Chapter 14
Computational Network Approaches and Their Applications for Complex Diseases

Ankita Shukla and Tiratha Raj Singh

Abstract Network biology has been widely used for the interaction studies and analysis in modern era. Studies associated with biological networks, their modeling, analysis, and visualization are imperative to the biological world. The advancements in network biology have helped us better understand the biomolecular complexities which in earlier times were difficult to study in vivo. High-throughput technologies have revolutionized the genomic-sequencing procedures for the generation of tremendous data that need to be analyzed and interpreted rigorously. However it has still been difficult to construe biological networks completely due to the complexity of the interactions that exist between its components. Great efforts have been made to disclose the maximum possible interactions that are significant to maintain the potential mechanisms with the help of network biology. With improvement in the analysis process, network biology has thought to be playing a key role in understanding the complex biological behavior of the networks. In this chapter we will cover the basics of network biology and its expansion in various disciplines and its implementations in complex diseases and disorders, current resources, and tools available for studying diverse forms of pathways (transcriptional regulation, protein–protein interaction, signal transduction, and metabolism). This chapter therefore deals with the core of network biology, its role in various disease studies, and the advancements introduced so far in this field.

Keywords Interactome • Cancers • Autism • Network motif • Cardiovascular • Aging • Diabetes

A. Shukla • T.R. Singh (✉)
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Waknaghat, Solan, India
e-mail: tiratharaj.singh@juit.ac.in

D.-Q. Wei et al. (eds.), *Translational Bioinformatics and Its Application*, Translational Medicine Research, DOI 10.1007/978-94-024-1045-7_14

14.1 Introduction

Network biology is an amalgamation of systems biology, graph theory, and computational and statistical analysis techniques wherein the topology of the graphs represents the molecular interaction (Barabási and Oltvai 2004). Network biology deals with biological complexities of the cellular components that comprise macromolecules (genes, RNA and DNA) and metabolites. Most biological characteristics arise from complex interactions between these cellular components. There are diverse forms of biological networks which include metabolic networks, cell signaling networks, kinase–substrate networks, gene regulatory networks, protein–protein interaction networks, epistasis interaction networks, disease–gene interaction networks, and drug interaction networks (Ma'ayan 2011; Zhu et al. 2007; Winterbach et al. 2013). In network biology, a whole system is summarized in the form of nodes and edges, wherein nodes represent the biological component and edges represent interactions between them. These interaction studies can prove to be helpful to unlock the mystery behind interrelatedness of biological pathways. The use of standardized and efficient approaches is the first step in knowledge casting to provide unbiased maps of functional interactions. The next challenge in network biology is to examine the resulting interactome and to excavate global or local network properties. These graph properties can be used to improve understanding of biological processes that are crucial for organism's livelihood. Advancement in network biology offers a novel conceptual scaffold that will possibly revolutionize our view of biology and disease pathologies (Barabási and Oltvai 2004).

It has been seen from various studies that behavior of the cell is not the result of functioning of single pathway but of multiple ones that work together. So the rule of unity can also be seen in biological environment wherein if there is a problem with one pathway, it will not directly affect the system as the interrelated pathways try to compensate for it. Therefore, a key challenge today is to understand the structure and the dynamics of the complex intercellular web of interactions that contribute to the structure and functioning of a living cell. Numerous experimental techniques have been developed to date for identifying the physical interactions as well as the state of the cell at any point of time. It includes protein chips, yeast two-hybrid screens for physical interactions, and high-throughput screening for determining the status of the cell (Srinivasa et al. 2014). On the other hand, network biology proposes a quantifiable description of the networks that helps researchers characterize diverse biological systems.

To understand the network topology, one needs to know the properties related to the nodes and edges that should be considered while performing network analysis. Properties of nodes include (a) connectivity degree, the number of links each node possesses; (b) node betweenness centrality, the number of the shortest paths among all the shortest paths between all possible pairs of nodes; (c) closeness centrality, the average shortest path from one node to all other nodes; and (d) eigenvector centrality, a more sophisticated centrality measure that assesses the closeness to

highly connected nodes. Properties of edges include (a) betweenness centrality, the number of the shortest paths among all possible shortest paths between all pairs of nodes; (b) types of relationship, for example, edges may represent activating or inhibiting relationships between a pair of nodes; and (c) edge directionality, the upstream and downstream nodes connected by a particular link.

Properties of network include:

(a) Connectivity distribution: quantitative links between nodes and edges
(b) Characteristic path length: the average shortest path between all pairs of nodes
(c) Clustering coefficient: local density of interactions measured by the connectivity of neighbors for each node averaged over the entire network
(d) Grid coefficient: extends the clustering coefficient by only looking at first neighbors to also examine second neighbors
(e) Network diameter: represents the longest of the shortest paths
(f) Assortativity: assesses whether nodes prefer to attach to other nodes on the basis of common nodal properties (Barabási and Oltvai 2004; Ma'ayan 2011).

