ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/42857298

A Roadmap of Cancer Systems Biology

Article *in* Nature Precedings · January 2010

DOI: 10.1201/9781439811863-c1 · Source: OAI

CITATIONS
21

READS
125

1 author:

Edwin Wang
National Research Council Canada
119 PUBLICATIONS **2,878** CITATIONS

SEE PROFILE

All content following this page was uploaded by Edwin Wang on 19 December 2013.

The user has requested enhancement of the downloaded file.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

A Roadmap of Cancer Systems Biology

Edwin Wang[1,2]

Contents

1 Cancer Systems Biology and Personalized Medicine

1.1 Systems Biology Is Transforming Attitudes about Cancer Biology

1.2 Systems Biology Is the Tool for Personalized Medicine

2 Strategies for Cancer Systems Biology Study

2.1 Requirements for Experimental Models to Perform Cancer Systems Biology

2.2 Data Integration and Cancer Gene Network Construction

2.3 Network Visualization

2.4 Network Analysis

2.5 Dynamic Network Modeling

---

1. Biotechnology Research Institute, National Research Council Canada.

2. Center for Bioinformatics, McGill University, Montreal, Canada.

Corresponding to: Edwin.Wang@cnrc-nrc.gc.ca

***** This is the first chapter of the book: Cancer Systems Biology (Wang, E. ed), CRC Press, 2010***

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

1 Cancer Systems Biology and Personalized Medicine

1.1 Systems Biology Is Transforming Attitudes about Cancer Biology

When an accident occurs on a busy road during rush hour in a big city, such as Montreal or New York, traffic is blocked for a short time. Soon, however, drivers begin to turn around and use alternative roads to reach their destinations. A road map of a city is a web, a collection of intertwined roads that allows for identification of alternative routes. Increasing evidence (see Chapters 4-7) shows that, similar to roads, molecules in cells are also networked. This structure suggests that biochemical pathways are interconnected, which may allow cancer to bypass the effects of a drug.

Traditional approaches to biological studies rely mainly on linear verbal logic and illustrative descriptions without mathematical explanations. These approaches are only satisfactory for addressing mechanisms that involve a small number of elements or short chains of causality. Therefore, the approaches of traditional biology are unable to capture and unravel elaborate webs of molecular interactions. Most diseases, including cancer, involve a large number and variety of elements that interact via complex networks and, consequently, display highly nonlinear dynamics. Therefore, simply knocking out one target molecule in a biochemical pathway is not sufficient for treating a disease like cancer, because the cells often find alternative molecular routes to escape the blockage. This is one reason why current drug design strategies often fail. It is increasingly believed that a systems perspective, rather than the current gene-centric view, could solve these problems and open up entirely new options for cancer treatment.

The systems approach to biological studies combines empirical, mathematical and computational techniques to gain an understanding of complex biological and physiological phenomena. For example, hundreds of proteins might be involved in signaling processes that ensure proper functioning of a cell. If such a signaling network is disturbed or altered, a cancer phenotype could be generated. As we discussed in Chapters 4 and 5, systems biology helps to shed light on these complex phenomena by generating detailed route maps of the various kinds of cellular networks and by developing sophisticated mathematical, statistical and computational methods and tools to analyze these networks. Understanding the complex systems involved in cancer will make it possible to develop smarter therapeutic strategies, for example, by disrupting two or three

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

key intersections in a biochemical network at the same time. These approaches could lead to significant advances in the treatment of cancer and help in transforming traditional reductionism-based approaches into unbiased systems-level approaches for drug discovery.

The birth and growth of the field of systems biology have been driven by technological innovation in high-throughput techniques targeted to life science applications. Over the past few years, high-throughput techniques, such as next generation genome sequencing, RNA-seq, chip-on-chip, large-scale immunoprecipitation (ChIP-seq), microarrays and others, have been developed and used to measure gene expression and gene regulatory elements to identify genes that influence some interesting phenotype on a genome-wide scale. These technologies have triggered a dramatic change in the style of biological studies from a ‘one gene model’ (i.e., focusing on the identification of individual genes and proteins and pinpointing their roles in the cell) to a ‘multiple gene model’ (i.e., the belief that molecules almost never act alone and biological entities are 'systems' - collections of interacting parts) and have generated many ‘large-scale biology projects’. As these technologies become more affordable and accessible, the implementation of large-scale biological projects is becoming more popular and routine.

With the emergence of systems biology, huge amounts of biological data have been produced and this trend is expected to continue in the future. The nature of high-throughput data is more comprehensive and unbiased than one-on-one biological data. This high-throughput approach to research has greatly altered the field of cancer research. Scientists have quickly realized that the combination of data management, interpretation, and our ability to obtain insights into these data are now the bottleneck in systems science, because ‘real signals’ or molecular mechanisms and biological principles are buried in this flood of data.

The only way to deal with large amounts of data and the relationships within those datasets is through mathematical representation and computation. Systems biology tends to meet theses challenges by integrating many types of -omic data and developing effective computational tools to decipher the complex systems. Network and graph theory have been developed to describe, analyze and model the complexity of these biological

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

systems using a mathematical language. As shown in Chapters 2, 4, 6 and 8, by applying network theory to biological systems, we are able to transform the biological language into a mathematical language, which is computable and can deal with the huge number of relations in a biological dataset. In fact, the fundamental framework of systems biology is network biology, which involves the use of networks to represent complexity, compute and model biological relationships and seek to uncover biological principles and insights. A detailed discussion of network biology can be found in Chapter 2. Examples of cancer network studies can be found in Chapters 4-6 and 8.

This chapter illustrates strategies, procedures and computational techniques for the study of cancer systems biology by focusing on network reconstruction, network analysis and modeling. Meanwhile, to match the contents of these strategies and procedures, I will guide the readers to the relevant chapters of this book. Finally, certain challenges and hurdles in cancer systems biology will also be discussed.

1.2 Systems Biology Is the Tool for Personalized Medicine

Recent studies have determined that many drugs work well for less than half of the patients for whom they are prescribed. Furthermore, nearly 3 million incorrect or ineffective prescriptions are written annually and more than 100,000 people in the U.S. die each year from drug-related adverse events (Kirk *et al.*, 2008). These data strongly suggest that one-size-fits-all medicine and preventive care are not effective. Moreover, effective treatment of disease requires that the provider consider the effects of the patient's personal genetic background. Personalized medicine is a proposed approach to develop treatment regimes that take into account each patient's unique genetic profile, allowing the treatment to fit the specific needs of subpopulations of patients with different genetic backgrounds. Furthermore, this approach would help doctors to better evaluate the risk-to-reward scenarios and prescribe appropriate pharmaceuticals for different subpopulations of patients.

