Synthetic Gene Networks That Count

Ari E. Friedland, *et al.*

Science **324**, 1199 (2009);
DOI: 10.1126/science.1172005

Synthetic Gene Networks That Count

Ari E. Friedland,[1*] Timothy K. Lu,[1,2*] Xiao Wang,[1] David Shi,[1]
George Church,[2,3] James J. Collins[1]†

Synthetic gene networks can be constructed to emulate digital circuits and devices, giving one the ability to program and design cells with some of the principles of modern computing, such as counting. A cellular counter would enable complex synthetic programming and a variety of biotechnology applications. Here, we report two complementary synthetic genetic counters in *Escherichia coli* that can count up to three induction events: the first, a riboregulated transcriptional cascade, and the second, a recombinase-based cascade of memory units. These modular devices permit counting of varied user-defined inputs over a range of frequencies and can be expanded to count higher numbers.

A counter is a key component in digital circuits and computing that retains memory of events or objects, representing each number of such as a distinct state. Counters would also be useful in cells, which often must have accurate accounting of tightly controlled processes or biomolecules to effectively maintain metabolism and growth. Counting mechanisms have been reportedly found in telomere length regulation (*1*, *2*) and cell aggregation (*3*). These system behaviors appear to be the result of a threshold effect in which some critical molecule number or density must be reached for the observed phenotypic change.

In this study, we first developed a counter, termed the riboregulated transcriptional cascade (RTC) counter, which is based on a transcriptional cascade with additional translational regulation. Two such cascades are illustrated in Fig. 1, A and C that can count up to two and three, respectively (hence, the designations RTC two-counter and RTC three-counter). For the RTC two-counter, the constitutive promoter $P_{Ltet0\text{-}1}$ drives transcription of T7 RNA polymerase (RNAP), whose protein binds the T7 promoter and transcribes the downstream gene, in this case, green fluorescent protein (GFP). Both genes are additionally regulated by riboregulators (*4*), whose cis and trans elements silence and activate posttranscriptional gene expression, respectively. The cis-repressor sequence [cr in Fig. 1] is placed between the transcription start site and the ribosome-binding site (RBS), and its complementarity with the RBS causes a stem-loop structure to form upon transcription. This secondary structure prevents binding of the 30*S* ribosomal subunit to the RBS, which inhibits translation. A short, transactivating, noncoding RNA (taRNA), driven by the arabinose promoter P_{BAD}, binds to the cis repressor in trans, which relieves RBS repression and allows translation. With this riboregulation, each node (i.e., gene) in the cascade requires both independent transcription and translation for protein expression. This cascade is able to count brief arabinose pulses [for pulse definition, see (*5*)] by expressing a different protein in response to each pulse (Fig. 1A). With cis-repressed T7 RNAP

[1]Howard Hughes Medical Institute, Department of Biomedical Engineering, Center for BioDynamics and Center for Advanced Biotechnology, Boston University, Boston, MA 02215, USA. [2]Harvard-MIT Division of Health Sciences and Technology, 77 Massachusetts Avenue, Room E25-519, Cambridge, MA 02139, USA. [3]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: jcollins@bu.edu

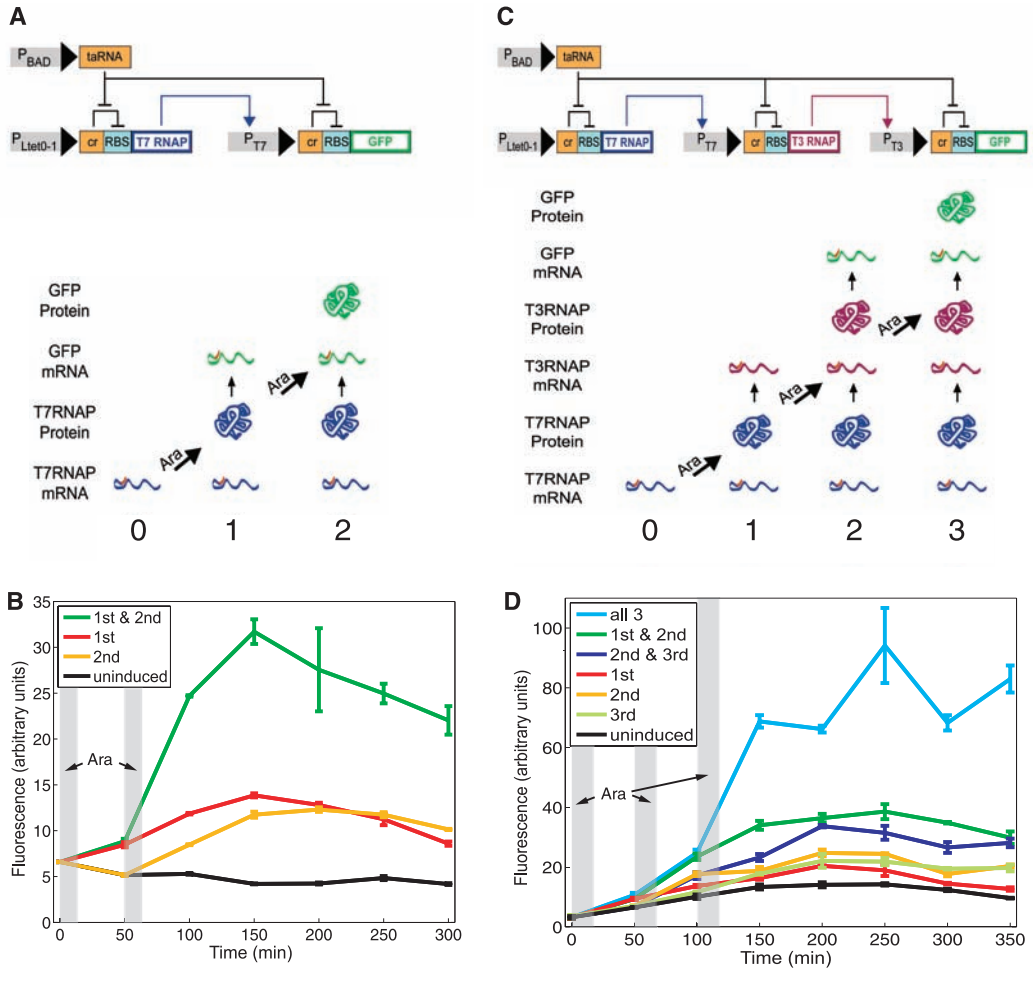


Fig. 1. The RTC two-counter and RTC three-counter construct designs and results. (**A**) The RTC two-counter is a transcriptional cascade with two nodes. Shown at the bottom are expected expression profiles after zero, one, and two arabinose (Ara) pulses. (**B**) Mean fluorescence of three replicates of RTC two-counter cell populations over time, measured by a flow cytometer. Shaded areas represent arabinose pulse duration. (**C**) The RTC three-counter is a transcriptional cascade with three nodes. Shown at the bottom are expected expression profiles after zero, one, two, and three arabinose pulses. (**D**) Mean fluorescence of three replicates of RTC three-counter cell populations over time, measured by a flow cytometer. Shaded areas represent arabinose pulse duration.

mRNAs in the cell, the first pulse of arabinose drives a short burst of taRNA production and, consequently, expression of T7 RNAP proteins. After the pulse is delivered, arabinose is removed from the cell environment, intracellular arabinose and taRNA are metabolized, and expression of T7 RNAP protein halts. The T7 RNAP proteins that have been translated then transcribe cis-repressed GFP transcripts, but few GFP proteins are made until the next arabinose pulse is delivered and translation is once again activated.

We built the RTC two-counter construct on a high-copy plasmid and transformed it into *Escherichia coli* strain K-12pro (*5*). Cells containing this construct were pulsed with the inducer arabinose, and fluorescence was measured over time (Fig. 1B). Uninduced cells show no increase in mean fluorescence, whereas cells that received either the first or second pulse show only small increases, indicating some degree of leakage—an effect in which the intended protein is expressed in each arabinose pulse, but also some unintended, downstream proteins are expressed as well. Cells that received both arabinose pulses show a substantial increase in fluorescence when the second pulse is delivered, precisely when the cells are expected to express GFP proteins.

To extend the RTC counter's capability to count to three, we built a second synthetic construct, the RTC three-counter, again with GFP as the quantitative readout. It is similar to the RTC two-counter but has three nodes in the cascade instead of two (Fig. 1C). T7 RNAP is the gene at the first node driving transcription of T3 RNAP, which ultimately drives transcription of GFP. All transcripts are likewise cis-repressed with the same riboregulator sequence. When pulsed with arabinose, this counter should primarily produce T7 RNAP proteins during the first pulse, T3 RNAP proteins during the second pulse, and GFP proteins during the third pulse (Fig. 1C).

Experimental results demonstrate that fluorescence increases substantially only when all three arabinose pulses are delivered (Fig. 1D). Flow cytometry measurements show this increase beginning at precisely the time of the third pulse, and the considerable slope at this juncture suggests that cells contain a high concentration of cis-repressed GFP transcripts ready for transactivation. The data also reveal slight leakage in cells that are pulsed only once or twice, but their fluorescence remains comparatively low. This result, in combination with the RTC two-counter evidence, shows that the temporal progression of RNA and protein species logically predicted by the counter network architecture is indeed responsible for the observed effect.

To further support these results, we constructed and analyzed a mathematical model based on the design of the RTC two-counter and three-counter constructs. This model, with fitted parameters [see section 6 of (*5*)], was able to match both the RTC two-counter and three-counter experimental results (Fig. 2, A and B). We used the model to investigate the effects of pulse frequency and pulse length on the performance of the RTC three-counter and to guide our experimental search for optimal combinations. The mathematical model predictions, shown as contour lines in Fig. 2C, indicate that maximum expression occurs with pulse lengths of ~20 to 30 min and pulse intervals of 10 to 40 min. The absolute difference in fluorescence after three pulses and two pulses is shown in Fig. 2D, with optimal counting behavior requiring similar pulse length and interval combinations noted above.

Experimentally, we sampled various pulse lengths and intervals, plotting these results as circles in Fig. 2, C and D. These results are consistent with the model predictions across a wide range of temporal conditions, and they confirm that the RTC three-counter has a sizable temporal region in which its counting behavior is robust. Within this region, the counter is also capable of counting irregular pulses; for example, it is able to distinguish between two short pulses followed by a long pulse and two long pulses, as predicted by the model (fig. S5). However, as indicated in Fig. 2, when pulse length or frequency is either too high or low, the RTC three-counter is unable to count properly, presumably because of the

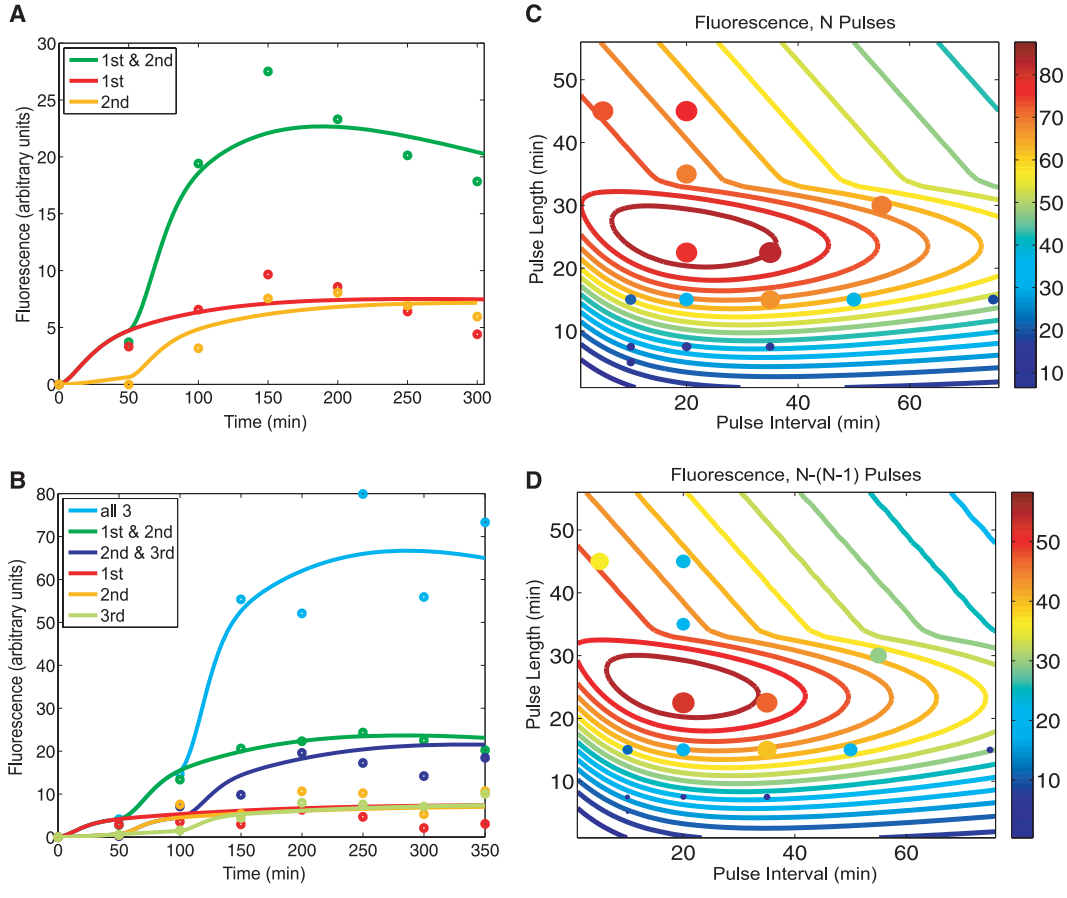


Fig. 2. Modeling predictions and RTC three-counter experimental characterization. (**A**) A model with fitted parameters captures the salient features of the normalized fluorescence results of the RTC two-counter. (**B**) An expanded model, again with fitted parameters, matches the normalized fluorescence results of the RTC three-counter. (**C**) On the basis of parameters fitted in (B), the model predicts expression output of the RTC three-counter (*N*) across a range of pulse lengths and intervals, and these calculations were used to generate the colored contour lines. Solid circles represent experimental results, with both color and size of circles indicating the level of expression. (**D**) Similar to (C), except that values shown are the difference in expression output after three (N) pulses and two (*N* – 1) pulses.

intrinsic kinetic limits of the biochemical processes involved, such as transcription and mRNA degradation.

Our second counter design, termed the DNA invertase cascade (DIC) counter, was built by chaining together modular DNA-based counting units (Fig. 3A). The DIC counter uses recombinases, such as Cre and flp_e (*6*), which can invert DNA between two oppositely oriented cognate recognition sites, such as *loxP* and flp_e-recombination target (*FRT*) sites, respectively. Recombinases have been used for numerous applications, including the creation of gene knockouts and inducible expression systems (*7*, *8*). In our counter design, each recombinase gene (*rec*) is downstream of an inverted promoter (P_{inv}), fused to an *ssrA* tag that causes rapid protein degradation (*9*), and followed by a transcriptional terminator (Term) (fig. S7). The P_{inv}-*rec*-*ssrA*-Term DNA sequences are placed between forward and reverse recombinase recognition sites (R_f and R_r) (fig. S7), forming a single counting unit that we have named a single invertase memory module (SIMM) (Fig. 3A and fig. S7). Upon expression of recombinase by an upstream promoter, the entire SIMM is inverted between the recognition sites. Because the recombinase gene is inverted with respect to the upstream promoter, further expression of recombinase protein ceases, and DNA orientation is fixed.

We developed a single-inducer DIC two-counter (fig. S8) and three-counter (Fig. 3A and fig. S9), which are composed of one and two SIMMs, respectively, and placed them on pBAC plasmids that are maintained as single-copy episomes (*10*). These circuits utilize P_{BAD}, so that pulses of arabinose constitute inputs to the circuit. Each pulse of arabinose results in promoter activation and expression of the next recombinase in the cascade, which then inverts the SIMM in which it is located. This allows the inverted promoter contained within that SIMM to be placed in a forward orientation to drive expression of the next SIMM stage. The single-inducer DIC two-counter shows high GFP output after two pulses of arabinose but only low GFP output after one pulse of arabinose, which shows that a single SIMM can be inverted to count events (fig. S11). In the single-inducer DIC three-counter, some premature flipping of the Cre recombinase–based SIMM did occur, which resulted in a small amount of leakage, e.g., fluorescence increased after only two arabinose pulses (Fig. 3B and fig. S12). However, this leakage was small compared with the high GFP output exhibited in response to three pulses of arabinose (Fig. 3B). In order to probe the temporal characteristics of the single-inducer DIC three-counter, we varied the pulse lengths and intervals, calculating the ratio of GFP output for cells exposed to three, versus two, pulses of arabinose (Fig. 3C). This ratio was at least 1.5 for most conditions tested, which demonstrated that the single-inducer DIC three-counter is able to successfully count pulses whose lengths and intervals range from 2 to 12 hours (Fig. 3C).

