

SCIENTIFIC REPORTS



OPEN

A kinetic model of multiple phenotypic states for breast cancer cells

Kang Qiu^{1,2}, Kai-fu Gao¹ , Li-jian Yang¹, Zhao-kang Zhang¹, Ran Wang¹, Hui-shu Ma¹ & Ya Jia¹

Quantitative modeling of microscopic genes regulatory mechanisms in an individual cell is a crucial step towards understanding various macroscopic physiological phenomena of cell populations. Based on the regulatory mechanisms of genes *zeb1* and *cdh1* in the growth and development of breast cancer cells, we propose a kinetic model at the level of single cell. By constructing the effective landscape of underlying stationary probability for the genes expressions, it is found that (i) each breast cancer cell has three phenotypic states (i.e., the stem-like, basal, and luminal states) which correspond to three attractions of the probability landscape. (ii) The interconversions between phenotypic states can be induced by the noise intensity and the property of phenotypic switching is quantified by the mean first-passage time. (iii) Under certain conditions, the probabilities of each cancer cell appearing in the three states are consistent with the macroscopic phenotypic equilibrium proportions in the breast cancer SUM159 cell line. (iv) Our kinetic model involving the TGF- β signal can also qualitatively explain several macroscopic physiological phenomena of breast cancer cells, such as the “TGF- β paradox” in tumor therapy, the five clinical subtypes of breast cancer cells, and the effects of transient TGF- β on breast cancer metastasis.

The regulation of cell phenotype decisions is critical for the survival of living cells. The clonal or stem cell was found with multiple phenotypic states, for example, the multiple states can arise in a cell with different gene expression states in *E. coli*¹. The cell state changes occur in response to microenvironmental signals and fluctuations^{2–17}. The multiple phenotypic states also exist in a variety of cancer cells^{18–24}, such as breast, colorectal cancers, etc.

Recent experimental observations²⁵ demonstrated there are three mammary epithelial cell phenotypic states (i.e., the stem-like, basal, and luminal states) in human breast cancer cell lines (the primary tumors SUM159 and SUM149), and the subpopulations of cancer cells purified for a given phenotypic state return towards equilibrium proportions of three phenotypes over time. It was found that the phenomenon of phenotypic proportions in human breast cancer cell lines is not due to differential growth rates of cells in the basal, stem-like, or luminal state but rather to interconversion between the three states, and a Markov model in which breast cancer cells transit stochastically between states was proposed to explain those experimental observations.

The observed breast cancer populations²⁵ are composed of a large number of cancer cells, although the cancer cells transition stochastically between three states, it is assumed that each cancer cell has the same gene regulatory pathways or kinetics of genes regulatory mechanisms *in vivo*. Understanding how the macroscopic phenotypic equilibrium proportions arise in each cancer cell and how the multiple states coexistence of an individual cell is mapped onto various macroscopic phenomena at the level of the whole cancer populations, implies that we ought to structure useful kinetic models of microscopic regulatory mechanisms at the single cell level. Therefore, the question of how the cell-state decisions of each cancer cell are made by genes regulatory mechanisms is critical outstanding. To our knowledge, however, the kinetic model of microscopic regulatory mechanisms for the multiple phenotypic states of an individual cancer cell is still unknown so far.

Although there are a large number of genes involved in the multiple phenotypic states of an individual cancer cell, a few key genes regulations might determine the cancer cell's phenotype or invasion and metastasis, and the cancer cell's response to microenvironmental signals (such as oestrogen, TGF- β , survival factors, cytokines and

¹Institute of Biophysics and Department of Physics, Central China Normal University, Wuhan, 430079, China.

²Department of Mathematics and Physics, Xuzhou Medical University, Xuzhou, 221004, China. Correspondence and requests for materials should be addressed to Y.J. (email: jiay@phy.ccnu.edu.cn)

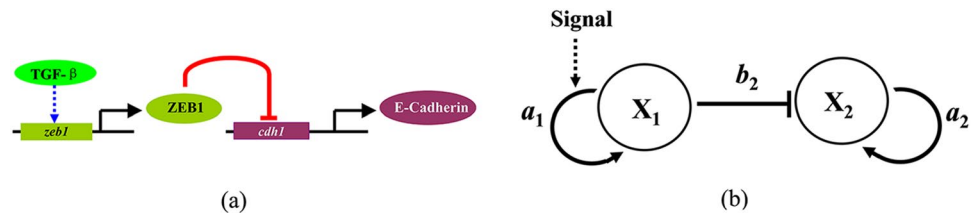


Figure 1. A schematic diagram of key genes regulations. (a) The microscopic regulatory mechanisms between genes *zeb1* and *cdh1* in breast cancer cells^{24,29,30,45}, where the expression of EMT-associated transcription factor ZEB1 can be induced by TGF- β signaling⁵¹. (b) A general kinetic model of the genes regulatory mechanisms, where a_1 and a_2 are the self-activation rates of genes X_1 and X_2 , b_2 is the strength of inhibition by the transcriptional factor of X_1 .

extracellular matrix)^{26–28}. For example, genes *zeb1* (the transcription factor ZEB1) and *cdh1* (encoding the protein E-Cadherin) play a vital role in cancer cells developmental processes^{29,30}, especially in the epithelial-mesenchymal transition (EMT) process, which is a key developmental program that is often activated during cancer invasion and metastasis. Thus, interesting questions now arise: Can the macroscopic phenotypic equilibrium phenomena at the level of the whole breast cancer populations be understood by the multiple states coexistence of each cancer cell at the level of single cell? What is the kinetic model of the key genes regulatory mechanism in a cancer cell?

In this paper, based on the transcriptional regulatory mechanisms between two key genes (*zeb1* and *cdh1*) in the developmental process of breast cancer cells, we proposed a general kinetic model of the genes regulation mechanisms for each cancer cell, our results showed that each cancer cell also exists three phenotypic states (i.e., the stem-like, basal, and luminal states), and there are interconversions between the three phenotypic states. In order to quantify the properties of phenotypic transition (or switching) between states, a theoretical formula of mean first-passage time is derived. Most interestingly, our general kinetic model of genes regulation mechanisms at the single cell level could help one to understand some macroscopic physiological phenomena at the level of whole breast cancer cell population, such as the phenotypic equilibrium in subpopulations of breast cancer lines²⁵, the “TGF- β paradox” in tumor therapy^{31–41}, the five clinical subtypes of breast cancer cells^{42,43}, and the effects of transient TGF- β on cancer metastasis⁴⁴.

The paper is arranged as follows. Firstly, a general kinetic model for the multiple phenotypic states of each breast cancer cell is proposed at the level of single cancer cell, and a theoretical formula of mean first-passage time for the phenotypic switching between states is derived by using an approximate Fokker-Planck equation. Secondly, we study the multiple phenotypic states coexistence and the phenotypic switching of an individual cancer cell. Then, the expression levels of genes *zeb1* and *cdh1* and the probabilities of an individual cancer cell appearing in three phenotypic states are compared with those of the human breast SUM159 line. Most interestingly, several clinical and therapeutic phenomena of breast tumors are qualitatively discussed by virtue of the general kinetic model. We end with the conclusions and discussions.