Over the past decade, cancer therapy has slowly begun to change from a one-size-fits-all approach to a more personalized approach. In a personalized approach, patients are treated based on the specific genetic defects present in their tumor. However, cancer is an extremely complex, heterogeneous disease. It is believed that crucial breakthroughs

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

in the treatment of cancer, in the framework of personalized medicine, rely on the achievements of the powerful scientific approach of systems biology. Therefore, more efforts in '-omics' and systems biology have been made in the cancer research community. As a result, a tremendous amount of money has been poured into the field of cancer research over the past few years. Relatively speaking, more high-throughput data have been generated in cancer biology than in any other field of biology. However, the complexity of cancer is a major obstacle preventing a comprehensive understanding of the underlying molecular mechanisms of tumorigenesis. To crack the cancer code, network approaches have been developed and applied to cellular networks of cancer.

The examination of the entire genome of tumors (i.e., for the identification of cancer driver-mutating genes) and the global profiling of -omic data for cell signaling (i.e., gene expression, epigenetic and metabolomic profiles, and signaling data such as phosphoproteomic profiles) will aid in the construction of patient-specific cancer signaling networks. Analysis of such tumor signaling networks could help in making individualized risk predictions and treatment decisions. The cost of sequencing an entire human genome is rapidly falling. The continual development of faster and cheaper DNA sequencing technologies, (for example, the next generation of DNA sequencing, which aims to decode a human genome for $1,000) could provide the ability to identify cancer driver-mutating genes in individual patients. Furthermore, profiling of tumor gene expression is also accessible and affordable.

Because these data can be generated in a routine clinical manner, it is possible to adopt a systems biology strategy for medical research and finally move forward into the era of personalized medicine. For example, construction and analysis of patient-specific tumor signaling maps will allow for the identification of key protein communication modules that are critical for development of a specific tumor. Modeling and simulation of such a patient-specific tumor signaling map will help to infer the molecular mechanisms responsible for the cancer and will aid in pinpointing the key targets of the tumor. Furthermore, the use of computational modeling and simulation would lessen the risk of therapeutic failure at clinical stages. Therefore, it is predicted that network analysis and modeling will become a mainstream tool in both the pharmaceutical and the biotech industries (Figure 1).

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

Three major aspects of cancer biology are expected to benefit from the application of a systems biology approach: (1) identification of prognostic and drug-response biomarkers of tumors by using a systems approach to link genomic data and medical records, such as blood samples, lifestyle questionnaires, and patient survival (see Chapter 4; (2) an understanding of network-oriented molecular mechanisms by building networks and computational models of different stages of cancer progression; (3) and an understanding of the network-based molecular mechanisms of metastasis and improved treatment of the later stages of tumors by comparative analysis of the networks of primary and metastatic tumors (see Chapter 5). Finally, cancer systems biology could provide new insights into the network-based molecular mechanisms that cause certain drugs to fail, thereby helping in the selection of multiple anti-cancer drugs and optimization of treatment strategies.

2 Strategies for Cancer Systems Biology Study

Recent tumor genome sequencing efforts have shown that there may be thousands of cancer driver-mutating genes. Moreover, cancer driver-mutating genes are diverse and have little overlap between different tumors. This diversity is seen among different types of tumors and between tumors that originate from the same tissue. These observations suggest that cancer is a phenotype that can be caused by a collection of many genetic paths. However, several functional modules, the hallmarks of cancer (details of cancer hallmarks have been described in Chapter 12), have been uncovered and documented. In general, cancer driver-mutating genes reflect cancer hallmarks or functional modules. Each hallmark, or functional module, is composed of a set of functionally linked pathways. Therefore, it is possible to map the functional modules and the mutating genes onto network modules, each of which is a subnetwork that contains the functionally linked pathways on the network. For example, an integrative analysis of the human signaling network and cancer driver-mutating genes has revealed network modules of this type (details of such modules have been described in Chapter 5).

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

The systems approach to cancer studies must build realistic network models of tumors (network construction) and identify network modules, as well as the key genes and other network features in each module, from these networks (network analysis and modeling). Ultimately, the results derived from this system biology approach must be experimentally validated in cancer cell lines and mouse models. Through this approach, cancer systems biology enables the integration of biological and clinical data at various levels and has the potential to provide insight into this complex disease (Figure 1).

Network construction focuses on reconstructing functional networks that reflect the relationships between genes and proteins under specific conditions, such as cancer gene signaling networks in metastasis. These networks can encode the links between the -omic data and the fundamental processes of cancer development and metastasis, i.e., cancer hallmarks, cell cycle, apoptosis, and immunological response. Constructing a series of networks that incorporate time-course data may reveal the dynamics of biological processes such as tumor progression. Network approaches ease computational analysis, simplify and reduce complex interactions, and allow for the identification and quantification of relationships between inputs and outputs. Furthermore, network analysis aids in uncovering the general principles that underlie systems. To reach these goals, network construction relies heavily on integrative approaches to combining -omic data and accumulated knowledge.

Network visualization is the process of providing tools to build intuition that is unsurpassed by analysis tools. These intuitions may help in forming ideas regarding network exploration using analytical tools.

Network analysis focuses on computational analysis of the constructed networks using mathematical and statistical tools. Analysis may be performed on a single network to identify the important nodes, key network modules/subnetworks, and high-order relations between modules, such as collaboration, co-expression or co-regulation of modules. Furthermore, functional principles of cancer can be inferred from this type of

