Supporting Information for: Quantifying the underlying landscape and paths of cancer

Chunhe Li[1], Jin Wang[1,2,*]
1 Department of Chemistry and Physics, State University of New York at Stony Brook, Stony Brook, NY, USA
2 State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin, China
∗ E-mail: jin.wang.1@stonybrook.edu

Supplementary Results

In order to show the landscape of the complete 32 dimensional system, we also used Langevin dynamics method to obtain landscape (Figure S1, three attractors from left to right separately represent cancer, normal and apoptosis state). For a 32 dimensional system, for visualization, we harnessed $RMSD$ (root mean squared distance) as the coordinate to reduce the dimensionality to 2 dimension ($RMSD = \sqrt{\sum_i^N (x_i - x_i^{ref})^2}$, N is the number of variables, and x_i^{ref} is the reference state, here we chose two potential minima of normal attractor and cancer attractor as the reference states). $RMSD$ represents the distance between a state point and reference point in state space. As a result, from 32-dimensional trajectory, we can generate two new coordinates $RMSD1$ and $RMSD2$, separately representing the distance from a state point to the reference state 1 (the potential minimum of normal attractor) and the reference state 2 (the potential minimum of cancer attractor). We can find that the landscapes using RMSD method based on Langevin dynamics (Fig. S1) possess the similar dynamics as using AKT and RB as the coordinates (Figure 2 in main text) based on the self consistent approximation. This shows that the two dimensional projection of landscape in AKT and RB state space can reflect the major dynamics of the full 32-dimensional gene network, and the 3 attractor landscape is not affected by choosing which gene pairs.

We'd like to emphasize that we used RMSD methods to calculate barriers and MFPTs in terms of 21 genes. We chose RMSD1 and RMSD2 as reaction coordinates. Since these coordinates measure the relative distance with respect to the normal and cancer states, we believe this is a reasonable choice of the reaction coordinates. We expect the barrier to be correlated with the MFPT when the choice of the reaction coordinates is reasonable. As we know, how to calculate the saddle point of high-dimensional system precisely is still challenging. Therefore, we believe that the RMSD method we employed in this manuscript might provide a possible route.

We calculated the mean first passage time (MFPT) from the trajectory based on Langevin dynamics. In Figure S1, three attractors from left to right respectively represent cancer, normal and apoptosis state, and this landscape is obtained by collecting the statistics through histogram or distribution of the temporal trajectories for 32 variables. Starting from a random initial state at Normal attractor, following the temporal evolution trajectory of system, we can find the time where the system first enters into the Cancer attractor. The difference between initial time and the final time is defined as the first passage time (FPT) for cancerization process (transition from normal to cancer). Repeating this process and obtaining the average of the FPT is defined as the mean first passage time (MFPT) for the cancerization process. In the same way, we can obtain the MFPT for the transition from Cancer attractor to Normal attractor.

Our simulation results showed that the landscape is critically influenced by the activation constant a or repression constant b. Fig. S2 show the landscape of the cancer network when activation constant a is changed, separately corresponding to $a = 0.4$, $a = 0.46$, $a = 0.48$, $a = 0.5$, $a = 0.52$, and $a = 0.6$. Landscape comparisons illustrated that with a gradually increasing from 0.4 to 0.6, the landscape of cancer network experiences a transition from monostable state (apoptosis), bistable state (apoptosis and normal coexist), to tristable state (normal, cancer and apoptosis coexist), and finally to another monostable state (cancer state). So, when a increases the cancer state will become more dominant. Similarly, in Figure S3, we give the change of landscape when repression strength b is changed. We can see that with b increased the landscape begins from a monostable apoptosis state, then become tristable state, and finally, get into a monostable normal state ($b = 1$).

Figure S4 show barrier changes when parameters $(M(j,i), i, j = 1, 2, ..., 32)$ are changed in terms of self consistent

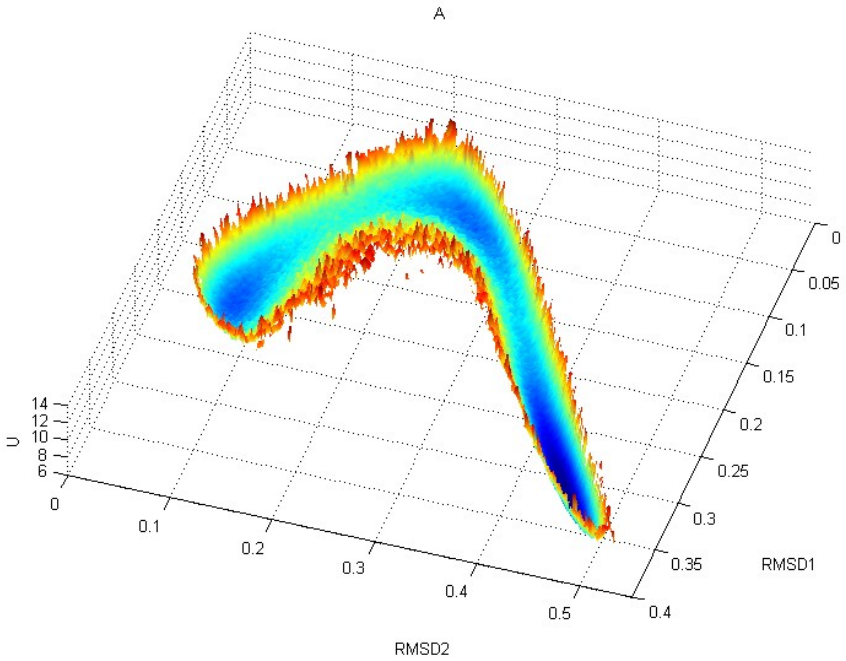

Figure S1. Landscape from Langevin dynamics method. Three attractors from left to right represent respectively cancer state, normal state and apoptosis state. RMSD1 represent the distance of states to the normal state, and RMSD2 represents the distance of states to the cancer state in terms of expression level of 21 major marker genes. Parameters are set as: a=0.5, b=0.5, k=1.

approximation method. The results in (A) are for 66 activation links, and The results in (B) are for 45 activation links. Blue stairs represent the barrier for normal state U_{SN} for different mutations, red stairs represent the barrier for cancer state U_{SC} for different mutations. X axis in (A) represent all 66 repression links, and X axis in (B) represent all 45 activation links. Choosing the parameters affecting barrier heights most, we acquired 25 most critical parameters or connections (15 of them are activation links and 10 of the others are repression links).

In the following, we employed the Langevin dynamics method to further obtain the change of barrier heights when these 25 parameters are changed, because by the Langevin dynamics method the landscape of the system can be acquired directly by the statistics of the trajectories of the system - not through approximation. The ODE governing the dynamics of the system has the form as:

$$F_i = -k * X_i + \frac{a * Xa_i}{S^n + Xa_i} + \frac{b * S^n}{S^n + Xb_i} \tag{1}$$

Here in Eq (1), i=1,2,...,32, so totally there are 32 equations. S represents the threshold (inflection point) of the explicitly sigmoidal functions, i.e., the strength of the regulatory interaction, and n is the Hill coefficient which determines the steepness of the sigmoidal function [1]. Here, parameters for Hill function are specified as: $S = 0.5, n = 4$. In addition, k is self-degradation constant, b is repression constant, and a is activation constant.

Here, Xa_i and Xb_i represent average regulation items separately for activation and repression from other nodes to certain node i. For every node i, Xa_i is defined as: $((X^n_{a(1)} * M(a(1), i) + X^n_{a(2)} * M(a(2), i) + ... + X^n_{a(m1)} * M(a(m1), i))/m1$, and Xb_i is defined as: $((X^n_{b(1)} * M(b(1), i) + X^n_{b(2)} * M(b(2), i) + ... + X^n_{b(m2)} * M(b(m2), i))/m2$. Here $a(1), a(2), ..., a(m1)$ is the number list of nodes which have activation interactions to node i, and $b(1), b(2), ..., b(m1)$ is the number list of nodes which have repression interactions to node i. $M(j, i)(i, j = 1, 2, ..., 32)$ is the element

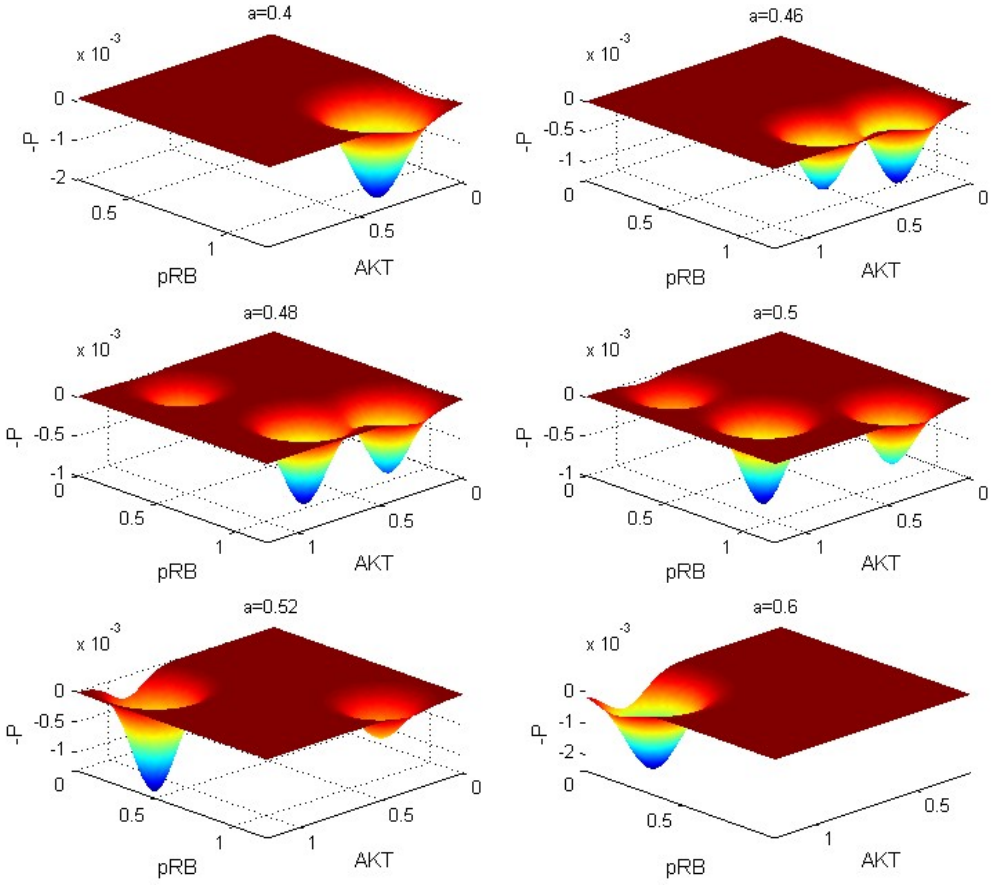

Figure S2. The change of landscape when activation strength a is changed (b is fixed at $b = 0.5$). The three attractors from left to right are separately corresponding to cancer, normal and apoptosis state. With activation strength a going up, the cancer state becomes more prominent, and the normal state and apoptosis state become less prominent.

