Quantifying the Waddington landscape and biological paths for development and differentiation

Jin Wang[a,b,1], Kun Zhang[a], Li Xu[a], and Erkang Wang[a,1]

[a]State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022, People's Republic of China; and [b]Department of Chemistry, Physics, and Applied Mathematics, State University of New York, Stony Brook, NY 11794-3400

Edited* by José N. Onuchic, University of California, La Jolla, CA, and approved March 18, 2011 (received for review November 11, 2010)

We developed a theoretical framework to prove the existence and quantify the Waddington landscape as well as chreode-biological paths for development and differentiation. The cells can have states with the higher probability ones giving the different cell types. Different cell types correspond to different basins of attractions of the probability landscape. We study how the cells develop from undifferentiated cells to differentiated cells from landscape perspectives. We quantified the Waddington landscape through construction of underlying probability landscape for cell development. We show the developmental process proceeds as moving from undifferentiated to the differentiated basins of attractions. The barrier height of the basins of attractions correlates with the escape time that determines the stability of cell types. We show that the developmental process can be quantitatively described and uncovered by the biological paths on the quantified Waddington landscape from undifferentiated to the differentiated cells. We found the dynamics of the developmental process is controlled by a combination of the gradient and curl force on the landscape. The biological paths often do not follow the steepest descent path on the landscape. The landscape framework also quantifies the possibility of reverse differentiation process such as cell reprogramming from differentiated cells back to the original stem cell. We show that the biological path of reverse differentiation is irreversible and different from the one for differentiation process. We found that the developmental process described by the underlying landscape and the associated biological paths is relatively stable and robust against the influences of environmental perturbations.

Cells are dynamical entities: They change their phenotype during development in an almost discontinuous manner, giving rise to discrete developmental stages (such progenitor and differentiated states) as well as discrete lineages and terminally differentiated types. This pattern of cell dynamics during development was already noted by C. Waddington in the 1940s, leading to his by now famous metaphor of the epigenetic landscape (1) (see Fig. 1) In this iconic picture a marble rolls down a surface (landscape), staying in valleys and seeking the lowest point. At watersheds, the valleys branch so that the marble takes one of two available paths. In Waddington's picture, the ball represents a developing cell in an embryo and the landscape epitomizes some more abstract set of constraints, thus clearly heralding the notions of stability and instability in the modern sense of dynamics (1). Indeed, it has recently become clear that Waddington's epigenetic landscape in principle represents the dynamics of a system of gene regulatory interactions that impose constraints to and drive cell development (2, 3), giving a metaphor for the qualitative understanding of developmental processes of cells. However, a detailed quantitative examination of how the dynamics of a gene regulatory circuit that governs binary cell fate decisions produces a generalized potential landscape that may recapitulate the epigenetic landscape has not been presented. In other words, it is not very clear on exactly what the Waddington landscape represents, how the qualitative picture of landscape can be quantified, and how the connection to the experiments can be made.

Here we develop a theoretical framework to show the existence of such a landscape as the formal representation of the dynamics of a gene circuit and quantify its detailed topography. We define Waddington's chreodes—the biological paths (or trajectories) of development. Herein, the entity that changes, and hence is embodied by the marble, is the gene expression pattern of a cell that reflects the network state of the genes in a particular network. The state $S(t) = (x_1, x_2, x_i, .x_N)$ thus reflects the vector consisting of the expression values (cellular concentration and activity of the products of gene i), x_1, x_2, in a gene regulatory network of N genes at a given time t. Because of the regulatory interactions not all states can be realized with equal ease and in a system with fluctuations, probabilities can be assigned to the states S. They can have a higher or lower probability of appearance that translates inversely into the elevation (potential) of the landscape (4, 5). The states with locally highest probability (lowest potential) represent attractor states of the gene regulatory network as a dynamical system, surrounded by their basins of attraction. Attractor states have been proposed to represent cell types (6).

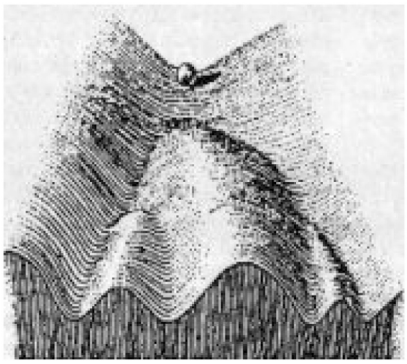

Fig. 1. The original artistical picture of Waddington epigenetic landscape (1).

The landscape metaphor has recently seen renewed interest with the arrival of cell reprogramming (5, 7, 8). If cell types are attractors, then reprogramming represents the transition between attractors. Thus the height of hills between two valleys, or the barrier heights between the attractors, can be correlated to the escape time from the basins, offering a quantitative measure for the nonlocal relative stability of the attractors or cell types. The landscape topography thus has a quantitative meaning.

Here we construct such a probability landscape quantitatively based on the underlying gene regulatory circuit for the biological process of a binary cell fate decision of a multipotent progenitor

Author contributions: J.W. and E.W. designed research; J.W., K.Z., and L.X. performed research; J.W. and E.W. contributed new reagents/analytic tools; J.W., K.Z., L.X., and E.W. analyzed data; and J.W., K.Z., L.X., and E.W. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence may be addressed. E-mail: jin.wang.1@stonybrook.edu or ekwang@ciac.jl.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1017017108/-/DCSupplemental.

cell. Fate decision and subsequent lineage commitment to a particular fate is described by a path in this landscape from the attractor of the progenitor cell to that of either one of two differentiated cell types. These paths can be identified and quantified through counting the weights of all the possible paths and selecting the optimal paths with largest weights. We will do so with a path integral approach (9–11). The advantage of the path integral approach is that it addresses the fundamental issues of biological paths directly. Furthermore, the weights associated with the biological paths can also be determined. By varying the configurational states, we can also explore the corresponding probability landscape topography and investigate the association of escape time with the barrier height for global stability analysis of the cell types. Thus, landscape, paths, kinetics, and stability that are all of practical interest for determining reprogramming strategies, can be evaluated in one model.

We will study an important example of cell developmental circuit (Fig. 2) (12) composed of a pair of self-activating and mutually inhibiting genes as a model, a gene regulatory motif that has been found in various tissues where a pluri/multipotent stem cells has to undergo a binary cell fate decision (8, 12). For instance, in the multipotent common myeloic progenitor cell (CMP), which faces the binary cell fate decision between the myeloid and the erythroid fate, the fate determining transcription factors (TF), PU.1, and GATA1, which promote the myeloid or the erythroid fates, respectively, form such a circuit. The relative expression levels x_1 (PU.1) and x_2 (GATA1) of these two reciprocal TFs can tilt the decision toward either lineage (7, 12).

We quantitatively constructed the probability landscape for the GATA1-PU.1 type of circuit, showing its similarity and difference with Waddington's epigenetic landscape, and quantified the developmental paths. We show that the cell development process proceeds from the basin of attraction of the undifferentiated state to that of the differentiated attractors. The heights of the barriers separating the basins of attractions correlate with the escape time that reflect the stability of cell types. Specifically, we demonstrate that the developmental process can be quantitatively described by the biological path on a quantified Waddington landscape and is governed by a combination of a gradient and a curl force on the landscape. The paths of differentiation do not follow the steepest descent path on the landscape (gradient) and the paths of the reverse differentiation process is different from the ones of the differentiation process indicating irreversibility.

Results and Discussions

Developmental Network and Cell Fate Decision Module. A gene regulatory circuit that governs binary cell fate decision module is shown in Fig. 2*A*. It consists of mutual regulation of two opposing fate determining master TF X_1 and X_2. The module has been shown to control developmental cell fate decision and commitment in several instances of multipotent stem or progenitor cells that faces a binary fate decision, (i.e., GATA1 and PU.1) (8, 12). X_1 and X_2 are coexpressed in the multipotent undecided cell and committed to either one of the two alternative lineages is associated with either one factor dominating over the others, leads to expression patterns in a mutually exclusive manner (13). Importantly, in many cases the genes X_1 and X_2 also self-activate (positive autoregulate) themselves (Fig. 2*A*). The circuit can be described by the following minimal system equations (12): $\frac{dx_1}{dt} = \frac{a_1 x_1^n}{S^n + x_1^n} + \frac{b_1 S^n}{S^n + x_2^n} - k_1 x_1 = F_1(x_1, x_2)$ and $\frac{dx_2}{dt} = \frac{a_2 x_2^n}{S^n + x_2^n} + \frac{b_2 S^n}{S^n + x_1^n} - k_2 x_2 = F_2(x_1, x_2)$. In vector form, $d\mathbf{x}/dt = \mathbf{F}(\mathbf{x}) = [F_1(x_1, x_2), F_2(x_1, x_2)]$. x_1 and x_2 represent the cellular expression or activation levels of the two lineage determining transcription factors X_1 and X_2, and a_1, a_2, b_1, b_2, k_1, k_2 are positive parameters that denote the strength of the following interactions or processes: The first expression represents, in the common formalization (12), a self-activation (of strength a_1, a_2) that obeys a sigmoidal transfer function, the second term represents mutual inhibition, given basal expression, (of strength b_1, b_2); and the last term is the first-order inactivation (degradation) of either factor with the rate k_1, k_2. For our purpose it suffice to consider the symmetric situation $a = a_1 = a_2$; $b = b_1 = b_2$; $k = k_1 = k_2$. The mutual regulations and self-activations are often separated and do not have to be simultaneously changed in a synchronized way in the experiments, therefore the underlying interactions described above follow an OR rather than AND logic. (Details in *SI Text*).

In this example, the self-activation strengths change due to the regulations on the transcription factors by other regulators such as Klf4 (14) during the developmental process. As we will see that parameter a of self-activation strength gives a sense of measure and direction of the development.

Landscape at a Given Stage of Development: Cross-Section of Waddington Landscape. Based on the above dynamics of the cell fate decision, we quantitatively mapped out landscape of the development at different stages in Fig. 2*B*. In the resulting landscape representation, the vertical axis represents the potential landscape U (inversely related to the probability landscape P: $U = -\ln P$) (5). The horizontal axis represents the expression level of either x_1 or x_2 (due to the symmetry between x_1 and x_2 as well as for easy visualization we only show one of them).

We see that the decision making circuit often has three basins of attractions. The center attractor C represents the undifferentiated undecided multipotent stem cell state with balanced expression of the two opposing fate determining transcription factors (13). The side attractors A and B represent the the differentiated states with mutually excluding expressions of x_1 and x_2 (12). In this view the three attractors become evident as potential wells.