General kinetic model of key genes regulations

The stochastic kinetic model. In the developmental process of breast cancer cells, it was found that the transcription factor ZEB1 can promote EMT through inhibiting the expression of gene *cdh1* (which encodes the adhesion protein E-Cadherin) as shown in Fig. 1(a)^{24,29,30,45}. The E-Cadherin is a kind of transmembrane protein and essential for the stable cell-cell adhesion, and plays an important role in cellular development and cancer metastasis through modulating the EMT and the mesenchymal-epithelial transition (MET)^{29,30,45}. The low expression of E-Cadherin (through allelic loss and methylation/hyper-methylation of 5’CpG sites of *cdh1*) can promote tumor metastasis and malignancy in the early stage of a tumor, the high expression of E-Cadherin can induce new tumors forming at distant organs in the late stage of the tumor^{46–48}. The activation of *zeb1* induces the stem-like cells by inhibiting the expression of mir-200 family members which repress the stemness-associated factors such as SOX2 and KLF4^{28,49,50}. The expression of EMT-associated transcription factors (such as SNAIL1, SNAIL2, ZEB1, ZEB2, and LEF1) can be induced by TGF- β signal⁵¹.

Above microscopic regulatory mechanisms of the key genes *zeb1* and *cdh1* in the developmental process of breast cancer cells can be described by a general regulatory model as shown in Fig. 1(b), where X_1 represents the gene *zeb1*, and X_2 represents the gene *cdh1*. In the deterministic description, the kinetic model of the genes regulatory mechanisms with a transcriptional negative regulation can be written as following ordinary differential equations in the dimensionless:

$$\frac{dx_1}{dt} = a_1 \frac{x_1^n}{\theta^n + x_1^n} + b_1 - k_1 x_1 = F_1(x_1) \quad (1)$$

$$\frac{dx_2}{dt} = a_2 \frac{x_2^n}{\theta^n + x_2^n} + b_2 \frac{\theta^n}{\theta^n + x_1^n} - k_2 x_2 = F_2(x_1, x_2) \quad (2)$$

where x_1 and x_2 represent the expression level of genes X_1 and X_2 , respectively. a_1 and a_2 are the self-activation rates of genes, b_1 is the basal expression rate of X_1 , and b_2 is the strength of inhibition by the transcriptional

factor of X_1 . k_1 and k_2 are the self-degradation rates. θ represents the threshold which is the critical value needed for appreciable changes, and n is the Hill coefficient which controls the steepness of the sigmoidal function. The parameter values are $k_1 = k_2 = 1.0$, $\theta = 0.5$, $n = 4$, $b_1 = 0.2$, and $b_2 = 1.0$ for simplicity. It is hypothesized that the transcriptional negative regulation on X_2 by X_1 and the self-activation of X_2 do not simultaneously occur and the regulations in equation (2) follow an “or” rather than “and” logic¹³.

The regulations and expressions of genes are the intracellular random biochemical events^{4–8, 10}, and the stochasticity plays an important role in regulating cell-state equilibria in subpopulations of cells²⁵. In the stochastic description, the equations (1) and (2) are described by following stochastic differential equations:

$$\frac{dx_1}{dt} = F_1(x_1) + \xi_1(t) \quad (3)$$

$$\frac{dx_2}{dt} = F_2(x_1, x_2) + \xi_2(t) \quad (4)$$

where $\xi_1(t)$ and $\xi_2(t)$ are Gaussian white noises with zero means and $\langle \xi_1(t)\xi_1(s) \rangle = 2D_1\delta(t-s)$, $\langle \xi_2(t)\xi_2(s) \rangle = 2D_2\delta(t-s)$. Here we consider a homogeneous and non-correlation situation $D_1 = D_2 \equiv D$ which represents the total effect of intrinsic and extrinsic noises. Hence, the probability distribution $P(x_1, x_2, t)$ of equations (3) and (4) obeys the Fokker-Planck equation^{52–54}:

$$\frac{\partial P(x_1, x_2, t)}{\partial t} = -\frac{\partial}{\partial x_1}[F_1(x_1)P(x_1, x_2, t)] - \frac{\partial}{\partial x_2}[F_2(x_1, x_2)P(x_1, x_2, t)] + D\left[\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right]P(x_1, x_2, t) \quad (5)$$

In the equilibrium case, the stationary probability $P_{st}(x_1, x_2)$ of equation (5) represents the states distribution of cancer cells. An effective potential function $U_{st}(x_1, x_2)$ for nonequilibrium system is defined by the stationary probability:

$$U_{st}(x_1, x_2) = -\ln[P_{st}(x_1, x_2)] \quad (6)$$

Each minimum of the potential function $U_{st}(x_1, x_2)$ corresponds to one state (or phenotype) of a cancer cell. The phenotypic switching of a cancer cell means that the state of the cancer cell moves from one minimum of potential landscape to another.

A formula of phenotypic switching. To quantify the properties of phenotypic switching between states in the case of multiple phenotypic states coexistence, one can calculate the escape time from one steady state of $U_{st}(x_1, x_2)$ to another. A rigorous definition of escape time out of y_1 is provided by the mean first-passage time (MFPT) τ of the stochastic process $y(t)$ to reach the point y_2 with initial condition $y(t=0) = y_1$. This is given by^{55, 56}

$$\tau = \int_{y_1}^{y_2} \frac{dy}{D(y)P_{st}(y)} \int_{-\infty}^y P_{st}(k)dk. \quad (7)$$

If the noise $\xi_1(t)$ in equations (3) is neglected, the stationary expression level x_1^{st} of gene X_1 can be obtained by setting $dx_1/dt = 0$, then one can substitute $x_1 = x_1^{st} \equiv x_0$ into equation (4), and

$$F_2(x_0, x_2) = a_2 \frac{x_2^4}{\theta^4 + x_2^4} + b_2 \frac{\theta^4}{\theta^4 + x_0^4} - k_2 x_2 \quad (8)$$

which corresponds to the potential function:

$$U(x_0, x_2) = \frac{a_2\theta}{2\sqrt{2}} \left[\arctg\left(\frac{x_2^2 - \theta^2}{\sqrt{2}\theta x_2}\right) - \frac{1}{2} \ln\left(\frac{x_2^2 - \sqrt{2}\theta x_2 + \theta^2}{x_2^2 + \sqrt{2}\theta x_2 + \theta^2}\right) \right] + \frac{1}{2} k_2 x_2^2 - \left(\frac{b_2\theta^4}{\theta^4 + x_0^4} + a_2 \right) x_2 \quad (9)$$

It is found that the potential function equation (9) is a bistable system with the given parameter values. Then, the probability distribution $P(x_0, x_2, t)$ of expression concentration of X_2 obeys the following Fokker-Planck equation:

$$\frac{\partial P(x_0, x_2, t)}{\partial t} = -\frac{\partial}{\partial x_2}[F_2(x_0, x_2)P(x_0, x_2, t)] + D\frac{\partial^2}{\partial x_2^2}P(x_0, x_2, t) \quad (10)$$

