**Supplementary Information:**
**The timing of transcriptional regulation in synthetic gene circuits**

Yu-Yu Cheng[1], Andrew J. Hirning[1], Krešimir Josić[1,2,3], and Matthew R. Bennett[1,4,†]

[1]Department of Biosciences, Rice University, Houston, TX 77005, USA.
[2]Department of Mathematics, University of Houston, Houston, TX 77204, USA.
[3]Department of Biosciences, University of Houston, Houston, TX 77204, USA.
[4]Department of Bioengineering, Rice University, Houston, TX 77005, USA.
[†]Corresponding author. E-mail: matthew.bennett@rice.edu

## Materials and Methods

### Plasmids

All circuits were constructed on plasmids (Fig. S1). *araC* and *lacI* are tagged with the LAA version of the *ssrA* degradation tag[1] in all circuit designs, except for *araC* in the $P_{Tet}$-*araC* circuit (Fig. S1D and S1F). *sfyfp* is also tagged with the LAA version of *ssrA* in the repression circuit, the cascade, the $P_{LlacO1}$-*sfyfp* reporter-only circuit, and the $P_{lac/ara-1}$-*sfyfp* reporter-only circuit (Fig. S1B, S1D, S1E and S1G). We used a bicistronic ribosome binding site[2] (BCD) for all genes, except for *araC* in the activation circuit (Fig. S1A), for which we used the B0034 ribosome binding site (http://partsregistry.org/Part:BBa_B0034). All circuits were constructed on plasmids with the p15A origin, except for $P_{Tet}$-*araC*, which is on a plasmid with the pSC101 origin (Fig. S1D and S1F).

### Strain construction

The three *E. coli* strains used in this study (JS006LT, JS006A, and JS006T) were derived from the JS006 strain (BW25113 *ΔlacIΔaraC*)[3]. The activation circuit (Fig. S1A) and $P_{LlacO1}$-*sfyfp* reporter-only circuit (Fig. S1E) were transformed into JS006LT, which contains constitutively expressed *lacI* and *tetR*. The repression circuit (Fig. S1B), $P_{BAD}$-*sfyfp* reporter-only circuit (Fig. S1C), and $P_{lac/ara-1}$-*sfyfp* reporter-only circuit (Fig. S1G) were transformed into JS006A, which contains constitutively expressed *araC*. The two-step genetic cascade (Fig. S1D) and the modified activation circuit (Fig. S1F) were transformed into JS006T, which contains constitutively expressed *tetR*. The genes were knocked into the genome via the lambda integrase method[4].

### Microfluidic device and microscopy setup

We used microfluidic devices to observe single cells over time with fluorescence

microscopy. Master molds and microfluidic chips were created with standard photolithography and soft lithography techniques[5]. The 'dial-a-wave' chip was designed to fine-tune the culture medium constituents for the cells (Fig. S2A). This device was created by integrating the dial-a-wave junction from Bennett *et al.* 2008[6] with the microfluidic chip used in Hussain *et al.* 2014[7]. The dye sulforhodamine 101 (Sigma-Aldrich) was used (at 1 $\mu$g/ml) to track the concentration of the inducers. The typical dye trajectory of the step function (Fig. S2B) shows that it needs 5 minutes for the inducer to reach maximal concentration. We defined the first time point at which inducer started to increase as time 0. We found that for *yfp* expressed from strong promoters, such as $P_{lac/ara}$ promoter, the *yfp* signal can be detected within 1 to 2 minutes (Fig. S2C).

For each microscopy experiment, cells were first loaded into the 'dial-a-wave' chip. When about 30 cells were trapped in the cell trapping region, phase contrast (100x magnification with additional 1.5x magnification) and fluorescence images (mCherry and YFP channels) were taken every 1 minute. The power of the excitation light was 20% power of a 200W light source (Lumen 200 Pro). The exposure time was 300ms and 100ms for sfYFP and mCherry, respectively. Cells remained healthy after 8 hours of image acquisition. The inducers were loaded into the chip after 10 minutes. We used the 10-minute background YFP signal to define the threshold for the activation time. Each experiment was performed at least twice to get a minimum of 60 single-cell trajectories for the statistics.

Image analysis and data acquisition

Images were analyzed using a semi-automatic tracking algorithm[5] developed using MATLAB (Mathworks, Inc.). Cells were manually segmented using phase-contrast images. The results of the segmentations were used by the algorithm to record fluorescence at the regions identified as individual cells. The output fluorescence is the summed fluorescence intensity divided by the number of pixels in the segmented cell region. When the cells divided, we randomly picked one of the daughter cells to follow. The script is available for download online (github.com/alanavc/rodtracker).

