Toward the dynamic interactome: it’s about time
Teresa M. Przytycka, Mona Singh and Donna K. Slonim
Submitted: 20th July 2009; Received (in revised form): 1st November 2009

Abstract
Dynamic molecular interactions play a central role in regulating the functioning of cells and organisms. The availability of experimentally determined large-scale cellular networks, along with other high-throughput experimental data sets that provide snapshots of biological systems at different times and conditions, is increasingly helpful in elucidating interaction dynamics. Here we review the beginnings of a new subfield within computational biology, one focused on the global inference and analysis of the dynamic interactome. This burgeoning research area, which entails a shift from static to dynamic network analysis, promises to be a major step forward in our ability to model and reason about cellular function and behavior.

Keywords: network analysis; network dynamics; interaction networks; systems biology

INTRODUCTION
Over the past decade, high-throughput experimental and computational methods have been developed to infer and predict the structure of gene and protein networks. As a result, large-scale cellular networks have been obtained for a wide range of organisms across the evolutionary spectrum [1–12]. Computation analyses of these networks have great potential in aiding our understanding of gene function, biological pathways and cellular organization. While significant progress has been made in computational analysis of proteome-scale cellular networks [13, 14], the dynamics inherent within these networks are often overlooked in computational network analysis. However, proper cellular functioning requires the precise coordination of a large number of events, and identifying the temporal and contextual signals underlying proposed interactions is a crucial part of understanding cellular function. Network dynamics can describe, for example, how cells respond to environmental cues or how an interaction network changes during development or differentiation.

We hold the view that modeling and analyzing interaction and network dynamics should be at the forefront of current research efforts in computational interactomics.

The first question, perhaps, is to define what we mean by interaction or network ‘dynamics’. In a simplified view, an interaction may occur or not depending upon spatial, temporal and/or contextual variation. Though temporal variation of interactions can be assessed on an evolutionary time scale, in this review we focus primarily on short-term interaction variation within a specific organism. Interactions may be constitutive or obligate, or may instead occur only in specific situations. Among these dynamically varying interactions (sometimes referred to as transient interactions), the variation may be either reactive (i.e. caused by exogenous factors, such as a response to some environmental stimulus) or programmed (i.e. due to endogenous signals, such as cell-cycle dynamics or developmental processes). Contextual variation overlaps heavily with temporal variation, but focuses more specifically on characterizing...
reactive variation and the conditions that cause it. Studying context may also encompass examining sequence or genetic variation within a population of contemporaries and exploring how that variation affects network interactions, topology and function.

Of course, biologists have studied dynamics in biological systems for many years. Traditionally, efforts have focused on individual genes or proteins as well as specific interactions in limited contexts. At a somewhat larger scale, metabolic networks, which inherently include temporal information and rely on the availability of specific metabolites as context, have also been mapped across numerous organisms [15]. However, within the last few years, we have been given an unprecedented opportunity to investigate molecular networks from a global perspective, using diverse high-throughput experimental data to inform our understanding of molecular interactions on a genomic scale. Our goal in this review is to focus specifically on research that links these global analyses to the dynamics and context of biological systems. We will highlight bioinformatics approaches that utilize genomic-scale experimental data sets for analyzing network dynamics rather than more established methods for quantitative modeling and simulation of dynamical systems [16–18] or their more recent extensions and implementations [19–26].

If dynamics play such a central role in biological systems, why have genome-scale computational analyses of interaction and network dynamics remained elusive? One central reason is that the most widely-applied large-scale technologies to determine protein interactions, such as yeast two-hybrid [27] and TAP-MS [28] to detect protein–protein interactions or in vitro proteome micro-arrays to detect phosphorylation interactions [12], do not provide spatial, temporal or contextual information about detected interactions. ChIP-chip [29] or ChIP-seq [30] approaches for uncovering regulatory interactions have been used to uncover reactive contextual variation [10] but have only recently started to be used to uncover temporal variation over a dynamic time course [31, 32]. Thus, even if all possible interactions in an organism could be determined using these technologies, for any given protein, it would generally not be known when and where each of its interactions occurs.

