Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals

Evan J Olson[1], Lucas A Hartsough[2], Brian P Landry[2], Raghav Shroff[2] & Jeffrey J Tabor[2,3]

Gene circuits are dynamical systems that regulate cellular behaviors, often using protein signals as inputs and outputs. Here we have developed an optogenetic 'function generator' method for programming tailor-made gene expression signals in live bacterial cells. We designed precomputed light sequences based on experimentally calibrated mathematical models of light-switchable two-component systems and used them to drive intracellular protein levels to match user-defined reference time courses. We used this approach to generate accelerated and linearized dynamics, sinusoidal oscillations with desired amplitudes and periods, and a complex waveform, all with unprecedented accuracy and precision. We also combined the function generator with a dual fluorescent protein reporter system, analogous to a dual-channel oscilloscope, to reveal that a synthetic repressible promoter linearly transforms repressor signals with an approximate 7-min delay. Our approach will enable a new generation of dynamical analyses of synthetic and natural gene circuits, providing an essential step toward the predictive design and rigorous understanding of biological systems.

To characterize the signal processing properties of a circuit, one must measure how it transforms a wide range of dynamical inputs into outputs. Recently, gene circuits have been dynamically characterized by using microfluidic devices to create step changes and waveforms of extracellular effector molecules while simultaneously monitoring intracellular responses[1–6]. There are two fundamental limitations to this approach. First, in the absence of a transmembrane receptor, chemicals must diffuse or be transported across cellular membranes before affecting the circuit under study. Transport processes introduce unknown delays, which confound circuit analysis, and low-pass filtering, which slows the timescale over which a signal can change[1,5]. Second, using chemical approaches, one can analyze gene circuits only with convenient effectors such as sugars or osmolytes. Alternatively, an optical method for creating on-demand protein signals in live cells would bypass these limitations and, in principle, enable the dynamical characterization of virtually any gene circuit that responds to changes in protein concentration.

Recently, *in silico* feedback control has been used to set and hold a desired protein expression level[7] and create basic expression dynamics[8] in *Saccharomyces cerevisiae*. In these studies, light- and sorbitol-responsive promoters were used to control expression of a fluorescent reporter protein, which was measured by flow cytometry or fluorescence microscopy. Experimental characterization data were collected to calibrate mathematical models, which were in turn used to design a series of input pulses to drive protein expression to follow desired reference signals. The experimental data were observed to deviate substantially from the model predictions, and a feedback controller was used to improve the fidelity of *in vivo* control. Nonetheless, the accuracy and degree of gene expression programmability of these methods remains limited[7,8]. We hypothesized that the combination of continuously applied analog light inputs and rapidly photoreversible promoters would permit the development of quantitatively predictive mathematical models, which would in turn enable more reliable programming and more complex gene expression signals.

Previously, we engineered two *Escherichia coli* two-component systems (TCSs) wherein transcription from an output promoter is controlled by different activating and inhibitory light wavelengths[9,10] (**Fig. 1a**,**b**, Online Methods and **Supplementary Fig. 1**). Each TCS comprises a light-switchable sensor histidine kinase (SK) containing an N-terminal phytochrome-family photosensory domain and a C-terminal bifunctional kinase-phosphatase signaling domain. In the first TCS (hereafter 'CcaS-CcaR'), the SK CcaS is produced in a green-absorbing ground state, termed Pg. Absorption of green light flips CcaS to a kinase-active red-absorbing state (Pr) that phosphorylates the response regulator CcaR, which then binds to the *cpcG2* promoter and activates transcription (**Fig. 1a**). Absorption of red light switches CcaS Pr back to Pg, which dephosphorylates phospho-CcaR, deactivating transcription (**Fig. 1a**). In the second TCS (hereafter 'Cph8-OmpR'), the SK Cph8 is produced in a kinase-active Pr state that phosphorylates the response regulator OmpR, activating transcription from the *ompC* promoter (**Fig. 1b**). Red light switches Cph8 Pr to a far red–absorbing state (Pfr), which dephosphorylates phospho-OmpR, deactivating transcription (**Fig. 1b**). In this study, we systematically characterized the steady state and dynamical properties of these light sensors in response to analog

[1]Graduate Program in Applied Physics, Rice University, Houston, Texas, USA. [2]Department of Bioengineering, Rice University, Houston, Texas, USA. [3]Department of Biochemistry and Cell Biology, Rice University, Houston, Texas, USA. Correspondence should be addressed to J.J.T. (jeff.tabor@rice.edu).

RECEIVED 15 AUGUST 2013; ACCEPTED 23 JANUARY 2014; PUBLISHED ONLINE 9 MARCH 2014; DOI:10.1038/NMETH.2884

light intensities and developed a predictive mathematical model that we used to program custom gene expression dynamics. We then applied this approach to control the expression dynamics of the transcriptional repressor TetR and incorporated two fluorescent protein probes to characterize the input-output signal processing properties of a widely used TetR-repressible promoter (PLtetO-1; ref. 11) in live *E. coli* cells.

RESULTS

Light tube array design and gene expression analysis

To control gene expression dynamics with light, we designed and constructed the light tube array (LTA), an instrument that contains independently programmable blue, green, red and far-red LEDs (**Fig. 1c**) that deliver calibrated intensities of each wavelength to each of 64 standard test tubes (**Fig. 1d** and **Supplementary Figs. 2** and **3**). Housed in a shaking incubator, the LTA allows optical signals to be sent into batch cultures undergoing exponential growth. We characterized expression from each of the two light-sensor output promoters using superfolder GFP[12] (sfGFP) (**Fig. 1a**,**b**). Bacteria were grown in the LTA under a constant or dynamical (**Fig. 1e**) light signal (Online Methods and **Supplementary Note 1**). At the end of the experiment, bacteria were harvested and placed in a transcriptional inhibitor solution, and sfGFP maturation was allowed to go to completion before measurement by flow cytometry (**Supplementary Note 2**).

Characterization and modeling of light sensors

We first used the LTA to characterize the relationship between light input and protein expression output for both sensors (i.e., their steady-state transfer functions). Our measurements revealed that expression from CcaS-CcaR increases with green light intensity up to the LTA maximum of 4.03 W/m^2 (**Fig. 2a**) with a response that is fit well by a Hill function (Hill coefficient $n = 2.8 \pm 0.4$; values are ± the standard error of the fit) containing an additional linear term (**Supplementary Fig. 4**). Measurements also revealed that red light competitively inhibits activation by green light (**Supplementary Fig. 5**). This feature allowed us to reduce the sensitivity of CcaS-CcaR to green light (half-maximal response, k, was raised from 0.010 W/m^2 to 0.13 W/m^2) while preserving the full output range (**Fig. 2a** and **Supplementary Note 3**). The Cph8-OmpR transfer function differs from that of CcaS-CcaR in several ways. First, gene expression output decreases with the intensity of red light in a manner described by a standard Hill function ($n = 1.4 \pm 0.1$; **Fig. 2b** and **Supplementary Fig. 6**). Additionally, far-red light has a minimal competitive effect, likely owing to rapid dark reversion of Cph8 from the Pfr to Pr state (**Fig. 1e** and **Supplementary Fig. 7**). Thus, the CcaS-CcaR system is a green:red light–ratio sensor, and the Cph8-OmpR system is a red light–intensity sensor.