We also developed a multiple-inducer DIC three-counter by replacing the P_{BAD} promoters in the single-inducer DIC three-counter with the inducible promoters $P_{Ltet0-1}$, P_{BAD}, and P_{A1lacO} (Fig. 4A and fig. S10). These promoters respond to anhydrotetracycline (aTc), arabinose, and isopropyl β-D-1-thiogalactopyranoside (IPTG), respectively (Fig. 4A). When exposed to aTc followed by arabinose, followed by IPTG, the multiple-inducer DIC three-counter produced a high GFP output (Fig. 4B). No other permutations of the three inducers produced a high output, although some did exhibit a small amount of leakage (Fig. 4, C and D). These results demonstrate that the circuit can be programmed to record only a desired sequence of events.

We have constructed and validated two complementary designs for synthetic counters that operate across a range of time scales. These counters are both highly modular and capable of functioning with multiple inducer-promoter pairs. In addition, the architectures of both counters allow for the tunable output expression of different protein species of interest at any number (up to three are shown) in the counting

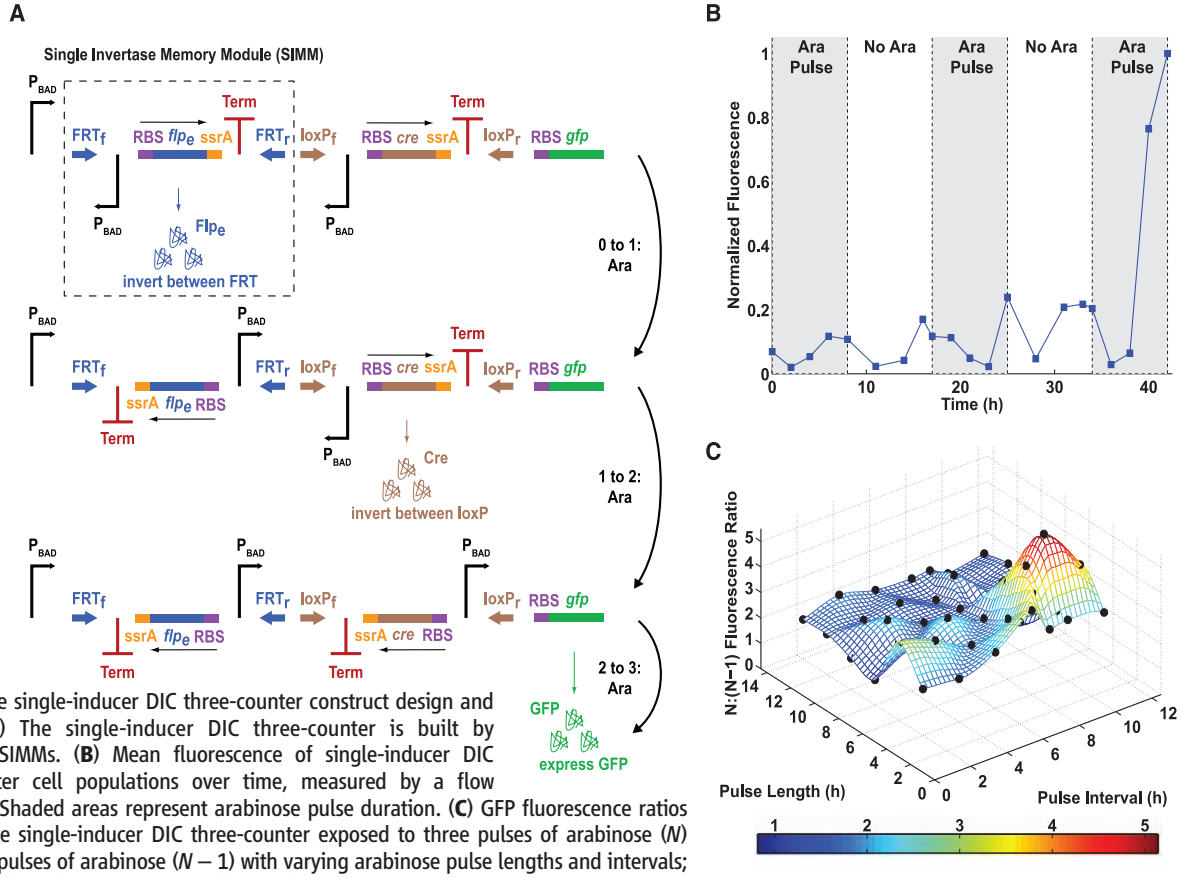


Fig. 3. The single-inducer DIC three-counter construct design and results. (**A**) The single-inducer DIC three-counter is built by cascading SIMMs. (**B**) Mean fluorescence of single-inducer DIC three-counter cell populations over time, measured by a flow cytometer. Shaded areas represent arabinose pulse duration. (**C**) GFP fluorescence ratios between the single-inducer DIC three-counter exposed to three pulses of arabinose (*N*) versus two pulses of arabinose (*N* − 1) with varying arabinose pulse lengths and intervals; experimental results are represented by black dots.

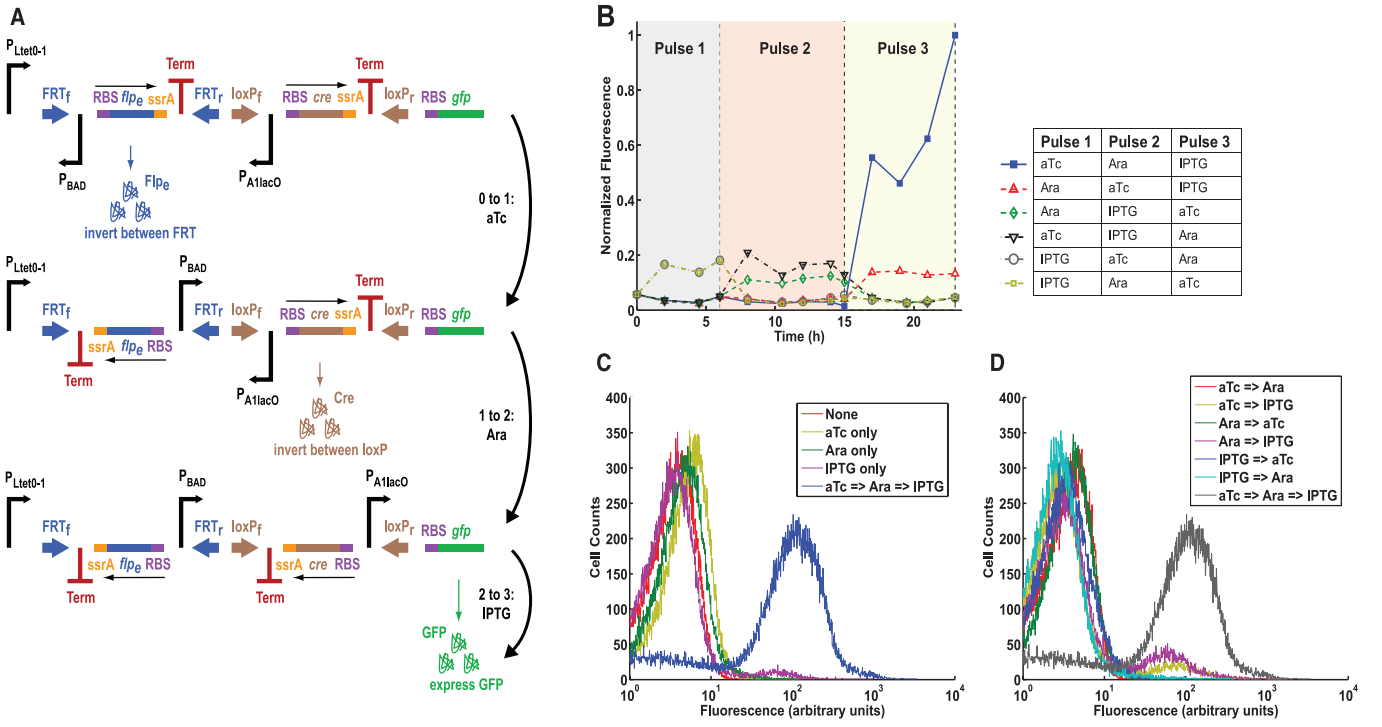


Fig. 4. The multiple-inducer DIC three-counter construct design and results. (**A**) The multiple-inducer DIC three-counter is similar to the single-inducer DIC three-counter in Fig. 3, except that each promoter is a unique inducible promoter: $P_{Ltet0\text{-}1}$, P_{BAD}, and P_{A1lacO}. These promoters respond to aTc, arabinose, and IPTG, respectively. (**B**) Mean fluorescence of multiple-inducer DIC three-counter cell populations over time, measured by a flow cytometer. Colored areas represent the durations of consecutive inducer pulses. (**C**) Flow cytometry population data showing the multiple-inducer DIC three-counter when exposed to the desired sequence of three inducers and to single inducers only. (**D**) Flow cytometry population data showing the multiple-inducer DIC three-counter when exposed to the desired sequence of three inducers and to all pairwise permutations of inducers.

process. Our constructs were built to count up to three events, but they both should be extensible with the use of other unique polymerases or recombinases, of which many are known (*5*). In addition to these shared qualities, each counter comes with its own set of properties. Our RTC counters demonstrate fast activation because of transcriptional and translational regulatory elements, which makes them useful for counting cellular events on the time scale of cell division. The DIC counters operate on time scales of hours (fig. S13) as a result of DNA recombination dynamics (*11*), and they are built with a novel SIMM design that retains counter state based on DNA orientation.

Synthetic gene circuits have enlarged the molecular tool set available to bioengineers and molecular biologists (*4*, *12*–*24*) and have enabled them to program novel cellular behaviors (*25*–*27*) and to construct therapeutic agents (*28*, *29*). Our synthetic counters represent complementary designs that can be used in different settings for a variety of purposes across a range of time scales. For example, if inputs to our RTC counter were coupled to the cell cycle, one might program cell death to occur after a user-defined number of cell divisions as a safety mechanism in engineered strains used for biosensing, bioremediation, or medical purposes. In addition, the multiple-inducer DIC counter might be used to study sequential events that occur in settings such as developmental biology and gene cascades; the single-inducer DIC counter could record events encountered in its environment (e.g., for biosensing); and our SIMM design could be used in other synthetic circuits to maintain genetic memory of low-frequency events, for therapeutic or other applications, such as studying neural circuits.

References and Notes

1. S. Marcand, E. Gilson, D. Shore, *Science* **275**, 986 (1997).
2. A. Ray, K. W. Runge, *Mol. Cell. Biol.* **19**, 31 (1999).
3. D. A. Brock, R. H. Gomer, *Genes Dev.* **13**, 1960 (1999).
4. F. J. Isaacs *et al.*, *Nat. Biotechnol.* **22**, 841 (2004).
5. Materials and methods are available as supporting material on *Science* Online.
6. F. Buchholz, P. O. Angrand, A. F. Stewart, *Nat. Biotechnol.* **16**, 657 (1998).
7. A. C. Groth, M. P. Calos, *J. Mol. Biol.* **335**, 667 (2004).
8. T. S. Ham, S. K. Lee, J. D. Keasling, A. P. Arkin, *Biotechnol. Bioeng.* **94**, 1 (2006).
9. J. B. Andersen *et al.*, *Appl. Environ. Microbiol.* **64**, 2240 (1998).
10. D. A. Wright *et al.*, *Nat. Protocols* **1**, 1637 (2006).
11. S. W. Santoro, P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4185 (2002).
12. M. B. Elowitz, S. Leibler, *Nature* **403**, 335 (2000).
13. T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **403**, 339 (2000).
14. T. S. Bayer, C. D. Smolke, *Nat. Biotechnol.* **23**, 337 (2005).
15. D. Baker *et al.*, *Sci. Am.* **294**, 44 (2006).
16. O. Rackham, J. W. Chin, *Nat. Chem. Biol.* **1**, 159 (2005).
17. K. Rinaudo *et al.*, *Nat. Biotechnol.* **25**, 795 (2007).
18. Y. Yokobayashi, R. Weiss, F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16587 (2002).
19. N. J. Guido *et al.*, *Nature* **439**, 856 (2006).
20. J. M. Pedraza, A. van Oudenaarden, *Science* **307**, 1965 (2005).
21. T. L. Deans, C. R. Cantor, J. J. Collins, *Cell* **130**, 363 (2007).
22. E. Fung *et al.*, *Nature* **435**, 118 (2005).
23. B. P. Kramer *et al.*, *Nat. Biotechnol.* **22**, 867 (2004).
24. J. Stricker *et al.*, *Nature* **456**, 516 (2008).
25. S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, *Nature* **434**, 1130 (2005).
26. H. Kobayashi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8414 (2004).
27. A. Levskaya *et al.*, *Nature* **438**, 441 (2005).
28. T. K. Lu, J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11197 (2007).
29. T. K. Lu, J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4629 (2009).

30. Supported by the NIH Director's Pioneer Award Program, the NSF Frontiers in Integrative Biological Research program, and the Howard Hughes Medical Institute. Authors' contributions are in supporting online materials.

Supporting Online Material

www.sciencemag.org/cgi/content/full/324/5931/1199/DC1
Materials and Methods
SOM Text
Figs. S1 to S13
Tables S1 and S2
References

9 February 2009; accepted 31 March 2009
10.1126/science.1172005

Supporting Online Material for

Synthetic Gene Networks That Count

Ari E. Friedland, Timothy K. Lu, Xiao Wang, David Shi, George Church, James J. Collins*

*To whom correspondence should be addressed. E-mail: jcollins@bu.edu

Published 29 May 2009, *Science* **324**, 1199 (2009)
DOI: 10.1126/science.1172005

This PDF file includes

SUPPLEMENTARY ONLINE MATERIALS

Authors' Contributions

AEF and JJC designed the RTC counters; AEF built and tested the RTC counters, with assistance from DS; XW conducted all of the modeling work; TKL, JJC and GC designed the DIC counters; TKL built and tested the DIC counters and contributed to building the RTC counters; AEF, TKL, XW, DS, GC and JJC analyzed the data and wrote the paper.

Materials and Methods

1. RTC Counter Plasmid Construction

RTC Counter plasmids were constructed using basic molecular cloning techniques (*S1*). New England Biolab's restriction endonucleases, T4 DNA Ligase, and Taq Polymerase were used as well as Invitrogen's PCR SuperMix High Fidelity. PCRs were carried out with an MJ Research PTC-200 Peltier Thermal Cycler. Synthetic oligonucleotides were made by Integrated DNA Technologies. For cloning, plasmids were transformed into *E. coli* strain DH5α (F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1

*hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-) with a standard heat shock protocol (*S1*), and isolated with Qiagen QIAprep Spin Miniprep Kits. Plasmid modifications were confirmed by restriction digests and sequencing by Agencourt.

2. RTC Counter Plasmid Design

Two plasmids – the RTC 2-Counter (Fig. S1) and RTC 3-Counter (Fig. S2) – were made, both derived from the riboregulator vector pZER21Y12α12G reported by Isaacs *et al.* (*4*), itself based strongly on the Lutz and Bujard pZE21 expression vector (*S2*). These contain kanamycin resistance, ColE1 origin of replication, the P_{BAD} promoter driving transcription of taRNA version taR12 (*4*), and the $P_{Ltet0\text{-}1}$ promoter. Both constructs were modified to have the $P_{Ltet0\text{-}1}$ promoter driving transcription of T7 RNA polymerase (NCBI Accession NC_001604.1). For the RTC 2-Counter construct, there is also the T7 promoter (TAATACGACTCACTATAGGGAGA) driving transcription of GFPmut3b; for the RTC 3-Counter construct, the T7 promoter drives transcription of T3 RNA polymerase (NCBI Accession NC_003298.1). The RTC 3-Counter additionally contains the T3 promoter 14.3m (ATTAACCCTCACTAAAGGGAGA) (*S3*), which drives transcription of GFPmut3b. All genes used in these constructs were engineered with the crR12 *cis*-repressor sequence upstream of the RBS (*4*). All promoters were paired with appropriate transcription terminators: P_{BAD} with the *E. coli* rrnB terminator, $P_{Ltet0\text{-}1}$ with the *E. coli* terminator T1 (of the rrnB terminator), P_{T7} with T7 transcription terminator Tphi, and P_{T3} with T3 transcription terminator Tphi.

