Interdiscip Sci Comput Life Sci (2010) 2: 140–144
DOI: 10.1007/s12539-010-0072-3

Towards Predictive Stochastic Dynamical Modeling of Cancer Genesis and Progression

P. AO[1,2*], D. GALAS[3], L. HOOD[3], L. YIN[4], X.M. ZHU[5]
[1](Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine of Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China)
[2](Departments of Mechanical Engineering and Physics, University of Washington, Seattle, WA 98195, USA)
[3](Institute for Systems Biology, 1441 N. 34 St., Seattle, WA 98103, USA)
[4](School of Physics, Peking University, Beijing 100871, China)
[5](GenMath, Corp. 5525 27th Ave.N.E., Seattle, WA 98105, USA)

Received 21 August 2009 / Revised 19 October 2009 / Accepted 23 October 2009

Abstract: Based on an innovative endogenous network hypothesis on cancer genesis and progression we have been working towards a quantitative cancer theory along the systems biology perspective. Here we give a brief report on our progress and illustrate that combing ideas from evolutionary and molecular biology, mathematics, engineering, and physics, such quantitative approach is feasible.
Key words: cancer, stochastic processes, systems biology, endogenous network, functional landscape.

1 Introduction

We have been advancing an evolutionary and stochastic dynamics formulation of carcinogenesis. The novel biological hypothesis behind such formulation has been stated in our previous publication (Ao *et al.*, 2008): Cancer as an intrinsic robust state of the endogenous network not optimized for the interest of whole organism. More explicitly, the molecular and cellular agents, such as oncogenes and suppressor genes, and related growth factors, hormones, cytokines, etc, form a nonlinear, stochastic, and collective dynamical network, the endogenous molecular – cellular network. This endogenous network may be specified by the expression or activity levels of a minimum set of endogenous agents, resulting in a high dimensional stochastic dynamical system. The nonlinear dynamical interactions among the endogenous agents can generate many locally stable states with obvious or non-obvious biological functions. The endogenous network may stay in any of such stable state for a considerably long time. In this manner the endogenous network is able to autonomously decide its operational functioning state. Some states may be normal, such as cell growth, apoptosis, arresting, etc. Others may be abnormal, such as growth with elevated immune response and high energy consumption, likely the signature of cancer, or of still useful functions to deal with occasional stressful situations. The stochasticity may accidentally cause a transition from one stable state to another. If with a given condition the endogenous network is in a state not optimized for the interest of whole organism, the organism is 'sick', though this state might be 'normal' under other conditions. Through the identifying agents of this endogenous network, the delineating of its wiring rules among endogenous agents, and the elucidating its global dynamical properties, a systems understanding of both normal and abnormal behaviors on how a tissue functions may be reached.

In this way, we envisage that the oncogenes and other molecular and cellular agents first form pathways and modules. The pathways and modules then cross talk to each other to form the endogenous network. Such a hierarchical structure is similar to the modular organization principle. The essential pathways and modules (Greaves, 2001; Weinberg, 2007) are, for example, cell cycle pathway, Myc-p53 pathway, immune response, Ras-MAPK pathway, invasion and metastasis, PTEN-Akt pathway, growth factors and their receptors, hormones and their receptors, metabolism, and apoptosis. In short, we believe that there is at least one more important, and nearly autonomous, layer of mechanism between genetic (genomics, etc) and environmental factors. It is consistent with the systems biology perspective (Auffray *et al.*, 2009) but differs in focus from the current mainstream genomic centric view

*Corresponding author.
E-mail: aoping@u.washington.edu, aoping@sjtu.edu.cn.

of "cancer as disease of genome" (Heng *et al.*, 2009). In the following sections the feasible key steps leading to such quantitative formulation are explained: determining endogenous network, establishing mathematical model, analyzing and predictions.

2 Determining key players of core endogenous network

Granted the validity of such hypothesis, the crucial question is on its quantitative realization by a mathematical model. In the rest of this extended abstract a procedure to implement such hypothesis is summarized: sorting through experimental data, establishing a set of stochastic differential equations, and generating testable predictions.

