Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 237 (2005) 401–411
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Lethality and synthetic lethality in the genome-wide metabolic network of *Escherichia coli*

Cheol-Min Ghim[a,1], Kwang-Il Goh[a], Byungnam Kahng[a,b,*]

[a]*School of Physics, Seoul National University NS50, Seoul 151-747, Republic of Korea*
[b]*Program in Bioinformatics, Seoul National University NS50, Seoul 151-747, Republic of Korea*

Received 5 December 2004; received in revised form 25 April 2005; accepted 26 April 2005
Available online 21 June 2005

Abstract

Recent genomic analyses on the cellular metabolic network show that reaction flux across enzymes are diverse and exhibit power-law behavior in its distribution. While intuition might suggest that the reactions with larger fluxes are more likely to be lethal under the blockade of its catalysing gene products or gene knockouts, we find, by in silico flux analysis, that the lethality rarely has correlations with the flux level owing to the widespread backup pathways innate in the genome-wide metabolism of *Escherichia coli*. Lethal reactions, of which the deletion generates cascading failure of following reactions up to the biomass reaction, are identified in terms of the Boolean network scheme as well as the flux balance analysis. The avalanche size of a reaction, defined as the number of subsequently blocked reactions after its removal, turns out to be a useful measure of lethality. As a means to elucidate phenotypic robustness to a single deletion, we investigate synthetic lethality in reaction level, where simultaneous deletion of a pair of nonlethal reactions leads to the failure of the biomass reaction. Synthetic lethals identified via flux balance and Boolean scheme are consistently shown to act in parallel pathways, working in such a way that the backup machinery is compromised.

Keywords: Metabolic network; Flux balance analysis; Synthetic lethality

1. Introduction

Complex machinery of cellular metabolism occurring in a living organism makes up a part of autocatalytic network of biochemical reaction pathways. The reaction network in itself constitutes an intricate web in such a way as sharing intermediates. Yet another dimension of complexity comes from the tight control of reactions by functional proteins which are again under transcriptional, translational control as well as degradative and other inductive regulations, making even a single pathway analysis formidable task. Only recently, advances in high-throughput experiments and the computing power incorporating diverse data sets collected in genomic research make it possible to construct cellular networks of metabolism in genome-wide perspectives. At the same time, many quantitative theoretical methods including graph theories and other mathematical tools developed from diverse disciplines attract much attention to tackle the large-scale networks (Xia et al., 2004; Barabási and Oltvai, 2004).

In the early graph-theoretic approaches to the metabolic network, attention has been paid to the so-called scale-free feature of topological structure (Jeong et al., 2000), small-world-ness (Wagner and Fell, 2001), modularity (Girvan and Newman, 2002) and hierarchical organization (Ravasz et al., 2002). Despite the immanent specificity in cellular functions of various

Abbreviation: FBA: flux balance analysis; CLS: cumulative lethality score

*Corresponding author. School of Physics, Seoul National University, Shillim-dong, Gwanak-gu, Seoul 151-747, Republic of Korea. Tel.: +82 2 880 1326; fax: +82 2 884 3002.
E-mail address: kahng@phya.snu.ac.kr (B. Kahng).
[1]Present address: Center for Theoretical Biological Physics, University of California, San Diego, La Jolla, CA 92093, USA.

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.04.025

organisms, the connectivity, or number of connections each node (metabolite or associated reactions) has, is generally far from homogeneous. In particular, this connectivity distribution of the metabolic network, as shared by many naturally occurring complex networks, follows a power law, meaning large deviations in spite of well-defined average value. It is this context that borrows the term *scale-free* network, where hubs, nodes with large number of connections, play essential roles. When such hubs are removed or turned off, the whole system becomes vulnerable. Indeed, it was found (Jeong et al., 2001) that, for the yeast protein interaction network, hub proteins are more likely to be lethal than the others.

