# Modular gene-circuit design takes two steps forward

Jeffrey J Tabor

Work from two independent groups opens the door to designing gene circuits with advanced regulatory functions.



Engineered organisms enable studies of the general organizing principles of life and have the potential to transform industries including medicine, agriculture and energy. Organisms can be programmed by combining genetically encoded parts, such as transcription factors and promoters, into synthetic regulatory circuits. However, a limited set of orthogonal (non–cross-reacting) parts, an inability to predict how parts will perform when assembled, and a slow design cycle based on DNA assembly limit the capabilities and performance of engineered circuits relative to those of their evolved counterparts. Two studies published in this issue should allow the construction of more reliable gene circuits with advanced regulatory functions. Arkin and colleagues combine orthogonal RNA regulators of translation with a transcription antitermination system to construct transcriptional attenuators and activators that can be predictably assembled into higher-order circuits[1]. Collins and colleagues describe a genetic circuit 'breadboarding' methodology that accelerates optimization and simplifies the repurposing of components from existing circuits[2]. These approaches could be combined with automated gene-circuit design and assembly methods[3] to dramatically accelerate the synthetic biology design cycle.

Regulatory RNAs can activate or repress translation of specific genes through base-pairing interactions in the 5′ untranslated region (UTR) of the mRNA that affect ribosome binding. The programmability and specificity of base-pairing interactions has made RNA an appealing molecule for genetic circuit design since the early days of synthetic biology[4]. Nonetheless, the goal of constructing large genetic circuits from libraries of orthogonal RNA parts has remained elusive.

Recently, Arkin and colleagues computationally redesigned the interaction between a 5′ UTR named RNA-IN and its antisense inhibitor, RNA-OUT (**Fig. 1a**)[5]. Groups of up to six mutually orthogonal pairs were engineered that retained the translational regulatory properties of the original system. However, these parts lack two features that make regulated promoters so useful for circuit design: a large dynamic range (that is, the ratio of output gene expression between the active and repressed states) and the ability to be composed upstream of a single gene for multi-input regulation. To improve the performance of RNA regulators of translation initiation, the group turned to the transcription antitermination system from the *Escherichia coli* tryptophanase (*tna*) operon. Here, a short leader peptide (TnaC) is translated from the 5′ end of the nascent mRNA. In the presence of tryptophan, TnaC inhibits its own release from the ribosome, which results in stalling. Stalling obscures a nearby binding site for the transcription termination factor Rho, allowing transcription to continue.

To link translational to transcriptional regulation, the group replaced the constitutive *tnaC* 5′ UTR with six different RNA-IN sequences (**Fig. 1a**) and a small RNA-activated 5′ UTR[6]. The composition was surprisingly modular: in the presence of tryptophan, each of the hybrid parts converts binding of the small RNA to regulation of downstream transcription, with no apparent dependence on secondary structure of the 5′ UTR. Moreover, the engineered RNA-IN/OUT pairs retained their orthogonality, and their dynamic range was increased to levels comparable to that of many regulated promoters.

In digital circuit design, NOR gates convert high levels of either of two binary inputs into low levels of output. NOR gates are referred to as 'functionally complete' because they can be interconnected to implement all possible digital logic operations. Pairs of ligand-inducible promoters have previously been combined upstream of a repressor gene to build NOR gates in *E. coli*[7]. The gates were then expressed in different strains and linked with membrane-diffusible chemicals to implement all two-input logic operations[7]. The promoter gates could not be combined in one cell because of a lack of orthogonal transcription factors, a problem that could be overcome using the Arkin system. To build RNA-mediated transcriptional NOR gates, the group concatenated up to four orthogonal RNA-IN sequences, each fused to a *tna* element, in the 5′ UTR of a reporter gene[1]. In all cases, expression of any one or any combination of the cognate RNA-OUT molecules represses transcription of the reporter gene, resulting in N-input NOR logic.

