# EVOLUTIONARY DESIGN OF GENETIC CIRCUITS AND CELL-CELL COMMUNICATIONS

YOHEI YOKOBAYASHI[*], CYNTHIA H. COLLINS[*], JARED R. LEADBETTER[†],
RON WEISS[‡] and FRANCES H. ARNOLD[*,**]

[*] *Division of Chemistry and Chemical Engineering,* [†] *Division of Engineering and Applied Science, California Institute of Technology 1200 E. California Blvd, Pasadena, CA 91125, USA*
[**] *frances@cheme.caltech.edu*

[‡] *Department of Electrical Engineering, Princeton University, B-312, E-Quad, Princeton, NJ 08544, USA*
[‡] *rweiss@princeton.edu*

Synthetic genetic circuits are artificial networks of transcriptional control elements inserted into living cells in order to 'program' cellular behavior. We can extend this application to programming population behavior by incorporating cell-cell communications capabilities. By designing and building such networks, cellular circuit engineers expect to gain insight into how natural genetic networks function with remarkable robustness, stability, and adaptability to changing environments. Programmed cells also have promising applications in biotechnology and medicine. A major challenge that biological circuit engineers face is the difficulty of predicting circuit performance at the design stage, with the consequence that actual construction requires significant experimental effort, even for very simple circuits. To address this fundamental obstacle we propose the use of laboratory evolution methods to create new circuit components and optimize circuit performance inside living cells.

*Keywords*: Genetic circuit; directed evolution; quorum sensing; cell-cell communications; genetic regulatory networks.

## 1. Introduction

Recent remarkable advances in genome sequencing, proteomics, bioinformatics, and biological information processing have inspired scientists to try to understand the complexity of biological systems *as they are* [11]. While these "top-down" approaches seem to have gained momentum with the vast amount of data that are being produced using technologies such as DNA sequencing and DNA chips, a great number of experimental and theoretical problems remain to be solved. Researchers are also studying biological systems by building networks from the bottom up. Gardner *et al.*, for example, designed a simple two-gene genetic network, synthesized it on a plasmid (autonomously replicating DNA in bacteria), introduced it into *Escherichia coli*, and demonstrated that the artificial network functioned as a bistable toggle switch [8]. Elowitz and Leibler constructed a three-gene network

that functions as an oscillator [5]. Synthetic networks such as these will provide new insights into the nature of genetic regulatory systems and biological complexity. By designing, simulating and building simple genetic circuits, we can study the origins and consequences of characteristics such as robustness and adaptability, as well as explore different evolutionary scenarios for the creation and diversification of regulatory networks. It is also foreseeable that synthetic genetic circuits will have significant practical implications in biotechnology and medicine, as we acquire the ability to program cells to perform complex tasks.

While biological systems offer great opportunities for engineering, they also pose tough challenges to the cellular programmer who would dare ask a cell to follow new instructions. Numerous biochemical parameters govern the behavior of a genetic regulatory circuit—among them are protein and mRNA stability, binding constants, formation of higher order structures, transcription and translation rates—and these are hard to measure precisely inside a living cell. Considering this and other factors such as the stochastic nature of gene expression and the complex behavior of the host cells in a changing environment (circuit performance is strongly context dependent), we must conclude that predicting the detailed performance of a circuit is difficult.

We believe that a partial solution to the design problem can come from mimicking the strategy Nature used to create the very systems we wish to engineer—evolution. By adapting the simple yet powerful algorithm of evolution, engineers have been able to alter the characteristics of individual nucleic acids [2] and proteins [1]. The same algorithm can also be applied to design of biosynthetic pathways [16] and even whole organisms [22]. Evolution in the laboratory, often called 'directed evolution' or 'molecular breeding' to draw attention to the fact that it is headed in a particular (functional) direction, can generate solutions to difficult biological design problems for which we do not have sufficient understanding to engineer by conventional 'rational' approaches. Here, we describe how directed evolution will be able to assist the forward engineering of genetic regulatory circuits and cell-cell communications systems.

