Role of feedback and dynamics in a gene regulatory network

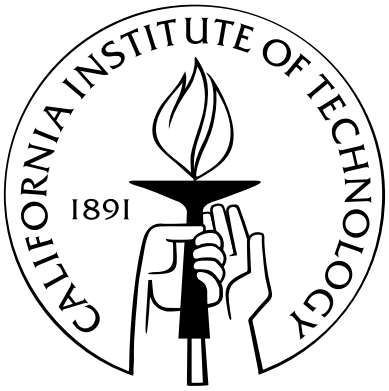
Thesis by

Ophelia S. Venturelli

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 30, 2013)

Acknowledgements

I am very fortunate to have had the opportunity of attending Caltech for graduate school. I appreciate many unique characteristics of Caltech including an intimate community where I can walk around campus and recognize many familiar faces, a relaxed yet challenging and stimulating intellectual environment and the absence of barriers for interdisciplinary research. I also had the privilege of spending almost two years at UCSF during my graduate studies. UCSF is a cutting-edge and exciting place to work and I have learned a tremendous amount during my tenure at UCSF.

I would like to especially thank my advisor Richard Murray for his guidance, support and for allowing me the freedom to explore interesting scientific questions. Richard has an incredible positive energy and enthusiasm that has reinforced my motivation about science over the years. By working with Richard, I have learned a quantitative and rigorous reasoning about biological systems. I have enjoyed interacting with members of the Murray group and appreciate their diverse backgrounds in math, computer science, engineering and biology.

I am most grateful to Hana El-Samad, who has been a second advisor to me, for the opportunity to pursue exciting research projects in her lab and for supporting my research. Hana is very detail-oriented and meticulous about scientific research and as a consequence I think my approach to scientific problems and writing has significantly improved. Hana also has excellent intuition for how to tie together a set of data to tell a compelling scientific story. Finally, Hana provided me with an enjoyable and intellectually stimulating environment. I have learned many lessons in her lab that will be indispensable for my future scientific career.

I would like to credit many colleagues at both Caltech and UCSF for their guidance, support and scientific feedback. I would like to especially recognize Ignacio Zuleta for our productive collaboration. I have greatly benefited from countless scientific discussions with Fiona Chandra, Elisa Franco and Marcella Gomez. I am appreciative for Andrés Aranda's help in the lab. My research has significantly benefited from my collaboration with Adam Rosenthal. I am thankful to David Soloveichik for helping me to think differently about theoretical research problems. Richard's group recently exhibited an period of exponential growth and I would like to thank the new members of his lab including Dan Siegal-Gaskins, Anu Thubagere, Jongmin Kim, Nadine Dabby, Emzo de los Santos, Joe Meyerowitz, Enoch Yeung and Vanessa Jönsson. Finally, I would like to extend my appreciation to other members of Hana's lab including Benjamin Heineike, Raj Bhatnagar, Michael Chevalier, Joanna Lipinski-Kruszka, Susan Chen, Patrick Harrigan and Graham Heimberg.

I am especially grateful to my committee including Michael Elowitz, Rob Phillips and Raymond Deshaies for providing valuable feedback. In particular, I am indebted to Michael Elowitz who allowed me to participate in his lab's group meetings and provide excellent scientific guidance. Michael's approach to biology has been an inspiration for my graduate work. I am thankful to Steve Mayo for providing me with lab space at Caltech.

I would like to recognize the Biochemistry and Molecular Biophysics (BMB) graduate program and especially thank Allison Ross and Doug Rees in organizing this program and supporting graduate students.

I would not be where I am today without the love and support of my family. My parents, Shalini and Peter Venturelli have provided me with unconditional support and I would like to thank them for always believing in me. My brother, Philip Venturelli, instilled in me a deep scientific curiosity. I am obliged to Louis and Shelley Romero for their assistance during my graduate studies. In

particular, my research has benefited significantly from many instructive conversations with Louis Romero about math and dynamical systems. It has been really fun to share my graduate experiences with my husband, Philip Romero. Philip's scientific feedback and support has been invaluable.

Abstract

Cells exhibit a diverse repertoire of dynamic behaviors. These dynamic functions are implemented by circuits of interacting biomolecules. Although these regulatory networks function deterministically by executing specific programs in response to extracellular signals, molecular interactions are inherently governed by stochastic fluctuations. This molecular noise can manifest as cell-to-cell phenotypic heterogeneity in a well-mixed environment. Single-cell variability may seem like a design flaw but the coexistence of diverse phenotypes in an isogenic population of cells can also serve a biological function by increasing the probability of survival of individual cells upon an abrupt change in environmental conditions. Decades of extensive molecular and biochemical characterization have revealed the connectivity and mechanisms that constitute regulatory networks. We are now confronted with the challenge of integrating this information to link the structure of these circuits to systems-level properties such as cellular decision making. To investigate cellular decision-making, we used the well studied galactose gene-regulatory network in *Saccharomyces cerevisiae*. We analyzed the mechanism and dynamics of the coexistence of two stable on and off states for pathway activity. We demonstrate that this bimodality in the pathway activity originates from two positive feedback loops that trigger bistability in the network. By measuring the dynamics of single-cells in a mixed sugar environment, we observe that the bimodality in gene expression is a transient phenomenon. Our experiments indicate that early pathway activation in a cohort of cells prior to galactose metabolism can accelerate galactose consumption and provide a transient increase in growth rate. Together these results provide important insights into strategies implemented by cells that may have been evolutionary advantageous in competitive environments.

Contents

Chapter 1

Introduction

Introduction

Biological systems exhibit remarkable phenotypic diversity. The myriad of phenotypes are achieved through networks of interacting biomolecules that produce a rich repertoire of dynamical functions at the cellular level. For example, these circuits have been shown to produce oscillations, excitability, adaptive responses and bistability [1, 2, 3, 4]. Due to the discrete nature of molecular interactions, the dynamic responses are inherently stochastic and display significant fluctuations in the concentrations of components. Indeed, single microbial and eukaryotic cells can exhibit significant phenotypic heterogeneity that does not stem from a genetic origin.

Although the inescapable noise in the levels and activities of biomolecules seems undesirable for cells, previous studies have shown that phenotypic variability can be functionally beneficial in specific environments. This bet-hedging phenomenon is characterized by more than one coexisting phenotype in a uniform environment that provides a temporary disadvantage for the population but can confer a long-term fitness advantage upon an abrupt environmental change [5]. In bacteria for example, competence, sporulation, and persistence in response to high doses of antibiotics have been shown to enhance population fitness by increasing the chance of survival upon a shift in the environmental conditions [6, 7, 8]. In all of these cases, noise in the levels of critical regulatory molecules influences a binary cellular decisions. Recently, yeast have also been shown to benefit from bet-hedging strategies by diversifying the range of growth rates or displaying asymmetric

growth behaviors in response to specific stress stimuli such as metal deprivation or heat [9, 10].

In several cases, the potential for significant phenotypic variability has been shown to arise from the feedback structure of the regulatory network [11, 12, 13, 2]. Feedback loops are defined as molecular interactions that link the output of a system back to the input [14]. These loops ubiquitous regulatory features of biological networks and can significantly modulate the dynamics and function of circuits. For example, negative feedback can quantitatively shape cellular responses by enhancing the system's response time, reducing phenotypic variability and generating transient dynamic behaviors [15, 16, 17]. Positive feedback loops can provide signal amplification and induce bistability if the positive feedback loop is coupled to a sufficiently ultrasensitive mechanism [4, 18, 19]. Natural biological circuits frequently contain many feedback loops and it is a challenging task to disentangle the roles of individual loops and understand how their activities are coordinated in a densely connected network [13].

In this work, we have explored the role of feedback loops, dynamics and biological function of single-cell variability using the galactose gene-regulatory network (GAL) in *Saccharomyces cerevisiae* (*S. cerevisiae*) as a model system. The galactose regulatory network is a very well analyzed eukaryotic gene circuit that provides cells with the capability to metabolize the galactose as a carbon source. Extensive biochemical and molecular characterization has elucidated the key molecular events that enable cells to turn this metabolic switch on and off in response to changes in environmental signals. Here, we built upon this foundation to understand how these complex molecular interactions can combine to produce system level properties. This gene regulatory network has two interesting behaviors: single-cells can exhibit coexisting all-or-none network activity for intermediate concentrations of galactose or combinations of glucose and galactose [19, 20, 18] and small variations in concentration of galactose can generate a large fold change in the downstream enzyme levels, referred to as ultrasensitivity. Our analysis of this system identified the mechanism that generates the switch-like bistable response and revealed how this bimodal strategy can be advantageous for a population of cells in a combinatorial environment.

In Chapter 2, we use a combination of experiments and computational modeling to analyze the

roles of the feedback loops on the bimodal response of the GAL system. We demonstrate that two positive feedback loops established by the bifunctional galatokinase and signal transducer Gal1p and signal transducer Gal3p collaborate to induce bistability in the system. Our computational analysis identifies molecular sequestration as a critical mechanism for generating the ultrasensitive stage necessary for robust bistability in the system.

In Chapter 3, we probe how the GAL network integrates two signals of glucose and galactose to generate a dynamic transcriptional response. By measuring single-cell dynamics over a long time scale, we observe transient bimodality for conditions of similar concentrations of the two sugars. In fact, after many cell generations, all cells in the population converge to a single monomodal on-state. We construct a computational model that captures the structure of the network with two inputs. Our analysis of the model reveals that the observed transient bimodality originates from bistability that vanishes when glucose is depleted beyond a threshold. Sugar measurements indicate that the delayed turn-on of the repressed subpopulation of cells occurs precisely when cells begin to consume galactose following glucose depletion. The early activation of the GAL genes in a fraction of cells is shown to reduce the transition time between carbon sources and provides a transient enhancement of cellular growth rate. Our experiments also reveal a cost to constitutive GAL gene expression that arises from a reduced glucose consumption rate, thus highlighting the intricate tradeoffs involved in the timing of GAL gene induction.

In Chapter 4, we investigate the mapping between promoter sequence and expression level using the bidirectional *GAL1-10* promoter. We construct a library of randomly mutated promoter sequences and model the effects of these mutations with a statistical model to identify the relative contributions of single nucleotides to the observed expression levels. By measuring the expression levels of both directions of this bidirectional promoter, our results indicate that individual nucleotides can differentially tune the promoter activity. Experimental characterization of critical nucleotides displays a strong correlation with the model's prediction, highlighting the predictive capabilities of the model.

In Chapter 5, using a computational model of the GAL network, we analyze the roles of the feed-

back loops and sequestration on a set of phenotypes including bistability, ultrasensitivity, switching threshold, dynamic range, response time and tunability.

Chapter 2

Dual feedback loops established by molecular sequestration generates robust bimodal response

A version of this chapter has been published as [19].

Introduction

Cells are continuously faced with the challenge of sensing signals in their environment and eliciting intracellular programs accordingly. While changes in some environmental cues engender graded and proportional responses, others induce decisive action whereby a cell exhibits a binary (on or off) phenotypic change. In the latter case, amplification of phenotypic heterogeneity may arise since single cells in a population make individual decisions based on their perception of the environmental stimulus, stochastic fluctuations in their molecular components, and memory of past conditions. This thresholded cellular response can manifest as a bimodal distribution in network activity across an isogenic cell population.

Feedback regulation, which links the output of a circuit back to its input, expands the set of possible biological properties, including robustness to uncertainty [14] and can produce single cell phenotypic heterogeneity in a uniform environment. Many features of individual positive and negative feedback loops have been elucidated, including enhancement of response time and reduction of gene expression noise by negative autoregulation, and signal amplification and bistability using

positive autoregulation [15, 16, 21, 13]. However, quantitative characterization of how multiple feedback pathways interact to regulate and fine-tune cellular decision-making presents many unresolved challenges.

The galactose gene-regulatory network *S. cerevisiae* (GAL) contains numerous feedback pathways. Isogenic single cells respond heterogeneously to a range of galactose concentrations, which manifests as a bimodal distribution of GAL gene expression across the cell population [18]. In contrast to a graded response, in which the mean of a unimodal distribution is continuously adjusted as the input is modulated, variations in the concentration of galactose within a range shifts the fraction of the cell population distributed between distinct metabolic states. Here, we focused on how the multiple feedback loops in the system shape this bimodal cellular decision-making strategy in response to galactose.

The GAL circuit consists of regulatory machinery (Gal2p, Gal3p, Gal80p, Gal4p) that dictates network activity and a set of enzymes required for metabolizing galactose (Gal1p, Gal7p, Gal10p). In the absence of galactose, GAL genes are repressed due to the sequestration of the potent transcriptional activator Gal4p by the repressor Gal80p (Fig. 2.1). [22]. In the presence of galactose, the membrane-bound permease transporter Gal2p significantly increases the rate of galactose uptake from the extracellular environment [23]. Galactose and ATP-dependent activation of the signal transducer Gal3p lead to repression of Gal80p by sequestration, thus liberating Gal4p [24]. The galactokinase Gal1p catalyzes the first step in galactose metabolism by phosphorylating galactose to form galactose 1-phosphate and has been shown to possess weak co-inducing functionality [25].

Galactose-dependent regulation of Gal2p, Gal3p and Gal80p form feedback loops because these proteins modulate network activity and are themselves transcriptionally regulated by Gal4p [27]. Gal2p and Gal3p form positive feedback loops since up-regulation of their expression levels leads to an increase in pathway activity, whereas Gal80p reduces pathway activity and thus forms a negative feedback loop.

