Short Communication

Identification of crucial nodes in biological networks

DeWu Ding
Department of Mathematics and Computer Science, Chizhou College, Chizhou 247000, China
E-mail: dwding@yahoo.com.cn

Received 3 June 2012; *Accepted 5 July 2012*; *Published online 1 September 2012*
IAEES

Abstract
This study showed that the crucial nodes in biological networks could be identified with network communities.

Keywords biological network; metabolic network; centralization; crucial nodes; identification.

1 Introduction

As the most powerful tool to understand complex systems, complex networks have attracted so much attentions in a number of topics recent years, e.g., social networks, transportation network, Internet, and particularly biological networks (Newman, 2010; Dormann, 2011; Ibrahim et al., 2011; Martinez-Antonio, 2011 Tacutu et al., 2011; Zhang, 2011). In generally, a large number of nodes and connections in these networks, and there have been developed a number of network molding methods, such as small-world, scale-free, and so on (Barabasi and Oltvai, 2004; Kininmonth et al., 2012; Zhang, 2012a, b).

Identification of crucial nodes is a fundamental problem in the study of these networks, and a number of centralization approaches have been developed to address this challenge (Junker et al., 2006). For example, degree centrality is used to identify the hubs, betweenness centrality depends on the number of shortest pathways going through the nodes, closeness centrality is helpful to identify if the nodes in the core of the network, and so on. It is suggested that single centrality approach is not sufficient, several approaches need be used together in biological networks (Ding et al., 2008). However, we herein show that the crucial nodes in biological networks could be identified with network community centrality, a new proposed method (Kovacs et al., 2010).

2 Method

Most networks are composed by a number of different elements, which would form communities (or modules). There are lots connections within these communities, but sparse connections between them (Newman, 2006; Fortunato, 2010). In a general way, expose the communities is useful in understanding the structure and function of the networks, which has inspired many empirical research and practical application, such as identification of protein complexes (Gavin et al., 2006; Krogan et al., 2006). Here, we show another practical application with network communities (Kovacs et al., 2010), the identification of crucial nodes in biological networks.

To achieve the study, we first construct a metabolite graph for recent reconstructed high-quality S. aureus metabolic network model, we then revised the metabolite graph, and extract the giant strong component for

network communities study (Ding and Li, 2009). To simplify the problem, we merge all self-loops to single undirected connections, there are 250 nodes and 331 connections in the final metabolite graph model. At last, using the ModuLand (Kovacs et al., 2010), we identify 11 communities and corresponding 11 crucial nodes with community centrality (see Fig. 1).

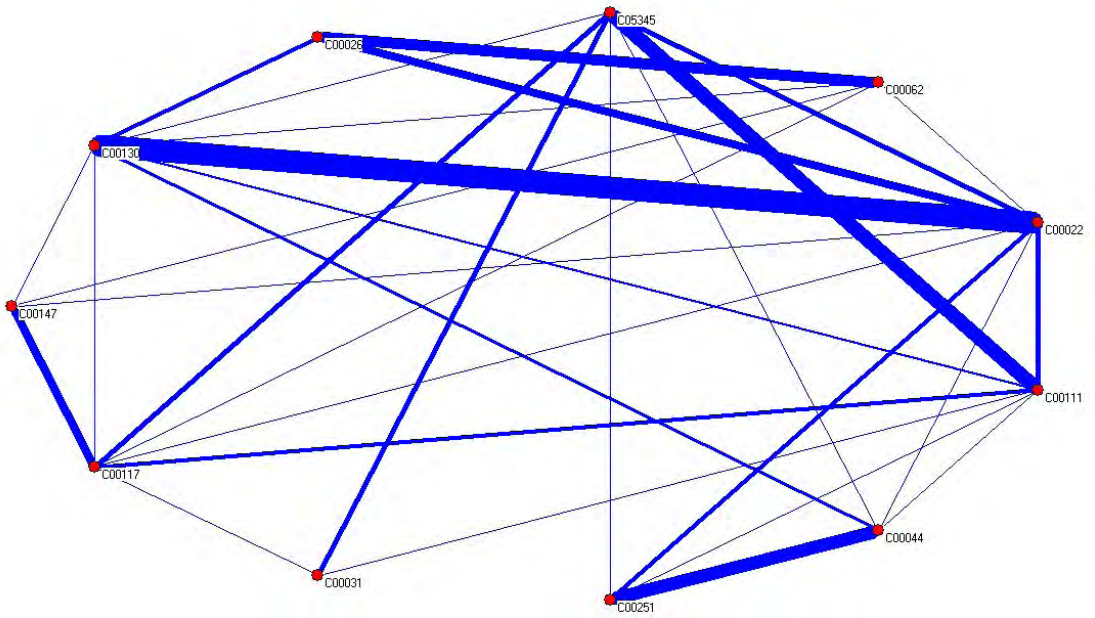

Fig. 1 The 11 crucial nodes in the giant strong component of *S. aureus* metabolic network.