of interaction matrix M characterizing the interaction type and the interaction strength from node j to node i. M is acquired by multiply interaction type matrix Mi (see Table S6) to interaction strength matrix Ms (Table S7): $M(j,i) = Mi(j,i) * Ms(j), i, j = 1, 2, ..., 32$. Here, we made an assumption that the regulation from one individual gene j to other genes has the same interaction strength which is determined by $Ms(j)$. Therefore, in Eq. (1), the first term represents self-degradation, the second term represents activation from other nodes (m1 nodes) to node i, and the last item denotes repression from other nodes (m2 nodes) to node i. For the nodes with only activation regulations and without repression regulations, a is set to $2 * a$ and b is set to 0. For the nodes with only repression regulations and without activation regulations, b is set to $2 * b$ and a is set to 0.

Then, we did sensitivity analysis for parameters by changing each activation or repression strength M_{ji} (increased by 50%).

We also quantified the effects of key parameters (15 activation parameters and 10 repression parameters) on the MFPT from cancer state to apoptosis state (τ_{CA}) from from stochastic dynamics approach (Figure S5A and B), in

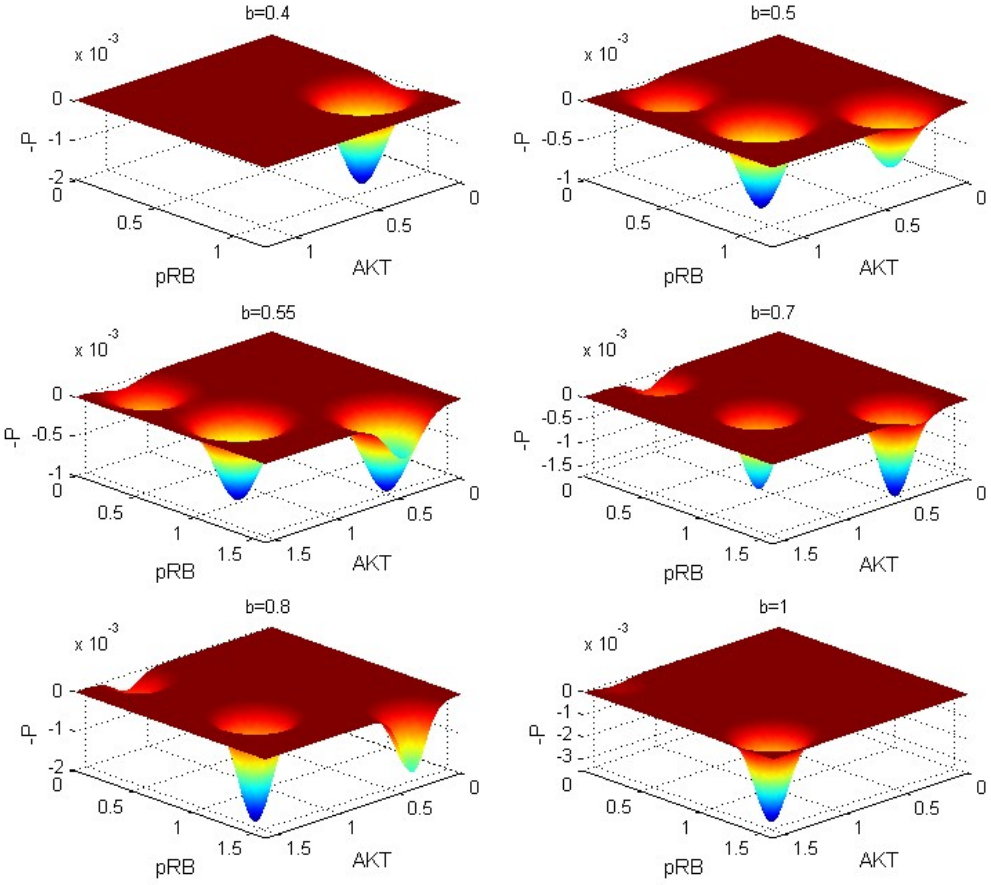

Figure S3. The change of landscape when the repression strength b is changed (a is fixed at $a = 0.5$). The three attractors from left to right are separately corresponding to cancer, normal and apoptosis state. With the repression strength b going up, the cancer state becomes more prominent. When b is very big ($b = 1$), the normal state becomes dominant.

order to uncover the key factors determining the transition from cancer attractor to apoptosis attractor (the death of cancer cells). This should provide some insight for killing cancer cells. In Figure S5, most of the mutations make τ_{CA} longer, which means that the cancer attractor becomes more stable and more time is needed for system to jump out from cancer attractor to apoptosis attractor. In the mean time, we find that link R9 ($PTEN \dashv AKT$) and R10 ($P53 \dashv VEGF$) can make τ_{CA} shorter significantly, meaning that cancer attractor lose their stability or death of cancer cells. This provides the second mechanism that AKT and VEGF serve as the valid target gene of anti-cancer, which is inducing caner cells to death by suppressing AKT or VEGF. In Figure S5 C and D, we show the effects of key parameters (15 activation parameters and 10 repression parameters) on the MFPT from normal state to apoptosis state (τ_{NA}). These uncover the the key factors determining the apoptosis process of normal cells. We find that link R9 ($PTEN \dashv AKT$) and R10 ($P53 \dashv VEGF$) can make τ_{NA} shorter significantly. This indicates that repressing AKT and VEGF can also lead the normal cells to death.

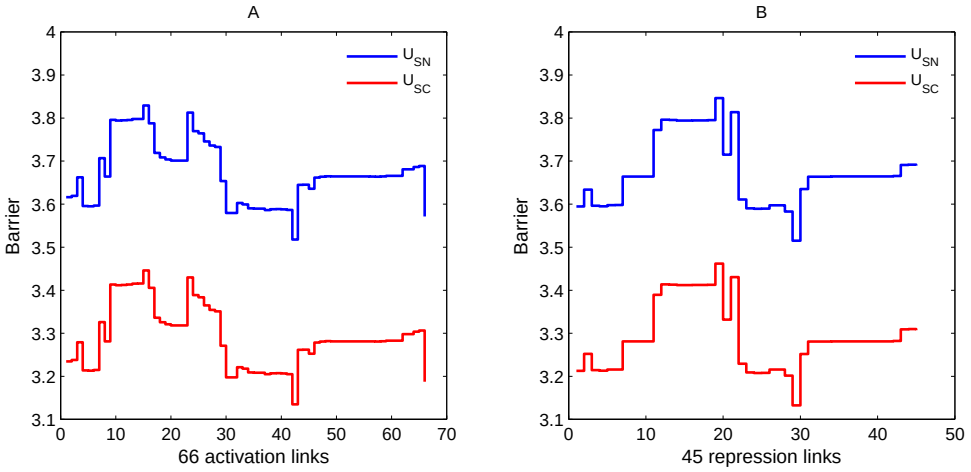

Figure S4. The effects of parameters (66 activation parameters and 45 repression parameters) on the barrier (U_{SN} and U_{SC}) from self consistent approximation approach. Here, U_{SN} represents the potential difference between the saddle point (between normal attractor and cancer attractor) and the local minimum in normal attractor. U_{SC} represents the potential difference between the saddle point (between normal attractor and cancer attractor) and the local minimum in cancer attractor.

Kinetic Transition Rates Between Cancer and Normal States

We also investigated the kinetics or speed of the transition between the two attractors (normal and cancer) according to the mean first passage time (MFPT), in order to further quantify the dynamics of the system. We calculated the mean first passage time (MFPT) from the trajectory based on Langevin dynamics. In Figure S1, left, middle and right attractor represent cancer, normal and apoptosis state respectively. This landscape is obtained by collecting the statistics through histogram or distribution of the temporal stochastic trajectories for 32 variables. Starting from a random initial state at normal attractor, following the temporal stochastic evolution trajectory of system, we can find the time where the system first enters into the cancer attractor. The time difference between initial time and the final time is defined as the first passage time (FPT) for the transition process from normal to cancer attractor. Repeating this process and obtaining the average of the FPT is defined as the MFPT for the transition process from normal to cancer attractor. In the same way, we can obtain the MFPT for the transition process from cancer to normal attractor.

