Natural Computing (2005) 4:245–254
DOI 10.1007/s11047-004-7442-x

A dual selection module for directed evolution of genetic circuits

YOHEI YOKOBAYASHI[‡] and FRANCES H. ARNOLD*
*Division of Chemistry and Chemical Engineering, California Institute of Technology, mail code 210-41, Pasadena, California 91125, USA; (*Author for correspondence, e-mail: frances@cheme.caltech.edu); [‡]Present address: Department of Biomedical Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA*

Abstract. We have developed a genetic selection module which enables us to select for positive and negative expression of a gene by controlling the expression of two antibiotic resistance genes. This module can be used to tune the performance of genetic circuits that yield simple ON/OFF gene expression outputs depending on the environment (input). We have demonstrated the utility of the module by selecting a genetic inverter from a 200 fold excess of nonfunctional inverters in two rounds.

Key words: cellular computation, directed evolution, genetic circuits, selection, synthetic biology, synthetic gene circuits

Abbreviations: BLIP – β - lactamase inhibitory protein; Cb – carbenicillin; IPTG – isopropyl-β-D-thiogalactopyranoside; LB – Luria-Bertani; Tc – tetracycline

1. Introduction

Design and construction of synthetic genetic networks that perform programmed functions is a major new opportunity, and challenge, for biological engineering (Hasty et a1., 2001; Hasty et a1., 2002; Kaern et a1., 2003; Yokobayashi et a1., 2003; Wall et a1., 2004). One of the key challenges in building such networks is controlling and optimizing the many parameters that influence circuit performance, including protein and RNA stabilities, affinities and specificities of molecular interactions, and the effects of temperature and host genotype, most of which are neither easily measured nor manipulated. Recently, we demonstrated that we could optimize a synthetic genetic circuit by directing its evolution inside the host cells, by screening randomly-

mutated circuits ($\sim 10^4$ clones) for a desired circuit output (Yokobayashi et a1., 2002). Evolutionary design complements imperfect ''rational'' circuit design approaches by providing a means to tune circuit performance, even when the underlying mechanism is poorly or not at all understood. Here we describe a generic selection system that can be used to identify rare functional circuits and to optimize those circuits by coupling the circuit's output to the survival or death of the host bacteria. In this way, we can use bacterial growth under appropriate conditions to select for circuits whose outputs match the performance criteria. With genetic selection, it will be possible to enrich functional circuits in significantly larger libraries of mutants (typically $\sim 10^7$, limited mainly by the ability to make such libraries).

2. Materials and methods

2.1. *Plasmid construction*

All plasmids were constructed by standard molecular biology protocols, similar to our previous report (Yokobayashi et a1., 2002). A gene encoding β-lactamase inhibitory protein (BLIP) was cloned from pTP154, kindly provided by Tim Palzkill (Huang et a1., 2000). We have extensively modified this gene using synthetic oligo DNA to lower the GC content and homology of the periplasmic leader sequence region with that of β-lactamase, in order to improve genetic stability in *vivo* without changing the amino acid sequence. The re-engineered BLIP coding gene sequence is: 5′-atgcttttatataaaatgtgtgacaatcaaaattatgg ggttacttacatgaagtttttattggcattttcgcttttaataccatccgtggtttttgcaagtagtgc aggtgttatgacaggagcaaaattcacgcagatccagtttggtatgacacgtcagcaggtcctc gacatagcaggtgctgagaactgtgagactggtggatcgttcggtgacagcatccattgtcgtgg acatgcagcaggagactattatgcatacgcaaccttcggcttcaccagcgcagctgcagacgca aaggtggattcgaaaagccaggaaaaactgcttgcaccaagcgcaccaactcttactcttgcta agttcaaccaagtcactgttggtatgactagagcacaagtacttgctaccgtcggacaggttctt gtaccacttggagtgagtactatccagcatatccatcgacggcaggagtgactctcagcctgtc ctgcttcgatgtggacggttactcgtcgaccggggttctaccgaggctcggcgcacctctggttc acggacggggtgcttcagggcaagcggcagtgggaccttgta-3′ The selection plasmid $P\lambda_{PRO12}$Select contains a ColEl replication origin; circuit plasmids pInverter, pAlwaysON, pAlwaysOFF contain a p15A replication origin allowing the selection and circuit plasmids to coexist in host cells.