Assuming that more mutually orthogonal RNAs can be designed, these NOR gates could be cascaded into circuits capable of next-generation cellular computations such as counting, set-reset latching and rewritable data storage[8]. Input RNAs could also be expressed under the control of myriad bacterial sensory systems, thereby allowing cells to decode complex environmental patterns such as those associated with specific diseased states, adjust metabolism in response to changing bioreactor conditions, or even function as smart probiotics that produce different beneficial compounds as needed. The small size of these RNA parts would produce a much smaller synthetic DNA footprint than would analogous protein circuits.

Synthetic biologists must often iterate through cycles of optimization when composing even well-understood parts[9]. To accelerate this process, Collins and colleagues[2] developed an iterative DNA

Jeffrey J. Tabor is in the Department of Bioengineering and Department of Biochemistry and Cell Biology at Rice University, Houston, Texas, USA.
e-mail: jeff.tabor@rice.edu

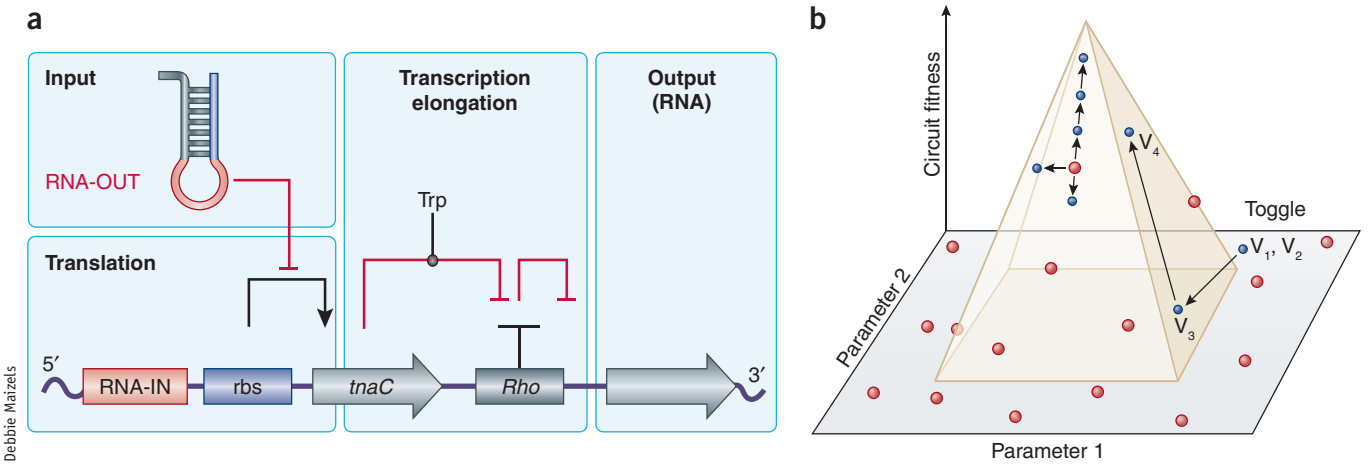


**Figure 1** | Modularity improves design in synthetic biology. (**a**) A system to convert regulation of translational initiation into regulation of transcriptional elongation[1]. (**b**) Plug-and-play part replacement (black arrows) can optimize the performance of circuits designed discretely (blue dots) or combinatorially (red dots).





assembly approach based on an *E. coli* vector with an array of 31 unique 6-base-pair sequences specifically recognized by standard restriction enzymes that work in common reaction conditions. A library of parts including promoters, transcription factors and fluorescent reporter genes lacking all 31 sites is then used for circuit design. Parts are targeted to specific positions in the array by PCR amplification with primers containing appropriate 5′ and 3′ restriction sites as overhangs. Parallel and hierarchical cycles of digestion and ligation are then used to speed construction.

