Insulating gene circuits from context by RNA processing

Caleb J Bashor & James J Collins

Two studies find that programmable RNA-processing tools counter the problem of context-dependence in the construction of synthetic biology circuits.

Engineered biological circuits rarely work as designed. In most cases, the performance of their molecular parts is highly dependent on cellular and sequence context and varies greatly from one system to the next[1–3]. Synthetic biology urgently requires strategies to limit such context-dependence[4], and two new reports in *Nature Biotechnology* describe progress toward this goal. Lou *et al.*[5] and Qi *et al.*[6] both use programmable RNA processing to improve the interoperability of coding sequences and their linked bacterial promoters. The studies demonstrate that broadly applicable engineering tools can be developed to simplify the construction of biological circuits.

Synthetic biologists aim to build programmable cells through the design and control of complex cellular functions. At present, the field still lacks essential features common to many engineering disciplines, such as assembly standards, broad tools for computer-aided design, and platforms for high-throughput fabrication and rapid prototyping of devices. Building gene circuits relies on slow, trial-and-error tinkering to create one-off prototype circuits. To advance beyond this artisanal stage, practitioners are attempting to develop modular, well-characterized molecular parts[1,2], which, when connected according to design specifications, would allow rapid assembly of circuits with desired functions.

One of the key hurdles in achieving this is that circuit parts show a context-dependence that derives from the chemical and physical

Caleb J. Bashor and James J. Collins are at the Howard Hughes Medical Institute, Department of Biomedical Engineering and Center for BioDynamics, Boston University, Boston, Massachusetts, USA.
e-mail: jcollins@bu.edu

constraints of the cellular environment. The interior of a cell behaves as a moderately well-stirred solution in which diffusing circuit components freely associate with constituents of the host cell, resulting in unforeseen interactions that can perturb circuit performance and/or host-cell function. Context effects can also arise from the physical arrangement of parts. For example, the order in which parts are positioned in a plasmid (or synthetic genome) may create unanticipated regulatory interactions between adjacent secondary structural elements. Such complex relationships between part context and function are difficult to predict based on biophysical principles.

For these reasons, workflows for constructing synthetic circuits typically get bogged down in iterative cycles of troubleshooting[4]. Labor-intensive diversification and screening approaches are often needed to select parts that behave well in a specific context. Although many useful parts and design rules have emerged[7], the circuit-design process remains suboptimal[8], and there is a pressing need for rational, generalizable strategies to minimize context-dependence.

The studies by Lou *et al.*[5] and Qi *et al.*[6] tackle one aspect of this problem, namely, context-dependent variation in the transcriptional strength of promoters (**Fig. 1**). Although their experimental approaches differ, both groups provide solutions in which a transcript's coding region is physically separated from surrounding sequences by engineered RNA processing. While constructing a simple NOT gate in *Escherichia coli*, Lou *et al.*[5] noted that the quantitative relationship between the activity of a transgene's input promoter and transgene expression—known as the 'transfer function'—was altered unexpectedly when the input promoter was changed. As the authors deduced, transfer-function

variability was a by-product of the sequence linking the input promoter and the coding sequence (in this case, the coding sequence of the cI repressor). Specifically, 5′ untranslated region (UTR) leader sequences that arise through transcription of sequences present downstream of various input promoters were interacting with the proximal region of the *cI* mRNA and reducing transcript amounts to different extents.

To buffer against transfer-function variability, the authors searched for 'insulator' elements. They screened a small library of 5′ UTRs chosen from the literature based on the elements' reported ability to increase transcript numbers. One of these elements, the ribozyme RiboJ, proved to be a highly effective buffer. RiboJ autocatalytically cleaves upstream sequences. Placing RiboJ between a promoter and the cI coding sequence resulted in cleavage between the 5′ UTR (which was transcribed downstream from the promoter) and the coding sequence, and thereby eliminated the promoter-dependent variability of the circuit's transfer function. The authors went on to show that the method was generalizable: RiboJ could buffer promoters in multiple genetic circuits, and several other 5′-cleaving ribozymes functioned as effective insulators. Given a sufficiently large assortment of riboregulators to choose from, this approach could reduce promoter-context effects in circuits containing multiple promoters.

