Modular Design of Artificial Tissue Homeostasis: Robust Control through Synthetic Cellular Heterogeneity

Miles Miller[1❡], **Marc Hafner**[1,2,3,4❡¤], **Eduardo Sontag**[5], **Noah Davidsohn**[1], **Sairam Subramanian**[6], **Priscilla E. M. Purnick**[7], **Douglas Lauffenburger**[1], **Ron Weiss**[1*]

1 Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **2** School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, **3** Department of Biochemistry, University of Zurich, Zurich, Switzerland, **4** Swiss Institute of Bioinformatics, Lausanne, Switzerland, **5** Department of Mathematics, Rutgers University, New Brunswick, New Jersey, United States of America, **6** Department of Electrical Engineering, Princeton University, Princeton, New Jersey, United States of America, **7** Department of Gene and Cell Medicine, Black Family Stem Cell Institute, Mount Sinai School of Medicine, New York, New York, United States of America

Abstract

Synthetic biology efforts have largely focused on small engineered gene networks, yet understanding how to integrate multiple synthetic modules and interface them with endogenous pathways remains a challenge. Here we present the design, system integration, and analysis of several large scale synthetic gene circuits for artificial tissue homeostasis. Diabetes therapy represents a possible application for engineered homeostasis, where genetically programmed stem cells maintain a steady population of β-cells despite continuous turnover. We develop a new iterative process that incorporates modular design principles with hierarchical performance optimization targeted for environments with uncertainty and incomplete information. We employ theoretical analysis and computational simulations of multicellular reaction/diffusion models to design and understand system behavior, and find that certain features often associated with robustness (e.g., multicellular synchronization and noise attenuation) are actually detrimental for tissue homeostasis. We overcome these problems by engineering a new class of genetic modules for 'synthetic cellular heterogeneity' that function to generate beneficial population diversity. We design two such modules (an asynchronous genetic oscillator and a signaling throttle mechanism), demonstrate their capacity for enhancing robust control, and provide guidance for experimental implementation with various computational techniques. We found that designing modules for synthetic heterogeneity can be complex, and in general requires a framework for non-linear and multifactorial analysis. Consequently, we adapt a 'phenotypic sensitivity analysis' method to determine how functional module behaviors combine to achieve optimal system performance. We ultimately combine this analysis with Bayesian network inference to extract critical, causal relationships between a module's biochemical rate-constants, its high level functional behavior in isolation, and its impact on overall system performance once integrated.

Citation: Miller M, Hafner M, Sontag E, Davidsohn N, Subramanian S, et al. (2012) Modular Design of Artificial Tissue Homeostasis: Robust Control through Synthetic Cellular Heterogeneity. PLoS Comput Biol 8(7): e1002579. doi:10.1371/journal.pcbi.1002579

Editor: Greg Tucker-Kellogg, National University of Singapore, Singapore

Received November 30, 2011; **Accepted** April 25, 2012; **Published** July 19, 2012

Copyright: © 2012 Miller et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by NIH NIGMS grant R01GM086881, NSF Award \#1001092 for RW, an NSF Graduate Research Fellowship Program for MM and the SystemsX.ch grant for the IPhD Project "Quantifying Robustness of Biochemical Modules to Parametric and Structural Perturbations" for MH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rweiss@mit.edu

❡ These authors contributed equally to this work.

¤ Current address: Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, United States of America

Introduction

One of the key challenges facing synthetic biology today is the ability to engineer large-scale, multicellular systems with sophisticated yet predictable and robust behaviors. Previous work in synthetic biology has successfully implemented and characterized a variety of relatively small synthetic gene networks including oscillators [1–4], toggle switches [5,6], and intercellular sender/receiver or quorum sensing (QS) communication systems [7–9]. Computational tools have encouragingly demonstrated an ability to guide experimental optimization of several of such modules [10,11], and some recent projects have successfully integrated a few of these 'standard modules' and interfaced them with endogenous pathways to program more sophisticated behaviors [12–18]. Ultimately, however, the path to success will require bridging the gap between specifying sophisticated systems-level objectives and a list of molecular parts and interactions that can be properly assembled to accomplish these objectives [19]. To address this challenge, here we present and apply a novel combination of computational methods to aid the iterative design and optimization of synthetic biological systems. Importantly, these tools address issues stemming from the incomplete and imprecise knowledge of rate constants and cellular context.

As a case study, we design a system to control tissue homeostasis, broadly defined as the property of balancing growth, death, and differentiation of multiple cell-types within a multicellular community. Tissue homeostasis represents an important class of problems in biology, and the ability to control it is fundamental to the success of a wide range of tissue engineering goals. At the same time the ability to create and analyze such a system may

Author Summary

Over the last decade several relatively small synthetic gene networks have been successfully implemented and characterized, including oscillators, toggle switches, and intercellular communication systems. However, the ability to engineer large-scale synthetic gene networks for controlling multicellular systems with predictable and robust behavior remains a challenge. Here we present a novel combination of computational methods to aid the iterative design and optimization of such synthetic biological systems. We apply these methods to the design and analysis of an artificial tissue homeostasis system that exhibits coordinated control of cellular proliferation, differentiation, and cell-death. Achieving artificial tissue homeostasis would be therapeutically relevant for diseases such as Type I diabetes, for instance by transplanting genetically engineered stem cells that stably maintain populations of insulin-producing beta-cells despite normal cell death and autoimmune attacks. To manage complexity in the design process, we employ principles of logic abstraction and modularity and investigate their limits in biological networks. In this work, we find factors often associated with robustness (e.g., multicellular synchronization and noise attenuation) to be actually detrimental, and overcome these problems by engineering genetic modules that generate beneficial population heterogeneity. A combination of computational methods elucidates how these modules function to enhance robust control, and provides guidance for experimental implementation.

provide insight into mechanisms of endogenous tissue homeostasis and its misregulation in diseases such as cancer and diabetes. For example, misregulation of tissue homeostasis plays a central role in Type I diabetes, in which natural populations of insulin-producing β-cells are destroyed due to autoimmune defects. Automated mechanical systems have been proposed for insulin control in diabetes but still face significant challenges including long-term efficacy [20]. Stem cell and β-cell transplantations have also been studied as possible solutions [21,22], but the last decade of results suggest that the transplanted cells fail to maintain homeostasis and become either tumorigenic or depleted within months [23].

Approach

As potential solutions for this problem, we propose several increasingly robust variants of a synthetic gene network that are designed to maintain a steady level of β-cells despite normal cell death and constant destruction of the β-cells by the immune system. The synthetic gene networks continuously direct proliferation, quiescence, and stem cell differentiation into insulin producing β-cells as needed (Figure 1A). The resulting engineered circuits may be employed to regulate tissue homeostasis both *in vitro* where the cell culture is removed from natural cues, and *in vivo* when natural systems fail or tissue is ectopically transplanted (for example, the Edmonton protocol involves implanting pancreatic islets including β-cells to the liver [24]).

The efforts described here are based on encouraging genetic engineering accomplishments that have demonstrated population control of bacteria and yeast [12,25], mammalian cell proliferation [26], and stem cell differentiation [27,28]. To mitigate some of the uncertainties involved in system construction, we restricted our designs to use only genetic parts and modules that have already been demonstrated experimentally. These include engineered cell-cell communication to determine population densities, a toggle switch, an oscillator, and a multi-input AND gate.

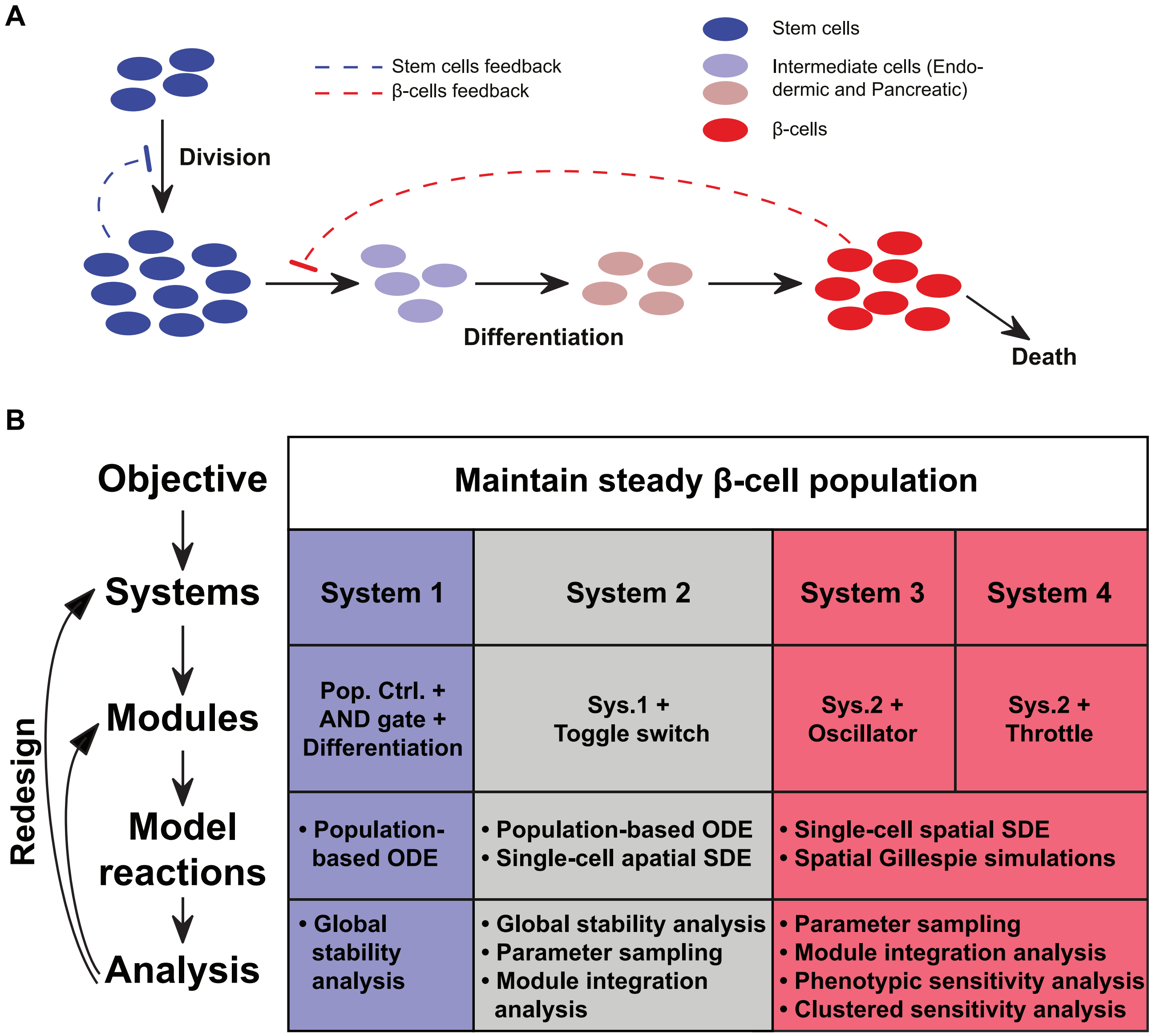
To gain a detailed understanding of our proposed synthetic gene networks, we carried out theoretical analysis and computational simulations using Ordinary Differential Equations (ODE's), Langevin, and Gillespie algorithms. The analysis revealed that while simple modular composition was useful for initial system design, various factors such as stochastic effects, feedback control, and module interdependence significantly impacted system function and hence had to be taken into account when evaluating system designs. Strikingly, we observed that system features typically associated with robustness, including cell-synchronization, noise attenuation, and rapid signal processing destabilized our systems. To overcome these problems, we propose and analyze mechanisms that generate population diversity, and through this symmetry breaking facilitate proportionate and homeostatic system response to population-wide cues. Endogenous mechanisms of cellular heterogeneity have been previously observed in many physiological processes, including differentiation [29]. In the synthetic biology context, however, these mechanisms may be either unavailable for integration into the synthetic genetic circuit or too poorly understood to fully utilize. As a result, we forward engineer modules to generate synthetic cellular heterogeneity. For example, we incorporate an asynchronous oscillator module into the design as an engineered generator of intrinsic variability. Ultimately, our analysis indicates that such modules greatly improve homeostatic robustness among an isogenic population of cells, and we identify several examples of natural analogs.

Key results

We found that the design and optimization of modules for synthetic heterogeneity is both non-intuitive and multifactorial, and in general requires a framework for non-linear and multivariate analysis. For example, with the asynchronous oscillator, we could not *a priori* define a simple objective or ideal 'phenotype' since oscillator properties such as period, dynamic range, and asynchronicity affected overall system performance in complex and interdependent manners. Furthermore, even if ideal module phenotypes are known, understanding the physical parameters required to achieve such phenotypes also represents a challenge. To address these issues, we developed a new framework using a hierarchy of computational tools to understand the optimal phenotypic and physical characteristics of the synthetic heterogeneity modules with respect to overall system behavior. We developed a 'phenotypic sensitivity analysis' method to determine how functional module behaviors combine to achieve optimal system performance. Parametric sensitivity analysis then captures the dependency of a module's phenotypes on its underlying physical rate constants. Ultimately, we integrated both analyses using Bayesian network inference to extract critical, causal relationships between a module's biochemical rate constants, its high level functional behavior in isolation, and its impact on overall system performance once integrated. Importantly, we anticipate that our hierarchical optimization strategy prescribes directions for system design that readily apply to experimental systems facing high degrees of uncertainty in rate constants and cellular environment.

Outline

We designed and modeled an artificial tissue homeostasis system where a population of self-renewing stem cells grow and differentiate in a regulated manner to sustain a steady population of adult cells which, in this case, are insulin-producing β-cells (Figure 1A). Here we present four iterations of system design, analysis, and redesign with increased sophistication for improved



Objective	Maintain steady β-cell population			
Systems	System 1	System 2	System 3	System 4
Modules	Pop. Ctrl. + AND gate + Differentiation	Sys.1 + Toggle switch	Sys.2 + Oscillator	Sys.2 + Throttle
Model reactions	• Population-based ODE	• Population-based ODE • Single-cell apatial SDE	• Single-cell spatial SDE • Spatial Gillespie simulations	
Analysis	• Global stability analysis	• Global stability analysis • Parameter sampling • Module integration analysis	• Parameter sampling • Module integration analysis • Phenotypic sensitivity analysis • Clustered sensitivity analysis	

Figure 1. Overview of system design. (A) The general tissue homeostasis design. Proliferation of stem cells (blue) is regulated by their population size through negative feedback (dashed blue line). Sequential differentiation into endodermic, pancreatic, and finally *β*-cells (red) occurs when the stem cell population has sufficient size, and is governed through negative feedback from differentiated cells (dashed red line). (B) Design workflow. Starting with a high-level objective, iterative design proceeds through a top-down decomposition into modules and then basic reactions of the system, followed by analysis and redesign (left). The table columns (right) show the four iterations of system designs presented in this work. Table rows describe the top-down decomposition for each system, and correspond to the workflow at left.
doi:10.1371/journal.pcbi.1002579.g001

robustness in controlling tissue homeostasis (Figure 1B). The initial model for artificial tissue homeostasis (System 1) comprises four integrated modules, and is analyzed using ODE simulation and global stability analysis. We incorporate a toggle switch in System 2 to minimize undesired *β*-cell population fluctuations observed in System 1, and analyze the improved design using stochastic differential equations (SDEs). Although System 2 represents an improvement, its homogeneous response to commitment cues results in poor performance, thereby motivating the incorporation of an oscillator module and a throttle module for Systems 3 and 4, respectively. Using SDE simulations, we optimize these modules and their integration into the full system. Throughout the discussion, we focus on several aspects of system design, including module integration, optimization of rate constants for individual modules, and optimization of module phenotypic behaviors.

Results

Iterative system design and analysis

Simple mathematical analysis suggested that feedback regulation between the two populations of stem cells and adult cells was necessary for robust homeostatic control, and recent work has explored the essential role of feedback control in stem cell biology (Text S1, Sec. 2.1, [30]). In all alternative system designs presented in this manuscript, we implemented feedback control through artificial cell-cell communication pathways. Our first design, System 1, allows differentiation only with a high density of stem cells and a low density of β-cells (Figure 2A). The "Stem Cell Population Control" (SPC) module allows for differentiation only when the population density of self-renewing cells lies above some threshold. We also designed the SPC to suppress proliferation through the expression of a growth arrest factor (GAF), currently under development in the Weiss lab. The "β-Cell Population Control" (BPC) module produces high output and inhibits differentiation when the density of β-cells reaches a threshold (Figure 2A). We based the cell-cell communication systems in the SPC and BPC modules on previously described communication systems [7–9]. As a proof of concept, Supplementary Figure S1 A–B presents results for a signal-receiver circuit based on the LuxR protein that responds to 3-oxo-hexanoyl-homoserine lactone (3OC6HSL), that has been experimentally implemented in human embryonic kidney (HEK293) cells.

We model stem cell differentiation as a multistage process that can take several weeks to complete [31]. For example, directed *in vitro* differentiation of hES cells into insulin-producing cells involves stepwise administration of growth factors to first induce endodermal cell fate, followed by pancreatic specialization, expansion, and maturation [32]. This general process is modeled by four cell types: stem cells (population size S) grow with a constant division rate k_b. Upon maturation, they proceed through two intermediate populations of endodermic (E) and pancreatic (P) cells before becoming β-cells (B), which die at a constant rate k_k. We describe the sequential maturation of S into E, P, and B as first-order reactions with rates k_{c1}, k_{c2}, and k_d. Feedback terms are modeled as Hill functions, where K_S and K_B represent the SPC and BPC module thresholds, respectively.

$$
\begin{aligned}
\frac{dS}{dt} &= k_b S(t)\cdot\frac{K_S^n}{K_S^n+S(t)^n} - k_{c1}S(t)\cdot\frac{S(t)^n}{K_S^n+S(t)^n}\cdot\frac{K_B^n}{K_B^n+B(t)^n},\\
\frac{dE}{dt} &= k_{c1}S(t)\cdot\frac{S(t)^n}{K_S^n+S(t)^n}\frac{K_B^n}{K_B^n+B(t)^n} - k_{c2}E(t),\\
\frac{dP}{dt} &= k_{c2}E(t) - k_d P(t),\\
\frac{dB}{dt} &= k_d P(t) - k_k B(t)
\end{aligned}
\qquad (1)
$$

The differentiation process is generally long *in vivo* (e.g., 20 days [32]. For System 1, such delay in the feedback could induce undesirable oscillations (Figure 2B–C). As a result, System 1 failed to maintain homeostasis for a large range of parameter values (Text S1, Sec. 2).

Toggle switch facilitates rapid feedback but neglects the heterogeneity requirement. System 2 minimizes feedback delay by using a 'commitment' module to decouple the BPC module from the slow differentiation process (Figure 3A). Commitment occurs through a one-way toggle switch, which we

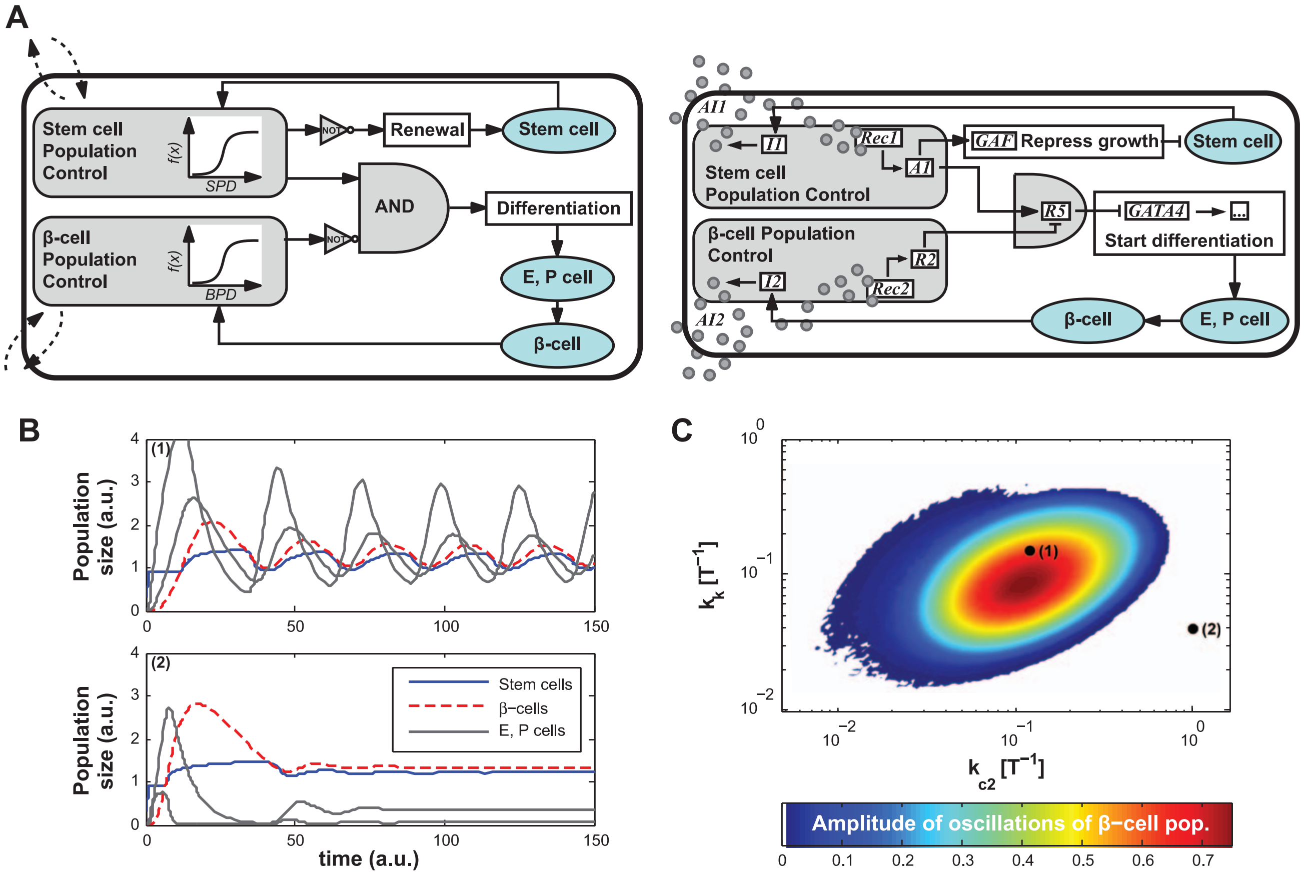

A
Stem cell Population Control
f(x)
SPD
β-cell Population Control
BPD
NOT
NOT
Renewal
Stem cell
AND
Differentiation
E, P cell
β-cell
AI1
I1
Rec1
A1
GAF
Repress growth
Stem cell
Stem cell Population Control
R5
GATA4
...
Start differentiation
β-cell Population Control
R2
I2
Rec2
AI2
β-cell
E, P cell
B
(1)
Population size (a.u.)
(2)
Stem cells
β–cells
E, P cells
time (a.u.)
0
50
100
150
0
1
2
3
4
C
10^0
10^{-1}
10^{-2}
k_k [T^{-1}]
k_{c2} [T^{-1}]
(1)
(2)
Amplitude of oscillations of β−cell pop.
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Figure 2. System 1. (A) Circuit diagram: two Population Control modules (in gray) sense the density of stem- and β-cells. The AND gate integrates the output of the modules to induce differentiation. Circles represent intercellular signaling molecules. (B) Two examples of population evolution showing sustained oscillations (point 1 in *C*) and a stable steady state (point 2 in *C*), with other parameters fixed (SI Sec. 2 and Figure S2). (C) A planar slice of the parameter space where population oscillations occur for System 1.
doi:10.1371/journal.pcbi.1002579.g002

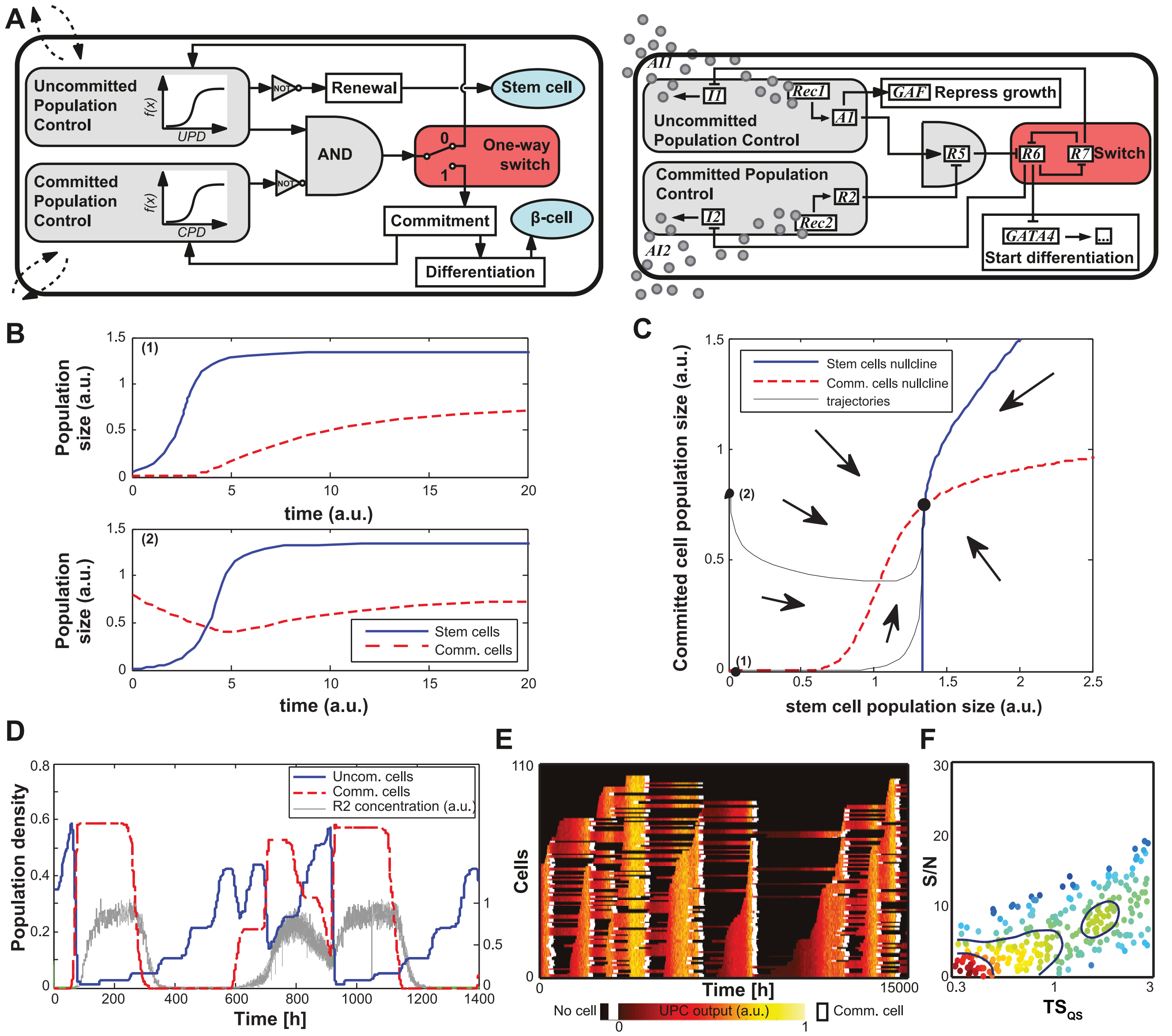


Figure 3. System 2. (A) Circuit diagram: two Population Control modules sense the density of stem and committed cells. The AND gate integrates the output of the modules to induce commitment through the switch state (red module). (B) Deterministic time trajectories for System 2 with two different initial conditions: both converge to the same equilibrium populations. (C) Phase space diagram: all trajectories converge to a unique equilibrium point. Black lines correspond to trajectories plotted in *B*. See Text S1, Sec. 2 and Figure S3 for other phase space diagrams. (D) Stochastic trajectories for a simulation starting with a small stem cell population, showing the output of the Committed Population Control module ($R2$) in representative uncommitted cells (right axis, a.u.). (E) Individual rows track the single-cell UPC module output ($A1$, shown as a heat map) in uncommitted cells within a population. White signifies single-cell commitment, followed by black "null space" that is filled by newly divided uncommitted cells. As soon as UPC output is high (yellow), stem cells commit *en masse*. (F) Overall system performance, S/N, as a function of the module time-scale for cell communication, TS_{QS}. Several hundred different sets of time-scales were tested, with all time-scale parameters simultaneously varied. Each point represents an individual set of time-scales. Color and contour lines indicate point density.

doi:10.1371/journal.pcbi.1002579.g003

designed to reflect earlier computational models [33] and an *E. coli* implementation [34]. As a first step and proof of concept, Supplementary Figure S1 C–D presents an experimental implementation of the proposed toggle switch in human cells. In System 2, the toggle activates both differentiation and population feedback, such that the feedback control is immediately downstream of the toggle switch rather than following the full differentiation process (Figure 3A). The state of the one-way switch defines whether or not the cell has irreversibly committed to differentiate, and this status feeds back into what we now term the "Uncommitted Population Control" (UPC) and "Committed Population Control" (CPC) modules. The density of cells in any stage of the differentiation process determines CPC module output. Consequently, we gained a faster feedback response in exchange for assuming that a relatively constant fraction of cells successfully differentiate upon commitment. Accordingly, in our model for System 2, the rate of the first stage of differentiation ($S \rightarrow E$ in Eq. (1)) is now (other equations remain the same):

$$\begin{aligned}\frac{dS}{dt} &= k_b S(t)\cdot\frac{K_S^n}{K_S^n+S(t)^n}\\ &\quad -k_{c1}S(t)\cdot\frac{S(t)^n}{K_S^n+S(t)^n}\cdot\frac{K_C^n}{K_C^n+(E(t)+P(t)+B(t))^n}\end{aligned} \tag{2}$$

Compared to System 1, the population sizes quickly equilibrated in System 2 (Supplementary Figure S2). We further tested different initial conditions and parameter vectors, and found the System 2 equilibrium point to be independent of the initial conditions (Text S1, Sec. 2.2).

For subsequent analyses, we simplified our model to a two-population system. Given that E, P and B populations are identical with respect to feedback, we merged them into the committed population C, resulting in the following equations for System 2:

$$\begin{aligned}\frac{dS}{dt} &= k_b S(t)\cdot\frac{K_S^n}{K_S^n+S(t)^n}-k_d S(t)\cdot\frac{S(t)^n}{K_S^n+S(t)^n}\cdot\frac{K_C^n}{K_C^n+C(t)^n},\\ \frac{dC}{dt} &= k_d S(t)\cdot\frac{S(t)^n}{K_S^n+S(t)^n}\cdot\frac{K_C^n}{K_C^n+C(t)^n}-\overline{k}_c C(t)\end{aligned} \tag{3}$$

In this ODE model, the actual β-cell population B is a fraction of this committed population C (Text S1, Sec. 2.3). We note that a two population system may not fully restore the complexity of a four population system, for example by precluding chaotic behavior. Nonetheless, the two population model showed a qualitatively similar behavior in the working range of our system. Henceforth, we focused on maintaining a constant population of committed rather than differentiated cells. Indeed, this system demonstrated a stable equilibrium point in a large range of conditions (Figure 3B–C, S3 and proof in Text S1, Sec. 2.4).

Our deterministic model of continuous population dynamics suggested that System 2 stabilized homeostasis sufficiently. However, low molecular count, small population size, and localized reaction/diffusion may constitute critical determinants of system dynamics [35,36]. To obtain an improved understanding of how these factors affect system behavior, we performed spatiotemporally-resolved simulations of multicellular populations using stochastic differential equations (Text S1, Sec. 3), assuming Hill functions for inhibition and activation relations.

These simulations revealed that phenotypic homogeneity within the isogenic stem cell population impedes system performance. More specifically, strong population-wide cues to commit may cause massive simultaneous commitment, thereby depleting the stem-cell pool and leading to homeostasis failure (Figure 3D–E). To quantify system performance, we employed a signal to noise ratio (S/N) metric (inverse of the coefficient of variation, see Text S1, Sec. 3) that denotes how steady the committed population density is maintained. As an initial analysis of overall system robustness, we explored how S/N was affected by variations in the time-scales with which individual modules operate. We lumped system parameters according to their module (Text S1, Sec. 5.2.1, and Table S2) and adjusted them in a coordinated manner to change only how fast a module processed incoming signals and produced the appropriate output, while keeping steady-state behavior of individual modules constant. Perturbing time-scales for modules such as the toggle switch and cell-cell communication randomly and simultaneously allowed us to observe how robust S/N was across the range of time-scales. For System 2, S/N was very sensitive to module time-scales, and most combinations of time-scales resulted in a poorly functional system (Figure 3F). Relative to other processes in the system, rapid feedback kinetics described by the 'quorum sensing' cell-cell communication time-scale (TS_{QS}) could decrease the simultaneous commitment observed in Figure 3D–E, but it may not be possible to implement such a fast response in practice. Moreover, significant environmental perturbations to the system, for example resulting from injury or elevated autoimmune response, could still provoke situations where System 2 fails to maintain homeostasis. We therefore implemented synthetic modules that generate phenotypic heterogeneity in an isogenic population. These modules desynchronize single-cell responses to population-wide signaling cues, thereby facilitating a proportionate and homeostatic system response and balancing the necessity for a fast quorum sensing.

Oscillator stabilizes through asynchronicity. In System 3, we incorporated an asynchronous oscillator (e.g. [3,4]) into the design as a generator of intrinsic heterogeneity (Figure 4 A). In this system, a cell's commitment to differentiation can only occur when its oscillator peaks (and $R4$ concentration is low). Stochasticity drives individual oscillators out of phase, and coupling the oscillator to cell-fate decisions prevents cells in a population from all simultaneously responding to homogeneous commitment signals. Simulations indicated that with the oscillator, our system maintained tissue homeostasis robustly despite the fact that feedback signaling cues to commit remained synchronized even after homeostasis was established (Figure 4B–C). Compared to System 2, System 3 behaved much more robustly to variations in module time-scales, with more than double the S/N of System 2 when averaged across all tested time-scales (Figure 4D). Although our analysis suggested that the oscillator would be a powerful addition to the system design, unforeseen experimental factors may hamper its successful implementation. Unaccounted for drivers of oscillator synchronization across a population, for example, could negatively impact system performance. To address this issue, we developed an alternate strategy for generating population diversity (System 4). Subsequent analysis of these systems then allows us to compare their specific advantages and disadvantages.

Commitment throttle stabilizes through local inhibition. System 4 achieves population heterogeneity through rapid lateral inhibition acting as a throttle on the commitment process during toggle switching (Figure 4E). Through this mechanism, a cell starting to commit blocks the commitment process of adjacent cells. The throttle approach requires a third intercellular signaling molecule that is synthesized transiently while the toggle switches and temporary inhibits neighboring cells from committing likewise. The rest of the circuit remains similar to previous systems. Simulations indicated that when populations reached their steady state values, the throttle mechanism prevented simultaneous commitment of too many cells and therefore maintained homeostasis (Figure 4F–G). Consequently, System 4, like System 3, behaved more robustly to variations in module time-scales compared to System 2 (Figure 4H). Although Systems 3 and 4 clearly outperformed System 2 in these simulations, appreciable differences in time-scale robustness among the three systems warrant further analysis, and the following section explores this from a multivariate perspective.

Robustness analysis and optimization

The integration of several network modules presents a challenge on multiple levels, especially in the context of uncertain biological environments and complex module dynamics. In the following sections, we introduce a framework composed of computational modeling and analysis techniques that addresses these issues in

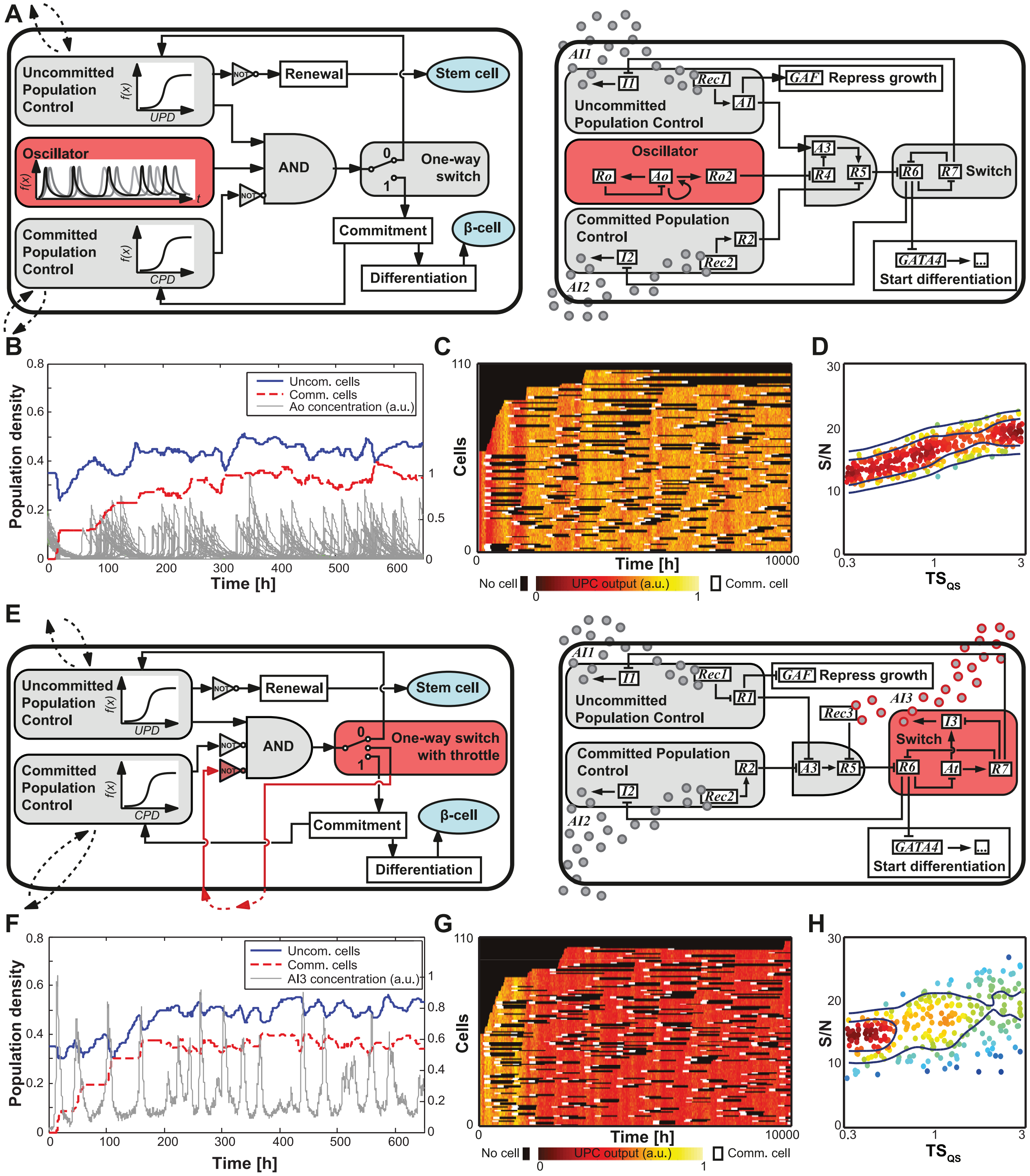

A
Uncommitted Population Control
f(x)
UPD
NOT
Renewal
Stem cell
Oscillator
f(x)
t
AND
NOT
0
1
One-way switch
Commitment
β-cell
Committed Population Control
f(x)
CPD
Differentiation
AI1
I1
Rec1
A1
GAF
Repress growth
Uncommitted Population Control
Oscillator
Ro
Ao
Ro2
A3
R4
R5
R6
R7
Switch
Committed Population Control
R2
I2
Rec2
GATA4
Start differentiation
AI2
B
Population density
Uncom. cells
Comm. cells
Ao concentration (a.u.)
Time [h]
C
Cells
Time [h]
No cell
UPC output (a.u.)
Comm. cell
D
S/N
TS_QS
E
Uncommitted Population Control
Renewal
Stem cell
AND
One-way switch with throttle
Commitment
β-cell
Committed Population Control
Differentiation
AI1
I1
Rec1
R1
GAF
Repress growth
Uncommitted Population Control
Rec3
AI3
I3
Switch
Committed Population Control
R2
A3
R5
R6
At
R7
I2
Rec2
AI2
GATA4
Start differentiation
F
Population density
Uncom. cells
Comm. cells
AI3 concentration (a.u.)
Time [h]
G
Cells
Time [h]
No cell
UPC output (a.u.)
Comm. cell
H
S/N
TS_QS

Figure 4. Systems 3 and 4. (A) Circuit diagram for System 3: in addition to System 2 modules, the AND gate integrates the output of the oscillator (red module) that allows commitment only when peaking. (B) Time trajectories for a simulation starting with a small stem cell population. The oscillator activator (Ao) is plotted for some representative stem cells (right axis, a.u.). (C) Individual rows track the single-cell UPC module output ($A1$, shown as a heat map) in uncommitted cells within a population. White signifies single-cell commitment, followed by black "null space" that is filled by newly divided uncommitted cells. Due to the oscillator, only a fraction of the cells commit when the $A1$ concentration is high. (D) Overall system performance, S/N, as a function of the module time-scale for cell communication, TS_{QS}. Several hundred different sets of time-scales were tested, with all time-scale parameters simultaneously varied. Each point represents an individual set of time-scales. Color and contour lines indicate point density. (E) Circuit diagram for System 4: a throttle mechanism (red module) activates during a cell's commitment and represses commitment in its neighbors. (F–G) Time trajectories for a simulation starting with a small stem cell population, where B shows the average throttle signaling component ($AI3$) in the external medium (right axis, a.u.) over time. (H) S/N as a function of the module time-scale for cell communication, TS_{QS}.
doi:10.1371/journal.pcbi.1002579.g004

optimizing Systems 2, 3 and 4. We first study overall system robustness to external parameters such as cell survival dynamics, and introduce time-scale analysis as a method for guiding module integration. We then optimize the population control module using a novel 'clustered sensitivity analysis' to comprehend global patterns of parametric sensitivity in the context of a detailed biochemical model. Finally, we analyze the synthetic heterogeneity modules with an approach that focuses on module phenotype rather than rate constants alone. Comparisons among the different system architectures ultimately provide guidance for experimental optimization.

