\sqrt{(1-a_3)(1+b_3)}\right]. \tag{27}$$

Similarly, plane wavefront ending with $\vec{X}_1$ propagates with a speed

$$C_1^{(4)} \equiv \max\left[2\Pi_{11}\sqrt{(a_3-a_1)(a_3+b_1)}, 2\Pi_{22}\sqrt{(a_3-a_2)(a_3+b_2)}\right]. \tag{28}$$

If $a_3 > a_2$, another fixed point $\mathbf{X}_2^* \equiv (0, n_2^{**}, n_3^{**}, 0, 0, 0)$ emerges and the wavefront that ends with $\vec{X}_2$ travels with the speed

$$C_2^{(4)} \equiv 2\Pi_{11}\sqrt{(a_2-a_1)(a_2+b_1)}. \tag{29}$$

Finally, if $a_3 > a_2 > a_1$, another fixed point $\mathbf{X}_3^* \equiv (n_1^{***}, n_2^{***}, n_3^{***}, 0, 0, 0)$ emerges. This fixed point can be shown to be highly unstable, and cannot set a speed limit on the wavefront flowing into it. Moreover, a wavefront ending with $\vec{X}_3$ has never been observed in the numerical solutions of the PDEs.

The steady wavefront profile shapes and the corresponding propagating wavefront speeds can all be analyzed and predicted using the general method outline in appendix C. Figure 8 illustrates some situations where the propagating wavefronts display a very rich variety as one of the parameter, b_3, is varied. The steady wave profiles and obtained from numerical solution of the coupled PDEs and the wavefront speeds are measured (denoted by symbols). The wavefront speeds all take on the corresponding minimal values which agrees exactly with the analytical results (the solid and dashed lines fig. 8) given above. The steady cell growth rates in the absence of de-differentiation can be calculated analytically in a similar way as in sect. 4.3. To illustrate this, consider the parameters as in fig. 8 which corresponds to the case of richest behavior with four fixed points ($a_3 > a_2 > a_1$) $\vec{X}_0$, $\vec{X}_1$, $\vec{X}_2$, and $\vec{X}_3$ defined above. Denoting $\vec{N} = (N_Z, N_A, N_{A'})$, one gets

$$\begin{aligned}
\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} &= 2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}\vec{X}_3, \\
&\quad \text{for} \quad 2\Pi_{11}\sqrt{(1-a_1)(1+b_1)} = C_0^{(4)}, \\
\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} &= C_1^{(4)}\vec{X}_3 + \left(2\Pi_{33}\sqrt{(1-a_3)(1+b_3)} - C_1^{(4)}\right)\vec{X}_1, \\
&\quad \text{for} \quad 2\Pi_{33}\sqrt{(1-a_3)(1+b_3)} = C_0^{(4)}. \qquad (30)
\end{aligned}$$

for the parameters corresponding to fig. 8a ($b_1 = b_2$). And

$$\begin{aligned}
\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} &= C_2^{(4)}\vec{X}_3 + \left(2\Pi_{22}\sqrt{(1-a_2)(1+b_2)} - C_2^{(4)}\right)\vec{X}_2, \\
&\quad \text{for} \quad 2\Pi_{22}\sqrt{(1-a_2)(1+b_2)} = C_0^{(4)}, \\
\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} &= C_1^{(4)}\vec{X}_3 + \left(2\Pi_{33}\sqrt{(1-a_3)(1+b_3)} - C_1^{(4)}\right)\vec{X}_1, \\
&\quad \text{for} \quad 2\Pi_{33}\sqrt{(1-a_3)(1+b_3)} = C_0^{(4)}, \qquad (31)
\end{aligned}$$

for the parameters corresponding to fig. 8b ($b_1 < b_2$), where $C_0^{(4)}$, $C_1^{(4)}$, $C_2^{(4)}$ are defined by eqs. (27), (28) and (29), respectively.

The growth in cell populations can also be measured directly from the numerical solutions. Figure 9 shows the variation of steady total cell growth rate, as a function of the parameter b_2 indicating a monotonic growth rate. The theoretical predictions, $\mathrm{d}N_Z/\mathrm{d}t + \mathrm{d}N_A/\mathrm{d}t + \mathrm{d}N_{A'}/\mathrm{d}t$ given by (30) and (31), show perfect agreement.

In the case with cell de-differentiation, the cell can flip backwards from the differentiated state: A to Z $\Pi_{21} \equiv \Pi_{AZ} > 0$, and/or A' to Z $\Pi_{31} \equiv \Pi_{A'Z} > 0$, and/or A' to A $\Pi_{32} \equiv \Pi_{A'A} > 0$. With the different combinations of non-zero de-differentiation Π's, the resulting wave profiles are very rich; these results will be published elsewhere [47].

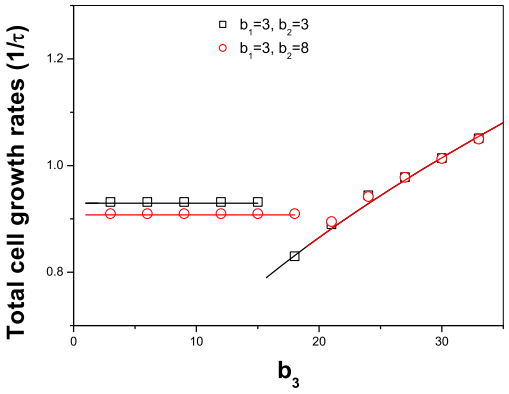


Fig. 9. Total cell growth rate for the 4-state model, defined as the increase in the number of A, A' and Z cells in one time interval τ, as a function of b_3. $\Pi_{11} = 0.5$, $\Pi_{22} = 0.4$, $\Pi_{33} = 0.3$, $\Pi_{12} = 0.1$, $\Pi_{23} = 0.1$, $a_1 = 0.4$, $a_2 = 0.5$, $a_3 = 0.6$. The symbols are total cell growth rates measured from the numerical solution of the PDEs. (□): $b_1 = 3$, $b_2 = 3$. (○): $b_1 = 3$, $b_2 = 8$. The solid curves are analytic results from eqs. (30) and (31).

