Biochemical and Biophysical Research Communications 424 (2012) 7–11

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Mini Review

A quick guide to biomolecular network studies: Construction, analysis, applications, and resources

Hailin Tang [a,b], Fan Zhong [b,*], Hongwei Xie [a,*]

[a] College of Mechanical & Electronic Engineering and Automatization, National University of Defense Technology, Changsha 410073, China
[b] Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

ARTICLE INFO

Article history:
Received 5 June 2012
Available online 23 June 2012

Keywords:
Network
Interaction
Model
Pathway

ABSTRACT

Over the past decade, a rapid increase in network data including signaling, transcription regulation, metabolic reaction, protein–protein interaction and genetic interaction has been observed. Many biology issues have been investigated by analyzing these diverse networks, providing new insights into biology. Networks also play an important role in disease studies including disease gene screening and clinical diagnosis. Large amounts of databases and software have been developed to facilitate the storage, exchange, integration, and analysis of network data and network analysis is becoming a routine procedure for biologists to infer biological information. In this review, several main aspects of network studies are discussed, including network construction, analysis, application, and resources.

1. Introduction

Biology has been explored in multiple levels, from the population, organism, tissue/organ, cell, and organelle, to the biomolecule. In the molecular level, researchers have realized that protein interactions are the foundation of biological functions. Numerous experimental and computational methods have been developed to depict functional and physical interactions, such as signal transduction, transcription regulation, metabolic reaction, protein–protein interaction (PPI), and genetic interaction, among others. Signal transduction focuses on the propagation of information by a series of chemical reactions, such as protein phosphorylation. Transcription regulation depicts the relationships between transcription factors and their target genes. Metabolic reactions mainly focus on small-molecule substrates and their modification enzymes. These three kinds of networks refer to real routes *in vivo* and are usually produced by molecular biological experiments. PPI generally represents the physical interactions between proteins [1], which can be produced by both traditional small-scale and high-throughput experiments, such as the yeast two hybrid (Y2H) technique and tandem affinity purification (TAP) followed by mass spectrometry. Genetic interactions represent a functional interaction between two genes. To illustrate, a certain phenotype occurs when two genes are perturbed, while the phenotype does not occur when only one gene is perturbed. These network data are dispersed in hundreds of public and private databases [2].

Many basic biology issues can be explored by these diverse networks, such as the sensitivity, robustness and evolution of biosystem. These networks also play an important role in diseases studies including disease gene screening and clinical diagnosis, to greatly improve our understanding of biology. In this review, the main aspects of network studies are systematically discussed, including network construction, analysis, application, and resources, aiming to provide a comprehensive picture of network studies for researchers.

2. Construction

The construction of a network is a prerequisite for any analysis. Network construction approaches can be classified based on different principles. Bruggeman et al. classified the construction approaches into "bottom-up" and "top-down" [3]. The bottom-up approach generates network by integrating scattered biological knowledge with interactions, while the top-down approach infers a network directly from omics data (e.g. microarray) by statistical methods [4]. Viswannathan et al. classified the construction approaches into Data-Driven Objective (DDO) and Knowledge-Driven Objective (KDO) models [5]. DDO generates a network based on a specific experiment, while KDO aims to develop a detailed network to depict a particular biological process such as cell differentiation. More biological knowledge is needed for performing a KDO construction. Comprehensive revalidation is required to achieve a high-quality network reconstruction because of the insufficiency of information about the network, such as the spatiotemporal specificity of interactions. For example, networks 1 and 2 occur in different biological conditions, but the merged network 3 is

* Corresponding authors. Address: Institutes of Biomedical Sciences, Fudan University, 138 Yixueyuan Rd., Shanghai 200032, China (F. Zhong).
E-mail addresses: zonefan@163.com (F. Zhong), xhwei_65@163.com (H. Xie).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2012.06.085