2.2. *Selection experiment*

Circuit plasmids were transformed into DH5αPRO cells (Clontech) harboring the selection plasmid ($p\lambda_{PRO12}$Select) using the *Z*-Competent *E. coli* Transformation Kit (Zymo Research). All Luria–Bertani (LB) broth plates were supplemented with kanamycin (50 μg/ml) to select for circuit plasmid-containing cells and spectinomycin (25 μg/ml) for maintaining the host genes. In addition, the plates contained tetracycline (50 μg/ml) for the positive selection or carbenicillin (100 μg/ml) and isopropyl-β-D-thiogalactopyranoside (IPTG, 1 mM) for the negative selection.

Selected cells were recovered by overlaying liquid LB medium over the agar plates, and cultured at 37°C overnight in the presence of kanamycin (50 μg/ml), spectinomycin (25 μg/ml), carbenicillin (100 μg/ml), and IPTG (1 mM) for subsequent DNA isolation by miniprep (QIAGEN). Approximately 70 ng of DNA was digested with 5 units of *Apa*LI at 37°C for 1 h and column purified (Zymo Research). The DNA was then used for a second round of selection.

Colony PCR was performed to identify pInverter plasmid-containing cells after each round of dual selection. The Colony Fast-Screen Kit (Epicentre) was used according to instructions, with two primers: Plac-F, 5′-GTTAGCTCACTCATTAGGCAC-3′; cI-R 5′-AAGCAATGCGGCGTTATAAGC-3′. Positive and negative controls (see below) were run simultaneously.

3. Results and discussion

Our goal was to develop a generic selection module which can be interfaced with any genetic circuit that yields simple ON/OFF gene expression as the output. Genetic circuits that fall in this category include a toggle switch(Gardner et al., 2000), various logic gates (Weiss et al., 1999; Buchler et al., 2003), and a bandpass filter (Basu et al., 2003). The selection module can be co-expressed in bacterial cells with the synthetic gene circuit, where the circuit's output serves as input to the selection module. Properly functioning circuits may then be selected from among a large number of nonfunctional or poorly functional circuits based solely on cell growth. To this end, we coupled both states, circuit ON and OFF, to cell survival. Selection for cells in the ON state can be achieved by coupling the expression of a gene required for survival (e.g. antibiotic resistance) to the ON output of the circuit. The

OFF state turns off expression of a gene that inhibits or interferes with a second gene required for survival. Our system builds on numerous genetic selection schemes developed for studying molecular interactions *in vivo* (protein-protein(Pelletier et al., 1999; Kim and Oakley, 2002), DNA-protein (Elledge et al., 1989; Mossing et al., 1991; Huang et al., 1994; Sera and Schultz, 1996), RNA-protein (Buskirk et al., 2003), etc.) by enabling selection for both the ON and OFF states of gene expression.

The specific implementation of the dual selection module is shown in Figure 1 and comprises two antibiotic resistance genes (*bla* and *tetA*) and a gene encoding the protein inhibitor of β-lactamase (*bli*). Genes *bli* and *tetA* are placed under the control of the output of the circuit (here represented by the λ_{PRO12} promoter), and *bla* is placed under a constitutive promoter. The selection module is conveniently integrated into a single plasmid, and the output promoter (λ_{PRO12}) can be replaced with an appropriate sequence for the circuit of interest. The ON state of the circuit results in the polycistronic expression of *bli* and *tetA* gene products from the output promoter. The *tetA* gene product confers resistance to tetracycline, while concomitant expression of BLIP encoded by *bli* which binds to and inhibits β-lactamase (Doran et al., 1990; Strynadka et al., 1996), resulting in the loss of resistance to carbenicillin. Thus cells with the circuit ON are sensitive to carbenicillin and resistant to tetracycline. Conversely, the circuit OFF state generates resistance to carbenicillin and sensitivity to tetracycline in the absence of

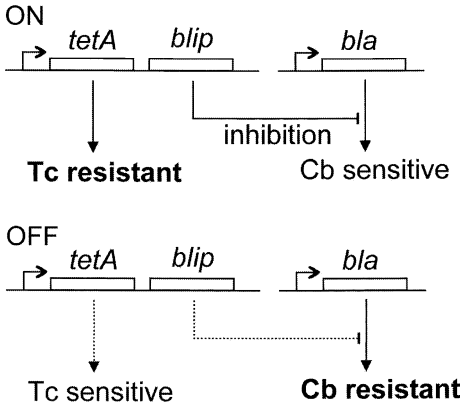

Figure 1. Schematic of the selection module. Two genes *bli* and *tetA* are placed under the control of the output promoter of the circuit of interest, while *bla* is expressed constitutively from its own promoter. Depending on the expression state of the output promoter, the host cell becomes resistant to tetracycline (Tc, ON state) or carbenicillin (Cb, OFF state). The dotted lines indicate the absence of expression in the OFF state.

tetA expression. A genetic circuit designer can select from a large pool of mutant circuits those that yield the desired output (ON/OFF) by exposing the pool to the appropriate antibiotic(s) and recovering the surviving cells.