## 2. Correction of a Mismatch within a Genetic Circuit

Natural genetic regulatory networks are comprised of modules that have specific network topologies and functions [17], and it is clear that this modularity will be important for building synthetic genetic circuits. However, engineering the interfaces between multiple genetic modules has proven to be a significant technical challenge. A good example is the simple genetic circuit depicted in Fig. 1 [18, 19].

Logic operations in genetic circuits are biochemically mediated by promoters. A stretch of DNA sequence located upstream of a gene, a promoter recruits RNA polymerase to transcribe the downstream genes. Some promoters are either activated or repressed by proteins (transcription factors) that bind to specific regions within promoters.
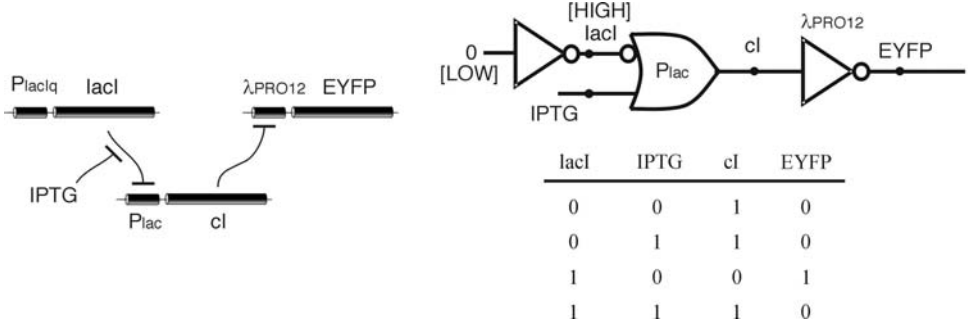
Fig. 1.   Schematic diagrams of the genetic circuit studied and the truth table outlining the logic states. Lac repressor (LacI) is constitutively expressed from the $P_{lacIq}$ promoter (always HIGH). LacI binds and negatively regulates transcription from the $P_{lac}$ promoter controlling the *cI* gene. Addition of IPTG inhibits LacI binding to $P_{lac}$ thereby allowing the expression of CI protein, which in turn represses transcription from $\lambda_{PRO12}$ (a synthetic promoter based on the natural $\lambda_{PR}$ from which OR3 is deleted). The observable output, EYFP expression, represents the activity of the $\lambda_{PRO12}$ promoter. The logical states of the circuit components as designed are summarized in the truth table.

There are three components, or gates, in this circuit. LacI is constitutively expressed from the $P_{lacIq}$ promoter, meaning LacI is synthesized regardless of the state of the circuit. The promoter is logically represented by an inverter (a NOT gate) having a constant input (LOW) and LacI as the output (always HIGH). The second (IMPLIES) logic gate, mediated by the $P_{lac}$ promoter, functions with two inputs, LacI protein and isopropyl $\beta$-D-thiogalactoside (IPTG), and an output in the form of CI protein. The expression level of CI protein follows the IMPLIES logic with respect to the LacI and IPTG inputs, as depicted in the truth table (Fig. 1). The third gate (of $\lambda_{PRO12}$) is an inverter in which CI protein is the input and yellow fluorescent protein (EYFP) is the output. EYFP is conveniently detected by its fluorescence. The fluorescence intensity of EYFP was measured under different concentrations (input) of IPTG. Since LacI was fixed to be in HIGH state, the EYFP output was expected to be inverted with respect to the IPTG level. Without IPTG, CI is supposed to be LOW and EYFP is HIGH. With IPTG, CI is supposed to be HIGH and EYFP is LOW.