Along with the abovesaid properties, there are two significant characteristics that are found in the network topology; one is network motifs which are recurring patterns composed of few nodes and edges. It has been observed that these network motifs appear in the regulatory networks much more frequently than in random or shuffled networks. There are also subcategories that underlie motifs including autoregulatory motifs, feedback loops, feedforward loops, bifans, diamonds, 3 and 4 chains, and other types of cycles that directly influence system's overall dynamics (Kim et al. 2011). These biologically significant subgraphs are vital to uncover the structural aspect of the complex network. Although network motifs provide insight to the functional properties, its detection is quite challenging (Masoudi-Nejad et al. 2012). Another characteristic of networks is their modularity, representing the modules or network clusters. These modules are dense regions of connectivity and are separated by low connectivity regions. Nearest neighbors clustering, Markov clustering, and betweenness centrality-based clustering which comes under the category of unsupervised clustering algorithms are used to identify the modules in networks (Ma'ayan 2011). Modularity is often used for detecting community structure in networks which have significant functionality at local levels. In the subcategory of modules come party hubs, nodes that interact with several proteins in one cellular compartment at a specific time, and date hubs, proteins that can be found in many places inside the cell and interact with diverse partners at different times (Chang et al. 2013). It has been observed from various studies that most of the biological molecular regulatory networks are scale-free, meaning that their degree distribution, i.e., distribution of edges per node, fits a power law. The overall functioning and homeostasis are being maintained by the scale-free architectures which make the network robust to the random failures.

Networks can be directed or undirected depending on the nature of the interactions. The directed network represents the interaction between any two nodes with a well-defined direction. These directionalities can be an inference of the material flow from a substrate to a product in a metabolic reaction or the information flow

from a transcription factor to the gene. However, in case of the undirected networks, the links do not show any assigned direction. Network motifs identified within directed or undirected networks are called graphlets. One of the examples of graphlets is those found in protein–protein interaction networks (Przulj et al. 2004). The fundamental method in understanding the biological networks is the network visualization that helps scientists in uncovering important properties of the underlying biochemical processes. The computational methods are proved helpful for analysis purpose, but the major disadvantage is in data peculiarity that makes the network interpretation difficult due to the complexity of the relationships.

As mentioned earlier, complex diseases are generally caused not from the malfunction of individual molecules but from the interplay of a group of correlated molecules or a network (Schadt 2009). Molecular biomarkers are widely employed today as they are helpful to discriminate normal vs. disease samples. However, there is a serious problem regarding their usage, i.e., they suffer from low coverage along with the high false-positive/false-negative rates and further limit their clinical applications. The limitation in traditional concept of biomarkers has been now conquered with the modern concept of network biomarkers (also called module biomarkers), and they achieve better performance because of the involvement of diverse interactions of the molecules. Networks are considered more robust to characterize the disease conditions than individual molecules. One drawback allied both to the molecular biomarkers and the network biomarkers is that they can only differentiate disease and normal conditions but cannot reliably identify predisease conditions, therefore lacking the ability for early diagnosis.

Regarding the abovementioned condition, a new concept of dynamical network biomarkers (DNBs) has been developed based on nonlinear dynamical theory and complex network theory. One of the advantages of the DNB is its ability to distinguish a predisease state from normal and disease states for even a small number of samples, providing great potential to achieve authentic early diagnosis for the complex diseases. Network biomarkers offer quantifiable and stable forms to characterize biomedical phenotypes or diseases in contrast to individual molecular biomarkers, which has inspired the development of systems medicine in the network level (Liu et al. 2012; Kitano 2002; Aryee et al. 2013). Unlike molecular biomarkers and network biomarkers, a DNB does not always contain a group of fixed members even for the same disease but might have different molecules depending on individual variations that can be identified by high-throughput data. As compared to the edge biomarkers (or network biomarkers), which exploit correlation or association information between molecules (expected to uncover better biomarkers relating genotypes to phenotypes), DNB explores dynamical information of data together with network information. An unavoidable problem for both edge biomarkers and network biomarkers is the requirement of multiple samples in the predicting step. Thus, DNB is used for detecting the predisease condition and therefore provides the early signal of a disease. DNB method is relatively easy to implement as it can be achieved with a smaller number of samples and is a model-free method (Liu et al. 2014). Therefore, network biomarkers can

exploit network information to unravel mechanisms of disease initiation and progression and thus improve the accuracy of diagnosis and prognosis.