6 Conclusion and outlook

In this paper, we develop the basic model to describe a system consisting of different cell types that possess diffusive motilities, proliferate and transform among different cell types. Our approach focuses on the macroscopic features of the dynamics resulting from the interplay of proliferation, diffusive motility and differentiation/de-differentiation. Effects of the detailed cellular biochemistry or biology are not modeled directly, the coarse features of proliferation, movement and differentiation are represented through the parameters in our model. Some of the assumptions in the model can be improved to be more biologically realistic, for example our model can be extended to include the case in which the proliferation rates for different cell states are state dependent and characterized by different flipping times τ_α (instead of a common τ_f as described in sect. 3). In appendix G, such an extended model is considered and it is shown that the resultant PDE can be cast into the same form as eq. (6), and hence can be described by similar dynamics. The resulting dependence of the wave propagation and wave speeds on the parameters governing proliferation, motility and differentiation are thus generic, and arise from the nonlinear dynamics of the system.

Although this work focused mainly on the analytic analysis of the reaction-diffusion type coupled PDEs and the associated wave propagation behavior in one dimension, our theory can be employed in a straightforward manner to simulate systems of many cell types and in a complicated spatial domain. Only a single lineage is considered in this paper, complicated situations of more lineages and further levels of differentiation can be investigated by numerical solution of the coupled reaction-diffusion system. Furthermore, simulation of the discretized version of the model can easily allow cell-cell in-

teractions to be taken into account. This will affect the pattern morphology and growth dynamics. This is under current investigation. Another important feature that is not considered in this work is the feedback regulation of proliferation, differentiation and motility by gene expression and signaling molecules. Presumably these regulation/control effects can be modeled by the parameters of our model (such as D_α, $\Pi_{\alpha,\beta}$ and g_α, etc.) whose dynamics are now coupled to the cell concentrations. These effects can be incorporated in the present model (see appendix F for incorporation of signaling molecules) and easily implemented in numerical simulations.

Our results on the wave propagation in the absence of de-differentiation can find applications in wound healing speed and morphology [1], especially in the early stage of healing. For example, dermal wound healing involves the interplay of cell differentiation, migration and proliferation [46], although in our present model the creation of new tissue and the associated mechanical deformation and stresses by the cells have yet to be taken into account. Our results indicate that the effect of finite de-differentiation will increase both the total cell growth rate and the propagation speed, resembling the acceleration of growth and spreading in cancer tumor. Carcinogenesis is often accompanied with backward differentiation to the more primitive stage of the lineage and the metastatic invasion of prostate cancer cells may be due to this de-differentiation. Also, the metastatic invasion of prostate cancer cells is often the source of the fast spread of the cancerous tumor [19]. Furthermore, the recent discovery of induced pluripotent stem cells (iPSC) [36] marked a breakthrough in stem cell biology and one anticipates that more detailed quantitative experimental results in cell differentiation/de-differentiation with controlled factors will be available. Hopefully our model can be modified and applied to the description of the transitions among various cell types.

Finally, we would like to remark that our model can be applied to the scenario of bacterial migration motion with mutation. The "switching of gear" in the wavefront speed in our model provides a mechanism for the bacteria to take up space more efficiently by sending energetic down-lineage descendent with a faster speed to occupy a new colony while the upper-lineage ancestors stayed back and move slowly. Such a strategy would be biologically advantageous for the whole bacterial community, and the interesting point is that it arises automatically from the dynamics of the system. The above situation can also be thought of in terms of human/animal population migration strategy: the young and aggressive descendants are sent out with a fast speed to conquer new territories. Such a "switch of gear" phenomenon has also been recently produced in the population dynamics of competitive Lotka-Volterra model with spatial diffusion [48].

This work has been supported by the NSC of ROC under the grant nos. NSC 100-2923-M-001-008-MY3, 101-2112-M-008-004-MY3, NCTS of Taiwan, and NSFC of China under grant no. 11047163/A05.

Appendix A. Mean-field analysis

Here we consider the mean-field approximation by averaging over the spatial dependence. Integrating over space in eq. (6) and denoting the fraction of α-type cell by $f_\alpha = \frac{1}{L^d}\int n_\alpha \mathrm{d}\vec{r} \equiv \langle n_\alpha \rangle$, and invoking no-flux boundary conditions, one gets

$$\frac{\mathrm{d}f_\alpha}{\mathrm{d}t} = -g_\alpha f_\alpha + \frac{1}{\tau_f}\sum_\sigma \Pi_{\sigma\alpha}\langle n_0 n_\sigma\rangle + D_f \sum_\sigma \Pi_{\sigma\alpha}\sum_{\mu\neq 0}\langle \nabla n_\mu \cdot \nabla n_\sigma\rangle, \quad \text{for} \quad \alpha \neq 0. \tag{A.1}$$

One replaces $n_\alpha(\vec{r},t)$ by its spatial average $f_\alpha(t)$ in mean-field approximation, and one gets a system of ODEs

$$\frac{\mathrm{d}f_\alpha}{\mathrm{d}t} = -g_\alpha f_\alpha + \frac{1}{\tau_f}\sum_\sigma \Pi_{\sigma\alpha} f_0 f_\sigma, \quad \text{for} \quad \alpha \neq 0. \tag{A.2}$$

However, in most situations the mean-field approximation will not be valid since the concentration gradient term, which represents the interfacial driven growth, is very important and cannot be ignored. This can be seen in the simple 2-states case ($q = 2$). Equation (A.1) gives

$$\frac{\mathrm{d}f_1}{\mathrm{d}t} = -g_1 f_1 + \frac{1}{\tau_f}\Pi_{11} f(1-f) + D_f \Pi_{11}\langle |\nabla n_1|^2\rangle. \tag{A.3}$$

The last term is always positive and interfacial growth will drive the system to grow in number until filling up the whole space. Dropping the last term in the mean-field approximation leads to the usual logistic growth governed by the fixed points $n_1 = 0$ and $n_1 = 1 - g_1\tau_1/\Pi_{11}$, which leads to incorrect predictions.

