ELSEVIER

Systems mapping of genes controlling chemotherapeutic drug efficiency for cancer stem cells

Weimiao Wu[1,3], Sisi Feng[1,3], Yaqun Wang[2], Ningtao Wang[2], Han Hao[2] and Rongling Wu[1,2]

[1] Center for Computational Biology, Beijing Forestry University, Beijing 100083, China
[2] Center for Statistical Genetics, The Pennsylvania State University, Hershey, PA 17033, USA

Cancer can be controlled effectively by using chemotherapeutic drugs to inhibit cancer stem cells, but there is considerable inter-patient variability regarding how these cells respond to drug intervention. Here, we describe a statistical framework for mapping genes that control tumor responses to chemotherapeutic drugs as well as the efficacy of treatments in arresting tumor growth. The framework integrates the mathematical aspects of the cancer stem cell hypothesis into genetic association studies, equipped with a capacity to quantify the magnitude and pattern of genetic effects on the kinetic decline of cancer stem cells in response to therapy. By quantifying how specific genes and their interactions govern drug response, the model provides essential information to tailor personalized drugs for individual patients.

Introduction

The discovery of cancer stem cells in malignancies of hematopoietic origin and in some solid tumors has changed our vision of the biological processes involved in carcinogenesis and chemotherapeutic practices. Just as normal cells are maintained by self-renewing stem cells, malignant tumors are produced through the mutations of stem cells and their subsequent proliferation [1–5]. For example, leukemia is believed to arise from a stem cell that gives rise to a large population of clones that proliferate into malignancies. Therefore, by developing specific therapies targeted at cancer stem cells, malignant tumors can be controlled and prevented and, finally, eradicated through blocking the recurrence of cancer cells [6–8].

To make it effective to treat cancer based on the cancer stem cell hypothesis, two essential questions need to be addressed. First, how can we distinguish cancer stem cells from cancer non-stem cells in terms of their origin, property and function [3]? Second, through which mechanisms do cancer stem cells respond to chemotherapeutic drugs [9]? The availability of genetic, genomic and proteomic expression data provides an unprecedented opportunity to detect and define expression patterns of cancer stem cells and predict the clinical outcome of patients who receive a particular drug therapy [10,11]. By contrast, mathematical modeling has exemplified increasing vitality to uncover and explain many still unknown aspects of cell behavior, tissue function and network organization [12–14]. More recently, an avalanching interest has emerged in applying differential equations to quantify the proliferation and differentiation of normal stem cells and cancer stem cells and detect the differences of these two types of cells [14,15].

Wang *et al.* [16] have for the first time integrated expression data with mathematical models to identify genes and proteins or their expression patterns that are linked with the formation, proliferation and programming of cancer stem cells. This integration can potentially lead to understanding of the genetic and molecular mechanisms of carcinogenesis and the complexity of its progress and dynamics. Here, we argue that the model described by Wang *et al.* can be reformed to map genes that control the response of cancer stem cells to chemotherapeutic drugs. The new model is constructed on a mapping approach – systems mapping – by incorporating chemotherapeutic drug efficacy that describes the kinetic reduction of abnormal cell populations in response to therapy [17,18]. It provides an analytical tool to test the temporal effects of genes on drug response and can be used to assess the efficacy of treatments in arresting tumor growth.

Corresponding author: Wu, R. (rwu@bjfu.edu.cn, RWu@phs.psu.edu, RWU@hes.hmc.psu.edu)
[3] These authors contributed to this work equally.

1359-6446/06/$ - see front matter http://dx.doi.org/10.1016/j.drudis.2013.12.010

Systems mapping

The formation of phenotypic traits is one of the most complex processes in nature. Traditional approaches for genetic dissection of complex traits is to associate genetic variation with phenotypic variation in a trait measured at a particular time point. These approaches have proven to be instrumental for identifying quantitative trait loci (QTLs), but they have not considered the complexity and dynamics of phenotypic formation. A new computational model, known as systems mapping, has been recently developed to enhance the biological relevance of QTL mapping [19]. Systems mapping views a complex phenotype as a dynamic system, dissects it into its underlying interconnected components and organizes and connects different components through mathematical equations in biological laws [20,21]. By mapping specific genes that govern each component and its mutual connections with other components, this model has a capacity to help understand not only the behavior of the components but also how these components act together to form the behavior of the whole. As a bottom–top model, a systems approach can identify specific QTLs that govern the developmental interactions of different components that lead to the function and behavior of the system. By estimating and testing mathematical parameters that specify the system, systems mapping enables the prediction or alteration of the physiological status of a phenotype based on the underlying genetic control mechanisms.