We next applied step increases and decreases of the control signal (green or red intensity, respectively) to cultures preconditioned to low, intermediate and high light sensor expression levels. The resulting sfGFP time courses (**Fig. 2c**,**d** and **Supplementary Figs. 8** and **9**) revealed three timescales underlying response dynamics. The dominant timescale arises from sfGFP dilution during cell growth and division, an expected result for a stable protein[13]. In addition, we observed a gene expression switching time, during which the sfGFP production rate transitions from the initial to final level. For step increases in gene expression from CcaS-CcaR, the switching time depends on the final green intensity, requiring 1–28 min for the production rate to transition to 50% of its final value (**Fig. 2c** and **Supplementary Fig. 8**). In contrast, the production rate of Cph8-OmpR has a constant switching timescale of 4 min for step increases in expression, regardless of the final red intensity (**Supplementary Fig. 9**). For step decreases in gene expression, the production rate from CcaS-CcaR switches halfway in 10 min, whereas Cph8-OmpR switches as quickly as 4 min in a manner dependent on the final red intensity (**Fig. 2d**

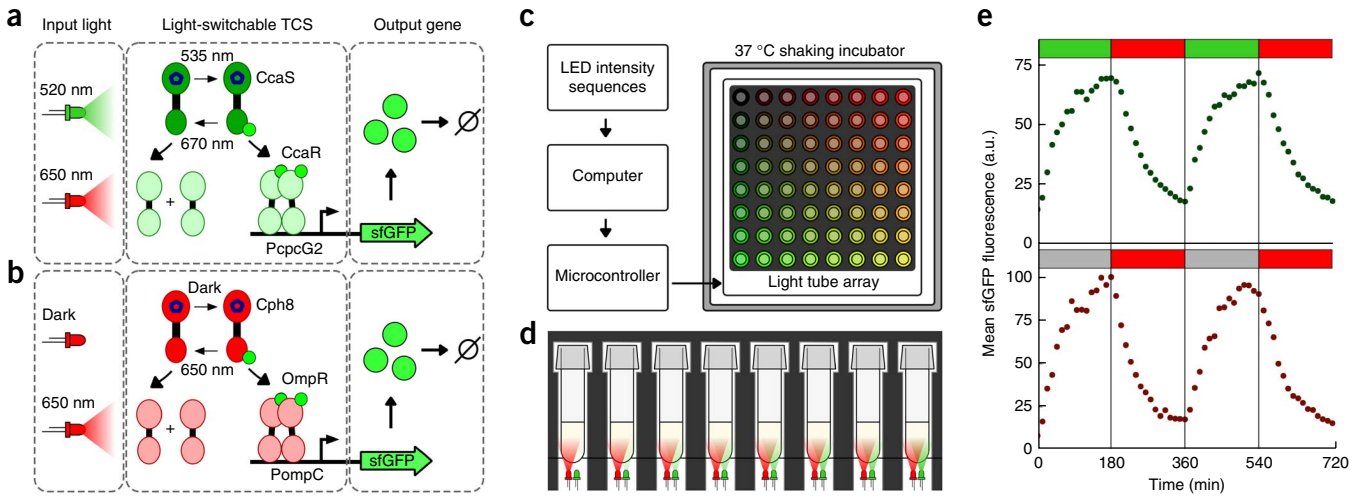

Figure 1 | Light-switchable two-component systems (TCSs) and light tube array (LTA). (**a**) Expression of sfGFP from the green-activated (λ_{max} = 535 nm), red-deactivated (λ_{max} = 670 nm) CcaS-CcaR two-component system is controlled by modulating the intensity of a green LED (λ_{max} = 520 nm) while a red LED (λ_{max} = 650 nm) is maintained at high intensity. (**b**) Expression from the dark-activated, red-deactivated (λ_{max} = 650 nm) Cph8-OmpR TCS is controlled by the intensity of the red LED. (**c**) Schematic of the LTA. An array of individually controlled LEDs is used to deliver programmed light inputs to exponentially growing bacterial cultures in a shaking incubator. (**d**) Each culture tube is optically isolated with opaque foam. (**e**) Cells preconditioned into the low-expression state were grown in the LTA and exposed to a series of 180-min light inputs (green, 4.05 W/m^2 emission from green LED; red, 1.05 W/m^2 from red LED; gray, LEDs off). Data points represent population means of single-cell fluorescence distributions (N = 2,000–5,000 cells) collected by flow cytometry.

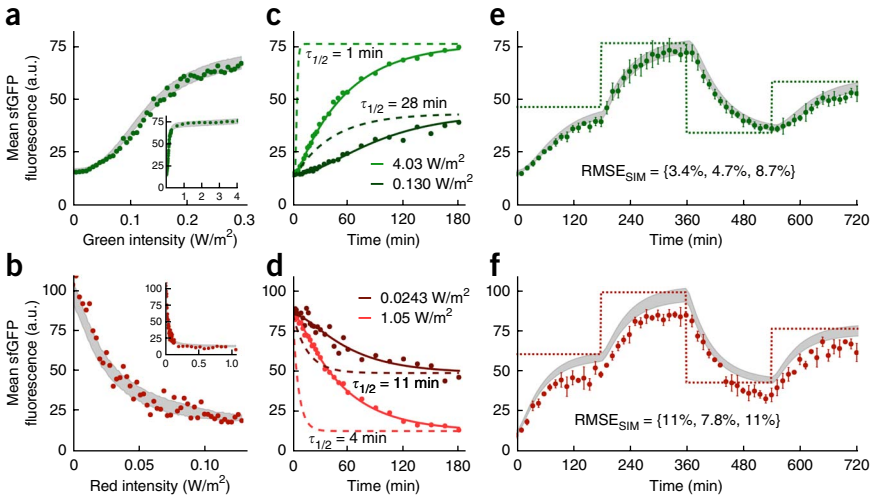

Figure 2 | Experimentally characterized TCS models predict input-output dynamics. (**a**,**b**) Steady-state protein expression from the CcaS-CcaR (**a**) and Cph8-OmpR (**b**) systems with increasing intensity of input light. Red light was held at 1.05 W/m^2 for the CcaS-CcaR system for all experiments. The mathematical model (gray) is shown as the interquartile range of $N = 500$ simulations produced by sampling the measured parameter uncertainties. Insets show the full intensity range available using the LTA. (**c**,**d**) TCS responses to step increases of the control wavelength from 0 to the intensities shown. The kinetic model is parameterized via a least-squares fit of the sfGFP fluorescence (solid) and production rate (dash) variables, with the switching time of production rate indicated ($\tau_{1/2}$). (**e**,**f**) Model validation comparing predicted response (gray) to a series of step changes in light input with experimental data. The input light signal (dotted) is shown converted to fluorescence units using the steady-state response of the system as measured in **a**,**b**. Markers and error bars represent the mean and s.d. of experiments on three separate days. Simulated region is determined as in **a**,**b**. r.m.s. errors ($RMSE_{SIM}$) compare predictions and experiments for each day and are expressed as a percentage of the full output range of each system.

and **Supplementary Fig. 8**). Finally, the third timescale observed in CcaS-CcaR, but not Cph8-OmpR, is an approximate 5-min delay before the sfGFP production rate begins to shift in response to a change in light input (**Supplementary Fig. 8**).

On the basis of these step-response data, we developed a two-dimensional ordinary differential equation model of the dynamics of the light sensors. The model incorporates two dynamic variables and three parameters describing the three observed timescales.

