BMC Biology
incorporating **Journal of Biology**

COMMENTARY   **Open Access**

Network integration meets network dynamics

Teresa M Przytycka* and Yoo-Ah Kim

Abstract

Molecular interaction networks provide a window on the workings of the cell. However, combining various types of networks into one coherent large-scale dynamic model remains a formidable challenge. A recent paper in *BMC Systems Biology* describes a promising step in this direction.

New high-throughput experimental techniques, complemented by recently developed computational methods, have facilitated the initial reconstructions of large-scale cellular networks. These reconstructions provide important clues about the topological organization of these networks and elucidate relationships between the topological characteristics and biological properties of the corresponding molecules. In particular, studies of protein-interaction networks have revealed complex relationships between vertex degree (number of neighbors in the network), network modularity (organization of the network into connected subnetworks), gene essentiality, gene pleiotropy, and so on. Importantly, despite considerable noise in the data, the utility of these networks goes beyond merely describing the rough landscape of biomolecular systems. They are being used increasingly to predict functionality of individual molecules in the network, membership in protein complexes, association with signaling pathways, disease-associated subnetworks, and so on (see [1] and references therein).

Network dynamics

Experimentally and computationally derived networks, such as protein-protein interaction networks, regulatory networks or metabolic networks, provide static depictions of the dynamically changing cellular environment. Therefore, their utility for modeling cellular dynamics might not be clear. However, it is now increasingly recognized that static network topology can be used as a scaffold for studies of network dynamics. In fact, some

*Correspondence: przytyck@ncbi.nlm.nih.gov
National Center of Biotechnology Information, NLM, NIH, 8000 Rockville Pike, Bethesda, MD 20814, USA

dynamical properties can be uncovered from network topology alone, or in combination with other types of data, such as gene expression. For example, an analysis of network connectivity in terms of possible ways in which information can be propagated has been used to predict the molecules perturbed as a result of gene knockouts [1,2]. A more recent study combined protein-protein interactions, protein-DNA interactions, and phosphorylation networks with gene-expression profiles to provide a link between causative copy number variations (genetic perturbation) and molecular pathways affected in cancer [3].

Although the above-mentioned approaches provide tools for predicting which molecules are likely to be affected by a perturbation, their power to predict the changes quantitatively is extremely limited. Such quantitative predictions require knowledge of the parameters of a molecular system that goes beyond simple network connectivity. There are a number of well-established methods for quantitative modeling of dynamical systems (for a review see [4]). However, such approaches typically require knowledge of experimentally determined parameters describing the individual reactions. Consequently, these models have been developed and applied to small-scale networks only, limiting such quantitative studies to the better understood subnetworks for which such measurements can be obtained. Because such detailed data are not available on a genome-wide scale, a dynamical analysis of large-scale networks must rely on less precise methods that can estimate the behavior of the systems without knowledge of reaction parameters. For example, flux balance analysis (a mathematical approach for analyzing the flow of metabolites through a metabolic network) is often used in analyzing metabolic networks; variants of Boolean logic (a way of combining activation/inhibition signals from individual parts of a network) are frequently applied to signaling networks; and a variety of different methods have been proposed for regulatory networks (for a review see [5]).

Modeling dynamics in large-scale heterogeneous networks

Until recently, the large-scale modeling of network dynamics has been focused on individual network types. However, within a cell, all network types are interrelated

and dynamics of any individual network has an impact on the behavior of other networks. Several recent studies have begun to address the challenge of coupling large-scale dynamical models for different network types to obtain one consistent dynamical network. Such methods have been spearheaded by approaches to combine metabolic and regulatory networks (see [6-8] and references therein). For example, to obtain a combined model of metabolic and regulatory networks, Covert *et al.* [6] used flux-balance analysis to model the metabolic network component while the transcriptional regulatory network was modeled as a Boolean network. The genes in the transcriptional network were assigned Boolean (binary) values indicating whether or not a given gene is being expressed. An interactive procedure was applied to ensure that the combined model satisfies both the metabolic and the regulatory constraints. A subsequent study used mixed integer linear programming (a general optimization framework for capturing problems with both discrete and continuous variables) to couple such metabolic and regulatory models [8].

In their recent paper in *BMC Biology*, Wang and Chen [9] propose a promising approach for integrating transcription regulation and protein-protein interactions using dynamic gene-expression data. They start with candidate gene regulatory and signaling networks obtained from genome-scale data. These candidate networks are then pruned and combined, utilizing gene-expression data at multiple time points, to obtain an integrated and focused network under a specific condition of interest. The transcriptional network is modeled as a dynamical system in which the expression of a target gene (a gene subject to regulation by transcription factors included in the network) is computed as a function of regulatory impact of the corresponding transaction factors, its expression at a previous time point, and mRNA degradation rate. The modeling of a signaling/protein-interaction network takes into account, among other factors, the activities of its neighbors in the network. The interaction rate between two neighboring proteins is assumed to be proportional to the product of their concentrations. An overview of the method used by Wang and Chen [9] is depicted in Figure 1 and further details are given in Figure 2.

