

Chapter XVI

Graphical Analysis and Visualization Tools for Protein Interaction Networks

Sirisha Gollapudi*

MyCIB, University of Nottingham, UK

Alex Marshall*

MyCIB, University of Nottingham, UK

Daniel Zadik*

MyCIB, University of Nottingham, UK

Charlie Hodgman

MyCIB, University of Nottingham, UK

** Joint first authors*

ABSTRACT

Software for the visualization and analysis of protein-protein interaction (PPI) networks can enable general exploration, as well as providing graph-theoretic algorithms for specific tasks. Analyses can include reduction of complexity or the scope of the network in order to make it more manageable, or increase in complexity by integration with other datasets, to represent biology more accurately. Two software approaches are outlined in this chapter: desktop applications and web services. Desktop applications have attractive user interfaces with a wide range of analysis tools, and often capabilities for integration of other bio-molecular data. Web services provide a newer approach to network analysis. They have the advantages of a broader range of potential functionalities and a more extensible framework than standalone desktop tools. However, their relative infancy means that they are not as well developed. This chapter provides an evaluation of some common desktop applications, compared to and contrasted with several examples of web services.

INTRODUCTION

In recent years, a predominant theme in biological research has been a shift from the reductionist approach to an integrative systems approach (Kitano *et al.*, 2002; Oltvai & Barabási, 2002; Han, 2008), which views the components of a cell as acting together in a network of reactions and interactions. This change has been aspired to for

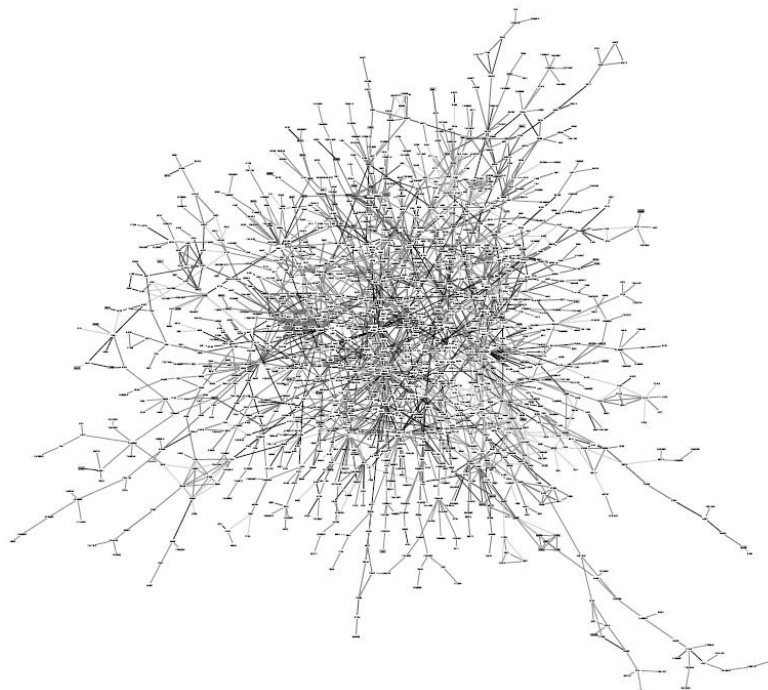
Graphical Analysis and Visualization Tools for Protein Interaction Networks

many years, as it is widely understood that observed cell behaviour is rarely attributed to one component acting alone. The continued growth in the processing power of computers and the increase in availability of biological data from high throughput technologies have allowed such an approach to become realistic.

This chapter describes software that uses algorithms taken from the field of graph theory to draw biological conclusions about protein-protein interaction (PPI) networks. It contains background on the subject, a method for the representation of networks, including bio-molecular networks, as a set of nodes connected by edges. In PPI networks, the nodes represent proteins, while the edges represent the interactions between them. Very often, PPI networks aim to represent the whole proteome of a species. As an example, Figure 1 shows a holistic PPI network for the budding yeast, *Saccharomyces cerevisiae* (Schwikowski, Uetz & Fields, 2000). Holistic PPI networks are usually constructed from two-hybrid screening (Fields & Song, 1989) and other experimental data, much of which is available in public databases (see Background).

Two-hybrid technology is particularly powerful as it identifies large numbers of putative interactions. To test if two proteins interact, the DNA-binding domain from a transcription factor is spliced onto the end of the gene for one, and the activating domain from the transcription factor is spliced onto the other. When the genes are transformed into a cell (usually in *S. cerevisiae* or *E. coli*) and expressed, two hybrid proteins are produced. The protein with the DNA-binding domain binds to an upstream activating sequence (UAS) of a reporter gene and the protein with the activating domain binds the remaining transcriptional machinery. If the two proteins bind each other then the transcriptional machinery will be brought into close proximity with the UAS, and the reporter gene will be expressed. Whole libraries of genes can be transformed into a cell suspension, to create a population with a very diverse set of combinations of hybrid genes. If a reporter gene is used that is essential for survival under particular conditions, then those conditions can be used to select only those cells that contain

Figure 1. An *S. cerevisiae* holistic PPI network, containing 1548 proteins and 2358 interactions, from a study carried out by Schwikowski, Uetz and Fields, (2000). (Permission to reproduce figure kindly provided by Peter Uetz, Institute of Genetics, University of Karlsruhe.)



24 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/graphical-analysis-visualization-tools-protein/5570

Related Content

Integrated Features Based on Gray-Level and Hu Moment-Invariants with Ant Colony System for Retinal Blood Vessels Segmentation

Ahmed H. Asad, Ahmad Taher Azarand Aboul Ella Hassanien (2012). *International Journal of Systems Biology and Biomedical Technologies* (pp. 60-73).

www.irma-international.org/article/integrated-features-based-gray-level/75154

Identification Methods of G Protein-Coupled Receptors

Meriem Zekri, Karima Alemand Labiba Souici-Meslati (2011). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 35-52).

www.irma-international.org/article/identification-methods-protein-coupled-receptors/73910

Molecular Dynamics Simulation of Interlayer Structure and Hydration Properties of Glycine Intercalated Layered Double Hydroxides

Guo-Xiang Pan, Feng Cao, Pei-Song Tang, Hai-Feng Chen, Zhe-Ming Ni, Jin-Tian Yang, Li-Geng Wangand Min-Hong Xu (2011). *Interdisciplinary Research and Applications in Bioinformatics, Computational Biology, and Environmental Sciences* (pp. 242-250).

www.irma-international.org/chapter/molecular-dynamics-simulation-interlayer-structure/48381

Ethics and Privacy Considerations for Systems Biology Applications in Predictive and Personalized Medicine

Jake Y. Chen, Heng Xu, Pan Shi, Adam Culbertsonand Eric M. Meslin (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 1378-1404).

www.irma-international.org/chapter/ethics-privacy-considerations-systems-biology/76124

Clustering Genes Using Heterogeneous Data Sources

Erliang Zeng, Chengyong Yang, Tao Liand Giri Narasimhan (2012). *Computational Knowledge Discovery for Bioinformatics Research* (pp. 67-83).

www.irma-international.org/chapter/clustering-genes-using-heterogeneous-data/66705