These 11 crucial nodes are all with important biological signification, for example: C00022 (pyruvate, PYR) is the most important intermediate in the glycolysis pathway; C00111 (glycerone phosphate, GlyP) plays a key role in many pathways, such as glycolysis pathway, fructose and mannose metabolism, glycerophospholipid metabolism, carbon fixation, nicotinate and nicotinamide metabolism; C00117 (D-ribose 5-phosphate, R5P) and C05345 (beta-D-fructose 6-phosphate, F6P) are important intermediates in the glycolysis pathway and pentose phosphate pathway; etc. Please see table 1 for their correlation.

Table 1 The correlation of the 11 crucial nodes

	C00022	C00111	C00117	C05345	C00062	C00026	C00130	C00044	C00251	C00147	C00031
C00022	1.000	-0.225	-0.267	-0.217	-0.283	-0.193	0.090	-0.246	-0.160	-0.182	-0.203
C00111	-0.225	1.000	0.443	0.759	-0.189	-0.249	-0.154	-0.103	-0.103	-0.147	0.284
C00117	-0.267	0.443	1.000	0.475	-0.054	-0.228	-0.185	-0.146	-0.146	0.510	0.173
C05345	-0.217	0.759	0.475	1.000	-0.178	-0.234	-0.193	-0.084	-0.084	-0.139	0.371
C00062	-0.283	-0.189	-0.054	-0.178	1.000	0.359	-0.067	-0.114	-0.114	0.074	-0.083
C00026	-0.193	-0.249	-0.228	-0.234	0.359	1.000	-0.041	-0.150	-0.150	-0.139	-0.109
C00130	0.090	-0.154	-0.185	-0.193	-0.067	-0.041	1.000	-0.006	-0.150	-0.103	-0.109
C00044	-0.246	-0.103	-0.146	-0.084	-0.114	-0.150	-0.006	1.000	0.690	-0.089	-0.070
C00251	-0.160	-0.103	-0.146	-0.084	-0.114	-0.150	-0.150	0.690	1.000	-0.089	-0.070
C00147	-0.182	-0.147	0.510	-0.139	0.074	-0.139	-0.103	-0.089	-0.089	1.000	-0.064
C00031	-0.203	0.284	0.173	0.371	-0.083	-0.109	-0.109	-0.070	-0.070	-0.064	1.000

Acknowledgement

The authors would like to thank the anonymous reviewers for their valuable comments.

References

Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. Nature Reviews Genetics, 5: 101-113

Ding DW, Liu T, Lu KZ. 2008. Centralization of complex networks: application to metabolic networks. Computers and Applied Chemistry, 25: 1508-1510

Ding DW, Li LN. 2009. Why giant strong component is so important for metabolic networks? Rivista di Biologia / Biology Forum, 102: 12-16

Dormann CF. 2011. How to be a specialist? Quantifying specialisation in pollination networks. Network Biology, 1(1): 1-20

Fortunato S. 2010. Community detection in graphs. Physics Reports, 486: 75-174

Gavin AC, Aloy P, Grandi P, et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. Nature, 440: 631-636

Ibrahim SS, Eldeeb MAR, Rady MAH, et al. 2011. The role of protein interaction domains in the human cancer network. Network Biology, 1(1): 59-71

Junker BH, Koschutzki D, Schreiber F. 2006. Exploration of biological network centralities with CentiBiN. BMC Bioinformatics, 7: 219

Kininmonth S, van Oppen M, Castine S, et al. 2012. The small genetic world of *Seriatopora hystrix*. Network Biology, 2(1): 1-15

Kovacs IA, Palotai R, Szalay MS, et al. 2010. Community landscapes: a novel, integrative approach for the determination of overlapping network modules. PLoS ONE, 5: e12528

Krogan NJ, Cagney G, Yu HY, et al. 2006. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature, 440: 637-643

Martinez-Antonio A. 2011. *Escherichia coli* transcriptional regulatory network. Network Biology, 1(1): 21-33

Newman MEJ. 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of USA, 103: 8577-8582

Newman MEJ. 2010. Networks: An Introduction. Oxford University Press, UK

Tacutu R, Budovsky A, Yanai H, et al. 2011.Immunoregulatory network and cancer-associated genes: molecular links and relevance to aging. Network Biology, 1(2): 112-120

Zhang WJ. 2011. Constructing ecological interaction networks by correlation analysis: hints from community sampling. Network Biology, 1(2): 81-98

Zhang WJ. 2012a. How to construct the statistic network? An association network of herbaceous plants constructed from field sampling. Network Biology, 2(2): 57-68

Zhang WJ. 2012b. Modeling community succession and assembly: A novel method for network evolution. Network Biology, 2(2): 69-78