In addition to Gal2p, Gal3p and Gal80p, there is evidence to suggest that Gal1p has a regulatory role beyond its vital enzymatic function for growth on galactose [28, 25, 29]. Gal1p is a close homolog

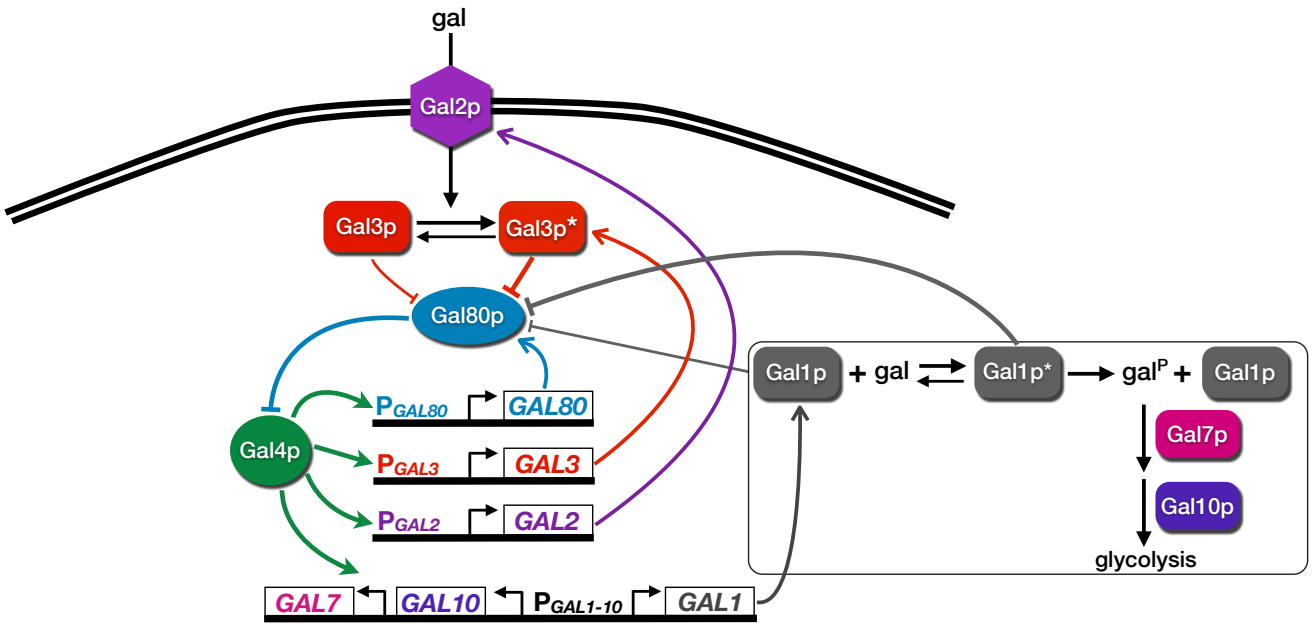


Figure 2.1: The galactose gene-regulatory network in *S. cerevisiae*. The permease Gal2p facilitates intracellular galactose transport. By binding to galactose, the signal transducer Gal3p becomes highly activated to sequester the transcriptional repressor Gal80p. In the absence of galactose, Gal3p can also inhibit Gal80p, presumably with lower affinity, leading to GAL gene induction [25]. Repression of Gal80p liberates the transcriptional activator Gal4p to up-regulate a set of target enzymatic and regulatory genes. A series of enzymatic reactions (interactions inside box) transforms galactose into glucose-6-phosphate for glycolysis through the activities of the galactokinase Gal1p, transferase Gal7p, and epimerase Gal10p. The regulatory proteins Gal2p, Gal3p and Gal80p, form positive, positive and negative feedback loops, respectively. Gal1p, a paralogue of Gal3p, has been shown to possess bifunctional activities by sequestering Gal80p in the presence and absence of galactose with different affinities, leading to GAL gene activation [25, 26]. *GAL1* and *GAL10* share a bidirectional promoter ($P_{GAL10\text{-}1}$).

of Gal3p and has been shown to interact with Gal80p with a weaker affinity than Gal3p [30, 31]. Furthermore, a *GAL3* deletion strain was shown to induce GAL gene expression at a significantly slower rate compared to wild-type whereas cells with combined *GAL1* and *GAL3* deletions fail to activate their GAL pathway [32]. A recent study demonstrated that cells initially grown in galactose and then transferred to glucose exhibit a faster induction response to a second galactose exposure than cells grown only in glucose, and that Gal1p was critical for this decrease in response time [33]. Finally, galactose induction was shown to consist of two stages, the first of which is dominated by rapid association of Gal3p to Gal80p and a delayed second stage consisting of dominance of the Gal1p-Gal80p complex [34].

In this paper, we use a combination of experimental measurements and computational modeling to demonstrate that the observed bimodality in the galactose metabolic pathway arises from an underlying bistability in the system and that this bimodal response relies on the synergistic interplay of the *GAL1* and *GAL3* feedback loops. These central mediators have unique mechanistic roles in the GAL system since they both regulate circuit activity by competitive molecular sequestration of Gal80p. While the bimodal response can be transformed into a graded response in the absence of the individual *GAL1* and *GAL3* feedback loops, this only occurs in a specific parameter regime in which the constitutive production rates of Gal1p and Gal3p are greater than a threshold. A mathematical model recapitulates the experimental results and provides crucial insights about the roles of the autoregulatory loops on bistability. More broadly, a simple mathematical model is used to identify generalizable properties of positive feedback loops created by molecular sequestration that implement robust switch-like responses.

Results

History-dependent response indicates that bimodality arises from underlying bistability and Gal1p significantly enhances sensitivity to galactose

The presence of bimodality does not necessarily imply bistability since a bimodal distribution can arise from stochastic effects [35, 36, 37]. Hysteresis is a characteristic feature of bistability, in which the system jumps from one branch of stable steady-states to a different branch of steady-states as a parameter is continuously increased, but jumps from the second branch of steady-states back to the first branch at a different value of the parameter as it is continuously decreased. This behavior stems from a difference in the local stability of multiple stable equilibria. To determine if bimodality in the GAL system was linked to bistability, we checked for a history-dependent response, which is an indicator of local equilibrium point stability. The bistable stochastic counterpart of a deterministic bistable system may not exhibit hysteresis due to an insufficient time-scale separation and a deterministic system can be bistable without displaying hysteresis [38, 39]. Here, we tested for a stochastic system that exhibits hysteresis, which would be consistent with an underlying bistability in a deterministic model of the system. Distinguishing whether bimodality arises from stochastic interactions or a deterministic bistability provides critical information about the operation of the system, including the types of molecular interactions that might be underlying this response and suggests a mathematical modeling framework for studying this phenotype.

We investigated the GAL system's history-dependent response by comparing the stability of its high and low metabolic states as a function of galactose. To measure relative expression state stability, we used flow cytometry to quantify the fluorescence distributions of a genome integrated *GAL10* promoter fusion to Venus (YFP) in wild-type (WT) single cells as an indicator of network activity [40] (P_{GAL10}Venus). The cells were grown first in the presence (E_H) and absence (E_L) of 2% galactose in 2% raffinose media. Cells from the two environments were then shifted to a second set of environments containing a wide range of galactose concentrations.

A history-dependent response existed if cell populations grown in the two environments (E_L

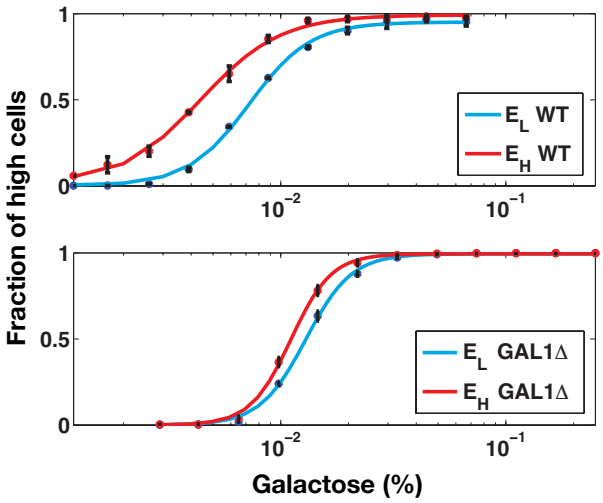


Figure 2.2: History response experiment indicates that GAL bimodal response arises from underlying bistability. **(A)** Conceptual diagram of the history response experiment in which the depth of the potential wells (expression level vs. stability) can be controlled by a bifurcation parameter (galactose concentration). In this experiment, isogenic cells were grown in two environments until steady-state, E_H and E_L. In E_H (2% galactose and 2% raffinose), the high expression state (H) has a lower potential energy and is strongly favored whereas the low expression state (L) is favored in E_L (2% raffinose). Cells are then transferred from the two initial environments to a new set of environments (E_1, ..., E_n) containing a range of galactose concentrations. A history-dependent response was present if cells from E_H and E_L were distributed differently between the high and low states for a range of galactose concentrations after 30 hours of induction. **(B)** History-dependent response experiment indicates that the GAL network is bistable. Experimental data showing history response region for wild-type (WT) cells (top) following the experimental procedure outlined in **A**. Cells that do not metabolize galactose (GAL1Δ) also displayed a history-dependent response (bottom). Each data point is the mean of the fraction of cells in the high expression state and the error bars represent one standard deviation (n=3). Lines represent fitted Hill functions.

and E_H) had a different fraction of cells distributed between the high and low expression states in a range of galactose concentrations after approximately 10 cell divisions post shift (30 hours). A 30-hour induction period was selected to allow a sufficient number of cell divisions for dilution of the fluorescent reporter from E_H cells (Section S2.1). Within a range of galactose concentrations, cells from E_L and E_H were distributed differently between the high and low metabolic states (Fig. 2.2), revealing a history-dependent response and corroborating the existence of bistability.

To exclude the possibility that the difference in the thresholds of the dose responses was due to variable consumption of galactose, the history-response experiment was performed using a *GAL1* deletion strain that is incapable of metabolizing galactose (GAL1Δ) [41]. The GAL1Δ strain was used since cells with gene deletions for the transferase *GAL7* and epimerase *GAL10* are unable to grow in the presence of galactose due to the toxic accumulation of phosphorylated galactose [42].

Investigation of history-dependence in the GAL1Δ strain revealed that its dose response threshold was approximately twofold higher than WT, demonstrating that Gal1p significantly contributes to galactose sensitivity. The GAL1Δ cells also exhibited a history-dependent difference in the galactose threshold. However, the area separating the activation response curves for GAL1Δ was smaller than WT, indicating a diminished history-dependent response. Taken together, these data corroborate bistability as the source of bimodality in the response of the GAL network to galactose and strongly suggest that Gal1p plays an important regulatory role in addition to its metabolic function.

Combined deletion of the *GAL1* and *GAL3* feedback loops produces a graded response, demonstrating the unique role of Gal1p and Gal3p in generating bistability

To further explore Gal1p as a regulatory component of the system and evaluate its role relative to the other autoregulatory loops, we constructed a series of feedback loop deletions involving different components of the system. To do so, we deleted the coding region of a given gene and integrated a single copy of this gene regulated by an inducible *TET* promoter or a constitutive promoter. The rate of production from the *TET* promoter could be adjusted by a doxycycline (dox) dependent

activation of rtTA, a reverse mutant of the transcription factor, TetR [43]. In this fashion, the expression of the gene involved in the feedback loop can be decoupled from the activity of the galactose pathway since the regulation of the constitutive or inducible promoter is external to the GAL regulatory circuit.

In order to compare the operation of the WT system and the different feedback mutants on equal footing, we selected the strength of constitutive expression of each gene by mapping it to the corresponding WT expression levels using quantitative real-time PCR (qPCR) (Table 1 and Supplementary Fig. 2.3). We also explored a range of *TET* promoter expression levels by scanning different dox concentrations to investigate the relationship between constitutive expression of each regulatory component and the steady-state dose response. The fluorescence distributions were classified as unimodal or bimodal using a Gaussian mixture model threshold (GMM) (see Materials and methods).

Eliminating the *GAL2* or *GAL80* feedback loops did not abolish the GAL system's bimodal response (Fig. 3B-1 and B-2). Instead, bimodality persisted for a range of expression levels for Gal2p and Gal80p (Supplementary Fig. 2.5B,C). Compared to WT, cells with a deleted *GAL80* feedback loop (*GAL80*Δ fb) displayed bimodality for a larger number of galactose concentrations. Contrary to a previous study [20], we observed that the *GAL3* feedback loop was not necessary for bimodality for WT expression levels of Gal3p (Fig. 3A-4 and Supplementary Fig. 2.2). However in the GAL3Δ fb cells, the bimodal response could be transformed into a graded response by driving the rate of constitutive Gal3p production beyond a critical threshold (Fig. 5B). We found that the discrepancy with the previous study [20] can be explained by constitutive Gal3p expression above this threshold (Section S2.2).

Since the *GAL2*, *GAL3* and *GAL80* feedback loops were not individually necessary for bimodality, we hypothesized that they either play compensatory roles or that bimodality relies on yet another uncharacterized feedback loop. To address the possibility that the feedback loops had overlapping or compensatory functions, we constructed combinations of feedback loop deletions of *GAL2*, *GAL3* and *GAL80* by constitutively expressing them from the *ADH1*, *TET* and *STE5* promoters, respec-

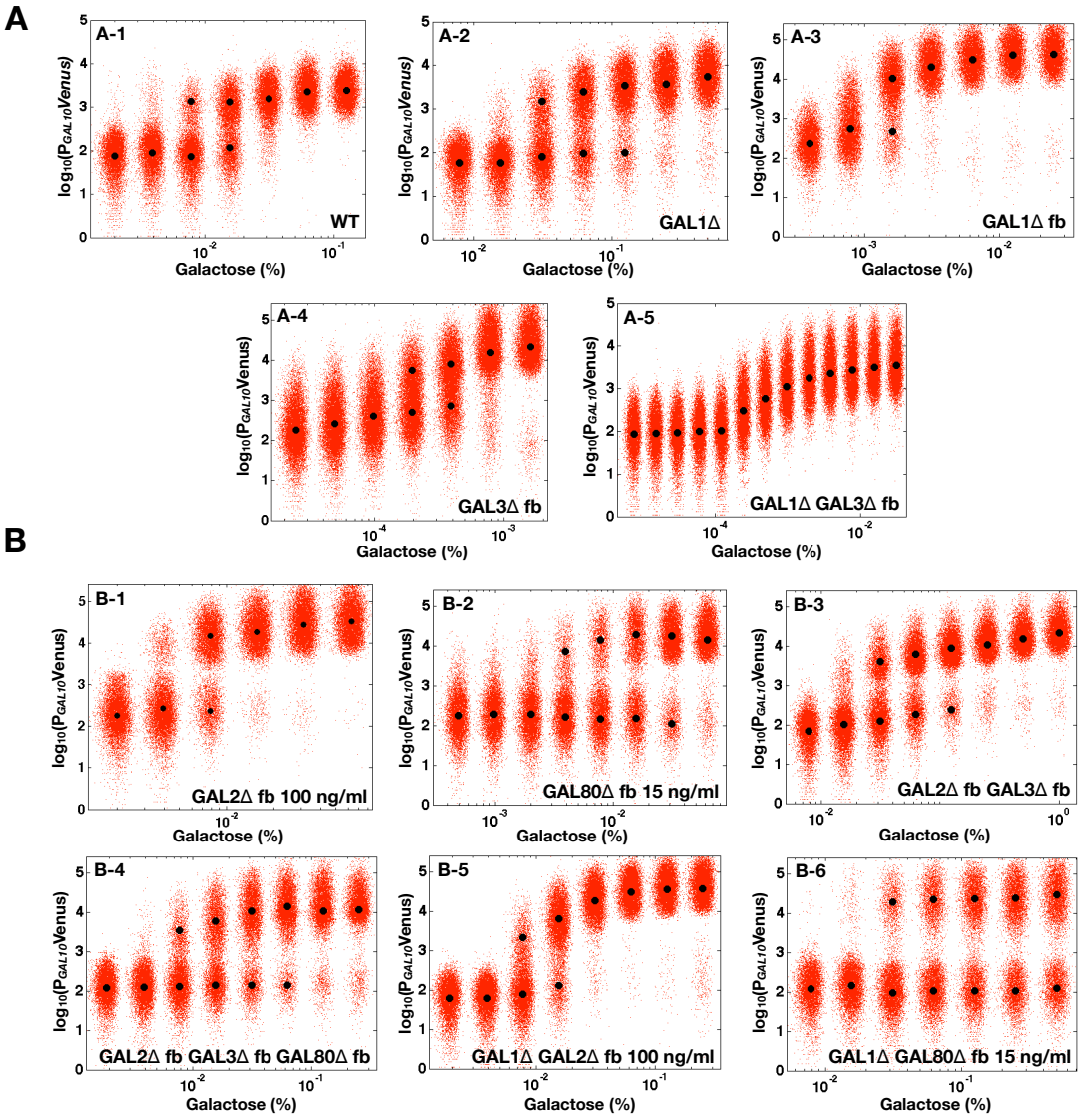


Figure 2.3: Double deletion of *GAL1* and the *GAL3* feedback loop abolishes bimodality. Representative steady-state flow cytometry data of P_{GAL10}Venus in wild-type (WT) and a set of single and multiple feedback loop deletions induced with a range of galactose concentrations. Each black circle indicates the mean of the distribution determined by a Gaussian mixture model (see Materials and methods). Small random deviations were added to each galactose concentration to highlight the spread of the fluorescence distributions. **(A)** Either the *GAL1* or the *GAL3* feedback loop is required for bimodality. The wild-type (WT), *GAL1* deletion (GAL1Δ), *GAL1* feedback deletion (GAL1Δ fb) and *GAL3* feedback deletion (GAL3Δ fb) strains displayed bimodality for at least one galactose concentration. GAL1Δ fb and GAL3Δ fb were not induced with doxycycline (dox). Eliminating the *GAL3* feedback loop in the absence of *GAL1* (GAL1Δ GAL3Δ fb) produced a graded response for the full range of galactose. **(B)** Bimodality was preserved for a series of feedback loop disruptions. The single *GAL2* (GAL2Δ fb) and *GAL80* (GAL80Δ fb) loop deletions were induced with 100 and 15 ng/ml dox, respectively. Bimodality persisted for a dual feedback loop disruption of *GAL2* and *GAL3* (GAL1Δ fb GAL3Δ fb) and a triple feedback deletion of *GAL2*, *GAL3* and *GAL80* (GAL2Δ fb GAL3Δ fb GAL80Δ fb). For these two strains, *GAL2*, *GAL3* and *GAL80* were expressed from an *ADH1*, *TET* and *STE5* promoter, respectively in the absence of dox. Deleting the *GAL2* (GAL1Δ GAL2Δ fb) and *GAL80* (GAL1Δ GAL80Δ fb) feedback loops individually in a strain lacking *GAL1* preserved bimodality. GAL1Δ GAL2Δ fb and GAL1Δ GAL80Δ fb were induced with 100 and 15 ng/ml dox, respectively.