Appendix B. Singular perturbation calculation of the wavefront profile

To look for propagation wave solution, we change to the moving frame of wave speed c using the new variable $z = x - ct$, and defining $\xi \equiv z/c = \sqrt{\epsilon}$, $U(z) \equiv u(\xi)$, $\epsilon \equiv 1/c^2$ as the small parameter for perturbation expansion. For notational convenience, we define $\alpha \equiv a_1$, $\beta \equiv 1/b_1$, and the origin of the profile can be shifted to the middle of the front such that $U(z=0) = \frac{1-\alpha}{2}$, then the boundary conditions become $u(-\infty) = 1-\alpha$ and $u(\infty) = 0$. Equation (9) reads

$$[1+\beta(1-u)]\epsilon u'' + u' + u(1-\alpha-u) = 0. \tag{B.1}$$

Writing

$$u(\xi) = u_0(\xi) + \epsilon u_1(\xi) + \ldots \tag{B.2}$$

and substituting into eq. (B.1), the $O(\epsilon^0)$ and $O(\epsilon^1)$ terms give

$$u_0' + u_0(1-\alpha-u_0) = 0, \tag{B.3}$$

$$[1+\beta(1-u_0)]u_0'' + u_1' + u_1(1-\alpha-2u_0) = 0. \tag{B.4}$$

Direct integration of (B.3) gives

$$u_0(\xi) = \frac{1-\alpha}{1+e^{(1-\alpha)\xi}}\,. \tag{B.5}$$

It is easy to show that eq. (B.4) can be written as

$$u_1' - \frac{u_0''}{u_0'}u_1 = -[1+\beta(1-u_0)]u_0'' \tag{B.6}$$

and with the integrating factor $|u_0'|^{-1}$, u_1 can be readily integrated to give

$$u_1 = -u_0'\left(\int [1+\beta(1-u_0)]\frac{u_0''}{u_0'}\,\mathrm{d}\xi + \text{const.}\right) \tag{B.7}$$

$$= -u_0'\Bigg(\ln|u_0'| + \beta\Bigg\{\alpha(1-\alpha)\xi - \frac{2(1-\alpha)}{1+e^{(1-\alpha)\xi}} - (1+\alpha)\ln(1+e^{(1-\alpha)\xi})\Bigg\} + \text{const.}\Bigg). \tag{B.8}$$

The integration constant is determined by the condition $u_1(0) = 0$, and after some algebra, one finally gets

$$u_1(\xi) = \frac{(1-\alpha)^2 e^{(1-\alpha)\xi}}{(1+e^{(1-\alpha)\xi})^2}\left(\ln\frac{4e^{(1-\alpha)\xi}}{(1+e^{(1-\alpha)\xi})^2} + \beta\left((1-\alpha)\left(\alpha\xi - \frac{1-e^{(1-\alpha)\xi}}{1+e^{(1-\alpha)\xi}}\right) + (1+\alpha)\ln\frac{2}{1+e^{(1-\alpha)\xi}}\right)\right). \tag{B.9}$$

With $y \equiv \frac{(1-\alpha)\xi}{2}$, u_1 can be expressed as

$$u_1 = -\left(\frac{(1-\alpha)}{2}\right)^2 \mathrm{sech}^2\, y\,([2+\beta(1+\alpha)]\ln\cosh y + \beta(1-\alpha)[\tanh y + y])\,, \tag{B.10}$$

and upon transforming back to the variable z, giving eq. (13).

Appendix C. General rules for determining the wavefront speed constraint and population growth rates

We consider a wavefront that connects two homogeneous states $\vec{X}$ and $\vec{Y}$. For local wavefront solutions, one writes $n_\alpha(x,t) = U_\alpha(x-ct)$, for $\alpha = 1,\ldots,q-1$, for some wavefront speed c to be determined. Then the reaction-diffusion type equations such as eq. (7) will lead to a system of ODEs with dynamics described by $(U_1,\ldots,U_{q-1},U_1',\ldots,U_{q-1}')$ in a $2(q-1)$-dimensional phase space. Suppose $\mathbf{X}^*$ and $\mathbf{Y}^*$ are two fixed points in the multi-dimensional phase space of the dynamical system that correspond to the homogeneous steady states X and Y respectively, such as the one in eq. (D.3). A propagating wavefront with speed c is characterized by a flow from $\mathbf{X}^*$ to $\mathbf{Y}^*$ as shown schematically in fig. 10. Suppose the first few components of the dynamical system represent the population/concentration profiles, with the physical requirement imposed that these population components to be non-negative. If the final fixed point $\mathbf{Y}^*$ includes some zero population components, *e.g.* $\mathbf{Y}^* = (0,*,*,\ldots)$, then the flow in phase space when approaching $\mathbf{Y}^*$ cannot spiral into $\mathbf{Y}^*$ otherwise some population components would be negative. This will impose a constraint on the eigenvalues of the Jacobian at $\mathbf{Y}^*$ and in turn result in a constraint on the speed $c \geq c_{\min}$, as demonstrated in various scenarios in this paper. Furthermore, in many situations (as in all the scenarios in this paper), the stable wavefront will select to propagate with the minimal speed $c_{\min}$. However, it should be noted that even though a wavefront propagation is possible if there is a flow connecting from $\mathbf{X}^*$ to $\mathbf{Y}^*$, whether this wavefront is stable or not is a separate issue. The stability of the wavefront can be tested by numerical solution of the PDEs, and in some situations stability analysis can be performed analytically. Furthermore, in all the numerical solutions we obtained, the wavefront always propagates with its minimal allowed speed and we never observe any steady wavefront that propagates with some speed not constrained by an upper limit. Therefore, we shall adopt the conjecture that the wavefront will propagate with its minimal constrained limit.

