Sensing Your Surroundings: How Transcription-Regulatory Networks of the Cell Discern Environmental Signals

Gábor Balázsi and Zoltán N. Oltvai*

(Published 3 May 2005)

You pick up the phone and dial. With this simple act, you enter into a vast network of electronic gadgets and wires that seamlessly establishes the connection between you and the person you called. But what if you also want the other person to see you? Modern cellular phones solve this problem with the use of technology specialized to capture images and transmit them to the receiver without interfering with the sound signal. If in the future we want to communicate smell or touch sensation as well, current cell phones will need to be developed even further, making them capable of capturing and transmitting multiple types of input signals at once. However, cells already solved this problem billions of years ago. Their biochemical regulatory networks, with which they sense and respond to environmental changes and internal cues, have accomplished a similar but much more complicated task, and have done so in a highly noisy environment.

A characteristic feature of intracellular information transfer is that the components of the various regulatory networks, functioning in a highly viscous cellular interior, operate over several orders of magnitude in time. The fastest of these constitute the various signal transduction networks, ranging from two-component systems of prokaryotes (*1*) to the highly complex signal transduction networks of mammalian cells (*2*–*4*). Fast signaling, however, is frequently followed by slower transcription-regulatory events, during which regulatory gene products such as transcription factors (TFs) (*5*) and regulatory RNAs (*6*, *7*) alter the rate of transcription of other genes, thus reorganizing gene expression to achieve new metabolic states or initiate cellular programs such as the cell cycle or sporulation.

Understanding the system-level properties of these networks requires both computational and experimental efforts that start with mapping potential regulatory interactions that can exist in a given cell type. In the yeast *Saccharomyces cerevisiae* and in the bacterium *Escherichia coli*, the "wiring diagrams" for potential TF-mediated interactions have been mapped out to such a degree (*8*–*10*) that their system-level characteristics and function can be investigated. Initial studies of the global and local properties of these transcription-regulatory (TR) networks revealed that their out-degree (the number of target genes regulated by each TF) follows a power-law distribution (*11*–*13*). This is a characteristic property of scale-free networks, in which a small number of hub nodes are connected to a very large number of other nodes (*14*). On a small scale, certain elementary interaction patterns (*15*), called "motifs," are significantly more abundant than expected by chance (*12*, *13*). Such findings imply that their evolutionary selection is driven by their information-processing capabilities (*12*, *13*, *16*–*18*).

Department of Pathology, University of Pittsburgh School of Medicine, 3550 Terrace Street, Pittsburgh, PA 15213, USA.

***Corresponding author. E-mail, oltvai@pitt.edu**

Although informative, the topological properties of potential regulatory interactions explain little of how a TR network functions. To learn more, we must decipher the "functional logic" by which various parts of these static regulatory networks operate dynamically in a given condition. Two recent studies (*19*, *20*) addressed this question by examining the TR network of *S. cerevisiae* from two complementary perspectives.

After compiling a static representation of known TF-mediated regulatory interactions, Luscombe *et al*. (*19*) used published microarray data to assemble condition-dependent TR networks under five specific environmental and developmental conditions. The networks were defined by linking TFs present in a given condition to their differentially expressed target genes. TFs were classified as "present" or "absent" on the basis of their abundance during the condition relative to their starting abundance. Although a TF might physically bind to its target site, the corresponding link was not considered active if the expression of the target gene did not change significantly, or if the TF abundance stayed low under the specific condition. Analyses of the resulting subnetworks revealed that the majority of regulatory interactions are condition-specific, and only a small subset (called "hot links") are active in four or more conditions. Also, on the basis of their topology, each of the five condition-dependent TR subnetworks could be classified as either exogenous (induced by diauxic shift, DNA damage, or stress response) or endogenous (induced by sporulation or cell cycle), the latter having a more complex architecture.

Transcriptional hubs have the potential to regulate a large number of target genes, and therefore they might be expected to maintain high connectivity in different conditions. Contrary to this expectation, the authors found that most hubs are "transient" (their expression is not maintained between conditions). Only a small percentage of hubs (called "permanent") maintain a high out-degree in all conditions. However, even permanent hubs switch their targets between conditions. As a result of link rewiring, the same TFs can be used in different conditions to regulate the expression of various sets of genes and to elicit a condition-dependent response, which implies that the response of the cell is commonly a result of combinatorial TF usage.

