

Chapter XI

Discovering Lethal Proteins in Protein Interaction Networks

Kar Leong Tew

Institute for Infocomm Research, Singapore

Xiao-Li Li

Institute for Infocomm Research, Singapore

ABSTRACT

This chapter introduces state-of-the-art computational methods which discover lethal proteins from Protein Interaction Networks (PINs). Lethal proteins are an interesting subject in understanding the minimal condition for cellular development and survival. A dysfunctional research subject or absence of a lethal protein would result in fatality of the cell. Biological experiments have been conducted to systematically detect such proteins. However, such processes are time consuming and requires huge amount of effort to conduct. The researchers have developed a series of computational methods which take advantage of the network properties of individual proteins to detect lethal proteins in PINs. In this chapter, each computational method is studied in depth with an analysis on its pros and cons. Finally, a discussion on the possible further research directions will conclude the chapter.

INTRODUCTION

Proteins play a key role in the operation and survival of a cell. They rarely function alone, but perform biological tasks through interactions with other biological entities such as DNA, RNA and other proteins. Interactions between proteins within an organism collectively form a **Protein Interaction Network (PIN)**. The most common representation of a PIN is a graph where proteins are depicted by nodes and an edge between two nodes indicates a detected interaction between the two proteins. A lethal (or essential) protein is one that renders the cell unviable upon its removal. Lethal proteins play an intricate role in the development and survival of the cell and their characteristics and detection is an interesting research topic in proteomics.

A protein's lethality has been used in various biological and medical researches in recent years. In the study of protein evolution and conservation, it was predicted that proteins differing in their importance (lethality) are subjected to different evolution rate (Kimura *et al.* 1974; Wilson *et al.* 1977). Their prediction has been put to test with the availability of large gene-knockout data by various research groups (Hurst *et al.* 1999; Liao *et al.* 2006;

Pal *et al.* 2003; Yang *et al.* 2003; Zhang *et al.* 2005). Due to its biological significance, even with a substantial number of genes with unknown lethality profiles (Jeong *et al.* 2003), lethal proteins are used as the focus of their work. The results of their investigations do agree that lethal proteins evolve at a different rate from normal proteins. However, there is a dispute over whether does lethal proteins evolve *faster* (to better adapt to changing environment) or *slower* (to avoid drastic changes).

Cross species studies have also utilized information pertaining to protein lethality. A study on the PIN of *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* indicates an association between protein evolution, centrality, and protein lethality (Hahn *et al.* 2005). Conservation of lethal proteins has also been detected between *Saccharomyces cerevisiae* and *Saccharomyces mikatae* (Seringhaus *et al.* 2006). Such work demonstrates the usefulness of lethal proteins in aiding biological research while investigating new species.

Associations between lethal proteins, disease, and human gene morbidity have also been established by various research groups. Steinmetz *et al.* (2002) used the *Saccharomyces cerevisiae* deletion mutants to identify 256 new human mitochondrial proteins with a fivefold greater selection than gene expression analysis. Kondrashov *et al.* (2004) first stated the close relationship between morbidity and protein lethality and further found that morbid genes are more similar to lethal proteins of *Drosophila melanogaster*. Furney *et al.* (2006) made a finer division of disease genes into dominant or recessive mutations and lethal proteins were found to have a higher correlation with dominant genes.

All the above mentioned research work reflects the biological insights and discovery related to lethal proteins. Currently lethal proteins are found based on gene knockout experiments which can be time-consuming and cost-prohibitive when the biological experiments are done on a large-scale basis. Because of this reason, a large number of proteins are still without known lethality information, making it desirable to formulate alternative detection approaches, e.g. computational methods, which can harness the knowledge of social network analysis, graph mining and biological information. Such computational approaches could help generating a ranked list of potential lethal proteins and the cost to detect lethal proteins is thus largely reduced.

The impact of removing a lethal protein from a PIN is similar to that of a “*Key Player*” in a social network where the node is strategically located within the network (Bonacich 1987; Everett *et al.* 1999; Stephen 2006). Their absence or dysfunctional would create an adverse disruption to the topological stability of the network which in the case of PIN – biological lethality. This laid the foundation where computation methods based on topological features are developed to better detect lethal proteins (Estrada 2006; Jeong *et al.* 2001; Joy *et al.* 2005; Schmith *et al.* 2005; Yu *et al.* 2004). Recently, there is also a new trend where biological information is integrated into their models (Saha *et al.* 2006; Tew *et al.* 2007; Ulitsky *et al.* 2007). The first category of computational methods can be broadly classified under “*Topological*” and the second category as “*Integrative*” approaches.

This chapter will provide an overview of the present state-of-the-art methods to detect the lethal proteins from protein interaction networks. The basic concepts for each individual technique will be first covered in the Methods section. This is followed by Definitions and Formulas Section which provides the detailed computational description for the different techniques and compares their strengths and weakness of each method. Finally, a discussion on the possible further research directions will conclude the chapter.

METHODS

There has been a series of work in developing computational methods to detect the lethal proteins from PIN. Each of this work is based on their own assumptions, i.e. because of some proteins have some important network properties or functionalities, their removal will thus disrupt the topological stability or impair the biological functions of the cell (they are thus essential to the cell). These network properties are also known as the **centrality** of the node. It accounts for the different node characteristics that permit them to be ranked in order of importance in the network.

Two broad categories can be used to classify the available computational prediction methods, a) “*Topological approaches*” which is solely based on the network characteristic, without using other biological information and b) “*Integrative approaches*” which make use of various biological information and properties. The details of each method would be discussed in this section while the comparisons will be done later.

18 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/discovering-lethal-proteins-protein-interaction/5565

Related Content

Perspective Wall Technique for Visualizing and Interpreting Medical Data

Hela Ltfi, Mounir Ben Ayed, Ghada Trabelsi and Adel M. Alimi (2012). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 45-61).

www.irma-international.org/article/perspective-wall-technique-visualizing-interpreting/77930

Recognition of Translation Initiation Sites in Arabidopsis Thaliana

Haitham Ashoor, Arturo M. Mora, Karim Awara, Boris R. Jankovic, Rajesh Chowdhary, John A.C. Archer and Vladimir B. Bajic (2012). *Systemic Approaches in Bioinformatics and Computational Systems Biology: Recent Advances* (pp. 105-116).

www.irma-international.org/chapter/recognition-translation-initiation-sites-arabidopsis/60830

Spatial Uncertainty Analysis in Ecological Biology

Stelios Zimeras and Yiannis Matsinos (2013). *International Journal of Systems Biology and Biomedical Technologies* (pp. 14-24).

www.irma-international.org/article/spatial-uncertainty-analysis-ecological-biology/78389

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

Ahmed H. Asad, Ahmad Taher Azar and 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

Deep Learning in Bioinformatics: Principles and Applications

Alessandra G. Cioletti, Frederico C. Carvalho, Lucas M. Dos Santos and Raquel C. M. Minardi (2025). *Effective Techniques for Bioinformatic Exploration* (pp. 137-168).

www.irma-international.org/chapter/deep-learning-in-bioinformatics/361321