3. RTC Counter Experimental Conditions

All experiments were conducted with the *E. coli* K-12pro strain (F^+, P_{N25}/tetR, P_{laciq}/lacI, Sp^r). For both the RTC 2-Counter and 3-Counter experiments, cells containing the counting vector were grown overnight in a Luria-Bertani (DIFCO) medium containing 30 μg/mL kanamycin, then diluted 1:100 and grown between 5 and 6.5 hours to an OD between 1.1 and 1.6 before being aliquoted into clear-bottom 24-well assay plates, 1mL per well. For the 2-Counter data shown in Figs. 1B and S3A, cells pulsed with arabinose had arabinose added to their wells for a final concentration of 0.001% at 0 minutes (immediately following the aliquot) and/or at 50 minutes. Pulses were left in the media for 10 minutes before cells were transferred into 1.5 mL tubes and spun for 1 minute at 8,000 rpm. Media was aspirated out of these tubes, and cells were resuspended in fresh media and transferred back to the plate. The 3-Counter experiments had 0.01% (final concentration) arabinose pulses delivered at varying times. All cells in 24-well plates were maintained at 37^{o}C throughout the course of the experiments, with shaking in between measurements. Experiments were performed in triplicate, and all data points shown are the mean values of these replicates.

4. RTC Counter Flow Cytometer Measurements

Data for Figs. 1B and 1D were collected with a Becton Dickinson FACSCalibur flow cytometer. Fluorescence was calibrated with Calibrite Beads (Becton Dickinson) and measured with a 488-nm argon laser excitation and a 515-nm to 545-nm emission filter. At each time point, 8 μL of cells were taken from the plate wells and diluted into 1 mL of filtered PBS, pH 7.2. Mean fluorescence measurements were calculated by BD

Biosciences' Cellquest Pro software, from samples containing at least 100,000 cells. No filters or gates were used on the cell populations.

5. RTC Counter Spectrophotometer Measurements

Data for Fig. S3 was collected with a Tecan SPECTRAFluor Plus spectrophotometer. Excitation and emission wavelengths were 485nm and 535nm, respectively, with a fixed gain set at 40.

6. RTC Counter Mathematical Modeling

Mathematical modeling was used to verify the logic-based predictions of our design, to investigate the effects of pulse frequency and pulse length on the performance of the RTC counters, and to explore the possibility of counting to higher numbers. We used ordinary differential equations (ODE) to describe the temporal trajectories of population averages for all biochemical species. Stochastic modeling was not included because of the population homogeneity demonstrated in Fig. S4. Details for the modeling of each of the two constructs are explained in the sections below.

6.1. The RTC 2-Counter Model

Based on the design of the RTC 2-Counter, we approximated the system dynamics using the following biochemical reactions:

$$\phi \underset{d\_taRNA}{\overset{s\_taRNA}{\rightleftharpoons}} taRNA \quad (1)$$

$$\phi \underset{d\_mT7}{\overset{s0\_mT7}{\rightleftharpoons}} mT7cr \quad (2)$$

$$\phi \underset{d\_mGFP}{\overset{s\_GmRNA}{\rightleftharpoons}} mGFPcr \tag{3}$$

$$mT7cr + taRNA \underset{kb1}{\overset{kf1}{\rightleftharpoons}} mT7 \tag{4}$$

$$mGFPcr + taRNA \underset{kb2}{\overset{kf2}{\rightleftharpoons}} mGFP \tag{5}$$

$$mT7 \xrightarrow{sT7} mT7 + pT7 \tag{6}$$

$$mGFP \xrightarrow{sGFP} mGFP + pGFP \tag{7}$$

$$pT7 \xrightarrow{d\_pT7} \phi \tag{8}$$

$$pGFP \xrightarrow{d\_pGFP} \phi \tag{9}$$

where Eqs. (1)-(3) represent the synthesis and degradation of trans-activator (*taRNA*), T7 RNA polymerase transcripts in cis-repressed form (*mT7cr*), and GFP transcripts in cis-repressed form (*mGFPcr*), respectively. Transcripts in cis-repressed form are indicated by "cr". Kinetic parameters are as indicated in the equations. Eqs. (4) and (5) represent the binding of *taRNA* with *mT7cr* and *mGFPcr* so that the transcripts can be translated; these repression-relieved transcripts are denoted as *mT7* and *mGFP*. Eqs. (6) and (7) represent the translations of *mT7* and *mGFP*, respectively, with *pT7* and *pGFP* as notations for these two proteins. Finally, Eqs. (8) and (9) represent the degradation of proteins. As shown in Fig. 2A of the main text, these biochemical reactions were sufficient to describe the system dynamics with high accuracy.

Based on these reactions, we wrote down the differential equations that describe the temporal evolution of all the species. Some of the parameters in the biochemical reactions are lumped parameters that are expanded to their explicit forms in the differential equations. The notations for all chemical species in this model (and the RTC 3-Counter

model) are simplified and listed in Table S1, with all parameter values listed in Table S2. The square brackets in these equations indicate chemical species concentration. Because the fluorescence data to which we directly fit the model (see below for details) have arbitrary units, GFP protein concentrations in the model are considered nondimensional. All other parameter values, except for degradation rates (min^{-1}) and *k_ara* (concentration), are nondimensional as well. The following five equations were used to capture the temporal dynamics of the system:

$$\frac{d[taRNA]}{dt} = sT\frac{[ara]}{[ara]+k\_ara} + s0\_taRNA - d\_taRNA\cdot[taRNA] \quad (10)$$

$$\frac{d[mT7cr]}{dt} = s0\_mT7cr - d\_mT7cr\cdot[mT7cr] \quad (11)$$

$$\frac{d[pT7]}{dt} = s_0\_pT7\cdot[mT7cr] + s\_pT7k\cdot[taRNA]\cdot[mT7cr] - d\_pT7\cdot[pT7] \quad (12)$$

$$\frac{d[mGFPcr]}{dt} = s0\_mGFPcr + k\_pT7\cdot\frac{[pT7]^n}{km7^n+[pT7]^n} - d\_mGFP\cdot[mGFPcr] \quad (13)$$

$$\frac{d[pGFP]}{dt} = s_0\_pGFP\cdot[mGFPcr] + s\_pGFPk\cdot[taRNA]\cdot[mGFPcr] - d\_pGFP\cdot[pGFP] \quad (14)$$

where on the right-hand side of Eq. (10), *taRNA* synthesis rate has two parts: the first part (*sT*[ara]/([ara]+kara)*) represents the synthesis rate induced by arabinose and the second part (s_0*_taRNA)* represents the basal production rate without any induction,. To simplify the system, we assumed that the arabinose induction effect has a Hill function form with a Hill coefficient equal to 1. The third term of Eq. (10) represents *taRNA*

degradation using a simple exponential decay with rate *d_taRNA*. In Eq. (11), cis-repressed T7 RNA polymerase transcripts (*mT7cr*) are constitutively expressed, with a constant production rate (*s0_mT7cr*) and exponential decay. Similarly, in Eq. (12), T7 RNA polymerase protein synthesis rate has two parts: s_0*_pT7*[mT7cr]* represents the translation rate of *mT7cr* without *taRNA* binding, and *s_pT7k*[taRNA]*[mT7cr]* represents the translation rate of *mT7cr* with *taRNA* binding. Here we assumed that *taRNA* binding and dissociation with mRNA [Eqs. (4) and (5)] have a much faster time scale than other reactions and reach equilibrium instantly. Thus, the parameter *s_pT7k*, for example, is also a lumped parameter with information about the binding reaction in Eq. (4) included. In Eq. (13), GFP mRNA synthesis depends on the basal transcription rate and on T7 RNA polymerase protein abundance. We used a general Hill function to describe this dependency: $k\_pT7*[pT7]^n/([pT7]^n+[km7]^n)$, where *n* accounts for any type of cooperativity caused by T7 RNA polymerase activation. In Eq. (14), GFP protein dynamics parallel that of the T7 RNA polymerase protein in Eq. (12).

6.2. Extension to the RTC 3-Counter

The RTC 3-Counter construct is similar to the RTC 2-Counter in design and topology (Fig. 1C of main text), and they have a number of components in common. So for the RTC 3-Counter model we used many of the same equations used for the RTC 2-Counter model. Besides the reactions in Eqs. (1)-(9), there are four additional reactions:

$$\phi \underset{d\_mT3}{\overset{s\_T3mRNA}{\rightleftharpoons}} mT3cr \tag{15}$$

$$mT3cr + taRNA \underset{kb1}{\overset{kf1}{\rightleftharpoons}} mT3 \tag{16}$$

$$mT3 \xrightarrow{sT3} mT3 + pT3 \quad (17)$$

$$pT3 \xrightarrow{d\_pT3} \phi \quad (18)$$

where Eq. (15) represents the synthesis and degradation of cis-repressed T3 RNA polymerase transcripts (*mT3cr*) and Eq. (16) represents the binding of *taRNA* with *mT3cr*. Eqs. (17) and (18) represent translation and degradation of T3 RNA polymerase protein, respectively. The differential equations describing the RTC 3-Counter construct are similar to the RTC 2-Counter differential equations, except Eq. (13) changes to:

$$\frac{d[mGFPcr]}{dt} = s0\_mGFPcr + k\_pT3 \cdot \frac{[pT3]^n}{km3^n + [pT3]^n} - d\_mGFP \cdot [mGFPcr] \quad (19)$$

and the following two equations are added:

$$\frac{d[mT3cr]}{dt} = s0\_mT3cr + k\_pT7 \cdot \frac{[pT7]^n}{km7^n + [pT7]^n} - d\_mT3 \cdot [mT3cr] \quad (20)$$

$$\frac{d[pT3]}{dt} = s_0\_pT3 \cdot [mT3cr] + s\_pT3k \cdot [taRNA] \cdot [mT3cr] - d\_pT3 \cdot [mT3cr] \quad (21)$$

Eqs. (20) and (21) describe the change of T3 RNA polymerase transcripts and proteins over time. They have the same forms as Eqs. (13) and (14), respectively, with similar parameter implications.

6.3. RTC Counter Arabinose Induction

Different counter strains were induced by different numbers of external arabinose pulses to test and verify the counting behavior. To account for the arabinose pulse dynamics, we modeled it with two differential equations. The first equation describes arabinose when it is present in the medium:

$$\frac{d[ara]}{dt} = -cAra \tag{22}$$

This represents a constant consumption rate of arabinose, when it is present in abundance. The second equation describes arabinose after the cells have been spun and resuspended in arabinose-free media. The leftover, mainly intracellular arabinose is modeled as an exponentially decaying chemical species:

$$\frac{d[ara]}{dt} = -dAra \cdot [ara] \tag{23}$$

In the simulations, Eqs. (22) and (23) were used alternately so as to be consistent with actual experimental conditions.

6.4. RTC Counter Fitting of Experimental Data

Matlab function *lsqcurvefit* was used to narrow down the model parameters by fitting the model equations to experimental measurements. The parameter set that resulted in the optimal data fitting among two hundred runs was chosen as the set for Fig. 2 of the main text, with fluorescence levels of uninduced samples subtracted from all other experimental data. Parameters values used for these figures are listed in Table S2. The experimental arabinose doses used in the RTC 3-Counter experiments were ten-fold higher than those in the RTC 2-Counter experiments; thus parameters *k_ara* and *cAra*

were adjusted ten-fold higher (as written in Table S2) for the RTC 3-Counter model to match the experimental results.

7. DIC Counter Plasmid Construction

DIC Counter plasmids were constructed using basic molecular cloning techniques (*S1*). New England Biolab's restriction endonucleases, T4 DNA Ligase, and NEB's Phusion PCR kits were used. PCRs were carried out with an MJ Research PTC-200 Peltier Thermal Cycler. Synthetic oligonucleotides were made by Integrated DNA Technologies. Single-inducer DIC Counter plasmids were transformed into *E. coli* strain DH5α (F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-). Multiple-inducer DIC Counter plasmids were transformed into *E. coli* strain DH5αPRO (F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-, P_{N25}/tetR, P_{laciq}/lacI, Sp^r). Transformations were carried out using standard electroporation protocols (*S1*) and isolated with Qiagen QIAprep Spin Miniprep Kits. Plasmid modifications were confirmed by restriction digests.

8. DIC Counter Plasmid Design

The single-inducer DIC 2-Counter (Fig. S8) and 3-Counter (Fig. S9) and multiple-inducer DIC 3-Counter (Fig. S10) were based on the single-copy pBAC platform (*9*). pBAC-lacZ (Addgene plasmid 13422) was obtained from Addgene (Cambridge, MA). We cloned all components for the DIC Counters in between ScaI and PciI restriction sites

in pBAC-lacZ. Custom sequences, including recombinase recognition sites, were constructed using sequential PCR with DNA obtained from Integrated DNA Technologies (Coralville, IA). Promoters $P_{Ltet0-1}$ and P_{A1lacO} and terminators were obtained from Ref. *S2*, while P_{BAD} was obtained from Ref. *4*. The *cre* gene was obtained from Ref. *20*. The flp_e gene was derived from pCAG-Flpe (Addgene plasmid 13787) and based on Ref. *S4*. The ribosome-binding sequences used in each stage were derived from Ref. *12*, while *ssrA*-based degradation tags were designed according to Ref. *S5*.

9. DIC Counter Experimental Conditions

All experiments were performed in Luria-Bertani media containing 30 μg/mL kanamycin. Prior to performing flow cytometer measurements on the DIC Counters, cells were grown overnight. To initiate experiments, cells were diluted 1:2000 in fresh media and grown at 37˚C and 300 rpm with inducers as indicated in the specific figures. Inducer concentrations were anhydrotetracycline = 700 ng/mL, arabinose = 0.1%, and IPTG = 10 mM except for Fig. 4B in which anhydrotetracycline = 100 ng/mL, arabinose = 10^{-3}%, and IPTG = 10 mM. At all inducer transitions (i.e., transitions from media with inducer to media without inducer, transitions from media without inducer to media with inducer, or transitions from media with one inducer to media with another inducer), cells were diluted 1:2000 in fresh media. Only in Fig. 3C with pulse intervals 2h and 4h were cells not diluted 1:2000 in fresh media for transitions from media without inducer to media with inducer due to low optical density of the cultures. Instead, inducer was added directly to the media; however, cells were still diluted 1:2000 in fresh media for transitions from media with inducer to media without inducer.

10. DIC Counter Flow Cytometer Measurements

Data for Figs. 3, 4, S11, S12, and S13 were collected with a Becton Dickinson FACSCalibur flow cytometer. Fluorescence was calibrated with Calibrite Beads (Becton Dickinson) and measured with a 488-nm argon laser excitation and a 515-nm to 545-nm emission filter. Before analysis, cells were diluted in sterile phosphate-buffered saline. Becton Dickinson Calibrite Beads were used for instrument calibration. 50,000 cells were collected for each sample, gated to ensure consistency between samples, and processed with MATLAB to calculate mean fluorescence data points (Mathworks, Natick, MA).

SOM Text

1. RTC Counter Characteristics and Improvements

1.1. Whole Population Measurements

To verify the counting behavior, we also analyzed the RTC 2-Counter and the RTC 3-Counter with a spectrophotometer, which measures the total fluorescence in a given cell population. These spectrophotometer results (Figs. S3A and S3B) corroborate the data from the flow cytometer (Figs. 1B and 1D). In the case of the RTC 2-Counter (Fig. S3A), the uninduced population similarly shows no increase in fluorescence, while populations that received either the first or the second arabinose pulse exhibit only some fluorescence. Cells that receive both pulses show a striking increase in fluorescence at 50 minutes, validating our design. The spectrophotometer measurements of the RTC 3-Counter reveal

a similar corroboration, in which only cells that are pulsed three times respond with sharp increases in fluorescence.

Flow cytometer and spectrophotometer data sets do diverge qualitatively, where flow cytometer measurements exhibit a peak in fluorescence and then decrease whereas spectrophotometer measurements exhibit a fluorescence plateau. The decrease is likely due to external factors such as cell division (*S6*), and is revealed in single-cell measurements of the flow cytometer. This effect is not seen in the spectrophotometer, where measurements are made on whole populations. Data presented in Figs. S3A and S3B are the mean of three replicates, and smoothed with a rolling window average.