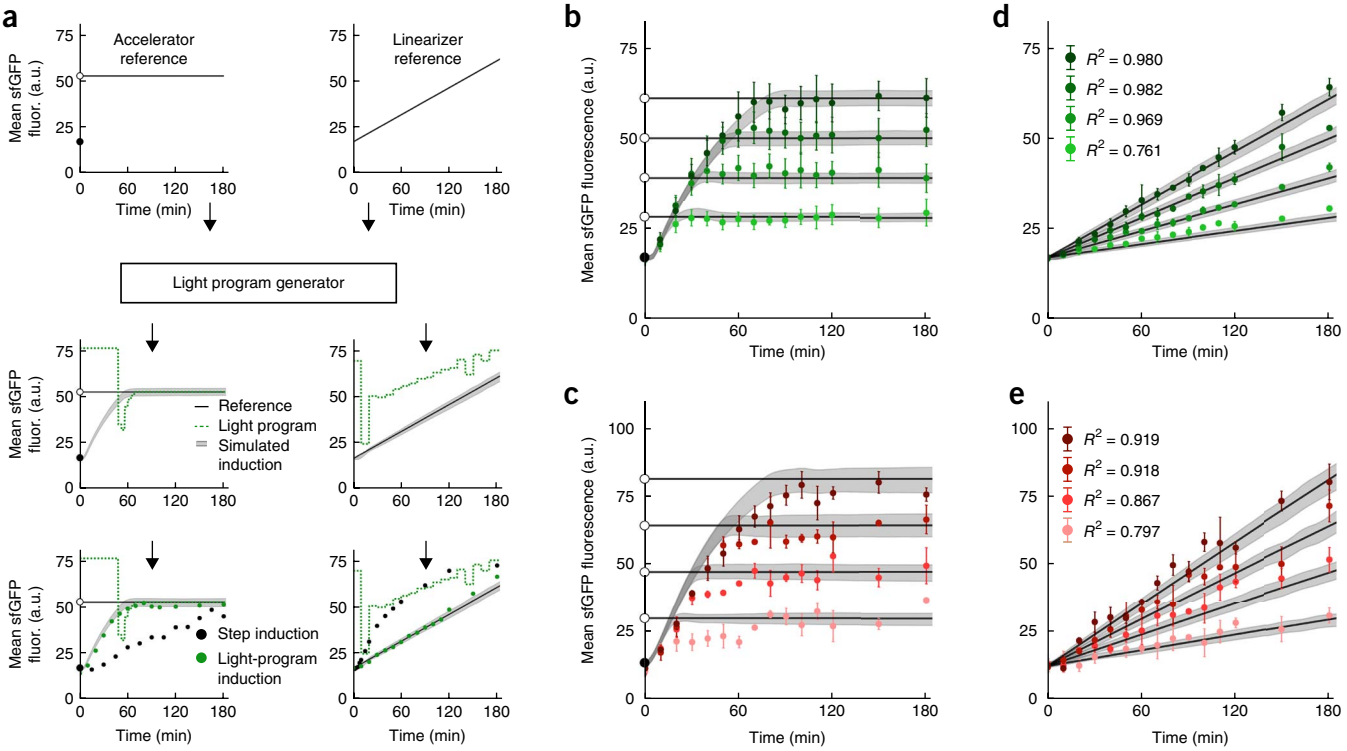

Figure 3 | Programming protein transition dynamics with a biological function generator. (**a**) A desired protein reference signal is used as the input to the algorithm that generates the light program (top row). Left, the reference signal transitions discontinuously from a low expression output (solid circle) to a high expression level (open circle, black line) at time $t = 0$. Right, the reference signal transitions linearly from low to high. These reference signals are used to compute programmed sequences of step changes in light intensity that minimize the square error between the predicted and reference protein expression signal (center row). The light programs are depicted as in **Figure 2**. Experimental protein-expression response of the CcaS-CcaR system (green dots) to the precomputed light sequence is shown alongside basic step input–driven transitions (black dots) (bottom row). (**b**–**e**) Accelerated (**b**,**c**) and linearized (**d**,**e**) transitions are shown with reference signals and predicted trajectories. R^2 values were calculated by comparing data to the linear reference signal. Error bars and simulations are as in **Figure 2**. Fluor., fluorescence.

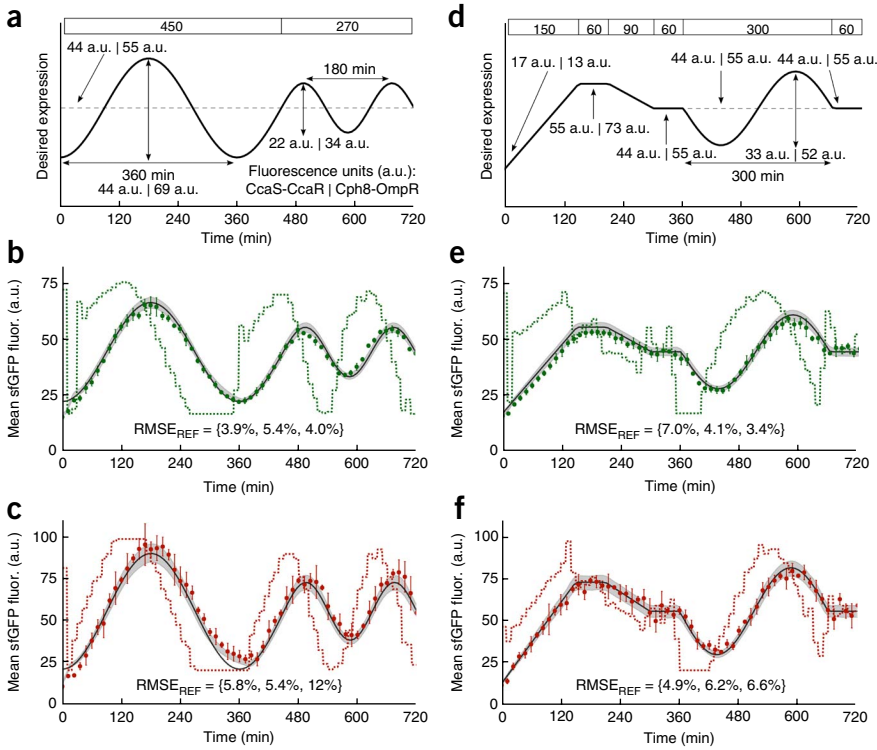

Figure 4 | Programming custom protein-expression signals with a biological function generator. (**a**) Sinusoidal reference expression signal with labels indicating absolute expression levels of the CcaS-CcaR and Cph8-OmpR systems for **b**,**e** and **c**,**f** respectively. Timings used in the reference are shown in minutes (top). (**b**,**c**) Experimental response to the precomputed light inputs for the sinusoidal reference signal. (**d**–**f**) Complex waveform generation as described in **a**–**c**. Light input, prediction and experimental data are as described for **Figure 2**. $RMSE_{REF}$ compares the experiment to the reference for each of the three trials. Fluor., fluorescence.

The two variables are the sfGFP production rate, $p(t)$, and the sfGFP abundance, $g(t)$. The parameters include a short delay in response to the control signal (τ_{delay}), a light intensity–dependent rate of change of the sfGFP production rate (k_p) and the sfGFP dilution rate due to cell growth (k_g) (Online Methods and **Supplementary Fig. 10**). We used least-squares minimization to parameterize the model against the experimental data sets (**Supplementary Figs. 8** and **9** and **Supplementary Table 1**). To validate our modeling approach, we then simulated the response of both sensors to a series of step changes in the control signal that we had not previously tested in the calibration experiments (**Fig. 2e**,**f**). We then exposed bacterial cultures expressing each sensor to the same series of step changes, and measured the resulting sfGFP expression dynamics. For both sensors, our simulations predicted the experimental dynamics accurately and reproducibly over experiments performed on three separate days (**Fig. 2e**,**f** and **Supplementary Table 2**).

Computational design of light control programs

We next investigated whether the model could be used to design time-varying light programs capable of driving gene expression to follow a user-defined reference signal. To this end, we wrote an *in silico* algorithm that accepts a reference gene expression time course and an initial light control program as inputs (**Fig. 3a** and **Supplementary Fig. 11**). The algorithm uses the dynamic model to simulate the gene expression response and iteratively optimizes the light control program until the error between the reference and the simulation is sufficiently small. We then tested the computationally optimized light control programs experimentally with bacteria grown in the LTA.

Programming gene expression transitions

First, we used our computational approach to dramatically reduce the time required to transition between two analog gene expression levels. The canonical strategy of applying a step increase in the concentration of an inducer resulted in relatively slow, exponential dynamics with a fixed shape and timescale (**Fig. 2c**,**d**). To accelerate gene expression response dynamics, we simply defined references that are equivalent to the desired final expression level for cells beginning at a different, initial expression level. For example, we specified a time-zero target of 54 arbitrary sfGFP fluorescence units (a.u.) for CcaS-CcaR–expressing cells preconditioned to the minimal sfGFP expression level of 17 a.u. The algorithm began by applying the maximal green intensity in order to quickly minimize the difference between the gene expression simulation and the reference (**Fig. 3a**). According to the model, this maximal step increase minimizes the gene expression switching time and increases sfGFP production to the maximal rate. As the simulated expression level approaches the reference, the algorithm decreases green light intensity to avoid unwanted overshooting. Finally, green intensity is increased to a final intermediate value over a series of smaller step changes (**Fig. 3a**).

Experimental tests of the above accelerator light program revealed that *in vivo* gene expression closely follows the simulation (**Fig. 3a**), thus dramatically reshaping response dynamics. Indeed, the transition to 90% of the target expression level was accelerated to 60 min, relative to 165 min for a step increase. We used this same approach to accelerate the CcaS-CcaR step transition from 17 to 28, 39, 50, and 61 a.u., as well as the transition of Cph8-OmpR from 13 to 30, 47, 64 and 81 a.u. (**Fig. 3b**,**c**). In all cases, the model predicted the accelerated response dynamics accurately, and the transition times were reduced by up to 83% (**Supplementary Note 4** and **Supplementary Table 3**). In addition, we used this strategy to accelerate gene expression decreases between many different analog expression levels (**Supplementary Fig. 12**).

Next we defined a set of linearly increasing reference signals with different slopes over a 180-min time window (**Fig. 3a**,**d**,**e**). The algorithm produced light control programs predicted to generate these linearized dynamics for each sensor. Once again, the experimental gene expression dynamics closely followed the references with low error, for both sensors, over three different days (coefficient of determination $R^2 = 0.761–0.980$; **Supplementary Table 2**). Our method also generated highly linear increases and decreases over all longer experimental trials examined, up to 12 h (**Supplementary Fig. 13**).

