

Cancer systems biology: a network modeling perspective

Pamela K.Kreeger and Douglas A.Lauffenburger^{1,*}

Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA and ¹Department of Biological Engineering, Massachusetts Institute of Technology, Building 16, Room 343, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

*To whom correspondence should be addressed. Tel: +1 617 252 1629;
Fax: +1 617 258 0204;
Email: lauffen@mit.edu

Cancer is now appreciated as not only a highly heterogeneous pathology with respect to cell type and tissue origin but also as a disease involving dysregulation of multiple pathways governing fundamental cell processes such as death, proliferation, differentiation and migration. Thus, the activities of molecular networks that execute metabolic or cytoskeletal processes, or regulate these by signal transduction, are altered in a complex manner by diverse genetic mutations in concert with the environmental context. A major challenge therefore is how to develop actionable understanding of this multivariate dysregulation, with respect both to how it arises from diverse genetic mutations and to how it may be ameliorated by prospective treatments. While high-throughput experimental platform technologies ranging from genomic sequencing to transcriptomic, proteomic and metabolomic profiling are now commonly used for molecular-level characterization of tumor cells and surrounding tissues, the resulting data sets defy straightforward intuitive interpretation with respect to potential therapeutic targets or the effects of perturbation. In this review article, we will discuss how significant advances can be obtained by applying computational modeling approaches to elucidate the pathways most critically involved in tumor formation and progression, impact of particular mutations on pathway operation, consequences of altered cell behavior in tissue environments and effects of molecular therapeutics.

Introduction

The advent of genomic sequencing and copy number evaluation has revealed that aspirations for understanding cancer in terms of mutation of some limited number of oncogenes and tumor suppressor genes are unlikely to be fulfilled; instead, even tumors of a particular tissue type bear highly heterogeneous sets of defects in dozens of different genes (1–3). Similarly, RNA interference studies show that a large number and wide spectrum of gene products contribute to tumor cell phenotype (4,5). The processes affected closely correspond to defined ‘hallmarks’ of cancer, such as resistance to cell death, extension of replicative potential, enhancement of invasiveness, and escape from immune surveillance, among others (6,7). Thus, the relationship between genomic information *per se* and malignant disease is more elusive than originally hoped, due to the broadly multivariate nature of the molecular-level changes involved in any given cancer (8). Moreover, even when a specific gene is identified to make a substantial contribution to pathology, that determination does not easily lead to an effective avenue for treatment because of the complex consequences propagated down transcriptional, translational and posttranslational circuits (7,9).

Abbreviations: EGFR, epidermal growth factor receptor; ERK, extracellular signal-regulated kinase; TNF α , tumor necrosis factor α .

Although the genomic mutations associated with any tumor are many and diverse, a promising organizing principle is emerging: that a recognizable cohort of key pathways governing cell phenotypic behaviors can be identified as pathologically altered by the various underlying genetic defects (10,11). This notion is encouraging for elucidating therapeutic targets because a pathway (or multiple pathways) can be targeted through any of numerous comprised components—whether or not they in fact suffer from genetic defect—rather than being restricted solely to mutated oncogenes or tumor suppressor genes and needing to find a unique approach for each gene product (7). An example of this is the class of mammalian target of rapamycin inhibitors (including temsirolimus and everolimus) that are being found effective for treatment of renal cell carcinoma; the predominant genetic defect in these tumors is loss of Von-Hippel Lindau tumor suppressor (thus reducing degradation of the hypoxia-inducible transcription factors HIF1/2/3) rather than any mutation related to the mammalian target of rapamycin gene itself (12).

A pathway-centric approach remains incomplete, however, because of the intricate cross talk among cell regulatory pathways (13). Indeed, a given molecular component can be identified to be associated with or interact with multiple signaling, transcriptional regulation, metabolic and/or cytoskeletal process pathways (14). Pathways thus cannot properly be considered to operate in isolation of one another, as an alteration of one pathway can lead directly (via protein–protein interactions) or indirectly (via transcriptional/translational influences) to changes in others. Accordingly, cancer—along with other complex diseases such as arthritis and diabetes—is most productively conceived of and strategized for treatment as a dysregulation of a multi-pathway network (15,16). Moreover, these networks connect to components beyond the tumor cells themselves, including other cells in the environment along with the extracellular milieu (17,18). This perspective yields at least two consequences. First, experimental characterization of the pathological dysregulation will need to be multivariate and quantitative. For instance, biomarkers in the form of single molecular components or qualitative component lists will probably be inadequate even for categorization of treatment outcomes. Second, predictive or mechanistic understanding of the pathology will almost certainly elude intuition unaided by computational analysis. Hence, a network perspective on cancer strongly motivates the application of computational modeling approaches (19–21). Methodologies for computational analysis can vary widely depending on the question being posed and the experimental data at hand, ranging from highly abstracted models using correlative regression to highly specified models using differential equations, with network component interaction and logic modeling techniques intermediate to these. A number of reviews have provided discussion of which of these methods are more or less appropriate for employment for various kinds of studies, outlining their respective strengths and weaknesses with respect to different applications (22–25).