Computational approaches provide one means for inferring and analyzing changes in interaction and network dynamics. Broadly speaking, dynamically changing interactions can be thought of as a means for transmitting ‘information’ through cellular networks, and much of the work we discuss here attempts to characterize this ‘flow’ of information, often without an explicit understanding of the full dynamic nature of individual interactions. Specific research goals toward computationally elucidating interaction and network dynamics are diverse. To date, these have included discovering dynamic pathway information from network data, inferring network structure utilizing dynamic data, discerning large-scale changes in network topology and function, and understanding network responses to engineered perturbations (such as knock-outs and knock-downs) or to evolutionary perturbations (such as sequence variation between individuals). We have organized our review with respect to this seemingly broad set of research goals. Ultimately, however, all of the research described here is intended to improve our understanding of biological and physiological processes at the molecular level by attempting to characterize and analyze the dynamic interactome.

PATHWAY INFERENCE FROM PROTEIN INTERACTION NETWORKS

By itself, a complete cellular interaction map gives no information about the ordering of proteins with respect to their regulatory or physical relationships. One way of inferring dynamics from static network data is to use prior biological knowledge to help order the interactions. With sufficient information, one can construct a pattern representing some important biological structure and then search for instances of that pattern in static networks. Signaling pathways provide excellent opportunities for such approaches.

Signal transduction pathways transfer information detected at the cell surface to regulatory factors in the nucleus in order to orchestrate a cellular response. Steffen et al. [33] introduced a computational approach for discovering signaling pathways from protein–protein interaction data. Their approach is based on enumerating relatively short linear paths starting at membrane proteins and ending with DNA-binding proteins. These pathways are then evaluated with respect to gene expression data, with the expectation that proteins in the same pathway should be expressed in the same conditions and at
approximately the same time. Interaction reliabilities can also be incorporated in this approach, along with improved algorithms for path enumeration \[34\]. While it is convenient to think of one signaling pathway at a time, numerous signals are detected essentially simultaneously by various sensory proteins at the cell surface. Supper et al. \[35\] suggest an approach for inferring signaling pathways based on an arbitrary number of sensor and regulatory proteins, and use Steiner tree formulations that favor bow tie architectures \[36\] corresponding to intermediate ‘integrator’ core proteins.

Signaling and regulatory pathways, as well as other complex network patterns, may also be described with respect to the properties of their constituent proteins—for example, kinases, transcription factors or descriptions of constituent interaction domains. Banks et al. \[37\] introduced network schemas to describe such patterns and developed fast algorithms for searching for their matches in interactomes. A network schema consists of descriptions of proteins (e.g. their molecular functions or putative domains) along with the desired topology and types of interactions (e.g. physical, phosphorylation or regulatory). Network schemas can thus describe the linear patterns underlying signaling pathways suggested in Steffen et al. as well as more complex pathways. In addition to searching for matches to particular network schemas, it is also possible to infer which network schemas are frequent and over-represented in networks \[38\] and to thereby uncover general recurring patterns underlying a range of biological processes.

While the above-mentioned approaches consider signaling pathways in terms of orderings of individual molecules, Zotenko et al. \[39\] proposed an approach that focuses on ordering overlapping groups of molecules, putatively corresponding to transient complexes or functional groups. They approximate signaling networks by chordal graphs \[40\] in which functional groups, corresponding to dense subgraphs in the original network, are represented as cliques. Then they use clique tree representations of chordal graphs \[41\] to elucidate (partial) orderings within these functional groups.

Comparative interactomics is an alternative approach for discovering pathways within cellular networks. The core idea is that if a pathway is known in one organism, searching for homologs of its component proteins, along with conserved patterns of interactions, is a powerful technique for pathway detection \[42\]. Much research in recent years has focused on such pathway querying as well as the more general task of uncovering conserved pathways and modules via network alignment \[42–47\].

**PATHWAY INFERENCE BY INTEGRATING DYNAMIC DATA WITH PROTEIN INTERACTION NETWORKS**

Protein–protein physical and regulatory interaction networks provide a ‘scaffold’ by which to interpret other types of dynamic data and thereby obtain hints about signaling and regulatory pathways. Since proteins along a biological pathway tend to be co-expressed, one approach for identifying pathways is to identify active subnetworks, connected regions in physical interaction networks that exhibit significant expression changes over some set of conditions \[48\]. These attempts to characterize contextual variation in networks have been further extended and improved \[49–51\]. Such context-specific subnetworks do not have temporal orderings between proteins; combining these subnetworks with the pathway-ordering approaches (such as those reviewed above) may be a promising approach for refining the flow of information through the network.

