Commentary

Design then mutate

Jeff Hasty*

Department of Bioengineering, University of California at San Diego, La Jolla, CA 92093-0412

The development of a systematic design process for synthetic gene regulatory networks is an intriguing prospect that could improve both our understanding of regulatory motifs and our techniques for constructing novel applications. Such a design process would entail analysis tools capable of reliably predicting the *in vivo* behavior resulting from a given regulatory architecture. As with the process of designing an electrical circuit, the design tools would take the form of a set of equations that could be numerically simulated to predict the output of a given device. This type of "device genetics" would have an important influence on the development of therapeutic applications, as one can envision the design of complex genetic circuits for monitoring or controlling cellular processes.

Although the pursuit of a systematic program for modeling genetic circuits is within the constraints of current biotechnology, the deduction of the governing dynamical processes that describe these regulatory networks is not straightforward. Many complications can confound a model description. For example, fluctuations in local concentrations or in the conformational properties of regulatory proteins are not easily reconciled in models that focus on phenomena at the regulatory level. Likewise, proteins can aggregate near the DNA or slide along it, and active processes, such as transport, are not easily incorporated. These context-dependent complications often result in a mismatch between the lump kinetic modeling parameters that are measured *in vitro* and those that are actually present *in vivo*. This mismatch can render difficult the forward design of useful genetic circuits. In this issue of PNAS, Yokobayashi *et al.* (1) address this difficulty by demonstrating how directed evolution can be used to tune a designed gene circuit to a desired behavior.

Directed evolution harnesses mechanisms used during natural selection to generate and then identify evolutionary adaptations to novel environmental demands. It typically consists of the introduction of mutations via random mutagenesis or recombination, followed by the selection or screening for properties specified by the experimenter. Although
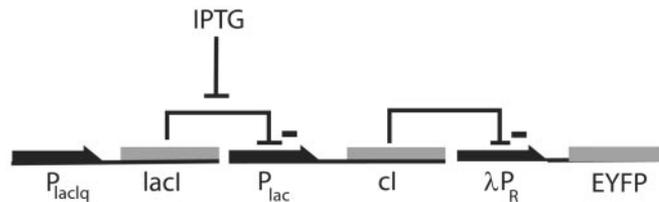

Fig. 1. The essential interactions of the three-gene regulatory circuit explored by Yokobayashi *et al.* (1).

this technique has been widely used to expand the library of enzymes for industrial and research purposes (2), its utility in the design process for gene regulatory networks has not been systematically investigated. In a series of experiments on an engineered gene network, Yokobayashi *et al.* show how directed evolution can be used to transform a nonfunctional genetic circuit into a library of "mutant" devices that are fully functional in the sense that their behavior is in accordance with the rules that govern logical gates. The starting point for their study was a three-gene network originally explored in ref. 3 by one of the authors of ref. 1 (Weiss). The essential interactions of this regulatory circuit are best understood from left to right in Fig. 1. The *lacI* gene is expressed from an unregulated promoter P_{lacIq}, and its product LacI represses the P_{lac} promoter. Likewise, the P_{lac} promoter controls the expression of the *cI* gene, and its product CI represses the P_R promoter. The P_R promoter controls the production of the enhanced yellow fluorescent protein (EYFP), which represents the "output" for the circuit, whereas the chemical inducer isopropyl β-D-thiogalactopyranoside (IPTG), which binds to LacI tetramers and renders them unable to repress P_{lac}, provides an external control over the network and thus represents the "input."

Guided by Fig. 1, heuristic reasoning would suggest the following two limiting input/output behaviors for this circuit: (*i*) with no IPTG, LacI shuts off the production of CI so that EYFP is produced, whereas (*ii*) for large amounts of IPTG, LacI cannot shut off CI and EYFP production is repressed. Thus, as the IPTG concentration is increased, one should observe a transition between these extremes as the output of EYFP proceeds from the "on" to "off" state. Although this logic seems reasonable, the devil turns up in the details as the leaky nature of the P_{lac} promoter leads to the repression of P_R regardless of IPTG concentrations, and thus no EYFP output for any input (3). In the parlance of electrical engineering, the first two promoter gene pairs form an implies gate, whereas the last pair is an inverter, and the lack of functionality arises because of a signal mismatch between the gates.