tively. Remarkably, bimodality was preserved in the absence of both the *GAL2* and *GAL3* feedback loops (GAL2Δ fb GAL3Δ fb) and also in a triple feedback loop deletion strain of *GAL2*, *GAL3* and *GAL80* (Fig. 3B-3 and B-4).

Therefore, combinations of *GAL2*, *GAL3* and *GAL80* feedback loops did not functionally overlap to create bimodality. Since Gal1p regulates both sensitivity and memory of the GAL network to galactose (Fig. 2.2), we explored the possibility that Gal1p could be an important component of the system's bimodality.

In contrast to Gal3p and Gal80p transcriptional regulation, Gal1p is tightly repressed in the absence and strongly induced in the presence of galactose. As a consequence, matching the open and closed loop production rates using the *TET* promoter was challenging. Similar to Gal3p, Gal1p has been shown to activate GAL genes independently of galactose, and a sufficiently strong constitutive Gal1p production rate could shift the operating point of the network [25]. We first explored the lowest regime of Gal1p expression using a *GAL1* gene deletion (GAL1Δ) and bimodality was detected in this strain for several galactose concentrations (Fig. 2.3A-2). The *GAL1* feedback loop deletion, P_{TET}GAL1 (GAL1Δ fb) was also bimodal in the absence of dox for at least one galactose concentration (Fig. 2.3A-3) but was graded in the presence of 10, 25, 50 and 100 ng/ml doxycycline (Supplementary Fig. 2.4).

We examined the combined effect of removing the *GAL2*, *GAL3* or *GAL80* in a strain lacking *GAL1*. As shown in Fig. 2.3B-5, B-6, the combined deletion of *GAL1* and the *GAL2* feedback loop (GAL1Δ GAL2Δ fb) and dual deletion of *GAL1* and the *GAL80* feedback loop (GAL1Δ GAL80Δ fb) displayed bimodality for at least two galactose concentrations.

By stark contrast, the simultaneous deletion of *GAL1* and the *GAL3* feedback loop (GAL1Δ GAL3Δ fb) produced a graded response for the entire range of galactose (Fig. 2.3A-5). Remarkably, this graded response persisted irrespective of the constitutive Gal3p production rate in contrast to the single *GAL3* feedback knockout that displayed bimodality for some range of constitutive Gal3p levels (Supplementary Fig. 2.5A). These data provide further evidence that *GAL1* is an active regulatory component of the circuit and that the interplay between the *GAL1* and *GAL3* feedback

loops is crucial for bimodality.

In addition to eliminating bimodality, our results revealed that removing *GAL1* and the *GAL3* feedback loop abolished ultrasensitivity in the dose-response to galactose, indicating a coupling between the mechanisms for ultrasensitivity and bistability in the GAL network. We found that the Hill coefficient for P_{GAL10}Venus in WT was approximately 3 whereas this same reporter exhibited a Hill coefficient of approximately 1.3 in the absence of *GAL1* and the *GAL3* feedback loop (GAL1Δ GAL3Δ fb) (Supplementary Fig. 2.5D). This link between ultrasensitivity and bimodality may arise due to the necessity of ultrasensitivity for bistability [44].

Cooperative Gal4p interactions at the promoter level does not generate bimodal response

Bimodality was not observed using the *GAL3* and *GAL80* promoters as reporters of GAL network activity in WT for any concentration of galactose (Supplementary Fig. 2.1C). In contrast to the *GAL10* promoter, these promoters each contain a single *GAL4* binding site. Multiple *GAL4* binding sites may augment the dynamic range of the reporter to provide a sufficient separation of the high and low expression states or cooperativity of Gal4 proteins at the promoter level may be an important parameter of the bimodal response. To test whether multiple *GAL4* binding sites are necessary for bimodality, a synthetic GAL promoter containing a single Gal4p binding site from the *GAL7* promoter driving the expression of a fluorescent reporter was constructed (see Materials and methods). This reporter had minimal cooperativity and yet bimodality was detected for two galactose concentrations at steady-state (Supplementary Fig. 2.6). These data demonstrate that bimodality is not an exclusive property of promoters with multiple *GAL4* binding sites but is instead a property of the upstream regulatory network.

Deterministic model of GAL network recapitulates experimental results and provides insights into the roles of feedback loops

To further probe the roles of the feedback loops, we constructed an ordinary differential equation (ODE) model of the system (Section S2.4) which takes into account the concentrations of Gal1p (G1), Gal3p (G3), Gal4p (G4) and Gal80p (G80). Since Gal1p and Gal3p can function as co-inducers of GAL gene expression independently of galactose, presumably with lower affinities than the galactose bound forms, these different forms were not differentiated in the model [25].

Based on these assumptions (see Section S2.4 for a full description), the set of differential equations for G1, G3, G4 and G80 that model the interactions shown in Fig. 2.1 is given by

$$\begin{aligned}
\frac{d[\mathrm{G1}]}{dt} &= \alpha_{gal}\epsilon + \frac{\alpha_{G1}[\mathrm{G4}]^3}{[\mathrm{G4}]^3 + K_{G1}^3} + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],\\
\frac{d[\mathrm{G3}]}{dt} &= \alpha_{gal} + \frac{\alpha_{G3}[\mathrm{G4}]^2}{[\mathrm{G4}]^2 + K_{G3}^2} + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],\\
\frac{d[\mathrm{G4}]}{dt} &= \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],\\
\frac{d[\mathrm{G80}]}{dt} &= \alpha_{oG80} + \frac{\alpha_{G80}[\mathrm{G4}]^2}{[\mathrm{G4}^2 + K_{G80}^2} + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}].
\end{aligned}$$

Here, α_{gal} represents galactose as a constant input rate. Parameters were approximated from experimental measurements and values from the literature (Section S2.5). Using these estimates, the Hill coefficients for the feedback functions involving *GAL1*, *GAL3* and *GAL80* were set to 3, 2 and 2, respectively, but our conclusions were not sensitive to variations in these values. Models of the individual *GAL1*, *GAL3*, *GAL80* and combined *GAL1* and *GAL3* feedback knockouts (GAL1Δ fb, GAL3Δ fb, GAL80Δ fb, GAL1Δ GAL3Δ fb, respectively) were constructed by modifying appropriate terms in the WT model (Section S2.6).

Bifurcation analysis of GAL model confirms that only the combined *GAL1* and *GAL3* feedback deletion eliminates bistability

At equilibrium, the concentration of Gal4p can be written as an eleventh-order polynomial as described in Section S2.4. Similarly, the individual feedback deletion models for *GAL1*, *GAL3* and

GAL80 and combined *GAL1*, *GAL3* were simplified to an eighth, ninth, ninth and sixth order polynomials, respectively (Section S2.6). The roots of these polynomials include the equilibrium concentrations of Gal4p, which represents the activity of the GAL network. All of the models had the potential for bistability for some region of parameter space since the degrees of the polynomials were larger than a quadratic. Indeed, models with individual feedback deletions were still capable of bistability as a function of α_{gal} (Fig. 2.4A-1, A-2, A-3 and A-4). By contrast, removing both the *GAL1* and *GAL3* feedback loops abolished bistability for the entire range of α_{gal}, consistent with experimental data (Fig. 2.4A-5 and A-6).

GAL1 and *GAL3* feedback loops combine synergistically to augment bistability

Using the model, we explored the effects of the *GAL1* and *GAL3* feedback loops on the range of α_{gal} for which the system exhibits bistability. We defined the hysteresis strength D_H as the difference between the bifurcation points of α_{gal} as shown in Fig. 2.4A-1 (see Materials and methods). D_H represents the range of conditions in which the system exhibits bistability and thus the robustness of bistability to parameter variations increases with D_H. The *GAL1* and *GAL3* feedback deletion models had approximately 48% and 31% D_H compared to WT (Fig. 2.4B). By contrast, removing the *GAL80* feedback significantly increased D_H to 166% compared to its WT value, indicating that this negative autoregulatory loop undermines bistability.

The generality of these results and the dependence on parameters were explored by comparing the D_H of the WT and feedback deletions using randomly generated parameter sets. 10,000 parameter sets were obtained by sampling a normal distribution with mean equal to the values of parameter set I and coefficient of variation equal to 0.1. All parameters were varied except for the constitutive production rates of Gal1p (α_{G1s}), Gal3p (α_{G3s}) and Gal80p (α_{G80s}). This computation confirmed that GAL80Δ fb had a larger D_H compared to WT and that the WT exhibited a larger D_H than either of the GAL1Δ fb or GAL3Δ fb models for all parameter sets (Fig. 2.4C). These findings are consistent with the experimental characterization of history-dependent response of the GAL1Δ strain (Fig. 2.2B) and the data showing that range of galactose concentrations that produced bimodality

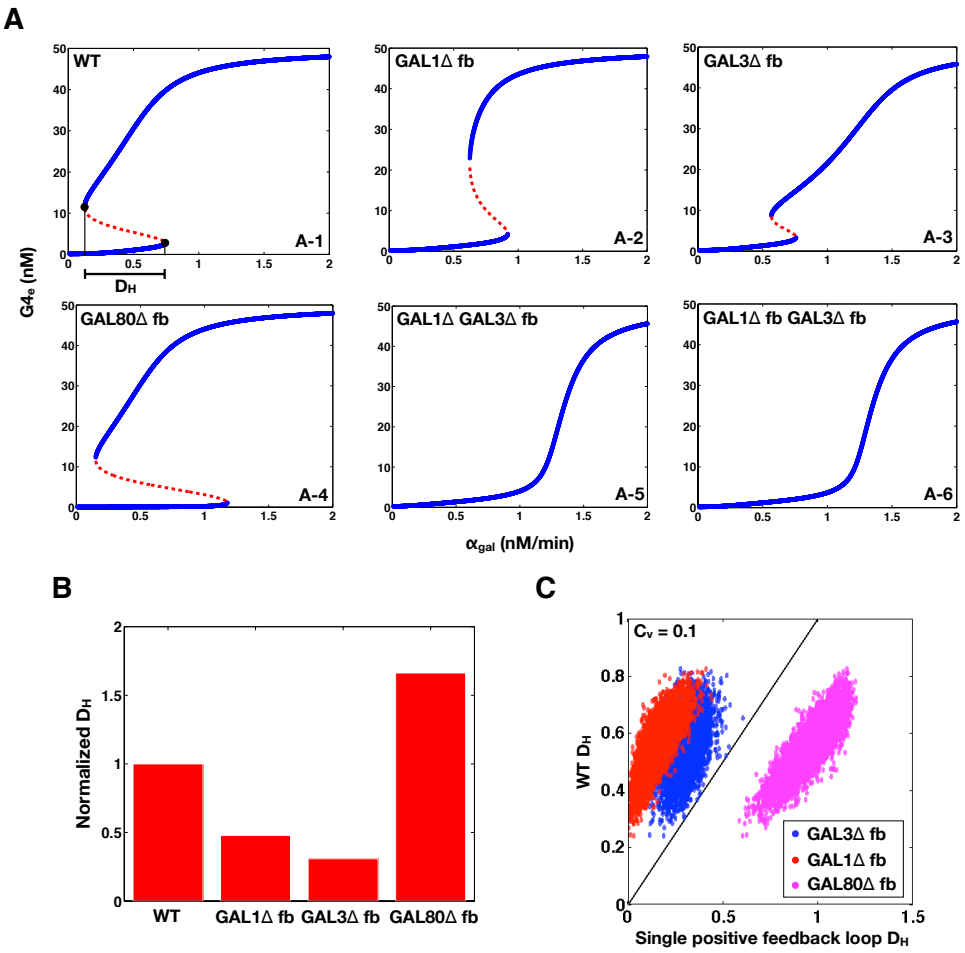


Figure 2.4: Bifurcation analysis of GAL model qualitatively matches experimental results and reveals that the *GAL1* and *GAL3* feedback loops combine synergistically to expand the parameter region for bistability. The bifurcation parameter α_{gal} represents galactose and the equilibrium value of Gal4p ($G4_e$) represents the activity level of the GAL network. **(A)** Bifurcation plots of wild-type (WT) and the feedback deletions of *GAL1* (GAL1Δ fb), *GAL3* (GAL3Δ fb), *GAL80* (GAL80Δ fb) and combined *GAL1* and *GAL3* (GAL1Δ GAL3Δ fb and GAL1Δ fb GAL3Δ fb). Blue and red represent stable and unstable equilibrium points, respectively. Reflecting the experimental results in Fig. 3, WT, GAL3Δ fb, GAL1Δ fb and GAL80Δ fb exhibits bistability whereas GAL1Δ GAL3Δ fb and GAL1Δ fb GAL3Δ fb are monostable for the full range of α_{gal}. A representative distance between the bifurcation points, D_H, is highlighted by a solid black line. **(B)** Quantification of the range of bistability for the WT and single feedback knockouts shown in **A**. Normalized D_H is equal to the range of α_{gal} that produces bistability relative to WT. **(C)** Comparison of D_H in the WT, GAL1Δ fb (blue), GAL3Δ fb (red) and GAL80Δ fb (magenta) models for 5000 representative randomly generated parameters sets sampled from a normal distribution with $C_v = 0.1$. Data points above the $x = y$ line (black) correspond to parameter sets where D_H is larger in WT compared to the single feedback loop knockouts.

was expanded in the absence of the *GAL80* feedback loop (Fig. 2.3B-2).