Promoter characterization

Overnight cultures of the reporter-only strains – $P_{BAD}$-*yfp* reporter-only circuit in JS006A, $P_{lac}$-*yfp* reporter-only circuit in JS006LT, and $P_{lac/ara}$-*yfp* reporter-only circuit in JS006A, were diluted 1:1000 into 96-well plates with varying levels of inducers. The cells were then grown at 37°C for 2 hours. Fluorescence and $OD_{600}$ were measured with an Infinite® 200 PRO fluorescence plate reader (TECAN). Each data point represents the average of

three measurements (Fig. 1B and 2A). Data is normalized because the $P_{lac/ara-1}$ promoter is approximately 10 times stronger than the $P_{BAD}$ promoter.

Threshold for the estimation of signaling time of the activation circuit

To estimate the time at which the YFP signal starts to increase for the activation circuit, we defined the threshold to be the mean plus N standard deviations of the 10-minute background fluorescence signal, where N is a positive integer. We found that when N<4, the estimated signaling time for some cells is shorter than the YFP maturation time. Figure S3 shows the case for N=3, which lead to an estimate of less than one minute for the expression of YFP in some cells. The YFP maturation time was estimated using the reporter-only circuit and found to be at least 4 minutes. Therefore, we chose N=4 to avoid underestimating the signaling time.

Estimation of the signaling time in the two-step genetic cascade

The measured times $T_1$, $T_y$, and $T_2$ were used to estimate the total signaling time, $T_{TOT}$, in the cascade (Fig. 3C). We first assumed that each of the measured times are independent random variables. To estimate the distribution of the total signaling time, $T_{TOT}$ (Fig. 3D), we sampled three times independently from the collection of measured values, $T_{1,i}$, $T_{y,i}$, and $T_{2,i}$, from the measured values of $T_1$, $T_y$, and $T_2$. We estimated the total signaling time from these three sampled values as $T_{TOT,i} = T_{1,i} - T_{y,i} + T_{2,i}$, and repeated the process 100,000 times (with replacement). The mean of distribution is very close (within 1 minute) to that found experimentally. However, the estimated standard deviation is slightly larger than what was measured (2.4 minutes versus 1.3 minutes). This might due to correlations between the three measurements. For example, *araC* and *lacI* expression could be negatively correlated since they compete for the same resources[8]. This will lead to a negative correlation between $T_1$ and $T_2$.

To take into account the possible correlation between $T_1$ and $T_2$, we first assume that they are normally distributed for simplicity. We fit the measurements of $T_1$ and $T_2$ assuming normality to obtain the mean $\mu_1$ and $\mu_2$ and standard deviation $\sigma_1$ and $\sigma_2$. Then $T_{1+2}$ can be assumed to follow a normal distribution with mean $\mu_1$ and standard deviation $\sqrt{{\sigma_1}^2 + {\sigma_2}^2 - 2\rho * \sigma_1 * \sigma_2}$, where $\rho$ is the correlation coefficient between $T_1$ and $T_2$. The same Monte Carlo method was used to calculate $T_{TOT}$ by subtracting $T_y$ from the newly estimated $T_{1+2}$. When $\rho < 0$, the standard deviation of estimated $T_{TOT}$ is smaller, with the mean not changed. Figure S4 shows the case for $\rho = -0.5$ for illustration.

**References**

1 Andersen, J. B. *et al.* New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240-2246 (1998).

2 Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Meth* **10**, 354-360 (2013).

3 Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-519 (2008).

4 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640-6645 (2000).

5 Veliz-Cuba, A. *et al.* Sources of Variability in a Synthetic Gene Oscillator. *PLoS Comput Biol* **11**, e1004674 (2015).

6 Bennett, M. R. *et al.* Metabolic gene regulation in a dynamically changing environment. *Nature* **454**, 1119-1122 (2008).

7 Hussain, F. *et al.* Engineered temperature compensation in a synthetic genetic clock. *Proc Natl Acad Sci USA* **111**, 972-977 (2014).

8 Qian, Y., Huang, H.-H., Jiménez, J. I. & Del Vecchio, D. Resource Competition Shapes the Response of Genetic Circuits. *Acs Synth Biol*, Epub ahead of print, doi:10.1021/acssynbio.6b00361 (2017).

