

Engineering of regulated stochastic cell fate determination

Min Wu^a, Ri-Qi Su^b, Xiaohui Li^a, Tom Ellis^{c,d}, Ying-Cheng Lai^b, and Xiao Wang^{a,1}

^aSchool of Biological and Health Systems Engineering and ^bSchool of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287; and ^cCentre for Synthetic Biology and Innovation, and ^dDepartment of Bioengineering, Imperial College London, London SW7 2AZ, United Kingdom

Edited* by Charles R. Cantor, Sequenom, Inc., San Diego, CA, and approved May 16, 2013 (received for review March 21, 2013)

Both microbes and multicellular organisms actively regulate their cell fate determination to cope with changing environments or to ensure proper development. Here, we use synthetic biology approaches to engineer bistable gene networks to demonstrate that stochastic and permanent cell fate determination can be achieved through initializing gene regulatory networks (GRNs) at the boundary between dynamic attractors. We realize this experimentally by linking a synthetic GRN to a natural output of galactose metabolism regulation in yeast. Combining mathematical modeling and flow cytometry, we show that our engineered systems are bistable and that inherent gene expression stochasticity does not induce spontaneous state transitioning at steady state. Mathematical analysis predicts that stochastic cell fate determination in this case can only be realized when gene expression fluctuation occurs on or near attractor basin boundaries (the points of instability). Guided by numerical simulations, experiments are designed and performed with quantitatively diverse gene networks to test model predictions, which are verified by both flow cytometry and single-cell microscopy. By interfacing rationally designed synthetic GRNs with background gene regulation mechanisms, this work investigates intricate properties of networks that illuminate possible regulatory mechanisms for cell differentiation and development that can be initiated from points of instability.

Bistability and the binary decision making it imparts have been widely observed and hypothesized as one of the possible mechanisms for cell fate determination (1–3). Previous studies of bistable systems have attributed this binary decision making to either (i) random and reversible state transitioning [i.e., cells spontaneously and randomly switching back and forth between two states without environmental perturbations (4–7)] or (ii) deterministic and irreversible state transitioning [i.e., cells uniformly and irreversibly choosing one of two states in response to external signals (8–10)]. Few mechanisms, however, have been proposed to explain the scenario of random and yet irreversible cell fate determination, which is commonly seen in development and cell differentiation. Recent studies (11–13) show examples of stochastic and irreversible cell differentiation in multicellular organisms, and they have identified that the central regulatory motif driving these stochastic differentiations is a mutual inhibitory gene regulatory network (GRN), a common topological module that can generate bistability. However, with neither fluctuating environmental cues nor spontaneous state transitioning identified in these cases, an understanding of how cells differentiate stochastically and irreversibly into distinct subpopulations remains elusive, especially when under the tight control of GRNs.

Synthetic gene networks provide an effective platform with which to probe the otherwise intractable properties of common network motifs and to uncover novel mechanisms for counterintuitive observations (14). Such investigations are impossible in their natural settings, where the complex interconnectivity of native GRNs acts as a major barrier to detailed analysis. Synthetic gene networks, on the other hand, are rationally designed and constructed to realize core topological modules of GRN in vivo without interference from auxiliary connections. They can therefore be studied in isolation to greater detail and reveal

novel insights into the design and working of biological systems and processes (15), such as gene expression noise (4, 16–21), multistability (8, 10, 22), oscillations (23–26), intracellular signaling (27, 28), intercellular communications (29, 30), and multicellular pattern formation (31, 32).

In a demonstration of how a GRN can be used to investigate an intricate network property, we use synthetic biology approaches here to explore possible mechanisms for stochastic and irreversible cell fate determinations in a multistable system. First, we combined experimental characterization and mathematical modeling to calibrate dynamic parameters of three different mutual inhibitory gene networks constructed using our previously developed promoter library approach (8). With the aim of initializing the cell population on the basin boundary, we used the natural regulation machinery for galactose metabolism in yeast to shut off the synthesis of all proteins completely in the synthetic network in glucose-supplemented media, and hence to initialize the system close to the basin boundary. The predicted stochastic differentiation was then experimentally verified in all three gene networks by moving the cells, after initial growth to steady state in glucose-supplemented media, to galactose-supplemented media with the Tn10.B tetracycline repressor (TetR) inhibitor anhydrotetracycline (ATc) to ensure bistability.

Results

Bistable Regions Located by Showing Hysteresis. Each multistable gene network can be viewed as an energy potential landscape with multiple local minima, each representing one specific cellular state [i.e., cell fate (33)]. Cells operate on these landscapes and eventually settle into one of the minima, choosing their cell fates until they transition to another state in response to a perturbation, signal, or even inherent noise. Which local minimum a cell settles into depends largely on where the cell starts its growth on this landscape (i.e., its initial conditions). Typically when inherent noise is not too strong, such cell fate determination is fairly predictable and deterministic, depending on initial conditions (green and red circles in Fig. 1A). However, when the initial conditions happen to be on the boundary between local minima, the eventual outcome of the cell could become stochastic, regardless of how weak the inherent noise is. This is analogous to uncertain marble movements when placed on the top of the barrier between two local minima (black circle in Fig. 1A). Such physical intuition has not, however, been observed or realized in either natural or engineered biological systems.