The initial construct did not function as expected: EYFP output was always LOW, both in the absence and presence of IPTG. The lack of function of the circuit was attributed to a mismatch between the two individually well-characterized gates. In the absence of IPTG (LOW), CI is also LOW. However, the LOW CI level from the first gate was actually interpreted as HIGH by the second gate, the inverter. In molecular terms, the small amount of CI repressor that was synthesized from $P_{lac}$ in the absence of IPTG was still sufficient to repress the $\lambda_{PRO12}$ promoter, so that no EYFP was produced. This case illustrates how important it is that the multiple components of a synthetic genetic circuit be properly matched. Weiss eventually succeeded in rectifying the circuit with a series of rational modifications guided

by biochemical data and semi-quantitative simulations. We recently demonstrated that the same circuit could be "debugged" efficiently by directed evolution [21].

To optimize the circuit by evolution, we applied random point mutagenesis to the *cI* gene that encodes CI protein, creating a pool of circuits with altered CI properties. CI is a repressor that binds to the $\lambda_{\text{PRO12}}$ promoter as a tetramer. By modifying DNA binding affinity, oligomerization constant, and/or transcription efficiency, we expected to discover appropriate sets of parameters that allow the circuit to function. It is extremely difficult, if not impossible, to predict the specific amino acid changes that would confer these properties. But, since the output of the circuit was coupled to a fluorescent protein, we could easily introduce random mutations and screen for the mutants that exhibited the desired function out of thousands of mutants.

After two rounds of mutagenesis and screening, we identified dozens of functioning circuits. Figure 2 shows detailed characterization (transfer curves) of two of these mutants, along with one of the rationally engineered functional circuits made by Weiss. The mutations in our functional circuits are assumed to decrease the cooperativity of oligomerization of the CI protein, and, through that, the DNA binding affinity. Thus, the directed evolution discovered solutions to the debugging problem (adjusting cooperativity) that were not considered during the rational engineering. It would have been very difficult to predict such mutations *a priori*.

During our investigations, we found that it is extremely important to control the experimental conditions to obtain reproducible circuit performance. Protein
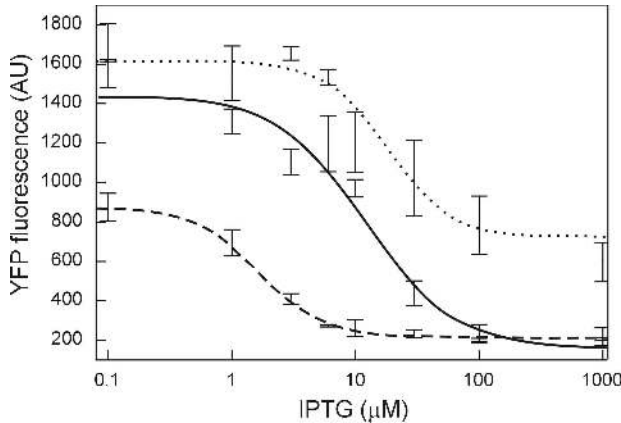


Fig. 2.    Transfer curves of selected mutants. Fluorescence of EYFP was measured after culture growth under various IPTG concentrations. The EYFP response of the original circuit (not shown) was low and completely unresponsive to variations in IPTG levels. The three curves (dashed line: R3 mutant rationally engineered by Weiss, solid line: C3 mutant discovered by directed evolution, dotted line: A4-04 mutant discovered by directed evolution) exhibit distinct device characteristics with different HIGH/LOW levels and transition from HIGH to LOW output at different IPTG thresholds [21].

expression levels are very sensitive to cell growth conditions such as temperature, aeration, cell density, and growth medium. The cells were grown overnight to stationary phase, which was diluted 1/250 into fresh medium (Lurian–Bertani) and grown for 5.5 hours to log phase. The cells were harvested after a second dilution of 1/250 into fresh medium and further incubation of 6.0 hours. The serial dilution steps were critical for achieving reasonable reproducibility by keeping the cells in a pseudo-steady state, where the protein production rate is balanced by the protein degradation and dilution rate. Other experimental parameters such as temperature (37°C), medium composition, and measurement protocols were carefully controlled.