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

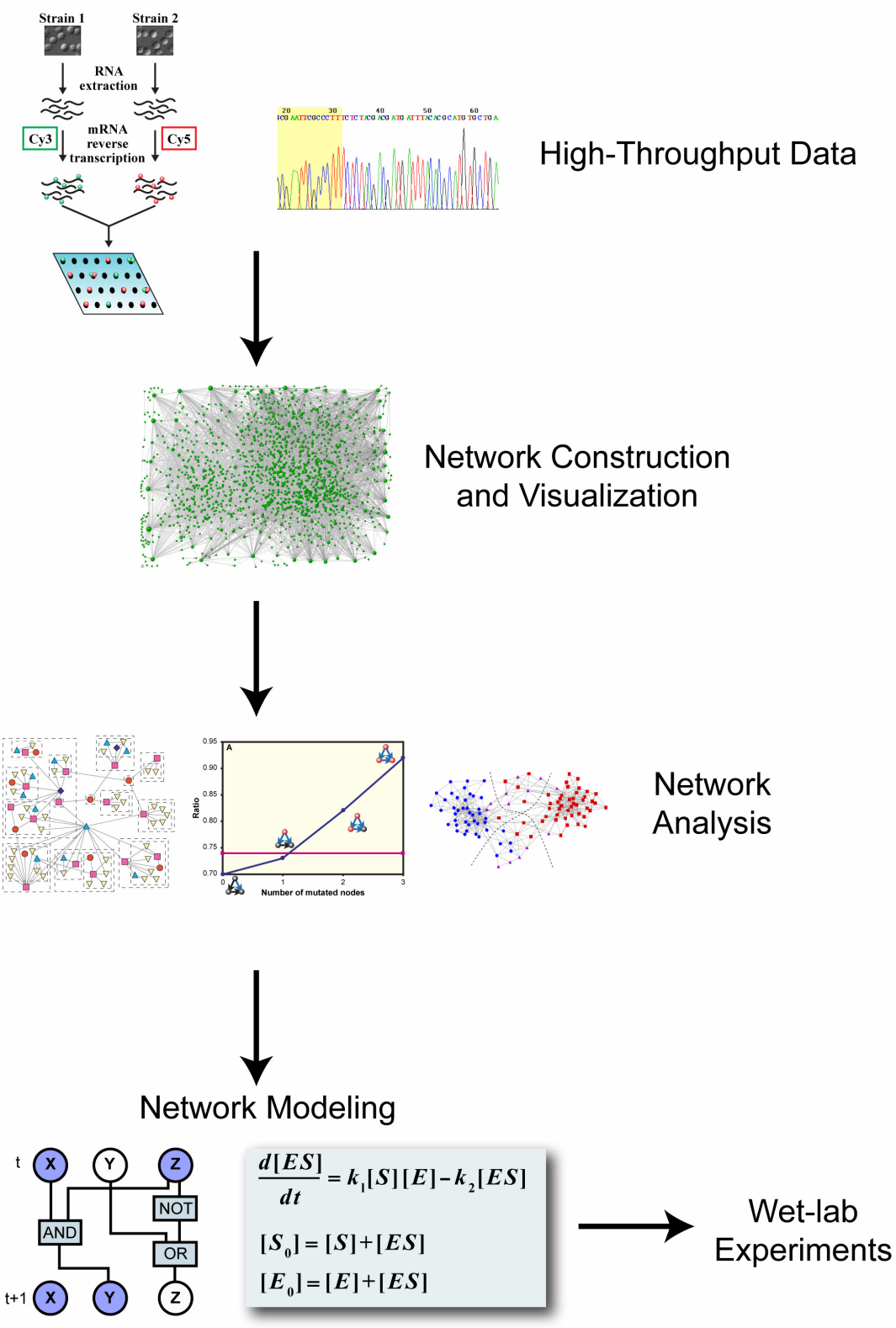

Figure 1. The strategy and procedures for cancer systems biology research.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

analysis. Therefore, hypotheses about mechanisms underlying cancer progression and metastasis can be generated through network analysis. Comparative analysis of time-course networks can highlight the dynamic nature of the functioning (or malfunctioning) of cells in the development and progression of diseases (i.e., cancer progression). Furthermore, comparative analysis can aid in the identification of key network components and their causal relationships during developmental stages. These analyses would capture the dynamic interactions between large numbers of components across different time scales, as well as the nonlinear nature of the systems. In addition, network analysis may lead to the identification of gene signatures that could be used for prognosis and drug response prediction through integration of gene expression or protein abundance profiles and clinical information about the cancer patients.

Network modeling involves the use of dynamic systems theory and mathematical tools to investigate complex biological systems in order to demonstrate nonlinear spatio-temporal behavior. However, the generation of experimental data that are suitable to parameterize, calibrate and validate such models is often time-consuming and expensive, or even impossible, with the technology available today. Regardless, the spatiotemporal dynamics of the system as a whole are of such complexity that understanding those dynamics challenges conventional approaches and makes mathematical modeling a necessity.

The behavior of complex cancer cell networks cannot be deduced by intuitive approaches. Instead, it requires sophisticated and elegant network models and computational analysis and simulation. Cancer cell network models will aid in the generation of experimentally testable hypotheses and discovery of the underlying mechanisms of tumorigenesis and metastasis. Network construction may provide insights into specifying the necessary components of a biological process, a subject that is highly related to the explicit hypotheses of cancer development, progression and metastasis. Both network analysis and modeling may improve our understanding of the cancer system and reveal hidden patterns or counter-intuitive mechanisms in cancer, uncovering critical points about which our understanding is still poor. Furthermore, both analysis and

modeling may help generate hypotheses that, in turn, can be tested in a wet lab. Finally, network analysis may help identify biomarkers useful in the clinical practice of personalized medicine (Figure 1).

2.1 Requirements for Experimental Models to Perform Cancer Systems Biology

Before applying systems biology approaches, we should focus on the types or subtypes of cancer that have high clinical relevance and are well studied in terms of molecular pathology (i.e., pathological features can be mapped onto gene signatures). Breast cancer is one example of this type of cancer. Gene microarray profiles can stratify breast cancer samples into four subtypes (Sotiriou and Pusztai, 2009). In fact, mammary epithelial cells of different origins can give rise to tumors with distinctly different phenotypes. Gene expression profiling of tumor types has shown that gene-expression changes depend on the nature of the precursor cells (Ince et al. 2007). Therefore, it is important to know the clinical subtypes of tumors when performing analyses.

For a particular type or subtype of cancer, it is essential to have high quality experimental mouse models (i.e., progression models for different cancer stages, for instance prostate cancer) and a set of targeted drugs available for treatment of the cancer type or subtype. In general, how closely the mouse models mimic human cancer types and subtypes will affect the usefulness of those models in understanding the molecular mechanisms and guiding therapeutic decisions. Therefore, it is desirable to have mouse models that mirror clinical outcomes in patients with specific types of cancer. Such models can be used to predict poor responses to chemotherapy in cancer patients and, therefore, might help in determining patient prognosis.

The selected cancer types or subtypes should have high quality cell line models that have been well characterized using genomic approaches. The breast cancer cell line MDA-MB-231 is a good example of one of these models. The MDA-MB-231 line has several derivative lines, which display the features of organ-specific (lung, bone or brain) metastases (Nguyen *et al.*, 2009).

For experimental models, large-scale -omic data, such as genome-wide gene expression, phosphoproteomics, epigenetics, cancer driver-mutating genes (i.e., via tumor genome sequencing), and metabolomics, can be generated. To produce the best data for

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

network construction, analysis and modeling, experimental biologists should interact with computational scientists to design experiments properly. Using this variety of data, construction, analysis, and modeling of tumor specific networks can be conducted. The networks can also be applied to mathematical frameworks for modeling and simulation.