Author contributions: X.W. designed research; M.W., R.-Q.S., and X.L. performed research; M.W., R.-Q.S., X.L., T.E., Y.-C.L., and X.W. analyzed data; and R.-Q.S., T.E., Y.-C.L., and X.W. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: xiaowang@asu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1305423110/-DCSupplemental.

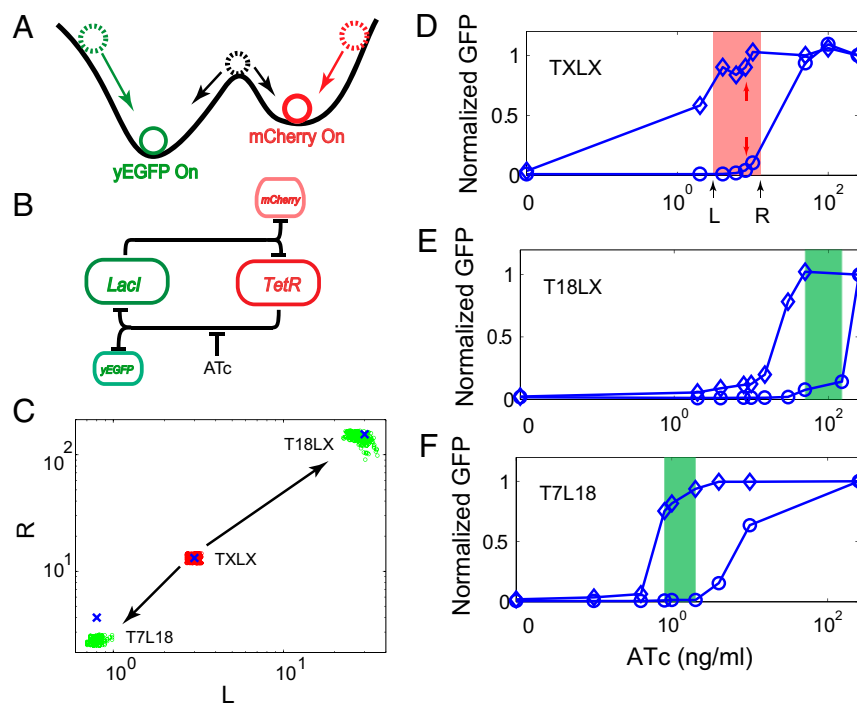


Fig. 1. Bistable systems experimentally verified by showing hysteresis. (A) When fluctuations of movements (inherent noise) are small, initial conditions of free-moving marbles (red and green dashed circles) on a bistable energy landscape can determine their final steady states (red and green solid circles, analogous to GFP ON and mCherry ON states of our synthetic gene network). However, when the initial conditions are on the tip of the barrier between two local minima (black dashed circle), the marble will drop into one of the two local minima randomly (suggested by black arrows), even with a minimal amount of noise. (B) Schematic diagram of an engineered yeast mutual inhibitory network. LacI and TetR proteins repress each other's expression. yEGFP and mCherry are also under control of these promoters, and hence indicate the abundance of LacI and TetR proteins. ATc inductions can be used to block TetR (details are provided in *SI Appendix*). (C) Left bound (L) of the bistable region is plotted vs. the right bound (R) on log scale for all three strains. Blue crosses represent experimentally observed values, and colored dots represent model-fitted (red) or -predicted (green) values. Each dot represents a prediction by one parameter set. Overlap of dots and crosses shows the accuracy of predictions. (D–F) Average green fluorescence of gated cells (pretreated with 250 ng/mL ATc) at steady states is plotted as diamonds for three different strains. Similarly, data for cultures first treated with no ATc are marked as circles. In all three strains, there is a range in which the same dose of ATc induction produces different levels of yEGFP, the sign of hysteresis. Colored, shaded regions indicate experimentally observed (D) or model-predicted (E and F) bistable regions for each strain. In D, red arrows point to two data points of TXLX with 8 ng/mL ATc induction, whose full histograms are shown in Fig. 2B in purple and green. Choices of L and R are illustrated on the x axis.

To investigate the possibility of initial condition-dependent random cell fate determination, we chose to use the simplest multistable biological systems, bistable gene networks, to test this hypothesis in cells. Based on our previous work in an engineered yeast GRN with a mutual inhibitory motif (8), three modified versions were constructed with both yeast-enhanced green fluorescent protein (yEGFP) and mCherry red fluorescent protein as reporters for Lac inhibitor (LacI) and TetR (Fig. 1B and *SI Appendix*, Fig. S1). To verify bistability and determine the optimal dosage of inducer for balanced steady states, we first tested the gene network constructed using the previously described TX and LX synthetic promoters (8), which we called TXLX. In this synthetic gene network, the repressors TetR and LacI inhibit the expression of each other by binding to their corresponding operator sites, TetR operator (*Otet*) and LacI operator (*Olac*), placed within engineered *GALI* promoters (8). The promoters were chosen from our previously generated promoter library (8). In this library, a *GALI* promoter with an *Otet* operator site (labeled TX) was engineered to form the foundation of 20 variants (labeled T1 to T20). Similarly, a library of LacI-controlled promoters (labeled LX and L1 to L20) was also engineered. All these promoters can be regulated by corresponding repressors and only differ in their maximal expression levels and leakage under repression. The choices of promoter combinations therefore fully determine network characteristics. As readouts, fluorescent protein reporters are under control of TX and LX promoters for all three strains tested

so that they can track LacI and TetR dynamics and also generate strong signals.