Genetic mapping of complex traits is constructed by a mixture model in which different mixture components are presented by QTL genotypes that are segregating among individuals in a mapping population [22,23]. Because QTLs cannot be observed directly, the proportions of mixture components are specified by conditional probabilities of QTL genotypes given observable marker genotypes. Phenotypic values of individuals carrying a particular QTL genotype are assumed to follow a distribution function, such as the normal distribution, characterized by expected mean (denoted as the genotypic mean) and variance. Systems mapping embeds a system of ordinary differential equations (ODEs) into a genetic mapping setting containing dynamic measures of phenotypic values. Unlike traditional approaches that estimate genetic effects directly, systems mapping specifies and estimates genotype-specific mean vectors by ODE parameters and a covariance matrix by a parsimonious statistical model. Mathematical tools, like the fourth-order Runge–Kutta algorithm, have been incorporated to estimate ODE parameters for individual QTL genotypes contained within a mixture-model framework [24,25]. Structural approaches have been used to model the covariance matrix for longitudinal traits, which include (i) parametric stationary [26], (ii) parametric nonstationary [26,27], (iii) nonparametric [28] and (iv) semiparametric models [28]. Each of these approaches has advantages and disadvantages regarding computing efficiency, flexibility and power.

Mapping QTLs for chemotherapeutic efficiency

Mathematical models for efficacy of a chemotherapeutic drug

Based on the cancer stem cell hypothesis [29], Ganguly and Puri [12] described a basic model for healthy and cancer stem cell pathways (Fig. 1). Normal stem cells (SC) are of two types, one that performs self-renewal with a probability, P_{SC}, and the other