Even for each pathway or module its mathematical description can be very complicated. A full quantitative account has indeed been very difficult and subjected to extensive current investigation. We thus further simplify each pathway or module to its most essential endogenous agents, whose number may typically be 3-5. For example, for the cell cycle pathway 5 endogenous agents, E2F, Myc, CyclinE/Cdk2, CyclinD/Cdk4,6, and Rb reasonably describe the main features of this pathway. Therefore, for above mentioned essential pathways and modules, we may expect about 40 endogenous agents, which, for a mechanistic modeling point of view, still need a large number with numerous parameters to be determined experimentally. A normalized procedure to reduce such dependence will be explicitly stated below.

3 From experimental data to mathematical model

The next step is how to construct a mathematical description corresponding to those endogenous agents based on experimental input. Evidently, there are at least two very different kinds of experimental data: those of high throughput, which may determine the whole structure of the mathematical equations governing the dynamical evolution of endogenous agents at once, and those of local and pair-wise classical molecular and biochemical experiments. We believe that the current high throughput data are not ideal to construct a reliable dynamical model because they are too noisy. Even if this problem would be overcome soon, we believe the amount of data alone in the foreseeable future would be not enough to determine the nonlinear and stochastic network with multiple stable states. Instead, we make an effort to construct the whole network solely based on local and targeted biochemical and molecular biology experimental data accumulated during past 6 decades, and will use the high throughput data as experimental check for theoretical predictions. In order to obtain a most consensus among experimental input, we further limit the experimental data to the most essential types which can be rather reliably established in a given experiment: those of activation and inhibition. Indeed, based on a reading of the literature we are able to construct the corresponding endogenous network (Zhu *et al.*, 2010). A sampling of such collection is shown in Table 1, and its graphical representation is in Fig. 1.

Functional diagram as depicted in Fig. 1 can be turned into a set of stochastic dynamics equations of the form,

$$\mathrm{d}\boldsymbol{x}/\mathrm{d}t = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{x}/\tau + \xi(\boldsymbol{x}, t). \tag{1}$$

For example, the deterministic force $\boldsymbol{f}$ can be obtained by modeling the activation as sigmoidal or S-shaped typical in engineering, threshold functions varying between 0 and 1 (Zhu *et al.*, 2010), a procedure similar to those in fuzzy cognitive maps (Kosco, 1997; Miao *et al.*, 2001; Weinreb *et al.*, 2006). Our own convenient choice is, for activation,

$$f_A(y) = ay^n/(1 + ay^n), \tag{2}$$

Table 1 A sampling directed interactions among endogenous agents from targeted pair-wise experimental data. Prostate cancer is the focus, represented by the presence of androgen receptor (AR) and the insulin-like growth factor receptor (IGF-1R).

Endogenous agent	Activated by	Inhibited by
Cyclin E/Cdk2	Myc, E2F	p21, p27
Myc	pRb(+), E2F, Akt, MAPK	P53, TGF-β
p53	Myc, PTEN	Akt
Cytochrome c	Caspase 3, Bad, Bax	Bcl-2, Bcl-xL
Bad		p21, Akt, MAPK
Bax	Myc, p53, Bim	
Ras	VEGF, IL-6, Integrin, Androgen R	
Akt	NF-κB, HIF, Ras, IGF-1R	PTEN
VEGF	Akt, HIF, COX-2, IL-6, Androgen R	
IGF-1R	Androgen R	p53
Androgen R(AR)	EGF, IL-6	PTEN
Integrin	EGF, TNF-α, VEGF	
E-cadherin		TNF-α, EGF, HIF, TGF-β
HIF	Akt	p53
TNF-α	NF-κB	IL-10
IL-10	TNF-α, Fas	IL-10
COX-2	NF-κB, MAPK	