By using the stationary solution of equation (10) and the steepest-descent approximation to equation (7), the MFPT can be given by

$$\tau = \frac{2\pi}{|U''_{st}(x_0, x_2^{st})U''_{st}(x_0, x_2^u)|^{1/2}} \exp\left[\frac{U_{st}(x_0, x_2^u) - U_{st}(x_0, x_2^{st})}{D}\right]. \quad (11)$$

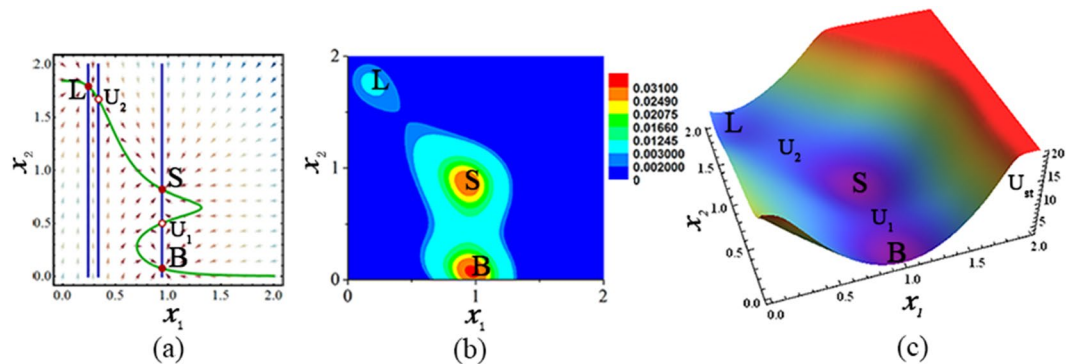


Figure 2. Deterministic trajectories, probability distribution, and potential landscape of model. (a)

Deterministic trajectories of kinetic model for an individual cancer cell. There are three steady states (filled circles) corresponding to three phenotypic states (the stem-like (S), basal (B), and luminal (L) states), and two unstable steady states (circles, U_1 and U_2). (b) Effective landscape of stationary probability distribution $P_{st}(x_1, x_2)$ in two dimensions. (c) Potential landscape $U_{st}(x_1, x_2)$ in three dimensions. The other parameters: $a_1 = 0.8$, $a_2 = 0.85$, and $D = 0.02$.

where x_2^{st} and x_2^u are the expression levels of gene X_2 at the steady state and unstable steady state, $U_{st}(x_0, x_2^{st})$ and $U_{st}(x_0, x_2^u)$ are the values of potential function at the steady state and unstable steady state, respectively.

Results and Discussions

Multiple phenotypes and phenotypic switching of a single breast cancer cell. In the last section, based on the regulatory mechanism of genes *zeb1* and *cdh1* in the growth and development of breast cancer cells, a general kinetic model was proposed. In this section, by using our kinetic model, it is shown that an individual breast cancer cell can exist in any of three possible phenotypic states (i.e., the stem-like, basal, and luminal states) which correspond to three basins of attractions of the probability landscape. The cell-state transition between the three states can be induced by the noise, and the properties of phenotypic switching are quantified by the mean first-passage time.

Deterministic trajectories, probability distribution, and potential landscape of model. By using the gene regulatory kinetic model, the multiple phenotypic states can arise in each breast cancer cell. Under the deterministic description equations (1) and (2), the deterministic trajectories of the kinetic model for each breast cancer cell show that there are three steady states and two unstable steady states as given by Fig. 2(a), and the three steady states correspond to the three phenotypes of each cancer cell: the stem-like (S), basal (B), and luminal (L) states.

Under the stochastic description of equations (3) and (4), the stationary probability $P_{st}(x_1, x_2)$ and the potential landscape $U_{st}(x_1, x_2)$ as given by Fig. 2(b,c) also show that a breast cancer cell can exist in any of three possible phenotypes (S, B, and L states) with different probabilities.

Multiple states coexistence and phase diagram. The variation of expression level of genes can be considered as the change of self-activation strength of genes. In our kinetic model, the self-activation strength a_1 of transcription factor ZEB1 can be induced by the microenvironmental signal (e.g., the TGF- β signal)⁵¹, the expression level of protein E-Cadherin determined by self-activation strength a_2 of gene *cdh1* can indicate the different stages of cancer. The expression of E-Cadherin in the early stage of some tumors is low (through allelic loss and methylation/hyper-methylation of 5'CpG sites of *cdh1*), while the expression is high in the late stage of the tumor^{46–48}.

A steady state of the kinetic model corresponds to a phenotypic state of a cancer cell. The steady state properties of the kinetic model show that there are the mono-stability (e.g., L or B), the bi-stability (e.g., LS, LB, or BS), and the tri-stability (e.g., LBS) under the different conditions. A phase diagram for the properties of phenotypic states of a single cancer cell is drawn in parameters (a_1, a_2) plane as shown by Fig. 3.

With the variation of parameter a_1 (or a_2), the phenotypic states of each cancer cell are very different. From the phase diagram, it can be found that, in the early stage of cancer (i.e., a_2 is small), a cancer cell is found in the L state at low level of TGF- β signal, and in the B state at high level of TGF- β signal. In the late stage of cancer (i.e., a_2 is large), however, a cancer cell is found in the multiple phenotypic states coexistence at high level of TGF- β signal, such as the LS, LB, LBS, and BS states.

Phenotypic switching between states due to noise. In the regions of multiple phenotypic states coexistence (e.g., the LS, LB, BS or LBS in Fig. 3), the cell-state transition can be induced by the noise, and the properties of phenotypic switching between phenotypic states are characterized by using the MFPTs (obtained by the theoretical formula equation (11) and the numerical simulation of stochastic process according to equations (3) and (4)). Furthermore, the barrier height of minima of potential function equation (9) can also be used to imply the properties of phenotypic switching, the height of barriers of two attractors (e.g., the B and S states) is defined by:

$$\Delta U_{u_1B} = U_{u_1} - U_B, \Delta U_{u_1S} = U_{u_1} - U_S. \quad (12)$$

where u_1 is the saddle point between B and S states.