In summary, collaboration between the *GAL1* and *GAL3* autoregulatory loops expands the region of bistability across a broad region of parameter space, suggesting that this synergy between dual positive feedback loops may be a consequence of the unique regulatory roles of Gal1p and Gal3p in the GAL circuit. In addition, we found that GAL1Δ fb GAL3Δfb and GAL1Δ fb GAL3Δfb were monostable for all 10,000 parameter sets, indicating that one of these autoregulatory loops is necessary for generating bistability across a broad region of parameter space.

Recently, a two-stage galactose induction model has been proposed whereby the Gal3p-Gal80p complex (C83) dominates initially and the Gal1p-Gal80 complex (C81) dominates at a later stage [34]. To check the consequences of including this feature in our model, we scanned over a wide range of parameters using the Latin hypercube sampling method [45] (Section S2.5) and identified sets of parameters that qualitatively matched all of our data in addition to the dynamic ordering response of C83 and C81 (Supplementary Fig. 2.7B,C). This new parameter set exhibited the same roles for the *GAL1* and *GAL3* feedback loops in enhancing D_H across a broad region of parameter space, further illustrating the generality of our results (Supplementary Fig. 2.7D).

Constitutive production of Gal1p and Gal3p can abolish bimodality in the absence of the individual *GAL1* and *GAL3* feedback loops

We next tested whether the model could predict and explain the disappearance of bimodality due to high unregulated levels of Gal1p and Gal3p in the absence of their individual feedback loops (Fig. 2.5, Supplementary Fig. 2.2 and Supplementary Fig. 2.4). The individual *GAL1* and *GAL3* feedback loop deletion models predicted the loss of bistability as the rates of constitutive production, α_{G1s} or α_{G3s}, was increased (Fig. 2.5C,D). An increase in α_{G1s} in the GAL1Δ fb model caused the bistable region to contract and vanish at a critical value ($\alpha_{G1s} = 4$) (Fig. 2.5C). In the GAL3Δ fb model, increasing α_{G3s} caused the bistable region to shift to smaller values of α_{gal} (Fig. 2.5D) and eventually move out the positive orthant to negative values of α_{gal} at a critical α_{G3s} ($\alpha_{G3s} = 1$), thus producing monostability for all physically realistic values of α_{gal}.

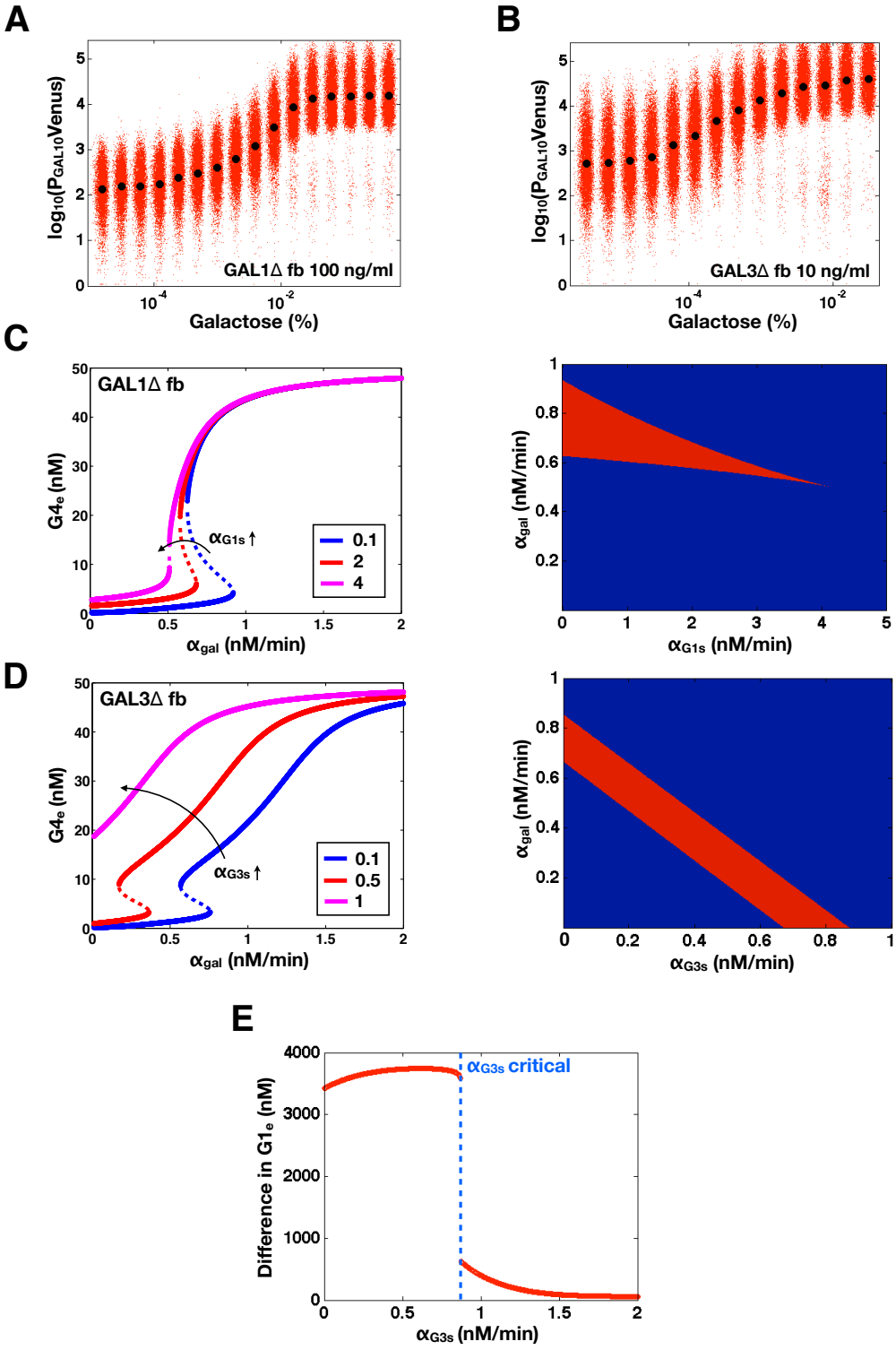

A
B
C
D
E
log10(PGAL10Venus)
GAL1Δ fb 100 ng/ml
GAL3Δ fb 10 ng/ml
Galactose (%)
GAL1Δ fb
G4e (nM)
αG1s ↑
0.1
2
4
αgal (nM/min)
αG1s (nM/min)
GAL3Δ fb
αG3s ↑
0.1
0.5
1
αG3s (nM/min)
Difference in G1e (nM)
αG3s critical

Figure 2.5

Figure 2.5: Model predicts that constitutive production of Gal1p or Gal3p above a threshold can abolish bistability in the absence of the individual *GAL1* or *GAL3* feedback loops (matching experimental data in Section S2.2, Supplementary Fig. 2.2 and Supplementary Fig. 2.4). **(A)** Critical constitutive level of Gal1p in the absence of the *GAL1* feedback loop produced a graded response. Flow cytometry measurements of P_{GAL10}Venus in a *GAL1* feedback deletion strain (GAL1Δ fb). In this strain, *GAL1* was expressed from a *TET* promoter and induced with 100 ng/ml doxycycline (dox), corresponding to approximately 20% of fully induced wild-type (WT) levels (Supplementary Fig. 2.3A-1). **(B)** Critical level of Gal3p in the absence of the *GAL3* feedback loop produced a graded response. Flow cytometry measurements of the *GAL3* feedback deletion strain (GAL3Δ fb). *GAL3* was expressed from a *TET* promoter and induced with 10 ng/ml dox, corresponding to approximately 63% of fully induced WT levels (Supplementary Fig. 2.3A-3). **(C)** In the GAL1Δ fb model, increasing the constitutive production rate of Gal1p (α_{G1s}) decreases the region of bistability and causes bistability to vanish at a critical value ($\alpha_{G1s} = 4$ nM/min). Regions of bistability (red) and monostability (blue) for different values of α_{G1s} and α_{gal} in GAL1Δ fb shows that the bistability parameter region contracts and eventually vanishes with increasing α_{G1s}. **(D)** In the GAL3Δ fb model, increasing the constitutive production rate of Gal3p (α_{G3s}) eliminates bistability by shifting the bistable region to smaller α_{gal} values. A critical threshold of α_{G3s} ($\alpha_{G3s} = 1$ nM/min) causes the bistable region to move out of the positive orthant, producing monostability for all physically realistic α_{gal} values. Regions of bistability (red) and monostability (blue) for different values of α_{G3s} and α_{gal}. **(E)** The *GAL1* feedback nonlinearity disappears with increasing α_{G3s} in the GAL3Δ fb model. The maximum difference in Gal1p steady-state concentration ($G1_e$) was computed across the full range of α_{gal} for a series of α_{G3s} values and represents the activity of the *GAL1* feedback loop. Above the critical α_{G3s} threshold (dashed blue line), $G1_e$ does not change in response α_{gal}, indicating that the *GAL1* autoregulatory loop is not active in this parameter regime.

Since Gal1p and Gal3p played an important role in generating bistability, we suspected that the disappearance of bistable behavior for α_{G3s} or α_{G1s} exceeding critical values could be the result of an indirect neutralization of the remaining loop. For example, it could be case that overexpression of Gal3p in a *GAL3* feedback deletion had the effect of neutralizing the *GAL1* feedback loop. The computational model afforded us the possibility of testing this hypothesis. For a given value of α_{G3s}, we defined the *GAL1* feedback activity as the maximum change in steady-state Gal1p concentration ($G1_e$) across the full range of galactose ($\alpha_{gal} = 0 - 2$ nM/min). As shown in Fig. 2.5E, the *GAL1* feedback was highly active for a range of α_{G3s} values but abruptly approached zero at a critical threshold of α_{G3s} (dashed blue line). Therefore, increasing the constitutive production rate of Gal3p was indeed equivalent to removing the *GAL1* feedback since a sufficiently large α_{G3s} mapped the *GAL1* feedback nonlinearity to a saturated (inactive) regime.

The insight generated by the model about the link between the constitutive production rates of Gal1p or Gal3p and the loss of bistability suggested that the graded response observed in GAL3Δ fb (Fig. 2.5B) should be the result of overexpressing the Gal3 protein. To test this possibility, we

compared the *GAL3* mRNA expressed from the *TET* promoter to that of WT induced with 0.005% and 0.05% galactose using qPCR. These data showed that the *GAL3* mRNA level in GAL3Δ fb induced with 10 ng/ml dox was overexpressed by 43% relative to WT induced with 0.05% galactose, significantly higher than *GAL3* mRNA levels for the bimodal range of WT and GAL3Δ fb (Supplementary Fig. 2.3A-3). These results argue that in order to study the functional contribution of feedback loops to a phenotype, the strength of constitutive expression needs to be carefully tuned in order to recapitulate the physiological operating point(s) of the wild-type circuit.

Properties of positive feedback loops established by molecular sequestration

Sequestration binding affinity of an activator and repressor can tune the range of conditions for bistability

To generalize our results further, we explored the principles by which the interactions of the positive feedback loops mediated by Gal1p and Gal3p generate bistability. Characterizing the set of essential molecular interactions that combine to generate bistability in the GAL system may be useful for analyzing other natural switch-like biological networks and for constructing robust and tunable bistable synthetic circuits. Gal1p and Gal3p competitively sequester a common protein, Gal80p. Competitive binding interactions and molecular sequestration can produce ultrasensitivity, which is crucial building block for a bistable system [46, 47, 48, 49]. Therefore, we suspected that the competitive sequestration of Gal80p by Gal1p and Gal3p may constitute a critical feature of the system.

To probe the functionalities provided by positive feedback loops linked to molecular sequestration, we examined a simple model of a single positive feedback loop that is implemented by an activator x that can form an inactive complex with a transcriptional repressor z. In this circuit, z transcriptionally represses the production of x and therefore a positive feedback loop is established by inhibition of the transcriptional repressor using molecular sequestration (Supplementary Fig. 2.8A). We first examined the parameter dependence of this system in the absence of transcriptional

cooperativity and found that this circuit could exhibit bistability depending on the value of the binding affinity of the activator and repressor (Supplementary Fig. 2.8B). Therefore, modifying this parameter is an alternative mechanism to induce bistability in the circuit without increasing the cooperativity.

Building on these results, we next investigated the roles of double positive feedback loops connected by molecular sequestration. We considered a three-state ODE model consisting of a transcriptional repressor z that directly regulates two activators, x_1 and x_2 with Hill coefficients of 3 and 2, respectively. x_1 and x_2 can form inactive heterodimers with z and hence x_1 and x_2 compete to bind z (Fig. 2.6A). In this model, the mechanisms of sequestration and positive feedback are triggered by an input (u) that represents a basal production rate of x_1 and x_2. The system of equations that model the interactions in Fig. 2.6A (see Section S2.7 for a full description) is

$$\begin{aligned}
\frac{dx_1}{dt} &= u + \frac{\alpha_1 K_1^3}{K_1^3 + z^3} + \beta_1 x_1 z - \gamma_1 x_1, \\
\frac{dx_2}{dt} &= u + \frac{\alpha_2 K_2^2}{K_2^2 + z^2} + \beta_2 x_2 z - \gamma_2 x_2, \\
\frac{dz}{dt} &= \alpha_z + \beta_1 x_1 z + \beta_2 x_2 z - \gamma_z z.
\end{aligned}$$

In the double positive feedback case, bistability could be induced in this system by adjusting the binding affinities K_{D1} and K_{D2} (which modify β_1 and β_2) as bifurcation parameters without changing the cooperativity of the transcriptional regulation (Fig. 2.6B,C). Setting $K_{D1} = K_{D2}$, we found that the range of the input that produced bistability was inversely related to the magnitude of the binding affinities (Fig. 2.6B). In addition, the range of the input that generated bistability was increased in a system with two positive feedback loops compared to a single positive feedback loop for the set of symmetrically varying K_{D1} and K_{D2} values (Fig. 2.6C).