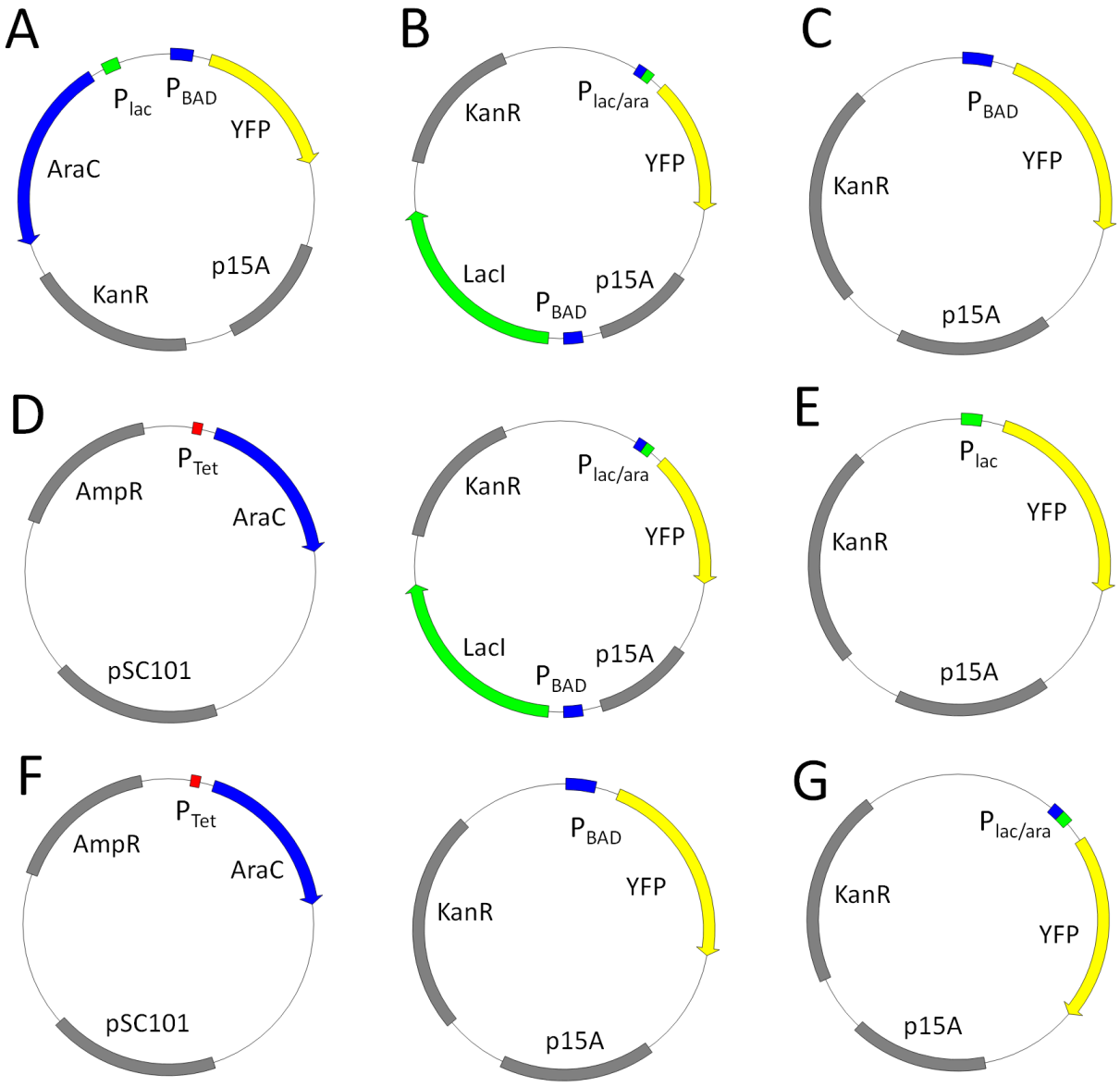


Figure. S1. Plasmids built for each genetic circuit. (A) Activation circuit. (B) Repression circuit. (C) $P_{BAD}$ reporter only circuit. (D) Two-step genetic cascade. (E) $P_{lac}$ reporter only circuit. (F) Modified activation circuit. (G) $P_{lac/ara}$ reporter only circuit. Plasmid maps were drawn with Savvy (http://bioinformatics.org/savvy/).