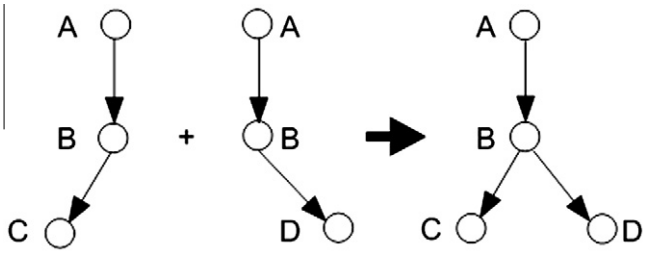

Fig. 1. The influence of spatiotemporal specificity for network construction. Networks 1 and 2 occur in different biological conditions. The merged network 3 is always constructed by ignoring spatiotemporal specificity, which cannot represent the truth precisely.

always built by ignoring spatiotemporal specificity (Fig. 1), which cannot represent the truth precisely. This problem currently exists in a number of databases.

3. Analysis

Structure and dynamics are the two fundamental aspects of network analysis. Some open questions, regarding the basic structural pattern of the biosystem, the mechanism of biosystem robustness, the molecular mechanism of a phenotype/disease, among others, may be answered through network analysis.

3.1. Structure

The structure of a biomolecular network is important to biosystem's function. Some structure patterns, such as scale-free, small-world, motif, modularity and hierarchy, which are found in many real world networks including WWW, instant messaging, email, movie actors, co-authorship, and citation networks, also exist in signaling, metabolic, PPI and other biological networks. Some specific structure patterns, such as bow-ties in metabolic networks [6], have also been discovered. Furthermore, several complex indices, such as information flow score [7] and SigFlux [8] have been developed to depict the structural characteristics of biomolecular networks, which can help in the discovery of novel patterns.

Graphs are the most widely used to representing network structures. The simple graph consists of nodes and edges, in which a node represents a biomolecule (e.g. protein and metabolite), while an edge represents the relationship between two biomolecules. Simple graphs are the most commonly used, but their simplicity makes precise depiction of complex biological information, such as protein complexes and metabolic reactions, difficult. As a result, some improved graphs, such as the hyper-graph [9] and power-graph [10], have been proposed. These graphs have a stronger capacity for facilitating a more precise presentation of biological knowledge. Hyper-graphs allow edges to connect more than two nodes, which are suitable for representation of biochemical reactions. Power-graphs are represented by power nodes and power edges. Power nodes are a set of nodes bought together that are suitable for depicting protein complexes. Novel discoveries may be found when using complex graphs instead of simple graphs [9].

How do structural features influence function? Take hub proteins as an example: Hub proteins can be classified into party-hub and date-hub proteins. The former interact with their partners simultaneously, while the latter interact at different times. Studies have shown that the party-hub tends to connect proteins within functional modules, whereas the date-hub usually makes a link between different modules [11]. Studies in the hierarchy of networks show that proteins in different layers have distinct function characteristics. Bhardwaj et al. suggested that more collaborations occur in the middle levels of various regulatory networks, the reason maybe that entities in the middle level must interact as many as possible to ensure organizational effectiveness [12]. Jothi et al. classified transcription factors (TFs) of yeast regulatory network into three hierarchical layers (i.e. top, core, and bottom), and found that the top-layer TFs are relatively abundant, long-lived, and noisy [13]. Oberdorf and Kortemme verified that complex topology rather than complex membership is a determinant of protein dosage sensitivity [14].

3.2. Dynamics

Network topology is usually insufficient to infer function, and information of network dynamics is required [15]. Network dynamics models can be classified into analytical method and statistical method-based models (Fig. 2). The analytical model, which focuses on depicting the molecular state quantitatively, could be eventually developed to simulate network behaviors [16]. Three steps are used to build an analytical model. The first step is to select and determine the state variables of the entities, for example, the phosphorylation state in the signaling pathway or the substrate concentration in the metabolic pathway. The second step is to select an appropriate mathematical model, such as the differential equation [17], random differential equation [18], rule-based [19], or probability-graph, to depict the relationships of the state variables. The third step is to evaluate and refine the coefficients of the mathematical models through experiment data so that the model matches the system behavior better. The development of an analytical model requires more accurate experiment information, and the model is currently only applied in small-scale and well-studied networks (e.g. EGFR [20], JAK/STAT pathway [21], metabolic network of microbes [22]). Thus, development of analytical models from novel high-throughput and more precise cell measurement techniques remains necessary. Some public databases (e.g. BiGG [23]) have been developed to store the quantitative model to support dynamics studies.