Two sets of experiments were designed to demonstrate that the system is capable of hysteresis, an indicator of bistability (5, 10). Because we previously showed that the networks exhibit the default state of TetR ON (GFP OFF) (8), where isopropyl- β -D-thiogalactopyranoside has no effect at steady state, we chose ATc inductions here as the method for tuning parameters to probe the system's bistability region. TXLX cultures treated with full ATc induction (250 ng/mL) in galactose-supplemented media for 48 h were rediluted into media containing 0, 2, 4, 6, 8, 10, 50, 100, and 250 ng/mL ATc. Using flow cytometry, measurements of yEGFP were taken after the fluorescence levels became stable in each condition (diamonds in Fig. 1D; experimental details are provided in *SI Appendix*). It can be seen that full ATc induction successfully tilts the balance toward LacI and produced LacI-dominant cell cultures (high yEGFP expression). It is also clearly demonstrated that the cultures remain LacI-dominant after growth in media with the ATc concentration as low as 4 ng/mL but fail to maintain the state with 0 and 2 ng/mL ATc induction. Meanwhile, similar experiments were also carried out for cultures treated with no ATc induction for 48 h. Cells were also rediluted into media containing various doses of ATc, and yEGFP expression was measured (circles in Fig. 1D). It can be seen that with no ATc induction, the gene network is TetR-dominant (low yEGFP expression). Even with increased doses of ATc induction, cell cultures with up to 10 ng/mL ATc induction remain low in yEGFP

expression, and only demonstrate increased yEGFP expression with greater than 10 ng/mL ATc induction. Taken together, it can be seen that with ATc induction between ~ 3 and 13 ng/mL, the gene network responds to ATc inductions in an initial condition-dependent fashion, demonstrating hysteresis, and hence verifying bistability (*SI Appendix, Fig. S4*).

Experimental data were used by the mathematical model to calibrate parameters to locate the bistable region accurately (blue cross and red dots in Fig. 1C; details about parameter fitting are provided in *SI Appendix, Section II*). This fully quantitative description of our gene networks makes it possible to predict bistable regions for other gene networks. Using calibrated parameters and only adjusting promoter strengths to reflect different constructs, bistable regions are predicted for network T18LX (green dots in Fig. 1C and green shaded area in Fig. 1E) and network T7L18 (green dots in Fig. 1C and green shaded area in Fig. 1F). To test the accuracy of model predictions, similar experiments were conducted to test hysteresis and to locate bistable regions for both constructs. As illustrated in Fig. 1E, the construct of T18LX shows a bistable region approximately between 30 and 150 ng/mL ATc, whereas the construct of T7L18 shows a bistable region approximately between 0.6 and 4 ng/mL ATc (Fig. 1F). The fact that both experimental measurements are consistent with model predictions demonstrates the predictive power of our model and also builds a solid foundation for further predictions with gene expression stochasticity taken into account.

Model Predicts Ways to Achieve Stochastic and Irreversible Cell Fate Determination. With parameters fitted against experimental data, the model is numerically mapped into a quasipotential that directs evolution of protein abundances (Fig. 2A). The potential at each point is defined as the trajectory length from this point to its final steady state without stochasticity (34). Analogous to a marble moving in a landscape in response to gravity, protein abundance changes are directed by the vector field (2D visualization is shown in *SI Appendix, Fig. S7*), which is different from gravitational forces as explained by Kim and Wang (35). This map visualizes a more complete picture of a bistable landscape, where the dark blue basin bottoms represent two stable steady

states and other regions represent slopes directing the system to the steady states. This saddle-like landscape has one ridge and one valley visualized as white dashed and solid lines, respectively, which are known mathematically as the manifolds of this dynamical system (36). Although the latitudes on the solid line form a double-well potential as depicted in Fig. 1A, the dashed line is the basin boundary, mathematically termed the separatrix, that divides the landscape into two basins with separate local minima. Cells initialized within each basin should eventually settle into their respective basin bottoms. The intersection of the ridge and valley (marked by a black arrow) corresponds to the barrier tip in Fig. 1A, mathematically known as the unstable steady state (USS) (36). As hypothesized above, cells initialized at the USS could show random cellular state determination even with a minimal level of stochasticity. However, such a specific initial condition is difficult to realize experimentally despite careful parameter calibrations because initialization at this point requires complete and accurate control of protein abundances in a living cell, which is nontrivial to achieve using chemical inductions.