Harbison *et al*. (*20*) addressed the question of differential utilization from a more biological standpoint. Extending their previous work (*9*), they determined the location of TF binding sites in gene promoters on a genome-wide scale for 203 TFs in rich medium, and for 84 TFs in at least one of 12 other environmental conditions. By compiling information from TF binding data, phylogenetically conserved sequences, and prior knowledge, they were able to map with high confidence the yeast TR network for 102 TFs. This method can be used to reliably detect "all-or-none" changes in TF-promoter binding, but not cases in which the amount of promoter binding increases or decreases without dropping to zero. On the basis of the presence of TF

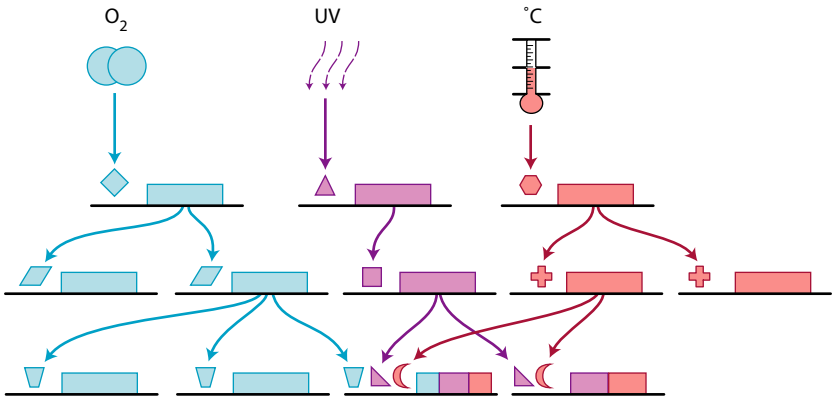

Fig. 1. Condition-dependent utilization of TR networks. Environmental signals are sensed by signal sensor proteins that regulate distinct cellular response in TR networks. Simple environmental signals affect only nodes within subnetworks (schematically colored blue, purple, and red) originating at regulators sensitive to the given environmental change. In contrast, complex environmental signals are decomposed by signal-specific sensors and reassembled deeper inside the network.

binding sites, the authors classified the identified promoter regions into four different architectures, called "single regulator," "repetitive motif," "multiple regulator," and "co-occurring regulator" types. The first two of these architectures are characterized by the presence of one or more binding sites for a single TF, whereas the other two contain binding sites for two or more TFs. The authors also classified promoter utilization patterns into four types: "condition invariant," in which the set of TF binding sites does not change; "condition enabled," in which TF binds in one growth condition but not in the other; "condition expanded," in which the set of binding sites in one condition includes those used in the other; and "condition altered," in which different sets of promoters are bound in the two conditions. What remains to be explored is the environment-dependent change of TF binding patterns. For example, are all binding sites within a "repetitive motif" used in one condition, versus only one (equivalent to a "single motif") in another? Does regulation by two or more TFs (equivalent to a "multiple regulator") shift to regulation by a single TF (equivalent to the "single regulator" or "repetitive motif") in certain conditions?

Taken together, the two studies show that *S. cerevisiae* uses largely disjointed parts of its TR network in different environmental conditions by activating various promoter regulatory modes. But what is the basis of such condition-dependent utilization? We suggest that part of the answer may relate to the mode in which cells perceive various signals. Just as we humans have specialized sensors for perceiving light, sound, and heat (the eye, the ear, and the skin, respectively), which use a signal transduction process (the afferent neural pathways) to affect a network (the central nervous system), individual cells are also able to perceive input from the environment and initiate a signal transduction process to affect a TR network. Cells use highly specialized "sensor" proteins to detect concentration changes of nutrients, oxygen, variations of temperature, or the damaging effect of ultraviolet radiation (Fig. 1). Just as the microphone in your phone receiver relays your voice into a telecommunication network, sensor proteins are specialized gateways through which environmental signals enter the TR network. Information could be quickly lost in the noisy cellular environment, were it not for highly specialized sets of molecules relaying different types of signals.