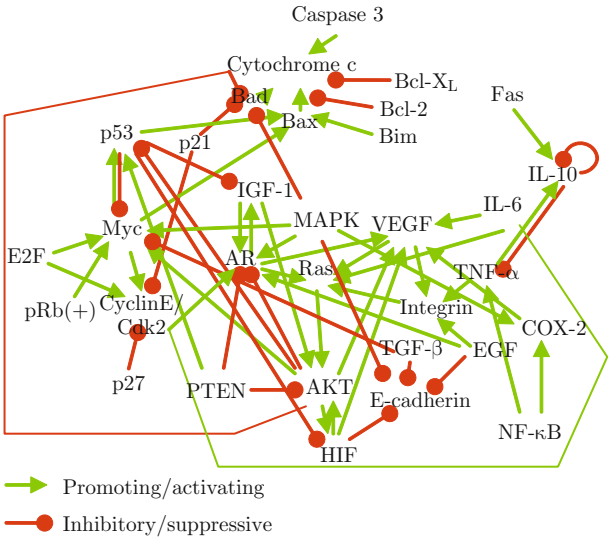

Fig. 1 The directed graph representation of endogenous interaction corresponding to Table 1. Biological experimental data are typically in this way, too.

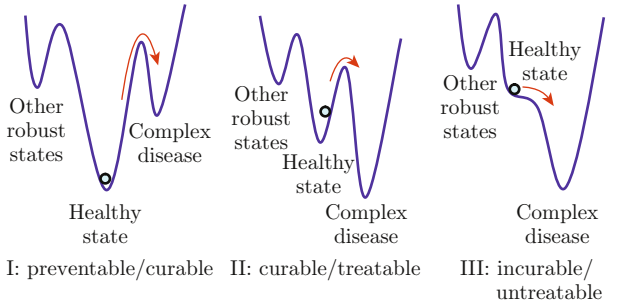

Fig. 2 Three typical situations of the functional landscape, modified from Fig. 2 in (Ao *et al.*, 2008). (a) The healthy state is a globally stable under normal conditions; (b) due to genetic and epidemiologic influence on the endogenous network, tumor or cancer states may become more stable than healthy 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. The vertical scale illustrates the relative stability of robust states, healthy, tumor and others, in the multiple dimensional state space, along an optimal trajectory passing through a "mountain pass". The ball indicates the state in the functional landscape.

and for inhibition,

$$f_I(y) = 1 - f_A(y) = 1/(1 + ay^n), \qquad (3)$$

with numerical values are chosen to be $a = 10$ and $n = 3$. This implies that the variables, largely the activity levels of endogenous agents, are normalized to minimize the demand in input parameters: the strength of interaction is varying between 0 and 1, as well as those dynamical variations. The degradation time τ is set to be constant to begin with. The stochastic term can be modeled by a diffusion matrix D dictated by biological considerations. For simplicity we will choose D to be diagonal.

4 Analyzing Stochastic Dynamical Model

Once such a set of stochastic differential equations is obtained (Zhu *et al.*, 2010), an adaptive landscape quantified by a potential function (a Lyapunov function) can be constructed by a procedure similar to that of recent study on phage lambda genetic switch and others (Zhu *et al.*, 2004; Wang *et al.*, 2006; Zhang *et al.*, 2006; Morelli *et al.*, 2008; Cao and Liang, 2008). The rudimental mathematical elaborations have been done recently (Ao, 2004; Kwon *et al.*, 2005; Yin and Ao, 2006; Ao *et al.*, 2007), involving feedbacks, multiple stability and other nonlinear stochastic dynamical features (Zhu *et al.*, 2004; Zhang *et al.*, 2006; Karmakar and Bose, 2007; Qian *et al.*, 2009). The dimension corresponding to the wiring diagram similar to Fig. 1 for finally obtained network is 37, which is not easy to be visualized. A schematic representation is given in Fig. 2.

Given the existence of multiple stable states, we may postulate that there are some states may correspond to healthy states under normal conditions, some to deal with rare stressful situations, and a few others would be the "disease" states or states as illustrated in Fig. 2. Table 2 lists the positions of such land markers in the functional landscape, minima, mountain pass, computed according the procedure outlined above, for endogenous agents in Table 1. For our minimum model, the values are given in Table 2. Here it may be worthwhile to mention that the landscape idea has already explored early on in biology in other contexts: for example, the adaptive landscape in population genetics and the developmental landscape in developmental biology. Such idea has already been considered quantitatively in biology (Onuchic *et al.*, 1999; Bar-Yam *et al.*, 2009; Ao, 2009).