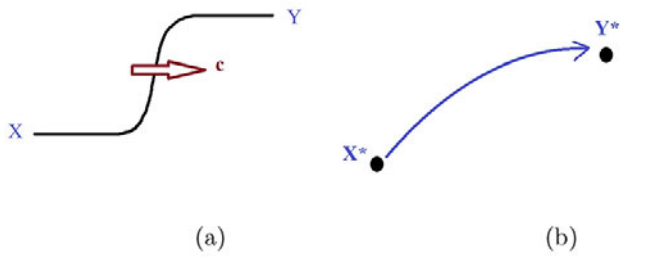


Fig. 10. Schematics illustrating the wavefront propagation for a wavefront obtained by connecting from the homogeneous steady states X to Y. (a) Schematic wavefront profile propagating with speed c. (b) $2(q-1)$-dimensional phase space flow from the fixed points $\mathbf{X}^*$ to $\mathbf{Y}^*$ resulting in a propagating wavefront in (a).

As an illustration, consider the 3-state OZA system with only pure differentiation ($\Pi_{21} = 0$). For the case of $a_2 \leq a_1$, there are only 2 fixed points $\mathbf{X}_0^* \equiv (0,0,0,0)$ and $\mathbf{X}_1^* \equiv (0,1-a_2,0,0)$. Since $\mathbf{X}_0^*$ is always stable, hence the only possibility is $\mathbf{X}_1^* \xrightarrow{C_0} \mathbf{X}_0^*$. This will result in wave profiles shown in fig. 4e or f. For the other case of $a_2 > a_1$, there are 3 possible scenarios: i) $\mathbf{X}_2^* \xrightarrow{C_0} \mathbf{X}_0^*$ resulting in stable wave profiles shown in fig. 4d. ii) $\mathbf{X}_2^* \xrightarrow{C_1} \mathbf{X}_1^* \xrightarrow{C_0} \mathbf{X}_0^*$ resulting in stable wave profiles shown in fig. 4g. iii) $\mathbf{X}_1^* \xrightarrow{C_0} \mathbf{X}_0^*$, however such a wavefront is stable only when n_1 is identically zero everywhere. Any finite perturbation in n_1 will destablize this wavefront. One anticipates that

the scenario of $\mathbf{X}_1^* \to \mathbf{X}_2^* \to \mathbf{X}_0^*$ is impossible since $\mathbf{X}_2^*$ cannot set a speed limit because its population components are all non-zero. This is further supported by that fact that $\mathbf{X}_2^*$ is highly unstable and the only unstable direction of $\mathbf{X}_1^*$ has a zero n_1 component.

Knowing the steady wavefront profiles and making use of the observation or conjecture that the wavefront speed always assumes its minimal value, the steady growth rates of the populations can also be calculated. Suppose in general the wave dynamics that connects $m+1$ fixed points in phase space is given by $\mathbf{X}^*_{\alpha_m} \xrightarrow{C_{\alpha_{m-1}}} \mathbf{X}^*_{\alpha_{m-1}} \ldots \mathbf{X}^*_{\alpha_1} \xrightarrow{C_{\alpha_0}} \mathbf{X}^*_{\alpha_0}$, and $\vec{X}$ is the $q-1$-dimensional vector that contains the fixed point values of $(n_1, n_2, \ldots, n_{q-1})$. Then it is easy to see that the population growth rates of the cells are given by

$$\frac{\mathrm{d}\vec{N}}{\mathrm{d}t} = \sum_{j=0}^{m-1} (\vec{X_{\alpha_{j+1}}} - \vec{X_{\alpha_j}}) C_{\alpha_j}, \tag{C.1}$$

where the cell populations $\{N_i\}|_{i=1,\ldots,q-1}$ are the components of the vector $\vec{N} \equiv (N_1, N_2, \ldots, N_{q-1})$.

Appendix D. Plane wavefront speeds calculations for the OZA (3-states) system

Since cell $Z(n_1)$ will differentiate into cell $A(n_2)$, under steady traveling wavefront conditions, these two wavefronts should have the same speed c. Therefore assuming local plane wavefronts and substituting $n_1(x,t) = U_1(x-ct)$ and $n_2(x,t) = U_2(x-ct)$ into (16), one obtains

$$\begin{aligned} &\Pi_{11}(1+b_1-U_1-U_2)U_1'' + \Pi_{21}(1-U_1-U_2)U_2'' + cU_1 \\ &+\Pi_{11}U_1(1-a_1-U_1-U_2) + \Pi_{21}U_2(1-U_1-U_2) = 0, \end{aligned} \tag{D.1}$$

$$\begin{aligned} &\Pi_{22}(1+b_2-U_1-U_2)U_2'' + \Pi_{12}(1-U_1-U_2)U_1'' + cU_2 \\ &+\Pi_{22}U_2(1-a_2-U_1-U_2) + \Pi_{12}U_1(1-U_1-U_2) = 0. \end{aligned} \tag{D.2}$$

Cell Differentiation: $\Pi_{21} \equiv \Pi_{AZ} = 0$

Defining $\tilde{c}_\alpha \equiv c/\Pi_{\alpha\alpha}$, eqs. (D.1) and (D.2) reduce to the 4-dimensional dynamical system

$$\begin{aligned} U_1' &= V_1, \\ U_2' &= V_2, \\ V_1' &= -\frac{\tilde{c}_1 V_1 + (1-U_1-U_2-a_1)U_1}{1-U_1-U_2+b_1}, \\ V_2' &= \\ &-\frac{\tilde{c}_2 V_2 + (1-U_1-U_2-a_2)U_2 + \frac{\Pi_{12}}{\Pi_{22}}(1-U_1-U_2)(U_1+V_1')}{1-U_1-U_2+b_2}, \end{aligned} \tag{D.3}$$

with two fixed points $(0,0,0,0)$, $(0,1-a_2,0,0)$ if $a_1 \geq a_2$ and one more fixed point $(n_1^{(S)}, n_2^{(S)}, 0, 0)$ emerges if $a_2 > a_1$.