MFPT reflects the average transition time of the system from one state to another state in the state space of gene networks, and thus can be used to quantify the ability of the system switching from one state to another state. The results of MFPT are shown in Fig. S6. It can be seen that for the normal to cancer process, the MFPT (τ_{NC}, blue line) become shorter when the activation strength a increases. This implies that when activation is larger the system need less time to jump from the normal state to the cancer state. Cells have more chances to stay in the cancer state, and therefore the increase of the activation is advantageous to the progress of normal to cancer transition. On the contrary, when the activation decreases, MFPT for normal to cancer become larger and MFPT for cancer to normal become smaller, which means that smaller activation leads to more stable normal attractor. We also explored the influence of changing the repression strength b on the landscape topography. Figure S6(C,D) show the MFPT results when the repression strength b is changed. It shows that when b is increased, the MFPT for normal to cancer process (τ_{NC}) increases, and the MFPT for cancer to normal process (τ_{CN}) declines. This demonstrates that the increase of repression strength also makes cancer attractor more stable and normal attractor less stable.

We also show how the fluctuations measured by diffusion coefficient D influence the barriers and MFPT (Fig. 5(E,F) and Fig. S6 (E,F)). It shows that at small noise level D, barrier is large and MFPT (escape time) is slower, which demonstrates the landscape of cancer system is stable against certain fluctuations. We can also see that when D increases, the barriers decline and the MFPT become faster. This indicates that larger noise destroy the stability

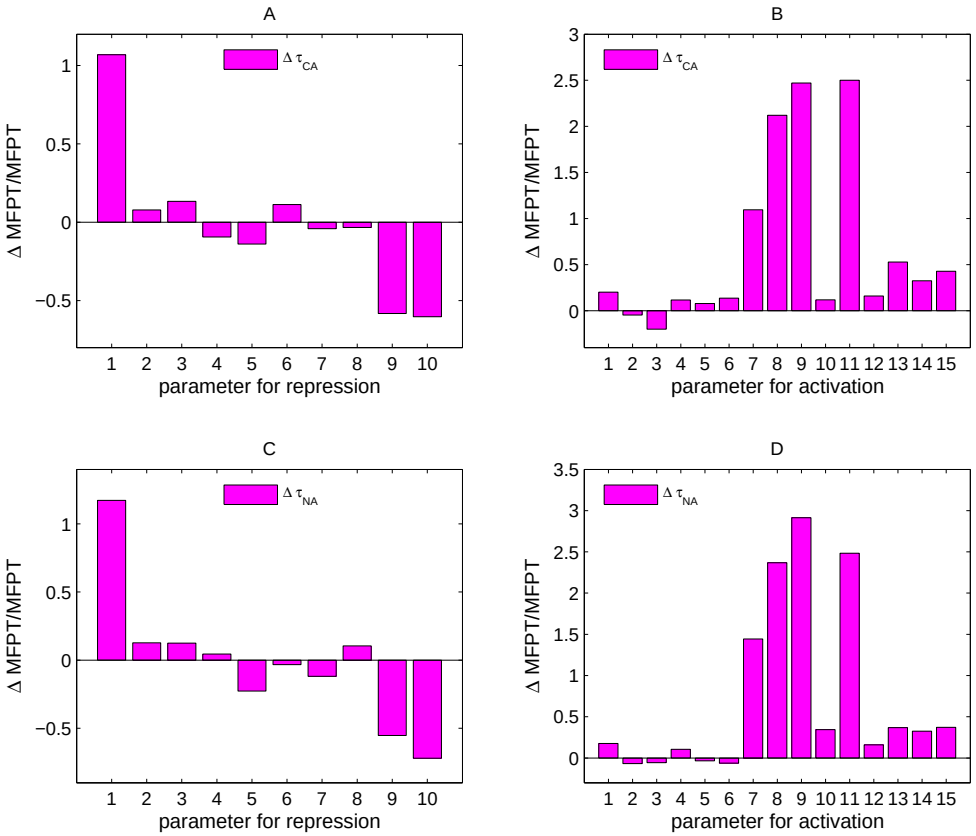

Figure S5. The effects of key parameters (15 activation parameters and 10 repression parameters) on the MFPT from cancer state to apoptosis state (τ_{CA}) as well as from normal state to apoptosis state (τ_{NA}) from Langevin dynamics approach. Y axis represents the percentage of MFPT changed. (A) and (B) are for from cancer state to apoptosis state. (C) and (D) are for from normal state to apoptosis state.

of system (landscape becomes flat) and transitions between different attractors become easy, so barriers and escape time between normal and cancer attractors both decline. Additionally, we also found that with noise D changing the barriers and MFPT do not change monotonically, which means that the noise have a complicated influence to the relative stability of normal and cancer attractor, i.e. we cannot say that the increase of noise is advantageous to normal or cancer state.

We can find that the global barrier heights and the MFPT have the similar trend for quantifying the transition between normal and cancer state, and both of them can serve as the quantitative measure to the global stability of the two attractors and kinetic speeds.

Construction of the Cancer Network

Cancer is a complex and heterogeneous disease displaying the degree of complexity at physiological, tissue and cellular levels. Recent tumor genome sequencing efforts have demonstrated that there may be thousands of cancer driver-mutating genes. These genes are diverse and have little overlap between different tumors. Nevertheless, most cancer

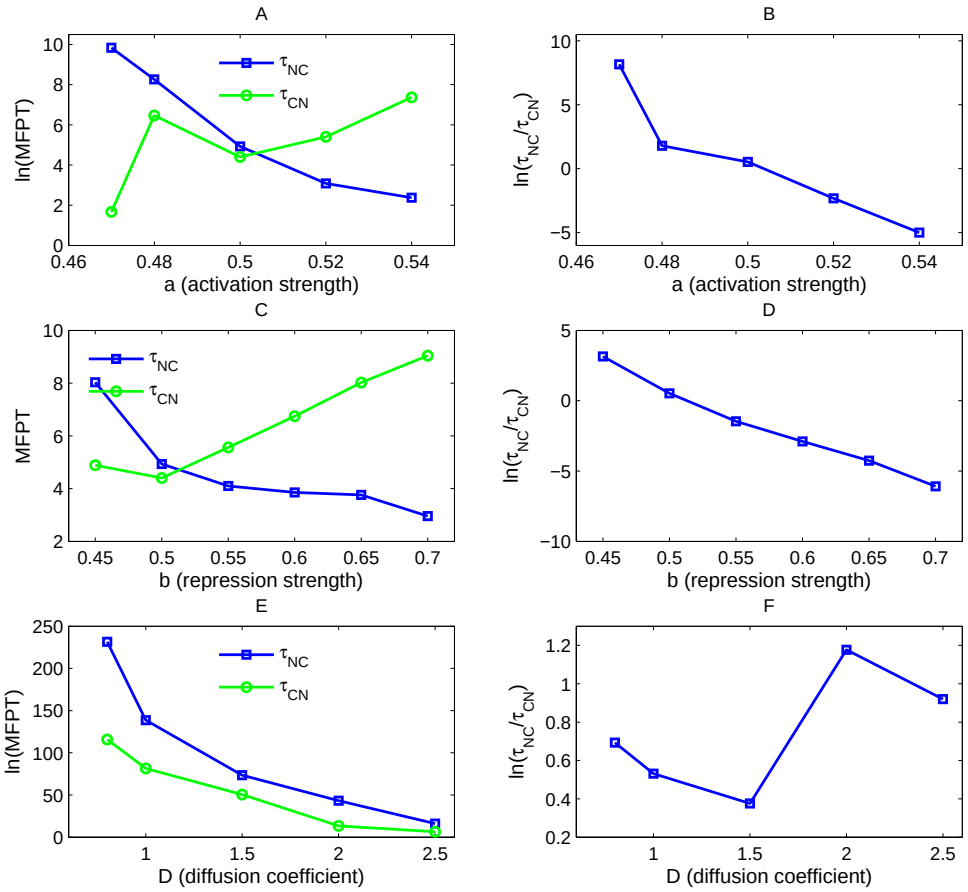

Figure S6. The change of MFPT when activation a, the repression strength b and noise level D are changed. Here, τ_{NC} represents the MFPT (first passage time) from normal attractor to cancer attractor. τ_{CN} represents the MFPT (first passage time) from the cancer attractor to the normal attractor).

driver-mutated genes reflect cancer hallmarks or functional modules. Each hallmark or functional module is composed of a set of functionally linked pathways. Therefore, it is possible to map the functional modules and the mutated genes onto a cancer network. By doing so, we can use the network to represent complexity, compute biological relationships and seek to uncover the biological principles and insights of cancer [2].

Weinberg et al. put forward the 10 hallmarks of cancer [3–5]. These hallmarks are characterized by some key cancer marker genes, such as EGFR for proliferative signal, VEGF for angiogenesis, HGF for metastasis, hTERT for unlimited replication, HIF1 for glycolysis, CDK2 and CDk4 for evading growth suppressors. Starting from these cancer marker genes and some critical tumor suppressor genes such as P53, RB, P21, and PTEN, we made an extensive literature search [6] for the interactions among these key genes as well as the interactions among other cancer associated genes, and constructed a cancer gene regulatory network with 32 gene nodes (Fig. 1). This cancer network includes 32 nodes (genes) and 111 edges (66 activation interactions and 45 repression interactions). In Fig. 1, arrows represent activation and short bars represent repression. The network mainly includes three kinds of marker genes: apoptosis marker genes (green nodes, including BAX, BAD, BCL2, and Caspase), cancer marker genes (magenta nodes, including AKT,

MDM2, CDK2, CDK4, CDK1, NFKB, hTERT, VEGF, HIF1, HGF, and EGFR), and tumor repressor genes (light blue nodes, including P53, RB, P21, PTEN, ARF, and CDH1). The brown nodes represent other genes. We provide the order numbers and names of 32 genes, as well related functions in Table S1. We also show the evidences for the connections in constructing the network in Table S2.