When the self-activation is strong (characterized by large value of a), the central basin of attraction is deep and therefore the cell

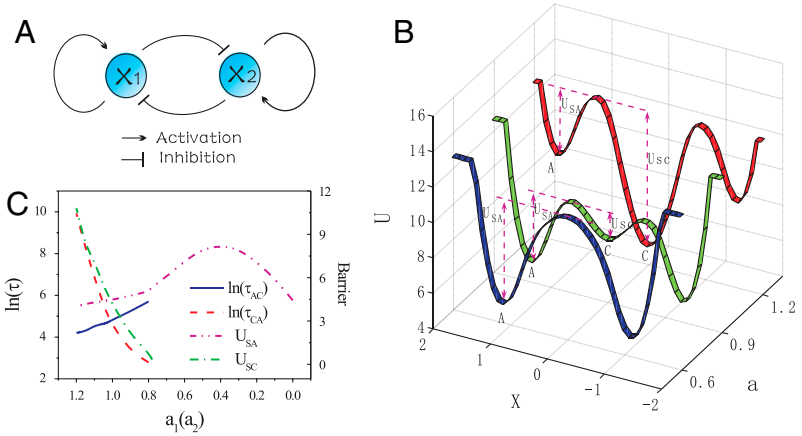

Fig. 2. (*A*) The illustration of cell fate decision module. (*B*) The landscape of the development at different stages or different paramenters. (*C*) Escape time under certain fluctuations versus the parameter a and the barrier height from the central undifferentiated state to differentiated state versus the parameter a. ($a = a_1 = a_2$, $b_1 = b_2 = 1$, $k_1 = k_2 = 1$, and $S = 0.5$, $n = 4$).

is more preferred to stay there, corresponding to undifferentiated conditions. As self-activation is weaker, the basins of attractions on the side become deeper, thus the differentiated states are preferred. At the extreme value of a where self-activation is weak, the central basins of attraction becomes hilltop and therefore unstable, only the differentiated states are preferred. Thus destabilization of the central progenitor attractor as the self-activation parameter a is gradually decreased (Fig. 2*B*) (12) represents one possible mechanism for the development undifferentiated cell C committed to either differentiated cell A or B.

Another possible way to formalize the fate commitment is to assume there is stochastic fluctuation-driven transition from the central undifferentiated attractor into either of the side differentiated attractors. The fluctuations can be intrinsic from small copy numbers of molecules or extrinsic from the environments. The strength of fluctuations is quantified by diffusion coefficient in method section. Experimental evidences support the role of both a destabilization of the progenitor undifferentiated attractor (12), as well as gene expression fluctuation-induced state transitions (15).

Fig. 2*C* shows the barrier height as a measure of the landscape topography) for the transition from the central undifferentiated state to the differentiated state (green) and the escape time (red) from the former state under certain fluctuations versus the parameter a mimicking the differentiation process. We can see that the escape time correlates with the barrier height. The barrier height decreases (increases) as parameter a decreases (increases) during development, and the escape time becomes faster (slower). This guarantees the stability of the developmental process to the differentiated states (when parameter a goes down) or stability of undifferentiated state (when parameter a goes up) (5).

As one can see, the chance for differentiation at initial stage of development when a is large is nonzero but small (long escape time). As development proceeds (a decreases) the chance for differentiation becomes larger and reaches the largest when the undifferentiated state becomes unstable. We see both induction for instability of undifferentiated state through the change of self-activation and fluctuations in action for the developmental processes in our picture.

Pathways at a Given Stage of Development. From our path integral approach, we can uncover quantitatively the developmental paths from the undifferentiated state to the differentiated state at certain stage of development (at certain value of self-activation parameter a). Fig. 3 shows the transcription factor expression level X_1/X_2 state space as two dimensional contour and three dimensional landscape and the paths from theoretical reversed multipotent undifferentiated state to differentiated state for development (and the paths from differentiated state to undifferentiate state) at certain developmental stage characterized by self-activation strength a.

Fig. 3 *A* and *B* show that effective developmental paths (from central basin of the undifferentiated state to side basin of the differentiated state and vice versa) and the undifferentiated ones do not follow the gradient paths, that is the path determined by following the steepest directions of U. As we pointed out earlier, (for details see method and supporting information), the dynamics or paths of such developmental circuit is determined by both the force from the gradient of the landscape and an additional term representing the curl flux (4). The additional dynamical driving force emanating from the curl flux causes the paths to deviate from the naively expected steepest descent path computed from the gradient of U. This quantitative picture of paths is different from what would follow from Waddington's picture where the developmental paths, symbolized by the rolling down of a marble, follow the gradient of the underlying landscape. By contrast, the real developmental paths do not simply go down the gradient but also are driven by a curl force leading to spiraling movements that can be quantitatively tested in experiments.

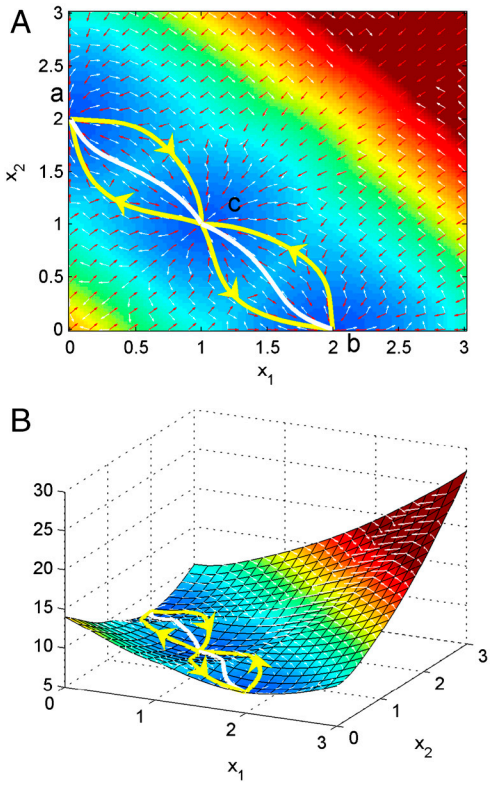

Fig. 3. (*A*) Two dimensional illustration of dominant kinetic path and flux between three basins of attraction in gene network. (*B*) Three dimensional illustration of dominant kinetic path and flux between three basins of attraction in gene network.

Furthermore, the forward developmental paths and backward retrodifferentiation paths are not identical. In other words, the developmental pathways are irreversible. The differentiation and the reverse process of retrodifferentiation follow different routes. This is unexpected from the original Waddington picture. The irreversibility of the developmental pathways is very fundamental and provides a unique prediction to test for developmental biology.

In Fig. 4 *A* and *B*, we show how fluctuations (defined in the method section as the strength of the autocorrelation function of protein concentrations quantitatively measured by the diffusion coefficient D) influence the (apparent) population barrier heights and the escape time for development. We see when the fluctuations increase, the population barrier height decreases, the escape time is faster. Nonzero but small fluctuations can help to accelerate the developmental processes. Large fluctuations can be damaging because the resulting population landscape becomes flat and therefore no essential preference or distinction among differentiated and undifferentiated state (equal probability) as well as other states. Therefore the landscape for development is stable against certain fluctuations.

In Fig. 4 *C* and *D*, we show how the fluctuations influence the paths. We can see that the consistency from the original paths for development decreases as fluctuations increase. For small fluctuations, the paths do not deviate much from the original ones. When fluctuations further increase, the barrier decreases, increasing the chances for different paths to go from undifferentiated state to the differentiated state. Therefore, the paths deviate more from the original paths when the fluctuations are larger. The paths for development is stable and robust against certain fluctuations.

Quantifying the Waddington Landscape and Paths of Development. From experimental evidences (5, 12), both mechanisms are in action for development: the instructive change of landscape via

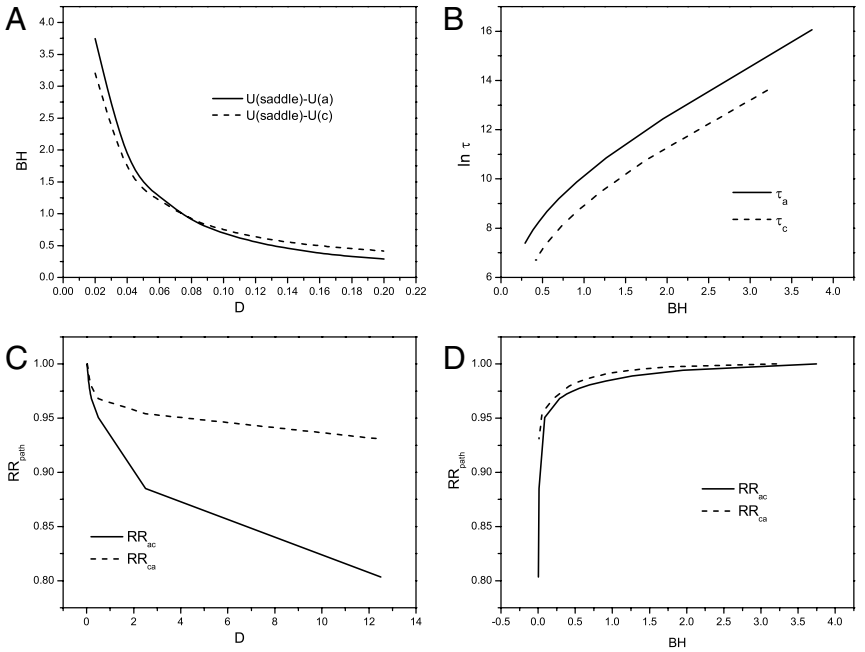

Fig. 4. (*A*) Barrier heights BH from side (center) to center (side) basin with solid line (dotted line) versus fluctuations via diffusion coefficients, *D*. (*B*) Logarithm of escape time τ from center (side minimum) versus barrier height, BH. (*C*) RR_{path} from from center (side) to side (center) versus *D*. RR_{path} represents the ratio of the weights between the path at fluctuation strength *D* compared with the low fluctuation strength $D_0 = 0.01$. (*D*) RR_{path} from versus the barrier height BH.

decrease of a and the stochastic state transitions in cell differentiation. To cover the whole picture of the developmental process, one needs to explore the dynamics of different developmental stages (from the stage where undifferentiated state is preferred to the stage where the differentiated states are preferred). Because in our model the different stages of development are characterized by a single parameter, the self-regulation strengths. We use that as our reaction coordinate for development, and the other coordinate (cross-section coordinate connecting undifferentiated to differentiated state, as in Fig. 3) in which the relationship between the undifferentiated state to differentiated state can be displayed and the expression level changes of the transcription factors can be monitored. The third and vertical axis is the height of the potential landscape (see Fig. 5).

We have discussed the dynamics at certain stage of development in the previous section. The full landscape and dynamical paths for the development requires the specification of the dynamics of self-regulation parameter *a*. We assume the self-activation strengths decrease in the developmental process due to the influences of the other regulators in a self-degraded fashion. Then the dynamical equations for the developmental process is controlled by the following equations: $\frac{dx_1}{dt} = \frac{ax_1^n}{S^n+x_1^n} + \frac{bS^n}{S^n+x_2^n} - k_1x_1 = F_1(x_1,x_2)$; and $\frac{dx_2}{dt} = \frac{ax_2^n}{S^n+x_2^n} + \frac{bS^n}{S^n+x_1^n} - k_2x_2 = F_2(x_1,x_2)$; and $\frac{da}{dt} = -\lambda a$. Here λ is the self-degradation rate. We assume self-activation strength *a* representing the developmental process changes relatively slowly compared to the dynamics of the x_1 and x_2 (For full time scale dynamics, see discussions in *SI Text*). In this way, in each developmental evolution stage, the system has time to relax to steady state among x_1 and x_2.

Based on the above developmental network circuit, we quantitatively obtain our probability landscape and associated paths for the developmental processes. This can be seen as the quantification of the Waddington landscape and associated chreodes.