Expression data can also be utilized to infer causality as well as information flow within cellular networks. A particularly illuminating source of dynamic data comes from knock-out experiments, where a gene is perturbed or removed from a genetic background and the expression levels of all other genes are measured. Yeang et al. \[52\] develop a probabilistic approach for explaining observed gene expression changes due to a knock-out by inferring molecular cascades of ‘flow’ through the interaction network. These molecular cascades correspond to paths beginning from the knock-out gene (‘cause’) and ending at the gene whose expression has changed (‘effect’). Additionally, interactions along the path are inferred to be either activating or repressing, and together must be consistent with the sign of the knock-out effect. The basic approach was later refined and experimentally validated in yeast \[53\]. Furthermore, it is possible to change the formulation of the problem so that it favors explaining many cause–effect pairs by few pathways \[54\].

RNA interference (RNAi) screens are a powerful technique for obtaining perturbation data in higher
organisms. Here, a known gene from a pathway of interest is chosen as a reporter gene, other genes in the genome are systematically knocked-down using RNAi, and the effect on the reporter is measured. Flow-based methods through protein interaction networks have been utilized to connect and order genes that affect the reporter [55]. Alternatively, genes that affect the reporter gene, or some other experimentally measurable phenotype of interest, can be used as ‘seeds’ within known protein interaction networks to uncover subnetworks related to the biological process of interest [56]. Methodologically, this approach is similar to the active subnetwork approaches outlined above. As RNA-based loss of function screens are increasingly being applied with automated image analysis to detect effects on specific processes or phenotypes [57, 58], these types of network analyses are likely to be relevant to study a broad range of interesting biological questions. Future challenges for all these types of approaches include incorporating additional types of data into these models (e.g. known annotations of individual proteins or other types of interaction data).

INFERRING NETWORK STRUCTURE FROM DYNAMIC DATA

The earliest attempts to incorporate dynamic data into models of gene or protein interaction networks were aimed primarily at inferring the network structure itself. (See [48, 59–63] for a general introduction to this area.) Most of these early network inference papers do not model dynamics. However, many used dynamic data, most commonly expression data from genome-wide microarrays, to infer static network structure. Such work includes efforts at reconstructing regulatory networks using mutual-information measures [64, 65] or Bayesian/graphical modeling methods [61, 66, 67]. Though reverse-engineering efforts have more frequently focused on inferring regulatory networks, Bayesian networks have been used to infer human T-cell signaling networks from single-cell measurements of phosphorylation in response to perturbations [68]. Parameterized biochemical modeling of responses, as measured by microarrays, to RNA interference and overexpression screens have also been utilized to infer signaling pathways in Drosophila [69].

Friedman et al. [70] and Murphy and Mian [71] were among the first to suggest using Dynamic Bayesian Networks (DBNs) for modeling regulatory networks because such models can capture time-dependent structures (such as feedback loops) impossible to express with traditional probabilistic networks. Ong [72] and Kim [73] also explored DBN approaches. However, the aim was typically to infer a static network structure from dynamic data, rather than using the dynamic model to assess the context of specific interactions.

Coexpression data can also be used to construct networks relating genes by the similarity of their expression patterns under different conditions [74, 75]. Coexpression networks have been applied to finding functional modules in expression data [75, 76]. While typically intended to reveal protein function, such networks are directly related to the context of the samples used in their construction. Thus, this information can be exploited explicitly to reveal context-specific patterns in expression data. Choi et al. [77] constructed two distinct coexpression networks using data sets representing different conditions (tumors and normal control samples) and identified context-specific differences by comparing the resulting ‘static’ networks. Kostka and Spang [78] proposed an approach for finding sets of genes whose coexpression patterns differed in different contexts. Fang et al. [79] have extended this idea to identify ‘subspace’ patterns of differential coexpression (subsets of samples and genes where the genes are coexpressed in most samples from one class but not coexpressed in most samples from the other class). Recent work has also used dynamic modeling to identify transient coexpression relationships capturing static but temporally related snapshots of the dynamic network [80, 81].