To explore asymmetry in the binding affinities, D_H was computed for a series of linearly spaced K_{D1} and K_{D2} values within the range of 0.5-80 nM (Fig. 2.6D). The largest range of bistability was obtained for the strongest binding affinities and D_H decreased monotonically with increasing K_{D1} or K_{D2}. In addition, fixing one K_D while varying the other (left column and bottom row)

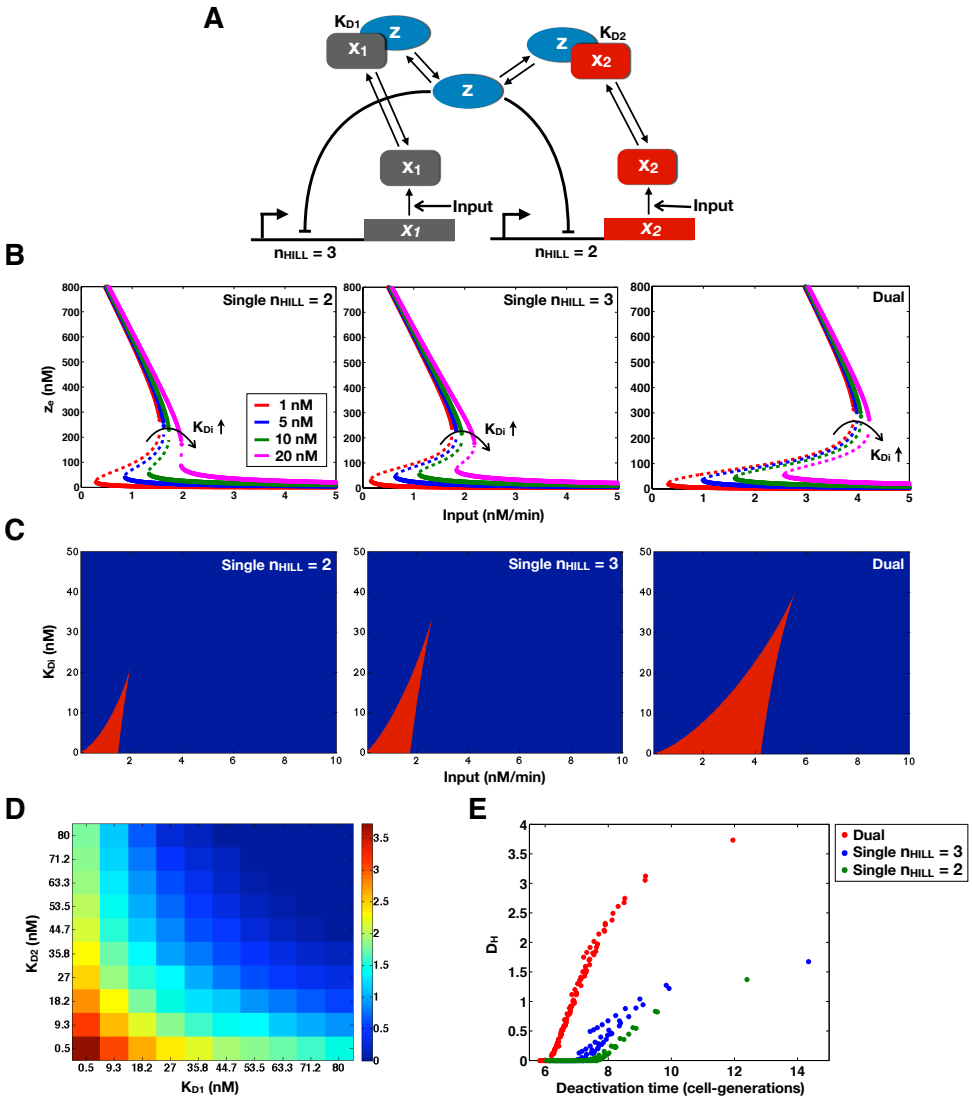


Figure 2.6: Molecular consequences of positive feedback loops established by molecular sequestration. Sequestration binding affinities (K_{D1} and K_{D2}) can tune the parameter region for bistability and the addition of a second positive feedback loop can reduce the deactivation response time and augment the range of conditions for bistability. **(A)** Circuit diagram for dual positive feedback loops mediated by the activators, x_1 and x_2 coupled by molecular sequestration to a transcriptional repressor (z). Transcriptional feedback regulation of x_1 and x_2 are modeled by Hill functions with Hill coefficients of 3 and 2. The single positive feedback loop models were obtained by removing the appropriate repression arrow from z to the promoter of x_1 or x_2 or equivalently, replacing the Hill functions with a constant production rate, α_{1s} or α_{2s}. **(B)** Bifurcation diagrams relating the input to the steady-state concentration of z (z_e) reveal that symmetrically weakening the binding affinities shrinks the region of bistability. **(C)** Parameter regions of bistability (red) and monostability (blue) for different values of the input and symmetrically varying K_{D1}, K_{D2} in the single and double feedback loop models. **(D)** Range of bistability (D_H) for a range of K_{D1} and K_{D2} values in the double feedback loop system. **(E)** Relationship between D_H and the deactivation response time measured in cell-generations (see Section S2.7). For a constant nonzero D_H, the dual feedback loop circuit exhibited a faster deactivation response time compared to the either of the single positive feedback loop models.

did not decrease D_H as significantly as symmetrically changing the two binding affinities together (diagonal). These results suggest that asymmetry in the binding affinity strengths whereby one activator interacts strongly and the other activator binds weakly to the same repressor can preserve bistability over a wide range of values for the weaker K_D, thus reducing the system's sensitivity to variations in this parameter.

Double positive feedback loops can produce larger range of bistability and a faster dynamic response than a single feedback loop

We suspected that modulating the binding affinities to induce bistability may concurrently alter other circuit functions such as the dynamic response time to a change in the input. To explore these relationships, we measured the response times of the circuits to switch from the low→high state (activation response time) and from the high→low state (deactivation response time). To do so, a step function increase or decrease in the input was applied and the delay for the circuit to adapt to this transition was quantified (see Section S2.7). The time required for an output species that was transcriptionally repressed by z (representing a fluorescent reporter) to increase or decay to half its maximum value was quantified in cell-generations.

In the double feedback loop system, the activation response time decreased with the strength of the binding affinities whereas the deactivation response time had the opposing relationship and increased with the strength of this binding affinities (Supplementary Fig. 2.8C,D). For a constant nonzero D_H, the dual feedback loop system could switch faster to the high state than either of the single feedback loop models (Supplementary Fig. 2.8E). Since both the D_H and the deactivation response time are inversely related to K_{D1} and K_{D2}, a tradeoff exists between increasing the range of conditions for bistability and decreasing the deactivation response time (Fig. 2.6E). A comparison of D_H and the response times for the single and double feedback loop systems revealed that dual feedback loops can produce a larger D_H over a narrower range of deactivation response times compared to the single feedback loop systems with Hill coefficients of 2 or 3. Taken together, the dual feedback loop system can produce a larger range of bistability and exhibit a faster response

time to abrupt changes in the environment compared to a single feedback loop system.

Discussion

A bimodal distribution of gene expression across a population of isogenic cells, which generates two distinct cellular states, can produce significant cell-to-cell heterogeneity. This bimodality can also lead to a switch-like response that filters out noise below a threshold and produces a large fold-change in the system's output if the input crosses this threshold [50]. In this work, we used the GAL gene-regulatory circuit as a model system to dissect and analyze the origins of bimodality in a natural biological network. We demonstrated that bistability underlies this bimodality and used a combination of experiments and computational modeling to identify two key features that produce bistability: (1) a threshold established by two positive feedback loops mediated by Gal1p and Gal3p and (2) an ultrasensitive stage produced by competitive molecular sequestration of Gal80p by Gal1p and Gal3p.

To unravel the molecular interactions critical for bistability in the GAL system, we performed a comprehensive exploration of multiple feedback loops. Our investigations revealed that the GAL bimodal response is remarkably robust to feedback loop perturbations. Indeed, individual elimination of the *GAL1*, *GAL2*, *GAL3*, and *GAL80* feedback loops was insufficient to abolish bimodality. Furthermore, bimodality persisted for multiple deletions of these loops and, surprisingly, only disappeared in a double deletion of *GAL1* and the *GAL3* feedback loop. We therefore identified Gal1p and Gal3p as central mediators of two synergistic positive feedback loops that generate bistability in the GAL gene-regulatory network. Multiple positive feedback loops can facilitate the bistable behavior of a circuit by expanding the range of conditions for bistability, which improves the robustness of bistability to parameter variations [50].

A previous study attributed bimodality in the GAL pathway to the activity of the *GAL3* feedback loop [20]. Here we demonstrate that cells with a deleted *GAL3* feedback loop are still capable of bimodality in their response to galactose for low levels of constitutive Gal3p expression. However, we found experimentally that bimodality vanishes when Gal3p is expressed at high and unregulated

levels. Our computational model explains this behavior by the loss of remaining *GAL1* feedback due to constitutive expression of Gal3p beyond a threshold. Interestingly, in this regime, the genetic wiring of the *GAL1* feedback loop is present, but the feedback loop was rendered inactive indirectly by constitutive Gal3p expression above a threshold.

These results underscore the challenges inherent in the interpretation of feedback deletion experiments in which the specific range of constitutive expression of the deleted link might become an important determinant of the system's properties and can mask the true functional roles of the feedback pathway. These findings also argue that the complete interpretation of feedback knockouts requires thorough investigation of active mechanisms and nonlinearities that are operational in a given circuit, beyond static snapshots of the circuit's topology as determined by genetics [51].

Stoichiometric binding interactions, for example, molecular sequestration of a repressor by an activator or inhibition of an enzyme by a small molecule, can produce ultrasensitivity in biological circuits [46, 47, 48, 49, 52]. Our computational model indicates that competitive molecular sequestration of Gal80p by Gal1p and Gal3p produces an ultrasensitive change in the concentration of free Gal4p in response to a small variations in extracellular galactose and this ultrasensitivity does not rely on cooperative binding of Gal4p to GAL promoters and/or oligomerization. These results suggest that the stoichiometric inhibition of Gal80p by Gal1p and Gal3p is a crucial source of ultrasensitivity in the GAL network that sets the stage for a robust bistable response to galactose since ultrasensitivity is required for bistability [53, 49].

Beyond the GAL system, we believe these results to be applicable to many bistable systems. We used a simple computational model to explore the general mechanisms by which positive feedback loops linked to competitive sequestration can produce ultrasensitivity and bistability. Using this model, we found that the positive feedback and sequestration topology can be used to build a bistable system in the absence of transcriptional cooperativity by adjusting the binding affinity parameter between the activator and inhibitor. If bistability confers an fitness advantage, this parameter could be adjusted through mutation of the protein-protein binding interface and may be more evolvable than modifying the cooperativity of transcriptional regulation through oligomerization or multiple

transcription factor binding sites. In addition, we identified a tradeoff between the range of bistability and the deactivation response time of this circuit. In response to an abrupt change in the stimulus, we found that a system with double positive feedback loops can switch faster to the low state compared to the single feedback loop system for a fixed range of bistability, highlighting a novel advantage of multiple positive feedback loops.

Positive feedback loops established by molecular sequestration may represent a general class of systems for implementing robust switch-like cellular responses. For example, the conserved regulatory network that controls cell-differentiation in *Drosophila* consists of similar molecular mechanisms to the GAL circuit including molecular sequestration and multiple feedback loops that implement a switch-like developmental program [54, 55]. Activation of this cell-differentiation circuit relies on molecular titration of a repressor, Extramacrochaetae (Emc) by the activators Daughterless (Da) and the Achaete-Scute Complex (As-c). Da and As-c transcriptionally autoregulate and thus form two positive feedback loops [56].

S. cerevisiae cells growing on galactose could benefit from bistability on a single-cell and population level. A bistable circuit can produce a decisive response to a slow variation in the stimulus [52]. This decoupling ensures that the abrupt change in the system's output is not dependent on the rate of change of the stimulus and is instead an intrinsic property of the circuit's dynamical system. In addition, bimodality due to an underlying bistability can produce stable lineages of cells with a memory of previous environmental conditions. As a consequence of hysteresis, cells with a history of the stimulus will respond differently to a second exposure due to a shift in the threshold of deactivation. This epigenetic memory of previous environments can fine-tune the switching threshold and provide an additional source of cell-to-cell heterogeneity in the perception of the stimulus.

There are also several potential advantages of bimodality at a population level. For example, significant single-cell phenotypic variation, generated by bimodality, can serve as a bet-hedging strategy for microbial populations in uncertain environments [7, 57]. Since *S. cerervisiae* grows poorly even in the presence of high concentrations of galactose and risks accumulation of the toxic intermediate galactose-1-phosphate, the bimodal response may serve as a population strategy to weigh the ener-

getic costs and benefits of activating the GAL regulon [42]. Another intriguing possibility is whether bimodality establishes a division-of-labor in which the high population metabolizes galactose and produces a byproduct that is used by the low population [58].

Feedback loops are ubiquitous in biological systems, and dissecting their precise quantitative roles is a crucial step for unraveling the organizational principles of cellular decision-making. While a single transcriptional positive feedback loop can generate bistability with cooperativity and precise parameter tuning, this study suggests that a single noncooperative positive feedback loop with sequestration can generate bistability and this bistability parameter region can be significantly augmented by the addition of a second positive feedback loop. These insights will be essential for pinpointing the operational principles of switch-like cellular responses, in addition to suggesting rules for designing robust synthetic circuits.

Materials and Methods

Strains

All plasmids used in this study were derived from a set of yeast single integration vectors constructed in the lab of Wendell Lim (UCSF). These vectors contain markers and targeting sequences for the *LEU2*, *HIS3*, *TRP1* and *URA3* loci. These vectors were linearized for transformation by digesting with PmeI and transformed using standard techniques. Promoters were cloned between the PspOMI and XhoI restriction sites and coding sequences were inserted between the XhoI and BamHI sites. These plasmids contained an *ADH1* terminator downstream of BamHI site. All strains were haploid with the exception of MA0182 and wild-type diploid [20]. In the haploid backgrounds, rtTA-M2 was expressed from a medium strength variant of the *TEF* promoter, *TEFm4* [43, 59]. Gene deletions were verified using polymerase chain reaction. A functional test for constitutive P_{GAL10}Venus expression in the absence of galactose was also used to verify successful deletion of *GAL80*. Strains are listed in Table SII. The sequences for the *GAL3*, *GAL10* and *GAL80* promoters were 1017, 646 and 283 base pairs upstream of the start codons, respectively. The *TET* promoter consisted

of a region of the *CYC1* promoter and two TetR operator binding sites [60]. The synthetic single *GAL4* binding site promoter, $P_{CYC1\text{-}G4BS}$, consisted of a binding site from the *GAL7* promoter (CGGACAACTGTTGACCG) upstream of the *CYC1* core promoter.

Growth conditions and flow cytometry

Cells were grown in appropriate dropout media supplemented with 2% filter-sterilized raffinose at 30°C. In 2% raffinose media supplemented with zero or small amounts of galactose, cell divisions occurred approximately every three hours during exponential growth phase. Steady-state measurements were performed after a 20 hour induction period. Cells were induced for 30 hours for hysteresis experiments as explained in Section S2.1. OD_{600} (cell density) was maintained below 0.1 to prevent significant changes in the galactose concentration for the duration of the experiment. Flow cytometry measurements were made using a MACSQuant VYB (Miltenyi Biotec) or LSRII analyzer (BD Biosciences). For both instruments, a blue (488 nm) laser was used to excite YFP. Emission was detected on the MACSQuant or LSRII using a 525/50 nm and 530/30 nm filter, respectively. At least 10,000 cells were collected for each measurement.