Statistical models only aim to reveal some dynamic characteristics of networks from limited time-resolved omics data (e.g. microarray, quantitative proteomic, or phosphoproteomic data) by statistical methods, and do not develop precise quantitative models due to insufficient of information, thus, such models are usually applied to large-scale networks. Integration of quantitative information into qualitative networks is an important issue for statistical models. Han et al. revealed dynamically organized modularity in the yeast PPI network by analyzing time-resolved transcriptome profiling with PPI data [11], while other research works have focused on the network dynamics of pathological processes. Teschendorff and Severini demonstrated that breast cancers metastasis is characterized by a small significant increase in the degree of randomness by comparing the metastatic and non-metastatic breast cancer transcriptomes in the PPI network [24].

4. Applications

4.1. Illustrating biomolecular mechanisms

Biology is an extremely complicated consequence of the various interactions of a large number of biomolecules. These biomolecular mechanisms can be described better through networks. Due to the diversity of biological processes, a fixed network analysis workflow to illustrate biomolecular mechanisms has yet to be achieved. Such an undertaking requires more biological knowledge than that yielded by the topology and dynamic characteristics of current networks. FANTOM consortium and the Riken Omics Science Center constructed a transcription network to yield a comprehensive

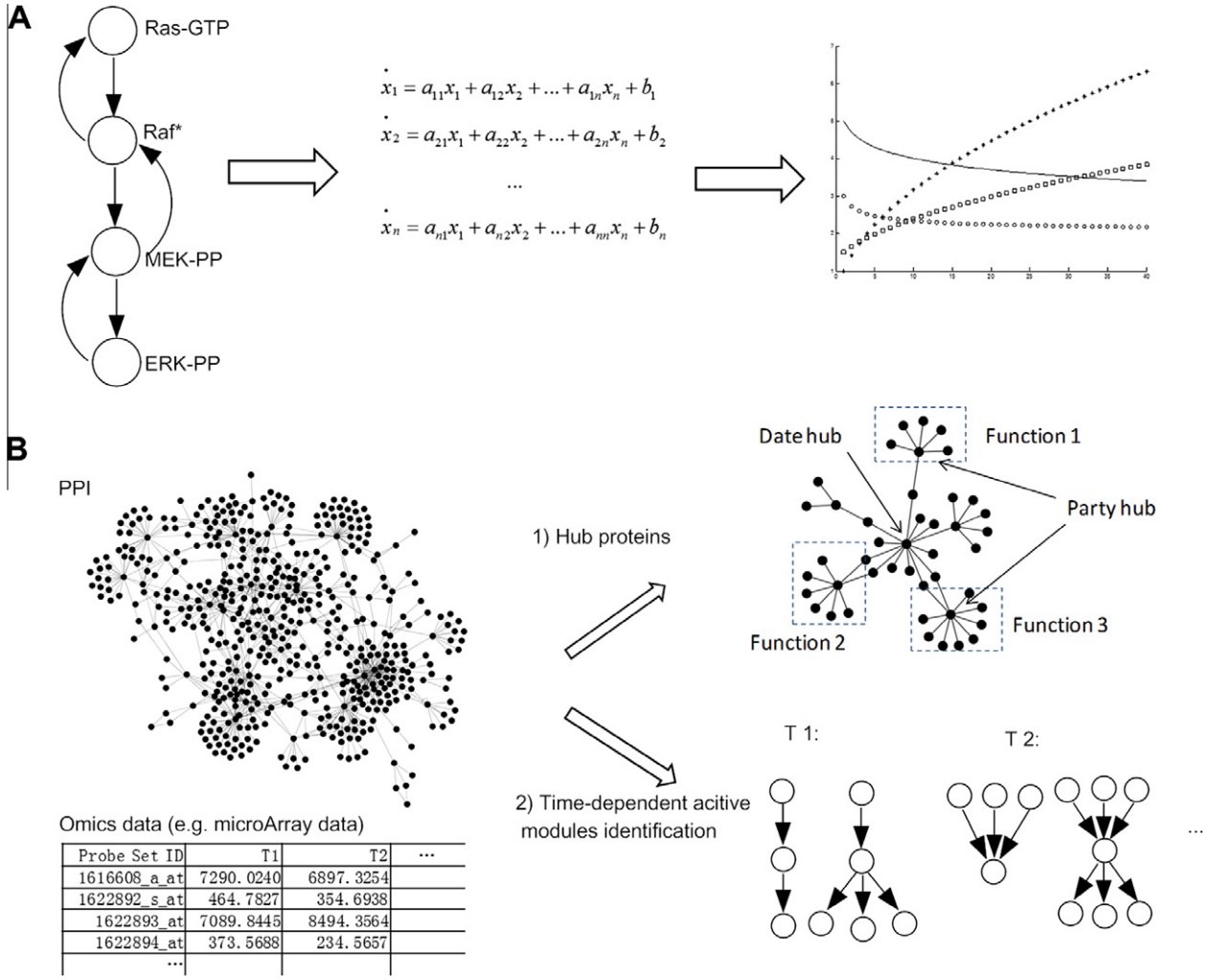

Fig. 2. Scheme of the two main network dynamics analysis models. (A) An analytical model is developed to simulate the behavior of a network. It is usually applied in small-scale and well-studied network. (B) Most statistical models aim only to reveal some dynamic characteristics (e.g. dynamic modularity) by statistical methods based on limited time-resolved omics data (e.g. microarray) and do not develop quantitative models due to insufficient of information. Such methods are often applied in large-scale networks.

analysis of growth arrest and differentiation in THP-1 cell. They identified the key transcription regulators with time-dependent activities and target genes, and showed that many transcription factors (TFs) involving in cell differentiation and cellular states are constrained by complex networks. These networks contain both positive and negative regulatory interactions among those TFs [25]. Zhu et al. reconstructed the causal gene networks by using genotype, transcriptome, transcription factor binding site (TFBS), and PPI data to predict causal regulators responsible for common genetic loci of gene expression activity in a yeast population, elucidating the mechanisms of how causal regulators give rise to larger-scale changes in gene expression activity [26].

