Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 248 (2007) 618–626
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Comparative study of the transcriptional regulatory networks of *E. coli* and yeast: Structural characteristics leading to marginal dynamic stability

Deok-Sun Lee*[,1], Heiko Rieger

Theoretische Physik, Universität des Saarlandes, 66041 Saarbrücken, Germany

Received 19 October 2006; received in revised form 18 May 2007; accepted 3 July 2007
Available online 14 July 2007

Abstract

Dynamical properties of the transcriptional regulatory network of *Escherichia coli* and *Saccharomyces cerevisiae* are studied within the framework of random Boolean functions. The dynamical response of these networks to a single point mutation is characterized by the number of mutated elements as a function of time and the distribution of the relaxation time to a new stationary state, which turn out to be different in both networks. Comparison with the behavior of randomized networks reveals relevant structural characteristics other than the mean connectivity, namely the organization of circuits and the functional form of the in-degree distribution. The abundance of single-element circuits in *E. coli* and the broad in-degree distribution of *S. cerevisiae* shift their dynamics towards marginal stability overcoming the restrictions imposed by their mean connectivities, which is argued to be related to the simultaneous presence of robustness and adaptivity in living organisms.

Keywords: Gene regulatory network; Boolean model; Dynamic stability

1. Introduction

Living organisms depend simultaneously on a stable internal environment and a capability to adapt to a fluctuating external environment (Causton et al., 2001). Since the biological characteristics of an organism are determined by the interplay between its gene repertoire and the regulatory apparatus (Babu et al., 2004), robustness and adaptiveness should be generic features of the molecular interactions composing the gene regulation machinery. The organization of the gene transcriptional regulatory network has been analyzed for numerous organisms, in particular for the prokaryote *Escherichia coli* (*E. coli*) (Thieffry et al., 1998; Dobrin et al., 2004; Shen-Orr et al., 2002) and the eukaryote *Saccharomyces cerevisiae* (*S. cerevisiae*) (Guelzim et al., 2002; Lee et al., 2002; Luscombe et al., 2004).

Adaptivity of an organism implies the production of different cell types with different functions from the same genome. This begins with a regulated transcription by certain proteins, transcriptional factor (TF) (Orphanides and Reinberg, 2002). The identification of the target genes for each TF allows the construction of a gene transcriptional regulatory network, where the nodes are the genes or operons that produce TF's or are regulated by TF's, and the directed edges indicate a regulatory dependence: A directed edge from node A to node B implies that a TF encoded by gene A is involved in the regulation of the expression of gene B. The expression level of each gene defines the dynamical state of the network. To achieve robustness and adaptiveness at the same time one expects the regulatory network dynamics to be neither chaotic nor fully insensitive to perturbations, but marginally stable. Structural characteristics of the network must support these dynamical features.

*Corresponding author.
E-mail address: deoksun.lee@gmail.com (D.-S. Lee).
[1]Present address: Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA.

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2007.07.001

Our study reveals specific topological features in the transcriptional regulatory network architecture of *E. coli* and *S. cerevisiae* that shift the dynamics towards marginal stability. *E. coli*'s network has a very low mean connectivity, the number of edges per node, which would lead to a high stability thus deteriorating adaptiveness in random networks, where all regulating rules are equally probable. But we find that single-element circuits which are anomalously rich in *E. coli*'s network help mutations triggered by random perturbations to persist, favoring an unstable dynamical behavior. *S. cerevisiae* on the other hand has a sufficiently high mean connectivity which favors chaotic dynamics in the random networks deteriorating stability. Here we find that *S. cerevisiae*'s network has a broad (algebraic) node degree distribution and we demonstrate the stabilizing effect of this feature upon the dynamics.

Practically, the information about the transcriptional regulatory network structure—which TF binds to which gene—is available, for example, via the chromatin-immunoprecipitation microarray experiments (Lee et al., 2002). The question, whether a specific TF enforces or inhibits the expression of a specific target gene has to be studied separately. However, those individual interactions do not necessarily occur independently and these regulatory interactions are often combinatorial (Buchler et al., 2003) and time-, cell cycle-, or environment-dependent, limiting the available information on the complete regulation profile. Generic dynamical features then have to be extracted using model interactions as suggested by Kauffman (1969, 1993): One digitizes the continuous expression level to a Boolean variable, 0 (inactive) and 1 (active), and assumes a random static regulation rule for each gene in the form of a random Boolean function for each gene determining its state at the next time step by the current states of its regulators. Here *random* means that the output value of these Boolean functions is 0 or 1 with equal probabilities.

