

REVIEW

From Phage lambda to human cancer: endogenous molecular-cellular network hypothesis

Gaowei Wang^{1,2}, Xiaomei Zhu³, Leroy Hood⁴ and Ping Ao^{1,2,5,*}

¹ Ministry of Education Key Laboratory of Systems Biomedicine, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China

² State Key Laboratory for Oncogenes and Related Genes, Shanghai Cancer Institute, Shanghai Jiao Tong University School of Medicine, Shanghai 200032, China

³ GenMath, Corp., Seattle, WA 98105, USA

⁴ Institute for Systems Biology, Seattle, WA 98103, USA

⁵ Department of Physics, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence: aoping@sjtu.edu.cn

Received November 11, 2012; Revised December 10, 2012; Accepted December 10, 2012

Experimental evidences and theoretical analyses have amply suggested that in cancer genesis and progression genetic information is very important but not the whole. Nevertheless, “cancer as a disease of the genome” is still currently the dominant doctrine. With such a background and based on the fundamental properties of biological systems, a new endogenous molecular-cellular network theory for cancer was recently proposed by us. Similar proposals were also made by others. The new theory attempts to incorporate both genetic and environmental effects into one single framework, with the possibility to give a quantitative and dynamical description. It is asserted that the complex regulatory machinery behind biological processes may be modeled by a nonlinear stochastic dynamical system similar to a noise perturbed Morse-Smale system. Both qualitative and quantitative descriptions may be obtained. The dynamical variables are specified by a set of endogenous molecular-cellular agents and the structure of the dynamical system by the interactions among those biological agents. Here we review this theory from a pedagogical angle which emphasizes the role of modularization, hierarchy and autonomous regulation. We discuss how the core set of assumptions is exemplified in detail in one of the simple, important and well studied model organisms, Phage lambda. With this concrete and quantitative example in hand, we show that the application of the hypothesized theory in human cancer, such as hepatocellular carcinoma (HCC), is plausible, and that it may provide a set of new insights on understanding cancer genesis and progression, and on strategies for cancer prevention, cure, and care.

CHALLENGES TO UNDERSTAND CANCER GENESIS AND PROGRESSION

Cancer, a popular generic term for malignant neoplasms, has been defined as an abnormal mass of tissue the growth of which exceeds and is uncoordinated with that of the normal tissues and persists in the same excessive manner even after the cessation of the stimuli which evoked the change, and the new growth has virulent or adverse properties in the body [1]. Its exact nature is still not well understood. Numerous theories and hypotheses on cancer

genesis and progression have been proposed during the long discourse on this disease and the attempts to cure it. In ancient times, physicians had blamed cancers on the gods [2]. From the Renaissance to the 19th century, scientific hypotheses about cancer had started to be formed, such as cancer as chronic irritation disease, trauma disease, and infectious disease and so on. These theories and hypothesizes are obviously subjected to their limited understanding of cancer at that time, and evidently not suitable hypotheses for today’s perspective. Nevertheless, those efforts are helpful to develop better

understanding of cancer by changing the paradigm in cancer research [3,4].

During the 20th century as the understanding of DNA and genes increased tremendously cancer has been hypothesized as a genetic disease, which is persistent to these days. This genetically centric hypothesis suggested that the genesis and progression of cancer is caused by genetic alterations, carcinogenic factors caused cancer by their damages to normal genome [5,6]. Since then, cancer research has focused on genetic and genomic aspects, such as gene sequencing [7], oncogenes [8], suppressor genes [9].

On the other hand, mountains of experimental evidence and theoretical analyses have suggested that genome is not the whole story on cancer genesis and progression. From experimental side, evidences show that some other factors such as microenvironment and inflammation cannot be ignored. One pronounced instance is the seed and soil hypothesis: In 1889, the English surgeon, Stephen Pagt, concluded his analyses of cancer histories borrowing a plant analogy. It states that when a plant goes to seeding, its seeds are carried in all directions; but they can only live and grow if they fall on congenial soil [10]. The seed in the modern usage has been reinterpreted as progenitor cell, cancer stem cell, or metastatic cell, and the soil as host factors, stoma, or the organ microenvironment [10,11]. The experimental evidence suggested that organ microenvironment cannot be ignored in cancer genesis and progression. Another more directly evidence is the study of the precursor of esophageal adenocarcinoma. By tracking its source of precursor, recent findings suggest that certain precancerous lesions, such as Barrett's, initiate not from genetic alterations but from competitive interactions among cell lineages driven by opportunity [12]. Based on this kind of phenomena as well as others one may readily concludes that besides genome other factors such as congenial soil (microenvironment) and inflammation also play key roles in cancer genesis and progression which cannot be ignored.

From both clinic and theoretical sides, evidences also suggest that the genetic and genomic information are important but not enough. Biological systems are characterized by the stochastic dynamical phenomena and concepts such as adaptation [13,14], robustness [15,16], phenotype switch [17,18]. In tumors such concepts correspond to well-documented drug resistance [19], tremendously difficult for cancer regression [20] and genesis and progression to tumor from normal tissue respectively. The phenomena and concepts arise from the complex regulatory machinery, the building blocks of the regulatory machinery including genetic switch, feedback loops [21], double-edge effect [22,23], et al. A desirable way to make the biological system's phenomenon and concepts clear is to understand and manipulate the

regulatory machinery quantitatively, it is also one of the biggest challenges for contemporary biology [24–28]. It is clear that we cannot achieve this goal just by using genetic information. Consensus starts to form that complete information of the DNA sequence of an organism will not enable us to reconstruct the regulatory machinery quantitatively because of the many gaps between genotype and phenotype. This is similar to the known situation that a complete list of English words cannot enable us to reconstruct a play by Shakespeare. Those examples clearly indicate that the genomic information is important but not enough. We need to reveal the regulatory machinery behind a biological phenomenon quantitatively, which would be of great importance for our understanding and manipulating biological phenomenon, such as cancer genesis and progression.

To meet this challenge, based on the current understanding of biological systems, the endogenous molecular-cellular network hypothesis for cancer genesis and progression has been proposed [3]. It is further shown that a set of stochastic differential equations for the endogenous network may be set up [29]. Such hypothesized theory allows us to grasp and manipulate the general regulatory machinery of cancer genesis and progression quantitatively and at the same time gives ways to incorporate both genetic and environmental factors into consideration. Based on this hypothesis a different set of strategies for cancer prevention, cure, and care, may be suggested.

In the following sections key aspects of the hypothesized theory will be reviewed and elaborated. In section *FORMING ENDOGENOUS MOLECULAR-CELLULAR NETWORK HYPOTHESIS* we will discuss the basic elements in the hypothesis and the essential requirements. In section *LEARNING FROM PHAGE LAMBDA GENETIC SWITCH* Phage lambda genetic switch, one of the most simple, well-studied, and important systems, is used to understand the hypothesis. We demonstrate that the endogenous molecular-cellular network of Phage lambda genetic switch quantitatively describes Phage lambda's two modes of life cycle. In section *MOVING UP IN COMPLEXITY: HEPATOCELLULAR CARCINOMA* we go further to take hepatocellular carcinoma (HCC) as an example. We use the endogenous molecular-cellular network hypothesis to quantify and understand the genesis and progression of cancer, which provides another perspective to find strategies for cancer prevention, cure, and care. The similarities and differences among similar proposals, along with a few dominant cancer theories, are discussed in section *RAMIFYING THE ENDOGENOUS NETWORK THEORY*. We conclude in section *PERSPECTIVE* that the endogenous network theory may provide the best candidate theory to understand cancer genesis and progression.

FORMING ENDOGENOUS MOLECULAR-CELLULAR NETWORK HYPOTHESIS

In this section we discuss what would be needed for the endogenous molecular-cellular network hypothesis on cancer genesis and progression. Three important aspects of this hypothesis will be examined in detail.

The hypothesis can be stated as the following. In order to maintain the normal physiological function and developmental process for tissue specific function shaped by evolution, a minimal set of fundamental functional modules or pathways, for example the cell cycle, cell death, inflammation, metabolism, cell adhesion, angiogenesis, are needed. Each module can accomplish a relatively autonomous function, and cross-talks between modules allow one function to influence another. At molecular-cellular level, it is hypothesized that the functional modules are deeply hierarchical and may be specified by important molecular and cellular agents, such as oncogenes, suppressor genes, and related growth factors, hormones, cytokines, etc. The interactions among these agents form an autonomous, nonlinear, stochastic, and collective dynamical network. We have tentatively named it as the endogenous molecular-cellular network. The endogenous network may generate many locally stable states with obvious or non-obvious biological function. Normal state and cancer state are assumed to be stable states of the endogenous molecular-cellular network. The endogenous network may stay in each stable state for a considerably long time, and in certain condition, stable states can switch between each other. In this hypothesis, the genesis and progression of cancer can be regarded as transition of the endogenous molecular-cellular network from normal stable state to cancer state [3].

With this hypothesis, biological systems will be greatly simplified, and make it feasible for us to grasp and manipulate the general regulatory machinery of cancer genesis and progression quantitatively. We now discuss the three basic tenets of this hypothesis, modularization, deep hierarchy and autonomous regulation, one by one.

1. *Modularization: Biological system is built by modules and cross-talk between modules, each module can accomplish a relatively autonomous function, and cross-talks between modules allow one function to influence another.*

A functional module is, by definition, a discrete entity whose function is separable from those of other modules [21]. The first evidence in support of this assumption is that functional modules are transplantable and conserved in different organisms. For example, the cell cycle module is conserved in different organisms, much of what we know about cell cycle can be traced back to basic studies in yeast cells [30,31], cell-specific functional modules can

be transplanted into different type of cells by several transcription factors [32,33]. These evidences directly indicated that many biological functions are relatively autonomous; the separated functional modules can accomplish its own function and of course the modules cross-talk with each other.

Another evidence is the theoretical models based on functional modules verifying their predictions match reality, such as cell cycle [15,34], apoptosis [35]. The quantitative model, by theoretical predications and experimental validation, is the best way to test our understanding of biological systems. The match of predications and reality suggest that we can adopt the concepts of autonomous functional modules.

Therefore, we argue here that biological system is built by modules and cross-talk between modules. Each module can accomplish a relatively autonomous function, and cross-talks between modules allow one function to influence another.