Herein, we will present selected examples of recent research contributions that are helping to establish the field of cancer systems biology. These studies demonstrate the unique advances in understanding and prediction that can be gained by integration of computational modeling with quantitative experimental data on molecular and cellular networks. Our examples emphasize systems at the level of dynamic protein operations, such as phosphorylation, because it is this level that most effectively integrates the convolution of genomic information and environmental context. As illustrated in Figure 1, environmental context strongly influences network operations

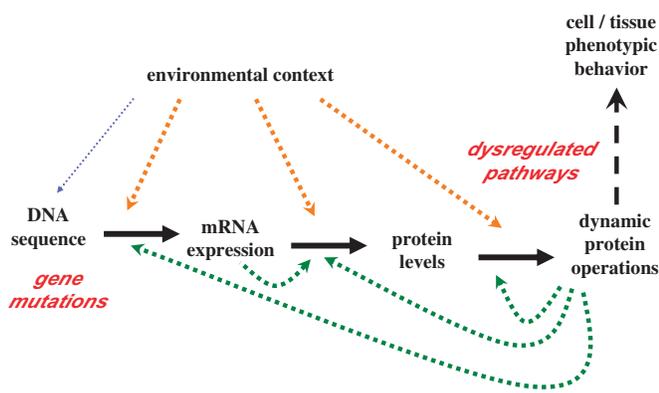


Fig. 1. Schematic illustration of molecular processes governing cell and tissue functional behavior, depicting how genetic alterations convolute with environmental context to yield ultimate pathos/physiological phenotypes. Environmental context influences transcriptional, translational and posttranslational processes and moreover can modulate genomic information (directly via DNA mutation or indirectly via epigenetic modulation). Thus, the molecular-level characterization containing the greatest amount of information concerning phenotypic behavior resides in the realm comprehending both genomic and environmental effects: dynamic protein network operations.

regulating transcription, translation and posttranslational processes, so that restricting experimental measurement information to DNA sequence, mRNA expression or even protein expression will miss important aspects of the molecular components and interactions governing cell and tissue behavior. We organize this presentation into five main categories, denoting the kinds of problems being addressed by the various studies: identifying dysregulated pathways, elucidating consequences of mutations on network activities, integrating network operation into cell behavioral functions, integrating cell behavior into tissue-level processes and predicting effects of molecular interventions.

Identifying dysregulated pathways

Historically, dysregulated pathways have been identified in cancers based on reductionist studies surrounding an identified mutation. The shift to examining dysregulated networks requires new techniques to identify pathways within the context of the intact cellular network. An especially compelling avenue for identifying affected pathways is the use of interactomes, which define the molecular interactions in a cell. Although still at an early developmental stage for human cells and tissues, interactomes include protein–protein and protein–DNA associations (26,27) and provide a framework for analysis of empirical data of various types, such as transcriptomic, phosphoproteomic and phenotypic assessments. The biochemical resolution of interactome information has been recently enhanced by directly incorporating phosphoproteomic data to indicate not only physical interactions but also kinase–substrate interactions (28). A further exciting breakthrough toward an increasingly powerful network framework for identifying pathways involved in responses to environmental conditions or in governing phenotypic behaviors is offered by a recent development of a Steiner tree computational algorithm, able to integrate protein–protein and protein–DNA interactomes (29). The initial application of this new methodology has been in yeast, demonstrating success in linking genetic data (e.g. knockouts and knockdowns) with gene expression data and in gleaning new insights concerning network activities characterizing the yeast pheromone response. These techniques promise to improve our ability to discern the complex networks impacted by oncogenic changes in human cells.

Utilizing an internally assembled interactome and a compendium of >200 transcriptional microarray profiles for normal and tumor-related B cells, Mani *et al.* (30) investigated the pathways dysregulated in

human B cells for three kinds of non-Hodgkin's lymphomas (follicular, Burkett's and mantle cell). A mutual information algorithm was applied to identify interactions exhibiting gain-of-correlation or loss-of-correlation with respect to tumor phenotype relative to normal. Dysregulated interactions were defined as those exhibiting mutual information (essentially, correlation) in all samples except for a particular phenotype (loss-of-correlation) or lack of mutual information in all samples except for a particular phenotype (gain-of-correlation). A fundamentally interesting finding was that ~80% of the roughly 65 000 network interactions appeared to be unaffected in any of the tumor phenotypes. That is, the interactions demonstrated similar mutual information whether in tumor or normal cellular backgrounds; the authors noted that this implies a network 'backbone' that operates consistently across various cellular backgrounds. Nonetheless, hundreds of interactions were discerned to be differentially correlated with each of the particular tumor phenotypes, cutting across multiple diverse pathways. A further important insight obtained was that dysregulated pathway interactions could arise even without mutation of genes explicitly corresponding to the pathway components involved.

Chang *et al.* (31) pursued an analogous study of the NCI-60 cancer cell line set (32), analyzing a combination of transcriptional profile data with human protein–protein interactome information but using a different computational method called statistical factor analysis. This method undertakes a linear regression of gene expression data with respect to particular gene signature vectors, determining score coefficients relating the strengths of contribution of corresponding vectors to a given set of transcriptional profiles. A major focus of this study was elucidation of how important different effector pathways downstream of RAS are in generating the expression profiles exhibited by certain cancer cell lines. Subsets of cell lines bearing stronger contributions of the extracellular signal-regulated kinase (ERK) effector pathway or the AKT effector pathway were identified and validated, at least in aggregate, by their comparative sensitivities to ERK pathway inhibitors versus AKT pathway inhibitors.

Models that include influence and/or logic aspects of network component interactions can also be constructed, with the advantage of understanding propagation of pathway activities following introduction of stimulatory or inhibitor cues. One example application of this kind of approach was recently contributed by Heiser *et al.* (33). The authors utilized Pathway Logic formalism to elucidate how the ERK pathway is activated downstream of ErbB family receptors in breast cancer cell lines. A network comprising 286 signaling nodes and 396 interaction rules was constructed via literature curation, and disparities in network topologies across different cell lines were identified from cell line-associated mRNA and/or protein expression data for 191 of these components. The expression levels were discretized into 'present' or 'absent' categories, thus yielding altered interaction rules among cell lines depending on which components were deemed in one category or the other. Across the 30 breast lines examined, roughly 40 components were found to be disparately expressed (with respect to present versus absent calls), resulting in >50% of the interaction rules varying concomitantly. Clustering the cell lines with respect to network features generated classification associated with pathway interactions, leading to identification of novel points such as the presence of PAK1 in association with strong activation of ERK. Explicit calculations on network information flows can be conducted using dynamic Boolean logic or fuzzy logic algorithms (34–37). To date, these algorithms have been mainly applied to hand-curated networks, but the prospect for being founded on formal interactome databases is clear. Indeed, a new contribution shows how a Boolean logic model appropriate for an actual cell type and particular set of contexts can be derived from combination of empirical data with a prior interactome database (38). Bayesian networks, which describe pathway interactions (whether direct or indirect) in terms of probabilistic influences of certain components upon other components, have been demonstrated useful for understanding operation of cell signaling networks but not yet applied to cancer biology problems (39,40).