Based on considerations of random Boolean networks with a fixed number of regulators k for every element, Kauffman (1969, 1993) hypothesized that distinct stationary states—limit cycles—correspond to different types of cells. This idea got some support from the agreement of the scaling behavior of the number of limit-cycles for $k = 2$-random Boolean networks and the number of cell types with respect to the genome size, but was also debated (Samuelsson and Troein, 2003; Klemm and Bornholdt, 2005). Among networks with fixed in-degree, $k = 2$ is a critical point distinguishing two different dynamical phases: stable and unstable against perturbations, suggesting that the regulatory network dynamics of living organisms is "on the edge" between order and chaos (Kauffman, 1969, 1993).

However, real regulatory networks do not have a fixed in-degree but a heterogeneous connectivity, even their average in-degree $\langle k \rangle$ is usually different from 2. Nevertheless the Boolean model itself is useful, and recently the effects of the nature of the regulating rules on the dynamical stability were studied within its framework (Harris et al., 2002; Kauffman et al., 2003, 2004), which will be discussed later in connection with our results. We propose that the network structure itself is also relevant for the stability/instability aspect mentioned before. Therefore we construct a network from the data for the transcriptional regulatory interactions for *E. coli* and *S. cerevisiae*, and study how a point mutation, i.e., an altered dynamical state of a single element, spreads over the whole network by inducing another mutation through regulatory interactions in the random Boolean functional form.

2. Method

2.1. Datasets

The transcriptional regulatory network in *E. coli* has long been studied and the obtained results are well integrated e.g., in RegulonDB database (http://regulondb.ccg.unam.mx). We used the dataset of Ref. Shen-Orr et al. (2002), which are based on the Regulon DB and enhanced by literature search. The resultant network consists of 418 operons and 573 interactions with 111 nodes having at least one outward edge. On the contrary, the transcriptional regulation of *S. cerevisiae* on the genome scale became available only very recently via the combination of chromatin-immunoprecipitation and DNA microarray analysis (Lee et al., 2002). We used the data of Ref. [Lee et al. (2002)] and chose the P value threshold 0.01 to yield a network of 4555 nodes and 12 455 directed edges with 112 nodes having at least one outward edge. Isolated nodes and those possessing only self-regulation have been excluded in both networks since they have no interaction with other elements.

2.2. Random Boolean functions

These experimental data establish a directed network G of N nodes, and we assign a dynamic Boolean variable σ_i (that can take on the values 0 or 1 only, corresponding to an inactive or active state, respectively) to each node i. These dynamical variables evolve synchronously via $\sigma_i(t+1) = f_i(\sigma_{i_1}(t), \sigma_{i_2}(t), \ldots, \sigma_{i_{k_i}}(t))$, with the nodes $i_1, i_2, \ldots, i_{k_i}$ having the outward edges incident on the node i and k_i the in-degree of the node i. The output value of f_i for each input configuration $\{\sigma_{i_1}(t), \sigma_{i_2}(t), \ldots, \sigma_{i_{k_i}}(t)\}$ is 1 with probability p or 0 with probability $1-p$. Once determined at the beginning, the regulating rule f_i does not change with time. If $k_i = 0$, σ_i is fixed at f_i; $\sigma_i(t+1) = f_i$ regardless of the value of $\sigma_i(t)$. Here we introduced a parameter p controlling the distribution of the regulating rules. If $p = 0$ (1), the output value should be 0 (1) for any input configuration, yielding $\sigma_i = 0(1)$ for all i. On the other hand, if $p = \frac{1}{2}$, an input configuration can lead to the output value 0 and 1 with the same probability $\frac{1}{2}$, and as a result, all $2^{2^{k_i}}$ regulating rules for a node with in-degree k_i