

Chapter 2.13

Module Finding Approaches for Protein Interaction Networks

Tero Aittokallio
University of Turku, Finland

ABSTRACT

This chapter provides an overview of the computational approaches developed for exploring the modular organization of protein interaction networks. A special emphasis is placed on the module finding tools implemented in three freely available software packages, VisANT, Cytoscape and MATISSE, as well as on their biomedical applications. The selected methods are presented in the broader context of module discovery options, ranging from approaches that rely merely on topological properties of the underlying network to those that take into account also other comple-

mentary data sources, such as the mRNA levels of the proteins. The author will also highlight some current limitations in the measured network data that should be understood when developing and applying module finding methodology, and discuss some key future trends and promising research directions with potential implications for clinical research.

INTRODUCTION

Recent advances in experimental technologies and computational methods have made it possible to measure and predict protein-protein interactions on a global scale (see Chapters I-VI and

DOI: 10.4018/978-1-60960-561-2.ch213

Shoemaker & Panchenko, 2007a, b). While the large-scale interaction datasets can provide an unprecedented glimpse into the cellular mechanisms underlying the behavior of various biological systems, the increasing sizes and densities of the protein interaction networks available today pose also many challenging computational problems. In particular, the inherent complexity of most biological processes and the large number of possible interactions involved can make it difficult to interpret and mine the interaction networks only by eye, even with the help of sophisticated visualization and layout tools available. Software packages that implement computational tools for more advanced network data mining can facilitate the explorative network analysis by identifying the key players and their interactions that contribute to the cellular processes of interest. This may allow e.g. to pinpoint errors in experimentally or computationally derived interaction links, identify proteins directly involved in the particular process, and to formulate hypotheses for follow-up experiments. It should be realized, however, that even if the rapidly developing complex network theory has successfully been applied to analysis of various social and technological networks, such as the Internet and computer chips, its impact on studying molecular interaction networks is still an emerging area of research and therefore all these methods should be considered as experimental. At the moment, the computational tools are best used together with network visualization and analysis software that enable interactive and fully-controlled mining of the complex protein interaction networks.

One of the most fundamental properties found in many biological networks is their modular organization (Hartwell et al., 1999). Consequently, the decomposition of large networks into a hierarchy of possible overlapping sub-networks (so-called *modules*) has become a principal analytical approach to deal with the complexity of large cellular networks (Barabási & Oltvai, 2004). In protein interaction networks, a functional module refers

to a group of physically connected proteins that work together to carry out a specific cellular function in a particular spatio-temporal context. A large number of computational tools of increasing complexity have recently been developed for investigating the modular organization of interaction networks. These tools cannot only identify whether a given network is modular or not, but also detect the modules and their inter-relationships in the underlying network. By relating the found sub-networks with complementary functional genomics or proteomics data, such as gene expression profiles from genome-wide microarray experiments or protein abundance measurements from mass-spectrometry-based assays, it is also possible to identify a hierarchy of connected groups of components that show coherent expression patterns. Such functionally organized modules cannot only emphasize the biological meaning of the modules discovered, but also allow us to gradually focus on the active subsystems of particular interest (*active modules*), which can lead to concrete hypotheses about the regulatory mechanisms and pathways most important for the given process (Ideker et al., 2002). Moreover, these modules can subsequently be used in predictive modeling studies, with the aim of suggesting new biological hypotheses, such as unexplored new interactions or the function of individual components, or even distinguishing different biomedical phenotypes (*discriminative modules*).

While several computational approaches to discovering functional modules in protein interactions networks are available to molecular biologists, and new ones are constantly being developed by the methodologists, there is no general consensus on how to choose between the different strategies and definitions. The choice of the network mining approach is an important aspect which depends both on the network and questions under the analysis. Therefore, basic understanding of the graph-theoretic concepts and algorithms behind the software packages

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/module-finding-approaches-protein-interaction/53599

Related Content

Wearable Kinesthetic System for Joint Knee Flexion-Extension Monitoring In Gait Analysis

Mario Tesconi, Enzo Pasquale Scilingo, Pierluigi Barba and Danilo De Rossi (2011). *Clinical Technologies: Concepts, Methodologies, Tools and Applications* (pp. 792-800).

www.irma-international.org/chapter/wearable-kinesthetic-system-joint-knee/53620

Healthcare Information Systems Research: Who is the Real User?

Alexander J. McLeod and Jan G. Clark (2011). *New Technologies for Advancing Healthcare and Clinical Practices* (pp. 192-210).

www.irma-international.org/chapter/healthcare-information-systems-research/55145

Magnetic Resonance Imaging and the Signal-Image Processing Techniques Developed Under the Umbrella of the Unifying Theory

Carlo Ciulla (2009). *Improved Signal and Image Interpolation in Biomedical Applications: The Case of Magnetic Resonance Imaging (MRI)* (pp. 1-21).

www.irma-international.org/chapter/magnetic-resonance-imaging-signal-image/22487

Computer Assisted Cervical Cytology

Liron Pantanowitz (2009). *Medical Informatics in Obstetrics and Gynecology* (pp. 160-184).

www.irma-international.org/chapter/computer-assisted-cervical-cytology/26191

Optimization of Medical Supervision, Management, and Reimbursement of contemporary Home Care

B. Spyropoulos, M. Botsivaly, A. Tzavaras and K. Koutsourakis (2011). *Clinical Technologies: Concepts, Methodologies, Tools and Applications* (pp. 1674-1683).

www.irma-international.org/chapter/optimization-medical-supervision-management-reimbursement/53674