The eigenvalues of the Jacobian at $(0,0,0,0)$ can be directly computed to give

$$\begin{aligned} 2\lambda_\pm^{(1)} &= -\frac{\tilde{c}_1}{1+b_1} \pm \sqrt{\left(\frac{\tilde{c}_1}{1+b_1}\right)^2 - \frac{4(1-a_1)}{1+b_1}}, \\ 2\lambda_\pm^{(2)} &= -\frac{\tilde{c}_2}{1+b_2} \pm \sqrt{\left(\frac{\tilde{c}_2}{1+b_2}\right)^2 - \frac{4(1-a_2)}{1+b_2}}. \end{aligned} \tag{D.4}$$

It is easy to see that all eigenvalues have negative real parts and the fixed point is stable. However, physical requirement of non-negative U_1 and U_2 imposed the constraints that the fixed point must be a stable node, but not stable focus. Hence the wave speeds for wavefront ending with $(n_1, n_2) = (0,0)$ must exceed the lower bounds given by

$$\begin{aligned} c \geq C_0 \equiv \max \Big[&2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}, \\ &2\Pi_{22}\sqrt{(1-a_2)(1+b_2)} \Big]. \end{aligned} \tag{D.5}$$

As in the case of the OZ case, one anticipates that the wavefronts will evolve and attain steady minimal speeds given by the RHS of (D.5).

The eigenvalues at $\mathbf{X}_1^* = (0, 1-a_2, 0, 0)$ can be similarly computed to be

$$\begin{aligned} 2\lambda_\pm^{(1)} &= -\frac{\tilde{c}_1}{a_2+b_1} \pm \sqrt{\left(\frac{\tilde{c}_1}{a_2+b_1}\right)^2 - \frac{4(a_2-a_1)}{a_2+b_1}}, \\ 2\lambda_\pm^{(2)} &= -\frac{\tilde{c}_2}{a_2+b_2} \pm \sqrt{\left(\frac{\tilde{c}_2}{a_2+b_2}\right)^2 + \frac{4(1-a_2)}{a_2+b_2}}. \end{aligned} \tag{D.6}$$

Note that $\lambda_\pm^{(2)}$ is always real with $\lambda_+^{(2)} > 0$ and $\lambda_-^{(2)} < 0$. If $a_1 \geq a_2$, then $\lambda_\pm^{(1)}$ is also real with $\lambda_+^{(1)} > 0$ and $\lambda_-^{(1)} < 0$. Thus for $a_1 \geq a_2$, this fixed point will not set any constraint on the wavefront speed.

On the other hand, the situation is different for $a_2 > a_1$ with the possibility of $\mathrm{Im}\{\lambda_\pm^{(1)}\} \neq 0$ resulting a focus. In this case, even though $\mathbf{X}_1^*$ is a saddle, it has 3 stable directions ($\mathrm{Re}\{\lambda_\pm^{(1)}\} < 0, \lambda_-^{(2)} < 0$) and only 1 unstable directions ($\lambda_+^{(2)} > 0$). Furthermore, with the presence of the $\mathbf{X}_2^* = (n_1^{(S)}, n_2^{(S)}, 0, 0)$ fixed point, a wavefront solution that flows from $\mathbf{X}_2^*$ to $\mathbf{X}_1^*$ is possible. But the physical requirement of non-negative n_1 imposes that the constraint that $\lambda_\pm^{(1)}$ must be real leads to

$$c \geq C_1 \equiv 2\Pi_{11}\sqrt{(a_2-a_1)(a_2+b_1)}, \tag{D.7}$$

for a wavefront that ends at $(n_1, n_2) = (0, 1-a_2)$. Note

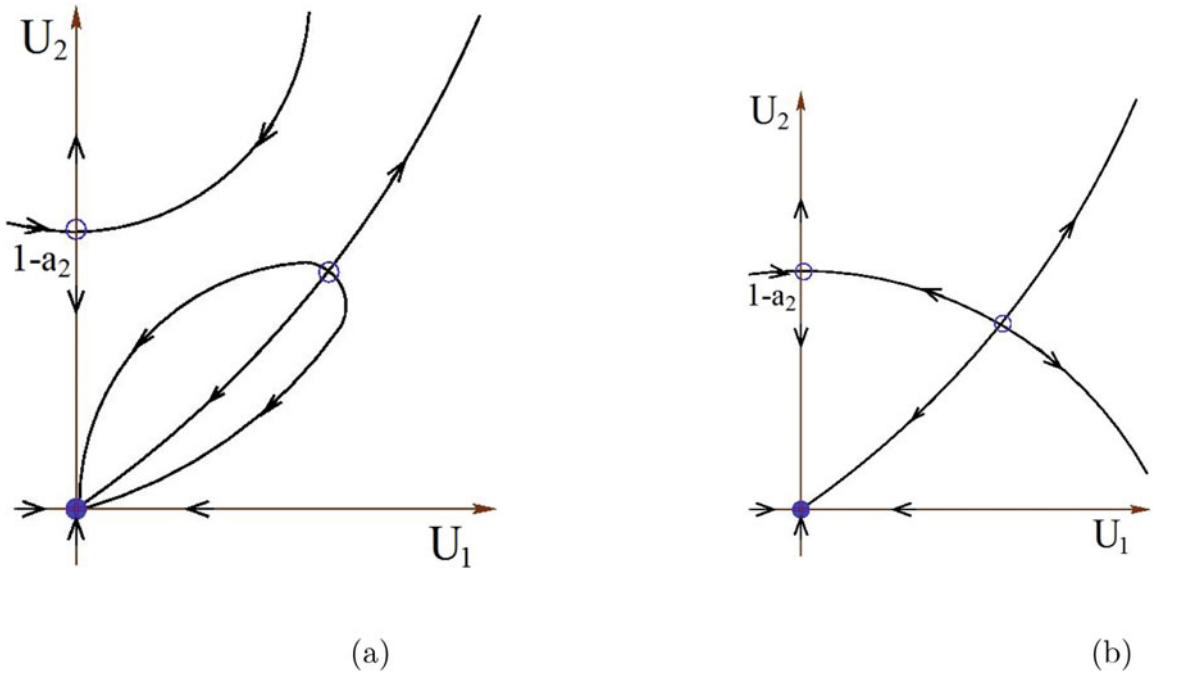


Fig. 11. Schematic flow in phase space for the case of $a_2 > a_1$. (a) small b_2 and (b) large b_2, in which the attraction of the stable direction of the fixed point $\mathbf{X}_1^* = (0, 1-a_2, 0, 0)$ becomes strong enough to attract the flow from $\mathbf{X}_2^* = (n_1^{(S)}, n_2^{(S)}, 0, 0)$, resulting in the propagation of wave profiles of the form shown in fig. 4g.

