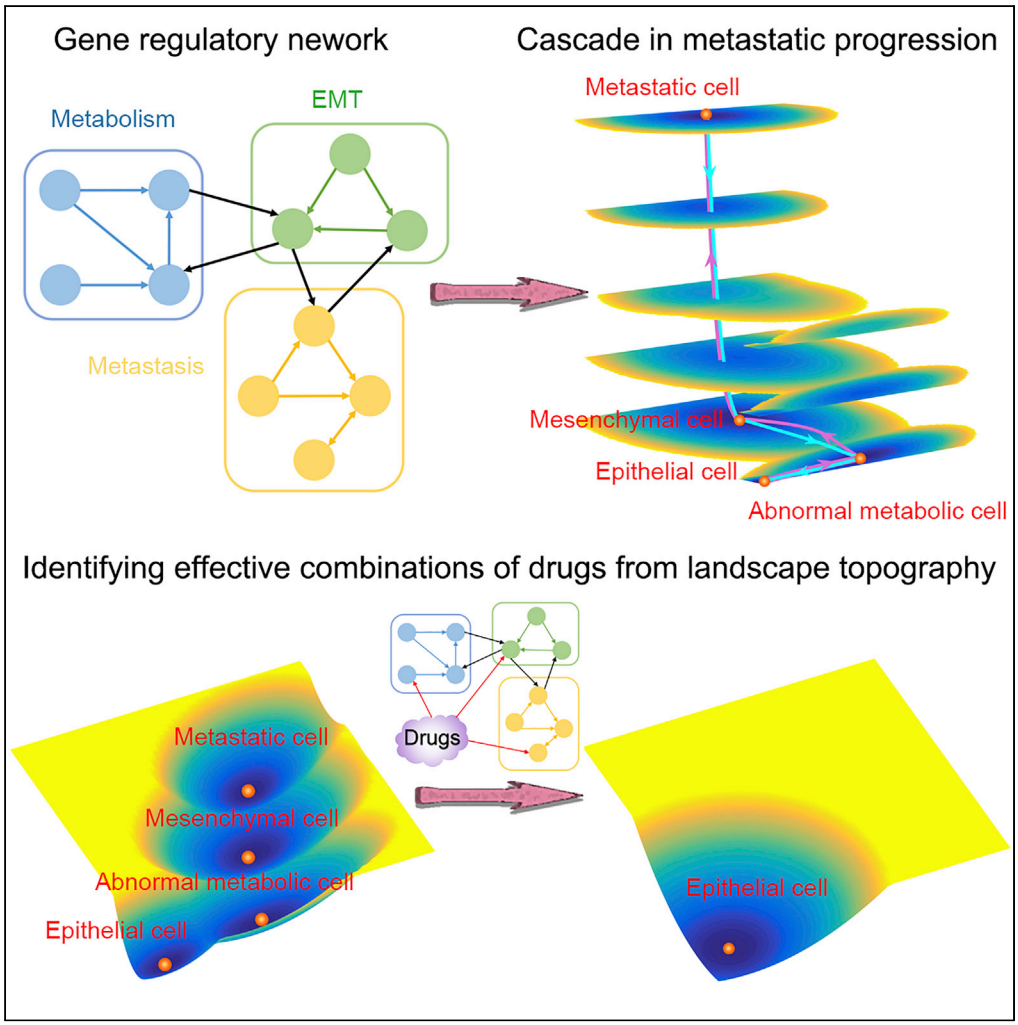
iScience

CellPress

Article

Exposing the Underlying Relationship of Cancer Metastasis to Metabolism and Epithelial-Mesenchymal Transitions



Xin Kang, Jin Wang, Chunhe Li

jin.wang.1@stonybrook.edu (J.W.)
chunheli@fudan.edu.cn (C.L.)

HIGHLIGHTS

We develop an approach to quantify the interplay among metabolism, EMT, and metastasis

An intermediate abnormal metabolic state exists in metastatic progression

Landscape and path indicate that the metastasis is a sequential process

Landscape suggests a potential therapeutic strategy by targeting BACH1 and metabolism

Kang et al., iScience *21*, 754–772
November 22, 2019
https://doi.org/10.1016/j.isci.2019.10.060

iScience

CellPress

Article

Exposing the Underlying Relationship of Cancer Metastasis to Metabolism and Epithelial-Mesenchymal Transitions

Xin Kang,[1,2] Jin Wang,[3,*] and Chunhe Li[1,4,5,*]

SUMMARY

Cancer is a disease governed by the underlying gene regulatory networks. The hallmarks of cancer have been proposed to characterize the cancerization, e.g., abnormal metabolism, epithelial to mesenchymal transition (EMT), and cancer metastasis. We constructed a metabolism-EMT-metastasis regulatory network and quantified its underlying landscape. We identified four attractors, characterizing epithelial, abnormal metabolic, mesenchymal, and metastatic cell states, respectively. Importantly, we identified an abnormal metabolic state. Based on the transition path theory, we quantified the kinetic transition paths among these different cell states. Our results for landscape and paths indicate that metastasis is a sequential process: cells tend to first change their metabolism, then activate the EMT and eventually reach the metastatic state. This demonstrates the importance of the temporal order for different gene circuits switching on or off during metastatic progression of cancer cells and underlines the cascading regulation of metastasis through an abnormal metabolic intermediate state.

INTRODUCTION

Cancer is a disease involving the changes in underlying gene regulatory networks. Lately, some hallmarks of cancer have been proposed to characterize the cancerization (Hanahan and Weinberg, 2000, 2011). For example, abnormal metabolism is a hallmark of cancer, which has been explored in recent studies (Hanahan and Weinberg, 2011; Boroughs and Deberardinis, 2015). Metastasis has been suggested to account for the great majority of cancer-related deaths, although its underlying mechanism remains elusive (Lambert et al., 2017). Epithelial to mesenchymal transition (EMT) has long been shown to be related to the acquisition of malignant cell traits, such as motility, invasiveness, and tumor-initiating potential, and therefore being associated with the progression of cancer metastasis (Brabletz et al., 2018; Lambert et al., 2017; Gibbons et al., 2009; Thiery et al., 2009). A critical question arises as to what the connections among these hallmarks of cancers are. For example, what is the relationship between EMT and cancer metabolism? And how do they contribute to the cancer metastasis, the most fatal stage of tumors?

Cancer cells have abnormal metabolism compared with normal cells. In the low oxygen and oxygen-free conditions, metabolisms of cells are mainly in glycolysis form that produces less ATP. However, under normoxic conditions, the metabolisms for cancer cells and normal cells are different. Cancer cells often use glycolysis to generate energy, whereas normal cells use glucose for oxidative phosphorylation (OXPHOS). This phenomenon is called Warburg effect (Munozpinedo et al., 2012). Recently, Yu et al. proposed a mathematical model to study the regulations for genes and metabolites on cancer metabolism (Yu et al., 2017), which has been extended to a more complete model in the following work (Jia et al., 2019). However, the global stability and stochastic dynamics of cancer metabolism remain to be elucidated. More importantly, the connection between metabolism and metastasis remains unclear. These issues are critical to the mechanistic understanding of tumorigenesis and have not been fully clarified from previous studies.

Here, we aim to apply the landscape approach to address the above-mentioned issues by investigating the stochastic dynamics of underlying gene regulatory networks for cancer metastasis. The classic Waddington landscape has been proposed as a metaphor to explain the development and differentiation of cells (Waddington, 1957). Recently, the epigenetic landscapes for biological networks have been quantified from various approaches (Wang et al., 2011; Liao and Lu, 2013; Lv et al., 2015; Li and Wang, 2014a; Lu et al., 2014b; Ge and Qian, 2012, 2016; Zhang and Wolynes, 2014; Huang et al., 2005) and employed to investigate the stochastic dynamics of embryonic development and cancer (Wang, 2015; Wang et al., 2011; Li and

[1]Shanghai Center for Mathematical Sciences, Fudan University, Shanghai 200438, China

[2]School of Mathematical Sciences, Fudan University, Shanghai 200433, China

[3]Department of Chemistry and Physics, State University of New York at Stony Brook, Stony Brook, NY 11794, USA

[4]Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China

[5]Lead Contact

*Correspondence: jin.wang.1@stonybrook.edu (J.W.), chunheli@fudan.edu.cn (C.L.)

https://doi.org/10.1016/j.isci.2019.10.060

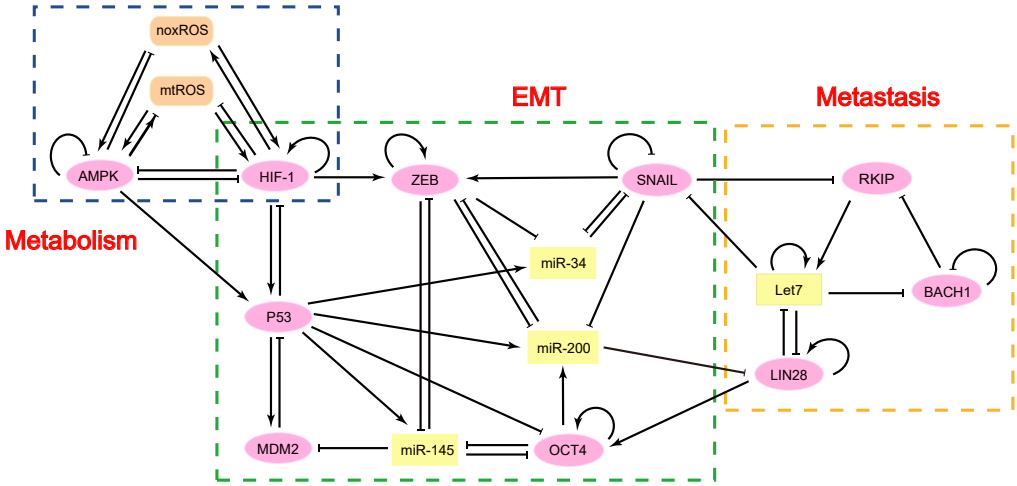


Figure 1. The Regulatory Network for the Interplay among EMT, Metabolism, and Cancer Metastasis including 16 Gene or Metabolite Nodes and 51 Regulation Links (22 Activations and 29 Inhibitions)

The network from left to right corresponds to the core circuit of metabolism, EMT, and cancer metastasis, respectively. The orange nodes, yellow nodes, and pink nodes represent metabolites, microRNAs, and genes, respectively. The arrows represent activations, and the short bars represent inhibitions.

Wang, 2013, 2014b; Chen et al., 2014, 2015; Huang, 2013; Li, 2017). From the landscape perspective, different cell types are described as the basins of attraction on a potential surface.

To uncover the mechanisms of the interplay among metabolism, EMT, and cancer metastasis, we constructed a metabolism-EMT-metastasis network by mining the experimental literature and combining these three gene regulatory circuits (Yu et al., 2017; Li and Wang, 2015; Lu et al., 2014a; Jolly et al., 2015b; Lee et al., 2014). The landscape for the metabolism-EMT-metastasis network displays four stable cell types quantified by attractors, characterizing epithelial cell state (E), abnormal metabolic cell state (A), mesenchymal cell state (M), and metastatic cell state (Met), respectively. Importantly, we identified an abnormal metabolic state. We also calculated the minimum action paths (MAPs) to quantify the most probable transition paths for cell fate decision processes in metastatic progression. The MAPs among the four cell states quantify the dynamical processes for how cells switch from the epithelial state to the metastatic state: cells need to first change their metabolism, then finish the EMT and finally enter the metastatic state. This demonstrates the importance of the temporal order for different gene circuits switching on or off in cell fate decision process. We found that the transition path for metastatic process and the transition path for anti-metastatic process are irreversible. The transition actions among attractors are correlated with the potential barriers for basins but provide more accurate descriptions for kinetic switching. By the global sensitivity analysis for the metabolism-EMT-metastasis network based on the transition actions, we identified some critical network elements governing the dynamics of cellular metabolism and metastasis.

To further shed light on the interplay specifically between metabolism and EMT, we explored a subnetwork of the whole network, i.e., an EMT-metabolism regulatory network (blue and green boxes in Figure 1), by neglecting the metastasis circuit (yellow box in Figure 1) and further quantified its potential landscape. The landscape for the EMT-metabolism network displays three stable states (attractors) including epithelial (E) state, mesenchymal (M) state and abnormal metabolic (A) state. To further verify our modeling results, we compared the gene expression data of three types of cancers (one for single-cell RNA sequencing [RNA-seq] data and two for population data) and the landscape, which supports the consistency between the gene expression data and the landscape results (Pastushenko et al., 2018; Cancer Genome Atlas Research Network et al., 2013; Agrawal et al., 2014).

Previously we have developed a partial self-consistent approximation (PSCA) approach to calculate the potential landscape for high-dimensional gene regulatory systems (Li and Wang, 2013, 2014a). One limitation of the PSCA approach is that it assumes weak correlations among gene variables (gene expression levels). This approach may not be accurate for the cases in which variables have strong correlations. Here, to

resolve this issue, we developed a truncated moment equations (TME) approach. This approach does not rely on assumptions of weak correlations between variables. By comparing the results of both the TME method and PSCA method with the simulation results, we showed that the TME approach provides a more precise description for the probability distribution of the system than that of the PSCA approach. In summary, the TME approach offers a new method to quantify the potential landscape for a high-dimensional gene regulatory system, and the results of landscape and kinetic paths provide new insights into the mechanistic understanding of the interplay among metabolism, EMT, and cancer metastasis in the process of cancer progression.

RESULTS

Regulatory Network Models for Metabolism, EMT, and Cancer Metastasis

Recently, Yu et al. constructed a computational model for cancer metabolism using ordinary differential equations (ODEs), which includes two genes (AMPK and HIF-1) and two metabolites (noxROS and mtROS) (Yu et al., 2017). The EMT has been suggested to be related to metastatic progression (Brabletz et al., 2018; Lambert et al., 2017; Gibbons et al., 2009). So, a critical question would be, are there any connections among EMT, metabolism, and cancer metastasis? To uncover the mechanisms underlying the interplay between EMT, metabolism, and metastasis, we established a metabolism-EMT-metastasis gene regulatory network by incorporating the core components for each process through mining the experimental literature (Figure 1). The metabolism-EMT-metastasis network involves 16 components, including ten genes, two metabolites, and four microRNAs. AMPK, HIF-1, mtROS, and noxROS are the core components controlling cellular metabolism (Figure 1, blue box). SNAIL, ZEB, OCT4, MDM2, miR-145, miR-200, miR-34, and P53 are the core components governing the EMT (Figure 1, green box). RKIP, BACH1, LIN28, and Let7 are the core components governing the metastasis (Figure 1, yellow box) (Lee et al., 2014). The experimental evidences for the interactions among different components are shown in Table S1.

Based on the network structure, we wrote down the ODEs describing the time evolution of relative expression levels for each of the 16 genes or metabolites. The ODEs depicting the dynamic evolution of the metabolism-EMT-metastasis system are shown in Equation 1:

$$\frac{dX}{dt} = g_X \cdot G - k_X \cdot K \cdot X \qquad \text{(Equation 1)}$$

Here, X represents the level of the gene expression or the metabolite. g_X and k_X denote the basal synthesis rate and degradation rate of X, respectively. G and K denote the regulation of other genes or metabolites on the synthesis and the degradation of X, respectively. The activation or inhibition regulations among different components are described by the product of the shifted Hill function $H^s(Y, S, \lambda, n) = 1 + (\lambda - 1)\frac{Y^n}{S^n + Y^n}$ (representing the regulation of component Y on component X). Here, λ represents the fold change for the regulations (λ>1 for activation and 0<λ<1 for inhibition). S represents the threshold of the sigmoidal function, and n is the Hill coefficient, which determines the steepness of the sigmoidal function (Li and Wang, 2013; Li, 2017) (see Supplemental Information Section S1 for more details on how to determine the parameters). The specific ODEs of the metabolism-EMT-metastasis model are shown in Supplemental Information Section S2 and corresponding parameter values are shown in Tables S2 and S3.

Landscape Reveals the Cascade for Metabolism, EMT, and Metastasis

Yu's work provides a basis for understanding the cancer metabolism from a deterministic ODE model (Yu et al., 2017). However, the stochastic dynamics and global properties of cancer metabolism remain to be elucidated, as the fluctuations have been suggested to play critical roles in biological systems, including intrinsic fluctuations from a finite number of molecules and external fluctuations from highly inhomogeneous environments (Swain et al., 2002; Kaern et al., 2005; Thattai and Van, 2001). Also, it is crucial to uncover the dynamical mechanisms for the interplay among metabolism, EMT, and cancer metastasis.

Previously, we have developed a partial self-consistent approximation approach to study the stochastic dynamics for high-dimensional systems by the potential landscape theory (Li and Wang, 2013, 2014a). One limitation of the PSCA approach is that it assumes the weak correlation between variables, which in some cases may not be accurate. Here, we improved previous methods and developed a Truncated Moment Equations (TME) approach (see Supplemental Information Section S3 and Section S4) to calculate the probability distribution of gene expression levels and obtain the potential landscape. By comparing the

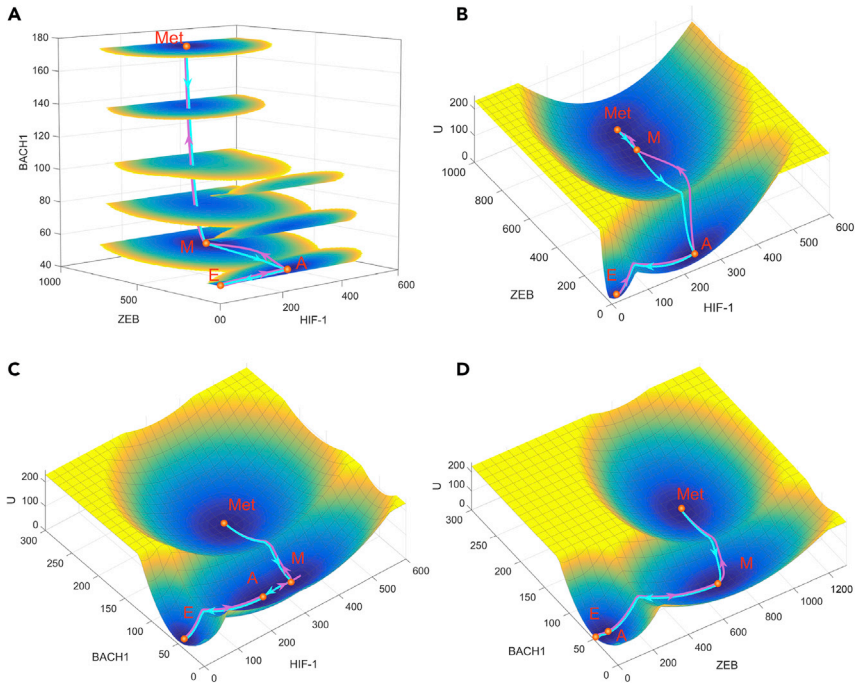


Figure 2. Landscape and Path for the Metabolism-EMT-Metastasis Model Shown in ZEB, HIF-1, and BACH1 Coordinates

(A) Landscape is shown in a four-dimensional picture. The blue regions represent higher probability or lower potential, and the yellow regions indicate lower probability or higher potential. Solid magenta lines represent transition paths from the E to A, M, and Met states. Solid cyan lines represent transition paths from the Met to M, A, and E states.

(B) Two-dimensional landscape and kinetic paths are displayed in HIF-1/ZEB coordinates.

(C) Two-dimensional landscape and kinetic paths are displayed in HIF-1/BACH1 coordinates.

(D) Two-dimensional landscape and kinetic paths are displayed in ZEB/BACH1 coordinates. E, epithelial state; A, abnormal metabolic state; M, mesenchymal state; Met, metastasis state.

See also Figure S4 and Table S4.

two approaches, we showed that the TME approach provides a more precise description for calculating the probability distribution of the systems (see Supplemental Information Section S5, Section S6, Figures S1 and S2 for details).

Based on the metabolism-EMT-metastasis regulatory network model (Figure 1), we calculated the steady state probability distribution of the system employing the TME approach and acquired the potential landscape by $U = -\ln P_{ss}$ (Wang et al., 2008, 2011; Li, 2017; Li and Wang, 2014a). Here P_{ss} represents the steady state probability distribution and U represents the dimensionless potential. Because we are dealing with a 16-dimensional potential landscape, it is hard for visualization. Here we pick three representative marker genes, HIF-1, ZEB, and BACH1, as three coordinates and project the 16-dimensional landscape into the three dimensional space (see Figure S3 and Section S7 for another way to show landscape). In this way, we have a four-dimensional landscape, as displayed in Figure 2A. It is worth noting that our major conclusions do not depend on the specific choice of the coordinates (see Figure S4 for landscapes with other pairs of variables as coordinates) because we also calculated the transition actions among different attractors, based on the 16-dimensional state space. On the landscape the blue region represents high probability or low potential and the yellow region represents low probability or high potential. We identified four stable states on the landscape, which characterize epithelial (E), abnormal metabolism (A), mesenchymal (M), and metastatic (Met) cell states, respectively (Figure 2, see Table S4 for stable state gene expression levels for the quadrastable landscape). Importantly, we identified a new intermediate state, which we defined as the abnormal metabolism (A) state, since it has an increased expression of the glycolysis marker gene HIF-1.

From the network structure of our metabolism-EMT-metastasis regulatory network, each of these three circuits studied here can be multistable, e.g., three states in EMT, three states in metabolism, and two states

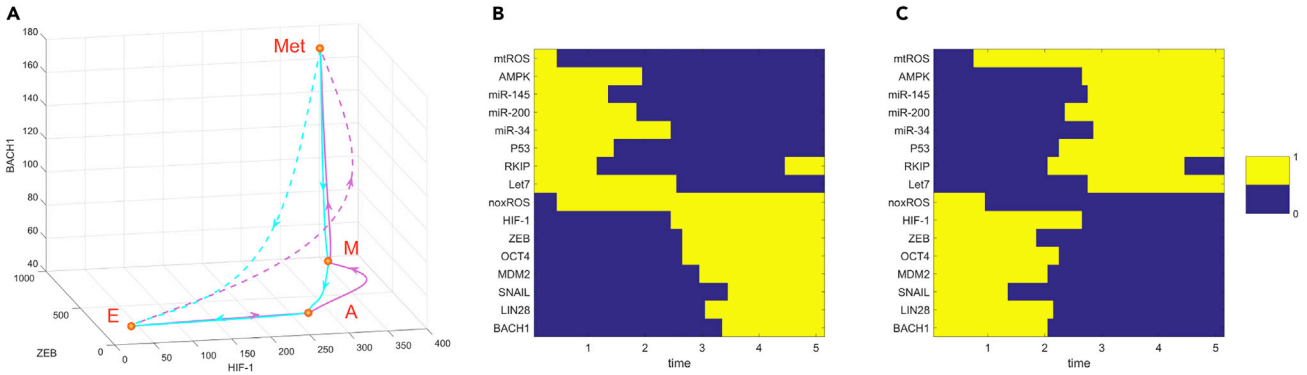


Figure 3. Transition Paths among Different Cell States

(A) Fixed points and kinetic transition paths between different cell states. Solid magenta lines represent transition paths between nearby states (from E to A, A to M, and M to Met state) in metastatic progression direction. Solid cyan lines represent transition paths between nearby states (from Met to M, M to A, and A to E state) in de-metastasis direction. The dashed lines represent the direct transition path from E to Met state and from Met to E state, respectively. (B and C) Discrete transition paths from E state to Met state (B) and from Met state to E state (C) in terms of expression levels of 16 genes and metabolites. Relative gene expression levels are discretized to 0 or 1; 1 represents that the corresponding genes are in the activated state and 0 represents that the corresponding genes are in the repressed state. X axis shows the time points along the transition path. E, epithelial state; A, abnormal metabolic state; M, mesenchymal state; Met, metastasis state.

See also Table S4.

in metastatic circuit. Therefore, if there is no coupling between these three circuits, there are a total of $3\times3\times2=18$ possible states. However, the coupling among these circuits can introduce correlations and thus lead to a much smaller number of stable states, as discussed in previous work (Jolly et al., 2015b; Bocci et al., 2018). We need to stress that, although we identified four stable states in our model, it is possible to discover more cell states (e.g., partial EMT state) by fine-tuning parameters, owing to the complexity involved in EMT and cancer metastasis.

Here, the E state has a low HIF-1, low ZEB, and low BACH1 expression. So, the marker genes for the abnormal metabolism, the EMT, and the metastasis are all off. The A state has a high HIF-1, low ZEB, and low BACH1 expression and therefore corresponds to an abnormal aggressive metabolic phenotype (i.e., aerobic glycolysis state) whereby cancer cells change their metabolism to produce energy more quickly. Besides, we have a mesenchymal state with a high HIF-1, high ZEB, and low BACH1 expression and a metastatic state with a high HIF-1, high ZEB, and high BACH1 expression. Previous work has shown experimentally that the metastatic state has high expression level of BACH1, low expression level of RKIP, and Let7 for breast cancer (Lee et al., 2014). Our metastatic attractor from the model is consistent with this experimental observation (see Table S4). To quantify the kinetic transitions among these states, we calculate the kinetic transition paths among different cell states by minimizing the transition actions $\mathcal{S}$, which are also called minimum action paths (MAPs) (see Table S5 for results of transition actions among four states and Supplemental Information Section S8 for how to obtain the transition paths). Figures 2A and 2B show the landscape and transition paths for the quadrastable system in four-dimensional and three-dimensional space, respectively. The magenta paths denote the transitions from the epithelial state to the metastatic state (metastatic progression process), whereas the cyan paths denote the transitions from the metastatic state to epithelial state (de-metastasis process). The transition paths for the E to metastasis state transition and the backward transition paths from metastasis to the E state are not identical, reflected by the disparity between the forward and backward kinetic transition paths. This irreversibility of MAPs is a consequence of non-gradient force, i.e., curl flux (Wang et al., 2008; Li and Wang, 2014a).