Wang and Chen applied their method to *Saccharomyces cerevisiae* (budding yeast) networks for three different stress responses - hyperosmotic stress, heat-shock stress and oxidative stress - and identified highly connected transcription factors and genes. Further analysis of the crosstalk between these three networks revealed the significance of some transcription factors in serving as the decision-making devices and in playing a role in rapid adaptation in the stress-response mechanism.

The authors also showed that their method can be used to predict gene-expression levels under different
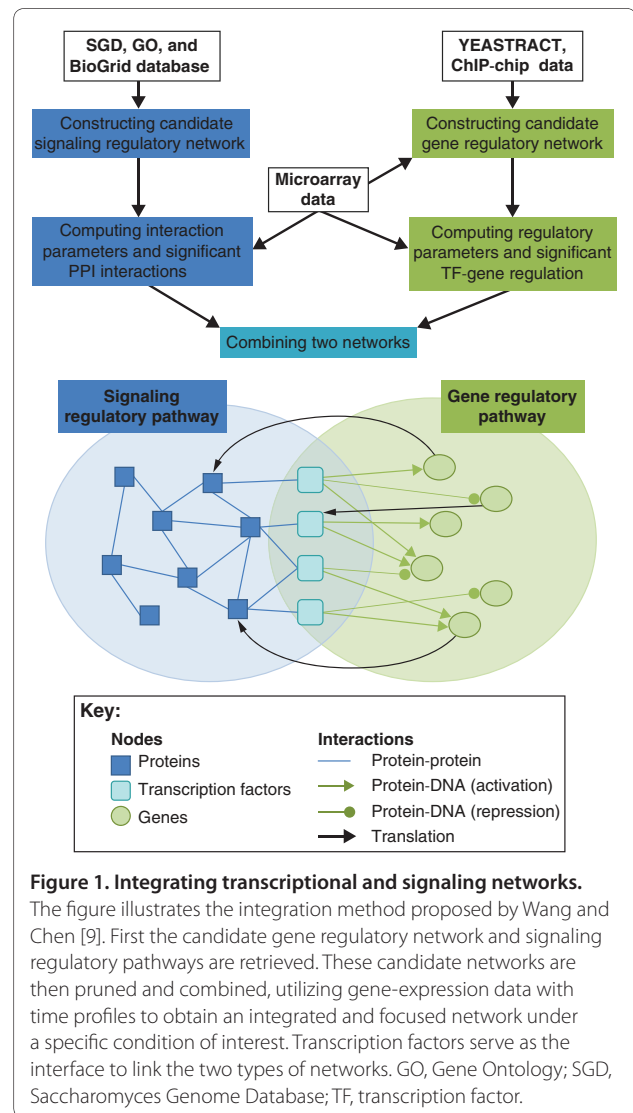

Figure 1. Integrating transcriptional and signaling networks. The figure illustrates the integration method proposed by Wang and Chen [9]. First the candidate gene regulatory network and signaling regulatory pathways are retrieved. These candidate networks are then pruned and combined, utilizing gene-expression data with time profiles to obtain an integrated and focused network under a specific condition of interest. Transcription factors serve as the interface to link the two types of networks. GO, Gene Ontology; SGD, Saccharomyces Genome Database; TF, transcription factor.

conditions. To do so, they first constructed the integrated network under heat-shock stress for the wild-type strain of yeast and then used the trained data to predict the expression level of the gene *HXT5* in the *yap1* mutant strain, which had been originally determined by Gasch *et al.* [10]. Their results suggest that various types of network models can be combined successfully to yield a predictive dynamic model of the heterogeneous system.

Challenges and future directions

Studies of large-scale biological networks are gradually shifting from the analysis of their organizational principles and guilt-by-association predictions of the function of individual network components towards examining cell dynamics. In such studies, experimentally determined static networks are often used as scaffolds for modeling of dynamical changes in the system.