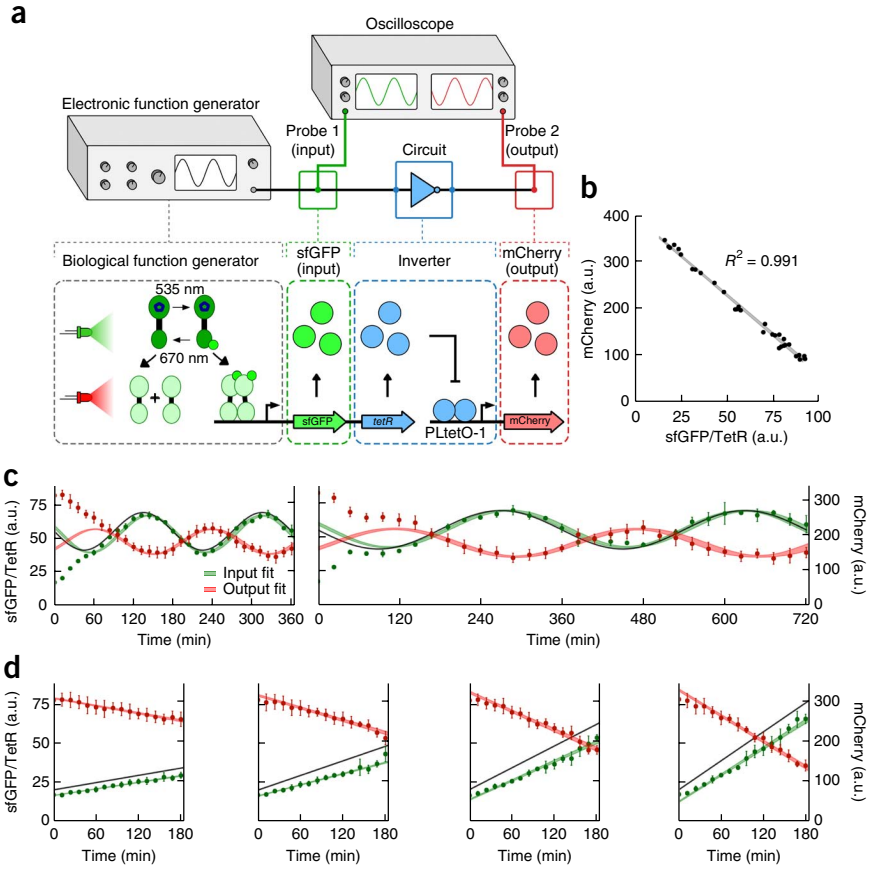

Figure 5 | Analysis of gene circuit signal processing using the biological function generator and dual-reporter oscilloscope system. (**a**) Top, a basic toolkit for characterizing signal propagation through electronic circuits consists of a function generator, which sends an input signal into a circuit of interest, and an oscilloscope, which uses two probes to measure the signal both before and after it has been transformed by the circuit. Bottom, an analogous toolkit for biological circuits. Here the CcaS-CcaR biological function generator sends a TetR input signal into the TetR/PLtetO-1 genetic inverter. A dual-reporter system, in which sfGFP is proportional to the input signal and mCherry is the output signal, is used to measure circuit signal-processing features. (**b**) Steady-state response of the inverter circuit. The markers indicate the measured means of the sfGFP and TetR ('sfGFP/TetR') and mCherry fluorescence distributions in response to different input light intensities. A least-squares fit (gray) is shown as the interquartile range of $N = 500$ Gaussian samplings of the parameter standard errors. (**c**,**d**) Two sinusoidal (**c**) and four linear (**d**) reference signals (black) were used to generate an sfGFP/TetR input signal (green dots), which is transformed by the inverter into a mCherry output signal (red dots). The sinusoidal reference signals have periods of 180 (left) and 360 min (right), and the linear reference signals drive the system with four different slopes over 180 min. Data are shown as previously described (**Fig. 2e**,**f**). Least-squares sine and linear fits (green and red shaded regions) were performed to the data following 120 min (**c**) and 24 min (**d**) so that the input signals could approach the reference. The fit regions are shown as the interquartile range of $N = 500$ simulations produced by sampling the fit parameter uncertainties.

Programming gene expression signals

A technology for programming tailor-made gene expression signals *in vivo* (i.e., a biological function generator) could transform our understanding of biological systems, in particular genetic circuits. We therefore investigated whether our approach could be used to generate user-defined protein waveforms. First we defined a 12-h sinusoidal reference that oscillates with a period of 360 min and has an amplitude of 80% and offset of 50% of the total CcaS-CcaR and Cph8-OmpR output ranges. After one oscillation, the period and amplitude are halved, and two further oscillations occur (**Fig. 4a**). The algorithm produced oscillatory light-control programs predicted to drive each sensor to closely follow this reference, and bacteria exposed to these light programs produced well-defined gene expression sinusoids (**Fig. 4b**,**c**). For both sensors, we observed no error propagation over the 12-h experiment, and sfGFP expression patterns quantitatively tracked the amplitude, period and offset values for both of the sine waves specified in the reference with the same standard of accuracy, precision and reproducibility achieved for the accelerated and linear transitions (**Supplementary Table 2**).

To examine the degree of gene expression programmability enabled by our method, we next challenged the algorithm with a 12-h reference composed of several different dynamics juxtaposed in series (**Fig. 4d**). For CcaS-CcaR, the reference begins with a linear increase from 17 to 56 a.u. over 150 min, followed by a 60-min hold, a 90-min linear decrease to 44 a.u. and a 60-min hold. Then the signal switches to a sinusoidal oscillation with a 300-min period, peak-to-peak amplitude of 50 a.u. and offset of 44 a.u. Finally, the signal holds constant for 60 min. The same reference is generated for the Cph8-OmpR system, with the same relative sfGFP values, scaled to the output range of the system (**Fig. 4d**). As before, the algorithm produced light control programs predicted to drive sfGFP dynamics that closely follow the reference. Despite the length and complexity, the light control program faithfully generated the desired waveform *in vivo* via both light sensors, with no loss in accuracy, precision or reproducibility compared to the other reference signals tested (**Fig. 4e**,**f**). The combined experimental r.m.s. error between the simulated and observed signals ($RMSE_{SIM}$) for all programming experiments was 5.0% for CcaS-CcaR ($N = 53$ trials) and 6.9% for Cph8-OmpR ($N = 33$ trials) (**Supplementary Table 2**).

Using the function generator to analyze gene circuits

To probe the utility of our biological function generator for characterizing gene circuits (**Fig. 5a**), we next connected the CcaS-CcaR system to the tetracycline repressor gene *tetR*, which we used to repress the PLtetO-1 promoter. This configuration results in a transcriptional 'inverter' circuit, wherein high levels of transcriptional input from CcaS-CcaR result in low levels of transcriptional output from PLtetO-1, and vice versa (Online Methods and **Supplementary Fig. 1**). To simultaneously monitor circuit input and output signals, we used a dual fluorescent reporter probe

system wherein sfGFP is expressed from the same mRNA as TetR, and mCherry is expressed from PLtetO-1 (ref. 11) (**Fig. 5a**). Dual-channel flow cytometry measurements revealed that the addition of this inverter circuit had a negligible effect on the performance of the CcaS-CcaR system (**Supplementary Note 5** and **Supplementary Table 1**). Furthermore, we observed a highly linear ($R^2 = 0.991$), inverse steady-state relationship between TetR and mCherry levels across the entire CcaS-CcaR output range (**Fig. 5b**, **Supplementary Note 6** and **Supplementary Table 4**).

We next used our approach to produce sinusoidal and linear TetR input signals (**Fig. 5c**,**d**). A simple recalibration of the experimentally measured cell division rate and steady-state parameters in the dynamical model (**Supplementary Note 5** and **Supplementary Table 1**) resulted in experimentally measured sfGFP and TetR dynamics that closely followed the reference sinusoids over the entire time course (**Fig. 5c**), demonstrating that the function generator is robust to these different genetic contexts. Dual-channel cytometry revealed that the circuit transforms 180- and 360-min-period sinusoidal inputs into inverted output oscillations with the same periods (**Fig. 5c**). Both the input and output signals are well-fit by sine functions, as evidenced by nonlinear least-squares fits, revealing that the circuit performs a linear transformation wherein the input signal is inverted and offset by a time delay of 7.0 ± 5.4 min (**Supplementary Note 7** and **Supplementary Table 5**).