The authors demonstrate breadboarding by optimizing a co-repressive toggle switch[10] design. Co-repressive toggles require balanced expression of two antagonistic transcription factors and are used to implement switch-like cell decisions in applications from sensing to differentiation. Here, the first design is based on the ligand-inhibited repressors LacI and TetR. Each is initially placed upstream of an associated fluorescent reporter on a polycistronic mRNA. The operons show poor reporter expression, which is then improved by 'plugging in' additional copies of the appropriate promoter upstream of each reporter. This increases reporter expression, but reveals that the circuit cannot reach the TetR-dominated state. The *tetR* promoter is then swapped for a stronger version, but this overcompensates for the problem making only the TetR state stable. A library of random *tetR* ribosome binding sites (RBSs) is then screened, and a variant that hits the bistable sweet spot is found (**Fig. 1b**). The plug-and-play DNA modifications of the breadboard allowed Toggle versions 2 and 3 to be constructed in 3 days each. This is not only faster than most methods, which require an unstandardized modification scheme based on primer synthesis, but also more scalable and amenable to future automation.

RBSs are famously nonmodular because their activity is sensitive to the presence of adjacent nucleotide sequences such as cloning sites and genes. However, a recent method allows the strength of an RBS to be tuned continuously by up to three orders of magnitude using a variable length spacer of nucleotide repeats between the Shine-Dalgarno region and start codon[11]. A spacer library could be constructed for a given RBS-gene fusion to allow tunable translation from a particular location within the breadboard. A second, nonvariable spacer could also be added to the 5′ end of each RBS[12] to buffer the impact of upstream cloning sites in the breadboard.

The breadboarding method also allows validated components to be reused in new circuits by plug-and-play modification rather than by DNA assembly. This is demonstrated by converting the final toggle design into a three- and then four-node coherent feed-forward loop[2] in only 5 days. Notably, both new circuits work without further optimization. Though genetic part composition can be unpredictable for many reasons, the ability to rewire validated modules within existing plasmids should reduce the amount of debugging in future designs. The value of breadboarding will therefore increase with time and with the number of users who adopt the method. Breadboarding and other methods of modular assembly[13] can also be used to combinatorially construct circuit variants from libraries of parts with different parameters. Characterization of partially randomized designs could generate coarse-grained maps of parameter sensitivity, and iterative part replacement could then be used to fine-tune circuit performance (**Fig. 1b**).

Not only will the Arkin and Collins methods improve biological design in the short term, they are also amenable to significant future development. As automated circuit design and assembly dovetails with iterative optimization, our ability to engineer circuits should extend beyond our ability to truly understand how they work. The tractability of modularly constructed synthetic circuits, however, should also feed back to accelerate the cycle of hypothesis generation and testing in systems biology.

**COMPETING FINANCIAL INTERESTS**
The author declares no competing financial interests.

1. Liu, C.C. *et al. Nat. Methods* **9**, 1087–1093 (2012).
2. Litcofsky, K.D., Afeyan, R.B., Krom, R.J., Khalil, A.S. & Collins, J.J. *Nat. Methods* **9**, 1077–1080 (2012).
3. Beal, J. *et al. ACS Synth. Biol.* **1**, 317–331 (2012).
4. Tabor, J.J. & Ellington, A.D. *Nat. Biotechnol.* **21**, 1013–1015 (2003).
5. Mutalik, V.K., Qi, L., Guimaraes, J.C., Lucks, J.B. & Arkin, A.P. *Nat. Chem. Biol.* **8**, 447–454 (2012).
6. Isaacs, F.J. *et al. Nat. Biotechnol.* **22**, 841–847 (2004).
7. Tamsir, A., Tabor, J.J. & Voigt, C.A. *Nature* **469**, 212–215 (2011).
8. Bonnet, J., Subsoontorn, P. & Endy, D. *Proc. Natl. Acad. Sci. USA* **109**, 8884–8889 (2012).
9. Slusarczyk, A.L., Lin, A. & Weiss, R. *Nat. Rev. Genet.* **13**, 406–420 (2012).
10. Gardner, T.S., Cantor, C.R. & Collins, J.J. *Nature* **403**, 339–342 (2000).
11. Egbert, R.G. & Klavins, E. *Proc. Natl. Acad. Sci. USA* published online, doi:10.1073/pnas.1205693109 (27 August 2012).
12. Temme, K., Zhao, D. & Voigt, C.A. *Proc. Natl. Acad. Sci. USA* **109**, 7085–7090 (2012).
13. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. *PLoS ONE* **6**, e16765 (2011).