TR networks are inherently directional and sparse (*12*, *21*) and tend to be acyclic (*21*, *22*). Thus, the genes affected by a specific environmental signal and the TR pathways along which the perturbation propagates away from sensor proteins can be predicted on the basis of the network's hierarchical and acyclic topology (Fig. 1). The sparseness of connections has its own benefit: If sensor TFs cross-regulate all potential target genes, the whole genome would be affected very quickly after any environmental change. Heat shock response would be automatically turned on as a result of minor changes in nutrients, and the DNA repair proteins would be expressed in anaerobic growth conditions. This would clearly be a wasteful and evolutionarily disadvantageous solution for the cell.

At the same time, cells must have the capacity to respond to most dynamical environmental changes, whether simple (involving the change of a single factor, such as oxygen or a nutrient) or complex (involving the simultaneous change of many factors). The network topology (*12*, *21*) indicates that, in order to properly process and respond to complex environmental changes, organisms are likely to use distinct TR subnetworks regulated by sets of sensors specialized to detect specific aspects of complex environmental stimuli. The question remains: Is it possible to predict a priori which part of the network will respond to a certain stimulus? The answer is probably yes, if we know the sensors for the stimulus and the full network topology. However, there is still much to be examined—for example, what is the "general stress response"? Is it due to the same sensors responding to multiple types of stress or to the convergence of various types of incoming perturbations on the same set of nodes?

Signal-specific sensors individually perceive the components of complex stimuli, whereas TR subnetworks reassemble the resulting pieces of processed information deeper inside the network (Fig. 1). Thus, the existence of signal-processing units may make information processing simple and economical for the cell. If a simple environmental signal is sensed only by one TF, only the subnetwork originating at the sensor may be dynamically affected while the rest of the network remains relatively dormant. In contrast, complex signals may undergo parallel processing in quasi-independent subnetworks before the development of an integrated response.

Combined theoretical and experimental studies of motif dynamics have already generated initial insights into the information-processing capabilities of TR networks (*23*, *24*). However, the position of motifs within condition-specific subnetworks and their frequent aggregation into larger topological structures (*25*) may substantially modify their behavior. Therefore, to understand cellular response to a dynamic environment, it will be necessary to complement small- and large-scale studies with a "medium-scale" understanding of the dynamics of subnetworks affected in specific conditions. Differential utilization is apparent in protein-protein interaction networks (*26*) as well as in metabolic networks (*27*). Thus, the analysis and modeling of integrated, condition-specific "cellular subnetworks" (*28*, *29*) will improve our understanding of cellular responses to a dynamic environment.