To further characterize the linearity of the inverter, we used the function generator to create the same four 180-min linear TetR signals that we previously demonstrated with sfGFP alone (**Fig. 3d**). The experimental sfGFP and TetR dynamics remained highly linear (**Fig. 5d**), albeit with a slight systematic discrepancy between the reference and *in vivo* signals (**Supplementary Note 8**). Despite this small difference, the dual probe approach allowed us to directly measure the *in vivo* transformation performed by the circuit. Fits again revealed that the circuit output is exceptionally linear regardless of the slope of the input (**Fig. 5d** and **Supplementary Note 7**).

DISCUSSION

Our overall approach is analogous to the standard electronic engineering technique of using a function generator to produce programmable voltage signals and an oscilloscope to monitor the signals before and after passing through a circuit of interest. In principle, our method could be used to study virtually any biological process that is dynamically affected by gene expression. We have demonstrated the utility of the approach by varying the expression level of the transcriptional repressor TetR linearly and sinusoidally and monitoring the protein expression output from the regulated promoter PLtetO-1. Despite numerous inherent sources of nonlinearity in transcriptional regulation, our experiments revealed an extended range over which PLtetO-1 responds linearly to TetR signals, with an approximate 7-min delay.

Linear gene circuits are desirable for synthetic biology because they are relatively simple to characterize and model and are likely to perform predictably when connected to other genetic devices. One could construct amplifiers, attenuators, adders and adaptors that correct signal-strength mismatches[14] between parts, as well as other advanced devices from linear gene circuits. Our function generator could also be combined with linear systems–identification methods from engineering, such as frequency analysis[15], to construct black-box models of natural or synthetic gene circuits. Several recent studies have used microfluidic devices to perform frequency analysis of gene circuits by creating extracellular signals of their native chemical effectors[1–3,6]. In contrast, our function generator uses a single input, light, to produce intracellular signals of different proteins, yielding a simpler and more generalizable approach.

Experimentally parameterized dynamical models of gene circuits have major implications in basic science and engineering. For natural circuits, black-box models can be developed quickly using our approach, thereby accelerating the generation of new hypotheses. These models can then be elaborated into mechanistic models in a rational or iterative manner, allowing rapid testing of the hypotheses. In particular, the high performance standard of our method is ideal for analyzing dynamic biological phenomena that are technically challenging to study[16], such as cell cycles[17], circadian clocks[18], stress responses[19,20] and differentiation[21–24]. Our approach could also facilitate more detailed studies of ordered assembly in multiprotein complexes[25,26] and of the dependence of gene circuit dynamics on host context[27,28], and it should improve the predictability with which synthetic gene circuits can be composed into higher-order systems[29]. Finally, our approach could be used to dynamically program metabolic enzyme expression, allowing faster optimization of engineered pathways, especially those wherein the timing of expression affects efficiency, toxicity or yield[30,31].

The two light-switchable TCSs used in this study have different action spectra and can be independently controlled when expressed in the same cell[9], making our approach amenable to multiplexing. The function-generator method should also be extensible to other phytochrome-based TCSs that respond from the UV to the near infrared[32,33] and to light-switchable eukaryotic gene expression systems[34,35]. The performance features of our TCSs, such as basal (leaky) expression and output dynamic range, could be improved with further engineering, thereby extending the applicability of our method to more biological problems. In addition, continuous culture instruments could be developed to eliminate well-to-well variability in the LTA, decrease handling requirements and increase data collection throughput. The addition of on-line fluorescent protein detectors would also enable *in silico* feedback, wherein the light control program could be adjusted to compensate for any variability in protein expression levels not accounted for in the model. Finally, an arrayed micromirror technique has recently been combined with a eukaryotic optogenetic tool[36] to produce user-defined signals of membrane localization of eukaryotic signaling proteins[37] and reveal new properties of a signal transduction pathway[38]. Our method should be compatible with this approach, allowing gene circuit dynamics to be characterized with unprecedented single-cell resolution.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Addgene: pJT119b plasmid expressing CcaS and CcaR with sfGFP under the PcpcG2 promoter, 50551; pCph8 plasmid expressing Cph8, 50552; pEO100c plasmid expressing

sfGFP under the PompC promoter, 50550; pBL4 plasmid expressing CcaS and CcaR with sfGFP-TetR under the PcpcG2 promoter, 50548; pBL6 plasmid expressing mCherry under the PLtetO-1 promoter, 50549.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This research was supported by the US National Science Foundation Biotechnology, Biochemical, and Biomass Engineering (BBBE) program (EFRI–1137266) and the Office of Naval Research MURI programs (N000141310074). L.A.H. was supported by the US National Aeronautics and Space Administration Office of the Chief Technologist's Space Technology Research Fellowship (NSTRF) (NNX11AN39H). We thank S. Schmidl (Rice University) for his contribution of the pPCBSE plasmid.

AUTHOR CONTRIBUTIONS

J.J.T. and E.J.O. conceived of the project, and J.J.T. supervised the project. E.J.O., L.A.H., B.P.L. and R.S. designed and performed experiments. E.J.O. and B.P.L. designed and constructed plasmids. E.J.O. designed and constructed the LTA, analyzed data, constructed the model and developed the light program sequence optimization algorithm. J.J.T. and E.J.O. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bennett, M.R. *et al.* Metabolic gene regulation in a dynamically changing environment. *Nature* **454**, 1119–1122 (2008).
2. Mettetal, J.T., Muzzey, D., Gómez-Uribe, C. & van Oudenaarden, A. The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* **319**, 482–484 (2008).
3. Wei, P. *et al.* Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells. *Nature* **488**, 384–388 (2012).
4. Muzzey, D., Gómez-Uribe, C.A., Mettetal, J.T. & van Oudenaarden, A. A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell* **138**, 160–171 (2009).
5. Tan, C. *et al.* The inoculum effect and band-pass bacterial response to periodic antibiotic treatment. *Mol. Syst. Biol.* **8**, 617 (2012).
6. Hersen, P., McClean, M.N., Mahadevan, L. & Ramanathan, S. Signal processing by the HOG MAP kinase pathway. *Proc. Natl. Acad. Sci. USA* **105**, 7165–7170 (2008).
7. Milias-Argeitis, A. *et al. In silico* feedback for *in vivo* regulation of a gene expression circuit. *Nat. Biotechnol.* **29**, 1114–1116 (2011).
8. Uhlendorf, J. *et al.* Long-term model predictive control of gene expression at the population and single-cell levels. *Proc. Natl. Acad. Sci. USA* **109**, 14271–14276 (2012).
9. Tabor, J.J., Levskaya, A. & Voigt, C.A. Multichromatic control of gene expression in *Escherichia coli*. *J. Mol. Biol.* **405**, 315–324 (2011).
10. Levskaya, A. *et al.* Synthetic biology: engineering *Escherichia coli* to see light. *Nature* **438**, 441–442 (2005).
11. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
12. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C. & Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88 (2006).
13. Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Networks* (CRC Press, 2007).
14. Yokobayashi, Y., Weiss, R. & Arnold, F.H. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA* **99**, 16587–16591 (2002).
15. Ang, J., Ingalls, B. & McMillen, D. Probing the input-output behavior of biochemical and genetic systems system identification methods from control theory. *Methods Enzymol.* **487**, 279–317 (2011).
16. Levine, J.H., Lin, Y. & Elowitz, M.B. Functional roles of pulsing in genetic circuits. *Science* **342**, 1193–1200 (2013).
17. McAdams, H.H. & Shapiro, L. A bacterial cell-cycle regulatory network operating in time and space. *Science* **301**, 1874–1877 (2003).
18. Bell-Pedersen, D. *et al.* Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat. Rev. Genet.* **6**, 544–556 (2005).
19. Young, J.W., Locke, J.C.W. & Elowitz, M.B. Rate of environmental change determines stress response specificity. *Proc. Natl. Acad. Sci. USA* **110**, 4140–4145 (2013).
20. El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A. & Khammash, M. Surviving heat shock: control strategies for robustness and performance. *Proc. Natl. Acad. Sci. USA* **102**, 2736–2741 (2005).
21. Vishnoi, M. *et al.* Triggering sporulation in *Bacillus subtilis* with artificial two-component systems reveals the importance of proper Spo0A activation dynamics. *Mol. Microbiol.* **90**, 181–194 (2013).
22. Levine, J.H., Fontes, M.E., Dworkin, J. & Elowitz, M.B. Pulsed feedback defers cellular differentiation. *PLoS Biol.* **10**, e1001252 (2012).
23. Kuchina, A. *et al.* Temporal competition between differentiation programs determines cell fate choice. *Mol. Syst. Biol.* **7**, 557 (2011).
24. Ray, J.C.J., Tabor, J.J. & Igoshin, O.A. Non-transcriptional regulatory processes shape transcriptional network dynamics. *Nat. Rev. Microbiol.* **9**, 817–828 (2011).
25. Kalir, S. *et al.* Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292**, 2080–2083 (2001).
26. Temme, K. *et al.* Induction and relaxation dynamics of the regulatory network controlling the type III secretion system encoded within *Salmonella* pathogenicity island 1. *J. Mol. Biol.* **377**, 47–61 (2008).
27. Cardinale, S., Joachimiak, M.P. & Arkin, A.P. Effects of genetic variation on the *E. coli* host-circuit interface. *Cell Rep.* **4**, 231–237 (2013).
28. Arkin, A.P. A wise consistency: engineering biology for conformity, reliability, predictability. *Curr. Opin. Chem. Biol.* **17**, 893–901 (2013).
29. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **26**, 787–793 (2008).
30. Zaslaver, A. *et al.* Just-in-time transcription program in metabolic pathways. *Nat. Genet.* **36**, 486–491 (2004).
31. Temme, K., Zhao, D. & Voigt, C.A. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci. USA* **109**, 7085–7090 (2012).
32. Rockwell, N.C., Martin, S.S., Feoktistova, K. & Lagarias, J.C. Diverse two-cysteine photocycles in phytochromes and cyanobacteriochromes. *Proc. Natl. Acad. Sci. USA* **108**, 11854–11859 (2011).
33. Jaubert, M. *et al.* Control of peripheral light-harvesting complex synthesis by a bacteriophytochrome in the aerobic photosynthetic bacterium *Bradyrhizobium* strain BTAi1. *J. Bacteriol.* **190**, 5824–5831 (2008).
34. Shimizu-Sato, S., Huq, E., Tepperman, J.M. & Quail, P.H. A light-switchable gene promoter system. *Nat. Biotechnol.* **20**, 1041–1044 (2002).
35. Müller, K. *et al.* Multi-chromatic control of mammalian gene expression and signaling. *Nucleic Acids Res.* **41**, e124 (2013).
36. Levskaya, A., Weiner, O.D., Lim, W.A. & Voigt, C.A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997–1001 (2009).
37. Toettcher, J.E., Gong, D., Lim, W.A. & Weiner, O.D. Light-based feedback for controlling intracellular signaling dynamics. *Nat. Methods* **8**, 837–839 (2011).
38. Toettcher, J.E., Weiner, O.D. & Lim, W.A. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* **155**, 1422–1434 (2013).