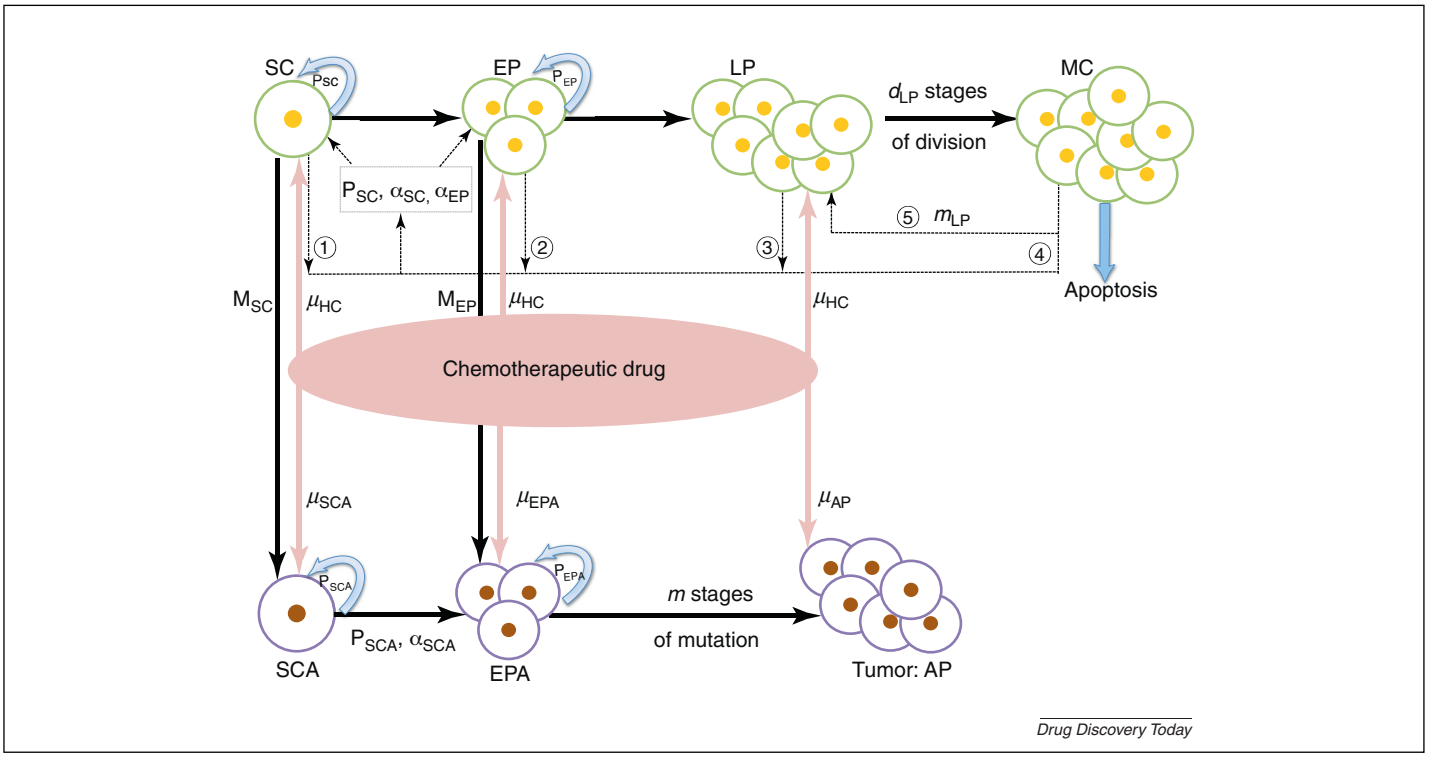

FIGURE 1

Cancer stem cell model showing the cell signaling pathway and the action of a chemotherapeutic drug. Arrowed dotted lines represent a direction of regulatory feedback signals of one process to others, numbered from (1) to (5). SC, normal stem cell; EP, early progenitor cells; LP, late progenitor cell; MC, mature cell; SCA, abnormal stem cell; EPA, abnormal early progenitor cell; AP, abnormal progeny. For the definitions of parameter symbols, see the text. Adapted, with permission, from Ganguly and Puri [30].

that differentiates into early progenitor cells (EP). The self-renewal probability of SC depends on itself as well as EP and late progenitor cell (LP) populations. Part of the EP self-renews a number of times. The LP undertakes d_{LP} stages of cell division and produces mature cells (MC). The number of cell division stages is controlled by a cell division regulatory feedback signal to sustain a steady population of MC. The MC also provides a feedback to the SC to influence their mitotic fraction and self-renewal rate.

Stem cells and progenitor cells can produce oncogenic mutations during DNA transcription, with the probabilities denoted as M_{SC} and M_{EP}, respectively. The mutation of SC forms abnormal stem cells (SCA). The formation of abnormal early progenitor cells (EPA) can be caused by either the differentiation of SCA or mutation in EP, or both. Through subsequent differentiation, EPA form abnormal progeny (AP), which is the precursor of malignancy. Each cell type is treated as a separate compartment with a particular rate of cell population growth. MC and AP can undertake apoptosis.

When a chemotherapeutic drug is used, its effect on the cell population size is reflected by its cell-kill rate μ [30] (Fig. 1). Each cell type is assumed to have its own carrying capacity Θ [31]. Thus, the dynamics of cell proliferation after drug therapy is expressed as:

$$\frac{dN}{dt} = \omega N\left(1 - \frac{N}{\Theta}\right) - \mu AN \quad (1)$$

where N denotes the cell population, $\omega = (\alpha/\tau)\ln(2)$ is the cell division rate, with α being the proliferative fraction and τ being the cell cycle time, and A is the average drug concentration in a cellular matrix. The net rate of change of A is determined in Eq. (2) by:

$$\frac{dA}{dt} = a(t) - (\lambda - \gamma N)A \quad (2)$$

where $a(t)$ is a function of the drug infusion rate, λ is the natural drug decay caused by chemical decomposition of the drug or its transport and γN is the rate at which the drug becomes ineffective as a result of cell-kill rate [31]. Based on Eq. (1), Ganguly and Puri [30] provided a group of differential equations that describe the rate of change in a cell population.

Clinical design for systems mapping of chemotherapeutic efficiency

Chemotherapeutic drugs have been used to treat and control cancer, but there is tremendous interpersonal variability in drug response, implicating a genetic component involved in chemotherapeutic efficiency [32]. In a recent study of cancer gene identification, Fugger *et al.* [33] used RNA interference to find a gene, called FBH1, which is crucial for some chemotherapeutics to kill cancer cells. Through a systematic search of such genes and a feasible method to activate their expression, cancer treatment can be improved. Integrated with differential equations for the dynamics of cancer stem cells treated by a chemotherapeutic drug [30], we reform systems mapping to allow genes controlling the chemotherapeutic efficacy of cancer to be identified.