It is also expected that circuit behavior will depend on the host bacterium's genotype. In fact, Guet *et al.* observed that the same circuits performed different logic functions under different genotypic environments [9]. This extreme sensitivity of biological systems to environment further underscores the utility of an evolutionary approach to fine-tuning circuit performance. We should be able, using directed evolution, to adjust for the small (or not so small) environmental changes that happen when a circuit is transplanted from one cell type to another.

The transfer curves in Fig. 2 also show the varied 'device physics' that the directed evolution experiment produced. The transfer curves indirectly represent the device physics of the $CI/\lambda_{PRO12}$ inverter because CI and IPTG levels can be correlated experimentally. The device physics can also be measured for other inverters (based on different repressor/promoter pairs) using the same LacI/IPTG/repressor gate of Fig. 1. We hope to use this system to characterize other inverters, with which we will be able to design genetic circuits more reliably than is currently possible.

## 3. Engineering Intercellular Communication

To increase the functionality of *de novo* genetic circuits for programming cell populations and coordinating behavior across a population, we want to incorporate the ability to pass messages from cell to cell. Intercellular communications will allow us to design and implement more complex circuits by circumventing limitations on the number of exogenous components we can put into a single cell. In some cases, it may also be useful to compartmentalize pieces of circuits to limit undesirable interactions and promote robust behavior. The resulting integrated communications and genetic regulatory circuits will allow us to program populations of genetically distinct cells which can cooperate to carry out tasks in a reliable and predictable manner.

Gram-negative bacteria use the detection of a class of small organic molecules, acyl-homoserine lactones (AHLs), to sense and respond to local changes in their environment by altering gene expression [6, 14]. AHLs are typically produced during the entire life cycle of these 'quorum-sensing' bacteria, and have been shown to be diffusible across cell membranes [10, 15]. When the bacteria grow to high population densities, the local concentration of the AHL increases [6]. Upon reaching a critical concentration, typically between 10 nM and 10 $\mu$M, the AHL signal interacts with

a specific DNA-binding protein inside the cell. This complex in turn activates or represses the expression of specific genes or sets of genes. Enzymatic activities that serve to inactivate these signals have also been identified [3, 12, 13]. This system is ideal as a starting template for generating intercellular communication circuits because components that synthesize, respond to and degrade a diffusible chemical message have all been identified. We have already started building intercellular communication systems using these mechanisms by transferring genetic elements across gram-negative bacteria and combining them with synthetic gene networks [20].

Unlike computer networks, where physical wires connect the disparate elements, biological circuits use chemical diffusion to carry data from one component to the next. Specific wiring requires specific interactions among molecules based on molecular affinities, so that the devices used to make up synthetic circuits must have not only the desired sensitivities but also enough specificity to prevent undesirable crosstalk. This means that a large number of biochemically distinct components will be required to implement complex circuits. These components can be generated rapidly by directed evolution. For example, genetic elements that perceive and respond to non-natural signal molecules will be extremely useful for building new intercellular signaling capabilities. To make these, we are screening libraries of mutant AHL-mediated activator proteins (LuxR) from *Vibrio fischeri* [6] for their ability to activate expression in the presence of non-natural AHLs.

To do this, we have set up an experimental system where the effects of externally added AHL on LuxR- mediated expression of green fluorescent protein (GFP) can be monitored rapidly in a large number of mutants (Fig. 3). We characterized the ability of LuxR to activate gene expression in the presence of a variety of different AHLs, both natural and not. We have now identified two LuxR mutants that have increased sensitivities towards our first target non-natural AHL (Fig. 4).
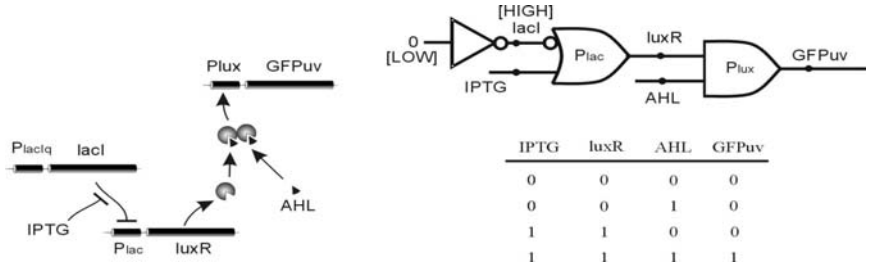