ONLINE METHODS

Plasmids and strains. For the CcaS-CcaR system, plasmids pJT119b and pPLPCB(S) (**Supplementary Fig. 1**) were constructed previously[9]. For the Cph8-OmpR system, pCph8 (ref. 10) and pPLPCB(S) were constructed previously. Plasmid pEO100c (**Supplementary Fig. 1**) was constructed by swapping in a new synthetic ribosome binding site (RBS; TCATATATAAATAAAAT AAGGTAGGTCAATAT) into pEO100b using the MEGAWHOP procedure[39]. The RBS was designed using the RBS Calculator[40] with a target translation rate of 100,000 a.u. pEO100b (not used for data collection in this study) was constructed using a two-part Gibson assembly[41] with pJT108 (ref. 42) as PCR template for the plasmid backbone and pJT119b as a PCR template for the sfGFP gene. Plasmid pBL4 (**Supplementary Fig. 1**) was constructed using a two-part Golden Gate assembly[43] with pJT119b as PCR template for the backbone and a lab stock of the *tetR* gene as PCR template for the *tetR* assembly fragment. *tetR* was translated from a synthetic RBS (TCACACAGGAAACCTACTAG) sourced from part BBa_B0031 (http://parts.igem.org/). Plasmid pBL6 (**Supplementary Fig. 1**) was constructed using a two-part Golden Gate assembly[43] with pEO100c (**Supplementary Fig. 1**) as PCR template for the backbone and a lab stock of the mCherry gene as PCR template for the mCherry assembly fragment, and PLtetO-1 (ref. 11) was embedded in the mCherry forward primer. Plasmid pPCBSE (graciously gifted by S. Schmidl) was used in place of pPLPCB(S) for the inverter circuit experiments (**Fig. 5**). pPCBSE was constructed using one-part Golden Gate assembly[43] with pPLPCB(S) as PCR template. The engineered promoter Plac/ara-1, which drives *ho1* and *pcyA* transcription on pPLPCB(S), was replaced by primer overhangs with the constitutive promoter J23108 (CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC, http://parts.igem.org/). The kanamycin-resistant *envZ*-deficient *E. coli* strain JT2 (ref. 9) was used for all experiments. Primers used during the construction of these plasmids have been listed (**Supplementary Table 6**). The plasmids and sequences are available at Addgene (http://www.addgene.org/).

Light tube array (LTA) design and construction. The LTA is constructed from layers of opaque foam and aluminum and a printed circuit board layer (**Supplementary Figs. 2** and **3**). The instrument provides an isolated optical environment for 64 standard 14-mL culture tubes. Described from the bottom up, an aluminum base-plate is used to fasten the device to the platform of a benchtop 37 °C shaking incubator (Thermo Fisher MaxQ4000). The next layer is a custom-designed printed circuit board (fabricated by Pad2Pad) containing red (Kingbright cat. #WP1503SRD), green (Kingbright cat. #WP7083ZGD), far-red (Epitex cat. L740-05AU) and blue (Vishay Semiconductors cat. #TLHB4400) LEDs, TLC5940 LED drivers (Texas Instruments), connections for signal-carrying wires, and traces to carry the electrical signals between the components. The LEDs are elevated by nylon standoffs so that they can be positioned closer to the cultures tubes. A foam tube surrounds each set of LEDs to optically isolate the wells. The next aluminum layer compresses the foam tubes and serves as the top layer of the light-emitting section of the LTA. These three layers are fixed by seven screws, which penetrate through holes placed across their surfaces.

The next layer is a 1-inch-thick foam block with punched holes, allowing light to pass up to the tubes above (**Supplementary Fig. 3**). The thickness of this layer can be used to control the spacing between the LEDs and the bottom of the tubes, which can be used to modulate light intensity in the samples. The aluminum layer resting above has holes precisely milled to 13.6 mm, a diameter slightly less than the 16-mm diameter of the culture tubes, ensuring that the tubes will stay in a fixed position during shaking.

The remaining seven layers of 1-inch and 0.5-inch foam with holes punched through serve as a tube rack (**Supplementary Fig. 3**). These layers hold the tubes in place and further serve to optically isolate the light environments of the wells from one another. The top layer is a 0.5-inch foam sheet without holes, used as a lid to block ambient room light. All of the layers are held in place at the corners by 2-inch-wide strips of 1/16-inch aluminum that have been bent lengthwise into a 90° angle.

Light intensity is controlled via a pulse width–modulated (PWM) signal generated by the LED drivers, each of which can control 16 LEDs. The PWM signals have a 12-bit resolution (4,096 levels) and a full PWM cycle frequency of 500 Hz. To control the 256 LEDs on the board, 16 TLC5940 devices are daisy chained. The signals to control the TLC5940s are provided by an Arduino microcontroller board located outside the incubator. A freely available, GPL-licensed library for the Arduino was used to drive the TLC5940s (https://code.google.com/p/tlc5940arduino/, written by A. Leone). A small wrapper was written for this library to enable simple commands to be used to drive the LEDs in the LTA, requiring the user to specify only a row, column, color and intensity.

Light intensity calibrations. LED intensity measurements are performed using a calibrated fiber-optic spectrophotometer (Stellar-Net EPP2000 UVN-SR-25). The fiber-optic sensor is placed into a culture tube and placed level with the 1-mL marking on the tube. The fiber is held in position inside the tube by foam stuffing. This tube-sensor is placed into each LTA well, and the intensity of each LED is measured. The LED intensities are then corrected to compensate for variations. A diffuser layer (Rosco, Roscolux #116 Tough White Diffusion) is positioned just above the LEDs while the calibration measurements are performed.