Synthetic heterogeneity enhances robustness to noise and cell survival times. We first explored the impact of stochasticity on homeostasis by adjusting the simulated cell volume, Ω, which is related to the number of molecules in each cell (Text S1, Sec. 3). Increasing noise, by decreasing Ω, impacted homeostasis performance both positively and negatively, depending on several factors. Without either the oscillator or throttle, System 2's S/N value decreased monotonically with decreased noise (Figure 5A). In contrast, S/N values for Systems 3 and 4 displayed biphasic dependency on Ω. For small Ω, Systems 3 and 4 showed the same

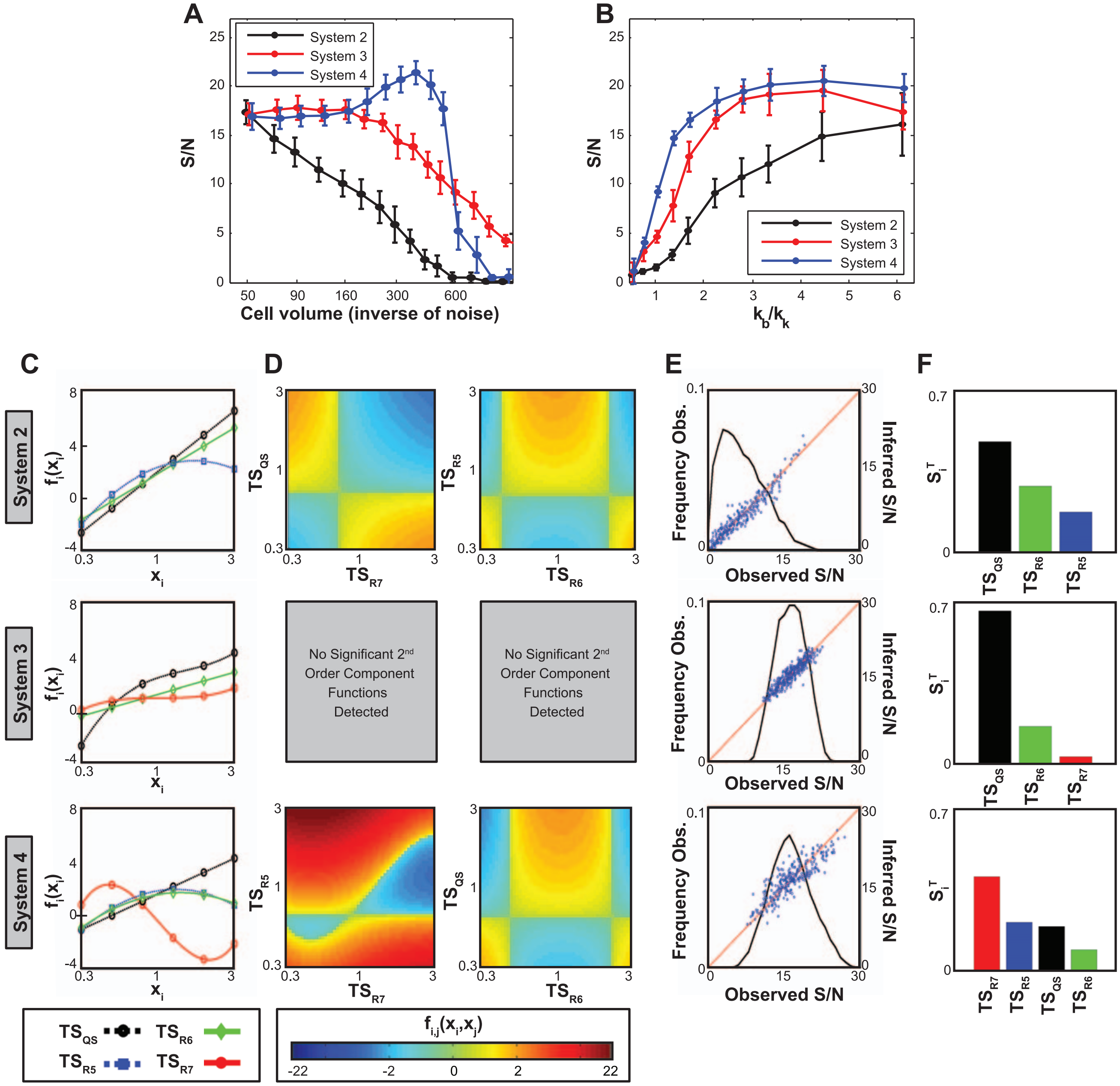


Figure 5. Robustness analyses and time-scale optimization for Systems 2–4. (A) S/N for different cell volume Ω, which corresponds to the number of molecules in each cell. (B) S/N for different ratios of stem cell division rate (k_b) and β-cell killing rate (k_k). (C–D) RS-HDMR analysis of Systems 2–4 to changes in the reaction time-scales of module components. (C) First- and (D) second-order RS-HDMR component functions describe the relationship between reaction time-scales (normalized to [0,1]) and the corresponding S/N observed in the overall system. (E) Distribution of S/N observed in response to time-scale parameter sampling (black) and RS-HDMR inference accuracy of that variation (blue). (F) Total sensitivity indices (S_i^T) of the module time-scales observed for each system.

doi:10.1371/journal.pcbi.1002579.g005

performance as System 2, largely because high noise obscured feedback signals. Intermediate values of Ω allowed the oscillator and throttle to generate optimally heterogeneous population responses. The S/N observed for large Ω was low for all systems due to more synchronous cellular commitment during the dynamic establishment of homeostasis, emphasizing the importance of stochasticity for generating heterogeneity in homeostasis regulation.

We analyzed the robustness of the systems to another external parameter, the average committed-cell survival time ($1/k_k$), which may fluctuate *in vivo*, by simulating system behavior with different ratios of uncommitted-cell division rate to committed-cell killing rate (k_b/k_k). In general, Systems 3 and 4 exhibited greater robustness to decreasing (k_b/k_k) compared to System 2 (Figure 5B and Text S1, Sec. 2.5). We also analyzed the effect of the parameter k_k on the homeostatic population size. Equilibrium populations remained near the desired homeostatic levels for high (k_b/k_k), but could decrease at lower ratios (Figure 5B). Ultimately, the robustness to noise and cell survival times underscores the need for heterogeneity within the population, and provides further evidence that the synthetic heterogeneity generated from the oscillator and throttle improves system performance over a range of parameters.

Intermodular time-scale matching reveals system dependent module coupling. In our system, accurate cell decision processing requires the appropriate integration of modules that generally have well defined behaviors in isolation. Even if we assume input-output behavior that meets our design specifications for each module (see Text S1, Sec. 3), integrating these modules together still presents a challenge. As introduced in Figures 3F and 4D,H, we explored system robustness to variations in the time-scales with which individual modules operate.

We used the Random-Sampling High Dimensional Model Representation (RS-HDMR) algorithm [11,37] (Text S1, Sec. 5.1) to understand both the individual and cooperative nonlinear effects of time-scale modulation on S/N (Figure 5C–F and S5). RS-HDMR describes the independent and cooperative effects of inputs, which in this analysis are module time-scales, on an output, the S/N value, in terms of a hierarchy of interpretable RS-HDMR component functions. Importantly, RS-HDMR supports global parametric sensitivity analysis, which is appropriate in this work where precise parameter values (time-scales in this case) may be highly uncertain. The first-order component function $f_i(x_i)$ describes the generally non-linear independent contribution of the i^{th} input variable to the output. For System 2, first-order RS-HDMR component functions showed that fast diffusion and a rapid toggle switch (through $R6$ dynamics) contribute to good system performance. Second-order RS-HDMR component functions indicated cooperative interactions among parameters. Here, parameters correspond to individual modules; therefore, we interpreted cooperative relationships as 'intermodular coupling' (Figure 5D). For example, having a fast toggle switch ($R7$) dynamics in System 2 offset the detrimental impact of slow diffusion. For System 3, the only significant correlations between performance and time scales were found for diffusion and, to a lesser extent, the toggle switch (Figure 5C). Interestingly, RS-HDMR detected no significant second-order component functions in System 3. These results indicated that the oscillator, in effect, decoupled the modules from each other, minimizing cooperative interactions between diffusion and the toggle switch by creating a buffer between the two. Compared to Systems 2 and 3, System 4 performance exhibited a more complex dependency on time-scale parameters, indicated by its significant second-order functions (Figure 5C–D). In particular, the cooperative interaction of slow $R7$ dynamics combined with fast $R5$ dynamics produced a strong synergistic improvement in S/N. This combined effect facilitates effective $AI3$-mediated lateral inhibition while the toggle switches. Total sensitivity indices represent the summed weight of first- and second-order RS-HDMR component functions for each parameter (Figure 5F). For Systems 2 and 3, observed S/N was most sensitive to changes in diffusion (TS_{QS}). In contrast, toggle-switch dynamics (TS_{R7}) most significantly affected performance in System 4. Of note, optimal time-scale matching yielded an improvement for all systems in robustness to molecular noise and cell survival dynamics, particularly under conditions of relatively fast cell death (Supplementary Figure S6). Overall, analysis of intermodular time-scale matching prescribes strategies for integrating modules and suggests ways in which module dynamics can be coordinately manipulated to yield improved system performance *in vitro*.

Clustered sensitivity analysis for targeted optimization. We also modeled System 3 using the Gillespie algorithm to explicitly account for binding and transcription events (for example, the binding of the receiver protein Rec1 to its inducer AI1, *Bind Rec1.AI1*, Text S1, Sec. 4). Results presented in the previous section were based on Langevin models that assume Hill functions for all inhibition and activation interactions, but our initial results with the Gillespie model suggested that achieving useful sigmoidal responses in the UPC module may be particularly challenging. Note that Systems 2 and 4 share the same UPC module as System 3 and the following results are valid for all systems. Figure 6A–B demonstrates how excess UPC output below the threshold (first row) or insufficient output above the threshold (second row) in suboptimal systems can lead to overactive commitment or proliferation, respectively. Consequently, we focused on optimizing the UPC module to obtain a step-like response to population density, ρ. We incorporated positive feedback in the UPC module, and then employed a genetic algorithm (GA) to optimize module parameters. The GA allowed us to efficiently navigate the high-dimensional parameter space and avoid local minima in the optimization process [38]. However, initial optimization of the module's rate constants only considered scenarios where the population densities increased ("forward response"). Unfortunately, this generated hysteresis, where high UPC output is maintained as ρ decreases below the threshold level (similar to [39]). Such hysteresis can lead to sub-optimal or even non-functional tissue homeostasis performance (Figure 6C–D, first row). Consequently, we also took into account the "reverse response" in the optimization process, which describes UPC output under conditions of decreasing cell density. Our GA optimization then successfully generated a diverse ensemble of rate constants, each yielding UPC networks with positive feedback that exhibited both step-like and non-hysteretic behavior (Figure 6C, second row). These optimized subnetworks produced stable homeostasis when integrated in the full-system Gillespie model (Figure 6D, second row).

We performed RS-HDMR analysis of the UPC subnetwork to understand how rate constants affect hysteresis, which would help guide the experimental construction of the system. We examined local parameter "neighborhoods" around each GA-generated vector of optimized parameters from Figure 6E (Text S1, Sec. 5.3). Our sensitivity analysis suggested that systems displaying similar UPC behavior can have drastically different responses to similar changes in rate constants: each parameter neighborhood that we analyzed had a distinct signature of parametric sensitivity (Figure 6F). We clustered parametric neighborhoods based on these signatures. Despite differences in individual sensitivities, the clustered sensitivity analysis revealed that the majority of

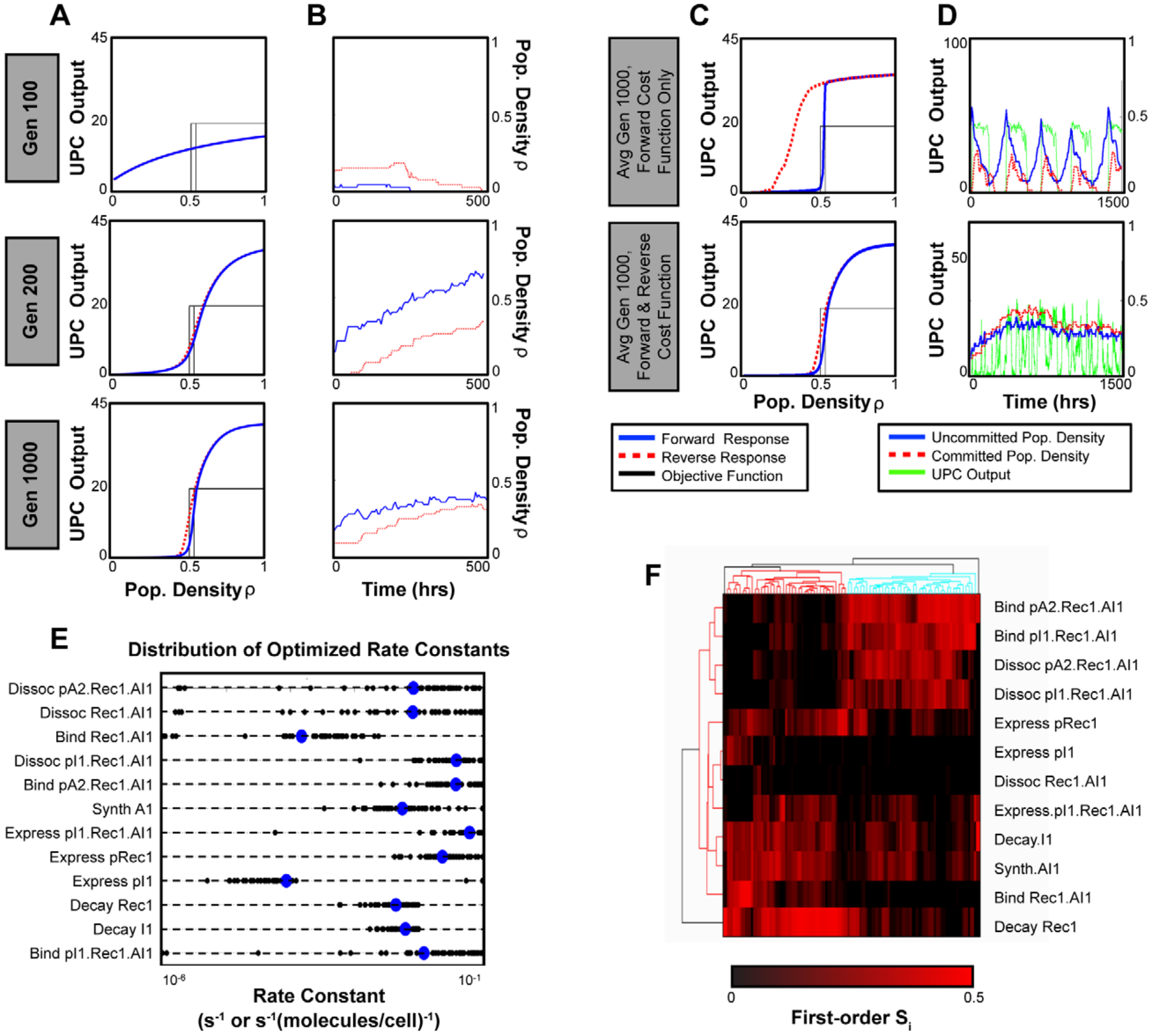


Figure 6. Parametric optimization of the UPC module. (A) GA optimization progress for three representative generations, using an ODE model of the UPC module. The GA objective function is a three-component step-function, with zero UPC activity below a defined threshold, an ignored transition region, and high activity above the transition region. (B) Gillespie simulations of System 3, corresponding to optimization progress in *A*. (C) Average UPC module transfer curves when the reverse response is either excluded or included in the subnetwork GA optimization. (D) Full system behavior corresponding by row to the module optimization results in *C*. (E) Distribution of rate constants for the optimized parameter vectors determined by 75 independent GA runs of 1000 generations each, using both forward and reverse response objective functions. (F) Clustered sensitivity analysis of the UPC Module. Each column corresponds to a "parameter sensitivity signature" for each of the 75 local parameter neighborhoods that we sampled; rows correspond to the analyzed parameters of the UPC module. First-order sensitivity values shown in the heat map range from 0.0 (black) to 0.5 (red).

doi:10.1371/journal.pcbi.1002579.g006

signatures fell into two main clusters, each with distinctive features. For example, in one cluster (red on the dendrogram) the decay rate of the receptor protein *Rec*1 (rate *Decay Rec1*) significantly affected hysteresis, while the binding and dissociation rates of *AI*1-bound *Rec*1 complex (*Rec*1.*AI*1) had little influence. The opposite was true for the other cluster (cyan on the dendrogram).

When building genetic networks experimentally, precise parameter values and their influence on system behavior may be unknown, presenting a challenge for optimization. Logistical constraints limit the number of parameters that can be reasonably manipulated, but clustered sensitivity analysis can act as a guide for iteratively prioritizing which parameters to mutate. In our system, for example, results suggest that we manipulate the most sensitive parameters from each of the two main clusters (*Decay Rec1* and the binding of the Rec1-AI1 complex to its promoter, *Bind pA2.Rec1.AI1*). At least one of these two parameter manipulations is likely to reduce hysteresis. Depending on which parameter is more sensitive, we may be able to deduce in which cluster the system lies, predict the sensitivity signature, and use this information for further optimization.

Parametric sensitivity analysis of synthetic heterogeneity modules. The impact of the oscillator and throttle modules on the performance of Systems 3 and 4 presents a particular challenge to understand and analyze (Figure 7A,G). As the two principle modules for generating synthetic heterogeneity, their ideal operating characteristics are complex and non-intuitive. Additionally, their non-trivial dynamics imply highly sensitive dependence on intramodular rate constants. As a first step to understand how to optimize these modules, we used RS-HDMR to investigate the sensitivity of S/N to random perturbations of the oscillator's and throttle's individual rate constants (Text S1, Sec. 5.2.4). As expected, results suggested highly complex and cooperative interactions among intramodular parameters, and no single parameter wholly determined system performance for either module (Figures 7, S7 and S8). Nonetheless, several parameters stood out as particularly important in governing performance.

For the oscillator, RS-HDMR indicated that the threshold at which *Ro* expression is activated by *Ao* (parameter H_{Ro}) had the largest impact on system behavior (Figure 7B–D). RS-HDMR also identified cooperative relationships among oscillator rates, the most significant being between H_{Ro} and H_{Ao-A} (threshold for *Ao* activation by itself), as shown in Figure 7E–F and S9. Although such correlations can classify 'good' performers (S/N >15) from 'bad' (S/N <2) with accuracy of roughly 95% (see Text S1, Sec. 6.2), analysis of the rate constants alone insufficiently described system behavior in a quantitative manner ($R^2 < 0.5$, Supplementary Figure S10).

For the throttle, results indicated that the thresholds for *At* repression by *R6* (H_{At}) and *R7* activation by *At* (H_{R7}) had the largest impact on system performance, and both interacted cooperatively to affect overall system performance such that low values of both parameters yielded the best S/N (Figure 7H–L and S11). As with System 3, our analysis of the rate constants alone failed to fully capture system performance in a quantitative manner ($R^2 < 0.5$, Supplementary Figure S10).

Phenotypic sensitivity analysis quantitatively informs system performance. Although a good first step, analysis of the module rate constants alone demonstrated two main drawbacks in this application. First, the statistical relationships between S/N and rate constants are highly convoluted and poorly captured by RS-HDMR. Second, focusing on rate constants can limit the analysis to a particularly defined network structure. To

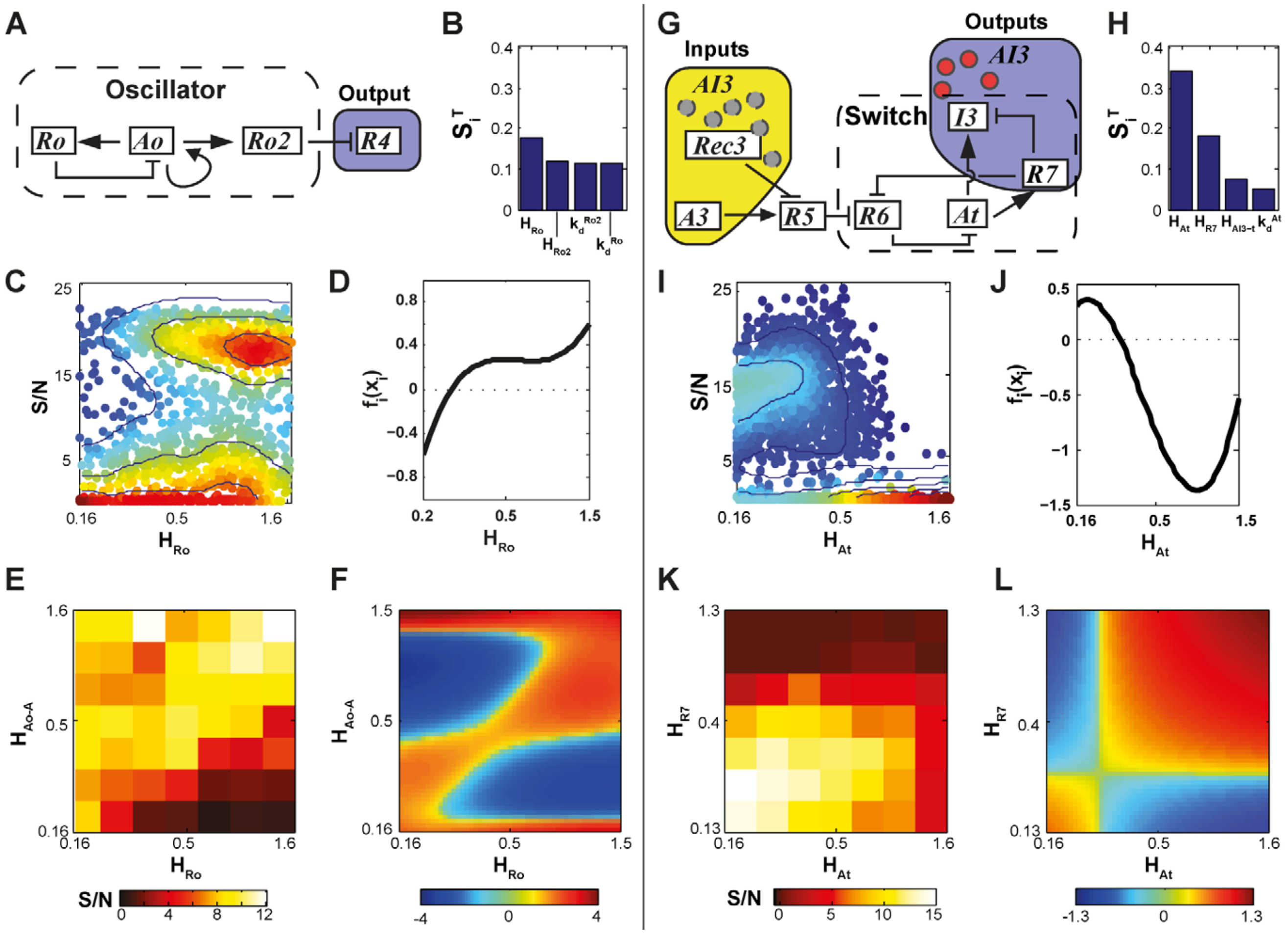


Figure 7. Parametric sensitivity analysis. (A,G) Circuit diagrams of the genetic components considered in (A) oscillator and (G) throttle optimization. (B,H) The most significant RS-HDMR sensitivity indices, S_i^T, for parametric variations of the oscillator and throttle, respectively. (C,I) Observed S/N values as a function of randomly sampled rate constant values. Around 2000 different parameter sets were tested, with all oscillator or throttle parameters simultaneously varied. Each point represents an individual parameter set. Warmer colors and contour lines indicate higher point density. (D,J) Inferred first-order RS-HDMR functions describing S/N as a function of the parameters sampled in *C* and *I*. (E,K) Heat map of the S/N values against the parameters resulting from the 2000 parameter sets tested in *C* and *I*. (F,L) RS-HDMR second-order functions describing the cooperative effects between rate constants, corresponding to *E* and *K*. Second-order RS-HDMR functions capture remaining variance after the first-order functions (see *D* and *J*) have been subtracted from the data.
doi:10.1371/journal.pcbi.1002579.g007

address these issues, we instead turned to analysis of high-level properties, or 'phenotypes', of the oscillator and throttle modules.

With the oscillator, examples of phenotypes include the average period of $R4$ oscillations and $R4$ dynamic range (Figure 8A and Supplementary Table S4). As with the rate constants, correlations between oscillator phenotype and system performance are multivariate by nature (Supplementary Figure S12): for example, the relationship between $\sigma_{Duration\ High}$ and S/N suggested a biphasic relationship, where optimal performance occurred with an intermediate level of variability (Figure 8B,C). We therefore used RS-HDMR to identify key phenotypic determinants of system performance. Interestingly, results indicated that metrics of oscillator heterogeneity (e.g., the coefficient of variation for the duration that the $R4$ concentration is low, CV^{P}_{Low}, and the standard deviation for the duration that the $R4$ concentration is high, $\sigma_{Duration\ High}$), are nearly as important as the concentrations within which the modules operate (i.e., the high and low oscillator values, Figure 8D). Ultimately, RS-HDMR results suggested that module phenotypes are far more predictive of S/N than the rate constants alone (Figure 8E and S10).

For the throttle, we defined phenotypes (Supplementary Table S5) of $R7$ and $AI3$ behavior as a function of the randomly-perturbed throttle rate constants. These phenotypes are more complex than those of the oscillator because the throttle module responds to two inputs, $R5$ and $AI3$ (Figure 8F). Consequently, we evaluated each throttle phenotype across combinations of both inputs, thereby producing an 'image' of throttle behavior over the two-input sampling space (Figure 8G and Supplementary Figure S13). For example, average images of *R7 T to St. St.* corresponding to different S/N values (Figure 8G) revealed that the best performing throttles show a clear pattern of activity: at low $A3$ or high $AI3$, no toggle switch occurs; at high $A3$ and low $AI3$, $R7$ stabilizes relatively quickly; lastly, inputs lying between these two regions cause toggle switching but with much slower (and heterogeneous) $R7$ dynamics. For a more systematic approach, we used feature extraction methods from image processing along

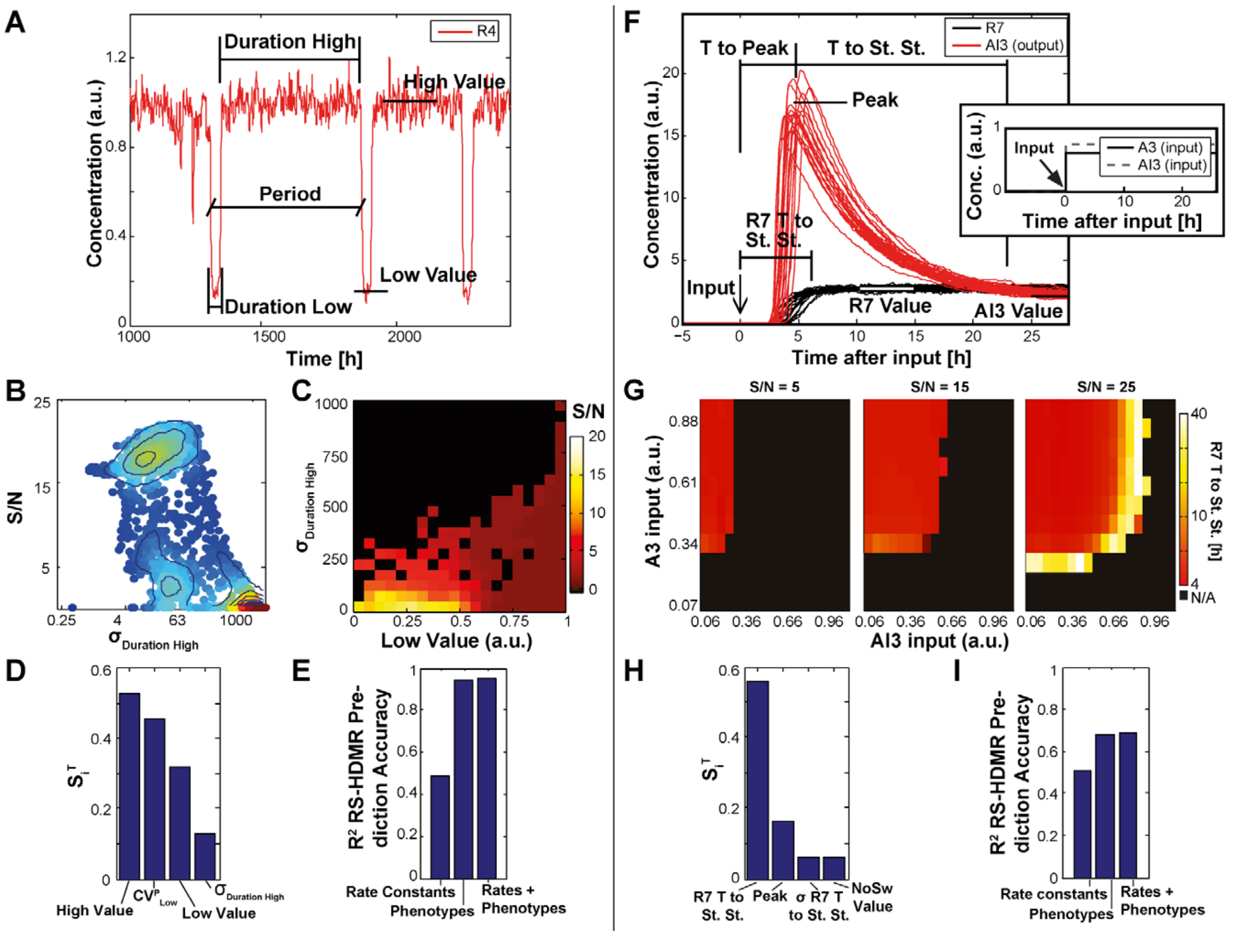


Figure 8. Phenotypic sensitivity analysis. (A,F) Phenotypic behavior of the oscillator (A) and throttle (F), when isolated from the full system. Roughly 2000 different sets of rate constants were tested, with all oscillator or throttle rate constants simultaneously varied. Module phenotypes were recorded for each set of rate constants. (B) Observed S/N values as a function of variance in the "duration high" of the oscillator. (C) Heat map of the S/N values against the phenotypes resulting from the random parameter sets. (G) Average 'images' for the phenotype *R7 T to St. St.*, observed from the random parameter sets yielding an S/N value of either 5, 15 or 25. Black represents regions where no switch occurs and no value for *R7 T to St. St.* is recorded. (D,H) The most significant RS-HDMR sensitivity indices, S_i^T, for phenotypic variations of the oscillator and throttle, respectively (see also Supplementary Table S8). (E,I) For the oscillator and throttle, respectively, RS-HDMR cross-validation predication accuracy using rate constants, phenotypes, or both.

doi:10.1371/journal.pcbi.1002579.g008

with RS-HDMR to identify key phenotypic determinants of system performance (Text S1, Sec. 5.2.4). As with the oscillator analysis, results indicated that module phenotypes predict overall system performance significantly better than rate constants alone, with the most significant phenotype being the time for $R7$ to reach steady state after receiving cues to commit (*R7 T to St. St.*, Figure 8H–I). Strikingly, RS-HDMR identified the variance with which $R7$ reaches steady state within this region to also be critically important for overall system performance (Supplementary Figure S8). Ultimately, phenotypic sensitivity analysis allowed for a more direct and accurate assessment of module performance compared to the analysis of rate constants alone, and did so while obviating concerns regarding the determination of rate constants that are tied to a particular system architecture (Figure 8I and S10).

Bayesian network analysis integrates rate constants and module phenotypes with overall system behavior. We applied Bayesian network inference to graphically represent the strong interdependencies of the module phenotypes and their relations with the rate constants that govern them and the S/N value (Figures 9, S15, S16 and Text S1, Sec. 5.2.4). Consistent with trends seen in Figures 8D and 8H, Bayesian network inference revealed that in general, module phenotypes more directly relate to overall system performance, and the effect of rate constants on overall S/N can be described in terms of their influence on the module phenotypic behavior. Nonetheless, in some cases the module phenotypes failed to adequately capture a rate constant's influence. For example, in the oscillator this led to a direct connection between the decay rate of the oscillator's repressor, k_d^{Ro}, and overall S/N. Remarkably Bayesian inference identified significant upstream effectors of S/N similar to those identified by RS-HDMR, while also suggesting a hierarchy of conditional dependencies (Figure 9). Multi-parent interactions identified by Bayesian networks supported RS-HDMR results; for example, the standard deviation of the time during which the oscillator is high ($\sigma_{Duration\ High}$) and the oscillator's *Low Value* showed significant cooperative interaction in both analyses (Figures 8C and 9A). Bayesian inference of the throttle relationships also agreed with RS-HDMR results, for example confirming relaxation kinetics (e.g., *R7 T to St. St.*) to be a significant influence on S/N, along with descriptors of its variability (Figures 8G–H and 9 B).

The integration of module phenotypes with the underlying rate constants ultimately allowed for efficient experimental optimization. Modules are likely to be experimentally implemented and phenotypically characterized in isolation before being integrated with each other. At this stage of optimization, Bayesian analysis can predict behavioral features of the individual module that will most directly influence performance in the fully integrated system, and such analysis may guide fine-tune adjustments of those module behaviors. In System 3, for example, Bayesian inference suggested that the oscillator's low value critically determined S/N, and that the threshold at which Ro expression is activated by Ao (parameter H_{Ro}) was the most direct parameter for modulating that phenotype. Although many module features and rate constants displayed covariation with overall S/N, Bayesian analysis distilled the most direct, causal influences on overall system behavior.

Discussion

System design and analogues from natural systems

In this work, we engineer mechanisms of robust control using synthetic generators of heterogeneity, and use a multi-faceted computational framework for design and optimization in the context of a relatively large-scale synthetic gene network. As a case study we chose tissue homeostasis control where individual cell decisions need to be coordinated to obtain desired multi-cellular behavior. To tackle this complex problem, we used top-down decomposition, achieving the overall task through the creation of interconnected modules, where each module has its own specific objective. Throughout this hierarchical optimization process we used different modeling approaches (population-based, Langevin and Gillespie simulations, see Figure 1B), while ensuring that the population-based results are consistent between the models (Supplementary Figure S17).

We designed System 1 by coupling four modules together, and simulated this system using a simplified ODE model. Computational analysis elucidated properties of global stability and demarcated regimes of steady vs. oscillatory homeostatic behavior in general tissue homeostasis systems. Analogous oscillatory homeostatic behavior from delayed feedback has been observed in natural mammalian systems, for example with hematopoiesis [40] and bacterial biofilms [41]. To mitigate the problem of population level oscillations, we created System 2 which includes a toggle switch module to implement faster feedback (Supplementary Table S9). Of note, various natural cell types regulate

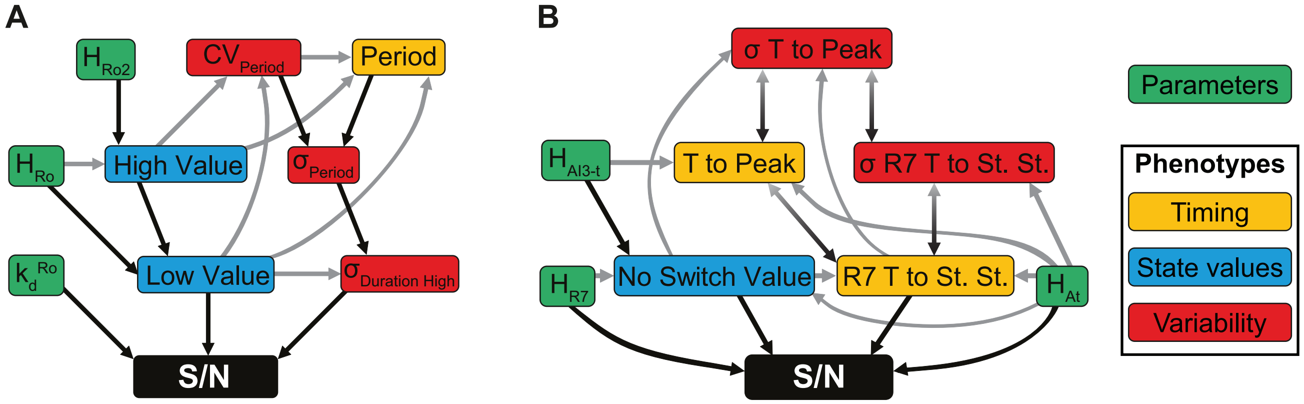


Figure 9. Bayesian networks of the impact of synthetic heterogeneity module phenotypes and rate constants on system performance value (S/N). (A) Bayesian network inference using oscillator rate constants and phenotypes. (B) Bayesian network inference using throttle rate constants and phenotypes. Black arrows indicate the most direct connections between a node and S/N. The Bayesian inference describes phenotype groupings relevant to state values (blue), timing (yellow), and variability (red), along with the rate constants that control these phenotypes (green).

doi:10.1371/journal.pcbi.1002579.g009

proliferation and differentiation by a switch similar in principle to that used in our system [42]. Analysis of System 2 using a stochastic Langevin model revealed how population-wide communication signals can be highly destabilizing to homeostasis, leading us to two new system designs. For Systems 3 and 4, the addition of the oscillator or the throttle module, respectively, provides more robust performance compared to System 2 (Figure 5) because these systems are less dependent on precise parameter values and are able to maintain sufficient population heterogeneity at lower levels of intrinsic molecular noise (Supplementary Table S9). Alternative mechanisms for generating population heterogeneity may exist. For example, the AND gate in System 2 could have been coupled with endogenously heterogeneous biological behavior such as Nanog expression (discussed below) [43]. Nonetheless, we chose to focus on the oscillator and throttle because they do not rely on potentially unpredictable endogenous mechanisms that would complicate computational modeling, and they represent two substantially distinct mechanisms for generating heterogeneity.

The design and analysis methods developed in this work attempt to identify relationships between rate constants, module phenotypes, and overall system performance, while maintaining an appreciation for the high degree of uncertainty and incomplete system knowledge in the experimental setting. For example, relating overall system performance directly to phenomenological definitions of module behavior frees the analysis from constraints to a particular module architecture or set of rate constants. Nonetheless, when more detailed information is desired we can apply global optimization strategies to capture patterns of parametric sensitivity that remain consistent across a broad range of rate constant values. For example, our analysis of the cell-cell communication module used a detailed biochemical reaction model with a large number of unknown rate constants. This level of granularity allowed us to analyze hysteretic response, which is not possible in the more abstract models. Ultimately, we addressed uncertainty by employing a novel technique, clustered sensitivity analysis, that revealed distinct patterns of relative parametric sensitivity for hysteresis that persisted across a wide range of rate constants. Previous reports have shown that bistability and hysteretic responses exist for both natural and engineered bacterial QS systems [44,45], and in this work such bistability drives undesired oscillations. Accordingly, we designed the population control module to avoid hysteretic response and identified specific properties affecting hysteresis in our system.

Synthetic and natural population heterogeneity

The synthetic heterogeneity modules in our systems display complex and multivariate behaviors that depend on the cooperative influence of multiple rate constants. Since existing experimental and computational biological circuit optimization methods do not scale well with system complexity, we decomposed the analysis and optimization processes for Systems 3 and 4 by characterizing modules first in isolation and then by relating their phenotypes to the performance of the overall system. We correlated module phenotypic behaviors with overall system performance, and found several significant correlations that were non-intuitive. Similarly, we identified dependencies between particular rate constants and the ability to maintain homeostasis. While Systems 3 and 4 exhibited comparable overall performances, further analyses revealed several distinguishing strengths and weaknesses (Supplementary Table S9). For example, the oscillator in System 3 appears to insulate modules from each other, while the throttle mechanism in System 4 amplifies their coupling strength (Figure 5 C–F). Our results suggest that the oscillator may mitigate problems associated with module integration, at least with respect to matching dynamics. However, the throttle mechanism is likely to be better suited for toggle switches with slow switching times (similar to the one we report on experimentally in Text S1, Sec. 1).