2. *Deep hierarchy: There are regulators that serve as decision-marking in modules to determine module's fate, then the regulators triggering the expression of a whole battery of downstream genes related to the decided fate. Modules and cross-talk between modules can be simplified and specified by interactions of important proteins, and the general principles will allow us to grasp and manipulate the modules.*

Multiple experimental evidences suggest the existence of key regulators. For example two regulatory protein CI and Cro regulate Phage lambda genetic switch [36], key regulators regulate developmental process of sea urchin [37], and four transcriptional factors induce human fibroblasts to pluripotent stem (iPS) cells [32].

Module components include genes (DNA), mRNA, microRNA, proteins, small molecules, etc. Among all the components proteins play special roles in biological system, because they are the main biological function executor, they can regulate gene expression, and they are key regulators in signaling transduction. It is clear that functions cannot be understood by studying the single proteins [27,38]. Important functions arise from the regulatory machinery. General modules and cross-talk between modules can be specified by proteins and their interactions in modeling [39].

3. *Autonomous network and intrinsic stable states: The regulatory proteins and their interactions form a closed and decision-making network which is responsible for the biological stable states. We named the closed network as endogenous molecular-cellular network. Endogenous molecular-cellular network is shaped by evolution, and normal tissue and tumor can be regarded as intrinsic stable states.*

There is a paradox in biological systems [39]: If

mRNAs are required to synthesize proteins and proteins are required, in turn, to regulate the expression of mRNAs, then what's the cause-and-effect relationship of the interdependent components? The paradox can be interpreted as the interdependent components of biological systems forming a closed network. Mathematically, the closed network forms a nonlinear autonomous dynamical system implying many locally stable states [40,41].

We further assume that the backbone and essential structure of the endogenous molecular-cellular network, in other words the regulatory machinery, is conserved, because it has been shaped by millions, or even billions, years of evolution. During the lifetime of an organism, there is a little chance of any major modification of the essential structure of the endogenous network for a viable cell [42]. And multiple evidences suggest that cancer is similar to many normal physiological processes such as wound healing, developmental process, and cancer is often related to un-healing wound [43], inflammation and aberrant development [44]. Based on the evidence we reason that cancer may be an intrinsic stable state in organisms shaped by evolution, cancer state is one of the stable states of endogenous molecular-cellular network. We proposed the endogenous molecular-cellular network hypothesis, in fact, the hypothesis suits all organisms. We will discuss how this hypothesis is exemplified in Phage lambda and how this hypothesis is used in human cancer in the following parts.

LEARNING FROM PHAGE LAMBDA GENETIC SWITCH

Before applying the endogenous molecular-cellular hypothesis to human cancer we analyze it against one of the most simple and important biological systems: Phage lambda. The hierarchy, modularity, and autonomous regulation are revealed clearly and quantitatively.

Life cycle of Phage lambda

Phage lambda played important roles in our understanding of molecular biology. As one of the simplest living organisms Phage lambda was discovered by Esther Lederberg in 1950 [45]. Since then Phage lambda has been used heavily as a model organism, and has been a rich source in microbial genetics, and later in molecular genetics in the mid-20th century. Phage lambda have made great contribution to molecular biology, many terms in molecular biology we use today are defined in the course of studying phage, such as gene regulatory protein, genetic switch [39]. Because of Phage lambda's great contribution, it has been arguably regarded as the origin of molecular biology [46]. Phage lambda is still one of the best model organisms to help us understand biological

concepts, precision biological quantification [17]. More important, because of evolution this type of simple organism can be viewed as a prototype for similar, but more complex higher eukaryote.

We focus on Phage lambda's genetic switch module here. Phage lambda as a functional organism is a virus particle including a head, a tail, and tail fibers. The head contains the Phage's double-stranded circular DNA genome. The phage particle recognizes and binds to its host, *E. coli*, causing DNA in the head of the phage to be injected into the cytoplasm of the bacterial cell. There are two possible fates of the *E. coli* after the injection of phage DNA (Figure 1A). One is lysis. The lambda DNA is replicated many times and the genes for head, tail and lytic proteins are expressed. This leads to assembly of multiple new phage particles releasing the *E. coli* contents into the environment. The other is lysogeny. The phage DNA may integrate itself into the host cell chromosome and stays resident within the host's genome without apparent harm to the host. In certain conditions, such as changed temperature or DNA, an integrated virus can switch from lysogeny to lysis. In the absence of such interference; however, the phage can maintain their state robustness [47].

As almost all parts of Phage lambda have been known during past 60 years (Figure 1B), we may systematically describe the precise mechanisms of Phage lambda's two stable states, lysis and lysogeny, and the switch from lysogeny to lysis at the molecular level. As one of the simplest living organisms, Phage genome is about 49k base pairs, encoding about 40 genes. The regulatory mechanisms and functions of these genes are well documented, and the knowledge lays a solid foundation for us to construct a detailed endogenous molecular-cellular network of Phage lambda [47].

Endogenous molecular-cellular network of Phage lambda

When we focus on the maintenance of lytic and lysogeny states and switch between the states, we have found that many factors (Figure 1B), such as nutrition, temperature, DNA damage, et al., can affect genetic switch by affecting regulatory mechanisms. The building block of regulatory mechanisms in Phage lambda is gene regulation. In gene transcription, RNA polymerase slides into a region on the DNA double helix called a promoter, the RNA polymerase moves stepwise along the DNA, indicating the starting point for RNA synthesis. Gene regulatory proteins can increase or decrease gene transcription dramatically by binding to a nearby site on the DNA and contact the RNA polymerase initiating transcription.

We take right operator O_R as an example, which is composed of three regions designated O_{R1} , O_{R2} and O_{R3} . The promoter P_R completely overlaps O_{R1} and partially

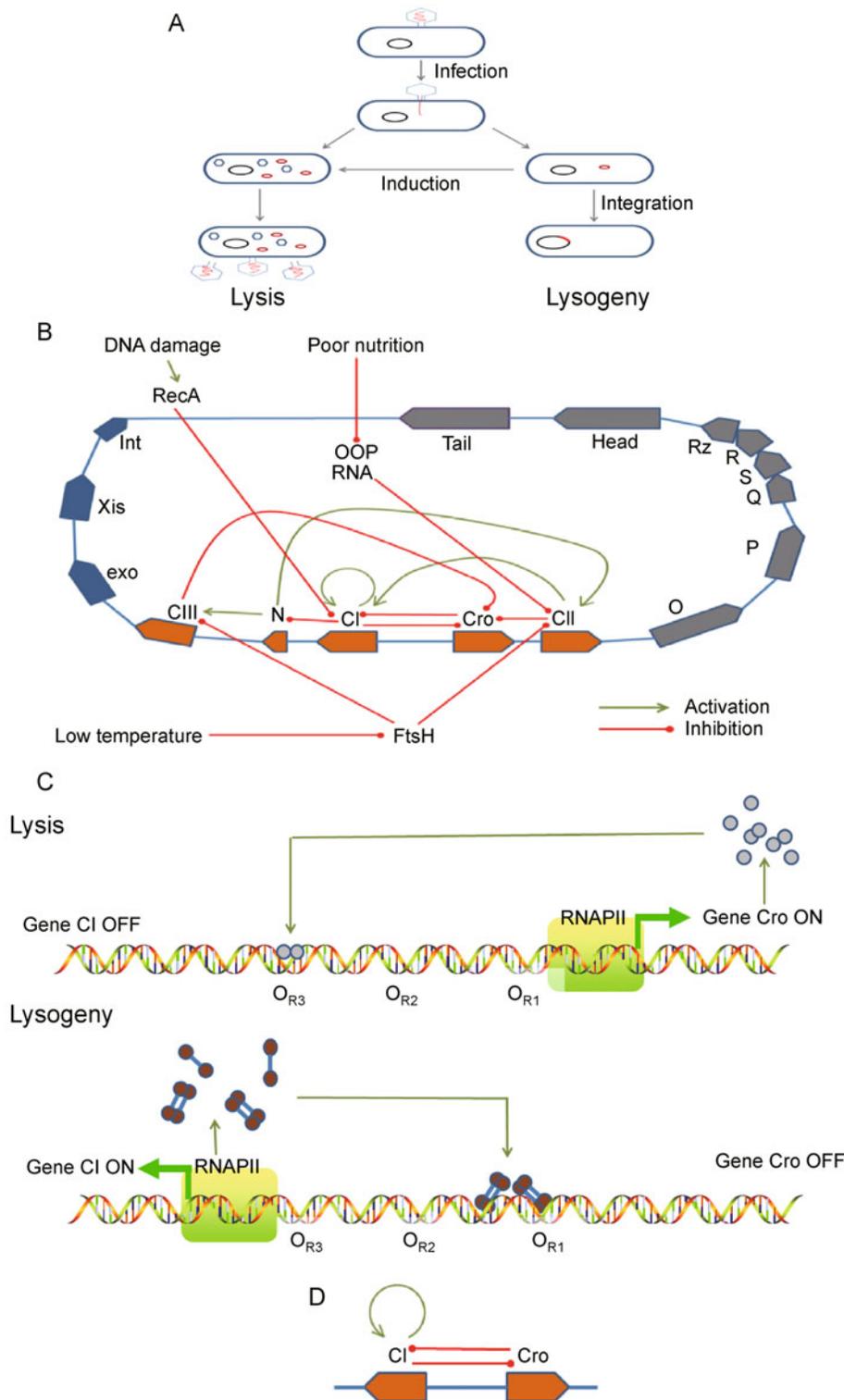


Figure 1. Life cycle, whole genome, regulatory network of Phage lambda. (A) After infection, there are two possible fates of the Phage lambda, one is lysis, Phage lambda replicating and releasing the *E. coli*. The other is lysogeny, the phage DNA integrated into the host cell chromosome. Phage lambda can switch from lysogeny to lysis under the influence of temperature, nutrients, and mutated genes. (B) Whole genome of Phage lambda containing a total of about 40 genes, signaling transduction pathways and gene regulatory networks. (C) configurations of right operator O_R at lysogeny state and lysis state. (D) Endogenous molecular-cellular network of Phage lambda, the core network.

overlaps O_{R2} . RNA polymerase enzymes that bind to the promoter P_R initiate transcription of gene Cro. Similarly, the promoter P_{RM} completely overlaps O_{R3} and partially overlaps O_{R2} . RNA polymerase enzymes bound to the promoter P_{RM} initiate transcription of the CI gene. Dimers of CI product (denoted CI_2) can bind to O_{R1} , O_{R2} preferentially, decrease the transcription of Cro. Conversely, dimers of Cro product (denoted Cro_2) bind to O_{R3} with the highest affinity, decrease the transcription of CI, CI_2 and Cro_2 were defined as gene regulatory proteins, because they can dramatically increase or decrease other genes. We summarized the regulation in Phage lambda in (Figure 1B), this regulatory network can help us understand phage developmental process, the maintenance of lytic and lysogeny states and switch between the states.