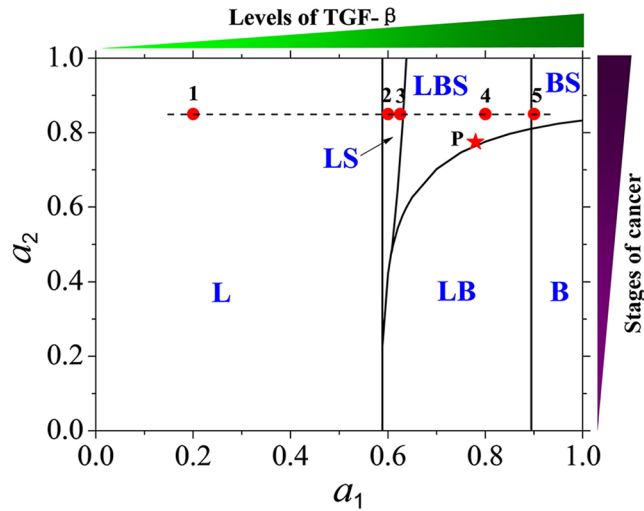


Figure 3. Phase diagram of multiple phenotypic states of a breast cancer cell in parameters (a_1, a_2) plane. The self-activation strength a_1 of gene *zeb1* can be induced by the TGF- β signal, and the self-activation strength a_2 of gene *cdh1* (encoding protein E-Cadherin) can indicate the different stages of cancers. The point P (the star) corresponds to phenotypic equilibrium distribution in the SUM159 breast cancer line²⁵. The five points (1, 2, 3, 4, 5) on the dashed line correspond to the five clinical subtypes of breast cancer cells⁴³.

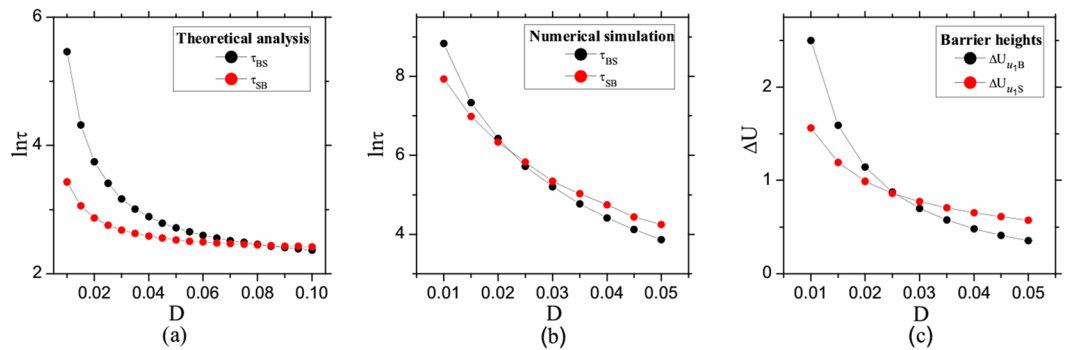


Figure 4. The properties of phenotypic switching due to noise in the multiple states coexistence regions. (a) The MFPTs (τ_{BS} and τ_{SB}) are obtained by the theoretical formula equation (11). (b) The MFPTs (τ_{BS} and τ_{SB}) are obtained by the numerical simulation of stochastic process of equations (3) and (4). (c) Relative barrier heights ΔU_{u_1B} and ΔU_{u_1S} are obtained by equation (12). The other parameters: $a_1 = 0.8$, $a_2 = 0.85$.

It is interesting to note that in equation (10) the dependence on the height of potential function between the steady state and unstable steady state is contained in the exponential factor. The higher the barrier height is, the larger the MFPT of the phenotype will be. The larger MFPT means that this phenotype is more difficult to switch to the other phenotype. Hence, the barrier heights of minima of potential function can also be used to imply the transition directionality of phenotypic switching. In the case of multiple phenotypic states coexistence, for example, the heights of barriers of two attractors (B and S states) are defined by equation (12).

For instance, taking into account the interconversions between B and S states, Fig. 4 shows that both the MFPTs (obtained by the theoretical formula and the numerical simulation) and the barrier heights of minima (the B and S states) of potential function are decreased with the increasing of noise intensity D , and there exists a threshold (the cross point) of noise intensity when the phenotype of cancer cells converts between B and S states.

With the increasing of noise intensity D , Fig. 4(a) shows that τ_{BS} and τ_{SB} are decreased, and the threshold of noise intensity is $D_c \approx 0.085$. When $D < D_c$, $\tau_{SB} < \tau_{BS}$, a cancer cell can change from S state to B state (i.e., $S \rightarrow B$), and the cancer cell has much larger probability to stay in B phenotype. However, when $D > D_c$, $\tau_{SB} > \tau_{BS}$, a cancer cell can change from B state to S state (i.e., $B \rightarrow S$), and the cancer cell has larger probability to stay in S phenotype.

It should be pointed out that, in the regions of multiple phenotypic states coexistence, the transitions between cell phenotypes can also be induced by the self-activation strength a_1 of ZEB1 through the increasing of the TGF- β signal, the self-activation strength a_2 of *cdh1* through the demethylation of 5' CpG sites of *cdh1*, and the repression strength b_2 of *cdh1* by ZEB1, respectively. Those data are not provided in this paper.

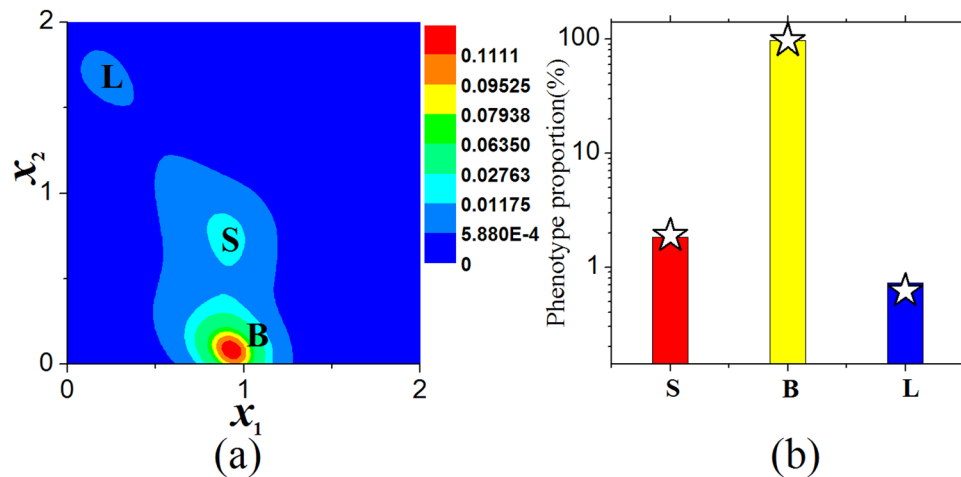


Figure 5. Phenotypic equilibrium in SUM159 line. (a) Effective landscape of stationary probability distribution $P_{st}(x_1, x_2)$ of three states (corresponding to point P in Fig. 3 with noise intensity $D = 0.01$) of an individual breast cancer cell. (b) Comparison the stationary probabilities proportions of three states (columns) of an individual cancer cell with the phenotypic equilibrium (stars represent the experimental data) in subpopulations of cancer cells in SUM159 line²⁵.

Phenotypic equilibrium in SUM159 cell line. The SUM159 populations are composed of a large number of breast cancer cells, in which the kinetics of genes regulatory mechanisms in each cancer cell is identical. It was found that²⁵ the phenomenon of phenotypic proportions in breast cancer cell lines is due to the interconversions between states. The multiple phenotypic states of each cancer cell at the level of single cancer cell can be mapped onto the macroscopic phenotypic equilibrium in subpopulations of cancer cells at the level of the whole cancer population, where the expression levels of genes *zeb1* and *cdh1* associated with different subpopulations are different.