1.2. Flow Cytometry Population Analysis

The data presented in Figs. 1B and 1D are mean fluorescence values of RTC counter cell populations, measured by a flow cytometer. In Fig. S4, we show the fluorescence profile of the entire RTC 3-Counter population when it is uninduced, after the second pulse, and after the third pulse. It is clear that the entire population shifts homogeneously following induction, with the greatest shift occurring as a result of the third pulse.

1.3. Verification of Discrete Counting

To verify that the counting response is driven by discrete induction pulses and not simply a summation of induction length, we took a fixed total length of induction and split it into two and three pulses. RTC 3-Counter cells were either given two short pulses followed by a long pulse or two long pulses, with total induction time equal for both sets of cells

(Fig. S5 inset). It can be seen in Fig. S5 that cells receiving three pulses (blue) generate significantly more GFP than cells receiving two pulses (red), demonstrating a true counting mechanism and not simply a summing effect. This supports our claim that the counter is able to distinguish between different numbers of pulses, even when total induction time is held constant. Additionally, our mathematical model accurately predicted the experimental results for both scenarios.

1.4. Higher Number Counters

To investigate the possibility of expanding our design to count higher numbers, we hypothetically expanded our system using mathematical modeling. We added extra genes to the cascade, each one an RNA polymerase whose downstream promoter regulates the transcription of the gene at the next node. We modeled cascades with up to ten nodes; in each case the first node is T7 RNA polymerase, the last node is GFP, and all nodes in between are polymerases with exactly the same kinetic properties as T3 RNA polymerase. With two additional differential equations for each node, we use mathematical modeling to predict the behavior of these higher number counters by comparing the fluorescence readout of n, n-1, and n-2 arabinose pulses for each n-node counter. As shown in Fig. S6A, the red line is the fluorescence result of n pulses, the green line of n-1 pulses, and the blue line n-2 pulses. It can be seen that the absolute difference in fluorescence levels between n and n-1 pulses increases with cascade length, suggesting the design can better distinguish different numbers of pulses as it is extended. Additionally, all three lines

increase as the construct is extended, due to signal propagation and the accumulation of long-lived proteins as more pulses are delivered.

This predicted accumulation effect results in the failure of this system to perform digitally as *n* increases, with ones and zeros no longer represented by high and low protein concentrations. However, by examining the temporal dynamics of all the chemical species in the cascades, we identified that it is the long half-life of GFP protein that causes the signal increase after *n*-1 and *n*-2 pulses. Figure S6B is the predicted counter output in which GFP protein has its half-life shortened to 8 minutes instead of the 231 minute half life used for Fig. S6A. Figure S6B illustrates that when the final output protein has a shortened half-life, the counter performance is remarkably robust as *n* increases. Counting from 2 to 10, output from *n* pulses increases almost exponentially while output from *n*-1 and *n*-2 pulses increases only marginally.

If shortening the final protein's half-life is not possible or desirable, an alternative method for digitizing the output signal would be to couple the counter to a toggle switch (12). By placing one of the toggle repressor proteins at the final node of the counter cascade, it would be possible to flip a toggle from one state to the other with expression from the counter. The sharp and tunable switching threshold of a toggle switch may be used to filter out counter leakage due to *n*-1 or *n*-2 pulses, switching states only when *n* pulses produces a concentration of repressor proteins in excess of the switching threshold.

2. Extending the DIC Counter

Each of our individual counting units requires only a single recombinase whereas the protein-based toggle switch utilizes two proteins (*12*). This allows our design to be extendable in a modular fashion using >100 identified recombinases to count to higher numbers (*6*). Recombinases can also be mutagenized to have altered site preferences or thermostabilities, allowing for increased diversity to create synthetic gene circuits. The availability of additional recombinases enables the DIC counter to be extended more readily than other systems that require rarer or more specialized components.

3. RTC and DIC Counter Designs: Possible Improvements

Compared to electronic counters, our biological counters are in an early stage of development and have some distinct limitations. Our counters scale linearly instead of exponentially as is the case with digital electronic circuits that count in binary (*S7*). Counter designs which count in binary require the addition of bit reset and carry operations (*S7*). The DIC counter is amenable to being adapted with advanced digital designs due to the ability of SIMMs to maintain memory and invert in both orientations. Reset operations could be carried out by downstream promoters which drive the transition of inverted SIMMs back to their original orientations. Carry operations could be achieved by components that act in *trans* to affect DNA orientation, insertions, or deletions on many different SIMMs or DIC counters; these *trans*-based components may include bacteriophage integrases and excisionases *(6)* or transcriptional activators. Future development of biological counters with exponential scaling will greatly expand the potential applications of biological counters.

An additional limitation of our counters is their inability to detect very high frequency inputs. Though there will invariably be upper limits to pulse frequencies that can be detected by counters, those limits may be improved by combining synthetic counters with pulse-generating circuits that can detect edge transitions with greater rapidity and/or with amplifiers that can enhance the magnitude of inputs. Pulse-generating circuits may also enable the RTC counter and the single-inducer DIC counter to record low-frequency events with greater fidelity.

Supplementary Figure Captions

Fig. S1. The RTC 2-Counter plasmid. Genes are denoted by arrows within the plasmid circle, promoters by arrows on the plasmid circle, transcriptional terminators by red rectangles, taRNA by a green rectangle, and origin by a blue rectangle.

Fig. S2. The RTC 3-Counter plasmid. Genes are denoted by arrows within the plasmid circle, promoters by arrows on the plasmid circle, transcriptional terminators by red rectangles, taRNA by a green rectangle, and origin by a blue rectangle.

Fig. S3. Spectrophotometer measurements of RTC 2-Counter and RTC 3-Counter fluorescence over time. (**A**) Spectrophotometer measurements show total population fluorescence of cells containing the RTC 2-Counter construct. Experimental conditions match exactly those used for Fig. 1B. (**B**) Spectrophotometer measurements of cells containing the RTC 3-Counter construct, with experimental conditions matching exactly those used for Fig. 1D.

Fig. S4. A histogram of cell counts and fluorescence in RTC 3-Counter cells. Shown in red is a population of uninduced cells, in green cells that have been pulsed twice, and in blue cells that have been pulsed three times.

Fig. S5. The RTC 3-Counter response to varying length pulses. The counter is induced by two sets of arabinose pulses, displayed as blue and red bars in the figure inset. With 20 minute intervals between all pulses, the first set (in blue) is induced for 11 minutes, then 11 minutes again, then 22 minutes. The second set (in red) is induced by two 22 minute pulses. Experimental data, plotted as a mean fluorescence value for cell populations as measured by flow cytometry, is represented by circles. The blue and red lines are mathematical model predictions corresponding to the pulse patterns in the inset.

Fig. S6. Model predictions of the fluorescence output of *n*-node RTC Counters in response to *n*, *n*-1, and *n*-2 arabinose pulses. (**A**) The numbers on the x-axis represent counters with *n* nodes, and for each counter we plot the fluorescence output due to *n*, *n*-1, and *n*-2 pulses. We use our best-fit parameter values in this figure, the same as those used for Fig. 2. (**B**) This is similar to Fig. S6A, except that GFP protein half-life has been reduced from 231 minutes to 8 minutes.

Fig. S7. Abstract design of the Single Invertase Memory Modules (SIMMs) used in the DIC counters. The SIMMs are composed of opposing recombinase recognition sites (R_f and R_r) which contain between them an inverted promoter (P_{inv}), a synthetic ribosome-binding-sequence (RBS), a recombinase gene (*rec*), an *ssrA*-based degradation tag, and a transcriptional terminator (Term). The SIMM maintains memory based on its DNA orientation, which can be inverted when the recombinase is expressed.

Fig. S8. The single-inducer DIC 2-Counter plasmid. Genes are denoted by arrows within the plasmid circle, promoters by arrows on the plasmid circle, transcriptional terminators by red rectangles, *ssrA*-based degradation tags by brown rectangles, and recombinase recognition sites by rectangles of other colors.

Fig. S9. The single-inducer DIC 3-Counter Plasmid. Genes are denoted by arrows within the plasmid circle, promoters by arrows on the plasmid circle, transcriptional terminators by red rectangles, *ssrA*-based degradation tags by brown rectangles, and recombinase recognition sites by rectangles of other colors.

Fig. S10. The multiple-inducer DIC 3-Counter Plasmid. Genes are denoted by arrows within the plasmid circle, promoters by arrows on the plasmid circle, transcriptional terminators by red rectangles, *ssrA*-based degradation tags by brown rectangles, and recombinase recognition sites by rectangles of other colors.

Fig. S11. The single-inducer DIC 2-Counter construct design and results. **(A)** The single-inducer DIC 2-Counter is characterized by a single Single Invertase Memory Module (SIMM) with P_{BAD} as the inducible upstream promoter and inducible inverted promoter within the SIMM. **(B)** Mean fluorescence of single-inducer DIC 2-Counter cell populations over time, measured by a flow cytometer, demonstrates a significant increase in GFP fluorescence after exposure to two pulses of arabinose. **(C)** Mean fluorescence of single-inducer DIC 2-Counter cell

populations over time, measured by a flow cytometer, demonstrates that cells grown with no inducer for 9 hours followed by a single pulse of arabinose lasting 7 hours did not show significant GFP expression. Mean fluorescence was normalized against the maximum fluorescence for cells obtained in Figs. S11B and S11C in order to allow comparison between the two plots.

Fig. S12. Flow cytometry population data for the single-inducer DIC 3-Counter exposed to zero, one, two, or three pulses of arabinose. Each arabinose pulse was 8 hours long and spaced by 9 hours of no arabinose exposure. The data demonstrate that there is no leakage with one pulse ("Ara" in the legend), a small degree of leakage with two pulses ("Ara => Ara" in the legend), and a large degree of activation after three pulses ("Ara => Ara => Ara" in the legend).

Fig. S13. Switching times for each SIMM stage in the multiple-inducer DIC 3-Counter were examined by varying the length of exposure to either anhydrotetracycline or arabinose while holding all other inputs constant (aTc followed by Ara followed by IPTG). When not being varied, aTc and Ara pulses were 18 hours in duration and IPTG pulses were 12 hours in duration. The last input to the multiple-inducer DIC 3-Counter, IPTG, does not drive an invertase stage and directly induces the transcriptional GFP output of the system. **(A)** The first SIMM stage responds to aTc within 6 hours of exposure. Very long aTc exposure times did not result in increased GFP fluorescence. **(B)** The second SIMM stage begins to respond to

Ara within 9 hours of exposure. Very long arabinose exposure times did not result in increased GFP fluorescence.

SOM Tables

Table S1. A summary of chemical species represented in the RTC Counter model and their notations.

Notation	**Chemical species**
taRNA	*trans*-activator
mT7cr	*cis*-repressed T7 RNA polymerase mRNA
mT7	T7 RNA polymerase mRNA
pT7	T7 RNA polymerase protein
mGFPcr	*cis*-repressed GFP mRNA
mGFP	GFP mRNA
pGFP	GFP protein
mT3cr	*cis*-repressed T3 RNA polymerase mRNA
mT3	T3 RNA polymerase mRNA
pT3	T3 RNA polymerase protein
ara	Arabinose

Table S2. A list of all parameter values used in the RTC Counter models.

k_ara*	0.0571	d_pT7	0.0056	km7	3.0455
s0_taRNA	0.0008	s_pGFPk	0.9923	k_pT3	3.006
d_taRNA	0.1177	d_pGFP	0.003	s0_mT3cr	0.0003
s0_mT7cr	0.0252	dAra	0.1201	d_mT3	0.0701
d_mT7	0.0706	s0_pT7	0.0003	s0_pT3	0
k_pT7	3.8009	s0_pGFP	0.1007	s_pT3k	0.0115

s0_mGFPcr	0.0123	sT	0.8467	d_pT3	0.0069
d_mGFP	0.07	cAra*	0.0003	n3	0.8892
s_pT7k	0.0766	n7	2.602	km3	7.9075

SOM References

S1. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, Plainview, New York, edn. 2, 1989).

S2. R. Lutz, H. Bujard, *Nucleic Acids Res* **25**, 1203 (1997).

S3. D. Sengupta, D. Chakravarti, U. Maitra, *J Biol Chem* **264**, 14246 (1989).

S4. F. Buchholz, P. O. Angrand, A. F. Stewart, *Nature biotechnology* **16**, 657 (Jul, 1998).

S5. J. B. Andersen et al., *Appl Environ Microbiol* **64**, 2240 (Jun, 1998).

S6. J. Roostalu, A. Joers, H. Luidalepp, N. Kaldalu, T. Tenson, *BMC Microbiol.* **8**, 68 (2008).

S7. P. Horowitz, W. Hill, *The Art of Electronics*. (Cambridge University Press, Cambridge, United Kingdom, ed. 2nd, 1989).

Figure S1

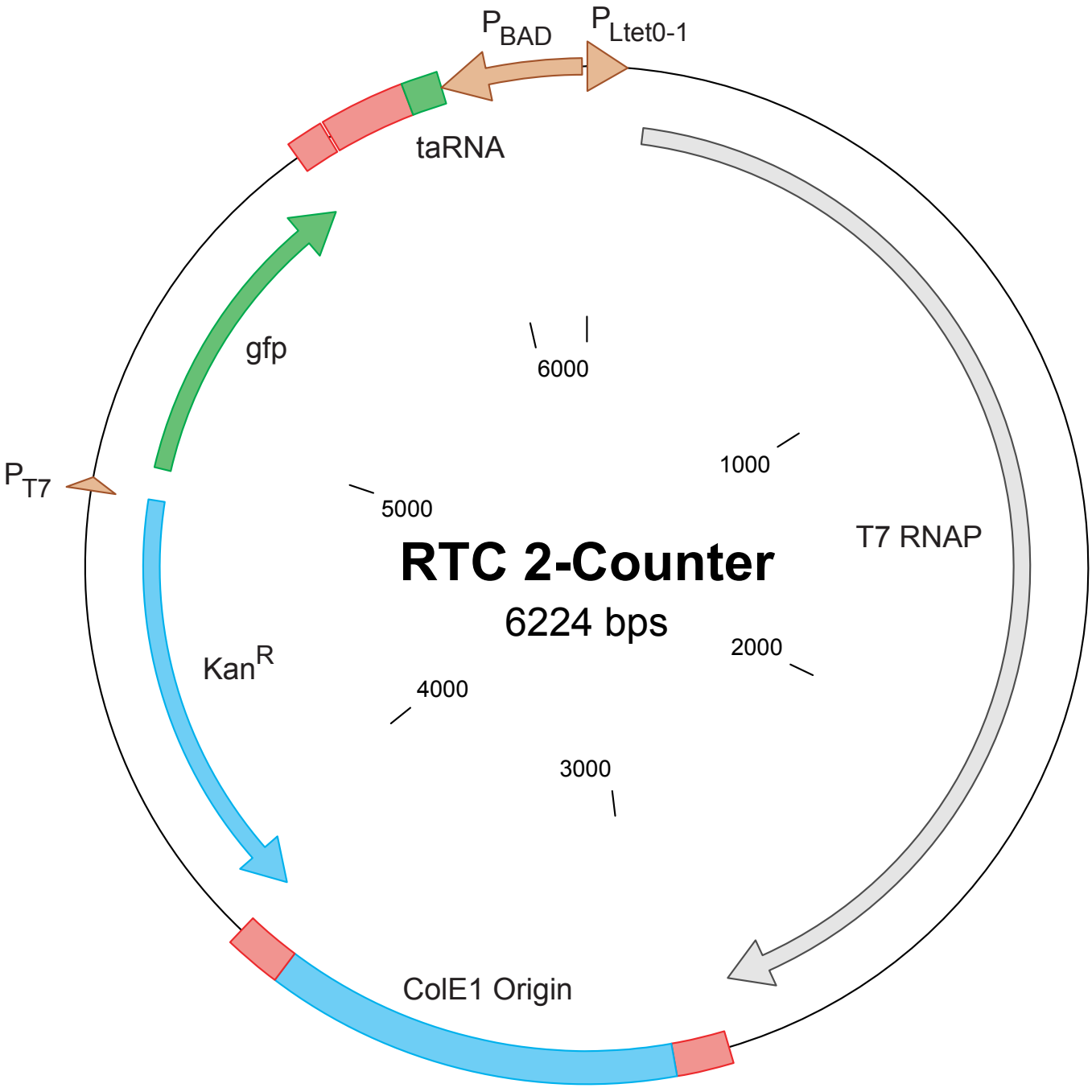

P BAD
P Ltet0-1
taRNA
gfp
6000
1000
P T7
5000
T7 RNAP
RTC 2-Counter
6224 bps
2000
Kan R
4000
3000
ColE1 Origin