In the current cancer network, some important biological pathways have been indicated. For example, CDK4-CDK2-E2F1-RB represents the main pathway of cell cycle. CDKs (Cyclin-dependent kinases, each consisting of a catalytic CDK subunit and an activating cyclin subunit) regulate the cell's progression through the phases of the cell cycle by modulating the activity of key substrates. Tumor repressor RB (pRB) gene represses cell cycle regulator E2F1, which promotes cell proliferation by activate CDK2. In the meantime, CDK4 and CDK2 both repress pRB, which forms a positive feedback loop.

BAD, BAX, Caspase, and BCL2 are main proteins for apoptosis. When facing DNA damage, P53 represses the expression of BCL2, BCL2 represses Caspase, so to induce cell apoptosis.

MYC-AKT pathway represents a positive regulator of cell proliferation progress [7,8]. Activated MYC has a wide range of effects on cell cycle by activate Cyclin E/CDK2 complex and Cyclin D/CDK4 complex, down-regulating P21 levels, and up-regulating E2F1 expression. AKT promote cell proliferation by repressing P21 tumor repressor, and antagonize cell apoptosis by repressing BAD and activating MDM2 oncogene. Oncogene NFKB can promote cell cycle by activating oncogene MYC and CDK4. Meanwhile, NFKB can also repress tumor repressor PTEN.

Determination of Parameters for Cancer Network

Here, we designated parameters value uniformly (Eq. (1)), i.e. all activation strength a is same, and also for the repression strength b, because so far we have no access for the information about the regulation strength — or the magnitude of activation and repression parameters — between different genes in the cancer network. The default parameter values are: $a = 0.5$, $b = 0.5$, $k = 1$. In this way, the value of gene expression level of every gene are restricted between 0 to 1.

About the value of parameters, we choose parameter values according to the following criteria:

1, We chose parameter values according to some previous work [1, 9, 10], and used Hill cooperative binding expressions for the regulations. In the work of Huang et al [1], they explored a two gene system, where the region for producing bistability is given as: threshold S=0.5-1.5, Hill coefficient n=4-8. Here, we choose parameter values as: S=0.5, n=4. For the degradation, activation and repression strength, we set them uniformly for different variables and set them in the same magnitude, because so far for cancer network there is no such information about regulation strength which should come from the detailed biochemistry reactions involved in cancer system.

2, We chose those parameters that can satisfy some biological constrains, including producing steady state solution as well as producing multi-stability, since our purpose is to explore the transition between normal state, cancer state and apoptosis state (3 key stable states).

3, About the interaction strength matrix Ms, the default value of $Ms(j), j = 1, 2, ..., 32$ is 1. We did an extensive parameter search by giving $Ms(j)$ a gaussian perturbation with 0 as the mean and 0.5 as the standard deviation, in order to explore the solutions for the broader interaction strength parameter space. At every specific $Ms(j)$, by giving different initial conditions for the ODE (100000 different initial conditions), we can obtain the distribution of the steady state solutions (monostable, bi-stable or tri-stable state). We choose the parameter for $Ms(j)$ producing tri-stability (Table S7), which is corresponding to normal, cancer and apoptosis state.

Analyzing the relative gene expression level for the three state (Table S3), we can find that the apoptosis attractor state has higher expression of tumor repressor gene RB,P21,PTEN, lower expression of oncogene AKT, EGFR, VEGF, HGF, HIF1, hTERT, MDM2, CDK2, CDK4, and higher expression level of apoptosis marker gene Caspase. The normal attractor state has higher expression level of tumor repressor gene RB,P21,PTEN, higher expression level of oncogene AKT, EGFR, VEGF, HGF, HIF1, hTERT, MDM2, CDK2, CDK4, and lower expression level of apoptosis marker gene Caspase. The cancer attractor state has lower expression level of tumor repressor gene RB,P21,PTEN, much more higher expression level of oncogene AKT, EGFR, VEGF, HGF, HIF1, hTERT, MDM2, CDK2, CDK4, and lower

expression level of apoptosis marker gene Caspase. Biologically, this is consistent with our understanding for normal, cancer and apoptosis state [3–5].

Additionally, in barrier height and MFPT section (changing activation strength a and repression strength b), as well as in global sensitivity analysis section, we did the perturbation for both activation strength and repression strength, which show the relative robustness of current parameter regions. We also provided the results of changing individual interaction strength from any gene i to gene j in the sensitivity analysis section.

Self Consistent Mean Field Approximation

The time evolution the dynamical systems are governed by the diffusion equations. Given the system state $P(X_1, X_2, ..., X_n, t)$, where $X_1, X_2, ..., X_n$ represent the concentration or populations of molecules or species, we expected to have N-coupled differential equations, which are difficult to solve. Following a self consistent mean field approach [9, 11, 12], we split the probability into the products of individual ones: $P(X_1, X_2, ..., X_n, t) \sim \prod_i^n P(X_i, t)$ and solve the probability self-consistently. This can effectively reduce the dimensionality from M^N to $M \times N$, and thus make the computation of the problem tractable.

However, for the multi-dimensional system, it is still hard to solve diffusion equations directly. We start from moment equations and simply assume specific probability distribution based on physical argument, i.e. we give some specific connections between moments. In principle, once we know all moments, we can acquire the probability distribution. For instance, the Poisson distribution has only one parameter, so we may calculate all other moments from the first moment, the mean. In this work, we use gaussian distribution as approximation, which means we need two moments, mean and variance.

When the diffusion coefficient D is small, the moment equations can be approximated to [13, 14]:

$$\dot{\bar{\mathbf{x}}}(t) = F[\bar{\mathbf{x}}(\mathbf{t})] \tag{2}$$

$$\dot{\sigma}(t) = \sigma(t)\mathbf{A^T}(\mathbf{t}) + \mathbf{A}(t)\sigma(\mathbf{t}) + 2\mathbf{D}[\bar{\mathbf{x}}(\mathbf{t})]. \tag{3}$$

Here, $\mathbf{x}$, $\sigma(t)$ and $\mathbf{A}(t)$ are vectors and tensors, and $\mathbf{A^T}(t)$ is the transpose of $\mathbf{A}(t)$. The matrix elements of A is $A_{ij} = \frac{\partial F_i[X(t)]}{\partial x_j(t)}$. According to this equation, we can solve $\mathbf{x}(t)$ and $\sigma(t)$. Here, we consider only diagonal elements of $\sigma(t)$ from mean field splitting approximation. Therefore, the evolution of probabilistic distribution for each variable could be acquired using the mean and variance based on gaussian approximation:

$$P(x,t) = \frac{1}{\sqrt{2\pi}\sigma(t)} \exp - \frac{[x - \bar{x}(t)]^2}{2\sigma(t)} \tag{4}$$

The probability obtained above corresponds to one fixed point or basin of attraction. If the system has multistability, then there are several probabilistic distributions localized at every basin of attraction, with different variations. Therefore, the total probability is the weighted sum of all these probability distributions. The weighting factors (w_1, w_2) are the size of the basin, representing the relative size of different basins of attractions. For example, for a bistable system, the probability distribution takes the form: $P(x,t) = w_1 P^a(x) + w_2 P^b(x)$, here $w_1 + w_2 = 1$. Here, we determine the weights w_i by giving a large number of random initial conditions for ODEs to find solution, and then collect the statistics for different solution. For example, for a bistable system, if 10% initial condition goes to the first steady state, and 90% initial condition goes to the second steady state, then the weight w_1 for the first basin is 0.1 and w_2 for the second basin is 0.9. By giving large number of random different initial conditions for ODEs, we can solve the equations at a fixed parameter set. By collecting the statistics of the solution, we can determine if the system is monostable or bistable or multi-stable at current parameter region. In this work, the solution of 32 ODEs giving multiple (100000) different random initial conditions produce the multi-stability (bistable or tri-stable). For tri-stability, the probability distribution takes the form: $P(x,t) = w_1 P^a(x) + w_2 P^b(x) + w_3 P^c(x)$, $w_1 + w_2 + w_3 = 1$. For 32 dimensional system, we can acquire 32 dimensional probability distribution. To exhibit the results in a 2-dimensional space, we integrated out the other 30 variables and left two variables AKT and RB.

Finally, once we have the total probability, we can construct the potential landscape by the relationship with the steady state probability: $U(x) = -lnP_{ss}(x)$. In the gene regulatory network system, every parameter or link contributes to the structure and dynamics of the system, which is encoded in the total probability distribution, or the underlying potential landscape.

For nonequilibrium gene regulatory systems, the driving force F can not be written as the gradient of potential U, like the equilibrium case. In general, F can be decomposed into a gradient of the potential and a curl flux force linking the steady state flux $\mathbf{J_{ss}}$ and the steady state probability P_{ss} [12, 15]($\mathbf{F} = +\mathbf{D}/P_{ss} \cdot \frac{\partial}{\partial \mathbf{x}} P_{ss} + \mathbf{J_{ss}}(\mathbf{x})/P_{ss} = -D\frac{\partial}{\partial \mathbf{x}} U + \mathbf{J_{ss}}(\mathbf{x})/P_{ss}$). P_{ss} denotes steady state probability and potential U is defined as $U = -lnP_{ss}$. The probability flux vector $\mathbf{J}$ of the system in concentration or gene expression level space $\mathbf{x}$ is defined as [13]: $\mathbf{J}(\mathbf{x}, t) = \mathbf{F}P - \mathbf{D} \cdot \frac{\partial}{\partial \mathbf{x}} P$.