The vector field ensures that trajectories initialized near the ridge will follow it approaching the USS (36). Therefore, we hypothesize that given some stochasticity, cells initialized on or near the separatrix would first approach the USS along the ridge and then diverge randomly to different cellular states. This eliminates the need to be initialized exactly on the USS for stochastic differentiation. To test this hypothesis computationally, the model was expanded to incorporate gene expression stochasticity using the Gillespie algorithm (37) to simulate the temporal dynamics of cellular state determination. It can be seen in Fig. 2A that two isogenic cells (illustrated as two solid lines) starting from the same initial condition (no LacI and no TetR) near the separatrix first follow similar trajectories approaching the USS, then diverge onto distinct trajectories, and finally diverge into different cellular states, two separate local minima colored in dark blue (Fig. 2A). The coloring of the solid lines represents the changing color of detectable fluorescence reporter signals. After reaching steady state, one cell will emit a strong red fluorescence signal and the other cell will emit a strong green fluorescence signal. This result computationally verified the hypothesis that as

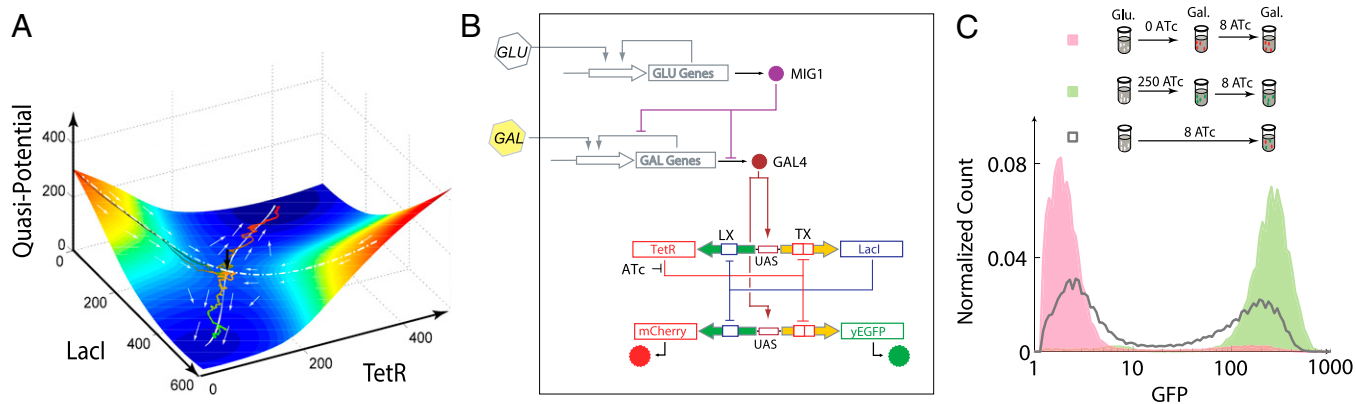


Fig. 2. Model-predicted stochastic cell fate determination and experimental verifications. (A) Based on the ordinary differential equation (ODE) model with 8 ng/mL ATc induction, derivatives of state variables are mapped onto an energy quasipotential landscape that directs the system's evolution. Altitude is color-mapped, with a cold color indicating a lower energy potential. The dashed and solid white lines illustrate the location of the ridge and valley in the landscape. The black arrow indicates the location of the USS. Trajectories of two identical stochastic simulations from the same (0,0) initial conditions are superimposed onto this landscape as solid lines, which are also color-coded to match the corresponding single-cell fluorescence signal as the cell's Lacl and TetR concentrations evolve along the landscape. White arrows illustrate the vector field. Landscapes with different doses of inductions are included in *SI Appendix, Fig. S9*. (B) Simplified schematics of complete inhibition of *GAL1* promoter and upstream activating sequence (UAS)-regulated genes by glucose culturing. (C) After initial growth in glucose (Glu.)-supplemented media to reach log phase, cultures were grown in galactose (Gal.)-supplemented media without ATc induction (pink) and with 250 ng/mL ATc (green) for 48 h before being moved into galactose-supplemented media with 8 ng/mL ATc until steady state. These two cultures showed different levels of fluorescence, suggesting hysteresis, but a similar homogeneous response with a unimodal distribution. Cells moved into the same final condition directly from glucose-supplemented media, however, showed a bimodal distribution (gray curve), suggesting an initially uniform population diverged into two distinct populations with either low- or high-fluorescence output, consistent with stochastic simulation predictions.

long as initial conditions are on or near the separatrix, isogenic cells can randomly settle into different cellular states even with low-level stochasticity.

To study cellular state determination further when initial conditions are not in close proximity to the basin boundary, stochastic simulations were carried out with initial conditions further away from the basin boundaries. From these initial conditions, it is shown that distinct and isogenic cells always settle into the same steady state despite the same model parameters and noise levels (*SI Appendix, Fig. S8*). This prediction is also consistent with experiments in the hysteresis experiments. Flow cytometry data collected in Fig. 1*D* all illustrate unimodal distributions, indicating that all cells homogeneously settled into one state (*SI Appendix, Fig. S10*).