2.2 Data Integration and Cancer Gene Network Construction

2.2.1 Cancer Gene Network Construction

The major objective of cancer systems biology is to create dynamic models of biological processes closely related to cancer initiation, progression and metastasis. Therefore, cancer networks should capture the important functional themes of cancer biology. Several fundamental biological processes play central roles in cancer. For example, the hallmarks of cancer are typical examples of these fundamental processes (detailed descriptions of cancer hallmarks can be read in Chapter 12). It should be noted that "new cancer hallmarks" might be added as understanding of cancer biology, especially cancer systems biology, increases. Indeed, some groups have recently proposed that inflammation may be a "new hallmark" for cancer (Mantovani, 2009). Cancer metastasis depends on both intrinsic properties of the tumor cells and factors in the tumor microenvironment. These factors provide tumors with blood vessels and an inflammatory environment, consisting of immune cells and their secretory products, which promote tumor growth (the tumor microenvironment and blood vessels have been described in detail in Chapters 14 and 15).

Cell cycle and division, differentiation, apoptosis, angiogenesis, insensitivity to inhibitors (robustness), tissue invasion and metastasis (i.e., EMT, epithelial-to-mesenchymal transition, more details about EMT have been described in Chapter 13), and inflammation are among the important properties of tumors. All of these processes are associated with cell signaling. Integrative analysis of the human signaling network with data about cancer driver-mutating genes suggests that cell cycle and apoptotic signaling are essential in all types of cancers (Cui *et al.*, 2007) (Chapter 5). Most of these biological themes should be captured in cancer gene networks. Furthermore, the subnetworks of these themes should be viewed as a priority for systems biology research. Separation of these subnetworks will allow different cellular processes to be studied in a

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

relatively isolated manner by network analysis and modeling. It should be noted that at least some of these processes are potentially inter-linked within the cell or the tumor microenvironment. Therefore, high-order relationships between these processes could be modeled after a careful and systematic study of each individual process.

There are three approaches for construction of cancer gene networks. The first approach is to infer or reverse engineer the cancer gene or signaling network using genome-wide datasets, such as gene expression profiles, RNAi knockout phenotype data, etc. For example, a gene regulatory network has been constructed using time course microarray profiles from a mouse epithelial breast cell line (BRI-JM01) (Wang et al. 2007), which undergoes an epithelial to mesenchymal transition (EMT) when treated with TGF-β (Lenferink *et al.*, 2004). Notably, clusterin, one of the genes that is up-regulated at the middle and late time-points, shows many regulatory links to other genes in the network. During the EMT process, clusterin is secreted by the BRI-JM01 cells. Interestingly, application of anti-clusterin antibodies to the TGF-β treated BRI-JM01 cells blocks the TGF-β induced EMT (Wang *et al.*, 2007). Chapter 3 describes additional computational methods and examples of reverse engineering of networks.

The second network construction approach is to extend a protein interaction or signaling network using high-throughput experimental approaches, such as protein interaction measurements (Chapter 19 describes an approach for extending the human signaling network). This extension method also allows for the construction of gene regulatory networks using large-scale ChIP-seq or ChIP-on-chip data. This approach works especially well for the construction of gene regulatory networks. For example, the application of this type of approach has constructed a P53 regulatory network containing 98 novel direct target genes of P53 (Wei *et al.*, 2006).

A cancer gene collaborative and mutually exclusive interaction network has been constructed by large-scale mutagenesis (i.e., retroviral insertional mutagenesis), to screen ARF- and P53-deficient as well as wild-type mice to identify genes that interact with one or the other of these tumor suppressors (Uren *et al.*, 2008). For this kind of network, it should be noted that cancer gene collaboration has two levels: (1) a gene tends to collaborate with another gene; or (2) one gene may have distinct sets of collaborators, based on different mutations of that gene (i.e., Notch1).

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

The third approach to network construction is to integrate data from high-throughput studies or manually curated literature databases onto current networks. For example, current signaling networks are largely constructed using manually curated data from the literature (Awan et al., 2007; Cui et al., 2007; Ma'ayan et al., 2005). This data integration approach is widely used in network construction because more high-throughput data are easily accessible. Integration of high-throughput data through computational approaches provides a powerful method to address and dissect the complexity of cancer at various levels in a systems manner. Quality of data is very important in the data integration approach. The decision regarding particular data sources used for network construction must be based on the questions to be addressed by the network analysis. A discussion of useful data sources for systems biology can be found in Chapters 16-18 and 20.

Directly mapping genes of interest (i.e., modulated genes between normal and cancer tissues) onto protein interaction networks and signaling networks will lead to the construction of cancer related subnetworks or network modules (i.e., modulated genes are connected together to form a subnetwork). Several such examples are described in Chapters 4-7.

Alternatively, we can map the genes of interest onto a network (i.e., a human signaling network) and extract all of the shortest paths between any two genes of interest. These shortest paths can then be merged to construct a network. The genes of interest used for this type of approach may be cancer-modulated genes, cancer driver-mutating genes or others groups.

We can examine which shortest paths of the entire network are important for a particular cellular condition using functional genomics data and collect these shortest paths to build networks. For example, we have constructed cancer cell line-specific signaling networks by collecting the shortest paths that are significantly enriched in the cell line gene expression profile of the human signaling network. Tumor gene co-expression networks can be constructed using different types or subtypes of tumors. Weighted gene coexpression networks have been constructed using gene microarray profiles from glioblastoma samples (Horvath *et al.*, 2006). Analysis of such networks provides a blueprint for leveraging genomic data to identify key control networks and

molecular targets in cancer.

Cancer molecular networks can also be constructed by linking the information between genotypes and phenotypes. For example, Quigley and colleagues crossed mice of two species, *Mus spretus* and *M. musculus*, which were either resistant (*Mus spretus*) or susceptible (*M. musculus*) to skin tumor development. Following the cross, they combined gene expression profiling with linkage analysis to construct a 'susceptibility network' of gene expression and regulation in the normal skin (Quigley *et al.*, 2009). This study highlights the power of a network approach for identifying genotype–phenotype relationships.

Cancer development and metastasis are dynamic processes with different timescales. Time series high-throughput data of cancer processes may be used to construct a series of networks that represent these scenarios, using the methods discussed above. To develop dynamic network models, we must clearly conceptualize the way time is encoded in networks. Analytically, time has two distinct forms: discrete and continuous. A discrete representation of time often consists of a series of snapshots of the network. Hence, longitudinal analysis focuses on the change from one network state to another. In such cases, a process is generally inferred from the total change in the network across time. A continuous representation of time consists of sequential events or interactions recorded with exact starting and ending times. Continuous representations of time enable the identification of overall network changes. However, most of the current experimental systems are only able to produce data for discrete representation of time in networks.