that in this case the eigenvector corresponding to $\lambda_\pm^{(2)}$ is given by

$$\hat{e}_\pm^{(2)} = \frac{1}{\sqrt{1+\lambda_\pm^{(2)^2}}}\left(0, 1, 0, \lambda_\pm^{(2)}\right). \tag{D.8}$$

Therefore with $c \geq C_1$, the saddle node $\mathbf{X}_1^*$ flows towards the stable node $\mathbf{X}_0^*$ along the unstable direction close the U_2 axis. And the saddle node becomes less unstable ($\lambda_+^{(2)}$ decreases) as b_2 increases.

Finally, the fixed point $\mathbf{X}_2^* = (n_1^{(S)}, n_2^{(S)}, 0, 0)$ exists if $a_2 > a_1$, and the eigenvalues cannot be expressed analytically, but they are given by the solution of a quartic polynomial equation and their numerical values of the eigenvalues and eigenvectors can be obtained precisely. $\mathbf{X}_2^*$ is a highly unstable saddle with 3 unstable and 1 stable directions. Note that even though it may be possible for phase space flow to $\mathbf{X}_2^*$ along its stable direction, $\mathbf{X}_2^*$ cannot set a speed limit of this wavefront. The only possibility of the above scenario is a flow from $\mathbf{X}_1^*$ to $\mathbf{X}_2^*$, but it is ruled out by examining the unstable direction of the $\mathbf{X}_1^*$ given by eq. (D.8), which has a large U_2 component towards $\mathbf{X}_0^*$ but a zero U_1-component (see fig. 11). This is further supported by our numerical solution in which wavefront connecting from $\mathbf{X}_1^*$ to $\mathbf{X}_2^*$ was possible only for n_1 identically zero and this wavefront is unstable upon local finite perturbation of $n_1 \neq 0$.

De-differentiation: $\Pi_{21} \equiv \Pi_{AZ} > 0$

In this case the 4-dimensional dynamical system reads

$$U_1' = V_1,$$

$$U_2' = V_2,$$

$$V_1' = -\frac{\tilde{c}_1 V_1 + (1-U_1-U_2-a_1)U_1 + \frac{\Pi_{21}}{\Pi_{11}}(1-U_1-U_2)(U_2+V_2')}{1-U_1-U_2+b_1}, \tag{D.9}$$

$$V_2' = -\frac{\tilde{c}_2 V_2 + (1-U_1-U_2-a_2)U_2 + \frac{\Pi_{12}}{\Pi_{22}}(1-U_1-U_2)(U_1+V_1')}{1-U_1-U_2+b_2}. \tag{D.10}$$

The above system has only two fixed points $(0,0,0,0)$ and $(n_1^*, n_2^*, 0, 0)$. Employing similar ideas as before, one can obtain the characteristic (quartic equation) for the eigenvalues. And by requiring that the fixed point $(0,0,0,0)$ is a stable node but not a stable focus, one can similarly obtain a minimal wavefront speed, $c_{\min}$, in terms of the parameters. Indeed numerical solutions of the PDEs indicate that steady wavefront propagate with $c_{\min}$ in this case.

Appendix E. Plane wavefront speeds calculations for the OZAA′ (4-states) system

Direct calculation of the Jacobian at $\mathbf{X}_0$ leads to the eigenvalues

$$\begin{aligned}
&-\frac{\tilde{c}_1}{1+b_1} \pm \sqrt{\left(\frac{\tilde{c}_1}{1+b_1}\right)^2 - \frac{4(1-a_1)}{1+b_1}}\,,\\
&-\frac{\tilde{c}_2}{1+b_2} \pm \sqrt{\left(\frac{\tilde{c}_2}{1+b_2}\right)^2 - \frac{4(1-a_2)}{1+b_2}}\,,\\
&-\frac{\tilde{c}_3}{1+b_3} \pm \sqrt{\left(\frac{\tilde{c}_3}{1+b_3}\right)^2 - \frac{4(1-a_3)}{1+b_3}}\,.
\end{aligned} \tag{E.1}$$

Notice that $\mathbf{X}_0^*$ is always stable. Using the physical requirement of non-negative n_i which constrains the $\mathbf{X}_0^*$ fixed point not to be a stable focus, revealed that a plane wavefront ending with $n_1 = n_2 = n_3 = 0$ can propagate for $n_i(x,t)$ with speed c above the minimal value:

$$\begin{aligned}
c \geq C_0^{(4)} \equiv \max\Big[&2\Pi_{11}\sqrt{(1-a_1)(1+b_1)}\,,\\
&2\Pi_{22}\sqrt{(1-a_2)(1+b_2)}\,,\\
&2\Pi_{33}\sqrt{(1-a_3)(1+b_3)}\Big].
\end{aligned} \tag{E.2}$$

Similarly, direct calculation of the Jacobian at $\mathbf{X}_1^*$ leads to the eigenvalues

$$\begin{aligned}
&-\frac{\tilde{c}_1}{a_3+b_1} \pm \sqrt{\left(\frac{\tilde{c}_1}{a_3+b_1}\right)^2 - \frac{4(a_3-a_1)}{a_3+b_1}}\,,\\
&-\frac{\tilde{c}_2}{a_3+b_2} \pm \sqrt{\left(\frac{\tilde{c}_2}{a_3+b_2}\right)^2 - \frac{4(a_3-a_2)}{a_3+b_2}}\,,\\
&-\frac{\tilde{c}_3}{a_3+b_3} \pm \sqrt{\left(\frac{\tilde{c}_3}{a_3+b_3}\right)^2 + \frac{4(1-a_3)}{a_3+b_3}}\,.
\end{aligned} \tag{E.3}$$

