Chapter 34

Modules in Biological Networks: Identification and Application

Bing Zhang
Vanderbilt University School of Medicine, USA

Zhiao Shi
Vanderbilt University, USA

ABSTRACT

One of the most prominent properties of networks representing complex systems is modularity. Network-based module identification has captured the attention of a diverse group of scientists from various domains and a variety of methods have been developed. The ability to decompose complex biological systems into modules allows the use of modules rather than individual genes as units in biological studies. A modular view is shaping research methods in biology. Module-based approaches have found broad applications in protein complex identification, protein function prediction, protein expression prediction, as well as disease studies. Compared to single gene-level analyses, module-level analyses offer higher robustness and sensitivity. More importantly, module-level analyses can lead to a better understanding of the design and organization of complex biological systems.

INTRODUCTION

The twentieth-century biology has been focused on individual cellular components and their functions. Despite the huge success of this approach, a discrete biological function can rarely be attributed to an individual molecule (Hartwell et al., 1999). It is increasingly clear that the cell can be understood as a complex network of interacting components (Barabasi & Oltvai, 2004). Unraveling the interactions between the components of a cell constitutes a major goal of the post-genomic era.

With recent advances in high-throughput experimental technologies, genomic data are now available for the reconstruction of large-scale biological networks, in which nodes are biological molecules (e.g. proteins, genes, metabolites, microRNAs, etc.) and edges are functional relationships among the molecules (e.g. protein interactions, genetic interactions, transcriptional regu-
modules. The ability to decompose complex biological systems into modules allows the use of modules rather than individual genes as units in biological studies. Module-based analyses have several advantages over gene-based methods, including improved robustness against the inherent noise that exists in the sample population and increased sensitivity in identifying patterns that are too subtle to discern when considering individual genes (Chuang et al., 2007; Mootha et al., 2003). More importantly, module-based analyses can achieve a higher-level understanding of the design and organization of biological systems (Gavin et al., 2006; Segal et al., 2004).

In this review, we first summarize module identification algorithms that have been developed in different research communities. Next, we provide examples of module-based applications in biological research, including protein complex identification, protein function prediction, protein expression prediction, and disease studies.

**COMPUTATIONAL METHODS FOR MODULE IDENTIFICATION**

Module identification in networks is the process of grouping network elements that are similar in some way into different sets, i.e. modules. Depending on different research communities, modules are also called communities or clusters. Loosely stated, a module is a cohesive set of nodes that are connected "more tightly" to each other than to other nodes in the graph. Despite of various attempts, there is no universally accepted quantitative definition of a module. In many cases, it is induced by the algorithms without a precise a priori definition. Not surprisingly, many module identification algorithms have been proposed based on their needs in different scientific domains (Fortunato, 2009). In this section, we present representative module identification methods categorized according to the approaches they take (Table 1).
Related Content

Normalizing Cross-Border Healthcare in Europe via New E-Prescription Paradigms
[www.igi-global.com/chapter/normalizing-cross-border-healthcare-in-europe-via-new-e-prescription-paradigms/78697?camid=4v1a](www.igi-global.com/chapter/normalizing-cross-border-healthcare-in-europe-via-new-e-prescription-paradigms/78697?camid=4v1a)

Efficient Computational Construction of Weighted Protein-Protein Interaction Networks Using Adaptive Filtering Techniques Combined with Natural Selection-Based Heuristic Algorithms
[www.igi-global.com/article/efficient-computational-construction-weighted-protein/67104?camid=4v1a](www.igi-global.com/article/efficient-computational-construction-weighted-protein/67104?camid=4v1a)

Pattern Differentiations and Formulations for Heterogeneous Genomic Data through Hybrid Approaches
[www.igi-global.com/chapter/pattern-differentiations-formulations-heterogeneous-genomic/4250?camid=4v1a](www.igi-global.com/chapter/pattern-differentiations-formulations-heterogeneous-genomic/4250?camid=4v1a)

Clustering Genes Using Heterogeneous Data Sources
[www.igi-global.com/article/clustering-genes-using-heterogeneous-data/45163?camid=4v1a](www.igi-global.com/article/clustering-genes-using-heterogeneous-data/45163?camid=4v1a)