All of the approaches discussed can be used to build tumor sample or cell line specific cancer networks. Construction of such networks simplifies and reduces complex interactions. In other words, it removes the “noisy information” from the global network and assembles various parts that are highly related in tumors. After constructing the cancer networks, it is necessary to check which cancer hallmarks have been captured. If none of the hallmarks can be found in the networks, it is worthwhile to check the data and network construction procedures.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

2.2.2 Data Integration in Systems Biology Drives New Concepts for Bioinformatics Analysis

Bioinformatics provides essential tools for data integration for network construction, analysis and modeling. Many bioinformatics methods and tools have been developed for large-scale data analysis. These methods and tools include statistical tests (i.e., for testing gene expression differences, genetic associations and gene expression correlations), data extraction from literature and databases (i.e., text mining), and procedures for pattern recognition and machine learning (i.e., clustering analysis).

One trend in bioinformatics analysis is the movement away from consideration of genes and proteins in an isolated manner. In the early days of microarray analysis, genes studied in gene expression profiles were examined statistically in an individual manner. Currently, the statistical significance of gene expression changes is assessed in a gene set-dependent manner (i.e., taking into account pathway genes or a set of genes in a biological process such as cell cycle). Indeed, these methods have led to new insights into cancer biology. For example, alteration of the expression correlations of protein network modules seems to be involved in cancer metastasis. Many methods following this trend have been developed and used, for example the Gene Set Enrichment Analysis (GSEA). More details on these methods are described in Chapter 18. However, there are still many challenges in developing systems-oriented bioinformatics methods. For example, the problem of how to dissect the multivariable factors in systems, given the fact that biological variables are highly intertwined and correlated, remains to be solved.

2.3 Network Visualization

Networks represent complex systems. Although we have developed (and are continuing to develop) mathematical concepts and computational tools for network analysis (more details in next section), we are still unable to fully decode complex systems. In certain contexts, the human brain is still more powerful than an analytical method in forming intuitions that can help guide network analysis.

Network visualization allows us to exploit the human mind's capacity for building intuition that is unsurpassed by analysis tools. Before conducting network visualization, we must carefully formalize questions that can be used to extract meaning and

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

implication from cancer molecular networks.

The effectiveness of network visualization differs depending on network size. Thus, visualization has different objectives based on size. Visualization of small networks focuses on detailed elements of the graph structure, whereas that of larger networks mainly captures a more global picture. Visual analysis of small networks allows for the formation of insights into molecularly interacting relationships, whereas the analysis of larger networks allows for judgments about high-order relations (more abstractive) between subnetworks (i.e., interplay between network modules or components such as cell cycle and cell death in cancer progression). It is possible to define network modules and subnetworks in large networks. The intrinsic structure and relationships in each network module can be examined using network visualization tools. In turn, the nodes in a network module can be collapsed into token nodes and used to explore high-order relationships between modules.

Color-coding of nodes or arcs based on molecular function, information flows or other related features closely related to the problems addressed, often helps in discovering network patterns and forming hypotheses to guide in depth analysis.

There are increasing efforts to produce network visualization beyond 'static' representations of cellular states, toward a more dynamic model of cellular processes. These efforts strive to incorporate high-throughput and functional data, such as time-series gene expression data, Gene Ontology terms and subcellular localization data. In this context, dynamic network visualization helps in capturing the dynamic features of a process, augmenting theoretical intuition and extracting meaningful patterns.

Two basic approaches to visualization have been developed. The first common visualization approach encodes all changes and transitivity between developmental stages (time series data) into a single network. For example, a signaling network encodes all the modulated genes from a time course dataset of cancer progression with different colors (i.e., representing different stages in cancer progression). Such a network consists of patterns of causal or collaborative gene relationships. However, it is necessary to identify the stages that substantively capture the nature of the relational events and the character of temporary cellular states that arise in the focal context. Changes in transitivity provide information about a single dimension of a network's structure. One might find that a

network reaches a high or low transitivity level, suggesting the potential importance of some stages. The clusters and information flows obtained by mixing different colorful nodes and links might suggest new network modules. This approach is commonly used to evaluate cancer progression or identify casual relationships between network modules.

One of the most effective methods to implement dynamic network visualization is to present sparse networks in a way that shows how the network emerges over time (i.e., modulated genes in a time-series manner for cancer progression models) by adding and color-coding nodes and relationships as they appear (i.e., different colors can be used for nodes and relationships in a time-dependent manner). It is important to organize the nodes and edges in the display plane based on the final stage of the network. The appearance of dynamic elements over time reveals key genes that play roles in different stages and suggests regulatory relationships between network modules at different stages.

The second common approach to network visualization is to explore separate networks at each time point. However, these networks are often difficult to interpret using pure visualization, because it is impossible to identify the sequential links between node positions from one network to the next. In this situation, comparative network analysis is a proper and powerful approach to these problems (see Section 2.4).

A number of network visualization tools, such as Cytoscape and VisANT, have been developed. VisANT has developed high-level abstraction of network relationships (more details about VisANT are described in Chapter 17). Chapter 20 lists many network visualization and analysis tools for systems biology studies.

2.4 Network Analysis

Genes and proteins are often used as nodes in networks, while relationships between them are usually represented by edges (undirected links) or arcs (directed links). A detailed description of network types is provided in Chapter 2 and 4. Reading Chapter 2 prior to this section is recommended to provide an understanding of basic concepts and terms in network biology.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

2.4.1 Different Biological Properties Are Encoded in Different Network Types

Although a set of common questions can be asked of cancer molecular networks, different questions can be addressed using different types of molecular networks, in which different network characteristics and relationships are encoded. Moreover, different types of data are necessary to construct different types of networks.

In gene regulatory networks, the length of regulatory cascades is often short, normally 3-5 steps from the first layer to the last layer of the network (Wang and Purisima, 2005), reflecting the quick regulation response of these systems. Hubs in gene regulatory networks play a major role in responding to stimuli and coordinating the regulated genes. In agreement with this mechanism, the transcripts of the hub transcription factors often display the property of rapid decay (Wang and Purisima, 2005). Local transcription factors often encode genes that take part in one or a few biological processes, whereas intermediate hub transcription factors encode the collaborative relationships (i.e., co-expression) between a few biological processes. The rapidly decaying transcripts of global hub transcription factors might encode "switch" functions, which are used under different conditions and stimuli. Most of the target genes of transcription factors are "workers," which directly perform the tasks of biological processes and do not have regulatory roles. Collaborative relationships between transcription factors can be also found. Therefore, gene regulatory networks are useful for identifying key regulators, co-expression of genes, and sets of "workers" involved in cancer processes.

Nodes in the human signaling network are sparsely connected. The length of regulatory cascades in signaling networks (normally 7-14 steps from receptors to transcription factors) (Cui *et al.*, 2009) is often longer than the length of cascades in gene regulatory networks. Along these signaling cascades in protein signaling networks, almost all of the nodes are "regulators". Therefore, logical regulatory relationships are extensively encoded in signaling networks. Cell signaling information flow propagates from a receptor to the nucleus. It is believed that a number of proteins scattered directly downstream of receptors are logical "organizers" that integrate signals. For example, hubs in signaling networks play a major role in integrating different signals and

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

pathways. Therefore, signaling networks are useful in identifying cancer causal genes and regulatory logic involved in cancer processes.