To see the transition path more clearly, we showed both the direct transition path (from E to Met attractor, dashed lines) and indirect transition path (for nearby attractors, solid lines) from minimizing corresponding transition actions (Figure 3A). We found that, for the E to Met transition, the direct transition path (dashed magenta line) is inclined to follow the similar mode as the indirect transition path is (solid magenta line), i.e., the E to Met transition is inclined to go through the attractor states in the middle of the transition. Here, the landscape and the transition paths are both from three-dimensional projection for the whole system (Figures 2 and 3A). To see the paths for more components, we visualized the 16-dimensional transition paths

between the E and the Met states by discretizing the gene expression levels of all 16 genes or metabolites (Figures 3B and 3C). For the metastatic progression process, different genes switch on or off in different orders. For example, microRNAs and tumor repressor P53 are first down-regulated. Then the HIF-1 level is increased, indicating a transition to the abnormal metabolic state, which is followed by the ZEB activation indicating the EMT occurs. Finally, the BACH1 is activated, indicating that cells enter the metastatic state. Therefore, both the continuous and discrete kinetic paths suggest that metastasis is a sequential process, i.e., abnormal metabolism precedes EMT, which precedes metastasis. Interestingly, we found that for both metastatic progression and de-metastasis process (Figures 3B and 3C), both mtROS and noxROS switch in the very early stage in the whole cell fate transition process. This suggests that reactive oxygen species (ROS) plays critical roles in regulating cancer metastasis (Ishikawa et al., 2008), presumably because of its prominent influence on the cellular metabolism.

The quadrastable landscape for metastasis and the kinetic transition paths between attractors demonstrate that the transitions from epithelial state to metastatic state follow a specific order, i.e., first increasing the expression level of HIF-1, then increasing the expression level of ZEB, and finally increasing the expression level of BACH1. These results suggest that the temporal order for different genes switching on or off is critical for the process of metastatic progression in cancer cells. In the first stage, the cells change their metabolism to an abnormal aggressive form (e.g., aerobic glycolysis). In the second stage, cells finish the EMT process where cells obtain mesenchymal features in numerous different settings (Kalluri, 2009). In the last stage, cells improve their metastatic ability by activating metastasis marker genes, such as BACH1. These results agree well with the observations that EMT induction in cancer cells results in the acquisition of invasive and metastatic properties (Singh and Settleman, 2010).

Global Sensitivity Analysis Identifies the Key Players for the Interplay among Metabolism, EMT, and Metastasis

To quantify the topography of landscape, we define the barrier height as the potential difference between the local minimum and the corresponding saddle point. To evaluate the relationship between the transition action and the barrier height, we change the λ_{hh} (the fold change for the self-activation of HIF-1) to see how the transition action changes with the potential barrier. For simplicity, we use the metabolism-EMT-metastasis network but constrain the parameter choices within the bistable regime (the E state and M state coexist) to calculate both barrier heights and transition actions. From the results (Figure 4), the barrier height and the corresponding transition action both increase as λ_{hh} increases (Figures 4A and 4B). However, the transition action from the M to E state increases faster than the transition action from the E to M state (Figure 4C, red line). This is because λ_{hh} quantifies the self-activation of HIF-1, which promotes the glycolysis state and M state. A faster increase of transition action from the M to E state means a more stable M state, which is consistent with the role of λ_{hh} for promoting HIF-1. This result is not reflected by the barrier height results (Figure 4C, blue line), although there is a statistical correlation between barrier height and transition action (Figure 4D). This suggests that the transition action provides a more precise description for the dynamics of the system than the barrier height, since the barrier heights are calculated from two-dimensional approximation, whereas the transition actions are calculated directly from the high-dimensional system.

To further study the effects of parameters on the dynamics of the whole network, we performed a global sensitivity analysis on parameters for the metabolism-EMT-metastasis model. For simplicity, we constrained the parameter values in the bistable regime (the E state and M state coexist). We calculated the transition actions between the E state and M state to quantify the feasibility in transitions for EMT and mesenchymal to epithelial transition (MET). We assume that the transition actions for EMT and MET should be critical to metastasis or de-metastasis progress because the EMT is a critical step for metastasis progression. We increase or decrease each parameter (here we focus on the parameters for the synthesis rate and regulatory strengths) individually by 10% to modulate the situations for gene over-expression or knockdown and for the change of regulatory strengths among genes or metabolites. Then we calculate how the transition actions between the E state and M state change after these perturbations. We also performed a global sensitivity analysis for the metabolic model to uncover the critical factors specifically for the metabolic model (see Supplemental Information Section S9 and Section S10).

From the results of sensitivity analysis, we picked top 25 parameters based on their sensitivity for detailed analysis (see Figure S5 for the complete sensitivity analysis of 61 parameters), which are shown in Figure 5.

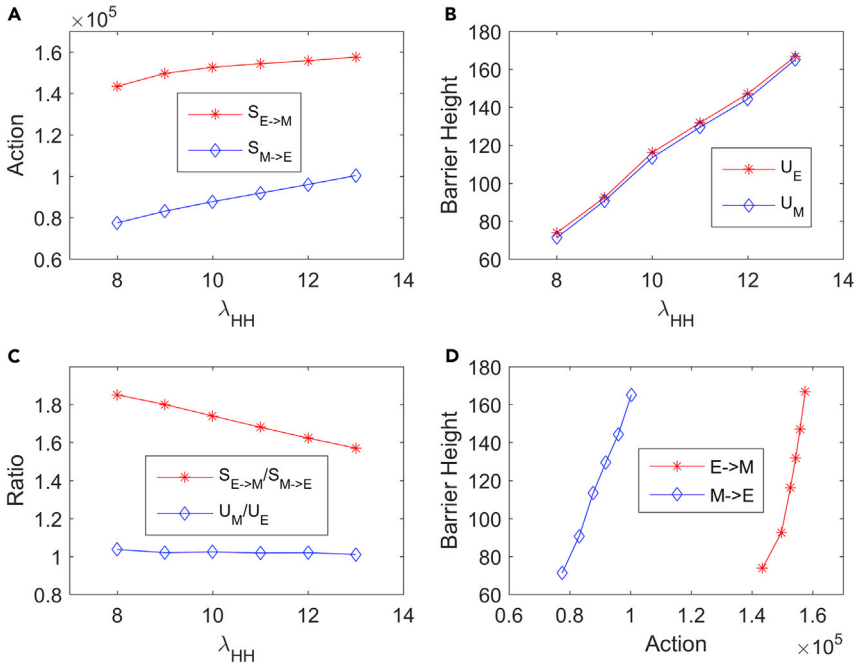


Figure 4. The Relation between Transition Action and Barrier Height Based on the Bistable Landscape of the Metabolism-EMT-Metastasis Model as λ_{hh} Increases (the Fold Change for the Self-Activation of HIF-1)

(A) The relation between λ_{hh} and the transition action (S).

(B) The relation between λ_{hh} and the barrier height (U).

(C) Relative changes for the transition action S (red line) and barrier heights U (blue line) as λ_{hh} increases.

(D) The barrier height changes as the transition action changes.

Among these top 25 paramount parameters, a few of the regulations are related with P53, for example, the activation of AMPK on P53 and the activation of P53 on miR-145. The increase of regulatory strengths for either of these two links will increase the transition actions from the E to M state (or the transition barrier from the E to M state), making the E state more stable. This embodies the critical roles of P53 in preventing cancer metastasis (Powell et al., 2014). This is consistent with our expectations from the network diagram because P53 activates multiple miRNAs and inhibits EMT-promoting factor OCT4, therefore playing a critical role in preventing cancer metastasis. Another prominent factor is the synthesis rate of mtROS, whose increase will greatly enhance the transition action from the E to M state, making the E state more stable. This indicates that mtROS could be a useful target for the cancer therapeutic strategy, probably because of its key influence on metabolism.

Additionally, increasing the fold change of inhibitory regulation of miR-145 on MDM2 (as well as on ZEB), i.e., reducing the inhibitory effects, results in the decrease of transition action from the E state to M state, promoting EMT (Figure 5A). This is because MDM2 inhibits P53 by promoting its degradation. Reducing the inhibition on MDM2 will promote MDM2 and inactivate P53, which will promote EMT owing to the inhibitory roles of P53 on EMT. Similarly, reducing the inhibition of miR-145 on ZEB will promote ZEB and further promote EMT. These predictions from our modeling are consistent with the experimental observations showing that miR-145 blocks the EMT (Zhao et al., 2016).

Another prediction from the sensitivity analysis is that the increase of the strength of the self-activation of ZEB λ_{ZZ} (or the increase of the synthesis rate of ZEB g_Z or the increase of the fold change for the activation regulation of HIF-1 on ZEB λ_{hZ}) can enhance the transition action from the M state to E state $S_{M\to E}$, making the transition from the M to E state less probable and the M state more stable (Figure 5A). This prediction is consistent with previous finding, showing that ZEB1 levels drive the dynamics of EMT (Jolly et al., 2018). Recently, Jia et al. demonstrated that miR-200/ZEB1 behaves as a three-way decision-making switch enabling transitions among the E, intermediate E/M, and M cell states by theoretical and experimental studies (Jia et al., 2017). Our prediction here is consistent with this previous work (Jia et al., 2017), by

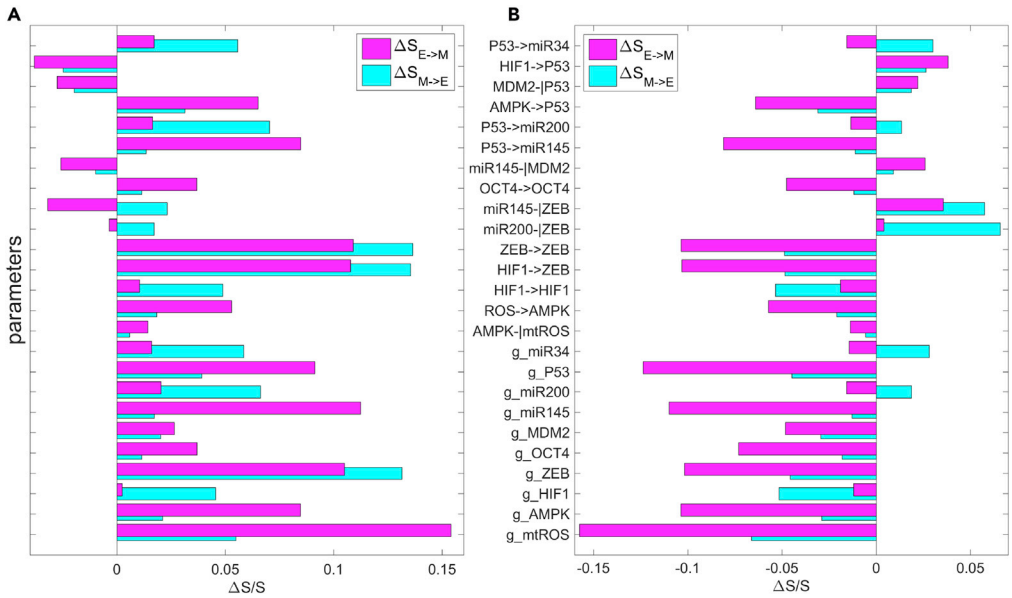


Figure 5. Global Sensitivity Analysis for Parameters Based on the Transition Action for the Metabolism-EMT-Metastasis Model

Y axis represents the 25 parameters. X axis represents the percentage of the change of the transition action ($\mathcal{S}$) relative to $\mathcal{S}$ with default parameters. Here, $S_{E\to M}$ represents the transition action from attractor E to attractor M (magenta bars) and $S_{M\to E}$ represents the transition action from attractor M to attractor E (cyan bars). The top 25 parameters are picked, in which the first 15 parameters represent the regulatory strength (λ) and the last 10 parameters represent the synthesis rate (g) for 10 proteins or metabolites. See Supplemental Information for the complete sensitivity analysis of 61 parameters.

(A) Each parameter is increased by 10%, individually.

(B) Each parameter is decreased by 10%, individually.

emphasizing the role of ZEB as a mesenchymal state promoter. From the network diagram, this is because the increase of the synthesis rate of ZEB or the activation strength of HIF-1 on ZEB will make ZEB activated, which further promotes EMT or the stability of M state. This also agrees with previous studies showing that the ZEB family of transcription factors are inducers of EMT, and P53, miR-200, and miR-139 affect the EMT through ZEB transcription factors (Ohashi et al., 2010; Korpal et al., 2008; Qiu et al., 2015; Sánchez-Tilló et al., 2011).

We also found that the over-expression of HIF-1 will increase the stability of the M state, reflected by the increase of transition action from the M to E state (Figure 5A). From the network topology, the over-expression of HIF-1 will promote ZEB and therefore make the M state more stable. This is consistent with the experimental study showing that the HIF-1 is a significant positive factor for regulating tumor progression and metastatic potential (Liao et al., 2007). Additionally, the inhibition of AMPK will promote EMT and metastasis, indicated by the decrease of transition action from the E to M state (Figure 5B). This is because AMPK promotes the expression of P53, which plays a critical role in inhibiting EMT. Therefore, the inhibition of AMPK will down-regulate P53 and promote EMT. This is also consistent with previous investigations showing that the loss of AMPK activation promotes the invasion and metastasis of pancreatic cancer (Chen et al., 2017).

Model Predictions Based on Landscape and Paths for Cancer Metastasis

Our computational results provide some guidance for experimental studies on the relationships among EMT, metabolism, and cancer metastasis. We summarized the experimental evidences that are consistent with our modeling results, as well as some predictions that could be tested from experimental studies (Table 1). First, we proposed a quadrastable landscape for the interplay among EMT, metabolism, and metastasis. Among the four stable states on the landscape, the epithelial state and mesenchymal state have been suggested from previous studies (Zhang et al., 2014; Lu et al., 2013; Hong et al., 2015). The metastatic state with high BACH1 expression was consistent with a previous study (Lee et al., 2014). Also, a

Model Predictions	Supporting References
In metastatic progression, there exists an abnormal metabolic state.	Prediction
Forward and reverse paths for EMT and metastasis are not identical.	(Zhang et al., 2014)
Metastasis is a stepwise progress, i.e., abnormal metabolism, EMT, and metastasis should be activated in a cascade.	Prediction
ROS changes the levels in the very early stage of the cell fate decision process toward EMT or metastasis.	Prediction
Over-expression of P53 attenuates EMT and metastasis progression.	(Powell et al., 2014)
Over-expression of ZEB promotes EMT.	(Sánchez-Tilló et al., 2011; Korpal et al., 2008)
Over-expression of miR-145 attenuates EMT.	(Wang et al., 2017)
Over-expression of HIF-1 promotes EMT and metastasis.	(Liao et al., 2007)
Inhibition of ROS promotes EMT and metastasis.	(Piskounova et al., 2015; Gill et al., 2016)
Inhibition of AMPK promotes EMT and metastasis.	(Chen et al., 2017)

Table 1. Major Model Predictions and Experimental Supports

recent work suggested that a combination therapy by targeting BACH1 and mitochondrial metabolism suppressed tumor growth and metastasis in triple-negative breast cancer (Lee et al., 2019), which provides another support that BACH1 could be a marker gene for cancer metastasis. However, the intermediate abnormal metabolic state we identified here has not been explicitly reported from experiments. Therefore, detailed molecular experiments and steady state measurements on cancer metastasis systems might be needed to test if this abnormal metabolic state, along with the multistability, exists.

Second, our results for kinetic transition path suggest that cancer metastatic progression is an irreversible process, which is indicated by the fact that the forward transition paths and the backward transition paths are not identical. We also found that the cancer metastatic progression is a sequential progress, i.e., the HIF-1 is first activated, and then the ZEB (characterizing EMT marker genes) is switched on followed by the activation of BACH1 (characterizing metastatic marker genes). These predictions can be tested experimentally.

Third, the discrete path results suggest that ROS changes the levels in the very early stage of the cell fate decision process toward metastasis (Figures 3B and 3C). Our results for the sensitivity analysis show that ROS significantly influences the transition actions for the E to M transition (Figure 5). This suggests that ROS might be a critical factor in the cancer metastatic progression by regulating metabolism. This is consistent with previous findings showing that antioxidants promote metastasis in melanoma (Piskounova et al., 2015; Gill et al., 2016). Previous studies showed that mtROS can also promote metastasis (Ishikawa et al., 2008; Porporato et al., 2014). These controversial conclusions might be related with the fact that cancer cells can use different ways of metabolism, e.g., use glycolysis or OXPHOS metabolism to activate metastasis (Yu et al., 2017).

Of note, the metastatic state identified in our model should represent a state with high expression level of HIF-1 (representing abnormal metabolism), high expression level of ZEB and other EMT marker genes (representing that EMT is activated), and high expression level of BACH1 (representing certain metastatic markers activated). This notion is consistent with the idea of hallmarks of cancer (Hanahan and Weinberg, 2000, 2011), i.e., the cancer metastasis should be a state in which cells have acquired different hallmarks of cancer. To test this type of cells experimentally, one should measure the expression level of relevant genes in the regulatory network and look for the states, for example, with a high level of HIF-1, ZEB, and SNAIL; low level of miR145, RKIP, and Let7; and high level of LIN28 and BACH1 (Table S4). We need to stress that, in our work, we focus on the interplay between metabolism, EMT, and metastasis, and have not explicitly considered the tristability properties of LIN28/Let7 circuit or the interplay between metabolism and

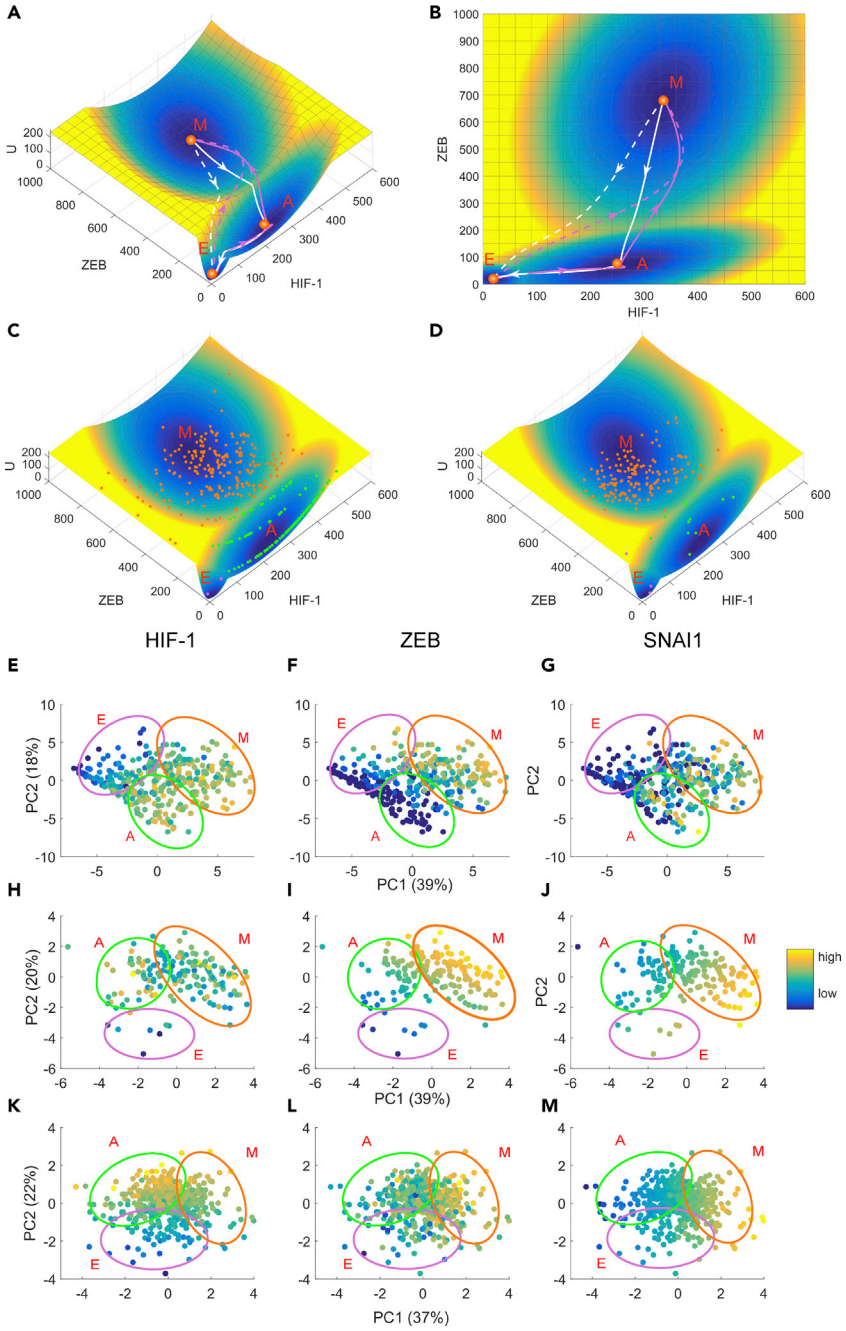


Figure 6. The Tristable Landscape for EMT-metabolism Model and Comparisons with Experimental Data

The EMT-metabolism model corresponds to the subnetwork of the whole network without consideration of metastasis circuit in Figure 1.

(A) Three dimensional landscape and transition paths. Solid magenta lines represent transition paths from the E to A, and M state. Solid white lines represent transition paths from the M to A and to E state. The dashed lines represent the direct transition path from E to M state and from M to E state, respectively.

(B) Two dimensional landscape and transition paths.

(C and D) Landscapes are compared with the gene expression data of single-cell RNA-seq data for a genetic mouse model of skin squamous cell carcinoma (SCC) undergoing EMT including 383 single cells (C) and clinically annotated adult cases of de novo AML from TCGA including 173 samples (D). The gene expression data have been rescaled to fit the landscape. Each point represents a gene expression pattern in ZEB and HIF-1 coordinates for one sample from experiments.

Figure 6. *Continued*

(E-G) PCA plots for the SCC data with respect to HIF-1 (E), ZEB (F), and SNAI1 (G).

(H-J) PCA plots for the AML data with respect to HIF-1 (H), ZEB (I), and SNAI1 (J).

(K-M) PCA plots for the PTC data with respect to HIF-1 (K), ZEB (L), and SNAI1 (M). Three clusters have been marked in the PCA coordinates, which correspond to E state (purple ovals, with low HIF-1/low ZEB expression), A state (green ovals, with high HIF-1/low ZEB expression), and M state (orange ovals, with high HIF-1/high ZEB expression), respectively. The colors in the PCA plots represent the expression levels of key marker genes (e.g., HIF-1, ZEB, and SNAIL) for E, A, and M phenotypes. E, epithelial state; M, mesenchymal state; A, abnormal metabolic state.

stemness. As proposed in previous work (Jolly et al., 2015b), the stemness and EMT can be coupled in a flexible way. So, our metastatic state with high expression level of LIN28 and low expression level of Let7 might be corresponding to the state with stemness and the ability for the tumor initiation (Jolly et al., 2015b). Future work can study how ROS influences the coupling among metabolism, EMT, metastasis, and stemness, which might help to resolve the controversial conclusions for the roles of ROS in metastasis.

Finally, from our sensitivity analysis, the over-expression of P53 will significantly increase the transition action for the transition from the E to M state. This is consistent with the key roles of P53 as a tumor repressor (Powell et al., 2014). However, we found that the over-expression of P53 will also increase the transition action for the transition from the M to E state, which means that it is not very effective to make an M cell switched back to an E cell. This indicates that P53 alone might not be enough to induce the transition of an M state cell back to an E state cell, whereas a better strategy could be targeting multiple genes simultaneously. Of note, recent studies have suggested a computational approach to identify the optimal combination of multiple anti-cancer targets (Li, 2017; Li and Balazsi, 2018).

Landscape and Kinetic Paths for EMT and Cancer Metabolism

To further reveal the relationship specifically between EMT and metabolism, we quantified the potential landscape of a subnetwork of the whole network by neglecting the metastasis circuit (yellow box in Figure 1), i.e., we consider only the coupling between EMT and metabolism (see Figures S6–S8 and Table S6 for the landscape analysis specifically for the metabolic model). Here we picked ZEB and HIF-1 as the marker genes of EMT and metabolism, respectively, and obtained the potential landscape (Figures 6A and 6B). On the landscape, the blue region denotes lower potential or higher probability and the yellow region denotes higher potential or lower probability. From the landscape of EMT-metabolism, we identified three attractors (cell states), which characterize the epithelial (E), the mesenchymal (M), and an intermediate abnormal metabolic (A) state, respectively. The multistability and the existence of intermediate states of EMT have been suggested from previous work (Lu et al., 2013; Zhang et al., 2014; Li et al., 2016). However, these previous works focus on small circuits and have not considered the influence of cellular metabolism on the EMT. By coupling the EMT circuit and metabolism circuit, here we are able to study the dynamics of the interplay between EMT and metabolism.

To investigate the transition dynamics for the EMT-metabolism system, we calculated the MAPs among different attractors by minimizing the transition actions $\mathcal{S}$. The MAPs for different transitions are shown on the landscape (Figures 6A and 6B). The MAPs from the epithelial state to the mesenchymal state are shown in magenta lines, and the reverse paths are shown in white lines.

The landscape and transition path results indicate that the transition process from the epithelial state to the mesenchymal state can be divided into two steps: the epithelial cells first switch to the intermediate state (or abnormal metabolic state A) with the increase of HIF-1 level and then are transformed to mesenchymal cells with the increase of ZEB level. These results demonstrate that, before cells are transformed to the mesenchymal state, they need to first change their metabolism. Of note, the intermediate abnormal metabolic state we identified here is different from the hybrid state suggested in EMT modeling (Lu et al., 2013) or the hybrid state proposed on cancer metabolism (Yu et al., 2017). The hybrid or intermediate state in EMT models (Lu et al., 2013) was defined as a state with half-epithelial and half-mesenchymal cell properties, and the hybrid or intermediate state in cancer metabolism model (Yu et al., 2017) was defined as a cell state that can use both glycolysis and OXPHOS metabolism. Neither of these two previous models has discussed the interplay between EMT and metabolism. Differently, our intermediate state suggested here corresponds to a state with abnormal metabolism (increased HIF-1 expression) but is still similar to an epithelial cell state (with the similar expression levels of epithelial marker genes as the epithelial cells).