Analysis of flow cytometry distributions

Bimodality classification

Flow cytometry distributions were analyzed using a Gaussian mixture model algorithm (GMM, MATLAB) [61]. The GMM assumes that the data is a mixture distribution where the probability density function is a linear combination with coefficients that sum to one ($\xi_1 + \xi_2 = 1$).

$$f(x) = \xi_1 N_1(\mu_1, \sigma_1) + \xi_2 N_2(\mu_2, \sigma_2).$$

The parameters for the GMM include the means, μ_1, μ_2, standard deviations, σ_1, σ_2 and mode weights ξ_1, ξ_2. A distribution was categorized bimodal if the following conditions were true

$$|\mu_1 - \mu_2| > 2\max(\sigma_1, \sigma_2),$$

$$\min(\xi_1, \xi_2) > 0.1.$$

Activation responses

Activation responses for bimodal transitions were analyzed using the fraction of high expressing cells ($\mathrm{F_H}$). The threshold was set to the minimum separating the two local maxima. $\mathrm{F_H} = \frac{n_\mathrm{H}}{n_\mathrm{tot}}$ where n_H and n_L are the number of high and low expressing cells, respectively ($n_\mathrm{tot} = n_\mathrm{H} + n_\mathrm{L}$). The activation level for a graded response was quantified using the normalized mean fluorescence level ($\mathrm{M_Y}$),

$$\mathrm{M_Y} = \frac{(\log_{10}(\mathrm{Y}) - \min(\log_{10}(\mathrm{Y}))}{(\max(\log_{10}(\mathrm{Y})) - \min(\log_{10}(\mathrm{Y}))}.$$

Quantitative real-time PCR

Total RNA was isolated using a YeaStar RNA Kit (Zymo Research Corp.). Oligonucleotides for quantitative real-time PCR (qPCR) were designed using Integrated DNA Technologies PrimeTime qPCR assay. 500 nanograms total RNA was reverse-transcribed using the iScript cDNA synthesis kit (Bio-Rad). The reaction mix contained 5 μl of SsoFast Probes SuperMix (Bio-Rad), 0.5 μl of primer probe corresponding to 250 nM primers and 125 nM probe (20X stock) and 0.5 μl cDNA. Three technical replicates for each sample were analyzed using the CFX96 real-time PCR machine (Bio-Rad). Relative expression levels were determined by the $2^{(-\Delta\Delta \mathrm{C_t})}$ method [62]. Each sample was normalized by the $\mathrm{C_t}$ geometric mean for the reference genes, *ACT1* and *UBC6* [63].

Computational modeling

Code for mathematical modeling was written in MATLAB (Mathworks) and Mathematica (Wolfram Research).

Bifurcation points

We identified turning, fold and saddle-node bifurcation points that can create bistability by computing the values of α_{gal} that caused a real eigenvalue of the Jacobian matrix to change from negative to positive, producing a singular Jacobian matrix at the point where the real part of the eigenvalue equaled zero. The bifurcation parameter ($\lambda = \alpha_{gal}$) appeared linearly in the polynomial equations for the equilibrium concentrations of Gal4p. To satisfy the conditions of a singular Jacobian matrix and equilibrium, there were two equations in two unknowns using the Gal4p polynomial ($x =$ Gal4p)

$$f(x) + \lambda g(x) = 0,$$

$$f'(x) + \lambda g'(x) = 0.$$

We solved the system of equations using the Sylvester resultant [64]. This resultant provides conditions for the coefficients of two polynomials of a single variable to have a root in common. Sylvester matrices A and B contained the coefficients of f, f' and g, g', respectively. The dimensions of A and B were $(d_1 + d_2)$ x $(d_1 + d_2)$ where d_1, d_2 are the degrees of highest polynomial of either f or g and correspondingly f' or g', respectively. The bifurcation points were computed by solving the generalized eigenvalue problem $(A + \lambda B)\,\phi = 0$.

Acknowledgments

The authors would like to thank Michael Elowitz, Marcella Gomez and Fiona Chandra for helpful discussions, Louis Romero for mathematical modeling insights, and Rochelle Diamond and Josh Verceles for assistance with flow cytometry. We are grateful to the laboratory of Christina Smolke and to Murat Acar for providing yeast strains used in this study. This research project was supported by the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office.

Author contributions

OSV, RMM and HES conceived the research and OSV did the experiments, analyzed the data, constructed the model, did the analysis and wrote the manuscript with substantial input from RMM and HES.

Supplementary Information

1. Distinguishing dilution memory from a history-dependent response

2. Characterization of the *GAL3* feedback loop on the bimodal response

3. Comparison of open and closed loop transcriptional circuits

4. GAL model description and steady-state solution

5. Estimation of GAL model parameters

6. GAL feedback loop deletion models

7. General model for positive feedback loops established by molecular sequestration

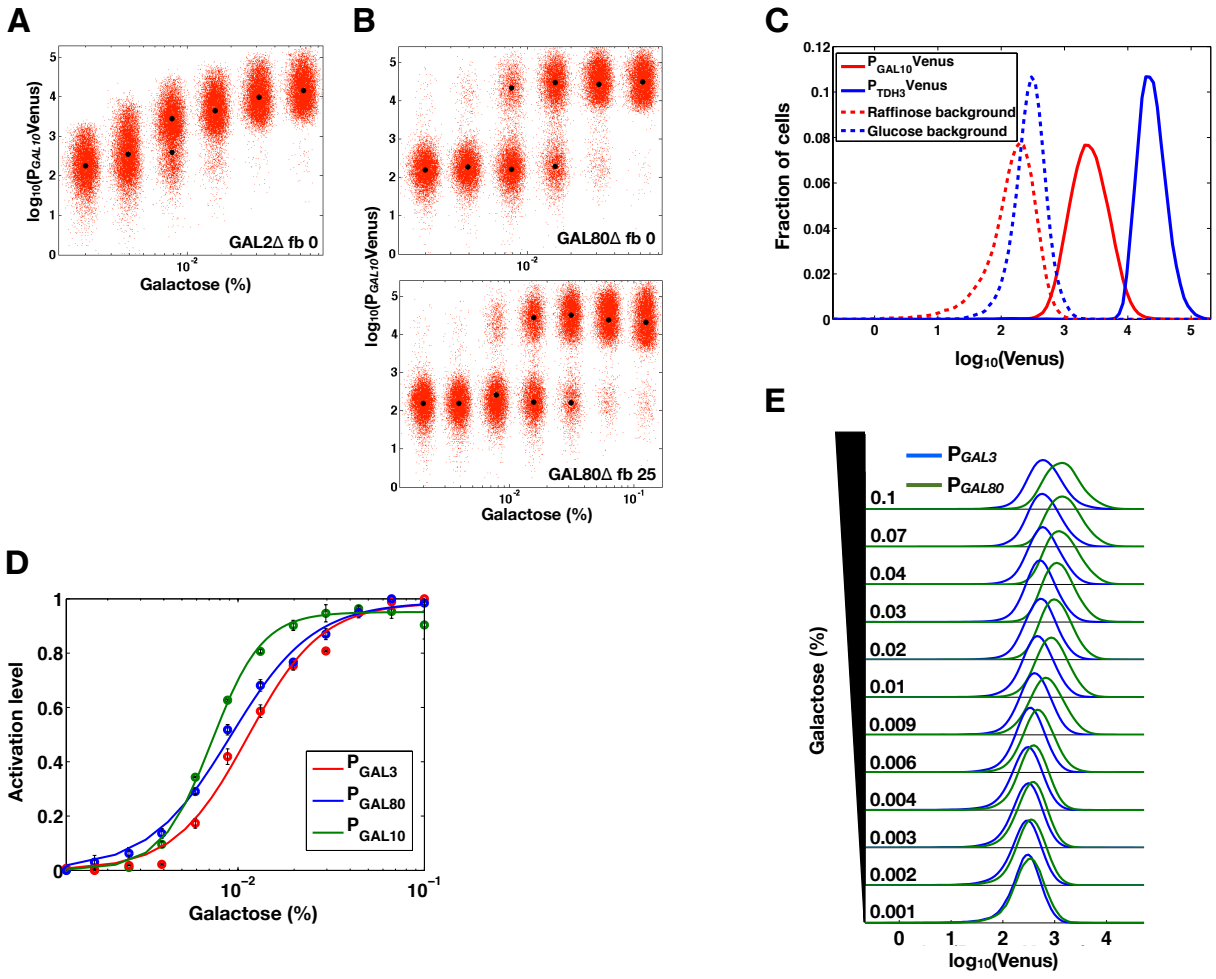
S2.1 Distinguishing dilution memory from a history-dependent response

YFP is highly stable and predominantly decreases through cell dilution [65]. As a consequence, activation of P_{GAL10}YFP is faster than deactivation, and distributions of cells from E_H and E_L will inevitably differ until cells equilibrate to a new steady-state in the second set of environments (E_1, ..., E_n). We estimated the amount of time and hence the number of cell divisions necessary to distinguish dilution memory from a history-dependent response (Fig. 2.2).

In the presence of 2% galactose, the steady-state YFP fluorescence expressed from the *GAL10* promoter was approximately 78% of YFP fluorescence expressed from the *TDH3* promoter after autofluorescent background subtraction (Supplementary Fig. 2.1A). There are approximately 169,000 Tdh3 proteins present in glucose conditions in a haploid *S. cerevisiae* background [66], corresponding to 131,820 Gal10 proteins (assuming promoter strength is proportional to the number of molecules). Therefore, the concentration of Gal10p is 7.5 μM at full galactose induction [67]. As a lower bound, we assumed that 150 molecules of YFP (8.6 nM) was indistinguishable from the autofluorescence background using flow cytometry [68]. Therefore, the number of cell divisions required to dilute YFP from full induction to background is $\log_2(7.5) - \log_2(0.0086) = 9.8$. In minimal dropout media supplemented with 2% raffinose, cells doubled approximately every three hours during exponential phase, which corresponds to 30 hours to distinguish dilution memory from a history-dependent response.

S2.2 Characterization of the *GAL3* feedback loop on the bimodal response

We found that bimodality persisted in the absence of the *GAL3* feedback loop for a range of wild-type (WT) *GAL3* levels as shown in Fig. 2.3A-4. These results are different from a previous study that attributed the observed bimodality of the GAL network to the activity of the *GAL3* feedback loop [20]. This study used a diploid *GAL3* feedback loop deletion strain where Gal3p was constitutively expressed with a *TET* inducible promoter. Using fluorescent Gal3 fusion proteins, the authors identified 50 ng/ml dox as equivalent to 80% of the WT *GAL3* levels induced with 0.5%



Supplementary Figure 2.1: Comparison of promoter strengths and ultrasensitivity. Venus (YFP) fusions to the *TDH3*, *GAL3*, *GAL10* and *GAL80* promoters in a wild-type (WT) background (P_{TDH3}Venus, P_{GAL3}Venus, P_{GAL10}Venus and P_{GAL80}Venus). **(A)** Comparison of *GAL10* and *TDH3* promoter strengths. Promoter were compared after subtracting the corresponding autofluoresence background (solid red and blue histograms). P_{GAL10}Venus and P_{TDH3}Venus were grown separately in 2% raffinose + 2% galactose or 2% glucose. The autofluoresence background values were obtained from a wild-type W303a strain lacking a fluorescent reporter grown separately in 2% raffinose or 2% glucose media (dashed red and blue histograms) **(B)** Activation level represents the fraction of high expressing cells for P_{GAL10} and the normalized mean of unimodal distributions for P_{GAL3} and P_{GAL80} (M_Y, see Materials and methods). Lines are fits of the data to Hill functions with Hill coefficients of 3.2, 2.2 and 2 for P_{GAL10}, P_{GAL3} and P_{GAL80}. Error bars represent one standard deviation (n=3). **(C)** Representative flow cytometry distributions of Venus from P_{GAL3} and P_{GAL80} for a range of galactose concentrations at steady-state. Since the *GAL3* and *GAL80* promoters are weaker than the *GAL10* promoter, the flow cytometry gain settings were increased for these strains to detect the full expression range.

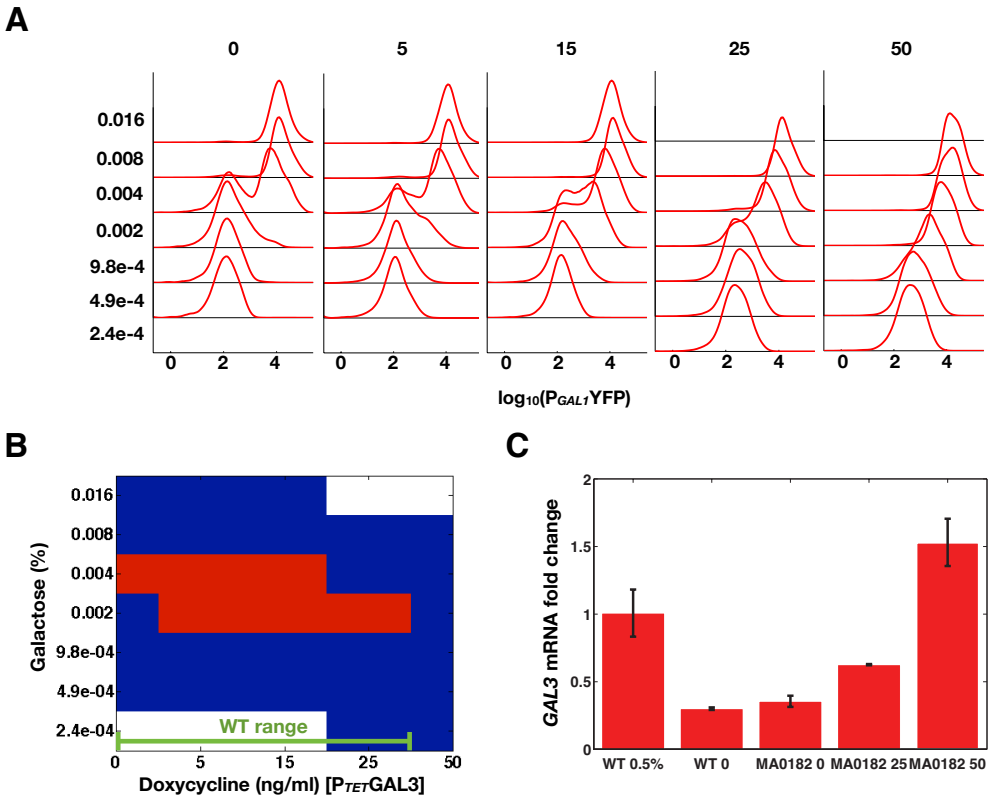
galactose.

To further explore the roles of the *GAL3* feedback loop on the bimodal response, we repeated the experiments from Acar *et al.* using the MA0182 strain. Following the authors' protocol, we observed bimodality in P_{GAL1}YFP expression after an induction of 27 hours for 0.004% galactose in the absence of doxycycline (dox) (Supplementary Fig. 2.2A). A Gaussian mixture model (GMM) was used to classify bimodality (see Materials and Methods). Using this criteria, these data showed the *GAL3* feedback loop was not necessary for bimodality for some range of *GAL3* levels (Supplementary Fig. 2.2B).

The galactose dose response was next measured for different *GAL3* levels by inducing MA0182 with a range of galactose and dox concentrations (Supplementary Fig. 2.2A). These data showed that MA0182 was bimodal for at least one galactose concentration between 0-25 ng/ml dox (Supplementary Fig. 2.2B). However, bimodality was not detected for 50 ng/ml dox.

We compared *GAL3* mRNA levels to WT *GAL3* expression using quantitative real-time PCR (qPCR). According to these results, 50 ng/ml dox corresponded to approximately 150% *GAL3* levels relative to WT induced with 0.5% galactose (Supplementary Fig. 2.2C). These results indicated that WT *GAL3* expression in MA0182 was between 0 ng/ml (36%) and approximately 35 ng/ml dox (100%). Acar *et al.* stated that MA0182 displayed a graded response for 5-300% of *GAL3* levels with respect to WT. In our experiments, the lower bound for *GAL3* levels in MA0182 was 36% of maximal WT levels due to leakiness of the *TET* inducible promoter system.

