

Functional Modules from Protein Networks of Kinome and Cell Cycle in *Saccharomyces cerevisiae*

Hee Young Kang and Hak Yong Kim

Division of Life Sciences, College of Natural Sciences, Chungbuk National University, 361-763
Republic of Korea hykim@chungbuk.ac.kr

Abstract

We have constructed protein networks of kinome and cell cycle from MIPS database. The networks show character of scale free network, in which a few highly connected proteins play a central role in most cells. Protein-protein interaction database obtained from a yeast two hybrid screen or a composite data set includes random false positives. To purify the database, we employed cellular localization and DNA expression profile and then reconstructed kinome and cell cycle networks. The kinome and cell cycle networks constructed from the purified database show characters of scale free network and hierarchical network, in which preserves its scale-free organization and displays inherent modularity of protein clusters. To find functional modules from the network, we propose a new technique that is based on multi-body correlations in kinome or cell cycle network. From the derived modules, we predicted and estimated tentative functions for unannotated proteins with high certainty.

1. Introduction

Proteins bind to each other to form stable complex and to contribute cellular responses. At a complex system, protein complexes interact with other proteins weakly, transiently, or conditionally to form a biological module that serves a specific function in cells and systems as well [1]. Driven by possession of whole genome scale data sets from biological complex systems, our conception of biomolecular organization is changed to hierarchical levels of organization and complex networks [2].

With the sufficiency of protein-protein interaction data obtained by genome projects, it is possible to create a widespread representation of the protein network of the yeast cells [2]. Protein networks are currently investigated in terms of topology, motifs, correlation structure, and modular properties that are

related to function [3]. A functional module is defined as a group of proteins that its function is separable from other modules. In this work, we have been constructed kinome and cell cycle networks and analyzed structural characterization and functional modules of the networks. Since large scale protein-protein interaction database contains random false positives, we employed cellular localization concept and DNA expression pattern analysis to construct highly purified kinome and cell cycle networks of the yeast cells in here. From functional modules of kinome and cell cycle networks, we derive tentative functions for unannotated proteins and identify several hub proteins and significant linkers regarding the lethality of the null mutants. Therefore, protein network is a useful tool for identifying unknown functions of proteins and prediction of hub or lethal proteins that highly connected with proteins to form modules.

2. Materials and Methods

2.1. DNA and protein database

Protein-protein interaction database obtained from the Munich Information Center for protein Sequences [MIPS, <http://mips.gsf.de/genre/proj/yeast/index.jsp>] database lists of proteins, localization, and functional category in January 2004. DNA expression profiles generated from hundreds of yeast deletion mutants obtained from author's supplemental data on the Cell website [<http://www.cell.com/cgi/content/full/101/7/109/DC1>, Ref. 4].

2.2. Network construction

For each protein network investigated, nodes (proper proteins) and links (protein-protein interactions) were assembled and represented protein network as a graph using InterViewer program. Each link in the network was assigned a length of 1. In a basic principle, if a protein interacts with its partner, the link was designed

as one. If, however, a protein does not interact with any proteins, zero was given in the link. For construction of network, the basic principle, therefore, follows adjacent matrix that is a matrix with rows and columns labeled by graph nodes with a 1 or 0 in position (i, j) according to whether i and j are adjacent or not. Protein complexes and modules are derived from clustering the protein interaction network.

3. Results

We have constructed protein networks of kinome (proteins: 2,130, interactions: 3,392, and clustering coefficient: 0.05362) and cell cycle (proteins: 2,099, interactions: 4,923, and clustering coefficient: 0.11668) using *Saccharomyces cerevisiae* genome data set taken from MIPS database. The networks show character of scale free network, in which a few highly connected proteins, called hub proteins, play a central role in most living organisms. However, protein-protein interaction database contains a high rate of false positives [5]. We, therefore, applied two considerations, cellular localization and DNA expression profile for obtaining highly purified database and then reconstructed kinome network (proteins: 1,161, interactions: 1,698, and clustering coefficient: 0.12204) and cell cycle network (proteins: 1,629, interactions: 4,155, and clustering coefficient: 0.14575) and as well. Newly constructed kinome and cell cycle networks show characters of scale free network and hierarchical network, in which preserves its scale-free organization and displays inherent modularity of protein clusters. For functional prediction, we validate four rules that i) unidentified proteins connected with hub are same as a hub protein in function; ii) the function of linker is considered by the function of connecting hubs; iii) if hub's function is different from components' function, unidentified proteins follow general function of components; and iv) unidentified hubs follow functional modules. We observed that clusters of the network are composed larger groups of modules and these act as functional modules in the protein networks (Fig. 1). The significant linkers are also determined as analyzing DNA expression patterns [4]. From the derived modules of kinome, we predict and estimate functions for 124 unidentified proteins. Although we are not able to estimate certainty of the prediction at this point, we convict that our predictions have high certainty, comparing result of probability expression method [5]. We also identify lethal hub and significant linker proteins from the protein networks of kinome and cell cycle.

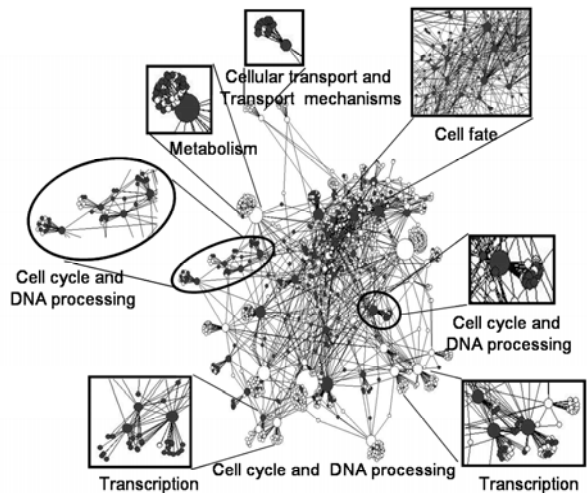


Figure 1. Functional modules of the kinome network.

4. Discussion

Protein networks are organized in a modular fashion, leading to predict unannotated proteins and hub and significant linker proteins. It is important that how reduces the complexity of the protein network to small units of modular structure and function. We applied this concept to yeast kinome and cell cycle and extracted modular network structures from protein interaction networks. Annotated proteins on the modules are major determinants of the functions of modules. Linker between modules is an important role as cross-talking point in the functional and modular communication. Therefore, identification of modules, linkers, and hub proteins from network could be valuable to understand complex systems in system level.

5. References

- [1] L.H. Hartwell, J.J. Hopfield, S. Leibler, and A.W. Murray, "From molecular to modular cell biology", *Nature*, London, December 1999, 402:C47-C52.
- [2] B. Schwikowski, P. Uetz, and S. Fields, "A network of protein-protein interactions in yeast", *Nat. Biotechnol.* London, December 2000, 18:1257-1261.
- [3] A.W. Rives and T. Galitski, "Modular organization of cellular networks", *PNAS*, February 2003, 100:1128-1133.
- [4] T.R. Hughes et al. "Functional discovery via a compendium of expression profiles", *Cell*, Cell Press, USA, July 2000, 102:109-126.
- [5] M.P. Samanta and S. Liang, "Predicting protein functions from redundancies in large-scale protein interaction networks", *PNAS*, USA, October 2003, 100:12579-12583.