By exploiting the modular nature of cellular function, it is possible to infer context-specific regulators from expression data. Groups of genes that are coexpressed in a subset of a conditions are first identified by biclustering approaches [82, 83], and then their condition-specific regulators can be inferred by either relating the expression levels of the genes comprising the module to those of its regulators [84, 85] or by integrating computational or experimental information about regulatory interactions [86]. De Bivort et al. take a coarser-grain view of regulatory networks by considering regulatory relationships between the modules themselves and relating expression levels. For one module to expression levels of another module after a time delay [87].

Integration of additional data types has improved the power of network inference methods [88].
Integration of static network and dynamic experimental data was essential for the work of Bromberg et al. [89]. Using a mouse network assembled from multiple sources, they predicted a role for the breast cancer protein BRCA1 in neuronal differentiation and suggested a new signaling pathway. Furthermore, by combining this data with transcription factor activation experiments, they have been able to propose a Boolean logic-type description of the neurite outgrowth network. Similarly, Baugh et al. [90] integrated temporal and spatial variation to identify a developmental regulatory network in a Caenorhabditis elegans model.

The integration of expression data and ChIP-chip data is one particularly promising approach for elucidating dynamic control in regulatory networks. Recently Seok et al. used Network Component Analysis (NCA) [91] to build a series of time-varying transcriptional networks for lipopolysaccharide response in humans [92]. NCA is a computational approach that can be used to predict dynamic transcription factor activity over time. It utilizes (partial) knowledge of a regulatory network as well as gene expression data, and its goal is to infer both the concentrations of active transcription factors as well as their promoter affinities. The dynamic networks built by Seok et al. captured changes in transcription factor activities and gene expression levels, as well as in signaling and regulatory interactions. Ye et al. [93] further extended NCA by developing a computational approach to model the impact of single nucleotide polymorphisms (SNPs) on the concentrations of transcription factors and their promoter affinities. This analysis helps explain how genetic variations linked to gene expression via eQTL analysis (described later in this review) might perturb the regulatory network.

An interesting approach for integrating expression and ChIP-chip data was taken by Ernst et al. [94]. They exploited time-course expression experiments to build a detailed dynamic regulatory map for yeast stress response by incorporating a hidden Markov model to identify dynamic ‘bifurcation points’ where expression of a subset of genes diverges from that of others. Recent work by Ucar et al. [95] further integrates context-specific ChIP-chip and expression data with context-invariant sequence and nucleosome occupancy data to predict which binding relationships are functional and to characterize the functional roles played by individual transcription factors.

**UNCOVERING ORGANIZATIONAL PRINCIPLES UNDERLYING NETWORK DYNAMICS**

Combining static regulatory or protein–protein interactions with dynamic data can lead to a better understanding of protein or gene function and can reveal global changes in network topology that hint at higher level cellular organizational principles and functions. The goals of such work often involve finding key regulators of transcription or of dynamic cellular responses in specific conditions. For example, by integrating transcriptional regulatory networks with gene expression data from different conditions, Luscombe et al. offer a functional characterization of yeast hub proteins as either ‘transient’ or constitutively active (‘permanent’) [96]. Among the transient hubs, they further distinguish between those responding to programmed changes (such as sporulation or the cell cycle), which tend to involve slower, multi-level response cascades, and those regulating reactive changes to environmental stimuli (such as DNA damage or stress response), which cause rapid activation of many targets.

A different approach to dynamic modeling of transcriptional networks is taken by Chechik et al. [97]. Their work focuses on identifying network motifs [98, 99], over-represented substructures (such as feed-forward loops) within the regulatory network that are thought to work together to perform specific functions. The basic approach is not unlike that of some of the pathway-discovery methods described above. Possible ‘timing activity motifs’ (a motif structure and a ‘wiring pattern’ indicating the order of expression onset of the associated elements) are enumerated, and motifs over-represented in a large collection of gene expression time series data sets are identified. Inferred response times are derived from the ‘impulse model’, an idealized model representing a common pattern of gene expression state changes [100]. Analysis of the resulting motifs results in identifying bifurcation points as well as feed-forward and ‘backward activation’ patterns thought to quickly regulate cellular response to changing environmental conditions.

Studies of network motifs can also be informative about the dynamics of signal propagation [101]. In the context of regulatory networks, network motifs provide important clues to how signals might be propagated and regulated [102]. Recently, Ma’ayan et al. [103] showed, in the context of hippocampal CA1 neurons, that similar motifs are
also formed as information flows through a signaling network. Interestingly, they observed that ligands causing rapid, transient changes are typically characterized by different motifs than ligands facilitating permanent changes.