We note that one alternative approach to using interactome information is the definition of gene modules based on Gene Ontology categories, which has been successfully applied to discerning types of pharmacological agents effective in killing tumor cells representative of cohorts classified by these modules (9). Another option not reliant on an interactome framework is direct application of mutual information-based algorithms, such as algorithm for the reconstructions of accurate cellular networks (41) to relate transcription factor activity to gene expression profiles. This methodology has been successfully demonstrated in work that ascertained how NOTCH1 and c-MYC work together to regulate growth of T-cell leukemia cells (42).

Elucidating consequences of mutations

Oncogenic mutations affect cell behavior by changing the cellular network (13). While these effects include obvious changes to proteins immediately downstream, systems biology studies are beginning to reveal how apparently small changes in the network have broad-reaching effects (43). Here, we will profile recent studies of three of the most common mutated pathways in cancer—p53, the ErbB family of receptors and RAS. While many reports have focused on the identification of biomarkers for tumor detection, our focus is on how the mutation changes the network and how this information could be used to identify new targets or treatment strategies.

p53

In response to DNA damage, a variety of pathways are activated in the cell to arrest cell cycle progress, repair the damage or initiate apoptosis in order to prevent generation of cells with mutations. Accordingly, many tumors have alterations in their DNA damage response pathways. A number of systems biology studies have focused on p53, the tumor suppressor mutated in at least half of all cancers. In single cells, levels of p53 oscillate in response to stress, which has been shown computationally to involve feedback circuits in the DNA damage-signaling network. Batchelor *et al.* (44) used a combination of dynamical systems modeling and quantitative single-cell experiments to determine that the p53/MDM2-negative feedback loop is critically involved in this phenomenon. Activation of signaling kinases ATM and CHK2 activation downstream of DNA damage drive p53 peaks, with WIP1-mediated feedback maintaining coherence during ensuing cycles. In an ensuing contribution employing stochastic physicochemical modeling, p14^{ARF} was identified as another key participant in the feedback (45). An important further step was provided by Toettcher *et al.* (46) in connecting these signals to effects on cell cycle control. By integrating the DNA damage-signaling pathways with cell cycle regulatory pathways, the authors were able to investigate the role of p53 in both initial and long-term maintenance of cell cycle arrest. Initial arrest was determined to be p53 independent; however, p53 was necessary to maintain arrest, indicating its key role in preventing accumulation of DNA damage and importance in tumor development. Moreover, a novel prediction of pathological endo-reduplication behavior enabled by defects in the p21 pathways was produced by the model and validated by direct experimental test. Linking model predictions to experimental tests is an important step to further the application of systems biology to the study of oncogenic impact.

ErbB receptors

The ErbB family of receptor tyrosine kinases and components of the downstream network are frequently mutated in cancer (47). The complexity of the ErbB system, with four receptor isoforms and >12 ligands, makes it an ideal network to analyze by systems biology methods (48). Not surprisingly, this system has been subject of numerous studies applying modeling to quantitative experimental data, with the most recent contributions incorporating multiple members of the receptor family (49,50).

One common perturbation to the ErbB network found in breast, lung and colon cancers is the overexpression of ErbB2 (HER2). A mass-action model accounting for ErbB1–4 dimerization and sig-

naling to ERK and AKT determined that overexpression of ErbB2 shifts the cell to a higher percentage of ErbB1–2 heterodimers in place of ErbB1 homodimers (51). This model suggests that since these heterodimers do not undergo ligand-induced degradation, increase of ErbB2 results in sustained signaling. Utilizing partial least squares regression modeling applied to phosphoproteomic data from human mammary epithelial cells expressing increasing levels of ErbB2, Kumar *et al.* (52) identified nine phosphorylation events that serve as a ‘network gauge’ to determine the impact of treatment proliferation or migration. The components of this gauge include some of the usual suspects such as PI3K signals and endocytosis proteins, but it is the quantitative integration of these signals, rather than any single event, that enables prediction of cell behavior.

In addition to ErbB2 overexpression, mutations to ErbB1 [epidermal growth factor receptor (EGFR)] are frequently identified in tumors. Phosphoproteomic analysis of cells expressing increasing levels of the constitutively active EGFRvIII mutant demonstrated that the active phosphorylation site of c-MET was highly responsive to EGFRvIII level (53). This suggested that EGFRvIII cross-activated c-MET; indeed, dual inhibition strategies against EGFR and c-MET proved to be more effective in killing cells than single treatments. Identifying cotreatment strategies such as these is one of the promising applications of systems biology. Additionally, modeling can help to explain clinical observations of why certain tumors are more susceptible to treatment by targeted inhibitors. Deletion of exon 19 and the point mutation L858R in EGFR are associated with elevated phosphorylated AKT and sensitivity to the EGFR tyrosine kinase inhibitor gefitinib. Using a mass-action model incorporating receptor internalization and activation of ERK and AKT, slower EGFR internalization was determined to be sufficient to explain the signaling differences observed between wild-type EGFR and these mutants (54). When receptor internalization was quantified, mutant cell lines were indeed found to have slower rates. The model suggests that delayed internalization leads to AKT addiction, which is then inhibited by gefitinib treatment, providing an explanation for the limited benefit for tumors with wild-type EGFR.