We emphasize that there should be a minimum set of endogenous molecular and cellular agents in order to give a comprehensive description of the endogenous network to describe the complexity of cancer phenomena. In our current minimum mathematical model (Zhu *et al.*, 2010), the number of endogenous agents is 37. Though still incomplete, we believe it represent a part of core endogenous agents in prostate cancer, and perhaps shared by most, if not all, other cancers.

Finally, we discuss two experimental implications in the present endogenous network cancer theory. Both Fig. 2 and Table 2 imply that there is a finite probability, though very small, for cancer to occur sponta-

neously, even without mutations and other harmful effects. Such suggestion implies that cancer is a property of the endogenous network, though not a welcome prediction. While this possibility may be difficult to check experimentally, its reverse effect, the spontaneous cancer regression is also implied. Careful examining experimental and clinical data indeed shows that such phenomena indeed exist (Dinulos *et al.*, 1997; Reynolds, 2002; Abedelrazeq, 2007). This may be regarded as a zero-one type validation of present nonlinear stochastic modeling.

Table 2 Three functional states in the functional landscape: normal state, tumor-like state, and saddle point configuration connection them. Values are given in terms of their "maximum" activities, respectively. Though the construction of the present mathematical model is based on local and targeted experimental data, the predictions given in the form of this table are clearly testable by both targeted and high-through put experiments.

	Normal growth	"Mountain pass"	Tumor-like state
Cytochrome c	0.08	0.01	0
Myc	0.84	0.60	0.53
CyclinE/CDK2	0.92	0.80	0.84
p53	0.16	0.05	0.01
Bad	0.26	0.11	0.06
Bax	0.29	0.10	0.07
Akt	0.02	0.37	0.63
HIF	0.00	0.34	0.72
TNF-α	0.16	0.37	0.44
Ras	0.18	0.73	0.81
COX-2	0.26	0.66	0.74
VEGF	0.19	0.83	0.93
IL-10	0.04	0.28	0.34
Integrin	0.39	0.57	0.65
Androgen R	0.13	0.36	0.54
IGF-1R	0.02	0.31	0.61
E-Cadherin	0.36	0.15	0.07

Because of normalized nature of the minimum cancer model, any two measurements of a set of endogenous agents are directly testable. For example, for p53, from normal to tumor-like state, its activity level is predicted to decrease, as listed in Table 2. The opposite trend can be said for androgen receptor. If agreeing with further experimental observation, it is an additional validation for the current modeling. If a disagreement would occur, further research would be naturally suggested, experimentally and/or theoretically. Regardless of outcome, it would be exciting because we may finally on the way to have a quantitative, predictive and mechanistic cancer theory interacting directly with both genetic and epigenetic experiments.

5 Conclusions

Rooting firmly in biological observations, with adaptive landscape concept from evolutionary biology, typical engineering modeling techniques, and the recent progresses in stochastic processes, we believe a feasible way for quantitative modeling of cancer genesis and progression is possible and the resulting predictions are experimentally testable. The preliminary results reported here support such vision.

Acknowledgments This work was supported in part by a grant from USA National Institutes of Health (No. K25-HG002894), by China 985 Project via Shanghai Jiao Tong University, and by 973 projects from Chinese Natural Science Foundation (No. 2007CB914700 and No. 2010CB529200).

References

[1] Abedelrazeq, A.S. 2007. Spontaneous regression of colorectal cancer: A review of cases from 1900 to 2005. Int J Colorectal Disease 22, 727–736.

[2] Ao, P. 2004. Potential in stochastic differential equations: novel construction. J Phys A37, L25–L30.

[3] Ao, P. 2009. Global view of bionetwork dynamics: Adaptive landscape. J Genetics and Genomics 36, 1–11.