We tested the module by selecting a genetic inverter similar to one studied previously (Yokobayashi et al., 2002) as shown in Figure 2. The core function of this circuit is a simple inversion of the genetic expression state with respect to the lambda repressor CI and the output gene expression from the λ_{PRO12} promoter. Expression of CI is driven by P_{lac} which is repressed by constitutively-expressed lac repressor (LacI). The actual input to the circuit is IPTG, which is added externally and binds to LacI, inducing expression of CI. Thus the logical expression of input/output states is as shown in Figure 2. Note that, with the exception of the output promoter λ_{PRO12}, the circuit is contained in a separate plasmid (pInverter) from the selection module. This enables efficient manipulation and construction of mutant circuit libraries and is important in managing unintended mutations (see below).

As controls, we prepared two "nonfunctional" inverter circuits, neither of which responds to the input (IPTG) (Figure 3). pAlwaysOFF is identical to pInverter except for the P_{lac} promoter, which is replaced with P22 phage P_{ant} promoter. Since LacI does not bind P_{ant}, pAl-

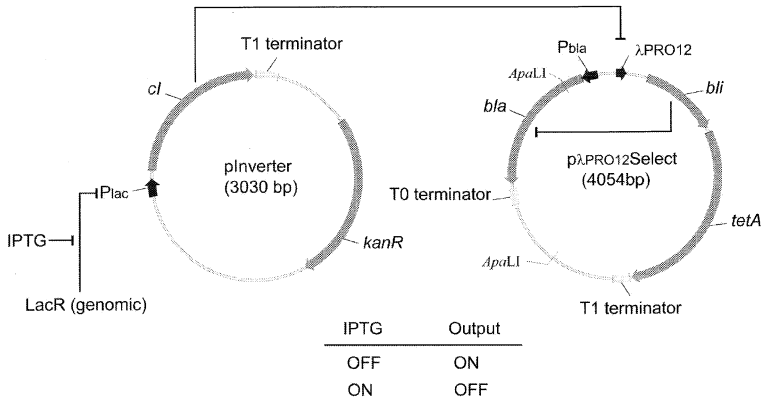

Figure 2. Schematic of the inverter circuit and selection module as used for the enrichment experiment. pInverter contains the lambda repressor gene *cI* under the control of the P_{lac} promoter, which is repressed by LacI encoded in the host genome (DH5αPRO). The selection module $p\lambda_{PRO12}$Select harbors the output promoter of the circuit λ_{PRO12} that is repressible by CI and the genes used for selection (*bli, tetA, bla*). The two plasmids have compatible origins of replications (pl5A for pInverter and ColEl for $p\lambda_{PRO12}$Select).

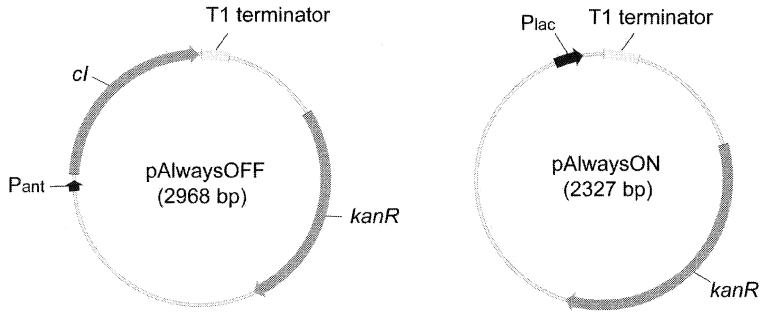

Figure 3. Non-functional circuit plasmids used as controls and to determine the background in the selection experiment. pAlwaysOFF constitutively expresses *c*I under P_{ant} promoter while pAlwaysON lacks the *c*I gene, resulting in constitutive expression from the output promoter in the selection module.

waysOFF constitutively expresses CI and results in OFF output regardless of the IPTG concentration. pAlwaysON is essentially a blank plasmid from which pInverter was derived and encodes no meaningful polypeptide under P_{lac} control, resulting in constitutive expression from the output promoter (ON output) regardless of the input.