Figure S2

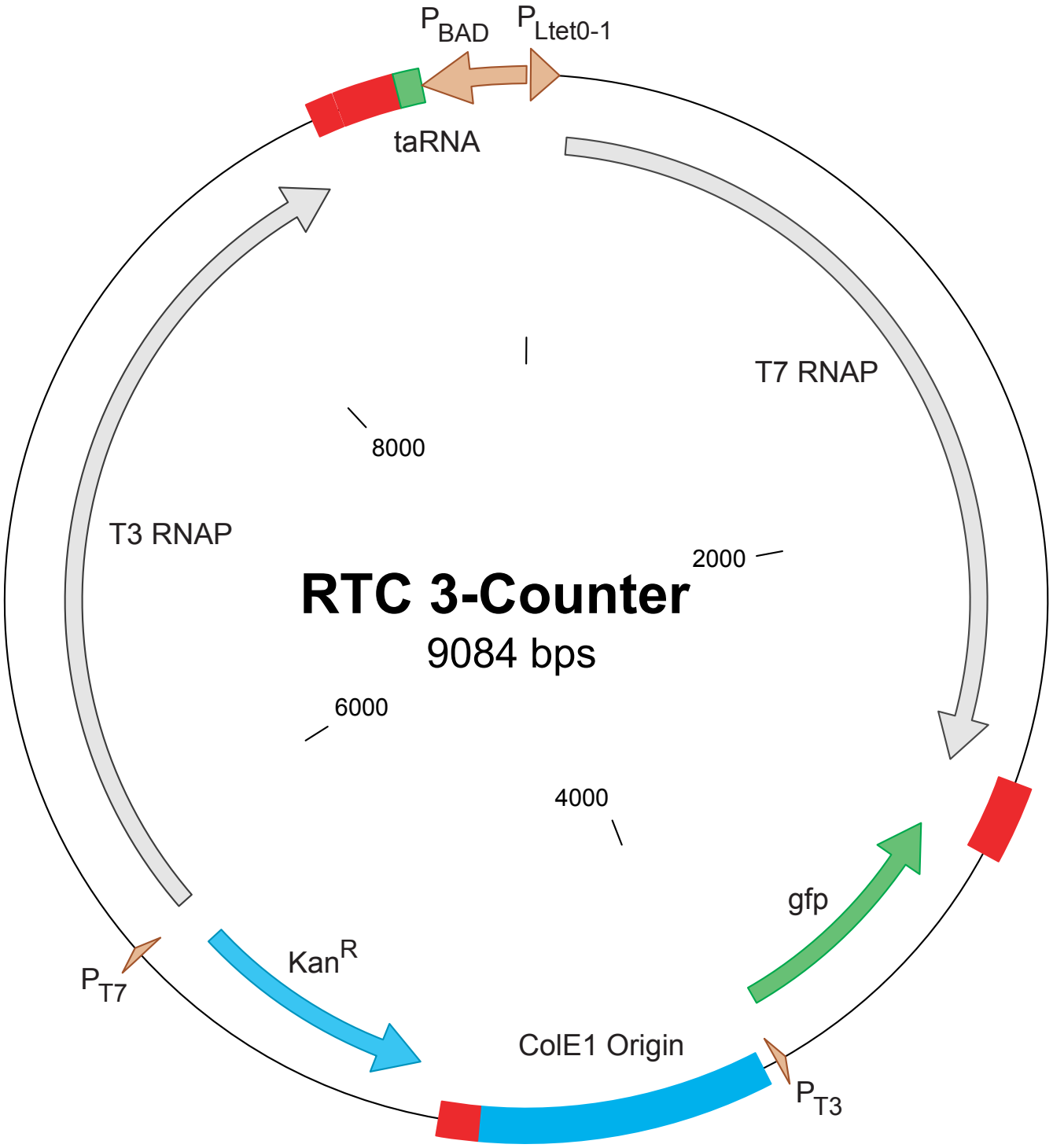
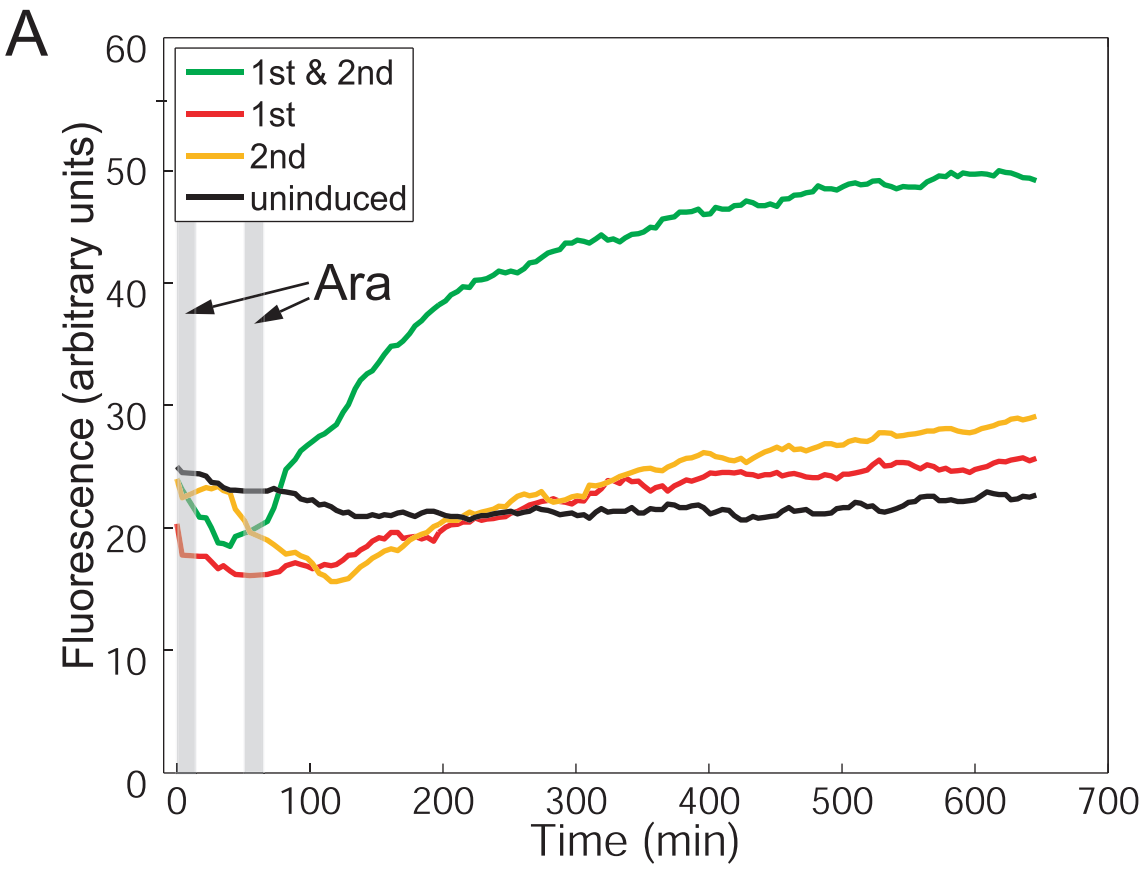

P BAD
P Ltet0-1
taRNA
T7 RNAP
8000
T3 RNAP
2000
RTC 3-Counter
9084 bps
6000
4000
gfp
Kan R
P T7
ColE1 Origin
P T3

Figure S3


A
1st & 2nd
1st
2nd
uninduced
Ara
Fluorescence (arbitrary units)
Time (min)
0
10
20
30
40
50
60
0
100
200
300
400
500
600
700

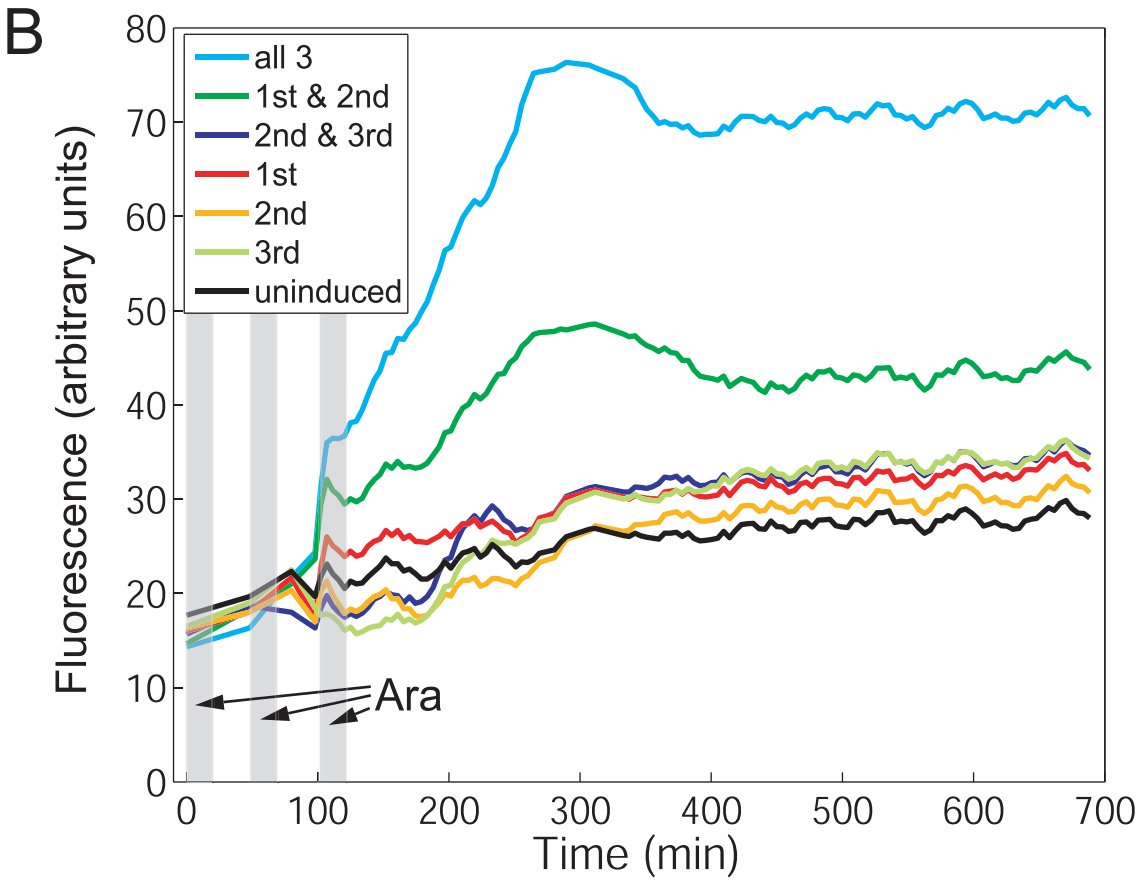

B
all 3
1st & 2nd
2nd & 3rd
1st
2nd
3rd
uninduced
Ara
Fluorescence (arbitrary units)
Time (min)
0
10
20
30
40
50
60
70
80
0
100
200
300
400
500
600
700

Figure S4

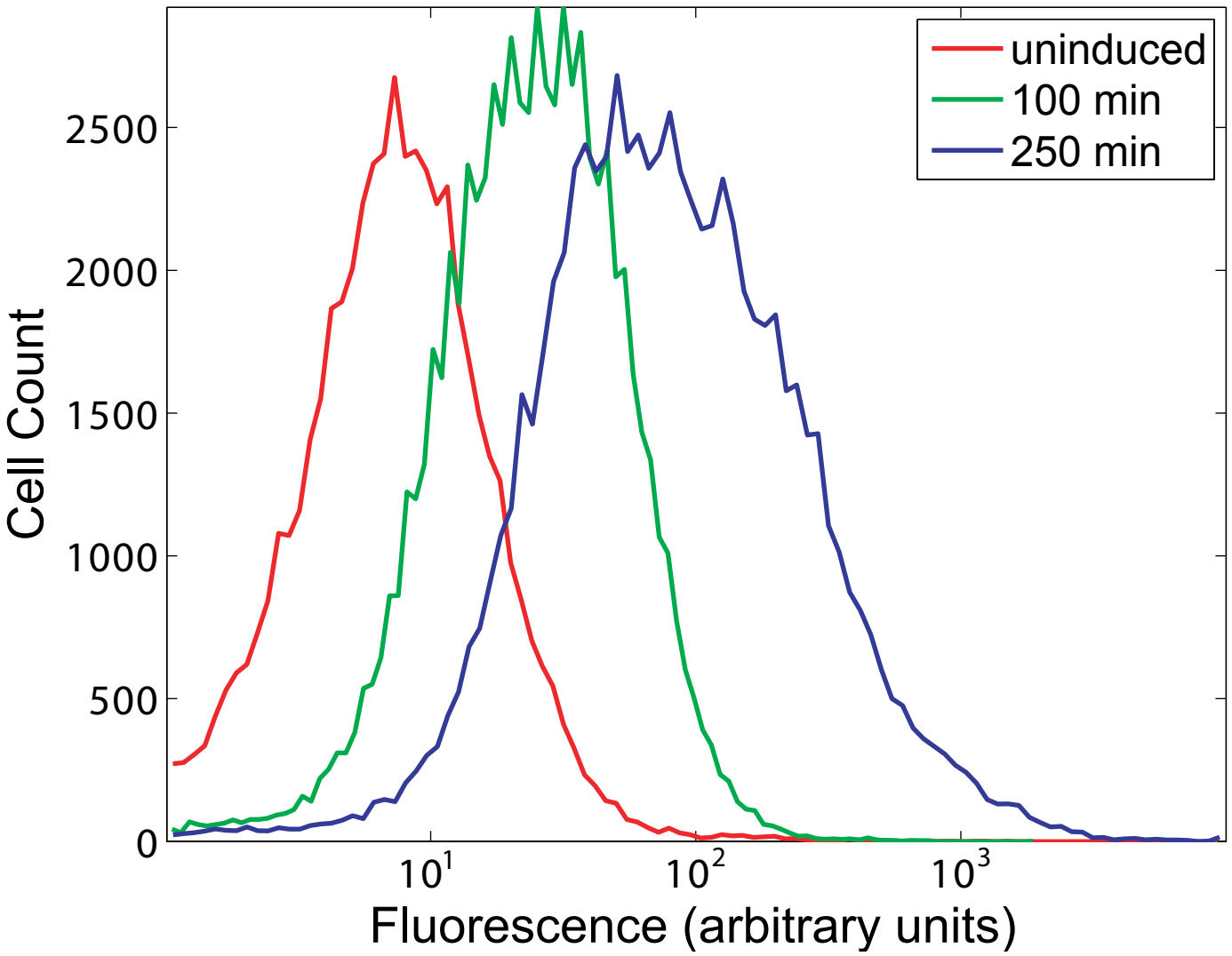

uninduced
100 min
250 min
Cell Count
2500
2000
1500
1000
500
0
10^1
10^2
10^3
Fluorescence (arbitrary units)