The relationship between a position of a gene in a network and its biological properties has been also explored in the context of transcriptional networks. An interesting property of transcriptional networks is their semi–hierarchical structure [104–107]. Therefore, one might ask whether the position of a gene in such structured networks is related to its biological properties. To answer this question, Jothi et al. divided the genes in the yeast regulatory network into three levels: ‘top’, ‘core’ and ‘bottom’ [107]. They demonstrated that genes within a network level have similar transcriptional half-lives, abundances and noise levels, but that these properties differ between levels. Elucidating such relationships between network structure and biological properties might provide new insights toward understanding cell dynamics.

Analogous to Luscombe et al.’s characterization of transcriptional regulatory networks [96], work by de Lichtenberg et al. integrates dynamic data and static protein–protein interaction networks for the purpose of characterizing protein complexes as either transient or constitutively expressed [108]. They find that most complexes consist of both types of proteins, and they suggest that the periodically expressed components control complex activity via a just-in-time assembly mechanism. In contrast, Komurov and White look at related data for condition-dependent and condition-independent (constitutive) expression patterns. They conclude that functional modules (not necessarily complexes) are predominantly comprised either of constitutive or of dynamic proteins, but rarely of both [109]. Differences in the data sources and the types of modules considered are likely to explain the different conclusions of these two papers.

Han et al. characterized hubs in the yeast protein–protein interaction network into two groups called ‘party’ and ‘date’ hubs [110]. The expression patterns of party hubs’ interaction partners have a high average correlation with the expression pattern of the hub itself, suggesting that all of the interactions may take place simultaneously or under similar conditions. Date hubs, which have a lower average expression correlation with their neighbors, are thought to interact with different partners under different conditions. The authors suggest that the date hubs are more likely to be global regulators linking lower-level functional modules comprised of party hubs and their neighbors. Importantly, there are significantly more date hubs in the binary network, whereas party hubs are prevalent in the co-complex networks [111]. There is also some debate about how strongly the party-date analysis depends on the exact expression data sets used [112, 113], suggesting that using additional information to identify static or dynamic hubs may be desirable.

Schmidt and McMahon emphasize that network-related concepts, such as the property of being a hub protein, must be placed in a functional and temporal context [114]. In their study, which focuses on the Clathrin-mediated endocytosis (CME) pathway, they demonstrate that hub connectivity changes over different stages of vesicle formation and suggest that the AP2 hub can only function as a hub when many AP2 molecules form a ‘hub assembly zone’. Arguably, hubs play important roles in network modularity and dynamics [101, 114].

**RELATING THE EFFECT OF GENE KNOCKOUTS TO NETWORK TOPOLOGY**

Are static, topological properties of a network related to the dynamic behavior of the cell in response to perturbations? Jeong et al. observed that high-degree nodes in a protein interaction network tend to correspond to proteins essential for the survival of yeast cells in optimum conditions [115]. This is tested by knocking down the corresponding genes and testing the viability of resulting mutant. Subsequently, the relationship between protein essentiality and global and local topological features of the protein interaction network has been broadly investigated by many authors [116–122]. To answer the question of why such hubs are essential, Jeong and colleagues [115] suggested that over-representation of essential proteins among high-degree nodes can be attributed to the central role hubs play in mediating interactions among numerous, less connected proteins. In contrast, He and Zhang proposed that the majority of proteins are essential due to their involvement in one or more essential protein–protein interactions that are distributed uniformly at random among the network edges [120].
If either of the above two views was confirmed, essential hubs would have played a very special role in network dynamics. However, a recent study of Zotenko et al. [122] rejected both of these explanations. Instead, they proposed that the majority of hubs are essential due to their involvement in Essential Complex Biological Modules (ECOBIMs), a group of densely connected proteins with shared biological function that are enriched in essential proteins. This and related studies [123–126] suggest that phenotypic traits such as essentiality have a modular nature. In particular, Zotenko et al. showed that, for most data sets and a number of centrality measures, after correcting for correlation between degree and essentiality, essentiality typically correlates with local but not global centrality measures. Subsequent studies of predicted protein complexes not only confirmed the relationship between essentiality and participation in complexes but also pointed to the relationship between the size of the complex and the essentiality of its members [127]. Interestingly, no relation between essentiality and the number of complexes that include a given gene has been found [126]. However, organizational properties may be related to individual transcript dynamics. While Janga and Babu observed no correlation between essentiality and protein half-life, they did find a correlation between centrality and mRNA half-life [128].