4.2. Network properties of disease genes

Exploring the characteristics of disease genes, such as function and tissue specific characteristics, among others, has long been a matter of concern. Do disease genes have special network characteristics? What are these special network characteristics? These questions have been explored carefully in the past decade. Some researches show that disease genes tend to form modules [27]. Genes with pleiotropic effects have more interactors than single disease associated genes [28]. Genes with intermediate connectivities have the highest probability of harboring germ-line disease mutations [29]. Cancer genes have a high degree and tend to be central hubs [30–32]. However, Goh et al. showed that disease genes do not tend to possess centrality and be essential [27]. Genes from some diseases (e.g. cancer) tend to centrality, while others do not. These network characteristics give us some novel views of diseases.

4.3. Disease gene screening

Screening of genes resulting in specific diseases has long been one of the major tasks in human genetics studies. Numerous methods based on network and gene-expression data have been proposed to screen potential disease genes. Three approaches are most widely used. The first is based on the "guilty by association" hypothesis that similar mutational phenotypes arise from functionally related genes, which has been proven to be effective in screening novel disease genes [33]. The second is based on the topology and function properties of disease genes [34–36]. For this approach, machine learning methods are often used to capture the topology or function features by training the known disease genes. Prior known disease genes are required for these two methods, so they are not suitable for the diseases without enough prior disease genes. The third involves determination of differential gene expression modules (including complex and pathway) [37,38]. Chuang et al. first provided a subnetwork/module method to indentify biomarkers [39]. This method has three main advantages: 1) The resulting

subnetworks provide the molecular mechanisms underlying the disease. 2) Genes that are potentially related to disease but without differential expression can also be detected. 3) The identified subnetworks are significantly more robust than individual marker without network information. A large number of subnetwork/module-based methods have been proposed to improve the accuracy or applicability of disease gene screening [40]. Recently, Emmert-Streib and Glazko proposed a basic protocol with which to determine differential networks [41]. In addition, lots of tools have been developed to facilitate prioritization of candidate disease genes [42].