We design a clinical trial for n cancer patients each infused by a chemotherapeutic drug with efficacies toward different cell populations. Cell-kill rates for different healthy and abnormal stem cell compartments are assumed [34]. As depicted by Fig. 1, the healthy stem cell compartment includes stem cells N_{SC}, EP [consisting of k subcompartments each with a population N_{EPl} ($l = 1, \ldots, k$)], total efflux of differentiated EP N_{EP}^{out} that enter the LP compartment, LP N_{LP} and MC N_{MC}. N_{EP}^{out} contains an efflux of EP that are mutated. The abnormal stem cell compartment is composed of SCA N_{SCA}, EPA with k subcompartments N_{EPAl} ($l = 1, \ldots, k$), total efflux of EPA N_{EPA}^{out} and AP N_{AP}. Ganguly and Puri [30] constructed two groups of ODEs for dynamic changes of the healthy and abnormal cell compartments, respectively. These equations can be incorporated by cell-kill rates by a chemotherapeutic drug (Fig. 1).

All n patients are infused by the same drug in a continuous or periodic manner. The continuous infusion uses a constant drug dose over the entire period of chemotherapy, whereas periodic infusion is made every t hours in which instantaneous drug concentration decays with time. Given the same total amount of infused drug, period infusion is designed with two schedules: short cycle ($t = 100$ h) and long cycle ($t = 250$ h). The overall efficacy of drug is characterized by its effect on abnormal progeny size. During chemotherapy under the continuous and periodic infusion, different compartments of healthy cells and abnormal cells are screened at different time points. For these patients, genotypes at different single nucleotide polymorphisms (SNPs) throughout the genome are collected, providing a fuel of genome-wide association studies (GWAS) with systems mapping.

For different genotypes at a particular SNP, systems mapping is equipped to fit dynamic changes of cell populations in different compartments using a system of ODEs by Ganguly and Puri [30]. Fu *et al.* [24] implemented the fourth-order Runge–Kutta algorithm to estimate the ODE parameters numerically. By comparing differences in these parameters between genotypes, systems mapping can test how a gene affects cell dynamics after chemotherapeutic treatment. Specifically, systems mapping can be used to address the following questions of fundamental importance:

i. How does the gene affect each compartment of healthy cells from stem cells to mature cells through early and late progenitor cells (upper pathway of Fig. 1)? This is addressed by testing whether a set of ODE parameters characterizing a particular compartment differs among genotypes.
ii. How does the gene affect each compartment of abnormal cells from cancer stem cells to tumor cells through abnormal early progenitor cells (lower pathway of Fig. 1)? The testing procedure for this question is similar to (i).
iii. How does the gene control the efflux of differentiated early progenitor cells as well the amount of mutated early progenitor cells?
iv. How does the gene control the efflux of abnormal early progenitor cells?
v. How does the gene affect the mutation rates of stem cells and early progenitor cells, respectively?
vi. How does the gene affect the cell-kill rates of stem cells and early progenitors?
vii. How does the gene affect the cell-kill rates of cancer stem cells and abnormal early progenitors?
viii. How does the gene affect the cell-kill rates of tumors? This question is directly related to the genetic control of chemotherapeutic efficacy.

