



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Physica A 357 (2005) 501–512

PHYSICA A

[www.elsevier.com/locate/physa](http://www.elsevier.com/locate/physa)

# Graph theoretic analysis of protein interaction networks of eukaryotes

K.-I. Goh<sup>a,\*</sup>, B. Kahng<sup>a,b</sup>, D. Kim<sup>a</sup>

<sup>a</sup>*School of Physics, Seoul National University, Seoul 151-747, Republic of Korea*

<sup>b</sup>*Program in Bioinformatics, Seoul National University, Seoul 151-747, Republic of Korea*

Received 2 March 2005

Available online 23 May 2005

---

## Abstract

Owing to the recent progress in high-throughput experimental techniques, the datasets of large-scale protein interactions of prototypical multicellular species, the nematode worm *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*, have been assayed. The datasets are obtained mainly by using the yeast hybrid method, which contains false-positive and false-negative simultaneously. Accordingly, while it is desirable to test such datasets through further wet experiments, here we invoke recent developed network theory to test such high-throughput datasets in a simple way. Based on the fact that the key biological processes indispensable to maintaining life are conserved across eukaryotic species, and the comparison of structural properties of the protein interaction networks (PINs) of the two species with those of the yeast PIN, we find that while the worm and yeast PIN datasets exhibit similar structural properties, the current fly dataset, though most comprehensively screened ever, does not reflect generic structural properties correctly as it is. The modularity is suppressed and the connectivity correlation is lacking. Addition of interologs to the current fly dataset increases the modularity and enhances the occurrence of triangular motifs as well. The connectivity correlation function of the fly, however, remains distinct under such interolog additions, for which we present a possible scenario through an in silico modeling.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Protein interaction network; Modularity; Degree correlation

---

\*Corresponding author.

*E-mail address:* [vroom@soeh.snu.ac.kr](mailto:vroom@soeh.snu.ac.kr) (K.-I. Goh).

## 1. Introduction

In the last few years, graph theoretic methods to understand complex biomolecular systems have been developed very rapidly [1]. Such a development has made advances toward uncovering the organizing principles of cellular networks in post-genomic biology. The cellular components such as genes, proteins, and other biological molecules, connected by all physiologically relevant interactions, form a full web-like molecular architecture in a cell. In such an architecture, genes which are expressed through proteins play a central role. Proteins rarely act alone, rather they cooperate with others to act physiologically. Thus, protein interactions play pivotal roles in various aspects of the structural and functional organizations and their complete description would be the first step toward a thorough understanding of the web of life. Proteins are viewed as nodes of a complex protein interaction network (PIN) in which two proteins are linked if they physically contact each other. The graph theoretic approach has been useful to understand intricate interwoven structures of the PIN [2–4]. The key biological processes indispensable to maintaining life are robust across eukaryotic species since many involved genes are evolutionarily conserved [5]. Using this property, one can test a newly discovered dataset if it really contains more or less complete information of protein interactions. Moreover, this *in silico* approach offers one the candidates of protein interaction pairs, of which the number is considerably reduced compared with the total combinatorial pairs. Thus, the graphic theoretic analysis would provide a useful guide for further wet studies of protein interactions.

Species with sequenced genome such as the yeast *Saccharomyces cerevisiae* provide important test beds for the study of the PIN. Owing to the recent progress in the high-throughput experimental techniques such as the yeast two-hybrid assay [6,7] and the mass spectroscopy [8,9], the dataset of the yeast PIN has been firmly established [10,11]. Very recently, large-scale protein interactions of multicellular species, the nematode worm *Caenorhabditis elegans* [12] and the fruit fly *Drosophila melanogaster* [13], have been assayed. While those datasets, mainly based on the yeast two-hybrid assay, need physiological proof, they contain large-scale proteins and protein interactions, making graph theoretic study possible. In this paper, we analyze those datasets and compare them with the more-established set of interactions in the budding yeast [11]. Our graph theoretic analysis suggests that the present interaction dataset of the fruit fly, based on the yeast two-hybrid (Y2H) assay, may have left out a significant part of protein interactions, though most comprehensively screened ever. Such a conclusion was reached by the comparison of the generic features of the PIN, the modularity and connectivity correlations, across the three species. For the fly, those quantities behave distinctively: The modularity is suppressed and the connectivity correlation is lacking. Such distinct behavior can be overcome partially by the addition of yeast interologs into the fly dataset.

## 2. Materials and methods

### 2.1. Graph theory terminology

Till recently, it has been well known that the number of connections of a given protein in the PIN is inhomogeneous and its distribution, called the degree

Download English Version:

<https://daneshyari.com/en/article/9727543>

Download Persian Version:

<https://daneshyari.com/article/9727543>

[Daneshyari.com](https://daneshyari.com)