At a high level, our work describes strategies to exploit stochastic effects for enhancing stability of tissue homeostasis. This concept has been recently explored in a number of reports emphasizing the role probabilistic strategies play in natural mechanisms of cell-decision processing, including differentiation [29,46,47]. Furthermore, attempts have been made at engineering inherently stochastic processes for functions such as enhanced cellular reprogramming into induced pluripotent stem (iPS) cells [48]. Nonetheless, to our knowledge no efforts have yet been made that combine advances in synthetic biology with an appreciation of stochastic processes to engineer homeostatic tissue from isogenic cellular populations. The asynchronous oscillator stabilizes our system by generating population heterogeneity during conditions of environmental homogeneity and exogenous perturbation. Among natural systems, recent work has highlighted the role multistable feedback systems and stochastic switching play in appropriately priming cells for differentiation [49]. For example, evidence indicates the Nanog-Sox2-Oct4 network functions in part to generate population diversity by stochastically interrupting differentiation signals. Oscillators have been described as mediating cell-decisions in other biological systems, for example with p53 and NF-κB oscillations in response to DNA damage or other stimulation. These oscillations are hypothesized to enable discrete single-cell decisions to achieve a proportionate population-wide response [50]. Intrinsic noise generated by the oscillator also affects spatiotemporal clustering in our system (Supplementary Figures S18 B,E and S19) and natural analogues of this phenomenon exist. For example, non-genetic sources of cell-cell variability can cause recently divided cells to react more similarly to pharmacological treatment [51]. Similarly, lateral inhibition as proposed in the throttle mechanism of System 4 has also been observed in biological systems, for example in pattern formation [52], segmentation [53] or in the Notch signaling pathway [54]. Consistent with these studies, our spatial simulations show strong bias towards closely spaced alternate cell types in System 4 (Supplementary Figure S18 C).

Our optimization process, as well as the different biological examples described above, aim at seemingly contradictory objectives: information has to be processed faithfully from the population control modules to a commitment signal while, at the same time, stochasticity has to be amplified to generate heterogeneity. To achieve the first objective, several of our modules exhibit digital-like behavior, allowing us to effectively match components such that downstream modules react appropriately and with relative certainty to changes in upstream module output, attenuating the effects of noise. At the same time, to generate population heterogeneity, we exploit stochasticity by amplifying its effects in nonlinear modules operating in a transient regime. As a consequence, our modules are optimized to exhibit nonlinear responses to their inputs and, depending on the objective of the module, are tuned to work far from the transition regions for robust processing of information, or near the transition region where the response is highly sensitive to stochastic effects and hence efficiently generates heterogeneity.

Conclusions and future directions

We present here an integrated framework for forward-engineering large scale synthetic genetic circuits that combines several distinct computational approaches, and demonstrate its application to the design, analysis, and optimization of systems for controlling artificial tissue homeostasis. This framework represents a conceptual advancement for guiding experimental implementation by introducing hierarchical strategies that coordinate detailed biochemical models with modular phenotypes and optimization of module integration, all while considering parametric uncertainty and incomplete knowledge of the underlying biological context. With regard to methods development, future work may consider how to incorporate iterations of computational design with stepwise experimental implementation. Experiments could be designed to determine rate constants or high-level properties such as module phenotypes that most critically impact system performance, according to the computational modeling. Future work may also explore the limits of design automation. Network-level modeling could benefit from an integration with molecular modeling for directed optimization of molecular rate constants. Importantly, the modular design principles described in this work have been developed in part to facilitate redesign for improved performance or alternative applications. Artificial homeostasis systems have a range of potential applications in lower organisms, including co-culture systems for biosynthetic chemical production [55], controlled microbial homeostasis for environmental applications [56], and maintenance of microbial bio-sensors [57]. Medical applications may include a range of stem cell therapies currently being researched for treatment of degenerative diseases and traumatic injuries [58,59]. Forward-engineering efforts such as those presented here may elucidate roles of heterogeneity and homeostasis in diseases such as cancer, where tumor diversity potentially contributes to chemoresistance and metastasis [60]. Beyond guiding experimental implementation of the systems described herein, we believe the design principles and control motifs revealed by our analyses may offer more general insights into the role of population heterogeneity for robust behavior, with implications for both synthetic and systems biology.

Methods

Experimental implementations of the toggle switch and the cell-cell communication receiver were performed using immortalized human embryonic kidney cells (HEK293FT; Invitrogen), further discussed in the Text S1, Sec. 1. Computational methods and models utilized a variety of software platforms. We examined Systems 1–2 using ODE stability analyses and simulations (described in Text S1, Sec. 2), performed in Maple (Maplesoft; Waterloo, ON, Canada) and Matlab (MathWorks; Natick, MA). Systems 2–4 were analyzed using stochastic simulations. Langevin chemical simulations [35] (Text S1, Sec. 3) were performed using custom C++ code based on the 2-stage stochastic Runge-Kutta integration method with optimized parameters as described in [61]. All equations and parameters are reported in the Text S1, Sec. 3 and Table S1, respectively. In addition to Langevin simulations, Gillespie simulations (Figure 6, Text S1, Sec. 4) were implemented for Systems 2–3 using a standard rate-equation approach and the Gibson-modified Gillespie algorithm [62]. Transition rates were chosen to match the dynamics of the Langevin implementations (Table S3). For both the Langevin and Gillespie simulations, systems were described using a previously reported multicellular spatiotemporal simulation environment [15,63]. The simulation platform (written in C++) tracks the temporal evolution of intracellular reactions within individual cells that grow and die on a 2D grid. Furthermore, the platform monitors the spatiotemporal evolution of the cells themselves and extracellular signaling molecules that diffuse among them (Text S1, Sec. 2 and 4). We utilized a two-compartment ODE model of the UPC module for the GA optimizations (Text S1, Sec. 5.3 and Table S7), and implemented the GA in C++ using a distributed computing cluster (n = 40 processor nodes). RS-HDMR (Text S1, Sec. 5.1) was implemented as reported elsewhere [37,64]. A version of RS-HDMR [64] can be found online at http://www.aerodyne.com (free for academic users). Partial least squares regression and support vector machine classification (Text S1, Sec. 6.2) were implemented using standard Matlab functions, and Bayesian network inference (Text S1, Sec. 5.2.4) was performed in Matlab using previously described software [65].

Supporting Information

Figure S1 Experimental design and implementation for the signaling receiver and the toggle switch in mammalian cells. (A) 3OC6HSL mammalian receiver circuit design: Lux activator is co-expressed with a red fluorescent protein. Addition of 3OC6HSL induces EGFP expression. (B) Dose-response of 293FT cells infected with receiver circuit to 3OC6HSL, as measured by FACS. (C) Toggle switch design: Tet inhibits lac, which is expressed along GFP. Lac inhibits tet expression, which is coupled to mCherry. (D) Bistability of the toggle switch for both activation and deactivation. The shaded gray areas denote incubation with $10\mu M$ aTc. Yellow shading denotes incubation with 0.1 mM IPTG.
(PDF)

Figure S2 Simulations with feedback from all committed cells on the four-population system. At top, heatmap shows k_{c2} & k_k influence on β-cell oscillations for System 1, with $k_b = 1.5$, $k_{c1} = 5$, $k_d = 0.1$ and $n = 16$. Below the heatmap are trajectories with feedback from the β-cells (left column) and all committed cells (middle column), corresponding to parameter vectors 1–3 in the heatmap. The right column shows an equivalent two-population system with stem cells (blue line) and committed cells (red lines). The approximate β-cell population was extrapolated according to Eq. S5 (see Text S1).
(PDF)

Figure S3 Nullclines of the reduced model. (A) Nontrivial component of nullcline X in the reduced two-population model. (B) Nullcline Y in the reduced two-population model. (C) Complete phase-plane in the reduced two-population model. (D) Nullclines for an example with three nonzero steady states in the reduced two-population model. (E) Nullcline Y for large Hill exponents in the reduced two-population model.
(PDF)

Figure S4 Gillespie implementation of System 3. Gillespie implementation of System 2 is identical, but without the oscillator module. Although similar, design details in the population control modules differ slightly from the Langevin implementation. Arrowed and barred connections represent transcriptional activation and repression, respectively. The dashed connection in the differentiation module represents indirect transcriptional activation.
(PDF)

Figure S5 Parametric sampling distribution for modular time-scale analysis. Time scale parameters were randomly and uniformly varied across one order of magnitude for the time-scale of each module or component to produce roughly 360 parameter sets for each System (2, 3, and 4). Simulations of each parameter set yielded a corresponding S/N value, which is plotted here as a function of the individual time-scale parameters. Each point represents an individual parameter set. Warmer colors indicate higher point density; contour lines also indicate point density. TS_{QS} describes the time-scale of the quorum signaling

molecules (including diffusion), TS_{QM} denotes the time-scale of the quorum sensing module ($A1$, $R2$, …), and other time-scales are specific to the components $R5$, $R6$, $R7$ and At.
(PDF)

Figure S6 Population level properties of time-scale optimized Systems 2, 3 and 4. (A) Signal to noise value (S/N) for different cell volume Ω. (B) Signal to noise value (S/N) for different ratio of stem cell division rate (k_b) and β-cell killing rate (k_k). With the time-scale optimization, all systems show an increase by ~ 5 units of their S/N value.
(PDF)

Figure S7 Oscillator rate constants (see Table S1) were randomly varied across one order of magnitude around initial values (uniform distribution in the log space) to produce roughly 2000 parameter sets. Simulations of each parameter set yielded a corresponding S/N value, which is plotted here as a function of the individual parameters. Each point represents an individual parameter set. Warmer colors and contour lines indicate higher point density.
(PDF)

Figure S8 Throttle rate constants (see Table S1) were randomly varied across one order of magnitude around initial values (uniform distribution in the log space) to produce roughly 6000 parameter sets. Simulations of each parameter set yielded a corresponding S/N value, which is plotted here as a function of the individual parameters. Each point represents an individual parameter set. Warmer colors indicate higher point density; contour lines also indicate point density.
(PDF)

Figure S9 RS-HDMR global parametric sensitivity analysis of oscillator module rate constants (see Figure S7), describing the influence of parameter variation on observed S/N. (A) RS-HDMR first-order component functions, in order of decreasing global sensitivity index S_i. (B) Second-order RS-HDMR component functions in order of decreasing global sensitivity index S_{ij}.
(PDF)

Figure S10 Inference of the S/N values for Systems 3 and 4. (A) RS-HDMR inference of System 3 S/N value using oscillator rate constants (A) or oscillator phenotypes (B), and RS-HDMR inference of System 4 S/N value using either throttle rate constants (C) or throttle phenotypes (D). The red curve indicates the distribution of S/N observed in response to parameter variation in either the oscillator or throttle. Black dots indicate observed vs. inferred S/N value for individual sets of oscillator or throttle parameter vectors. Inference accuracy corresponds to R^2 values reported in Figure 8E and 8I.
(PDF)

Figure S11 RS-HDMR parametric sensitivity analysis of the throttle module rate constants (see Figure S8), describing the influence of parameter variation on observed S/N. (A) RS-HDMR first-order component functions, in order of decreasing sensitivity index S_i. (B) Second-order RS-HDMR component functions in order of decreasing sensitivity index S_{ij}.
(PDF)

Figure S12 S/N values plotted against the different oscillator phenotypes (as described in Table S4) corresponding to the parameter sets of Figure S7. Multiple simulations of each parameter set yielded a phenotype in the isolated System (see Figure 8A) corresponding to the S/N value evaluated with the whole System 3. Each point represents an individual parameter set. Warmer colors indicate higher point density; contour lines also indicate point density.
(PDF)

Figure S13 Standard deviation of the time for $R7$ to reach steady state in the throttle module. (A) The standard deviation of the time for $R7$ to reach its steady state is measured for given levels of $A3$ and external $AI3$; the colorbar denotes the standard deviation for 100 independent simulations. (B) Time trajectories for different combinations of $A3$ and $AI3$: (1) the intermediate case exhibits high variability with switching behavior; (2) high $A3$ and low $AI3$ results in rapid and simultaneous toggle switching; (3) high $AI3$ and $A3$ results in no toggle switching (notice the different scale on the y-axis). Input $A3$ and $AI3$ doses are introduced into the system at time $t=0h$ as marked by the arrow.
(PDF)

Figure S14 Average heat map for different values of S/N for the throttle phenotypes (as described in Table S5). These maps are obtained as the average of the maps resulting from simulations of parameter sets having similar S/N values.
(PDF)

Figure S15 Scores for the edges of the Bayesian network of the oscillator module including module parameters and phenotypes (see Text S1, Sec. 5.2.4). Only the most significant phenotypes are taken as nodes of the network. For the Figure 9 A, only edges with scores above 0.8 are shown.
(PDF)

Figure S16 Scores for the edges of the Bayesian network of the throttle module including module parameters and phenotypes (see Text S1, Sec. 5.2.4). Only the most significant phenotypes are taken as nodes of the network. For the Figure 9 B, only edges with scores above 0.3 are shown.
(PDF)

Figure S17 Population density for different ratios of division and killing rate. Deterministic simulation with a two-population model (A,C) and stochastic simulations of the Systems 2, 3 and 4 (B–D) show qualitatively similar results. (A–B) The population of uncommitted cells remains constant with a small decrease for low rate ratio. (C–D) The population of committed cells follows a power law with an exponent near 1 for low ratio and close to $1/16=0.625$ for large ratio. Power laws in (D) are fitted on the results of System 2, the closest to the ODE model.
(PDF)

Figure S18 Spatial patterning and impact of molecular noise on the patterning. For a given uncommitted (blue) or committed (red) reference cell, the Z-score (see Text S1, Sec. 5.4) indicates the distribution bias of committed neighbors at a given distance (dashed lines, $p<0.01$). We performed simulations using the Langevin models with $\Omega=200$ (A–C) or $\Omega=500$ (D–F). For Systems 2 and 4, committed cells are not likely to have committed neighbors (A,C), whereas System 3 has no significant bias for short distances. With lower noise (D,E), committed cells in Systems 2 and 3 tend to cluster, such that committed cells bias to have committed cell neighbors. (F) System 4 demonstrates enhanced lateral inhibition, and committed cells bias to not have committed cell neighbors.
(PDF)

Figure S19 Spatiotemporal analysis of System 3 using the Gillespie model. We define activity for the "Population Control" (PC) module as the level of Rec2.AI2 complex-bound promoter for the R5 repressor (*pR5.Rec2.AI2*). (A) The thick lines represent the PC activity for uncommitted cells as a function of

distance from uncommitted (blue) and committed (red) neighboring cells, averaged over all cells and all time points for a given simulation. Thin lines represent PC activity +/− the standard error of the mean at each distance. (B) Average PC activity for all uncommitted cells over all time points for a given simulation are shown as a function of the number of committed neighbors at one (ordinate) and two (abscissa) grid units away. (C) We measured the time difference between nearest oscillation peaks of dimerized R1 (R1D) for all pairs of coexistent uncommitted cells throughout a given simulation. For example, if four uncommitted cells are alive at a given time point, we would calculate the phase difference among all of the pairs of cells (six in this case). Average phase difference increases as the distance between neighboring cells increases (blue line). The lower and upper black dashed lines represent the first and third quartiles of the phase difference, respectively. Phase difference increases as a function of distance because cells closer together are more likely to have originated from the same parent cell. (D) For a given uncommitted or committed reference cell, the Z-score (see Text S1, Sec. 5.4) indicates the distribution bias of committed and uncommitted neighbors at a given distance (dashed lines, $p < 0.01$). Patterning was examined for Systems 2 and 3.
(PDF)

Table S1 Parameters for the Langevin models of Systems 2 to 4.
(PDF)

Table S2 Scaled parameters for the time-scale analysis. The kinetics parameters (k_p^α and k_d^α) from Table S1 are scaled by the time-scale parameters TS_α according to their module. For each combination of time-scale parameters, the $\overline{k_p^\alpha}$ and $\overline{k_d^\alpha}$ parameters are used for the Langevin simulations.
(PDF)

Table S3 List of reactions for the full multicellular model of System 3. Depending on whether the reactions are associative or dissociative, reaction rates are in units of (molecules per cell)$^{-1}$(s)$^{-1}$ or s^{-1}.
(PDF)

Table S4 Phenotypes for the oscillator module (see Figure 8A).
(PDF)

Table S5 Phenotypes for the throttle module (see Figure 8F).
(PDF)

Table S6 Features used to analyze throttle behavior. These features were measured for each throttle phenotype (see Table S5), where "image" refers to the observed phenotype as a response to the two inputs, $A3$ and $AI3$ (see Figure S14).
(PDF)

Table S7 Rate constants for two-compartment model of the UPC module.
(PDF)

Table S8 Top RS-HDMR identified throttle features and their corresponding RS-HDMR sensitivity indices, S_i^T (see Figure 8 H).
(PDF)

Table S9 Summary of the advantages and disadvantages of Systems 1–4.
(PDF)

Text S1 Supporting text. Subsections include the following: (1) experimental proof of concept, (2) methods for the ODE modeling of Systems 1–2 and related analytical proofs, (3) methods for the Langevin modeling of Systems 1–2, (4) methods for the Gillespie modeling of Systems 2–3, (5) methods for results analyses of Systems 2–4, (6) additional results.
(PDF)

Acknowledgments

We would like to thank Joel Wagner and other members of the Weiss and Lauffenburger labs, along with Dr. Xiao-Jeng Feng and Prof. Herschel Rabitz (Princeton University) for helpful discussions.

Author Contributions

Conceived and designed the experiments: MM MH PEMP RW. Performed the experiments: MM MH ND SS. Analyzed the data: MM MH ES RW. Contributed reagents/materials/analysis tools: MM MH ES DL. Wrote the paper: MM MH ES RW.

References

1. Elowitz M, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335–8.
2. Swinburne I, Miguez D, Landgraf D, Silver P (2008) Intron length increases oscillatory periods of gene expression in animal cells. Genes Dev 22: 2342–2346.
3. Tigges M, Marquez-Lago T, Stelling J, Fussenegger M (2009) A tunable synthetic mammalian oscillator. Nature 457: 309–312.
4. Stricker J, Cookson S, Bennett M, Mather W, Tsimring L, et al. (2008) A fast, robust and tunable synthetic gene oscillator. Nature 456: 516–519.
5. Gardner T, Cantor C, Collins J (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339–42.
6. Kramer BP, Usseglio Viretta A, Daoud-El Baba M, Aubel D, Weber W, et al. (2004) An engineered epigenetic transgene switch in mammalian cells. Nat Biotechnol 22: 867–870.
7. Weiss R, Knight T (2000) Engineered communications for microbial robotics. In: DNA6: Sixth International Workshop on DNA-Based Computers, DNA2000. Berlin: Springer. pp. 1–16.
8. Weber W, Schuetz M, Dnervaud N, FusseneggerM(2009) A synthetic metabolite-based mammalian inter-cell signaling system. Mol BioSyst 5: 757–763.
9. Chen M, Weiss R (2005) Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. Nat Biotechnol 23: 1551–1555.
10. Ellis T, Wang X, Collins J (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. Nat Biotechnol 27: 465–471.
11. Feng XJ, Hooshangi S, Chen D, Li G, Weiss R, et al. (2004) Optimizing genetic circuits by global sensitivity analysis. Biophys J 87: 2195–2202.
12. You L, Cox III R, Weiss R (2004) Programmed population control by cell- cell communication and regulated killing. Nature 428: 868–871.
13. Weber W, Daoud-El Baba M, Fussenegger M (2007) Synthetic ecosystems based on airborne interand intrakingdom communication. Proc Natl Acad Sci USA 104: 10435–10440.
14. Tabor JJ, Salis HM, Booth Simpson Z, Chevalier AA, Levskaya A, et al. (2009) A synthetic genetic edge detection program. Cell 137: 1272–1281.
15. Basu S, Mehreja R, Thiberge S, Chen M, Weiss R (2004) Spatiotemporal control of gene expression with pulse-generating networks. Proc Natl Acad Sci USA 101: 6355–6360.
16. Basu S, Gerchman Y, Collins C, Arnold F, Weiss R (2005) A synthetic multicellular system for programmed pattern formation. Nature 434: 1130–1134.
17. Kobayashi H, Kaern M, Araki M, Chung K, Gardner T, et al. (2004) Programmable cells: Interfacing natural and engineered gene networks. Proc Natl Acad Sci USA 101: 8414–8419.
18. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y (2011) Multi-input rnai-based logic circuit for identification of specific cancer cells. Science 333: 1307–1311.
19. Purnick P, Weiss R (2009) The second wave of synthetic biology: from modules to systems. Nat Rev Mol Cell Biol 10: 410–422.
20. Dassau E, Zisser H, Palerm C, Buckingham B, Jovanović L, et al. (2008) Modular artificial β-cell system: a prototype for clinical research. J Diabetes Sci Technol 2: 863.
21. Fujikawa T, Oh S, Pi L, Hatch H, Shupe T, et al. (2005) Teratoma Formation Leads to Failure of Treatment for Type I Diabetes Using Embryonic Stem Cell-Derived Insulin-Producing Cells. Am J Pathol 166: 1781–1791.
22. Ryan EA, Paty BW, A SP, Bigam D, Eman A, et al. (2005) Five-year follow-up after clinical islet transplantation. Diabetes 54: 2060–2069.
23. Robertson R (2010) Islet transplantation a decade later and strategies for filling a half-full glass. Diabetes 59: 1285.

24. Shapiro A, Ricordi C, Hering B, Auchincloss H, Lindblad R, et al. (2006) International trial of the edmonton protocol for islet transplantation. N Engl J Med 355: 1318–1330.
25. Rangarajan A, Talora C, Okuyama R, Nicolas M, Mammucari C, et al. (2001) Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation. EMBO J 20: 3427–3436.
26. Cattaneo E, McKay R (1990) Proliferation and differentiation of neuronal stem cells regulated by nerve growth factor. Nature 347: 762–765.
27. Coleman M, DeMayo F, Yin K, Lee H, Geske R, et al. (2005) Myogenic Vector Expression of Insulin-like Growth Factor I Stimulates Muscle Cell Differentiation and Myofiber Hypertrophy in Transgenic Mice. J Biol Chem 146: 5455–5462.
28. Song S, Gannon M, Washington M, Scoggins C, Meszoely I, et al. (1999) Expansion of Pdx1-expressing pancreatic epithelium and islet neogenesis in transgenic mice overexpressing transforming growth factor α. Gastroenterology 117: 1416–1426.
29. Chang H, Hemberg M, Barahona M, Ingber D, Huang S (2008) Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature 453: 544–547.
30. Lander A, Gokoffski K, Wan F, Nie Q, Calof A (2009) Cell lineages and the logic of proliferative control. PLoS Biol 7: e1000015.
31. Leon-Quinto T, Jones J, Skoudy A, Burcin M, Soria B (2004) In vitro directed differentiation of mouse embryonic stem cells into insulin-producing cells. Diabetologia 47: 1442–1451.
32. Zhang D, Jiang W, Liu M, Sui X, Yin X, et al. (2009) Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. Cell Res 19: 429–438.
33. Weiss R, Homsy GE, Knight Jr TF (1999). Toward in vivo digital circuits. DIMACS Workshop on Evolution as Computation.
34. Gradner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339–342.
35. Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58: 35–55.
36. Shvartsman S, Wiley H, Deen W, Lauffenburger D (2001) Spatial range of autocrine signaling: modeling and computational analysis. Biophys J 81: 1854–1867.
37. Miller M, Feng X, Li G, Rabitz H (2012) Identifying biological network structure, predicting network behavior, and classifying network state with high dimensional model representation (hdmr). PLoS ONE 7: e37664.
38. Leardi R (2001) Genetic algorithms in chemometrics and chemistry: a review. J Chemom 15: 559–569.
39. Angeli D, Ferrell JE, Sontag ED (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. Proc Natl Acad Sci USA 101: 1822–1827.
40. Saker S (2003) Oscillation and global attractivity in hematopoiesis model with delay time. Appl Math Comput 136: 241–250.
41. Yarwood J, Bartels D, Volper E, Greenberg E (2004) Quorum Sensing in Staphylococcus aureus Biofilms. J Bacteriol 186: 1838–1850.
42. Richardson R, Dixon J, Malhotra S, Hardman M, Knowles L, et al. (2006) Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch. Nat Genet 38: 1329–1334.
43. Kalmar T, Lim C, Hayward P, Muñoz-Descalzo S, Nichols J, et al. (2009) Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol 7: e1000149.
44. Williams J, Cui X, Levchenko A, Stevens A (2008) Robust and sensitive control of a quorum-sensing circuit by two interlocked feedback loops. Mol Sys Biol 4 : 234
45. Haseltine E, Arnold F (2008) Implications of rewiring bacterial quorum sensing. Appl Environ Microbiol 74: 437.
46. Acar M, Mettetal J, van Oudenaarden A (2008) Stochastic switching as a survival strategy in fluctuating environments. Nat Genet 40: 471–475.
47. Suel G, Garcia-Ojalvo J, Liberman L, Elowitz M (2006) An excitable gene regulatory circuit induces transient cellular differentiation. Nature 440: 545–550.
48. Hanna J, Saha K, Pando B, Van Zon J, Lengner C, et al. (2009) Direct cell reprogramming is a stochastic process amenable to acceleration. Nature 462: 595–601.
49. Kalmar T, Lim C, Hayward P, Muñoz-Descalzo S, Nichols J, et al. (2009) Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol 7: e1000149.
50. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine A, et al. (2004) Dynamics of the p 53-Mdm 2 feedback loop in individual cells. Nat Genet 36: 147–150.
51. Spencer S, Gaudet S, Albeck J, Burke J, Sorger P (2009) Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature 459: 428–432.
52. Meinhardt H, Gierer A (2000) Pattern formation by local self-activation and lateral inhibition. BioEssays 22: 753–760.
53. von Dassow G, Meir E, Munro EM, Odell GM (2000) The segment polarity network is a robust developmental module. Nature 406: 188–192.
54. Axelrod JD (2010) Delivering the lateral inhibition punchline: It's all about the timing. Sci Signal 3: pe38.
55. Ding J, Liu B, Ren N, Xing D, Guo W, et al. (2009) Hydrogen production from glucose by co-culture of clostridium butyricum and immobilized rhodopseudomonas faecalis RLD-53. Int J Hydrogen Energy 34: 3647–3652.
56. Kumar M, Leon V, De Sisto Materano A, Ilzins O (2006) Enhancement of oil degradation by co-culture of hydrocarbon degrading and biosurfactant producing bacteria. Pol J Microbiol 55: 139–146.
57. Gu X, Trybi lo M, Ramsay S, Jensen M, Fulton R, et al. (2010) Engineering a novel self-powering electrochemical biosensor. Syst Synth Biol 4: 203–14.
58. Segers V, Lee R (2008) Stem-cell therapy for cardiac disease. Nature 451: 937–942.
59. Lindvall O, Kokaia Z (2009) Prospects of stem cell therapy for replacing dopamine neurons in parkinson's disease. Trends Pharmacol Sci 30: 260–267.
60. Clevers H (2011) The cancer stem cell: premises, promises and challenges. Nature Med 17: 313–319.
61. Burrage K, Burrage PM (1996) High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. Appl Numer Math 22: 81–101.
62. Gibson M, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem A 104: 1876–1889.
63. Hsu A, Vijayan V, Fomundam L, Gerchman Y, Basu S, et al. (2005) Dynamic Control in a Coordinated Multi-Cellular Maze Solving System (I). In: American Control Conference. volume 6, pp. 4399–4404.
64. Li G, Rabitz H, Yelvington P, Oluwole O, Bacon F, et al. (2010) Global sensitivity analysis for systems with independent and/or correlated inputs. J Phys Chem A 114: 6022–6032.
65. Eaton D, Murphy K (2007) Exact Bayesian structure learning from uncertain interventions. In: AI & Statistics. Citeseer, volume 2, pp. 107–114.

Supporting text for
“Modular design of artificial tissue homeostasis: robust control through cellular heterogeneity”

Miles Miller[1†], Marc Hafner[1,2,3,4†,♯], Eduardo Sontag[5], Noah Davidsohn[1], Sairam Subramanian[6],
Priscilla E. M. Purnick[7], Douglas Lauffenburger[1], and Ron Weiss[1,*]
1 Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
2 School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
3 Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland
4 Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland
5 Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA
6 Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA
7 Department of Gene and Cell Medicine, Black Family Stem Cell Institute, Mount Sinai School of Medicine, New York, NY 10029, USA
† equal contribution
∗ E-mail: rweiss@mit.edu

Contents

1 Experimental proof of concept 2
- 1.1 Cell-cell communication . . . 2
- 1.2 Toggle switch . . . 2
- 1.3 Experimental methods . . . 2

2 Methods for the ODE modeling of Systems 1-2 and related analytical results 4
- 2.1 System 0: differentiation only . . . 4
- 2.2 Convergence and stability in Systems 1 and 2 . . . 4
- 2.3 Reduced two-population model . . . 5
- 2.4 Stability and convergence in the two-population model . . . 6
- 2.5 Robustness to the rate k_k . . . 11
- 2.6 Simulations . . . 12

3 Methods for the Langevin modeling of Systems 2-4 14
- 3.1 Quorum sensing module . . . 14
- 3.2 AND gate and toggle switch . . . 16
- 3.3 Cell fate . . . 16
- 3.4 System 3 – implementation of an oscillator . . . 17
- 3.5 System 4 – implementation of a throttle . . . 18
- 3.6 Spatial simulations . . . 18

4 Methods for the Gillespie modeling of Systems 2-3 20
- 4.1 Gene network design . . . 20

5 Methods for results analyses of Systems 2-4 23
- 5.1 RS-HDMR sensitivity analysis . . . 23
- 5.2 Langevin model analysis . . . 24
- 5.3 Analysis of a two-compartment ODE model of the UPC module . . . 25
- 5.4 Patterning and Neighbor Density Analysis . . . 28

6 Additional results 29
- 6.1 Detailed results for time-scale analysis . . . 29
- 6.2 Detailed results for the oscillator and throttle module analyses . . . 29
- 6.3 Population size when varying β-cell killing rate . . . 29

1 Experimental proof of concept

In recent years, synthetic biology efforts have produced a sizeable number of functional and characterized elements, ranging from repressors and activators to modules such as the toggle switch, oscillator, and cell communication systems. The Registry of Standard Biological Parts (http://partsregistry.org/) represents a collection of such elements [1, 2]. As a proof of concept, here we present experimental results for two critical components of the systems we aim to build. First, for cell-cell communications we engineered a mammalian receiver based on the LuxR protein that responds to 3-oxo-hexanoyl-homoserine lactone (3OC6HSL). Second, we employed two transcription factors, LacI and TetR, to create the toggle switch used in Systems 2-4 (Figures 3 and 4). The Weiss lab is currently developing two other modules needed for these systems: a mammalian 3OC6HSL sender based on LuxI, and a module to direct stem cell differentiation into insulin-producing pancreatic β-like cells. The differentiation module functions through stepwise expression of cell-fate regulators. *Gata4* expression in stem cells stimulates differentiation into endodermal cells, which activates an alpha-fetoprotein (AFP) promoter [3]. Preliminary results suggest that *ngn3* and *pdx1*, when fused to the AFP promoter, stimulate further differentiation into insulin-producing cells (data not shown).

1.1 Cell-cell communication

The mammalian receiver we built consists of a mammalian-optimized LuxR based signal transducer that binds 3OC6HSL and activates transcription from a novel mammalian optimized *lux* promoter (Figure S1A). We designed the signal transducer by fusing a p65 activation domain from the mammalian ReLa protein [4] to a helical linker H4 [5] and the N-terminus of a mammalian codon-optimized LuxRF, a hypersensitive LuxR mutant [6]. We also appended a nuclear localization signal (NLS) to the C-terminus of this protein. To test the redesigned receiver circuit (Figure S1A), HEK293FT cells were co-infected with a lentivirus constitutively expressing p65-H4-LuxRF/DsRed2 and a lentivirus with EGFP under PluxO7 control, demonstrating a highly functional mammalian 3OC6HSL receiver with an half maximal effective concentration (EC_{50}) of roughly 10μM (Figure S1B).

1.2 Toggle switch

Our toggle switch consists of two transcription factors, LacI and TetR (Figure S1C), that cross-repress each other. We fused a Krupple associated box (KRAB) domain to each of the LacI and TetR proteins to ensure efficient repression. The network design was based on earlier computational models [7] and an *E. coli* implementation [8]. Our experimental results indicate that the toggle switch state can be flipped with transient administration of IPTG and aTc. The system is able to maintain long term stability (>3 days), and the time required for the switch to reach 50% fluorescence is roughly 34 hours with the addition of aTc and 55 hours upon IPTG induction (Figure S1D).

1.3 Experimental methods

E. coli XL10-Gold cells were used to clone and propagate plasmids (Agilent; Santa Clara, CA). Cells were grown in LB broth (Difco, Detroit, MI) with 100μg/mL ampicillin (Sigma, St. Louis, MO) and 50 μg/mL kanamycin (Shelton Scientific, Shelton, CT) when appropriate. AHL 3-oxohexanoyl-homoserine lactone (3OC6HSL) was acquired from Sigma-Aldrich.

1.3.1 Mammalian cell culture

293-TetON (Clontech) cells were used to assay plasmids and viruses dependent on rtTA expression. NIH3T3 cells (ATCC) were used to assay viruses and determine their titer. Polybrene (Sigma) was used at a concentration of 10 μg/ml for infecting cells. All cells were grown at 37 °C and 5% CO2 in a sterile tissue culture incubator. Media for culturing 293FT/NIH3T3 cells was composed of DMEM (Hyclone), 10% Tet-approved Fetal Bovine

Serum (Clontech), 1% Penicillin-Streptomycin (Hyclone), and 0.1% Fungin (Invivogen) filtered through a 0.45 μ filter (Nalgene).

Lentivirus production and infection protocols were adapted from [9] using HEK293FT cells, packaging plasmids [10], and Superfect transfection reagent (Qiagen). Collected virus was concentrated either by ultrafiltration using Centricon Plus-70 100 kDa spin filters (Millipore) or by ultracentrifugation at 50000 g for 2.5 hours.

2 Methods for the ODE modeling of Systems 1-2 and related analytical results

In this section we discuss results related to analysis of the homeostasis system using ordinary differential equations (ODEs). We first describe the tissue homeostasis system using a four-population ODE model, but without any feedback control (termed "System 0"). We then show simulation results for the Systems 0, 1 (includes feedback) and 2 (includes toggle switch) with the four-population model and a corresponding two-population reduced model, which is equivalent to System 2. In the reduced system, we prove that equilibrium points exist under certain circumstances. We also show that the committed cell population remains robust to variations of the killing rate, k_k. The proofs are written in a general way and are valid for a broader scope than the present application.

2.1 System 0: differentiation only

The simplest tissue homeostasis system involves a mechanism that causes cells to differentiate, which we describe as the differentiation module. We model this system in terms of four cell types. Stem cells (population size is S) grow at a constant rate k_b and mature a constant rate k_{c1} into endodermic cells (E). Endodermic cells mature into pancreatic cells (P) at a rate k_{c2}. Finally, pancreatic cells differentiate at a rate k_d into β-cells (B), which then die at a constant rate k_k.

$$\begin{aligned}
\frac{dS}{dt} &= k_b \cdot S - k_{c1} \cdot S \\
\frac{dE}{dt} &= k_{c1} \cdot S - k_{c2} \cdot E \\
\frac{dP}{dt} &= k_{c2} \cdot E - k_d \cdot P \\
\frac{dB}{dt} &= k_d \cdot P - k_k \cdot B
\end{aligned} \tag{S1}$$

For this system, a non-zero equilibrium exists only if $k_b = k_{c1}$, for any sized equilibrium population $S_0 > 0$. Any deviation of $S_0 \cdot k_b/k_{c1}$ results in unabated proliferation or depletion of S. Moreover, S_0 and the equilibrium β-cell population ($B_0 = S_0 \cdot k_{c1}/k_k$) are sensitive to any deviation in k_{c1}/k_k.

We may also consider an external threshold on cell growth K_S, for example representative of nutrient limitations or contact inhibition:

$$\begin{aligned}
\frac{dS}{dt} &= k_b \cdot S \frac{K_S}{K_S + S} - k_{c1} \cdot S \\
\frac{dE}{dt} &= k_{c1} \cdot S - k_{c2} \cdot E \\
\frac{dP}{dt} &= k_{c2} E - k_d \cdot P \\
\frac{dB}{dt} &= k_d \cdot P - k_k \cdot B
\end{aligned} \tag{S2}$$

This system yields a non-zero stable equilibrium at $S_0 = \frac{k_b - k_{c1}}{k_{c1}} K_S$, and $B_0 = S_0 \cdot k_{c1}/k_k$, so long as $k_b > k_{c1}$. Even if the stem cell population may be controlled in this scenario, B nevertheless remains highly dependent on system parameters k_b, k_{c1}, and k_k. Such sensitive systems represent incomplete solutions to the problem of tissue homeostasis and are hardly ever observed *in vivo*; feedbacks ultimately remain critical components of a robust homeostasis system.

2.2 Convergence and stability in Systems 1 and 2

In System 1, the combination of a long delay (low values of k_{c1}, k_{c2} and k_d) and a nonlinear feedback (large n) induces undesirable oscillations. As discussed in the main text, reducing the delay in the feedback can suppress the oscillations, but even if we engineer feedback within intermediate maturing populations (e.g. E), there realistically remains at least a two day delay. In System 2, including feedback through the toggle switch addresses this issue. We examined the difference between feedback control from either the β-cell population alone (System 1) or all committed cells together, i.e. endodermic, pancreatic and β-cells (System 2, see main text). Figure S2 demonstrates that including all committed cells in the feedback signal can further stabilize System 2 compared to System 1. To show the existence of a unique stable equilibrium point for System 2, we sampled all 625 combinations of five different initial values for each population ($S \in \{0.5, 1, 1.5, 2, 2.5\}$, $\{E, P, B\} \in \{0, 0.5, 1, 1.5, 2\}$) for 625 different parameter sets. All trajectories converged to the same equilibrium point for a given parameter set.

2.3 Reduced two-population model

We reduce the ODE model from a four-population to a two-population abstraction in order to simplify global stability analysis. We introduce the committed population as a variable $C = E + P + B$ and reduce the system to the two populations, S and C. Figure S2(third column) reveals consistent dynamics between the two- and four- population models under certain parameter sets. We can describe the system of four populations as the following:

$$\begin{aligned}
\frac{dS}{dt} &= f_b(S) \cdot S - f_c(S,C) \cdot S \\
\frac{dE}{dt} &= f_c(S,C) \cdot S - k_{c2} \cdot E \\
\frac{dP}{dt} &= k_{c2} \cdot E - k_d \cdot P \\
\frac{dB}{dt} &= k_d \cdot P - k_k \cdot B
\end{aligned} \tag{S3}$$

where $f_b(S)$ represents the control of stem cell division as a function of the number of stem cells and $f_c(S, C)$ represents the control of stem cell commitment as a function of the number of stem and committed cells. With $C = E + P + B$, the second equation of (S3) is reduced to:

$$\frac{dC}{dt} = \frac{dE}{dt} + \frac{dP}{dt} + \frac{dB}{dt} = f_c(S,C) \cdot S - k_k \cdot B \ . \tag{S4}$$

At steady state B, the β-cells population, can be expressed as the fraction of C. The following equations describe the partition of the committed cells for a given equilibrium point ($S_0, C_0 = E_0 + P_0 + B_0$):

$$\begin{aligned}
E_0 &= \frac{k_d k_k}{k_{c2} k_d + k_{c2} k_k + k_d k_k} C_0 \\
P_0 &= \frac{k_k k_{c2}}{k_{c2} k_d + k_{c2} k_k + k_d k_k} C_0 \\
B_0 &= \frac{k_d k_{c2}}{k_{c2} k_d + k_{c2} k_k + k_d k_k} C_0
\end{aligned} \tag{S5}$$

Although the fractional composition of C with regards to E, P and β-cells may change dynamically (Figure S2), it remains a good approximation except for a short transient. Therefore, using the equilibrium populations (S5), the system (S3) can be written as

$$\begin{aligned}
\frac{dS}{dt} &= f_b(S) \cdot S - f_c(S,C) \cdot S \\
\frac{dC}{dt} &= f_c(S,C) \cdot S - \overline{k_k} \cdot C \ ,
\end{aligned} \tag{S6}$$

where

$$\overline{k_k} = \frac{k_k k_d k_{c2}}{k_{c2} k_d + k_{c2} k_k + k_d k_k} \ . \tag{S7}$$

2.4 Stability and convergence in the two-population model

The following section proves the existence of equilibrium points in a general system with two populations and Hill-function feedbacks. With the relations in (S5), we can make a 1:1 correspondence between existing equilibrium points of the two- and four-population systems. Nevertheless, initial transient responses may differ between the models (see Figure S2).