Figure S5

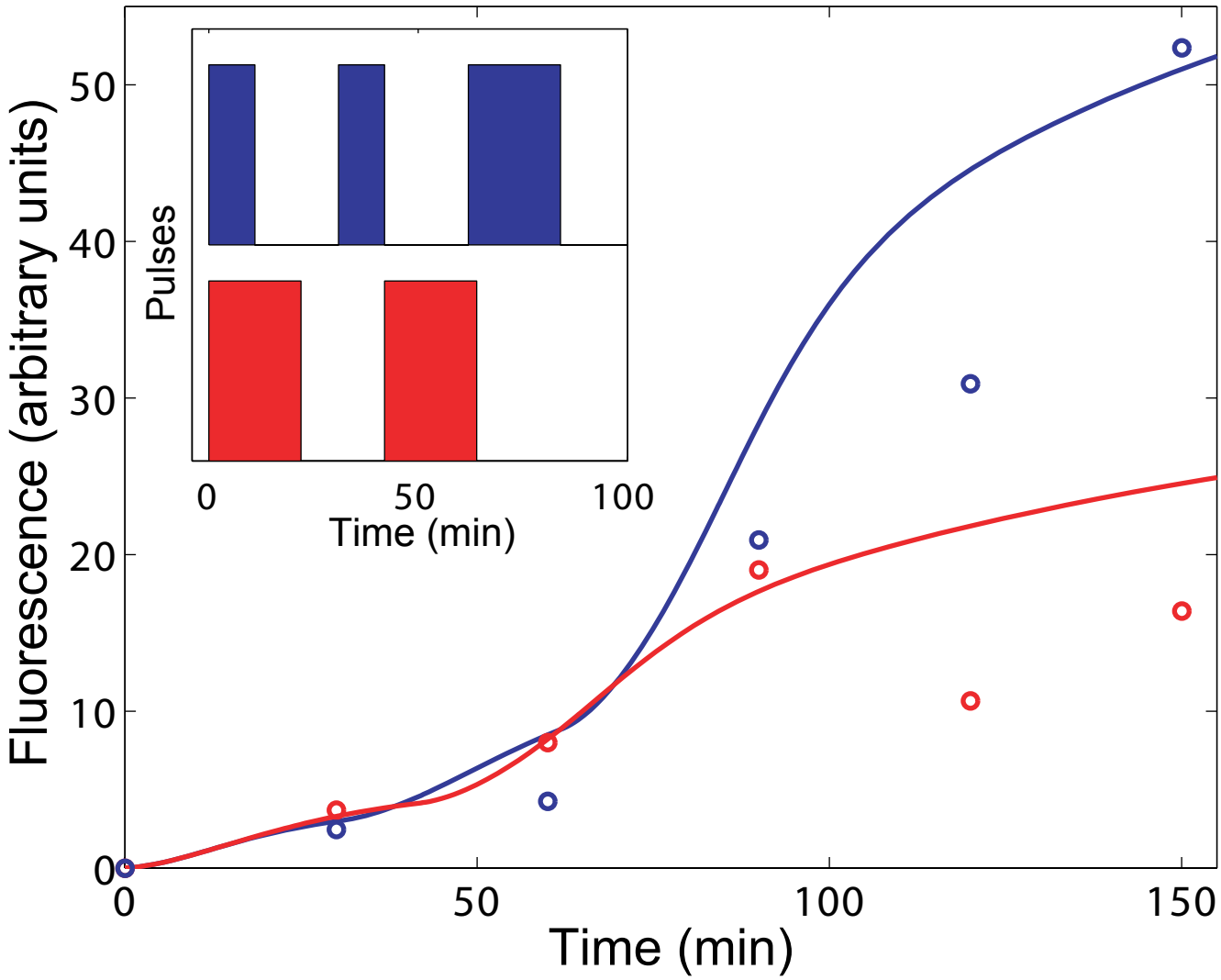

Fluorescence (arbitrary units)
0
10
20
30
40
50
0
50
100
150
Time (min)
Pulses
0
50
100
Time (min)

Figure S6

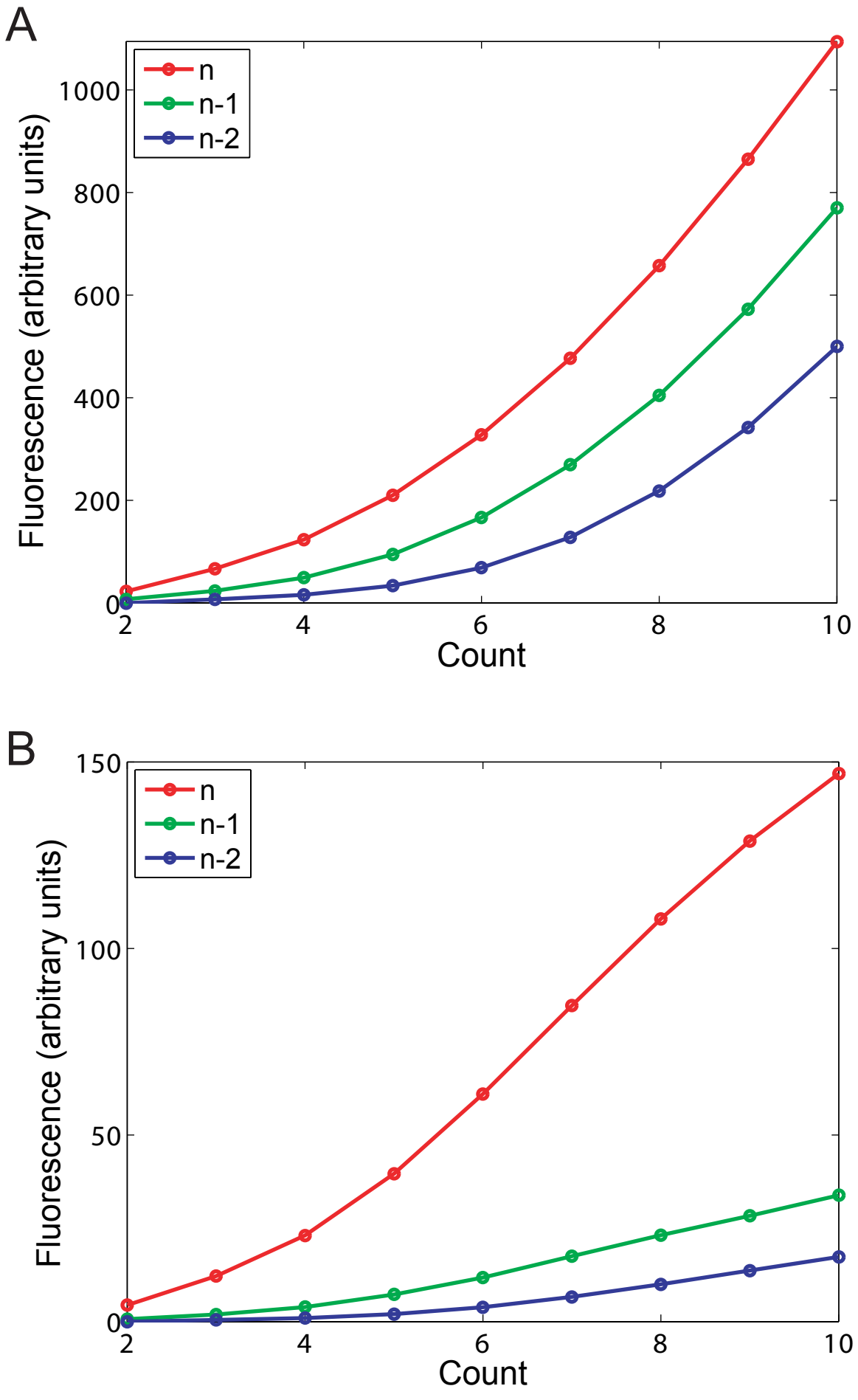

A
n
n-1
n-2
Fluorescence (arbitrary units)
0
200
400
600
800
1000
2
4
6
8
10
Count
B
n
n-1
n-2
Fluorescence (arbitrary units)
0
50
100
150
2
4
6
8
10
Count

Figure S7

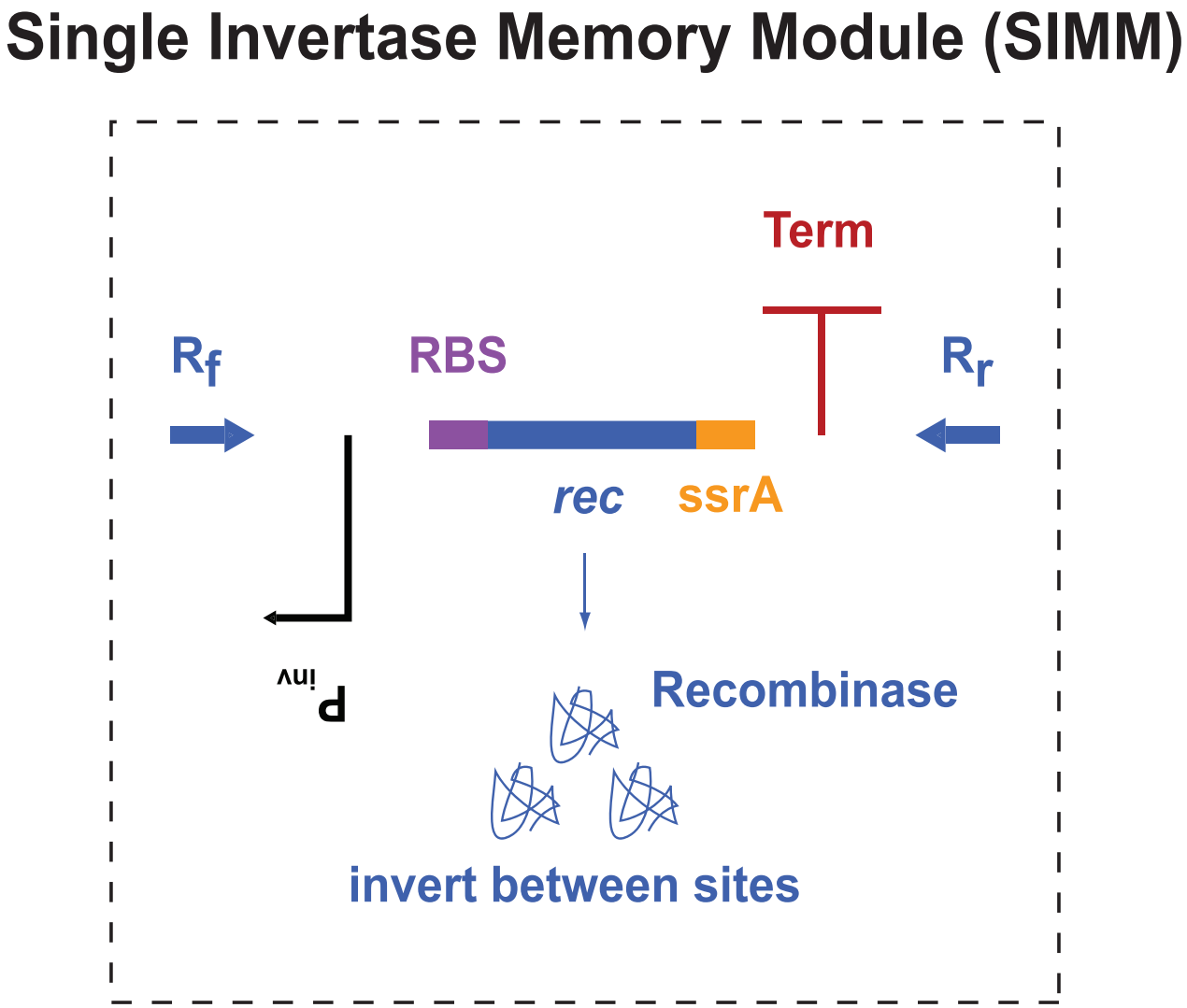

Single Invertase Memory Module (SIMM)
Term
R_f
RBS
R_r
rec
ssrA
P_{inv}
Recombinase
invert between sites

Figure S8

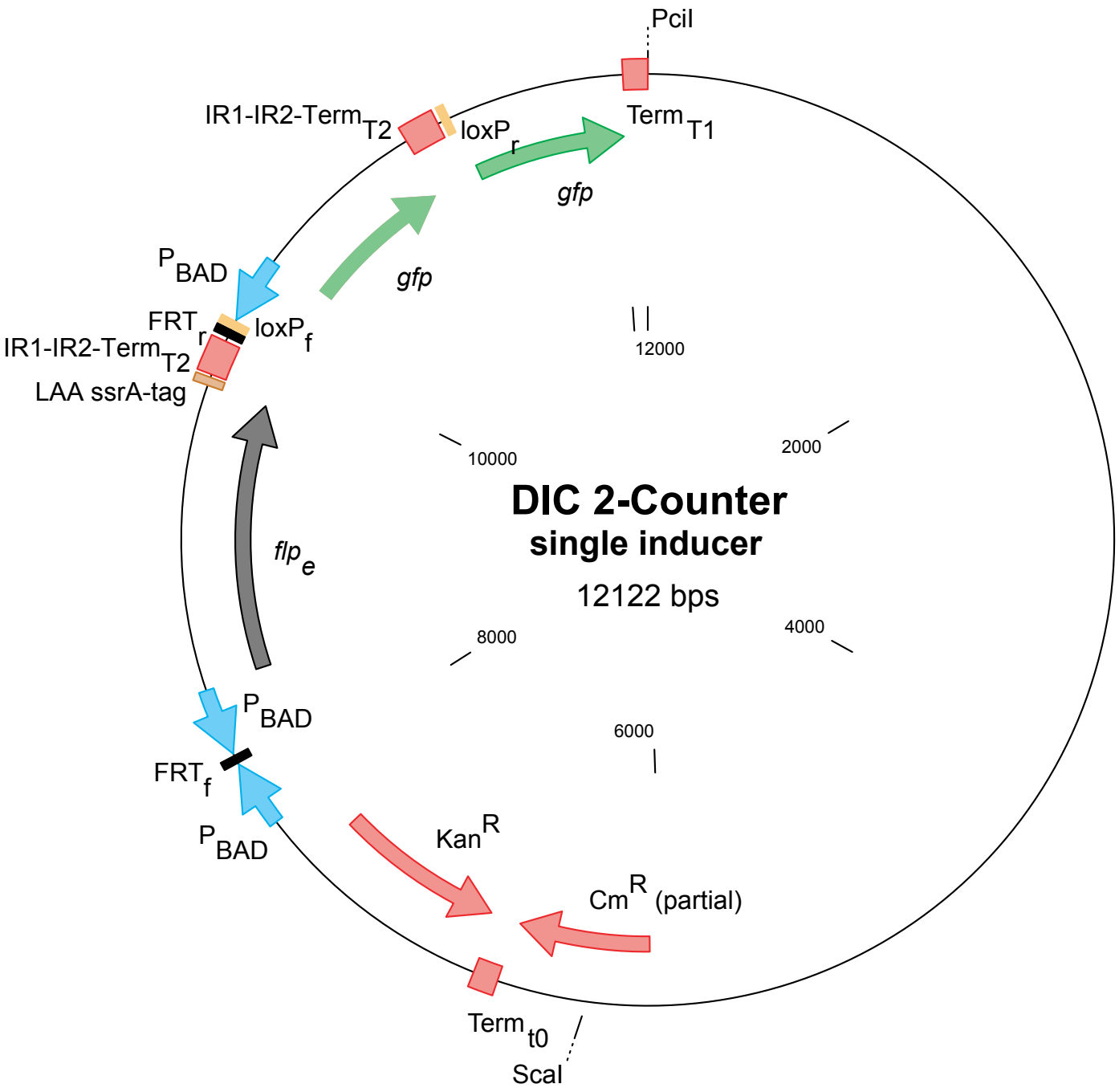

PciI
Term T1
gfp
loxP r
IR1-IR2-Term T2
P BAD
gfp
FRT r
loxP f
IR1-IR2-Term T2
LAA ssrA-tag
12000
10000
2000
DIC 2-Counter
single inducer
12122 bps
flp e
8000
4000
P BAD
6000
FRT f
P BAD
Kan R
Cm R (partial)
Term t0
ScaI