Therefore, this builds a connection between EMT and cancer metabolism. As a matter of fact, different studies have suggested that aberrant metabolism plays a key role in developing EMT and cancer metastasis (Huang and Zong, 2017; Sciacovelli and Frezza, 2017; Jiang et al., 2015). We need to stress that our intermediate metabolic state identified here is not in conflict with the previous hybrid state in EMT (Lu et al., 2013; Pastushenko et al., 2018). This difference is due to the disparate resolution of the two models because our models here focus on the interplay between metabolism, EMT, and metastasis, which is more complicated and inclusive than previous models only focusing on EMT. In principle, it is possible to discover both the hybrid EMT state and the abnormal metabolic state by fine-tuning parameters in our models.

Landscape Results Are Supported by Experimental Data

To further validate our modeling results, we compare our EMT-metabolism model with the experimental data (see Supplemental Information Section S11 for a detailed approach of data analysis). Since our model can be considered as a representative EMT-metabolism network (not for some specific types of cancer), we collected three different types of gene expression data from experiments. The first dataset we acquired is the single-cell RNA-seq data for a genetic mouse model of skin squamous cell carcinoma (SCC) undergoing EMT (383 single cells, available from the NCBI Gene Expression Omnibus under accession number GSE110357) (Pastushenko et al., 2018). We also collected another dataset of gene expression data (acute myeloid leukemia [AML]) from The Cancer Genome Atlas (TCGA) (173 samples for clinically annotated adult cases of de novo AML) (Cancer Genome Atlas Research Network et al., 2013). AML is the second most common leukemia diagnosed in both adults and children, and it has been proposed that leukemia cells are inherently metastatic compared with solid tumors (Trendowski, 2015). The regulation of EMT has been suggested to play critical roles in AML development (Meyer, 2017; Carmichael et al., 2016; Stavropoulou et al., 2016). The third dataset we collected is for papillary thyroid carcinoma (PTC) from TCGA including 496 PTC samples (Agrawal et al., 2014). All of these three types of data are time independent and therefore should correspond to the steady-state data. We extracted the expression data for the relevant genes that appear in our network (Figure 1) from these datasets and neglected the miRNAs and metabolites that are not available from the data.

To see if our modeling results agree with these data, we rescaled the expression data to the range of the basins obtained from our models by linearly transforming the expression values, so that they match the landscape basins approximately. We believe that this normalization of the data is reasonable because some corresponding gene expression values from different datasets differ from each other by many folds and they represent different types of cancer in different conditions. By mapping these gene expression data (the gene expression data of HIF-1 and ZEB after normalization) on the landscape (Figures 6C and 6D), we found reasonable consistency between experimental data and our landscape results, i.e., the experimental data also display some clusters, which could correspond to the three stable states (cell types) on the landscape. Here, each point represents a gene expression pattern in the ZEB and HIF-1 coordinates for one sample from experiments (Figures 6C and 6D). The experimental data points can be classified into three different cell states (the E state represented by magenta points, the M state represented by orange points, and the A state represented by green points), which is consistent with our modeling results.

We also compared our modeling results with the experimental data in another way. We performed the principal component analysis (PCA) of the three types of experimental data, which are shown in Figures 6E–6G (SCC data), Figures 6H–6J (AML data), and Figures 6K–6M (PTC data), respectively. In the PCA plots of the SCC data, we can identify three clusters, which are consistent with our landscape results. These three clusters have been marked in the PCA coordinates that correspond to the E state (purple ovals, with low HIF-1/low ZEB expression), A state (green ovals, with high HIF-1/low ZEB expression), and M state (orange ovals, with high HIF-1/high ZEB expression). Here, the colors in the PCA plots represent the expression levels of key marker genes (e.g., HIF-1, ZEB, and SNAIL) for the E, A, and M phenotypes. Importantly, we found that, from the cluster E to cluster A, the HIF1 expression increases (from blue to yellow), whereas the ZEB (and SNAIL) expression does not increase (keeping blue) (Figures 6E–6G). Also, from the cluster A to cluster M, the ZEB (and SNAIL) expression increases significantly (from blue to yellow) (Figures 6E–6G). These experimental data provide a support to the landscape attractors we identified, which is that the E, A, and M states coexist. For the PCA plots of AML and PTC data, we also find similar modes for the three attractor states (E, A, and M) in terms of HIF-1 and ZEB

(Figures 6H–6M). Both these ways of validations show good consistency between modeling results and experimental data. It is worth noting that our metabolism-EMT regulatory network is constructed purely by mining the experimental literature and the gene expression data are obtained from a different database, separately. So, the consistency between our modeling results and the experimental data is inspiring, which also confirms the predictive power of our models.

Here, for metabolism and EMT circuit, we found certain consistency between experimental data and our models. But we have not found significant consistency for the metastatic circuit in the SCC dataset we used. This is probably because our metastatic circuit is only a simplified module to represent metastasis (e.g., only one metastasis marker gene BACH1 is involved). To resolve this issue, a more inclusive metastasis model involving more genes is needed. Nevertheless, for the cascading regulation of EMT and metastasis, our previous work by combining modeling and single cell data provides support for this conclusion (Li and Balazsi, 2018).

Modeling Anti-Cancer Therapeutic Strategies

Targeting cellular metabolism has been shown to be a promising route against cancer. Several metabolic drugs have been uncovered and shown to be effective against cancers in certain cases, e.g., 3-bromopyruvate (3BP) (Ganapathy-Kanniappan et al., 2010) and metformin (Dowling et al., 2011). It has also been suggested that the combined drugs may be more effective than single drugs (Cheong et al., 2011). However, the underlying mechanisms for why the designed drugs work in certain cases but not in others have yet to be clarified. Here, we propose that these drugs take effects by changing the underlying landscape topography of the metabolism-EMT-metastasis gene regulatory network. We model the drug effects by making perturbations to the key nodes in the network and trace the changes on the landscape shape. Here, the green nodes (Figure 7A) represent the hypothetical drugs, and by changing the levels of drugs we can study how the landscape is influenced (see Supplemental Information Section S12 for detailed ODEs describing the drug influences and Table S7 for parameter values). In Figures 7B–7D, we showed how the landscape shape changes as the drug doses increase for 3BP (Figure 7B), metformin (Figure 7C), and the combined 3BP and metformin therapy (Figure 7D), respectively. We can see that all three therapies are effective on destabilizing the metastatic state, as indicated by the changes in the landscape topograph where the metastatic attractor gradually disappears as the drug doses increase. We found that the combined 3BP and metformin therapy is more effective than either of 3BP or metformin acting alone. This is because, when the drug doses increase to certain level, the combined 3BP and metformin therapy will lead to the disappearance of the A, M, and Met attractors. This does not happen for the single 3BP or metformin therapy (Figures 7B–7D). Therefore, these results are consistent with the experimental studies showing that the combinations of metformin and 2DG (whose effect is similar to 3BP) are more effective than each drug in action alone (Cheong et al., 2011).

We also found that even very large doses of any of these drugs (last column in Figures 7B–7D) alone cannot induce the cells to the E state completely because there are always other abnormal cell states present (Met2 in Figures 7B and 7D, M in Figure 7C). The Met2 state here corresponds to a state with low HIF-1, low ZEB, and high BACH1 expression level, which should be thought of as a metastasis-like state. This can explain why designed drugs work in certain cases but not in the others. It is because cells after drug treatments can still transform to the M state (Figure 7C) or the Met2 state (Figures 7B and 7D), which should preserve certain features of metastasis cells. Therefore, based on these changes in landscape topography, we suggest another therapeutic strategy, which is the combination of 3BP (or metformin) and certain types of BACH1 inhibitor (Figure 7E). The purpose of adding BACH1 inhibitor is to destabilize the Met2 state, so that the cells will be more prone to switch to the E state. As indicated in Figure 7E, the new combination of drugs suggested here can induce the change in landscape topography from four states coexistence to monostable E state alone. The new combination strategy suggests a better treatment for cancer metastasis. In fact, the BACH1 as one of the new targets for suggested combination drugs mentioned earlier has been suggested as a novel individual candidate target for cancer therapy (Davudian et al., 2016). More interestingly, a recent work suggested that a combination therapy by targeting BACH1 and mitochondrial metabolism suppressed tumor growth and metastasis in triple-negative breast cancer (TNBC) (Lee et al., 2019). These experiments provide a strong support for our predictions on the combination drugs.

Recently, an integrated computational and experimental approach was designed to identify effective therapeutic strategies based on temporal sequencing of multiple drugs from deterministic models

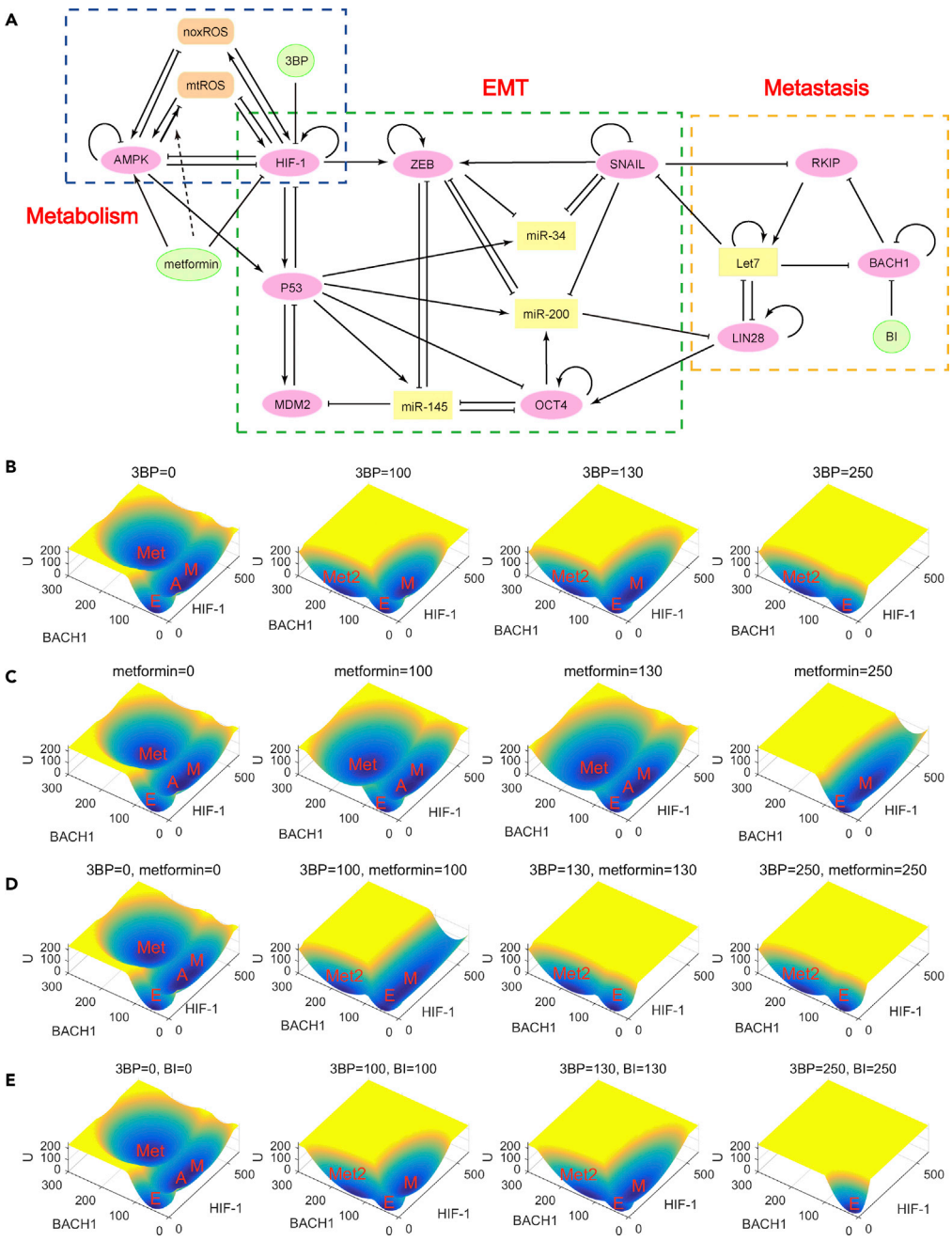


Figure 7. Landscape in Terms of HIF-1 and BACH1 in Response to Different Drugs in Different Levels

(A–E) (A) The metabolism-EMT-metastasis network with new added green nodes representing the hypothetical drugs. The simulated drugs include 3BP (B), metformin (C), combined 3BP and metformin therapy (D), and combined 3BP and BACH1-inhibitor therapy (E). In B–E, the hypothetical drug levels increase from left to right, which represent the drug level of 0, 100, 130, and 250, respectively. E, epithelial state; M, mesenchymal state; A, abnormal metabolic state; Met, metastasis state; Met2, metastasis-like state; BI, BACH1 inhibitor.

(Goldman et al., 2015, 2019). Of note, our current landscape approach is based on the steady-state distribution, where we explored multiple initial conditions from the ODEs. From the landscape perspective, drugs play roles by reshaping the landscape (steady states) and changing the relative stability of different cell states. To see the effects of temporal sequencing of multiple drugs, an approach for calculating temporal landscape (landscape at different time points) has yet to be developed.

DISCUSSION

Abnormal metabolism, EMT, and cancer metastasis are three important processes in the progression of cancer. These processes have been extensively explored individually both from experiments and modeling. However, the underlying mechanisms connecting these processes remain elusive. In this work, we aim to establish a theoretical framework for modeling the interplay among metabolism, EMT, and metastasis to reveal the possible principles for cancer metastasis. We develop a TME approach to quantify the potential landscape. We constructed a metabolism-EMT-metastasis regulatory network and identified four stable states from the landscape: epithelial (E), abnormal metabolic (A), mesenchymal (M), and metastatic (Met) cell states. Importantly, we identified a new abnormal metabolic cell state. We further calculate the MAPs to quantify the most probable transition paths for cell fate decision processes in metastatic progression. The results of landscape and MAPs indicate that the epithelial cells prefer to first change their metabolism, then finish EMT, and eventually enter the metastatic state. By performing a global sensitivity analysis from the transition actions, we identified some key parameters for the metastatic progression, which agree well with experimental observations.

Our landscape results have a few indications for cancer metastasis. First, for metastatic progression, cells prefer to follow a specific sequential order: change their metabolism, finish EMT, and proceed into the metastatic state. This feature indicates that there are interplays among these three processes to generate a temporal order. Second, the intermediate states identified from our models may account for the heterogeneity observed in experiments for tumors. Third, the landscape view considers metastasis state as an attractor. This means that, in principle, metastatic cancer cells can be destroyed by altering the underlying landscape such that the metastasis state is no longer a stable state. To achieve this, a possible effective way should be either intervening genes or regulatory interactions among genes in cancer gene regulatory networks. In fact, recent experimental work showed that EMT-derived breast cancer cells can be induced to differentiate into post-mitotic adipocytes through a combination therapy (Ishay-Ronen et al., 2019). Also, the reverse transitions from M cells to E cells or from hybrid E/M cells to E cells have been observed in experiments (Zhang et al., 2014; Hong et al., 2015). These experiments indicate the possibility for inducing the conversion of metastatic cancer cells to non-invasive cells (e.g., E cells or intermediate state cells) *in vivo*. On the contrary, the traditional strategy for killing metastatic cells does not change the stability of metastasis attractor. This might be one reason why we often see clinically that after the chemotherapy or radiotherapy the cancer relapses.

Recent studies have identified intermediate hybrid phenotypes both at the single-cell level and population levels across different cancer types (Pastushenko et al., 2018; Jolly et al., 2015a). Our models for the interplay among metabolism, EMT, and metastasis give rise to a quadrastable landscape. This supports the existence of intermediate states in EMT. Additionally, a critical prediction from our model is that the cancer metastatic progression might be a sequential process. Specifically, cancer cells tend to first change their metabolism. Then, with the abnormal metabolism (possibly providing energy more quickly), the EMT circuit is activated. This further leads to the cancer metastasis. This conclusion is partially supported by multiple types of gene expression data (Figures 6C–6H), although gene expression data with dynamical information are needed to further test this prediction. This sequential procedure resembles previous studies suggesting that metastasis appears mainly by a sequential, multi-step process that can be viewed as a cascade of the invasion-metastasis (Lambert et al., 2017). These results indicate that the temporal order is critical for different genes (characterizing different functions or hallmarks of cancer) switching on or off in metastatic progression process and different hallmarks of cancer can interact with each other and cooperate. Our landscape results also indicate that the multistable landscape and the existence of intermediate states play critical roles in metastatic progression. For example, the existence of intermediate states or sequential progression can increase cellular plasticity. As a matter of fact, the epigenetic landscape governing the stability of epithelial-mesenchymal plasticity has been proposed as an illustration (Tam and Weinberg, 2013). Our results offer quantitative supports for the epigenetic landscape of EMT and metastasis.

Our work has certain implications on the origins of cancer. Genetic mutations, metabolic dysfunction, and environmental factors are commonly considered to play distinct roles in the development of tumors. Our work suggests that these factors may be connected to each other intimately, and the debate on which one is dominant can be reconciled in a quantitative way. In fact, different circuits are intertwined with feedbacks to each other, as shown in Figure 1. First, the metabolic disorder undoubtedly plays critical roles in tumor development because from our models the E to A state transition (Figure 2) is a crucial step for tumor

development. The metabolic disorder can be caused by the environmental stress (e.g., hypoxia) or the genetic mutations (e.g., oncogene RAS, MYC) (Hanahan and Weinberg, 2011), or combinations of both.

However, the abnormal metabolic state does not necessarily lead to cancer, if the system stays at the A state stably. To become tumor cells or obtain the invasive ability, the cells need to activate other circuits (such as EMT circuit) to induce the A to M state transition (Figure 2). In a similar way, the genetic mutations (inherited or acquired) and the environmental stress (such as UV radiation) can contribute to this step. In fact, this can be considered as a process of accumulations of mutations (e.g., the mutations related to EMT marker genes ZEB and SNAIL). Finally, even in the M state, the cells still have the opportunity of not becoming malignant tumor cells if no M to metastasis state transition is induced. In principle, an M cell can transform back to an A or E cell in certain conditions. However, with the further accumulations of mutations (related with metastasis circuit), the M state cells can be further transformed to metastatic cells. This completes the journey of the cascade for the progression of cancer metastasis. Therefore, the quantitative landscape models offer a possible explanation for the origin of cancer, which supports the notion that cancer originates from the combined effects of accumulated mutations and environmental stress.

The sequential cascade in metastatic progression also suggests that a more effective therapeutic strategy might be targeting multiple circuits (representing different functions) of the cancer network, e.g., targeting metabolism, EMT, and metastasis simultaneously. In fact, many studies have suggested that EMT activation can cause tumor relapse and enhanced tumorigenesis in different human cancer cell lines (Moody et al., 2005; Creighton et al., 2010; Ye and Weinberg, 2015). By modeling the landscape changes with several designed drugs (such as 3BP and metformin), we found that these drugs can weaken the metastatic state in certain circumstances, but they also induce new metastasis-like states. This can explain why some drugs can work in certain cases against cancer but not in others. Our results suggest that the combinations of different drugs should be more effective than single drugs (Cheong et al., 2011). Based on the changes of landscape topography upon drug treatment, we suggest a new therapeutic strategy (combining 3BP/metformin and BACH1 inhibitor), which can lead to a better treatment. Remarkably, our predictions on the combination drugs are partially supported by a recent work showing that a combination therapy by targeting BACH1 and mitochondrial metabolism suppressed tumor growth and metastasis in TNBC (Lee et al., 2019).

Yu et al. proposed that cancer cells use both glycolysis and OXPHOS for metabolism (Yu et al., 2017). In our work, the abnormal metabolic state is more related to glycolysis state marked by higher expression of HIF-1. Our modeling results indicate that the abnormal metabolism (such as aerobic glycolysis) is critical for activating metastasis. This is supported by previous work, e.g., cancer metastasis has long been related to the switch from OXPHOS to glycolysis metabolism, i.e., the Warburg effects (Han et al., 2013; Gatenby and Gillies, 2004; Kamarajugadda et al., 2012). Additionally, both our modeling results and gene expression data from experiments show that HIF-1 is activated in metastatic cancer cells (Figure 6). In fact, previous work has suggested that HIF-1 activation is critical for EMT and HIF-1 activation by acidic microenvironment contributes to tumorigenesis and metastasis (Han et al., 2013). This is consistent with our conclusion that glycolysis metabolism (HIF-1 activation) promotes EMT and then activates metastasis. However, we cannot rule out the possibility that cancer cells could also use OXPHOS metabolism to activate metastasis, which warrants further computational and experimental explorations.

In summary, the TME approach developed in this work provides an effective approach to quantify the stochastic dynamics of high-dimensional gene regulatory systems, and our results provide new insights into the mechanistic understanding of the interplay among metabolism, EMT, and cancer metastasis in the progression of cancer. The landscape and path approach can be used to explore the interplay among other hallmarks of cancer and to study the stochastic dynamics of other biological networks.

Limitations of the Study

One limitation of the current work is that we compare our modeling results only with a few types of cancer data. These data partially support the existence of an abnormal metabolic state. To our best knowledge, it remains challenging to integrate "model driven" approach (such as our gene network models) and "data driven" approach (such as single cell data analysis) in systems biology (Cho et al., 2017). Therefore, we still lack effective ways to compare modeling results and single cell data of cancer, especially for different types of cancer owing to the heterogeneity in many tumors. In this work, we performed some initial attempts for

this task. Another limitation is that we considered only certain genes, microRNAs, and metabolites that are important for the processes of metabolism, EMT, and cancer metastasis. In reality, other gene circuits or metabolites can also be important to the whole-cell fate decision processes in cancer metastasis. Third, cancer is a complex disease, which involves many hallmarks (Hanahan and Weinberg, 2000, 2011). Owing to the complexity, in this work we modeled a few hallmarks of cancer, including abnormal metabolism, EMT, and metastasis. Future work can incorporate other critical genes or metabolic circuits (and/or other hallmarks of cancer) into the models of underlying cancer gene regulatory networks, which may lead to the larger complexity but will provide more insights into underlying regulatory mechanisms for cancer metastasis.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.10.060.

ACKNOWLEDGMENTS

C.L. would like to thank Prof. Fengzhu Sun for valuable discussions. C.L. is supported by the National Natural Science Foundation of China (11771098), Natural Science Foundation of Shanghai, China (17ZR1444500), and Thousand Youth Talents Program. J.W. thanks for the supports from NSF-PHY-76066 and NSF-CHE-1808474.

AUTHOR CONTRIBUTIONS

C.L. conceived and designed the research. X.K. and C.L. performed the research. X.K., J.W., and C.L. analyzed the results and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 2, 2019
Revised: September 21, 2019
Accepted: October 28, 2019
Published: November 22, 2019

REFERENCES

Agrawal, N., Akbani, R., Aksoy, B.A., Ally, A., Arachchi, H., Asa, S.L., Auman, J.T., Balasundaram, M., Balu, S., Baylin, S.B., et al. (2014). Integrated genomic characterization of papillary thyroid carcinoma. Cell *159*, 676–690.

Bocci, F., Jolly, M.K., George, J.T., Levine, H., and Onuchic, J.N. (2018). A mechanism-based computational model to capture the interconnections among epithelial-mesenchymal transition, cancer stem cells and notch-jagged signaling. Oncotarget *9*, 29906.

Boroughs, L.K., and Deberardinis, R.J. (2015). Metabolic pathways promoting cancer cell survival and growth. Nat. Cell Biol. *17*, 351.

Brabletz, T., Kalluri, R., Nieto, M.A., and Weinberg, R.A. (2018). Emt in cancer. Nat. Rev. Cancer *18*, 128.

Cancer Genome Atlas Research Network, Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Carmichael, C.L., Goossens, S., Wang, J., Nguyen, T., Haigh, K., Berx, G., Kile, B., and Haigh, J.J. (2016). The EMT modulator snai1 drives AML development via its interaction with the chromatin modulator lsd1. Blood *128*, 2688.

Chen, C., Baumann, W., Xing, J., Xu, L., Clarke, R., and Tyson, J. (2014). Mathematical models of the transitions between endocrine therapy responsive and resistant states in breast cancer. J. R. Soc. Interface *96*, 20140206.

Chen, K., Qian, W., Li, J., Jiang, Z., Cheng, L., Yan, B., Cao, J., Sun, L., Zhou, C., Lei, M., et al. (2017). Loss of AMPK activation promotes the invasion and metastasis of pancreatic cancer through an hsf1-dependent pathway. Mol. Oncol. *11*, 1475–1492.

Cheong, J.H., Park, E.S., Liang, J., Dennison, J.B., Tsavachidou, D., Nguyen-Charles, C., Cheng, K.W., Hall, H., Zhang, D., Lu, Y., et al. (2011). Dual inhibition of tumor energy pathway by 2-deoxy glucose and metformin is effective against a broad spectrum of preclinical cancer models. Mol. Cancer Ther. *10*, 2350–2362.

Cho, K.H., Lee, S., Kim, D., Shin, D., Joo, J.I., and Park, S.M. (2017). Cancer reversion, a renewed challenge in systems biology. Curr. Opin. Syst. Biol. *2*, 49–58.

Creighton, C.J., Chang, J.C., and Rosen, J.M. (2010). Epithelial-mesenchymal transition (EMT) in tumor-initiating cells and its clinical implications in breast cancer. J. Mammary Gland Biol. Neoplasia *15*, 253–260.

Davudian, S., Mansoori, B., Shajari, N., Mohammadi, A., and Baradaran, B. (2016). Bach1, the master regulator gene: a novel candidate target for cancer therapy. Gene *588*, 30–37.

Dowling, R.J., Goodwin, P.J., and Stambolic, V. (2011). Understanding the benefit of metformin use in cancer treatment. BMC Med. *9*, 33.