4.4. Clinical diagnosis

Microarray data are widely used to predict biomarkers by scoring each individual gene. However, this method has low robustness due to genetic, physiological differences, random fluctuations and other factors [43]. Chuang et al. improved the prediction robustness of biomarkers by identifying differential expression subnetwork/module [39]. More complex variables have been constructed to characterize the phenotypes through integrating the gene expression and PPI data. For example, Teschendorff and Severini developed the "entropy" parameter to predict the metastasis ability of breast cancer [24]. Recent overviews have also discussed the application of network in disease studies [44,45].

5. Resources

Research has shown that the datasets may be error-prone and possibly of lower quality than commonly assumed [46]. It is thus necessary for researchers to realize the advantages and disadvantages of various databases and corresponding data standards to select appropriate data and avoid misuse.

5.1. Data standards

Data standards aim to facilitate the storage, exchange, extraction and analysis of data by unifying the modes of presentation. SBML [47], CellML [48], PSI-MI [49], BioPAX [50] and SBGN [51] are the most widely used standards. Excepting for SBGN, these standards all adopt eXtensible Markup Language (XML), which has been proven to be an effective tool for the management of data from diverse resources. SBML and CellML support dynamic simulations of networks. PSI-MI mainly emphasizes on the depiction of the information of molecular interaction pairs, including the interaction type, experimental method, specie, literature, and author. BioPAX provides more complete functions, which includes all the PSI-MI items and the most of SBML. But BioPAX cannot support simulation like SBML, and is more complicated than PSI-MI when used to store PPI data. Stromback and Lambrix have made a systematic comparison of these three standards [52]. SBGN is completely different from the four other standards. SBGN aims to unify the representation of network graphs, while the other four standards aim to provide universal models for storing network information. In order to demand diverse needs, SBGN is designed to consist of three complementary sublanguages, including the process diagram, the entity relationship diagram, and the activity flow diagram.

5.2. Databases

Network data are spread throughout hundreds of databases. Pathguide systematically sorts them out and provides basic directions [2]. Some other systematic comparisons and assessments of these databases have been made to provide directions [53–58]. Results show that PPI databases have larges variation in the interactors and interaction scales, and depth of annotations [54]. There is poor consistency and compatibility among several popular pathway databases [57]. Thus, it is necessary to integrate these databases to provide consistent information. Several integrated databases have been developed [59–61]. However, integration of these heterogeneous data remains largely incomplete.

5.3. Software tools

Network analysis tools can be classified into six functional types: 1) Data format transition among different standards of diverse databases [62,63]. 2) Network graph construction and visualization [64–66]. There are two network graph models: one is based on user-provided interactions, and another is based on software-provided interactions, when users only need to provide a list of genes/proteins. Gehlenborg et al. made a detailed discussion about visualization of omics data for systems biology [67]. 3) Topology analysis of networks, including node degree, shortest path, cluster coefficient, and specific structure detection, such as motif and module detecting, among others. [68,69]. 4) Dynamic modeling based on mathematical models [70]. 5) Enrichment analysis in canonical pathways or gene sets [71–73]. 6) Network merging of multiple subnetworks to obtain more comprehensive views [69].

6. Discussion

In the past decade, great progress has been made in molecular interaction detection and network analysis, providing new insights into biology. Network study is still in its infancy, and there remain numerous problems to overcome, such as incompleteness, low precision of interaction data, and lack of spatiotemporal information of interactions. It is expected that network studies will maintain a rapid pace of development in the near future, especially in the field of network dynamics.

Funding

The study was supported by the National Basic Research Programs [2010CB912700], National Natural Science Foundation Projects [31000379, 31000591, and 30900257], Shanghai Health Bureau Project [2010Y005], Chinese State Key Project Specialized for Infectious Diseases (2012ZX10002012–006), and Natural Science Foundation of Shanghai [09ZR1404100].