Experimental Validations Exploit Natural Yeast Metabolism Regulatory Mechanisms. The challenge of testing our hypothesis is experimental realization of specific initial conditions on the basin boundary. Imperfect regulation by inducers and some leaky expression of all genes in the network when in galactose-supplemented media make it difficult to realize such initial conditions. Here, we chose to use the natural yeast glucose-galactose metabolism switch mechanism to help us achieve the specific initial condition of no LacI and no TetR in the cell, which is (0,0) in Fig. 2*A* and, based on the model predictions, resides on the basin boundary of the landscape.

The promoter library core to our GRN is based on the *GALI* promoter, which has been characterized extensively by others (38) and has the ability to be tightly inhibited when glucose is present as the only carbon source. In the presence of glucose and the absence of galactose, galactose metabolism in yeast is completely turned off by the natural GAL metabolic regulatory network, which coordinates expression and repression of GAL promoters via upstream activating sequences (Fig. 2*B*). For our GRN, glucose strongly represses all promoters and results in no expression of any gene within our engineered network. As experimentally illustrated in *SI Appendix, Fig. S11*, expression of yEGFP is fully repressed, comparable to blank control, in the presence of glucose as the only carbon source, with significantly reduced detection of fluorescence compared with when it is repressed in galactose-supplemented media without any ATc. Growing our engineered yeast constructs in glucose-supplemented media therefore essentially places these systems on the (0,0) coordinate of LacI and TetR levels, and this will become a point on the basin boundary when the cells are transitioned into galactose-supplemented media with appropriate ATc induction concentrations.

In addition to galactose in the media, it is shown in Fig. 1*D* that a specific range of ATc concentrations is needed for our systems to be bistable. Numerical simulations also suggest that the stochastic cell fate determination is the most pronounced (i.e., with the highest possibility of being experimentally realized) when the separatrix divides the whole LacI–TetR space into two basins with roughly equal areas. Such requirements typically can be achieved through tuning of levels of inducers. Therefore, we chose 8 ng/mL as the ATc concentration to test the hypothesis because it is about one-half of the way between the lower and upper bounds of bistability. TXLX cells initially grown in glucose-supplemented media were washed and directly moved into galactose-supplemented media with 8 ng/mL ATc induction. Flow cytometry measurements were taken after 60 h of growth (gray curve in Fig. 2*C*). It can be seen that isogenic cells from the same initial conditions formed two distinct populations expressing completely different levels of yEGFP, one with low expression and another with high expression. In comparison, cells grown in galactose-supplemented media with and without ATc induction before being moved into galactose-supplemented media plus 8 ng/mL ATc induction only have homogeneous expression of yEGFP (green and pink histograms in Fig. 2*C*), demonstrating

the irreversibility of cell fate choices. Prolonged growth for all three strains was also carried out and further verified the permanency of such cell fate decisions (*SI Appendix, Fig. S12*). Despite the same inductions of galactose plus 8 ng/mL ATc for all three samples at the end, cells chose clearly different states depending completely on their initial conditions. Starting from high yEGFP expression, cells will remain in this state; starting from low yEGFP expression, cells will remain in the low state; and starting from the basin boundary, cells will choose the high- or low-expression state stochastically.

To verify that this stochastic cell fate determination is not dependent on the specifics of the gene network, similar experiments were carried out for two other versions of the yeast bistable gene network: T7L18 and T18LX. As suggested by hysteresis experiments and model predictions, 1.5 ng/mL and 80 ng/mL were chosen for T7L18 and T18LX, respectively, as the ATc dosage to form bistable cell fate landscapes. Starting from (0,0) initial conditions realized by growth in glucose-supplemented media, both networks showed well-pronounced bimodal distribution after growth in galactose with the respective ATc dosages (*SI Appendix, Fig. S13*), suggesting stochastic cell fate choices. This verifies that as long as cells are initialized on the basin boundary of a multistable system, stochastic cell fate determination can be robustly demonstrated. In addition, each network was tested with ATc induction outside the bistability region and showed only unimodal distribution (*SI Appendix, Fig. S13*). This verifies that a proper multistable landscape is also a necessary condition for random cell fate choices. In addition, it can be seen that even with a large abundance of proteins, intracellular stochasticity can be amplified to dictate cell fate due to the nature of initial conditions and the underlying nonlinear system. This complements the common theory of low molecule abundance causing intracellular stochasticity (39).