Yokobayashi *et al.* introduced random mutations to the nonfunctional device and successively screened the mutant circuits for those that obeyed the two limiting output behaviors for low and high IPTG inputs. They first used the process of error-prone PCR to introduce mutations targeting the *cI* gene. They then grew colonies in the absence of IPTG and observed that approximately half of the mutants were fluorescent, indicative of the on state for a functional device. They then selected these functional colonies, grew them in the presence of IPTG, and screened for those that were in the nonfluorescent off state ($\approx$5–10% were observed). Then, because the output of these mutant circuits had responded properly to both limits of the input (with and without IPTG), they were deemed candidates for a functional device and selected for further testing. In this final test, which involved the measurement of fluorescence for each mutant circuit as the concentration of IPTG was increased, they observed circuit responses beginning in the on state then transitioning to the off state. The results demonstrated that a number of

www.pnas.org/cgi/doi/10.1073/pnas.012694899

functional circuits with differing response curves had indeed been generated.

Once a set of functional circuits was generated, Yokobayashi *et al.* then explored the nature of the mutations that lead to these successful constructs. This analysis revealed that only a small number of mutations were required for the generation of an operational circuit. They then characterized the mutations in terms of protein functionality and found that the mutations were related to CI dimerization, dimer–dimer cooperativity, and dimer–DNA binding. As expected, these functional processes all are related to the degree in which CI is able to repress the P_R promoter, so that the mutations effectively used these avenues to tune the CI output to a range appropriate for input to P_R.

In light of the biological importance of robust genetic regulatory behavior, the demonstration that very subtle mutations can result in dramatic changes in the output characteristics of a genetic device is interesting. From an engineering perspective, perhaps it should come as no surprise that the designed gene circuit would be quite fragile to a small number of mutations because the design included no control circuitry. This observation supports the notion that real genomic wiring should have a significant battery of control circuitry designed to compensate for the fragility of its essential subnetworks (ref. 4 and references therein). Indeed the approach of Yokobayashi *et al.* of carefully analyzing the effect of mutation on network response could shed light on the types of control mechanisms that are used by real genomes. In this regard, future studies might focus on gene circuits constructed with network control designed to mitigate the effect of mutations.

Although one emphasis of the Yokobayashi *et al.* study is that directed evolution can be used as an additional step for fine-tuning a rationally engineered genetic circuit, their approach may also play an important role in the development of quantitative models. In this regard, structure–function insights obtained from the directed evolution experiments might be incorporated into the formulation of modeling schemes and lead to improved predictive capability. Yokobayashi *et al.* explored a related point in the context of mutations that altered the dimerization of the CI protein. Because the directed evolution experiments generated one of the functional circuits by altering the sequence that influences CI oligomerization, they tested the effect of changing the dimerization rate in a mathematical model for the circuit. As demonstrated in the experiment, they found that decreasing the dimerization rate for CI leads to a functional model circuit. This ability to connect mutations to functional changes
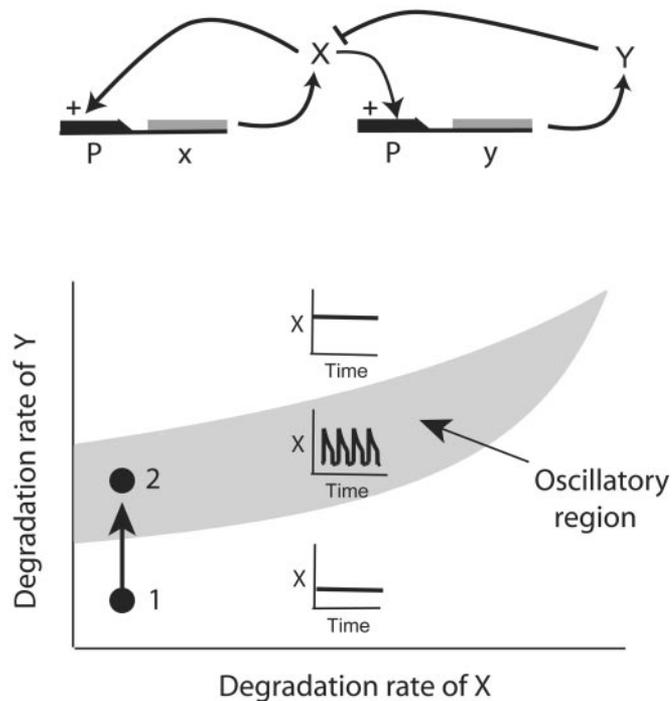
Fig. 2. The coupling of mutagenesis with epigenetic type behavior. (*Upper*) A generic two-gene network consisting of the genes *x* and *y*. Both proteins (X and Y) are under the control of a promoter that is activated by the protein X, and protein Y is a protease for X. Oscillations can arise because Y degrades X and thus reduces its own expression level (because X activates transcription of y). (*Lower*) Depending on the parameter values, the network can exhibit different phenotypic behaviors. A phase plane diagram indicates the degradation rates for X and Y that lead to either equilibrium or oscillatory behavior for the network. The path from point 1 to point 2 depicts how mutations that increase the degradation rate of Y could result in an abrupt change in the output behavior of X (and thus Y) from equilibrium to oscillatory behavior. See ref. 14 for a detailed analysis of this model network and ref. 10 for an experimental demonstration of oscillations in a synthetic gene network.