In short, we prove that this system has a non-trivial stable equilibrium point when $k_b > k_d$. Moreover, at this equilibrium, we have that $S \geq K_S$ and $B \geq K_B$, provided that the parameters satisfy $\frac{k_d}{k_k} \geq 4\frac{K_B}{K_S}$. In the case when the condition $k_b > k_d$ fails, multiple non-trivial equilibria may exist; however, our system is monotone [11, 12], which insures global convergence to equilibrium even in that case. K_S and K_B control equilibrium population levels, and when $k_d > 4k_k$, $B = K_B$. Note that the analytical results suggest the necessity of having a nonlinear function for the feedback (Hill term), which could be biologically realized through cooperative binding of the signaling elements or a signal cascade.

2.4.1 Model and statements of results

We consider the following general system of two differential equations defined for $x = x(t) \geq 0$ and $y = y(t) \geq 0$:

$$\begin{aligned} \dot{x} &= f(x,y) &&= k_b[1-\theta_x(x)]x - k_d\theta_x(x)[1-\theta_y(y)]x \\ \dot{y} &= g(x,y) &&= k_d\theta_x(x)[1-\theta_y(y)]x - k_k y \end{aligned} \tag{S8}$$

where k_b, k_d, k_k are positive constants. In the reduced model that we are considering for tissue homeostasis, x is the stem cell population, y the committed cell population and the constants correspond respectively to the birth, differentiation and effective killing rates k_k. The continuously differentiable functions

$$\theta_x, \theta_y : [0,\infty) \to [0,1)$$

are assumed to satisfy:

$$\theta_x'(x) > 0 \text{ and } \theta_y'(y) > 0 \qquad \text{for all } x > 0,\, y > 0$$

and

$$\theta_x(1) = \theta_y(1) = \frac{1}{2}\,.$$

For the main conclusions, we will specialize to the normalized Hill functions:

$$\theta_x(x) = \frac{x^{n_x}}{1+x^{n_x}} \tag{S9}$$

and

$$\theta_y(y) = \frac{y^{n_y}}{1+y^{n_y}} \tag{S10}$$

where

$$n_x \geq 1\,, \quad n_y > 0\,. \tag{S11}$$

With these functions, and for large n_x and n_y, we have that $\theta_x(x) \approx 0$ if $x < 1$ and ≈ 1 if $x > 1$, and $\theta_y(y) \approx 0$ if $y < 1$ and ≈ 1 if $y > 1$. Thus, the first (growth) term in the definition of $\dot{x}$ in (S8) will be zero when the population x is larger than 1, while the second term, which represents the flux from the x to the y population, will be only nonzero if both $x > 1$ and $y < 1$. Intuitively, one would expect a homeostatic behaviour, which attempts to bring the value of y to a target of 1 while keeping the x population from extinction.

One could consider, more generally, functions θ_x and θ_y of the following form, with V_x, V_y, K_x, K_y not necessarily equal to 1:

$$\theta(s) = \frac{Vs^n}{K^n + s^n}\,,$$

and replace the terms $1-\theta(s)$ by $V-\theta(s)$. This more general situation corresponds to desired values of $x \approx K_x$ and $y \approx K_y$. We remark that the main conclusions also hold for this more general model. Indeed, (1) the coefficients V_x and V_y can be absorbed into the constants k_b, k_d, k_k, and (2) rescaling x and y to, respectively, $K_x x$ and $K_y y$, we may take $K_x = K_y = 1$ without loss of generality, except that the parameter k_d in the equation for y has to be replaced by $k_d' = \frac{k_d K_S}{K_B}$. However, our results below only rely upon the algebraic form of the nullclines, the qualitative directions of the flow, and the location of the steady states. Such results remain invariant when multiplying the equation for y by the constant k_d/k_d'. Thus, we may assume $k_d = k_d'$, provided that we replace k_k by $\frac{k_k k_d}{k_d'} = \frac{k_k K_B}{K_S}$. Note, that for the same reasons, we could as well make $k_k = 1$, replacing k_b by k_b/k_k and k_d by k_d/k_k.

The equilibria are the points at which the x and y nullclines:

$$X = \{(x,y) \,|\, f(x,y) = 0\}\,, \quad Y = \{(x,y) \,|\, g(x,y) = 0\}$$

intersect. Note that there is always an equilibrium at $x = y = 0$. We are interested in nonzero equilibria.

The main results will be as follows; they are proved in the next section.

Lemma 2.1 The x nullcline X is the union of the line $x = 0$ and the graph of a strictly increasing and onto function

$$\psi \,:\, [\xi, \infty) \to [0, \infty)\,,$$

where

$$\xi \,:=\, \theta_x^{-1}\left(\frac{k_b}{k_b + k_d}\right)\,.$$

See Figure S3A; the arrows in the figure indicate the sign of the x-component of the vector field. Moreover:

- If θ_x is as in (S9) then $\xi \to 1$ as $n_x \to \infty$.
- $\xi \geq 1$ if and only if $k_b \geq k_d$.
- If θ_y is as in (S10) and $n_y > 1$, then ψ has a vertical tangent at ξ.
- If both θ_x is as in (S9) and θ_y is as in (S10), then for large x the function ψ has the asymptotic form:

$$y = \psi(x) \approx c_1 x^{\frac{n_x}{n_y}} \quad \text{with } c_1 = \left(\frac{k_d}{k_b}\right)^{\frac{1}{n_y}} \tag{S12}$$

Lemma 2.2 The y nullcline Y is the graph of a strictly increasing and onto function

$$\gamma \,:\, [0, \infty) \to [0, \infty)\,.$$

See Figure S3B; the arrows in the figure indicate the sign of the y-component of the vector field. Moreover:

- For all x:

  $$\gamma(x) \;\geq\; \min\left\{1, \frac{k_d}{2k_k}\theta_x(x)x\right\}$$

  and, in particular

  $$\gamma(x) \;\geq\; \min\left\{1, \frac{k_d}{4k_k}\right\} \qquad \text{for all } x \geq 1\,.$$

- If $k_b \geq k_d$, then

  $$\gamma(\xi) \;\geq\; \min\left\{1, \frac{k_d}{4k_k}\right\}\,.$$

- If both θ_x is as in (S9) and θ_y is as in (S10), then for large x the function γ has the asymptotic form:

$$y = \gamma(x) \approx c_2 x^{\frac{1}{n_y+1}}\,, \quad \text{with } c_2 = \left(\frac{k_d}{k_k}\right)^{\frac{1}{n_y+1}}\,. \tag{S13}$$

These results suggest that the complete phase diagram is qualitatively as shown in Figure S3C. If there is a unique positive equilibrium, as in the figure, then the direction of the arrows shows that every trajectory starting from an initial condition with nonzero $x(0)$ must converge to that positive equilibrium.

However, the figure is misleading. Figure S3D shows a situation where more than one positive steady state exists. This example has θ_x as in (S9) and θ_y is in (S10), with $n_x = 2$, $n_y = 1$, and the following parameters: $k_b = 0.15$, $k_d = 4$, $k_k = 0.04$.

Thus, to prove that Figure S3C is indeed the correct picture, and global stability to a unique positive equilibrium holds, we need to impose some constraints on parameters.

Corollary 2.3 If the following condition is satisfied:

$$k_b \ \geq \ k_d \ \geq 4k_k \, , \tag{S14}$$

then every equilibrium point (x, y) different from $(0, 0)$ has the property that $x \geq 1$ and $y \geq 1$.

Corollary 2.4 If both θ_x is as in (S9) and θ_y is as in (S10), then there exists at least one positive equilibrium.

Corollary 2.5 In addition to the hypotheses of Corollary 2.4, suppose that $k_b \geq k_d$ and $n_x \geq 2$. Then, there is at most one positive equilibrium.

We summarize as follows.

Theorem 1 *Suppose that:*

- *θ_x is as in (S9) and θ_y is as in (S10),*
- *$n_x \geq 2$,*
- *$k_b \ \geq \ k_d$.*

Then, there is a unique positive equilibrium $(\bar{x}, \bar{y})$. All trajectories, except for those starting with $x(0) = 0$, converge to $(\bar{x}, \bar{y})$. Moreover, if also

- *$k_d \ \geq 4k_k$*

then $\bar{x} \geq 1$ and $\bar{y} \geq 1$. ∎

Remark. It is worth noting that, if θ_x is as in (S9) and θ_y is as in (S10), then as $n_x, n_y \to \infty$, the set Y takes the limiting form shown in Figure S3E.

To be precise, we show that, as $n \to \infty$, (1) for each fixed $x < 1$, $\gamma(x, n) \to 0$ and (2) for each fixed $x > 1$, $\gamma(x, n) \to 1$. To verify (1), we pick any $x < 1$. As $\theta_x(x) \to 0$ when $n_x \to \infty$, also $(k_d/k_k)\theta_x(x)x \to 0$; since $G^{-1}(0) = 0$, we conclude as claimed. To show (2), we pick $x > 1$, and pick y so that $G(y) = (k_d/k_k)\theta_x(x)x$. If $y > 1$, then $G(y) \to +\infty$ (because $\theta_y(y) \to 1$) as $n_y \to \infty$, so for large enough n_y, $G(y) > (k_d/k_k)\theta_x(x)x$, a contradiction. If instead $y < 1$, then $1 - \theta_y(y) \approx 1$, and thus (using that $k_d > k_k$, and thus also $(k_d/k_k)\theta_x(x)x \approx (k_d/k_k)x > 1$), as $n_x \to \infty$ we have that $y \approx G(y) \approx (k_d/k_k)x > 1$, contradicting the assumption $y < 1$. It follows that $y = 1$ in the limit, as claimed. This means that, for large Hill exponents, one may expect the value of y at nonzero steady states to be approximately 1. ∎

2.4.2 Proofs

Proof of Lemma 2.1

The set X is the union of the line $x = 0$ and the solution set of

$$k_b(1-\theta_x(x)) - k_d\theta_x(x)[1-\theta_y(y)] = 0 \tag{S15}$$

and we are interested in characterizing this latter solution set. We may rewrite the above equation as

$$1-\theta_y(y) \;=\; \frac{k_b}{k_d}\frac{1-\theta_x(x)}{\theta_x(x)} \;=\; \frac{k_b}{k_d}\left(\frac{1}{\theta_x(x)}-1\right). \tag{S16}$$

A solution $y = \psi(x)$ exists if and only if the right hand side is in the range $(0, 1]$, which amounts to saying (since

$$\frac{k_b}{k_d}\left(\frac{1}{\theta_x(x)}-1\right) \;>\; 0$$

because $\theta_x(x) < 1$) that

$$\frac{k_b}{k_d}\left(\frac{1}{\theta_x(x)}-1\right) \;\leq\; 1\,.$$

This property is equivalent to $\frac{1}{\theta_x(x)} - 1 \leq \frac{k_d}{k_b}$, that is,

$$\theta_x(x) \;\geq\; \frac{k_b}{k_b+k_d} = \theta_x(\xi)$$

which is the same as asking $x \geq \xi$. Thus $y = \psi(x)$ is defined for $x \geq \xi$. As both the left and right-hand sides of (S16) are strictly decreasing functions of their arguments, it follows that ψ is increasing, and it is clearly differentiable by the same reasoning. Note that $\psi(\xi) = 0$, by definition of ξ, and that $\psi(x) \to \infty$ as $x \to \infty$, because $\theta_x(x) \to 1$ as $x \to \infty$, which implies that the right-hand side of (S16) converges to zero, and thus $\theta_y(y) \to 1$.

The direction of the vector field is clear from the fact that, for any fixed y, the expression

$$\frac{1}{\theta_x(x)} - 1 - \frac{k_d}{k_b}\left[1-\theta_y(y)\right]$$

is positive when x is very small and negative when x is very large.

Suppose that θ_x is as in (S9), and use here a subscript n_x to indicate its dependence on n_x. Pick any $\varepsilon > 0$. There is uniform convergence $\theta_{x,n_x}(z) \to 0$ for $z \leq 1-\varepsilon$ and $\theta_{x,n_x}(z) \to 1$ for $z \geq 1+\varepsilon$ as $n_x \to \infty$. Therefore, for each fixed number $\eta \in (0, 1)$, it follows that $1-\varepsilon < \theta_{x,n_x}^{-1}(\eta) < 1+\varepsilon$ Since ε was arbitrary, this means that $\theta_{x,n_x}^{-1}(\eta) \to 1$ as $n_x \to \infty$. In particular, applied to $\eta = \xi$, we have that $\xi \to 1$ as $n_x \to \infty$.

Observe that, $k_b \geq k_d$ if and only if $\frac{k_b}{k_b+k_d} \geq \frac{1}{2} = \theta_x(1)$. Thus, since θ_x is an increasing function, $k_b \geq k_d$ if and only if $\xi = \theta_x^{-1}(\frac{k_b}{k_b+k_d}) \geq \theta_x^{-1}(\theta_x(1)) = 1$, as claimed.

We next show that if θ_y is as in (S10) and $n_y > 1$, then ψ has a vertical tangent at ξ. Taking implicit derivative with respect to x and then a limit as $x \searrow \xi$ and $y \searrow 0$ in (S16), we have that:

$$\psi'(x) \;\to\; \frac{1}{\theta_y'(0)}\,\frac{k_b}{k_d}\,\frac{\theta_x'(\xi)}{\theta_x(\xi)^2}$$

as $x \searrow \xi$, and therefore this limit is $+\infty$ if $\theta_y'(0) = 0$.

The asymptotic form of the nullcline, for large x, when both θ_x is as in (S9) and θ_y is as in (S10), is shown as follows. The equality in (S16) becomes:

$$\frac{1}{1+y^{n_y}} \;=\; \frac{k_b}{k_d}\frac{1}{x^{n_x}}$$

which means that

$$\psi(x) \ = \ \left(\frac{k_d}{k_b}x^{n_x} - 1\right)^{\frac{1}{n_y}} \ \approx \ c_1 x^{\frac{n_x}{n_y}}$$

with $c_1 = \left(\frac{k_d}{k_b}\right)^{\frac{1}{n_y}}$.

This completes the proof of Lemma 2.1. ∎

Proof of Lemma 2.2

The y nullcline set Y consists of the solutions of

$$G(y) = \frac{y}{1-\theta_y(y)} = (k_d/k_k)\theta_x(x)x\,.$$

The function G is continuous and strictly increasing (because θ_y is strictly increasing), $G(0) = 0$, and $G(y) \to \infty$ as $y \to \infty$. Therefore G is invertible, and thus Y is the graph of the strictly increasing function

$$y = \gamma(x) = G^{-1}((k_d/k_k)\theta_x(x)x)$$

which clearly satisfies that $\gamma(0) = 0$. As $x \to \infty$, $\theta_x(x) \to 1$, so $\theta_x(x)x \to \infty$, which implies that $\gamma(x) \to \infty$ as well.

The direction of the vector field is clear from the fact that

$$g(x, y) \ = \ k_d\theta_x(x)[1-\theta_y(y)]x - k_k y \ > \ 0$$

if and only if

$$G(y) \ < \ (k_d/k_k)\theta_x(x)x$$

which is the same as $y < \gamma(x)$, and that the expression is < 0 if and only if $y > \gamma(x)$.

We claim that:

$$\gamma(x) < 1 \quad \text{if and only if} \quad \gamma(x) > \frac{k_d}{2k_k}\theta_x(x)x\,.$$

Indeed, suppose that $y = \gamma(x)$, that is $\frac{y}{1-\theta_y(y)} = (k_d/k_k)\theta_x(x)x$. Note that $y < 1$ is equivalent to $\theta_y(y) < 1/2$, which is the same as $\frac{y}{1-\theta_y(y)} < 2y$. Thus

$$y < 1 \quad \text{if and only if} \quad 2y > \frac{k_d}{k_k}\theta_x(x)x$$

as claimed. It follows that

$$\gamma(x) \ \geq \ \min\left\{1, \frac{k_d}{2k_k}\theta_x(x)x\right\}\,.$$

In particular, when $x \geq 1$, $\theta_x(x) \geq 1/2$, so $\theta_x(x)x \geq 1/2$ and therefore $\frac{k_d}{2k_k}\theta_x(x)x \geq \frac{k_d}{4k_k}$

If in addition $k_b \geq k_d$, then $\xi \geq 1$ by Lemma 2.1, and thus $\gamma(\xi) \geq \min\left\{1, \frac{k_d}{4k_k}\right\}$.

Finally, suppose that θ_x is as in (S9) and θ_y is as in (S10). We have that $y = \gamma(x)$ must satisfy:

$$y + y^{n_y+1} = \frac{k_d}{k_k}\frac{x^{n_x+1}}{1+x^{n_x}}$$

When $x \to \infty$, also $y \to \infty$, so $\frac{k_d}{k_k}\frac{x^{n_x+1}}{1+x^{n_x}} \approx \frac{k_d}{k_k}x$ and $y + y^{n_y+1} \approx y^{n_y+1}$. Therefore, $y \approx c_2 x^{\frac{1}{ry+1}}$ with $c_2 = \left(\frac{k_d}{k_k}\right)^{\frac{1}{ry+1}}$ as $x \to \infty$ (and so also $y \to \infty$). ∎

Proof of Corollary 2.3

When $k_b \geq k_d$, Lemma 2.1 insures that $\xi \geq 1$ and Lemma 2.2 insures that $\gamma(\xi) \geq \min\left\{1, \frac{k_d}{4k_k}\right\}$. Thus, if also $k_d \geq 4k_k$, it follows that $\gamma(\xi) \geq 1$. Given the forms of the nullclines for x and y, any equilibrium should have $x \geq \xi \geq 1$. Thus $y = \gamma(x) \geq \gamma(\xi) \geq 1$. ∎

Proof of Corollary 2.4

By Lemma 2.1, the function ψ that describes the nontrivial branch of the x nullcline is defined for $x \geq \xi > 0$, with $\psi(\xi) = 0$, and has the asymptotic form $\psi(x) \approx c_1 x^{\frac{n_x}{n_y}}$. By Lemma 2.2, the function γ that describes the y nullcline has $\gamma(0) = 0$ and has the asymptotic form $\gamma(x) \approx c_2 x^{\frac{1}{n_y+1}}$. To show a nonzero intersection between the graphs of these two functions, it suffices to know that $\psi(x) > \gamma(x)$ for some x. This is clear because, for some positive constant c:

$$\frac{\psi(x)}{\gamma(x)} \approx c\,\frac{x^{\frac{n_x}{n_y}}}{x^{\frac{1}{n_y+1}}} = c\,x^{\frac{n_x}{n_y}-\frac{1}{n_y+1}} \to \infty \quad \text{as } x \to \infty$$

since $\frac{n_x}{n_y} - \frac{1}{n_y+1} > 0$ (using $n_x \geq 1$ and $n_y > 0$). ∎

Proof of Corollary 2.5

We must show that there is only one intersection of the zero sets of

$$k_b(1-\theta_x(x)) = k_d\theta_x(x)[1-\theta_y(y)]$$

and

$$k_d\theta_x(x)[1-\theta_y(y)]x = k_k y\,.$$

We can equally well replace the second equation by:

$$k_b(1-\theta_x(x))x = k_k y\,.$$

Thus, it is enough to show that the graphs of these two functions:

$$\alpha(x) = \frac{k_b}{k_k}(1-\theta_x(x))x$$

$$\beta(x) = \theta_y^{-1}\left(1 - \frac{k_b}{k_d}\left[\frac{1}{\theta_x(x)} - 1\right]\right)$$

(defined for $x \geq \xi$, where, as before, $\theta_x(\xi) = \frac{k_b}{k_b+k_d}$) intersect at only one point. Since β is a strictly increasing function defined for $x \geq \xi \geq 1$ (this last inequality uses that $k_b \geq k_d$), with $\beta(0) = 0$ and $\alpha(x) > 0$ for all x, it is sufficient to show that $\alpha'(x) \leq 0$ for $x \geq 1$. Indeed, $[1-\theta_x(x)]x = \frac{x}{1+x^{n_x}}$, so $\alpha'(x)$ vanishes only at $x = \bar{x} = (n_x-1)^{-1/n_x}$ and is negative for $x > \bar{x}$. Since $n_x \geq 2$, $n_x - 1 \geq 1$, from which it follows that $\bar{x} \leq 1$. Thus, as required, $\alpha'(x) \leq 0$ for $x \geq 1 \geq \bar{x}$. ∎

2.5 Robustness to the rate k_k

We present an informal argument to estimate the order of the dependence of the steady state on the degradation rate k_k of y, when all other parameters are kept constant. This is an important property of the homeostasis system as y, the committed cell population should be independent of external perturbations. We assume that θ_x is as in (S9) and θ_y is as in (S10) and that we are in the asymptotic regime. The nullclines have the forms in (S12)-(S13):

$$\psi(x) \approx c_1 x^{\frac{n_x}{n_y}} \quad \text{with } c_1 = \left(\frac{k_d}{k_b}\right)^{\frac{1}{n_y}}$$

$$\gamma(x) \approx c_2 x^{\frac{1}{n_y+1}}\,, \quad \text{with } c_2 = \left(\frac{k_d}{k_k}\right)^{\frac{1}{n_y+1}}\,.$$

At equilibrium, $\psi(x) = \gamma(x)$ gives that x has order ${k_k}^p$ and therefore $y = c_1 x^{\frac{n_x}{n_y}}$ has order ${k_k}^q$, where

$$p = \frac{n_y}{n_y - n_x n_y - n_x}\,, \quad q = \frac{n_x}{n_y - n_x n_y - n_x}\,.$$

If $n_x = n_y = n$, these simplify to $p = q = -1/n$. For example, with $n_x = n_y = 4$, we expect that x and y will grow like ${k_k}^{-1/4}$.

2.6 Simulations

We show next nullclines and simulations for these values:

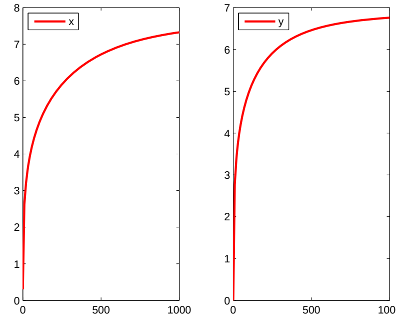
$$k_b = 3,\ k_d = 2,\ n = 4,$$

and k_k varying by 3 orders of magnitude:

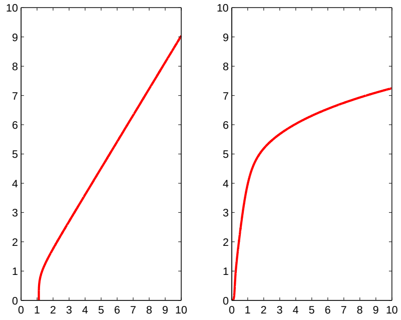
$$k_k = 0.001, 0.01, 0.1, 1\,.$$

Shown below are simulations for $x(t)$ and $y(t)$ (with initial states $x(0) = 0.3$, $y(0) = 0$), followed by the respective x and y nullclines, for values of k_k respectively as above. Observe the very weak dependence of the steady state values $\bar{x}$ and $\bar{y}$ on the rate k_k, consistently with their order being ${k_k}^{-1/4}$, since $0.001^{-1/4} \approx 5.6$, $0.01^{-1/4} \approx 3.16$, $0.1^{-1/4} \approx 1.78$, and $1^{-1/4} = 1$.

$k_k = 0.001$:

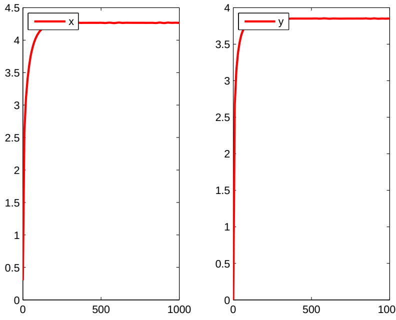


$x(t)$ and $y(t)$

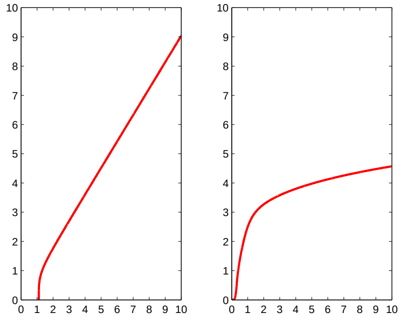


x and y nullclines

$k_k = 0.01$:

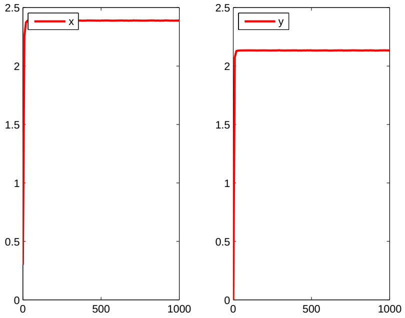


$x(t)$ and $y(t)$

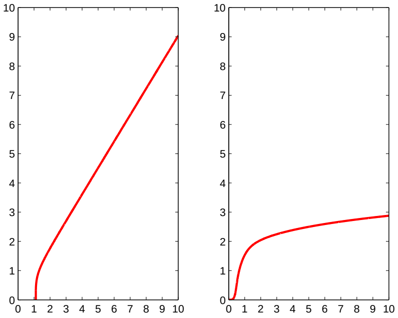


x and y nullclines

$k_k = 0.1$:

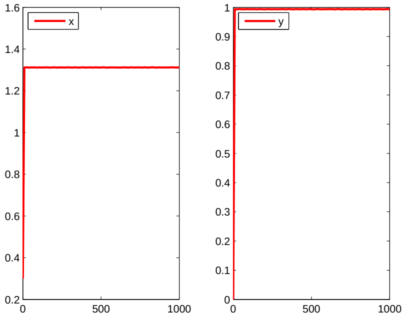


$x(t)$ and $y(t)$

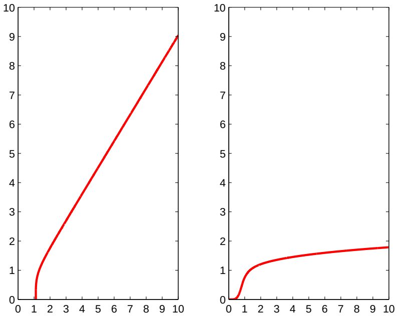


x and y nullclines

$k_k = 1$:



$x(t)$ and $y(t)$



x and y nullclines

3 Methods for the Langevin modeling of Systems 2-4

We perform chemical Langevin simulations [13] of the various system designs (Figures 3D-F, 4) and use these implementations to perform various analyses (Figures 5, 7 and 8). We add modulated white noise to each reaction at each integration step, with noise amplitude controlled by the parameter Ω. We refer to Ω as 'cell volume', because it relates concentration to molecular count. For example, a volume of 200 denotes that a particular molecule's concentration of 1 (which is the average value for most of the active components in our models) corresponds to a total of 200 molecules per cell. All the equations in the following are written as ODE to lighten the notation, but are simulated as chemical Langevin equations [15]. For example, for an ODE system written as

$$\frac{d\,\mathbf{X}}{dt} = \sum_{j=1}^{m} \boldsymbol{\nu}_j r_j(\mathbf{X}) \ ,$$

where $\mathbf{X}$ is the concentration of the different species, $\boldsymbol{\nu}_j$ the stoichiometric vector corresponding to the j-th column of the stoichiometric matrix and r_j the rate function of the j-th reaction. The increment of $\mathbf{X}$ for a time interval τ using the corresponding chemical Langevin equation will be

$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \sum_{j=1}^{m} \boldsymbol{\nu}_j r_j(\mathbf{X})\,\tau + \sum_{j=1}^{m} \boldsymbol{\nu}_j \sqrt{\frac{r_j(\mathbf{X})}{\Omega}\,\tau}\, N_j \ ,$$

where $N_j \sim \mathcal{N}(0,1)$ is a normal random variable with mean 0 and variance 1. Note that for the calculus of $r_j(\mathbf{X})$, we used a multiple state procedure as described in [14] for a higher precision.

We aim to optimize the systems such that the number of committed cells remains constant. We define the objective function as the signal to noise ratio S/N of the fraction of committed cells ρ_c in a simulation of duration T:

$$S/N = \frac{\overline{\rho_c}}{\sqrt{1/T \int_0^T \left(\rho_c(t) - \overline{\rho_c}\right)^2 dt}} \qquad \text{where} \quad \overline{\rho_c} = 1/T \int_0^T \rho_c(t) dt \tag{S17}$$

Note that the fraction of committed cells is similar in all systems due to the quorum sensing module (see Sec. 3.1) and therefore, the S/N value is not biased by large difference in $\overline{\rho_c}$ between different systems. S/N measurement begins after simulations have been allowed to somewhat equilibrate (generally after 500 hours of simulation).

The following sections detail the equations for Systems 2 to 4. Table S1 summarizes the parameter values for the three systems. We approximate the evolution of each component in each model (Figures 3-4) to follow a Hill kinetic with a coefficient of $n = 4$. k_p^α and H_α denote the maximum rate and half-rate constant, respectively, for the component α. Degradation follows mass-action kinetics (rate k_d^α) for all components. Finally, we describe diffusion as a linear function (rate k_{diff}) of the difference between internal and external (AI) concentrations. We constrain the maximum number of cells in the simulation (N_{max}) to be less than 150.

3.1 Quorum sensing module

In Systems 2-4, uncommitted cells signal through $AI1$. The toggle switch, comprised of $R6$ and $R7$ cross-inhibition, regulates $I1$ expression. $I1$ subsequently controls $AI1$ production. $R7$, which is produced only in committed cells, inhibits $AI1$ production. In contrast, $R6$, which is produced only in uncommitted cells, inhibits the signal for committed cells ($AI2$). We approximate $AI1$ and $AI2$ as being directly dependent upon $R6$ and

$R7$, and we describe the concentration of these components according to the following equations:

$$\begin{aligned}\frac{d\ AI1}{dt} &= k_p^{AI1}\frac{H_{AI1}^n}{H_{AI1}^n + R7^n}\\ &-k_d^{AI1}AI1\\ &+k_{\text{diff}}(AI1_{out} - AI1) \qquad (\text{S18})\\ \frac{d\ AI2}{dt} &= k_p^{AI2}\frac{H_{AI2}^n}{H_{AI2}^n + R6^n}\\ &-k_d^{AI2}AI2\\ &+k_{\text{diff}}(AI2_{out} - AI2) \qquad (\text{S19})\end{aligned}$$

The following equations describe the homogeneous extracellular AI concentration:

$$\begin{aligned}\frac{d\ AI1_{out}}{dt} &= -\sum_{Cells}\frac{k_{\text{diff}}}{N_{max}}(AI1_{out} - AI1_{Cells})\\ &-k_d^{AI1}AI1_{out} \qquad (\text{S20})\\ \frac{d\ AI2_{out}}{dt} &= -\sum_{Cells}\frac{k_{\text{diff}}}{N_{max}}(AI2_{out} - AI2_{Cells})\\ &-k_d^{AI2}AI2_{out} \qquad (\text{S21})\end{aligned}$$

The equilibrium concentrations in uncommitted cells of $AI1$ and $AI2$, $\overline{AI1}$ and resp. $\overline{AI2}$, for given population sizes ρ_u and ρ_c can be evaluated assuming that the production rate in eq. (S18) and (S19) is either zero or maximal ($k_p^{AI\alpha}$) depending on the state of each cell. If the size of both cell populations is expressed as a fraction of N_{max}: $\rho_u = \frac{N_u}{N_{max}}$ for uncommitted cells and $\rho_c = \frac{N_c}{N_{max}}$ for the committed cells, $\overline{AI1}$ and $\overline{AI2}$ can be written as:

$$\begin{aligned}\overline{AI1}(\rho_u,\rho_c) &= \frac{k_p^{AI1}}{k_d^{AI1}}\frac{(k_d^{AI1}+k_{\text{diff}})(k_d^{AI1}+\rho_u k_{\text{diff}}) + \rho_c k_{\text{diff}} k_d^{AI1}}{(k_{\text{diff}}+k_d^{AI1})(k_{\text{diff}}+k_d^{AI1}+\rho_u k_{\text{diff}}+\rho_c k_{\text{diff}})} \qquad (\text{S22})\\ \overline{AI2}(\rho_u,\rho_c) &= \frac{k_p^{AI2}}{k_d^{AI2}}\frac{\rho_c k_{\text{diff}}^2}{(k_{\text{diff}}+k_d^{AI2})(k_{\text{diff}}+k_d^{AI2}+\rho_u k_{\text{diff}}+\rho_c k_{\text{diff}})} \qquad (\text{S23})\end{aligned}$$

The production of the component $A1$ (or for System 4, the repressor $R1$) is controlled by the receptor $Rec1$ to which $AI1$ binds. Similarly, $AI2$ binds to $Rec2$ and activates the production of the repressor $R2$. In this model, we simplify the expressions of $A1/R1$ and $R2$ as depending directly on the concentration of $AI1$ or $AI2$ following a Hill-type equation. In all systems, the half-rate constants for the production terms of $A1/R1$ and $R2$ within the population control modules, is adjusted to trigger the cell-decision processes for a threshold of the fraction of uncommitted cells ρ_u around 0.45 and of committed cells ρ_c around 0.4. It means that the half-rates are equal to $\overline{AI1}(\rho_u = 0.45, \rho_c = 0.4)$ for $A1/R1$ production and $\overline{AI2}(\rho_u = 0.45, \rho_c = 0.4)$ for $R2$ production as defined above. Therefore,

$$\begin{aligned}\frac{d\ A1}{dt} &= k_p^{A1}\frac{AI1^n}{H_{A1}^n + AI1^n} &&\text{with}\quad H_{R1} = \overline{AI1}(0.45, 0.4)\\ &-k_d^{A1}A1 &&\qquad\text{for Systems 2 and 3} \qquad (\text{S24})\\ \frac{d\ R1}{dt} &= k_p^{R1}\frac{H_{R1}^n}{H_{R1}^n + AI1^n} &&\text{with}\quad H_{R1} = \overline{AI1}(0.45, 0.4)\\ &-k_d^{R1}R1 &&\qquad\text{for System 4} \qquad (\text{S25})\\ \frac{d\ R2}{dt} &= k_p^{R2}\frac{H_{R2}^n}{H_{R2}^n + AI2^n} &&\text{with}\quad H_{R2} = \overline{AI2}(0.45, 0.4)\\ &-k_d^{R2}R2 && \qquad (\text{S26})\end{aligned}$$

3.2 AND gate and toggle switch

The AND gate integrates the information from the two quorum sensing modules (and the oscillator or the throttle in Systems 3 and 4, respectively). The AND gate in Systems 3 and 4 include an activator $A3$, and System 3 also includes an additional repressor ($R4$). $R5$ serves as the output and interacts with downstream modules (e.g., the toggle switch in Systems 2-4). The System 2 AND gate only contains $R5$, which we describe by the following equation:

$$\begin{aligned}\frac{d\ R5}{dt} &= k_p^{R5}\frac{A1^n}{H_{R5-1}^n + A1^n}\frac{H_{R5-2}^n}{H_{R5-2}^n + R2^n} \\ &\quad -k_d^{R5} R5 \end{aligned} \tag{S27}$$

In System 3, the AND gate includes two additional elements ($A3$ and $R4$):

$$\begin{aligned}\frac{d\ A3}{dt} &= k_p^{A3}\frac{A1^n}{H_{A3-1}^n + A1^n}\frac{H_{A3-4}^n}{H_{A3-4}^n + R4^n} \\ &\quad -k_d^{A3} A3 \end{aligned} \tag{S28}$$

$$\begin{aligned}\frac{d\ R4}{dt} &= k_p^{R4}\frac{H_{R4}^n}{H_{R4}^n + Ro2^n} \\ &\quad -k_d^{R4} R4 \end{aligned} \tag{S29}$$

$$\begin{aligned}\frac{d\ R5}{dt} &= k_p^{R5}\frac{A3^n}{H_{R5-3}^n + A3^n}\frac{H_{R5-2}^n}{H_{R5-2}^n + R2^n} \\ &\quad -k_d^{R5} R5 \end{aligned} \tag{S30}$$

In System 4, the throttle acts on $R5$ through the signaling molecule $AI3$. Furthermore, System 4 includes $A3$:

$$\begin{aligned}\frac{d\ A3}{dt} &= k_p^{A3}\frac{H_{A3-1}^n}{H_{A3-1}^n + R1^n}\frac{H_{A3-2}^n}{H_{A3-2}^n + R2^n} \\ &\quad -k_d^{A3} A3 \end{aligned} \tag{S31}$$

$$\begin{aligned}\frac{d\ R5}{dt} &= k_p^{R5}\frac{A3^n}{H_{R5-3}^n + A3^n}\frac{H_{R5-t}^n}{H_{R5-t}^n + AI3^n} \\ &\quad -k_d^{R5} R5 \end{aligned} \tag{S32}$$

The toggle switch consists of the two repressors, $R6$ and $R7$, which inhibit each other. Cells initially have the toggle in the uncommitted state, with high $R6$ and low $R7$. $R5$ controls toggle switching by inhibiting $R6$. The equations for $R6$ and $R7$ are the following:

$$\begin{aligned}\frac{d\ R6}{dt} &= k_p^{R6}\frac{H_{R6-5}^n}{H_{R6-5}^n + R5^n}\frac{H_{R6-7}^n}{H_{R6-7}^n + R7^n} \\ &\quad -k_d^{R6} R6 \end{aligned} \tag{S33}$$

$$\begin{aligned}\frac{d\ R7}{dt} &= k_p^{R7}\frac{H_{R7}^n}{H_{R7}^n + R6^n} \\ &\quad -k_d^{R7} R7 \end{aligned} \tag{S34}$$

The equation for $R7$ differs for System 4, and is discussed below.

3.3 Cell fate

We describe uncommitted cells as potentially proliferative and immortal. In contrast, committed cells cannot grow but rather die. We model the control of proliferation in uncommitted cells using growth arrest factor

(GAF), which is regulated by quorum sensing activity. We model differentiation as potentiated by the expression of a transcription factor such as $GATA4$. Once above a certain threshold, $GATA4$ initiates slow cell death as approximated by a kinetic model described below.

3.3.1 Division process

GAF is controlled by $A1/R1$ (quorum sensing of the uncommitted cells):

$$\begin{aligned} \frac{d\ GAF}{dt} &= k_p^{GAF} \frac{A1^n}{H_{GAF}^n + A1^n} \\ &\quad - k_d^{GAF} GAF \qquad \text{for Systems 2 and 3} \qquad \text{(S35)} \\ \frac{d\ GAF}{dt} &= k_p^{GAF} \frac{H_{GAF}^n}{H_{GAF}^n + R1^n} \\ &\quad - k_d^{GAF} GAF \qquad \text{for System 4} \qquad \text{(S36)} \end{aligned}$$

The cell grows if GAF lies below a threshold th_{GAF}. To model growth, we use an integrator for the division depending on GAF level,

$$\frac{d\ Div}{dt} = k_b \Theta(th_{GAF} - GAF) - \frac{k_b}{3} \Theta(GAF - th_{GAF})$$

with $Div = 0$ at the time at cell division and Θ represents the Heaviside function ($\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$). k_b denotes the division rate. We choose $k_b = \frac{1}{96} h^{-1}$, such that the average time for division (in absence of GAF), is set to $96h$. Division occurs when $Div(t) \geq 1$. The two daughter cells inherit all concentrations of the mother cell's components except Div, which is reset to zero.

3.3.2 Commitment process

$R6$ inhibits $GATA$, and $GATA$ production begins when the toggle switches to low $R6$:

$$\begin{aligned} \frac{d\ GATA}{dt} &= k_p^{GATA} \frac{H_{GATA}^n}{H_{GATA}^n + R6^n} \\ &\quad - k_d^{GATA} GATA \qquad \text{(S37)} \end{aligned}$$

When $GATA$ concentration reaches a threshold, the cell becomes differentiated and no longer proliferates. We describe the lifetime of committed cells by the following:

$$\frac{d\ Death}{dt} = k_k$$

with $Death = 0$ at the time of differentiation and k_k denotes the death rate. A cell dies when $Death$ reaches 1. We choose $k_k = \frac{1}{200} h^{-1}$, such that the average time for the death of a β-cell is set to $200h$.