And the condition of non-negative n_1 leads to

$$\begin{aligned}
c \geq C_1^{(4)} \equiv \max\Big[&2\Pi_{11}\sqrt{(a_3-a_1)(a_3+b_1)}\,,\\
&2\Pi_{22}\sqrt{(a_3-a_2)(a_3+b_2)}\Big].
\end{aligned} \tag{E.4}$$

If $a_3 > a_2$, another fixed point $\mathbf{X}_2^* \equiv (0, n_2^{**}, n_3^{**}, 0, 0, 0)$ emerges and the Jacobian at $\mathbf{X}_2$ leads to the equation for the eigenvalues as

$$\begin{aligned}
&-\frac{\tilde{c}_1}{a_2+b_1} \pm \sqrt{\left(\frac{\tilde{c}_1}{a_2+b_1}\right)^2 - \frac{4(a_2-a_1)}{a_2+b_1}}\,,\\
&-\frac{\tilde{c}_2}{a_2+b_2} \pm \sqrt{\left(\frac{\tilde{c}_2}{a_2+b_2}\right)^2 + \frac{4n_2^{**}}{a_2+b_2}}\,,\\
&-\frac{\tilde{c}_3}{a_2+b_3} \pm \sqrt{\left(\frac{\tilde{c}_3}{a_2+b_3}\right)^2 + \frac{4n_3^{**}}{a_2+b_3}}\,.
\end{aligned} \tag{E.5}$$

And the condition of non-negative n_1 leads to

$$c \geq C_2^{(4)} \equiv 2\Pi_{11}\sqrt{(a_2-a_1)(a_2+b_1)}. \tag{E.6}$$

Finally, if $a_3 > a_2 > a_1$, another fixed point $\mathbf{X}_3^* \equiv (n_1^{***}, n_2^{***}, n_3^{***}, 0, 0, 0)$ emerges, however in this case the eigenvalues cannot be expressed analytically, but they are given by the solution of a polynomial equation of degree six and the numerical values of the eigenvalues and eigenvectors can be obtained precisely. This fixed point can be shown to be highly unstable, and cannot set a speed limit on the wavefront flowing into it.

Appendix F. Modeling interaction via signaling chemicals

Long-range chemical signaling is the most important mechanism for cell-cell communication that can lead to collective directed motion and quorum sensing in cell population [41–44]. The effect of communication via signaling molecules can be incorporated in the context of our model by introducing the concentration of the chemotactic signaling molecules $c(\vec{r},t)$. Equation (3) then becomes

$$\begin{aligned}
\frac{\partial n_\alpha}{\partial t} &= D_\alpha \nabla^2 n_\alpha - g_\alpha n_\alpha + \frac{n_0 p_\alpha(\{n_\sigma\}, c)}{\tau_\alpha},\\
&\quad + f_{ctx}(n_\alpha, c(\vec{r},t), \nabla c, \ldots), \quad \text{for} \quad \alpha \neq 0,
\end{aligned} \tag{F.1}$$

where f_{ctx} is some function which models the chemotactic signaling process. Note that the transition probability p_α becomes dependent on c to include the possibility of affecting the cell differentiation fate by the signaling molecule. $c(\vec{r},t)$ in turn can be modeled by reaction-diffusion type PDE as

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c - \gamma c + F(\{n_\sigma\}), \tag{F.2}$$

where D_c and γ are the diffusivity and decay rate of the signaling molecules, respectively, and F is some reaction function that represents the secretion of the signaling chemicals by the cells. For example, simple chemotactic motion that senses the chemical gradient and moves towards high concentration of c can be modeled with $f_{ctx} = -\mu\nabla\cdot(n_\alpha \nabla c)$ [49], where μ is a parameter that measures the chemotactic strength. And the secretion of c by the cells can be modeled as $F = \sum_{\sigma\neq 0} \frac{S_\sigma n_\sigma}{\beta_\alpha + n_\sigma}$, where S_σ and β_σ are, respectively, the secretion rate and Michaelis-Menton constant associated with the signaling molecule production of the σ-type cell.

Appendix G. Model with state dependent proliferation rates

In this appendix, we consider the situation in which the cell proliferation rate is dependent on the cell state, *i.e.*

instead of a common τ_f, the cell at state α has a proliferation rate τ_α ($\alpha \neq 0$). Then eq. (3) becomes

$$\frac{\partial n_\alpha}{\partial t} = D_\alpha \nabla^2 n_\alpha - g_\alpha n_\alpha + \frac{n_0 p_\alpha}{\tau_\alpha}, \quad \text{for} \quad \alpha \neq 0. \quad \text{(G.1)}$$

With $p_\alpha = \sum_\sigma \Pi_{\sigma\alpha} n_\sigma + \frac{h^2}{4} \sum_\sigma \Pi_{\sigma\alpha} \nabla^2 n_\sigma$ as before, one can adsorb τ_α into the proliferation probabilities by first defining a dimensionless proliferation/flipping rate, $\gamma_\alpha \equiv \tau_\alpha / T$ in terms of a common time unit T. For instance, T can be taken to be the proliferation/flipping rate of a particular cell state. And with $\mathcal{P}_{\sigma\alpha} \equiv \Pi_{\sigma\alpha}/\gamma_\alpha$, then $\mathcal{P}_{\sigma\alpha}/T = \Pi_{\sigma\alpha}/\tau_\alpha$ is the probability per unit time that an α-cell will proliferate a σ-cell into a neighboring site. The reaction-diffusion equation becomes

$$\frac{\partial n_\alpha}{\partial t} = D_\alpha \nabla^2 n_\alpha - g_\alpha n_\alpha + \frac{n_0}{T} \sum_\sigma \mathcal{P}_{\sigma\alpha} n_\sigma + n_0 D_f \sum_\sigma \mathcal{P}_{\sigma\alpha} \nabla^2 n_\sigma, \quad \text{for} \quad \alpha \neq 0, \quad \text{(G.2)}$$

where $D_f \equiv \frac{h^2}{4T}$. Rescale the space and time via $x \to x/\sqrt{D_f T}$ and $t \to t/T$, and defining $a_\alpha \equiv g_\alpha T/\mathcal{P}_{\alpha\alpha}$ and $b_\alpha \equiv D_\alpha/(D_f \mathcal{P}_{\alpha\alpha})$, one finally has

$$\frac{\partial n_\alpha}{\partial t} = \sum_\sigma \mathcal{P}_{\sigma\alpha} \Big[(n_0 + b_\alpha \delta_{\sigma\alpha}) \nabla^2 n_\sigma + (n_0 - a_\alpha \delta_{\sigma\alpha}) n_\sigma \Big], \quad \text{for} \quad \alpha \neq 0, \quad \text{(G.3)}$$

which is of the same form as in eq. (6).