In a similar RNA-processing strategy developed by Qi *et al.*[6], coding regions are cleaved from their 5′ UTRs using the clustered regularly interspaced short palindromic repeat (CRISPR) system. Processing is accomplished in *trans*, by expression of a *Pseudomonas*-derived, CRISPR-associated endo-RNase that specifically cleaves a 28-nucleotide cleavage site that is cloned upstream of the ribosome
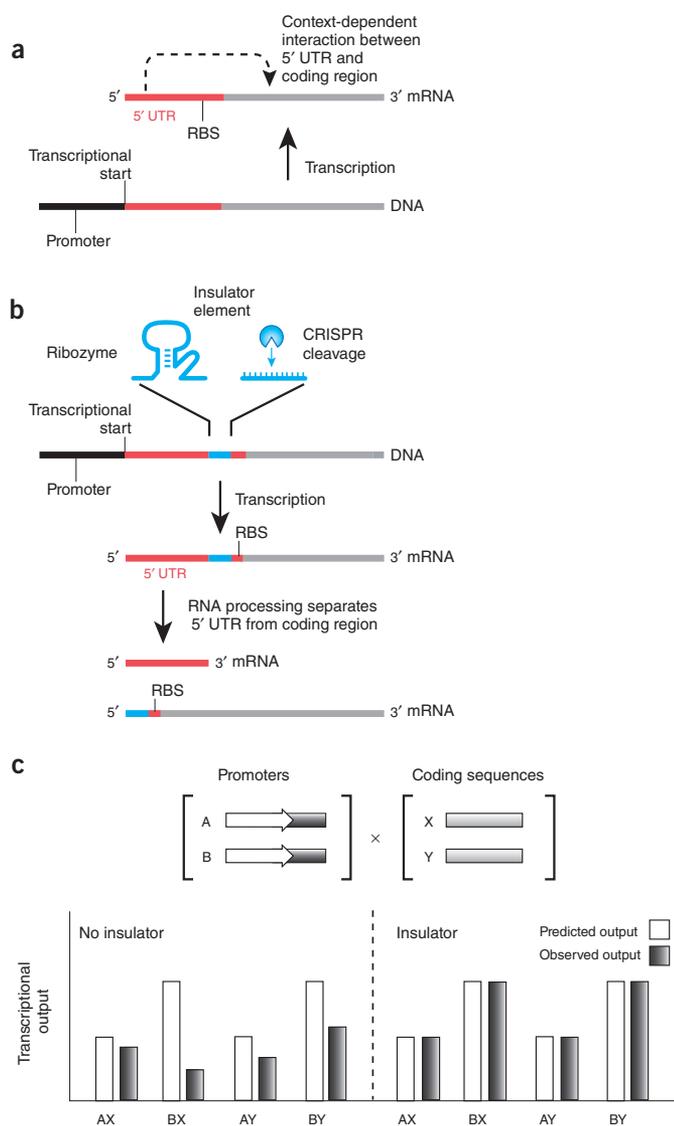
Figure 1 Synthetic RNA processing buffers against context-dependent variability of circuit output. (**a**) During construction of a synthetic circuit, unexpected interactions between the 5′ UTR (present in a transcript) and the coding region can alter gene expression in unpredictable ways. (**b**) RNA processing elements, such as ribozymes and CRISPR target sequences, inserted between the 5′ UTR and the ribosome binding site (RBS) act as insulators against context effects because the 5′ UTRs are cleaved from adjoining coding regions after transcription. RBS, ribosome binding site. (**c**) Insulator elements program more predictable promoter behavior. For example, for combinations of hypothetical promoters and coding regions, the inclusion of insulator elements can mitigate contextual effects on transcriptional output.

of complex synthetic regulatory systems containing many promoters.