In the framework of networks, metabolic reactions and participating metabolites can be mapped into alternating nodes, where the outward(inward) connections from a reaction node indicate that those metabolites are produced(consumed) as a result of the reaction. Once constructing a directed bipartite graph in this way, we calculate graph-theoretic quantities that characterize the global topology and give a clue to assessing lethality of metabolic reactions. Then, we examine the correlations between the metabolic flux level and the lethality of each metabolic reaction using the flux balance analysis[FBA] (Edwards and Pålsson). Here, by lethal, we mean the organism could rarely synthesize the indispensable biomass, or the flux of the biomass reaction is significantly reduced when that reaction is blocked or removed from the network, mimicking gene knockout experiments[2]. One seemingly counterintuitive result is the absence of correlation between flux level and lethality (Fig. 2), which is related with the fact that the high-flux reactions have abundant bypasses or backup pathways.

We also invoke the Boolean network scheme, an idealization of the metabolic network as a wiring of binary logic gates to elucidate the pathway structure of the network on the logical basis. Considering the knockout and consequent cascading failure in the metabolic reaction network as an avalanche, we investigate the distribution of avalanche size defined as the number of reactions subsequently turned off on account of the removal of a target reaction to find it a good measure of lethality. The distribution follows a power law with the characteristic exponent around 2.5, pervasive throughout disparate model systems having self-organized criticality (Bak, 1996).

In the latter part, we investigate the effects of simultaneous deletion of multiple reactions with a view to elucidating the interaction between them. Though only a small portion of reactions lead to distinctive phenotype under their single deletion, it is non-trivial to make a prediction on how destructive a dual deletion of non-lethal reactions. Actually, the effect can be aggravating or alleviating as well as simple sum of each, depending on their role played in the otherwise intact metabolic network. In particular, when a pair of non-lethal reactions are deleted to make no growth of cell, we call it synthetic-lethal, metabolic homologue of the same term in genetics. Synthetic lethality in metabolic network is a manifestation of their complementary nature responsible for the buffering between alternative parallel pathways. We show the synthetic-lethal pairs are distributed over distinct the avalanche size of a pair of reactions is strongly correlated with its synthetic lethality.

2. Materials and methods

We use, with minor curation, the recent revision of in silico model *E. coli* (Reed et al., 2003), which was obtained by searching databases, such as LIGAND (http://www.genome.jp/kegg/ligand.html), EcoCyc (http://www.ecocyc.org), TC-DB (http://tcdb.ucsd.edu/), and referring to updated literatures on sequence annotation (Serres et al., 2001). To mimic random or targeted mutation strains, a specific reaction is removed from the network and the resultant metabolic capabilities are to be assessed. For this purpose, we introduce a single pivotal reaction, the biomass reaction, originally formulated as a linear combination of essential metabolic reactions giving rise to the growth of the organism (Neidhardt and Umbarger, 1996). Throughout the study, lethality of a certain reaction is determined by the flux of this biomass production, which is contingent to the ansatz of optimality that the selection pressure has imposed in the long history of evolution.

2.1. Metabolic network as a graph

The overall map of metabolic reactions we study is a bipartite graph, composed of two different types of nodes, 627 metabolites and their participating 1074 metabolic reactions including transport and exchange events. One type of nodes connect only to the other type of nodes in the networks[3]. Each link between a pair of a metabolite and a reaction is directed, reflecting the metabolite is either consumed (substrate) or produced (product) or both in reactions. Of the 1074 reactions,

[2]More recently, Segrè et al. proposed alternative scheme phrased as "minimization of metabolic adjustment [MOMA]" for the phenotypic prediction of deletion mutants. Instead of assuming optimality in growth yield of deletion mutants, MOMA approximates metabolic phenotype by performing distance minimization in flux space, whereby the correlation with experimental results are improved (See Segrè et al., 2002).

[3]Depending upon the objectives, it can be projected to recover either of the single-mode metabolite network or reaction (enzyme) network.