In the 32-dimensional protein concentration space, it's hard to visualize 32-dimensional probabilistic flux. Approximately, we explored the associated 2-dimensional projection of flux vector: $J_1(x_1, x_2, t) = F_1(x_1, x_2)P - D\frac{\partial}{\partial x_1} P$ and $J_2(x_1, x_2, t) = F_2(x_1, x_2)P - D\frac{\partial}{\partial x_2} P$.

Langevin Dynamics Method

A network of chemical reactions in noisy fluctuating environments can be addressed by $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \zeta$. Here, $\mathbf{x}(t) = (x_1(t), x_2(t), ..., x_{32}(t))$ represents the vector of concentration or molecular number of species. $\mathbf{F}(\mathbf{x})$ is the vector for the driving force of chemical reaction. In the cell, there are external noise and intrinsic noise, which can be significant to the dynamics of the system [16, 17], so the noise term ζ is added to force item for which Gaussian distribution is assumed, since the force $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ depict only the averaged dynamics of the system. The noise item is satisfied with: $< \zeta_j(\mathbf{x}, t) >= 0$ and $< \zeta_i(\mathbf{x}, t)\zeta_j(\mathbf{x}, t') >= 2D_{ij}\delta_{ij}\delta(t - t')(\delta_{ij} = 1$ for $i = j$, and $\delta_{ij} = 0$ for $i \neq j$). Here $\delta(t)$ is the Dirac delta function, and D is diffusion coefficient matrix. The noise term is associated with the intensity of cellular fluctuations either from the environmental external fluctuations or intrinsic fluctuations. Under large n expansions, the process follows Brownian dynamics.

Following the Brownian dynamical trajectories with multiple different initial conditions by solving the above SDE (stochastic differential equations) iteratively, we can obtain the steady state distribution function $P(x)$ for the state variable $\mathbf{x}$ (relative gene expression value in the gene regulatory network), which is relevant to the potential energy function $U(x)$ as $P_0(\mathbf{x}) = \frac{1}{Z}\ \exp\{-U(\mathbf{x})/D\}$. Here the partition function $Z = \int d\mathbf{x}\ \exp\{-U(\mathbf{x})/D\}$. In this way, we map out the potential energy landscape.

Hamilton-Jacobian (HJ) Framework for Path Integral.

From our path integral formalism, we can evaluate the weights of the kinetic paths. The most probable trajectory can be acquired when the action $S(x)$ is minimized directly. The Lagrangian is written as:

$$L(\mathbf{x}) = \frac{1}{4D}\dot{\mathbf{x}}^2 + V(\mathbf{x}) - \frac{1}{2D}\mathbf{F}(\mathbf{x}) \cdot \dot{\mathbf{x}} \tag{5}$$

and thus the generalized momentum can be written out as: $\mathbf{P}(\mathbf{x}) = \frac{\partial L}{\partial x} = \frac{1}{2D}(\dot{\mathbf{x}} - \mathbf{F}(\mathbf{x}))$. In the kinetic system, the Hamiltonian of the system has the form:

$$H(\mathbf{x}) = -L(\mathbf{x}) + \mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} = E_{eff} \tag{6}$$

According to the above equation, we can obtain $\frac{1}{4D}\dot{\mathbf{x}}^2 - V(\mathbf{x}) = E_{eff}$ and $|\dot{\mathbf{x}}| = \sqrt{4D(E_{eff} + V(\mathbf{x}))}$. After substituting Eq. S2 into the action, we can obtain $S(\mathbf{x}) = \int(\mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} - H(\mathbf{x}))dt$. We can see that the action characterizing the weights of the paths depends on the values of the Hamiltonian. Specific values of the Hamiltonian correspond to specific values of the final time T. For a fixed Hamiltonian, a corresponding optimal path exits when minimizing the action $S(\mathbf{x})$.

From the least action principle, if the Hamiltonian of the system is constant, the variation of the action, for given initial and final coordinates and initial and final time, is zero. Allowing a variation of the final time T and leaving

the initial and the final coordinates fixed, we have $\delta S = -H\delta t$. For a constant Hamiltonian, $\delta S = -E\delta t$. We define $S_0 = \int \mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} dt$, since $S(\mathbf{x}) = \int (\mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} - H(\mathbf{x}))$. We find $\delta S_0 = 0$. Thus, the action S_0 is minimized with respect to all the paths satisfying the constant Hamiltonian and passing through the final point at any instant.

For multidimensional questions, the action depends not only the initial and final coordinates but also on the initial and final time. In the HJ framework, we can transform the formulations into a different representation in x space: $S_0 = S_{HJ}(\mathbf{x}) = \int \sum_i \frac{1}{2D}(\dot{\mathbf{x}}_\mathbf{i} - \mathbf{F}_\mathbf{i}) dx_i = \int \sum_i p_i(\mathbf{x}) dx_i$. Here p_i is the associated momentum. Now the action only depends on the initial and final coordinates. This action can be further simplified and is equivalent to a line integral along a particular one dimensional path l so that $S_{HJ}(\mathbf{x}) = \int \sum_i p_i(\mathbf{x}) dx_i = \int p_l dl$ where $p_l = \sqrt{(E_{eff} + V(\mathbf{x}))/D} - \frac{1}{2D} F_l$. This switch from the time-dependent to the Hamiltonian-dependent HJ description [10, 18, 19]. The dominant path connection given initial and final states is obtained by minimizing the action in the HJ representation $S_{HJ} = \int_{x_i}^{x_f} (\sqrt{(E_{eff} + V(\mathbf{x}))/D} - \frac{1}{2D} F_l) dl$, where dl is an infinitesimal displacement along the path trajectory. E_{eff} is a free parameter that determines the total time elapsed during the transition.

In the current work, for simplification we chose $E_{eff} = -V_{min}(x)$, which is the effective potential by minimizing $V(\mathbf{x})$, and corresponding to the longest kinetic time. Finally, the optimal paths were obtained by minimizing the discrete target function:

$$S_{HJ} = \sum_n^{N-1} (\sqrt{(E_{eff} + V(n))/D} - \frac{1}{2D} F_l(n)) \Delta l_{n,n+1} + \lambda P \tag{7}$$

where

$$\begin{aligned} P &= \sum_i^{N-1} (\Delta l_{i,i+1} - < \Delta l >)^2 \\ (\Delta l)^2_{n,n+1} &= \sum_i (\mathbf{x}_i(n+1) - \mathbf{x}_i(n))^2 \\ F_l(n) &= \sum_i \mathbf{F}_\mathbf{i}(\mathbf{x}(n))(\mathbf{x}_\mathbf{i}(n+1) - \mathbf{x}_i(n))/\Delta l_{n,n+1} \\ V(n) &= \sum_i (\frac{1}{4D} \mathbf{F}^2(\mathbf{x}_\mathbf{i}) + \frac{1}{2} \sum_j \frac{\partial \mathbf{F}_\mathbf{j}(\mathbf{x}_\mathbf{i})}{\partial \mathbf{x}_\mathbf{j}}) \end{aligned} \tag{8}$$

Here, $\Delta l_{n,n+1}$ is the Euclidean measure of the nth elementary path step, and P is a penalty function, which keeps all the length elements close to their average and becomes irrelevant in the continuum limit. The minimization of the discrete HJ effective action was performed by applying a simulated annealing algorithm or the conjugate gradient algorithm. In this study, we chose the discrete steps n as 20, and the diffusion coefficient is chosen as 0.01. In Figure 2 and Figure 6 of main text, we projected the 32-dimensional path to 2-dimensional state space with respect to gene AKT and RB as the biological paths.

Dynamic Paths from Discrete Dynamics

The landscape in Figure 2 only is the 2-dimensional projection of the whole 32 dimensional state space. In order to demonstrate the cell states and the transitions between different cell types in the complete state space, we projected the expression level of the 32 gene variables to binary states, and acquired discretized dynamics results of the network (Figure 4 in main text).

We first used the Langevin dynamics to obtain the stochastic dimensionless trajectories of the 32 dimensional system. Then the trajectory is converted to discrete trajectories by setting the value (0.5) of every variable as the cutoff (gene expression level is between 0 to 1), i.e. the value higher than the cutoff is set to 1 (indicating high expression), while the value lower than the cutoff is set to 0 (indicating low expression). So, we can obtain the discrete trajectories for 32 variables of the system. For a 32 dimension system, there will be 2^{32} states even in discrete case

(every variable has two value, 1 represent high expression, 0 represent low expression), which cannot be handled computationally. So, we chose the major 21 marker genes to present the discrete system, which has $2^{21} = 2097152$ states. For example, the normal state is represented by the binary number 001011010101110000000 (representing expression level from gene 1 to gene 21, 1 for high expression, 0 for low expression), and for the cancer state, it is represented by 100010000101110111100. For the apoptosis state, it is represented by 001101011010000000001. By the statistics for the discrete trajectory, we can obtain the appearing probability separately for 2^{21} different states. To present the results, we set a probability cutoff 0.0002 (only states with higher probability than 0.0005 are chosen, the cutoff is chosen so that the major states can be presented in a figure, not too many or too few states, i.e. we only demonstrate the states and paths with higher probability). Figure 4 in main text shows the discrete landscape and paths for cancer network represented by 247 cell states (nodes) and 334 transition jumps (edges) between the different cell states. The sizes of nodes and edges are separately proportional to the occurrence probability of the corresponding states and paths. Blue nodes represent states which are closer to normal cell states, and red nodes represent states which are closer to cancer states. Especially, we acquired the dominant kinetic paths as the biological paths from path integral formulism, which are shown as green and magenta paths separately for normal to cancer and cancer to normal process.