The transcription regulation part of the integrated network is considered as a system where transcription factors (TFs) regulate target genes and the dynamic model of gene x is described by a stochastic equation as follows:

$$x_g[t] = B_g + (1-\lambda_g)x_g[t-1] + \sum_{i=1}^{N} R_{ig}z_i[t-1] + \varepsilon_g[t-1] \qquad (1)$$

where $x_g[t]$ is the mRNA expression level of gene g at time t. It is modeled as a function of the basal level B_g, the expression level at the prior time point $x_g[t-1]$, the regulation level $z_i[t-1]$ of TF i, and stochastic noise ε_g. R_{ig} represents the regulatory parameter of transcription factor i to target gene g, which indicates activation if positive and repression otherwise. λ_g denotes the degradation effects. The regulation level z_i is defined as the sigmoid function $f_i(y_i[t])$ of the protein activity y_i, which is described below to model signaling regulatory pathway.

For the protein-protein interactions, the dynamic model of the activity $y_n[t]$ for a target protein n is described by the following equation

$$y_n[t] = B'_n + (1-\lambda'_n)y_n[t-1] + \sum_{j=1}^{N} I_{nj}y_n[t-1]y_j[t-1] + T_n x_n[t-1] + \varepsilon'_n[t-1] \qquad (2)$$

where B'_n is the basal activity level of protein n, λ'_n indicates the degradation effect of the protein, and I_{nj} denotes the interaction parameter of protein n and j. In addition, the protein activity level is also affected by the expression level of the corresponding gene $x_n[t-1]$ and T_n represents translation effect from mRNA to protein. ε'_n denotes the stochastic noise.

Note that the two networks interact in two ways - in equation (1), the protein activities of TFs regulate their target genes and in equation (2), genes affect the activities of their corresponding proteins through translation effect. Regulatory parameters R and interaction parameters I are computed, based on gene expression data collected at multiple time points, by solving the constrained least square parameter estimation problem. Given the regulatory parameters, significant interactions and regulations in the integrated network are identified via Akaike Information Criterion (AIC, measure of goodness of fit of a statistical model) and the Student's t-test.

Figure 2. Details of the integration method proposed by Wang and Chen [9].

Information about dynamics can be provided, for example, by measurements of gene expression at different time points or in different conditions. The methods used by Wang and Chen [9] to construct such dynamically coupled models provide important steps in this direction, but there is still much more to be done. First, the power and limitations of the new methods need to be investigated more extensively. Next, note that the combined model of protein-interaction networks and transactional network proposed by Wang and Chen [9] incorporated the impact of protein degradation, which was not included in the basic flux-balance model. In contrast, the work of Covert *et al.* [6] more accurately captured the functionality of the transcriptional network through an application of Boolean logic. Obviously, both

approaches are important and should be considered in future models. Finally, an approach that combines all three networks is still lacking. Indeed, much has to be done before genome-scale models will be able to approximate cell dynamics with a precision close to what is expected from differential equation methods. But keeping in mind that a great deal was learned from the static models alone, we expect that even simple genome-wide scale dynamical models will bring further interesting discoveries.

Authors' information
TMP is an Investigator in the National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), National Institutes of Health (NIH), where she heads a research group focused on algorithmic and graph theory methods in computational and systems biology. YK is a research fellow in NCBI\NLM\NIH where she works on developing computational methods for the genome-wide, systems-level analysis of cell biology.

References
1. Przytycka TM, Singh M, Slonim DK: **Toward the dynamic interactome: it's about time.** *Brief Bioinform* 2010, **11**:15-29.
2. Missiuro PV, Liu K, Zou L, Ross BC, Zhao G, Liu JS, Ge H: **Information flow analysis of interactome networks.** *PLoS Comput Biol* 2009, **5**:e1000350.
3. Kim Y-A, Wuchty S, Przytycka TM: **Simultaneous identification of causal genes and dysregulated pathways in complex diseases.** *Lecture Notes Bioinformatics* 2010, in press.
4. Tyson JJ, Chen K, Novak B: **Network dynamics and cell physiology.** *Nat Rev Mol Cell Biol* 2001, **2**:908-916.
5. Karlebach G, Shamir R: **Modelling and analysis of gene regulatory networks.** *Nat Rev Mol Cell Biol* 2008, **9**:770-780.
6. Covert MW, Schilling CH, Palsson B: **Regulation of gene expression in flux balance models of metabolism.** *J Theor Biol* 2001, **213**:73-88.
7. Covert MW, Palsson BO: **Constraints-based models: regulation of gene expression reduces the steady-state solution space.** *J Theor Biol* 2003, **221**:309-325.
8. Shlomi T, Eisenberg Y, Sharan R, Ruppin E: **A genome-scale computational study of the interplay between transcriptional regulation and metabolism.** *Mol Syst Biol* 2007, **3**:101.
9. Wang YC, Chen BS: **Integrated cellular network of transcription regulations and protein-protein interactions.** *BMC Syst Biol*, **4**:20.
10. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO: **Genomic expression programs in the response of yeast cells to environmental changes.** *Mol Biol Cell* 2000, **11**:4241-4257.