Expressions level of genes zeb1 and cdh1 in SUM159 cell line. In SUM159 sorted cell subpopulations, the quantitative RT-PCR showed that (see the Fig. 1E in ref. 25) the expression level of gene *cdh1* (E-Cadherin) associated with stem and luminal states are specifically high, and the expression level of gene *cdh1* associated with basal state is low. However, the expression level of gene *zeb1* associated with stem state is same as that associated with basal state, and the expression level of gene *zeb1* associated with luminal state is lower than that associated with basal or stem state.

By utilizing our kinetic model, under the deterministic description, Fig. 2(a) shows that the relative expression levels of genes *zeb1* (i.e., x_1) and *cdh1* (i.e., x_2) at the three phenotypic states of each breast cancer cell are consistent with the experimental data of quantitative RT-PCR of genes associated with the stem-like, basal, and luminal states in SUM159 line (see the expression levels of Zeb1 and E-Cadherin in Fig. 1E of ref. 25).

Phenotypic equilibrium in SUM159 cell line. Under certain conditions (for example, at point P in the phase diagram Fig. 3), the cell-state equilibria in subpopulations of cancer cells²⁵ can be explained by our stochastic kinetic model.

Figure 5(a) shows the probability distributions of three phenotypes of a single cancer cell under certain noise intensity. Figure 5(b) shows that the cell-state proportions of three states of each cancer cell are consistent with those of phenotypic equilibrium in subpopulations of cancer cells (see the experimental data in Fig. 2B of ref. 25). The probability distribution is independent of the initial phenotype of each cancer cell, but depends on the microenvironmental fluctuations.

Our stochastic model can also predict that, with the increasing of noise intensity, the probabilities of each cancer cell in both L and S states become large, yet that of each cancer cell in B state becomes small as shown in Fig. 6.

Some macroscopic physiological phenomena of breast tumors. “Most cancer patients die from their disease as a result of metastasis”²⁷. Cancer cells in distinct phenotypic states exhibit differences in functional properties. In this section, it is showed that some macroscopic phenomena of breast cancer cells at the level of the whole cancer populations can also be qualitatively understood by using of the microscopic genes regulatory kinetic model at the level of single cancer cell.

The “TGF- β paradox” in breast tumor therapy. TGF- β is a multifunctional cytokine, and plays an essential role in modulation of cellular growth, maturation, differentiation, apoptosis, adhesion, and microenvironment. In tumor therapy, the effects of TGF- β on cancer cells are quite different.

In the early stage of cancers, it can induce the epithelial cell cycle arrest and promote apoptosis through its canonical signaling pathway via SMAD protein. In the late stage of cancers, however, it is linked with supporting cancer progression, such as higher cell motility, cancer metastasis, and immune evasion through the

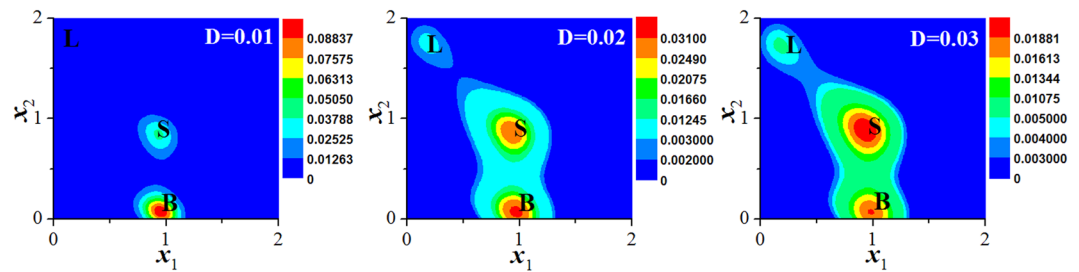


Figure 6. Effects of noises on phenotypic switching in the region of LBS coexistence. When noise intensity $D=0.01$, the stationary probability distribution of L state is too small to visualize. The other parameters: $a_1=0.8$ and $a_2=0.85$.

non-canonical signaling pathway. These contrasting, dichotomous TGF- β behaviors in cancer development and progression are referred to as the “TGF- β paradox”^{31–41}.

In fact, the contrasting, dichotomous TGF- β behaviors in the breast tumor therapy can be easily understood by using the phase diagram (Fig. 3) and the differences in functional properties of cancer cells within distinct phenotypic states.

In the early stage of cancers (i.e., a_2 is small), with the increasing of TGF- β , Fig. 3 shows that a cancer cell can transform from one phenotypic state (L) to the multiple phenotypic states coexistence (LB, or LB and LS), and then to another one phenotypic state (B). When the cancer cell is in L phenotype, it stays in a tumor, in which it cannot be identified and killed by immune cells located around that tumor. With the increasing of the TGF- β level, the cancer cell is possible to switch to B state which has larger motility and can be identified and killed by the immune cells. Thus, the increasing of TGF- β can inhibit tumor metastasis in the early stage of cancers.

However, in the late stage of cancers (i.e., a_2 is large), with the increasing of TGF- β signal, Fig. 3 shows that a cancer cell can transform from one state (L) to a successive multiple states coexistence (LS, LBS, and BS) process. In these coexisting states, the cancer cell has a certain probability of transition into S state, where it has the tumor-seeding ability, drug resistance, and the greatest cancer initiating capacity. Thus, the increasing of TGF- β can promote cancer progression in the late stage of cancers.

The five clinical subtypes of breast cancer cells. It was observed that there are three types of cancer stem-like cells (CSCs) in breast cancers (the basal, luminal, and basal-luminal CSCs)⁴². The basal CSCs are the mesenchymal-like state which comes from EMT, the luminal CSCs are the epithelial-like state which comes from MET, and the basal-luminal CSCs in which the two surface makers of both basal and luminal CSCs are simultaneously expressed.

More recently, Brooks *et al.*⁴³ found that these three phenotypes steady proportion distributions were apparently different between the five clinical breast cancer subtypes which are luminal A, luminal B, Her2 positive, basal-like, and triple negative (TN, the clinically aggressive claudin-low subtype). The breast cancer clinical classification is based on the cellular surface markers expression of estrogen and progesterone receptor as well as the growth factor receptor HER2. The TN is similar to basal-like, which are both the most difficult ones to cure^{58, 59}.

Our kinetic model of an individual cancer cell can be used to illustrate the five clinical subtypes of breast cancer cells since three types of CSCs have a similar pattern of gene expression and share a common regulatory pathway^{42, 43}.

For each breast cancer cell, it is found that the five subtypes observed by clinical trials could be respectively corresponded to the five points 1, 2, 3, 4, 5 in the phase diagram of phenotypes (Fig. 3) in the late stage of cancers. The probability distribution of the five points are shown in Fig. 7, which are similar to those of clinical classification of breast cancers (see the Fig. 2 of ref. 43).