References

[1] J. De Las Rivas, C. Fontanillo, Protein–protein interactions essentials: key concepts to building and analyzing interactome networks, PLoS Comput. Biol. 6 (2010) e1000807.
[2] G.D. Bader, M.P. Cary, C. Sander, Pathguide: a pathway resource list, Nucleic Acids Res. 34 (2006) D504–D506.
[3] F.J. Bruggeman, H.V. Westerhoff, The nature of systems biology, Trends Microbiol. 15 (2007) 45–50.
[4] M. Bansal, V. Belcastro, A. Ambesi-Impiombato, et al., How to infer gene networks from expression profiles, Mol. BiolSyst. 3 (2007) 78.
[5] G.A. Viswanathan, J. Seto, S. Patil, et al., Getting started in biological pathway construction and analysis, PLoS Comput. Biol. 4 (2008) e16.
[6] J. zhao, H. Yu, J. Luo, et al., Hierarchical modularity of nested bow-ties in metabolic networks, BMC Bioinformatics 7 (2006) 386.
[7] P.V. Missiuro, K. Liu, L. Zou, et al., Information flow analysis of interactome networks, PLoS Comput. Biol. 5 (2009) e1000350.
[8] W. Liu, D. Li, Z. zhang, et al., SigFlux: A novel network feature to evaluate the importance of proteins in signal transduction networks, BMC Bioinformatics 7 (2006) 515.
[9] S. Klamt, U.U. Haus, F. Theis, Hypergraphs and cellular networks, PLoS Comput. Biol. 5 (2009) e1000385.
[10] L. Royer, M. Reimann, B. Andreopoulos, et al., Unraveling protein networks with power graph analysis, PLoS Comput. Biol. 4 (2008) e1000108.