Fig. 3. Schematic of the genetic circuit used in screening for protein-signal interactions and the truth table outlining the logic states. LuxR expression is induced from the $P_{lac}$ promoter by adding IPTG (always present when screening, always HIGH). LuxR binds to an AHL molecule and positively regulates transcription of GFPuv. The observable output, GFPuv expression, represents activity of the $P_{lux}$ promoter. The amount of GFPuv expression is dependent upon the fatty acid chain moiety of the AHL, with greatest activity seen with LuxR and its cognate signal, 3-oxo-hexanoyl-homoserine lactone (3OC6HSL) [4, 7]. The logical states of the circuit components as designed are summarized in the truth table.
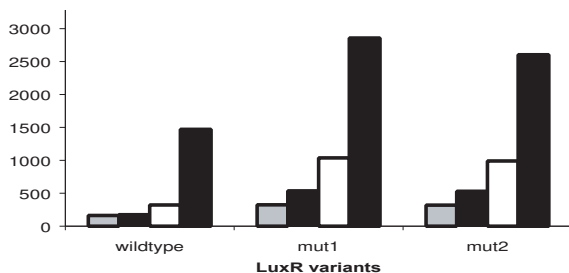
Fig. 4. AHL-mediated gene activation of selected LuxR mutants. The fluorescence of GFPuv in response to wildtype LuxR or each of two mutants identified in initial screens and in the presence of 3-oxo-hexanoyl-homoserine lactone (3OC6HSL) or hexanoyl-homoserine lactone (C6HSL) are shown (grey: 100nM C6HSL, checked: 1 $\mu$M C6HSL, white: 100 nM 3OC6HSL, black: 1 $\mu$M 3OC6HSL). 3OC6HSL is the AHL recognized by wild-type LuxR. C6HSL, which lacks the ketone moiety at C3, is our target for evolution of a LuxR that responds to the signal.

Our ideal toolbox of intercellular signaling components will contain LuxR proteins with different AHL sensitivities and specificities. In some cases it will be useful to engineer responses to micromolar (i.e. high) concentrations of a given signal, while in others a hair trigger might be desired, i.e. in the form of a protein that activates transcription in response to low nanomolar signal concentrations. Evolving the narrow signal binding specificity of these proteins will become important for minimizing undesired interactions with signaling machinery that has already evolved in Nature. By screening libraries of LuxR mutants against arrays of signaling molecules, we expect to be able to identify and characterize, with rapidity, LuxR variants that have a variety of such properties. Promising alleles will be used as stock for subsequent rounds of evolution, resulting in a collection of LuxR mutants with not only an array of desired properties, but also with unique combinations of those properties.

## 4. Conclusions

Synthetic genetic circuits are ideal targets for directed evolution because the output of a genetic circuit can be coupled to a fluorescent signal or a selectable marker. This makes screening or selection, which is the most critical step of a directed evolution experiment, particularly convenient. However, evolutionary design of complex biological systems is still quite challenging, and workable evolutionary strategies that balance library complexity against screening capabilities will have to be developed. In the examples described here, mutations were focused on specific biochemical components (proteins). While more 'blind' applications have also successfully altered metabolic pathways and even whole organisms [22], the cost is a explosion in combinatorial possibilities. To find useful mutants requires an evolutionary strategy that creates them at sufficient frequency that screening can uncover them. Creative screening or selection strategies will also be required to evolve circuits that exhibit

more complicated behaviors such as oscillations or shorter switching times.

## Acknowledgments

## References

[1]  Arnold, F. H., Design by directed evolution, *Acc. Chem. Res.* **31**, 125–131 (1998).

[2]  Bittker, J. A., Phillips, K. J. and Liu, D. R., Recent advances in the in vitro evolution of nucleic acids, *Curr. Opin. Chem. Biol.* **6**, 367–374 (2002).

