Molecular Cell
Previews

Gene Regulation: Hacking the Network on a Sugar High

Tom Ellis,[1,2] Xiao Wang,[1,2] and James J. Collins[1,2,*]
[1]Center for BioDynamics
[2]Department of Biomedical Engineering
Boston University, 44 Cummington Street, Boston, MA 02115, USA
*Correspondence: jcollins@bu.edu
DOI 10.1016/j.molcel.2008.03.005

In a recent issue of *Molecular Cell*, Kaplan et al. (2008) determine the input functions for 19 *E. coli* sugar-utilization genes by using a two-dimensional high-throughput approach. The resulting input-function map reveals that gene network regulation follows non-Boolean, and often nonmonotonic, logic.

If you had just spent $500 and a day standing in line to get your hands on the latest gadget, would you immediately take a hammer to it? It's probably the last thing a normal person would do, yet last June this is precisely what a handful of curious individuals did to Apple's iPhone (BBC News Online, 2007). Their motivation? To uncover the technology hidden within and, in doing so, to help a community of hackers and bootleggers get started on modifications and imitations. The approach is called reverse engineering—where detailed analyses of the workings of an existing system are undertaken to help build independent copies, new devices, or simply just to understand the system's logic without knowing its blueprint beforehand.

Reverse engineering is a method that is now being employed by systems biologists to understand the workings of biological systems. Much like the curious individuals dissecting their latest purchases, systems biologists use approaches that effectively open up a system they are keen to explore. Perturbations are applied to the system, and its functional architecture can be inferred from how its individual parts behave in response to different conditions. In the case of gene networks, this top-down approach is used to parse out the logical interactions between known parts, and much like in electrical engineering, a circuit diagram of Boolean gates like AND, OR, and NOR can typically be put together to describe the system.

The top-down method is relatively new to biology and has had some significant successes in the past few years (Gardner et al., 2003; Ma'ayan et al., 2005; Yeger-Lotem et al., 2004). However, this approach often suffers from the low resolution of indirect experimental measurements due to technological restrictions. Now, Kaplan et al. (2008) avoid this limitation by exploring a relatively complex regulatory system in two dimensions with high-resolution, high-throughput measurements (Figure 1). Their equivalent of the iPhone is *E. coli*, and the parts of interest within their model organism are 19 sugar-utilization genes. These genes are known to be regulated by a network motif described as dense overlapping regulons (DOR), where a set of regulators combinatorially control a set of output genes (Shen-Orr et al., 2002).

The output of each gene within the network is measured with a library of fluorescent reporter strains, where in each strain a copy of the promoter of one gene of interest is used to drive expression of green fluorescent protein (GFP). The GFP levels provide an accurate readout of transcriptional activity for that gene in vivo, and the logic within the network can be revealed by measuring each gene's activity under different conditions. The process is closely analogous to an electrical engineer testing each part of a circuit with a voltmeter to see which components respond to different inputs and by how much. From this, the authors are able to assign a two-dimensional input function to each gene that describes how its output is controlled by the multiple input signals. Once determined, the input functions of the system can be compared and used to map the logic underlying the network.

It is generally believed that bacterial sugar-utilization genes are regulated in a qualitatively similar manner. But the input functions inferred by Kaplan et al. (2008) show great diversity and cannot be explained by simple combinations of Boolean gates, challenging the usual analogy made between electronic circuits and gene regulatory networks. The most surprising results come from the input functions for genes involved in galactose metabolism. For the genes *galP* and *galETKM*, the input functions are nonmonotonic with respect to the input signals; as an input signal increases, transcription rates rise to a peak and then decrease. The authors hypothesize that this might be due to the fact that, unlike other sugars in this study, galactose has a dual use—it is used as both a carbon source and a cell-wall component. This functional duality and its relationship to the unexpected input functions may warrant further studies to provide a molecular-level understanding of nonmonotonic behavior.

The authors also show that many of the two-dimensional input functions can be decomposed into the products of two one-dimensional functions. This finding could greatly reduce the experimental workload needed for high-resolution data, and the decomposition of input functions could give us some hints about the existence of general regulatory principles. However, there are exceptions such as the regulator of the fucose operon, *fucR*, the input function of which displayed two peaks and could not be decomposed into the products of two one-dimensional functions. Molecular reasoning will be needed to understand such anomalies;