Temporal Measurements Further Illustrate Unique Dynamics. To verify further that the evolution of fluorescence signals from our GRNs actually follows the trajectories predicted in Fig. 2*A*, both flow cytometry and fluorescence microscopy measurements were taken at different time points after each strain was moved into galactose-supplemented media. It can be seen that flow cytometry measurements of TXLX at 0 h have very low yEGFP expression (0-h flow cytometry in Fig. 3 and black curve in *SI Appendix, Fig. S14*). This is consistent with the microscopy image showing only background fluorescence (imaging details and parameters are provided in *SI Appendix*). This time point represents the initial condition of no LacI or TetR expression in our cells. After 12 h of growth in the right media, it can be seen that the whole population of cells show increased yEGFP expression in a homogeneous fashion (12-h flow cytometry in Fig. 3 and green curve in *SI Appendix, Fig. S14A*). The microscope image (at 12 h in Fig. 3) also illustrates increased signals for both yEGFP and mCherry. Interestingly, many cells emit both green and red fluorescence simultaneously at the single-cell level (labeled by white arrows in Fig. 3). These cells show a wide range of fluorescence levels but cannot be grouped into distinct populations. This is also corroborated by the broad but unimodal distribution of flow cytometry data at 12 h. This time point corresponds to the time point in Fig. 2*A*, where cells are approaching the USS, express an increased amount of both mCherry and yEGFP, and hence result in a homogeneous population. The broad distribution and varied fluorescence levels suggest that cells take different amounts of time to travel from the beginning to the USS due to stochasticity. By the time cells have spent 36 h in the media, the pattern of two populations starts to emerge (36-h flow cytometry in Fig. 3 and blue curve in *SI Appendix, Fig. S14A*). In addition, one of the peaks shows lower yEGFP expression than 24 h earlier. This clearly demonstrates that the cells start to diverge to different steady states. Cells moving

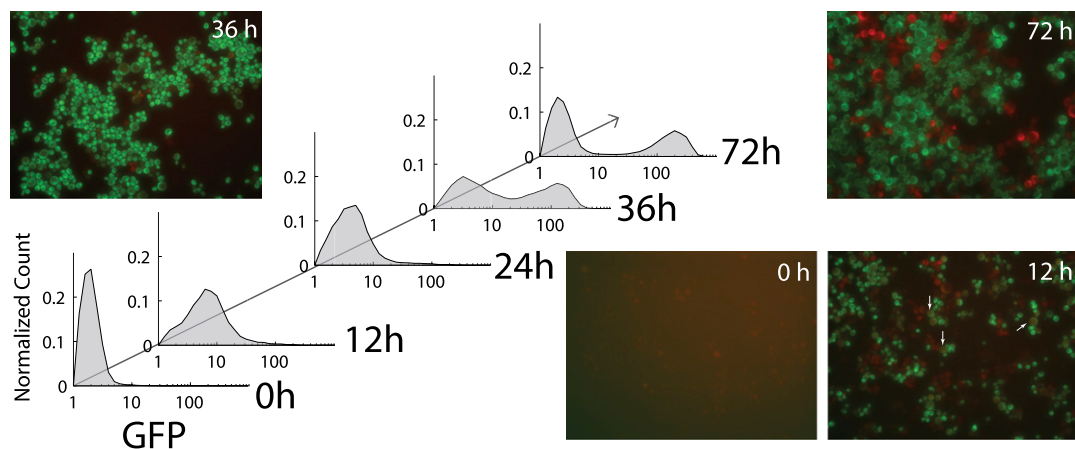


Fig. 3. Temporal dynamics of random cellular state differentiation demonstrated using both flow cytometry and microscopy imaging. Flow cytometry measurements (histograms) were taken at different time points after TXLX cultures were moved directly into galactose-supplemented media with 8 ng/mL ATc media from glucose-supplemented media. The population gradually increases its fluorescence signal homogeneously until after 36 h, when it starts to differentiate into two distinct populations. This observation is consistent with microscopic images taken at these same time points. After 72 h of growth in galactose-supplemented media with 8 ng/mL ATc induction, cultured cells stably differentiated into two distinct populations, as evidenced by bimodal distribution of flow cytometry results and differently colored cells in microscopic images. (Magnification: 40 \times .) White arrows in the 12-h microscopy image point to cells significantly expressing both GFP and mCherry.

toward the LacI-dominant state keep increasing their yEGFP expression, whereas cells moving toward the TetR-dominant state start to produce more mCherry and inhibit production of yEGFP, eventually making its level even smaller than 24 h earlier. This temporal nonmonotonic expression of yEGFP by a subpopulation verifies that cells indeed follow a trajectory of approaching the saddle point and then diverging onto two distinct states. Microscopy imaging is consistent with flow cytometry results, showing cells expressing either yEGFP or mCherry signal. By the time of 72 h of growth in galactose-supplemented media with 8 ng/mL ATc, the cells have clearly formed two populations, illustrated by the two peaks in flow cytometry (72-h flow cytometry in Fig. 3 and red curve in *SI Appendix, Fig. S14A*). Correspondingly, microscopy imaging also shows that cells express either yEGFP or mCherry strongly in a mutually exclusive fashion. The corroboration between flow cytometry and microscopy measurements further supports our predicted temporal dynamics of cells that were started from the (0,0) initial condition on the basin boundary, namely, the transient nonmonotonic expression of both fluorescent proteins and eventual strong but random expression of either one.

Temporal flow cytometry measurements were also carried out for T18LX and T7L18. Both show a transient increase and then decrease of yEGFP expression of a subpopulation (*SI Appendix, Fig. S14 B and C*), verifying that the unique dynamics are not dependent on the strain or network.