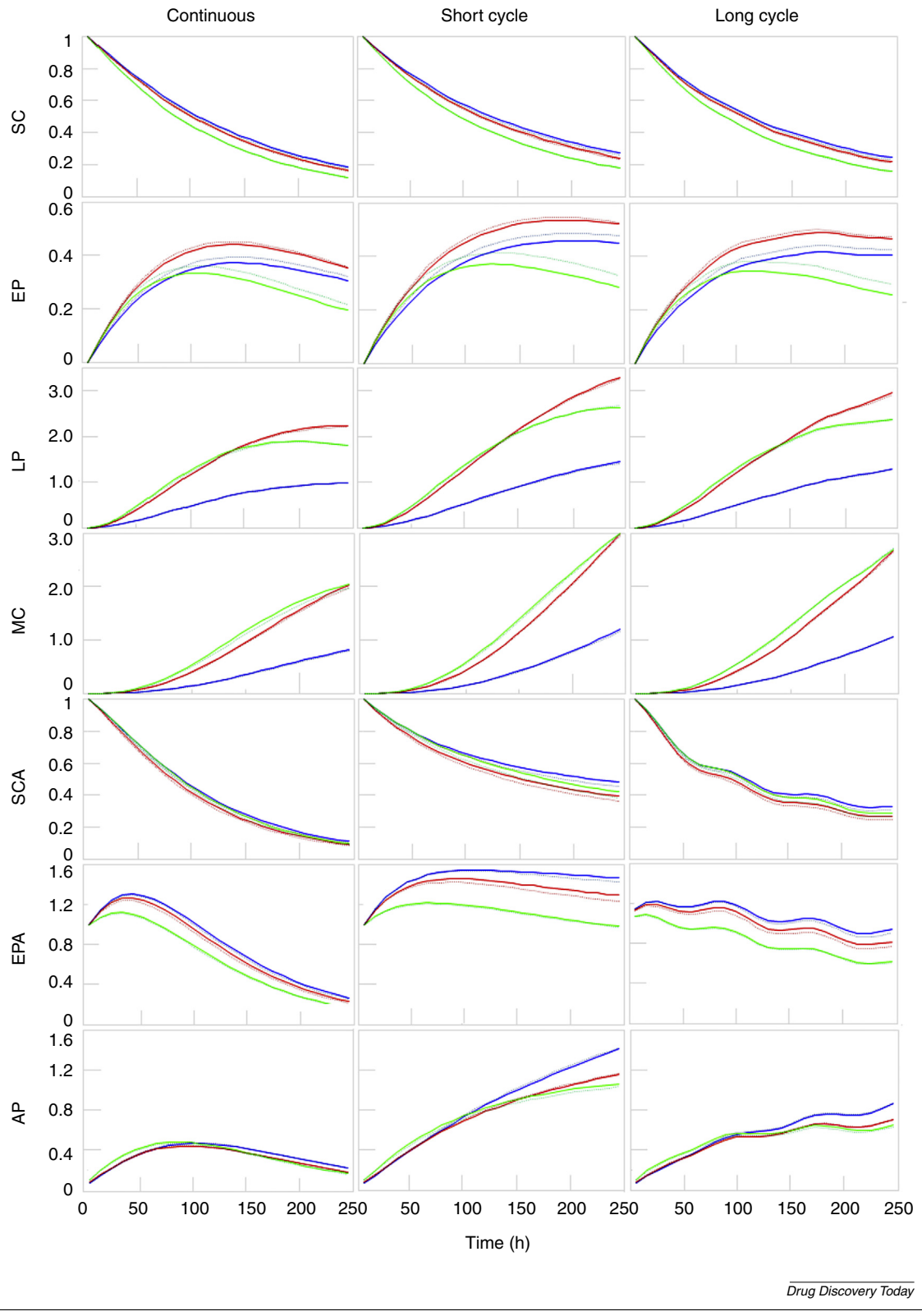

FIGURE 2

Genetic differences in the dynamic behavior of healthy stem cell and abnormal stem cell compartments as a complex system composed of seven biological variables (SC, EP, LP, MC, SCA, EPA and AP) of cell proliferation characterized by Ganguly and Puri's [30] ODEs. SC, normal stem cell; EP, early progenitor cells; LP, late progenitor cell; MC, mature cell; SCA, abnormal stem cell; EPA, abnormal early progenitor cell; AP, abnormal progeny. Each curve is denoted as one of three genotypes: *AA* (red), *Aa* (blue) and *aa* (green). Solid and dotted curves are true and estimated curves, respectively. The sample size and heritability used are assumed to be 400 and 0.1, respectively, for this simulation.

Questions (i), (iii) and (vi) are related to the genetic control of drug toxicity, whereas questions (ii), (iv), (vii) and (viii) are concerned with the genetic control of chemotherapeutic efficacy. Question (v) tests the genetic control regarding how healthy cells are mutated to be abnormal under a particular environmental condition.