14.2 Importance of Network Biology Toward Disease Prevention and Cure

The study of networks has emerged in diverse disciplines for analyzing complex relationships. The analysis of biological networks with respect to the human diseases has led to the discovery of field known as network medicine. Network biology has revolutionized the disease interrogation by uncovering many complex linkages that reflect perturbations in the biological networks. The functional interdependencies play major roles in maintaining the potential mechanisms which if not worked properly could lead to a disease condition. It has been found that disease is rarely a consequence of abnormality in single gene; the majority is due to the irregularities in multiple linked genes. The advent of network medicine leads to the emergence of a variety of tools which provide platforms to explore not only the molecular complexity of a particular disease but also the molecular relationships that exist between distinct phenotypes (Barabási et al. 2011). Advances in this direction are vital for the detection of new disease genes and for revealing the biological significance of disease-associated mutations. Since network also influences functioning of the other related networks, this interconnectivity implies that the impact of a specific genetic abnormality is not only limited to the activity of the gene product that holds it but can extend along the links of the network. These anomalies therefore alter the activity of gene products that otherwise had no defects initially. Hence, identifying phenotypic impact of a defect is not solely dependent on the known function of the mutated gene but also on the functions of components with which the gene and its products interact.

From the field of network medicine, it was found that it is the essential genes that are encoding hubs and not the diseased ones (Barabási et al. 2011). This statement is justified in terms of evolutionary perception because if we assume that mutations disrupt hubs, the absence of hubs will create so many disruptions that the host may not survive long enough to evolve and reproduce. Thus, only mutations that impair functioning lie at the periphery, accounting for the numerous disease conditions (Park et al. 2008). To determine the network-based position of disease genes, we need to understand three distinct phenomena which comprise (a) topological module, represents a locally dense neighborhood in a network, (b) functional module, represents the aggregation of nodes of related function, and (c) disease module, represents a group of network components that together contribute to a cellular function and their disruption results in a particular disease phenotype (Vidal et al. 2011). These three concepts are interrelated given that the cellular components that form a topological module have closely related functions and thus correspond to a

functional module, and a disease is an outcome of the breakdown in a particular functional module.

Network-based approaches to human diseases can have numerous biological and clinical applications. It therefore provides a better understanding of the implications of cellular interconnectedness on disease progression which offers better targets for drug development by the identification of disease genes and pathways. These advances will possibly reshape clinical practice, through the discovery of better and more accurate biomarkers for better disease classification, paving the way to personalized treatments and therapies.

14.3 Network-Based Computational Approaches from Network Biology Available for Complex Diseases

In recent years, network-based approaches emerged as powerful tools for studying complex diseases. Computational biologists are working continuously in utilizing various approaches for understanding implicated pathways in complex diseases. This leads to the expansion of diverse algorithmic approaches to expose a range of facet of network biology. Today there is a dire need to enhance this field so as to tackle the most challenging diseases that fall under the category of complex one (named so as currently there is no effective therapy available to treat them). Many diseases fall in this category including cancer, autism, diabetes, obesity, Alzheimer's disease (AD), and cardiovascular diseases (CVDs). Genetic unbalancing is the root cause in complex diseases along with internal and external perturbations (Cho et al. 2012; Mitchell 2012). Primary difficulties to deal with complex diseases are that each of them might be caused by different genetic conditions. In addition, if a disease is caused by a combinatorial effect of many mutations, the individual effects of each mutation might be small and therefore hard to discover. According to the study, autism is considered to be one of the most inbred complex disorders, but its principal genetic causes are still largely unknown (Diaz-Beltran et al. 2013). Rare genetic disparity and its heterogeneity aid in the emergence of the complex disease (Kristiansson et al. 2008). Therefore, in case of the complex diseases, researchers are gradually focusing more on groups of related genes, referred to as modules or subnetworks. Typically, these topologies hold information like whether a given molecule acts as an activator or inhibitor (Mitra et al. 2013). An important advantage of working with modules rather than individual gene is that it is often easier to predict the function of a module than the function of a gene.

It is important to keep in mind that there are some disadvantages pertaining to the modules; those identified from high-throughput techniques are noisy, containing both false-negative and false-positive edges (Cho et al. 2012). Also, many times they skip information about the nature of an interaction. Not only the experimental ones but the computationally identified network modules also lack a

mechanistic explanation of pathway activities. Therefore, selection of data and associated methods of analysis has to be chosen carefully. Network biology enlightens the diverse ways to deal with complex forms of the disease.

Aging Aging is the most prominent factor allied to the more complex forms of diseases, such as cancer, diabetes, CVDs, and neurodegenerative disorders (Cevenini et al. 2010). Aging phenotypes are coupled to the large and complex networks where cross talk occurs between assorted components. The main challenge in post-genomic aging research will be the dissection and analysis of the complex gene regulatory networks involved in aging processes. Structure and behavior studies are helpful in deducing the phenotypic effects of the responses that take place inside the cell. Since it is hard to infer logic of genetic networks experimentally, the union of new experiments and computational modeling techniques has been pursued currently.