Experimental evidence suggests that Phage lambda genetic switch module is deeply hierarchical. CI and Cro play key roles in determining the heritable states of Phage lambda (Figure 1C). In lysogeny state, CI occupies the operator, blocking the synthesis of Cro and also activating its own synthesis. Expression of a critical gene regulatory protein, such as CI and Cro, can trigger the expression of a whole battery of downstream genes. In lytic state, the protein Cro occupies a different site in the operator, blocking the synthesis of CI but allowing synthesis itself. In the lysogeny state most of the DNA of the stably integrated bacteriophage is not transcribed; in the lytic state, this DNA is extensively transcribed, replicated, packaged into new bacteriophage, and released by host cell lysis [47].

The heart of this autonomous regulatory machinery are two gene regulatory proteins, CI and Cro, repressing the expression of each other (Figure 1D), and this form the endogenous molecular-cellular network for Phage lambda genetic switch, an arrangement that gives rise to just two heritable stable states. We assume this simplified version of the regulatory system can determine the fate of Phage lambda in the *E. coli* host cell, and we will verify this assumption by comparing quantitative predications and experimental data.

Quantification of endogenous network of Phage lambda

The above simplification and interpretation of Phage lambda's two stable states and genetic switch is just descriptive. To test our biological understanding and provide precision quantitative understanding in defining the concepts and mechanisms, the best way is to give a mathematical formulation of these processes [16,48].

The modeling of Phage lambda gene regulation is based on one elegant physical-chemical model [48]. With the biological constraints, there are 40 possible configurations of the combination of CI, Cro and RNA_p bound at the right operator. The probability of each configuration

may be determined by equilibrium statistical thermodynamic. And the transcription rate of CI and Cro at each configuration can be measured by *in vitro* or *in vivo* experiment.

For each of the 40 states, its probability is a function of states' free energy, concentrations of CI, Cro and RNA_p according to principles of statistical thermodynamics. The probability in configuration s may be written as

$$P(s) = Z^{-1} [CI]^{i(s)} [Cro]^{j(s)} [RNA_p]^{k(s)} \exp\left(-\frac{\Delta G(s)}{RT}\right), \quad (1)$$

where $Z = \sum_s [CI]^{i(s)} [Cro]^{j(s)} [RNA_p]^{k(s)} \exp\left(-\frac{\Delta G(s)}{RT}\right)$,

$\Delta G(s)$ is the Gibbs free energy of the s configuration, the determination of the energies form a major cornerstone of the model, R is the gas constant and T the absolute temperature, $[CI]$, $[Cro]$ and $[RNA_p]$ are the concentrations of unbound CI dimer, Cro dimer and RNA_p molecules, indices $i(s)$, $j(s)$ and $k(s)$ indicate the number of protein bound to an operator in the s configuration, and have the value (0,1,2,3). Table 1 gives the 40 configurations corresponding to right operator.

According to the probability distribution of each configuration, and corresponding transcription rate and degradation rate of CI mRNA and Cro mRNA, the dynamical change of CI mRNA and Cro mRNA number with respect to time can be written as

$$\frac{dCI(t)}{dt} = \sum_s k_{CI}(s)P(s) - \frac{CI(t)}{\tau_{CI}}, \quad (2)$$

$$\frac{dCro(t)}{dt} = \sum_s k_{Cro}(s)P(s) - \frac{Cro(t)}{\tau_{Cro}}. \quad (3)$$

Stochasticity is ubiquitous in biological systems, for the present modeling the fluctuation in the number is an order of $N^{1/2}$, if the numbers of CI and Cro were macroscopically large then deterministic model would be an entirely accurate description of the dynamics, the numbers are however only in the range of hundreds. Hence the fluctuation is not negligible. Therefore, we consider two independent standard Gaussian and white noise sources:

$$\frac{dCI(t)}{dt} = \sum_s k_{CI}(s)P(s) - \frac{CI(t)}{\tau_{CI}} + \zeta_{CI}(t), \quad (4)$$

$$\frac{dCro(t)}{dt} = \sum_s k_{Cro}(s)P(s) - \frac{Cro(t)}{\tau_{Cro}} + \zeta_{Cro}(t). \quad (5)$$

We further assume:

Table 1. The 40 configurations corresponding to right operator^{a)}.

State	P _{RM}	O _{R1}	O _{R2}	O _{R3}	P _R	<i>i</i> (<i>s</i>)	<i>j</i> (<i>s</i>)	<i>k</i> (<i>s</i>)
1						0	0	0
2				R ₂		1	0	0
3			R ₂			1	0	0
4		R ₂				1	0	0
5			R ₂	R ₂		2	0	0
6		R ₂	R ₂			2	0	0
7		R ₂		R ₂		2	0	0
8		R ₂	R ₂	R ₂		3	0	0
9				C ₂		0	1	0
10			C ₂			0	1	0
11		C ₂				0	1	0
12			C ₂	C ₂		0	2	0
13		C ₂	C ₂			0	2	0
14		C ₂		C ₂		0	2	0
15		C ₂	C ₂	C ₂		0	3	0
16			C ₂	R ₂		1	1	0
17		C ₂		R ₂		1	1	0
18		C ₂	C ₂	R ₂		1	2	0
19		C ₂	R ₂			1	1	0
20			R ₂	C ₂		1	1	0
21		C ₂	R ₂	C ₂		1	2	0
22		R ₂	C ₂			1	1	0
23		R ₂		C ₂		1	1	0
24		R ₂	C ₂	C ₂		1	2	0
25		C ₂	R ₂	R ₂		2	1	0
26		R ₂	R ₂	C ₂		2	1	0
27		R ₂	C ₂	R ₂		2	1	0
28					RNA _p	0	0	1
29		R ₂			RNA _p	1	0	1
30		C ₂			RNA _p	0	1	1
31	RNA _p		R ₂			1	0	1
32	RNA _p		R ₂	R ₂		2	0	1
33	RNA _p		R ₂	C ₂		1	1	1
34	RNA _p					0	0	1
35	RNA _p			R ₂		1	0	1
36	RNA _p			C ₂		0	1	1
37	RNA _p		C ₂			0	1	1
38	RNA _p		C ₂	R ₂		1	1	1
39	RNA _p		C ₂	C ₂		0	2	1
40	RNA _p				RNA _p	0	0	2

^{a)} R₂ stands for CI (λ repressor) dimer and C₂ for Cro dimer, indices *i*(*s*), *j*(*s*) and *k*(*s*) indicate the number of CI dimer, Cro dimer and RNA_p molecules bound to operators in the *s* configuration, have the value (0,1,2,3).

$$\langle \zeta_{CI}(t) \zeta_{CI}(t') \rangle = 2D_{CI} \delta(t-t'), \quad (6)$$

$$\langle \zeta_{Cro}(t) \zeta_{Cro}(t') \rangle = 2D_{Cro} \delta(t-t'), \quad (7)$$

$$\langle \zeta_{CI}(t) \zeta_{Cro}(t') \rangle = 0. \quad (8)$$

Here $\frac{dCI(t)}{dt}$ and $\frac{dCro(t)}{dt}$ mean the rate CI and Cro

number changes, $P(s)$ is the probability in configuration s , $k_{CI}(s)$ and $k_{Cro}(s)$ mean the transcript rate of CI and Cro in configuration s , $\sum_s k_{CI}(s)P(s)$ and $\sum_s k_{Cro}(s)P(s)$ mean the production term of CI and Cro, τ_{CI} and τ_{Cro} mean the decay constant of CI and Cro, $\zeta_{CI}(t)$ and $\zeta_{Cro}(t)$ are

Gaussian and white noise, $\langle \rangle$ denotes the average over noise, Eqs. (6) and (7) define 2×2 diffusion matrix D_{CI} and D_{Cro} .

The noise strength may contain contributions from the production and decay rates, assuming each is dominated by one single independent reaction. Such a noise may be called the ‘intrinsic’ noise. Other noise sources, ‘extrinsic’ noises, also exist. We therefore obtained two coupled stochastic differential equations as minimal quantitative model for Phage lambda genetic switch.

For the quantitative and visualized description of the robustness and stability of Phage lambda’s two stable states and switch, adaptive landscape is an ideal tool [14]. The powerful adaptive landscape concept was first proposed by several biologists independently a long time ago [49,50]. Similar ones have been repeatedly and independently proposed in biology since then in many fields such as population genetics [50], developmental biology [49,51], gene regulation and genetic switch [16,52], protein folding [53]. However, this concept has suffered certain conceptual and theoretical problems, such as being used just as a metaphor, and has been nearly forgotten by modern molecular biologists. Until recently, it has been demonstrated that there exists a unique decomposition such that from the stochastic differential equation, we can obtain the adaptive landscape of the system [54,55].

Recent progresses allow us to construct the adaptive landscape based on the endogenous molecular-cellular

network of Phage lambda. The dynamic state of the network is represented by a particle whose position is given by instantaneous protein numbers. The potential function maps a landscape in the protein number space. For a genetic switch (Figure 2), there are two potential minima corresponding to two epigenetic states. The area around each of the minima forms the attractive basin. The state of the network always tends to relax to one of the minima. The phage sees two minima and one saddle point in the potential landscape. Those two minima correspond to the lytic and lysogenic states. Once the phage is at one of the minima, the probability rate for it to move into another minimum is related to the potential barrier height.