The findings of Zotenko et al. explain why only a very weak centrality–lethality relationship [119, 122] was observed in the network resulting from the yeast two-hybrid screen of Ito et al. [129]. In addition, a recently constructed, high-quality yeast two-hybrid map showed highly significant clustering between essential proteins [111] but no significant correlation between essentiality and vertex degree. Indeed, one should keep in mind that the yeast two-hybrid assay reports binary interactions only (i.e. direct physical contact as opposed to membership in the same complex), so the size of a complex to which a given protein belongs has relatively low impact on the degree of the corresponding node. Furthermore, large complexes are not always detected by the yeast two-hybrid technique. If the observed relationship between vertex degree and lethality was mostly due to membership in large essential complexes, it is not surprising that the relationship is much weaker in a network comprised only of yeast two-hybrid data.

In contrast to the weak correlation between essentiality and vertex degree, Yu et al. [121] observed that in their large yeast two-hybrid protein interaction network, vertex degree correlates with pleiotropy, the number of phenotypes observed as a consequence of a gene knock-out. Interestingly, such a correlation between the degree and number of phenotypes is not observed in networks constructed from co-purification experiments. These results underscore the importance of considering the experimental technique used to obtain an interaction network when interpreting a network’s dynamic properties. Further exploring the relationship between topology and dynamic function, Missiuro et al. [116] assess the total information flow through each protein in a network and demonstrate that this information flow score correlates with both essentiality and pleiotropy.

## RELATING DOUBLE MUTANT PERTURBATIONS WITH THE PHYSICAL INTERACTOME

Different types of experimental assays may reflect different types of dynamic information. For example, genetic perturbation phenotypes may be inherently condition-specific (e.g. temperature-sensitive lethal mutations), whereas yeast two-hybrid systems create an artificial context to assess the possibility of physical interactions between proteins that may never even appear in proximity to each other in vivo. Integrating multiple types of interaction data may help overcome the contextual biases or omissions of particular methods and provide more reliable temporal and contextual information about molecular relationships.

Genetic (or epistatic) interactions occur when the phenotypic impact of one gene is modified by another gene or set of genes. In this review, we will touch upon only symmetric binary interactions, such as synthetic lethals, identified by observing that the composite effect of the simultaneous mutation of two genes has a different phenotypic effect than expected from modeling the combined independent effect of such mutations. Genetic interactions are one means for revealing active information flow in cellular networks. Central methods for uncovering genetic interactions include screens for synthetic lethality and specialized expression Quantitative Trait Loci (eQTL) analysis aimed at identifying genes regulated by several loci. However, one should keep in mind that while synthetic lethal interactions are always between genes, eQTL-based analysis uncovers interacting genetic loci which may...
include non-coding regions of the genome. In this section, we survey recent results related to synthetic lethal interactions, and we defer the discussion of interactions between loci to the following section.

A synthetic lethal (or sick) relation between two non-essential genes occurs when simultaneous knockout of the two genes results in cell death (or aggravating genetic impact on cell fitness). Genetic interactions can provide valuable information about protein function and the buffering properties of cellular networks [130]. Considering a hybrid network of genetic and physical interactions provides additional insights into protein function, modular organization of cellular processes and clues as to how these two networks complement each other to achieve the desired robustness to perturbation or response to changing conditions [125, 131–137].

In particular, it has been noted that pairwise interactions between genes often reflect larger relations between whole sets of genes. To capture such dependencies, Kelley and Ideker [131] introduced the between pathway model (BPM). This model consists of two sets of proteins that have many genetic interactions between the two sets, but few within each set, and conversely many physical interactions within each set but few between them. Such a pattern of synthetic–lethal interactions can be explained by a model in which the proteins within each of the two sets function as a pathway (or a protein complex), working together to achieve some goal, and where one set can functionally compensate for the other. Recent work [138] incorporates dynamic information from gene expression data to evaluate the functional coherence and compensation of putative BPMs, focusing attention on the strongest candidates with evident compensatory roles. By integrating temporal data with these topologically derived models, their approach provides additional functional information about the relevant gene sets and the relationships between them.