Cardiovascular Diseases CVD covers a wide variety of disorders which influence different parts of cardiovascular system and includes coronary diseases, carotid diseases, peripheral arterial diseases, and aneurysms (Sarajlić and Pržulj 2014). There are also other forms of CVDs that are Mendelian disorders resulting from a mutation of the single gene. Genome-wide association studies have revealed that cardiovascular diseases, like the majority of complex diseases, have amazing complex genetic architecture, and they actually do not possess any major genes. Researchers have tried to investigate whether basic topological information such as connectivity of the nodes can be interrelated with biological properties which underlie CVD onset and progression. Module-based approaches are applied to determine functional modules related to the disease and to discover new associations between genes and disease. Various types of network biology approaches have been applied so far for CVD like in analysis done by Diez et al. who had created a combined gene association and correlation network (Diez et al. 2010). Rende et al. used topological features of PPI networks in search of genes common to CVDs and other diseases by identifying functional modules of genes (Rende et al. 2011). Approaches based on the biomolecular interaction networks provide better insight into network topology of the disease and thus could help researchers discover novel CVD genes and pathways.

Autism Autism is an early onset complex neurodevelopment disorder manifested in a broad phenotypic range (Diaz-Beltran et al. 2013). Although it is recognized as a highly heritable disorder, it is still uncertain whether the genetic variations are due to few common variants or because of many rare ones. Multifactorial nature of the complex disease makes use of systems biology perspective to embrace the complexity of the biological processes and the vast variety of molecular interactions that take place. Results showed that more than half of the published autism genes have been also allied to related neurological disorders (Wall et al. 2009). These findings provide evidence of molecular overlapping and indicate that these disorders might share molecular mechanisms which will perhaps help us understand the etiology of this complex disorder. The main objective of exploring the autism

network is that the researchers will be able to locate those genes that cause the disorder and pave the way toward faster diagnosis as well as treatment.

Diabetes About 90% of the diabetic population is affected with type 2 diabetes which poses serious health issues to the society (Bergholdt et al. 2007). Elevated blood glucose level is the primary marker which occurs as a consequence of declined insulin activity. The adverse form of the disease could lead to the cardiovascular, renal, neurological, and organ complications. To date, a key challenge has been to identify the biological processes or signaling pathways that play significant roles in the disorder. There are various system-level studies done for diabetes like the one contributed by Davis et al. who found the loci contributing to diabetes-related traits along with the candidate genes with variation in gene expression (Davis et al. 2012). Integrating high-throughput microarray studies, with protein–protein interaction networks, seems to give the benefit in elucidating the underlying biological processes associated with chronic diabetic conditions. Therefore, there is a need to place more emphasis on the network biology methods to envisage the alteration caused by the summation of disordered pathways.

Cancers Cancer is caused by deleterious mutations leading to the abnormal functioning of a complex network. In cancer, dysregulation of multiple pathways, which govern fundamental cell processes, leads to different consequences like cell death, proliferation, differentiation, and migration. Like in case of other complex diseases, interrelatedness of biological interactions affects multiple cellular functioning. A major challenge is to find how diverse genetic mutations could help researchers build actionable understanding of this multivariate dysregulation. Therefore, the availability of diverse forms of networks which are amenable for computational analysis offers successful application of bioinformatics and systems biology methods for analysis of high-throughput data in cancer research. However, the key challenge is how significant advances can be made by applying computational modeling approaches to expose the pathways most critically involved in tumor formation and succession (Shukla et al. 2015, 2016; Sehgal et al. 2015).

Alzheimer's Disease Alzheimer's disease also abbreviated as AD is another category of complex disease of the central nervous system that occurs as a result of abnormal increase in levels of beta-amyloid (Aβ) and hyperphosphorylation of the tau protein (Mondragón-Rodríguez et al. 2012). Although AD is the most common form of dementia, its pathogenesis is still not well understood. Network modeling offers a unique opportunity for better understanding of AD by combining the current knowledge with a quantitative framework. The use of network biology in elucidating AD markers has been widely reported by various researchers like in one study performed by Ray et al. where severity across brain regions was examined by topological analysis of gene co-expression (Ray and Zhang 2010).

14.4 Available Databases/Resources/Computational Tools/ Servers for the Network Analysis

One of the challenges in network analysis is data visualization. Here we present several useful databases and software tools that exist for network analysis. First we introduce the databases that we categorized in four groups: transcriptional regulation pathways, protein–protein interaction pathways, signal transduction pathways, and metabolic pathways (Table 14.1). Thereafter we present network biology tools for network analysis and visualization (Table 14.2). With the advancement of techniques and availability of the data, a myriad of such resources and tools have been developed as shown in Fig. 14.1. We have compiled the list based upon accuracy and applications.