In a signaling network, paths represent signaling information flow and regulatory logics. In a gene regulatory network, paths represent regulatory hierarchy. In contrast, paths in a protein interaction network have no clear biological implications. Network paths also have different evolutionary features (Cui *et al.*, 2009). For example, in the case of directed shortest paths, the more distance between two proteins, the less chance they share similar evolutionary rates. However, such a correlation was not observed with respect to the neutral shortest path. It has been shown that the evolutionary rate of proteins decreases along the signaling information flow from the extracellular space (input layer) to the intracellular space to the nucleus (output layer) (Cui *et al.*, 2009).

The expression levels of major regulators (i.e., kinases) in signaling networks do not necessarily change dramatically during cancer progression and metastasis. The major regulatory reactions are modulated via protein modification (i.e., phosphorylation and dephosphorylation), not via modulation of gene expression. Therefore, direct functional consequences of cancer driver-mutating genes are difficult to address in gene regulatory networks. Most of the cancer driver-mutating genes are signaling genes. Furthermore, these mutating genes do not simply increase expression levels of their targets, but increase or decrease the activity of their targets (Cui *et al.*, 2007). Monitoring the dynamics of phosphorylation and dephosphorylation is essential to decode signaling networks.

Network motifs in protein interaction networks represent protein complexes, whereas they represent information processing units and regulatory loops in signaling and gene regulatory networks. In protein interaction networks, network modules represent protein interaction communities associated with particular biological processes, whereas in signaling networks, they represent blocks of regulatory logics and information processing.

Compared to the human signaling network, nodes in the human protein interaction network are densely connected. Regulatory logics are difficult to identify in protein interaction networks. However, network modules or network communities are encoded in protein interaction networks. Therefore, such networks are suitable for integration of

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

gene expression profiles to determine subnetworks (teams of protein 'workers') that perform certain functions at different stages of cancer development, progression and metastasis. Furthermore, such subnetworks could be used as biomarkers in a clinical setting.

Generally, signaling networks are sparse and full of logical codes of regulation, whereas protein interaction networks are dense and do not code for logic of regulation. Gene regulatory networks encode both regulatory logic and gene "workers".

It should be noted that post-transcriptional and post-translational regulation are both prevalent in cells. It is important to consider these aspects in terms of network construction, analysis and modeling. Ubiquitination is applicable to a wide range of human proteins (Yen *et al.*, 2008) (see Chapter 6). In the human signaling network, ubiquitin-mediated regulation is enriched in receptors and ligands, the signal initiating portion of the network (Fu *et al.*, 2009). Initiated signals can be immediately organized and processed in the upstream region of the network, which resides in intracellular space close to the cell membrane. This network region is enriched for many built-in negative feedback loops (Legewie *et al.*, 2008). In contrast, MicroRNAs (miRNAs) regulation (negative regulation) focuses on the downstream regions of the network (Cui *et al.*, 2006).

Post- transcriptional and translational modifications of genes and their products provide feedback mechanisms in gene regulatory networks. miRNAs tend to post-transcriptional regulate transcription factors. Nearly half of the human transcription factors are regulated by miRNAs (Cui *et al.*, 2006). Furthermore, hub transcription factors tend to regulate more miRNAs (Chapter 7).

2.4.2 Network Analysis Using Network Biology Methods

Evolution is the central law of cancer cells. Similar to the laws of physics and chemistry, the design principles that constrain cancer biology are all amenable to discovery and modeling. "Core design principles" in biology must be modeled to express the mechanistic rules easily and efficiently. Abstraction is the most critical process required to uncover the design principles of biological systems. A proper abstraction aids in data examination from different perspectives and helps to extract meaningful knowledge from

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

the data. Graph theory allows for representation of the abstraction of biological relationships, analysis of the information, and extraction of insights. Evidence shows that biological insights have been encoded in network properties (Wang *et al.*, 2007). Therefore, network property analysis of integrated networks (i.e., signaling networks incorporated with cancer related high-throughput -omic data) will provide new biological insights.

The core concepts of network analysis are non-linear and network perspectives. Emergent biological properties may be discovered from non-linear thinking. The results of linear thinking are often predictable and expected, whereas the results of network analyses are often non-linear and unexpected. In theory, network analysis could lead to more unexpected and, therefore, exciting results.

Network properties range from local (i.e., single node or edge, network bottlenecks, network motifs and modules) to global or network-wide (i.e., whether all nodes are connected, network diameter, shortest path, density, average links, clustering coefficient, network centrality, degree centrality, closeness centrality, radiality, betweenness and pageRank, minimum spanning trees, and network flows). A detailed survey of network measurements and properties has been described by Costa et al. (http://arxiv.org/abs/cond-mat/0505185). Intrinsic relationships exist between local and global properties, such that sometimes a perturbation of a small number of linked nodes can result in widespread consequences. For example, a collection of protein network modules with gene co-expression alterations leads to breast cancer metastasis (Taylor *et al.*, 2009). Further details about network biology concepts such as graph theory, network measurements and analysis are described in Chapters 2 and 4. The terms and concepts in Chapters 2 and 4 will be used in the following examples to illustrate network analysis in cancer biology.

Based on the specific questions being addressed, different methods of network analysis can be applied. The architectural structure of cellular networks provides a framework to illustrate the logics and mechanisms of cancer biology. Regarding global network features, for example, the following cancer biology questions could be addressed: extracting a subnetwork of cancer signaling that reflects functionality of cancer development or metastasis; uncovering the mechanisms by which genetic and

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

epigenetic events affect cancer cell signaling and tumor progression; identifying central players (i.e., network hubs that are cancer genes) in cancer cellular networks; identifying subnetworks that are functionally targeted cancer hallmarks, such as cell cycle and apoptosis. These analyses allow for the narrowing down of the scale of these complex networks. Further, they capture the communications between the core molecular processes in cancer and uncover the molecular mechanisms responsible.

Many questions can also be addressed using local network features. For example, the enrichment of cancer driver-mutating genes (oncogenes) in positive feedback signaling loops (network motifs) suggests that oncogenes gain function due to mutation, whereas the enrichment of cancer methylated genes in negative regulatory signaling loops suggests that loss of function by gene methylation promotes tumorgenesis (Cui *et al.*, 2007). The cancer signaling network, extracted from the human signaling network by integrating cancer driver mutated genes and cancer methylated genes, has been decomposed into 12 modules (network modules or communities). Furthermore, high-order collaborative relationships between these modules have been identified in different types of cancers (Cui *et al.*, 2007).