The analysis of robustness of Phage lambda genetic switch demonstrates that its epigenetic states are stable against the variations in parameters and robust against major changes in terms of mutations. Then how does the switching take place? From theoretical point of view, there are two channels that the phage can be induced from lysogenic growth to lytic growth. In reality phage seems to use both of these strategies. For clarity, we begin by discussing these two channels separately.

The first channel is through the deterministic terms. For example, when damaging the DNA of host *E. coli*, protein Rec would be active to cleave CI, which makes the Phage switch from lysogenic state to lytic state more easily than wild type. Mathematically, introducing CI monomer cleavage is equivalent to substituting N_{CI} with aN_{CI} . a is a factor representing cleavage strength. If a is small, say 0.1, meaning 90% CI monomers are cleaved, the lysogenic state is still stable with a lifetime almost unchanged. If a is smaller than 0.02, we find that lysogenic state is no longer stable, i.e., no longer a potential minimum (Figure 2B). The interpretation of such a small a is that almost each of the CI monomer is

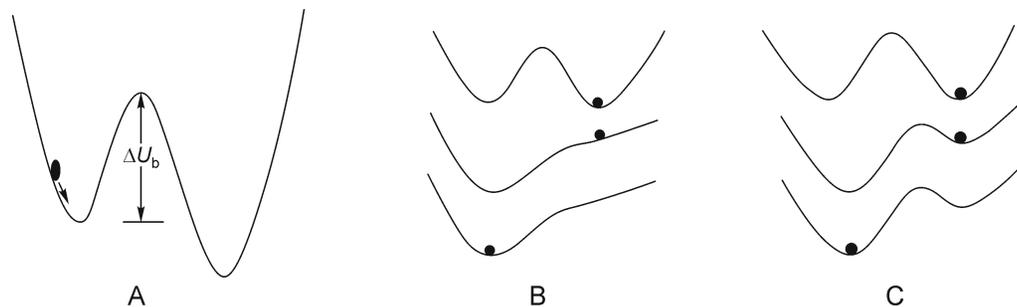


Figure 2. Illustration of adaptive landscape of Phage lambda genetic switch. (A) The dynamic state of the network is represented by a particle whose position is given by instantaneous protein numbers. The potential function maps a landscape in the protein number space. For Phage lambda genetic switch, there are two potential minima corresponding to two epigenetic states. The area around each of the minima forms the attractive basin. The state of the network always tends to relax to one of the minima. The fluctuation may bring the network from one minimum to another. (B) Before switching the Phage grows in lysogenic state. The potential barrier separating the lysogenic state and the lytic state is high. When *recA* is activated, this barrier is lowered. The lifetime of lysogenic state reduces drastically and the Phage switches to lytic state. (C) When stochastic effect is included, the switching happens when the lysogenic potential minimum become too shallow to confine fluctuation.

cleaved. Apparently a uniform CI cleavage alone without introducing extra noise to CI levels is not an efficient way for the induction of the switch. Many other factors can affect switch efficiency, such as mutation of CI or Cro or operator (O_R). These changes will affect the parameters in SDE.

The second channel of induction is to change the noise level of the SDE. For example, we increase the noise level of CI protein number while keeping all the other conditions intact. Mathematically it means to increase $\zeta_{CI}(t)$ and D_{CI} , while keeping all the other terms unchanged. As a result, the potential energy U is also changed. Therefore for different noise level, the phage moves in different potential landscape. Such a change of the noise level has a drastic effect. It changes the minima of the potential well of lysogen by making it shallower (Figure 2C). As a good approximation, the barrier height of the lysogenic potential well scales inversely with the noise strength. Doubling the noise level reduces the potential barrier by half. As a result, the increased noise level drastically decreases the lifetime of lysogenic state. The lifetime of the lytic state, on the other hand, remains unchanged. The combination of these two changes in the potential landscape brings the phage to lytic growth efficiently.

Validation of endogenous molecular-cellular network in Phage lambda

The quantitative model based on the endogenous molecular-cellular hypothesis of Phage lambda genetic switch got positive results by comparing with experiment, thus it lays a solid foundation for our understanding of this biological process.

The modeling results of Phage lambda genetic switch validated the assumptions we proposed for biological systems in section *FORMING ENDOGENOUS MOLECULAR-CELLULAR NETWORK HYPOTHESIS*. We have showed that Phage lambda can complete switch from lysogeny to lysis autonomously, and the switch is affected by nutrition, temperature, DNA damage. The regulatory mechanisms of Phage lambda genetic switch

can be simplified and specified by interactions of two important proteins, CI and Cro by repressing the expression of each other the two proteins formed a closed autonomous network, which implies Phage lambda's two stable states lysogeny and lysis. Thus, we argue that Phage lambda genetic switch has the properties of modularization, deep hierarchy, and regulatory mechanisms can be simplified and specified by endogenous molecular-cellular network.

We also obtained that lytic and lysogenic states are inherited and intrinsic states shaped by evolution. Genetic switch can be regarded as switch from one intrinsic state to the other intrinsic state, and there are two channels that Phage lambda can be induced from lysogenic growth to lytic growth. One is through the topology changes of endogenous molecular-cellular network, such as mutated CI and Cro; the other is stochastic noise which can switch the stable states of endogenous molecular-cellular network from one to the other.

More importantly, such small core network permits us make non-trivial predictions. It was predicted that there would be a very large corporation energy needed to explain the stability [16]. At almost the same time the experiment revealed that the looping effect provides exactly the same amount of 'corporation' energy [17]. Another major prediction is on existence of large extrinsic noise. Later experiment completely validated such prediction [56].

We have now compared the quantitative results obtained from theoretical model with experiment (Table 2), and we got positive conclusion about the modeling. Further, we make predications based on this quantitative model. Hence, the endogenous molecular-cellular network hypothesis is validated in Phage lambda. With such a concrete and quantitative example, we will apply this hypothesis to human cancer in next section.

MOVING UP IN COMPLEXITY: HEPATOCELLULAR CARCINOMA

The endogenous molecular-cellular network hypothesis has been validated in Phage lambda genetic switch. Now,

Table 2. Comparison between the calculation and the experiment data (adapted from Ref. [16])^b.

Phage genotype	Relative CI level in lysogen	Relative Cro level in lysis	Switching frequency to lytic state $recA^-$ per minute	Switching frequency to lytic state $recA^+$ per minute
	Theoretical (Experimental)	Theoretical	Theoretical (Experimental)	Theoretical
λ^+	100% (100%)	100%	1×10^{-9}	1×10^{-5}
λO_R121	20% (25%–30%)	100%	3×10^{-9}	3×10^{-5}
λO_R323	70% (60%–75%)	70%	7×10^{-9}	1×10^{-5}
$\lambda O_R3'23'$	50% (50%–60%)	130%	1×10^{-9}	2×10^{-5}

^b The relative CI level in lysogen and switching frequency to lytic state per minute ($recA^+$) calculated from model were compared with experimental data. The relative Cro level in lysis and switching frequency to lytic state per minute ($recA^-$) are predications.

we apply this hypothesis to Human cancer, we take Hepatocellular carcinoma (HCC) as an example. HCC is the main primary liver tumor, accounting for 85%–90% of primary liver cancers diagnosed [57]. The prevalence varies greatly, Asia is a high prevalence area and more than half of the world's HCC cases occur in China [58,59]. We take hepatocellular carcinoma as an example using the endogenous molecular-cellular network hypothesis to quantify and understand the genesis and progression of cancer.

The endogenous molecular-cellular network of HCC hypothesizes that in order to maintain the normal physiological function and developmental process of liver, many fundamental functional modules are needed and have been formed by evolution. These modules and cross-talks between modules may be simplified and specified by regulatory proteins. The interactions between these proteins form a nonlinear, stochastic, and collective dynamical network. We tentatively name it as the liver endogenous molecular-cellular network. The endogenous molecular-cellular network may generate many locally stable states with obvious or non-obvious biological function. Normal liver and HCC are assumed to be stable states of the endogenous molecular-cellular network.

In the following, we will first briefly summary the fundamental modules and its key regulator proteins in liver, then present the work-flows to establish and quantify endogenous molecular-cellular network of HCC. At last we provide some straightforward insights on human cancer and strategies for cancer prevention, cure, and care based on our current hypothesis.

Functional modules

Firstly, we briefly discuss the fundamental and conserved modules and molecular-cellular agents of each module. The fundamental modules include the cell cycle, cell death, inflammation, metabolism, angiogenesis, cell adhesion, liver specific functions (Figure 3A).

Cell cycle

Cell cycle is fundamental to life, it can compensate hepatocytes death. Adult liver is normally quiescent and has a very low level of hepatocyte cell division [60]. However, most hepatocytes rapidly proliferate in response to a reduction in liver mass caused by physical, chemical, nutritional, vascular, virus-induced liver injury or partial hepatectomy [61,62]. Earlier studies have demonstrated that the regeneration afterwards involves all hepatocytes in the liver to enter into cell proliferation [62]. In molecular-cellular level, quiescence and proliferation of hepatocyte is regulated by proteins including

cyclin D-CDK4/6, cyclin E-CDK2 [63], Rb, E2F [9], p21 [64], C/EBP α [63].

Cell death

Cell death can move damaged cell. Together with the cell division this process ensues continuously and seamlessly from birth to death homeostatic [65]. Cell death is an indispensable and cardinal element to life. One important mode of cell death is apoptosis. Apoptosis can be triggered by extracellular or intracellular pathways, extracellular pathway can be activated by death receptors [66]. Intracellular (mitochondrial) pathway of apoptosis can be activated by a myriad of cellular stressors, e.g., metabolic stress including oxidative stress, DNA damage [67]. At molecular-cellular level, the regulators and executor of apoptosis include Bcl-2 [68], tBid [69], XIAP [70], Bax [71], Cytochrome c [72], Caspase 9, Caspase 8, Caspase 3 [73].

Immune response

The immune system can protect the proper functions of liver by detecting a wide variety of agents, such as virus. Inflammation is one of the first responses of the immune system to infection. Inflammation is a protective attempt by the organism to remove the injurious stimuli and to initiate the healing process. Without inflammation, wounds and infections would never heal. However, inflammation can also lead to a host of diseases, such as cancer [74,75]. Regulators and cytokines related to inflammation include NF- κ B [76], TNF α [77], IL-1 [78], IL-6, stat3 [79], IL-10 [80].