14.5 Current Status of These Tools and Their Future Enhancements

We have discussed a wide variety of tools that encompass features essential for network visualization. As observed from various studies, three major challenges are faced while performing data visualization, i.e., large amount of data, heterogeneous data integration, and the representation of multiple linkages between nodes with heterogeneous biological meaning (Pavlopoulos et al. 2008). Each visualization tool differs from others in terms of specific features they possess and therefore tackles the aforementioned challenges in its own way. Regarding the heterogeneity tools, Ondex, Pivot, or Medusa offers some possible solutions. However, Medusa and other tools that can handle multi-edged networks are also used when working with systems biology data such as in case of highly interlinked nodes. Cytoscape or BioLayout Express3D tools have good resolution and scaling features which further augment the visualization process. Pajek is ideal for pattern recognition and for studying the properties like density, centrality, and frequency of nodes. Osprey is suitable for comparative biological analysis. The tools presented in this chapter have a wide range of applicability in the network-related problems.

Besides a wide range of applicability of the tools, there are some drawbacks associated with them. Firstly, the majority of tools can handle datasets only up to a certain limit. As the size of datasets increases rapidly, there is a need for new generations of visualization tools that can withstand this problem. Secondly, there is a scaling problem posting a challenge to this field. Although many algorithms have been developed, they are still not able to deal with the layout problem and follow a heuristic approach instead of exhaustive ones. Therefore, there is an urgent need to develop fast and more efficient algorithms for speedy analysis of large-scale networks. One possible way to evade this problem is the parallel processing that makes use of powerful machines which greatly speed up the process of visualization and hence reduce the computational load. In addition, layout can be extended

Table 14.1 Tools and resources for the analysis of transcriptional regulation pathway, protein–protein interaction pathways, signal transduction pathways, and metabolic pathways

Transcriptional regulation pathway		
PAZAR	http://www.pazar.info/	A public database of transcription factor and regulatory sequence annotation
RegulonDB	regulondb.ccg.unam.mx	A reference database of *Escherichia coli* K-12 offering curated knowledge of the regulatory network and operon organization
TRANSFAC	http://transfac.gbf.de/	A manually curated database of eukaryotic transcription factors, their genomic-binding sites and DNA-binding profiles
TRRUST	http://www.grnpedia.org/trrust/	A reference database of human transcriptional regulatory interactions
YTRP (Yeast Transcriptional Regulatory Pathway) Database	http://cosbi3.ee.ncku.edu.tw/YTRP/	A repository for yeast transcriptional regulatory pathways
Protein–protein interaction pathways		
BIND/BOND (Biomolecular Interaction Network Database)	http://bind.ca	Archives biomolecular interaction, complex and pathway information
BioGRID (Biological General Repository for Interaction Datasets)	http://thebiogrid.org/	A repository for set of physical and genetic interactions that include interactions, chemical associations, and post-translational modifications (PTM)
CYGD (Comprehensive Yeast Genome Database)	http://mips.gsf.de/genre/proj/yeast/index.jsp	Present information on the molecular structure and functional network of the entirely sequenced, well-studied model eukaryote, the budding yeast *Saccharomyces cerevisiae*
DIP (Database of Interacting Proteins)	http://dip.doembi.ucla.edu/	Catalogs experimentally determined interactions between proteins
HPRD (Human Protein Reference Database)	http://www.hprd.org/	A web-based resource for protein–protein interactions, posttranslational modifications, enzyme–substrate relationships, and disease associations
MINT (Molecular INTeraction) Database	http://mint.bio.uniroma2.it/mint/Welcome.do	A public repository for protein–protein interactions (PPI)
STRING (Search Tool for the Retrieval of Interacting Genes/Proteins)	http://string-db.org/	A database of known and predicted protein interactions
Signal transduction pathways		
BBID (Biological Biochemical Image Database)	http://bbid.grc.nia.nih.gov/	A searchable database of images of putative biological pathways, macromolecular structures, gene families, and cellular relationships

(continued)

Table 14.1 (continued)