The effects of transient TGF- β on cancer metastasis. It was demonstrated that⁴⁴ the consecutive high level of TGF- β can enhance the motility and intravasation of breast cancer cells by switching from cohesive to single cell motility but with low efficiency in forming new tumors at distant organs like the lung. However, the transient expression TGF- β can induce subsequent new tumor growth in the lungs. Thus, localized and reversible TGF- β signaling switches breast cancer cells from single cell motility to cohesive.

By virtue of the kinetic model, in the late stage of cancers (e.g., $a_2=0.8$), when TGF- β signal is consecutively at high level (e.g., $a_1 > 0.9$), Fig. 8 shows that the cancer cell is in the B state, in this case, the motility of the cancer cell is enhanced by the consecutive increasing of TGF- β .

However, when TGF- β is instantaneously increased to a high level (e.g., $a_1 > 0.9$) and decreased subsequently, with the decreasing of TGF- β , Fig. 8 shows that a cancer cell converts from B (Fig. 9(e)) state to LB (Fig. 9(d)), LBS (Fig. 9(c)), LS (Fig. 9(b)) states, and ends in L (Fig. 9(a)) state, respectively. In this process, the cancer cell gradually loses the ability to metastasize since the cancer cell can transfer into S or L state with a certain probability, and more and more of cancer cells have the ability to stick to distant organs and become resistant to immune cells, radiotherapy, and chemotherapy.

The motility capacities of cancer cells is decreased step by step as shown by the size of arrows in Fig. 8. The consecutive high expression of TGF- β can enhance the metastasis ability through promoting EMT, but the

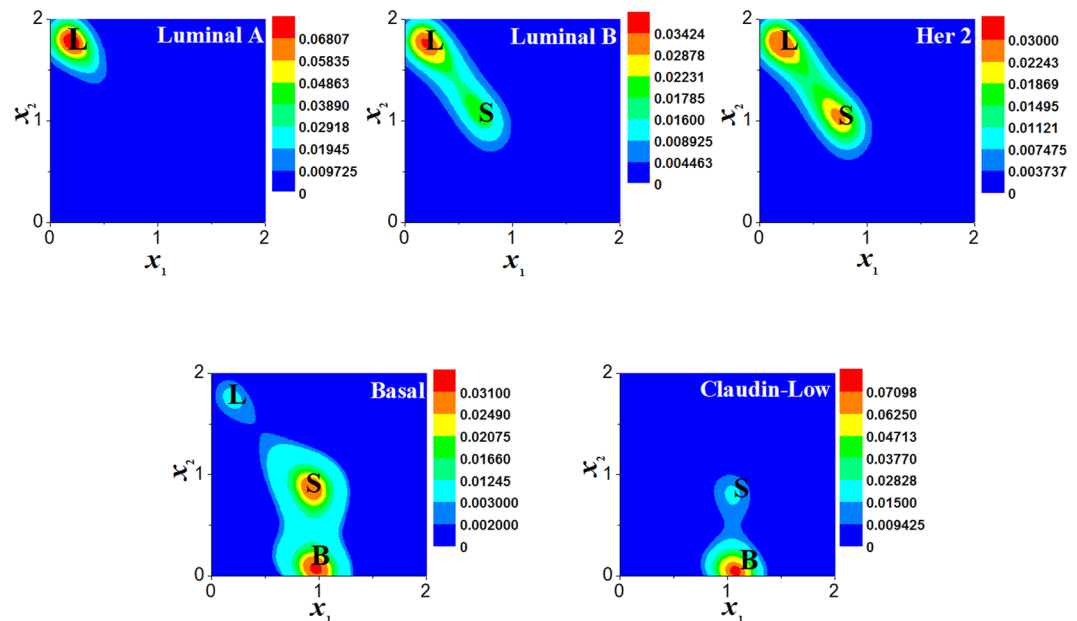


Figure 7. Effective landscape of stationary probability distributions of the three phenotypes. The probability distributions $P_{st}(x_1, x_2)$ corresponding to the five points 1, 2, 3, 4, and 5 in Fig. 3 with noise intensity $D = 0.02$ are similar to those of the five clinical subtypes (Luminal A, Luminal B, Her2, Basal, and Claudin-Low) of breast cancer⁴³.

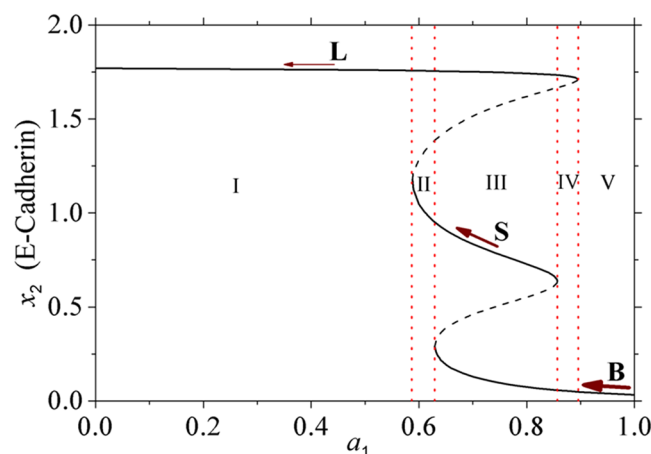


Figure 8. Bifurcation diagram of the expression level x_2 of E-Cadherin as a function of a_1 . Parameter a_1 represents the TGF- β signal since the self-activation strength a_1 of gene *zeb1* can be induced by the TGF- β signal. The motility capacity of cancer cells is decreased step by step as shown by the size of arrow in different regions. The parameter: $a_2 = 0.8$.

transient expression of TGF- β can induce the reversible process MET, which makes it easier for the cancer cells to stick to new sites through enhancing the expression of E-Cadherin, and form new tumors.

Conclusions and Discussions

In this paper, a general kinetic model of microscopic regulatory mechanisms between two genes (*zeb1* and *cdh1*) with a transcriptional negative regulation at the level of single cancer cell was proposed to uncover several interesting macroscopic physiological phenomena of cancer cells observed by experiments and clinical trials, such as the phenotypic equilibrium in populations of breast cancer cell lines, the “TGF- β paradox” in tumor therapy, the five clinical subtypes of breast cancer cells, and the effects of transient TGF- β on cancer metastasis.