Numerical simulation

We performed simulation studies to test how systems mapping functions to detect genes for chemotherapeutic efficacy in terms of the cancer stem cell hypothesis. Our simulation was designed according to three different scenarios of drug infusion: continuous, short periodic cycle (100 h) and long periodic cycle (250 h). In each scenario, different sample sizes (400 and 800) and different heritability levels (0.1 and 0.2) are assumed. In the simulated population, we assume that there is a segregating QTL with three genotypes *AA*, *Aa* and *aa* in frequencies of 0.35, 0.45 and 0.20. These three genotypes each bear different sets of mathematical parameters that define Ganguly and Puri's [30] ODEs for healthy and abnormal cell dynamics. The patients are treated with a chemotherapeutic drug under three different infusion scenarios. Phenotypic compartment cell data were simulated for seven components by assuming a heritability value for each component at a middle point. The time-dependent residual covariance matrix for each component variable is structured by a first-order autoregressive model. To facilitate computing, no residual correlations are assumed to exist among component variables, although the interdependence of different variables is reflected by ODE.

The model was used to analyze the simulated data. The model can precisely estimate allele frequencies and linkage disequilibrium between the marker and QTL (results not given), suggesting that an underlying QTL can be identified by a marker when a modest sample size (400–800) is used. For given sample sizes and heritabilities, the model can precisely estimate ODE parameters that specify the dynamic system of healthy and abnormal cells (Fig. 1) under three simulation scenarios (results not given). Figure 2 illustrates an example in which the estimated curves of different compartments (comprising healthy and abnormal cell populations) after the treatment of continuous and periodic infusion are compared with the respective true curves, assuming the sample size of 400 and heritability of 0.1. In general, the estimated and true curves are consistent, suggesting that systems mapping can serve as a powerful approach for identifying specific genes involved in cancer initiation and further estimating the temporal pattern of their control over cancer formation. There is not much difference in the accuracy of curve estimation under three scenarios of infusion, given a modest sample size and modest heritability.

By looking at results from different sample sizes and heritability levels, the overall conclusion is that increasing sample and heritability sizes can increase the precision of parameter estimation, but estimation precision can be improved better by increasing trait heritability through reducing phenotyping errors rather than by a simple increase of sample size. Simulation studies have also been performed to study the power of gene identification by systems mapping. It appears that the model indicates good power (>0.90) even when there is a modest sample size (400) and heritability (0.10). The model has an acceptable false-positive rate (<0.05).

Concluding remarks

Increasing evidence indicates that cancer arises from cancer stem cells mutated from normal stem cells [1–5]. Genes are thought to play an important part in initiating cancer stem cells and guiding their division and proliferation [35]. Given their dynamic properties, Wang *et al.* [16] integrated mathematical aspects of cancer stem cells [12] into systems mapping, a model designed to map dynamic genes [19,24,25], to identify the genetic control of cancer formation and progression. Because of our growing recognition of cancer stem cells, Wang *et al.*'s model could help to redesign current cancer-killing therapies by first identifying this type of cell from normal stem cells, followed by the eradication of the former through developing specified chemotherapies.

In this article, we put forward a theoretical framework by which to characterize specific QTLs involved in a pathway toward chemotherapeutic efficacy. This framework takes advantage of systems mapping by viewing the process of cancer stem cell formation and division as a dynamic system. Within the system, the pathways of cell compartments and their interactions are quantified by a system of differential equations [30]. The genetic control of cancer cell formation can be understood in terms of the molecular mechanisms of neoplastic processes. The model allows the test of numerous biologically and clinically meaningful hypotheses about pleiotropic control of QTLs for different steps of cancer susceptibility. The statistical properties of the model were investigated through computer simulation, validating the model's statistical usefulness and utility.

To be more clinically relevant, the framework described should be filled by the latest discoveries of genetic control for cancer stem cell growth and response to drug interventions. The model should be innovated to characterize a network of genes that interact in a coordinated manner [36] to determine cancer stem cells and their responsiveness to chemotherapeutic drugs. Genetic imprinting, a phenomenon where the expression of a gene relies on the paternal or maternal origin of an allele [37–39], is caused by epigenetic marks, and has been thought to be important for cancer susceptibility [40–42]. The molecular mechanisms for genetic imprinting arise from epigenetic marks by switching the genetic information on and off [40,43]. A new design based on family structure [44] can be adopted to take into account the functional role of genetic imprinting and epigenetic marks in initiating cancer stem cells.