The maximum intensity levels of the red and green LEDs used in this study are 1.05 W/m^2 and 4.03 W/m^2, respectively (measured without diffuser). These intensities are reached when the PWM value is set to its maximum of 4,095. If the PWM value is set to 0, the LEDs emit no light. For intermediate PWM values, a linear interpolation can be performed between the end points. The relationships between green and red intensities measured in W/m^2 ($I_{g,W/m^2}$ and $I_{r,W/m^2}$) and intensities referenced by PWM value ($I_{g,PWM}$ and $I_{r,PWM}$) are:

$$I_{g,W/m^2} = 9.841 \times 10^{-4} I_{g,PWM}$$

$$I_{r,W/m^2} = 2.563 \times 10^{-4} I_{r,PWM}$$

Optical control schemes. The control scheme chosen for the CcaS-CcaR system is to apply the maximum LTA red intensity of 1.05 W/m^2 and to control gene expression by modulating green intensity between 0 and 4.03 W/m^2. This results in a larger number of accessible analog states and also faster response kinetics (**Supplementary Note 3** and **Supplementary Fig. 14**). For the Cph8-OmpR system, we chose to modulate expression via

doi:10.1038/nmeth.2884

red light intensity alone, which simplifies the control scheme (**Supplementary Note 3**).

***E. coli* growth, light exposure and harvesting protocol.** The protocol is performed over two consecutive days.

1. Late in the day, start a 37 °C, shaking overnight culture from a −80 °C stock in a tube containing 3 mL LB medium and the appropriate antibiotics (50 μg/mL kanamycin, 100 μg/mL spectinomycin and 34 μg/mL chloramphenicol for the CcaS-CcaR system, and the same antibiotics plus 50 μg/mL ampicillin for the Cph8-OmpR system).
2. After the overnight culture has grown for 10–12 h, prepare 200 mL M9 medium (151.58 mL autoclaved, distilled H_2O, 40 mL 5× M9 salts, 4 mL 10% casamino acids, 4 mL 20% glucose, 400 μL 1 M $MgSO_4$, 20 μL $CaCl_2$). Add appropriate antibiotics to medium. Shake/stir the container to ensure the antibiotics are mixed well in the medium.
3. For dual-reporter inverter experiments, add anhydrotetracycline (aTc) to bring the medium to a concentration of 40 ng/mL. (See "Note 1" at the end of the protocol.) Shake/stir the container to ensure the inducer is mixed well in the medium.
4. Measure the OD_{600} of the overnight culture.
5. Dilute the overnight culture into the M9 + antibiotics, bringing the OD_{600} to 0.0001. (See "Note 2" at the end of the protocol.) Shake/stir the container to ensure the cells are mixed well in the medium.
6. Distribute 3 mL of inoculated medium into each of 64 BD Falcon round-bottom 14-mL polypropylene test tubes (BD Biosciences catalog #352006).
7. Place tubes in the LTA and grow at 37 °C with shaking at 250 r.p.m. for 8 h. Randomize the time points of the light program (**Supplementary Note 1** and **Supplementary Fig. 15**) throughout the 8 × 8 array to avoid reproduction of systematic errors due to slight variations in LED intensity. Randomize anew for each trial.
8. After 8 h of growth, harvest all test tubes by immediately transferring them into an ice-water bath. Wait 10 min for the cultures to equilibrate to the cold temperature and for gene expression to stop.
9. Approximately 1.5 h before stopping the experimental cultures, begin preparing a solution of phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH to 7.4) + 500 μg/mL of the transcription inhibitor rifampicin (Rif, Tokyo Chemical Industry, cat. #R0079). Prepare at least 1 mL for each culture to be measured via flow cytometry. Rif dissolves slowly, so allow at least 45 min of stirring. Also at this time, begin preparing a 37 °C water bath.
10. Filter the dissolved solution of PBS + Rif through a 0.22-μm 20-mL syringe filter.
11. Transfer 1 mL of the filtered PBS + Rif into one 5-mL cytometer tube per LTA culture sample, and chill tubes in a rack in an ice-water bath.
12. Transfer 50 μL of each chilled culture from step 7 into the chilled PBS + Rif solution.
13. Incubate the rack(s) of PBS + Rif + culture tubes in a 37 °C water bath for 1 h. Our measurements reveal that this allows sfGFP maturation to go to completion (**Supplementary Note 2**).
14. Transfer the rack(s) back into ice-water bath.
15. Wait 15 min, and then begin measuring each tube on a flow cytometer.

Note 1. To achieve a high degree of repeatability in the concentration of aTc, we created a master stock at 10,000× working concentration and divided this stock into aliquots containing 10% more volume than is required for a single run into individual tubes. These tubes were wrapped in foil and stored at −20 °C until use.

Note 2. The starting OD_{600} of 0.0001 was determined empirically to ensure that the cultures remain in exponential phase (final OD_{600} ~0.3) throughout the 8-h growth experiment. For 12-h growth experiments, the starting density is decreased to 10^{-6}. Over this range, the starting density has no impact on the response to light (**Supplementary Fig. 16**). Additionally, we have observed that increased expression of sfGFP results in a slight slowdown of growth (**Supplementary Fig. 17**), however we maintained the same inoculation density of cell cultures regardless of the light conditions they experienced.

Flow cytometry data acquisition and analysis. Cytometry acquisition was performed using a BD FACScan flow cytometer with the original laser system replaced by blue (488 nm, 30 mW) and yellow (561 nm, 50 mW) solid-state lasers (Cytek). The FL1 (sfGFP) acquisition channel emission filter was also replaced with a 510/21-nm filter. The FL3 (mCherry) acquisition channel emission filter (650 nm long pass) is original to the instrument. The cytometer is calibrated using Spherotech cat. #RCP-30-5A beads approximately once per week. Acquisition is performed with typical count rates of 1,000–2,000 events/s. Approximately 50,000 events are stored for each sample. A SSC threshold is used to eliminate instrument noise events that are clearly not due to cell scattering. The cytometer settings used for the two light-switchable systems have been listed (**Supplementary Table 7**).

After acquisition, the raw cytometry data are processed using custom Python scripts (**Supplementary Software**). First, the first 250 and last 100 events are removed from the data set to avoid transient errors introduced owing to uneven pressurization of the sample tube. Then the highest and lowest measured histogram channel for each of the measured values (FSC, SSC, FL1 and FL3) are removed, as the events in these channels have an undetermined fluorescence value. Then a small elliptical gate centered at the median FSC and SSC values is used to isolate a uniformly sized population of cells. The elliptical gate was chosen to have a semi-major length of 64 channels and a semi-minor axis of 32 channels. The gate is tilted so that the angle between the FSC and semi-major axes is 28°, to align with the observed cell populations. The gating procedure leaves $N = 2{,}000–5{,}000$ events. Finally, a trim is performed on FL1 to remove a small number of apparent noncellular events. For the CcaS-CcaR system, the gate is set to a channel value of 127, whereas for the Cph8-OmpR system it is set to a value of 32.

The extracted data set is then used to calculate the statistics of the measured FL1 distribution. However, before the sample statistics are calculated, the extracted values are transformed from their measured log-scale channel units (10 bit, ranging from 0 to 1,023) to linear-scale fluorescence units (ranging from 1 to 9,910) using the transformation $FL1_{LIN} = 10^{(FL1\_LOG/256)}$. The arithmetic mean of the resulting histograms (**Supplementary Fig. 18**) is the primary statistic of interest and is the value used

to represent the population mean of each tube throughout this work. Finally, the measured autofluorescence *E. coli* JT2 value of 9.583 a.u. is subtracted from the Cph8-OmpR system means. This blanking procedure is not possible for the CcaS-CcaR system, as the non-sfGFP-expressing cells were below the detection threshold of the cytometer at the low-sensitivity FL1 gain setting used. Fluorescence measurements of mCherry for the dual-reporter inverter system required compensation for sfGFP emission bleed-through into the FL3 detector (**Supplementary Note 9**).

Scatter in the steady-state transfer function. We used the steady-state transfer characterization data to estimate the deviation from the steady-state transfer function model (**Supplementary Figs. 4** and **6**) at different light input values. This measured scatter is used to assign uncertainties to individual data points in the model fitting process to reduce the possibility of individual data points having a disproportionate impact on the error minimization during the fit routine. To determine the scatter of the points around the fit as a function of light intensity, we calculated a seven-point-wide moving s.d. across the absolute value of the residuals to the transfer function fit of the data set. A second-order polynomial fit is used to provide an empirical relation between the input light intensity and observed sfGFP scatter (**Supplementary Note 10**).