By using the effective landscape through construction of underlying stationary probability, it is found that each breast cancer cell can also exist in any of three possible phenotypic states (i.e., the stem-like, basal, and luminal states) which correspond to three basins of attractions of the probability landscape. The transitions between the three states are induced by the noise (or the self-activation strength, or the repression strength of genes). The property of phenotypic switching is quantified by the mean first-passage time. Under certain conditions, the

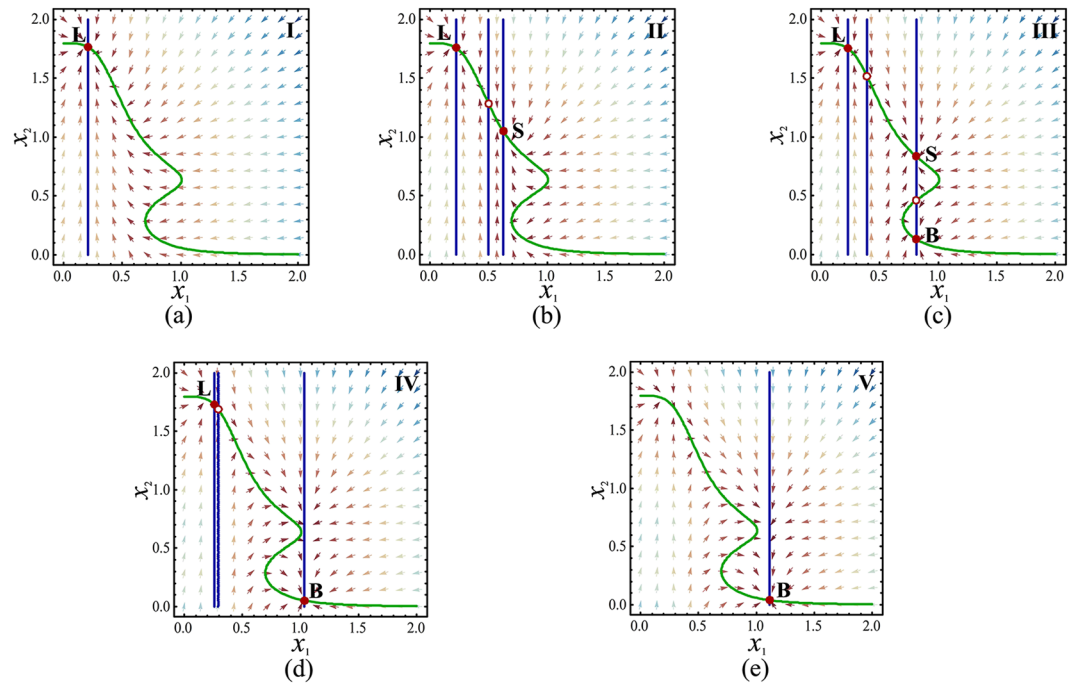


Figure 9. Deterministic trajectories of the five regions (I-V) in Fig. 8. The filled circle corresponds to the stable state, and the circle corresponds to the unstable steady states. (a) $a_1 = 0.3$. (b) $a_1 = 0.6$. (c) $a_1 = 0.7$. (d) $a_1 = 0.88$. (e) $a_1 = 0.95$.

probabilities of each breast cancer cell appearing in the three states are consistent with the macroscopic phenotypic equilibrium in subpopulations of breast cancer cells observed by the experiments²⁵. Comparing with the Markov model²⁵ proposed by Gupta *et al.* which is a macroscopic model, our kinetic model which depends on microscopic regulatory mechanisms between two genes is more fundamental and can explain more physiological phenomena of breast cancers observed by experiments and clinical trials.

The phase diagram of the deterministic kinetic model in parameters (a_1, a_2) plane given by Fig. 3 shows that there are multiple phenotypic states coexistence regions (e.g., LB, LBS, LS, BS). The five breast cancer clinical subtypes^{42,43} can be explained by different proportion distributions of three cell phenotypes of each cancer cell, and the “TGF- β paradox”^{31–41} in the tumor therapy can be understood by the phase diagram. With the increasing of TGF- β signal, the motility of cancer cells is increased. While the motility of cancer cells is decreased by decreasing TGF- β , then the cancer cells can reach and form new tumors at distant organs⁴⁴. Thus, high level of TGF- β signal is worse prognosis for tumors in the late stage of cancers.

In order to broadly explore the parameters used in the model, we first drew different phase diagrams with different parameters combinations, such as $(a_1, b_2), (b_1, a_2)$ etc., and found they all had the similar phase diagrams which contain the same multiple phenotypic states coexistence regions (e.g., LB, LBS, LS, BS). Second, we changed the Hill power parameter n and found that there always exist three steady states which correspond to the stem-like, basal, luminal states except $n < 2.24$.

Our kinetic model also predicts that there exists a threshold of noise intensity when the phenotype of a cancer cell transits between B state and S state. Due to the complexities of the equations which is highly nonlinear and have two unknowns, it is difficult to calculate the corresponding potential functions between any two states, such as L and B or L and S, except B and S states which coincidentally have the same value of x_1 . However, the role of noise on state conversions is also numerically studied between the states of L, B and L, S in the LBS region and we found that higher noise intensity induce the cancer cell state switching from L to B or S state which means enhancing noise intensity can promote the breast cancer metastasis. Although our model can reveal the multiple phenotypic states and phenotypic switching of breast cancer cells, it also should be mentioned that the real regulatory network of cell phenotype decisions is much more complex and there are probably other genes taking part in the dynamics of phenotypic switching. Above results reveal that the increasing of TGF- β can promote the metastasis ability of tumors through the EMT process, whereas the enhancing of the E-Cadherin expression, the noise intensity, and the transitory TGF- β signal can induce the forming of new tumors at distant organs through the opposite process MET.

In conclusion, by using a general kinetic model of microscopic regulatory mechanisms between two key genes, we demonstrated that the multiple phenotypic states of each cancer cell at the level of single cancer cell can be mapped onto some macroscopic physiological phenomena of breast cancer cells observed by experiments and clinical trials. Our results could provide new insights into the roles of microenvironmental signals and fluctuations at different stages of cancer cells, and the kinetic model might give some insights for various tumors clinical therapy strategies.