Metabolism

To meet these needs of cell division, proliferation cells acquire alterations to the metabolism of all four major classes of macromolecules: carbohydrates, proteins, lipids and nucleic acids. Most proliferation cancer cells rely on aerobic glycolysis [81], a phenomenon termed 'the Warburg effect'. In fact, all proliferating cells are adapted to facilitate the uptake and incorporation of nutrients into the biomass needed to produce a new cell. Metabolic regulation is considered as component of the cell growth machinery [82]. The regulators of metabolism include HIF [83], Myc, Akt [84].

Angiogenesis

Angiogenesis is the physiological process involving the growth of new blood vessels from pre-existing vessels.

Angiogenesis is a normal and vital process in growth and development, as well as in wound healing and in granulation tissue. However, it is also a fundamental step in the transition of tumor from a dormant state to a malignant one [85]. VEGF is one of the most important regulators of angiogenesis [86].

Cell adhesion

Collective cell migration plays a key role in developmental processes that range from gastrulation to organogenesis [87]. Wound healing and immune response also require the orchestrated movement of cells in particular directions of specific locations. Errors during this process have serious consequences, including mental retardation, vascular disease, tumor formation invasion and metastasis. The regulators of Invasion and metastasis include TGF- β , E-cadherin, Integrin [88].

These fundamental modules are almost conserved for all organ, even for different organisms. And because of evolution, the core biological modules for different organs and organisms use essentially the same proteins to carry out certain tasks.

Liver specific modules

Besides the conserved modules for all organs, liver has specific modules to maintain its physiological function. Experiments have revealed that the transcriptional control is vital to maintain normal physiological hepatocyte function, and liver environment is vital to liver to deal with outside interference.

Liver specific modules include hepatocyte specific transcriptional factors and specific liver microenvironment. In order to maintain liver's functions, such as energy storage and metabolic processing, the unique gene expression that defines the hepatocyte conforms to a set of transcription factors. By over-expressing core factors, fibroblasts can be induced to functional hepatocyte-like cells [33]. Experiments have revealed three hepatocyte specific transcription factors, FOXA2, HNF4a, and C/EBPa [89].

In liver microenvironment, many cytokines and factors can affect liver regeneration, proliferation, survival, immune response. These cytokines and factors include bile acids [90].

Endogenous molecular-cellular network of HCC

After the physiological function and developmental process of liver have been specified by modules and molecular-cellular agents, the endogenous molecular-cellular network arises from the interactions among the molecular-cellular agents (Figure 3B). The activation and

inhibition of molecular-cellular agents are obtained from biological experiments. For example, the inhibition of Cyclin D-CDK4/6 and Cyclin E-CDK2 by p21 in endogenous molecular-cellular network is based on the fact that p21 inhibit the activity of Cyclin D-CDK4/6 and Cyclin E-CDK2 by direct binding to these complexes [64].

Quantification of endogenous molecular-cellular network of HCC

There are many mathematical frameworks for the quantified biological network [91–93]. Here we adopt one widely used in engineering [29]. The endogenous molecular-cellular network can be quantified by a set of stochastic differential equations. Each protein, for instance protein one, can be quantified as

$$\frac{dx_1(t)}{dt} = f_1(X) - \frac{x_1(t)}{\tau} + \zeta(X, t). \quad (9)$$

In this equation, $x_1(t)$ means the content or activity of protein one; $f_1(X)$ means the production of protein one is influenced by other agents according to endogenous molecular-cellular network, $X = (x_1, x_2, x_3, \dots, x_n)$ means all the agents in endogenous molecular-cellular network; $x_1(t)/\tau$ means the natural protein degradation, τ is the degradation rate; $\zeta(X, t)$ means the noise in biological systems.

Coarse quantitative modeling, that means $x_1(t)$ is not the absolute content or activity of protein one, but their relative content or activity, allows us to simplify our equations, but we can only decide relative change of each protein, up-regulation, down-regulation or unchanged regulation. We assume the stochastic noise can be temporarily ignored. Eq. (9) can be simplified as

$$\frac{dx_1(t)}{dt} = f_1(X) - \frac{x_1(t)}{\tau}. \quad (10)$$

The production of protein one $f_1(X)$ is determined by specific endogenous molecular-cellular network. The activating or up-regulating may be quantified as

$$f_1(X) = \frac{ax_2^n}{1 + ax_2^n}. \quad (11)$$

The inhibiting or down-regulating as

$$f_1(X) = \frac{1}{1 + ax_2^n}. \quad (12)$$

If there are multiple inputs to affect the output content or activity of protein one, for example if x_2 and x_3 will activate or up-regulate x_1 when either x_2 or x_3 is activated, and x_4 and x_5 will inhibit or down-regulate x_1 when either

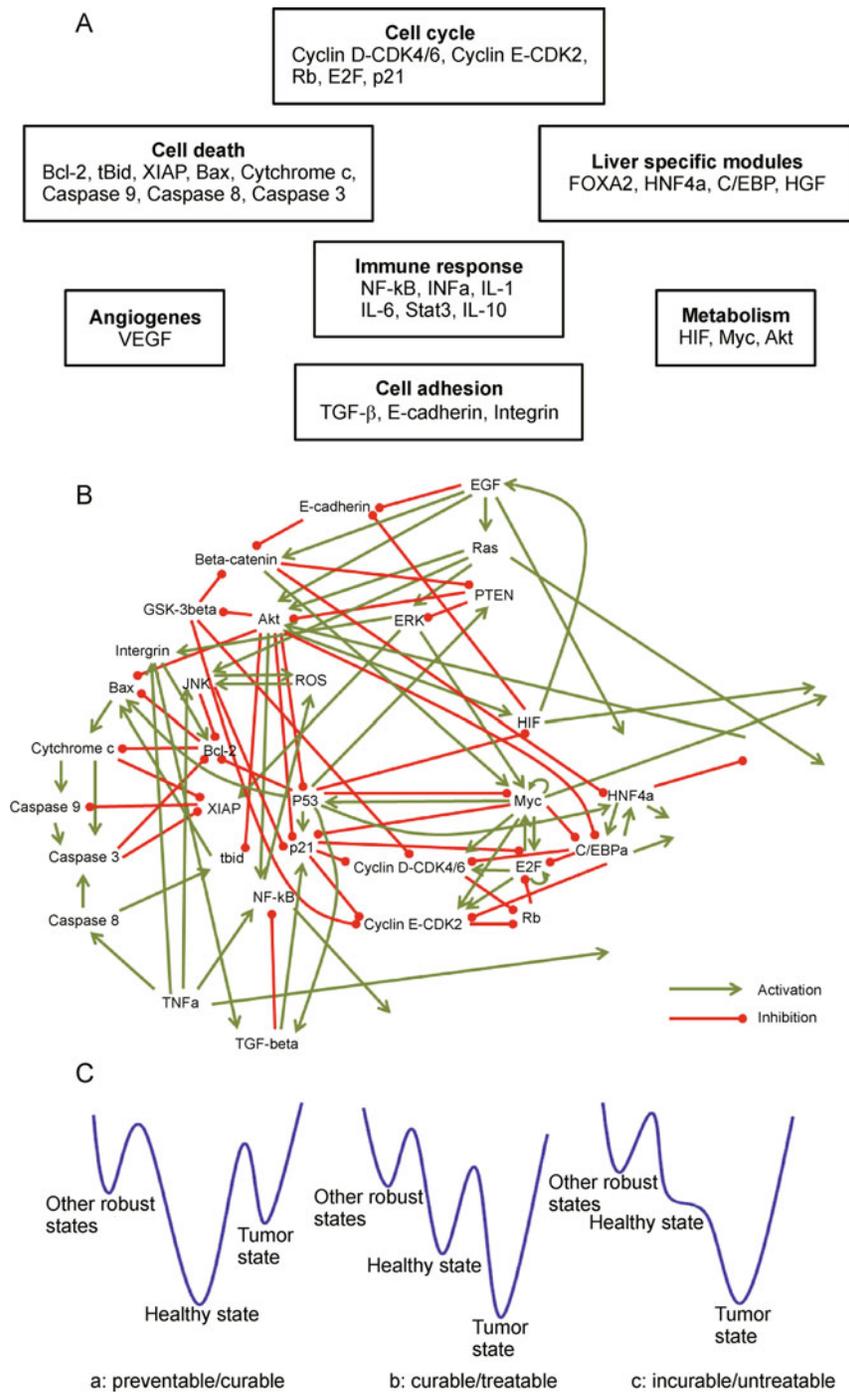


Figure 3. Minimum set of modules of HCC and typical situations of the adaptive landscape. (A) The minimum modules include cell proliferation, cell death, inflammation, metabolism, cell adhesion, angiogenesis, liver specific functions, in molecular-cellular level, these modules and cross-talks between modules may be simplified and specified by regulatory proteins, modules cross talk to each other to form the endogenous molecular-cellular network. (B) Endogenous molecular-cellular network of HCC (incomplete). (C) Three typical situations of the functional landscape, the vertical scale illustrates the relative stability of robust states, healthy, tumor and others, in the multiple dimensional state space. (a) The healthy state is a globally stable under normal conditions; (b) due to genetic and epidemiologic influence on the endogenous network, tumor or cancer states may become more stable than healthy state. Such metastable healthy state may still have a long life time for the whole organism being viable; (c) a very ‘damaged’ endogenous network may not be able to produce a locally stable healthy state.

x_2 or x_3 is activated, and protein one is dominated by inhibition, then the specific is

$$f_1(X) = \frac{a(x_2^n + x_3^n)}{1 + a(x_2^n + x_3^n)} \times \frac{1}{1 + a(x_4^n + x_5^n)}. \quad (13)$$

The nonlinear dynamical interactions among the endogenous agents can generate many other locally stable states with obvious or non-obvious biological functions. Normal liver and HCC are assumed to be stable states of the endogenous molecular-cellular network. HCC genesis and progression is hypothesized as that transition from normal liver state to HCC state in endogenous molecular-cellular network, similar to the genetic switch of Phage lambda. Based on our endogenous molecular-cellular network and its quantification by SDE, we can construct adaptive landscape of liver (Figure 3C). We show it as the following: For a healthy liver, its adaptive landscape is Figure 3C(a): The health state is its global stable state, it is hard for normal cell to pass through the barrier to tumor state; however, if some oncogenes mutation, or hepatitis occurs, adaptive landscape will become Figure 3C(b): The tumor state is the global stable state, and for normal cell it is still needed to overcome a barrier to become tumor cell, however, the barrier is lower than the health state; the worst is Figure 3C(c): The healthy state becomes totally unstable, there is no barrier for normal tissue to become tumor.