Dynamical characterization and 2D ODE model. The observation of the three timescales in the step response motivated the construction of a two-dimensional ordinary differential equation (2D ODE) model.

$$\frac{dp(t)}{dt} = k_p(I_g, I_r) \times \left(c(t - \tau_{delay}) - p(t)\right)$$

$$\frac{dg(t)}{dt} = k_g \times \left(p(t) - g(t)\right)$$

$$p_0 = g_0 = c(t < \tau_{delay}) = c_{precondition}$$

The two dynamic variables in this model are the production rate of sfGFP, $p(t)$, and the sfGFP abundance, $g(t)$. The dynamics of these two variables are determined by three model parameters each corresponding to each of the observed timescales. The time delay observed between a light input and a change in $p(t)$ is given by τ_{delay}, the rate of change of the $p(t)$ is given by k_p, and the dilution rate of the cells which governs the dynamics of $g(t)$ is given by k_g. For the CcaS-CcaR system, the k_p parameter exhibits a dependence on the controlling light intensities I_r and I_g, which correspond to red and green light, respectively. The set point for the system $c(t)$ is determined by mapping the light intensities at time $t - \tau_{delay}$ through the steady-state transfer function.

To further explain the model, it is illustrative to consider its response to a step change in light intensity (**Supplementary Fig. 6**). First, the system is preconditioned to an initial set point $c_{precondition}$ by the light intensities $I_{r,precondition}$ and $I_{g,precondition}$. Then, at time $t = 0$, the light input levels change to I_r and I_g. This drives the set-point of the system $c(I_r, I_g)$ to a new level that can be determined from the steady-state transfer function (**Supplementary Figs. 4** and **6**). The sfGFP production rate $p(t)$ of the light-induced promoter begins to respond to the new c (after τ_{delay} has elapsed) with first-order kinetics with rate k_p. We allow that k_p itself can depend on the light intensities I_r and I_g. The form of $k_p(I_r, I_g)$ was empirically determined for each system. For Cph8-OmpR, k_p is simply a constant, whereas for the CcaS-CcaR system the functional form used is

$$k_p(c) = \begin{cases} k_{p,o} & : c < p_0 \\ k_{p,on}(c) & : c \geq p_0 \end{cases}$$

$$k_{p,on}(c) = \begin{cases} k_{p,l} & : I_g < k_{p,k} \\ k_{p,l} + (k_{p,m} - k_{p,l})\dfrac{I_g(c) - k_{p,k}}{I_{g,max} - k_{p,k}} & : I_g \geq k_{p,k} \end{cases}$$

This form results in a rate that depends upon the whether the transition produces an increase or decrease in expression. For decreasing transitions, the rate is constant at $k_{p,o}$. For increasing transitions, the rate is given by $k_{p,on}$, which is detailed in the second equation. For $k_{p,on}$, the rate is held constant at $k_{p,l}$ for I_g less than a threshold intensity $k_{p,k}$ and increases linearly thereafter with intensity to a rate of $k_{p,m}$ when the light is at maximum intensity.

Continuing with the step-change response, the sfGFP fluorescence $g(t)$ follows the production rate $p(t)$ with first-order kinetics at a rate k_g, and the system eventually comes to equilibrium with $g(t) = c(I_r, I_g)$. All model parameter fit values, standard errors and units for the CcaS-CcaR, Cph8-OmpR and CcaS-CcaR + inverter systems have been listed (**Supplementary Table 1**).

Model fitting for CcaS-CcaR and Cph8-OmpR predictive models. We performed least-squares fitting with custom Python scripts (**Supplementary Software**) that utilize routines from a freely available Python library (Scipy.optimize.fmin). The fits were performed against data from multiple runs consisting of both single and multiple step functions to identify the optimal model parameter values for each system. The fits were performed across all of the calibration data sets simultaneously, using the scatter in the steady-state transfer function (see "Scatter in the steady-state transfer function" above) as an estimate for the uncertainty in the population mean for each data point (**Supplementary Figs. 8** and **9** and **Supplementary Table 1**). The fitting procedure allowed different steady-state a, b and τ_{delay} values for each day, as we observed some day-to-day variation in these parameters (**Supplementary Note 11**). The predictive model uses the mean of these daily parameter fits. The remaining model parameters were fit globally to single values for the calibration data set. The uncertainties for the a, b and τ_{delay} parameters for the CcaS-CcaR and Cph8-OmpR systems are the standard deviations of the daily fits for each of these parameters. The uncertainties for the remaining parameters are computed as standard errors from the diagonal elements of the covariance matrix resulting from the fit.

Time-course simulations. To simulate the time-course responses of the light-switchable systems, we first calculate an analytical solution to the 2D ODE model. If we use $t' = t - t_{delay}$, we find that

$$p(t') = \begin{cases} p_0 & : t' < 0 \\ c - (c - p_0)e^{-k_p(c)t'} & : t' \geq 0 \end{cases}, \text{ and}$$

$$g(t') = \begin{cases} g_0 & : t' < 0 \\ c - (c - g_0)e^{-k_g t'} - (c - p_0)\dfrac{k_g}{k_p(c) - k_g}(e^{-k_g t'} - e^{-k_p(c)t'}) & : t' \geq 0 \end{cases}$$

doi:10.1038/nmeth.2884

The initial fluorescence and fluorescence production rate are given by g_0 and p_0 and are determined by the preconditioning light intensities. To simulate the response of the system to a sequence of step changes in light, we stitch together a series of individual step-change responses, evaluating each successive p_0 and g_0 at the point where the system begins to respond to the new step change. Thus, if t_{step} is the time at which a step function occurs, then we evaluate the new p_0 and g_0 at $t_{step} + \tau_{delay}$, ensuring the stitched solution will be continuous.

We then perform 500 simulations with model parameters sampling the standard error of the fit. The sampling is performed on a Gaussian distribution of the parameter values with the mean set to the best-fit value and s.d. set to the standard error of the fit. The interquartile range of the expression levels for each minute of the simulations was used to determine the boundaries of the gray shaded envelopes in **Figures 2**–**4**. Model simulations are generated via custom Python scripts (**Supplementary Software**).

Precomputation of light control sequences. The algorithm uses model simulations of the light-switchable systems to compute a sequence of step changes in light intensity that will drive each system to follow a desired reference trajectory (**Supplementary Fig. 11**). The process is implemented using custom Python scripts (**Supplementary Software**). The algorithm accepts as inputs a list of time points at which the step changes should occur, a list of intensities allowed at each time point, the desired reference trajectory and an initial light control program from which to optimize. Then an iterative procedure begins in which a simulation of the response to the current set of light step changes is calculated from the beginning of the first step change up to the end of the next step change plus a user-defined extra time horizon. The integrated square error between this simulation and the reference is then calculated over this same range. The algorithm then modifies the intensity of the first step change and repeats the error calculation given the new intensity. This process is repeated until the intensity that minimizes the error is identified. The algorithm then moves on to the next step change and repeats this intensity-optimization process.

After the intensities of all step changes have been optimized, the integrated square error between the simulation and the reference is calculated for the entire time course. Following this calculation, the algorithm repeats the entire process of optimizing the intensity at each step change, using the previously computed series of step changes as the initial input light sequence. After completing the procedure, the integrated error is again calculated across the entire time course. This process continues repeating until the difference between total integrated square errors for two successive runs is less than 0.5%, or until a user-defined maximum number of iterations have occurred. Typically, the error convergence threshold is satisfied after 3–5 iterations.

Error analysis and model validation. For function-generator experiments, r.m.s. errors (RMSEs) are calculated between the experimental data and the reference signal as well as between the experimental data and the model simulations (**Supplementary Note 12**). For model validation experiments, RMSEs are calculated between the experimental data and simulation. RMSE values are reported as a percentage of the entire output range of the corresponding light-switchable system, allowing better appreciation of the magnitude of the error. RMSEs for all function-generator experiments have been calculated (**Supplementary Table 1**).

39. Miyazaki, K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol.* **498**, 399–406 (2011).
40. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
41. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
42. Tabor, J.J. *et al.* A synthetic genetic edge detection program. *Cell* **137**, 1272–1281 (2009).
43. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).