[4] Ao, P., Kwon, C., Qian, H. 2007. On the existence of potential landscape in the evolution of complex systems. Complexity 12, 19–27.

[5] Ao, P., Galas, D., Hood, L., Zhu, X.-M. 2008. Cancer as Robust Intrinsic State of Endogenous Molecular-Cellular Network Shaped by Evolution. Medical Hypotheses 70, 678–684.

[6] Auffray, C., Chen, Z., Hood, L. 2009. Systems medicine: The future of medical genomics and healthcare. Genome Medicine 1, article No. 2.

[7] Bar-Yam, Y., Harmon, D., de Bivort, B. 2009. Systems biology: Attractors and democratic dynamics. Science 323, 1016–1017.

[8] Cao, Y.F., Liang, J. 2008. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. BMC Syst Biol 2, article No. 30.

[9] Dinulos, J.G., Hawkins, D.S., Clark, B.S., Francis, J.S. 1997. Spontaneous remission of congenital leukemia. J Pediatrics 131, 300–303.

[10] Greaves, M. 2001. Cancer: The Evolutionary Legacy. Oxford University Press, Oxford.

[11] Heng, H.H.Q., Bremer, S.W., Stevens, J.B., Ye, K.J., Liu, G., Ye, C.J. 2009. Genetic and epigenetic heterogeneity in cancer: A genome-centric perspective. J Cellular Physiology 220, 538–547.

[12] Karmakar, R., Bose, I. 2007. Positive feedback, stochasticity and genetic competence. Physical Biology 4, 29–37.

[13] Kosko, B. 1997. Fuzzy Engineering. Prentice-Hall, Upper Saddle River.

[14] Kwon, C., Ao, P., Thouless, D.J. 2005. Structure of stochastic dynamics near fixed points. Proc Natl Acad Sci (USA) 102, 13029–13033.

[15] Miao, Y., Liu, Z.Q., Siew, C.K., Miao, C.Y. 2001. Dynamical cognitive network - an extension of fuzzy cognitive map. IEEE T Fuzzy Systems 9, 760–770.

[16] Morelli, M.J., Tanase-Nicola, S., Allen, R.J., ten Wolde, P.R. 2008. Reaction coordinates for the flipping of genetic switches. Biophys J 94, 3413–3423.

[17] Onuchic, J.N., Wang, J., Wolynes, P.G. 1999. Analyzing single molecule trajectories on complex energy landscapes using replica correlation functions. Chem Phys 247, 175–184.

[18] Qian, H., Shi, P.Z., Xing, J.H. 2009. Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. Phys Chem Chem Phys 11, 4861–4870.

[19] Reynolds, C.P. 2002. Ras and Seppuku in neuroibalstorma. J Natl Cancer Inst 94, 319–321.

[20] Wang, J., Huang, B., Xia, X.F., Sun, Z.R. 2006. Funneled landscape leads to robustness of cell networks: Yeast cell cycle. PLoS Comput Biol 2, 1385–1394.

[21] Weinberg, R.A. 2007. The Biology of Cancer. Taylor and Francis, New York.

[22] Weinreb, G.E., Elston, T.C., Jacobson, K. 2006. Causal mapping as a tool to mechanistically interpret phenomena in cell motility: Application to cortical oscillations in spreading cells. Cell Mot Cytosk 63, 523–532.

[23] Yin, L., Ao, P. 2006. Existence and construction of dynamical potential in nonequilibrium processes without detailed balance. J Phy A39, 8593–8601.

[24] Zhang, Y.P., Qian, M.P., Ouyang, Q., Deng, M.H., Li, F.T., Tang, C. 2006. Stochastic model of yeast cell-cycle network. Physica D219, 35–39.

[25] Zhu, X.M., Yin, L., Hood, L., Ao, P. 2004. Calculating robustness of epigenetic states in phage lambda life cycle. Funct Integr Genom 4, 188–195.

[26] Zhu, X.M., Hood, L., Galas, D., Ao, P. 2010. Construction of a molecular-cellular endogenous network prostate cancer model. To be published.