that, in turn, can be incorporated into model formulation could aid in the development of reliable large-scale models for whole genome networks. As techniques for reconstructing regulatory networks become increasingly reliable (5, 6) the need for such models is becoming evident as researchers struggle to understand how dynamical processes stemming from regulatory network interactions lead to phenotypic behavior (7–9).

Although the directed evolution approach used in this study screened for static on-off properties of a logical network, it sets the stage for an investigation into the screening for dynamical processes. For example, the technique might be applied to a synthetic network that generates oscillations (10) to select for a particular period or amplitude, or to a library of synthetic switches (11) to develop a very fast switch. In the case of switches, one could imagine replacing the visual screening technique with a process that uses cellular communication[†] such

that the faster switching cells trigger the death of the slow. In this way, the cells themselves would perform the screen, so that the experimenter need only introduce the mutations as was done by Yokobayashi *et al.*

An engineered gene circuit can be constructed to have the potential for multiple types of phenotypic behavior (10–12), and these phenotypes show epigenetic heritability. In such cases, the actual observed behavior depends on the particular value of the input control or the initial state of the network. In the language of dynamical systems theory, such a system is characterized by multiple attractors, whereas in biology, it is said to have epigenetic properties, because different phenotypic states can arise from two genetically identical cells and be stably inherited. In such systems it is possible for the network behavior to abruptly change from one epigenetic state (attractor) to another as a particular property of the system is changed. For example, the dynamics of protein concentration in a genetic network could suddenly change from a constant value to oscillations as a particular transcription factor is degraded beyond a threshold value (Fig. 2). Although such dramatic

changes can be generated by externally changing the input to a given network, they could also be achieved through mutation. In other words, mutagenesis could drive a genetic network to drastic change. In the context of naturally occurring gene regulatory networks, the existence of such an effect has been speculated (13), but because of the complexities of real genomic networks, systematic experiments demonstrating its existence are difficult. The work of Yokobayashi *et al.* establishes a methodology for exploring this important avenue in the context of an engineered gene circuit. As noted above, their findings show that only a few mutations were required to achieve functionality in a logical circuit. Perhaps, likewise in a circuit with epigenetic potential, dramatic changes in the output could be shown to be only a few mutations away. Such a demonstration would elucidate how small genetic changes can drastically alter the behavior of a gene network and could potentially lead to insights pertaining to the facilitation of evolution by regulatory network dynamics.

1. Yokobayashi, Y., Weiss, R. & Arnold, F. H. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 16587–16591.
2. Farinas, E., Tulter, T. & Arnold, F. H. (2001) *Curr. Opin. Biotechnol.* **6,** 545–551.
3. Weiss, R. (2001) Ph.D. thesis (Massachusetts Institute of Technology, Cambridge).
4. Csete, M. E. & Doyle, J. C. (2002) *Science* **295,** 1664–1669.
5. Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlet, D. R., Aebersold, R. & Hood, L. (2001) *Science* **292,** 929–934.
6. Covert, M. W., Schilling, C. H., Famili, I., Edwards, J. S., Goryanin, I. I., Selkov, E. & Palsson, B. O. (2001) *Trends Biochem. Sci.* **26,** 179–186.
7. Hasty, J., McMillen, D. & Collins, J. J. (2002) *Nature* **420,** 224–230.
8. Oltvai, Z. & Barabasi, A. (2002) *Science* **298,** 763–764.
9. Ibarra, R. U., Edwards, J. S. & Palsson, B. O. (2002) *Nature* **420,** 186–189.
10. Elowitz, M. B. & Leibler, S. (2000) *Nature* **403,** 335–338.
11. Gardner, T. S., Cantor, C. R. & Collins, J. J. (2000) *Nature* **403,** 339–342.
12. Becskei, A., Séraphin, B. & Serrano, L. (2001) *EMBO J.* **20,** 2528–2535.
13. Glass, L. & Kauffman, S. A. (1972) *J. Theor. Biol.* **34,** 219–237.
14. Hasty, J., Isaacs, F., Dolnik, M., McMillen, D. & Collins, J. J. (2001) *Chaos* **11,** 207–220.