nature biotechnology

In silico feedback for *in vivo* regulation of a gene expression circuit

Andreas Milias-Argeitis[1,4], Sean Summers[1,4], Jacob Stewart-Ornstein[2,4], Ignacio Zuleta[2], David Pincus[2], Hana El-Samad[2], Mustafa Khammash[3] & John Lygeros[1]

We show that difficulties in regulating cellular behavior with synthetic biological circuits may be circumvented using *in silico* feedback control. By tracking a circuit's output in *Saccharomyces cerevisiae* in real time, we precisely control its behavior using an *in silico* feedback algorithm to compute regulatory inputs implemented through a genetically encoded light-responsive module. Moving control functions outside the cell should enable more sophisticated manipulation of cellular processes whenever real-time measurements of cellular variables are possible.

Regulating a dynamic system using feedback control involves processing measurements of its output in real time to determine appropriate inputs designed to drive its behavior to follow a desired pattern. Because feedback control enables robust regulation in the face of uncertainty and disturbances, it is a recurring theme at every level of organization throughout biology and engineering. Although evolved biological feedback mechanisms found in nature appear robust, engineering feedback control schemes in cells to produce new functions has proven to be a tedious, iterative process[1–6] that has been achieved with limited success. Here we propose the concept of *in silico* feedback control as a complement to feedback circuitry built from biological components. *In silico* feedback uses computational control algorithms running on a digital computer and updated with real-time measurement data. The algorithms prescribe external inputs that achieve and maintain a desired circuit behavior while automatically compensating for circuit variability.

The implementation of *in silico* feedback requires an external input and biological modules that can respond to that input. Working with *S. cerevisiae*, we took advantage of a system that has been used to control diverse biological processes[7–9], the light-responsive Phy/PIF module[10] (**Fig. 1a**,**b**). Upon ligation to the small-molecule chromophore phycocyanobilin (PCB), the plant photoreceptor chromoprotein PhyB undergoes a light-gated interaction with phytochrome interacting factor (PIF). Two fusion constructs—the photosensory domain of PhyB fused to the Gal4 DNA-binding domain (PhyB-GBD) and PIF3 fused to the Gal4 activation domain (PIF3-GAD)—allow one to use red (650 nm) and far-red (730 nm) pulses of light to switch on and off, respectively, the transcription of Gal4-responsive genes in *S. cerevisiae*[10]. In particular, we used cells containing a YFP reporter driven by the Gal1 promoter, which contains Gal4 binding sites.

To probe the dynamic behavior of the system and to devise a computational model for *in silico* feedback control, we modeled the dynamics of the Phy/PIF/Gal system using a simple fourth-order linear ordinary differential equation. We then performed several time-course 'identification experiments' to excite the crucial responses of the system and estimate the model's five free parameters (**Fig. 1c–h** and **Supplementary Methods**). These experiments involved stimulating cells with different trains of red and far-red light pulses delivered using a custom-built 'light pulser' (**Supplementary Fig. 1**) and then obtaining single-cell fluorescence measurements by flow cytometry. In each experiment, we reproducibly initialized the system at its 'stationary condition', characterized by constant growth rate and basal fluorescence level, and we computed the fold change as induction of the system above this initial basal fluorescence (**Supplementary Figs. 2** and **3**). Despite the intrinsic ability of the system to achieve rapid activation and shutoff, the slow maturation dynamics and long half-life of the YFP reporter protein generate a delayed and time-integrated snapshot of the overall system dynamics, making control of the system particularly challenging. The mathematical model with optimized parameters captured the essential features of the data (**Fig. 1d**,**f**,**h**).

Using the model as a starting point, we investigated strategies to robustly regulate *in vivo* the average YFP fluorescence of the Phy/PIF/Gal system to desired levels, or 'set points', in this case seven- and four-fold above basal levels over a 7-h period (**Fig. 2**). First nonfeedback strategies (also known as open-loop control) were investigated. One such strategy is to select a control signal that, according to the model, will achieve the desired objective and then to apply this signal as an input to the process. Following this strategy, we computed trains of light pulses using the data-calibrated model and found that in computer simulations these pulses achieved accurate regulation of YFP (**Fig. 2b**,**c**). Yet in contrast to the simulation results, when these same trains of light pulses were experimentally administered to cells, they failed to regulate YFP intensity to the desired set points. This failure is the result of inaccuracies of the model and inevitable intracellular fluctuations.