CSNDB (Cell Signaling Networks Database)	http://geo.nihs.go.jp/csndb/	A data and knowledge base for signaling pathways of human cells. It compiles the information on biological molecules, sequences, structures, functions, and biological reactions which transfer the cellular signals
SPAD (Signaling PAthway Database)	http://www.grt.kyushu-u.ac.jp/spad/	An integrated database for genetic information and signal transduction systems
TransPath	http://transpath.gbf.de/	A database system about gene regulatory networks that combines encyclopedic information on signal transduction with tools for visualization and analysis
Metabolic pathways		
BioCyc	http://www.biocyc.org/	The BioCyc database collection is a set of more than 7600 pathway/genome databases (PGDBs) describing many sequenced genomes
BIOPATH (Biochemical Pathways) Database	http://www.mol-net.de/databases/biopath.html	A database of biochemical pathways that provides access to metabolic transformations and cellular regulations
ECMDB (E. coli Metabolome database)	http://ecmdb.ca/	An expertly curated database containing extensive metabolomic data and metabolic pathway diagrams about *Escherichia coli* (strain K12, MG1655)
EMP (Enzymes and Metabolic Pathways) Database	http://emp.mcs.anl.gov/	A database on the biochemistry of some 1800 different organisms
GMD (Golm Metabolome Database)	http://gmd.mpimp-golm.mpg.de/	Facilitates the search for and dissemination of reference mass spectra from biologically active metabolites quantified using gas chromatography (GC) coupled to mass spectrometry (MS)
HMDB (Human Metabolome Database)	http://www.hmdb.ca/	A freely available electronic database containing detailed information about small molecule metabolites found in the human body
KEGG (Kyoto Encyclopedia of Genes and Genomes)	http://www.genome.ad.jp/kegg/	A collection of databases dealing with genomes, biological pathways, diseases, drugs, and chemical substances
Metacyc	http://metacyc.org/	A curated database of experimentally elucidated metabolic pathways from all domains of life
MANET (Molecular Ancestry Network)	http://www.manet.illinois.edu/index.php	A database tracing the evolution of protein architecture in metabolic networks
PathCase (Pathways Database System)	http://nashua.case.edu/pathwaysweb/	Store, query, and visualize metabolic pathways, in addition to their specialized tasks

(continued)

Table 14.1 (continued)

PMN (Plant Metabolic Network)	http://www.plantcyc.org/	A broad network of plant metabolic pathway databases that contain curated information from the literature and computational analyses about the genes, enzymes, compounds, reactions, and pathways involved in primary and secondary metabolism in plants
UniPathway	http://www.unipathway.org/	A resource for the exploration of metabolic pathways

The left most column has been arranged alphabetically

Table 14.2 Important and popular tools and resources for the analysis of biological networks

Network biology tools and resources		
Arena3D	http://arena3d.org/	Use multilayered graphs to visualize biological networks. In such a way, heterogeneous data will be distinguished between each other
BioLayout Express3D	http://www.biolayout.org/	Offers different analytical approaches to microarray data analysis
BioTapestry	http://www.biotapestry.org/	A tool to visualize the dynamic properties of gene regulatory networks
CellDesigner	http://www.celldesigner.org/	A structured diagram editor for drawing gene regulatory and biochemical networks
CellML	https://www.cellml.org/	An XML-based markup language for describing mathematical models
COPASI	http://copasi.org/	COPASI is a software application for simulation and analysis of biochemical networks and their dynamics
CSB.DB	http://csbdb.mpimp-golm.mpg.de/csbdb/dbcor/cor.html	A comprehensive systems biology database
Cytoscape	http://www.cytoscape.org/	Incorporates statistical analysis of the network, and it makes it easy to cluster or detect highly interconnected regions
EAWAG-BBD	http://eawag-bbd.ethz.ch/	Contains information on microbial biocatalytic reactions and biodegradation pathways for primarily xenobiotic, chemical compounds
E-Cell	http://www.e-cell.org/	Develops general technologies and theoretical supports for computational biology with the grand aim to make precise whole cell simulation at the molecular level possible
FANMOD	http://theinf1.informatik.uni-jena.de/~wernicke/motifs/index.html	A tool for fast network motif detection
Genes2Networks	http://actin.pharm.mssm.edu/genes2networks/	Powerful web-based software that can help experimental biologists to interpret lists of genes and proteins such as those commonly produced through genomic and proteomic experiments, as well as lists of genes and proteins associated with disease processes

(continued)

Table 14.2 (continued)

Network biology tools and resources		
GEPHI	https://gephi.org/	Interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs
Igraph	http://igraph.org/	A collection of network analysis tools with the emphasis on efficiency, portability, and ease of use
JWS Online	http://omictools.com/jws-online-tool	A systems biology tool for the construction, modification, and simulation of kinetic models and for the storage of curated models
Medusa	http://coot.embl.de/medusa/	Medusa is optimized for protein–protein interaction data as taken from STRING or protein–chemical and chemical–chemical interactions as taken from STITCH
Ondex	http://ondex.sourceforge.net/	Ondex main strength is the ability to combine heterogeneous data types into one network. It is suitable for text mining and sequence and data integration analysis
Osprey	http://biodata.mshri.on.ca/osprey/servlet/Index	The ability to incorporate new interactions into an already existing network
Pajek	http://pajek.imfm.si/doku.php?id=pajek	Main strength is the variety of layout algorithms which greatly facilitate exploration and pattern identification within networks
PATIKA (Acquisition)	http://www.patika.org/	Integrated software environment designed to provide researchers a complete solution for modeling and analyzing cellular processes
PIVOT	http://acgt.cs.tau.ac.il/pivot/	Best suited for visualizing protein–protein interactions and identifying relationships between them
ProViz	http://cbi.labri.fr/eng/proviz.htm	Performs protein–protein interaction and their analysis using arbitrary properties, like for example annotations or taxonomic identifier
Reactome	http://www.reactome.org/	A free, open-source, curated, and peer-reviewed pathway database
VisANT	http://visant.bu.edu	An online visualization and analysis tool for biological interaction data