Fig. 5 shows the developmental landscape and paths. The developmental reaction coordinate is parameter a. At large *a*, the undifferentiated basin of attraction dominates and stable as shown in the cross-section coordinate X linking side minimum-central minimum-side minimum representing the gene expression level. As the developmental process progresses (parameter a decreases), the undifferentiated state becomes less and less stable, the differentiated state becomes more and more stable forming basins of attractions. The qualitative similarities between our quantified landscape/paths in Fig. 5 and Waddington's landscape/paths in Fig. 1 are obvious. So the basic picture and features conjectured by Waddington for development and differentiation do exist and can now be quantified based on the underlying gene regulatory circuit. The quantitative description of the landscape and paths now allow for predictions. The Waddington landscape is no longer a metaphor. It is physical and quantifiable by the underlying probability landscape.

However, we need to point out to the differences between the quantified Waddington landscape/paths in Fig. 5 and one suggested by Waddington's marble shown in Fig. 1. Waddington describes the developmental processes as the downhill rolling of the marble. The undifferentiated state is unstable, which triggers the differentiation process. Whereas it is true that in vivo during embryonic development the stem cells are not stable (except when taken into culture and provided with the appropriate fac-

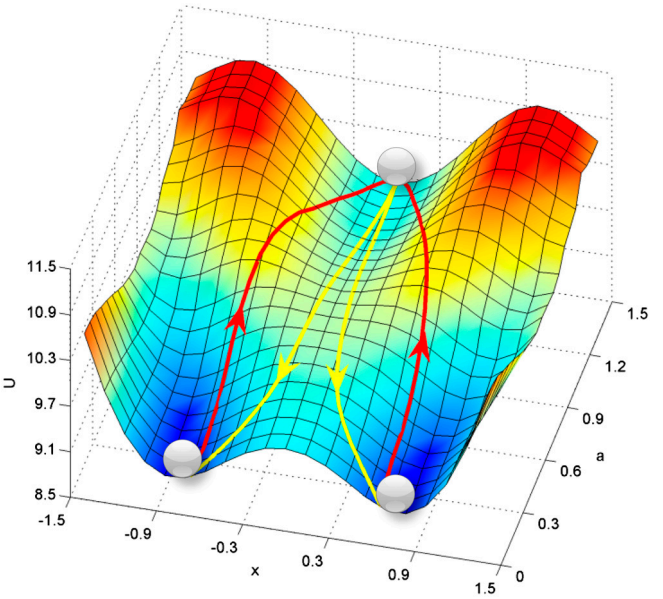

Fig. 5. The quantified Waddington developmental landscape and pathways ($a = a_1 = a_2$, $b_1 = b_2 = 1$, $k_1 = k_2 = 1$, $S = 0.5$, $n = 4$, and $\lambda = 0.01$).

tors), in the adult the multipotent undifferentiated cells are stable. In contrast to Waddington, our model allows for temporal stabilization of the undecided stem or progenitor state. This is not captured by the original Waddington picture of landscape. Our quantified landscape covers all stages of development process. Before or near the very early developmental stage, the undifferentiated state can still be stable but has a small but finite chance to climb up (induced from fluctuations) from the basin of attraction for escaping to the differentiated states. This process can happen even before the undifferentiated state become unstable.

The cell fate decision making process for development is also different in our picture compared with the Waddington picture. Whereas Waddington believes that the cell fate decision happens at the hill top, in our potential landscape we see that even before climbing up (induced by fluctuations) to the barrier top, the developmental paths (yellow lines) have already started to depart or bifurcate from each other. The bifurcation starts from the undifferentiated state even when it is still stable (down in the basin of attraction). The random directions of climbing out of the undifferentiated basin can bias to the choice of the different cell fates later on. This is the second quantitative difference between our quantified Waddington landscape (Fig. 5) and the original Waddington picture (Fig. 1).

We also quantified the dominant course of developmental paths from undifferentiated states (central basins of attraction) to the differentiated states (side basins of attractions) shown in yellow lines (Fig. 5). The developmental paths start from the undifferentiated states, climb up from the basin to reach the top, and then gradually bifurcate to two distinct pathways along the landscape valley to the differentiated states. As shown in Fig. 5, the developmental paths clearly do not follow gradient paths that the gravity driven metaphor of Waddingon would predict. This is the third quantitative deviation of our pathways (Fig. 5) from the Waddington paths or chreodes (Fig. 1). This is due to the fact that dynamics is controlled by both the force from landscape gradient and the force from the curl flux. The curl flux force makes the developmental path deviate from the steepest descent gradient path. This provides testable predictions.

Furthermore, there is a forth difference between our developmental paths (Fig. 5) and the Waddington paths (Fig. 1). The reverse paths in the original Waddington picture is supposed to be the same as the forward path. In other words, the retrodifferentiation process would follow exactly the same paths as differentiation except with opposite direction. By contrast, in our quantitative path picture, the paths are no longer reversible. The developmental paths (yellow lines) are clearly distinct from the retrodifferentiation paths (shown in red lines). This could be tested by studying the effective developmental path for instance in the iPS generation process. The study of paths on a potential landscape of development suggests a way to reverse the developmental process. We can increase the parameter a by increasing the self-activation strength of the transcription factors through the regulations or over expression of both transcription factors X_1 and X_2, as of the ongoing efforts in the stem cell research.

One should keep in mind that the quantified Waddington landscape topology depend on the structure of the underlying gene regulatory circuits. Accordingly, genetic mutations that change the network architecture through the wiring node, resulting rewiring (i.e., coupling strength b), change the resulting potential landscape as well as the paths for development. A slight distortion of the landscape may result in biasing paths to entire distinct, but independently robust valleys. This would straightforwardly explain homeotic mutations that Waddington noticed from studying drosophila. On the other hand, the landscape topology can also be altered by the links and connectivity strengths even when there is no missing nodes or genetic mutations. So through the epigenetics with environmental and local chemical reaction changes as Waddington envisioned, the network connection strengths can be altered, resulting the changes of the Waddington landscape and associated paths for development. Therefore both genetic changes (mutations on gene nodes) and epigenetic changes (link strengths between genes) are important and can alter the topology of landscape and paths for development.

Our study provides a framework for studying the global nature of the binary fate decision and commitment of a multipotent cell into one of two cell fates (8). The two fate options for CMP in hematopoietic system correspond to the erythroid fate and the myeloid fate attractor (side minimum). Measurement of PU.1 and GATA1 are confirming the predicted symmetric gene expression configuration in the progenitor state, which exhibits the coexpression of both fate determining factors at intermediate levels (center minimum) (12). Further experimental verifications and design can be studied by the use of two different GFPs for PU.1 and GATA1 and collecting the expression data. The joint histogram will directly give the information of the underlying landscape, confirming the tristable basins for cell decision making. Further experimental exploration can be carried out through the modulation of the transcription factor Klf4 for controlling the self-regulation strength and therefore the differentiation/undifferentiation process. The real time trajectories of expressions can provide us information on dynamics (jumps between basins) and kinetic paths, linking to the underlying landscape topography such as basins and barrier heights in between (through joint histograms). Previous experimental and theoretical studies (12) explored the natures of the cell fate decision systems using flow cytometry and dynamical system analysis at certain stages of development. In the present work, we focus on the underlying global landscape, transitions between the attractors and associated kinetic paths for the dynamical developmental process.

Methods and Models

Potential and Flux Landscape for Development. We first evaluated the global dynamics of the developmental network motif so that we can quantitatively assign potentials to the attractor states and determine their distinct relative depth within the same frame of reference. The idea of a potential landscape describes how forces acting in a system relate to its global behavior. They are particularly useful for systems of interacting components, such as chemical reactions and protein dynamics, motion, and folding (16). However, these applications deal with equilibrium systems where the potential function is a priori knowable. For nonequilibrium systems, such as gene circuits that exhibit stable stationary states in higher ($N > 1$) dimensional state space, the intuition of some form of a potential is still warranted and widely used metaphorically (1); however, its functional form is not easy to obtain. In the gene circuit $d\mathbf{x}/dt = \mathbf{F}(\mathbf{x})$, exemplified in Eq. [**S2**] of *SI Text*, the vector $\mathbf{F}(\mathbf{x})$ is the force that drives the system. However, $\mathbf{F}(\mathbf{x})$ cannot in general be written as a gradient of a potential U: $\mathbf{F}(\mathbf{x}) \neq -\text{grad}(U)$ for systems of more than one dimensions. For stochastic systems, the system is not determined by the local trajectories but by the probability distributions. The probability distribution follows master equations for discrete state space or diffusion equations for continuous case.

In continuous representation, the diffusion equation for probability evolution can be written in the form of the probability conservation: $\partial P/\partial t + \nabla \cdot J = 0$ where J is the probability flux $J = FP - D\nabla P$. This means flux in and out of the system gives the increase or decrease of the local probability. Whereas in nonequilibrium systems at steady-state the divergence of the probability flux $\mathbf{J}_{ss}$ vanishes $\nabla \cdot J = 0$, the flux itself need not vanish. When local flux is equal to zero, the detailed balance is preserved and the system is in equilibrium state. When local flux is not equal to zero, the detailed balance is broken and the system is in nonequilibrium state. We found (details in *SI Text*) (4, 5) that $\mathbf{F} = -D \cdot \nabla U + \mathbf{J}_{ss}/P_{ss}$. We have decomposed the force driving the dynamics of the system into two terms, the gradient of the

potential U where U is linked with the steady-state probability by $U = -\ln(P_{ss})$ and curl flux force linking the divergence-free steady-state (long time limit) probability flux $\mathbf{J}_{ss}$ (velocity current) and the steady-state probability P_{ss} (density). Divergence-free flux has no place to start or end. It is in this sense that the flux has a curl nature.

We treated x_1 and x_2 as independent variables. For the complete model of development, both concentration variables x_1 and x_2 characterizing the gene expression and self-regulation variable a are dynamical (as shown in Fig. 5). For simplicity of the treatment, we assume here the D is a constant independent of concentrations. Detailed discussions are in *SI Text*.

Developmental Pathways Through Path Integral. The dynamics of the cellular network has often been studied by the chemical reaction rate equations of various local protein concentration species. However, both internal statistical fluctuations from finite number of molecules within the cell and external fluctuations from cellular environments can be significant (17). Therefore it is more appropriate to formulate the dynamics of the chemical rate equations in the noisy fluctuating environments. The dynamics therefore can be formulated as: $\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) + \eta$ where $\mathbf{x}$ is the protein concentration vector, $\mathbf{F}(\mathbf{x})$ is the chemical rate flux vector. η is Gaussian noise term where its autocorrelation function is $\langle \eta(\mathbf{x},t)\eta(\mathbf{x},0)\rangle = 2D\delta(t)$. D is the diffusion coefficient. The noise term is related to the intensity of cellular fluctuations either from the environmental external fluctuations or intrinsic fluctuations (under large number expansions, the process follows Brownian dynamics and diffusion coefficient is often concentration dependent).