Pharmacogenes, genes that regulate drug response, act usually in a way of interacting with metabolic and environmental signals. It is essential that the model is integrated with metabolic and environmental factors; equipped to characterize key pathways of interactions between pharmacogenes and these factors toward the outcome of drug reactions. While inhibiting the formation and growth of cancer cells, some drugs, such as temozolomide, can induce genetic mutations of host cells [45]. How these mutations affect the efficacy and toxicity of chemotherapeutic drugs should be quantified and incorporated into the model framework.

Currently, GWAS have been used as a routine tool to scan functional or causal polymorphisms from 300 000 to 1 000 000 SNPs [46]. Other genetic variants, such as copy number variation (CNV), have also been increasingly recognized as important contributors to drug response. A powerful protocol allows genome-wide CNV identification from single nuclei isolated from a mixed population of cells [47], providing a new avenue for detecting

pharmacogenes. Recent developments in statistical variable selection modeling have enabled a comprehensive analysis of all the markers and CNV that cover the entire genome [48], facilitating the search of all possible QTLs. Thus, the integration of our systems mapping and GWAS through variable selection can potentially promote our ability to elucidate the picture of the genetic architecture of chemotherapeutic efficacies for cancer stem cells.

Acknowledgments

This work is partly supported by the Changjiang Scholars Award, 'One-Thousand Person' Award and NIH/UL1RR0330184.