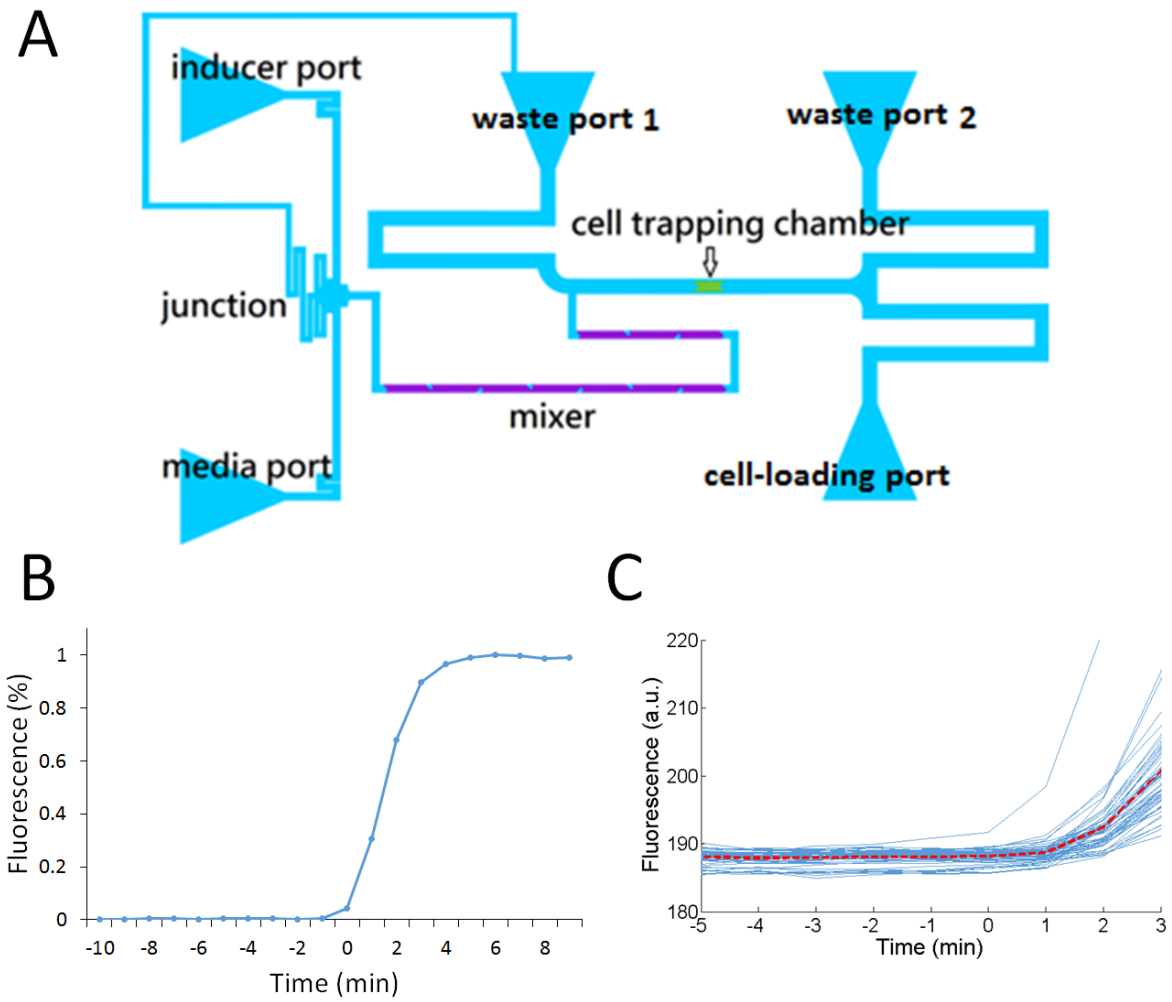


Figure. S2. The 'dial-a-wave' chip and its characterization. (A) The chip has 5 ports connected to external culture media. The media port is for LB media. The inducer port is for inducers and the dye Sulforhodamine 101. The two waste ports are for the outlet of the culture media. The junction and mixer are designed to fine-tune the inducer concentration with more precision. The cell-trapping chamber is 100 μm wide, 300 μm long, and 0.95 μm high. (B) The typical dye trajectory in the microfluidic experiments. It shows that inducers need 5 minutes to reach maximum concentration in the chamber after loading. Time 0 is defined at the point that the fluorescence starts to increase. (C) YFP trajectories for the repression circuit at 2% ARA (Figure 2B). YFP increase can be detected 1 to 2 minutes after inducer concentration starts to increase. The dashed red line is the average fluorescence.