To test whether the selection module can recover the functional inverter from a background of nonfunctional ones, the three plasmids pInverter, pAlwaysOFF, and pAlwaysON were mixed in 1:100:100 molar ratio and transformed into *Escherichia coli* DH5αPRO cells harboring the selection module $p\lambda_{PRO12}$Select. The transformed cells were plated on tetracycline-containing LB plates without IPTG to select for cells in the ON state (those containing pInverter or pAlwaysON). Approximately 2900 colonies grew after overnight incubation at 37°C. The cells were recovered by overlaying liquid LB medium, and portions were immediately plated on carbenicillin-containing LB plates with 1 mM IPTG and incubated overnight at 37°C in order to select for cells that successfully switched to the OFF state. Cells were recovered from plates (>2000 total colonies) by overlaying liquid LB medium and liquid cultured in the presence of carbenicillin (100 μg/ml), IPTG (1 mM), kanamycin (50 μg/ml), and spectinomycin (25 μg/ml). The plasmids were recovered by miniprep.

The plasmid mixture recovered from this dual selection is expected to contain enriched pInverter, as well as some amount of carried-over pAlwaysON and pAlwaysOFF, in addition to $p\lambda_{PRO12}$Select. The mixture was digested with restriction enzyme *Apa*LI, which only cleaves $p\lambda_{PRO12}$Select and removes it from the pool. The treated plasmid pool was

column-purified and transformed into DH5αPRO cells containing $p\lambda_{PRO12}$ Select for a second round of selection following the same protocol.

From carbenicillin/IPTG plates of the first and second rounds of selection, 48 colonies were randomly picked and tested for the presence of pInverter by colony PCR. Thirty-seven out of 48 (77%) colonies were confirmed to contain pInverter from the first selection, and all 48 colonies contained pInverter after the second round. The result translates to a 155 fold enrichment of pInverter in the first round of selection (the maximum enrichment attainable in this experiment is 200 fold). The 11 colonies not harboring pInverter after the first round of selection could represent cells containing pAlwaysON or pAlwaysOFF that were carried over from the initial selection due to incomplete killing or cells containing mutated $P\lambda_{PRO12}$Select, as discussed below.

As with all genetic selection and screening methods, false positives were observed during initial development of the system (Huang et al., 1994; Mossing et al., 1991; Sera and Schultz, 1996; Fromknecht et al., 2003). Depending on the experimental conditions, up to 0.1 % of the total transformed cells appeared to be false positives (data not shown). Some of the selection module plasmids recovered from these cells were found to have larger than expected size, suggesting recombination arising from the two-plasmid environment. Many other mutations can be anticipated that disrupt the promoter sequences controlling the antibiotic resistance genes or that inactivate the *bli* gene and lead to false positives. These limitations are inherent to almost all genetic selection systems that confer antibiotic resistance or repress toxic gene expression upon a particular molecular event (e.g. DNA binding). Isolating the selected circuit plasmid pool and retransforming it into a fresh batch of cells containing the selection module minimizes this problem. While the probability of false positives emerging is expected to remain constant with each round of selection, the functional circuit is exponentially enriched.

With careful monitoring of the frequency of emergence of false positives (by transforming nonfunctional control circuit plasmid(s) side-by-side with the target mutant library), this selection module will be able to isolate rare functional circuits from large libraries. For the inverter studied here, three rounds of selection should allow a functional circuit to be isolated from a pool with more than 3 million-fold background of nonfunctional ones; actual numbers for any experiment, however, will depend on the circuit and detailed selection protocol.

We expect that the dual selection module will be most useful in the initial stages of constructing complex circuits, particularly when only a

small fraction of a large pool of candidates may be functional. It is also possible to select for specific circuit performance criteria (ON and OFF gene expression levels, switching thresholds, etc) by adjusting antibiotic concentrations and expression levels of the selection markers. Qualitatively functional circuits may also be fine-tuned (perhaps more efficiently) by direct screening of smaller libraries (Yokobayashi et al., 2002). To implement this module, it is desirable to have control circuits that yield suitable ON and OFF outputs with the appropriate output promoter so that the selection conditions can be optimized (i.e. antibiotic concentrations, incubation time, etc), because the gene expression levels of ON and OFF states depend on the specific output promoter as well as translation efficiencies of the mRNAs.

This selection module can be used with a variety of host genotypes. The host bacterium must not contain the antibiotic resistance genes used in the selection module (i.e. *bla* and *tetA*). While *bla* is not replaceable because it works with *bli*, *tetA* can be replaced with other genes such as kanamycin resistance (kan^R, aminoglycoside 3′-phosphotransferase) or chloramphenicol acetyltransferase (*cat*).