Figure S9

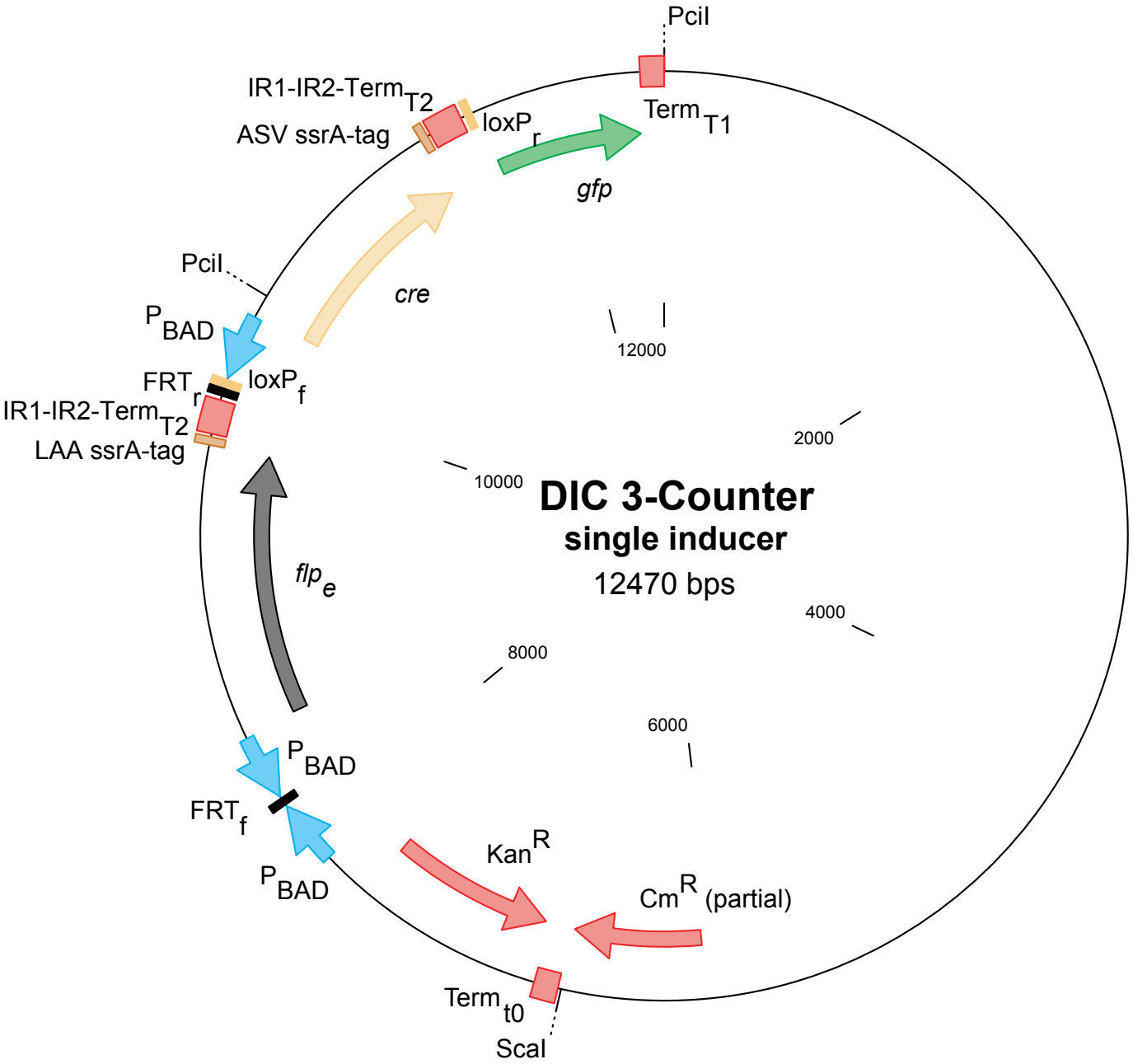

PciI
IR1-IR2-Term T2
ASV ssrA-tag
loxP r
Term T1
gfp
PciI
cre
P BAD
12000
FRT r
loxP f
IR1-IR2-Term T2
LAA ssrA-tag
2000
10000
DIC 3-Counter
single inducer
12470 bps
flp e
4000
8000
6000
P BAD
FRT f
Kan R
P BAD
Cm R (partial)
Term t0
ScaI

Figure S10

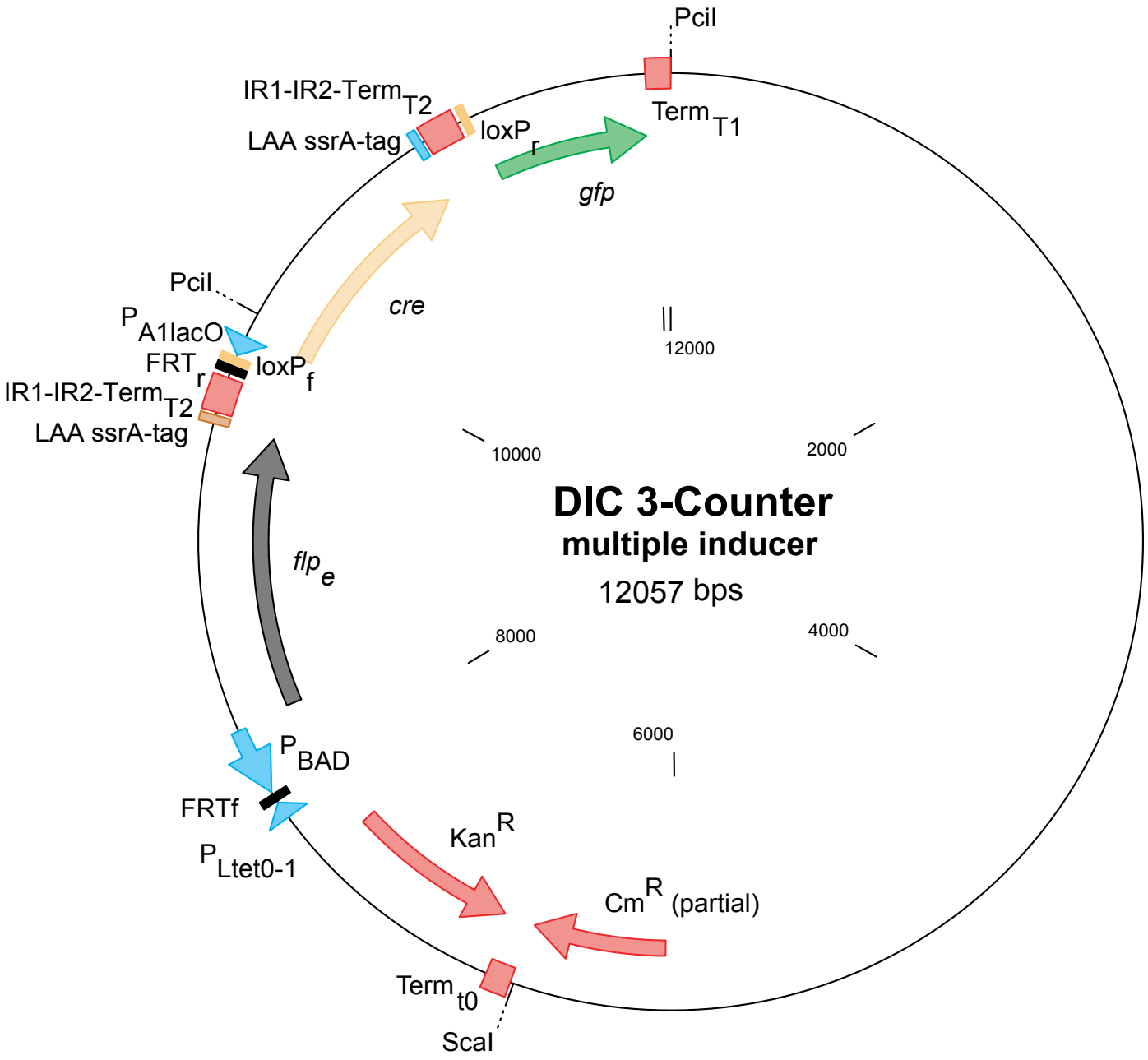

PciI
IR1-IR2-Term T2
LAA ssrA-tag
loxP r
Term T1
gfp
PciI
cre
P A1lacO
FRT r
loxP f
IR1-IR2-Term T2
LAA ssrA-tag
12000
10000
2000
DIC 3-Counter
multiple inducer
12057 bps
flp e
8000
4000
6000
P BAD
FRTf
P Ltet0-1
Kan R
Cm R (partial)
Term t0
ScaI

Figure S11

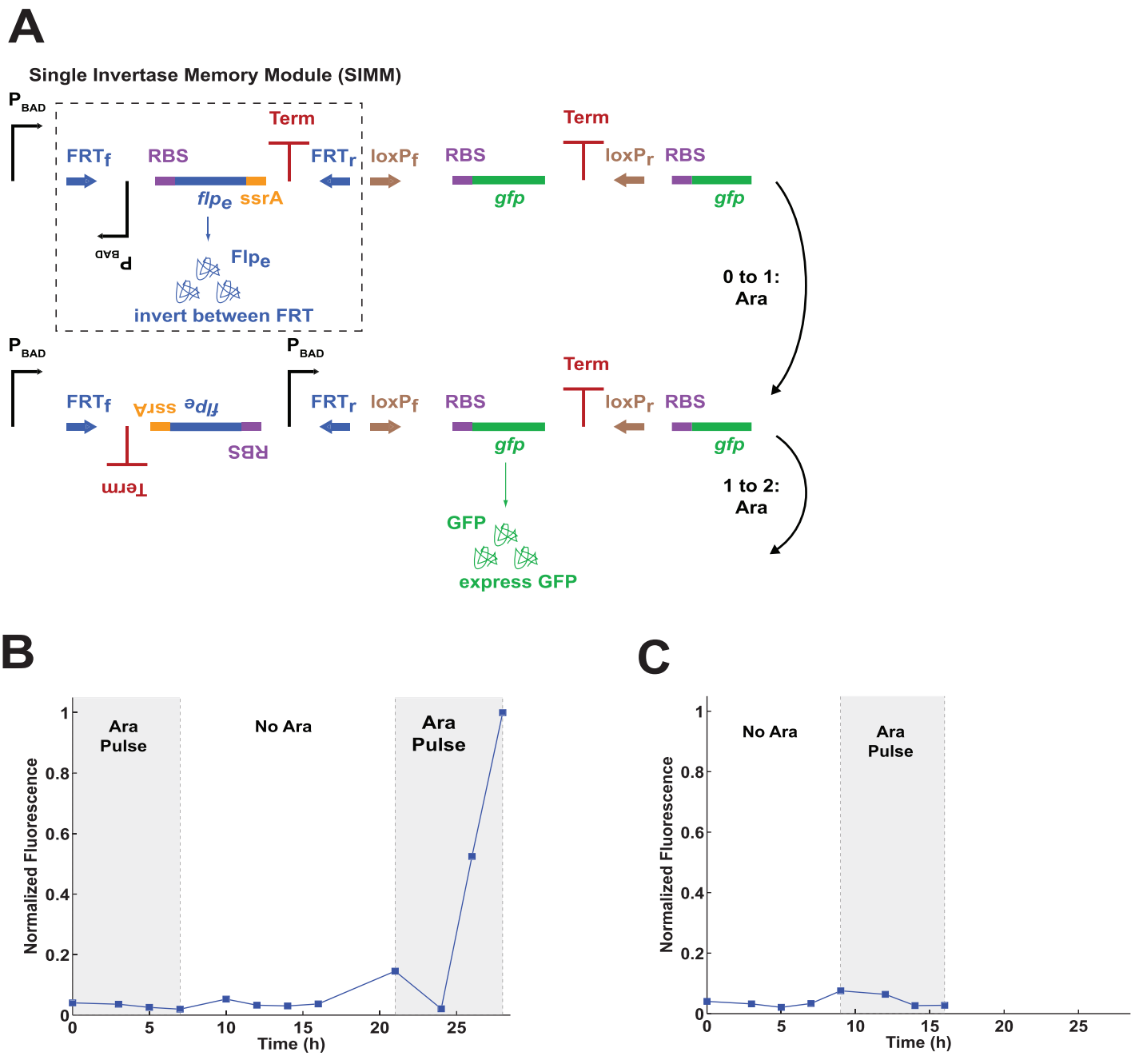

A
Single Invertase Memory Module (SIMM)
P BAD
Term
FRTf
RBS
FRTr
loxPf
RBS
Term
loxPr
RBS
flpe ssrA
gfp
gfp
P BAD
Flpe
invert between FRT
0 to 1:
Ara
P BAD
P BAD
Term
FRTf
flpe ssrA
FRTr
loxPf
RBS
loxPr
RBS
RBS
gfp
gfp
Term
1 to 2:
Ara
GFP
express GFP
B
Ara
Pulse
No Ara
Ara
Pulse
Normalized Fluorescence
Time (h)
C
No Ara
Ara
Pulse
Normalized Fluorescence
Time (h)

Figure S12

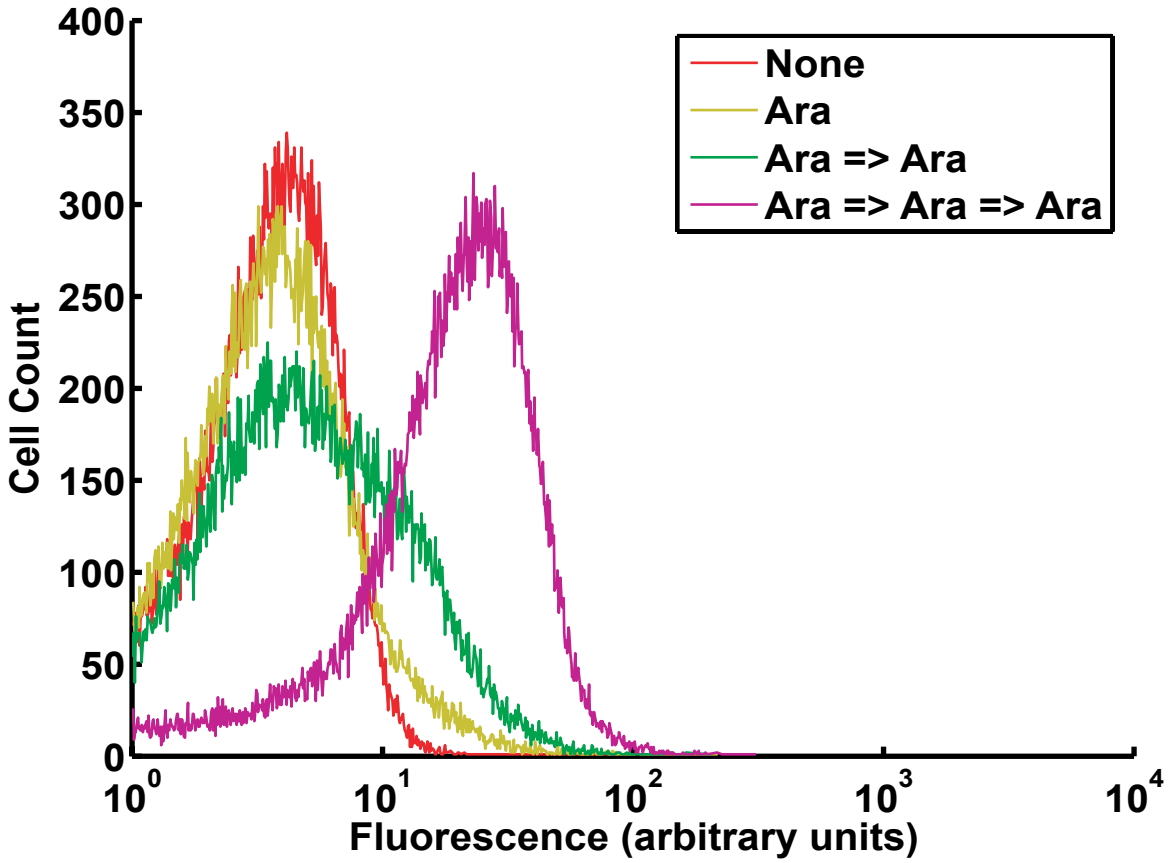
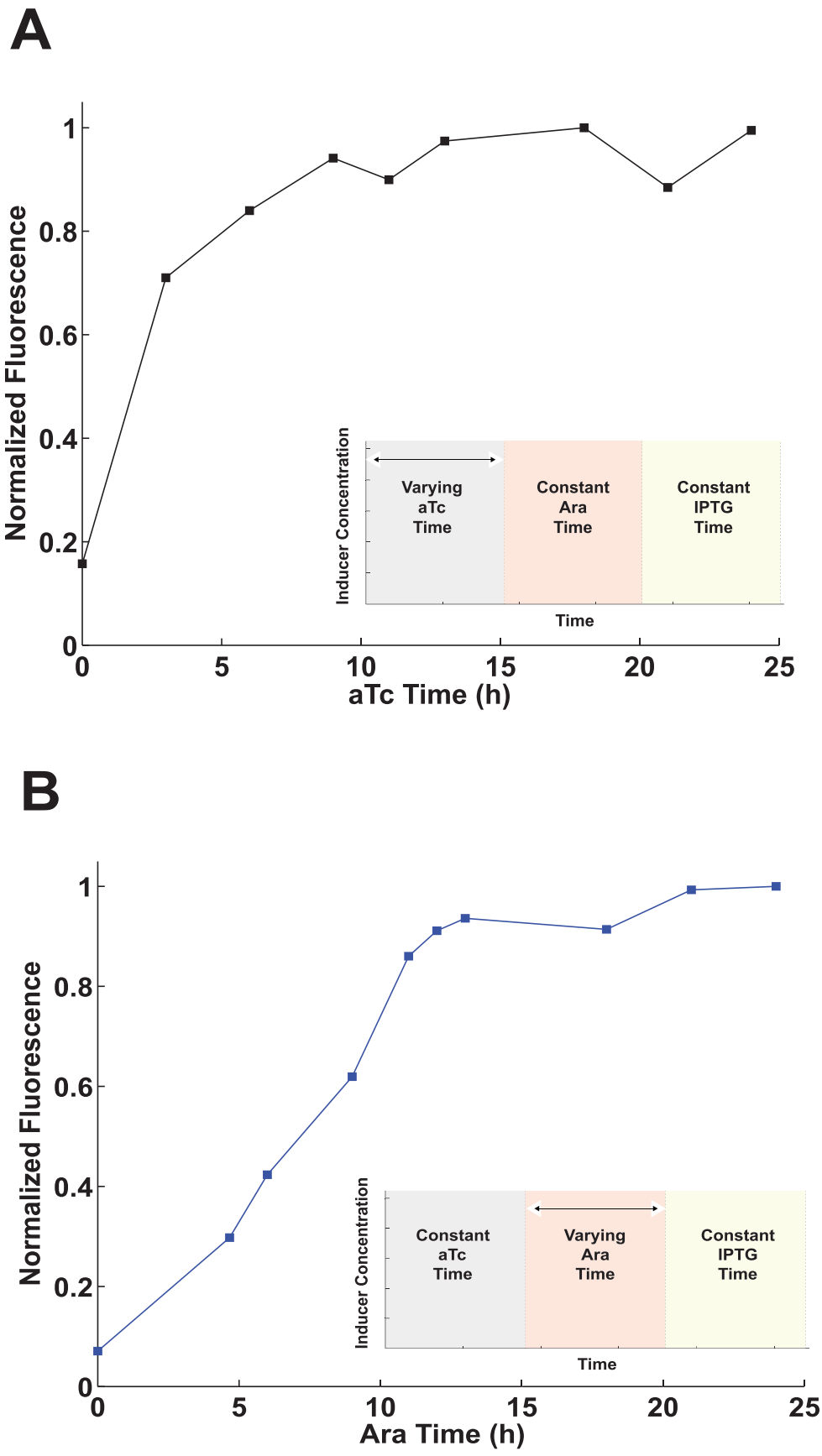