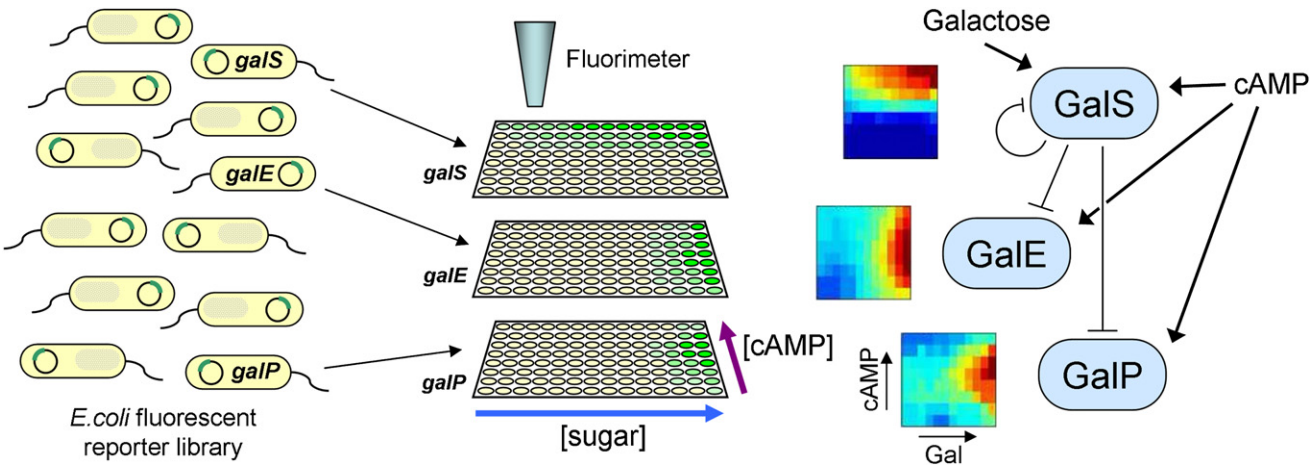

Figure 1. A High-Throughput Method to Determine the Input Functions within a Gene Regulatory Network
An existing *E. coli* fluorescent reporter library is used to provide individual clones, each with a promoter of interest driving green fluorescent protein (GFP) expression. The fluorescence emitted from the clone when grown in a range of cAMP and sugar (e.g., galactose) concentrations represents the transcriptional output of the promoter in response to the inputs. The input functions can be related to the known circuitry of the gene network and used to build an input-function map that describes how the cAMP concentration affects each gene directly and how the sugar concentration affects each gene through the network.

for example, unlike other genes in this study, *fucR* is controlled by an internal promoter in addition to an upstream promoter.

The high-resolution input functions described by Kaplan et al. (2008) provide a great tool for reverse-engineering gene regulation. A natural extension would be the addition of more dimensions. Other possible external inducers that could be studied include growth inhibition, heat shock, and DNA damage.

The authors further raise a question of whether the input function diversity originates from the various promoter structures or from differences in upper-stream circuitry. An intriguing way to address this question would be to engineer equivalent gene regulatory networks through synthetic biology (Guido et al., 2006) or to modify gene regulation within the existing network by replacing the natural promoters with synthetic promoters generated from different libraries (Alper et al., 2005; Cox et al., 2007; Murphy et al., 2007; Solem and Jensen, 2002). Bottom-up approaches like these would no doubt complement the successes of the top-down approach detailed here.

REFERENCES

Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005). Proc. Natl. Acad. Sci. USA *102*, 12678–12683.

BBC News Online (2007). Hackers lift the bonnet on iPhone (http://news.bbc.co.uk/1/hi/technology/6269014.stm).

Cox, R.S., 3rd, Surette, M.G., and Elowitz, M.B. (2007). Mol. Syst. Biol. (Stevenage) *3*, 145.

Gardner, T.S., di Bernardo, D., Lorenz, D., and Collins, J.J. (2003). Science *301*, 102–105.

Guido, N.J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C.R., Elston, T.C., and Collins, J.J. (2006). Nature *439*, 856–860.

Kaplan, S., Bren, A., Zaslaver, A., Dekel, E., and Alon, U. (2008). Mol. Cell *29*, 786–792.

Ma'ayan, A., Jenkins, S.L., Neves, S., Hasseldine, A., Grace, E., Dubin-Thaler, B., Eungdamrong, N.J., Weng, G., Ram, P.T., Rice, J.J., et al. (2005). Science *309*, 1078–1083.

Murphy, K.F., Balazsi, G., and Collins, J.J. (2007). Proc. Natl. Acad. Sci. USA *104*, 12726–12731.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Nat. Genet. *31*, 64–68.

Solem, C., and Jensen, P.R. (2002). Appl. Environ. Microbiol. *68*, 2397–2403.

Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S., Milo, R., Pinter, R.Y., Alon, U., and Margalit, H. (2004). Proc. Natl. Acad. Sci. USA *101*, 5934–5939.