Although cancer is considered a very heterogeneous disease, querying mutated genes in tumor samples using the network modules defined by a human cancer signaling map reveals that one common network module occurs in most tumor samples. Specifically, breast and lung cancers show more complex collaborative patterns of oncogenic signaling modules than the other cancer types examined, highlighting their heterogeneous nature (Cui *et al.*, 2007). These examples demonstrate that network biology is a powerful tool that elegantly provides new insights into biology. Moreover, most of these insights cannot be drawn from traditional biological approaches, which are dominated by linear thinking.

Network analysis also provides a powerful tool for generating testable hypotheses. For example, Fu et al. found that ubiquitin-mediated proteins are enriched in positive loops in the human signaling network. Gene Ontology enrichment analysis of the ubiquitin-mediated proteins in these positive loops suggests that the biological process apoptosis is enriched in this group of proteins. Furthermore, more than 85% of the ubiquitin-mediated-apoptotic proteins in these positive loops are cancer-associated genes.

These observations led to the hypothesis that the ubiquitination machinery, such as the 26S, could be more highly expressed in tumor cells than in normal cells (high expression of ubiquitination machinery genes will block apoptosis, an essential block in cancer signaling). Using microarray data from both tumor and normal samples, Fu et al. provide evidence that this hypothesis is true (Fu *et al.*, 2009). More examples of cancer network analyses are discussed extensively in Chapters 4-7.

2.4.3 Network Dynamics Analysis and Gene Markers for Diagnosis and Prognosis

Genetic variation and somatic mutations in human populations, and even in tumor samples from the same individual, make tumors a very heterogeneous tissue type. The heterogeneous nature of tumors leads to different responses from different patients with the same type of cancer to treatment with the same drug. To address this problem, personalized medicine proposes to identify molecular markers for drug responses.

Another issue in cancer treatment is how to predict which patients with cancer should receive extra therapy after surgery. Currently, most cancer patients undergo surgery. If physicians know which patients are in the earliest stages of cancer, they could better predict which patients might benefit from additional treatment. However, physicians cannot predict which patients with cancer should receive extra therapy after surgery. At present, it is very difficult to predict which patients will be cured by surgery alone, the single treatment most patients receive, and which patients might benefit from the addition of chemotherapy. Therefore, it is critical to identify those genes that can be used as tools to predict survival after diagnosis of cancer and those genes that can guide how oncologists should treat the cancer to obtain the best outcome.

Practically, we are facing the challenge of identifying robust and highly accurate molecular markers for drug response and survival prediction (prognosis). Enormous efforts have been made for more than 10 years to identify such biomarkers from gene microarray profiles. In fact, PubMed contains more than 3,000 publications on this subject. However, no robust biomarkers have yet been identified for cancer. Specifically, the current so-called breast cancer biomarkers are ineffective when used in a different set of breast tumor samples.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

It is expected that the use of a systems approach could extract more accurate and mechanism-based markers of patient response to drug treatment by capturing the system dynamics. This approach requires integration of cellular networks and the alterations of gene expression, genetic mutation, methylation, and protein modifications with clinical information such as drug response, patient outcomes, etc. These efforts will fundamentally change both the health care system and the management strategies for cancer patients.

Two recent studies demonstrated that the network biology approach offers promising results toward finding better markers for cancer prognosis. Mapping tumor-expressed genes onto a human protein interaction network allows for the identification of subnetworks as cancer biomarkers. The resulting subnetwork markers are more reproducible than individual marker genes, which are selected without protein interaction information. These subnetwork markers also achieve higher accuracy in classification of metastatic versus non-metastatic tumors (Chuang *et al.*, 2007). Alteration of the co-expression of genes that are organized as protein network modules is associated with cancer metastasis, suggesting that dynamic rewiring of protein interaction modules is implicated in metastasis. Based on this discovery, network module markers have been shown to reach higher survival prediction for breast cancer patients than other markers selected from gene clustering, an approach that does not consider genes as interacting modules (Taylor *et al.*, 2009). We hope that network biology approaches will be applied to the discovery of drug response markers in the future.

2.5 Dynamic Network Modeling

Biology is currently experiencing a high level of interest in developing an understanding of system dynamics, specifically in studying systems made up of communicating parts and machines, information processing (cell signaling), and interconnected computational and functional units. In this perspective, organisms are viewed as information manipulators and processors.

Network modeling uses methods from dynamical systems theory to model and simulate networks to decode the information processing machines and test hypotheses about the mechanisms that underlie the function of cancer cells. In network modeling, the

behavior of cancer cells is represented in terms of quantitative changes in the levels of gene transcripts or enzyme activities (i.e., kinase activity).

Network modeling provides conceptual and computational tools with which to perform and iterate dry-lab experiments. Network modeling enables simulation-based research within a quantitative reference framework that connects *in silico* replica and real systems by means of quantitative conceptual and computational tools. In the long term, network modeling approaches could replace many time intensive or expensive wet-lab experiments. In this context, the growing field of systems biology is expected to lead to fundamental breakthroughs in cancer biology.

Two basic approaches (qualitative and quantitative) to dynamic network modeling are often used. Qualitative network modeling considers the states (i.e., gene expression values, protein concentrations, active, or non-active) of the network nodes in a finite number of values (i.e., ON and OFF, higher or lower than threshold, etc), whereas quantitative network modeling considers the states of nodes over a wide range of values. In addition, quantitative modeling also considers probabilistic, deterministic or stochastic characteristics of the network.

Dynamic network models are composed of three basic elements: the cellular network (i.e., networks of gene regulation, protein interaction or signaling in a given cellular context), the initial state of each node, and the transfer functions that describe the state dependencies of each node in terms of its regulators. Node states can be modeled in either a continuous or a discrete manner, whereas the transfer functions can be modeled in either a deterministic or a stochastic fashion. Therefore, there are four methods (continuous deterministic model, continuous stochastic model, discrete deterministic models and discrete stochastic model) for dynamic network modeling. A discrete state approach may ‘precisely’ describe the system's behavior, whereas continuous state equations describe the "average behavior" of the system

In theory, a continuous stochastic model describes the system more accurately and more closely reflects the real system. However, high quality experimental data are required to apply this modeling method. The limited availability of high quality quantitative data forms a major bottleneck for the application of continuous stochastic models. Compared with the continuous stochastic model, a continuous deterministic

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

model describes the system without taking into account the stochastic (noise) nature of the system. A discrete stochastic model usually accounts for noise in the transfer functions (differential equations) and models the nodes with two or a few states.