In summary, MA0182 exhibited a bimodal response for a range of WT *GAL3* expression levels and was graded when Gal3p was overexpressed. This transformation of the GAL dose response from bimodal to graded by tuning the concentration of Gal3p corroborates the importance of comparing feedback loop knockouts at similar operating point(s) to fully understand the contribution of these regulatory connections to a phenotype (Section S2.3) [69].



Supplementary Figure 2.2: Experimental characterization of the diploid *GAL3* feedback loop knockout strain MA0182 from [20]. **(A)** Flow cytometry histograms of YFP fluorescence for a range of doxycycline (dox, horizontal axis) and galactose (percent, vertical axis) concentrations. **(B)** Representation of flow cytometry distributions in (A) as bimodal (red) and unimodal (blue) classified using a Gaussian mixture model (see Materials and methods). The concentrations of galactose that yielded bimodal distributions shifted to lower galactose concentrations as the concentration of *GAL3* was increased, qualitatively reflecting the decrease in the bistability region for the *GAL3* feedback deletion model (Fig. 2.5D). The dose response was graded for 50 ng/ml dox. The concentrations of dox that map *GAL3* levels in MA0182 to wild-type (WT) expression are indicated by a green line (0-25 ng/ml dox). **(C)** Quantitative real-time PCR measurements comparing *GAL3* mRNA levels in MA0182 to a diploid WT. This WT strain was induced with 0% and 0.5% galactose and MA0182 was induced with 0, 25 and 50 ng/ml dox. In comparison to WT induced with 0.5% galactose, *GAL3* levels in MA0182 were between 0 (36% with respect to wild-type) and approximately 35 ng/ml dox (100% with respect to WT). Error bars represent one standard deviation (n=3).

S2.3 Comparison of open and closed loop transcriptional circuits

In engineering, closed and open loop systems are frequently compared to determine the advantages of feedback control on performance [14]. Similarly, in biology, a controlled comparison for open and closed loop systems may provide insight about the role of a feedback loop [69]. One approach to creating the open loop system is to delete the gene involved in the loop. However, deleting a gene is an aggressive approach that may significantly shift the operating point of the circuit, making it difficult to attribute the changes in phenotype to the function of the feedback loop. Deleting the coding region of the gene involved in the loop and expressing this gene from a constitutive promoter is a superior approach for evaluating the function of a feedback loop. The constitutive promoter strength is an important parameter to adjust since a comparison of the open and closed loop systems should be made in the neighborhood of the wild-type equilibrium point(s).

Consider a bistable transcriptional circuit modeled by an ordinary differential equation (ODE) that has two stable steady-states for a specific range of an input parameter, u, $\frac{dx}{dt} = u + H(x,\theta) - \gamma x$. We are interested in the role of a positive feedback loop of protein, x. In the closed loop system (wild-type), $H(x,\theta)$ represents transcriptional feedback regulation where

$$H(x,\theta) = \frac{\alpha x^n}{x^n + K^n}.$$

For $u = u_1$, $\frac{dx}{dt} = 0 \Rightarrow x = x_{ei}$ where i corresponds to the particular equilibrium point ($i = 1, 2$ within the bistable parameter regime). Given $u = u_1$ in the bistable region, the open loop system should be evaluated at $\alpha_{OL1} = H(x_{e1},\theta)\big|_{u=u_1}$ and $\alpha_{OL2} = H(x_{e2},\theta)\big|_{u=u_1}$,
where α_{OL1} and α_{OL2} represents the constitutive (open-loop) production rates. Experimentally mapping the open and closed loop production rates for a range of inputs, $u = u_1, \ldots, u_n$, may be challenging due to a limited number of well-characterized constitutive promoters and restricted dynamic ranges of inducible promoter systems. To circumvent this, an intermediate α'_{OL} can be chosen within the WT expression range

$$\alpha'_{OL} \in [H(\min(x_{ei}), \theta), H(\max(x_{ei}), \theta)] \text{ for } u_1, \ldots, u_n.$$

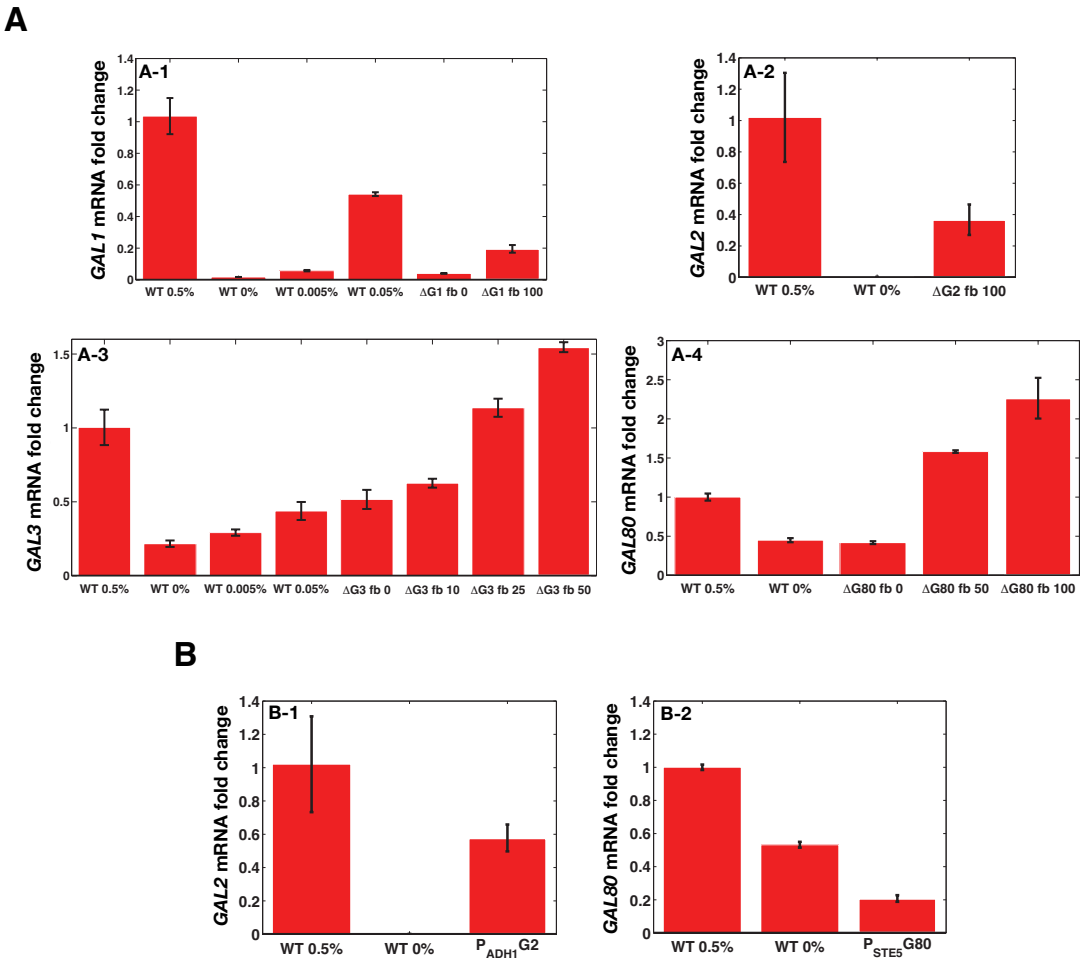
The caveat for this approximation is that α'_{OL} produces a higher and/or lower open loop expression level compared to WT for each value of u. Using this approach, it is therefore important to check that the role of the feedback does not depend on the specific value α'_{OL} by scanning several values within WT range (Supplementary Fig. 2.2 and Supplementary Fig. 2.3).

S2.4 Model description and steady-state solution

An ODE model of the GAL gene-regulatory circuit was constructed based on the interactions shown in Fig. 2.1. This model was able to provide explanations for experimental data and insights about the interplay of feedback loops. We assumed the following:

- Intracellular galactose concentration is constant.
- Since the *GAL2* feedback is not necessary for bimodality, Gal2p was not modeled for simplicity (Fig. 2.3B-1).
- No distinction was made between Gal1p, Gal1p bound to galactose (Gal1p*) and Gal3p, Gal3 bound to galactose (Gal3p*) since both the galactose bound and unbound forms can function as co-inducers of GAL gene expression, presumably with different affinities [25].
- We did not differentiate between nuclear and cytoplasmic partitioning of the GAL proteins because this is a subject of debate [70, 71, 72].
- Dimerization of Gal4p and Gal80p was not modeled for simplicity [73, 74].

For constant galactose concentrations, conversion of Gal1p, Gal3p into Gal1p*, Gal3p* is a first order reaction. This first order reaction was approximated as a zeroth order reaction using a constant input rate (α_{gal}). The protein concentrations of Gal1p (G1), Gal3p (G3), Gal4p (G4) and Gal80p (G80) were modeled. The Hill coefficients for G1 (n_1), G3 (n_3) and G80 (n_{80}) were estimated as 3, 2 and 2 based on experimental measurements (Supplementary Fig. 2.1B).



Supplementary Figure 2.3: Quantitative real-time PCR (qPCR) comparing constitutive and wild-type (WT) mRNA levels of *GAL1* (A-1), *GAL2* (A-2), *GAL3* (A-3), *GAL80* (A-4). The mRNA expression level for each gene was compared to the corresponding expression level of this gene in WT induced with 0.5% galactose. **(A)** Comparison of *TET* promoter and WT expression ranges. *GAL1* expressed from the *TET* promoter and induced with 0-100 ng/ml was within the range of WT *GAL1* expression. However, P_{TET}GAL1 induced with 0 and 100 ng/ml dox was overexpressed relative to WT induced with 0% and 0.005% galactose, respectively. P_{TET}GAL2 induced with 100 ng/ml dox corresponded to 37% of saturated WT *GAL2* levels. WT *GAL3* levels corresponded to 0-20 ng/ml dox for *GAL3* expressed from the *TET* promoter. However, P_{TET}GAL3 induced with 10 ng/ml dox was overexpressed relative to 0.05% galactose. P_{TET}GAL80 induced with 0-25 ng/ml dox corresponded to WT *GAL80* expression. **(B)** mRNA levels of *GAL2* (B-1) and *GAL80* (B-2) regulated by the *ADH1* and *STE5* promoters. *GAL2* and *GAL80* levels were approximately 58% and 20% of the corresponding gene in WT induced with 0.5% galactose, respectively. Error bars represent one standard deviation (n=3).

Based on these assumptions, the model that captures the set of critical molecular interactions for bistability in the wild-type (WT) GAL network is

$$\begin{aligned}
\frac{d[\mathrm{G1}]}{dt} &= \epsilon\alpha_{gal} + \alpha_{G1}\left(\frac{[\mathrm{G4}]^{n_1}}{K_{G1}^{n_1} + [\mathrm{G4}]^{n_1}}\right) - k_{f81}[\mathrm{G1}][\mathrm{G80}] + k_{r81}[\mathrm{C81}] - \gamma_{G1}[\mathrm{G1}],\\
\frac{d[\mathrm{G3}]}{dt} &= \alpha_{gal} + \alpha_{G3}\left(\frac{[\mathrm{G4}]^{n_3}}{K_{G3}^{n_3} + [\mathrm{G4}]^{n_3}}\right) - k_{f83}[\mathrm{G3}][\mathrm{G80}] + k_{r83}[\mathrm{C83}] - \gamma_{G3}[\mathrm{G3}],\\
\frac{d[\mathrm{G4}]}{dt} &= \alpha_{G4} - k_{f84}[\mathrm{G4}][\mathrm{G80}] + k_{r84}[\mathrm{C84}] - \gamma_{G4}[\mathrm{G4}],\\
\frac{d[\mathrm{G80}]}{dt} &= \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^{n_{80}}}{K_{G80}^{n_{80}} + [\mathrm{G4}]^{n_{80}}}\right) - k_{f81}[\mathrm{G1}][\mathrm{G80}] + k_{r81}[\mathrm{C81}]\\
&\qquad - k_{f83}[\mathrm{G3}][\mathrm{G80}] + k_{r83}[\mathrm{C83}] - k_{f84}[\mathrm{G4}][\mathrm{G80}] + k_{r84}[\mathrm{C84}] - \gamma_{G80}[\mathrm{G80}],\\
\frac{d[\mathrm{C81}]}{dt} &= k_{f81}[\mathrm{G1}][\mathrm{G80}] - k_{r81}[\mathrm{C81}] - \gamma_{C81}[\mathrm{C81}],\\
\frac{d[\mathrm{C83}]}{dt} &= k_{f83}[\mathrm{G3}][\mathrm{G80}] - k_{r83}[\mathrm{C83}] - \gamma_{C83}[\mathrm{C83}],\\
\frac{d[\mathrm{C84}]}{dt} &= k_{f84}[\mathrm{G4}][\mathrm{G80}] - k_{r84}[\mathrm{C84}] - \gamma_{C84}[\mathrm{C84}].
\end{aligned}$$

Using the quasi-steady-state assumption, the concentrations of the complexes, Gal1p-Gal80p (C81), Gal3p-Gal80p (C83) and Gal4p-Gal80p (C84) reached their respective equilibria significantly faster the dynamics of G1, G3, G4 and G80 ($\frac{d[\mathrm{C81}]}{dt} = \frac{d[\mathrm{C83}]}{dt} = \frac{d[\mathrm{C84}]}{dt} = 0$), yielding

$$[\mathrm{C81}] = \frac{k_{f81}[\mathrm{G1}][\mathrm{G80}]}{k_{r81} + \gamma_{C81}},\ \ [\mathrm{C83}] = \frac{k_{f83}[\mathrm{G3}][\mathrm{G80}]}{k_{r83} + \gamma_{C83}},\ \ [\mathrm{C84}] = \frac{k_{f84}[\mathrm{G4}][\mathrm{G80}]}{k_{r84} + \gamma_{C84}}.$$

This assumption was used to simplify the system of equations to the following four ODEs

$$\begin{aligned}
\frac{d[\mathrm{G1}]}{dt} &= \alpha_{gal}\epsilon + \alpha_{G1}\left(\frac{[\mathrm{G4}]^{n_1}}{K_{G1}^{n_1}[\mathrm{G4}]^{n_1}}\right) + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],\\
\frac{d[\mathrm{G3}]}{dt} &= \alpha_{gal} + \alpha_{G3}\left(\frac{[\mathrm{G4}]^{n_3}}{K_{G3}^{n_3} + [\mathrm{G4}]^{n_3}}\right) + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],\\
\frac{d[\mathrm{G4}]}{dt} &= \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],\\
\frac{d[\mathrm{G80}]}{dt} &= \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^{n_{80}}}{K_{G80}^{n_{80}} + [\mathrm{G4}]^{n_{80}}}\right) + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}],
\end{aligned}$$

where

$$\omega = \frac{k_{r81}k_{f81}}{k_{r81}+\gamma_{C81}} - k_{f81},\ \ \delta = \frac{k_{r83}k_{f83}}{k_{r83}+\gamma_{C83}} - k_{f83},\ \ \beta = \frac{k_{r84}k_{f84}}{k_{r84}+\gamma_{C84}} - k_{f84}.$$