Oncogenic RAS

Downstream of various receptor tyrosine kinases, members of the RAS family of GTPases are frequently mutated in human cancers. RAS mutations in tumors are primarily point mutations in one isoform and result in insensitivity to GTPase-activating proteins that increase GTP hydrolysis. These mutations result in increased levels of active RAS–GTP in cells, whereas additional non-mutant RAS isoforms are subject to ligand-induced activation (55). The RAS pathway has been modeled downstream of various receptors (50,51,56), and the RAS activation–deactivation cycle has been analyzed quantitatively (55). However, these contributions have not addressed the effects of RAS mutation on downstream effectors, which are not as straightforward as simply enhanced effector activities. Two recent systems biology studies have focused on this issue and have demonstrated that oncogenic RAS impacts multiple targets in complex ways. In the first study, an inducible H-RAS^{G12V} construct was activated and rapid induction of two phosphatases, DUSP1 and DUSP6, was observed (57). To examine this behavior, multiple model structures were characterized to see which model structure best explained the experimental data. The best-fit model incorporated ultrasensitive activation of ERK in response to RAS activation, ERK induction of DUSP6 and not DUSP1 and feedback from DUSP6 against phosphorylated ERK. Additionally, we have recently reported the impact of mutations to two different isoforms of RAS, K-RAS and N-RAS, on the response of colon carcinoma cell lines to tumor necrosis factor α (TNF α) (58). Using a quantitative phosphoproteomic data set, we determined that K-RAS-mutant cells have decreased activation of ERK due to a depressed TNF α -induced release of autocrine transforming growth factor- α and that N-RAS-mutant cells do not induce DUSP5 as robustly as either wild-type or K-RAS-mutant cells, resulting in sustained activation of ERK. Studies such as these demonstrate how

small perturbations impact the cell network by affecting positive and negative feedback mechanisms.

Integrating network operations into cell behavior

The process of developing from a single cell with an oncogenic profile to a metastatic cancer is exceedingly complex and multivariate. However, there are several key processes common to most cancers, including excessive proliferation, resistance to apoptosis, angiogenesis and metastasis (6). Systems biology modeling techniques promise to improve our understanding of each of these processes and ultimately, how they interact to drive tumor progression.

Excessive proliferation is perhaps the phenotype most associated with cancer progression. The ErbB-signaling network plays an important role in dysregulated proliferation in many tumors and has been extensively modeled (59). Recent models have highlighted the important multivariate characteristics of the ErbB-signaling network in proliferation as the ErbB signal transduction network interacts with a variety of other signaling pathways. Modeling approaches have proven useful in identifying dominant mechanisms of pathway cross talk. For example, a mass-action kinetic model examining the cross talk between the ErbB network and insulin signaling indicated a role for GAB1 in the synergism between treatment with epidermal growth factor and insulin (60). Additionally, identifying multi-pathway interactions may provide insight into potential new therapeutic targets. Sahin *et al.* (61) developed a Boolean model from the literature linking ErbB1–3 to phosphorylated retinoblastoma protein (a surrogate for cell cycle progression) in breast cancer resistant to ErbB2-targeted inhibitors. Following refinement, the model was able to predict a variety of new conditions that resulted from knocking down proteins in the network. Importantly, simulations indicated that inhibiting ErbB receptors would be insufficient to halt cell cycle progression in resistant cells and suggested c-MYC as a potential alternative target.

In normal development, apoptosis provides a counterpart to proliferation by removing damaged or unnecessary cells. To evade restriction of tumor growth, cancer cells have devised multiple mechanisms to provide sustained pro-growth stimuli or to counteract apoptosis. Apoptosis can be induced by stress and the mitochondrial pathway (intrinsic) or through activation of death receptors (extrinsic) (62). In the intrinsic pathways, stimuli such as DNA damage activate the apoptotic machinery. To model this process, one approach has been to develop mass-action models that decompose the intrinsic apoptotic pathway into subsections of the molecular network for analysis (63). Using such models to examine the impact of DNA damage on p53 phosphorylation suggested that transient DNA damage leads to a level of phosphorylated p53 that will induce cell cycle arrest, whereas sustained damage will induce apoptosis.

A novel mass-action model of tumor necrosis related apoptosis-inducing ligand-induced (extrinsic) apoptosis and mitochondrial outer membrane permeabilization in single cells has recently been employed to examine several mechanistic questions about the regulation of cell death in the HeLa cancer model (64,65). Model parameters were fit to experimental data from tumor necrosis related apoptosis-inducing ligand dose-response curves and small interfering RNA/overexpression perturbations to the network. A particularly striking result of the model is the impact of compartmentalization—once the steady changes in the cell network lead to mitochondrial compartment opening, there is a drastic release of SMAC down the concentration gradient and an all-or-nothing decision for apoptosis results (64). Additional work demonstrated that XIAP- and proteasome-dependent degradation of effector caspases is required to prevent caspase activity before the cell has committed to apoptosis by mitochondrial outer membrane permeabilization (65). This suggests that if this control is overridden, caspases may be activated and cause damage to target substrates, while the cell is in an 'undead' state. The developed model was also used to address the causes of cell-to-cell variation in timing of TRAIL-induced death (66). Using experimental measures of the mean and variation of five key proteins in the model, simulation results closely match the time-to-death variation observed. Combined

with other experimental observations, this suggests that cell-to-cell variation is a result of changes in protein concentrations rather than genetic or stochastic mechanisms.

It is important to remember that for a tumor to develop, it is the balance between apoptosis and proliferation that matters. In models of single colonic crypts composed of stem cells, differentiated cells and transit cells, increased proliferation, decreased differentiation or decreased apoptosis all lead to a net increase in cell number (67). Alternatively, it has been argued that resistance to apoptosis may actually decrease the ability for a tumor to expand as it would reduce the number of cell divisions before a tumor reached a critical size and lower the probability of forming the mutant combinations necessary for expansion from that stage (68). At the molecular level, recent experimental studies have demonstrated several mechanisms by which cells respond to both pro-death and pro-proliferative stimuli (69). Using principal components analysis to examine HT-29 apoptosis in response to TNF α , insulin or epidermal growth factor, a series of pro-survival and pro-death autocrine loops were revealed as a result of treatment with TNF α (70). These molecular balancing mechanisms help to explain why cytotoxic ligands are unable to kill all cells, which will be essential to consider when administering targeted chemotherapeutics.