Both Lou *et al.*[5] and Qi *et al.*[6], using the common strategy of programmed RNA processing, demonstrate effective buffering against promoter context-dependence. The key to their success—the physical separation of the transcript's 5′ UTR from its coding sequence—can be thought of as a further layer of design modularity that disengages the regulatory function of the promoter (and associated 5′ UTR) from the ribosome binding site property of the transcript. The studies are also notable in that they harness a well-studied cellular process that had not been previously explored for synthetic biology applications. One outstanding question raised by both groups' work is the scalability of their approaches to metazoan organisms. The role of promoter context-dependence in mammalian cell circuit engineering is not well characterized, and how synthetically imposed RNA processing might interact with complex native RNA processing is unclear.

To maintain forward momentum, synthetic biologists must seek out untapped cellular processes and information currencies[9,10]. New engineering strategies are needed to meet challenges such as creating faster-timescale circuit dynamics, achieving spatial regulation in three dimensions using post-translational modifications of signaling proteins, and programming cells with more sophisticated computational and information-storage capabilities. An investigation into how diverse regulatory processes such as phosphorylation, protein degradation and chromatin modification can be incorporated into circuit design will enlarge the pool of available parts for circuit engineering and enable a broader search for additional solutions to issues of part interoperability and context-dependence.

1. Lucks, J.B., Qi, L., Whitaker, W.R. & Arkin, A.P. *Curr. Opin. Microbiol.* **11**, 567–573 (2008).
2. Khalil, A., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, K. & Collins, J.J. *Cell* **150**, 647–658 (2012).
3. Cardinale, S. & Arkin, A.P. *Biotechnol. J.* **7**, 856–866 (2012).
4. Lu, T.K., Khalil, A.S. & Collins, J.J. *Nat. Biotechnol.* **27**, 1139–1150 (2009).
5. Lou, C., Stanton, B., Chen, Y.J., Munsky, B. & Voigt, C.A. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
6. Qi, L., Haruwitz, R.E., Shao, W., Doudna, J.A. & Arkin, A.P. *Nat. Biotechnol.* **30**, 1002–1006 (2012).
7. Ellis, T., Wang, X. & Collins, J.J. *Nat. Biotechnol.* **27**, 465–471 (2009).
8. Clancy, K. & Voigt, C.A. *Curr. Opin. Biotechnol.* **4**, 572–581 (2010).
9. Bashor, C.J., Horwitz, A.A., Peisajovich, S.G. & Lim, W.A. *Annu. Rev. Biophys.* **39**, 515–537 (2010).
10. Wieland, M. & Fussenegger, M. *Annu. Rev. Chem. Biomol. Eng.* **3**, 209–234 (2012).

binding site and coding sequence. Cleavage of the precursor mRNA removed 5′ UTR leader sequences and thereby reduced the context-dependence of reporter gene expression.

The studies of Qi *et al.*[6] in different genetic contexts suggest that this solution is generalizable. In one experiment, variability in gene expression for a library of randomized 5′ UTRs appended to a promoter was much reduced through addition of the CRISPR target sequence and cleavage of the transcript. In another experiment, a set of synthetic genes was combinatorially assembled from constituent elements (promoter, ribosome binding site and reporter gene). Reporter gene expression for different combinations varied widely, but introducing the CRISPR target sequence yielded uniform gene expression. The authors also applied CRISPR processing to a two-gene synthetic operon and eliminated the effects of cistron position on transcriptional output. Finally, they showed that their method is portable by using it to modulate transcriptional outputs in various synthetic regulatory circuits, even demonstrating that it works in *Bacillus subtilis* and yeast. As the approach relies on a single enzyme to cleave potentially many sites, it offers a way to achieve predictable function