An alternative strategy is to exploit the real-time fluorescence measurements to compute the control signal online, as the process evolves. To investigate this *in silico* closed-loop feedback control strategy (**Fig. 2a**), we used YFP signal measurements obtained every 30 min to compute red or far-red pulses to be applied every 15 min. Briefly, we first used a Kalman filter[11]—which uses information from the most recent YFP

[1]Department of Electrical Engineering, ETH Zurich, Automatic Control Laboratory, Zurich, Switzerland. [2]Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California, USA. [3]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. [4]These authors contributed equally to this work. Correspondence should be addressed to H.E.-S. (hana.el-samad@ucsf.edu) or M.K. (mustafa.khammash@bsse.ethz.ch) or J.L. (lygeros@control.ee.ethz.ch).

Received 1 July; accepted 27 September; published online 6 November 2011; doi:10.1038/nbt.2018

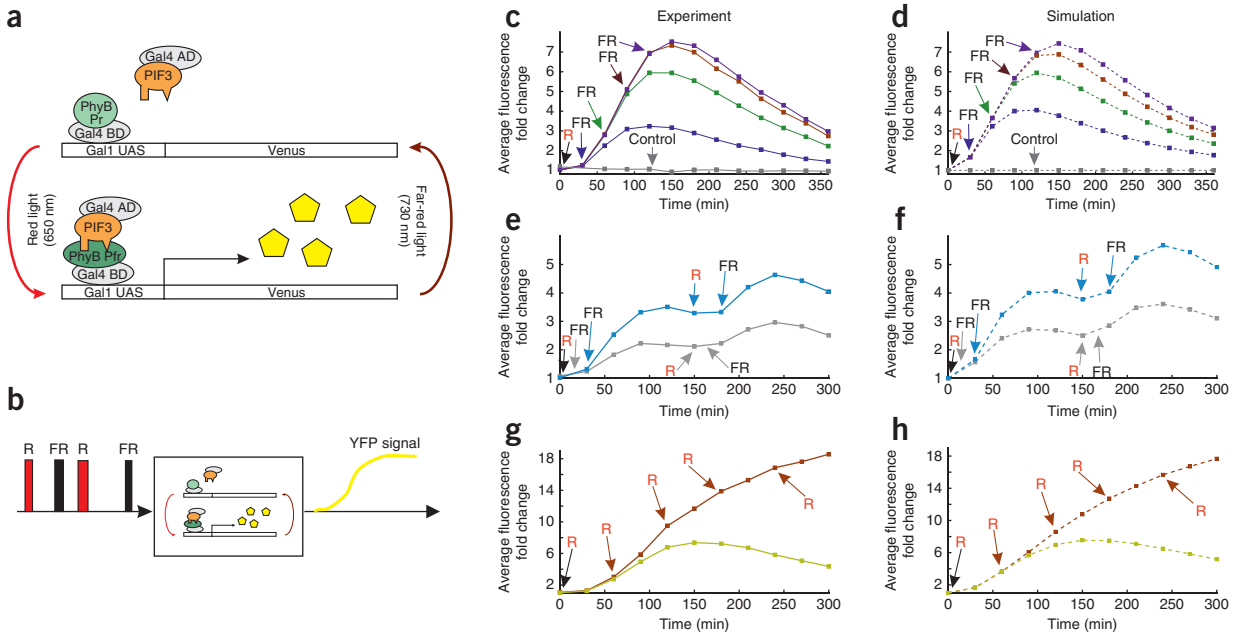

Figure 1 Characterization of the light-switched system. (**a**) Light-switchable gene system based on PhyB-PIF3 interaction. Transformed cells grown in darkness and incubated with the chromophore phycocyanobilin (PCB) synthesize both PhyB(Pr)-GBD and PIF3-GAD fusion proteins. Because PIF3 interacts only with the activated form of PhyB (Pfr), the Gal1 target gene is initially off. Upon exposure to red light, PhyB is rapidly converted into its active Pfr form and binds the PIF3 moiety of PIF3-GAD. The transcription activation domain of Gal4 is therefore recruited to the promoter and induces transcription of the target gene. Exposure to far-red light switches off gene expression by rapidly converting PhyB into its inactive Pr form, causing its dissociation from PIF3-GAD. (**b**) The expression of YFP driven by the Gal1 promoter can be repeatedly switched on and off using a train of red (R) and far-red (FR) pulses. The trains of light pulses can serve as a control input, and the amount of YFP plays the role of a controlled output. (**c**) Experimental dynamics of cell fluorescence in response to a red pulse followed by a far-red pulse. All pulses have 1-min duration. YFP flow cytometry measurements (squares) were taken every 30 min. Each set of matched colored arrow and output squares and curve represent a distinct experiment in which the far-red input was applied at different times. Spontaneous transition of PhyB Pfr to Pr takes place in the dark (a phenomenon known as 'dark reversion') resulting in dissociation of PIF3 from PhyB. Consequently, cell fluorescence reaches a peak and then declines, as mRNA decays over time. Gray squares and line correspond to a control experiment with chromophore addition and no light exposure. (**d**) Simulation of the response to the same input as in **c**. The model reproduces several essential features of the experimental responses, including peak times and decay dynamics. Slight differences between simulated and experimental responses are due to nonlinear effects and delays that are not captured by the model. (**e**) Reversibility of the PhyB-PIF3 interaction. The system does not lose its responsiveness to light over several on-off cycles. (**f**) Simulation results for the same input as in **e**. (**g**) Response to multiple red pulses. Multiple applications of red light drive the system to higher expression levels than a single red pulse. (**h**) Simulation results for multiple- compared with single-pulse responses.