Integrating cell functions into tissue processes

An ultimate challenge is the development of models that incorporate the behavior of the entire tumor. In the early-stages of tumor development, cancer cells are clustered together into an avascular structure, which has been modeled *in vitro* with multi-cellular spheroids (71) and *in silico* with mathematical balances of proliferating and dying cells that are coupled to physical constraints such as nutrient diffusion. In general, these models are limited to describing tumor morphology and distribution of necrotic cells (72). More recently, multiscale modeling was used to analyze avascular multicellular tumors (73). This model incorporated proliferation, diffusion and consumption of nutrients, diffusion and production of wastes, adhesion and cell-environmental interactions. A notable improvement over previous models that relied on probabilities for cellular decisions was the inclusion of a Boolean logic model for the G₁ to S progression, which incorporated molecular components such as transforming growth factor- β , p27, p21 and cyclins. Model simulations closely matched *in vitro* spheroid morphology, size and cell cycle distribution over a 20 day period (73). Further refinement of detailed models such as this should allow investigators to test the impact of molecular interventions on early-stage tumors.

As a tumor increases in size, the center core becomes necrotic and the tumor needs to develop its own vasculature to continue to grow. Therefore, targeting tumor angiogenesis is an attractive strategy to treat cancer (74). Computational models are beginning to provide tools to address how the tumor microenvironment and growth factor signaling regulate these events. Although numerous models for angiogenesis have been proffered, the complexity of the system presents a challenge regarding how to connect the behavior of the blood vessels to the tumor and how to incorporate molecular control mechanisms. Two recent models link the behavior of the tumor, growing blood vessel and cellular microenvironment. Macklin *et al.* (75) modeled tumor growth in response to the altered oxygen profile resulting from new vascular growth. In this model, mechanical forces directed blood vessel growth, and simulations demonstrated that when heterogeneous oxygen profiles were formed from the new vessels, tumor cell proliferation was heterogeneous resulting in invasive tumor morphologies. In the second model, blood vessel growth was modeled as a response to local gradients in angiogenic factors such as VEGF (76). By changing the various rules, these model forms can mimic different environmental or genetic conditions; however, they do not allow for the direct test of molecular interventions on angiogenesis.

As an alternative approach, another recent model linked the cancer cell cycle to sensing of environmental conditions such as oxygen levels, with the assumption that suboptimal oxygen levels result in the production of VEGF (77). In contrast to more phenomenological

models, the impact of VEGF on endothelial cell proliferation and migration was determined by the pharmacological Emax model and additional pro- and anti-angiogenic molecules were included. Simulations of the effect of endostatin induced by gene therapy indicated there is a critical rate of production needed; below this rate, longer treatment times were predicted to ‘rebound’ more quickly, whereas above this rate, the time to rebound increased with the duration of treatment (77). Expansion of the molecular components of multiscale models of angiogenesis will be essential to utilize them to determine drug targets and optimal treatment regimens.

An alternative method for a tumor to continue its growth is for cells to leave the primary tumor and implant in other tissues—the metastases that result from this process are responsible for the majority of cancer deaths. It is becoming increasingly apparent that metastasis is more than a random process, with additional mutations required for cancer cells to leave the primary tumor and metastatic cells demonstrating ‘preferences’ for target organs (78). Similar to angiogenesis, models of tumor invasion and metastasis must incorporate multiple scales and are primarily constructed using approximations of cellular signaling. For example, in one virtual tumor model, cells respond to microenvironmental changes (e.g. oxygen level, cell proximity) and decide to proliferate or die (79). At the same time, cells undergo random inheritable mutations that change the ‘phenotype’ of the cell (e.g. the likelihood to proliferate). Changes to the microenvironment to incorporate harsher conditions such as hypoxia or heterogenous matrix led to selection for cells with more aggressive traits and invasion tumor shapes. The impact of the matrix on cancer cell invasion is supported by a recent model of invadopodia, the cellular extensions believed to degrade extracellular matrix as tumor cells invade (80). In this model a single cancer cell sits on top of a matrix, which is modeled from known dimensions and characteristics of extracellular matrix proteins. A series of rules describe how invadopodia invade and interact with the extracellular matrix. The simulations suggested that dense matrix (such as basement membrane) forms an effective barrier to prevent invadopodia penetration and matrix degradation, whereas looser matrices (such as stroma) have gaps sufficient to allow invadopodia penetration and matrix degradation.

In order to metastasize to distant organs, cancer cells must invade into the blood or lymph to be transported to other parts of the body. The first stage in this process is for the cancer cell to invade the endothelial layer. To model this process, a multiscale model of transendothelial migration was developed, incorporating both endothelial cells, cancer cells and the cadherin interactions between the distinct cells (81). In the simulated invasion, cancer cells attach to endothelial cells by N-cadherins, and the endothelial layer is connected by VE-cadherin bonds, which the cancer cell must break in order to squeeze through. The model follows protein concentrations within each cell through a series of ordinary differential equations, cell–cell forces by a modified Hertz model and cell movement according to Langevin equations. As cancer cells contact endothelial cells, a N-cadherin bond forms between the cells, leading to competition for β -catenin and dissolution of the existing VE-cadherin bonds. The observed behaviors are similar to experimental results that noted the loss of N-cadherin slows transendothelial migration (82).