We can now formulate the dynamics for the probability of starting from initial configuration $\mathbf{x}_{\text{initial}}$ at $t = 0$ and end at the final configuration of $\mathbf{x}_{\text{final}}$ at time t, with the Onsager–Machlup functional (11) as $P(\mathbf{x}_{\mathbf{final}},t,\mathbf{x}_{\mathbf{initial}},0) = \int D\mathbf{x}\exp[-\int dt(\frac{1}{2}\nabla\cdot\mathbf{F}(\mathbf{x}) + \frac{1}{4}(d\mathbf{x}/dt - \mathbf{F}(\mathbf{x}))\cdot\frac{1}{\mathbf{D}(\mathbf{x})}\cdot(d\mathbf{x}/dt - \mathbf{F}(\mathbf{x})))] = \int D\mathbf{x}\exp[-S(\mathbf{x})] = \int D\mathbf{x}\exp[-\int L(\mathbf{x}(t))dt]$.

The integral over $D\mathbf{x}$ represents the sum over all possible paths connecting $\mathbf{x}_{\mathbf{initial}}$ at time $t = 0$ to $\mathbf{x}_{\mathbf{final}}$ at time t. $\mathbf{D}(\mathbf{x})$ is the diffusion coefficient matrix tensor. The second term of the exponent represents the weight contribution from specific trajectory path due to the underlying Gaussian noise. The first term of the exponent represents the contribution due to the variable transformation from the Gaussian noise η to the trajectory path x (Jacobian). The exponential factor gives the weight of each path. So the probability of network dynamics from initial configurations $\mathbf{x}_{\mathbf{initial}}$ to the final state $\mathbf{x}_{\mathbf{final}}$ is equal to the sum of all possible paths with different weights. The $S(\mathbf{x})$ is the action and $L(\mathbf{x}(t))$ is the Lagrangian or the weight for each path (Fig. 2).

Notice that not all the paths give the same contribution. We can approximate the path integrals with a set of dominant paths. Because each path is exponentially weighted, the other subleading path contributions are often small and can be ignored. One can easily use this observation to find the paths with the optimal weights. We identify the optimal paths as biological paths or developmental pathways in our case.

Once the paths are known, we can substitute back to the path integral formulation to calculate the probability evolution in time. We can obtain the rate or speed of kinetics from one state to another. One can also use the long time limit to infer the weights of states and therefore map out the landscape. Details are given in *SI Text*.

We further notice that if the force F is a gradient, then the term $F\cdot dx$ in the weight functional above is a constant depending only on ending points. When force F is not purely a gradient (nonequilibrium with no detailed balance), then the curl flux component of the force leads to path dependent weights $\int F\cdot dx$. It will create a topological winding (nonzero) contribution to the weights of the paths going back to itself ($\oint F\cdot dx \neq 0$). This will result the deviation of the pathways from the pure gradient paths and furthermore the forward path and backward path will not be the same ($\int_{\text{forward}} F\cdot dx \neq -\int_{\text{backward}} F\cdot dx$) and therefore the corresponding developmental pathways are irreversible. Details are shown in *SI Text*. We point out that curl flux as a result of detailed balance breaking of nonequilibrium system is the origin of optimal kinetic paths deviating from the steepest descent one, which was not previously known (9, 10).

Conclusion

We developed a theoretical framework to quantify the Waddington landscape and biological paths for development and differentiation. We quantified the Waddington landscape through the construction of the underlying probability landscape for the cell development. We show the developmental process proceeds as moving from undifferentiated basin of attraction to the differentiated attractors. The barrier heights between the basins of attractions correlate with the escape time that determines the stability of cell types.

We show that the developmental process can be quantitatively described and uncovered by the biological paths on the quantified Waddington landscape and the dynamics of the developmental process is controlled by a combination of the gradient and curl force on the landscape. The biological paths do not follow the normally expected steepest descent path on the landscape. We show that the biological paths of the reverse differentiation process or reprogramming are irreversible and different from the ones of the differentiation process.

We found that the developmental process described by the underlying landscape and the associated biological paths is stable and robust against the influences of environmental perturbations.

ACKNOWLEDGMENTS. We thank Dr. Sui Huang for help discussions and comments. We thank Virtual Cell (supported by National Institutes of Health Grant P41RR013186 from the National Center For Research Resources) for software support. We acknowledge support from the National Science Foundation and National Natural Science Foundation of China Grants 20735003 and 90713022, and 973 Project Grants 2009CB930100 and 2010CB933600.

1. Waddington CH (1957) *The Strategy of the Genes* (Allen &Unwin, London).
2. Huang S (2009) Reprogramming cell fates: reconciling rarity with robustness. *Bioessays* 31:546–560.
3. MacArthur BD, Ma'ayan A, Lemischka IR (2009) Systems biology of stem cell fate and cellular reprogramming. *Nat Rev Mol Cell Biol* 10:672–681.
4. Wang J, Xu L, Wang EK (2008) Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proc Natl Acad Sci USA* 105:12271–12276.
5. Wang J, Xu L, Wang EK, Huang S (2010) The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophys J* 99:29–39.
6. Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22:437–467.
7. Graf T, Enver T (2009) Forcing cells to change lineages. *Nature* 462:587–590.
8. Zhou JX, Huang S (2010) Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends Genet* 27:55–62.
9. Weinan E, Vanden-Eijnden E (2010) Transition-oath theory and path-finding algorithms for the study of rare events. *Annu Rev Phys Chem* 61:391–420.
10. Roma DM, O'Flanagan RA, Ruckenstein AE, Sengupta AM, Mukhopadhyay R (2005) Optimal path to epigentic switching. *Phys Rev E* 71(011902):1–5.
11. 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.
12. Huang S, Guo YP, May G, Enver T (2007) Bifurcation dynamics of cell fate decision in bipotent progenitor cells. *Dev Biol* 305:695–713.
13. Hu M, et al. (1997) Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 11:774–785.
14. Jiang J, et al. (2008) A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* 10:353–360.
15. Kalmar T, et al. (2009) Regulated fluctuations in Nanog expression mediate cell fate decisions in ES and EC cells. *PLoS Biol* 7:e1000149.
16. Wolynes PG, Onuchic JNO, Thirumalai D (1995) Navigating the folding routes. *Science* 267:1619–1622.
17. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* 99:12795–12800.

Supporting Information

Wang et al. 10.1073/pnas.1017017108

SI Methods and Models

Network Module for Development. We will study a relatively simple yet important example of cell fate decision for development (Fig. S1) (1). The GATA1-PU.1 circuit can be used to study a bistable cell fate decision system for differentiation of multipotent stem cell (hematopoietic stem cell in the bone marrow). The relative levels of transcription factors GATA1 and PU.1 translate to all or none fate decision. Over expression of PU.1 leads to suppression of GATA1 and to the differentiation of the macrophage/monocyte lineage. Over expression of GATA1 causes suppression of PU.1 and promotes differentiation into megakaryocyte/erythrocyte precursor cells. GATA1 and PU.1 not only mutually repress each other, but also auto-stimulate themselves. Therefore, excess of PU.1 leads the balance toward myelomonocytic fate and excess of GATA1 leads the balance toward the erythroid fate (1).

The simplified developmental network motif regulating the cell fate behavior is often composed of two proteins. x_1 and x_2 denote two transcription factor proteins GATA1 and PU.1 with mutual-inhibition and self-activation. The Hill functions explicitly quantify the degree of the activator and inhibitor influences as sigmoidal function. The kinetic equations are given as

$$\begin{aligned}\frac{dx_1}{dt} &= \frac{a_1 x_1^n}{S^n + x_1^n} + \frac{b_1 S^n}{S^n + x_2^n} - k_1 x_1.\\ \frac{dx_2}{dt} &= \frac{a_2 x_2^n}{S^n + x_2^n} + \frac{b_2 S^n}{S^n + x_1^n} - k_2 x_2.\end{aligned} \qquad \text{[S1]}$$

Where the parameters a_1 and a_2 express the relative strength of the self-activation of GATA1 and PU.1, b_1 and b_2 represent the strength of the mutual-inhibition on the basal activation. k_1 and k_2 show the rate of degradation. The parameters S represent the inflection point of the explicitly sigmoidal functions and n is Hill coefficient determined the steepness of the sigmoidal function or binding cooperativity.

The additive regulation is motivated from the experimental considerations. Because self-activation strengths can be modulated by gene regulations through the increase of the concentrations of transcription factors Klf4 whereas the mutual regulation is relatively fixed. In other words, the mutual regulations and self-activations are some what separated. They do not have to be simultaneously changed in a synchronized way. If we do include the AND logic by multiplying the mutual regulation and self-activation together, we often see bistability instead of tristability, which seems not the best model for explaining the stem cell fate decision.

It is interesting to consider what happened when only one of self-regulation strength is changed. We actually did some explorations on that. We found that when one of the self-regulation strengths decreases, both the center and the side minimum corresponding to the gene expression with changing self-regulation strength gradually diminish together. If we decrease the mutual repression strength of the corresponding gene also, we can see that first the side minimum disappears and then the center minimum disappears. However, in the experiments, this seems still a challenge to realize. One has to block the effect of transcription factor for example, Klf4 on one gene and not the other.

The regulation dynamics discussed is for fixed self-regulation strength. It can describe the regulation dynamics at certain development stage. When including the self-regulation dynamics, we can monitor the full dynamical developmental process rather than only the stages of development in a steady way,

$$\frac{da}{dt} = -\lambda a, \qquad \text{[S2]}$$

where a is the self-regulation strength (we assume symmetric case when $a = a_1 = a_2$ for simplicity) mimicking the developmental process. We learn something more from keeping track of both concentration and self-regulation dimension. Concentrations quantify the gene expression level dynamics whereas the self-regulation quantifies the developmental process. In other words, for fixed self-regulation strength, we have a fixed wiring of gene regulation, which provides a picture and landscape of gene regulation dynamics at certain stage of the development. However, as the developmental process proceeds, the gene network wiring is not fixed, the connections is changed by the gene regulations and in this example through the self-regulations. So only through monitoring both concentrations and self-regulations one can see the whole picture of the cell development and differentiation process.

We treated x_1 and x_2 as independent variables. In the results, because we also want to show the landscape (as for example, z axis) and monitor the developmental process characterized by the self-regulation strength a (as for example, y axis), we use a and the reaction coordinate (as for example, x axis) we defined as the one connecting the side-center-side minimum. For the complete model of development, both concentration variables x_1 and x_2 characterizing the gene expression and self-regulation variable a are dynamical (as shown in Fig. 5 in the main text). For simplicity of the treatment, we assume here the D is a constant independent of concentrations, which is usually a reasonable approximation for external fluctuations. In principle, the cell fate decision can be influenced by both external fluctuations and intrinsic fluctuations. In reality, the intrinsic fluctuations are only important when the number of the molecules is small. This may or may not be case for development and differentiation because it is not clear in the multicelluar developmental process the number of molecules is necessarily small. Nevertheless, we will return for the more complicated concentration dependent D for intrinsic fluctuations in the future study.

Path Integrals. The dynamics of the cellular network has often been studied by the chemical reaction rate equations of various local protein concentration species. However, both internal statistical fluctuations from finite number of molecules within the cell and external fluctuations from cellular environments can be significant (2–7). Therefore it is more appropriate to formulate the dynamics of the chemical rate equations in the noisy fluctuating environments. The dynamics therefore can be formulated as

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) + \eta. \qquad \text{[S3]}$$

Where $\mathbf{x}$ is the protein concentration vector, $\mathbf{F}(\mathbf{x})$ is the chemical rate flux vector. η is Gaussian noise term where its correlation is $\langle \eta(\mathbf{x},t)\eta(\mathbf{x},0)\rangle = 2D\delta(t)$. D is the diffusion coefficient. The noise term is related to the intensity of cellular fluctuations either from the environmental external fluctuations or intrinsic fluctuations (under large n expansions, the process follows Brownian dynamics and diffusion coefficient is often concentration dependent).