[11] J.D.J. Han, N. Bertin, T. Hao, et al., Evidence for dynamically organized modularity in the yeast protein–protein interaction network, Nature 430 (2004) 88–93.
[12] N. Bhardwaj, K.K. Yan, M.B. Gerstein, Analysis of diverse regulatory networks in a hierarchical context shows consistent tendencies for collaboration in the middle levels, Proc. Natl. Acad. Sci. U S A. 107 (2010) 6841–6846.
[13] R. Jothi, S. Balaji, A. Wuster, et al., Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture, Mol. BioISyst. 5 (2009) 294.
[14] R. Oberdorf, T. Kortemme, Complex topology rather than complex membership is a determinant of protein dosage sensitivity, Mol. BioISyst 5 (2009) 253.
[15] R.P. Alexander, P.M. Kim, T. Emonet, et al., Understanding modularity in molecular networks requires dynamics, Science Signaling 2 (2009) pe44.
[16] J.J. Hughey, T.K. Lee, M.W. Covert, Computational modeling of mammalian signaling networks, Wiley Interdiscip. Rev Syst. Biol. Med 2 (2010) 194–209.
[17] J. Lequieu, A. Chakrabarti, S. Nayak, et al., Computational modeling and analysis of insulin induced eukaryotic translation initiation, PLoS Comput. Biol. 7 (2011) e1002263.
[18] T. Meng, S. Somani, D. P, Modeling and simulation of biological systems with stochasticity, In Silico Biol. (2004) 209–293.
[19] J. Colvin, M.I. Monine, J.R. Faeder, et al., Simulation of large-scale rule-based models, Bioinformatics 25 (2009) 910–917.
[20] D.Y.Q. Wang, L. Cardelli, A. Philip, et al., Computational modeling of the EGFR network elucidates control mechanisms regulating signal dynamics, BMC Syst. Biol. 3 (2009) 118.
[21] J. Bachmann, A. Raue, M. Schilling, et al., Division of labor by dual feedback regulators controls JAK2/STAT5 signaling over broad ligand range, Mol. BioISyst. 7 (2011) 516.
[22] A.M. Feist, M.J. Herrgard, I. Thiele, et al., Reconstruction of biochemical networks in microorganisms, Nat. Rev. Microbiol. 7 (2009) 129–143.
[23] J. Schellenberger, J.O. Park, T.M. Conrad, et al., BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions, BMC Bioinformatics 11 (2010) 213.
[24] A.E. Teschendorff, S. Severini, Increased entropy of signal transduction in the cancer metastasis phenotype, BMC Syst. Biol. 4 (2010) 104.
[25] H. Suzuki, A.R. Forrest, E. van Nimwegen, et al., The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line, Nat. Genet. 41 (2009) 553–562.
[26] J. Zhu, B. Zhang, E.N. Smith, et al., Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks, Nat. Genet. 40 (2008) 854–861.
[27] K.I. Goh, M.E. Cusick, D. Valle, et al., The human disease network, Proc. Natl. Acad. Sci. U S A 104 (2007) 8685–8690.
[28] S. Chavali, F. Barrenas, K. Kanduri, et al., Network properties of human disease genes with pleiotropic effects, BMC Syst. Biol. 4 (2010) 78.
[29] I. Feldman, A. Rzhetsky, D. Vitkup, Network properties of genes harboring inherited disease mutations, Proc. Natl. Acad. Sci. U S A 105 (2008) 4323–4328.
[30] P.F. Jonsson, P.A. Bates, Global topological features of cancer proteins in the human interactome, Bioinformatics 22 (2006) 2291–2297.
[31] Q. Cui, Y. Ma, M. Jaramillo, et al., A map of human cancer signaling, Mol. BioISyst. 3 (2007) 152.
[32] J. Sun, Z. Zhao, A comparative study of cancer proteins in the human protein–protein interaction network, BMC Genet. 11 (Suppl 3) (2010) S5.
[33] S. Navlakha, C. Kingsford, The power of protein interaction networks for associating genes with diseases, Bioinformatics 26 (2010) 1057–1063.
[34] C. Ortutay, M. Vihinen, Identification of candidate disease genes by integrating Gene Ontologies and protein-interaction networks: case study of primary immunodeficiencies, Nucleic Acids Res. 37 (2009) 622–628.
[35] L. Li, K. Zhang, J. Lee, et al., Discovering cancer genes by integrating network and functional properties, BMC Med. Genet. 2 (2009) 61.
[36] J. Xu, Y. Li, Discovering disease-genes by topological features in human protein-protein interaction network, Bioinformatics 22 (2006) 2800–2805.
[37] A. Keller, C. Backes, A. Gerasch, et al., A novel algorithm for detecting differentially regulated paths based on gene set enrichment analysis, Bioinformatics 25 (2009) 2787–2794.
[38] J.H. Hung, T.W. Whitfield, T.H. Yang, et al., Identification of functional modules that correlate with phenotypic difference. the influence of network topology, Genome Biol. 11 (2010) R23.
[39] H.Y. Chuang, E. Lee, Y.T. Liu, et al., Network-based classification of breast cancer metastasis, Mol. BioISyst. 3 (2007) 140.
[40] T. Ideker, N.J. Krogan, Differential network biology, Mol. BioISyst 8 (2012) 565.
[41] F. Emmert-Streib, G.V. Glazko, Pathway analysis of expression data: deciphering functional building blocks of complex diseases, PLoS Comput. Biol. 7 (2011) e1002053.