3.4 System 3 – implementation of an oscillator

To break symmetry between individual cells in the population, oscillations are introduced through a relaxation oscillator with a design similar to previous experimental studies in prokaryotic and eukaryotic cells [16, 17, 18, 19, 20, 21]. The simplest possible oscillator is made of a component Ao that activates itself (autopositive feedback) and regulates the expression of a repressor Ro that inhibits Ao. Two successive components integrate oscillator dynamics with the AND gate. A second repressor $Ro2$ acts on $R4$, which in turn represses commitment. With proper parameter values, this system generates short intervals of low $R4$, with an irregular latency where $R4$

concentration remains high. The equations of the oscillator are as follows:

$$\begin{aligned}
\frac{d\ Ao}{dt} &= k_p^{Ao}\left(k_0^{Ao}+\frac{Aosc^n}{H_{Ao-A}^n+Ao^n}\right)\frac{H_{Ao-R}^n}{H_{Ao-R}^n+Ro^n}\\
&\quad -k_d^{Ao}Aosc && \text{(S38)}\\
\frac{d\ Ro}{dt} &= k_p^{Ro}\frac{Ao^n}{H_{Ro}^n+Ao^n}\\
&\quad -k_d^{Ro}Ro && \text{(S39)}\\
\frac{d\ Ro2}{dt} &= k_p^{Ro2}\frac{Ao^n}{H_{Ro2}^n+Ao^n}\\
&\quad -k_d^{Ro2}Ro2 && \text{(S40)}
\end{aligned}$$

Linking the oscillator and the quorum sensing module to the toggle switch occurs in two steps as described above. The element $A3$ is controlled by both $A1$ and $R4$, therefore $A3$ is produced only when $A1$ is high (enough uncommitted cells) and $R4$ is low.

3.5 System 4 – implementation of a throttle

We implement a throttle design involving a third quorum sensing molecule, $AI3$, as another mechanism to inhibit the simultaneous commitment of cells in a population. We design the activator At to control the production of $AI3$, such that $AI3$ is produced transiently when the toggle is switching. $AI3$ diffuses through the membrane and gives rise to an external concentration $AI3_{out}$ that enters other cells. The throttle equations are as follows:

$$\begin{aligned}
\frac{d\ At}{dt} &= k_p^{At}\frac{H_{At-6}^n}{H_{At-6}^n+R6^n}\\
&\quad -k_d^{At}At && \text{(S41)}\\
\frac{d\ AI3}{dt} &= k_p^{AI3}\frac{At^n}{H_{AI3-t}^n+At^n}\frac{H_{AI3-7}^n}{H_{AI3-7}^n+R7^n}\\
&\quad -k_d^{AI3}AI3\\
&\quad +k_{\text{diff}}(AI3_{out}-AI3) && \text{(S42)}\\
\frac{d\ AI3_{out}}{dt} &= -\sum_{Cells}\frac{k_{\text{diff}}}{N_{max}}(AI3_{out}-AI3_{Cells})\\
&\quad -k_d^{AI3}AI3_{out} && \text{(S43)}
\end{aligned}$$

The remaining elements are similar to System 2, except for two key differences: first, the addition of $A3$ (see Eq. S31 and S32), and second, the control of $R7$ by At instead of $R6$:

$$\begin{aligned}
\frac{d\ R7}{dt} &= k_p^{R7}\frac{At^n}{H_{R7}^n+At^n}\\
&\quad -k_d^{R7}R7 && \text{(S44)}
\end{aligned}$$

3.6 Spatial simulations

We implement spatially resolved multicellular simulations using the Langevin model to analyze spatiotemporal commitment patterns (discussed further in Supplementary Text 5.4). We represent the spatial distribution of individual cells in a manner similar to that employed in the Gillespie-simulations (discussed in the following section). The extracellular volume of the system is divided into a 6x5x5 grid with $N_{max} = 150$ boxes, each with the same volume as a cell. Each cell occupies one box on the grid. Diffusion can occur between the cell and its box or between the boxes. As diffusion is physically faster than any other process in the cell, we simulate it as an ODE without the Langevin noise term. Diffusion occurs from one edge to edge (periodic boundary conditions) in order to mimic a larger system and avoid border effects. A dividing cell can push neighboring cells into the nearest empty box; therefore, daughter cells are adjacent at time of division.

4 Methods for the Gillespie modeling of Systems 2-3

In addition to the Langevin model, we implement more mechanistically detailed models of Systems 2 and 3 that do not make quasi-steady-state assumptions for many of the network components. We use these models primarily for analysis of hysteresis in the UPC module (see section *Clustered sensitivity analysis* in the main text), but also use the Gillespie models to confirm several results observed in the Langevin simulations (including patterning analysis). We use a standard rate-equation approach and monitor the spatiotemporal evolution of cells, proteins, and signaling molecules with a Gibson-modified Gillespie algorithm. We describe the full tissue homeostasis system as a set of discrete stochastic reactions occurring in cells using a previously described multicellular spatiotemporal simulation environment [22, 23]. The simulation platform tracks the temporal evolution of intracellular reactions within individual cells that grow on a 2D grid, as well as the spatiotemporal evolution of the cells themselves and extracellular signaling molecules that diffuse among them. We model cell growth as a stochastic buildup of a species "Volume" that triggers cell division upon reaching a threshold. We model growth inhibition as the binding of GAF to a "Volume" precursor, thereby inhibiting "Volume" accumulation. "Volume" levels divide with cell division, and newly created cells form adjacent to their parents on the 2D grid. If needed, neighboring cells are "pushed" to adjacent grid positions in order to make room for newly divided cells.

4.1 Gene network design

The overall designs of Systems 2 and 3 are generally similar to the Langevin implementations in network topology (Figure 3-4) but contain more details about the receptor and the UPC model (Figure S4).

We model repressors and activators binding to a gene's promoter, thus changing transcription and translation rates. Repressors such as TetR-KRAB fusion protein and activators such as VP16-AraC fusion protein fit this model and have been previously implemented in mammalian synthetic genetic networks [24]. Our model of engineered cell-cell communication in a mammalian system is based on bacterial two-gene QS systems, such as *rhlI/rhlR* in *Pseudomonas (P.) aeruginosa* [25] and *luxI/luxR* in *Vibrio (V.) fischeri* [26]. QS systems have previously been used in synthetic gene networks to engineer cell-cell communication in both bacteria [27] and mammalian cells [28]. Generally, one protein (LuxI or RhlI, modeled generically as I1 and I2) catalyzes synthesis of a freely diffusing small molecule, known as an autoinducer (AI), specific to a particular QS system (modeled here as $AI1$ and $AI2$). A receptor protein (LuxR or RhlR, modeled generically as $Rec1$ and $Rec2$) binds the appropriate intracellular autoinducer. The resulting complex acts as an activating transcription factor similar to activators described above, and is engineered to reach high concentration when population density reaches a "QS threshold." Other artificial signaling pathways, for example those previously engineered in yeast [29], may also be implemented. As with the Langevin models, the proposed circuit design can be described in terms of six key modules: "Uncommitted Population Control" (UPC), the "Oscillator", the "Committed Population," the "AND" gate module, the "Toggle Switch" module, and the "Differentiation" module (see Figure 4A). Model reactions and rates are listed in Table S3.

The following section discusses the genes that comprise these modules and qualitatively discusses system dynamics. Each of the fourteen individual genes comprising the genetic modules belong to one of five categories: quorum sensing (QS) genes, repressors, activators, growth arrest factor, and cell-fate regulators. As discussed in the main part of the paper, repressors and activators are referred to generically as (Ro, $Ro2$,...) and (Ao, $A3$,...). QS molecular species adapted to a mammalian host from gram-negative bacterial communication are referred to as LuxR homologues ($Rec1$ and $Rec2$), LuxI homologues ($I1$ and $I2$), and cognate autoinducers ($AI1$ and $AI2$).

4.1.1 The UPC Module

In the UPC module, population control of uncommitted cells is controlled through a cell-cell communication system involving two proteins, $I1$ and $Rec1$. I1 catalyzes synthesis of $AI1$, which diffuses freely across cell membranes and acts as an intercellular biochemical signal. $AI1$ binds $Rec1$, a receptor protein. The resultant complex is an activating transcription factor ($Rec1.AI1$) that can bind specifically engineered promoters of various genes [30]. In our system, $Rec1.AI1$ binds the promoter of $I1$, causing activated transcription of that gene. The positive feedback in $I1$ expression produces nonlinear response to increasing $AI1$, ultimately leading to a step-like function in $Rec1/I1$ expression as population density increases. Within the UPC module, $Rec1.AI1$ binds the promoter for growth arrest factor (GAF) and activates its transcription when uncommitted population density is above the QS threshold. Once cells commit to differentiation via switching of the toggle switch (discussed later), expression of $I1$ is inhibited by $R7$. This repression allows $I1$ expression, and consequently $AI1$ concentration, to be reflective only of uncommitted population density.

4.1.2 The Committed Population

With the switch to commitment and to high $R7$ expression, gene $I2$ is no longer repressed by $R6$. $I2$ catalyzes synthesis of $AI2$, which binds $Rec2$ to form the complex $Rec2.AI2$. $Rec2.AI2$ activates both $I2$ and $Rec2$ expression through a double-positive feedback mechanism. As with the $Rec1/I1$ QS system used in the population control module, positive feedback creates a sharp gain in $AI2$ production as the population density of committed cells increases.

4.1.3 Symmetry Breaking Oscillator

To break symmetry between individual cells in the population, oscillations are introduced through a relaxation oscillator with a design similar to previous experimental studies in prokaryotic and eukaryotic cells [16, 17, 18, 19, 20, 21]. Ao activates expression of itself and two repressors, Ro and $Ro2$. Ro in turn represses Ao expression. Ro and $Ro2$ are activated as Ao levels increase. Ro subsequently represses further expression of Ao and, as Ao levels decrease, Ro expression is inactivated. As Ro levels fall to a sufficiently low level (due to inactive transcription), Ao again is freely expressed. Ro is modeled as a slowly degrading protein to maximize oscillation period, and $Ro2$ is designed as a relatively fast decaying protein to minimize periods of high $Ro2$.

4.1.4 The AND Gate

The AND gate module regulates the commitment of individual cells to differentiate. $Rec1.AI1$ interfaces with the commitment module through binding the $A3$ promoter and activating $A3$ expression. $A3$ expression is high only when the uncommitted population is high ($Rec1.AI1$ is bound to the $A3$ promoter at high levels) and the oscillator peaks (Ao is high). $Ro2$ output from the oscillator (high when Ao is high) represses $R4$, which in turn represses $A3$. Even when bound by $Rec1.AI1$, $A3$ expression is repressed when $R4$ is bound. $A3$ activates expression of the repressor $R5$. $R5$ subsequently interacts directly with the toggle switch. The "committed population" module interfaces with the AND module through $Rec2.AI2$ activation of $Ro2$. $Ro2$ represses $R5$ expression regardless of whether $A3$ is bound to the $R5$ promoter. $Ro2$ is high when the population of committed cells is high, thereby repressing further commitment. $R5$ expression is activated only when (1) uncommitted population density is high, (2) the oscillator is high, and (3) committed population density is low.

4.1.5 The Toggle Switch

The toggle switch defines whether an individual cell is uncommitted or committed. The bi-stable toggle switch consists of two mutually inhibitory repressors: $R6$ and $R7$. High expression of one repressor inhibits expression of the other. The switch is initially set to high $R6$ levels in uncommitted cells. This can be accomplished in engineered systems through the introduction of an inducer that inactivates $R7$. For example, if $R7$ was LacI, transient addition of IPTG would stably set the toggle switch to high $R6$. When the AND module output is high, $R5$ represses $R6$ expression, thus allowing $R7$ to be expressed. $R7$ then represses $R6$, stabilizing the toggle switch in the high $R7$ state. To increase the bi-stability of the switch, multiple repressor binding sites are incorporated into each of the promoters of the switch. The transition from high $R6$ to high $R7$ expression in the toggle switch signifies commitment to differentiation. $R7$ represses $I1$ expression in the UPC module, causing repression of $AI1$ expression.

4.1.6 The Differentiation Module

With the switch to commitment, subsequently low $R6$ levels allow differentiation to occur through unrepressed expression of cell fate regulators. For differentiation into pancreatic β-cells, for instance, the unrepressed cell fate regulator could be an endoderm-directing cell fate regulator such as Gata4. In order to describe Gata4 as inducing differentiation into endoderm cells, cells are modeled as having a one way toggle switch involving the factors that sustain either the undifferentiated or differentiated state. As a cell-fate regulator, Gata4 is modeled as inducing the switch from the undifferentiated to differentiated state. Because differentiated cells exhibit significantly lower growth rates and higher death rates as compared to stem cells, differentiated cells are modeled as accumulating killer protein ("E") and being unable to grow. When killer protein levels reach a given threshold, the cell dies.

Differentiation may be successfully directed if guided by stepwise expression of various cell-fate regulators at critical points in the differentiation pathway. Promoters have been identified which are only active in certain cell types. For example, the Alpha-FetoProtein promoter (pAFP) is only active in cells that have differentiated into endoderm cells. Such promoters could be used for the sequential expression of key transcription factors at specific points along the differentiation pathway. As described in this circuit, the pancreas specific cell-fate regulator genes *pdx1* and *ngn3* are fused downstream of pAFP. Thus once the cell reaches the endoderm stage, *pdx1* and *ngn3* expression is activated, and differentiation is further directed into a β-cell fate.

5 Methods for results analyses of Systems 2-4

5.1 RS-HDMR sensitivity analysis

RS-HDMR is a tool to deduce non-linear interactions between a set of inputs and an output [31]. In this work, we use RS-HDMR in multiple distinct applications. RS-HDMR describes the independent and cooperative effects of n parameters $\mathbf{x} = (x_1, x_2, ...x_n)$ on an output, $y = f(\mathbf{x})$, in terms of a hierarchy of RS-HDMR component functions:

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \le i < j \le n} f_{ij}(x_i, x_j) + ... + f_{12...n}(x_1, x_2, ...x_n) \tag{S45}$$

Here f_0 represents the mean value of $f(\mathbf{x})$ over the sample space, the first-order component function $f_i(x_i)$ describes the generally non-linear independent contribution of the i^{th} input variable to the output, the second-order component function $f_{ij}(x_i, x_j)$ describes the pairwise cooperative contribution of x_i and x_j, and further terms describe higher order cooperative contributions. In this work, we generally consider first-, second-, and third-order RS-HDMR component functions. We approximate RS-HDMR component functions as weighted orthonormal basis functions, which take the following form:

$$f_i(x_i) \approx \sum_{r=1}^{k} \alpha_r^i \varphi_r^i(x_i) \tag{S46}$$

where k is an integer (generally ≤ 3 for most applications), $\{\alpha\}$ are constant weighting coefficients to be determined, and the basis functions $\{\varphi\}$ are optimized from the distribution of sample data points to follow conditions of orthogonality [31]. Basis functions are approximated here as non-linear polynomials, where

$$\varphi_1^i(x_i) = a_1 x_i + a_0 \qquad \varphi_2^i(x_i) = b_2 x_i^2 + b_1 x_i + b_0 \qquad \varphi_3^i(x_i) = c_3 x_i^3 + c_2 x_i^2 + c_1 x_i + c_0 \tag{S47}$$

The coefficients a_0,a_1,b_0,...c_3 are calculated using Monte Carlo integration under constraints of orthogonality, such that when integrated over all data points,

$$\int \varphi_r(x)dx \approx 0 \ \ \forall\, r \qquad \int \varphi_r^2(x)dx \approx 1 \ \ \forall\, r \qquad \int \varphi_p(x)\varphi_q(x)dx \approx 0 \ \ (p \ne q) \tag{S48}$$

Optimal basis functions are weighted by coefficients (α_r^i), which are calculated from least-squares regression. Only inputs and their respective component functions measured as significant by the statistical F-test were included in RS-HDMR expansions [32]. The resultant expansion in Eq. S45 serves both as a predictive model of network response due to its parametric interactions and as a statistical representation of the underlying system.

The relative strength of response to parametric changes can be quantitatively determined through sensitivity analysis based on the respective RS-HDMR component functions. A global sensitivity analysis may be calculated from the RS-HDMR expansion through a decomposition of the total variance σ^2 of an output species, $f(x)$, into hierarchical contributions from the individual RS-HDMR component functions. For each RS-HDMR expansion, the total sensitivity/variance σ of the output $f(\mathbf{x})$ is decomposed into hierarchical contributions $(\sigma_i, \sigma_{i,j}, \ldots)$ from the individual RS-HDMR component functions of the remaining input variables:

$$1 = \sum_{i=1}^{n} S_i + \sum_{1 \le i < j \le n} S_{ij} + ... + S_{\epsilon} \tag{S49}$$

In Eq. (51), $S_i = \sigma_i^2/\sigma^2$ is defined as the sensitivity index of the corresponding RS-HDMR component function, $f_i(x_i)$. $S_{ij} = \sigma_{i,j}^2/\sigma^2$ is the sensitivity index of the corresponding second-order component function, $f_{ij}(x_i, x_j)$. S_ϵ is the sensitivity index of the residual variation of the model. The collection of sensitivity indices S_i,

$\sum_{j \neq i}^{n} S_{ij}$, $\sum_{j \neq i}^{n} \sum_{k \neq i,j}^{n} S_{ijk}$ corresponding to first, second, and third order component functions of the input variable x_i can then be summed into an index $S_i^T (i = 1, 2...n)$, describing both independent and higher-order effects of x_i on an output. The magnitudes of $S_i^T (i = 1, 2, ..., n)$ can be used to quantify the relative interaction strength between the outputs and the inputs.

5.2 Langevin model analysis

5.2.1 Time-scale optimization

Time-scale optimization involves multiplying rate constants by a scalar TS parameter while preserving the ratio of closely related rates (see Table S1). We scale the rate k_p (production rate), k_d (degradation rate) and k_{diff} (diffusion rate) of all components in the quorum signaling module by a factor $TS_{QS} \in [\frac{1}{3}, 3]$. We perform analogous scaling for the quorum sensing module ($TS_{QM} \in [\frac{1}{3}, 3]$). For the commitment module, we independently analyze three individual components. $TS_{R5} \in [\frac{1}{3}, 3]$ scales k_p^{R5} and k_d^{R5}. $TS_{R6} \in [\frac{1}{3}, 3]$ and $TS_{R7} \in [\frac{1}{3}, 3]$ denote analogously lumped parameters. For System 4, we include an additional parameter for the dynamics of the component At ($TS_{At} \in [\frac{1}{3}, 3]$). For Systems 2 & 3, we analyze five total time-scale parameters. We analyze six time-scale parameters in System 4 (see Table S2). We randomly sample the time-scale space from a log-uniform distribution over one order of magnitude for each parameter, ultimately generating 360 independent parameter sets (Figure S5). For each time-scale set, we average the observed S/N over eight simulations. We perform RS-HDMR analysis to map the input-output relationships between the time-scale parameters (inputs) and their corresponding S/N, the output (Figure 5C-F). We excluded 30% of the dataset to cross-validate RS-HDMR inference and used the remaining 70% of the data as the training set. RS-HDMR inference results indicated consistent fitting accuracy between the test and training sets.

5.2.2 Robustness to variations of molecular noise amplitude

To test the influence of stochastic fluctuations in the models, we changed the volume from 40 to 1400 as shown in Figures 5A and S6. For each value of Ω, 24 simulations are performed and the average S/N value is recorded. Error bars in these plots indicate standard deviation of the results for the 24 simulations.

5.2.3 Robustness to variation of the killing rate k_k

To test the influence of the killing rate in the models, we change the value of k_k from 45 to 550 hours (corresponding to a ratio k_b/k_k of 0.5 to 6) as shown in Figures 5B and S6. For each value of k_k, 24 simulations are performed and we record the average S/N value and the average committed population density. Error bars in these plots indicate standard deviation of the results for the 24 simulations.

5.2.4 Module optimization

We randomly and independently modified twelve parameters involving the oscillator in System 3 (Figure 7A), and nine parameters involving the throttle in System 4 (Figure 7G), to understand the impact of parameter variation on module properties and, ultimately overall S/N. We randomly sampled the parameter space from a log-uniform distribution, one order of magnitude around the nominal values for each parameter (see Table S1). We generated 2000 independent parameter sets for the oscillator and roughly 6000 parameter sets for the throttle. For each parameter set, we simulated system behavior with the Langevin models and recorded the observed S/N value (average of 16 independent simulations). We also performed simulations where the module is isolated from the system and focus on different phenotypes of the module. For System 3, we isolate the oscillatory module (Figure 7A) and record the properties for the $R4$ component (output of the module) as described in Figure 8A and Table S4. For System 4, we isolate the throttle (Figure 7G) by analyzing the toggle response ($AI3$ and $R7$ concentrations) to independently (exogenously) modulated inputs of $A3$ (signal from the QS modules) and $AI3$ (signal from adjacent cells), as shown in Figure 8F. For the isolated throttle, produced

$AI3$ is kept separate from the input of $AI3$ and therefore does not act on $R5$. We focus on changes in $R7$ and $AI3$ concentrations in response to exogenously controlled combinations of $A3$ and $AI3$ as inputs (see Table S5). For the oscillator, each tested parameter set yields a single value for each phenotype. In contrast, we test the isolated throttle with combinations of different $A3$ and $AI3$ inputs. Therefore, for each throttle parameter set, we obtain a 2-dimensional grid for each phenotype with 11 sampled values of $A3$ $(0.07, 0.16, \dots, 0.97)$ and 18 sampled values for $AI3$ $(0.06, 0.12, \dots, 1.08)$. These input values were chosen based on $A3$ and $AI3$ concentrations observed in the full system simulations.

We performed several types of analysis on both Systems 3 and 4 to understand the relationships between the module rate constants, the module phenotypes, and corresponding S/N of the full system (Figures 7, 8 and S7-S14). We perform RS-HDMR analysis using either rate constants, module phenotypes, or both as inputs to describe the system output, S/N. Because throttle phenotypes are described as a function of two inputs ($A3$ and $AI3$), we describe the two-input functions ("images") of each phenotype (see Figures S13-S14) as a set of features (see Table S6). For features extracted using `regionprops()`, images were thresholded at various levels, digitized accordingly, and then analyzed. `regionprops()` calculates properties such as "Centroid," and "Filled Area" for the region of an the image falling above a particular threshold; we used both relative thresholds, such as $> 90\%$ max value, and absolute thresholds for this analysis. For features extracted using `graycoprops`, images were first converted to gray-level co-occurrence matrices. Overall, we extracted roughly 10,000 features from the original 16 phenotype images. We performed partial least squares regression (PLSR, `plsregress()`, Matlab, Natick, MA) using the 10,000 features as inputs, and ranked features by their variable importance in the projection (VIP) [33]. The 20 most significant variables were then analyzed by RS-HDMR. Table S4 shows the most significant features identified by RS-HDMR. Even with the reduced number of features, RS-HDMR inference performs with roughly equal or better accuracy compared to partial least squares regression in all cases tested.

We used Bayesian network analysis to produce graphical representations of the conditional probabilistic dependencies of the module rate-constants and phenotypes on each other and on overall S/N. Directed graph structures produced by Bayesian network inference consist of "nodes", which in this application are the module phenotypes and S/N, and "edges", which represent conditional probabilistic relationships between the nodes. Bayesian network inference was performed as described previously [34]. Briefly, we derived consensus directed graph structures from exact Bayesian network model averaging over all directed acyclic graph (DAG) structures having at most four parents per node [35, 36]. Consensus networks for Systems 3 and 4 (Figure 9) only include those edges with a score > 0.8 and > 0.3, respectively, where an edge score of 0.8, for example, denotes that 80% of the derived Bayesian networks tested over all iterations of optimization include that edge. Although our implementation of Bayesian inference cannot capture cyclical feedback, bi-directional edges can be observed with model averaging when using a significance threshold below 0.5. The bi-directional edges reported in System 4 (Figure 9B) arise from observation of both edge directions above the 0.3 threshold. For System 3, we ranked phenotypes and rate-constants by their VIP score after performing PLSR as referred to above, and analyzed the top 15 most significant rate-constants and phenotypes using Bayesian inference. For System 4, we analyzed the 9 most significant module phenotypes (as determined by their VIP), along with all 9 module rate-constants. For both systems, we only considered DAGs with directed edges from module phenotypes to the "S/N" node (i.e., we defined module phenotypes as strictly upstream of S/N), and we specified rate-constants as strictly upstream of both the module phenotypes and S/N. From these constraints, we consequently inferred directionality from the remaining edges. We iteratively removed nodes that were not upstream of S/N in the consensus network, and recalculated the consensus network structure. As a result, the final network structures are comprised of S/N and only those nodes upstream of S/N at the chosen significance threshold. Prior to network inference, nodes were discretized using two, three, and four-level k-means clustering with squared Euclidean distance metric and 50 iterations. Bayesian inference results across the three types of k-means clustering were then averaged. Software used for Bayesian network inference has been previously described [35].

5.3 Analysis of a two-compartment ODE model of the UPC module

We implement the full model (Systems 2 and 3) as a system of discrete stochastic reactions using an established multicellular, spatiotemporal simulation platform [22, 23] (see section 4). Because computational costs of running this platform are high, we can turn to a simpler simulation framework for high-throughput sampling and optimization of rate-constants within individual subnetworks of the full system model. We employ a deterministic, two-compartment ordinary differential equation (ODE) model of the UPC module for efficient genetic algorithm (GA) optimization and sensitivity analysis of a portion of the full system stochastic model. (See the following page for specific equations and a table of rate constants.) Intracellular reactions of the ODE system use identical rate constants to those in the stochastic model, with the standard correction for dimerization reactions [37]. We approximate individual cells within the population as identical to reduce the complexity of the ODE system. We model diffusion as occurring between a variable population density and a homogeneous extracellular volume, thus neglecting the effect of spatial signaling gradients. Using k_{diff} as the rate of diffusion across a cell membrane, k_c as the rate of AI decay, and ρ_p as the population density of cells in the culture, concentration of AI outside of the cell ($[AI_{out}]$) is described by the following equation:

$$\frac{d\left[AI_{out}\right]}{dt} = \rho_p k_{diff}\left(\left[AI_{in}\right] - \left[AI_{out}\right]\right) - k_c\left[AI_{out}\right] \tag{S50}$$

where $[AI_{in}]$ represents the concentration of AI within the cells. As with the stochastic simulator, cell density and concentration of species within cells are approximated as uniform. In the ODE model, we also approximate the concentration of extracellular AI as homogeneous. We describe diffusion in and out of individual cells as independent of population density (at least explicitly, by exclusion of ρ_p), such that

$$\frac{d\left[AI_{in}\right]}{dt} = \sum_i f_{in}(\mathbf{x}) + k_{diff}\left(\left[AI_{out}\right] - \left[AI_{in}\right]\right) - k_c\left[AI_{in}\right] \tag{S51}$$

where $\sum_i f_{in}(\mathbf{x})$ represents the sum of all $i = 12$ intracellular reactions described on the following page and in Table S6.

$$
\begin{aligned}
\frac{d\,[pI1]}{dt} &= k_2\,[pI1.Rec1.AI1] + k_3\,[pI1.Rec1.AI1] - k_1\,[pI1]\,[Rec1.AI1] \\
\frac{d\,[pRec1]}{dt} &= 0.0 \\
\frac{d\,[LuxI]}{dt} &= k_0\,[pI1] + k_4\,[pI1.Rec1.AI1] - k_{10}\,[LuxI] \\
\frac{d\,[Rec1]}{dt} &= k_9\,[pRec1] - k_{11}\,[Rec1] - k_7\,[Rec1]\,[AI1_{in}] + k_8\,[Rec1.AI1] \\
\frac{d\,[AI1_{in}]}{dt} &= -k_7\,[Rec1]\,[AI1_{in}] + k_8\,[Rec1.AI1] + k_5\,[LuxI] \\
& \quad -k_6\,[AI1_{in}] + k_{17}([AI1_{out}] - [AI1_{in}]) \\
\frac{d\,[Rec1.AI1]}{dt} &= k_7\,[Rec1]\,[AI1_{in}] - k_8\,[Rec1.AI1] - k_3\,[Rec1.AI1] \\
& \quad +k_2\,[pI1.Rec1.AI1] - k_1\,[pI1]\,[Rec1.AI1] \\
& \quad -k_{15}\,[pA2]\,[Rec1.AI1] + k_{16}\,[pA2.Rec1.AI1] \\
\frac{d\,[pI1.Rec1.AI1]}{dt} &= -k_2\,[pI1.Rec1.AI1] + k_1\,[pI1]\,[Rec1.AI1] - k_3\,[pI1.Rec1.AI1] \\
\frac{d\,[pA2]}{dt} &= -k_{15}\,[pA2]\,[Rec1.AI1] + k_{16}\,[pA2.Rec1.AI1] + k_3\,[pA2.Rec1.AI1] \\
\frac{d\,[pA2.Rec1.AI1]}{dt} &= k_{15}\,[pA2]\,[Rec1.AI1] - k_{16}\,[pA2.Rec1.AI1] - k_3\,[pA2.Rec1.AI1] \\
\frac{d\,[A2]}{dt} &= k_{12}\,[pA2] + k_{13}\,[pA2.Rec1.AI1] - k_{14}\,[A2] \\
\frac{d\rho_p}{dt} &= k_{34} \\
\frac{d\,[AI1_{out}]}{dt} &= k_{17}\rho([AI1_{in}] - [AI1_{out}]) - k_6\,[AI1_{out}]
\end{aligned}
$$

5.3.1 Genetic algorithm

For the GA optimization, parameter vectors in initial generations consist of random points within biologically reasonable ranges of parameter space. To calculate the “forward QS response” in the ODE model, initial $\rho_p = 0$ and ρ_p increases at rate k_{34} that is much slower than other reactions in the system, such that module output, defined as $[pA2.Rec1.AI1]$, maintains quasi-steady-state. The “reverse QS response” in the ODE model is similarly calculated, but using the initial value $\rho_p = 1$, and having the population decrease at rate $d\rho_p/dt = -k_{34}$. The cost function for the GA defines the forward and reverse responses of $[pA2.Rec1.AI1]$ as a function of ρ_p to be a least-squares fit to a three component step function, described in the main text. Over 1000 generations (100 individual parameter vectors to a generation), system behavior evolves from a relatively flat response to a more optimal digital-like step function response.

5.3.2 RS-HDMR analysis of hysteresis

We use RS-HDMR to understand the impact of parametric variation on system performance in the two-compartment model of the UPC module. Input-output relationships are defined as the effects of parametric variation on hysteresis of the UPC module response to fluctuations in population density. More specifically, we define hysteresis as the difference between the forward and reverse response values of population density (ρ_p), where the UPC module’s output ($[pA2.Rec1.AI1]$) is 50% of maximum, or (max output + min output)/2. We focus on absolute levels of hysteresis rather than normalizing to the average population density threshold, because we focus on systems with similar average thresholds and consider absolute changes in population density to be a relevant optimization feature of our system in the context of its biomedical application.

We performed RS-HDMR sensitivity analysis on datasets describing neighborhoods of parameter space around optimal parameter vectors obtained from GA runs. We generated two sets of 75 GA optimizations: the first set considered both the forward and reverse QS responses to changing population density, and the second set only considered forward response. Random sampling around each optimal parameter vector was from a normal distribution $N(\mu, \sigma)$ where μ is the optimized parameter's value and $\sigma = \mu/20$. Empirical evidence suggested that significantly broader sampling resulted in too many parameter sets that did not yield QS behavior. Sample size of the training set was 2000, and the resultant model was tested on unsampled points for validation purposes. In this application, we only considered the first-order RS-HDMR component functions in order to perform efficient high-throughput analyses of local parameter "neighborhoods."

5.4 Patterning and Neighbor Density Analysis

We analyzed distances between pairs of committed and uncommitted cells to identify patterning between the two cell-types. For a given committed or uncommitted reference cell, the ratio of committed to uncommitted neighbors at a given distance was calculated for all distances. We define $p(c)_{i,d,t}$ as the observed fraction of committed cell neighbors at distance d for the i^{th} cell at time t. The normalized score $Z(c)_{i,d,t}$ for that observation is then described by

$$Z(c)_{i,d,t} = \frac{p(c)_{i,d,t} - \mu_t}{\sigma_{i,t}} \tag{S52}$$

where μ_t represents the overall fraction of committed cells at time t and $\sigma_{i,t}$ describes the standard deviation of the observed probability given by the standard form:

$$\sigma_{i,t} = \sqrt{\frac{\mu_t(1-\mu_t)}{n_{i,t}}} \tag{S53}$$

where $n_{i,t}$ is the number of total neighbors observed for the i^{th} cell at time t. Normalized Z-scores are combined into an average Z-score, $\overline{Z(c)_d}$, for each distance value,

$$\overline{Z(c)_d} = \sum_i \sum_t Z(c)_{i,d,t}/\sqrt{N_d} \tag{S54}$$

where N_d is the total number of sample Z-scores $Z(c)_{i,d,t}$ for each distance. In Figures S18 and S19, the "Z-score" metric refers to $\overline{Z(c)_d}$.

6 Additional results

6.1 Detailed results for time-scale analysis

With the results of the time-scale analysis (see the section *Intermodular time-scale* in main text), we selected individually optimized values for Systems 2–4. The optimized parameter values below are an average of the 10% top performing parameter sets tested.

	TS_{QS}	TS_{QM}	TS_{R5}	TS_{R6}	TS_{R7}	TS_{At}
System 2	2.31	1.14	1.29	1.67	1.24	–
System 3	2.14	0.91	1.08	1.84	1.11	–
System 4	1.95	0.88	1.38	1.54	0.72	1.01

Such parameter optimization increases system performance and robustness to variations in the killing rate k_k and the cell volume Ω (Figure S6). On average, the optimization affords a gain of 5 units for the S/N value in all systems for all conditions. The major qualitative difference is the ability of the time-scale optimized System 4 to function even with low molecular noise ($\Omega > 1000$) in a comparable way to System 3. Results can be compared to Figure 5A-B in the main text.

6.2 Detailed results for the oscillator and throttle module analyses

We employ RS-HDMR to predictively model the relationship between module rate-constants, phenotypes, and overall system performance (see Figures 7 and 8). RS-HDMR infers and predicts precise S/N values with little accuracy when only the module rate-constants are employed as predictive variables (Figure S10A,C). $R^2 \approx 0.5$ for the analysis of both Systems 3 (Figure 8E) and 4 (Figure 8I). This relatively poor fit arises from a highly uneven distribution of observed S/N values (>50% of parameter sets have S/N<2) and what are likely to be significant higher (greater than third) order RS-HDMR component functions, which we do not account for in this application. Nonetheless, RS-HDMR can classify system performance as 'good' or 'bad' based on the rate-constants alone with an area under the ROC-curve (AUROC) of 0.97 (we define 'good' performers as parameter sets with observed S/N > 15 and 'bad' ones as parameter sets with S/N < 2). When module phenotypes are used to classify system performance (Figure S10B,D) in the same manner, RS-HDMR predicts with AUROC>0.98 for both Systems 3 (Figure 8E) and 4 (Figure 8I). Prediction with both the rate-constants and module phenotypes marginally improves this accuracy to an AUROC>0.99 for both systems. This trend in accuracy mirrors the R^2 values reported Figure 8E,I, where inference using module phenotypes generally out-performs inference using rate-constants alone. To compare RS-HDMR classification accuracy with another algorithm, we also implemented SVM classification using MATLAB (R2009a, The MathWorks, Natick, MA), with a two-norm soft-margin SVM classifier and linear kernel. For both Systems 3 and 4, RS-HDMR outperforms SVM in classification accuracy in the three scenarios (rate-constants only, phenotypes only, and rate-constants and phenotypes combined).

6.3 Population size when varying β-cell killing rate

In the previous section describing the ODE model (section 2), we found three important results concerning the sensitivity of the population to variations of the β-cell killing rate (k_k). First, the population of uncommitted cells is well controlled and remains constant even for low ratio of division rate over killing rate (k_b/k_k). Second, for high ratio, the population of committed cells follows a power law with an exponent $1/n$ where n is the Hill coefficient in the feedback function. Third, for low ratios, on the contrary, the population decreases linearly with the killing rate. It means that the population of committed cells follows a power law with exponent 1 when plotted against the ratio of division over killing rates.

The results of simulations with the ODE model with two populations (see section 2.5) are consistent with these theoretical results (Figure S17). For an ODE model with a Hill coefficient $n = 16$, the uncommitted population is very robust to variations of the killing rate. The committed population is also robust (exponent of 0.07, close the theoretical value $1/n = 0.0625$) for high ratio, but follows exactly a linear dependence (exponent of 1.00) for high killing rate.

But more interestingly, the results of the stochastic simulations with the Langevin model are qualitatively similar (Figure S17). If the three systems show small differences for low ratio k_b/k_k, the fits of System 2 – which is the closest to the ODE model and have power laws with exponents 0.06 and 1.14 for respectively high and low ratio. These values are very close to both the theoretical analysis and the ODE simulations of the simplified model.

References

[1] Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. Nat Biotechnol 26: 787–793.

[2] Shetty R, Endy D, Knight Jr T (2008) Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2: 5.

[3] Fujikura J, Yamato E, Yonemura S, Hosoda K, Masui S, et al. (2002) Differentiation of embryonic stem cells is induced by GATA factors. Genes Dev 16: 784-789.

[4] Schmitz ML, Baeuerle PA (1991) The p65 subunit is responsible for the strong transcription activating potential of nf-kappa b. EMBO J 10: 3805-3817.

[5] Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Eng 14: 529-532.

[6] Collins CH, Arnold FH, Leadbetter JR (2005) Directed evolution of vibrio fischeri luxr for increased sensitivity to a broad spectrum of acyl-homoserine lactones. Mol Microbiol 55: 712-723.

[7] Weiss R, Homsy GE, Knight Jr TF (1999). Toward in vivo digital circuits. DIMACS Workshop on Evolution as Computation.

[8] Gradner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in escherichia coli. Nature 403: 339-342.

[9] Coleman JE, Huentelman MJ, Kasparov S, Metcalfe BL, Paton JFR, et al. (2003) Efficient large-scale production and concentration of hiv-1-based lentiviral vectors for use in vivo. Physiol Genomics 12: 221-228.

[10] Miyoshi H, Blomer U, Takahashi M, Gage FH, Verma IM (1998) Development of a self-inactivating lentivirus vector. J Virol 72: 8150-8157.

[11] Smith H (1995) Monotone Dynamical Systems. Providence, RI: AMS.

[12] Angeli D, Sontag E (2004) Multistability in monotone input/output systems. Syst Control Lett 51: 185–202.

[13] Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58: 35-55.

[14] Burrage K, Burrage PM (1996) High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. Appl Numer Math 22: 81 - 101.

[15] Gillespie DT (2000) The chemical langevin equation. J Chem Phys 113: 297-306.

[16] Hasty J, Isaacs F, Dolnik M, McMillen D, Collins J (2001) Designer gene networks: Towards fundamental cellular control. Chaos 11: 207.

[17] Hasty J, McMillen D, Collins J (2002) Engineered gene circuits. Nature 420: 224–230.

[18] McMillen D, Kopell N, Hasty J, Collins J (2002) Synchronizing genetic relaxation oscillators by intercell signaling. Proc Natl Acad Sci USA 99: 679.

[19] Stricker J, Cookson S, Bennett M, Mather W, Tsimring L, et al. (2008) A fast, robust and tunable synthetic gene oscillator. Nature 456: 516–519.

[20] Swinburne I, Miguez D, Landgraf D, Silver P (2008) Intron length increases oscillatory periods of gene expression in animal cells. Genes Dev 22: 2342–2346.

[21] Tigges M, Marquez-Lago T, Stelling J, Fussenegger M (2009) A tunable synthetic mammalian oscillator. Nature 457: 309–312.

[22] Hsu A, Vijayan V, Fomundam L, Gerchman Y, Basu S, et al. (2005) Dynamic Control in a Coordinated Multi-Cellular Maze Solving System (I). In: American Control Conference. volume 6, pp. 4399-4404.

[23] Basu S, Mehreja R, Thiberge S, Chen M, Weiss R (2004) Spatiotemporal control of gene expression with pulse-generating networks. Proc Natl Acad Sci USA 101: 6355-6360.

[24] Greber D, Fussenegger M (2007) Mammalian synthetic biology: Engineering of sophisticated gene networks. J Biotechnol 130: 329-345.

[25] Brint J, Ohman D (1995) Synthesis of multiple exoproducts in Pseudomonas aeruginosa is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family. J Bacteriol 177: 7155.

[26] Engebrecht J, Nealson K, Silverman M (1983) Bacterial bioluminescence: isolation and genetic analysis of functions from Vibrio fischeri. Cell 32: 773–81.

[27] You L, Cox III R, Weiss R (2004) Programmed population control by cell- cell communication and regulated killing. Nature 428: 868–871.

[28] Weber W, Schoenmakers R, Spielmann M, El-Baba M, Folcher M, et al. (2003) Streptomyces-derived quorum-sensing systems engineered for adjustable transgene expression in mammalian cells and mice. Nucleic Acids Res 31: e71.

[29] Chen M, Weiss R (2005) Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. Nat Biotechnol 23: 1551–1555.

[30] Brenner K, Karig D, Weiss R, Arnold F (2007) Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. Proc Natl Acad Sci USA 104: 17300.

[31] Li G, Hu J, Wang SW, Georgopoulos PG, Schoendorf J, et al. (2006) Random sampling-high dimensional model representation (rs-hdmr) and orthogonality of its different order component functions. J Phys Chem A 110: 2474-2485.