References

1. Huang S, Guo Y, May G, Enver T (2007) Bifurcation dynamics of cell fate decision lineage-commitment in bipotent progenitor cells. Dev Biol 305: 695-713.
2. Wang E, Lenferink A, O'Connor-McCourt M (2007) Cancer systems biology: exploring cancer-associated genes on cellular networks. Cell Mol Life Sci 64: 1752-62.
3. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646-74.
4. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100: 57-70.
5. Weinberg RA (2007) The Biology of Cancer. New York: Garland Science, 310-520 pp.
6. Van Landeghem S, Bjorne J, C-H W, Hakala K, Pyysalo S, et al. (2013) Large-scale event extraction from literature with multi-level gene normalization. PLoS One 8: e55814.
7. Seoane J, Le HV, Massague J (2002) Myc suppression of the p21(cip1) cdk inhibitor influences the outcome of the p53 response to dna damage. Nature 419: 729-34.
8. Dang CV (1999) c-myc target genes involved in cell growth, apoptosis, and metabolism. Molecular and Cellular Biology 19: 1-11.
9. Li CH, Wang J (2013) Quantifying cell fate decisions for differentiation and reprogramming of a human stem cell network: Landscape and biological paths. PLoS Comput Biol 9: e1003165.
10. Wang J, Zhang K, Xu L, Wang EK (2011) Quantifying the waddington landscape and biological paths for development and differentiation. Proc Natl Acad Sci USA 108: 8257-8262.
11. Sasai M, Wolynes P (2003) Stochastic gene expression as a many-body problem. Proc Natl Acad Sci U S A 100: 2374-2379.
12. Wang J, Li CH, Wang EK (2010) Potential and flux landscapes quantify the stability and robustness of budding yeast cell cycle network. Proc Natl Acad Sci USA 107: 8195-8200.
13. Van Kampen NG (1992) Stochastic processes in Chemistry and Physics. Amsterdam: North Holland, 1 edition, 120-127 pp.

14. Hu G (1994) Stochastic Forces and Nonlinear Systems. Shanghai: Shanghai Scientific and Technological Education Press, 68-74 pp.

15. Wang J, Xu L, Wang EK (2008) Potential landscape and flux framework of non-equilibrium networks: Robustness, dissipation and coherence of biochemical oscillations. Proc Natl Acad Sci USA 105: 12271-12276.

16. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci U S A 99: 12795-12800.

17. Kaern M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 6: 451-64.

18. Faccioli P, Sega M, Pederiva F, Orland H (2006) Dominant pathways in protein folding. Phys Rev Lett 97: 1-4.

19. Wang J, Zhang K, Wang EK (2010) Kinetic paths, time scale, and underlying landscapes: A path integral framework to study global natures of nonequilibrium systems and networks. J Chem Phys 133: 125103:1-13.

20. Vogelstein B, Kinzler KW (2004) Cancer genes and the pathways they control. Nature Medicine 10: 789-799.

21. Weinberg RA (2009) The Biology of Cancer. Beijing: Science Press, 1st edition, 89-337 pp.

22. Grard C, Goldbeter A (2009) Temporal self-organization of the cyclin/cdk network driving the mammalian cell cycle. Proc Natl Acad Sci USA 106: 21643-21648.

23. Lin WC, Lin FT, Nevins JR (2001) Selective induction of e2f1 in response to dna damage, mediated by atm-dependent phosphorylation. Genes Dev 15: 1833-44.

24. Khosravi R, Maya R, Gottlieb T, Oren M, Shiloh Y, et al. (1999) Rapid atm-dependent phosphorylation of mdm2 precedes p53 accumulation in response to dna damage. Proc Natl Acad Sci USA 96: 14973-7.

25. Khanna KK, Keating KE, Kozlov S, Scott S, Gatei M, et al. (1998) Atm associates with and phosphorylates p53: mapping the region of interaction. Nat Genet 20: 398-400.

26. Seoane J, Le H, Massagu J (2002) Myc suppression of the p21(cip1) cdk inhibitor influences the outcome of the p53 response to dna damage. Nature 419: 729-34.

27. Sherr C, Roberts J (1999) Cdk inhibitors: positive and negative regulators of g1-phase progression. Genes Dev 13: 1501-12.

28. Budihardjo I, Oliver H, Lutter M, Luo X, Wang X (1999) Biochemical pathways of caspase activation during apoptosis. Annu Rev Cell Dev Biol 15: 269-90.

29. Soucie E, Annis M, Sedivy J, Filmus J, Leber B, et al. (2001) Myc potentiates apoptosis by stimulating bax activity at the mitochondria. Mol Cell Biol 21: 4725-36.

30. Semenza G (2000) Hif-1: Using two hands to flip the angiogenic switch. Cancer and Metastasis Reviews 19: 59C65.

31. Sullu Y, Gun S, Atmaca S, Karagoz F, Kandemir B (2010) Poor prognostic clinicopathologic features correlate with vegf expression but not with pten expression in squamous cell carcinoma of the larynx. Diagn Pathol 14: 5:35.

32. Stambolic V, MacPherson D, Sas D, Lin Y, Snow B, et al. (2001) Regulation of pten transcription by p53. Mol Cell 8: 317-25.

33. Lin H, Hu Y, Lee D, Chang C (2004) Regulation of androgen receptor signaling by pten (phosphatase and tensin homolog deleted on chromosome 10) tumor suppressor through distinct mechanisms in prostate cancer cells. Mol Endocrinol 18: 2409-23.

34. Wu Z, Conaway M, Gioeli D, Weber M, Theodorescu D (2006) Conditional expression of pten alters the androgen responsiveness of prostate cancer cells. Prostate 66: 1114-23.

35. Vasudevan K, Gurumurthy S, Rangnekar V (2004) Suppression of pten expression by nf-kappa b prevents apoptosis. Mol Cell Biol 24: 1007-21.

36. Leung K, Ng H, Tang M, Wong C, Wong R, et al. (2011) Ginsenoside-rg1 mediates a hypoxia-independent upregulation of hypoxia-inducible factor-1 to promote angiogenesis. Angiogenesis 14: 515-22.

37. Lin H, Yeh S, Kang H, Chang C (2001) Akt suppresses androgen-induced apoptosis by phosphorylating and inhibiting androgen receptor. Proc Natl Acad Sci USA 98: 7200-5.

38. Lu S, Jenster G, Epner D (2000) Androgen induction of cyclin-dependent kinase inhibitor p21 gene: role of androgen receptor and transcription factor sp1 complex. Mol Endocrinol 14: 753-60.

39. Yeh S, Miyamoto H, Nishimura K, Kang H, Ludlow J, et al. (1998) Retinoblastoma, a tumor suppressor, is a coactivator for the androgen receptor in human prostate cancer du145 cells. Biochem Biophys Res Commun 248: 361-7.

40. Gaughan L, Logan I, Neal D, Robson C (2005) Regulation of androgen receptor and histone deacetylase 1 by mdm2-mediated ubiquitylation. Nucleic Acids Res 33: 13-26.

41. Luo Y, He D, Ning L, Shen S, Li L, et al. (2006) Over-expression of hypoxia-inducible factor-1alpha increases the invasive potency of lncap cells in vitro. BJU Int 98: 1315-9.

42. Choi M, Shi J, Jung SH, Chen X, Cho KH (2012) Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to dna damage. Sci Signal 5: ra83.

43. Damiano V, Caputo R, Garofalo S, Bianco R, Rosa R, et al. (2007) Tlr9 agonist acts by different mechanisms synergizing with bevacizumab in sensitive and cetuximab-resistant colon cancer xenografts. Proc Natl Acad Sci USA 104: 12468-73.

44. Jaganathan S, Yue P, Paladino D, Bogdanovic J, Huo Q, et al. (2011) A functional nuclear epidermal growth factor receptor, src and stat3 heteromeric complex in pancreatic cancer cells. PLoS One 6: e19605.

45. Javle M, Hsueh C (2009) Updates in gastrointestinal oncology - insights from the 2008 44th annual meeting of the american society of clinical oncology. J Hematol Oncol 2: 9.

46. Grunt T, Wagner R, Grusch M, Berger W, Singer C, et al. (2009) Interaction between fatty acid synthase- and erbb-systems in ovarian cancer cells. Biochem Biophys Res Commun 385: 454-9.

47. Lee I, Lin C, Wu Y, Yang C (2010) Tnf-alpha induces matrix metalloproteinase-9 expression in a549 cells: role of tnfr1/traf2/pkcalpha-dependent signaling pathways. J Cell Physiol 224: 454-64.

48. McCulloch R, Walker C, Chakera A, Jazayeri J, Leedman P (1998) Regulation of egf-receptor expression by egf and tgf alpha in epidermoid cancer cells is cell type-specific. Int J Biochem Cell Biol 30: 1265-78.

49. Juuti-Uusitalo K, Kaukinen K, Maki M, Tuimala J, Kainulainen H (2006) Gene expression in tgfbeta-induced epithelial cell differentiation in a three-dimensional intestinal epithelial cell differentiation model. BMC Genomics 7: 279.

50. Zhou L, Yang H (2011) The von hippel-lindau tumor suppressor protein promotes c-cbl-independent poly-ubiquitylation and degradation of the activated egfr. PLoS One 6: e23936.

51. Hayashi H, Nakagami H, Takami Y, Koriyama H, Mori M, et al. (2009) Fhl-2 suppresses vegf-induced phosphatidylinositol 3-kinase/akt activation via interaction with sphingosine kinase-1. Arterioscler Thromb Vasc Biol 29: 909-14.

52. Eppenberger M, Zlobec I, Baumhoer D, Terracciano L, Lugli A (2010) Role of the vegf ligand to receptor ratio in the progression of mismatch repair-proficient colorectal cancer. BMC Cancer 10: 93.

53. Bos R, van Diest P, van der Groep P, Shvarts A, Greijer A, et al. (2004) Expression of hypoxia-inducible factor-1alpha and cell cycle proteins in invasive breast cancer are estrogen receptor related. Breast Cancer Res 6: R450-9.