Ganapathy-Kanniappan, S., Vali, M., Kunjithapatham, R., Buijs, M., Syed, L., Rao, P., Ota, S., Kwak, B., Loffroy, R., and Geschwind, J. (2010). 3-bromopyruvate: a new targeted

antiglycolytic agent and a promise for cancer therapy. Curr. Pharm. Biotechnol. *11*, 510–517.

Gatenby, R.A., and Gillies, R.J. (2004). Why do cancers have high aerobic glycolysis? Nat. Rev. Cancer *4*, 891.

Ge, H., and Qian, H. (2012). Landscapes of non-gradient dynamics without detailed balance: stable limit cycles and multiple attractors. Chaos *22*, 023140.

Ge, H., and Qian, H. (2016). Mesoscopic kinetic basis of macroscopic chemical thermodynamics: a mathematical theory. Phys. Rev. E *94*, 052150.

Gibbons, D.L., Lin, W., Creighton, C.J., Rizvi, Z.H., Gregory, P.A., Goodall, G.J., Thilaganathan, N., Du, L., Zhang, Y., and Pertsemlidis, A. (2009). Contextual extracellular cues promote tumor cell EMT and metastasis by regulating mir-200 family expression. Genes Dev. *23*, 2140–2151.

Gill, J.G., Piskounova, E., and Morrison, S.J. (2016). Cancer, oxidative stress, and metastasis. Cold Spring Harb. Symp. Quant. Biol. *81*, 163–175.

Goldman, A., Majumder, B., Dhawan, A., Ravi, S., Goldman, D., Kohandel, M., Majumder, P.K., and Sengupta, S. (2015). Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. Nat. Commun. *6*, 6139.

Goldman, A., Khiste, S., Freinkman, E., Dhawan, A., Majumder, B., Mondal, J., Pinkerton, A.B., Eton, E., Medhi, R., Chandrasekar, V., et al. (2019). Targeting tumor phenotypic plasticity and metabolic remodeling in adaptive cross-drug tolerance. Sci. Signal. *12*, eaas8779.

Han, T., Kang, D., Ji, D., Wang, X., Zhan, W., Fu, M., Xin, H.B., and Wang, J.B. (2013). How does cancer cell metabolism affect tumor migration and invasion? Cell Adh. Migr. *7*, 395–403.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hong, T., Watanabe, K., Ta, C.H., Villarreal-Ponce, A., Nie, Q., and Dai, X. (2015). An ovol2-zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. PLoS Comput. Biol. *11*, e1004569.

Huang, S. (2013). Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells. Cancer Metastasis Rev. *32*, 423–448.

Huang, R., and Zong, X. (2017). Aberrant cancer metabolism in epithelial–mesenchymal transition and cancer metastasis: mechanisms in cancer progression. Crit. Rev. Oncol. Hematol. *115*, 13–22.

Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D.E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys. Rev. Lett. *94*, 128701.

Ishay-Ronen, D., Diepenbruck, M., Kalathur, R.K.R., Sugiyama, N., Tiede, S., Ivanek, R., Bantug, G., Morini, M.F., Wang, J., Hess, C., et al. (2019). Gain fat lose metastasis: converting invasive breast cancer cells into adipocytes inhibits cancer metastasis. Cancer Cell *35*, 17–32.

Ishikawa, K., Takenaga, K., Akimoto, M., Koshikawa, N., Yamaguchi, A., Imanishi, H., Nakada, K., Honma, Y., and Hayashi, J.I. (2008). Ros-generating mitochondrial DNA mutations can regulate tumor cell metastasis. Science *320*, 661–664.

Jia, D., Jolly, M.K., Tripathi, S.C., Den Hollander, P., Huang, B., Lu, M., Celiktas, M., Ramirez-Peña, E., Ben-Jacob, E., Onuchic, J.N., et al. (2017). Distinguishing mechanisms underlying EMT tristability. Cancer Converg. *1*, 2.

Jia, D., Lu, M., Jung, K.H., Park, J.H., Yu, L., Onuchic, J.N., Kaipparettu, B.A., and Levine, H. (2019). Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. Proc. Natl. Acad. Sci. U S A *116*, 3909–3918.

Jiang, L., Xiao, L., Sugiura, H., Huang, X., Ali, A., Kuro-o, M., Deberardinis, R.J., and Boothman, D.A. (2015). Metabolic reprogramming during tgfβ 1-induced epithelial-to-mesenchymal transition. Oncogene *34*, 3908–3916.

Jolly, M.K., Boareto, M., Huang, B., Jia, D., Lu, M., Ben-Jacob, E., Onuchic, J.N., and Levine, H. (2015a). Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. Front. Oncol. *5*, 155.

Jolly, M.K., Jia, D., Boareto, M., Mani, S.A., Pienta, K.J., Ben-Jacob, E., and Levine, H. (2015b). Coupling the modules of EMT and stemness: a tunable 'stemness window' model. Oncotarget *6*, 25161–25174.

Jolly, M.K., Preca, B.T., Tripathi, S.C., Jia, D., George, J.T., Hanash, S.M., Brabletz, T., Stemmler, M.P., Maurer, J., and Levine, H. (2018). Interconnected feedback loops among esrp1, has2, and cd44 regulate epithelial-mesenchymal plasticity in cancer. APL Bioeng. *2*, 031908.

Kaern, M., Elston, T.C., Blake, W.J., and Collins, J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. *6*, 451–464.

Kalluri, R. (2009). EMT: when epithelial cells decide to become mesenchymal-like cells. J. Clin. Invest. *119*, 1417–1419.

Kamarajugadda, S., Stemboroski, L., Cai, Q., Simpson, N.E., Nayak, S., Tan, M., and Lu, J. (2012). Glucose oxidation modulates anoikis and tumor metastasis. Mol. Cell. Biol. *32*, 1893–1907.

Korpal, M., Lee, E.S., Hu, G., and Kang, Y. (2008). The mir-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of e-cadherin transcriptional repressors zeb1 and zeb2. J. Biol. Chem. *283*, 14910–14914.

Lambert, A.W., Pattabiraman, D.R., and Weinberg, R.A. (2017). Emerging biological principles of metastasis. Cell *168*, 670–691.

Lee, J., Lee, J., Farquhar, K.S., Yun, J., Frankenberger, C.A., Bevilacqua, E., Yeung, K., Kim, E.J., Balazsi, G., and Rosner, M.R. (2014). Network of mutually repressive metastasis regulators can promote cell heterogeneity and metastatic transitions. Proc. Natl. Acad. Sci. U S A *111*, E364–E373.

Lee, J., Yesilkanal, A.E., Wynne, J.P., Frankenberger, C., Liu, J., Yan, J., Elbaz, M., Rabe, D.C., Rustandy, F.D., Tiwari, P., et al. (2019). Effective breast cancer combination therapy targeting bach1 and mitochondrial metabolism. Nature *568*, 254.

Li, C. (2017). Identifying the optimal anticancer targets from the landscape of a cancer–immunity interaction network. Phys. Chem. Chem. Phys. *19*, 7642–7651.

Li, C., and Balazsi, G. (2018). A landscape view on the interplay between EMT and cancer metastasis. NPJ Syst. Biol. Appl. *4*, 34.

Li, C., and 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.

Li, C., and Wang, J. (2014a). Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. Proc. Natl. Acad. Sci. U S A *111*, 14130–14135.

Li, C., and Wang, J. (2014b). Quantifying the underlying landscape and paths of cancer. J. R. Soc. Interface *10*, 20140774.

Li, C., and Wang, J. (2015). Quantifying the landscape for development and cancer from a core cancer stem cell circuit. Cancer Res. *75*, 2607–2618.

Li, C., Hong, T., and Nie, Q. (2016). Quantifying the landscape and kinetic paths for epithelial–mesenchymal transition from a core circuit. Phys. Chem. Chem. Phys. *18*, 17949–17956.

Liao, C., and Lu, T. (2013). A minimal transcriptional controlling network of regulatory t cell development. J. Phys. Chem. B *117*, 12995–13004.

Liao, D., Corle, C., Seagroves, T.N., and Johnson, R.S. (2007). Hypoxia-inducible factor-1α is a key regulator of metastasis in a transgenic model of cancer initiation and progression. Cancer Res. *67*, 563–572.

Lu, M., Jolly, H., Levine, H., Onuchic, J., and Ben-Jacob, E. (2013). MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. Proc. Natl. Acad. Sci. U S A *110*, 18144–18149.

Lu, M., Jolly, M.K., Onuchic, J., and Ben-Jacob, E. (2014a). Toward decoding the principles of cancer metastasis circuits. Cancer Res. *74*, 4574–4587.

Lu, M., Onuchic, J., and Ben-Jacob, E. (2014b). Construction of an effective landscape for multistate genetic switches. Phys. Rev. Lett. *113*, 078102.

Lv, C., Li, X., Li, F., and Li, T. (2015). Energy landscape reveals that the budding yeast cell cycle is a robust and adaptive multi-stage process. PLoS Comput. Biol. *11*, e1004156.

Meyer, S.E. (2017). From EMT to HSC to AML: Zeb2 is a cell fate switch. Blood *129*, 400–401.

Moody, S.E., Perez, D., Pan, T.C., Sarkisian, C.J., Portocarrero, C.P., Sterner, C.J., Notorfrancesco, K.L., Cardiff, R.D., and Chodosh, L.A. (2005). The transcriptional repressor snail promotes mammary tumor recurrence. Cancer Cell *8*, 197–209.

Munozpinedo, C., Mjiyad, N.E., and Ricci, J. (2012). Cancer metabolism: current perspectives and future directions. Cell Death Dis. *3*, e248.

Ohashi, S., Natsuizaka, M., Wong, G.S., Michaylira, C.Z., Grugan, K.D., Stairs, D.B., Kalabis, J., Vega, M.E., Kalman, R.A., and Nakagawa, M. (2010). Egfr and mutant p53 expand esophageal cellular subpopulation capable of epithelial-to-mesenchymal transition through zeb transcription factors. Cancer Res. *70*, 4174.

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S., et al. (2018). Identification of the tumour transition states occurring during EMT. Nature *556*, 463.

Piskounova, E., Agathocleous, M., Murphy, M.M., Hu, Z., Huddlestun, S.E., Zhao, Z., Leitch, A.M., Johnson, T.M., DeBerardinis, R.J., and Morrison, S.J. (2015). Oxidative stress inhibits distant metastasis by human melanoma cells. Nature *527*, 186.

Porporato, P.E., Payen, V.L., Pérez-Escuredo, J., De Saedeleer, C.J., Danhier, P., Copetti, T., Dhup, S., Tardy, M., Vazeille, T., Bouzin, C., et al. (2014). A mitochondrial switch promotes tumor metastasis. Cell Rep. *8*, 754–766.

Powell, E., Piwnica-Worms, D., and Piwnica-Worms, H. (2014). Contribution of p53 to metastasis. Cancer Discov. *4*, 405–414.

Qiu, G., Lin, Y., Zhang, H., and Wu, D. (2015). mir-139-5p inhibits epithelial–mesenchymal transition, migration and invasion of hepatocellular carcinoma cells by targeting zeb1 and zeb2. Biochem. Biophys. Res. Commun. *463*, 315–321.

Sánchez-Tilló, E., Siles, L., De Barrios, O., Cuatrecasas, M., Vaquero, E.C., Castells, A., and Postigo, A. (2011). Expanding roles of zeb factors in tumorigenesis and tumor progression. Am. J. Cancer Res. *1*, 897–912.

Sciacovelli, M., and Frezza, C. (2017). Metabolic reprogramming and epithelial-to-mesenchymal transition in cancer. FEBS J. *284*, 3132–3144.

Singh, A., and Settleman, J. (2010). Emt, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene *29*, 4741–4751.

Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M.A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I.J., Milne, T.A., et al. (2016). Mll-af9 expression in hematopoietic stem cells drives a highly invasive AML expressing EMT-related genes linked to poor outcome. Cancer Cell *30*, 43–58.

Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc. Natl. Acad. Sci. U S A *99*, 12795–12800.

Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat. Med. *19*, 1438.

Thattai, M., and Van, O.A. (2001). Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. U S A *98*, 8614–8619.

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

Trendowski, M. (2015). The inherent metastasis of leukaemia and its exploitation by sonodynamic therapy. Crit. Rev. Oncol. Hematol. *94*, 149–163.

Waddington, C.H. (1957). The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology (Allen and Unwin).

Wang, J. (2015). Landscape and flux theory of non-equilibrium dynamical systems with application to biology. Adv. Phys. *64*, 1–137.

Wang, J., Xu, L., and Wang, E.K. (2008). Potential landscape and flux framework of non-equilibrium networks: robustness, dissipation and coherence of biochemical oscillations. Proc. Natl. Acad. Sci. U S A *105*, 12271–12276.

Wang, J., Zhang, K., Xu, L., and Wang, E.K. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. Proc. Natl. Acad. Sci. U S A *108*, 8257–8262.

Wang, L., Wu, X., Wang, B., Wang, Q., and Han, L. (2017). Mechanisms of mir-145 regulating invasion and metastasis of ovarian carcinoma. Am. J. Transl. Res. *9*, 3443.

Ye, X., and Weinberg, R.A. (2015). Epithelial–mesenchymal plasticity: a central regulator of cancer progression. Trends Cell Biol. *25*, 675–686.

Yu, L., Lu, M., Jia, D., Ma, J., Ben-Jacob, E., Levine, H., Kaipparettu, B.A., and Onuchic, J.N. (2017). Modeling the genetic regulation of cancer metabolism: interplay between glycolysis and oxidative phosphorylation. Cancer Res. *77*, 1564–1574.

Zhang, B., and Wolynes, P.G. (2014). Stem cell differentiation as a many-body problem. Proc. Natl. Acad. Sci. U S A *111*, 10185–10190.

Zhang, J., Tian, X., Zhang, H., Teng, Y., Li, R., Bai, F., Elankumaran, S., and Xing, J. (2014). TGF-β-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. Sci. Signal. *7*, ra91.

Zhao, H., Kang, X., Xia, X., Wo, L., Gu, X., Hu, Y., Xie, X., Chang, H., Lou, L., and Shen, X. (2016). mir-145 suppresses breast cancer cell migration by targeting FSCN-1 and inhibiting epithelial-mesenchymal transition. Am. J. Transl. Res. *8*, 3106.

ISCI, Volume *21*

Supplemental Information

Exposing the Underlying Relationship of Cancer Metastasis to Metabolism and Epithelial-Mesenchymal Transitions

Xin Kang, Jin Wang, and Chunhe Li

Supplemental Information

Supplemental Figures:

Supplemental Tables

Transparent Methods

Supplemental References

Supplemental Figures

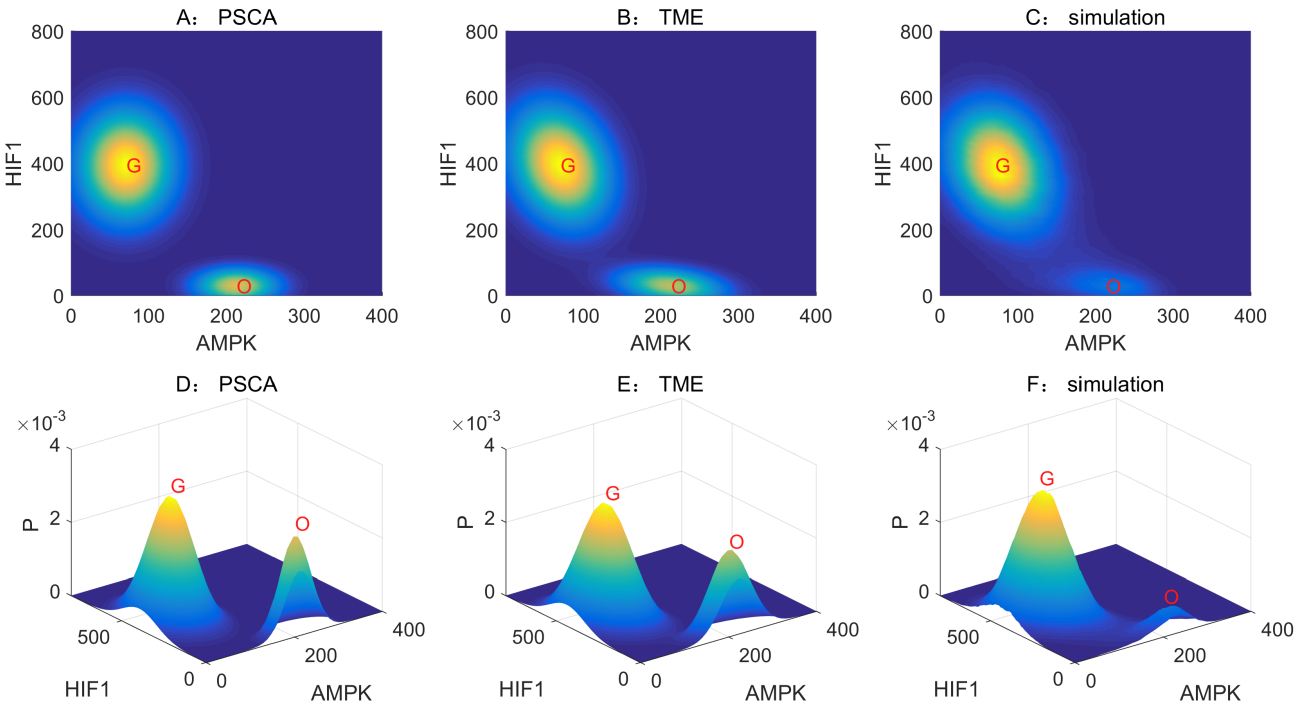


Figure S1: **Comparisons of different methods for calculating the probability distribution (P) of the metabolic model (related to Figure 2).** The distributions calculated from PSCA, TME and simulation methods are shown in 2-dimensional (A-C) and 3-dimensional (D-F) pictures. The relative distance and relative entropy between PSCA and simulations are $\sigma = 0.0140$ and $D_{KL} = 0.1636$; the relative distance and relative entropy between TME and simulations are $\sigma = 0.0109$ and $D_{KL} = 0.0990$. The relative distance between TME and simulations is 22.1% less than that between PSCA and simulations, and the relative entropy between TME and simulation is 39.5% less than that between PSCA and simulations. G: Glycolysis state, O: OXPHOS state.

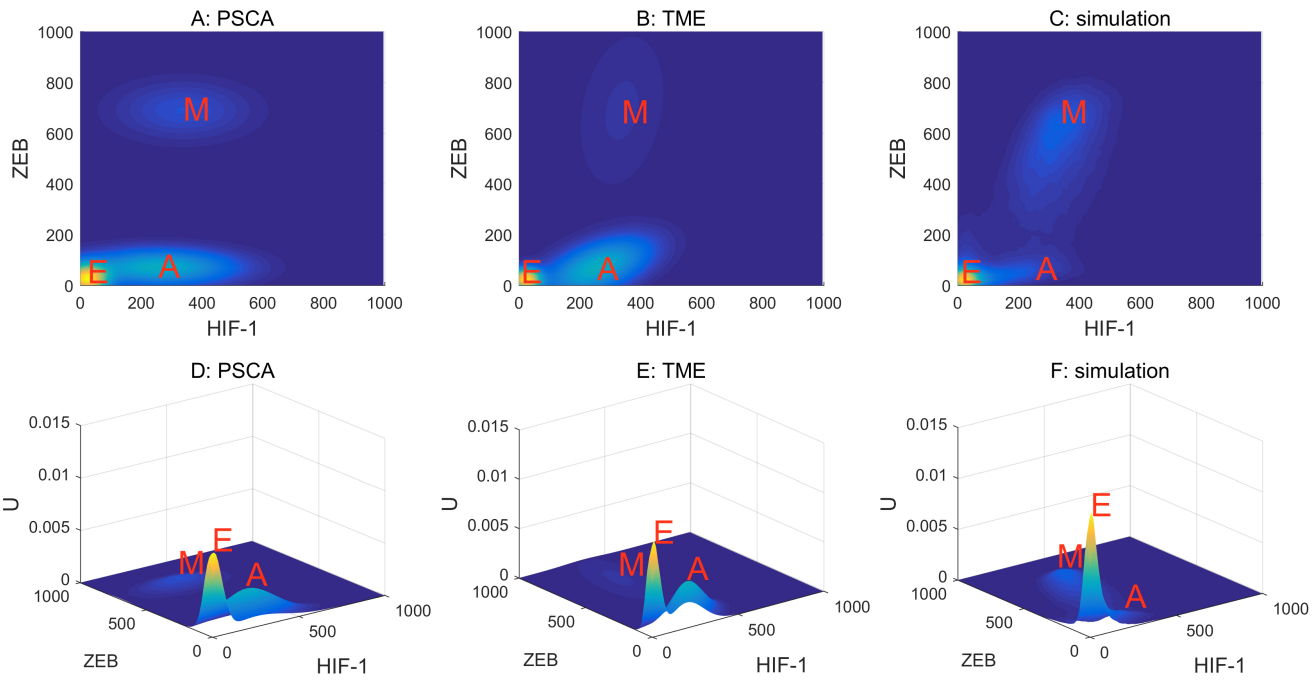


Figure S2: **Comparisons of different methods for calculating the probability distribution (P) of EMT-metabolism model (related to Figure 2).** The distributions calculated from PSCA, TME and simulation methods are shown in 2-dimensional (A-C) and 3-dimensional (D-F) pictures. The relative distance and relative entropy between PSCA and simulations are $\sigma = 0.3350$ and $D_{KL} = 1.7706$; the relative distance and relative entropy between TME and simulations are $\sigma = 0.3240$ and $D_{KL} = 0.6348$. The relative distance between TME and simulations is 3.40% less than that between PSCA and simulations, and the relative entropy between TME and simulation is 64.15% less than that between PSCA and simulations. E: Epithelial state; A: Abnormal metabolic state; M: Mesenchymal state.

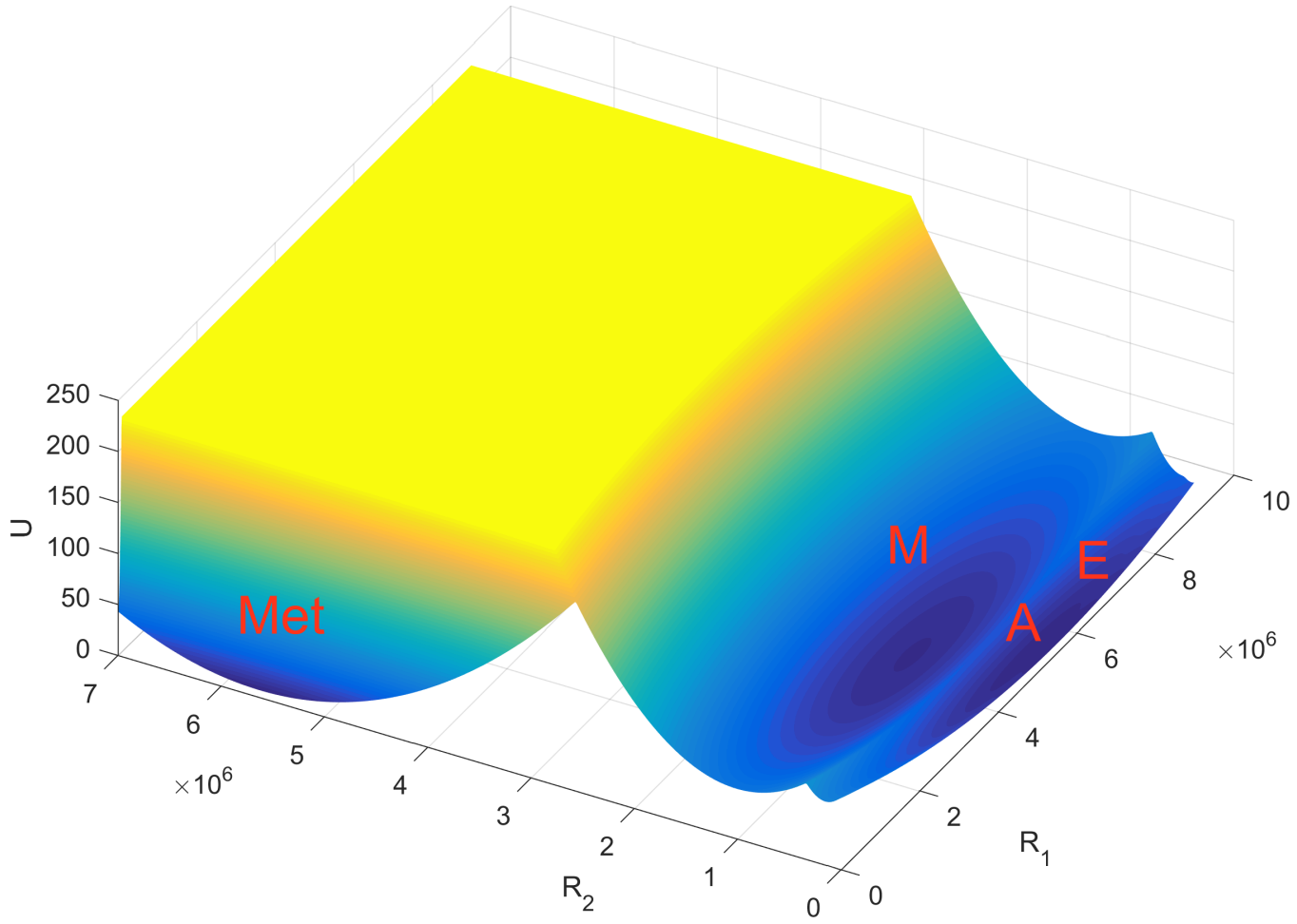


Figure S3: **The landscape in metabolism-EMT-metastasis model for four stable states using R_1 and R_2 as coordinates (related to Figure 2).** R_1: the square of the distance between the independent variable X and Met state, R_2: the square of the distance between the independent variable X and E state. E: Epithelial state; A: Abnormal metabolic state; M: Mesenchymal state; Met: Metastasis state.