INTERPERSONAL GENETIC VARIATION AS PERTURBATIONS OF CELLULAR SYSTEMS

Natural genetic variations within and between species can also be viewed as genome-wide ‘perturbations’. These perturbations can occur anywhere in the genomic sequence, not necessarily just within genes. Correlation of genotypic variations with observed phenotypes can provide clues about possible causal relationships between genetic mutations and the corresponding phenotypic variations. Many recent association results utilize whole genome gene expression data as the phenotype, and thereby such eQTL analysis allows observing the most direct consequences of genetic polymorphism, namely its effects on gene expression. Since the pioneering work of Jansen and Nap [139], eQTL analysis has been applied to a large number of species including yeast [140–143], mouse [143, 144] and human [145–147]. With eQTL analysis, it is possible to elucidate several properties of gene regulation including putative causative relationships between expression variation and genetic loci [148–151], co-regulated gene modules [140, 150, 151–156] and whole regulatory networks [157–161]. Integrative eQTL studies are widely used to identify disease causing mutations [162].

It is important to keep in mind that the large number of gene expression traits and genomic loci poses challenges for both computational efficiency and statistical power. Traditionally, an eQTL study tests the linkage between all genes’ expression and all loci, adding up to millions of individual statistical tests. This leads to unique challenges related to multiple testing adjustment [163]. Recently, Yang et al. proposed a graph-theoretical method for performing an eQTL analysis [164] that is able to uncover associations missed by traditional approaches. More work is still needed in developing methods with the statistical power to identify such context-dependent relationships, particularly those that involve multiple genes simultaneously.

To strengthen the signal and to explore additional dependencies between genes and their expression, several of the above-mentioned methods integrate additional information, including protein–protein interaction data, transcription factor-binding sites, co-expression and sequence conservation [154, 155, 161, 165]. Some utilize variants of information flow approaches to uncover pathways connecting putative causal genes and their targets [149,151].

As mentioned above, simultaneous knockdowns of pairs of non-essential genes allow discovery of genetic interactions between genes. Similarly, gene expression variation analyzed in the context of natural genetic variation present in the population is being increasingly used to discover interactions between loci. Epistatic interactions are typically revealed by testing whether the expression of a gene is better explained by a model of genetic variation at one locus or at multiple loci. Epistatic
interactions appear to be persistent among expression trait loci. Brem et al. [141] estimated that 16% of heritable transcripts in yeast exhibit epistasis. From a statistical and computational perspective, discovering epistatic interactions on a genome-wide scale is even more challenging than discovering individual eQTLs. Thus, despite significant progress [166], finding epistatic interactions on a genome-wide scale remains a major challenge. Currently, little is known about the dynamics of such interactions. While some work has been done toward identifying gene–environment or gene–context interactions involving individual genes [167], future work relating epistatic genetic relationships with response to exogenous stimuli such as environmental stressors is essential.

CONCLUSIONS

Since the dynamics of protein interactions play a central role in regulating the functioning of cells and organisms, and since high-throughput experimental techniques are increasingly helpful in elucidating interaction dynamics, we expect computational approaches for inferring and analyzing network dynamics to be a cornerstone in future efforts to understand biological systems. Here we have briefly reviewed some recent computational attempts to obtain a dynamic view of the interactome. Many of these approaches have built upon existing static molecular interaction data sets and integrated other types of information to help infer interaction dynamics, both to identify the paths by which information flows through cellular networks and to obtain a more global view of network organizational principles. These other sources have included gene expression data, phenotypic responses to perturbations such as gene knock-outs and information about eQTLs. In the near future, we anticipate that incorporating additional data sources, resulting from advances in mass spectrometry, next-generation sequencing and other high-throughput experimental methods, will help in further elucidating protein interaction and network dynamics. Novel data may include, for example, quantitative protein expression levels, protein localization and modification data, non-coding RNAs and detailed information about epigenetic regulation and the transcriptome. Looking further ahead, it is possible that incorporating future genome-scale data about dynamic responses to changes in DNA structure (e.g. supercoiling and its regulatory effects [168–170]) will require further expanding our dynamic models of regulatory control [171].