54. Berns E, Klijn J, Look M, Grebenchtchikov N, R V (2003) Combined vascular endothelial growth factor and tp53 status predicts poor response to tamoxifen therapy in estrogen receptor-positive advanced breast cancer. Clin Cancer Res 9: 1253-8.

55. Mezquita P, Parghi S, Brandvold K, Ruddell A (2005) Myc regulates vegf production in b cells by stimulating initiation of vegf mrna translation. Oncogene 24: 889-901.

56. Desiderio M, Pogliaghi G, Dansi P (1998) Hepatocyte growth factor-induced expression of ornithine decarboxylase, c-met, and c-myc is differently affected by protein kinase inhibitors in human hepatoma cells hepg2. Exp Cell Res 242: 401-9.

57. Franke R, Mller M, Wundrack N, Gilles E, Naumann M (2008) Host-pathogen systems biology: logical modelling of hepatocyte growth factor and helicobacter pylori induced c-met signal transduction. BMC Syst Biol 14: 2:4.

58. Lee K, Kim S, Kim J (2009) Reactive oxygen species regulate urokinase plasminogen activator expression and cell invasion via mitogen-activated protein kinase pathways after treatment with hepatocyte growth factor in stomach cancer cells. J Exp Clin Cancer Res 4: 28:73.

59. Xu K, FS Y (2007) Cross talk between c-met and epidermal growth factor receptor during retinal pigment epithelial wound healing. Invest Ophthalmol Vis Sci 48: 2242-8.

60. Tacchini L, Dansi P, Matteucci E, MA D (2001) Hepatocyte growth factor signalling stimulates hypoxia inducible factor-1 (hif-1) activity in hepg2 hepatoma cells. Carcinogenesis 22: 1363-71.

61. Nakagami H, Morishita R, Yamamoto K, Taniyama Y (2002) Hepatocyte growth factor prevents endothelial cell death through inhibition of bax translocation from cytosol to mitochondrial membrane. Diabetes 51: 2604-11.

62. Worden B, Yang X, Lee T, Bagain L, Yeh N, et al. (2005) Hepatocyte growth factor/scatter factor differentially regulates expression of proangiogenic factors through egr-1 in head and neck squamous cell carcinoma. Cancer Res 65: 7071-80.

63. Xie L, Bian L, Z L, Li Y, Li Z, et al. (2010) Altered expression of e-cadherin by hepatocyte growth factor and effect on the prognosis of nasopharyngeal carcinoma. Ann Surg Oncol 17: 1927-36.

64. Zisa D, Shabbir A, Mastri M, Suzuki G, Lee T (2009) Intramuscular vegf repairs the failing heart: role of host-derived growth factors and mobilization of progenitor cells. Am J Physiol Regul Integr Comp Physiol 297: R1503-15.

65. Yang Y, Ge P, Jiang L, Li F, QY Z (2011) Modulation of growth and angiogenic potential of oral squamous carcinoma cells in vitro using salvianolic acid b. BMC Complementary and Alternative Medicine 11: 54.

66. Takahashi H, Tsuji H, Hashimoto Y, Ishida-Yamamoto A, Iizuka H (2009) Cell proliferation and cytokine induction by tnf-alpha of psoriatic keratinocytes are not different from normal keratinocytes in vitro. Indian J Dermatol 54: 237-9.

67. Cheong R, Hoffmann A, Levchenko A (2008) Understanding nf-kappab signaling via mathematical modeling. Mol Syst Biol 4: 192.

68. Zhou L, Zheng D, Wang M, Cong Y (2009) Telomerase reverse transcriptase activates the expression of vascular endothelial growth factor independent of telomerase activity. Biochem Biophys Res Commun 386: 739-43.

69. Mamchaoui K, Trollet C, Bigot A, Negroni E, Chaouch S, et al. (2011) Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. Skelet Muscle 1: 34.

70. Sanduja S, Kaza V, Dixon D (2009) The mrna decay factor tristetraprolin (ttp) induces senescence in human papillomavirus-transformed cervical cancer cells by targeting e6-ap ubiquitin ligase. Aging (Albany NY) 1: 803-17.

71. Yu R, Chen E, Kong R, Ng P, Mok H, et al. (2006) Hypoxia induces telomerase reverse transcriptase (tert) gene expression in non-tumor fish tissues in vivo: the marine medaka (oryzias melastigma) model. BMC Mol Biol 7: 27.

72. Lacerte A, Korah J, Roy M, Yang X, Lemay S, et al. (2008) Transforming growth factor-beta inhibits telomerase through smad3 and e2f transcription factors. Cell Signal 20: 50-9.

73. Sermeus A, Cosse J, Crespin M, Mainfroid V, Michiels C (2008) Hypoxia induces protection against etoposide-induced apoptosis: molecular profiling of changes in gene expression and transcription factor activity. Mol Cancer 7: 27.

74. Treins C, Giorgetti-Peraldi S, Murdaca J, Semenza G, Van Obberghen E (2002) Insulin stimulates hypoxia-inducible factor 1 through a phosphatidylinositol 3-kinase/target of rapamycin-dependent signaling pathway. J Biol Chem 277: 27975-81.

75. Gorospe M, Tominaga K, Wu X, Ivan M (2011) Post-transcriptional control of the hypoxic response by rna-binding proteins and micrornas. Front Mol Neurosci 4: 7.

76. Golinska M, Troy H, Chung Y, McSheehy P, Mayr M, et al. (2011) Adaptation to hif-1 deficiency by upregulation of the amp/atp ratio and phosphofructokinase activation in hepatomas. BMC Cancer 11: 198.

77. Albina J, Mastrofrancesco B, Vessella J, Louis C, Henry WJ, et al. (2001) Hif-1 expression in healing wounds: Hif-1alpha induction in primary inflammatory cells by tnf-alpha. Am J Physiol Cell Physiol 281: C1971-7.

78. Cheng Z, Sun B, Wang S, Gao Y, Zhang Y (2011) Nuclear factor-b-dependent epithelial to mesenchymal transition induced by hif-1 activation in pancreatic cancer cells under hypoxic conditions. PLoS One 6: e23752.

Table S1. 32 gene names in the cancer network and the corresponding function description.

Gene	Functions
ATM=1	coordinate DNA repair by activating enzymes that fix the broken strands
P53=2	Tumor suppressor gene
P21=3	Tumor suppressor gene
PTEN=4	Tumor suppressor gene
CDH1=5	Tumor suppressor gene
RB=6	Tumor suppressor gene, control G1 checkpoint
ARF=7	Tumor suppressor gene, response to sustained mitogenic stimulation
AR=8	androgen receptor, oncogene
MYC=9	oncogene, promote proliferation
AKT=10	oncogene, promote proliferation, suppress apoptosis
EGFR=11	oncogene, release proliferative signal
VEGF=12	oncogene, promote angiogenesis
HGF=13	oncogene, promote metastasis
HIF1=14	oncogene, promote glycolysis
hTERT=15	oncogene, promote the product of enzyme telomerase
MDM2=16	oncogene, repress P53
CDK2=17	Cyclin-dependent kinases, cell cycle gene
CDK4=18	Cyclin-dependent kinases, cell cycle gene
CDK1=19	Cyclin-dependent kinases, cell cycle gene
E2F1=20	activate cyclins and ckds, promote proliferation
Caspase=21	apoptosis gene
BAX=22	Apoptosis regulator, promote apoptosis by binding to and antagonizing the Bcl2 protein
BAD=23	pro-apoptotic member of the Bcl-2 gene family which is involved in initiating apoptosis
BCL2=24	oncogene, inhibit apoptosis
NFkB=25	oncogene, repress tumor repressor PTEN
RAS=26	promote cell growth and division
TGFα=27	growth factor, activate cell proliferation, differentiation and development
TNFα=28	regulate immune cells, induce apoptosis
TGFβ=29	controls proliferation, differentiation
Wee1=30	repress cell cycle
MdmX=31	repress P53
Wip1=32	negative regulators of cell stress response pathways

Table S2. Connections and corresponding evidences in the network (32 nodes, 111 connections including 66 activation interactions and 45 repression interactions). For the column of interaction type, 1 represents activation, and -1 represents repression.

Source Node	Target Node	Interaction Type	References
RB	E2F1	-1	[20]
CDK2	RB	-1	[21]
Wee1	CDK1	-1	[22]
NFKB	MYC	1	[21]
ATM	E2F1	1	[23]
E2F1	ARF	1	[21]
ARF	MDM2	-1	[21]
ATM	MDM2	-1	[24]
ATM	P53	1	[25]
P53	P21	1	[20]
P53	PBCL2	-1	[20]
P53	BAX	1	[20]
BCL2	Caspase	-1	[20]
P21	CDK2	-1	[20]
MDM2	P53	-1	[21]
MYC	CDK2	1	[21]
MYC	CDK4	1	[21]
CDK4	RB	-1	[21]
AKT	MDM2	1	[21]
AKT	P21	-1	[21]
AKT	BAD	-1	[20]
PTEN	AKT	-1	[21]
MYC	P21	-1	[21, 26]
E2F1	CDK2	1	[22, 27]
CDK2	E2F1	-1	[22]
MYC	P53	1	[21]
BAX	Caspase	1	[20, 28]
BAD	Caspase	1	[20, 28]
MYC	BAX	1	[21, 29]
AKT	NFKB	1	[21]
AKT	HIF1	1	[30]
AKT	VEGF	1	[31]
P53	PTEN	1	[32]
PTEN	NFKB	-1	[33, 34]
NFKB	PTEN	-1	[35]
HIF1	VEGF	1	[36]
AKT	AR	-1	[37]
PTEN	AR	-1	[33, 34]
AR	P21	1	[38]
RB	AR	1	[39]
MDM2	AR	-1	[40]
HIF1	CDH1	-1	[41]
CDK2	CDH1	-1	[22]

Continued...