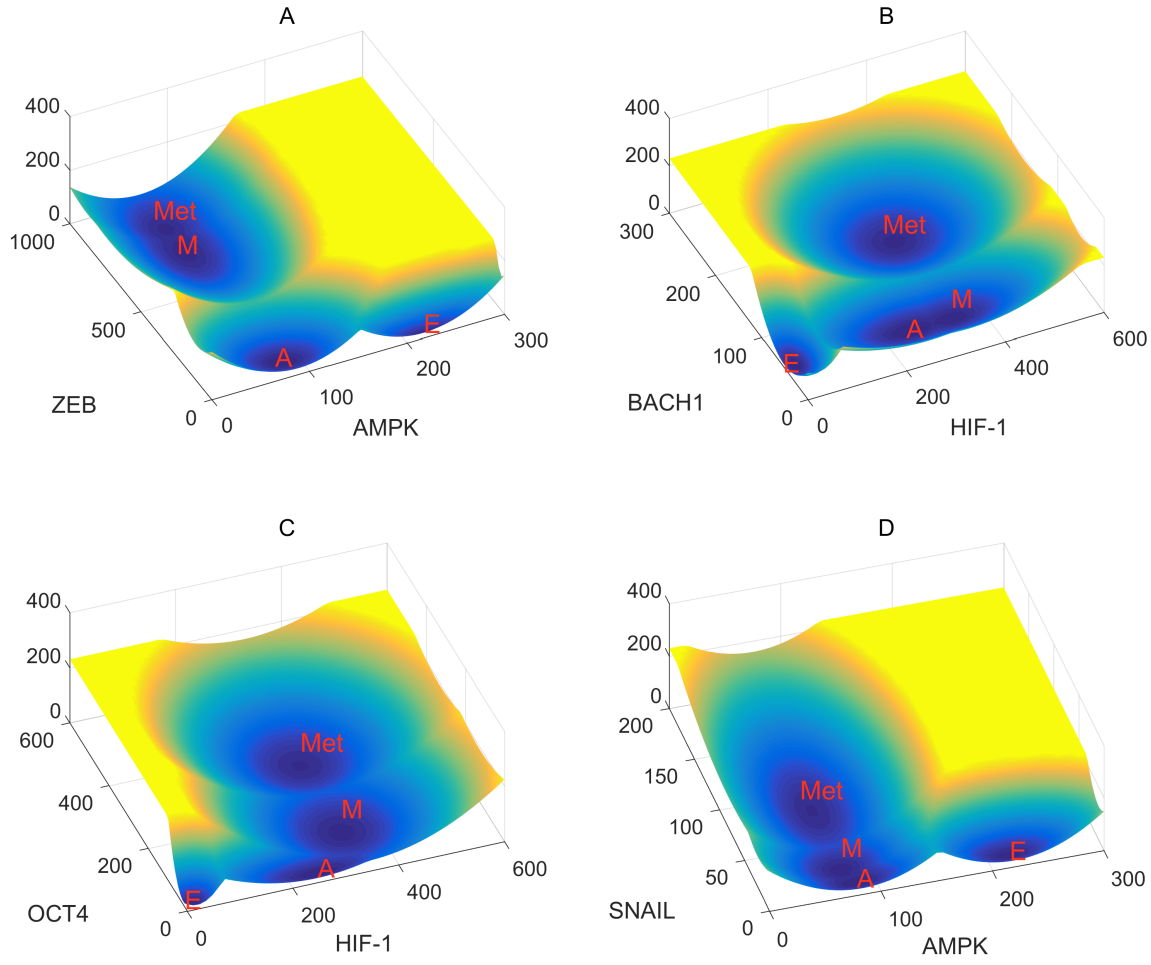


Figure S4: **Landscape using different pairs of variables in metabolism-EMT-metastasis model (related to Figure 2).** A: Landscape using AMPK and ZEB as the coordinates; B: Landscape using HIF-1 and BACH1 as the coordinates; C: Landscape using HIF-1 and OCT4 as the coordinates; D: Landscape using AMPK and SNAIL as the coordinates. E: Epithelial state; A: Abnormal metabolic state; M: Mesenchymal state; Met: Metastasis state.

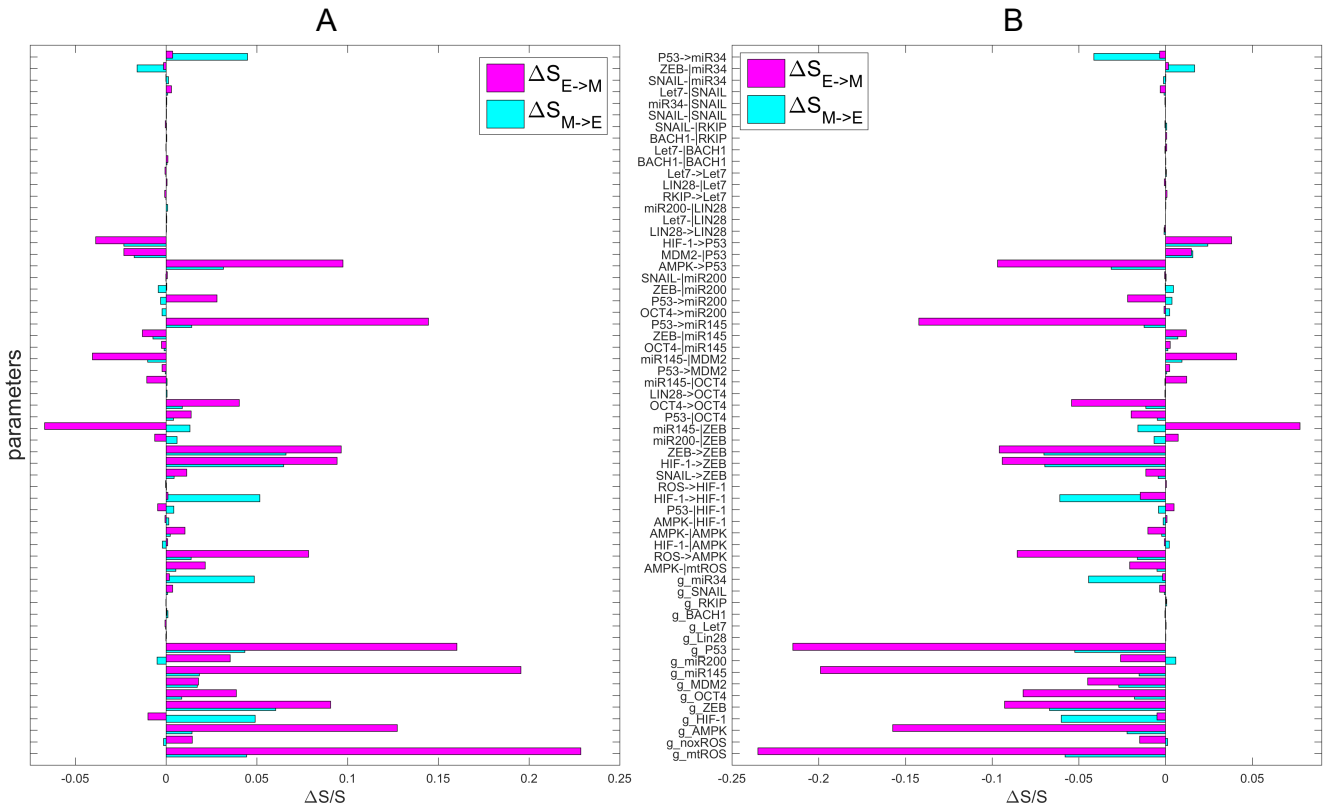


Figure S5: **Sensitivity analysis for the 61 key parameters in metabolism-EMT-metastasis model (16 production rate constants, 45 fold changes for different components including 18 activation constants and 27 inhibition constants) on the transition action (related to Figure 5).** Y-axis represents the 61 key parameters. X-axis represents the percentage of the relative change of the transition action S. Here, $\mathrm{S_{E->M}}$ represents the transition action from Epithelial state to Mesenchymal state (magenta bars), and $\mathrm{S_{M->E}}$ represents the transition action from Mesenchymal state to Epithelial state (cyan bars). '$->$' represents activation and '$-|$' represents inhibition. (A) Each parameter is increased by 10%, individually. (B) Each parameter is decreased by 10%, individually.

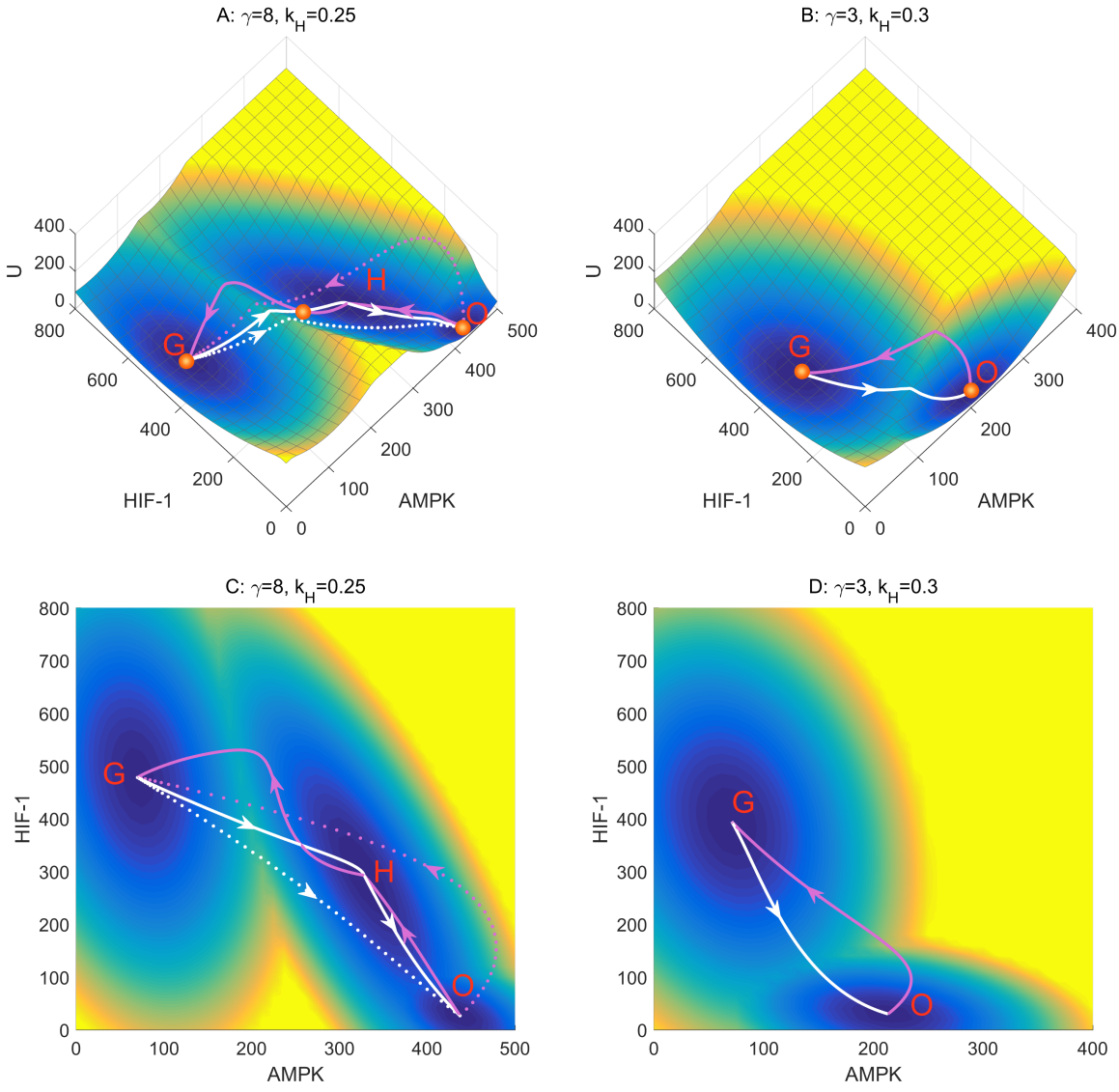


Figure S6: **Landscape and corresponding minimum action paths (MAP) when $d = 15$ for metabolic model are shown in 3-dimensional (A and B) and 2-dimensional figures (C and D) (related to Figure 2).** A and C: The metabolic model has 3 stable states ($\gamma = 8, k_H = 0.25$). B and D: The metabolic model has 2 stable states ($\gamma = 3, k_H = 0.3$). Magenta solid lines represent the MAP from O state to H states, and from H state to G state, and the white solid lines represent the MAP from G state to H state, and from H state to O state. The dashed lines represent the direct MAP from O to G and from G to O states, respectively. O: Oxidative Phosphorylation, G: Glycolysis, H: Hybrid phenotype.

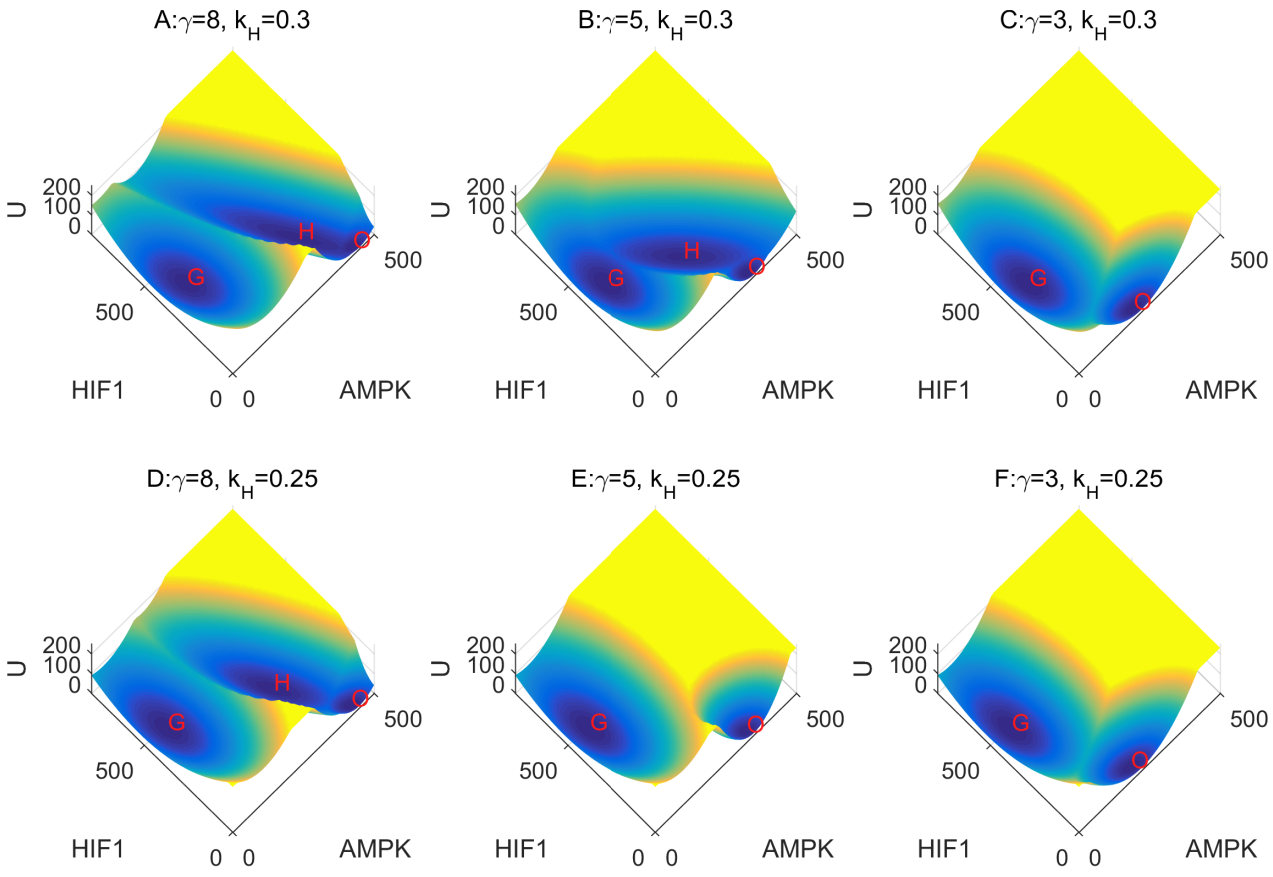


Figure S7: **Landscape of metabolic model changes as γ and k_H change when $d = 15$, which are shown in 3-dimensional space (related to Figure 2).** O: Oxidative Phosphorylation, G: Glycolysis, H: Hybrid phenotype.

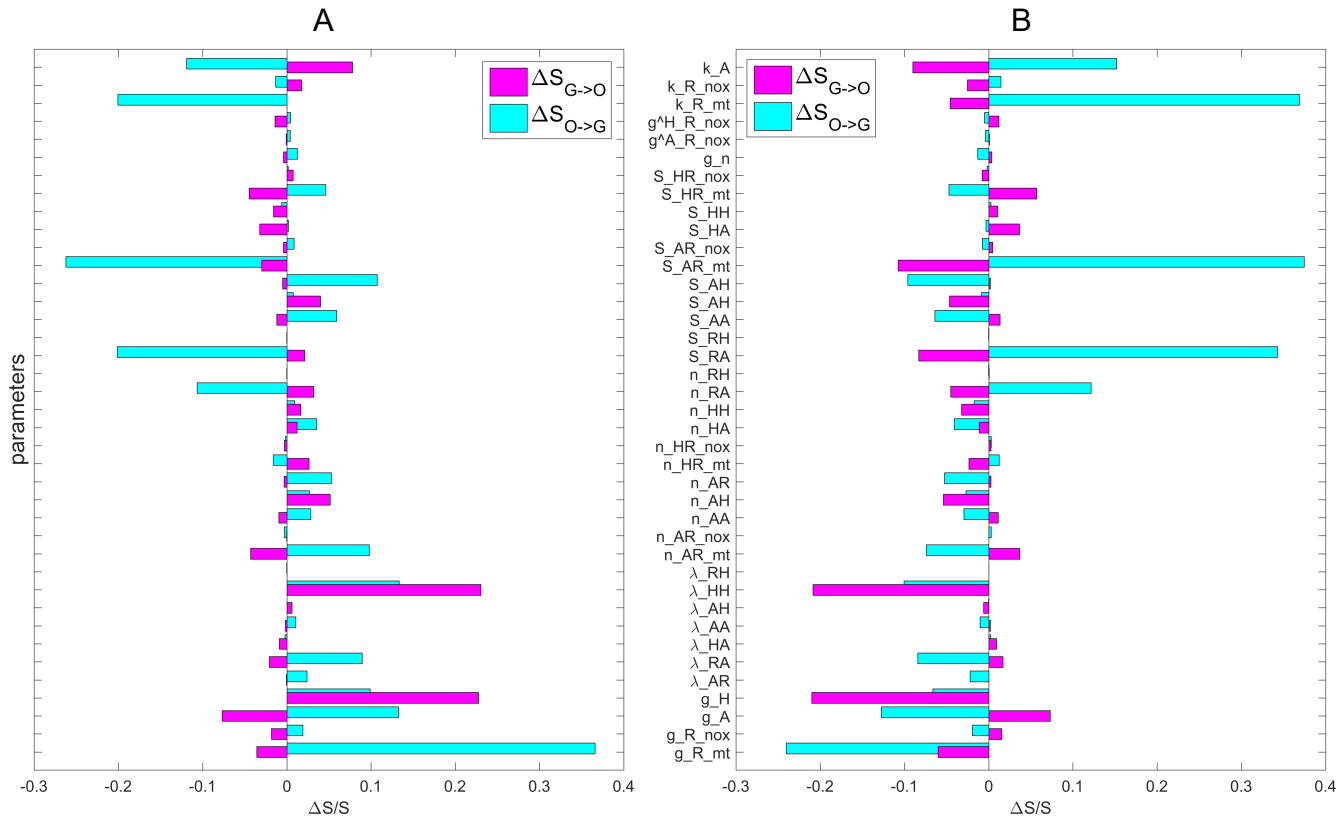


Figure S8: **Sensitivity analysis for the 39 parameters based on the transition action for the metabolic model (related to Figure 5).** Here, $\mathrm{S_{G->O}}$ represents the transition action from Glycolysis state to Oxidative Phosphorylation state (magenta bars), and $\mathrm{S_{O->G}}$ represents the transition action from Oxidative Phosphorylation state to Glycolysis state (cyan bars). Y-axis represents the 39 parameters. X-axis represents the percentage of the relative change of the transition action S. A: Each parameter is increased by 10%, individually. B: Each parameter is decreased by 10%, individually.

Supplemental Tables

Table S1: **Connections and corresponding evidences for the whole network (related to Figure 1)**. There are 22 activation regulations and 29 inhibition regulations. For the column of interaction type, 1 represents activation, and -1 represents inhibition.

Source Node	Target Node	Interaction Type	References
AMPK	noxROS	-1	(Yu et al., 2017)
HIF-1	noxROS	1	(Yu et al., 2017)
AMPK	mtROS	1	(Yu et al., 2017)
AMPK	mtROS	-1	(Yu et al., 2017)
HIF-1	mtROS	-1	(Yu et al., 2017)
noxROS	AMPK	1	(Yu et al., 2017)
mtROS	AMPK	1	(Yu et al., 2017)
AMPK	AMPK	-1	(Yu et al., 2017)
HIF-1	AMPK	-1	(Yu et al., 2017)
noxROS	HIF-1	1	(Yu et al., 2017)
mtROS	HIF-1	1	(Yu et al., 2017)
AMPK	HIF-1	-1	(Yu et al., 2017)
HIF-1	HIF-1	1	(Yu et al., 2017)
P53	HIF-1	-1	(Zhou et al., 2015; Sermeus and Michiels, 2011)
AMPK	P53	1	(Seppet et al., 2009)
HIF-1	P53	1	(Zhou et al., 2015; Sermeus and Michiels, 2011)
MDM2	P53	-1	(Abida et al., 2007)
P53	MDM2	1	(Robbins and Zhao, 2012)
miR-145	MDM2	-1	(Zhang et al., 2013)
miR-145	OCT4	-1	(Xu et al., 2009; Lu et al., 2014)
P53	OCT4	-1	(Li et al., 2012)
OCT4	OCT4	1	(Lu et al., 2014)
LIN28	OCT4	1	(Lu et al., 2014)
HIF-1	ZEB	1	(Zhang et al., 2015)
ZEB	ZEB	1	(Lu et al., 2014)
miR-145	ZEB	-1	(Ren et al., 2014; Lu et al., 2014)
miR-200	ZEB	-1	(Kim et al., 2011; Kong et al., 2010; Lu et al., 2014)
SNAIL	ZEB	1	(Lu et al., 2014)
P53	miR-145	1	(Zhang et al., 2013)
OCT4	miR-145	-1	(Xu et al., 2009; Lu et al., 2014)
ZEB	miR-145	-1	(Ren et al., 2014; Lu et al., 2014)
ZEB	miR-200	-1	(Brabletz and Brabletz, 2010; Lu et al., 2014)
P53	miR-200	1	(Kim et al., 2011)
OCT4	miR-200	1	(Lu et al., 2014; Wang et al., 2013)
SNAIL	miR-200	-1	(Lu et al., 2014)
SNAIL	SNAIL	-1	(Lu et al., 2014)
miR-34	SNAIL	-1	(Lu et al., 2014)
Let7	SNAIL	-1	(Lu et al., 2014)

P53	miR-34	1	(Liang et al., 2009)
ZEB	miR-34	-1	(Lu et al., 2014)
SNAIL	miR-34	-1	(Lu et al., 2014)
Let7	Let7	1	(Lu et al., 2014; Zisoulis et al., 2012)
RKIP	Let7	1	(Lee et al., 2014)
LIN28	Let7	-1	(Zisoulis et al., 2012; Lu et al., 2014)
SNAIL	RKIP	-1	(Lu et al., 2014)
BACH1	RKIP	-1	(Lee et al., 2014)
miR-200	LIN28	-1	(Kong et al., 2010)
Let7	LIN28	-1	(Lu et al., 2014)
LIN28	LIN28	1	(Lu et al., 2014)
Let7	BACH1	-1	(Lee et al., 2014)
BACH1	BACH1	-1	(Lee et al., 2014)

Table S2: **Parameters for the metabolic model(Yu et al., 2017) (related to Figure 1)**. The parameters k_H (degradation rate of HIF-1) and γ (fold change for mtROS activation) are different for normal cells and cancer cells. When the values of these two parameters change from$\{0.25, 8\}$ to $\{0.3, 3\}$, the system will change from a tristable state (G, H and O state, cancer cells) to a bistable state (G and O state, normal cells).