measurements, the system model and knowledge of the pulse history—to generate an estimate of the unmeasured states of the gene expression model. Based on this estimate, the feedback algorithm then calculates, in real time, the train of light pulses that will minimize the deviation of the model-predicted output from the desired fluorescence set point. Following the principle of model predictive control[12], only the first pulse of this pulse train is applied to the system, and the process is repeated. This scheme successfully achieved the desired regulation levels as evidenced by the robust seven- and fourfold induction (**Fig. 2b**,**c** and **Supplementary Fig. 4**). Thus, in contrast to the open-loop strategy, *in silico* closed-loop feedback robustly achieved and maintained the desired set points despite modeling errors and biological fluctuations.

To further demonstrate the robustness of the *in silico* feedback design, we perturbed the system with a train of red and far-red pulses over a period of 3 h to drive YFP expression to an unknown initial condition. We then activated the *in silico* feedback controller with the task of regulating YFP fluorescence to fivefold above basal over a 7-h period; in all cases, regulation was robustly achieved (**Fig. 2d** and **Supplementary Fig. 4**). Notably, in this case, an open-loop train of light pulses, even one based on a high-fidelity model of the process[13], could not possibly be used because the initial state of the system was unknown.

Until feedback control of biological systems can be genetically encoded *in vivo* to behave deterministically and without cross-talk, biology is far from being a predictable engineering medium like electronics. But by interfacing electronic control with biological responses, *in silico* feedback provides an approach for unprecedented, quantitative control over the activity of living cells. This technology should find wide application in systems and synthetic biology. Probing endogenous biological circuits often yields unpredictable results because of compensatory cellular regulation that alters the expression in dynamic and complex ways. In such scenarios, an *in silico* feedback module could be used to ensure that protein expression remains tightly regulated at the desired level, independent of other intervening processes. The magnitude of the *in silico* feedback required to maintain a constant protein level could be used as a surrogate for the endogenous feedback present in the pathway and as a measure of its homeostatic capabilities. The ability to exert precise *in silico* control of engineered biological circuits based on direct readings of intracellular states will also facilitate the use of synthetic systems for biotechnological applications. For example, in the production of biofuels or small-molecule drugs, an *in silico* feedback module could be deployed to regulate the levels of toxic by-products that invariably ensue from manipulation of metabolic pathways. Finally, *in silico* feedback offers exciting therapeutic opportunities. As real-time physiological readouts become available, one can envision the possibility of using real-time closed-loop control to achieve more regulated and precise interventions.