References

- Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
- Ozbudak, E. M., Thattai, M., Lim, H. N., Shraiman, B. I. & van Oudenaarden, A. Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**, 737–740 (2004).
- Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461 (2007).
- Cagatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J. & Suel, G. M. Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* **139**, 512–522 (2009).
- Süel, G. M., Garcia-Ojalvo, J., Liberman, L. M. & Elowitz, M. B. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545–550 (2006).
- Süel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J. & Elowitz, M. B. Tunability and noise dependence in differentiation dynamics. *Science* **315**, 1716–1719 (2007).
- Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078 (2005).
- Acar, M., Mettetal, J. T. & van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* **40**, 471–475 (2008).
- Jopling, C., Boue, S. & Belmonte, J. C. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. *Nat. Rev. Mol. Cell Biol.* **12**, 79–89 (2011).
- Thattai, M. & van Oudenaarden, A. Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530 (2004).
- Wolf, D. M., Vazirani, V. V. & Arkin, A. P. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J. Theor. Biol.* **234**, 227–253 (2005).
- Di Talia, S., Skotheim, J. M., Bean, J. M., Siggia, E. D. & Cross, F. R. The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* **448**, 947–951 (2007).
- Wang, J., Zhang, K., Xu, L. & Wang, E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA* **108**, 8257–8262 (2011).
- Zheng, X. D., Yang, X. Q. & Tao, Y. Bistability, probability transition rate and first-passage time in an autoactivating positive-feedback loop. *PLOS ONE* **6**, e17104 (2011).
- Lu, M. Y., Onuchic, J. & Ben-Jacob, E. Construction of an effective landscape for multistate genetic switches. *Phys. Rev. Lett.* **113**, 078102 (2014).
- Ge, H., Qian, H. & Xie, X. S. Stochastic phenotype transition of a single cell in an intermediate region of gene state switching. *Phys. Rev. Lett.* **114**, 078101 (2015).
- Antebi, Y. E. *et al.* Mapping differentiation under mixed culture conditions reveals a tunable continuum of T cell fates. *PLOS Biology* **11**, e1001616 (2013).
- Jordan, C. T., Guzman, M. L. & Noble, M. Cancer stem cells. *N. Engl. J. Med.* **355**, 1253–1261 (2006).
- Dalerba, P., Cho, R. W. & Clarke, M. F. Cancer stem cells: models and concepts. *Annu. Rev. Med.* **58**, 267–284 (2007).
- Rosen, J. M. & Jordan, C. T. The increasing complexity of the cancer stem cell paradigm. *Science* **324**, 1670–1673 (2009).
- Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nat. Rev. Cancer* **5**, 275–284 (2005).
- Fillmore, C. M. & Kuperwasser, C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res.* **10**, R25 (2008).
- Qiu, K. *et al.* A van der Waals-like transition between normal and cancerous phases in cell populations dynamics of colorectal cancer. *Scientific Reports* **6**, 36620 (2016).
- Chaffer, C. L. *et al.* Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74 (2013).
- Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
- Tyson, J. J. *et al.* Dynamic modelling of oestrogen signalling and cell fate in breast cancer cells. *Nat. Rev. Cancer* **11**, 523–532 (2011).
- Galván, J. A. *et al.* Expression of E-cadherin repressors SNAIL, ZEB1 and ZEB2 by tumour and stromal cells influences tumour-budding phenotype and suggests heterogeneity of stromal cells in pancreatic cancer. *Br. J. Cancer* **112**, 1944–1950 (2015).
- Gregory, P. A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10**, 593–601 (2008).
- Singh, A. B. *et al.* Claudin-1 up-regulates the repressor ZEB-1 to inhibit E-cadherin expression in colon cancer cells. *Gastroenterology* **141**, 2140–2153 (2011).
- Hashiguchi, M. *et al.* Clinical implication of ZEB-1 and E-cadherin expression in hepatocellular carcinoma (HCC). *BMC Cancer* **13**, 572 (2013).
- Massagué, J. TGF β In Cancer. *Cell* **134**, 215–230 (2008).
- Heldin, C. H., Landström, M. & Moustakas, A. Mechanism of TGF-beta signaling to growth arrest, apoptosis, and epithelial-mesenchymal transition. *Curr. Opin. Cell Biol.* **21**, 166–176 (2009).
- Yang, L., Pang, Y. & Moses, H. L. TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol.* **31**, 220–227 (2010).
- Ikushima, H. & Miyazono, K. TGFbeta signalling: a complex web in cancer progression. *Nat. Rev. Cancer* **10**, 415–424 (2010).
- de Visser, K. E., Eichten, A. & Coussens, L. M. Paradoxical roles of the immune system during cancer development. *Nat. Rev. Cancer* **6**, 24–37 (2006).
- Meulmeester, E. & Ten Dijke, P. The dynamic roles of TGF- β in cancer. *J. Pathol.* **223**, 205–218 (2011).
- Zarzynska, J. M. Two faces of TGF-beta1 in breast cancer. *Mediators Inflamm.* **2014**, 141747 (2014).
- Parvani, J. G., Taylor, M. A. & Schiemann, W. P. Noncanonical TGF- β signaling during mammary tumorigenesis. *J. Mammary Gland Biol. Neoplasia* **16**, 127–146 (2011).
- Morrison, C. D., Parvani, J. G. & Schiemann, W. P. The relevance of the “TGF- β paradox” to EMT-MET programs. *Cancer Lett.* **341**, 30–40 (2013).
- Risolino, M. *et al.* Transcription factor PREP1 induces EMT and metastasis by controlling the TGF- β -SMAD3 pathway in non-small cell lung adenocarcinoma. *Proc. Natl. Acad. Sci. USA* **111**, E3775–3784 (2014).
- Gong, C. *et al.* MiR-106b expression determines the proliferation paradox of TGF- β in breast cancer cells. *Oncogene* **34**, 84–93 (2015).
- Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports* **2**, 78–91 (2013).
- Brooks, M. D., Burness, M. L. & Wicha, M. S. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. *Cell Stem Cell* **17**, 260–271 (2015).
- Giampieri, S. *et al.* Localized and reversible TGF-beta signalling switches breast cancer cells from cohesive to single cell motility. *Nat. Cell Biol.* **11**, 1287–1296 (2009).
- Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).

46. Graff, J. R., Gabrielson, E., Fujii, H., Baylin, S. B. & Herman, J. G. Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. *J. Biol. Chem.* **275**, 2727–2732 (2000).
47. Matsumura, T., Makino, R. & Mitamura, K. Frequent down-regulation of E-cadherin by genetic and epigenetic changes in the malignant progression of hepatocellular carcinomas. *Clin. Cancer Res.* **7**, 594–599 (2001).
48. Lewis-Tuffin, L. J. *et al.* Misregulated E-cadherin expression associated with an aggressive brain tumor phenotype. *PLoS One* **5**, e13665 (2010).
49. Wellner, U. *et al.* The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* **11**, 1487–1495 (2009).
50. Brabletz, S. & Brabletz, T. The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO Rep.* **11**, 670–677 (2010).
51. Pickup, M., Novitskiy, S. & Moses, H. L. The roles of TGF- β in the tumour microenvironment. *Nat. Rev. Cancer* **13**, 788–799 (2013).
52. van Kampen, N. G. *Stochastic processes in physics and chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).
53. Pei, Q. M. *et al.* Fluctuations of cell population in a colonic crypt. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **89**, 032715 (2014).
54. Pei, Q. M. *et al.* Fluctuation and noise propagation in phenotypic transition cascades of clonal populations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **92**, 012721 (2015).
55. Stratonovich, R. L. *Topics in the Theory of Random Noise*, Vol. 1. (Gordon and Breach, New York, 1963).
56. Jia, Y. & Li, J. R. Transient properties of a bistable kinetic model with correlations between additive and multiplicative noises: Mean first-passage time. *Phys. Rev. E* **53**, 5764–5768 (1996).
57. From the editors, *Nat. Rev. Cancer* **11**, 309 (2011).
58. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
59. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* **98**, 10869–10874 (2001).

Acknowledgements

We thank Virtual Cell (supported by National Institutes of Health Grant P41RR013186 from the National Center for Research Resources) for software support. This work was supported by the National Natural Science Foundation of China under 11175068 and 11474117.

Author Contributions

K.Q. and Y.J. designed the project, K.Q., K.G., L.Y. Z.Z., R.W., H.M. performed the calculations, K.Q. and Y.J. analyzed the results. K.Q. and Y.J. wrote the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017