Parameters	Value	Unit	Description	Reference(s)
AMPK				
g_A	30	nM/h	Production rate of AMPK	(Yu et al., 2017)
k_A	0.2	/h	Degradation rate of AMPK	(Yu et al., 2017)
S_{AA}	350	nM	Threshold for self-inhibition	(Yu et al., 2017)
S_{AH}	250	nM	Threshold for HIF-1 inhibition	(Yu et al., 2017)
S_{AR}	350	nM	Threshold for mtROS inhibition	(Yu et al., 2017)
$S_{AR_{mt}}$	150	nM	Threshold for mtROS activation	(Yu et al., 2017)
$S_{AR_{nox}}$	150	nM	Threshold for noxROS inhibition	(Yu et al., 2017)
λ_{AA}	0.2	-	Fold change for self-inhibition	(Yu et al., 2017)
λ_{AH}	0.1	-	Fold change for HIF-1 inhibition	(Yu et al., 2017)
λ_{AH}	0.25	-	Fold change for mtROS inhibition	(Yu et al., 2017)
γ	0.3	-	Fold change for mtROS inhibition	(Yu et al., 2017)
$g^A_{R_{nox}}$	0.2	-	Fold change for noxROS inhibition	(Yu et al., 2017)
n_{AA}	2	-	Hill coefficient for self-inhibition	(Yu et al., 2017)
n_{AH}	1	-	Hill coefficient for HIF-1 inhibition	(Yu et al., 2017)
n_{AR}	2	-	Hill coefficient for mtROS inhibition	(Yu et al., 2017)
$n_{AR_{mt}}$	4	-	Hill coefficient for mtROS activation	(Yu et al., 2017)
$n_{AR_{nox}}$	2	-	Hill coefficient for noxROS inhibition	(Yu et al., 2017)
HIF-1				
g_H	15	nM/h	Production rate of HIF-1	(Yu et al., 2017)
S_{HH}	80	nM	Threshold for HIF-1 self-activation	(Yu et al., 2017)
S_{HA}	250	nM	Threshold for AMPK inhibition	(Yu et al., 2017)
$S_{HR_{mt}}$	200	nM	Threshold for mtROS inhibition	(Yu et al., 2017)
$S_{HR_{nox}}$	250	nM	Threshold for noxROS activation	(Yu et al., 2017)
λ_{HH}	10	-	Fold change for HIF-1 self-activation	(Yu et al., 2017)
λ_{HA}	0.1	-	Fold change for AMPK inhibition	(Yu et al., 2017)
$g^H_{R_{nox}}$	5	-	Fold change for noxROS activation	(Yu et al., 2017)

n_{HH}	4	-	Hill coefficient for HIF-1 self-activation	(Yu et al., 2017)
n_{HA}	0.1	-	Hill coefficient for AMPK inhibition	(Yu et al., 2017)
$n_{HR_{mt}}$	2	-	Hill coefficient for mtROS inhibition	(Yu et al., 2017)
$n_{HR_{nox}}$	2	-	Hill coefficient for noxROS inhibition	(Yu et al., 2017)
k_H	0.3	-	Degradation rate of HIF-1	(Yu et al., 2017)
ROS				
$g_{R_{mt}}$	150	μM/min	Production rate of mitochondrial ROS	(Yu et al., 2017)
$g_{R_{nox}}$	40	μM/min	Production rate of NOX derived ROS	(Yu et al., 2017)
g_n	0.2	-	Basal cytosol ROS	(Yu et al., 2017)
$k_{R_{mt}}$	5.0	/min	Degradation rate of mitochondrial ROS	(Yu et al., 2017)
$k_{R_{nox}}$	5.0	/min	Degradation rate of NOX derived ROS	(Yu et al., 2017)
S_{RA}	100	μM	Threshold for AMPK activation	(Yu et al., 2017)
S_{RH}	300	μM	Threshold for HIF-1 activation	(Yu et al., 2017)
λ_{RA}	8	-	Fold change for AMPK activation	(Yu et al., 2017)
λ_{RH}	5	-	Fold change for HIF-1 activation	(Yu et al., 2017)
n_{RA}	4	-	Hill coefficient for AMPK activation	(Yu et al., 2017)
n_{RH}	4	-	Hill coefficient for HIF-1 activation	(Yu et al., 2017)
g_0	1	-		

Table S3: **Parameters for the metabolism-EMT-metastasis model (related to Figure 1)**. The system has a quadrastable state (E, A, M and Met State).

Parameter	Value	Unit	Description	Reference(s)
AMPK				
S_{AP}	200	nM	Threshold for P53 activation	
n_{AP}	300	nM	Hill coefficient for P53 activation	
λ_{AP}	3	-	Fold change for P53 activation	
HIF-1				
S_{HZ}	200	nM	Threshold for ZEB activation	
S_{HP}	300	nM	Threshold for P53 activation	
λ_{HZ}	3	-	Fold change for ZEB activation	
λ_{HP}	10/7	-	Fold change for P53 activation	
n_{HZ}	4	-	Hill coefficient for ZEB activation	
n_{HP}	4	-	Hill coefficient for P53 activation	
ZEB				
S_{ZZ}	200	nM	Threshold for ZEB self-activation	
S_{Zm1}	100	nM	Threshold for miR-145 inhibition	
S_{Zm2}	300	nM	Threshold for miR-200 inhibition	
S_{Zm3}	200	nM	Threshold for miR-34 inhibition	
λ_{Zm2}	0.2	-	Fold change for miR-200 inhibition	
λ_{ZZ}	3	-	Fold change for ZEB self-activation	
λ_{Zm1}	0.2	-	Fold change for miR-145 inhibition	
λ_{Zm3}	0.3	-	Fold change for miR-34 inhibition	
n_{ZZ}	2	-	Hill coefficient for ZEB self-activation	(Lu et al., 2014)
n_{Zm1}	3	-	Hill coefficient for miR-145 inhibition	
n_{Zm2}	3	-	Hill coefficient for miR-200	(Lu et al., 2014)
n_{Zm3}	2	-	Hill coefficient for miR-34 inhibition	

g_Z	10	nM/h	Production rate of ZEB	
k_Z	0.1	/h	Degradation rate of ZEB	
P53				
S_{Pm1}	250	nM	Threshold for miR-145 inhibition	
S_{PH}	200	nM	Threshold for HIF-1 inhibition	
S_{PM}	500	nM	Threshold for MDM2 activation	
S_{PO}	100	nM	Threshold for OCT4 inhibition	
S_{Pm2}	200	nM	Threshold for miR-200 activation	
λ_{Pm2}	3	-	Fold change for miR-200 activation	
λ_{PO}	0.3	-	Fold change for OCT4 inhibition	
λ_{Pm1}	3	-	Fold change for miR-145 activation	
λ_{PM}	2	-	Fold change for MDM2 activation	
λ_{PH}	0.5	-	Fold change for HIF-1 inhibition	
n_{Pm2}	3	-	Hill coefficient for miR-200 activation	
n_{PH}	4	-	Hill coefficient for HIF-1 inhibition	
n_{Pm1}	4	-	Hill coefficient for miR-145 activation	
n_{PM}	4	-	Hill coefficient for MDM2 activation	
n_{PO}	3	-	Hill coefficient for OCT4 inhibition	
g_P	30	nM/h	Production rate of P53	
k_P	0.1	/h	Degradation rate of P53	
miR-200				
S_{m2Z}	300	nM	Threshold for ZEB inhibition	
S_{m2Li}	400	nM	Threshold for LIN28 inhibition	
λ_{m2Z}	1/3	-	Fold change for ZEB inhibition	
λ_{m2Li}	0.5	-	Fold change for LIN28 inhibition	
n_{m2Z}	6	-	Hill coefficient for ZEB inhibition	(Lu et al., 2014)
n_{m2Li}	4	-	Hill coefficient for LIN28 inhibition	
g_{m2}	10	nM/h	Production rate of miR-200	
k_{m2}	0.1	/h	Degradation rate of miR-200	
miR-145				
S_{m1Z}	200	nM	Threshold for ZEB inhibition	
S_{m1O}	200	nM	Threshold for OCT4 inhibition	
S_{m1M}	200	nM	Threshold for MDM2 inhibition	
λ_{m1O}	0.25	-	Fold change for OCT4 inhibition	
λ_{m1M}	0.2	-	Fold change for MDM2 inhibition	
λ_{m1Z}	0.2	-	Fold change for ZEB inhibition	
n_{m1M}	4	-	Hill coefficient for MDM2 inhibition	
n_{m1Z}	2	-	Hill coefficient for ZEB inhibition	
n_{m1O}	3	-	Hill coefficient for OCT4 inhibition	
g_{m1}	20	nM/h	Production rate of miR-145	
k_{m1}	0.1	/h	Degradation rate of miR-145	
MDM2				
S_{MP}	400	nM	Threshold for P53 inhibition	
λ_{MP}	1/6	-	Fold change for P53 inhibition	
n_{MP}	2	-	Hill coefficient for P53 inhibition	
k_M	0.1	/h	Degradation rate of MDM2	
g_M	20	nM/h	Production rate of MDM2	

OCT4				
S_{Om1}	200	nM	Threshold for miR-145 inhibition	
S_{Om2}	300	nM	Threshold for miR-200 activation	
S_{OO}	200	nM	Threshold for OCT4 self-activation	
S_{OLi}	200	nM	Threshold for LIN28 activation	
λ_{Om1}	0.2	-	Fold change for miR-145 inhibition	
λ_{OO}	3	-	Fold change for OCT4 self-activation	
λ_{Om2}	2	-	Fold change for miR-200 activation	
λ_{OLi}	4	-	Fold change for LIN28 activation	
n_{OO}	3	-	Hill coefficient for OCT4 self-activation	
n_{Om2}	3	-	Hill coefficient for miR-200 activation	
n_{Om1}	2	-	Hill coefficient for miR-145 inhibition	
n_{OLi}	4	-	Hill coefficient for LIN28 activation	
g_O	20	nM/h	Production rate of OCT4	
k_O	0.1	/h	Degradation rate of OCT4	
Let7				
S_{LeLe}	450	nM	Threshold for Let7 self-activation	
S_{LeLi}	200	nM	Threshold for LIN28 inhibition	
S_{LeS}	200	nM	Threshold for SNAIL inhibition	
S_{LeB}	300	nM	Threshold for BACH1 inhibition	
n_{LeLi}	4	-	Fold change for LIN28 inhibition	
n_{LeLe}	4	-	Fold change for Let7 self-activation	
n_{LeB}	4	-	Hill coefficient for BACH1 inhibition	
n_{LeS}	4	-	Hill coefficient for SNAIL inhibition	
k_{Le}	0.1	/h	Degradation rate of Let7	
g_{Le}	20	nM/h	Production rate of Let7	
λ_{LeLi}	0.25	-	Fold change for LIN28 inhibition	
λ_{LeLe}	4	-	Fold change for Let7 self-activation	
λ_{LeB}	0.25	-	Fold change for BACH1 inhibition	
λ_{LeS}	0.25	-	Fold change for SNAIL inhibition	
LIN28				
S_{LiLi}	200	nM	Fold change for LIN28 self-activation	
S_{LiLe}	200	nM	Fold change for Let7 inhibition	
S_{LiO}	400	nM	Fold change for OCT4 activation	
n_{LiLi}	4	-	Hill coefficient for LIN28 self-activation	
n_{LiO}	4	-	FHill coefficient for OCT4 activation	
n_{LiLe}	4	-	Hill coefficient for Let7 inhibition	
g_{Li}	20	Mn/h	Production rate of LIN28	
k_{Li}	0.1	/h	Degradation rate of LIN28	
λ_{LiLe}	0.3	-	Fold change for Let7 inhibition	
λ_{LiO}	1.4	-	Fold change for OCT4 activation	
λ_{LiLi}	5	-	Fold change for LIN28 self-activation	
BACH1				
S_{BB}	200	nM	Fold change for BACH1 self-inhibition	
S_{BRK}	200	nM	Fold change for PRKIP inhibition	
n_{BRK}	4	-	Hill coefficient for RKIP inhibition	
n_{BB}	3	-	Hill coefficient for BACH1 self-inhibition	

k_B	0.1	/h	Degradation rate of BACH1	
g_B	20	nM/h	Production rate of BACH1	
λ_{BB}	0.7	-	Fold change for BACH1 self-inhibition	
λ_{BRK}	0.2	-	Fold change for RKIP inhibition	
RKIP				
S_{RKLe}	200	nM	Threshold for Let7 activation	
n_{RKLe}	5	-	Hill coefficient for Let7 activation	
k_{RK}	0.1	/h	Degradation rate of RKIP	
g_{RK}	20	nM/h	Production rate of RKIP	
λ_{RKLe}	4	-	Fold change for Let7 activation	
SNAIL				
S_{SZ}	400	nM	Threshold for ZEB activation	
S_{Sm2}	400	nM	Threshold for miR-200 inhibition	
S_{Sm3}	200	nM	Threshold for miR-34 inhibition	
S_{SS}	100	nM	Threshold for SNAIL self-inhibition	
S_{SR}	200	nM	Threshold for RKIP inhibition	
n_{SR}	4	-	Hill coefficient for RKIP inhibition	
n_{Sm2}	1	-	Hill coefficient for miR-200 inhibition	
n_{Sm3}	1	-	Hill coefficient for miR-34 inhibition	
n_{SZ}	1	-	Hill coefficient for ZEB activation	
n_{SS}	1	-	Hill coefficient for SNAIL self-inhibition	
k_S	0.1	/h	Degradation rate of SNAIL	
g_S	20	nM/h	Production rate of SNAIL	
λ_{Sm2}	0.8	-	Fold change for miR-200 inhibition	
λ_{Sm3}	0.7	-	Fold change for miR-34 inhibition	
λ_{SR}	0.3	-	Fold change for RKIP inhibition	
λ_{SS}	0.3	-	Fold change for SNAIL self-inhibition	
λ_{SZ}	1.5	-	Fold change for ZEB activation	
miR-34				
S_{m3S}	200	nM	Threshold for SNAIL inhibition	
n_{m3S}	2	nM	Hill coefficient for SNAIL inhibition	
k_{m3}	0.1	/h	Degradation rate of miR-34	
g_{m3}	15	nM/h	Production rate of miR-34	
λ_{m3S}	0.2	-	Fold change for SNAIL inhibition	

Table S4: **The gene expression levels of the stable steady states for the quadrastable landscape (related to Figure 2)**. E: Epithelial state; A: Abnormal metabolic state; M: Mesenchymal state; Met: Metastasis state.

Gene	E	A	M	Met
mtROS	60.70	9.31	5.87	5.86
noxROS	3.70	21.39	26.42	26.44
AMPK	219.96	83.08	75.22	75.17
HIF-1	19.45	255.68	336.83	337.37
ZEB	19.44	73.24	710.26	837.88
OCT4	19.44	23.87	121.91	333.22
MDM2	33.48	39.03	150.38	151.19
SNAIL	10.56	10.99	24.68	82.66
miR-145	373.05	260.53	39.30	20.53
miR-200	223.55	206.91	43.31	57.55
miR-34	535.46	443.52	101.98	92.03
P53	236.61	214.32	151.99	151.53
RKIP	199.39	199.39	199.36	139.85
Let7	1982.80	1982.43	1980.57	86.61
LIN28	46.46	47.47	50.83	905.40
BACH1	49.79	49.79	49.79	172.88

Table S5: **Transition actions for the transitions among 4 states for the metabolism-EMT-metastasis model (related to Figure 2).** The transition actions are represented by a 4 $\times$ 4 matrix S, with S^{ij} representing the transition action for the transition from i state to j state. N denotes not applicable.

S^{ij}	E state	A state	M state	Met state
E state	N	11780($S_{E->A}$)	118676($S_{E->M}$)	389111($S_{E->Met}$)
A state	6788($S_{A->E}$)	N	49365($S_{A->M}$)	323344($S_{A->Met}$)
M state	54160($S_{M->E}$)	31300($S_{M->A}$)	N	267886($S_{M->Met}$)
Met state	294393($S_{Met->E}$)	272921($S_{Met->A}$)	246635($S_{Met->M}$)	N

Table S6: **The evidences for the conclusion in the sensitivity analysis for metabolic model (related to Figure 5)**.

Gene	Model predictions	Experimental phenotype	Reference(s)
mtROS	The increasing of production rate of mtROS $g_{r_{mt}}$(or the decreasing in the degradation rate of mtROS $k_{r_{mt}}$ or the decreasing in threshold for the regulation from AMPK to mtROS) can increase the transition action from O state to G state, which means cells tend to stay in O state.	Generation of mtROS mainly takes place at the ETC (mitochondrial electron transport chain) located on the inner mitochondrial membrane during the process of oxidative phosphorylation (OXPHOS).	(Li et al., 2013)
AMPK	Increasing in production rate of AMPK g_A, decreasing in degradation rate of AMPK or decreasing threshold for the effect from mtROS and noxROS to AMPK lead to the changes of transition action: the transition action from O state to G state increases, while the transition action from G state to O state decreases, making the O state more stable.	Loss of AMPKα promotes a glycolytic signature and increases the expression of HIF-1α.	(Faubert et al., 2013)
HIF-1	Increasing in production rate of HIF-1 g_H can increase the transition action from G state to O state, making G state more stable.	Experiment data indicates that HIF-1 is both necessary and sufficient for reducing mitochondrial oxygen consumption in hypoxia.	(Papandreou et al., 2006) (Kim et al., 2006)

Table S7: **Parameters for modeling anti-cancer therapeutic strategies (related to Figure 7).**

Parameters	Value	Unit	Description	Reference(s)
λ_{MA}	2	-	Fold change for AMPK activation	(Yu et al., 2017)
λ_{HA}	0.5	-	Fold change for HIF-1 inhibition	(Yu et al., 2017)
n_{MA}	2	-	Hill coefficient for AMPK activation	(Yu et al., 2017)
n_{HA}	2	-	Hill coefficient for HIF-1 inhibition	(Yu et al., 2017)
S_{MA}	300	μM	Threshold for AMPK activation	(Yu et al., 2017)
S_{MH}	300	μM	Threshold for HIF-1 inhibition	(Yu et al., 2017)
$\alpha_{metformin}$	0.025	$/\mu M$	Influence on mtROS maximum fold change	(Yu et al., 2017)
α_{3BP}	0.0005	$/\mu M$	Influence on HIF-1 fold change due to self-activation	(Yu et al., 2017)
α_{BI}	0.04	$/\mu M$	Influence on BACH1 maximum fold change	

Transparent Methods

Section S1: Parameter setting for the models

We construct a metabolism-EMT-metastasis gene regulatory network by mining literatures (Figure 1 in main text, the connections and corresponding evidences are shown in Table S1), which includes 16 components and 51 regulations. The network consists of three circuits: metabolic circuit, EMT circuit and metastasis circuit. In this study, we analyze the dynamical behavior of the metabolism-EMT-metastasis regulatory network by solving ordinary differential equations (ODEs). Typically, the deterministic rate equation for each component has a general form:

$$\frac{dX}{dt} = g_X \cdot G - k_X \cdot X \cdot K,$$

where X represents the expression level of a gene or a metabolite. g_X and k_X denote the basal production rate and degradation rate of $X(t)$, respectively. G and K denote regulations of other components on the production and degradation of X. The regulation of component Y on component X can be described by a nonlinear function, namely the shifted Hill function: $H^s(Y, S, \lambda, n) = 1 + (\lambda - 1)\frac{Y^n}{S^n + Y^n}$(Yu et al., 2017). Here, λ represents the fold change for regulation, S represents the threshold of a sigmoidal function, and n is the Hill coefficient, which determines the steepness of the sigmoidal function. Therefore, the Hill function has the following property:

$$H^s(Y, S, \lambda, n) \begin{cases} < 1 & 0 < \lambda < 1, \\ = 1 & \lambda = 1, \\ > 1 & \lambda > 1. \end{cases}$$

When Y regulates the production rate of X, we use $H^s(Y, S, \lambda, n)$ to denote this regulation. When Y regulates the degradation rate of X, we use $H^s(Y, S, 1/\lambda, n)$ to express this regulation. By doing so, λ in the function has the consistent meaning, i.e., λ decides whether the regulation is activation or inhibition ($\lambda > 1$ for activation and $0 < \lambda < 1$ for inhibition).

When the production (or degradation) of X is regulated by two components Y and Z simultaneously, G (or K) can be expressed as(Yu et al., 2017):

$$G(\text{or } K) = \begin{cases} H^s(Y, S_{YX}, \lambda_{YX}, n_{YX}) H^s(Z, S_{ZX}, \lambda_{ZX}, n_{ZX}) & \text{Y and Z are independent,} \\ C^{comp}(\gamma, Y, S_{YX}, \lambda_{YX}, n_{YX}, Z, S_{ZX}, \lambda_{ZX}, n_{ZX}) & \text{Y and Z are competitive.} \end{cases}$$

The specific form of C^{comp} can be found in Section S2.

We determine the ODEs and the parameter values based on the following:

1. We choose parameter values according to some previous works on gene regulatory networks (Li and Wang, 2013; Lu et al., 2013; Huang et al., 2007). The ODEs and parameters for the metabolic model (Figure 1 in main text, blue box) are based on the model in (Yu et al., 2017). The specific ODEs are shown in section S2 and the parameter values for this model are shown in Table S2.
2. To reduce the complexity of the model, we set most of the parameters uniformly, since so far for the metabolism-EMT-metastasis regulatory network there are little information about the regulatory strengths among different components. For example, we set the same degradation rate for most of gene variables, and we set basal production rate for different genes in similar range.
3. The value of Hill coefficient n is determined based on experimental data for certain regulations, e.g. the self-activation of ZEB, the inhibition of miR-200 on ZEB and the inhibition of ZEB on miR-200 (Lu et al., 2014). For the cases without experimental data, we choose Hill coefficient $n \in \{2, 3, 4\}$ to represent high nonlinearity of gene regulations following previous works (Lu et al., 2013; Li and Wang, 2013).
4. We carefully choose the parameter values in the metabolism-EMT-metastasis model to satisfy certain biological constraints, such as generating multistability, which are consistent with biological observations. We perform the sensitivity analysis to the parameters, which supports the robustness of current parameter values for our models.

5. Most of the parameters in the current model have not been determined by experiments. In this work, we focus on the dynamical implications of the regulatory structure of the metabolism-EMT-metastasis model, i.e. the topology of the network. We believe that the topology of the network determines operating principles of networks, as suggested by previous work (Huang et al., 2017; Gérard and Goldbeter, 2009).

Section S2: ODEs of the model

Metabolic model

In metabolic model, AMPK and HIF-1 competitively regulate the production of ROS (both mitochondria ROS and cytosol ROS). ROS and oxygen competitively regulate the degradation of HIF-1(Yu et al., 2017). The specific ODEs of the metabolic model (Figure 1 in main text, blue box) are shown as follows(Yu et al., 2017):

$$\dot{R}_{mt} = g_{R_{mt}} * H^{s-}(A, S_{AR}, \lambda_{AR}, n_{AR}) * C^{comp}_{R_{mt}}(\gamma, g_n, H, S_{HR_{mt}}, n_{HR_{mt}}, A, S_{AR_{mt}}, n_{AR_{mt}}) - k_{R_{mt}} * R_{mt}, \tag{S1}$$

$$\dot{R}_{nox} = g_{R_{nox}} * C^{comp}_{R_{nox}}(g_0, H, S_{HR_{nox}}, n_{HR_{nox}}, g^H_{R_{nox}}, A, S_{AR_{nox}}, g^A_{R_{nox}}, n_{AR_{nox}}) - k_{R_{nox}} * R_{nox}, \tag{S2}$$

$$R = R_{mt} + R_{nox}, \tag{S3}$$

$$\dot{A} = g_A * H^{s+}(R, S_{RA}, \lambda_{RA}, n_{RA}) * H^{s-}(H, S_{HA}, \lambda_{HA}, n_{HA}) * H^{s-}(A, S_{AA}, \lambda_{AA}, n_{AA}) - k_A * A, \tag{S4}$$

$$\dot{H} = g_H * H^{s-}(A, S_{AH}, \lambda_{AH}, n_{AH}) - k_H * H * H^{s-}(H, S_{HH}, 1/\lambda_{HH}, n_{HH}) * H^{s-}(R, S_{RH}, 1/\lambda_{RH}, n_{RH}).$$

Here, R_{mt}, R_{nox}, A, and H represent the levels of mtROS, noxROS, AMPK and HIF-1 respectively. The concrete form of competitive regulations C^{comp} are expressed as(Yu et al., 2017):

$$C^{comp}_{R_{mt}}(\gamma, g_n, H, S_{R_{mt}}, n_{HR_{mt}}, A, S_{AR_{mt}}, n_{AR_{mt}}) = \frac{\gamma(g_n + (\frac{A}{S_{AR_{mt}}})^{n_{AR_{mt}}})}{1 + (\frac{H}{S_{HR_{mt}}})^{n_{HR_{mt}}} + (\frac{A}{S_{AR_{mt}}})^{n_{AR_{mt}}}},$$

$$C^{comp}_{R_{nox}}(g_0, H, S_{HR_{nox}}, n_{HR_{nox}}, g^H_{R_{nox}}, A, S_{AR_{nox}}, g^A_{R_{nox}}, n_{AR_{nox}}) = \frac{g_0 + g^H_{R_{nox}}(\frac{H}{S_{HR_{nox}}})^{n_{HR_{nox}}} + g^A_{R_{nox}}(\frac{A}{S_{AR_{nox}}})^{n_{AR_{nox}}}}{1 + (\frac{H}{S_{HR_{nox}}})^{n_{HR_{nox}}} + (\frac{A}{S_{AR_{nox}}})^{n_{AR_{nox}}}},$$

where H^{s+} and H^{s-} denote the $\lambda > 1$ and $0 < \lambda < 1$ in H^s correspondingly, which indicates the activation and the inhibition regulation of Y on X, respectively.