[42] L.C. Tranchevent, F.B. Capdevila, D. Nitsch, et al., A guide to web tools to prioritize candidate genes, Brief Bioinformatics 12 (2011) 22–32.
[43] L. Ein-Dor, O. Zuk, E. Domany, Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer, Proc. Natl. Acad. Sci. U S A 103 (2006) 5923–5928.
[44] A.L. Barabasi, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nat. Rev. Genet. 12 (2011) 56–68.
[45] T. Ideker, R. Sharan, Protein networks in disease, Genome Res. 18 (2008) 644–652.
[46] M.E. Cusick, H. Yu, A. Smolyar, et al., Literature-curated protein interaction datasets, Nat. Methods 6 (2009) 39–46.
[47] M. Hucka, A. Finney, H.M. Sauro, et al., The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models, Bioinformatics 19 (2003) 524–531.
[48] C.M. Lloyd, M.D. Halstead, P.F. Nielsen, CellML: its future, present and past, Prog. Biophys. Mol. Biol. 85 (2004) 433–450.
[49] H. Hermjakob, The HUPO PSI's molecular interaction format—a community standard for the representation of protein interaction data, Nat. Biotechnol. (2005) 177–183.
[50] E. Demir, M.P. Cary, S. Paley, et al., The BioPAX community standard for pathway data sharing, Nat. Biotechnol. 28 (2010) 935–942.
[51] N. Le Novere, M. Hucka, H. Mi, et al., The Systems Biology Graphical Notation, Nat. Biotechnol. 27 (2009) 735–741.
[52] L. Stromback, P. Lambrix, Representations of molecular pathways: an evaluation of SBML, PSI MI and BioPAX, Bioinformatics 21 (2005) 4401–4407.
[53] G. Chaurasia, H. Herzel, M.E. Futschik, Comparison and integration of large-scale human protein–protein interaction maps, FEBS J. 274 (2007) 300.
[54] S. Mathivanan, B. Periaswamy, T.K.B. Gandhi, et al., An evaluation of human protein–protein interaction data in the public domain, BMC Bioinformatics 7 (2006) S19.
[55] M.E. Futschik, G. Chaurasia, H. Herzel, Comparison of human protein-protein interaction maps, Bioinformatics 23 (2007) 605–611.
[56] A. Bauer-Mehren, L.I. Furlong, F. Sanz, Pathway databases and tools for their exploitation: benefits, current limitations and challenges, Mol. BioISyst. 5 (2009) 290.
[57] D. Soh, D. Dong, Y. Guo, et al., Consistency, comprehensiveness, and compatibility of pathway databases, BMC Bioinformatics 11 (2010) 449.
[58] E. Shmelkov, Z. Tang, I. Aifantis, et al., Assessing quality and completeness of human transcriptional regulatory pathways on a genome-wide scale, Biol. Direct 6 (2011) 15.
[59] L.J. Jensen, M. Kuhn, M. Stark, et al., STRING 8-a global view on proteins and their functional interactions in 630 organisms, Nucleic Acids Res. 37 (2009) D412–D416.
[60] E.G. Cerami, B.E. Gross, E. Demir, et al., Pathway Commons, a web resource for biological pathway data, Nucleic Acids Res. 39 (2011) D685–D690.
[61] A. Kamburov, K. Pentchev, H. Galicka, et al., ConsensusPathDB: toward a more complete picture of cell biology, Nucleic Acids Res. 39 (2011) D712–D717.
[62] M.J. Schilstra, L. Li, J. Matthews, et al., CellML2SBML: conversion of CellML into SBML, Bioinformatics 22 (2006) 1018–1020.
[63] K. Moutselos, I. Kanaris, A. Chatziioannou, et al., KEGGconverter: a tool for the in-silico modelling of metabolic networks of the KEGG Pathways database, BMC Bioinformatics 10 (2009) 324.
[64] M.P. van Iersel, T. Kelder, A.R. Pico, et al., Presenting and exploring biological pathways with PathVisio, BMC Bioinformatics 9 (2008) 399.
[65] M.E. Smoot, K. Ono, J. Ruscheinski, et al., Cytoscape 2.8: new features for data integration and network visualization, Bioinformatics 27 (2011) 431–432.
[66] Z. Hu, J.H. Hung, Y. Wang, et al., VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology, Nucleic Acids Res. 37 (2009) W115–W121.
[67] N. Gehlenborg, S. IO'Donoghue, N.S. Baliga, et al., Visualization of omics data for systems biology, Nat. Methods 7 (2010) S56–S58.
[68] K.Y. Yip, H.Y. Yu, P.M. Kim, et al., The tYNA platform for comparative interactomics: a web tool for managing, comparing and mining multiple networks (vol 22, pg 2968, 2006), Bioinformatics 23 (2007) 1048.
[69] P. Shannon, A. Markiel, O. Ozier, et al., Cytoscape: A software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.
[70] A. Funahashi, Y. Matsuoka, A. Jouraku, et al., Cell Designer 3.5: A versatile modeling tool for biochemical networks, Proc. IEEE 96 (2008) 1254–1265.
[71] J. Xia, N. Psychogios, N. Young, et al., MetaboAnalyst: a web server for metabolomic data analysis and interpretation, Nucleic Acids Res. 37 (2009) W652–W660.
[72] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57.
[73] J. Shi, M.G. Walker, Gene set enrichment analysis (GSEA) for interpreting gene expression profiles, Curr. Bioinform. 2 (2007) 133–137.