[3]  Dong, Y. H., Xu, J. L., Li, X. Z. and Zhang, L. H., AiiA, an enzyme that inactivates the acylhomoserine lactone quorum-sensing signal and attenuates the virulence of *Erwinia carotovora*, *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3526–3531 (2000).

[4]  Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H. and Oppenheimer, N. J., Structural identification of autoinducer of photobacterium-fischeri luciferase, *Biochemistry* **20**, 2444–2449 (1981).

[5]  Elowitz, M. B. and Leibler, S., A synthetic oscillatory network of transcriptional regulators, *Nature* **403**, 335–338 (2000).

[6]  Fuqua, C., Parsek, M. R. and Greenberg, E. P., Regulation of gene expression by cell-to-cell communication: Acyl-homoserine lactone quorum sensing, *Ann. Rev. Genet.* **35**, 439–468 (2001).

[7]  Fuqua, C., Winans, S. C. and Greenberg, E. P., Census and consensus in bacterial ecosystems: The LuxR-LuxI family of quorum-sensing transcriptional regulators, *Ann. Rev. Microbiol.* **50**, 727–751 (1996).

[8]  Gardner, T. S., Cantor, C. R. and Collins, J. J., Construction of a genetic toggle switch in Escherichia coli, *Nature* **403**, 339–342 (2000).

[9]  Guet, C. C., Elowitz, M. B., Hsing, W. and Leibler, S., Combinatorial synthesis of genetic networks, *Science* **296**, 1466–1470 (2002).

[10]  Kaplan, H. B. and Greenberg, E. P., Diffusion of autoinducer is involved in regulation of the Vibrio-Fischeri luminescence system, *J. Bacteriol.* **163**, 1210–1214 (1985).

[11]  Kitano, H., Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).

[12]  Leadbetter, J. R., News and Views: plant microbiology - quieting the raucous crowd, *Nature* **411**, 748–749 (2001).

[13]  Leadbetter, J. R. and Greenberg, E. P., Metabolism of acyl-homoserine lactone quorum-sensing signals by *Variovorax paradoxus*, *J. Bacteriol.* **182**, 6921–6926 (2000).

[14]  Miller, M. B. and Bassler, B. L., Quorum sensing in bacteria, *Ann. Rev. of Microbiol.* **55**, 165–199 (2001).

[15]  Pearson, J. P., Van Delden, C. and Iglewski, B. H., Active efflux and diffusion are involved in transport of Pseudomonas aeruginosa cell-to-cell signals, *J. Bacteriol.* **181**, 1203–1210 (1999).

[16] Schmidt-Dannert, C., Umeno, D. and Arnold, F. H., Molecular breeding of carotenoid biosynthetic pathways, *Nat. Biotechnol.* **18**, 750–753 (2000).

[17] Shen-Orr, S. S., Milo, R., Mangan, S. and Alon, U., Network motifs in the transcriptional regulation network of Escherichia coli, *Nat. Genet.* **31**, 64–68 (2002).

[18] Weiss, R. and Basu, S. (2002). *First Workshop on Non-Silicon Computing, Cambridge, Massachusetts,* 54–61, http://www.hpcaconf.org/hpca8/.

[19] Weiss, R., Homsy, G. E. and Knight, T. F., Jr. (1999). *Dimacs Workshop on Evolution as Computation, Princeton, NJ*, Springer Verlag, 275–295,

[20] Weiss, R. and Knight, T. F., Jr. (2000). *DNA6: Sixth International Meeting on DNA Based Computers, Leiden, The Netherlands*, Springer-Verlag, 1–16,

[21] Yokobayashi, Y., Weiss, R. and Arnold, F. H., Directed evolution of a genetic circuit, *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16587–16591 (2002).

[22] Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P. C. and del Cardayre, S. B., Genome shuffling leads to rapid phenotypic improvement in bacteria, *Nature* **415**, 644–646 (2002).