Metabolism-EMT-metastasis model

To uncover the underlying mechanisms of the interplay among metabolism, EMT and metastasis, we combine these three circuits, including EMT (marked by ZEB, OCT4, MDM2, miR-145, miR-200 miR-34, SNAIL and P53), metastasis (marked by BACH1, RKIP, Let7 and LIN28) and the metabolic model, and construct a metabolism-EMT-metastasis regulatory network (Figure 1 in main text, whole diagram). We use the generic form of deterministic rate equation in (Yu et al., 2017) to construct the metabolism-EMT-metastasis model. The ODEs of AMPK, mtROS and noxROS are the same as Eq. (S1)-(S4), and the modified equations for HIF-1 and the equations for the other components in the whole network are shown as follows:

$$\dot{H} = g_H * H^{s-}(A, S_{AH}, \lambda_{AH}, n_{AH}) * H^{s-}(P, S_{PH}, \lambda_{PH}, n_{PH}) - k_H * H * H^{s-}(H, S_{HH}, 1/\lambda_{HH}, n_{HH}) * H^{s-}(R, S_{RH}, 1/\lambda_{RH}, n_{RH}),$$

$$\dot{Z} = g_Z * H^{s+}(S, S_{SZ}, \lambda_{SZ}, n_{SZ}) * H^{s+}(H, S_{HZ}, \lambda_{HZ}, n_{HZ}) * H^{s+}(Z, S_{ZZ}, \lambda_{ZZ}, n_{ZZ}) - k_Z * Z \\ *H^{s+}(m2, S_{m2Z}, 1/\lambda_{m2Z}, n_{m2Z}) * H^{s+}(m1, S_{m1Z}, 1/\lambda_{m1Z}, n_{m1Z}),$$

$$\dot{O} = g_O * H^{s-}(P, S_{PO}, \lambda_{PO}, n_{PO}) * H^{s+}(O, S_{OO}, \lambda_{OO}, n_{OO}) - k_O * O * H^{s+}(m1, S_{m1O}, 1/\lambda_{m1O}, n_{m1O}),$$

$$\dot{M} = g_M * H^{s+}(P, S_{PM}, \lambda_{PM}, n_{PM}) - k_M * M * H^{s+}(m1, S_{m1M}, 1/\lambda_{m1M}, n_{m1M}),$$

$$\dot{m1} = g_{m1} * H^{s-}(O, S_{Om1}, \lambda_{Om1}, n_{Om1}) * H^{s-}(Z, S_{Zm1}, \lambda_{Zm1}, n_{Zm1}) * H^{s+}(P, S_{Pm1}, \lambda_{Pm1}, n_{Pm1}) - k_{m1} * m1,$$

$$\dot{m2} = g_{m2} * H^{s+}(O, S_{Om2}, \lambda_{Om2}, n_{Om2}) * H^{s+}(P, S_{Pm2}, \lambda_{Pm2}, n_{Pm2}) * H^{s-}(Z, S_{Zm2}, \lambda_{Zm2}, n_{Zm2}) \\ *H^{s-}(S, S_{Sm2}, \lambda_{Sm2}, n_{Sm2}) - k_{m2} * m2,$$

$$\dot{P} = g_P * H^{s+}(A, S_{AP}, \lambda_{AP}, n_{AP}) - k_P * P * H^{s+}(M, S_{MP}, 1/\lambda_{MP}, n_{MP}) * H^{s-}(H, S_{HP}, 1/\lambda_{HP}, n_{HP}),$$

$$\dot{Li} = g_{Li} * H^{s+}(O, S_{OLi}, \lambda_{OLi}, n_{OLi}) * H^{s+}(Li, S_{LiLi}, \lambda_{LiLi}, n_{LiLi}) - k_{Li} * Li * H^{s+}(Le, S_{LeLi}, 1/\lambda_{LeLi}, n_{LeLi}) \\ *H^{s+}(m2, S_{m2Li}, 1/\lambda_{m2Li}, n_{m2Li}),$$

$$\dot{Le} = g_{Le} * H^{s+}(RK, S_{RKLe}, \lambda_{RKLe}, n_{RKLe}) * H^{s-}(Li, S_{LiLe}, \lambda_{LiLe}, n_{LiLe}) * H^{s+}(Le, S_{LeLe}, \lambda_{LeLe}, n_{LeLe}) \\ -k_{Le} * Le,$$

$$\dot{B} = g_B * H^{s+}(B, S_{BB}, \lambda_{BB}, n_{BB}) - k_B * B * H^{s+}(Le, S_{LeB}, 1/\lambda_{LeB}, n_{LeB}),$$

$$\dot{RK} = g_{RK} * H^{s-}(B, S_{BRK}, \lambda_{BRK}, n_{BRK}) * H^{s-}(S, S_{SRK}, \lambda_{SRK}, n_{SRK}) - k_{RK} * RK,$$

$$\dot{S} = g_S * H^{s-}(S, S_{SS}, \lambda_{SS}, n_{SS}) - k_S * S * H^{s+}(m3, S_{m3S}, 1/\lambda_{m3S}, n_{m3S}) * H^{s+}(Le, S_{LeS}, 1/\lambda_{LeS}, n_{LeS}),$$

$$\dot{m3} = g_{m3} * H^{s-}(S, S_{Sm3}, \lambda_{Sm3}, n_{Sm3}) * H^{s-}(Z, S_{Zm3}, \lambda_{Zm3}, n_{Zm3}) * H^{s+}(P, S_{Pm3}, \lambda_{Pm3}, n_{Pm3}) - k_{m3} * m3,$$

where H, Z, O, M, $m1$, $m2$, P, Li, Le, B, RK, S, $m3$ represent the expression levels of HIF-1, ZEB, OCT4, MDM2, miR-145, miR-200, P53, Lin28, Let7, Bach1, RKIP, SNAIL and miR-34, respectively.

Section S3: Truncated Moment Equations (TME)

The time evolution of the gene expression level can be studied as a complex dynamical system. The gene expression level $\mathbf{x}(t) = (x_1(t), x_2(t), \cdots, x_n(t))^T$ can be regarded as a stochastic process, and $\mathbf{f}(\mathbf{x}(t))$ represents the driving force of the system. Then the Langevin equations describing the dynamics of gene expression levels take the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{\Gamma}(t), \tag{S5}$$

where $\mathbf{g}(\mathbf{x}(t))\mathbf{\Gamma}(t)$ is the noise term, and $\mathbf{\Gamma}(t) = (\Gamma_1(t), \Gamma_2(t), \cdots, \Gamma_n(t))^T$ is n-dimensional independent Gaussian white noise (the generalized mean-square derivative of Brownian motion), which means:

$$\mathbb{E}[\Gamma_i(t)] = 0,$$

$$\mathbb{E}[\Gamma_i(t)\Gamma_j(t')] = 2D\delta_{ij}\delta_0(t-t'),$$

where D is a constant diffusion coefficient and

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Furthermore, $\mathbf{g}(\mathbf{x})$ is a matrix function defined on R^n:

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_{11}(\mathbf{x}) & g_{12}(\mathbf{x}) & \cdots & g_{1n}(\mathbf{x}) \\ g_{21}(\mathbf{x}) & g_{22}(\mathbf{x}) & \cdots & g_{2n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1}(\mathbf{x}) & g_{n2}(\mathbf{x}) & \cdots & g_{nn}(\mathbf{x}) \end{pmatrix}.$$

It indicates that the noise in the system depends on the state of the system. The corresponding Fokker-Planck equation (FPE) describing the time evolution of the density function of this system has the form:

$$\frac{\partial \rho(\mathbf{x},t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}[W_i(\mathbf{x},t)\rho(\mathbf{x},t)] + D\sum_i\sum_j \frac{\partial^2}{\partial x_i \partial x_j}[Q_{ij}(\mathbf{x},t)\rho(\mathbf{x},t)],$$

where

$$W_i(\mathbf{x},t) = f_i(\mathbf{x}) + D\sum_k\sum_j g_{kj}(\mathbf{x})\frac{\partial}{\partial x_k} g_{ij}(\mathbf{x}),$$

$$Q_{ij}(\mathbf{x},t) = \sum_k g_{ik}(\mathbf{x}) g_{jk}(\mathbf{x}).$$

When the drift part $W_i(\mathbf{x},t)$ is linear to $\mathbf{x}$ and the diffusion part $Q_{ij}(\mathbf{x},t)$ is constant, the FPE diffusion equation has an exact solution(Hu, 1994). Generally, when the drift part $W_i(\mathbf{x},t)$ is nonlinear, it's hard to solve the diffusion equation directly. Nonetheless, the solution of nonlinear system and linear system have similar forms when the diffusion coefficient $D << 1$. Therefore, when the diffusion coefficient $D << 1$, the solution can be approximated by a Gaussian distribution (Hu, 1994; Van Kampen, 1981):

$$\rho(\mathbf{x},t) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\sigma}(t)|^{1/2}}\exp\{-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}}(t))^T\boldsymbol{\sigma}^{-1}(t)(\mathbf{x}-\bar{\mathbf{x}}(t))\}, \tag{S6}$$

where, $\overline{\mathbf{x}}(t)$ and $\boldsymbol{\sigma}(t)$ are the first two moments of the Gaussian distribution, the mean and the covariance. They can be approximated to (Hu, 1994; Van Kampen, 1981):

$$\dot{\bar{\mathbf{x}}}(t) = \mathbf{f}(\bar{\mathbf{x}}(t)),$$

$$\dot{\boldsymbol{\sigma}}(t) = \boldsymbol{\sigma}(t)\mathbf{A}^T(t) + \mathbf{A}(t)\boldsymbol{\sigma}(t) + 2\mathbf{D}(\bar{\mathbf{x}}(t)). \tag{S7}$$

Here, $\boldsymbol{\sigma}(t)$ denotes the covariance matrix in time t, and $\mathbf{A}(t)$ is the jacobian matrix of $\mathbf{f}(\mathbf{x})$ when $\mathbf{x}$ is equal to $\bar{\mathbf{x}}(t)$, i.e., $A_{ij}(t) = \frac{\partial f_i(\mathbf{x})}{\partial x_j}|_{\mathbf{x}=\bar{\mathbf{x}}(t)}$. Furthermore, $\mathbf{D}(\bar{\mathbf{x}}(t))$ is a $n \times n$ matrix, which is equal to $D \cdot \mathbf{g}(\bar{\mathbf{x}}(t))\mathbf{g}(\bar{\mathbf{x}}(t))^T$.

In this work, we assume that the noise is homogeneous, and only consider the external noise. Therefore, $\mathbf{g}(\mathbf{x})$ is equal to $\mathbf{I}$, which is an identity matrix in n-dimensional space. Then the FPE reduces to:

$$\frac{\partial \rho(\mathbf{x},t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}[f_i(\mathbf{x},t)\rho(\mathbf{x},t)] + D\sum_i\sum_j \frac{\partial^2}{\partial x_i \partial x_j}[\rho(\mathbf{x},t)],$$

and then Eq. (S7) is modified to:

$$\dot{\boldsymbol{\sigma}}(t) = \boldsymbol{\sigma}(t)\mathbf{A}^T(t) + \mathbf{A}(t)\boldsymbol{\sigma}(t) + 2D \cdot \mathbf{I}.$$

In the previous method of partial self-consistent approximation (PSCA), $x_i(t)$, $i = 1, 2, \cdots, n$ are assumed to be weakly correlated, and only the diagonal elements in covariance matrix are calculated. In this study, the correlations between $x_i(t)$ and $x_j(t)$, $i, j = 1, 2, \cdots, n$ are taken into account to obtain a more accurate landscape for the system. To obtain the steady-state solution of FPE, we need to let $t \to +\infty$ in Eq. (S6). Then the mean $\bar{\mathbf{x}} = \lim_{t\to+\infty} \bar{\mathbf{x}}(t)$ and the covariance $\boldsymbol{\sigma} = \lim_{t\to+\infty} \boldsymbol{\sigma}(t)$ in steady states can be solved through numerical methods (see Section S4 for another way to calculate $\boldsymbol{\sigma}$). Then we can acquire the steady state density function $\rho(\mathbf{x})$ for the system.

The probability distribution acquired above corresponds to one stable steady state. If the system has multiple steady states, there should be several probability distributions localized at each basin with different covariance. To address this problem, we choose multiple random initial conditions to obtain different stable states, and use the frequency of each stable state to approximate the weight. The total probability is the weighted sum of all these probability distributions. For example, when the system has n stable states $x_i^s, i = 1, \cdots, n$, the frequency of x_i^s is b_i, and the corresponding density function is $\rho_i(\mathbf{x})$, we can obtain the stable density function of the system by $p(\mathbf{x}) = \sum b_i \rho_i(\mathbf{x})$.

If we only focus on the distribution of two variables such as x_1 and x_2, we can get the marginal distribution by integrating the other variables in $p(\mathbf{x})$. Then the steady state density function in x_1 and x_2 becomes:

$$p(x_1, x_2) = \int\int \cdots \int p(\mathbf{x}) dx_3 dx_4 \cdots dx_n.$$

Finally, with the steady state density function $p(\mathbf{x})$, we can get the steady state probability distribution $P_{ss}(\mathbf{x})$ after normalizing the density function, and further map out the potential landscape by $U(\mathbf{x}) = -\ln P_{ss}(\mathbf{x})$ (Li, 2017; Li and Wang, 2013, 2014). Here, $U(\mathbf{x})$ is dimensionless potential.

Section S4: Method of calculating σ

For the numerical calculation of ODEs, we can use Euler method by selecting the random initial conditions to obtain the stable states of $\bar{\mathbf{x}}$ and $\boldsymbol{\sigma}$. However, the number of the variable σ_{ij} is equal to the square of n (the dimension of $\mathbf{x}$). Therefore, instead of solving $\boldsymbol{\sigma}(t)$, we can solve $\boldsymbol{\sigma}$ in steady states analytically. This will dramatically reduce the computational cost. This method is feasible under the situation that the coefficient matrix in calculation of $\boldsymbol{\sigma}$ is nonsingular.

When the system reaches the steady state, $\dot{\bar{x}}(t)$ and $\dot{\boldsymbol{\sigma}}(t)$ must equal zero. It can be easily concluded that under the stable point $\mathbf{x}^0 = (x_1^0, x_2^0, \cdots, x_n^0)^T$, the following equations need to be satisfied:

$$\boldsymbol{\sigma}^0 \cdot (\mathbf{A}^0)^T + \mathbf{A}^0 \cdot \boldsymbol{\sigma}^0 + 2D \cdot \mathbf{I} = 0,$$

here, $\boldsymbol{\sigma}^0$ represents the matrix of $\boldsymbol{\sigma}$ at stable point $\mathbf{x}^0$, and $\mathbf{A}^0 = \mathbf{A}|_{\mathbf{x}^0}$. Note

$$\mathbf{A}^0 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix},$$

we can write these equations into the form of linear equations

$$\mathbf{L}\overline{\boldsymbol{\sigma}} = \mathbf{B},$$

where $\overline{\boldsymbol{\sigma}}$ and $\mathbf{B}$ are $\boldsymbol{\sigma}^0$ and $-\mathbf{2D}$ in row-major order form, which means:

$$\overline{\boldsymbol{\sigma}} = (\sigma_{11}^0, \sigma_{12}^0, \cdots, \sigma_{1n}^0, \sigma_{21}^0, \sigma_{22}^0, \cdots, \sigma_{2n}^0, \cdots, \sigma_{nn}^0)^T,$$

$$\mathbf{B} = -2D(1, 0, \cdots, 0, 0, 1, \cdots, 0, \cdots, 1)^T,$$

$\mathbf{L}$ can be seen as the sum of two block matrices,

$$L = \begin{pmatrix} \mathbf{A}^0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^0 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}^0 \end{pmatrix} + \begin{pmatrix} a_{11}\mathbf{I}_n & a_{12}\mathbf{I}_n & \cdots & a_{1n}\mathbf{I}_n \\ a_{21}\mathbf{I}_n & a_{22}\mathbf{I}_n & \cdots & a_{2n}\mathbf{I}_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{I}_n & a_{n2}\mathbf{I}_n & \cdots & a_{nn}\mathbf{I}_n \end{pmatrix},$$

where $\mathbf{I}_n$ is an identity matrix in n-dimensional space:

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

When the coefficient matrix $\mathbf{L}$ is nonsingular (in most cases this condition is satisfied), the only value of $\overline{\boldsymbol{\sigma}} = \mathbf{L}^{-1}\mathbf{B}$ can be obtained.

Section S5: Comparisons between TME and PSCA methods

Previously, a partial self-consistent mean field approximation (PSCA) method has been developed to calculate the probability distribution for multi-dimensional network systems (Li and Wang, 2013, 2014; Wang et al., 2010). With PSCA, we can split the probability into the products of the individual probabilities: $P(X_1, X_2, ..., X_N, t) \sim \prod_i^n P_i(X_i, t)$ and solve the probability self-consistently. In this way, we effectively reduce the dimensionality of the system from M^N to MN, and thus make the computation of the problem tractable. To verify the TME method we developed in this work is more accurate than PSCA, we perform simulations by Langevin dynamics method (see Section S6 for simulation method) and obtain the probability distribution calculated from the simulations. Then we calculate the distribution using PSCA and TME, separately. Therefore, we can obtain the corresponding deviations between the distribution calculated from PSCA (or TME) and that from simulations. This deviation measures the accuracy of corresponding methods for calculating probability distributions.

In the simulation of metabolic model, we choose $h = 0.1$, $D = d = 300$ and make the comparisons among PSCA, TME and Langevin simulations (Figure S1). It can be seen that the result of TME method is closer to the simulation result than the PSCA method.

To quantitatively compare the deviation of distributions, we applied relative distance and relative entropy(Cover and Thomas, 2012) to measure the deviation of the distributions. These two measures are respectively defined as:

1. Relative distance:
$$\sigma = \sqrt{\frac{\sum_{ij}(P_{ij}^{simulation} - P_{ij}^{a})^2}{(\sum_{ij} P^{simulation})^2}},$$
2. Relative entropy:
$$D_{KL} = \sum_{ij} P_{ij}^{simulation} log(\frac{P_{ij}^{simulation}}{P_{ij}^{a}}),$$

where P^a represents the probability distribution calculated from PSCA or TME method. Relative distance is a distance, which measures the distance between two distributions using Euclidean distance. Relative entropy is derived from information theory, which represents the information loss when we use an estimated distribution to fit the real distribution. When the deviation of two distribution increases, the relative entropy will also increase, and the relative entropy is zero when the two distributions are equal.

The comparison results (Figure S1) show that the relative distance for the results between TME method and simulations is 22.1% less than that between PSCA method and simulations, and the relative entropy for the result between TME method and simulations is 39.5% less than that between PSCA method and simulations. Both measures demonstrate that the distribution calculated from TME method fits the distribution from simulation better than the one of PSCA (Also see the comparisons for EMT-metabolism model in Figure S2).

Section S6: Simulation method based on Langevin dynamics

The Langevin equations (Eq. (S5)) can be written into the following form when $\mathbf{g}(\mathbf{x}) = \mathbf{I}$:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t)) + \mathbf{\Gamma}(t),$$

where $\mathbf{\Gamma}(t) = (\Gamma_1(t), \Gamma_2(t), \cdots, \Gamma_n(t))^T$ is Langevin force. It follows that $E[\Gamma_i(t)] = 0$ and $E[\Gamma_i(t)\Gamma_j(t')] = 2D \cdot \delta_{ij}\delta(t - t')$ $(i, j = 1, 2, \cdots, n)$ for arbitrary t and t'. Here, $\mathbf{x}(t)$ denotes the stochastic variables in n-dimensional space R^n at time t, and $\mathbf{f}(\mathbf{x}(t))$ represents the driving force of the system.

Under the general assumption that the stochastic process $\mathbf{\Gamma}(t)$ obeys a Gaussian probability distribution (Gaussian white noise), the Langevin force can be seen as a constant multiple distribution derivative of the n-dimensional standard Brownian motion in time t. Then the Langevin equations can be transformed into

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t))dt + \sqrt{2D}d\mathbf{B}(t),$$

and $\mathbf{B}(t) = (B_1(t), B_2(t), \cdots, B_n(t))^T$ denotes n-dimensional standard Brownian motion, which means $B_1(t), \cdots, B_n(t)$ are independent and satisfy the following several properties:

1. $B_i(t)$ is a stochastic process, which can be seen as a map from $(\Omega, \mathcal{F}, \mathbb{P}) \times ([0, +\infty), \mathcal{B})$ to $(R, \mathcal{B})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is probability space with a filtration $\{\mathcal{F}\}_{t\geq 0}$ and $\mathcal{B}$ is borel set.
2. $B_i(t)$ is independent increment process for $1 \leq i \leq n$. That is to say, for any time series $0 \leq t_1 < t_2 < \cdots < t_m$, random variables $B(t_1)$, $B(t_2) - B(t_1)$, $\cdots$, $B(t_m) - B(t_{m-1})$ are independent.
3. For any $t > s \geq 0$, random variable $B_i(t) - B_i(s)$ follows normal distribution $N(0, t - s)$ for $1 \leq i \leq n$.
4. The sample path of $B_i(t)$ is continuous almost surely in Ω.

The boundedness of fold change for self-inhibition λ and the value of Hill coefficient for inhibition or activation result in the Lipschitz continuity and the uniform boundedness of the function $\mathbf{f}(\mathbf{x}(t))$. Then the solution $\mathbf{x}(t)$ of these stochastic differential equations exists and is unique in the $\mathcal{M}^2([0, T], R^n) = \{x(\cdot)|x(t) \in \mathcal{F}_t, \forall t \in (0, T) \text{ and } \int_0^T |x(t)|^2 < +\infty\}$ for any finite T. However, since the complexity of $\mathbf{f}(\mathbf{x}(t))$ leads to the difficulty to show the analytical solution of these stochastic differential equations in a specific space, we apply Euler-Maruyama method to solve it:

$$\mathbf{x}(t + h) = \mathbf{x}(t) + \mathbf{f}(\mathbf{x}(t))h + \sqrt{2D}(\mathbf{B}(t + h) - \mathbf{B}(t)).$$

From the property of Brownian motion, the equations can be simplify as follows:

$$\mathbf{x}(t + h) = \mathbf{x}(t) + \mathbf{f}(\mathbf{x}(t))h + \sqrt{2D}\boldsymbol{\xi},$$

where h is time step and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)^T$, which satisfies that $\xi_1, \cdots, \xi_n$ are independent and follow $N(0, h)$ for $1 \leq i \leq n$. In this way, after a long enough time T, $\mathbf{x}(t)$ will reach a steady state. Therefore, We can approximate the probability of $\mathbf{x}(\mathbf{t})$ with the frequency of $\mathbf{x}(\mathbf{t})$ in R_+^n.

Section S7: Extended method to get potential landscape

In the main text, we choose two key variables as coordinates to obtain the landscape because it is hard to visualize the landscape in the high-dimensional space. An alternative way is to calculate the joint distribution of R_1 and R_2, where R_1 and R_2 are the square of the distance between the independent variable X and two chosen stable points. We first choose two stable fixed points of the system, and define them as μ_1 and μ_2, respectively. So we have

$$R_1 = (X - \mu_1)^T(X - \mu_1),$$

$$R_2 = (X - \mu_2)^T(X - \mu_2),$$

where, X is the random variable following the distribution calculated by TME method, which is the weighted sum of two Gaussian distributions. So we can firstly deal with the situation that X follows the Gaussian distribution. Therefore, we have the following claim.

Claim: Assume that the random vector $X = (x_1, x_2, \cdots, x_n)$ follows multi-dimensional normal distribution with expectation vector μ and invertible covariance matrix Σ. Let $a, b \in R^n$ be two arbitrary n-dimensional vectors and $R_i = (X - i)^T(X - i), i = a, b$. When the dimension n tends to infinity, we have

$$\mathbb{E}(R_a, R_b)^T = \tilde{\mu},$$

$$Cov(R_a, R_b) = \tilde{\Sigma},$$

where,

$$\tilde{\mu} = \begin{pmatrix} \tilde{\mu}_a \\ \tilde{\mu}_b \end{pmatrix} = \begin{pmatrix} tr(\Lambda) + (\mu - a)^T(\mu - a) \\ tr(\Lambda) + (\mu - b)^T(\mu - b) \end{pmatrix},$$

$$\tilde{\Sigma} = \begin{pmatrix} \tilde{\sigma}_{aa} & \tilde{\sigma}_{ab} \\ \tilde{\sigma}_{ab} & \tilde{\sigma}_{bb} \end{pmatrix} = \begin{pmatrix} 2tr(\Lambda^T\Lambda) + 4(\mu - a)^T\Sigma(\mu - a) & 2tr(\Lambda^T\Lambda) + 4(\mu - a)^T\Sigma(\mu - b) \\ 2tr(\Lambda^T\Lambda) + 4(\mu - a)^T\Sigma(\mu - b) & 2tr(\Lambda^T\Lambda) + 4(\mu - b)^T\Sigma(\mu - b) \end{pmatrix},$$

and Λ is a diagonal matrix whose diagonal elements are the eigenvalues of Σ.

Proof: Define $Y_a = X - a$, $\tilde{a} = \mu - a$, and $Z = \Sigma^{-1/2}(X - \mu)$, which follows Gaussian distribution $N(0, I_n)$. Then we obtain that $Z = \Sigma^{-1/2}(Y_a - \tilde{a})$. Hence, $Y_a = \Sigma^{1/2}(Z + \Sigma^{-1/2}\tilde{a})$.

Now,

$$R_a = Y_a^T Y_a = (Z + \Sigma^{-1/2}\tilde{a})^T\Sigma(Z + \Sigma^{-1/2}\tilde{a}).$$

Since the Σ is symmetric matrix, we can write that $\Sigma = H^T\Lambda H$ where H is an orthogonal matrix ($H^T H = I_n$) and Λ is a diagonal matrix with diagonal elements $\lambda_1, \lambda_2, \cdots, \lambda_n$. Write $U = HZ$ so that $U = (u_1, u_2, \cdots, u_n)$ also follows Gaussian distribution $N(0, I_n)$. Then we have

$$\begin{aligned} R_a &= (Z + \Sigma^{-1/2}\tilde{a})^T H^T\Lambda H(Z + \Sigma^{-1/2}\tilde{a}) \\ &= (HZ + H\Sigma^{-1/2}\tilde{a})^T\Lambda(HZ + H\Sigma^{-1/2}\tilde{a}) \\ &= (U + \alpha)^T\Lambda(U + \alpha) \\ &= \sum_{i=1}^{n} \lambda_i(u_i + \alpha_i)^2, \end{aligned}$$

where $\alpha = H\Sigma^{-1/2}\tilde{a}$. Therefore, R_a can be represented by the sum of independent random variables based on the independence of $u_1, u_2, \cdots, u_n$. Then we can calculate the expectation vector $\tilde{\mu}$ and covariance matrix $\tilde{\Sigma}$.

$$\begin{aligned} \tilde{\mu} = \mathbb{E}R_a &= \mathbb{E}(\sum_{i=1}^{n} \lambda_i(u_i + \alpha_i)^2) = \sum_{i=1}^{n} \lambda_i\mathbb{E}(u_i + \alpha_i)^2 \\ &= \sum_{i=1}^{n} \lambda_i\mathbb{E}(u_i^2 + 2\alpha_i u_i + \alpha_i^2) = \sum_{i=1}^{n} \lambda_i(1 + \alpha_i^2) \\ &= tr(\Lambda) + (\mu - a)^T(\mu - a), \end{aligned}$$

Similarly, we have $ER_b = tr(\Lambda) + (\mu - b)^T(\mu - b)$. Since u_i, u_j are independent for $i \neq j$, we also obtain

$$\begin{aligned} \tilde{\Sigma} = Cov(R_a, R_b) &= \mathbb{E}(R_a R_b) - \mathbb{E}R_a\mathbb{E}R_b \\ &= \mathbb{E}(\sum_{i=1}^{n} \lambda_i(u_i + \alpha_i)^2 \sum_{i=1}^{n} \lambda_i(u_i + \beta_i)^2) - \mathbb{E}(\sum_{i=1}^{n} \lambda_i(u_i + \alpha_i)^2)\mathbb{E}(\sum_{i=1}^{n} \lambda_i(u_i + \beta_i)^2) \\ &= \sum_{i=1}^{n} \lambda_i^2(2 + 4\alpha_i\beta_i) = 2tr[\Lambda^T\Lambda] + 4(\mu - a)^T\Sigma(\mu - b). \end{aligned}$$

Then the claim is proved.