We can now formulate the dynamics for the probability of starting from initial configuration $\mathbf{x}_{\text{initial}}$ at $t = 0$ and end at the final configuration of $\mathbf{x}_{\text{final}}$ at time t, with the Onsager–Machlup functional (8–16) as

$$P(\mathbf{x}_{\mathbf{final}},t,\mathbf{x}_{\mathbf{initial}},0) = \int D\mathbf{x}\exp\left[-\int dt\left(\frac{1}{2}\nabla \cdot \mathbf{F}(\mathbf{x}) + \frac{1}{4}(d\mathbf{x}/dt - \mathbf{F}(\mathbf{x}))\cdot\frac{1}{\mathbf{D}(\mathbf{x})}\cdot(d\mathbf{x}/dt - \mathbf{F}(\mathbf{x}))\right)\right] = \int D\mathbf{x}\exp[-S(\mathbf{x})] = \int D\mathbf{x}\exp\left[-\int L(\mathbf{x}(t))dt\right]. \quad \textbf{[S4]}$$

The integral over $D\mathbf{x}$ represents the sum over all possible paths connecting $\mathbf{x}_{\mathbf{initial}}$ at time $t = 0$ to $\mathbf{x}_{\mathbf{final}}$ at time t. $\mathbf{D}(\mathbf{x})$ is the diffusion coefficient matrix tensor. The second term of the exponent represents the weight contribution from specific trajectory path due to the underlying Gaussian noise. The first term of the exponent represents the contribution due to the variable transformation from the Gaussian noise η to the trajectory path x (Jacobian). The exponential factor gives the weight of each path. So the probability of network dynamics from initial configurations $\mathbf{x}_{\mathbf{initial}}$ to the final state $\mathbf{x}_{\mathbf{final}}$ is equal to the sum of all possible paths with different weights. The $S(\mathbf{x})$ is the action and $L(\mathbf{x}(t))$ is the Lagrangian or the weight for each path (Fig. S2).

Dominant Paths: Pathways. Notice that not all the paths give the same contribution. We can approximate the path integrals with a set of dominant paths. Because each path is exponentially weighted, the other subleading path contributions are often small and can be ignored. One can easily use this observation to find the paths with the optimal weights. The dominant paths should satisfy the Euler–Lagrangian equation from the minimization of the action or Lagrangian (see Fig. S2): $\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}} - \frac{\partial L}{\partial \mathbf{x}} = 0$. If a scalar potential $U(\mathbf{x})$ exists, it makes the equation, $\mathbf{F}(\mathbf{x}) = -\nabla U(\mathbf{x})$, can be obtain. Then, the E–L equation is simplified as: $\ddot{\mathbf{x}} - \frac{1}{2}\frac{\frac{\partial \mathbf{D}(x)}{\partial \mathbf{x}}}{\mathbf{D}(\mathbf{x})}\dot{\mathbf{x}}^2 = 2\mathbf{D}(\mathbf{x})\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}$, where $V(\mathbf{x}) = \frac{1}{4D}\mathbf{F}^2 + \frac{1}{2}\nabla\cdot\mathbf{F}(\mathbf{x})$. The equation of motion of x has the acceleration term $\ddot{x}$, the frictional (positive and negative) term $\frac{1}{2}\frac{\frac{\partial \mathbf{D}(\mathbf{x})}{\partial \mathbf{x}}}{\mathbf{D}(\mathbf{x})}\dot{\mathbf{x}}^2$ and the force term $2\mathbf{D}(\mathbf{x})\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}$. The problem becomes one of a n-dimensional particle moving in a force with friction (n is the number of different protein species in the cell).

But in general the potential $U(\mathbf{x})$ is not present for nonequilibrium networks, we can obtain the equation of dominant path as follows:

$$\ddot{\mathbf{x}} - \frac{1}{2}\frac{\frac{\partial \mathbf{D}(\mathbf{x})}{\partial \mathbf{x}}}{\mathbf{D}(\mathbf{x})}\dot{\mathbf{x}}^2 = \mathbf{E} - \nabla(\dot{\mathbf{x}}\cdot\mathbf{F}) + (\dot{\mathbf{x}}\cdot\nabla)\mathbf{F}, \quad \textbf{[S5]}$$

where $\mathbf{E} = 2\mathbf{D}(\mathbf{x})\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}$. When $\mathbf{D}(\mathbf{x})$ is a constant, the friction term is zero. For easy understanding, we consider the situation of three dimension. Therefore, the right of the equation can be written as: $\mathbf{E} + \dot{\mathbf{x}}\times\mathbf{B}$ (the form can only be written in three dimension). where $\mathbf{B} = \nabla\times(-\mathbf{F})$. As we can see, if the driving force of the dynamics F is a gradient force, then $B = 0$, and there is no curl component of the driving force. If F can not be written as a gradient force, then in general $B \neq 0$, and there is a curl current component of the driving force.

In general, the dynamic driving force F can be decomposed into a gradient of a potential and a curl flow flux ($\mathbf{F} = \mathbf{D}/P_{ss}\cdot\frac{\partial}{\partial \mathbf{x}}P_{ss} + \mathbf{J}_{ss}(\mathbf{x},t)/P_{ss} = -D\frac{\partial}{\partial \mathbf{x}}U + \mathbf{J}_{ss}(\mathbf{x},t)/P_{ss}$. P_{ss} represent steady state. With detailed balance, the gradient of potential energy controls the underlying dynamics as the driving force. Whereas for nonequilibrium systems, the gradient of potential landscape and flux of probability determine the dynamics and global properties together. The dynamics of a nonequilibrium network can be described as a spiral, along the gradient direction, not like the case of the equilibrium state only following the gradient. It is similar to electrons moving in both electric and magnetic field (17). Here we will see the dominant path is determined by both gradient and curl components of the force.

For simplicity, we assume the diffusion coefficient tensor matrix is diagonal whereas the nondiagonal elements to account for the kinetic coupling between different protein species are zero. Considering in three dimensions, the equation of motion in the scaling form is given as

$$\ddot{\mathbf{x}} = 2D_{11}\frac{\partial V}{\partial x} + \dot{y}\left(\frac{\partial F_x}{\partial y} - \frac{\partial F_y}{\partial x}\right) + \dot{z}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)$$
$$\ddot{\mathbf{y}} = 2D_{22}\frac{\partial V}{\partial y} + \dot{z}\left(\frac{\partial F_y}{\partial z} - \frac{\partial F_z}{\partial y}\right) + \dot{x}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)$$
$$\ddot{\mathbf{z}} = 2D_{33}\frac{\partial V}{\partial z} + \dot{x}\left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right) + \dot{y}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right).$$

It is analogous to the electrons moving in the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$.

In the two dimension system, the equation of motion is given as follows:

$$\ddot{\mathbf{x}} = 2D_{11}\frac{\partial V}{\partial x} + \dot{y}\left(\frac{\partial F_x}{\partial y} - \frac{\partial F_y}{\partial x}\right) \qquad \ddot{\mathbf{y}} = 2D_{22}\frac{\partial V}{\partial y} + \dot{x}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right).$$

The solution of the dominant path equations is important for understanding the kinetic mechanisms of the nonequilibrium biological networks. We can directly solve the equations of motion by the boundary condition and obtain the dominant kinetic path from the initial to the final state. But it is not always easy to solve the two boundary value problem, especially when the dimension n is enormous.

Hamilton–Jacobian (HJ) Framework for Path Integral. On the other hand, we can evaluate the weights of the kinetic paths from our path integral formalism. When the action $S(x)$ is minimized directly, the most probable trajectory can be obtained. the Lagrangian is written as

$$L(\mathbf{x}) = \frac{1}{4D}\dot{\mathbf{x}}^2 + V(\mathbf{x}) - \frac{1}{2D}\mathbf{F}(\mathbf{x})\cdot\dot{\mathbf{x}}, \quad \textbf{[S6]}$$

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

$$H(\mathbf{x}) = -L(\mathbf{x}) + \mathbf{P}(\mathbf{x})\cdot\dot{\mathbf{x}} = E_{\text{eff}}. \quad \textbf{[S7]}$$

In the light of the above equation,we can obtain

$$\frac{\dot{\mathbf{x}}^2}{4D} - V(\mathbf{x}) = E_{\text{eff}} \quad \textbf{[S8]}$$

$$|\dot{\mathbf{x}}| = \sqrt{4D(E_{\text{eff}} + V(\mathbf{x}))}. \quad \textbf{[S9]}$$

When we substitute Eq. **S7** into the action, we obtain $S(\mathbf{x}) = \int(\mathbf{P}(\mathbf{x})\cdot\dot{\mathbf{x}} - H(\mathbf{x}))dt$. We can see the action characterizing the weights of the paths depend on the values of the Hamiltonians. Specific values of the Hamiltonians correspond to specific values of the final time T. For a fixed Hamiltonian, there is a corresponding optimal path to it minimizing the action.

According to the least action principle, if the Hamiltonian of the system is a constant, then the variation of the action, for given initial and final coordinates and times (initial and final) is zero. If we allow a variation of the final time T and leave the initial coordinate and final coordinate fixed, then we have

$$\delta S = -H\delta t. \quad \text{[S10]}$$

For constant Hamiltonian, $\delta S = -E\delta t$. Because $S(\mathbf{x}) = \int(\mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} - H(\mathbf{x}))dt$, we can define $S_0 = \int(\mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}}dt$. We find that

$$\delta S_0 = 0. \quad \text{[S11]}$$

Therefore, the action S_0 is minimized with respect to all the paths that satisfy the constant Hamiltonian and pass through the final point at any instant.

For multidimensional problems, the action in terms of multidimensional integral is still not easy to characterize because the results depend on not only the initial and final space but also on initial and final time. In the HJ framework, we notice that we can transform the formulations into a different representation in x space: $S_0 = S_{\mathrm{HJ}}(\mathbf{x}) = \int\sum_i \frac{1}{2D}(\dot{\mathbf{x}}_\mathbf{i} - F_i)dx_i = \int\sum_i p_i(\mathbf{x})dx_i$, where p_i is the associated momentum. Now the action only depends on the initial and final states. So the problem of optimization of action becomes the optimization of accumulated momentum along the trajectories connecting the initial and final points in state space.

Determining the dominant paths for realistic large networks has technical challenges with different final times, recrossings and the two fixed end boundary-value rather than the relatively simple initial-value problems.