References

1. M. Poujade *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104**, 15988 (2007).
2. P. Rosen, D.S. Misfeldt, Proc. Natl. Acad. Sci. U.S.A. **77**, 4760 (1980).
3. A.D. Lander, Biophys. J. **99**, 3145 (2010).
4. S. Sell, Environ Health Perspect. **101**, 15 (1993).
5. G.Q. Daley, Cold Spring Harb. Symp. Quant. Biol. **73**, 171 (2008).
6. T. Kawamura *et al.*, Nature **460**, 1140 (2009).
7. C.N. Shen *et al.*, Mech. Dev. **120**, 107 (2003).
8. M.C. Harrishingh *et al.*, EMBO J. **23**, 3061 (2004).
9. D.J. Pearson, Y. Yang, D. Dhouailly, Proc. Natl. Acad. Sci. U.S.A. **102**, 3714 (2005).
10. S. Zhuang *et al.*, J. Biol. Chem. **280**, 21036 (2005).
11. X.S. Zhang *et al.*, Endocrinology **147**, 1237 (2006).
12. A. Duckmanton *et al.*, Chem. Biol. **12**, 1058 (2005).
13. S. Cai *et al.*, J. Health Sci. **55**, 709 (2009).
14. M. Ben Amar, C. Chatelain, P. Ciarletta, Phys. Rev. Lett. **106**, 148101 (2011).
15. S. Fedotov, A. Iomin, Phys. Rev. Lett. **98**, 118101 (2007).
16. S. Fedotov, A. Iomin, Phys. Rev. E **77**, 031911 (2008).
17. C. Deroulers, M. Aubert, M. Badoual, B. Grammaticos, Phys. Rev. E **79**, 031917 (2009).
18. S.A. Menchon, C.A. Condat, Phys. Rev. E **78**, 022901 (2008).
19. L. Liu *et al.*, Proc. Natl. Acad. Sci. U.S.A. **108**, 6853 (2011).
20. J.V. Bonventre, J. Am. Soc. Nephrol. **14**, S55 (2003).
21. X. Fu *et al.*, Lancet **358**, 1067 (2001).
22. C. Zhang *et al.*, J. Cell. Mol. Med. **14**, 1135 (2010).
23. X. Sun *et al.*, Biol. Pharm. Bull. **34**, 1037 (2011).
24. C. Zhang *et al.*, Aging Cell **11**, 14 (2012).
25. J.Y. Chang, P.Y. Lai, Phys. Rev. E **85**, 041926 (2012).
26. F. Graner, J.A. Glazier, Phys. Rev. Lett. **69**, 2013 (1992).
27. R.M.H. Merks, J.A. Glazier, Physica A **352**, 113 (2005).
28. N. Chen *et al.*, Comput. Phys. Commun. **176**, 670 (2007).
29. N. Bellomo, A. Bellouquid, J. Nieto, J. Soler, Math. Models Methods Appl. Sci. **20**, 1179 (2010).
30. N. Bellomo, A. Bellouquid, J. Nieto, J. Soler, Math. Models Methods Appl. Sci. **22**, 1130001 (2012).
31. P. Hogeweg, Biosystems **64**, 97 (2002).
32. E.L. Bearer *et al.*, Cancer Res. **69**, 4493 (2009).
33. N. Bellomo, M. Delitala, Phys. Life Rev. **5**, 183 (2008).
34. J.T. Oden, A. Hawkins, S. Prudhomme, Math. Models Methods Appl. Sci. **20**, 477 (2010).
35. C.H. Waddington, in *The Strategy of the Genes. A Discussion of some Aspects of Theoretical Biology* (Alen & Unwin, 1957).
36. K. Takahashi, S. Yamanaka, Cell **126**, 663 (2006).
37. K. Takahashi *et al.*, Cell **131**, 861 (2007).
38. S. Yamanaka, Nature **460**, 08180 (2009).
39. J.-P. Capp, BioEssays **34**, 170 (2012).
40. A. Bellouquid, M. Delitala, *Mathematical Modeling of Complex Biological Systems* (Birkhäuser, Inc., Boston, 2006).
41. A. Bagorda, C.A. Parent, J. Cell Sci. **121**, 2621 (2008).
42. K.F. Swaney, C.H. Huang, P.N. Devreotes, Annu. Rev. Biophys. **39**, 265 (2010).
43. E.T. Roussos, J.S. Condeelis, A. Patsialou, Nature Rev. Cancer **11**, 573 (2011).
44. W.Y. Chiang, Y.X. Li, P.Y. Lai, Phys. Rev. E **84**, 041921 (2011).
45. R.A. Fisher, Ann. Eugen. **7**, 353 (1937) A. Kolmogorov, I. Petrovskii, N. Piscounov, *A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem*, in *Selected Works of A.N. Kolmogorov I*, edited by V.M. Tikhomirov (Kluwer, 1991) pp. 248–270.
46. J.D. Murray, *Mathematical Biology*, 3rd edition (Springer, New York, 2002).
47. M.X. Wang, P.Y. Lai, in preparation.
48. M.X. Wang, P.Y. Lai, Phys. Rev. E **86**, 051908 (2012).
49. G.F. Oster, J.D. Murray, J. Exp. Zool. **251**, 186 (1989).