Based on the above claim and central limit theorem, as the dimension n tends to infinity, both limit distribution of R_a and R_b are normal distribution under some general conditions. Therefore, the joint distribution of R_a and R_b also follows multi-dimension normal distribution $N(\tilde{\mu}, \tilde{\Sigma})$. In the case of multiple stable states, we can estimate the distribution of R_a and R_b for different stable point μ and the weight average of all distribution can be seen as the approximate joint distribution of R_a and R_b.

Figure S3 shows the landscape using R_1 (the square of distance between the independent variable X and Met state) and R_2 (the square of distance between the independent variable X and E state) as the coordinates when the network has four stable states (Epithelial state, Abnormal metabolic state, Mesenchymal state and Metastasis state).

Section S8: Minimum Action Paths

In cells, there are external noise from highly dynamical and inhomogeneous environments, and intrinsic noise from statistical fluctuations of the finite number of molecules, which can be critical to the dynamics of the system(Swain et al., 2002; Kaern et al., 2005; Thattai and Van Oudenaarden, 2001). A dynamical system under fluctuations can be addressed by: $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \zeta$, where $\mathbf{x} = (x_1(t), x_2(t), ..., x_{10}(t))$ represents the vector of gene expression levels, and $\mathbf{F}(\mathbf{x})$ is the vector for the driving force of chemical reactions. ζ is Gaussian noise term whose autocorrelation function is $< \zeta_i(\mathbf{x}, t)\zeta_j(\mathbf{x}, 0) >= 2D\delta(t)$, and D is diffusion coefficient matrix, quantifying the magnitude of fluctuations.

Assuming that the initial time is 0 and the terminal time is T, we define the path between ith attractor $\mathbf{x}^i$ and jth attractor $\mathbf{x}^j$ as $\mathbf{x}^{ij}(t) = (x_1^{ij}(t), x_2^{ij}(t), \cdots, x_n^{ij}(t))^T$ for $t \in [0, T]$, where the path $\mathbf{x}^{ij}(t)$ is satisfied with:

$$\begin{cases} (x_1^{ij}(0), x_2^{ij}(0), \cdots, x_n^{ij}(0))^T = \mathbf{x}^i, \\ (x_1^{ij}(T), x_2^{ij}(T), \cdots, x_n^{ij}(T))^T = \mathbf{x}^j. \end{cases}$$

We also define transition action $\mathcal{S}^{ij}$ as integral of the Lagrangian L^{ij} between time 0 and T, where L^{ij} is the distance between the driving force $\mathbf{F}$ and the velocity along the path. Then we can obtain the minimum action paths (MAPs) between ith point and jth point by solving this optimal problem. Following the approaches (E et al., 2004; Zhou et al., 2008; Heymann and Vanden-Eijnden, 2008) based on the Wentzell-Freidlin theory (Freidlin and Weber, 2004), the most probable transition path from ith attractor at time 0 to jth attractor at time T, can be acquired through minimizing the transition action functional over all possible paths:

$$\min_{\mathbf{x}^{ij}(\cdot)} \mathcal{S}^{ij}(\mathbf{x}^{ij}(\cdot)) = \min_{\mathbf{x}^{ij}(\cdot)} \int_0^T L^{ij}(t, \mathbf{x}^{ij}(t), \frac{d\mathbf{x}^{ij}(t)}{dt})dt,$$

where

$$L^{ij}(t, \mathbf{x}^{ij}(t), \frac{d\mathbf{x}^{ij}(t)}{dt}) = \left|\left|(\frac{dx_1^{ij}(t)}{dt}, \frac{dx_2^{ij}(t)}{dt}, \cdots, \frac{dx_n^{ij}(t)}{dt})^T - \mathbf{F}(\mathbf{x}^{ij}(t))\right|\right|_2.$$

We calculated MAPs numerically by applying minimum action methods used in (Zhou et al., 2008; Li, 2017), and treated the MAPs as the biological paths in our models. In this work, T is set to be 10 and we verified that larger values of T would not improve accuracy of simulations significantly.

Section S9: Landscape and kinetic path for cancer metabolism

Yu's work provides a basis for understanding the cancer metabolism from a deterministic ODE model. However, the stochastic dynamics and global properties of cancer metabolism remain to be elucidated, as the fluctuations have been suggested to play critical roles in biological systems, including intrinsic fluctuations from the finite number of molecules, and external fluctuations from highly inhomogeneous environments (Swain et al., 2002; Kaern et al., 2005; Thattai and Van Oudenaarden, 2001). Therefore, it is important to investigate the effects of fluctuations on the dynamics of the cancer metabolism system. Previously, we have developed a partial self-consistent approximation method to study the stochastic dynamics for high-dimensional systems by the energy landscape theory based on

$U = -\ln P_{ss}$ (Wang et al., 2008, 2010, 2011; Li, 2017; Li and Wang, 2014). One limitation of PSCA method is that it assumes the weak correlation between variables, which in some cases may not be accurate. Here, we improved previous methods and developed a Truncated Moment Equations (TME) method to calculate the probability distribution of gene expression level and obtain the potential landscape. We use the TME method to quantify the potential landscape for the metabolic model. For the purpose of visualization, we chose AMPK and HIF-1 as the two coordinates to display the landscape, since AMPK and HIF-1 are the two primary regulators of OXPHOS and glycolysis, respectively (Yu et al., 2017). By comparing the two methods, we showed that the TME method provides a more precise description for calculating the probability distribution of the systems (see Section S5).

Figure S6 shows the tristable potential landscape for cancer cell (Figure S6A and C) and bistable potential landscape for normal cell (Figure S6B and D). On the landscape, the blue region denotes the low potential and high probability, and the yellow region denotes the high potential and low probability. The three stable states (attractors) on the landscape respectively characterize three different states of cellular metabolism based on the relative expression level of AMPK and HIF-1, i.e. oxidative phosphorylation (O) state, hybrid (H) state and glycolysis (G) state (Figure S6A and C). Of note, the tristable landscape is consistent with previous metabolic model (Yu et al., 2017), where three stable states for O, G and H are obtained based on a deterministic model. Beyond the tristability results, here we quantified the stochastic properties and global stability from landscape perspective, i.e. the relative stabilities for different basins (states) are quantified by the potential barrier heights from the landscape of cancer metabolism. Furthermore, due to reprogrammed mitochondria, cancer cells have higher mtROS production(Porporato et al., 2014); and due to abnormally rapid proliferation, there are less available oxygen for each cell, so cancer cells have more stable HIF-1. Thus, cancer cells produce more mtROS (represented by γ, mtROS maximum fold change by the AMPK activation) and have a lower HIF-1 degradation rate (expressed as K_h) compared to normal cells(Yu et al., 2017). We investigate the influences of γ and k_H on cancer metabolism, we obtained the landscapes at different γ and k_H (Figure S7). As can be seen, for fixed k_H (the first row or the second row of Figure S7), the landscape topography changes and the system changes gradually from a tristable landscape to a bistable landscape as γ decreases. This indicates that the increase of γ promotes the transition from normal cells to cancer cells.

To study how cell transforms between different metabolic states, we calculated the kinetic transition paths among different attractors by minimizing the transition actions (see Methods in main text for how to obtain the transition path), which are also called minimum action paths (MAPs). The MAPs for different transitions are shown on the landscape (Figure S6). The magenta paths represent the MAPs from the O state to the H state and then to the G state and the white paths represent the MAPs from the G state to the H state and then to the O state, with the arrows denoting the directions of the transitions. The MAP for the transition process from O to G and the MAP from G to O are irreversible, reflected by the disparity between the forward and backward kinetic transition paths. This irreversibility of MAPs is a consequence of non-gradient force, i.e. curl flux (Wang et al., 2008, 2010; Li and Wang, 2014). Additionally, the dashed lines on the landscape represent the direct transition paths by minimizing the transition actions from O to G states. We found that the direct transition path from O to G and the backward transition path from G to O are inclined to go through the Hybrid (H) state. This indicates the critical roles of the Hybrid state in the process for cancer cells to change their metabolism. This also suggests that when a cancer cell change its metabolism from OXPHOS to glycolysis, it needs to go through a Hybrid state. The hybrid state can play a role of promoting plasticity and speeding the cancer metabolism.

Section S10: Global sensitivity analysis identifies the key players for cellular metabolism

To evaluate the degree of the influence of different factors on the transition actions between O state and G state, we performed the global sensitivity analysis for metabolism system. To reduce the complexity, we constrained the model parameters within the bistable state region ($\gamma = 3$ and $k_H = 0.3$). We changed each parameter individually and kept the other parameters unchanged. Here each parameter was increased 10% (Figure S8A) or decreased 10% (Figure S8B). Then sensitivity analysis identifies some critical parameters in the metabolic model including $g_{R_{mt}}$, $S_{AR_{mt}}$, g_H, S_{RA}, $k_{R_{mt}}$, λ_{HH}, k_A and g_A.

We can classify these parameters into three groups based on their different roles affecting the expression level of mtROS, AMPK or HIF-1, respectively. From sensitivity analysis (Figure S8), we see some consistent changes in

transition action. For example, enhancing $g_{R_{mt}}$, reducing $k_{R_{mt}}$ or $S_{AR_{mt}}$ (all these changes increase the expression level of mtROS (R_{mt})) will increase the transition action from O state to G state (S^{OG}) and decrease the transition action from G state to O state (S^{GO}). This indicates that the cell tends to stay in O state rather than G state, i.e. O state becomes more stable. These results are consistent with the experimental observations that the generation of mtROS mainly takes place at the ETC (mitochondrial electron transport chain) located on the inner mitochondrial membrane during the process of OXPHOS (Li et al., 2013). Similarly, increasing g_A, or decreasing k_A or S_{RA} (these changes will result in an enhancement of the expression level of AMPK) will lead to the increase of transition action from O state to G state and the decrease of transition action from G state to O state decreased, i.e., lead O state to be more stable. These results agree with the fact that the loss of AMPK promotes a glycolytic signature(Faubert et al., 2013). As for the parameters which affect HIF-1 (λ_{HH} and g_H), their influences on transition action are also consistent with the experimental observations, which indicates that HIF-1 is both necessary and sufficient for reducing mitochondrial oxygen consumption in hypoxia (Papandreou et al., 2006; Kim et al., 2006) (see Table S6 for details on the experimental evidences).

Section S11: Method of comparing modelling results with experimental data

To validate our models, we compare our modelling results to the experimental data. Since our model can be considered as a representative metabolism-EMT-metastasis model (not only for some specific types of cancer), we collect three different types of gene expression data from experiments. The first dataset is the single cell RNAseq data for a genetic mouse model of skin squamous cell carcinoma (SCC) undergoing EMT (383 single cells, available from the NCBI Gene Expression Omnibus under accession number GSE110357) (Pastushenko et al., 2018). We also collect another dataset of gene expression data of acute myeloid leukemia (AML) from the TCGA (173 samples for clinically annotated adult cases of de novo AML) (Ley et al., 2013). The third dataset we collect is for papillary thyroid carcinoma (PTC) from TCGA including 496 PTC samples (Agrawal et al., 2014). All of these three types of data are time independent, and therefore should correspond to the steady state data.

First, we extracted the gene expression data for the relevant genes that appear in our network (Figure 1) from these datasets, and neglected the mirRNAs and metabolites that are not available from the data. Then we perform the normalization by taking logarithmic transformation for the data. To match the landscape basins approximately, a linear transformation is performed to the normalized expression level of HIF1 and ZEB. We believe that this normalization of the data is reasonable because some corresponding gene expression values from these datasets differ from each other by folds and they represent different types of cancer in different conditions. In this way, we map the normalized data on the landscape in the coordinates of ZEB and HIF-1 (Figure 6C and D).

To see how the expression data for other genes match the modelling results, we also adopted another way to make comparisons. For each dataset (SCC, AML and PTC), we perform PCA to the normalized data and project the normalized data onto the first two PCA components (Figure 6E-M). The colors in the PCA plots represent the expression levels of HIF1, ZEB and SNAI1, individually. Then we cluster these data points into three groups by visualization based on the expression level of these three marker genes, which are denoted by purple, green and orange ovals, respectively. The points in the purple ovals have low HIF1 and Low ZEB expression level, the points in the green ovals have high HIF1 and low ZEB expression level, and the points in the orange ovals have high HIF1 and ZEB expression level. These three clusters from PCA plots are qualitatively similar to the three attractor states (E, A and M) identified from our models. Therefore, these three types of cancer data support the existence of abnormal metabolic state (A) identified from our landscape model.

Section S12: Modeling anti-cancer therapeutic strategies

Metformin-based therapy

Metformin not only activates AMPK but also inhibits HIF-1 in an AMPK-independent way (Kalender et al., 2010). Thus, we used shifted Hill functions to model the influences of metformin on HIF-1 and AMPK (Yu et al., 2017). The modified equations for HIF-1 and AMPK under the influence of metformin are shown as follows:

$$\dot{H} = g_H * H^{s-}(A, S_{AH}, \lambda_{AH}, n_{AH}) * H^{s-}(Metformin, S_{MH}, \lambda_{MH}, n_{MH}) * H^{s-}(P, S_{PH}, \lambda_{PH}, n_{PH}) - k_H * H$$
$$* H^{s-}(H, S_{HH}, 1/\lambda_{HH}, n_{HH}) * H^{s-}(R, S_{RH}, 1/\lambda_{RH}, n_{RH}),$$

$$\dot{A} = g_A * H^{s+}(R, S_{RA}, \lambda_{RA}, n_{RA}) * H^{s+}(Metformin, S_{MA}, \lambda_{MA}, n_{MA}) * H^{s-}(H, S_{HA}, \lambda_{HA}, n_{HA}) \\ *H^{s-}(A, S_{AA}, \lambda_{AA}, n_{AA}) - k_A * A.$$

Additionally, metformin can also induce high mtROS production through altering the mitochondrial potential (Cheong et al., 2011). So the increase in the mtROS production rate is modeled to be proportional to the level of metformin:

$$(g_{R_{mt}})_{metformin} = g_{R_{mt}} + \alpha_{metformin} * Metformin$$

Here, $(g_{R_{mt}})_{metformin}$ is the generation rate of mtROS under the influence of metformin, and $\alpha_{metformin}$ is a weighting factor for the mtROS production.

3BP therapy

3BP inhibits glycolysis and decreases the self-activation strength of HIF-1 (Yu et al., 2017). Therefore, the influence of 3BP is modelled as:

$$1/(\lambda_{HH})_{3BP} = 1/\lambda_{HH} + \alpha_{3BP} * 3BP$$

Here, $(\lambda_{HH})_{3BP}$ is the maximum fold change due to HIF-1 self-activation under the influence of 3BP, and α_{3BP} is a constant denoting the extent of the glycolysis reduction.

BACH1-inhibitor therapy

Based on the landscape results with different drug effects (Figure 7B-D in main text), we suggest a new therapeutic strategy against cancer metastasis, which is the combination of 3BP (or metformin) and certain type of BACH1 inhibitor. The purpose of adding BACH1 inhibitor is to eliminate the Met2 state (metastasis-like state with high expression of BACH1). Here, the specific influence of BACH1 inhibitor (BI) decreases the production rate of BACH1. Then the modified g_B takes the form as:

$$(g_B)_{BI} = g_B - \alpha_{BI} * BI$$

Here, $(g_B)_{BI}$ is the production rate of BACH1 under the influence of BI, and α_{BI} is a weighting factor for the BACH1 production.

Supplemental References

Abida, W.M., Nikolaev, A., Zhao, W., Zhang, W., Gu, W., 2007. Fbxo11 promotes the neddylation of p53 and inhibits its transcriptional activity. J. Biol. Chem. 282, 1797–1804.

Agrawal, N., Akbani, R., Aksoy, B.A., Ally, A., Arachchi, H., Asa, S.L., Auman, J.T., Balasundaram, M., Balu, S., Baylin, S.B., et al., 2014. Integrated genomic characterization of papillary thyroid carcinoma. Cell 159, 676–690.

Brabletz, S., Brabletz, T., 2010. The zeb/mir-200 feedback loopła motor of cellular plasticity in development and cancer? EMBO Rep. 11, 670–677.

Cheong, J.H., Park, E.S., Liang, J., Dennison, J.B., Tsavachidou, D., Nguyen-Charles, C., Cheng, K.W., Hall, H., Zhang, D., Lu, Y., et al., 2011. Dual inhibition of tumor energy pathway by 2-deoxyglucose and metformin is effective against a broad spectrum of preclinical cancer models. Mol. Cancer Ther. 10, 2350–2362.

Cover, T.M., Thomas, J.A., 2012. Elements of information theory. John Wiley & Sons.

E, W., Ren, W., Vanden-Eijnden, E., 2004. Minimum action method for the study of rare events. Commun. Pur. Appl. Math 57, 637–656.

Faubert, B., Boily, G., Izreig, S., Griss, T., Samborska, B., Dong, Z., Dupuy, F., Chambers, C., Fuerth, B.J., Viollet, B., et al., 2013. Ampk is a negative regulator of the warburg effect and suppresses tumor growth in vivo. Cell Metab. 17, 113–124.

Freidlin, M., Weber, M., 2004. Random perturbations of dynamical systems and diffusion processes with conservation laws. Probab. Theory Rel. 128, 441–466.

Gérard, C., Goldbeter, A., 2009. Temporal self-organization of the cyclin/cdk network driving the mammalian cell cycle. Proc. Natl. Acad. Sci. U S A 106, 21643–21648.

Heymann, M., Vanden-Eijnden, E., 2008. The geometric minimum action method: A least action principle on the space of curves. Commun. Pur. Appl. Math 61, 1052–1117.

Hu, G., 1994. Stochastic forces and nonlinear systems.

Huang, B., Lu, M., Jia, D., Benjacob, E., Levine, H., Onuchic, J.N., 2017. Interrogating the topological robustness of gene regulatory circuits by randomization. PLoS Comput. Biol. 13, 1005456.

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.

Kaern, M., Elston, T.C., Blake, W.J., Collins, J.J., 2005. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 6, 451–64.

Kalender, A., Selvaraj, A., Kim, S.Y., Gulati, P., Brule, S., Viollet, B., Kemp, B.E., Bardeesy, N., Dennis, P., Schlager, J.J., 2010. Metformin, independent of ampk, inhibits mtorc1 in a rag gtpase-dependent manner. Cell Metab. 11, 390–401.

Kim, J.w., Tchernyshyov, I., Semenza, G.L., Dang, C.V., 2006. Hif-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. Cell Metab. 3, 177–185.

Kim, T., Veronese, A., Pichiorri, F., Lee, T.J., Jeon, Y.J., Volinia, S., Pineau, P., Marchio, A., Palatini, J., Suh, S.S., et al., 2011. p53 regulates epithelial–mesenchymal transition through microrna targeting zeb1 and zeb2. J. Exp. Med. 208, 875–883.

Kong, D., Banerjee, S., Ahmad, A., Li, Y., Wang, Z., Sethi, S., Sarkar, F.H., 2010. Epithelial to mesenchymal transition is mechanistically linked with stem cell signatures in prostate cancer cells. PloS One 5, e12445.

Lee, J., Lee, J., Farquhar, K.S., Yun, J., Frankenberger, C.A., Bevilacqua, E., Yeung, K., Kim, E.J., Balázsi, G., Rosner, M.R., 2014. Network of mutually repressive metastasis regulators can promote cell heterogeneity and metastatic transitions. Proc. Natl. Acad. Sci. U S A 111, E364–E373.

Ley, T.J., Miller, C.A., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, G., Hoadley, K.A., Triche, T.J., Laird, P.W., Baty, J., et al., 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. New Engl. J. Med. 368, 2059–2074.

Li, C., 2017. Identifying the optimal anticancer targets from the landscape of a cancer–immunity interaction network. Phy. Chem. Chem. Phys. 19, 7642–7651.

Li, C., 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.

Li, C., Wang, J., 2014. Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. Proc. Natl. Acad. Sci. U S A 111, 14130–14135.

Li, M., He, Y., Dubois, W., Wu, X., Shi, J., Huang, J., 2012. Distinct regulatory mechanisms and functions for p53-activated and p53-repressed dna damage response genes in embryonic stem cells. Mol. cell 46, 30–42.

Li, X., Fang, P., Mai, J., Choi, E.T., Wang, H., Yang, X.f., 2013. Targeting mitochondrial reactive oxygen species as novel therapy for inflammatory diseases and cancers. J. Hematol. Oncol. 6, 19.

Liang, R., Bates, D.J., Wang, E., 2009. Epigenetic control of microrna expression and aging. Curr. Genomics 10, 184–193.

Lu, M., Jolly, H., Levine, H., Onuchic, J., Ben-Jacob, E., 2013. Microrna-based regulation of epithelial-hybrid-mesenchymal fate determination. Proc. Natl. Acad. Sci. USA. 110, 18144–18149.

Lu, M., Jolly, M.K., Ben-Jacob, E., et al., 2014. Toward decoding the principles of cancer metastasis circuits. Cancer Res. 74, 4574–4587.

Papandreou, I., Cairns, R.A., Fontana, L., Lim, A.L., Denko, N.C., 2006. Hif-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. Cell Metab. 3, 187–197.

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S., et al., 2018. Identification of the tumour transition states occurring during emt. Nature 556, 463.

Porporato, P.E., Payen, V.L., Pérez-Escuredo, J., De Saedeleer, C.J., Danhier, P., Copetti, T., Dhup, S., Tardy, M., Vazeille, T., Bouzin, C., et al., 2014. A mitochondrial switch promotes tumor metastasis. Cell Rep. 8, 754–766.

Ren, D., Wang, M., Guo, W., Huang, S., Wang, Z., Zhao, X., Du, H., Song, L., Peng, X., 2014. Double-negative feedback loop between zeb2 and mir-145 regulates epithelial-mesenchymal transition and stem cell properties in prostate cancer cells. Cell Tissue Res. 358, 763–778.

Robbins, D., Zhao, Y., 2012. Oxidative stress induced by mnsod-p53 interaction: pro-or anti-tumorigenic? J. Signal Transduct. 2012, 101465–101465.

Seppet, E., Gruno, M., Peetsalu, A., Gizatullina, Z., Nguyen, H.P., Vielhaber, S., Wussling, M.H., Trumbeckaite, S., Arandarcikaite, O., Jerzembeck, D., et al., 2009. Mitochondria and energetic depression in cell pathophysiology. Int. J. Mol. Sci. 10, 2252–2303.

Sermeus, A., Michiels, C., 2011. Reciprocal influence of the p53 and the hypoxic pathways. Cell Death Dis. 2, e164.

Swain, P.S., Elowitz, M.B., Siggia, E.D., 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc. Natl. Acad. Sci. U S A 99, 12795–12800.

Thattai, M., Van Oudenaarden, A., 2001. Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. U S A 98, 8614–8619.

Van Kampen, N., 1981. Stochastic processes in chemistry and physics. volume 1.

Wang, G., Guo, X., Hong, W., Liu, Q., Wei, T., Lu, C., Gao, L., Ye, D., Zhou, Y., Chen, J., et al., 2013. Critical regulation of mir-200/zeb2 pathway in oct4/sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. Proc. Natl. Acad. Sci. U S A 110, 2858–2863.

Wang, J., Li, C., Wang, E., 2010. Potential and flux landscapes quantify the stability and robustness of budding yeast cell cycle network. Proc. Natl. Acad. Sci. U S A 107, 8195–8200.

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. U S A 105, 12271–12276.

Wang, J., Zhang, K., Xu, L., Wang, E., 2011. Quantifying the waddington landscape and biological paths for development and differentiation. Proc. Natl. Acad. Sci. U S A 108, 8257–8262.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., Kosik, K.S., 2009. Microrna-145 regulates oct4, sox2, and klf4 and represses pluripotency in human embryonic stem cells. Cell 137, 647–658.

Yu, L., Lu, M., Jia, D., Ma, J., Ben-Jacob, E., Levine, H., Kaipparettu, B.A., Onuchic, J.N., 2017. Modeling the genetic regulation of cancer metabolism: Interplay between glycolysis and oxidative phosphorylation. Cancer Res. 77, 1564–1574.

Zhang, J., Sun, Q., Zhang, Z., Ge, S., Han, Z., Chen, W., 2013. Loss of microrna-143/145 disturbs cellular growth and apoptosis of human epithelial cancers by impairing the mdm2-p53 feedback loop. Oncogene 32, 61.

Zhang, W., Shi, X., Peng, Y., Wu, M., Zhang, P., Xie, R., Wu, Y., Yan, Q., Liu, S., Wang, J., 2015. Hif-1α promotes epithelial-mesenchymal transition and metastasis through direct regulation of zeb1 in colorectal cancer. PLoS One 10, e0129603.

Zhou, C.H., Zhang, X.P., Liu, F., Wang, W., 2015. Modeling the interplay between the hif-1 and p53 pathways in hypoxia. Sci. Rep. 5, 13834.

Zhou, X., Ren, W., Weinan, E., et al., 2008. Adaptive minimum action method for the study of rare events. J. Chem. Phys. 128, 104111.

Zisoulis, D.G., Kai, Z.S., Chang, R.K., Pasquinelli, A.E., 2012. Autoregulation of microrna biogenesis by let-7 and argonaute. Nature 486, 541.