References

1 Dick, J.E. (2003) Breast cancer stem cells revealed. *Proc. Natl. Acad. Sci. U. S. A.* 100, 3547–3549
2 Alison, M.R. (2005) Liver stem cells: implications for hepatocarcinogenesis. *Stem Cell Rev.* 1, 253–260
3 Lobo, N.A. *et al.* (2007) The biology of cancer stem cells. *Ann. Rev. Cell Dev. Biol.* 23, 675–699
4 Baumann, M. *et al.* (2008) Exploring the role of cancer stem cells in radioresistance. *Nat. Rev. Cancer* 8, 545–554
5 Visvader, J.E. and Lindeman, G.J. (2008) Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* 8, 755–768
6 Al-Hajj, M. *et al.* (2004) Therapeutic implications of cancer stem cells. *Curr. Opin. Genet. Dev.* 14, 43–47
7 Clarke, M.F. and Fuller, M. (2006) Stem cells and cancer: two faces of eve. *Cell* 124, 1111–1115
8 Lang, J.Y. *et al.* (2011) BikDD eliminates breast cancer initiating cells and synergizes with lapatinib for breast cancer treatment. *Cancer Cell* 20, 341–356
9 Dean, M. *et al.* (2005) Tumour stem cells and drug resistance. *Nat. Rev. Cancer* 5, 275–284
10 Clarke, R. *et al.* (2008) The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. *Nat. Rev. Cancer* 8, 37–49
11 Cohen, A.A. *et al.* (2008) Dynamic proteomics of individual cancer cells in response to a drug. *Science* 322, 1511–1516
12 Ganguly, R. and Puri, I.K. (2006) Mathematical model for the cancer stem cell hypothesis. *Cell Prolif.* 39, 3–14
13 Anderson, A.R.A. and Quaranta, V. (2008) Integrative mathematical oncology. *Nat. Rev. Cancer* 8, 227–233
14 Piotrowska, M.J. *et al.* (2008) Mathematical modeling of stem cells related to cancer. In *Cancer and Stem Cells* (Dittmar, T. and Zanker, K.S., eds), pp. 1–25, Nova Science Publishers, New York
15 Adimy, M. *et al.* (2008) Numerical integration of a mathematical model of hematopoietic stem cell dynamics. *Comp. Math. Appl.* 56, 560–594
16 Wang, Z. *et al.* (2012) Dynamic mapping of genes controlling cancer stem cell proliferation. *Front. Genet.* 3, 84
17 Huntly, B.J. and Gilliland, D.G. (2005) Summing up cancer stem cells. *Nature* 435, 1169–1170
18 Michor, F. (2008) Mathematical models of cancer stem cells. *J. Clin. Oncol.* 26, 2854–2860
19 Wu, R.L. *et al.* (2011) Systems mapping: how to improve the genetic mapping of complex traits through design principles of biological systems. *BMC Syst. Biol.* 5, 84
20 Kitano, H. (2002) Systems biology: a brief overview. *Science* 295, 1662–1664
21 Alon, U., ed. (2007) *An Introduction to Systems Biology: Design Principles of Biological Circuits.*, Chapman and Hall/CRC, Boca Raton
22 Wu, R.L. and et, al., eds) (2007) *Statistical Genetics of Quantitative Traits: Linkage, Maps, and QTL*, Springer-Verlag, New York
23 Wu, R.L. and Lin, M., eds) (2008) *Statistical and Computational Pharmacogenomics*, Chapman & Hall/CRC, London
24 Fu, G.F. *et al.* (2010) A dynamic model for functional mapping of biological rhythms. *J. Biol. Dynamics* 4, 1–10
25 Fu, G.F. *et al.* (2011) A mathematical framework for functional mapping of complex systems using delay differential equations. *J. Theor. Biol.* 289, 206–216
26 Zimmerman, D.L. and Nunez-Anton, V. (2001) Parametric modeling of growth curve data: an overview (with discussion). *Test* 10, 1–73
27 Zhao, W. *et al.* (2005) A nonstationary model for functional mapping of complex traits. *Bioinformatics* 21, 2469–2477
28 Yap, J.S. *et al.* (2009) Nonparametric modeling of covariance structure in functional mapping of quantitative trait loci. *Biometrics* 65, 1068–1077
29 Al-Hajj, M. and Clarke, M.F. (2004) Self-renewal and solid tumor cells. *Oncogene* 23, 7274–7282
30 Ganguly, R. and Puri, I.K. (2007) Mathematical model for chemotherapeutic drug efficacy in arresting tumour growth based on the cancer stem cell hypothesis. *Cell Prolif.* 40, 338–354
31 Byrne, H.M. (2003) Modeling avascular tumour growth. In *Cancer Modeling and Simulation* (Preziosi, L., ed.), p. 75, Chapman & Hall/CRC, London
32 Wang, L.W. *et al.* (2011) Genomics and drug response. *N. Engl. J. Med.* 364, 1144–1153
33 Fugger, K. *et al.* (2013) FBH1 co-operates with MUS81 in inducing DNA double-strand breaks and cell death following replication stress. *Nat. Commun.* 4, 1423
34 Huff, C.A. *et al.* (2006) Strategies to eliminate cancer stem cells: clinical applications. *Eur. J. Cancer* 42, 1293–1297
35 Merlos-Suárez, A. *et al.* (2011) The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* 8, 511–524
36 Li, Y. and Wu, R.L. (2009) Modeling host–cancer genetic interactions with multilocus sequence data. *J. Comp. Sci. Syst. Biol.* 2, 24–43
37 Reik, W. and Lewis, A. (2005) Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat. Rev. Genet.* 6, 403–410
38 Jirtle, R.L. and Skinner, M.K. (2007) Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* 8, 253–262
39 Li, Y. *et al.* (2011) A statistical design for testing transgenerational genomic imprinting in natural human populations. *PLoS One* 6, e16858
40 Feinberg, A.P. and Tycko, B. (2004) The history of cancer epigenetics. *Nat. Rev. Cancer* 4, 143–153
41 Secko, D. (2005) How an imprint can lead to cancer. *Can. Med. Assoc. J.* 172, 1286
42 Sasaki, H. and Matsui, Y. (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* 9, 129–140
43 Sha, K. (2008) A mechanistic view of genomic imprinting. *Ann. Res. Genomics Hum. Genet.* 9, 197–216
44 Li, Q. and Wu, R.L. (2009) A multilocus model for constructing a linkage disequilibrium map in human populations. *Stat. Appl. Genet. Mol. Biol.* 8, 18
45 Taverna, P. *et al.* (1992) Influence of O6-methylguanine on DNA damage and cytotoxicity of temozolomide in L1210 mouse leukemia sensitive and resistant to chloroethylnitrosoureas. *Anti-Cancer Drugs* 3, 401–405
46 Pearson, T.A. and Manolio, T.A. (2008) How to interpret a genome-wide association study. *J. Am. Med. Assoc.* 299, 1335–1344
47 Baslan, T. *et al.* (2012) Genome-wide copy number analysis of single cells. *Nat. Protoc.* 7, 1024–1041
48 Li, J.H. *et al.* (2011) Bayesian lasso for genome-wide association studies. *Bioinformatics* 27, 516–523