Notice that the action is in terms of the integral $S_{\mathrm{HJ}} = \int\sum_i p_i(\mathbf{x})dx_i$, a dot product in multidimensional space of $\mathbf{x}$, which is still hard for calculations and visualizations. This action can be simplified furthermore and is equivalent to a line integral along a particular one dimensional path l so that $S_{\mathrm{HJ}} = \int\sum_i p_i(\mathbf{x})dx_i = \int p_l dl$ where $p_l = \sqrt{(E_{\mathrm{eff}} + V(\mathbf{x}))/D} - \frac{1}{2D}F_l$. This allows to switch from the time-dependent to the Hamiltonian-dependent HJ description (18–21). The dominant pathway connecting given initial and final positions is obtained by minimizing the action in the HJ representation

$$S_{\mathrm{HJ}} = \int_{x_i}^{x_f}\left(\sqrt{(E_{\mathrm{eff}} + V(\mathbf{x}))/D} - \frac{1}{2D}F_l\right)dl, \quad \text{[S12]}$$

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, according to

$$t_f - t_i = \int_{x_i}^{x_f} dl\sqrt{\frac{1}{4D(E_{\mathrm{eff}} + V(\mathbf{x})}}. \quad \text{[S13]}$$

In the present work, we adopted the simple choice $E_{\mathrm{eff}} = -V(x_{\mathrm{min}})$ ($-V(x_{\mathrm{min}})$is the minimum of the effective potential at location x_{min}), which corresponds to the longest kinetic time. The HJ formulation of the dynamics in length space is much more efficient to avoid spending long times in meta-stable traps than conventional approach by considering intervals of fixed displacements rather than fixed time length. The numerical advantages of the HJ formalism for describing long-time dynamics at constant energy were first pointed out in (18–21) and applied to equilibrium system where dynamics is determined by gradient energy. In this work, the computational advantages can also be achieved for nonequilibrium stochastic dynamics of networks, in which the effects of fluctuations and dissipations are consistently taken into account and the dynamics of the system can not be written as a pure gradient of the underlying landscape (the driving force can be decomposed as a gradient of the underlying landscape and a curl force from the nonequilibrium current).

Discrete Space HJ Paths and Underlying Potential Landscape. The optimal paths was obtained by minimizing the discretized target function

$$S_{\mathrm{HJ}} = \sum_n^{N-1}(\sqrt{(E_{\mathrm{eff}} + V(n))/D} - \frac{1}{2D}F_l(n))\Delta l_{n,n+1} + \lambda P. \quad \text{[S14]}$$

Where $P = \sum_i^{N-1}(\Delta l_{i,i+1} - \langle\Delta l\rangle)^2$ and

$$\begin{aligned}(\Delta l)^2_{n,n+1} &= \sum_i(\mathbf{x}_i(n+1) - \mathbf{x}_i(n))^2\\ F_l(n) &= \sum_i \mathbf{F}_i(\mathbf{x}(n))(\mathbf{x}_i(n+1) - \mathbf{x}_i(n))/\Delta l_{n,n+1}\\ V(n) &= \sum_i\left\{\frac{1}{4D}\mathbf{F}^2(\mathbf{x}_i) + \frac{1}{2}\sum_j\frac{\partial\mathbf{F}_j(\mathbf{x}_i)}{\partial x_j}\right\},\end{aligned} \quad \text{[S15]}$$

$\Delta l_{n,n+1}$ 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 (18–21) and becomes irrelevant in the continuum limit. We have checked that, with 20 discretization steps, simulations performed on a wide range of λ lead to consistent results. The minimization of the discretized HJ effective action was performed by applying a simulated annealing algorithm or the conjugate gradient algorithm. we can project it (them) into the protein concentration space to visualize the biological path (s). Furthermore, we can also substitute the biological path into the the expression of effective action to obtain the weight associated with the path(s). So in this way, we can fully quantitatively characterize the biological path in terms of identifications and quantification of the weights.

In the small diffusion coefficient or low noise limit, we can identify the fixed points of the deterministic equations. The stable fixed point will most likely correspond to the basin of attraction. Furthermore, when the system has one basin of attraction, then we can choose the position of the basin of attraction as the reference state (final state). By exploring the protein concentration space through varying the initial states, we can obtain the effective action $\mathbf{S}$ or weight $\mathbf{W}$ associated with each state relative to the reference state $W(\mathbf{x}) = \exp[-S(\mathbf{x})]$ (using dominant path approximations). In this way, we can map out the generalized potential energy landscape associated with the weights or probability for a network with monostability $U(\mathbf{x}) = -\ln W(\mathbf{x}) = S(\mathbf{x})$.

When the system has multiple basins of attraction, then we can choose all the positions of the basins of attraction as the reference states (final states, $x_{f1}, x_{f2}, \ldots x_{fN}$). For any arbitrary initial state, we can calculate the actions $\mathbf{S}$ relative to each of the basin of attractions: $S_1(\mathbf{x}), S_2(\mathbf{x}), \ldots, S_N(\mathbf{x})$, as well as the corresponding weights associated, $\mathbf{W}$: $W_1 = \exp[-S_1(\mathbf{x})]$, $W_2 = \exp[-S_2(\mathbf{x})], \ldots$, $W_N = \exp[-S_N(\mathbf{x})]$. Then we choose the optimal (largest) weight $\mathbf{W}$ or optimal (smallest) action $\mathbf{S}$ for this particular state to be the measure of the action or weight relative to all the basins of attraction reference points. By exploring the protein concentration space through varying the initial states, we can finally obtain the optimal action or weight associated with each state relative to the reference states of the basins of attractions $x_{f1}, x_{f2}, \ldots x_{fN}$. In this way, we can map out the generalized potential energy landscape associated with the weights or probabilities for the whole network with multistability $U(\mathbf{x}) = -\ln S_{\mathrm{optimal}}(\mathbf{x})$.

For this system, we only have two basins of attraction. We can follow the above procedure and obtain both the dominant biological path as well as the corresponding generalized potential landscape associated with the weights or probability in state space.

In a nonequilibrium open system, there are constant exchange in energy and information. This results the dissipation of energy.

The dissipation gives a global physical characterization of the nonequilibrium system. In the steady state, the dissipation of energy is closely associated with the entropy production rate. The entropy formula is well-known (22),

$$S = -k_{\mathrm{B}} \int P(\mathbf{x},t) \ln P(\mathbf{x},t) dx. \quad \text{[S16]}$$

By differentiating the above function, the increase of the entropy at constant temperature T is shown as follows:

$$T\dot{S} = k_{\mathrm{B}} * T \int (\ln P + 1) \nabla \cdot \mathbf{J} dx = -\int (k_{\mathrm{B}} T \nabla \ln P - \mathbf{F}) \cdot \mathbf{J} dx - \int \mathbf{F} \cdot \mathbf{J} dx, \quad \text{[S17]}$$

where $-\int (k_{\mathrm{B}} T \nabla \ln P - \mathbf{F}) \cdot \mathbf{J}\ dx = e_p$ is the entropy production rate following Onsager (8), and $\int \mathbf{F} \cdot \mathbf{J}\ dx = h_d$ is the mean rate of the heat dissipation. For a steady state, the $\dot{S} = 0$, and the entropy production e_p equal to the heat dissipation h_d.

Path Integral with General Inhomogeneous and Anisotropic Diffusion Tensor. The Lagrangian

$$\mathfrak{L} = \frac{1}{4} \sum_{i,j} (\dot{x}_i - F_i)(\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) + \frac{1}{2} \sum_i \frac{\partial F_i}{\partial x_i}. \quad \text{[S18]}$$

Then the Euler–Lagrange equation gives

$$\begin{aligned} 0 &= \frac{\partial \mathfrak{L}}{\partial x_\ell} - \frac{d}{dt}\left(\frac{\partial \mathfrak{L}}{\partial \dot{x}_\ell}\right) \\ &= -\frac{1}{2} \sum_{i,j} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) + \frac{1}{4} \sum_{i,j} (\dot{x}_i - F_i) \frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} (\dot{x}_j - F_j) \\ &\quad + \frac{1}{2} \sum_i \frac{\partial F_i}{\partial x_i \partial x_\ell} - \frac{1}{2} \frac{d}{dt} \left[\sum_j (\mathbf{D}^{-1})_{\ell j} (\dot{x}_j - F_j) \right] \\ &= -\frac{1}{2} \sum_{i,j} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) + \frac{1}{4} \sum_{i,j} (\dot{x}_i - F_i) \frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} (\dot{x}_j - F_j) \\ &\quad + \frac{1}{2} \sum_i \frac{\partial F_i}{\partial x_i \partial x_\ell} \\ &\quad - \frac{1}{2} \sum_j \left[\sum_i \frac{\partial (\mathbf{D}^{-1})_{\ell j}}{\partial x_i} \dot{x}_i (\dot{x}_j - F_j) + (\mathbf{D}^{-1})_{\ell j} \left(\ddot{x}_j - \sum_i \frac{\partial F_j}{\partial x_i} \dot{x}_i \right) \right]. \end{aligned} \quad \text{[S19]}$$

Therefore, we have the equation of motion

$$\begin{aligned} \sum_j (\mathbf{D}^{-1})_{\ell j} \ddot{x}_j &= -\sum_{i,j} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) \\ &\quad + \frac{1}{2} \sum_{i,j} (\dot{x}_i - F_i) \frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} (\dot{x}_j - F_j) + \sum_i \frac{\partial F_i}{\partial x_i \partial x_\ell} \\ &\quad - \sum_j \left[\sum_i \frac{\partial (\mathbf{D}^{-1})_{\ell j}}{\partial x_i} \dot{x}_i (\dot{x}_j - F_j) - (\mathbf{D}^{-1})_{\ell j} \sum_i \frac{\partial F_j}{\partial x_i} \dot{x}_i \right], \end{aligned} \quad \text{[S20]}$$

$$\begin{aligned} \ddot{x}_k &= -\sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) \\ &\quad + \frac{1}{2} \sum_{i,j,\ell} \mathbf{D}_{k\ell} (\dot{x}_i - F_i) \frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} (\dot{x}_j - F_j) \\ &\quad - \sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial (\mathbf{D}^{-1})_{\ell j}}{\partial x_i} \dot{x}_i (\dot{x}_j - F_j) + \sum_i \frac{\partial F_k}{\partial x_i} \dot{x}_i + \sum_{i,l} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_i \partial x_\ell}, \end{aligned} \quad \text{[S21]}$$

$$\begin{aligned} &\ddot{x}_k - \frac{1}{2} \sum_{i,j,\ell} \mathbf{D}_{k\ell} \dot{x}_i \left(\frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} - 2 \frac{\partial (\mathbf{D}^{-1})_{\ell j}}{\partial x_i} \right) \dot{x}_j \\ &= -\sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) \\ &\quad + \frac{1}{2} \sum_{i,j,\ell} \mathbf{D}_{k\ell} (F_i F_j - F_i \dot{x}_j - F_j \dot{x}_i) \frac{\partial (\mathbf{D}^{-1})_{ij}}{\partial x_\ell} + \sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial (\mathbf{D}^{-1})_{\ell j}}{\partial x_i} \dot{x}_i F_j \\ &\quad + \sum_i \frac{\partial F_k}{\partial x_i} \dot{x}_i + \sum_{i,l} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_i \partial x_\ell}. \end{aligned} \quad \text{[S22]}$$