Experimental technologies for uncovering protein interaction dynamics are also being developed and are being applied to increasingly large systems. With respect to detecting protein–protein interactions, promising techniques include fluorescence resonance energy transfer (FRET) [172], which has been used, for example, to map protein interaction in the *Escherichia coli* chemotaxis pathway and determine stimulation induced changes [173]; florescent tagging of proteins, followed by rapid isolations and mass spectrometry identification of interactions, which has been used to uncover the dynamics of an alphavirus protein’s host protein interactions [174]; and a protein-fragment complementation assay that has allowed detection of yeast protein interactions within their proper cellular contexts [175]. For regulatory interactions, ChIP-seq technology is likely to be increasingly applied to uncover contextual and temporal variation. As existing technical difficulties with these and other technologies are overcome and larger dynamic networks are elucidated, computational approaches will increasingly switch focus from attempts to infer network dynamics to the analysis of experimentally determined dynamic networks. We expect that the emerging field of dynamic network analysis, which includes algorithmic approaches for modeling and analysis, mostly in the context of social and information networks, will be a good source of analytical methods relevant to dynamic biological networks as well [176–181].

As increasingly diverse high-throughput experimental data sets are gathered, it is likely that novel computational models, or new applications of existing models, will be needed to incorporate the data into predictive quantitative models of the dynamic interactome. While many probabilistic network models have already been applied to incorporate disparate data sets in a single framework, new methods that can intelligently relate different types of interaction data with different sorts of temporal and contextual biases may be helpful. For example, hypergraph models might be used for explicit representations of different types of interaction data. Important questions arising from efforts to merge disparate data types include how to propagate different types of dynamic information among an integrated data set, how to represent explicit bias and how to identify and resolve discrepancies.
An important consideration generally overlooked in the work reviewed here is that temporal variations affecting network dynamics can range over several orders of magnitude. This wide range not only raises questions about increasing and expanding the temporal resolution of experimental methods but also demands the development of appropriate multi-timescale computational methods. Another simplification made by many of the approaches that we have reviewed is that they typically model biological networks with a single node for each protein; in reality, each node represents a population of proteins that may be in different phosphorylation, activation or interaction states. Dynamical systems approaches for analysis and simulation [17, 18] provide a framework for handling these complexities, and integrating these approaches with more high-throughput graph-theoretic methods may be particularly fruitful.

To conclude, we believe that the shift from static to dynamic network analysis is essential for further understanding of molecular systems. Furthermore, we postulate that ignoring dynamic information in future computational network analysis research could turn out to be perilous. Indeed, while the analysis of static protein interaction networks has led to many hypotheses concerning cellular organization, pathways and function, some of these may be challenged by analyses incorporating network dynamics. Thus, we encourage computational biologists who are studying molecular interactions and networks to develop methods that incorporate information about the dynamic nature of cellular systems. After nearly a decade of work focusing on static interactomes, we think it’s about time.

**Key Points**
- Temporally, spatially and contextually dependent molecular interactions play a key role in regulating the functioning of cells and organisms. Incorporating and analyzing these dynamics are the next major challenges for computational interactomics.
- While most large-scale technologies to determine protein interactions do not provide spatial, temporal or contextual information about detected interactions, dynamic views of the interactome can be obtained by integrating static molecular interaction data sets with other types of dynamic data, such as gene expression data, phenotypic responses to perturbations such as gene knock-outs or knock-downs, and information about expression Quantitative Trait Loci (eQTLs).
- It is essential that computational biologists studying molecular networks increasingly shift their focus to develop methods that incorporate information about the dynamic nature of cellular systems.

**Acknowledgements**
This article was inspired by our involvement in the Pacific Symposium on Biocomputing session ‘Dynamics of Biological Networks’, which we are co-organizing with Tanya Berger-Wolf (University of Illinois). We thank her for many illuminating discussions, and the organizers of the PSB conference for supporting our session. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**FUNDING**
Intramural Research Program of the National Institutes of Health and National Library of Medicine (to T.M.P.); National Science Foundation grant CCF-0542187, National Institutes of Health grant GM076275 and National Institutes of Health Center of Excellence grant P50 GM071508 (M.S.); National Institutes of Health grants LM009411 and HD058880 (to D.K.S.).

**References**