Implications of endogenous molecular-cellular network of HCC

The endogenous molecular-cellular network hypothesis suggested that the regulatory machinery of cancer genesis and progression may be quantitatively modeled by a nonlinear stochastic dynamical system. Based on the quantitative modeling additional insights of cancer genesis and progression can be suggested.

Two remarks are in order before proceeding. First, one may wonder that from genetic switch of bi-stable states to endogenous network of multiple stable states it is a bit too quick in moving up in complexity. Technically such concern is reasonable and on the other hand it is not as serious as it would appear. A major support for our approach is that the modules and pathways have been mathematically modeled in similar fashion as ours [51]. The landscape and related ideas indeed work in such intermediate level of network. More importantly, mathematically in order to transit among multiple states, an autonomous and closed network is needed. This would force the modeling to directly face the network level. It is what we have reasoned and carried out. Second, one may also wonder where are the roles of high throughput data so common in systems biology study in our above elaboration [94,95]. Due to embedded large measurement

errors currently the high throughput data sets are not ideal for constructing precise endogenous core networks. Nevertheless, we point out that there are at least two good roles such high throughput data sets may play: at the beginning of network construction they can be used to roughly define the endogenous network and at the end of prediction validation they can be used to test the predictions from such mathematical modeling.

Now we come to the implications. The first possible implication of endogenous molecular-cellular network hypothesis is that cancer is an intrinsic stable state shaped by evolution. The genesis and progression of cancer is the transition from one intrinsic state to another, and eventually to an intrinsic cancer state. The transition is affected by genetic alterations, carcinogen, random events in cell signaling or gene expression, et al. The straightforward predication of cancer being an intrinsic stable state is that there is nonzero probability that cancer be formed without genetic defect. Hence, within this framework a possible dynamical trajectory for cancer is suggested. In the same vein of reasoning, there is a nonzero probability of cancer regression without any major interventions. This predication is testable and directly contrasts to the hypothesis that cancer is a genetic disease.

A second implication is that the progression of cancer is not arbitrary. In adaptive landscape there are one or more saddle points between normal state and cancer state. This inference is certainly consistent with known seed and soil summary of cancer. From theoretical point of view, transition from normal to cancer needs to pass through critical saddle points. The quantitative modeling of cancer genesis and progression suggests that the quantitative standard for early state cancer detection is possible. Also, quantitative modeling of cancer provides us new tools to find ideal drug targets.

Having discussed the main features of the endogenous network theory for cancer, we come to the very first problem which a quantitative framework for cancer genesis and progression need to solve: effects of both genetic and environmental factors. Here we point out how to start to handle such problem. On the genetic level, the removing and adding functional genes via mutations can be simply represented by decreasing and increasing numbers of molecular and cellular agents in the network, hence the dynamical variables. On the mathematical level such manipulation can be handled by standard procedures which pose no additional conceptual issues. As for the environmental factors, including possibly micro RNA's, heat shock proteins, a main effect is to modify the regulation strengths, leaving the backbone of the network intact. Such effects can be readily represented by changing parameters in the mathematical equations, Eq. (9). This has been quantitatively demonstrated in case of

Phage lambda. We should point out that both genetic and environmental effects can change the relative stability of states, as well as their absolute stability. Such possibility opens new ways for both genetic and chemical interference hence to new scenarios to understand and treat cancer.

RAMIFYING THE ENDOGENOUS NETWORK THEORY

Evolution and adaptive landscape

The endogenous molecular-cellular network of Phage lambda and HCC would be conserved, because they both have been shaped by millions, or even billions, years of evolution. Human cancer network may borrow evolutionary stable elements from processes of wound healing, development and stress response. Such network may have a relative few classes of stable states, which may have at least two types of biological interpretations: the strategies of an organism to respond to various challenges and the number of cell types to make up an organism. It is known that cancer genesis is rare and is hard to eliminate. Cancer would be an intrinsic stable state of the endogenous molecular-cellular network dynamics. Mathematically, the structure of the endogenous molecular-cellular network model is similar to the Morse-Smale dynamical system [96] perturbed by noise: It is structurally stable with a finite number of attractors and the transitions between attractors are possible in the presence of noise. A suitable quantity to describe such process is the potential function, as the quantification of the adaptive landscape. Recent progress on stochastic dynamical systems allow us to construct the adaptive landscape based on the endogenous molecular-cellular network [14,54,55]. Such important biological applications may well move the study of Morse-Smale system into stochastic domain.

Within the endogenous network theory formulation, it has been reasoned that the development processes have been automatically included [97]. Some important phenomena difficult to be explained within genetic centric theories, such as orders of magnitude change in phenotype rates caused by simple mutations, may be readily and quantitatively explained [98]. In order to incorporate environmental effects into the endogenous network theory it is necessary to consider the dynamics of global metabolic networks. Related studies have been carried out in similar quantitative manner [99–101] to acknowledge such need. As has been reasoned, biological processes are highly hierarchical. In addition to the need of metabolic network, higher levels, such as hormones, also play an important role. As our understanding goes further, the endogenous network will be expanded further, too.

Theories and hypotheses similar to the endogenous network hypothesis have been also proposed by others. The cracks in cancer genetic mutation theory was noticed and a similar theory predicting the intrinsic inevitability of cancer was proposed [102]. They have also been pushing along a similar qualification road [103]. In a different study it was proposed that ionizing radiation can give rise to similar effects through two distinct and independent routes, genetic, and epigenetic and that phenotype is represented by a stable attractor [104]. From the biological perspective all those theories share the same set of considerations. One major difference may be in quantitative formulations: To the best knowledge of present authors a complete and consistent framework has been developed for the endogenous network theory. In addition, along our quantitative development the outstanding controversy on the adaptive landscape in evolutionary biology has been resolved [42].

More specific and related theories have also been proposed. It was proposed that cancer is an atavistic condition that occurs when genetic or epigenetic malfunction unlocks an ancient ‘toolkit’ of pre-existing adaptations [105,106]. Such theory is clearly consistent with the evolutionary dynamics structure embedded in the present endogenous network theory. During last few decades there has been a consistent effort to reveal the regulatory mechanism of developmental process in sea urchin [37]. Such study is another important support to the construction of endogenous network.

It should be pointed out that after many years of theoretical and experimental studies, key concepts of endogenous network theory, such as landscape and states, start to get into mainstream biological and medical research [41,107,108], even in the light of genomic centric theories [109].

Comparison with other proposed frameworks to understand cancer

It is one of the biggest challenges for contemporary biology to understand and to manipulate the complex regulatory machinery underlying biological systems quantitatively, such as cancer. Many other high throughput based frameworks also have been proposed to meet this challenge, such as genomic sequencing [7], transcriptomics, ENCODE project, et al. [110]. Those frameworks play an important role in accelerating our understanding of biological systems. Nevertheless, those frameworks may have inherent defects in revealing the regulatory machinery of biological systems, though they may certainly bring out important insights. For genomic sequencing, though important, it is clear that one cannot construct the regulatory machinery of biological systems only from genome data, because of the many gaps

between genotype and phenotype [109]. The quantitative analysis of transcriptomics is based too much on statistics, less on biological mechanics. It is likely not the appropriate mathematical approach to describe dynamical regulatory machinery. The ENCODE project examines the principles of the human transcriptional regulatory network high throughput data. They reveal the hierarchical property, feedback loops, evolution. However, it appears that the regulatory derived from ENCODE data cannot reveal the regulatory machinery quantitatively because they are generally descriptive topology networks. As a comparison, it seems that the endogenous molecular-cellular network hypothesis is a more ideal framework than those high throughput based frameworks in quantitatively revealing the regulatory machinery of biological systems.

Beyond Phage lambda and human cancer

The endogenous molecular-cellular network hypothesis certainly captures some essential aspects of complex biological processes. It can be used beyond Phage lambda and human cancer. The three important aspects of the endogenous molecular-cellular network hypothesis, modularization, deep hierarchy and autonomous regulation, are clearly not specifically designed for Phage lambda and human cancer. In fact, those properties suit various hierarchical levels, from proteins to cells, organs, individuals and populations. The endogenous network and adaptive landscape concept has been widely used in various fields, such as stem cell [111], population genetics [50], protein fold [53]. It appears to be applicable to single cell organisms, too. It is very likely the endogenous network is behind the working of all living organisms on Earth.

PERSPECTIVE

We have discussed the basic biological considerations and the key mathematical features behind the cancer endogenous network theory. It is an effort to understand cancer genesis and progression in detail. The hypothesized theory's capturing of the underlying autonomous regulatory machinery is reviewed against two biological systems. By using one of the simplest and most important organisms, the Phage lambda, we have validated hierarchical nature for the endogenous molecular-cellular network of Phage lambda genetic switch. The core network can quantitatively describe the general regulatory machinery for Phage lambda's two important modes. Those validated insights may be used in the more complex systems.

One of the most important applications of this hypothesis is to understand the regulatory machinery

underlying cancer genesis and progression. Work-flow to construct the endogenous molecular-cellular network of HCC and quantitative analysis of the network were provided. Endogenous molecular-cellular network hypothesis suggests that cancer is an intrinsic state shaped by evolution. The genesis and progression of cancer is the transition from intrinsic normal state to intrinsic cancer state and the progression of cancer is not arbitrary. Transition from normal to cancer needs to pass through the critical saddle points, which may provide us quantitative standard for early state cancer detection. With its capacity to take both genetic and environmental effect into consideration the endogenous network theory may provide a best candidate, both qualitatively and quantitatively, to understand cancer genesis and progression.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude for the helpful discussions and or comments with Jianren Gu, Hongyang Wang, Shu Zheng, Sui Huang. One of us (XMZ) thanks the hospitality of SCSB which enabled the collaboration. This work was supported in part by No. 2010CB529200 (PA and GWW).