The left most column has been arranged alphabetically

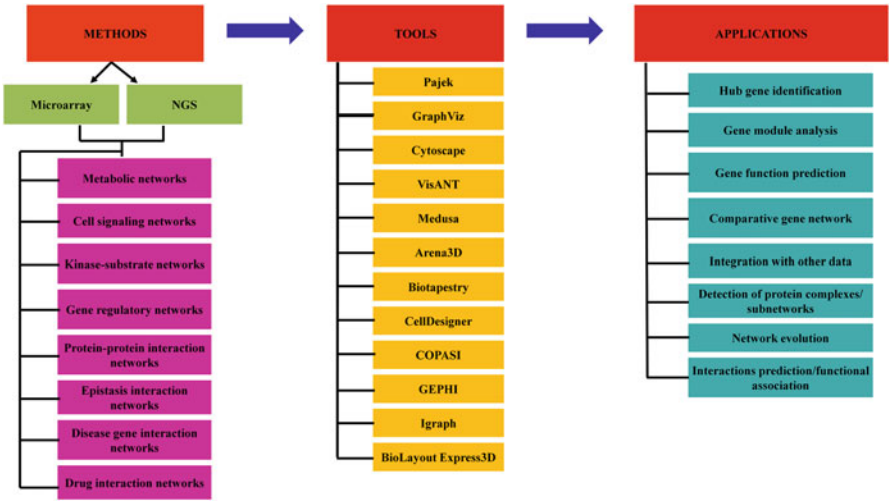

Fig. 14.1 Flowchart depicting the methods, tools, and applications of network biology

by adding a third dimension which would allow a clearer visualization of structures and strongly facilitate a better navigation within the network. Currently, most of the network tools only generate static snapshots of the interactions and provide no methods to visualize a time series data (Suderman and Hallett 2007). By introducing time series data, the process of network visualization would thus achieve a more complete picture of complex and highly dynamic biological systems. It is highly expected that this will provide breakthroughs in the pathway analysis process or the observation of interaction at different time points of cell cycles.

Systems memory is another important issue that should be taken care of while performing computational visualization and analysis. The limited functionalities of existing visualization tools make it necessary to constantly switch between different applications to complete different levels of analysis. Frequent information and data sharing between different tools has become possible with the availability of standard file formats. The visualization tools designed by taking care of aforementioned functionalities would greatly simplify large-scale research in molecular biology and would significantly cut down time and effort spent on data processing and analysis (Pavlopoulos et al. 2008).

14.6 Conclusion

Network biology is the revolution in the field of life science as it provides information not only on the significant interactions but also on the functionalities allied to them. The network complexities can be studied with the help of a variety of

methods available in network biology, and the most significant of them is the module-based approach. Modularity focuses on the local significant regions that share high-functional relationships comparative to the rest of the network. These methods not only uncover the complex biological mysteries but also help investigators understand various disease networks or their interactions and hence provide the potential therapeutic targets. Network biology has accelerated the biomolecular analysis process by offering different tools that have diverse applications depending on the number of features embed in them. Network-based methods also have several limitations including the lack of mechanistic explanations. Despite the limitations, network analysis has been applied successfully to understand the complexities of many disease states. It is anticipated that with rapid advancements, network biology will serve as an excellent research complement to annotate biomolecules at a system level and will also help in the generation of more biologically meaningful information.

References

Aryee MJ, Liu W, Engelmann JC, Nuhn P, Gurel M, Haffner MC. DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. Sci Transl Med. 2013;5(169):169ra10.

Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004;5(2):101–13.

Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011;12(1):56–68.

Bergholdt R, Størling ZM, Lage K, Karlberg EO, Olason PI, Aalund M. Integrative analysis for finding genes and networks involved in diabetes and other complex diseases. Genome Biol. 2007;8(11):R253.

Cevenini E, Bellavista E, Tieri P, Castellani G, Lescai F, Francesconi M. Systems biology and longevity: an emerging approach to identify innovative anti-aging targets and strategies. Curr Pharm Des. 2010;16(7):802–13.

Chang X, Xu T, Li Y, Wang K. Dynamic modular architecture of protein-protein interaction networks beyond the dichotomy of 'date' and 'party' hubs. Sci Rep. 2013;3:1691.