[32] Chatterjee S, Price B (1977) Regression analysis by example. Wiley New York.

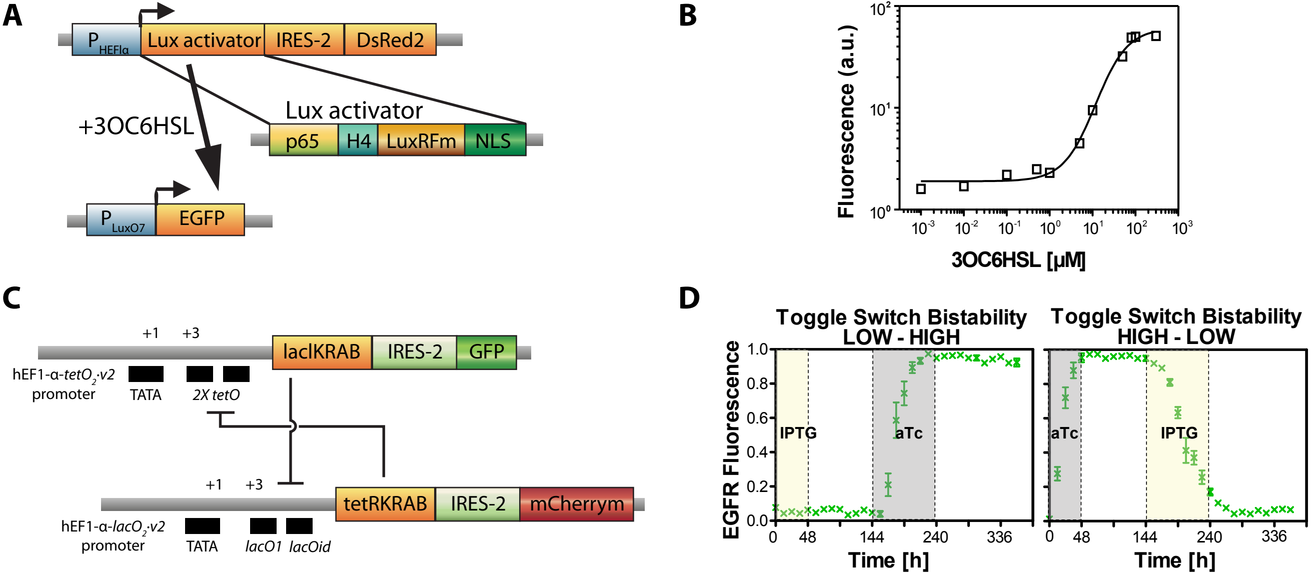
[33] Chong I, Jun C (2005) Performance of some variable selection methods when multicollinearity is present. Chemometr Intell Lab 78: 103–112.

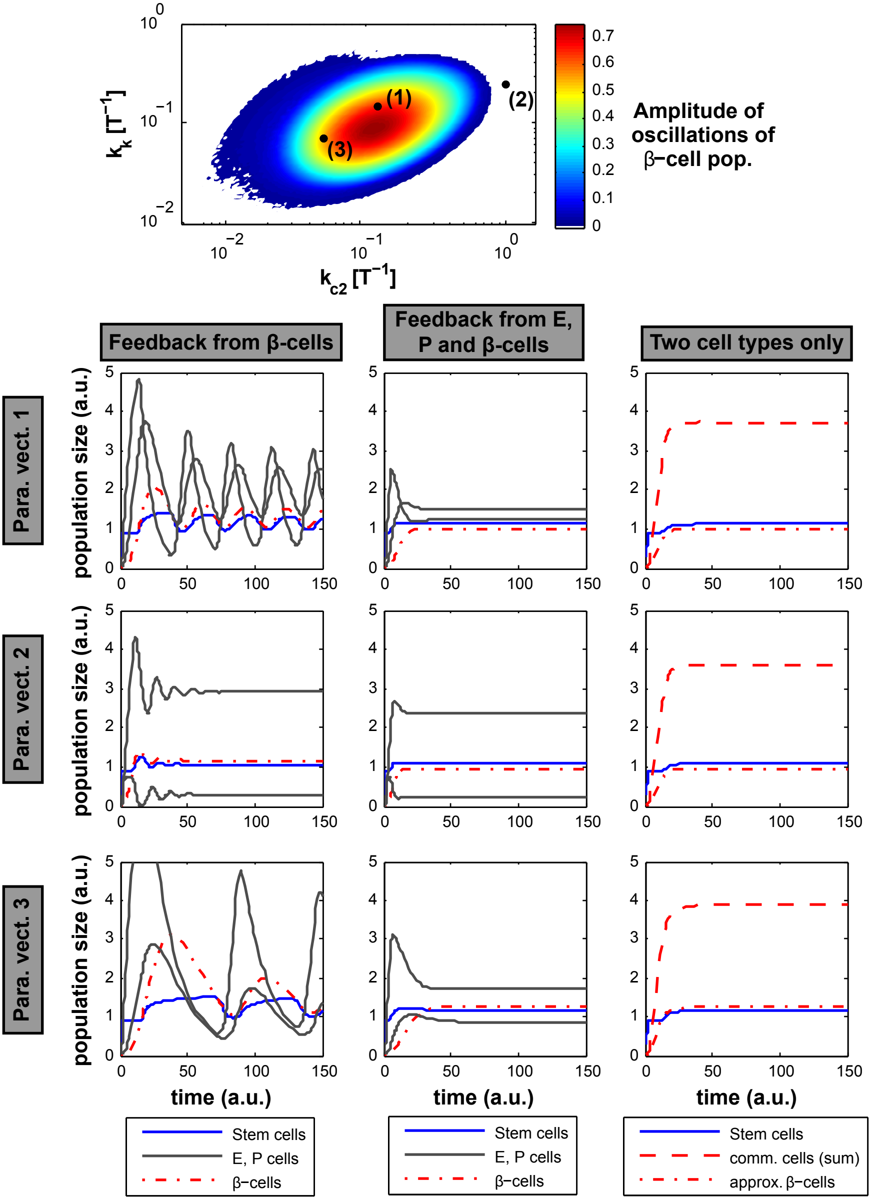
[34] Ciaccio M, Wagner J, Chuu C, Lauffenburger D, Jones R (2010) Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nat Methods 7: 148–155.

[35] Eaton D, Murphy K (2007) Exact Bayesian structure learning from uncertain interventions. In: AI & Statistics. Citeseer, volume 2, pp. 107–114.

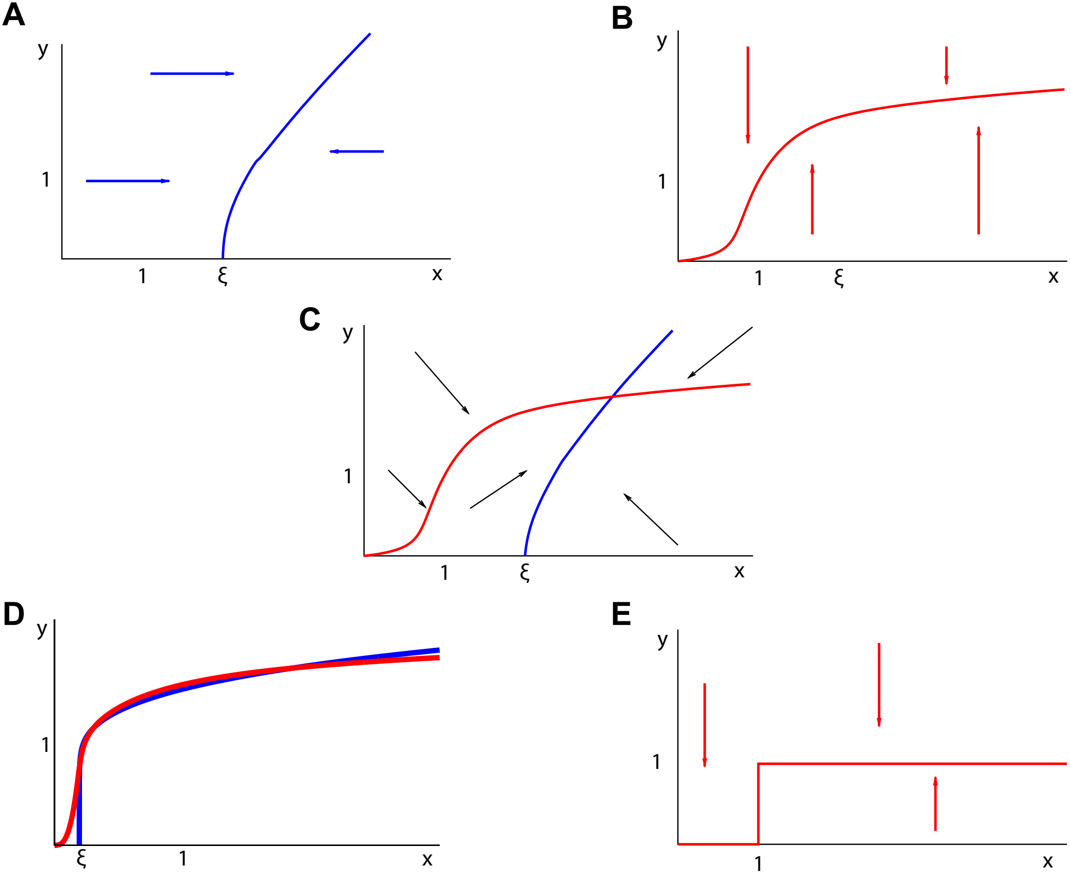
[36] Koivisto M, Sood K (2004) Exact Bayesian structure discovery in Bayesian networks. J Mach Learn Res 5: 573.

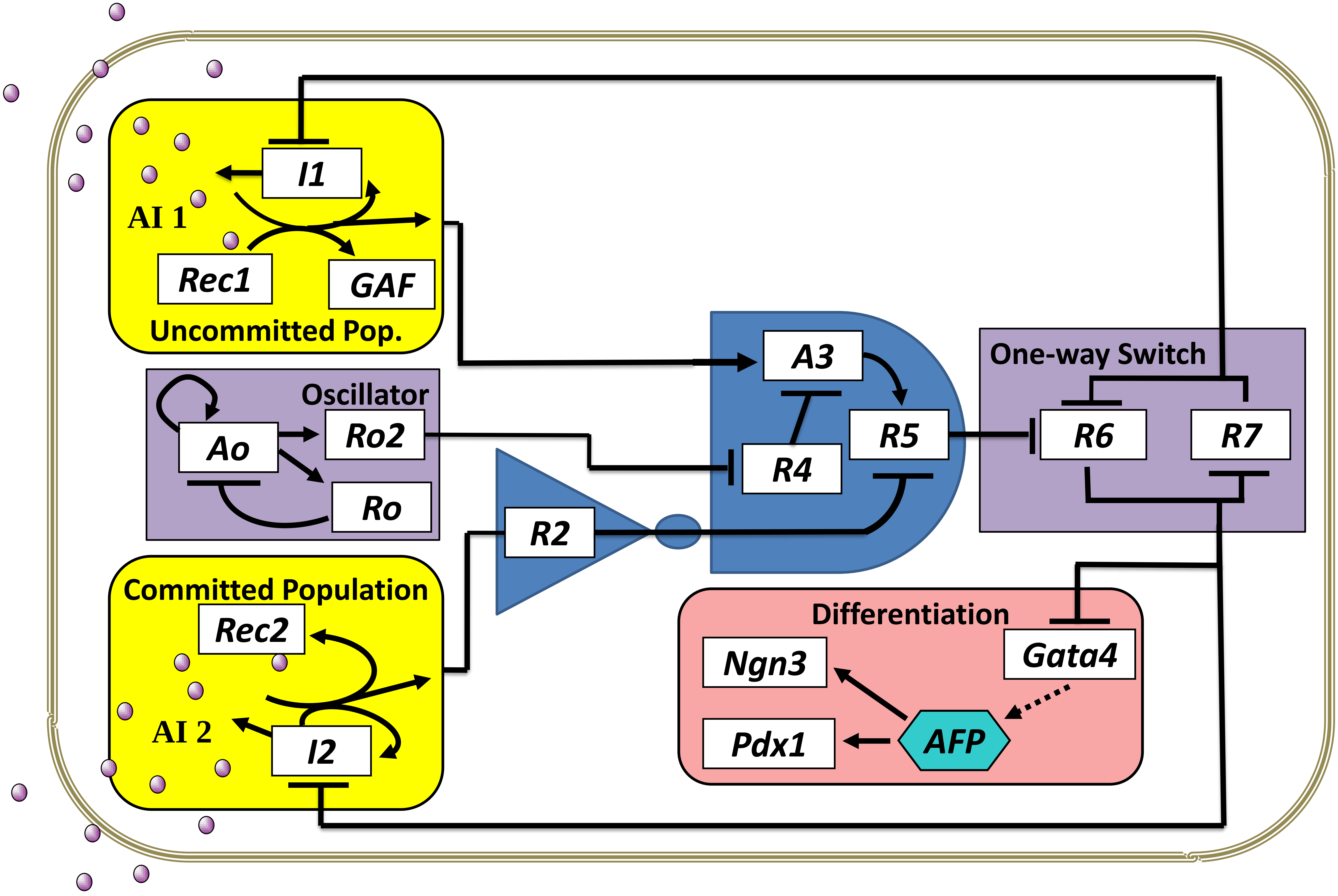
[37] Gillespie D (1977) Exact stochastic simulation of coupled chemical reactions. J Chem Phys 81: 2340–2361.

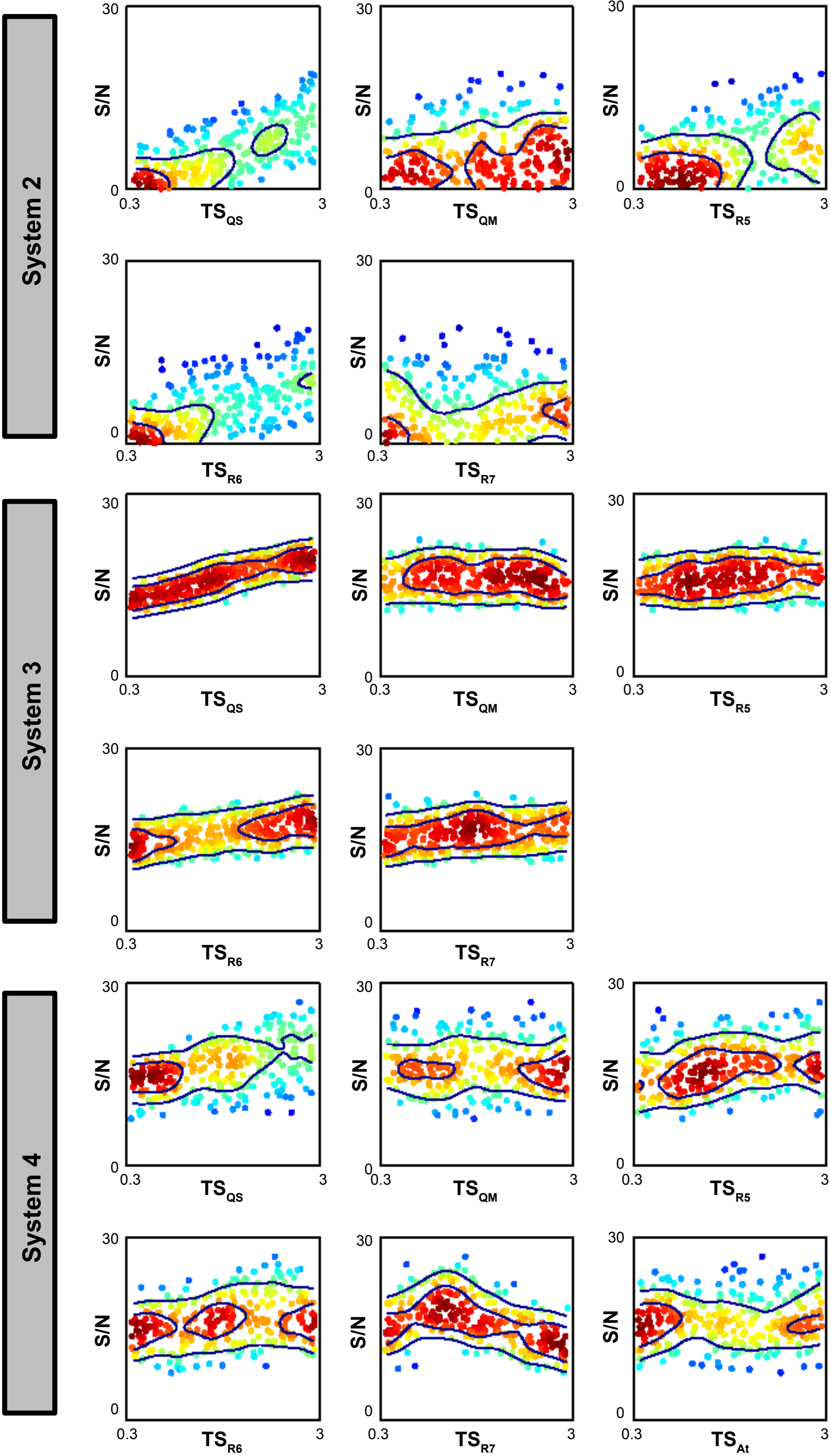

A
$P_{HEF1\alpha}$
Lux activator
IRES-2
DsRed2
+3OC6HSL
Lux activator
p65
H4
LuxRFm
NLS
P_{LuxO7}
EGFP
B
Fluorescence (a.u.)
10^2
10^1
10^0
10^{-3}
10^{-2}
10^{-1}
10^0
10^1
10^2
10^3
3OC6HSL [μM]
C
+1
+3
laclKRAB
IRES-2
GFP
hEF1-α-tetO$_2$·v2 promoter
TATA
2X tetO
+1
+3
tetRKRAB
IRES-2
mCherrym
hEF1-α-lacO$_2$·v2 promoter
TATA
lacO1
lacOid
D
Toggle Switch Bistability LOW - HIGH
Toggle Switch Bistability HIGH - LOW
EGFR Fluorescence
1.0
0.8
0.6
0.4
0.2
0.0
IPTG
aTc
aTc
IPTG
0
48
144
240
336
Time [h]
0
48
144
240
336
Time [h]

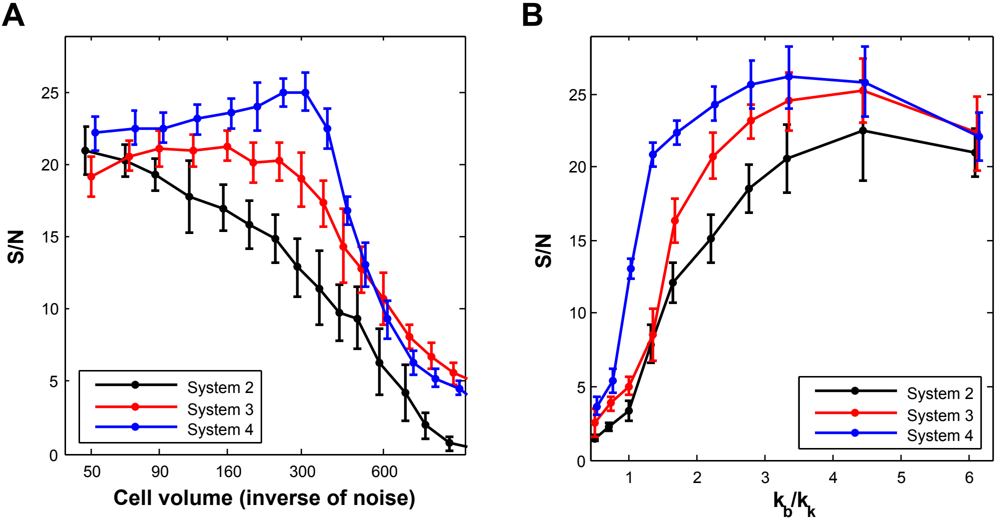


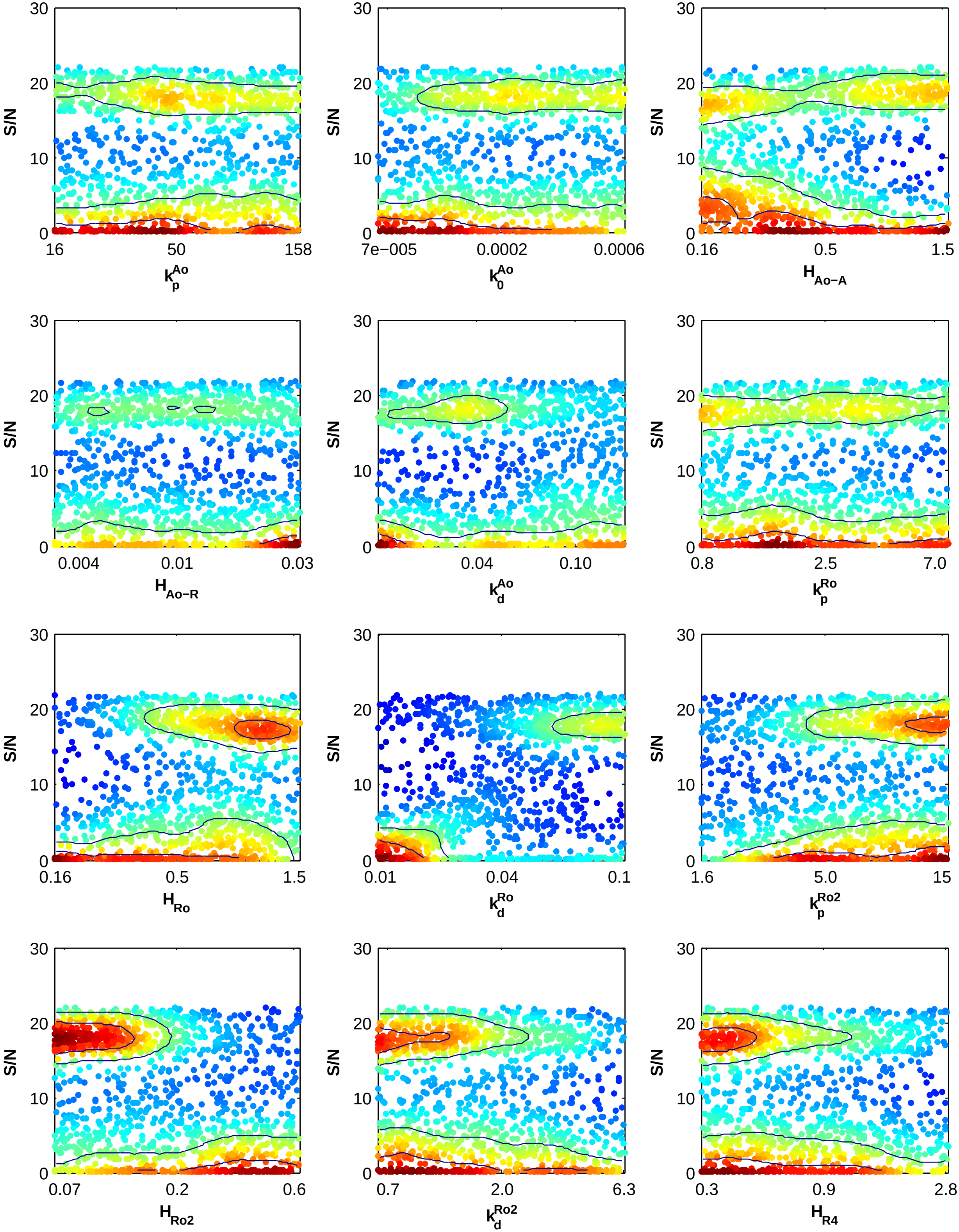
10^0
10^{-1}
10^{-2}
$k_k\ [T^{-1}]$
$k_{c2}\ [T^{-1}]$
(1)
(2)
(3)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Amplitude of oscillations of β−cell pop.
Feedback from β-cells
Feedback from E, P and β-cells
Two cell types only
Para. vect. 1
Para. vect. 2
Para. vect. 3
population size (a.u.)
time (a.u.)
0
1
2
3
4
5
0
50
100
150
Stem cells
E, P cells
β−cells
Stem cells
E, P cells
β−cells
Stem cells
comm. cells (sum)
approx. β−cells


A
y
1
1
ξ
x
B
y
1
1
ξ
x
C
y
1
1
ξ
x
D
y
1
ξ
1
x
E
y
1
1
x

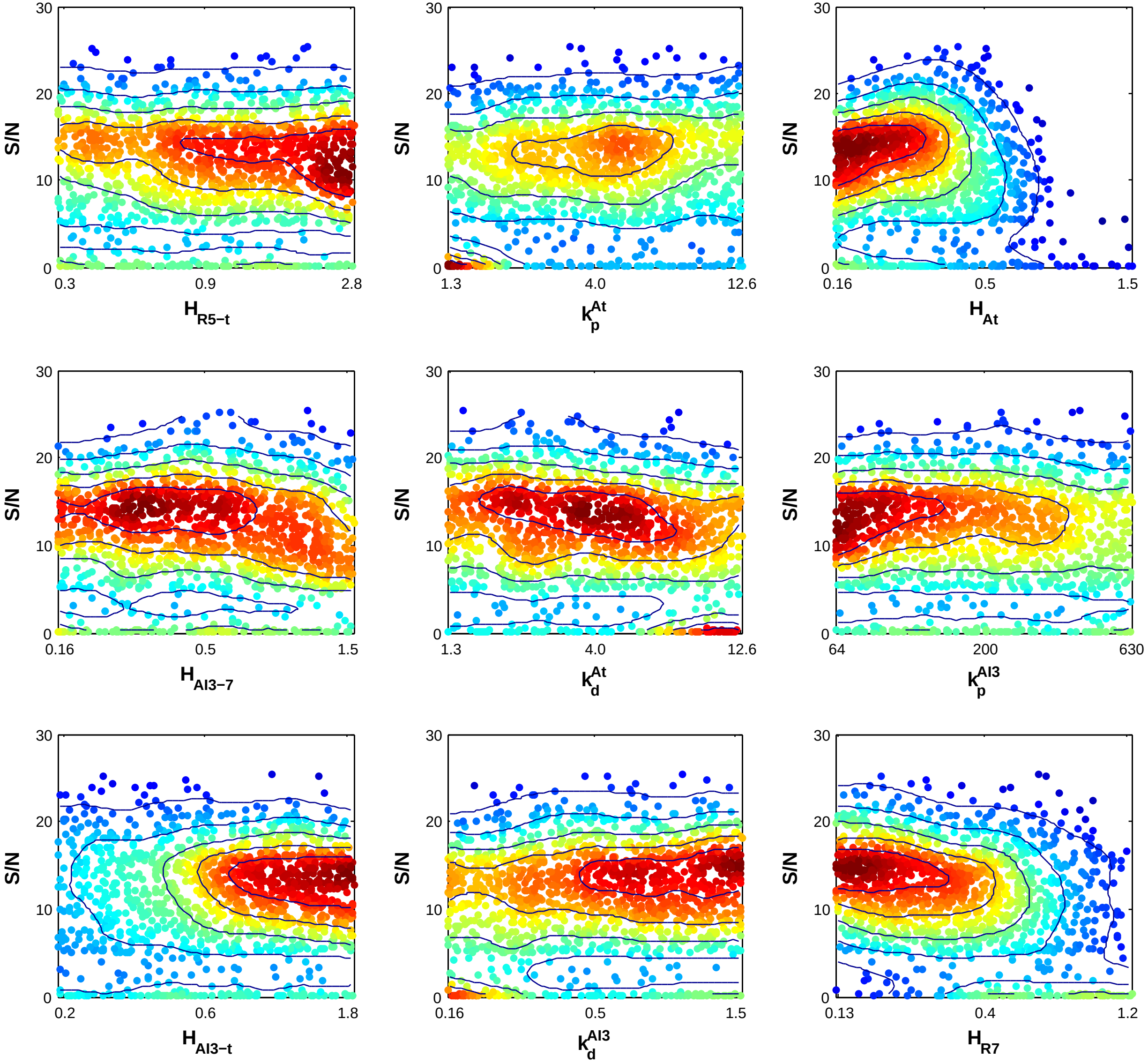

I1
AI 1
Rec1
GAF
Uncommitted Pop.
Oscillator
Ao
Ro2
Ro
A3
R4
R5
R2
One-way Switch
R6
R7
Committed Population
Rec2
AI 2
I2
Differentiation
Gata4
Ngn3
Pdx1
AFP

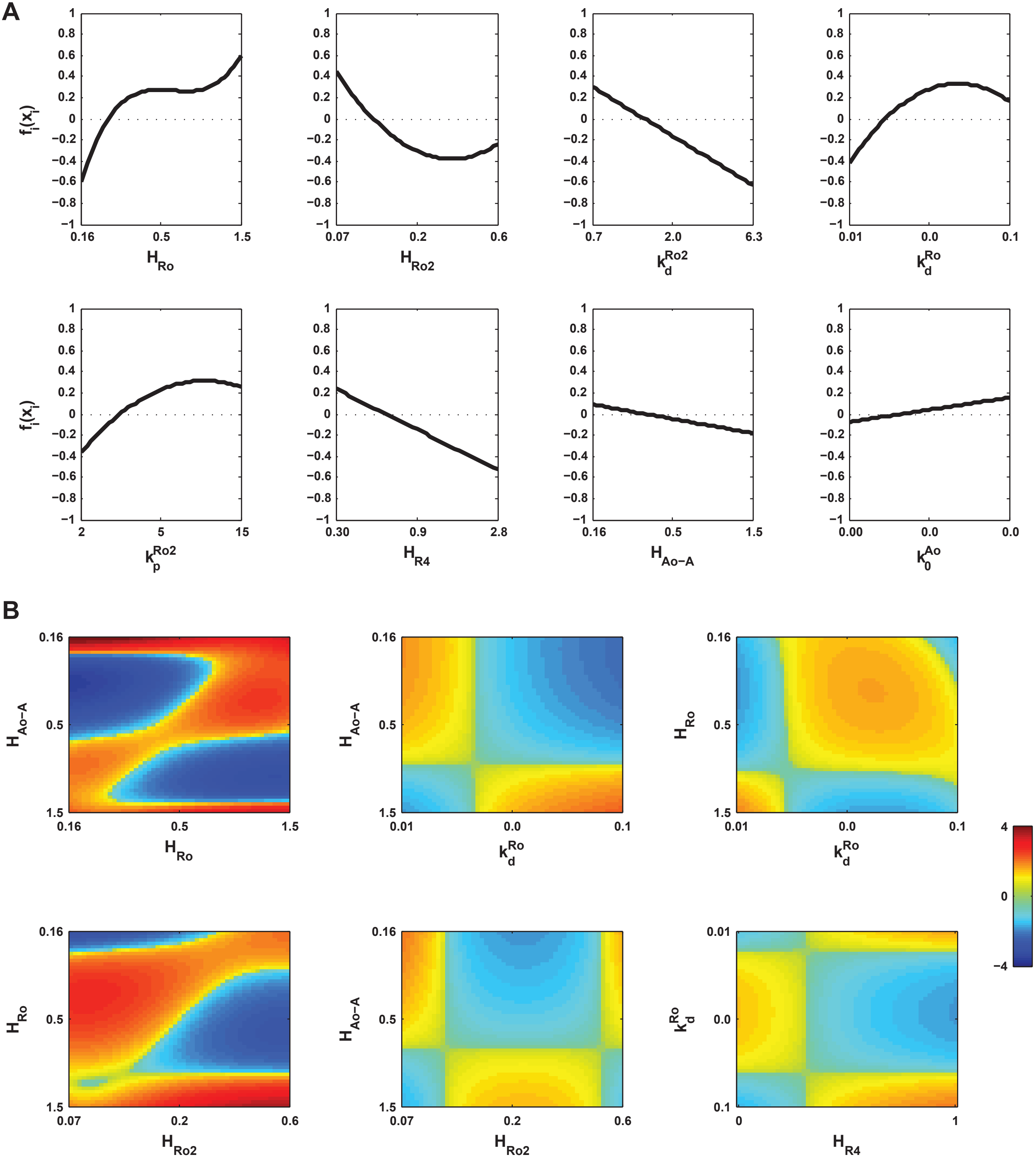

System 2
System 3
System 4
S/N
30
0
0.3
3
TS_QS
TS_QM
TS_R5
TS_R6
TS_R7
TS_At

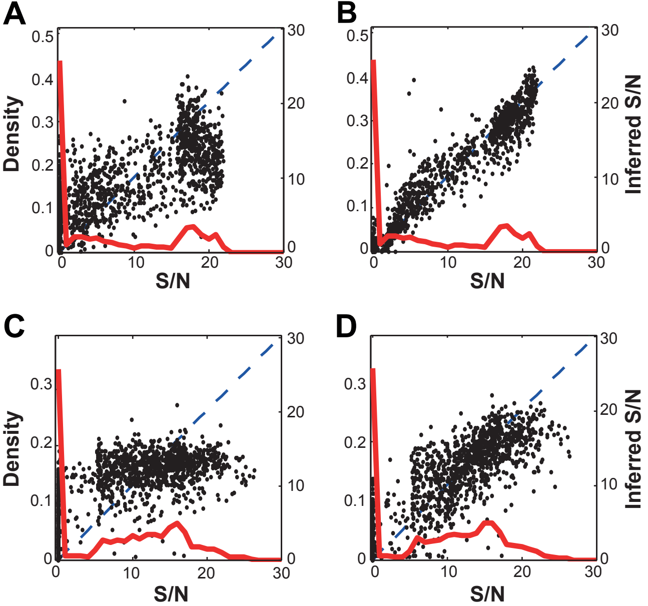

A
S/N
25
20
15
10
5
0
50
90
160
300
600
Cell volume (inverse of noise)
System 2
System 3
System 4
B
S/N
25
20
15
10
5
0
1
2
3
4
5
6
k_b/k_k
System 2
System 3
System 4

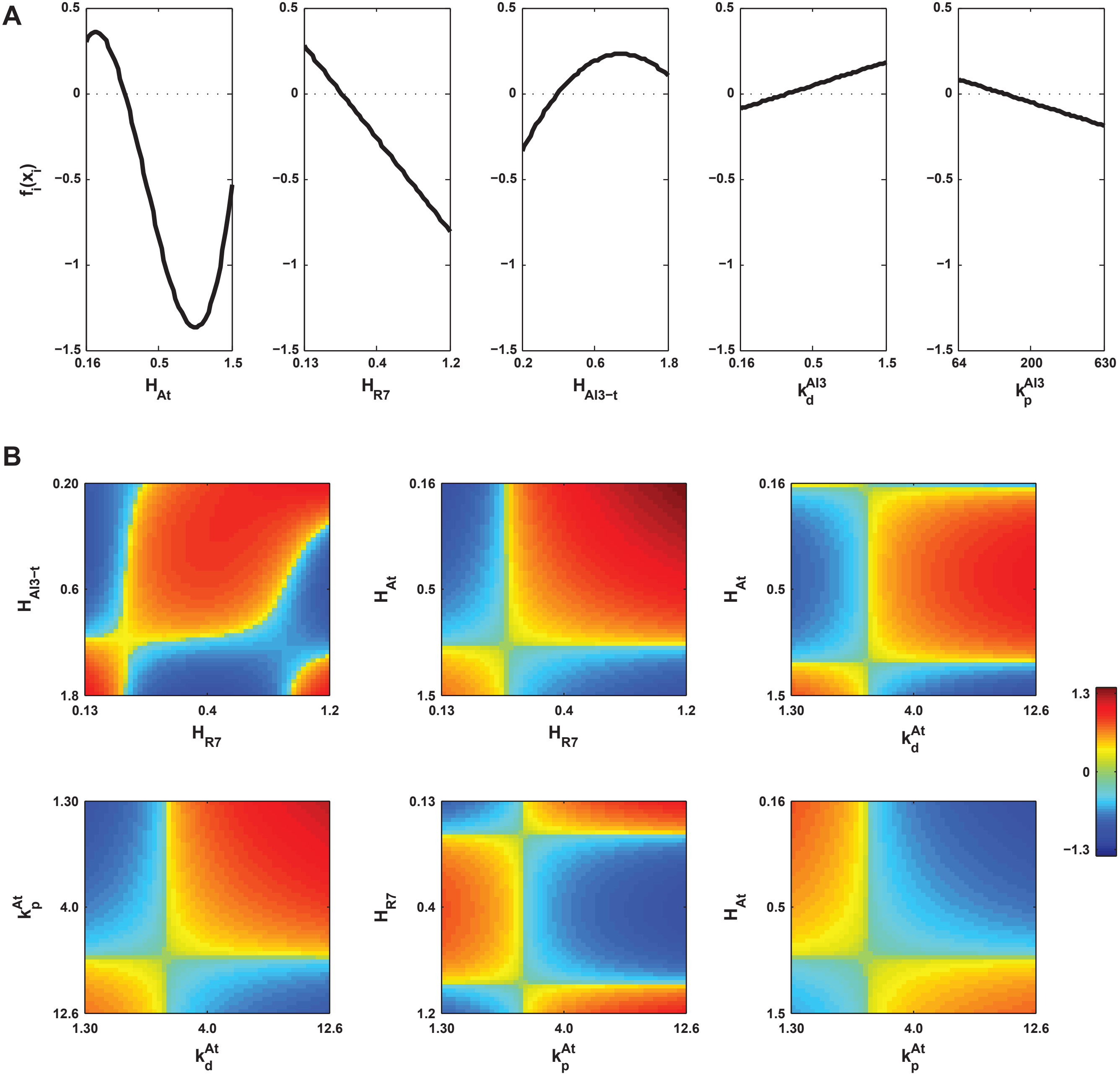


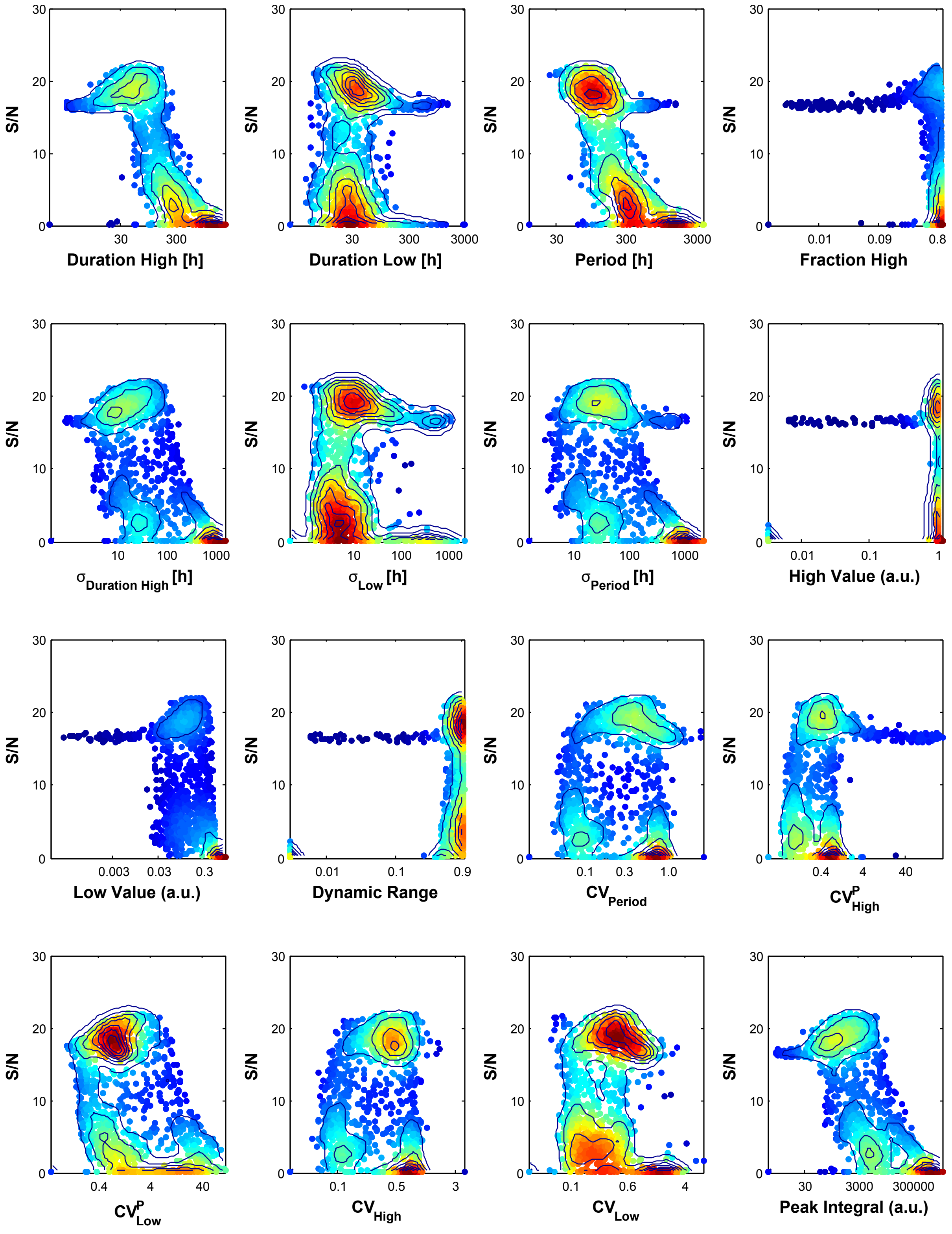
S/N
0
10
20
30
16
50
158
k_p^{Ao}
7e−005
0.0002
0.0006
k_0^{Ao}
0.16
0.5
1.5
H_{Ao-A}
0.004
0.01
0.03
H_{Ao-R}
0.04
0.10
k_d^{Ao}
0.8
2.5
7.0
k_p^{Ro}
0.16
0.5
1.5
H_{Ro}
0.01
0.04
0.1
k_d^{Ro}
1.6
5.0
15
k_p^{Ro2}
0.07
0.2
0.6
H_{Ro2}
0.7
2.0
6.3
k_d^{Ro2}
0.3
0.9
2.8
H_{R4}