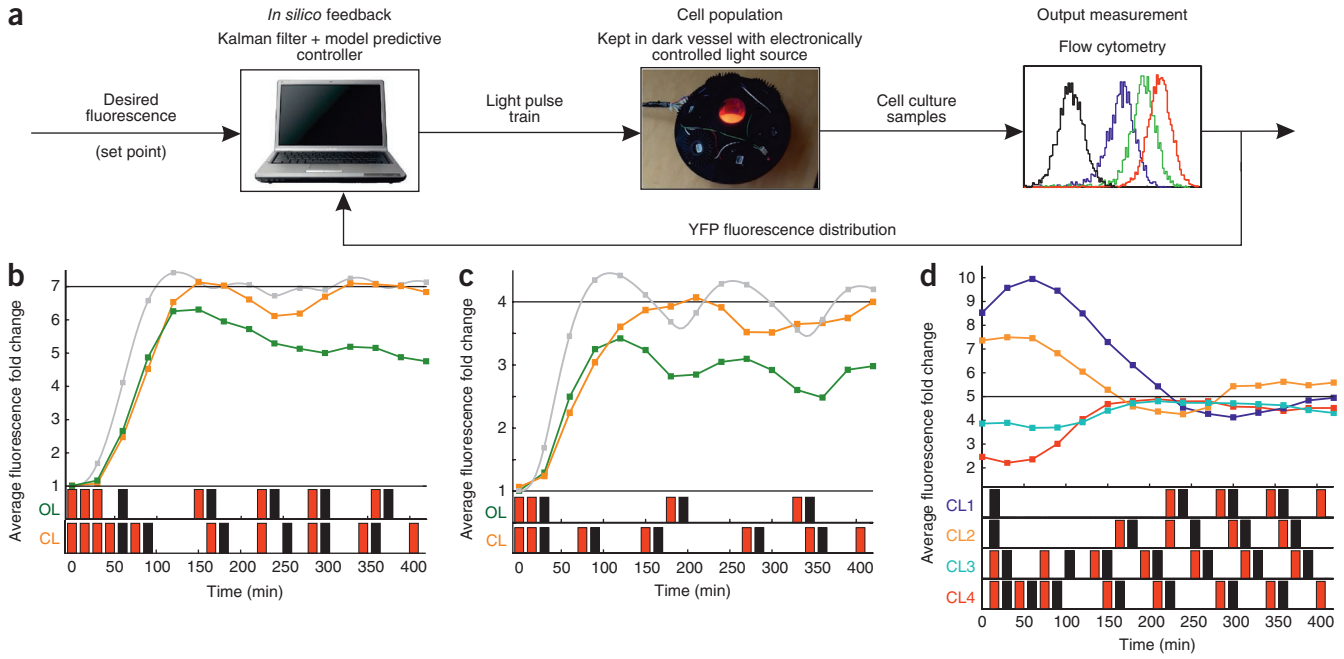

Figure 2 *In silico* feedback achieves robust regulation of gene expression fold change. (**a**) *In silico* feedback control scheme for the light-activated gene system. (**b**) Regulation of average YFP fluorescence to sevenfold over a 7-h period using *in silico* feedback (orange). A pre-computed light pulse train that achieves set point regulation when applied to the mathematical model (gray) did not achieve the desired fold induction when applied in open loop to the biological construct (green). In contrast, closed-loop feedback control achieves the desired fold induction. OL and CL denotes open- and closed-loop control, respectively. (**c**) Regulation of average YFP fluorescence to fourfold above basal over a 7-h period. Open- and closed-loop pulse trains determined as in **b**. (**d**) Regulation of average YFP to fivefold above basal over a 7-h period, starting from a randomly perturbed culture. Closed-loop control achieves the desired set point, irrespective of the initial conditions of the system.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We would like to thank P. Quail (UC, Berkeley) for the generous gift of the PIF3 construct and W. Lim (UCSF) and J. Stelling (ETH, Zurich) for providing PCB. The work was supported by National Science Foundation grant CCF-0943385 (H.E.-S.), ECCS-0835847 (M.K.), and MoVeS FP7-ICT-2009-257005 (J.L.).

AUTHORS CONTRIBUTIONS

M.K. originated the concept. A.M.-A., S.S., J.S.-O., H.E.-S., J.L. and M.K. carried out the project design. A.M.-A. and S.S. did the control design. J.S.-O. did the cloning and strain construction. S.S., A.M.-A. and I.Z. designed and built the electronics. S.S. and A.M.-A. performed the experiments with J.S.-O. and with help from D.P.; H.E.-S., M.K. and J.L. supervised the project; all authors contributed to the writing of the manuscript. Corresponding authors are listed in the affiliations in alphabetical order.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nbt/index.html.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Lim, W.A. *Nat. Rev. Mol. Cell Biol.* **11**, 393–403 (2010).
2. Andrianantoandro, E., Basu, S., Karig, D.K. & Weiss, R. *Mol. Syst. Biol.* **2** 2006.0028 (2006).
3. Benner, S.A. & Sismour, A.M. *Nat. Rev. Genet.* **6**, 533–543 (2005).
4. Khalil, A.S. & Collins, J.J. *Nat. Rev. Genet.* **11**, 367–379 (2010).
5. Purnick, P.E. & Weiss, R. *Nat. Rev. Mol. Cell Biol.* **10**, 410–422 (2009).
6. Haynes, K.A. & Silver, P.A. *J. Cell Biol.* **187**, 589–596 (2009).
7. Levskaya, A., Weiner, O.D., Lim, W.A. & Voigt, C.A. *Nature* **461**, 997–1001 (2009).
8. Tyszkiewicz, A.B. & Muir, T.W. *Nat. Methods* **5**, 303–305 (2008).
9. Leung, D.W., Otomo, C., Chory, J. & Rosen, M.K. *Proc. Natl. Acad. Sci. USA* **105**, 12797–12802 (2008).
10. Shimizu-Sato, S., Huq, E., Tepperman, J.M. & Quail, P.H. *Nat. Biotechnol.* **20**, 1041–1044 (2002).
11. Kalman, R.E. *J. Basic Eng.* **82**, 35–45 (1960).
12. Morari, M. *Comput. Chem. Eng.* **23**, 667–682 (1999).
13. Sorokina, O. *et al. J. Biol. Eng.* **3**, 15 (2009).