Cho DY, Kim YA, Przytycka TM. Chapter 5: network biology approach to complex diseases. PLoS Comput Biol. 2012;8(12):e1002820.

Davis RC, van Nas A, Castellani LW, Zhao Y, Zhou Z, Wen P. Systems genetics of susceptibility to obesity-induced diabetes in mice. Physiol Genomics. 2012;44(1):1–13.

Diaz-Beltran L, Cano C, Wall DP, Esteban FJ. Systems biology as a comparative approach to understand complex gene expression in neurological diseases. Behav Sci (Basel). 2013;3(2):253–72.

Diez D, Wheelock AM, Goto S, Haeggström JZ, Paulsson-Berne G, Hansson GK. The use of network analyses for elucidating mechanisms in cardiovascular disease. Mol BioSyst. 2010;6(2):289–304.

Kim W, Li M, Wang J, Pan Y. Biological network motif detection and evaluation. BMC Syst Biol. 2011;5(Suppl 3):S5.

Kitano H. Systems biology: a brief overview. Science. 2002;295(5560):1662–4.

Kristiansson K, Naukkarinen J, Peltonen L. Isolated populations and complex disease gene identification. Genome Biol. 2008;9(8):109.

Liu ZP, Wang Y, Zhang XS, Chen L. Network-based analysis of complex diseases. IET Syst Biol. 2012;6(1):22–33.
Liu R, Wang X, Aihara K, Chen L. Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. Med Res Rev. 2014;34(3):455–78.
Ma'ayan A. Introduction to network analysis in systems biology. Sci Signal. 2011;4(190):tr5.
Masoudi-Nejad A, Schreiber F, Kashani ZR. Building blocks of biological networks: a review on major network motif discovery algorithms. IET Syst Biol. 2012;6(5):164–74.
Mitchell KJ. What is complex about complex disorders? Genome Biol. 2012;13(1):237.
Mitra K, Carvunis AR, Ramesh SK, Ideker T. Integrative approaches for finding modular structure in biological networks. Nat Rev Genet. 2013;14(10):719–32.
Mondragón-Rodríguez S, Perry G, Zhu X, Boehm J. Amyloid beta and tau proteins as therapeutic targets for Alzheimer's disease treatment: rethinking the current strategy. Int J Alzheimers Dis. 2012;2012:630182.
Park D, Park J, Park SG, Park T, Choi SS. Analysis of human disease genes in the context of gene essentiality. Genomics. 2008;92(6):414–8.
Pavlopoulos GA, Wegener AL, Schneider R. A survey of visualization tools for biological network analysis. BioData Min. 2008;1:12.
Przulj N, Corneil DG, Jurisica I. Modeling interactome: scale-free or geometric? Bioinformatics. 2004;20(18):3508–15.
Ray M, Zhang W. Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. BMC Syst Biol. 2010;4:136.
Rende D, Baysal N, Kirdar B. A novel integrative network approach to understand the interplay between cardiovascular disease and other complex disorders. Mol BioSyst. 2011;7(7):2205–19.
Sarajlić A, Pržulj N. Survey of network-based approaches to research of cardiovascular diseases. Biomed Res Int. 2014;2014:527029.
Schadt EE. Molecular networks as sensors and drivers of common human diseases. Nature. 2009;461(7261):218–23.
Sehgal M, Gupta R, Moussa A, Singh TR. An integrative approach for mapping differentially expressed genes and network components using novel parameters to elucidate key regulatory genes in colorectal cancer. PLoS One. 2015;10(7):e0133901.
Shukla A, Sehgal M, Singh TR. Hydroxymethylation and its potential implication in DNA repair system: a review and future perspectives. Gene. 2015;564(2):109–18.
Shukla A, Moussa A, Singh TR. DREMECELS: a curated database for base excision and mismatch repair mechanisms associated human malignancies. PLoS One. 2016;11(6):e0157031.
Srinivasa RV, Srinivas K, Sujini NG, Kumar SNG. Protein-protein interaction detection: methods and analysis. Int J Proteomics. 2014;2014:12. Article ID 147648.
Suderman M, Hallett M. Tools for visually exploring biological networks. Bioinformatics. 2007;23(20):2651–9.
Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. Cell. 2011;144(6):986–98.
Wall DP, Esteban FJ, Deluca TF, Huyck M, Monaghan T, Velez de Mendizabal N. Comparative analysis of neurological disorders focuses genome-wide search for autism genes. Genomics. 2009;93(2):120–9.
Winterbach W, Van Mieghem P, Reinders M, Wang H, de Ridder D. Topology of molecular interaction networks. BMC Syst Biol. 2013;7:90. doi:10.1186/1752-0509-7-90.
Zhu X, Gerstein M, Snyder M. Getting connected: analysis and principles of biological networks. Genes Dev. 2007;21(9):1010–24.