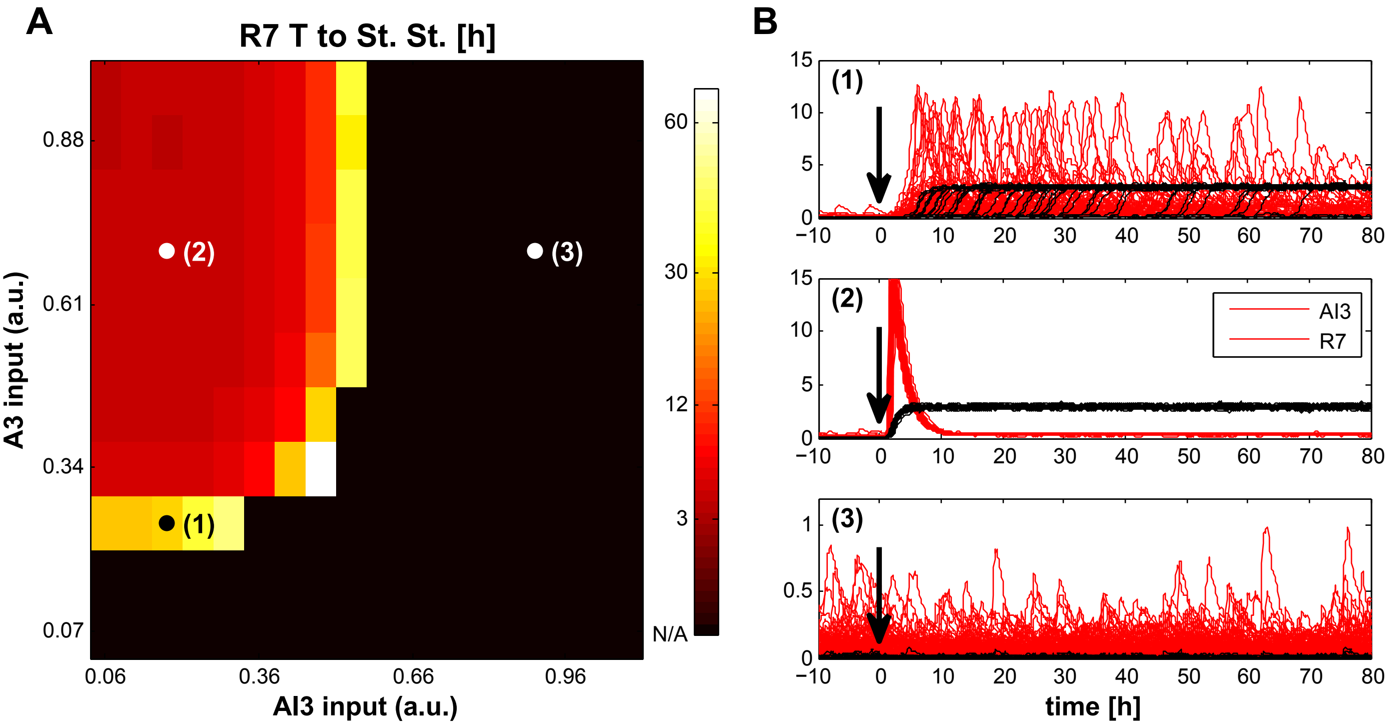

S/N
30
20
10
0
0.3
0.9
2.8
H_{R5-t}
1.3
4.0
12.6
k_p^{At}
0.16
0.5
1.5
H_{At}
0.16
0.5
1.5
H_{Al3-7}
1.3
4.0
12.6
k_d^{At}
64
200
630
k_p^{Al3}
0.2
0.6
1.8
H_{Al3-t}
0.16
0.5
1.5
k_d^{Al3}
0.13
0.4
1.2
H_{R7}

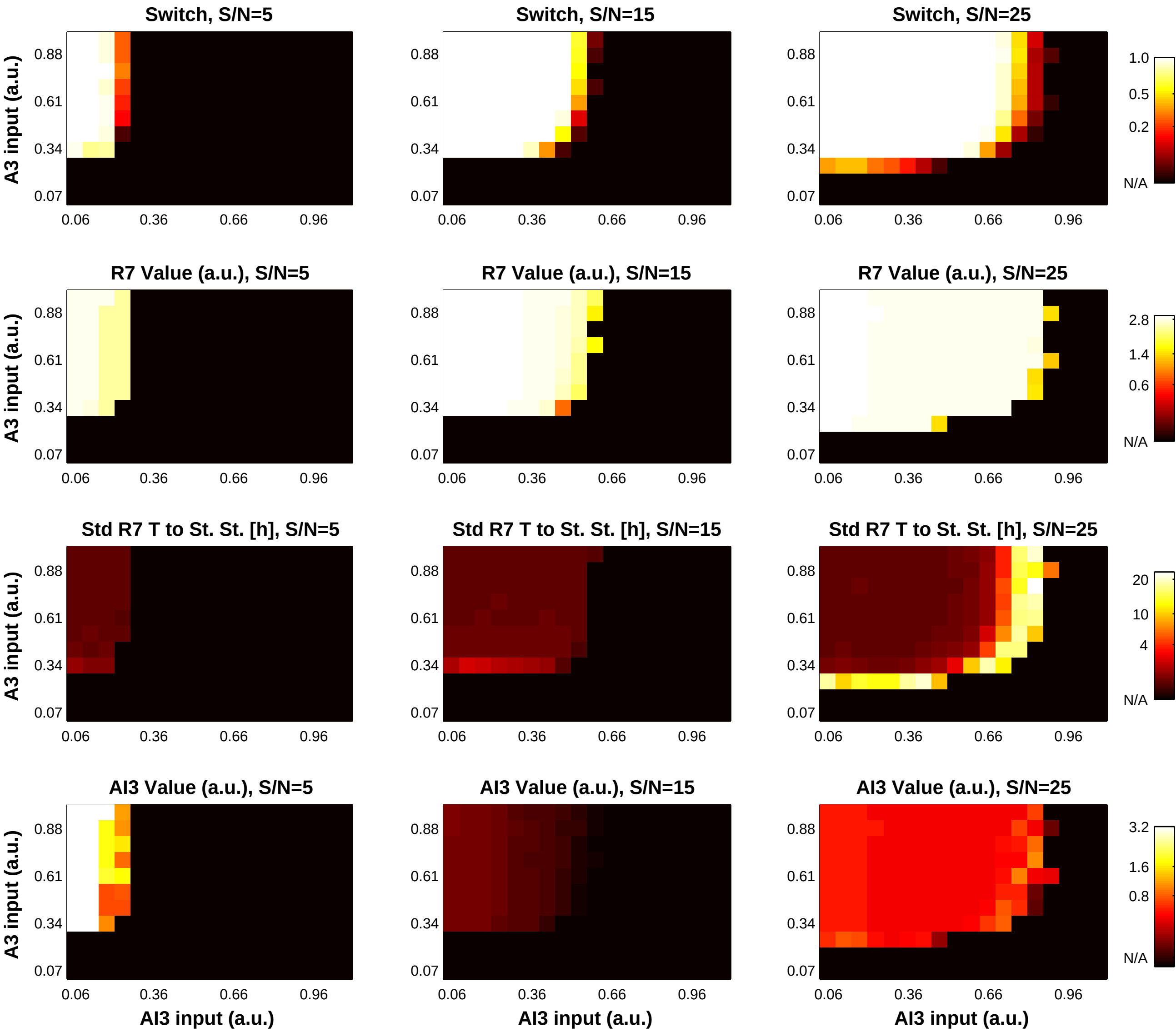

A
f_i(x_i)
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0.16
0.5
1.5
H_Ro
0.07
0.2
0.6
H_Ro2
0.7
2.0
6.3
k_d^Ro2
0.01
0.0
0.1
k_d^Ro
2
5
15
k_p^Ro2
0.30
0.9
2.8
H_R4
0.16
0.5
1.5
H_Ao−A
0.00
0.0
0.0
k_0^Ao
B
H_Ao−A
H_Ro
H_Ro2
k_d^Ro
H_R4
0.16
0.5
1.5
0.07
0.2
0.6
0.01
0.0
0.1
0
1
4
0
−4

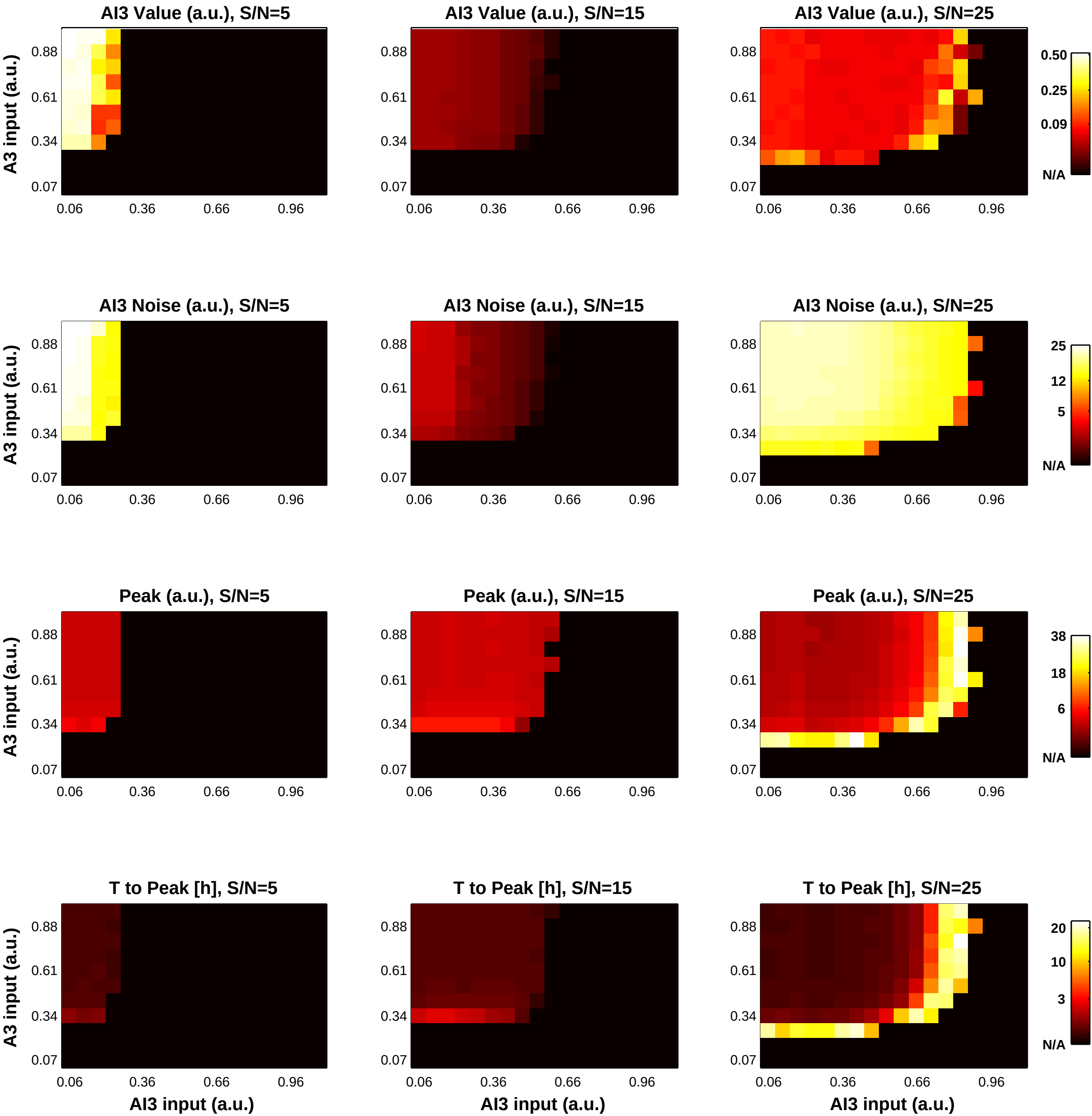

A
B
C
D
Density
Inferred S/N
S/N
0
0.1
0.2
0.3
0.4
0.5
10
20
30

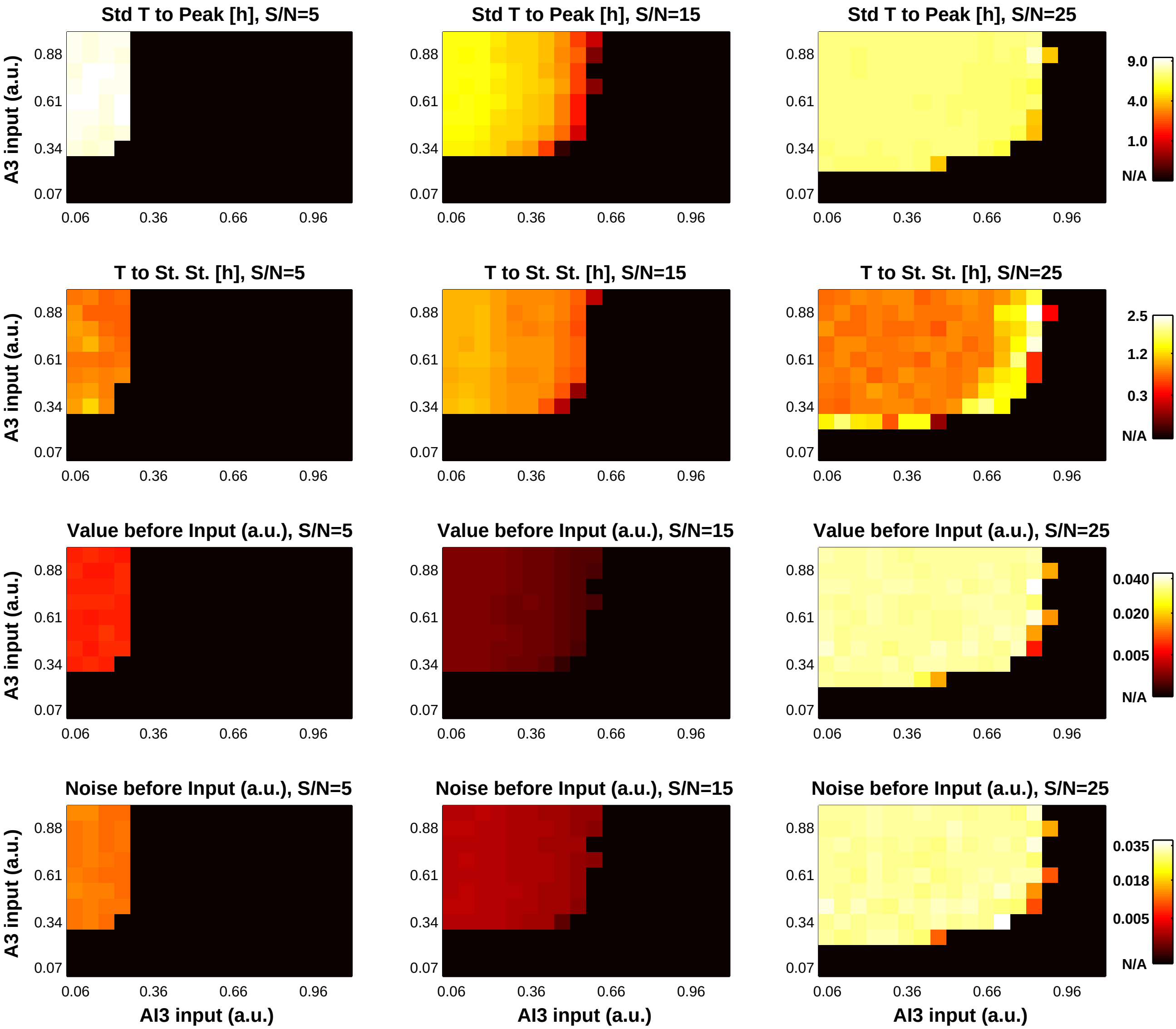

A
$f_i(x_i)$
0.5
0
−0.5
−1
−1.5
0.16
0.5
1.5
H_{At}
0.13
0.4
1.2
H_{R7}
0.2
0.6
1.8
H_{Al3-t}
0.16
0.5
1.5
k_d^{Al3}
64
200
630
k_p^{Al3}
B
0.20
0.6
1.8
H_{Al3-t}
0.13
0.4
1.2
H_{R7}
0.16
0.5
1.5
H_{At}
0.13
0.4
1.2
H_{R7}
0.16
0.5
1.5
H_{At}
1.30
4.0
12.6
k_d^{At}
1.30
4.0
12.6
k_p^{At}
1.30
4.0
12.6
k_d^{At}
0.13
0.4
1.2
H_{R7}
1.30
4.0
12.6
k_p^{At}
0.16
0.5
1.5
H_{At}
1.30
4.0
12.6
k_p^{At}
1.3
0
−1.3

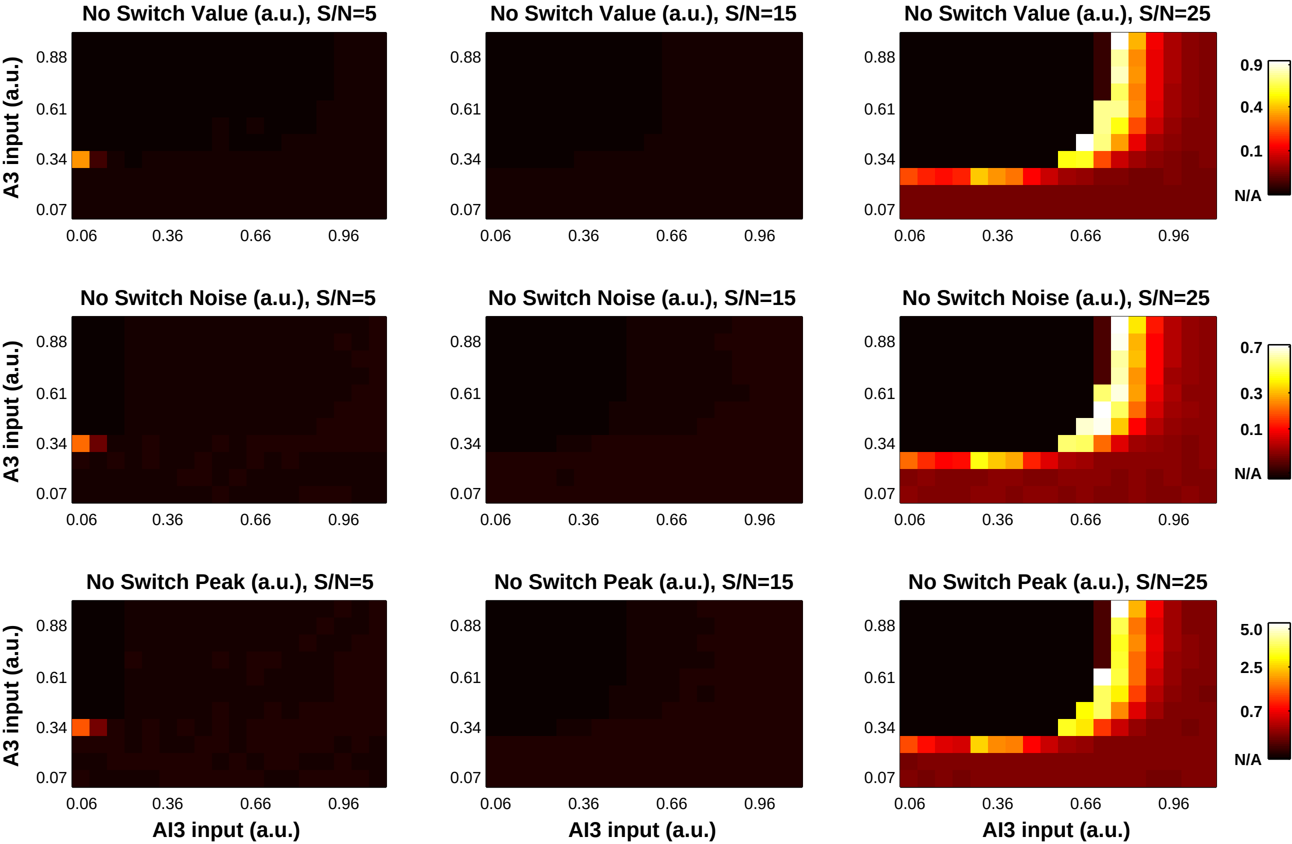

S/N
0
10
20
30
30
300
Duration High [h]
3000
Duration Low [h]
Period [h]
0.01
0.09
0.8
Fraction High
10
100
1000
$\sigma_{\text{Duration High}}$ [h]
σ_{Low} [h]
σ_{Period} [h]
0.01
0.1
1
High Value (a.u.)
0.003
0.03
0.3
Low Value (a.u.)
0.01
0.1
0.9
Dynamic Range
0.1
0.3
1.0
CV_{Period}
0.4
4
40
CV^{P}_{High}
CV^{P}_{Low}
0.1
0.5
3
CV_{High}
0.1
0.6
4
CV_{Low}
30
3000
300000
Peak Integral (a.u.)

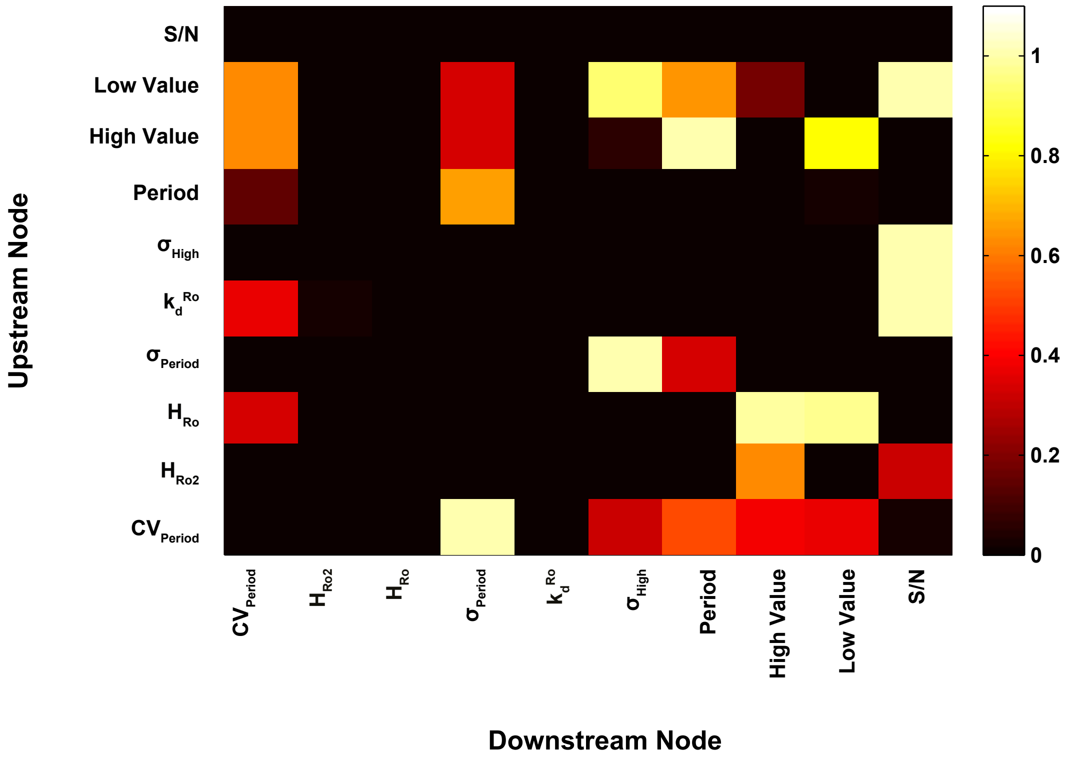

A
R7 T to St. St. [h]
A3 input (a.u.)
0.88
0.61
0.34
0.07
AI3 input (a.u.)
0.06
0.36
0.66
0.96
(1)
(2)
(3)
60
30
12
3
N/A
B
(1)
(2)
(3)
15
10
5
0
1
0.5
−10
0
10
20
30
40
50
60
70
80
AI3
R7
time [h]

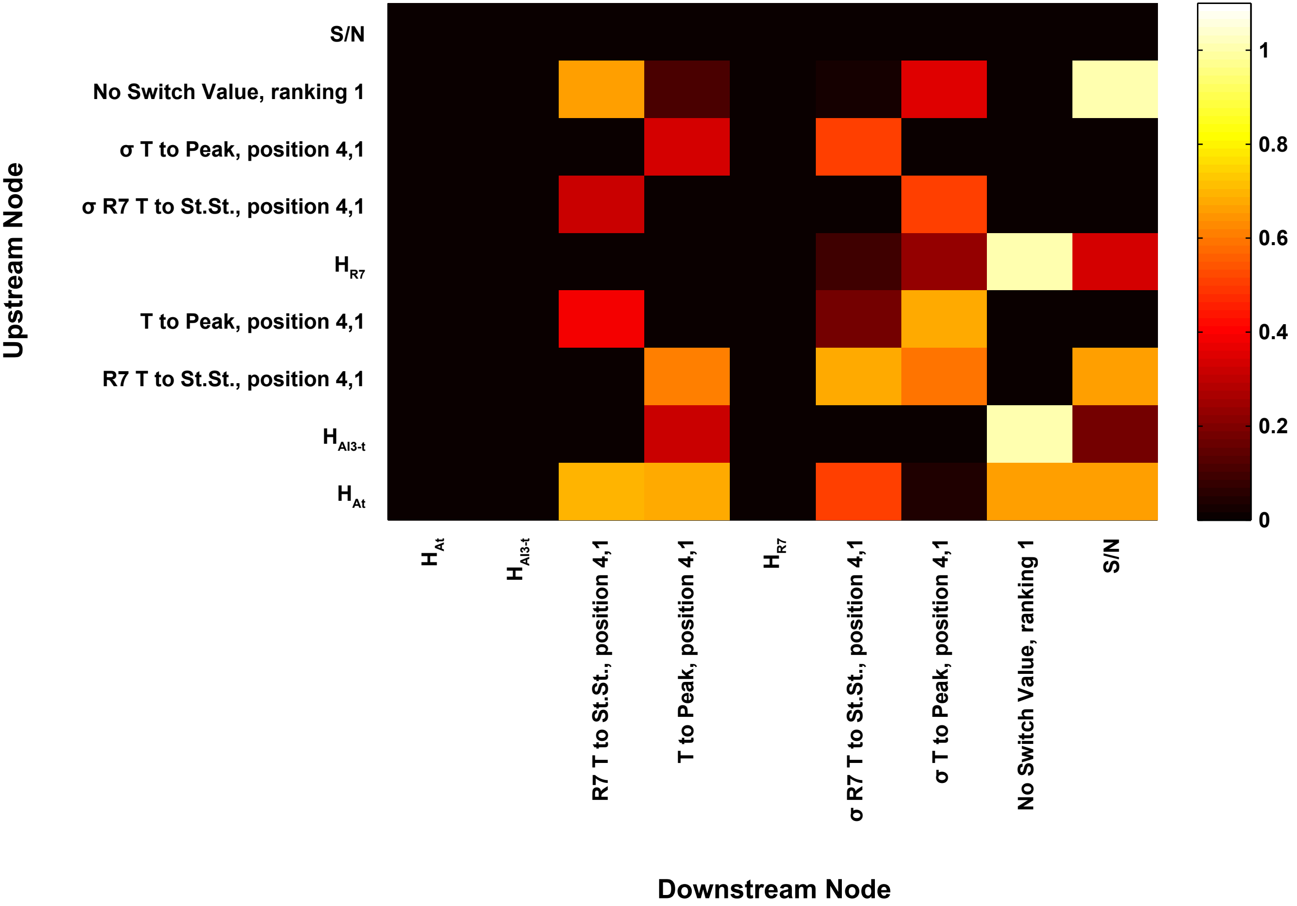

Switch, S/N=5
Switch, S/N=15
Switch, S/N=25
1.0
0.5
0.2
N/A
R7 Value (a.u.), S/N=5
R7 Value (a.u.), S/N=15
R7 Value (a.u.), S/N=25
2.8
1.4
0.6
N/A
Std R7 T to St. St. [h], S/N=5
Std R7 T to St. St. [h], S/N=15
Std R7 T to St. St. [h], S/N=25
20
10
4
N/A
AI3 Value (a.u.), S/N=5
AI3 Value (a.u.), S/N=15
AI3 Value (a.u.), S/N=25
3.2
1.6
0.8
N/A
A3 input (a.u.)
0.88
0.61
0.34
0.07
0.06
0.36
0.66
0.96
AI3 input (a.u.)

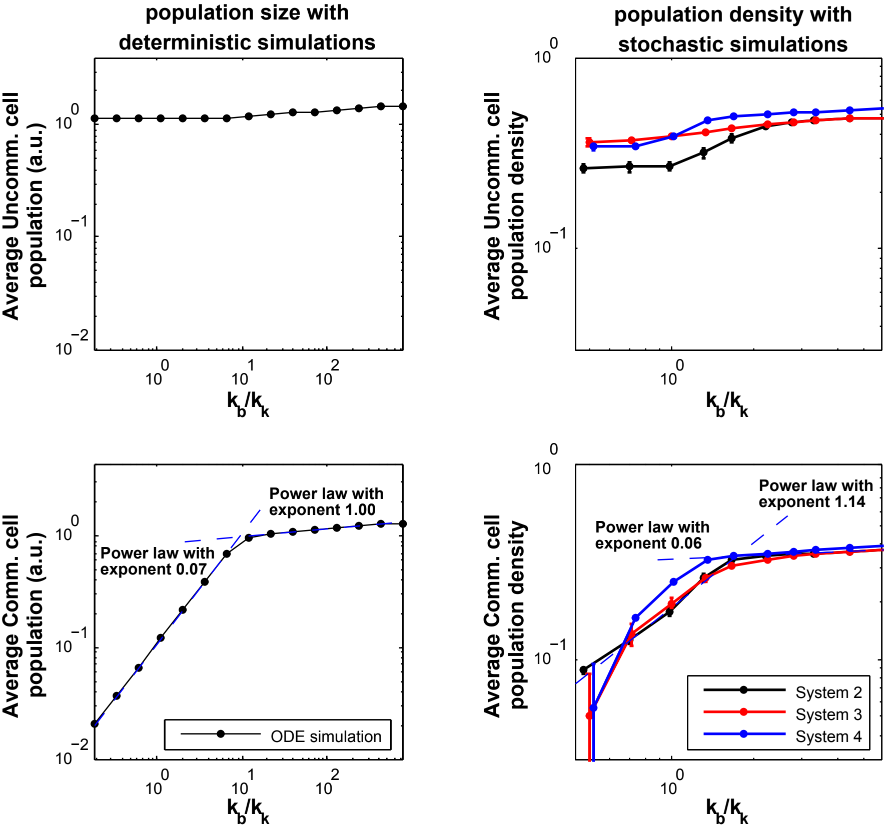

AI3 Value (a.u.), S/N=5
AI3 Value (a.u.), S/N=15
AI3 Value (a.u.), S/N=25
0.50
0.25
0.09
N/A
AI3 Noise (a.u.), S/N=5
AI3 Noise (a.u.), S/N=15
AI3 Noise (a.u.), S/N=25
25
12
5
N/A
Peak (a.u.), S/N=5
Peak (a.u.), S/N=15
Peak (a.u.), S/N=25
38
18
6
N/A
T to Peak [h], S/N=5
T to Peak [h], S/N=15
T to Peak [h], S/N=25
20
10
3
N/A
A3 input (a.u.)
0.88
0.61
0.34
0.07
AI3 input (a.u.)
0.06
0.36
0.66
0.96

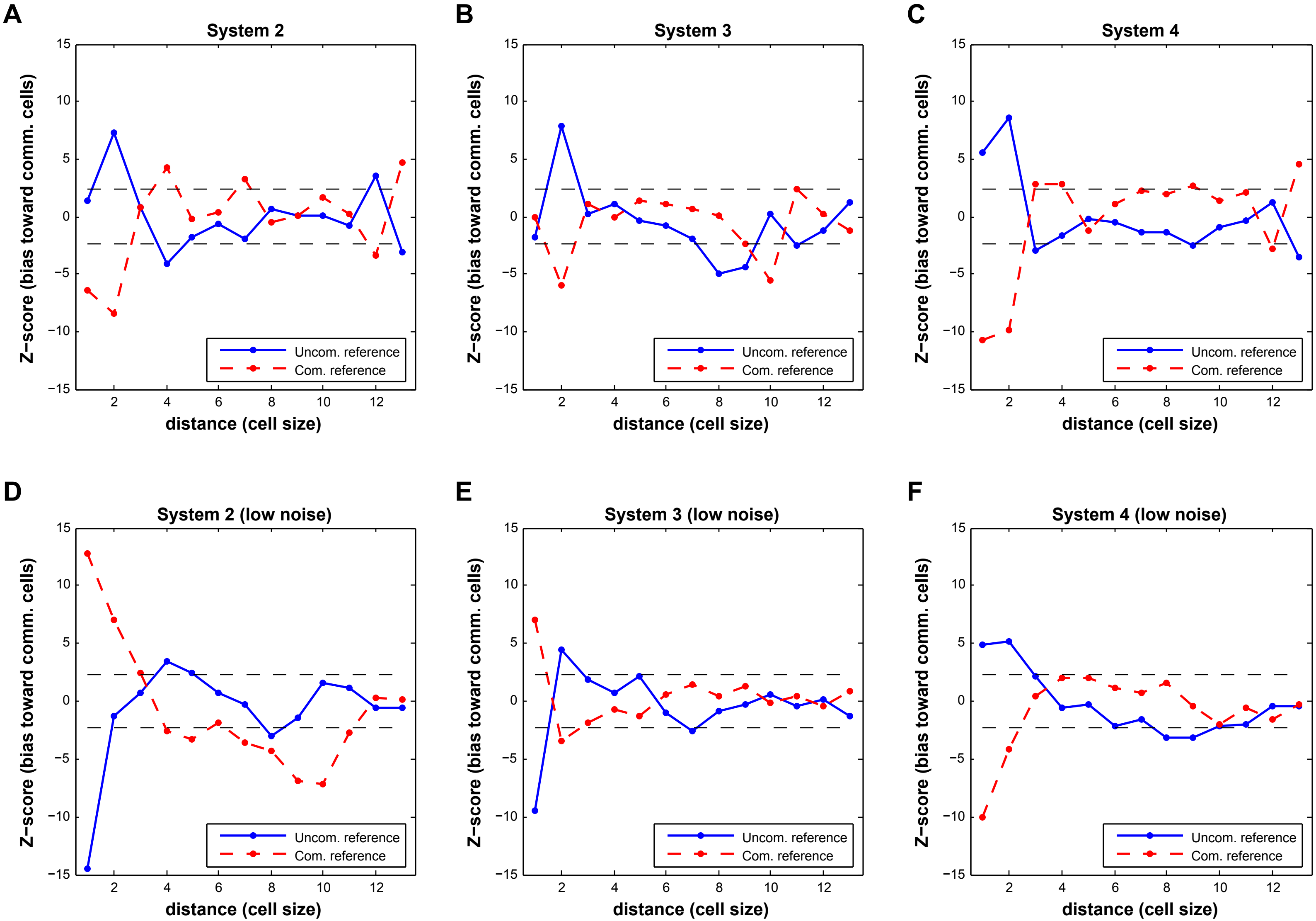

Std T to Peak [h], S/N=5
Std T to Peak [h], S/N=15
Std T to Peak [h], S/N=25
9.0
4.0
1.0
N/A
T to St. St. [h], S/N=5
T to St. St. [h], S/N=15
T to St. St. [h], S/N=25
2.5
1.2
0.3
N/A
Value before Input (a.u.), S/N=5
Value before Input (a.u.), S/N=15
Value before Input (a.u.), S/N=25
0.040
0.020
0.005
N/A
Noise before Input (a.u.), S/N=5
Noise before Input (a.u.), S/N=15
Noise before Input (a.u.), S/N=25
0.035
0.018
0.005
N/A
A3 input (a.u.)
0.88
0.61
0.34
0.07
0.06
0.36
0.66
0.96
AI3 input (a.u.)