Source Node	Target Node	Interaction Type	References
P53	MDM2	1	[21]
AKT	CDK2	1	[21]
NFKB	BCL2	1	[21]
NFKB	CDK4	1	[21]
RAS	MDM2	1	[21]
RAS	CDK4	1	[21]
TGFβ	P21	1	[21]
TGFβ	MYC	-1	[21]
TGFα	RAS	1	[21]
RAS	MYC	1	[21]
BCL2	BAX	-1	[42]
AKT	Caspase	-1	[21]
Caspase	BCL2	-1	[28, 42]
MdmX	P53	-1	[42]
P53	Wip1	1	[42]
Wip1	ATM	-1	[42]
Wip1	MdmX	1	[42]
Wip1	ARF	-1	[42]
ARF	MdmX	-1	[42]
Wip1	MDM2	1	[42]
MDM2	P21	-1	[42]
RB	MDM2	-1	[42]
ATM	MdmX	-1	[42]
AKT	MdmX	1	[42]
MDM2	MdmX	-1	[42]
RB	CDK4	-1	[22]
RB	CDK2	-1	[22]
P21	CDK1	-1	[22]
CDH1	CDK1	-1	[22]
CDK1	CDH1	-1	[22]
CDK1	Wee1	1	[22]
EGFR	VEGF	1	[43]
EGFR	MYC	1	[44]
EGFR	AKT	1	[45]
EGFR	RAS	1	[45]
AKT	EGFR	1	[46]
TNFα	EGFR	1	[47]
TGFα	EGFR	1	[48]
TGFβ	EGFR	-1	[49]
HIF1	EGFR	1	[50]
VEGF	AKT	1	[51]
VEGF	RAS	1	[52]
VEGF	BCL2	1	[53]
P53	VEGF	-1	[54]
MYC	VEGF	1	[55]
HGF	MYC	1	[56]
HGF	NFKB	1	[57]

Continued...

Source Node	Target Node	Interaction Type	References
HGF	AKT	1	[58]
HGF	EGFR	1	[59]
HGF	HIF1	1	[60]
HGF	BCL2	1	[61]
HGF	VEGF	1	[62]
HGF	CDH1	1	[63]
VEGF	HGF	1	[64]
HIF1	HGF	1	[65]
TNFα	HGF	1	[66]
TNFα	TGFα	1	[67]
hTERT	VEGF	1	[68]
hTERT	ARF	-1	[69]
P53	hTERT	-1	[70]
HIF1	hTERT	1	[71]
TGFβ	hTERT	-1	[72]
HIF1	P53	1	[73]
HIF1	VEGF	1	[30, 74]
HIF1	TGFβ	1	[75]
P53	HIF1	-1	[30]
RAS	HIF1	1	[76]
TNFα	HIF1	1	[77]
NFKB	HIF1	1	[78]

Table S3. Steady state relative gene expression level of three attractor state ($a = 0.5, b = 0.5, k = 1$) respectively representing apoptosis, normal and cancer states. Only the key 21 genes are shown. The gene expression level is in the range from 0 to 1.

Gene Numer	Gene	Apoptosis	Normal	Cancer
1	ATM	0.36340	0.42696	0.53399
2	P53	0.48017	0.46429	0.44160
3	P21	0.76646	0.50843	0.46523
4	PTEN	0.75901	0.29755	0.26638
5	CDH1	0.41957	0.63140	0.54350
6	RB	0.99767	0.71799	0.19438
7	ARF	0.25245	0.26346	0.36517
8	AR	0.68444	0.57478	0.11816
9	MYC	0.65714	0.47597	0.47606
10	AKT	0.28785	0.80542	0.85373
11	EGFR	0.52032	0.46055	0.46668
12	VEGF	0.39121	0.62358	0.67241
13	HGF	0.12002	0.59033	0.65474
14	HIF1	0.27710	0.66288	0.70476
15	hTERT	0.35362	0.46797	0.47319
16	MDM2	0.26235	0.48755	0.70843
17	CDK2	0.12482	0.45654	0.79604
18	CDK4	0.12593	0.42539	0.86807
19	CDK1	0.43442	0.46666	0.54233
20	E2F1	0.13813	0.25876	0.40906
21	Caspase	0.89277	0.06891	0.05561

Table S4. Repression link names in the sensitivity analysis and the corresponding regulations they represent. The order numbers for causal and target genes are shown, which are corresponding to the gene name in Table S1.

Link Name	Causal Genes	Target Genes
R1	17	6
R2	18	6
R3	15	7
R4	32	7
R5	4	8
R6	10	8
R7	16	8
R8	29	9
R9	4	10
R10	2	12

Table S5. Activation link names in sensitivity analysis and the corresponding regulations they represent. The order numbers for causal and target genes are shown, which are corresponding to the gene name in Table S1.

Link Name	Causal Genes	Target Genes
A1	20	7
A2	6	8
A3	11	9
A4	13	9
A5	25	9
A6	26	9
A7	11	10
A8	12	10
A9	13	10
A10	9	12
A11	10	12
A12	11	12
A13	13	12
A14	14	12
A15	15	12

Table S6. Interaction type matrix Mi. The element $Mi(j,i)$ (the jth row and the ith column) represents the interaction type from node j to node i. 1 represents activation, -1 represents repression, and 0 represents no interaction.

Gene	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	−1	0
2	0	0	3	1	0	0	0	0	0	0	0	−1	0	−1	−1	1	0	0	0	0	0	1	0	−1	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	−1	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	−1	−1	−1	0	−1	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0
8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	1	−1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
10	0	0	−1	0	0	0	0	−1	0	0	1	1	0	1	0	1	1	0	0	0	−1	0	−1	0	1	0	0	0	0	0	1	0
11	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
13	0	0	0	0	1	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
14	0	1	0	0	−1	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
15	0	0	0	0	0	0	−1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	−1	−1	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0
17	0	0	0	0	−1	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
20	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	−1	0	0	0	0	0	0	0	0	0	0
25	0	0	0	−1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
29	0	0	1	0	0	0	0	0	−1	0	−1	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	−1	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Table S7. Interaction strength Matrix Ms. The value of interaction strength in jth row ($Ms(j)$) represent the interaction strength of gene j to any other genes.

Gene	interaction strength
1	1.028
2	1.281
3	0.598
4	0.827
5	1.020
6	2.244
7	1.037
8	1.507
9	0.203
10	1.299
11	1.337
12	0.921
13	2.044
14	0.678
15	0.924
16	0.467
17	0.947
18	0.242
19	0.941
20	0.480
21	0.666
22	0.232
23	1.156
24	0.951
25	0.849
26	0.698
27	1.288
28	0.966
29	1.419
30	0.401
31	0.293
32	1.479

Table S8. Transition path (produced by path integral) from normal state to cancer state characterized by high/low expression level of 21 marker genes in Figure 3 of main text. The row of gene ID represent the corresponding genes in Table S1. From the row normal to row cancer, every line represents a cellular states. Normal represent the normal states, and cancer represents cancer states. 1 denotes high expression level, and 0 represents low expression level.

gene ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Normal	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0
	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	1	0	0	0	0	0
	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	1	1	0	0	0	0
	0	0	0	0	1	1	0	1	0	1	0	1	1	1	0	1	1	1	0	0	0
	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
Cancer	1	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0

Table S9. Transition path (produced by path integral) from cancer state to normal state characterized by high/low expression level of 21 marker genes in Figure 3 of main text. The row of gene ID represent the corresponding genes in Table S1. From the row normal to row cancer, every line represents a cellular states. Normal represent the normal states, and cancer represents cancer states. 1 denotes high expression level, and 0 represents low expression level.

gene ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Cancer	1	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	1	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	1	0	0	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	1	1	1	0	0	0
	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0
	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	1	0	0	0	0	0
Normal	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0

Table S10. Transition path (produced by path integral) from normal state to apoptosis state characterized by high/low expression level of 21 marker genes in Figure 3 of main text. The row of gene ID represent the corresponding genes in Table S1. From the row normal to row cancer, every line represents a cellular states. Normal represent the normal state, and Apoptosis represents apoptosis states. 1 denotes high expression level, and 0 represents low expression level.

gene ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Normal	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0
	0	1	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	1	0	0
	0	1	1	0	1	1	0	1	1	1	0	1	1	1	0	0	0	0	1	0	0
	0	1	1	0	1	1	0	1	1	1	0	1	0	1	0	0	0	0	1	0	0
	0	1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1	0	0
	0	1	1	1	1	1	0	1	1	0	0	0	0	1	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	0	0	0	1	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
	0	1	1	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1
Apoptosis	0	0	1	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1

Table S11. Transition path (produced by path integral) from cancer state to apoptosis state characterized by high/low expression level of 21 marker genes in Figure 3 of main text. The row of gene ID represent the corresponding genes in Table S1. From the row normal to row cancer, every line represents a cellular states. Cancer represent the cancer state, and Apoptosis represents apoptosis states. 1 denotes high expression level, and 0 represents low expression level.

gene ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Cancer	1	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	1	0	0	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
	0	1	1	0	1	0	0	0	1	1	0	1	1	1	0	1	1	1	1	0	0
	0	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	1	1	1	0	0
	0	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	1	1	0	0
	0	1	1	0	1	1	0	0	1	0	0	0	1	1	0	1	0	1	1	0	0
	0	1	1	0	0	1	0	0	1	0	0	0	1	1	0	1	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	0	0	0	1	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	1	0	1
	0	1	1	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1
Apoptosis	0	0	1	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1