400
350
300
250
200
150
100
50
0
Cell Count
None
Ara
Ara => Ara
Ara => Ara => Ara
10^0
10^1
10^2
10^3
10^4
Fluorescence (arbitrary units)

Figure S13


A
Normalized Fluorescence
aTc Time (h)
Inducer Concentration
Varying aTc Time
Constant Ara Time
Constant IPTG Time
Time
B
Normalized Fluorescence
Ara Time (h)
Inducer Concentration
Constant aTc Time
Varying Ara Time
Constant IPTG Time
Time

tissue-specific membrane trafficking pathway for GLUT4. Specifically, CHC22 copurifies with proteins important for sorting GLUT4 out of endosomes or the trans-Golgi network and into storage vesicles in human muscle cells. CHC22 colocalizes with GLUT4 in these vesicles, and its depletion results in the apparent loss of these vesicles. Conversely, expression of the human gene encoding CHC22 in mice resulted in abnormal glucose homeostasis. The number of GLUT4 storage vesicles appears to be increased. GLUT4 was poorly mobilized by insulin, as seen in humans with type 2 diabetes (see the figure).

The findings by Vassilopoulos *et al.* highlight the possibility that altered vesicle trafficking may contribute to diabetes pathophysiology, independent of impaired insulin signaling (*8*). CHC22-coated storage vesicles harboring GLUT4 may not be targeted by the insulin signal. In cultured adipocytes, for example, only about two-thirds of GLUT4 storage vesicles are translocated by maximal insulin stimulation (*5*, *9*). Why is the remainder of this pool inaccessible? Is the fraction of GLUT4 that can be translocated reduced in diabetes? Alternatively, GLUT4 may accumulate within storage vesicles because of deficient insulin signaling. Yet in the mice engineered to express human CHC22, activation of Akt, an enzyme that functions in one signaling pathway that affects GLUT4 translocation, was increased. Other signaling pathways have been implicated in both GLUT4 translocation and diabetes pathophysiology, and these may be important to mobilize vesicles formed by CHC22 (*10*–*12*).

Muscle contraction causes translocation of GLUT4 to enhance glucose uptake, similar to insulin, and it will be interesting to learn if the vesicles that are mobilized also have CHC22 coats. CHC22 appears to participate in trafficking GLUT4 storage vesicles in adipose tissue as well as in muscle. An interesting possibility is that GLUT4 storage vesicles shed their clathrin coats in response to insulin. Finally, it will be important to learn if obesity alters the abundance of CHC22 or of associated proteins, in either muscle or adipose tissue, to curtail insulin action.

Most studies on GLUT4 trafficking have used rodent models, in which CHC17 forms the GLUT4 storage vesicles (*13*, *14*), so it's not clear how CHC22 functions differently. For example, why does the diabetes-like phenotype of mice that express CHC22 become apparent with age; would a high-fat diet exacerbate the defect? A detailed analysis of the metabolic phenotype in these mice, possibly by infusions of insulin and a glucose tracer, should provide some answers.

The work of Vassilopoulos *et al.* highlights the differences between humans and mice by providing the first molecular identification of human-specific GLUT4 sorting. The exact role CHC22 may play in human diabetes remains uncertain. However, this work reminds investigators studying diabetes pathogenesis that it is important to think outside the vesicle.

References

1. D. B. Savage, K. F. Petersen, G. I. Shulman, *Physiol. Rev.* **87**, 507 (2007).
2. S. Huang, M. P. Czech, *Cell Metab.* **5**, 237 (2007).
3. M. Larance, G. Ramm, D. E. James, *Mol. Endocrinol.* **22**, 226 (2008).
4. S. Vassilopoulos *et al.*, *Science* **324**, 1192 (2009).
5. B. R. Rubin, J. S. Bogan, *Vitam. Horm.* **80**, 155 (2009).
6. P. F. Pilch, *Acta Physiol. (Oxf.)* **192**, 89 (2008).
7. M. A. Edeling, C. Smith, D. Owen, *Nat. Rev. Mol. Cell Biol.* **7**, 32 (2006).
8. W. T. Garvey *et al.*, *J. Clin. Invest.* **101**, 2377 (1998).
9. R. Govers, A. C. Coster, D. E. James, *Mol. Cell. Biol.* **24**, 6456 (2004).
10. L. Chang, S. H. Chiang, A. R. Saltiel, *Endocrinology* **148**, 27 (2007).
11. K. L. Hoehn *et al.*, *Cell Metab.* **7**, 421 (2008).
12. K. J. Gaulton *et al.*, *Diabetes* **57**, 3136 (2008).
13. D. J. Fazakerley, S. P. Lawrence, V. A. Lizunov, S. W. Cushman, G. D. Holman, *J. Cell Sci.* **122**, 727 (2009).
14. J. Li *et al.*, *J. Cell Biol.* **178**, 453 (2007).

10.1126/science.1174841

CELL BIOLOGY

It's the DNA That Counts

A simple genetic circuit that counts molecular events may be further developed to program complex cell behaviors.

Christina D. Smolke

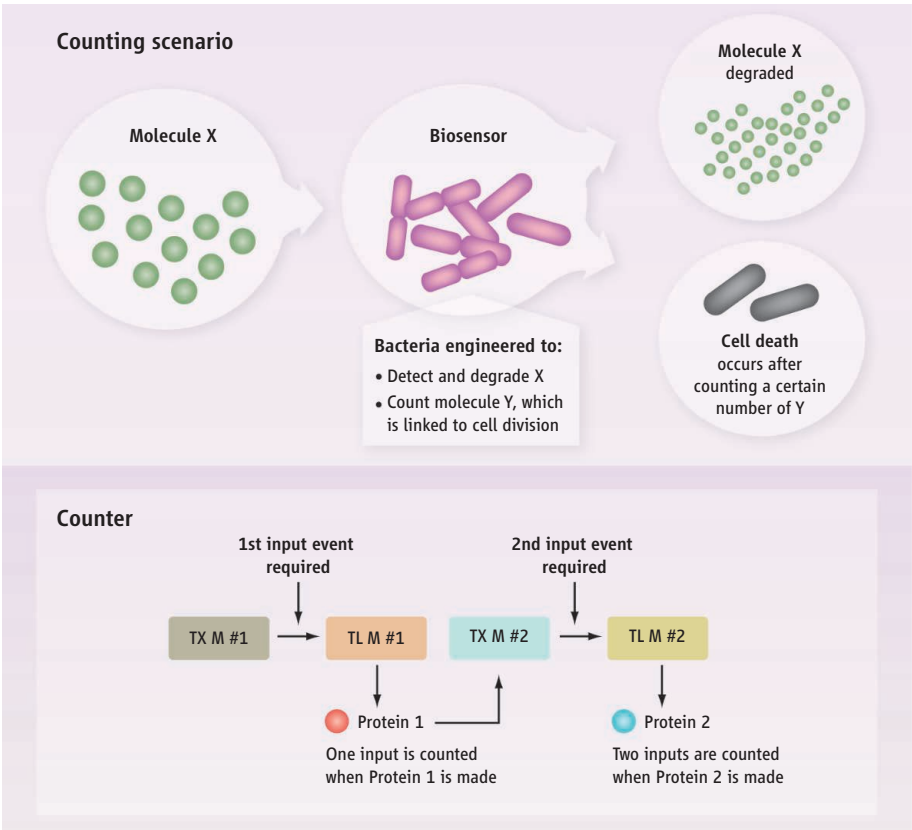
Natural biological systems have evolved genetic programs that control complex activities through the coordinated processing of signals received from their environments. The engineering of synthetic biological systems to perform programmed information processing and computational functions has remained a challenge. Counters represent one class of information-processing systems and can be used to trigger events in response to a series of detected signals that are integrated and processed over time. Engineered biological counters would enable many applications, such as regulating cell death after a specified number of cell division cycles, controlling cell differentiation in response to temporal cues, noninvasive monitoring of aging, and recording the frequency of environmental events. On page 1199 in this issue, Friedland *et al.* (*1*) report an important step toward the construction of genetically encoded counters.

Counters can be assembled from a variety of simpler functions, such as signal detection and processing (the ability to respond to an input signal), time delay (the ability to integrate signals and trigger events after a delay from the initial detection event), and memory (the ability to remember and track earlier detection events). Genetic circuits that encode these basic operations have been demonstrated, including systems that use both protein-based transcriptional regulators (*2*–*4*) and RNA-based posttranscriptional regulators (*5*–*7*). For example, time delays have been encoded in cascades of transcription (the production of RNA from corresponding DNA) that trigger the production (or repression) of a protein (via translation of the corresponding RNA) upon the detection of an initial signal. This event subsequently activates the production (or repression) of the next protein. A linear sequence of such events, in the form of repeating modules, can be designed in which the last module triggers the production of the desired protein output (*8*, *9*).

Transcriptional cascades introduce delays in triggering the final protein output through the latency associated with expressing each intermediate protein in the series. Such cascading systems exhibit other properties including signal amplification, signal filtering, sensitivity to detecting the signal, and modulation of variation across cell populations (*8*). Other genetic mechanisms for introducing time delays in protein production are based on increasing the time associated with RNA processing steps (*10*).

Memory has been encoded in genetic circuits by means of feedback loops that lock a system in one state following a signal that sets that state. For example, a circuit based on the incorporation of two mutually inhibitory tran-

Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. E-mail: csmolke@stanford.edu



Why count, and how? (**Top**) In the hypothetical counting scenario shown, bacteria can be engineered to sense molecule X and degrade it (such as in bioremediation). At the same time, as a safety mechanism, the bacteria are engineered to count molecule Y, which is associated with cell division. After counting a certain number of molecule Y, the bacteria execute a cell death function. (**Bottom**) An example of a simple genetic counter that counts up to two, based on a conditional transcriptional cascade. The cascade includes transcription modules (TX M) and translation modules (TL M).

scriptional negative-feedback loops acted as a genetic toggle switch, in which transient input signals moved the system between two states (*11*). Each state was associated with the production of a protein that inhibited the system from moving into the other state unless the appropriate input signal was applied. In a different memory circuit, based on an autoregulatory transcriptional positive-feedback loop (*12*), a transient input signal triggered the production of a protein that subsequently activated its own production, such that the system remained in the activated state after removal of the signal.

Early attempts to build heritable memory systems that do not require the sustained production of proteins were based on enzymatic mechanisms that allow the system state to be written directly into the structure of DNA (*13*). To date, these systems have been developed with recombinases, enzymes that can invert, insert, or remove specified DNA sequences and thus dynamically rewrite genetic programs in response to specified signals. For example, a genetic circuit based on two overlapping inversion systems encoded multiple output states depending on the order in which recombinases processed the DNA (*14*).

Friedland *et al.* take important steps to build biological counters by integrating these functional operations. The authors propose two different circuit architectures for encoding counters that trigger the expression of a desired protein following the processing of two or three input signal pulses. Each counter combines a time-delay operation (triggered by the initial detection event—setting the "1" state) with conditional regulation linked to the immediate detection of subsequent signals (allowing the counting of additional detection events "2" and "3"). In one system, the time delay is encoded through a transcriptional cascade operation. Conditional regulation is achieved through an RNA molecule that is expressed only when the input signal is present. The expression of this RNA is under the control of a protein-based transcriptional regulator and is required for the expression of the protein output from the transcriptional cascade modules. Because of the delay associated with protein accumulation from each transcription module and difference in decay rates between the protein and RNA components, the RNA regulator decays before a sufficient amount of the protein output from the previous module has accumulated to trigger activation of the next module. This resets the conditional regulation operation and allows counting of input pulses (see the figure).

In their second system, the time-delay operation is encoded within a heritable memory cascade. Each memory module encodes a protein that flips a segment of DNA within that module to turn off its own expression and primes the next module to be activated by the input signal by correctly orienting a protein-based transcriptional regulatory element. The time required for each memory module to prime the next module is longer than the signal pulse length such that counting is achieved. By incorporating memory modules, this second architecture allows signals to be integrated over longer time frames, as the priming of the module is "hard-coded" into the DNA. In addition, this architecture could allow identical or different input signals to be counted via the choice of each module's conditional regulatory element.

Friedland *et al.* provide examples from which to build more sophisticated counters. An important next step will be to develop circuits that can report on intermediate states, in addition to the final state. Another critical counter property will require the development of circuits that allow counters to distinguish between continuous and transient signals. Given the broad applications of counters that operate inside living cells, the continued development of next-generation genetically encoded counters will be of critical importance to synthetic biology (*15*).

References

1. A. E. Friedland *et al.*, *Science* **324**, 1199 (2009).
2. C. C. Guet, M. B. Elowitz, W. Hsing, S. Leibler, *Nature* **296**, 1466 (2002).
3. J. C. Anderson, C. A. Voigt, A. P. Arkin, *Mol. Syst. Biol.* **3,** 133 (2007).
4. R. S. Cox III, M. G. Surette, M. B. Elowitz, *Mol. Syst. Biol.* **3**, 145 (2007).
5. F. J. Isaacs *et al.*, *Nat. Biotechnol.* **22**, 841 (2004).
6. M. N. Win, C. D. Smolke, *Science* **322**, 456 (2008).
7. M. N. Win, J. C. Liang, C. D. Smolke, *Chem. Biol.* **16**, 298 (2009).
8. S. Hooshangi, S. Thiberge, R. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3581 (2005).
9. W. Weber, B. P. Kramer, M. Fussenegger, *Biotechnol. Bioeng.* **98,** 894 (2007).
10. I. A. Swinburne *et al.*, *Genes Dev.* **22**, 2342 (2008).
11. T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **403**, 339 (2000).
12. C. M. Ajo-Franklin *et al.*, *Genes Dev.* **21**, 2271 (2007).
13. T. S. Ham, S. K. Lee, J. D. Keasling, A. P. Arkin, *Biotechnol. Bioeng.* **94**, 1 (2006).
14. T. S. Ham *et al. PLoS One* **3**, e2815 (2008).
15. D. Endy, *Nature* **438**, 449 (2005).

10.1126/science.1174843

CREDIT: N. KEVITYAGALA/*SCIENCE*