REFERENCES

1. Kumar, V., Robbins, S., Zhai, Q. and Chen, J. (2009) Textbook of Pathology. Peking University Medical Press.
2. Hajdu, S. I. (2011) A note from history: landmarks in history of cancer, part 1. *Cancer*, 117, 1097–1102.
3. Ao, P., Galas, D., Hood, L. and Zhu, X. (2008) Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution. *Med. Hypotheses*, 70, 678–684.
4. Hanahan, D. and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell*, 144, 646–674.
5. Nowell, P. C. (1976) The clonal evolution of tumor cell populations. *Science*, 194, 23–28.
6. Greaves, M. and Maley, C. C. (2012) Clonal evolution in cancer. *Nature*, 481, 306–313.
7. Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., Li, F., Wu, K., Liang, J., Shao, D., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell*, 148, 873–885.
8. Land, H., Parada, L. F. and Weinberg, R. A. (1983) Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature*, 304, 596–602.
9. Weinberg, R. A. (1995) The retinoblastoma protein and cell cycle control. *Cell*, 81, 323–330.
10. Paget, S. (1989) The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev.*, 8, 98–101.
11. Fidler, I. J. and Poste, G. (2008) The “seed and soil” hypothesis revisited. *Lancet Oncol.*, 9, 808.
12. Wang, X., Ouyang, H., Yamamoto, Y., Kumar, P. A., Wei, T. S., Dagher, R., Vincent, M., Lu, X., Bellizzi, A. M., Ho, K. Y., et al. (2011) Residual embryonic cells as precursors of a Barret's-like metaplasia. *Cell*, 145, 1023–1035.

13. Koshland, D. E. Jr, Goldbeter, A. and Stock, J. B. (1982) Amplification and adaptation in regulatory and sensory systems. *Science*, 217, 220–225.
14. Ao, P. (2009) Global view of bionetwork dynamics: adaptive landscape. *J. Genet. Genomics*, 36, 63–73.
15. Li, F., Long, T., Lu, Y., Ouyang, Q. and Tang, C. (2004) The yeast cell-cycle network is robustly designed. *Proc. Natl. Acad. Sci. USA*, 101, 4781–4786.
16. Zhu, X. M., Yin, L., Hood, L. and Ao, P. (2004) Robustness, stability and efficiency of phage lambda genetic switch: dynamical structure analysis. *J. Bioinform. Comput. Biol.*, 2, 785–817.
17. Zhu, X., Yin, L., Hood, L., Galas, D. and Ao, P. (2007) Efficiency, robustness, and stochasticity of gene regulatory networks in systems biology: λ switch as a working example. In Choi, S. (ed.), *Introduction to Systems Biology*. Humana Press, 336–371.
18. Bizzarri, M., Giuliani, A., Cucina, A., D'Anselmi, F., Soto, A. M. and Sonnenschein, C. (2011) Fractal analysis in a systems biology approach to cancer. *Semin. Cancer Biol.*, 21, 175–182.
19. Pastan, I. and Gottesman, M. (1987) Multiple-drug resistance in human cancer. *N. Engl. J. Med.*, 316, 1388–1393.
20. Gimbrone, M. A. Jr, Leapman, S. B., Cotran, R. S. and Folkman, J. (1972) Tumor dormancy *in vivo* by prevention of neovascularization. *J. Exp. Med.*, 136, 261–276.
21. Hartwell, L. H., Hopfield, J. J., Leibler, S. and Murray, A. W. (1999) From molecular to modular cell biology. *Nature*, 402, C47–C52.
22. Akhurst, R. J. and Derynck, R. (2001) TGF- β signaling in cancer — a double-edged sword. *Trends Cell Biol.*, 11, S44–S51.
23. Feng, G.-S. (2012) Conflicting roles of molecules in hepatocarcinogenesis: paradigm or paradox. *Cancer Cell*, 21, 150–154.
24. Kauffman, S. (2008) Control circuits for determination and transdetermination: interpreting positional information in a binary epigenetic code. In Porter, R. and Rivers J. (eds.), *Ciba Foundation Symposium 29 - Cell Patterning*. John Wiley & Sons, Ltd., 201–221.
25. Auffray, C., Chen, Z. and Hood, L. (2009) Systems medicine: the future of medical genomics and healthcare. *Genome Med.*, 1, 2.
26. Chalancon, G., Ravarani, C. N. J., Balaji, S., Martinez-Arias, A., Aravind, L., Jothi, R. and Babu, M. M. (2012) Interplay between gene expression noise and regulatory network architecture. *Trends Genet.*, 28, 221–232.
27. Hood, L. and Flores, M. (2012) A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *New Biotechnol.*, 29, 613–624.
28. Vital-Lopez, F., Memišević, V. and Dutta, B. (2012) Tutorial on biological networks. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 2, 298–325.
29. Ao, P., Galas, D., Hood, L., Yin, L. and Zhu, X. M. (2010) Towards predictive stochastic dynamical modeling of cancer genesis and progression. *Interdiscip. Sci.*, 2, 140–144.
30. Nurse, P. (2000) A long twentieth century of the cell cycle and beyond. *Cell*, 100, 71–78.
31. Garber, K. (2001) Beyond the Nobel Prize: cell cycle research offers new view of cancer. *J. Natl. Cancer Inst.*, 93, 1766–1768.
32. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131, 861–872.
33. Huang, P., He, Z., Ji, S., Sun, H., Xiang, D., Liu, C., Hu, Y., Wang, X. and Hui, L. (2011) Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature*, 475, 386–389.
34. Tyson, J. J. and Novak, B. (2008) Temporal organization of the cell cycle. *Curr. Biol.*, 18, R759–R768.
35. Spencer, S. L. and Sorger, P. K. (2011) Measuring and modeling apoptosis in single cells. *Cell*, 144, 926–939.
36. Meyer, B. J., Maurer, R. and Ptashne, M. (1980) Gene regulation at the right operator (O_R) of bacteriophage λ . II. O_{R1} , O_{R2} , and O_{R3} : their roles in mediating the effects of repressor and cro. *J. Mol. Biol.*, 139, 163–194.
37. Yuh, C. H., Bolouri, H. and Davidson, E. H. (1998) Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science*, 279, 1896–1902.
38. Baker, S. G. and Kramer, B. S. (2011) Systems biology and cancer: promises and perils. *Prog. Biophys. Mol. Biol.*, 106, 410–413.
39. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2007) *Molecular Biology of the Cell* (5th ed.). Garland Science.
40. Smale, S., Hirsch, M. and Devaney, R. (2003) *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier Science.
41. Bar-Yam, Y., Harmon, D. and de Bivort, B. (2009) Attractors and democratic dynamics. *Science*, 323, 1016–1017.
42. Ao, P. (2005) Laws in Darwinian evolutionary theory. *Phys. Life Rev.*, 2, 117–156.
43. Schäfer, M. and Werner, S. (2008) Cancer as an overheating wound: an old hypothesis revisited. *Nat. Rev. Mol. Cell Biol.*, 9, 628–638.
44. Williams, C. S., Mann, M. and DuBois, R. N. (1999) The role of cyclooxygenases in inflammation, cancer, and development. *Oncogene*, 18, 7908–7916.
45. Neidhardt, F. (1987) *Escherichia Coli and Salmonella Typhimurium: Cellular and Molecular Biology*. American Society for Microbiology.
46. Cairns, J., Stent, G. S. and Watson, J. D. (2007) *Phage and the Origins of Molecular Biology* (The Centennial Edition). Cold Spring Harbor Laboratory Press.
47. Ptashne, M. (2004) *A Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press.
48. Shea, M. A. and Ackers, G. K. (1985) The O_R control system of bacteriophage lambda: a physical-chemical model for gene regulation. *J. Mol. Biol.*, 181, 211–230.
49. Waddington, C. (1957) *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Allen & Unwin.
50. Wright, S. (1982) The shifting balance theory and macroevolution. *Annu. Rev. Genet.*, 16, 1–19.
51. Wang, J., Zhang, K., Xu, L. and Wang, E. (2011) Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA*, 108, 8257–8262.
52. Cao, Y., Lu, H. M. and Liang, J. (2010) Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda. *Proc. Natl. Acad. Sci. USA*, 107, 18445–18450.
53. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Struct. Func. Bioinform.*, 21, 167–195.
54. Ao, P., Kwon, C. and Qian, H. (2007) On the existence of potential landscape in the evolution of complex systems. *Complexity*, 12, 19–27.
55. Yuan, R. S. and Ao, P. (2012) Beyond Itô versus Stratonovich. *J. Stat. Mech.*, 2012, P07010.
56. Tian, W., Zhu, H., Lei, X. and Ao, P. (2011) Extrinsic vs. intrinsic noises in Phage lambda genetic switch. In *Proceedings of 2011 IEEE Conference on Systems Biology*. 67–71.