Discussion

Combining rational engineering and natural biological regulation, we successfully demonstrate here synthetic stochastic and irreversible cell fate determination in eukaryotic cells. Bistability, stochasticity, and the resulting binary cellular decision making have been extensively studied (2, 3, 5, 6, 8, 10, 40, 41). These studies have illustrated both irreversible and uniform (5, 40, 41) and random yet reversible (6, 39) binary decision making at the single-cell level. However, the case of random and irreversible decision making has not been studied or demonstrated. Guided by stochastic simulations on a nonlinear potential landscape, we experimentally initialized cell cultures on the separatrix of our engineered bistable system and demonstrated stochastic and irreversible binary cellular state determination. Independent cultures grown under the same condition but from different

initial conditions showed completely different responses. This illustrates the complexity of dynamics of multistable gene networks when the effects of initial conditions and stochasticity are taken into account, more than just hysteresis.

With a more complete understanding of multistable gene networks, synthetic GRNs can serve as topological prototypes for their natural counterparts and provide novel insights not easily available through the study of natural systems. By demonstrating stochastic and irreversible cell fate determination, we have been able to shed light on the role of stochasticity in cell differentiation (42, 43). For example, during fruit fly eye development, precursor cells in a specific area of the eye differentiate into cells with two types of photoreceptor and maintain their cell fate. It has also been reported that the differentiation of these photoreceptors is purely stochastic and independent (11). Recent studies (12, 13) have identified a mutual inhibitory gene regulation motif as the core system driving the stochastic differentiation, but the exact mechanism of the stochastic differentiation is still unknown. The similarities between this natural system and our engineered system, both in terms of observation and underlying GRN, suggest that stochastic and irreversible cellular developments in the fruit fly eye could be due to initialization of cells near basin boundaries of a multistable network. The initialization may well be regulated by a metabolic event, like in our system, or by other mechanisms, such as epigenetics or micro-RNA regulation.

Finally, this work also demonstrates the power of linking synthetic GRNs to the outputs of the host cell's natural GRNs. Much effort in synthetic biology has been focused on using synthetic GRNs to drive natural GRNs toward desired responses (27, 44). Here, by linking outputs of natural metabolism regulations to synthetic GRNs, we were able to realize an initial condition that is difficult to achieve through engineering alone. Such accurate initiation of the cell culture is the key to our demonstrated stochastic and random cell fate determination. This concept of harvesting natural regulatory machineries to tune synthetic GRNs will greatly increase the quality and quantity of possible perturbations that can be applied to engineered systems, hence making it possible to engineer future biological devices that require more sophisticated and accurate controls.

Materials and Methods

Yeast Strains and Plasmid Constructions. Three yeast strains were used in the experiments and contained the TXLX, T18LX, or T7L18 network with the 714 red/green reporter construct. All were integrated into *Saccharomyces cerevisiae* strain YPH500 (a, *ura3-52*, *lys2-801*, *ade2-101*, *trp1D63*, *his3D200*, *leu2D1*) (Stratagene) with genomic integrations specifically targeted to the *ura3-52* locus (8). These strains produce yEGFP from the TX promoter and mCherry red fluorescent protein from the LX promoter (*SI Appendix, Fig. S1*). The plasmid construction methods and plasmid maps are described in detail in *SI Appendix*.

Flow Cytometry. Flow cytometry data acquisition was performed with a Becton Dickinson FACScan Analyzer. This machine is equipped for GFP measurements. The detector for forward scatter (FSC) used an E00 channel with side scatter (SSC) at 378 V and fluorescence channel 1 (FL1) at 436 V. All data were collected in a log mode. Samples were carried out at a medium flow rate until 100,000 cells had been collected. Data files were analyzed

using MATLAB (MathWorks) with gating. For the bistable region determination experiments, the fluorescence levels were monitored every 24 h until they became stable. For the stochastic cellular state determinations, samples were taken every 12 h for measurement until the fluorescence levels became stable.

Mathematical Modeling. Ordinary differential equation (ODE) models were solved and analyzed using MATLAB and XPPAUT (45). Stochastic simulations were written in C and run on a standard personal computer (details are provided in *SI Appendix*).

ACKNOWLEDGMENTS. We thank William Blake for stimulating discussions and Pamela Marshall and Joseph Chao for technical assistance. We also thank Jim Collins for yeast strains and helpful comments on the manuscript. This study was financially supported by National Science Foundation (NSF) Grants DMS-1100309 (to X.W.) and CDI-1026710 (to Y.-C.L.) and by American Heart Association Grant 11BGIA7440101 (to X.W.).