An alternative implementation of the dual selection module that we considered was to express, instead of BLIP, a protein that would repress the expression of an antibiotic resistance gene (e.g. β-lactamase).This scheme has at least two disadvantages compared to direct inhibition of β-lactamase via protein–protein interaction with BLIP. (1) It would take longer for the selection module to reach steady state with the additional protein expression step. (2) Adding another promoter/operator site upstream of an antibiotic resistance gene adds another potential mutation site that could result in false positives.

Selection strategies for dynamic behaviors such as oscillation, and optimization of dynamic characteristics such as switching time await development. However, some of these circuits can be broken down into sub-modules or components (e.g. toggle switches, logic gates, amplifiers, filters) which can be optimized for steady-state activities using a variant of the selection strategy described here (Atkinson et al., 2003).

Acknowledgements

We thank Tim Palzkill for providing pTP154 from which *bli* gene was engineered and the National Science Foundation (Award # 0130613) for financial support.

References

Atkinson MR, Savageau MA, Myers JT and Ninfa AJ (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. Cell 113: 597–607.

Basu S, Karig D and Weiss R (2003) Engineering signal processing in cells: Towards molecular concentration band detection. Natural Computing 2: 463–478.

Buchler NE, Gerland U and Hwa T (2003) On schemes of combinatorial transcription logic. Proceedings of the National Academy of Sciences USA 100: 5136–5141.

Buskirk AR, Kehayova PD, Landrigan A and Liu DR (2003) In vivo evolution of an RNA-based transcriptional activator. Chemistry and Biology 10: 533–540.

Doran JL, Leskiw BK, Aippersbach S and Jensen SE (1990) Isolation and characterization of a beta-1actamase-inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. Journal of Bacteriology 172: 4909–4918.

Elledge SJ, Sugiono P, Guarente L and Davis RW (1989) Genetic selection for genes encoding sequence-specific DNA-binding proteins. Proceedings of the National Academy of Sciences USA 86: 3689–3693.

Fromknecht K, Vogel PD and Wise JG (2003) Combinatorial redesign of the DNA binding specificity of a prokaryotic helix-turn-helix repressor. Journal of Bacteriology 185: 475–481.

Gardner TS, Cantor CR and Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339–342.

Hasty J, Isaacs F, Dolnik M, McMillen D and Collins JJ (2001) Designer gene networks: Towards fundamental cellular control. Chaos 11: 207–220.

Hasty J, McMillen D and Collins JJ (2002) Engineered gene circuits. Nature 420: 224–230.

Huang L, Sera T and Schultz PG (1994) A permutational approach toward protein-DNA recognition. Proceedings of the National Academy of Sciences USA 91: 3969–3973.

Huang W, Zhang Z and Palzkill T (2000) Design of potent beta-lactamase inhibitors by phage display of beta-lactamase inhibitory protein. Journal of Biological Chemistry 275: 14964–14968.

Kaern M, Blake WJ and Collins JJ (2003) The engineering of gene regulatory networks. Annual Review of Biomedical Engineering 5: 179–206.

Kim BM and Oakley MG (2002) A general method for selection and screening of coiled coils on the basis of relative helix orientation. Journal of the American Chemical Society 124: 8237–8244.

Mossing MC, Bowie JU and Sauer RT (1991) A streptomycin selection for DNA-binding activity. Methods in Enzymology 208: 604–619.

Pelletier JN, Arndt KM, Pluckthun A and Michnick SW (1999) An *in vivo* library-versus-library selection of optimized protein-protein interactions. Nature Biotechnology 17: 683–690.

Sera T and Schultz PG (1996) In vivo selection of basic region-leucine zipper proteins with altered DNA-binding specificities. Proceedings of the National Academy of Sciences USA 93: 2920–2925.

Strynadka NC, Jensen SE, Alzari PM and James MN (1996) A potent new mode of beta-lactamase inhibition revealed by the 1.7 Å X-ray crystallographic structure of the TEM-1-BLIP complex. Nature Structural Biology 3: 290–297.

Wall ME, Hlavacek WS and Savageau MA (2004) Design of gene circuits: Lessons from bacteria. Nature Reviews Genetics 5: 34–42.

Weiss R, Homsy GE and Knight TF Jr (1999) Toward *in vivo* Digital Circuits. DIMACS Workshop on Evolution as Computation, Springer, New York

Yokobayashi Y, Collins CH, Leadbetter JR, Arnold FH and Weiss R (2003) Evolutionary design of genetic circuits and cell-cell communications. Advances in Complex Systems 6: 37–45.

Yokobayashi Y, Weiss R and Arnold FH (2002) Directed evolution of a genetic circuit. Proceedings of the National Academy of Sciences USA 99: 16587–16591.