Ultimately, models such as these must be built upon to balance multiple signaling pathways and phenotypic processes in order to be used for identification of drug targets. A novel multiscale model of non-small-cell lung cancer provides an example of how such a process can be implemented. In an early version, ErbB activation of ERK and PLC γ in each cell in a tumor was modeled by a series of ordinary differential equations (83). Phosphorylated ERK and PLC γ are used as readouts for proliferation and migration, respectively, and cells migrated or proliferated in a virtual two-dimensional space. In this early version, overexpression of ErbB1 led to a migration dominant phenotype and accelerated tumor expansion. Building on this initial work, the group expanded the virtual space and tumor to three dimensions and added transforming growth factor- β activation of RAS to each cell’s signaling module (84). By comparing the results of simulations with single and cotreatments, the expanded model demon-

strated conditions where targeting a single pathway would be ineffective in deterring tumor expansion.

Predicting effects of molecular interventions

Importantly, computational network models have begun to identify novel targets in regulatory networks with the goal of more effectively treating tumors. Schoeberl *et al.* developed a mass-action kinetic model of ErbB1–3 incorporating ligand binding (betacellulin and heregulin1- β), receptor dimerization, internalization, degradation, recycling and downstream signaling through PI3K to phosphorylated AKT (85). Sensitivity analysis of the model revealed a dominant role of ErbB3 in the level of phosphorylated AKT, and *in silico* tests of a monoclonal antibody against ErbB3 demonstrated its effectiveness at decreasing phosphorylated AKT over a range of ErbB1/2 levels. Based on the predictions of their systems biology model, the group developed MM-121, a fully human IgG2 monoclonal antibody against ErbB3, and found that, as in the model, blockade of ErbB3 was effective in blocking phosphorylation of AKT (85). Additionally, when mice implanted with ACHN xenografts were treated with MM-121, tumor growth was slowed even after the end of the treatment period and MM-121 has entered Phase I Trials. This study is particularly provocative as the model suggested targeting the kinase-dead ErbB3 rather than a mutated or overexpressed protein, indicating that excessive growth may be dependent on the multivariate receptor and ligand environment rather than a single oncogenic change.

However, this work—like most systems modeling efforts aimed at predicting effects of therapeutic perturbations of cell regulatory networks—restricts its attention to predicting molecular-level processes (in this particular case, AKT phosphorylation). What is vital, of course, is to predict effects of these perturbations on cell phenotypic functions at the very least; predicting effects of integrated tissue properties would be an even more distant goal. The most difficult problem is connecting molecular-level network activities to cell-level functional behavior, even in absence of therapeutic perturbations. Relational modeling methods, such as partial least squares regression, currently represent one of the most effective approaches for this goal. This approach was tested directly by Kumar *et al.* (86), who used partial least -squares regression modeling to predict how modulation of key kinase pathways such as ERK, AKT and p38 by pharmacological inhibitors alters migration behavior of ErbB2-overexpressing mammary epithelial cell migration in response to epidermal growth factor and Heregulin. The most important insight gained from this contribution was that while a quantitative combination of five phosphorylation sites on four key proteins could successfully interpret the effects of kinase inhibitors on cell phenotypic behavior across multiple treatment conditions, no individual signal could predict the phenotypic behavior by itself. This finding emphasizes that signal-to-response relationships will in general require multiple signaling pathways to be included in the model and that attempts to predict how cells will behave on the basis of alterations (whether genetic or therapeutic) to a single component or pathway will most probably be in vain.

A major issue to emphasize as we conclude this discussion is the challenge posed by the gap between *in vitro* studies and *in vivo*. This, of course, is a crucial problem for the entire cancer biology field, not merely the systems biology approach. We submit that a systems perspective, in which multiple variables are considered integratively in explicit manner, is at least as likely or more so to find some significant success in bridging this gap than a focus restricted to intuitive predictions based on individual component effects. A compelling demonstration of this notion has recently been provided by Jiang *et al.* (87). These authors showed that assessment of both ATM and p53 together is required in order to move from cell culture experiments to organism studies, with the goal of predicting the efficacy of a DNA-damaging chemotherapeutic agent on destruction of lymphoma tumors in mice and survival of human breast cancer patients. While this particular effort did not incorporate formal mathematical analysis, a schematic logic model (involving DNA-PK and Chk2 along with ATM and p53) was found helpful in explicating the integrative

network operation. Extension of this kind of perspective to an increased number of components and wider set of pathways will be facilitated by the kinds of computational approaches we have noted, in order to enable understanding and prediction beyond even skilled intuition. We emphasize the work by Schoeberl *et al.* (85) again here as a promising example of employing computational modeling for effective prediction of novel and non-intuitive cancer therapeutic drug targets, in this case ErbB3, demonstrating promising translation from *in vitro* studies to *in vivo* validation.

Funding

National Cancer Institute Integrative Cancer Biology Program (U54-CA112967-03 to D.A.L.); American Cancer Society postdoctoral fellowship (PF-08-026-01-CCG to P.K.K.).

Acknowledgements

We gratefully acknowledge the helpful comments from J.Pritchard and our anonymous reviewers.

Conflict of Interest Statement: D.A.L. notes his Scientific Advisory Board relationship to Merrimack Pharmaceuticals.