- Ferrell JE, Jr. (2002) Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* 14(2):140–148.
- Xiong W, Ferrell JE, Jr. (2003) A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* 426(6965):460–465.
- Biggar SR, Crabtree GR (2001) Cell signaling can direct either binary or graded transcriptional responses. *EMBO J* 20(12):3167–3176.
- Nevozhay D, Adams RM, Van Itallie E, Bennett MR, Balázs G (2012) Mapping the environmental fitness landscape of a synthetic gene circuit. *PLoS Comput Biol* 8(4):e1002480.
- Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435(7039):228–232.
- Becskei A, S raphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: Cell differentiation by graded to binary response conversion. *EMBO J* 20(10):2528–2535.
- Balázs G, van Oudenaarden A, Collins JJ (2011) Cellular decision making and biological noise: From microbes to mammals. *Cell* 144(6):910–925.
- Ellis T, Wang X, Collins JJ (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* 27(5):465–471.
- Louis M, Becskei A (2002) Binary and graded responses in gene networks. *Sci STKE* 2002(143):pe33.
- Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403(6767):339–342.
- Bell ML, Earl JB, Britt SG (2007) Two types of *Drosophila* R7 photoreceptor cells are arranged randomly: A model for stochastic cell-fate determination. *J Comp Neurol* 502(1):75–85.
- Mikeladze-Dvali T, et al. (2005) The growth regulators warts/lats and melted interact in a bistable loop to specify opposite fates in *Drosophila* R8 photoreceptors. *Cell* 122(5):775–787.
- Miller AC, Seymour H, King C, Herman TG (2008) Loss of seven-up from *Drosophila* R1/R6 photoreceptors reveals a stochastic fate choice that is normally biased by Notch. *Development* 135(4):707–715.
- Milo R, et al. (2002) Network motifs: Simple building blocks of complex networks. *Science* 298(5594):824–827.
- Mukherji S, van Oudenaarden A (2009) Synthetic biology: Understanding biological design from synthetic circuits. *Nat Rev Genet* 10(12):859–871.
- Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297(5584):1183–1186.
- Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nat Genet* 31(1):69–73.
- Raj A, Rifkin SA, Andersen E, van Oudenaarden A (2010) Variability in gene expression underlies incomplete penetrance. *Nature* 463(7283):913–918.
- Blake WJ, KAern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. *Nature* 422(6932):633–637.
- Nevozhay D, Adams RM, Murphy KF, Josic K, Balázs G (2009) Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc Natl Acad Sci USA* 106(13):5123–5128.
- Blake WJ, et al. (2006) Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* 24(6):853–865.
- Isaacs FJ, Hasty J, Cantor CR, Collins JJ (2003) Prediction and measurement of an autoregulatory genetic module. *Proc Natl Acad Sci USA* 100(13):7714–7719.
- Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403(6767):335–338.
- Danino T, Mondrag n-Palomino O, Tsimring L, Hasty J (2010) A synchronized quorum of genetic clocks. *Nature* 463(7279):326–330.
- Mondrag n-Palomino O, Danino T, Selimkhanov J, Tsimring L, Hasty J (2011) Entrainment of a population of synthetic genetic oscillators. *Science* 333(6047):1315–1319.
- Stricker J, et al. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* 456(7221):516–519.
- Levkaya A, Weiner OD, Lim WA, Voigt CA (2009) Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461(7266):997–1001.
- Bashor CJ, Helman NC, Yan S, Lim WA (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* 319(5869):1539–1543.
- You L, Cox RS, 3rd, Weiss R, Arnold FH (2004) Programmed population control by cell-cell communication and regulated killing. *Nature* 428(6985):868–871.
- Balagadd  FK, et al. (2008) A synthetic *Escherichia coli* predator-prey ecosystem. *Mol Syst Biol* 4:187.
- Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R (2005) A synthetic multicellular system for programmed pattern formation. *Nature* 434(7037):1130–1134.
- Tabor JJ, et al. (2009) A synthetic genetic edge detection program. *Cell* 137(7):1272–1281.
- Dari A, Kia B, Wang X, Bulsara AR, Ditto W (2011) Noise-aided computation within a synthetic gene network through morphable and robust logic gates. *Phys Rev E Stat Nonlin Soft Matter Phys* 83(4 Pt 1):041909.
- Bhattacharya S, Zhang Q, Andersen ME (2011) A deterministic map of Waddington's epigenetic landscape for cell fate specification. *BMC Syst Biol* 5:85.
- Kim KY, Wang J (2007) Potential energy landscape and robustness of a gene regulatory network: Toggle switch. *PLoS Comput Biol* 3(3):e60.
- Strogatz SH (1994) *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Addison-Wesley, Reading, MA).
- Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 22(4):403–434.
- Johnston M (1987) A model fungal gene regulatory mechanism: The GAL genes of *Saccharomyces cerevisiae*. *Microbiol Rev* 51(4):458–476.
- Kaern M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: From theories to phenotypes. *Nat Rev Genet* 6(6):451–464.
- Yao G, Lee TJ, Mori S, Nevins JR, You L (2008) A bistable Rb-E2F switch underlies the restriction point. *Nat Cell Biol* 10(4):476–482.
- Ajo-Franklin CM, et al. (2007) Rational design of memory in eukaryotic cells. *Genes Dev* 21(18):2271–2276.
- Pedraza JM, van Oudenaarden A (2005) Noise propagation in gene networks. *Science* 307(5717):1965–1969.
- Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* 135(2):216–226.
- Prindle A, et al. (2012) A sensing array of radically coupled genetic 'biopixels'. *Nature* 481(7379):39–44.
- Ermentrout B (2002) *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students* (Society for Industrial and Applied Mathematics, Philadelphia).