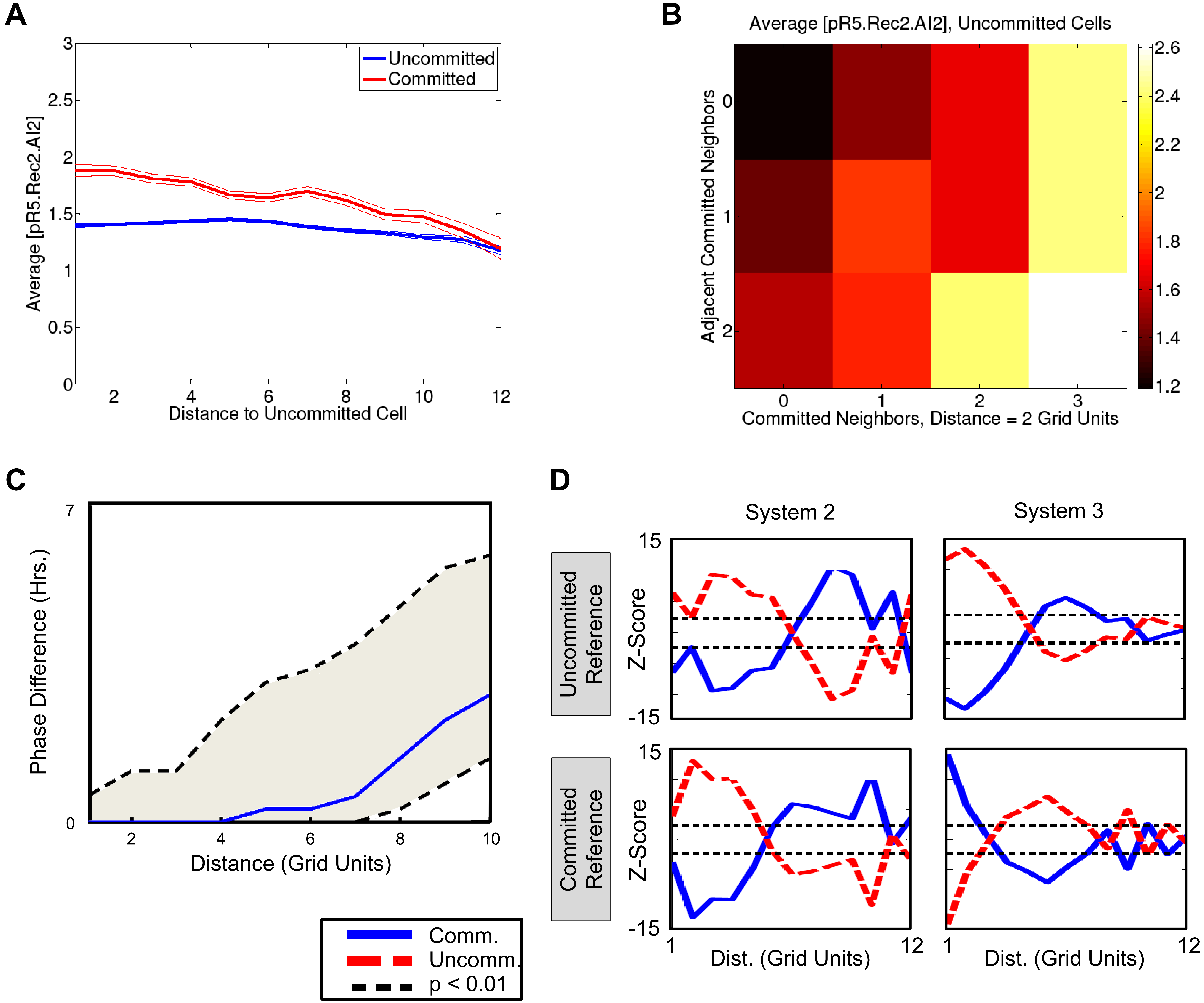

No Switch Value (a.u.), S/N=5
No Switch Value (a.u.), S/N=15
No Switch Value (a.u.), S/N=25
No Switch Noise (a.u.), S/N=5
No Switch Noise (a.u.), S/N=15
No Switch Noise (a.u.), S/N=25
No Switch Peak (a.u.), S/N=5
No Switch Peak (a.u.), S/N=15
No Switch Peak (a.u.), S/N=25
A3 input (a.u.)
0.88
0.61
0.34
0.07
AI3 input (a.u.)
0.06
0.36
0.66
0.96
0.9
0.4
0.1
N/A
0.7
0.3
0.1
N/A
5.0
2.5
0.7
N/A


Upstream Node
S/N
Low Value
High Value
Period
σ_{High}
k_d^{Ro}
σ_{Period}
H_{Ro}
H_{Ro2}
CV_{Period}
Downstream Node
CV_{Period}
H_{Ro2}
H_{Ro}
σ_{Period}
k_d^{Ro}
σ_{High}
Period
High Value
Low Value
S/N
1
0.8
0.6
0.4
0.2
0


Upstream Node
S/N
No Switch Value, ranking 1
σ T to Peak, position 4,1
σ R7 T to St.St., position 4,1
H_{R7}
T to Peak, position 4,1
R7 T to St.St., position 4,1
H_{Al3-t}
H_{At}
H_{At}
H_{Al3-t}
R7 T to St.St., position 4,1
T to Peak, position 4,1
H_{R7}
σ R7 T to St.St., position 4,1
σ T to Peak, position 4,1
No Switch Value, ranking 1
S/N
Downstream Node
1
0.8
0.6
0.4
0.2
0


population size with deterministic simulations
population density with stochastic simulations
Average Uncomm. cell population (a.u.)
Average Uncomm. cell population density
Average Comm. cell population (a.u.)
Average Comm. cell population density
10^{0}
10^{-1}
10^{-2}
10^{1}
10^{2}
k_b/k_k
Power law with exponent 1.00
Power law with exponent 0.07
ODE simulation
Power law with exponent 1.14
Power law with exponent 0.06
System 2
System 3
System 4


A
B
C
D
E
F
System 2
System 3
System 4
System 2 (low noise)
System 3 (low noise)
System 4 (low noise)
Z−score (bias toward comm. cells)
distance (cell size)
15
10
5
0
−5
−10
−15
2
4
6
8
10
12
Uncom. reference
Com. reference


A
Uncommitted
Committed
Average [pR5.Rec2.AI2]
Distance to Uncommitted Cell
0
0.5
1
1.5
2
2.5
3
2
4
6
8
10
12
B
Average [pR5.Rec2.AI2], Uncommitted Cells
Adjacent Committed Neighbors
Committed Neighbors, Distance = 2 Grid Units
0
1
2
3
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
C
Phase Difference (Hrs.)
Distance (Grid Units)
7
0
2
4
6
8
10
Comm.
Uncomm.
p < 0.01
D
System 2
System 3
Uncommitted Reference
Committed Reference
Z-Score
15
-15
1
12
Dist. (Grid Units)

	System 2			System 3			System 4		
Module	Name	para. #	Value	Name	para. #	Value	Name	para. #	Value
Quorum singaling	k_p^{AI1}	1	2.0	identical			identical		
	H_{AI1}	2	1.5	identical			identical		
	k_d^{AI1}	3	0.2	identical			identical		
	k_p^{AI2}	4	2.0	identical			identical		
	H_{AI2}	5	0.45	identical			identical		
	k_d^{AI2}	6	0.2	identical			identical		
	k_{diff}	7	5.0	identical			identical		
Quorum sensing module	k_p^{A1}	8	1.0	identical to Syst. 2			k_p^{R1}	8	1.0
	H_{A1}	9	2.7	identical to Syst. 2			H_{R1}	9	2.7
	k_d^{A1}	10	1.0	identical to Syst. 2			k_d^{R1}	10	1.0
	k_p^{R2}	11	1.0	identical			identical		
	H_{R2}	12	2.0	identical			identical		
	k_d^{R2}	13	1.0	identical			identical		
	–			k_p^{A3}	14	2.0	identical		
	–			H_{A3-1}	15	0.5	identical		
	–			H_{A3-4}	16	0.5	H_{A3-2}	16	0.5
	–			k_d^{A3}	17	2.0	identical		
Commitment module	k_p^{R5}	18	1.0	identical			identical		
	H_{R5-1}	19	0.6	H_{R5-3}	19	0.6	identical		
	H_{R5-2}	20	0.5	identical to Syst. 2			H_{R5-t}	20	0.9
	k_d^{R5}	21	1.0	identical			identical		
	k_p^{R6}	22	1.0	identical			identical		
	H_{R6-5}	23	0.5	identical			identical		
	H_{R6-7}	24	0.5	identical			identical		
	k_d^{R6}	25	1.0	identical			identical		
	k_p^{R7}	26	3.0	identical			identical		
	H_{R7}	27	0.4	identical			identical		
	k_d^{R7}	28	1.0	identical			identical		
Cell fate parameters	k_p^{GAF}	29	1.7	identical			identical		
	H_{GAF}	30	0.7	identical to Syst. 2			H_{GAF}	30	0.3
	k_d^{GAF}	31	1.0	identical			identical		
	k_p^{GATA}	32	1.0	identical			identical		
	H_{GATA}	33	0.2	identical			identical		
	k_d^{GATA}	34	1.0	identical			identical		
Additional modules	–			k_p^{Ao}	35	50	k_p^{At}	35	4.0
	–			k_0^{Ao}	36	0.0002	H_{At}	36	0.5
	–			H_{Ao-A}	37	0.5	k_d^{At}	38	4.0
	–			H_{Ao-R}	38	0.01	k_p^{AI3}	39	200
	–			k_d^{Ao}	39	0.1	H_{AI3-t}	40	0.6
	–			k_p^{Ro}	40	2.5	H_{AI3-7}	37	0.5
	–			H_{Ro}	41	0.5	k_d^{AI3}	41	0.5
	–			k_d^{Ro}	42	0.04	–		
	–			k_p^{Ro2}	43	5	–		
	–			H_{Ro2}	44	0.2	–		
	–			k_d^{Ro2}	45	2.0	–		
	–			k_p^{R4}	46	1.0	–		
	–			H_{R4}	47	0.9	–		
	–			k_d^{R4}	48	1.0	–		

Table S1: Parameters for the Langevin models of Systems 2 to 4

Quorum signaling module	Quorum sensing module	Toggle switch
$\overline{k_p^{AI1}} = TS_{QS} k_p^{AI1}$ $\overline{k_d^{AI1}} = TS_{QS} k_d^{AI1}$	$\overline{k_p^{A1}} = TS_{QM} k_p^{A1}$ $\overline{k_d^{A1}} = TS_{QM} k_d^{A1}$	$\overline{k_p^{R5}} = TS_{R5} k_p^{R5}$ $\overline{k_d^{R5}} = TS_{R5} k_d^{R5}$
$\overline{k_p^{AI2}} = TS_{QS} k_p^{AI2}$ $\overline{k_d^{AI2}} = TS_{QS} k_d^{AI2}$	$\overline{k_p^{R2}} = TS_{QM} k_p^{R2}$ $\overline{k_d^{R2}} = TS_{QM} k_d^{R2}$	$\overline{k_p^{R6}} = TS_{R6} k_p^{R6}$ $\overline{k_d^{R6}} = TS_{R6} k_d^{R6}$
$\overline{k_{diff}} = TS_{QS} k_{diff}$	(in Systems 3 & 4) $\overline{k_p^{A3}} = TS_{QM} k_p^{A3}$ $\overline{k_d^{A3}} = TS_{QM} k_d^{A3}$	$\overline{k_p^{R7}} = TS_{R7} k_p^{R7}$ $\overline{k_d^{R7}} = TS_{R7} k_d^{R7}$
(in System 4) $\overline{k_p^{AI3}} = TS_{QS} k_p^{AI3}$ $\overline{k_d^{AI3}} = TS_{QS} k_d^{AI3}$		(in System 4) $\overline{k_p^{At}} = TS_{At} k_p^{At}$ $\overline{k_d^{At}} = TS_{At} k_d^{At}$

Table S2: Scaled parameters for the time-scale analysis. The kinetics parameters (k_p^{α} and k_d^{α}) from table S1 are scaled by the time-scale parameters TS_{α} according to their module. For each combination of time-scale parameters, the $\overline{k_p^{\alpha}}$ and $\overline{k_d^{\alpha}}$ parameters are used for the Langevin simulations.

Reactants	**Products**	**Rate**
pI1	pI1+I1	$5.7 \cdot 10^{-5}$
pI1+R7$_2$	pI1.R7$_2$	$1.0 \cdot 10^{-3}$
pI1.R7$_2$	pI1+R7$_2$	$1.0 \cdot 10^{-3}$
pI1.R7$_2$	pI1	$1.0 \cdot 10^{-3}$
pI1.R7$_2$	pI1.R7$_2$+I1	$1.0 \cdot 10^{-6}$
pI1+Rec1.AI1	pI1.Rec1.AI1	$1.0 \cdot 10^{-1}$
pI1.Rec1.AI1	pI1+Rec1.AI1	$1.6 \cdot 10^{-2}$
pI1.Rec1.AI1	pI1	$2.5 \cdot 10^{-3}$
pI1.Rec1.AI1	pI1.Rec1.AI1+I1	$1.0 \cdot 10^{-1}$
pI1.Rec1.AI1+R7$_2$	pI1.R7$_2$.Rec1.AI1	$1.0 \cdot 10^{-3}$
pI1.R7$_2$+Rec1.AI1	pI1.R7$_2$.Rec1.AI1	$1.0 \cdot 10^{-1}$
pI1.R7$_2$.Rec1.AI1	pI1.Rec1.AI1+R7$_2$	$1.0 \cdot 10^{-3}$
pI1.R7$_2$.Rec1.AI1	pI1.R7$_2$+Rec1.AI1	$1.6 \cdot 10^{-2}$
pI1.R7$_2$.Rec1.AI1	pI1.R7$_2$	$2.5 \cdot 10^{-3}$
pI1.R7$_2$.Rec1.AI1	pI1.Rec1.AI1	$1.0 \cdot 10^{-3}$
pI1.R7$_2$.Rec1.AI1	pI1.R7$_2$.Rec1.AI1+I1	$1.0 \cdot 10^{-6}$
pGAF+Rec1.AI1	pGAF.Rec1.AI1	$8.7 \cdot 10^{-2}$
pGAF.Rec1.AI1	pGAF+Rec1.AI1	$8.0 \cdot 10^{-3}$
pGAF.Rec1.AI1	pGAF	$2.5 \cdot 10^{-3}$
pGAF	pGAF+GAF	$1.0 \cdot 10^{-7}$
pGAF.Rec1.AI1	pGAF.Rec1.AI1+GAF	$2.0 \cdot 10^{-2}$
pRec1	pRec1+Rec1	$1.5 \cdot 10^{-2}$
I1	I1+AI1	$2.1 \cdot 10^{-3}$
AI1		$3.0 \cdot 10^{-3}$
Rec1+AI1	Rec1.AI1	$5.6 \cdot 10^{-5}$
Rec1.AI1	Rec1+AI1	$1.4 \cdot 10^{-3}$
Rec1.AI1		$2.5 \cdot 10^{-3}$
GAF+GAF	GAF$_2$	$5.0 \cdot 10^{-3}$
GAF$_2$	GAF+GAF	$1.0 \cdot 10^{-3}$
GAF		$5.0 \cdot 10^{-4}$
I1		$5.7 \cdot 10^{-3}$
Rec1		$2.5 \cdot 10^{-3}$
GAF$_2$		$1.0 \cdot 10^{-3}$
pAo	pAo+Ao	$1.0 \cdot 10^{-5}$
pAo+Ro$_2$	pAo.Ro$_2$	$1.0 \cdot 10^{-3}$
pAo.Ro$_2$	pAo+Ro$_2$	$1.0 \cdot 10^{-7}$
pAo.Ro$_2$	pAo	$1.0 \cdot 10^{-5}$
pAo.Ro$_2$	pAo.Ro$_2$+Ao	$1.0 \cdot 10^{-7}$
pAo+Ao$_2$	pAo.Ao$_2$	$1.0 \cdot 10^{-3}$
pAo.Ao$_2$	pAo+Ao$_2$	$5.0 \cdot 10^{-4}$
pAo.Ao$_2$	pAo	$1.0 \cdot 10^{-4}$
pAo.Ao$_2$	pAo.Ao$_2$+Ao	$4.0 \cdot 10^{-2}$
pAo.Ao$_2$+Ro$_2$	pAo.Ro$_2$.Ao$_2$	$1.0 \cdot 10^{-3}$
pAo.Ro$_2$+Ao$_2$	pAo.Ro$_2$.Ao$_2$	$1.0 \cdot 10^{-3}$
pAo.Ro$_2$.Ao$_2$	pAo.Ao$_2$+Ro$_2$	$1.0 \cdot 10^{-7}$
pAo.Ro$_2$.Ao$_2$	pAo.Ro$_2$+Ao$_2$	$5.0 \cdot 10^{-4}$
pAo.Ro$_2$.Ao$_2$	pAo.Ro$_2$	$1.0 \cdot 10^{-4}$
pAo.Ro$_2$.Ao$_2$	pAo.Ao$_2$	$1.0 \cdot 10^{-5}$
pAo.Ro$_2$.Ao$_2$	pAo.Ro$_2$.Ao$_2$+Ao	$1.0 \cdot 10^{-7}$

Table S3: List of reactions for the full multicellular model of System 3. Depending on whether the reactions are associative or dissociative, reaction rates are in units of (molecules per cell)$^{-1}$(s)$^{-1}$ or s^{-1}.

Reactants	**Products**	**Rate**
pRo	pRo+Ro+Ro2	$1.0 \cdot 10^{-5}$
pRo+Ao$_2$	pRo.Ao$_2$	$1.0 \cdot 10^{-3}$
pRo.Ao$_2$	pRo+Ao$_2$	$5.0 \cdot 10^{-4}$
pRo.Ao$_2$	pRo	$1.0 \cdot 10^{-4}$
pRo.Ao$_2$	pRo.Ao$_2$+Ro+Ro2	$5.0 \cdot 10^{-3}$
pRo.Ao$_2$+Ao$_2$	pRo.Ao$_2$.Ao$_2$	$1.0 \cdot 10^{-3}$
pRo.Ao$_2$.Ao$_2$	pRo.Ao$_2$+Ao$_2$	$5.0 \cdot 10^{-4}$
pRo.Ao$_2$.Ao$_2$	pRo.Ao$_2$	$1.0 \cdot 10^{-4}$
pRo.Ao$_2$.Ao$_2$	pRo.Ao$_2$.Ao$_2$+Ro+Ro2	$5.0 \cdot 10^{-3}$
Ro		$4.0 \cdot 10^{-5}$
Ro2		$1.0 \cdot 10^{-3}$
Ao		$1.0 \cdot 10^{-4}$
pR4	pR4+R4	$1.0 \cdot 10^{-2}$
pR4+Ro2$_2$	pR4.Ro2$_2$	$7.0 \cdot 10^{-4}$
pR4.Ro2$_2$	pR4+Ro2$_2$	$1.0 \cdot 10^{-3}$
pR4.Ro2$_2$	pR4	$1.0 \cdot 10^{-3}$
pR4.Ro2$_2$	pR4.Ro2$_2$+R4	$1.0 \cdot 10^{-3}$
pR4.Ro2$_2$+Ro2$_2$	pR4.Ro2$_2$.Ro2$_2$	$7.0 \cdot 10^{-4}$
pR4.Ro2$_2$.Ro2$_2$	pR4.Ro2$_2$+Ro2$_2$	$1.0 \cdot 10^{-3}$
pR4.Ro2$_2$.Ro2$_2$	pR4.Ro2$_2$	$1.0 \cdot 10^{-3}$
pR4.Ro2$_2$.Ro2$_2$	pR4.Ro2$_2$.Ro2$_2$+R4	$1.0 \cdot 10^{-3}$
R4		$1.0 \cdot 10^{-3}$
pA3	pA3+A3	$1.0 \cdot 10^{-6}$
pA3+R4$_2$	pA3.R4$_2$	$2.0 \cdot 10^{-4}$
pA3.R4$_2$	pA3+R4$_2$	$1.0 \cdot 10^{-3}$
pA3.R4$_2$	pA3	$1.0 \cdot 10^{-3}$
pA3.R4$_2$	pA3.R4$_2$+A3	$1.0 \cdot 10^{-6}$
pA3.R4$_2$+R4$_2$	pA3.R4$_2$.R4$_2$	$2.0 \cdot 10^{-4}$
pA3.R4$_2$.R4$_2$	pA3.R4$_2$+R4$_2$	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$	pA3.R4$_2$	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$	pA3.R4$_2$.R4$_2$+A3	$1.0 \cdot 10^{-6}$
pA3+Rec1.AI1	pA3.Rec1.AI1	$8.7 \cdot 10^{-2}$
pA3.Rec1.AI1	pA3+Rec1.AI1	$8.0 \cdot 10^{-3}$
pA3.Rec1.AI1	pA3	$2.5 \cdot 10^{-3}$
pA3.Rec1.AI1	pA3.Rec1.AI1+A3	$2.0 \cdot 10^{-2}$
pA3.Rec1.AI1+R4$_2$	pA3.R4$_2$.Rec1.AI1	$2.0 \cdot 10^{-4}$
pA3.R4$_2$+Rec1.AI1	pA3.R4$_2$.Rec1.AI1	$8.7 \cdot 10^{-2}$
pA3.R4$_2$.Rec1.AI1	pA3.Rec1.AI1+R4$_2$	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.Rec1.AI1	pA3.R4$_2$+Rec1.AI1	$8.0 \cdot 10^{-3}$
pA3.R4$_2$.Rec1.AI1	pA3.R4$_2$	$2.5 \cdot 10^{-3}$
pA3.R4$_2$.Rec1.AI1	pA3.Rec1.AI1	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.Rec1.AI1	pA3.R4$_2$.Rec1.AI1+A3	$1.0 \cdot 10^{-6}$
pA3.R4$_2$.Rec1.AI1+R4$_2$	pA3.R4$_2$.R4$_2$.Rec1.AI1	$2.0 \cdot 10^{-4}$
pA3.R4$_2$.R4$_2$+Rec1.AI1	pA3.R4$_2$.R4$_2$.Rec1.AI1	$8.7 \cdot 10^{-2}$
pA3.R4$_2$.R4$_2$.Rec1.AI1	pA3.R4$_2$.Rec1.AI1+R4$_2$	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$.Rec1.AI1	pA3.R4$_2$.R4$_2$+Rec1.AI1	$8.0 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$.RhlR.AI1	pA3.R4$_2$.R4$_2$	$2.5 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$.Rec1.AI1	pA3.R4$_2$.Rec1.AI1	$1.0 \cdot 10^{-3}$
pA3.R4$_2$.R4$_2$.Rec1.AI1	pA3.R4$_2$.R4$_2$.Rec1.AI1+A3	$1.0 \cdot 10^{-6}$
A3		$1.0 \cdot 10^{-3}$

Table S3 continued.

Reactants	**Products**	**Rate**
pR5	pR5+R5	$1.0 \cdot 10^{-6}$
pR5+A3$_2$	pR5.A3$_2$	$1.0 \cdot 10^{-4}$
pR5.A3$_2$	pR5+A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$	pR5	$5.0 \cdot 10^{-4}$
pR5.A3$_2$	pR5.A3$_2$+R5	$1.0 \cdot 10^{-5}$
pR5.A3$_2$+A3$_2$	pR5.A3$_2$.A3$_2$	$1.0 \cdot 10^{-4}$
pR5.A3$_2$.A3$_2$	pR5.A3$_2$+A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.A3$_2$	pR5.A3$_2$	$5.0 \cdot 10^{-4}$
pR5.A3$_2$.A3$_2$	pR5.A3$_2$.A3$_2$+R5	$1.0 \cdot 10^{-2}$
pR5+R2$_2$	pR5.R2$_2$	$3.0 \cdot 10^{-4}$
pR5.R2$_2$	pR5+R2$_2$	$1.0 \cdot 10^{-3}$
pR5.R2$_2$	pR5	$1.0 \cdot 10^{-3}$
pR5.R2$_2$	pR5.R2$_2$+R5	$1.0 \cdot 10^{-6}$
pR5.R2$_2$+A3$_2$	pR5.A3$_2$.R2$_2$	$1.0 \cdot 10^{-4}$
pR5.A3$_2$+R2$_2$	pR5.A3$_2$.R2$_2$	$3.0 \cdot 10^{-4}$
pR5.A3$_2$.R2$_2$	pR5.R2$_2$+A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.R2$_2$	pR5.A3$_2$+R2$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.R2$_2$	pR5.A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.R2$_2$	pR5.R2$_2$	$5.0 \cdot 10^{-4}$
pR5.A3$_2$.R2$_2$	pR5.A3$_2$.R2$_2$+R5	$1.0 \cdot 10^{-6}$
pR5.A3$_2$.R2$_2$+A3$_2$	pR5.A3$_2$.A3$_2$.R2$_2$	$1.0 \cdot 10^{-4}$
pR5.A3$_2$.A3$_2$+R2$_2$	pR5.A3$_2$.A3$_2$.R2$_2$	$3.0 \cdot 10^{-4}$
pR5.A3$_2$.A3$_2$.R2$_2$	pR5.A3$_2$.R2$_2$+A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.A3$_2$.R2$_2$	pR5.A3$_2$.A3$_2$+R2$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.A3$_2$.R2$_2$	pR5.A3$_2$.A3$_2$	$1.0 \cdot 10^{-3}$
pR5.A3$_2$.A3$_2$.R2$_2$	pR5.A3$_2$.R2$_2$	$5.0 \cdot 10^{-4}$
pR5.A3$_2$.A3$_2$.R2$_2$	pR5.A3$_2$.A3$_2$.R2$_2$+R5	$1.0 \cdot 10^{-6}$
R5		$1.0 \cdot 10^{-3}$
pR2	pR2+R2	$1.0 \cdot 10^{-5}$
pR2+Rec2.AI2	pR2.Rec2.AI2	$3.0 \cdot 10^{-4}$
pR2.Rec2.AI2	pR2+Rec2.AI2	$1.0 \cdot 10^{-3}$
pR2.Rec2.AI2	pR2	$1.0 \cdot 10^{-3}$
pR2.Rec2.AI2	pR2.Rec2.AI2+R2	$1.0 \cdot 10^{-2}$
R2		$1.0 \cdot 10^{-3}$
R2+R2	R2$_2$	$5.0 \cdot 10^{-4}$
R2$_2$	R2+R2	$1.0 \cdot 10^{-3}$
R2$_2$		$1.0 \cdot 10^{-3}$
pI2	pI2+I2	$2.0 \cdot 10^{-3}$
pI2+R6$_2$	pI2.R6$_2$	$1.0 \cdot 10^{-3}$
pI2.R6$_2$	pI2+R6$_2$	$1.0 \cdot 10^{-3}$
pI2.R6$_2$	pI2	$1.0 \cdot 10^{-3}$
pI2.R6$_2$	pI2.R6$_2$+I2	$1.0 \cdot 10^{-5}$
pI2+Rec2.AI2	pI2.Rec2.AI2	$3.0 \cdot 10^{-4}$
pI2.Rec2.AI2	pI2+Rec2.AI2	$1.0 \cdot 10^{-3}$
pI2.Rec2.AI2	pI2	$1.0 \cdot 10^{-3}$
pI2.Rec2.AI2	pI2.Rec2.AI2+I2	$1.0 \cdot 10^{-2}$
pI2.Rec2.AI2+R6$_2$	pI2.R6$_2$.Rec2.AI2	$1.0 \cdot 10^{-3}$
pI2.R6$_2$+Rec2.AI2	pI2.R6$_2$.Rec2.AI2	$3.0 \cdot 10^{-4}$
pI2.R6$_2$.Rec2.AI2	pI2.Rec2.AI2+R6$_2$	$1.0 \cdot 10^{-3}$
pI2.R6$_2$.Rec2.AI2	pI2.R6$_2$+Rec2.AI2	$1.0 \cdot 10^{-3}$
pI2.R6$_2$.Rec2.AI2	pI2.R6$_2$	$1.0 \cdot 10^{-3}$
pI2.R6$_2$.Rec2.AI2	pI2.Rec2.AI2	$1.0 \cdot 10^{-3}$
pI2.R6$_2$.Rec2.AI2	pI2.R6$_2$.Rec2.AI2+I2	$1.0 \cdot 10^{-5}$

Table S3 continued.

Reactants	**Products**	**Rate**
I2	I2+AI2	$1.2 \cdot 10^{-3}$
AI2		$1.0 \cdot 10^{-3}$
Rec2+AI2	Rec2.AI2	$3.0 \cdot 10^{-4}$
Rec2.AI2	Rec2+AI2	$1.0 \cdot 10^{-3}$
Rec2.AI2		$1.0 \cdot 10^{-3}$
pRec2	pRec2+Rec2	$4.4 \cdot 10^{-4}$
pRec2+Rec2.AI2	pRec2.Rec2.AI2	$3.3 \cdot 10^{-4}$
pRec2.Rec2.AI2	pRec2+Rec2.AI2	$1.0 \cdot 10^{-3}$
pRec2.Rec2.AI2	pRec2	$1.0 \cdot 10^{-3}$
pRec2.Rec2.AI2	pRec2.Rec2.AI2+Rec2	$5.3 \cdot 10^{-3}$
I2		$1.0 \cdot 10^{-3}$
Rec2		$1.7 \cdot 10^{-3}$
pR6	pR6+R6	$2.0 \cdot 10^{-2}$
pR6+R5$_2$	pR6.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$	pR6+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$	pR6	$1.0 \cdot 10^{-3}$
pR6.R5$_2$	pR6.R5$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R5$_2$+R5$_2$	pR6.R5$_2$.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$	pR6.R5$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$	pR6.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$	pR6.R5$_2$.R5$_2$+R6	$1.0 \cdot 10^{-6}$
pR6+R7$_2$	pR6.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R7$_2$	pR6+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R7$_2$	pR6	$1.0 \cdot 10^{-3}$
pR6.R7$_2$	pR6.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R7$_2$+R5$_2$	pR6.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$+R7$_2$	pR6.R5$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R5$_2$.R7$_2$	pR6.R7$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$	pR6.R5$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$	pR6.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$	pR6.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$	pR6.R5$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R5$_2$.R7$_2$+R5$_2$	pR6.R5$_2$.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$+R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R7$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R7$_2$+R7$_2$	pR6.R7$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R7$_2$.R7$_2$	pR6.R7$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R7$_2$.R7$_2$	pR6.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R7$_2$.R7$_2$	pR6.R7$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R7$_2$.R7$_2$+R5$_2$	pR6.R5$_2$.R7$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$+R7$_2$	pR6.R5$_2$.R7$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R5$_2$.R7$_2$.R7$_2$	pR6.R7$_2$.R7$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R7$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$.R7$_2$	pR6.R7$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R7$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R5$_2$.R7$_2$.R7$_2$+R5$_2$	pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	$1.0 \cdot 10^{-3}$

Table S3 continued.

Reactants	**Products**	**Rate**
pR6.R5$_2$.R5$_2$.R7$_2$+R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R7$_2$.R7$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R7$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
pR6.R5$_2$.R7$_2$+R5$_2$	pR6.R5$_2$.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$+R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$	$8.0 \cdot 10^{-4}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R7$_2$+R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$+R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R7$_2$	$1.0 \cdot 10^{-3}$
pR6.R5$_2$.R5$_2$.R7$_2$	pR6.R5$_2$.R5$_2$.R7$_2$+R6	$1.0 \cdot 10^{-6}$
R6		$4.0 \cdot 10^{-3}$
R6+R6	R6$_2$	$1.0 \cdot 10^{-3}$
R6$_2$	R6+R6	$1.0 \cdot 10^{-3}$
R6$_2$		$2.0 \cdot 10^{-3}$
pR7	pR7+R7	$2.0 \cdot 10^{-2}$
pR7+R6$_2$	pR7.R6$_2$	$1.0 \cdot 10^{-3}$
pR7.R6$_2$	pR7+R6$_2$	$1.0 \cdot 10^{-3}$
pR7.R6$_2$	pR7	$1.0 \cdot 10^{-3}$
pR7.R6$_2$	pR7.R6$_2$+R7	$1.0 \cdot 10^{-6}$
pR7.R6$_2$+R6$_2$	pR7.R6$_2$.R6$_2$	$1.0 \cdot 10^{-3}$
pR7.R6$_2$.R6$_2$	pR7.R6$_2$+R6$_2$	$1.0 \cdot 10^{-3}$
pR7.R6$_2$.R6$_2$	pR7.R6$_2$	$1.0 \cdot 10^{-3}$
pR7.R6$_2$.R6$_2$	pR7.R6$_2$.R6$_2$+R7	$1.0 \cdot 10^{-6}$
R7		$4.0 \cdot 10^{-3}$
R7+R7	R7$_2$	$1.0 \cdot 10^{-3}$
R7$_2$	R7+R7	$1.0 \cdot 10^{-3}$
R7$_2$		$2.0 \cdot 10^{-3}$
pGATA4	pGATA4+GATA4	$5.0 \cdot 10^{-3}$
pGATA4.R6$_2$	pGATA4.R6$_2$+GATA4	$1.0 \cdot 10^{-6}$
pGATA4+R6$_2$	pGATA4.R6$_2$	$1.0 \cdot 10^{-3}$
pGATA4.R6$_2$	pGATA4+R6$_2$	$1.0 \cdot 10^{-3}$
GATA4		$1.0 \cdot 10^{-3}$
pGATA4.R6$_2$	pGATA4	$1.0 \cdot 10^{-3}$
pNotEndo	pNotEndo+NotEndo	$1.0 \cdot 10^{-2}$
pNotEndo+GATA4	pNotEndo.GATA4	$7.0 \cdot 10^{-4}$
pNotEndo.GATA4	pNotEndo+GATA4	$1.0 \cdot 10^{-3}$
pNotEndo.GATA4	pNotEndo	$1.0 \cdot 10^{-3}$
pNotEndo+Endo	pNotEndo.Endo	$1.0 \cdot 10^{-3}$
pNotEndo.Endo	pNotEndo+Endo	$1.0 \cdot 10^{-3}$
pNotEndo.Endo	pNotEndo	$1.0 \cdot 10^{-3}$
NotEndo		$1.0 \cdot 10^{-3}$

Table S3 continued.

Reactants	**Products**	**Rate**
pEndo	pEndo+Endo+GAF2	$5.0 \cdot 10^{-3}$
pEndo.NotEndo	pEndo.NotEndo+Endo+GAF2	$1.0 \cdot 10^{-7}$
pEndo	pEndo+E	$3.0 \cdot 10^{-5}$
pEndo+NotEndo	pEndo.NotEndo	$1.0 \cdot 10^{-3}$
pEndo.NotEndo	pEndo+NotEndo	$1.0 \cdot 10^{-3}$
pEndo.NotEndo	pEndo	$1.0 \cdot 10^{-3}$
Endo		$1.0 \cdot 10^{-3}$
Ao+Ao	Ao_2	$1.0 \cdot 10^{-3}$
Ao_2	Ao+Ao	$5.0 \cdot 10^{-3}$
Ao_2		$1.0 \cdot 10^{-4}$
Ro+Ro	Ro_2	$1.0 \cdot 10^{-3}$
Ro_2	Ro+Ro	$2.5 \cdot 10^{-2}$
Ro_2		$2.0 \cdot 10^{-5}$
Ro2+Ro2	Ro2_2	$1.0 \cdot 10^{-3}$
Ro2_2	Ro2+Ro2	$1.0 \cdot 10^{-3}$
Ro2_2		$1.0 \cdot 10^{-3}$
R4+R4	R4_2	$1.0 \cdot 10^{-3}$
R4_2	R4+R4	$1.0 \cdot 10^{-3}$
R4_2		$1.0 \cdot 10^{-3}$
R5+R5	R5_2	$1.0 \cdot 10^{-3}$
R5_2	R5+R5	$1.0 \cdot 10^{-3}$
R5_2		$1.0 \cdot 10^{-3}$
A3+A3	A3_2	$5.0 \cdot 10^{-3}$
A3_2	A3+A3	$1.0 \cdot 10^{-3}$
A3_2		$5.0 \cdot 10^{-4}$
CELL AI1	GRID AI1	$1.0 \cdot 10^{-1}$
GRID AI1	CELL AI1	$1.0 \cdot 10^{-1}$
CELL AI2	GRID AI2	$1.0 \cdot 10^{-1}$
GRID AI2	CELL AI2	$1.0 \cdot 10^{-1}$
GRID AI1		$3.0 \cdot 10^{-3}$
GRID AI2		$1.0 \cdot 10^{-3}$

Table S3 continued.

Phenotype Name	**Description**
Duration High	Average duration of the intervals where $R4$ concentration is above 50% of the dynamic range.
Duration Low	Average duration of the intervals where $R4$ concentration is below 50% of the dynamic range.
Period	Average time interval between two switches from high to low concentrations of $R4$.
Fraction High	Fraction of the time when $R4$ concentration is high ($Duration\ High/Period$).
$\sigma_{\text{Duration High}}$	Standard deviation of the duration of the intervals where $R4$ concentration is high.
$\sigma_{\text{Duration Low}}$	Standard deviation of the duration of the intervals where $R4$ concentration is high.
σ_{Period}	Standard deviation of the time interval between two switches from high to low concentrations of $R4$.
High Value	Average concentration of $R4$ when $R4$ has a high concentration.
Low Value	Average concentration of $R4$ when $R4$ has a low concentration.
Dynamic Range	Difference between the high and low values divided by the high value.
$\text{CV}_{\text{Period}}$	Coefficient of variation of the Period: standard deviation of the Period divided by its mean ($\sigma_{\text{Period}}/Period$).
$\text{CV}^{\text{P}}_{\text{High}}$	Standard deviation of the Duration High divided by the mean Period ($\sigma_{\text{Duration High}}/Period$).
$\text{CV}^{\text{P}}_{\text{Low}}$	Standard deviation of the Duration Low divided by the mean Period ($\sigma_{\text{Duration Low}}/Period$).
CV_{High}	Standard deviation of the duration of the intervals where $R4$ concentration is high divided by its mean ($\sigma_{\text{Duration High}}/Duration\ High$).
CV_{Low}	Standard deviation of the duration of the intervals where $R4$ concentration is low divided by its mean ($\sigma_{\text{Duration Low}}/Duration\ Low$).
Peak Integral	*Duration High* multiplied by *High Value*.

Table S4: Phenotypes for the oscillator module (see Figure 8A).

Phenotype Name	**Description**
Switch	Fraction of toggles that switch after the input (a switch occurs when $R7$ reaches a steady state with a concentration above 2 a.u.).
$R7$ Value	Average concentration of $R7$ at steady state (concentration above 2 a.u.) after a switch.
$R7$ T to St. St.	Average time for $R7$ to reach steady state (defined as the time when $R7$ concentration reaches $R7\ \mathit{Value} - 2 \cdot \mathit{R7\ Noise}$) after input (when a switch occurs).
σ $R7$ T to St. St.	Standard deviation of the time for $R7$ to reach steady state after input (when a switch occurs).
$AI3$ Value	Average concentration of $AI3$ at steady state (defines as 1 hour after $R7$ reaches steady state) after a switch.
$AI3$ Noise	Standard deviation of $AI3$ concentration at steady state after a switch.
Peak	Average peak (maximal) value of $AI3$ concentration after input (when a switch occurs).
T to Peak	Average time for $AI3$ to reach its maximal concentration after input (when a switch occurs).
σ T to Peak	Standard deviation of the time for $AI3$ to reach its maximal concentration after input (when a switch occurs).
T to St. St.	Average time for $AI3$ to reach steady state (defined as the time when $AI3$ concentration decreases below $AI3\ \mathit{Value} + 2 \cdot AI3\ \mathit{Noise}$) after peak (when a switch occurs).
σ T to St. St.	Standard deviation of the time for $AI3$ to reach steady state after peak (when a switch occurs).
Value before Input	Average concentration of $AI3$ at steady state before input.
Noise before Input	Standard deviation of $AI3$ concentration at steady state before input.
No Switch Value	Average concentration of $AI3$ at steady state after input when no switch occurs ($R7$ remains below 2 a.u. after input).
No Switch Noise	Standard deviation of the concentration of $AI3$ at steady state after input when no switch occurs.
No Switch Peak	Average maximal concentration of $AI3$ at steady state after input when no switch occurs.

Table S5: Phenotypes for the throttle module (see also Figure 8F).

Feature Name	**Description**
Value$_{x=[AI3],y=[A3]}$	Phenotype value at a particular pair of [AI3],[A3] input values.
Integral	Sum of all values over the image.
Max	Max value over the image.
Min	Min value over the image.
Mean	Mean value over the image.
σ	Standard deviation over the image.
Entropy	Entropy over the image.
Contrast	See `graycoprops()`, Matlab (Natick, MA)
Correlation	See `graycoprops()`, Matlab (Natick, MA)
Energy	See `graycoprops()`, Matlab (Natick, MA)
Homogeneity	See `graycoprops()`, Matlab (Natick, MA)
Area	See `regionprops()`, Matlab (Natick, MA)
Centriod	See `regionprops()`, Matlab (Natick, MA)
Bounding Box	See `regionprops()`, Matlab (Natick, MA)
Major Axis Length	See `regionprops()`, Matlab (Natick, MA)
Minor Axis Length	See `regionprops()`, Matlab (Natick, MA)
Eccentricity	See `regionprops()`, Matlab (Natick, MA)
Orientation	See `regionprops()`, Matlab (Natick, MA)
Convex Area	See `regionprops()`, Matlab (Natick, MA)
Filled Area	See `regionprops()`, Matlab (Natick, MA)
Euler Number	See `regionprops()`, Matlab (Natick, MA)
Equivalent Diameter	See `regionprops()`, Matlab (Natick, MA)
Solidity	See `regionprops()`, Matlab (Natick, MA)
Extent	See `regionprops()`, Matlab (Natick, MA)
Perimeter	See `regionprops()`, Matlab (Natick, MA)
Pattern Matching	Compare two-input function to mean functions calculated for every S/N (binned by integer) in a training set. Record S/N at which pattern most closely fits according to least-squares.
Pattern Matching$_{i=S/N}$	Compare image to mean function calculated for S/N=i (binned by integer) in a training set. Record the sum-of-squares difference calculated over the two-input space.
Pattern Matching$_{Norm}$	Same as "Pattern Matching," but images (and the averages to which they are compared) are mean-centered and variance-normalized.

Table S6: Features used to analyze throttle behavior. These features were measured for each throttle phenotype (see Table S5), where "image" refers to the observed phenotype as a response to the two inputs, $A3$ and $AI3$ (see Figure S14).

Rate Label	**Reaction**
k_0	Express pI1
k_1	Bind pI1.Rec1.AI1
k_2	Dissoc.pI1.Rec12.AI1
k_3	Decay.Rec1.AI1 (Same as Decay.Rec1)
k_4	Express.pI1.Rec1.AI1
k_5	Synth.AI1
k_6	Decay.AI1
k_7	Bind.Rec1.AI1
k_8	Dissoc.Rec1.AI1
k_9	Express.pRec1
k_{10}	Decay.LuxI
k_{11}	Decay.Rec1
k_{12}	Express.pA2
k_{13}	Express.pA2.Rec1.AI1
k_{14}	Decay.A2
k_{15}	Bind.pA2.Rec1.AI1
k_{16}	Dissoc.pA2.Rec1.AI1
k_{17}	AI1 diffusion across cell membrane

Table S7: Rate constants for two-compartment model of the UPC module.

S_i^T	**Phenotype/Rates**	**Feature Name**
0.55	R7 T to St. St.	Value at $A3 = 0.34$, $AI3 = 0.06$
0.16	Peak	Extent when image is thresholded at 90% max value
0.06	σ R7 T to St. St.	Bounding Box (corner of high $A3$, low $AI3$) when image is thresholded at 10% max value
0.06	No Switch Value	Pattern Matching
0.04	No Switch Value	Pattern Matching$_{Norm}$
0.03	H_{At}	
0.03	σ T to St. St.	Pattern Matching$_{Norm}$

Table S8: Top RS-HDMR identified throttle features and their corresponding RS-HDMR sensitivity indices, S_i^T (table supplements Figure 8H).

	System 1	System 2	System 3	System 4
Advantages	• Simplicity • Analytical solutions	• Intermediate simplicity • Rapid feedback for reduced oscillations and improved homeostasis	• Oscillator generates population heterogeneity. • Oscillator behavior and its ability to maintain homeostasis is well insulated from parameter values of other modules. • Improves homeostasis when intercellular signaling is slow.	• Throttle generates population heterogeneity. • Improves homeostasis when toggle switching times are slow. • Works well for various intercellular signaling rates. • Better overall performance relative to System 3 with intermediate molecular noise levels.
Disadvantages	• Poor population heterogeneity • Undesired oscillations in β-cell population levels due to delayed feedback	• Poor population heterogeneity • Highly dependent on reaction rates: requires rapid toggle switching and intercellular signaling.	• Reduced performance relative to System 4 with slow toggle switching and intercellular signaling. • Requires significant molecular noise to operate well relative to System 4.	• Requires a third intercellular signal. • Optimal performance requires parameter fine-tuning to match other modules.