References

- Wood, L.D. *et al.* (2007) The genomic landscapes of human breast and colorectal cancers. *Science*, **318**, 1108–1113.
- Network, Cancer Genome Atlas Research Network. (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*, **455**, 1061–1068.
- Jones, S. *et al.* (2008) Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*, **321**, 1801–1806.
- Ngo, V.N. *et al.* (2006) A loss-of-function RNA interference screen for molecular targets in cancer. *Nature*, **441**, 106–110.
- Silva, J.M. *et al.* (2008) Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science*, **319**, 617–620.
- Hanahan, D. *et al.* (2000) The hallmarks of cancer. *Cell*, **100**, 57–70.
- Luo, J. *et al.* (2009) Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell*, **136**, 823–837.
- Fox, E.J. *et al.* (2009) Cancer genome sequencing—an interim analysis. *Cancer Res.*, **69**, 4948–4950.
- Wong, D.J. *et al.* (2008) Revealing targeted therapy for human cancer by gene module maps. *Cancer Res.*, **68**, 369–378.
- Bild, A.H. *et al.* (2006) Linking oncogenic pathways with therapeutic opportunities. *Nat. Rev. Cancer*, **6**, 735–741.
- Itadani, H. *et al.* (2008) Can systems biology understand pathway activation? Gene expression signatures as surrogate markers for understanding the complexity of pathway activation. *Curr. Genomics*, **9**, 349–360.
- Kapoor, A. *et al.* (2009) Targeted inhibition of mammalian target of rapamycin for the treatment of advanced renal cell carcinoma. *Cancer*, **115**, 3618–3630.
- Pawson, T. *et al.* (2007) Oncogenic re-wiring of cellular signaling pathways. *Oncogene*, **26**, 1268–1275.
- Huang, Y.J. *et al.* (2008) Targeting the human cancer pathway protein interaction network by structural genomics. *Mol. Cell. Proteomics*, **7**, 2048–2060.
- Pawson, T. *et al.* (2008) Network medicine. *FEBS Lett.*, **582**, 1266–1270.
- Tan, C.S. *et al.* (2009) Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.*, **2**, ra39.
- Kenny, P.A. *et al.* (2007) Targeting the tumor microenvironment. *Front. Biosci.*, **12**, 3468–3474.
- Weinberg, R.A. (2008) Mechanisms of malignant progression. *Carcinogenesis*, **29**, 1092–1095.
- Hornberg, J.J. *et al.* (2006) Cancer: a systems biology disease. *Biosystems*, **83**, 81–90.
- Anderson, A.R. *et al.* (2008) Integrative mathematical oncology. *Nat. Rev. Cancer*, **8**, 227–234.
- Faratian, D. *et al.* (2009) Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab. *Cancer Res.*, **69**, 6713–6720.
- Ideker, T. *et al.* (2003) Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotechnol.*, **21**, 255–262.
- Cho, C.R. *et al.* (2006) The application of systems biology to drug discovery. *Curr. Opin. Chem. Biol.*, **10**, 294–302.
- Aldridge, B.B. *et al.* (2006) Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.*, **8**, 1195–1203.
- Janes, K.A. *et al.* (2006) Data-driven modelling of signal-transduction networks. *Nat. Rev. Mol. Cell Biol.*, **7**, 820–828.
- Harbison, C.T. *et al.* (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature*, **431**, 99–104.
- Rual, J.F. *et al.* (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, **437**, 1173–1178.
- Linding, R. *et al.* (2007) Systematic discovery of *in vivo* phosphorylation networks. *Cell*, **129**, 1415–1426.
- Huang, S.S. *et al.* (2009) Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks. *Sci. Signal.*, **2**, ra40.
- Mani, K.M. *et al.* (2008) A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol. Syst. Biol.*, **4**, 169.
- Chang, J.T. *et al.* (2009) A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Mol. Cell*, **34**, 104–14.
- Scherf, U. *et al.* (2000) A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.*, **24**, 236–244.
- Heiser, L.M. *et al.* (2009) Integrated analysis of breast cancer cell lines reveals unique signaling pathways. *Genome Biol.*, **10**, R31.
- Bosl, W.J. (2007) Systems biology by the rules: hybrid intelligent systems for pathway modeling and discovery. *BMC Syst. Biol.*, **1**, 13.
- Saez-Rodriguez, J. *et al.* (2007) A logical model provides insights into T cell receptor signaling. *PLoS Comput. Biol.*, **3**, e163.
- Samaga, R. *et al.* (2009) The logic of EGFR/ErbB signaling: theoretical properties and analysis of high-throughput data. *PLoS Comput. Biol.*, **5**, e1000438.
- Aldridge, B.B. *et al.* (2009) Fuzzy logic analysis of kinase pathway crosstalk in TNF/EGF/insulin-induced signaling. *PLoS Comput. Biol.*, **5**, e1000340.
- Saez-Rodriguez, J. *et al.* Discrete logic modeling to link pathway maps and functional analysis of mammalian signal transduction. under review. *Mol. Syst. Biol.*, in press.
- Wolf, P.J. *et al.* (2005) Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. *Bioinformatics*, **21**, 741–53.
- Sachs, K. *et al.* (2005) Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, **308**, 523–529.
- Basso, K. *et al.* (2005) Reverse engineering of regulatory networks in human B cells. *Nat. Genet.*, **37**, 382–390.
- Palomero, T. *et al.* (2006) NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc. Natl Acad. Sci. USA*, **103**, 18261–18266.
- Bizzarri, M. *et al.* (2008) Beyond the oncogene paradigm: understanding complexity in cancerogenesis. *Acta Biotheor.*, **56**, 173–196.
- Batchelor, E. *et al.* (2008) Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. *Mol. Cell*, **30**, 277–289.
- Proctor, C.J. *et al.* (2008) Explaining oscillations and variability in the p53-Mdm2 system. *BMC Syst. Biol.*, **2**, 75.
- Toettcher, J.E. *et al.* (2009) Distinct mechanisms act in concert to mediate cell cycle arrest. *Proc. Natl Acad. Sci. USA*, **106**, 785–790.
- Leary, R.J. *et al.* (2008) Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc. Natl Acad. Sci. USA*, **105**, 16224–16229.
- Citri, A. *et al.* (2006) EGF-ERBB signalling: towards the systems level. *Nat. Rev. Mol. Cell Biol.*, **7**, 505–516.
- Hendriks, B.S. *et al.* (2005) Parsing ERK activation reveals quantitatively equivalent contributions from epidermal growth factor receptor and HER2 in human mammary epithelial cells. *J. Biol. Chem.*, **280**, 6157–6169.
- Chen, W.W. *et al.* (2009) Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.*, **5**, 239.
- Birtwistle, M.R. *et al.* (2007) Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol. Syst. Biol.*, **3**, 144.
- Kumar, N. *et al.* (2007) Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput. Biol.*, **3**, e4.
- Huang, P.H. *et al.* (2007) Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proc. Natl Acad. Sci. USA*, **104**, 12867–72.
- Hendriks, B.S. *et al.* (2006) Decreased internalisation of erbB1 mutants in lung cancer is linked with a mechanism conferring sensitivity to gefitinib. *Syst. Biol. (Stevenage)*, **153**, 457–466.
- Stites, E.C. *et al.* (2007) Network analysis of oncogenic Ras activation in cancer. *Science*, **318**, 463–467.