57. El-Serag, H. B. and Rudolph, K. L. (2007) Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology*, 132, 2557–2576.
58. Parkin, D. M., Bray, F., Ferlay, J. and Pisani, P. (2005) Global cancer statistics, 2002. *CA Cancer J. Clin.*, 55, 74–108.
59. Monga, S. and Cagle, P. (2010) *Molecular Pathology of Liver Diseases*. Springer.
60. Michalopoulos, G. (2011) Liver regeneration. In 机构Monga, S. P. S. (ed.), *Molecular Pathology of Liver Diseases*. Springer US, 5, 261–278.
61. Higgins, G. M. (1931) Experimental pathology of the liver: I. Restoration of the liver of the white rat following surgical removal. *Arch. Pathol. (Chic)*, 12, 186–202.
62. Grisham, J. W. (1962) A morphologic study of deoxyribonucleic acid synthesis and cell proliferation in regenerating rat liver; autoradiography with thymidine-H3. *Cancer Res.*, 22, 842–849.
63. Wang, H., Iakova, P., Wilde, M., Welm, A., Goode, T., Roesler, W. J. and Timchenko, N. A. (2001) C/EBP α arrests cell proliferation through direct inhibition of Cdk2 and Cdk4. *Mol. Cell*, 8, 817–828.
64. Ilyin, G. P., Glaise, D., Gilot, D., Baffet, G. and Guguen-Guillouzo, C. (2003) Regulation and role of p21 and p27 cyclin-dependent kinase inhibitors during hepatocyte differentiation and growth. *Am. J. Physiol. Gastrointest. Liver Physiol.*, 285, G115–G127.
65. Papakyriakou, P., Tzardi, M., Valatas, V., Kanavaros, P., Karydi, E., Notas, G., Xidakis, C. and Kouroumalis, E. (2002) Apoptosis and apoptosis related proteins in chronic viral liver disease. *Apoptosis*, 7, 133–141.
66. Wilson, N. S., Dixit, V. and Ashkenazi, A. (2009) Death receptor signal transducers: nodes of coordination in immune signaling networks. *Nat. Immunol.*, 10, 348–355.
67. Green, D. R. and Kroemer, G. (2004) The pathophysiology of mitochondrial cell death. *Science*, 305, 626–629.
68. Adams, J. M. and Cory, S. (1998) The Bcl-2 protein family: arbiters of cell survival. *Science*, 281, 1322–1326.
69. Walensky, L. D., Pitter, K., Morash, J., Oh, K. J., Barbuto, S., Fisher, J., Smith, E., Verdine, G. L. and Korsmeyer, S. J. (2006) A stapled BID BH3 helix directly binds and activates BAX. *Mol. Cell*, 24, 199–210.
70. Jost, P. J., Grabow, S., Gray, D., McKenzie, M. D., Nachbur, U., Huang, D. C., Bouillet, P., Thomas, H. E., Borner, C., Silke, J., et al. (2009) XIAP discriminates between type I and type II FAS-induced apoptosis. *Nature*, 460, 1035–1039.
71. Wei, M. C., Zong, W. X., Cheng, E. H., Lindsten, T., Panoutsakopoulou, V., Ross, A. J., Roth, K. A., MacGregor, G. R., Thompson, C. B. and Korsmeyer, S. J. (2001) Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. *Science*, 292, 727–730.
72. Du, C., Fang, M., Li, Y., Li, L. and Wang, X. (2000) Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell*, 102, 33–42.
73. Zhao, Z., Yang, P., Eckert, R. L. and Reece, E. A. (2009) Caspase-8: a key role in the pathogenesis of diabetic embryopathy. *Birth Defects Res. B Dev. Reprod. Toxicol.*, 86, 72–77.
74. Trautwein, C., Böker, K. and Manns, M. P. (1994) Hepatocyte and immune system: acute phase reaction as a contribution to early defence mechanisms. *Gut*, 35, 1163–1166.
75. Diehl, A. M. (2000) Cytokine regulation of liver injury and repair. *Immunol. Rev.*, 174, 160–171.
76. Ben-Neriah, Y. and Karin, M. (2011) Inflammation meets cancer, with NF- κ B as the matchmaker. *Nat. Immunol.*, 12, 715–723.
77. Schwabe, R. F. and Brenner, D. A. (2006) Mechanisms of Liver Injury. I. TNF- α -induced liver injury: role of IKK, JNK, and ROS pathways. *Am. J. Physiol. Gastrointest. Liver Physiol.*, 290, G583–G589.
78. Sakurai, T., He, G., Matsuzawa, A., Yu, G.-Y., Maeda, S., Hardiman, G. and Karin, M. (2008) Hepatocyte necrosis induced by oxidative stress and IL-1 α release mediate carcinogen-induced compensatory proliferation and liver tumorigenesis. *Cancer Cell*, 14, 156–165.
79. Bromberg, J. and Wang, T. C. (2009) Inflammation and cancer: IL-6 and STAT3 complete the link. *Cancer Cell*, 15, 79–80.
80. Louis, H., Le Moine, O., Goldman, M. and Devière, J. (2003) Modulation of liver injury by interleukin-10. *Acta Gastroenterol. Belg.*, 66, 7–14.
81. Cairns, R. A., Harris, I. S. and Mak, T. W. (2011) Regulation of cancer cell metabolism. *Nat. Rev. Cancer*, 11, 85–95.
82. Vander Heiden, M. G., Cantley, L. C. and Thompson, C. B. (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science*, 324, 1029–1033.
83. Semenza, G. (2010) HIF-1: upstream and downstream of cancer metabolism. *Curr. Opin. Genet. Dev.*, 20, 51–56.
84. Lau, C. K., Yang, Z. F., Ho, D. W., Ng, M. N., Yeoh, G. C., Poon, R. T. and Fan, S. T. (2009) An Akt/hypoxia-inducible factor-1 α /platelet-derived growth factor-BB autocrine loop mediates hypoxia-induced chemoresistance in liver cancer cells and tumorigenic hepatic progenitor cells. *Clin. Cancer Res.*, 15, 3462–3471.
85. Adams, R. H. and Alitalo, K. (2007) Molecular regulation of angiogenesis and lymphangiogenesis. *Nat. Rev. Mol. Cell Biol.*, 8, 464–478.
86. Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., Fahrigh, M., Vandenhoec, A., Harpal, K., Eberhardt, C., et al. (1996) Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature*, 380, 435–439.
87. Massagué, J. (1998) TGF- β signal transduction. *Annu. Rev. Biochem.*, 67, 753–791.
88. Lu, Z., Ghosh, S., Wang, Z. and Hunter, T. (2003) Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of β -catenin, and enhanced tumor cell invasion. *Cancer Cell*, 4, 499–515.
89. Odom, D. T., Dowell, R. D., Jacobsen, E. S., Nekludova, L., Rolfe, P. A., Danford, T. W., Gifford, D. K., Fraenkel, E., Bell, G. I. and Young, R. A. (2006) Core transcriptional regulatory circuitry in human hepatocytes. *Mol. Syst. Biol.*, 2, 2006.0017.
90. Lu, T. T., Makishima, M., Repa, J. J., Schoonjans, K., Kerr, T. A., Auwerx, J. and Mangelsdorf, D. J. (2000) Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol. Cell*, 6, 507–515.
91. Vellela, M. and Qian, H. (2009) Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited. *J. R. Soc. Interface*, 6, 925–940.
92. Rosenfeld, S. (2011) Mathematical descriptions of biochemical networks: stability, stochasticity, evolution. *Prog. Biophys. Mol. Biol.*, 106, 400–409.
93. Villareal, C., Padilla-Longoria, P., and Alvarez-Buylla, E. R. (2012) General theory of genotype to phenotype mapping: derivation of epigenetic landscapes from N-node complex gene regulatory networks. *Phys. Rev. Lett.*, 109, 118102.

94. Xu, X. R., Huang, J., Xu, Z. G., Qian, B. Z., Zhu, Z. D., Yan, Q., Cai, T., Zhang, X., Xiao, H. S., Qu, J., et al. (2001) Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. *Proc. Natl. Acad. Sci. USA*, 98, 15089–15094.
95. Yang, W., Yan, H. X., Chen, L., Liu, Q., He, Y. Q., Yu, L. X., Zhang, S. H., Huang, D. D., Tang, L., Kong, X. N., et al. (2008) Wnt/ β -catenin signaling contributes to activation of normal and tumorigenic liver progenitor cells. *Cancer Res.*, 68, 4287–4295.
96. Holmes, P. (2005) Ninety plus thirty years of nonlinear dynamics: less is more and more is different. *Int. J. Bifurcat. Chaos*, 15, 2703–2716.
97. Ao, P. (2007) Darwinian dynamics implies developmental ascendancy. *Biol. Theory*, 2, 113–115.
98. Ao, P. (2007) Orders of magnitude change in phenotype rate caused by mutation. *Cell. Oncol.*, 29, 67–69.
99. Lee, L., Yin, L., Zhu, X. and Ao, P. (2007) Generic enzymatic rate equation under living conditions. *J. Biol. Syst.*, 15, 495–514.
100. Ao, P., Lee, L. W., Lidstrom, M. E., Yin, L. and Zhu, X. (2008) Towards kinetic modeling of global metabolic networks: methylotrophic bacterium *extorquens* AM1 growth as validation. *Chin. J. Biotechnol.*, 24, 980–994.
101. Liu, C., Fan, D., Shi, Y. and Zhou, Q. (2012) A glimpse of enzymology within the idea of systems. *Sci. China Life Sci.*, 55, 826–833.
102. Huang, S. (2011) On the intrinsic inevitability of cancer: from foetal to fatal attraction. *Semin. Cancer Biol.*, 21, 183–199.
103. Zhou, J. X., Aliyu, M. D., Aurell, E. and Huang, S. (2012) Quasi-potential landscape in complex multi-stable systems. *J. R. Soc. Interface*, 9, 3539–3553.
104. Baverstock, K. and Karotki, A. V. (2011) Towards a unifying theory of late stochastic effects of ionizing radiation. *Mutat. Res.*, 718, 1–9.
105. Davies, P. C. W. and Lineweaver, C. H. (2011) Cancer tumors as Metazoa 1.0: tapping genes of ancient ancestors. *Phys. Biol.*, 8, 015001.
106. Vincent, M. D. (2011) Cancer: beyond speciation. *Adv. Cancer Res.*, 112, 283–350.
107. Chaffer, C. L., Brueckmann, I., Scheel, C., Kaestli, A. J., Wiggins, P. A., Rodrigues, L. O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., et al. (2011) Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc. Natl. Acad. Sci. USA*, 108, 7950–7955.
108. Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D., Tao, K., Kuperwasser, C. and Lander, E. S. (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146, 633–644.
109. Heng, H. H. Q., Liu, G., Stevens, J. B., Bremer, S. W., Ye, K. J., Abdallah, B. Y., Horne, S. D. and Ye, C. J. (2011) Decoding the genome beyond sequencing: the new phase of genomic research. *Genomics*, 98, 242–252.
110. Gerstein, M. B., Kundaje, A., Hariharan, M., Landt, S. G., Yan, K.-K., Cheng, C., Mu, X. J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012) Architecture of the human regulatory network derived from ENCODE data. *Nature*, 489, 91–100.
111. Yamanaka, S. (2009) Elite and stochastic models for induced pluripotent stem cell generation. *Nature*, 460, 49–52.