A discrete deterministic model has a high level of abstraction of node states (i.e., node states are assigned into only a few categories, even two binary states such as ON and OFF). The transfer functions are encoded as logic functions, such as "and," "or," and "not". Boolean models are one representative group of the discrete deterministic modeling method. This method requires relatively little detailed input information. Therefore, this method is more attractive and feasible, because the data generated from current experimental systems are suitable for this modeling method. One of the disadvantages of this method is that the predictions are generally more macro-scale and less quantitative.

A comprehensive survey of the methods and computational tools for dynamic network modeling can be found in Chapter 16. An application of Boolean models to cancer cell death signaling networks has been documented in Chapter 8. Other modeling examples can be found in Chapters 9-11.

In summary, systems biology is leading to fundamental changes in cancer biology. In terms of dynamic network modeling, high quality quantitative data (i.e., quantitative proteomic data for signaling networks) are still requoired. Current methods for producing high quality data for modeling are still expensive and time consuming. Development of new methods for network analysis and modeling are also needed. Most importantly, more scientific investigators need to be trained to think in a network fashion, rather than in a traditional linear fashion.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

References

1. Awan, A., Bari, H., Yan, F., Mokin, S., Yang, S., Chowdhury, Q., Yu, Z., Purisima, E.O., and Wang, E. (2007). Regulatory network motifs and hotspots of cancer genes in a mammalian cellular signaling network. IET Syst Biol 1, 292-297.

2. Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D., and Ideker, T. (2007). Network-based classification of breast cancer metastasis. Mol. Syst. Biol. 3, 140.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

3. Cui, Q., Ma, Y., Jaramillo, M., Bari, H., Awan, A., Yang, S., Zhang, S., Liu, L., Lu, M., O'Connor-McCourt, M., Purisima, E.O., and Wang, E. (2007). A map of human cancer signaling. Mol. Syst. Biol. 3, 152.

4. Cui, Q., Purisima, E.O., and Wang, E. (2009). Protein evolution on a human signaling network. BMC Syst Biol 3, 21.

5. Cui, Q., Yu, Z., Purisima, E.O., and Wang, E. (2006). Principles of microRNA regulation of a human cellular signaling network. Mol. Syst. Biol. 2, 46.

6. Fu, C., Li, J., and Wang, E. (2009). Signaling network analysis of ubiquitin-mediated proteins suggests correlations between the 26S proteasome and tumor progression. Mol. Biosyst. 5, 1809-1816.

7. Horvath, S., Zhang, B., Carlson, M., Lu, K.V., Zhu, S., Felciano, R.M., Laurance, M.F., Zhao, W., Qi, S., Chen, Z., Lee, Y., Scheck, A.C., Liau, L.M., Wu, H., Geschwind, D.H., Febbo, P.G., Kornblum, H.I., Cloughesy, T.F., Nelson, S.F., and Mischel, P.S. (2006). Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. Proc. Natl. Acad. Sci. U. S. A 103, 17402-17407.

8. Kirk, R.J., Hung, J.L., Horner, S.R., and Perez, J.T. (2008). Implications of pharmacogenomics for drug development. Exp. Biol. Med. (Maywood.) 233, 1484-1497.

9. Legewie, S., Herzel, H., Westerhoff, H.V., and Bluthgen, N. (2008). Recurrent design patterns in the feedback regulation of the mammalian signalling network. Mol Syst Biol 4, 190.

10. Lenferink, A.E., Magoon, J., Cantin, C., and O'Connor-McCourt, M.D. (2004). Investigation of three new mouse mammary tumor cell lines as models for transforming growth factor (TGF)-beta and Neu pathway signaling studies: identification of a novel model for TGF-beta-induced epithelial-to-mesenchymal transition. Breast Cancer Res. 6, R514-R530.

11. Ma'ayan, A., Jenkins, S.L., Neves, S., Hasseldine, A., Grace, E., Dubin-Thaler, B., Eungdamrong, N.J., Weng, G., Ram, P.T., Rice, J.J., Kershenbaum, A., Stolovitzky, G.A., Blitzer, R.D., and Iyengar, R. (2005). Formation of regulatory patterns during signal propagation in a Mammalian cellular network. Science 309, 1078-1083.

12. Mantovani, A. (2009). Cancer: Inflaming metastasis. Nature 457, 36-37.

13. Nguyen, D.X., Bos, P.D., and Massague, J. (2009). Metastasis: from dissemination to organ-specific colonization. Nat. Rev. Cancer 9, 274-284.

14. Quigley, D.A., To, M.D., Perez-Losada, J., Pelorosso, F.G., Mao, J.H., Nagase, H., Ginzinger, D.G., and Balmain, A. (2009). Genetic architecture of mouse skin inflammation and tumour susceptibility. Nature 458, 505-508.

15. Sotiriou, C., and Pusztai, L. (2009). Gene-expression signatures in breast cancer. N. Engl. J. Med. 360, 790-800.

16. Taylor, I.W., Linding, R., Warde-Farley, D., Liu, Y., Pesquita, C., Faria, D., Bull, S., Pawson, T., Morris, Q., and Wrana, J.L. (2009). Dynamic modularity in protein interaction networks predicts breast cancer outcome. Nat. Biotechnol. 27, 199-204.

17. Uren, A.G., Kool, J., Matentzoglu, K., de, R.J., Mattison, J., van, U.M., Lagcher, W., Sie, D., Tanger, E., Cox, T., Reinders, M., Hubbard, T.J., Rogers, J., Jonkers, J., Wessels, L., Adams, D.J., van, L.M., and Berns, A. (2008). Large-scale mutagenesis in p19(ARF)- and p53-deficient mice identifies cancer genes and their collaborative networks. Cell 133, 727-741.

18. Wang, E., Lenferink, A., and O'Connor-McCourt, M. (2007). Cancer systems biology: exploring cancer-associated genes on cellular networks. Cell Mol Life Sci 64, 1752-1762.

19. Wang, E., and Purisima, E. (2005). Network motifs are enriched with transcription factors whose transcripts have short half-lives. Trends Genet 21, 492-495.

20. Wei, C.L., Wu, Q., Vega, V.B., Chiu, K.P., Ng, P., Zhang, T., Shahab, A., Yong, H.C., Fu, Y., Weng, Z., Liu, J., Zhao, X.D., Chew, J.L., Lee, Y.L., Kuznetsov, V.A., Sung, W.K., Miller, L.D., Lim, B., Liu, E.T., Yu, Q., Ng, H.H., and Ruan, Y. (2006). A global map of p53 transcription-factor binding sites in the human genome. Cell 124, 207-219.

21. Yen, H.C., Xu, Q., Chou, D.M., Zhao, Z., and Elledge, S.J. (2008). Global protein stability profiling in mammalian cells. Science 322, 918-923.

Nature Precedings : hdl:10101/npre.2010.4322.2 : Posted 2 Apr 2010

View publication stats