56. Wang,C.C. *et al.* (2009) PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk. *Mol. Syst. Biol.*, **5**, 246.
57. Bluthgen,N. *et al.* (2009) A systems biological approach suggests that transcriptional feedback regulation by dual-specificity phosphatase 6 shapes extracellular signal-related kinase activity in RAS-transformed fibroblasts. *FEBS J.*, **276**, 1024–35.
58. Kreeger,P.K. *et al.* (2009) RAS mutations affect tumor necrosis factor-induced apoptosis in colon carcinoma cells via ERK-modulatory negative and positive feedback circuits along with non-ERK pathway effects. *Cancer Res.*, **69**, 8191–8199.
59. Lazzara,M.J. *et al.* (2009) Quantitative modeling perspectives on the ErbB system of cell regulatory processes. *Exp. Cell Res.*, **315**, 717–725.
60. Borisov,N. *et al.* (2009) Systems-level interactions between insulin-EGF networks amplify mitogenic signaling. *Mol. Syst. Biol.*, **5**, 256.
61. Sahin,O. *et al.* (2009) Modeling ERBB receptor-regulated G1/S transition to find novel targets for *de novo* trastuzumab resistance. *BMC Syst. Biol.*, **3**, 1.
62. Chowdhury,I. *et al.* (2006) Current concepts in apoptosis: the physiological suicide program revisited. *Cell. Mol. Biol. Lett.*, **11**, 506–525.
63. Zhang,T. *et al.* (2009) Computational analysis of dynamical responses to the intrinsic pathway of programmed cell death. *Biophys. J.*, **97**, 415–434.
64. Albeck,J.G. *et al.* (2008) Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLoS Biol.*, **6**, 2831–2852.
65. Albeck,J.G. *et al.* (2008) Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Mol. Cell*, **30**, 11–25.
66. Spencer,S.L. *et al.* (2009) Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, **459**, 428–432.
67. Johnston,M.D. *et al.* (2007) Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc. Natl Acad. Sci. USA*, **104**, 4008–4013.
68. Wodarz,D. *et al.* (2007) Can loss of apoptosis protect against cancer? *Trends Genet.*, **23**, 232–237.
69. Gaudet,S. *et al.* (2005) A compendium of signals and responses triggered by prodeath and prosurvival cytokines. *Mol. Cell. Proteomics*, **4**, 1569–1590.
70. Janes,K.A. *et al.* (2006) The response of human epithelial cells to TNF involves an inducible autocrine cascade. *Cell*, **124**, 1225–1239.
71. Sutherland,R.M. (1988) Cell and environment interactions in tumor micro-regions: the multicell spheroid model. *Science*, **240**, 177–184.
72. Ferreira,S.C.Jr *et al.* (2002) Reaction-diffusion model for the growth of avascular tumor. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, **65**, 021907.
73. Jiang,Y. *et al.* (2005) A multiscale model for avascular tumor growth. *Biophys. J.*, **89**, 3884–94.
74. Bergers,G. *et al.* (2008) Modes of resistance to anti-angiogenic therapy. *Nat. Rev. Cancer*, **8**, 592–603.
75. Macklin,P. *et al.* (2009) Multiscale modelling and nonlinear simulation of vascular tumour growth. *J. Math. Biol.*, **58**, 765–798.
76. Qutub,A.A. *et al.* (2009) Elongation, proliferation & migration differentiate endothelial cell phenotypes and determine capillary sprouting. *BMC Syst. Biol.*, **3**, 13.
77. Billy,F. *et al.* (2009) A pharmacologically based multiscale mathematical model of angiogenesis and its use in investigating the efficacy of a new cancer treatment strategy. *J. Theor. Biol.*, **260**, 545–562.
78. Nguyen,D.X. *et al.* (2009) Metastasis: from dissemination to organ-specific colonization. *Nat. Rev. Cancer*, **9**, 274–84.
79. Anderson,A.R. *et al.* (2006) Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*, **127**, 905–915.
80. Enderling,H. *et al.* (2008) Dependence of invadopodia function on collagen fiber spacing and cross-linking: computational modeling and experimental evidence. *Biophys. J.*, **95**, 2203–2218.
81. Ramis-Conde,I. *et al.* (2009) Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Phys. Biol.*, **6**, 16008.
82. Qi,J. *et al.* (2006) Involvement of Src family kinases in N-cadherin phosphorylation and beta-catenin dissociation during transendothelial migration of melanoma cells. *Mol. Biol. Cell*, **17**, 1261–1272.
83. Wang,Z. *et al.* (2007) Simulating non-small cell lung cancer with a multi-scale agent-based model. *Theor. Biol. Med. Model.*, **4**, 50.
84. Wang,Z. *et al.* (2009) Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model. *Bioinformatics*, **25**, 2389–2396.
85. Schoeberl,B. *et al.* (2009) Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci. Signal.*, **2**, ra31.
86. Kumar,N. *et al.* (2008) Multipathway model enables prediction of kinase inhibitor cross-talk effects on migration of Her2-overexpressing mammary epithelial cells. *Mol. Pharmacol.*, **73**, 1668–78.
87. Jiang,H. *et al.* (2009) The combined status of ATM and p53 link tumor development with therapeutic response. *Genes Dev.*, **23**, 1895–1909.

Received August 19, 2009; revised October 17, 2009;
accepted October 18, 2009