References

1. J. J. Bijlsma, E. A. Groisman, Making informed decisions: Regulatory interactions between two-component systems. *Trends Microbiol.* **11**, 359 (2003).
2. G. Weng, U. S. Bhalla, R. Iyengar, Complexity in biological signaling systems. *Science* **284**, 92 (1999).
3. J. A. Papin, T. Hunter, B. O. Palsson, S. Subramaniam, Reconstruction of cellular signalling networks and analysis of their properties. *Nat. Rev. Mol. Cell Biol.* **6**, 99 (2005).
4. M. A. White, R. G. Anderson, Signaling networks in living cells. *Annu. Rev. Pharmacol. Toxicol.* **45**, 587 (2005).
5. M. M. Babu, N. M. Luscombe, L. Aravind, M. Gerstein, S. A. Teichmann, Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* **14**, 283 (2004).
6. J. S. Mattick, M. J. Gagen, The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* **18**, 1611 (2001).
7. S. Gottesman, The small RNA regulators of *Escherichia coli*: Roles and mechanisms. *Annu. Rev. Microbiol.* **58**, 303 (2004).
8. V. V. Svetlov, T. G. Cooper, Review: Compilation and characteristics of dedicated transcription factors in *Saccharomyces cerevisiae. Yeast* **11**, 1439 (1995).
9. T. I. Lee, N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, R. A. Young, Transcriptional regulatory networks in *Saccharomyces cerevisiae. Science* **298**, 799 (2002).
10. H. Salgado, S. Gama-Castro, A. Martinez-Antonio, E. Diaz-Peredo, F. Sanchez-Solano, M. Peralta-Gil, D. Garcia-Alonso, V. Jimenez-Jacinto, A. Santos-Zavaleta, C. Bonavides-Martinez, J. Collado-Vides, RegulonDB (version 4.0): Transcriptional regulation, operon organization and growth conditions in *Escherichia coli* K-12. *Nucleic Acids Res.* **32**, D303 (2004).
11. N. Guelzim, S. Bottani, P. Bourgine, F. Kepes, Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* **31**, 60 (2002).
12. S. Shen-Orr, R. Milo, S. Mangan, U. Alon, Network motifs in the transcriptional regulation network of *Escherichia coli. Nat. Genet.* **31**, 64 (2002).
13. R. Milo, S. S. Shen-Orr, S. Itzkovitz, N. Kashtan, U. Alon, Network motifs: Simple building blocks of complex networks. *Science* **298**, 824 (2002).
14. A.-L. Barabási, Z. N. Oltvai, Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101 (2004).
15. A. Vázquez, R. Dobrin, D. Sergi, J.-P. Eckmann, Z. N. Oltvai, A.-L. Barabási, The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17940 (2004).
16. G. C. Conant, A. Wagner, Convergent evolution of gene circuits. *Nat. Genet.* **34**, 264 (2003).
17. V. F. Hinman, A. T. Nguyen, R. A. Cameron, E. H. Davidson, Developmental gene regulatory network architecture across 500 million years of echinoderm evolution. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13356 (2003).
18. S. Mangan, U. Alon, Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11980 (2003).
19. N. M. Luscombe, M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann, M. Gerstein, Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308 (2004).
20. C. T. Harbison, D. B. Gordon, T. I. Lee, N. J. Rinaldi, K. D. Macisaac, T. W. Danford, N. M. Hannett, J. B. Tagne, D. B. Reynolds, J. Yoo, E. G. Jennings, J. Zeitlinger, D. K. Pokholok, M. Kellis, P. A. Rolfe, K. T. Takusagawa, E. S. Lander, D. K. Gifford, E. Fraenkel, R. A. Young, Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99 (2004).
21. H. W. Ma, J. Buer, A. P. Zeng, Hierarchical structure and modules in the *Escherichia coli* transcriptional regulatory network revealed by a new top-down approach. *BMC Bioinformatics* **5**, 199 (2004).
22. D. Thieffry, A. M. Huerta, E. Perez-Rueda, J. Collado-Vides, From specific gene regulation to genomic networks: A global analysis of transcriptional regulation in *Escherichia coli. Bioessays* **20**, 433 (1998).
23. S. Mangan, A. Zaslaver, U. Alon, The coherent feedforward loop serves as a sign-sensitive delay element in transcription network. *J. Mol. Biol.* **334**, 197 (2003).
24. S. Kalir, U. Alon, Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell* **117**, 713 (2004).
25. R. Dobrin, Q. K. Beg, A.-L. Barabási, Z. N. Oltvai, Aggregation of topological motifs in the *E. coli* transcriptional regulatory network. *BMC Bioinformat.* **5**, 10 (2004).
26. J. D. Han, N. Bertin, T. Hao, D. S. Goldberg, G. F. Berriz, L. V. Zhang, D. Dupuy, A. J. Walhout, M. E. Cusick, F. P. Roth, M. Vidal, Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88 (2004).
27. E. Almaas, B. Kovács, T. Vicsek, Z. N. Oltvai, A.-L. Barabási, Global organization of metabolic fluxes in *E. coli. Nature* **427**, 839 (2004).
28. T. Ideker, V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, L. Hood, Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929 (2001).
29. M. W. Covert, E. M. Knight, J. L. Reed, M. J. Herrgard, B. O. Palsson, Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429**, 92 (2004).

Citation: G. Balázsi, Z. N. Oltvai, Sensing your surroundings: How transcription-regulatory networks of the cell discern environmental signals. *Sci. STKE* **2005**, pe20 (2005).