If $\mathbf{D}$ is a constant matrix independent of $\mathbf{x}$, then we have

$$\begin{aligned} \ddot{x}_k &= -\sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) + \sum_i \frac{\partial F_k}{\partial x_i} \dot{x}_i + \sum_{i,l} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_i \partial x_\ell} \\ &= -\sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} (\dot{x}_j - F_j) + \sum_i \frac{\partial F_k}{\partial x_i} \dot{x}_i + \sum_{i,l} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_i \partial x_\ell} \\ &= \sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} F_j - \sum_{i,j,\ell} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_\ell} (\mathbf{D}^{-1})_{ij} \dot{x}_j + \sum_i \frac{\partial F_k}{\partial x_i} \dot{x}_i \\ &\quad + \sum_{i,l} \mathbf{D}_{k\ell} \frac{\partial F_i}{\partial x_i \partial x_\ell}. \end{aligned} \quad \text{[S23]}$$

When a system is three dimensional and $\mathbf{D}$ is diagonal matrix, then,

$$\begin{aligned} \ddot{x}_1 &= \sum_i \mathbf{D}_{11} \frac{\partial F_i}{\partial x_1} \mathbf{D}_{ii}^{-1} F_i - \sum_i \mathbf{D}_{11} \frac{\partial F_i}{\partial x_1} \mathbf{D}_{ii}^{-1} \dot{x}_i + \sum_i \frac{\partial F_1}{\partial x_i} \dot{x}_i \\ &\quad + \sum_i \mathbf{D}_{11} \frac{\partial F_i}{\partial x_i \partial x_1} \\ &= \frac{\mathbf{D}_{11}}{2} \frac{\partial}{\partial x_1} \sum_i \left(\mathbf{D}_{ii}^{-1} F_i^2 + 2 \frac{\partial F_i}{\partial x_i} \right) + \left(\frac{\partial F_1}{\partial x_2} - \frac{\mathbf{D}_{11}}{\mathbf{D}_{22}} \frac{\partial F_2}{\partial x_1} \right) \dot{x}_2 \\ &\quad + \left(\frac{\partial F_1}{\partial x_3} - \frac{\mathbf{D}_{11}}{\mathbf{D}_{33}} \frac{\partial F_3}{\partial x_1} \right) \dot{x}_3, \end{aligned} \quad \text{[S24]}$$

$$\ddot{x}_2 = \sum_i \mathbf{D}_{22}\frac{\partial F_i}{\partial x_2}\mathbf{D}_{ii}^{-1}F_i - \sum_i \mathbf{D}_{22}\frac{\partial F_i}{\partial x_2}\mathbf{D}_{ii}^{-1}\dot{x}_i + \sum_i \frac{\partial F_2}{\partial x_i}\dot{x}_i + \sum_i \mathbf{D}_{22}\frac{\partial F_i}{\partial x_i \partial x_2}$$
$$= \frac{\mathbf{D}_{22}}{2}\frac{\partial}{\partial x_2}\sum_i\left(\mathbf{D}_{ii}^{-1}F_i^2 + 2\frac{\partial F_i}{\partial x_i}\right) + \left(\frac{\partial F_2}{\partial x_1} - \frac{\mathbf{D}_{22}}{\mathbf{D}_{11}}\frac{\partial F_1}{\partial x_2}\right)\dot{x}_1 + \left(\frac{\partial F_2}{\partial x_3} - \frac{\mathbf{D}_{22}}{\mathbf{D}_{33}}\frac{\partial F_3}{\partial x_2}\right)\dot{x}_3, \quad [S25]$$

$$\ddot{x}_3 = \sum_i \mathbf{D}_{33}\frac{\partial F_i}{\partial x_3}\mathbf{D}_{ii}^{-1}F_i - \sum_i \mathbf{D}_{33}\frac{\partial F_i}{\partial x_3}\mathbf{D}_{ii}^{-1}\dot{x}_i + \sum_i \frac{\partial F_3}{\partial x_i}\dot{x}_i + \sum_i \mathbf{D}_{33}\frac{\partial F_i}{\partial x_i \partial x_3}$$
$$= \frac{\mathbf{D}_{33}}{2}\frac{\partial}{\partial x_3}\sum_i\left(\mathbf{D}_{ii}^{-1}F_i^2 + 2\frac{\partial F_i}{\partial x_i}\right) + \left(\frac{\partial F_3}{\partial x_1} - \frac{\mathbf{D}_{33}}{\mathbf{D}_{11}}\frac{\partial F_1}{\partial x_3}\right)\dot{x}_1 + \left(\frac{\partial F_3}{\partial x_2} - \frac{\mathbf{D}_{33}}{\mathbf{D}_{22}}\frac{\partial F_2}{\partial x_3}\right)\dot{x}_2. \quad [S26]$$

The Hamiltonian function

$$H = -\mathfrak{L} + P(x)\cdot\dot{x}$$
$$= -\frac{1}{4}\sum_{ij}(\dot{x}_i - F_i)(\mathbf{D}^{-1})_{ij}(\dot{x}_j - F_j) - \frac{1}{2}\sum_i \frac{\partial F_i}{\partial x_i} + \frac{1}{2}\sum_{ij}(\mathbf{D}^{-1})_{ij}(\dot{x}_i - F_i)\dot{x}_j$$
$$= \frac{1}{4}\sum_{ij}\dot{x}_i(\mathbf{D}^{-1})_{ij}\dot{x}_j - \frac{1}{4}\sum_{ij}F_i(\mathbf{D}^{-1})_{ij}F_j - \frac{1}{2}\sum_i\frac{\partial F_i}{\partial x_i}. \quad [S27]$$

Hamilton–Jacobian Approach to Path Integral. If D is a constant diagonal matrix independent of x, then

$$H = \frac{1}{4}\sum_i \mathbf{D}_{ii}^{-1}\dot{x}_i^2 - \sum_i\left(\frac{1}{4}\mathbf{D}_{ii}^{-1}F_i^2 + \frac{1}{2}\frac{\partial F_i}{\partial x_i}\right) = E_{\text{eff}}. \quad [S28]$$

We defined

$$\mathbf{x}'_{\mathbf{k}} = \sqrt{\mathbf{D}_{\mathbf{kk}}^{-1}}\mathbf{x}_{\mathbf{k}} \quad [S29]$$

$$V = \sum_i\left(\frac{1}{4}\mathbf{D}_{ii}^{-1}F_i^2 + \frac{1}{2}\frac{\partial F_i}{\partial x_i}\right), \quad [S30]$$

then

$$\frac{1}{4}\sum_i \mathbf{D}_{ii}^{-1}\dot{x}_i^2 - V = E_{\text{eff}} \quad [S31]$$

$$|\mathbf{x}'| = \sqrt{\mathbf{4}(\mathbf{V} + \mathbf{E}_{\mathbf{eff}})}. \quad [S32]$$

If we allow a variation of the final time T and leave the initial coordinate and final coordinate fixed, then we have

$$\delta S = -H\delta t. \quad [S33]$$

For constant effective Hamiltonian, $\delta S = -E\delta t$, because $S(x) = \int(\mathbf{P}(\mathbf{x})\cdot\dot{x} - \mathbf{H}(\mathbf{x}))\mathbf{dt}$, we can define $S_0 = \int(\mathbf{P}(\mathbf{x})\cdot\dot{\mathbf{x}}\mathbf{dt}$. We find that $\delta S_0 = 0$.

In the Hamiltonian–Jacobian framework, we notice that we can transform the formulations into a different representation in x space:

$$S_0 = S_{\text{HJ}}(x) = \int\sum_i\frac{1}{2}\mathbf{D}_{ii}^{-1}(\dot{x}_i - F_i)dx_i = \int\sum_i\frac{1}{2}\dot{x}'_i dx'_i - \int\frac{1}{2}\sum_i\mathbf{D}_{ii}^{-1}F_i dx_i = \int\sqrt{V + E_{\text{eff}}}dl' - \frac{1}{2}\int F'_l dl. \quad [S34]$$

The optimal path was obtained by minimizing the discretized target function

$$S_{\text{HJ}} = \sum_n^{N-1}\sqrt{(E_{\text{eff}} + V(n))}\,\Delta l'_{n,n+1} - \frac{1}{2}F'_l(n)\Delta l_{n,n+1} + \lambda P. \quad [S35]$$

Where $P = \sum_i^{N-1}(\Delta l_{i,i+1} - \langle\Delta l\rangle)^2$ and

$$(\Delta l)^2_{n,n+1} = \sum_i(\mathbf{x}_i(n+1) - \mathbf{x}_i(n))^2$$
$$(\Delta l')^2_{n,n+1} = \sum_i\mathbf{D}_{ii}^{-1}(\mathbf{x}_i(n+1) - \mathbf{x}_i(n))^2$$
$$F'_l(n) = \sum_i\mathbf{D}_{ii}^{-1}\mathbf{F}_i(\mathbf{x}(n))(\mathbf{x}_i(n+1) - \mathbf{x}_i(n))/\Delta l_{n,n+1} \quad [S36]$$
$$V(n) = \sum_i\left\{\frac{1}{4}\mathbf{D}_{ii}^{-1}\mathbf{F}^2(\mathbf{x}_i) + \frac{1}{2}\sum_j\frac{\partial\mathbf{F}_j(\mathbf{x}_i)}{\partial x_j}\right\}.$$

$\Delta l_{n,n+1}$ is the Euclidean measure of the nth elementary path step, and P is a penalty function that keeps all the length elements close to their average (18–21) and becomes irrelevant in the continuum limit.

Pathways and Transition Rates from One Basin of Attraction to Another: Instantons and Antiinstantons. Observing Eq. **S33**, we find it is an energy conservation equation with kinetic energy term ($\frac{1}{4}\sum_i\mathbf{D}_{ii}^{-1}\dot{x}_i^2$) and the potential ($U = -V$), and the total energy (E_{eff}). The system have certain similarities with a particle moving in a potential U. In the long time limit, presenting the possible path go back and forth many times in the effective potential U. The path corresponds to the oscillating solutions that are called the instanton solutions (23–26, 13, 14). Each instanton (antiinstanton) shows the one transition from a steady state or basin of attraction to other one (see Fig. S3). To obtain the total contribution to the probability of the transition, we assume no instanton–instanton interactions, then the contribution can be summed in the dilute gas approximation. The contribution of one instanton to the weight is written as

$$R = R_0\exp[-\gamma] = R_0\exp[-\int(L(x_s(t)) - L(x_{\min}))dt]. \quad [S37]$$

Where R_0 is a constant and x_s is the instanton path ($x_s(t)$) and the integral is from the beginning of one instanton at $t = 0$ and at $x_{\min}$ to the end of one instanton at the bounce back point $x_{\max}$ and at

$t_{\max}$. $x_{\min}$ and $x_{\max}$ are approximately equal to the minimum of V near the fixed points, where the value of V at the bounce back point is that of the minimum of V. The probability is calculated from the optimal paths contributed by the sum of the multi instanton contribution. It can be given by

$$P(t) \sim C\sum_{n=0}^{\infty} R_{\text{ins}}^n R_{\text{ant}}^n \int dt_1 \int dt_2 \ldots \int dt_{2n} \exp[-V(x_{\min})((t_1 - t_0) + (t_3 - t_2) + \ldots + (t - t_{2n}))] \exp[-V(x_{\max})((t_2 - t_1) + (t_4 - t_3) + \ldots + (t_{2n} - t_{2n-1}))]. \quad \text{[S38]}$$