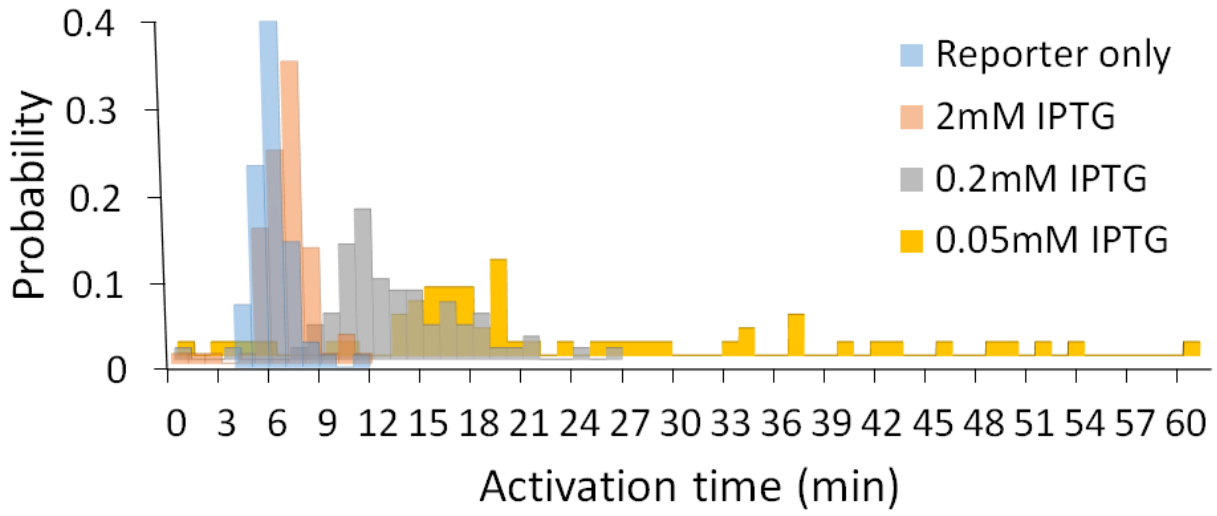


Figure S3. The estimated signaling time for the activation circuit when the threshold is defined as the mean plus three standard deviations of the background fluorescence signals. Some cells are estimated to express YFP in less than the estimated maturation time.

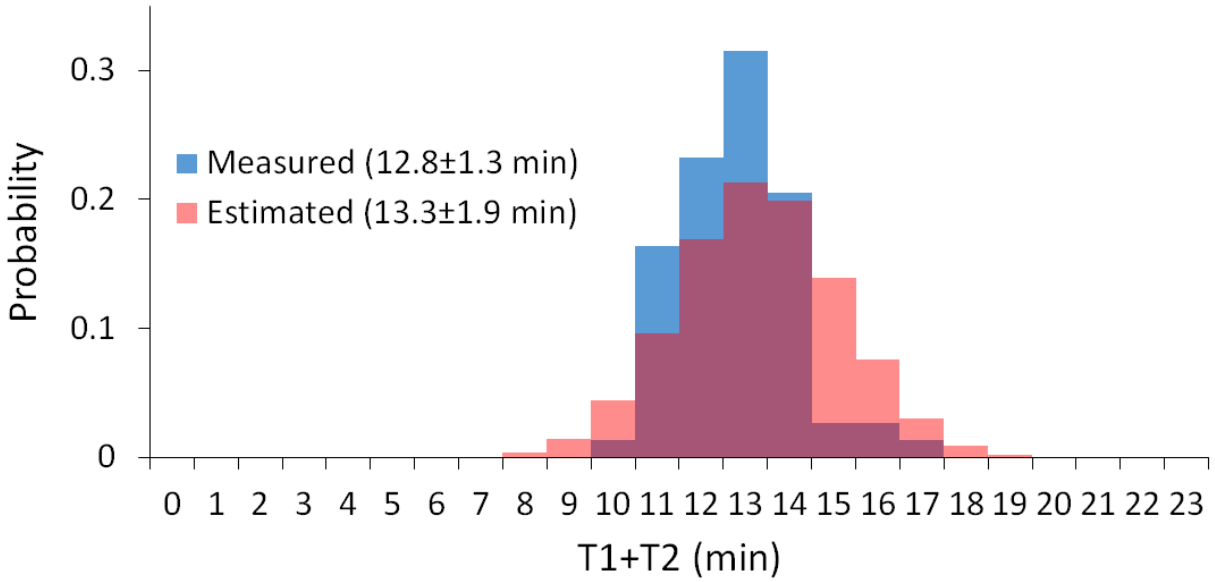


Figure S4. The comparison of the estimated signaling time with the measured signaling time of the cascade when $T_1$ and $T_2$ have correlation coefficient -0.5. The standard deviation of the estimated distribution is reduced from 2.4 minutes (Fig. 6C) to 1.9 minutes, with the mean not changed.