Where $t_1 - t_0, t_3 - t_2 \ldots t - t_{2n}$ are the time intervals staying at the position of the minimum of V near a steady states. $t_2 - t_1, t_4 - t_3 \ldots t_{2n} - t_{2n-1}$ are the time intervals staying at the another minimum of V. The R_{ins} or R_{ant} is the one-instanton or antiinstanton contribution to the probability $P(t)$. Considering $R_{\text{ins}}R_{\text{ant}}$ in the expression, the contribution to the probability for the trajectory is from one minimum to another minimum of V and back. This brings one instanton and one antiinstanton. n in the equation is the number of times the trajectory is traversing from one minimum of $V(x_{\min})$ to another $V(x_{\max})$ and back.

we can easily transform the above expression into the Laplace representation s:

$$P(s) = C\sum_{n=0}^{\infty} R_{\text{ins}}^n R_{\text{ant}}^n (s + V(x_{\min}))^{-n-1} (s + V(x_{\max}))^{-n} = C(s + V(x_{\max})) \frac{1}{(s + V(x_{\min}))(s + V(x_{\max})) - R_{\text{ins}}R_{\text{ant}}}. \quad \text{[S39]}$$

By the Laplace transformation, we obtain

$$P(t) = C_+ \exp[\lambda_+ t] + C_- \exp[\lambda_- t], \quad \text{[S40]}$$

where λ_+ and λ_- are written as

$$\lambda_\pm = -\frac{1}{2}(V(x_{\min}) + V(x_{\max})) \pm \frac{1}{2}\sqrt{(V(x_{\min}) - V(x_{\max}))^2 + 4R_{\text{ins}}R_{\text{ant}}}. \quad \text{[S41]}$$

When $V(x_{\min}) = V(x_{\max})$ as is the case in instantons, the equation is shown as

$$\lambda_\pm = -V(x_{\min}) \pm R. \quad \text{[S42]}$$

Because the $\exp[-V(x_{\min})t]$ term express the probability of staying in one of the minimum. The real transition frequency from one minimum to the other is constrained by R ($\exp[-V(x_{\min})t]$ term is normalized out). Thus the kinetic rate k is determined by R ($k = R$) and $R = \sqrt{R_{\text{ins}}R_{\text{ant}}}$ (13, 14, 21). Notice that in the equilibrium situation, the instantons and antiinstantons have the same contribution to the transition rate from A to B and B to A because they share the same path $R = R_{\text{ins}} = R_{\text{ant}}$. On the other hand, for nonequilibrium system, the instantons from basin A to B and antiinstantons from B to A have different paths (no reversal symmetry). The instanton contribution to the transition rate from A to B R_{ins} and antiinstanton contribution to the transition rate from B to A R_{ant} are not equal. This means the transition rate from basin of attraction A to B is in general different from transition rate of the reverse direction from basin of attraction B to A.

We would like to point out that our work provides the origin and underlying mechanism of the kinetic paths deviating from the steepest descent paths: the curl flux originated from breaking the detailed balance (21), whereas it was previously unknown. Furthermore, the path integral formulation we developed is a more complete one including the Jacobian transformation term (divergence of the force) that is often neglected by the previous approaches (27, 28). This term can be ignored if the diffusion term characterizing either the intrinsic or external fluctuations is small but can contribute significantly when the diffusion term becomes large.

Waddington Landscape With Fast Development Relative to Gene Regulations. When the dynamics of developmental process is comparable or faster than inherent gene regulation time scales, the landscape as shown in Fig. S4 here preserves the main features as the case when the developmental process is relatively slow compared with the regulations as shown in Fig. 5 of the main text, but gives somewhat different qualitative and quantitative pictures. The development process still proceeds in a downhill fashion (in the direction of a). However, the central undifferentiated basin becomes unstable (as compared to a relatively stable basin in the slow developmental process relative to the gene regulations) and bifurcates toward the two differentiated basins later on in the development.

1. Huang S, Guo YP, May G, Enver T (2007) Bifurcation dynamics of cell fate decision in bipotent progenitor cells. *Dev Biol* 305:695–713.
2. McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA* 94:814–819.
3. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338.
4. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* 99:12795–12800.
5. Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci USA* 98:8614–8619.
6. Vilar JMG, Guet CC, Leibler S (2003) Modeling network dynamics: The Lac operon, a case study. *J Cell Biol* 161:471–476.
7. Paulsson J (2004) Summing up the noise in gene networks. *Nature* 427:415–418.
8. Onsager L, Machlup S (1953) Fluctuations and irreversible processes. *Phys Rev* 91:1505–1512.
9. Wiener N (1964) *Generalized Harmonic Analysis and Tauberian Theorems* (MIT Press, Cambridge, MA).
10. Feynman RP, Hibbs AR (1965) *Quantum Mechanics and Path Integrals* (McGraw–Hill, New York).
11. Hunt KLC, Ross J (1981) Path integral solutions of stochastic equations for nonlinear irreversible processes: The uniqueness of the thermodynamic Lagrangian. *J Chem Phys* 75:976–984.
12. Wang J, Onuchic J, Wolynes P (1996) Statistics of kinetic pathways on biased rough energy landscapes with applications to protein folding. *Phys Rev Lett* 76:4861–4864.
13. Wang J, Wolynes PG (1996) Instantons and the fluctuating path description of reactions in complex environments. *J Phys Chem* 100:1129–1136.
14. Wang J, Zhang K, Lu HY, Wang EK (2005) Quantifying kinetic paths of protein folding. *Biophys J* 89:1612–1620.
15. Wang J, Zhang K, Lu HY, Wang EK (2006) Identifying the kinetic paths on biomolecular binding-folding energy landscape. *Phys Rev Lett* 96(168101):1–4.
16. Wang J, Zhang K, Lu HY, Wang EK (2006) Quantifying the kinetic paths of flexible biomolecular recognition. *Biophys J* 91:866–872.
17. Wang J, Xu L, Wang E (2008) Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc Natl Acad Sci USA* 105:12271–12276.
18. Olender R, Elber R (1996) Calculation of classical trajectories with a very large time step: Formalism and numerical examples. *J Chem Phys* 105:9299–9315.
19. Elber R, Meller J, Olender R (1999) Stochastic path approach to compute atomically detailed trajectories: Application to the folding of c peptide. *J Phys Chem B* 103:899–911.
20. Faccioli P, Sega M, Pederiva F, and Orland H (2006) Dominant pathways in protein folding. *Phys Rev Lett* 97(108101):1–4.
21. 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.
22. Qian H (2001) Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Phys Rev E* 65:0161021–0161025.
23. Langer JS (1967) Theory of the condensation point. *Ann Phys* 41:108–157.
24. Coleman, S (1977) Fate of the false vacuum: Semiclassical theory. *Phys Rev D* 15:2929–2936.
25. Callen CG, Coleman S (1977) Fate of the false vacuum. II. First quantum corrections. *Phys Rev D* 16:1762–1768.

26. Jalickee JB, Wiegel FW, Vezzetti DJ (1971) Role of droplets and bubbles in the "condensation" of one-dimensional gas of Kac, Uhlenbeck, and Hemmer. *Phys Fluids* 14:1041–1048.
27. Weinan E, Vanden–Eijnden E (2010) Transition-path theory and path-finding algorithms for the study of rare events. *Annu Rev Phys Chem* 61:391–420.
28. Roma DM, Flanagan RA, Ruckenstein AE, Sengupta AM, Mukhopadhyay R (2005) Optimal path to epigentic switching. *Phys Rev E* 71:0119021–0119025.

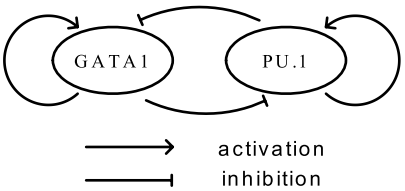

Fig. S1. The illustration of cell fate decision circuit for development.

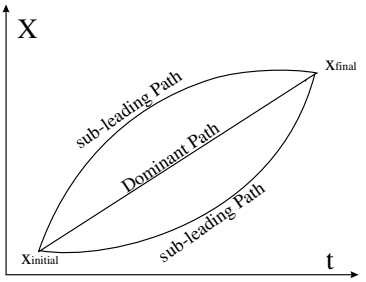

Fig. S2. Dominant path and subleading path.

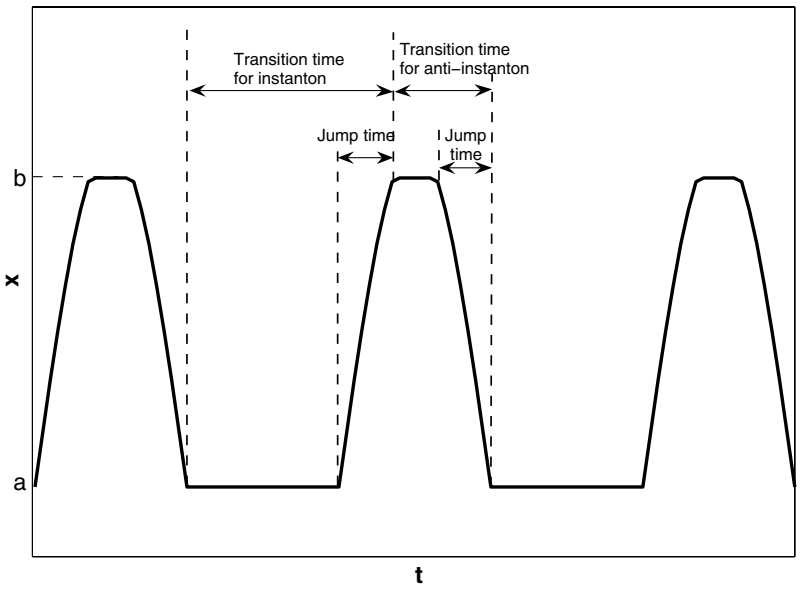

Fig. S3. Instantons and antiinstantons.

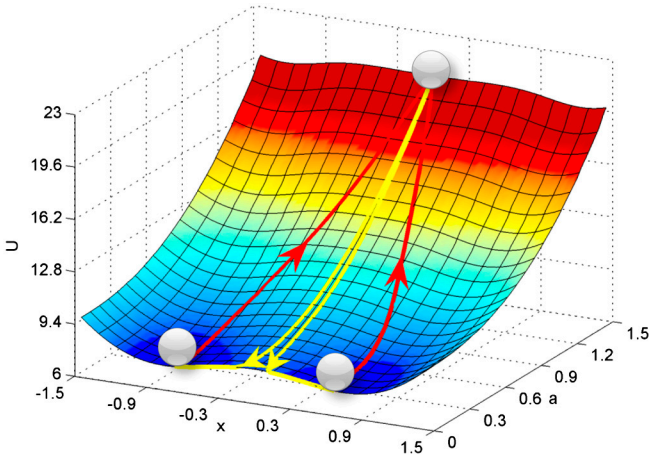

Fig. S4. The quantified Waddington developmental landscape and pathways when the developmental process is comparable or relatively fast compared with the gene regulation dynamics. ($a = a_1 = a_2$, $b_1 = b_2 = 1$, $k_1 = k_2 = 1$, $S = 0.5$, $n = 4$, and $\lambda = 1$).