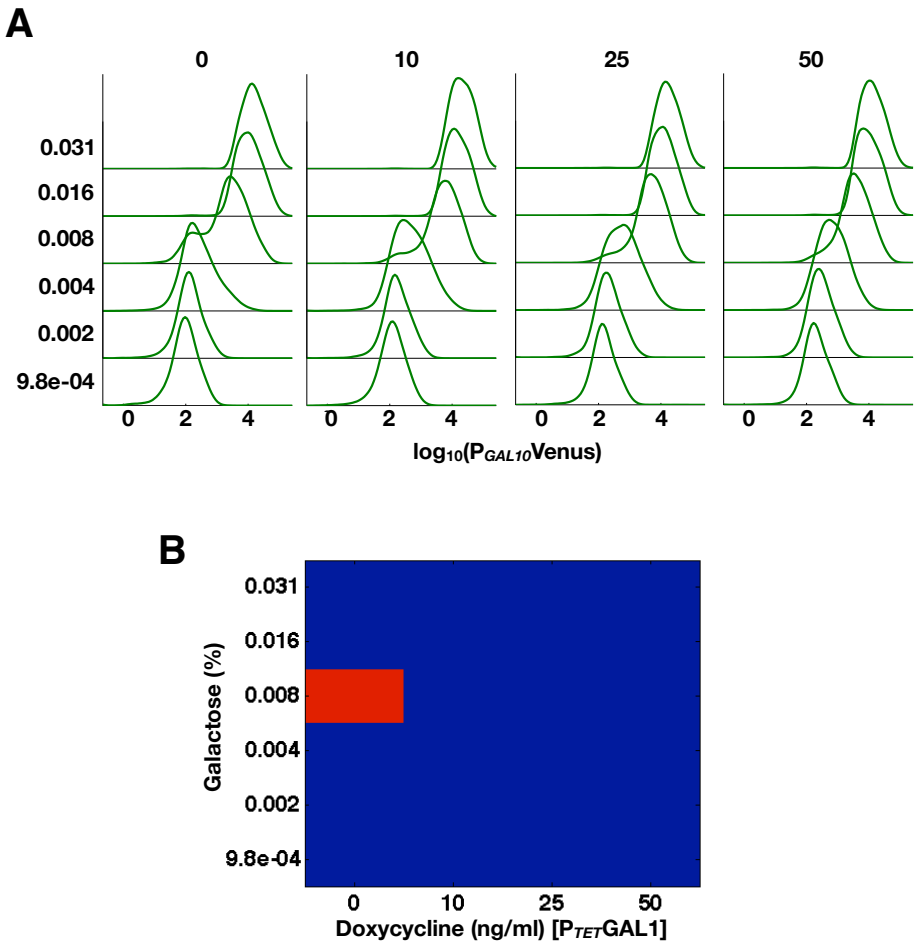
At steady-state, $\frac{d[\mathrm{G1}]}{dt} = \frac{d[\mathrm{G3}]}{dt} = \frac{d[\mathrm{G4}]}{dt} = \frac{d[\mathrm{G80}]}{dt} = 0$ and the equilibrium concentrations are

$$\begin{aligned}
\mathrm{G1}_e &= \frac{-\epsilon\alpha_{gal} - H_{G1}(\mathrm{G4}_e, \theta_{G1})}{\omega \mathrm{G80}_e - \gamma_{G1}}, \\
\mathrm{G3}_e &= \frac{-\alpha_{gal} - H_{G3}(\mathrm{G4}_e, \theta_{G3})}{\delta \mathrm{G80}_e - \gamma_{G1}}, \\
\mathrm{G80}_e &= \frac{-\alpha_{G4} + \gamma_{G4}\mathrm{G4}_e}{\beta \mathrm{G4}_e},
\end{aligned}$$

where $\mathrm{G1}_e, \mathrm{G3}_e, \mathrm{G4}_e$ and $\mathrm{G80}_e$ are the equilibrium values of G1, G3, G80 and G4. $\mathrm{G4}_e$ was determined by computing the roots of a eleventh order polynomial

$$a_0 + a_1\mathrm{G4}_e + \cdots + a_{11}\mathrm{G4}_e^{11} = 0,$$

where the coefficients, a_i, are functions of the model parameters. The *GAL1*, *GAL3* and *GAL80* feedback deletion models listed in Section S2.6 were solved by applying the same procedure. The stability of the equilibrium points was determined by computing the eigenvalues of the Jacobian matrix of the system of equations [14].

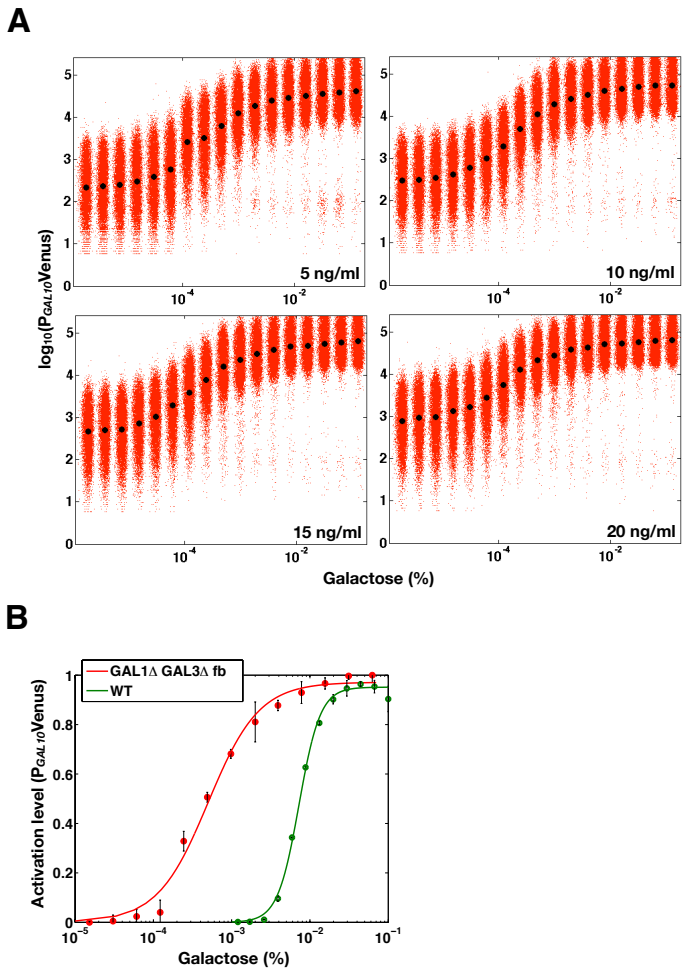


Supplementary Figure 2.4: Experimental characterization of the *GAL1* feedback loop knockout strain (GAL1Δ fb). **(A)** Flow cytometry histograms of P_{GAL10}Venus for a range of doxycycline (dox, horizontal axis) and galactose (percent, vertical axis). **(B)** Representation of flow cytometry data in **A** as bimodal (red) and unimodal (blue) determined by a Gaussian mixture model (see Materials and methods). Bimodality was detected for 0 ng/ml dox and vanished for 10, 25 and 50 ng/ml dox.

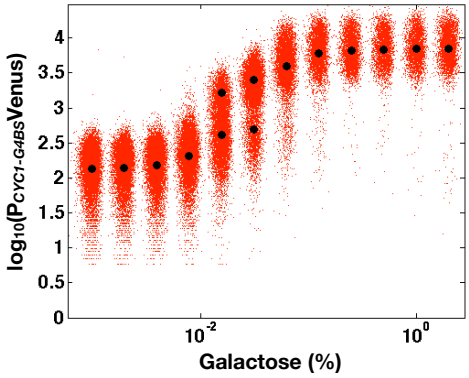
S2.5 Estimation of model parameters

S2.5.1 Parameter set I

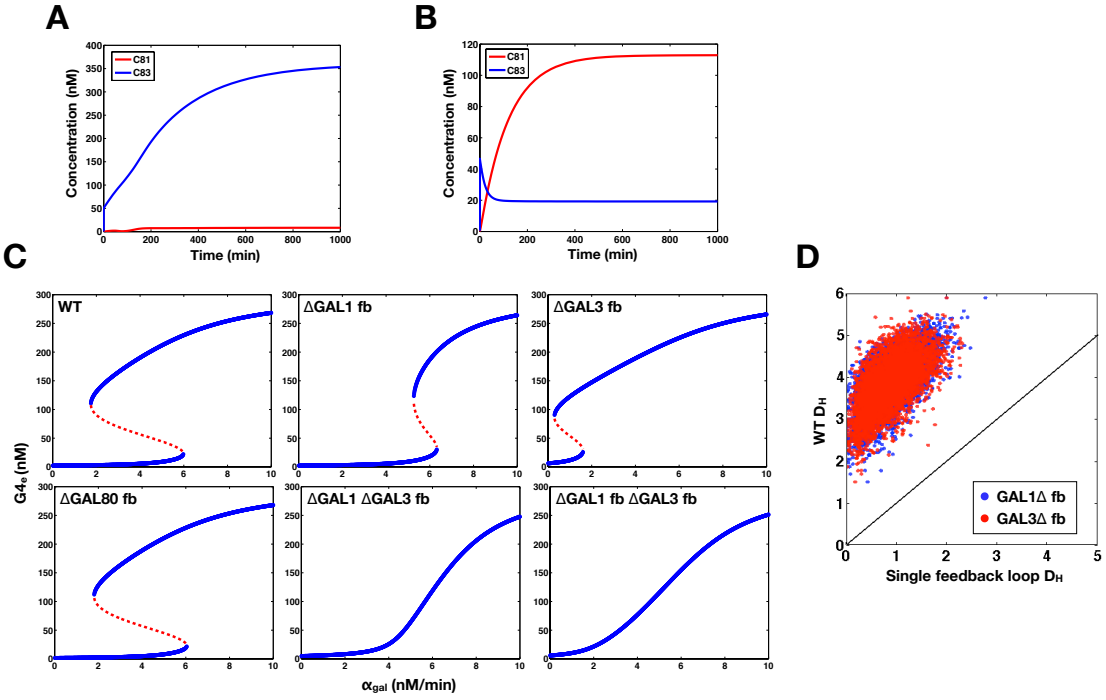
Parameters for the model were estimated from experimental measurements and previous studies (Table SI). *GAL1* and *GAL10* share a bidirectional promoter ($P_{GAL1\text{-}10}$). As a consequence, these genes exhibit highly similar galactose induction responses. The *GAL3* and *GAL80* promoters each have a single Gal4p binding site and are produced at a basal rate in the absence of galactose. Multiple Gal4p binding sites in the *GAL2*, *GAL7* and *GAL1-10* promoters stabilize Gal80p dimers



Supplementary Figure 2.5: Deletion of *GAL1* and the *GAL3* feedback loop (GAL1Δ GAL3Δ fb) produced a graded response irrespective of the concentration of *GAL3*. Bimodality persists in the absence of the individual *GAL2* and *GAL80* feedback loops for a wide range of constitutive *GAL2* and *GAL80* levels. **(A)** Steady-state flow cytometry measurements of P_{GAL10}Venus in GAL1Δ GAL3Δ fb cells where *GAL3* was expressed from a *TET* promoter induced with 5, 10, 15 and 20 ng/ml doxycycline (dox). These concentrations of dox correspond to wild-type *GAL3* levels (Supplementary Fig. 2.3A-3). These measurements were taken on an LSRII analyzer. **(B)** *GAL2* feedback deletion (GAL2Δ fb) displayed bimodality in the absence of doxycycline (dox). **(C)** *GAL80* feedback deletion (GAL80Δ fb) exhibited bimodal distributions for 0 and 25 ng/ml dox. These concentrations of dox correspond to 40% and 100% of fully induced WT *GAL80* mRNA levels, respectively. **(D)** Steady-state activation responses of P_{GAL10}Venus in the WT and GAL1Δ GAL3Δ fb. The Hill coefficient for the WT was approximately 3 whereas the Hill coefficient for GAL1Δ GAL3Δ fb was approximately 1.3. Each data point for GAL1Δ GAL3Δ fb represents the normalized mean of fluorescence (M_Y, see Materials and methods) and the error bars represent one standard deviation (n=3). Each data point for the WT represents mean of the fraction cells in the high expression state and the error bars represent one standard deviation (n=3).



Supplementary Figure 2.6: GAL bimodal response does not require multiple *GAL4* binding sites. A synthetic GAL regulated promoter fusion to Venus with a single *GAL4* binding site was bimodal for two galactose concentrations at steady-state. Bimodality was determined using a Gaussian mixture model (GMM, see Materials and methods). Black circles represent the means of the fluorescence distributions.



Supplementary Figure 2.7: Parameter set II captures dynamic interplay of Gal1p and Gal3p complex with Gal80p (C81 and C83) and feedback loop knockout experimental results. **(A)** Parameter set I indicates that the complex of Gal3p with Gal80p (C83) dominates transiently and at steady-state compared to C81 (complex of Gal1p with Gal80p). **(B)** Random parameter sampling (see Materials and Methods) was used to identify a new parameter set that exhibits transient dominance of the C83 complex and steady-state dominance of the C81 complex. **(C)** Parameter set II qualitatively matches the feedback loop knockout experimental data showing bistability in all feedback loop deletions except the double deletion of *GAL1* and the *GAL3* feedback loop. **(D)** WT model exhibits a larger range of bistability (D_H) compared to the single positive feedback loop systems (*GAL1* and *GAL3*) across a broad region of parameter space ($C_v = 0.1$). Parameter sets I and II are listed in Table SI.

on DNA, augmenting the strength of repression and the maximum production rate [75]. As a result, promoters with multiple Gal4p binding sites have a significantly larger dynamic range of expression.

Flow cytometry measurements of *GAL3*, *GAL10* and *GAL80* promoter fusions to Venus in response to galactose were used to compare relative promoter strengths and cooperativity. The *GAL3* (P_{GAL3}) and *GAL80* (P_{GAL80}) promoter fusions exhibited a graded response whereas the *GAL10* promoter fusion had a bimodal response as shown in Fig. 2.3A-1. The Hill coefficients for the Gal4p dependent feedback terms were approximated by fitting the means of the graded response distributions (M_Y as described in the Materials and methods) and the fraction of high expressing cells for the bimodal response (P_{GAL10}) to Hill functions.

A Hill function fit to the means of the distributions for P_{GAL3}Venus and P_{GAL80}Venus in response to galactose generated Hill coefficients of 2.2 and 2. Gal4p binds to DNA as a dimer and has been shown to interact cooperatively [73, 76]. Based on these results, we assumed that the Hill coefficients for the *GAL3* and *GAL80* transcriptional feedback terms were two [73, 76].

Fitting the fraction of high expression cells for P_{GAL10} produced a Hill coefficient of approximately 3.2. The *GAL1-10* promoter has four Gal4p binding sites which have been shown to increase cooperativity. Therefore, we set the Hill coefficient of the *GAL1* feedback to three [76]. We note that the main conclusions about the roles of the *GAL1*, *GAL3* and *GAL80* feedback loops do not change if the Hill coefficients of the feedback terms for Gal1p, Gal3p and Gal80p and are set to 4, 1, 1 or 3, 1, 1.

The constitutive and feedback production rates were approximated using the number of proteins per cell [66]. Gal4p is weakly expressed and its constitutive production rate (α_{G4}) was selected to reflect this observation [77]. The mean expression levels for P_{GAL3}Venus and P_{GAL80}Venus were similar in response to galactose as shown in Supplementary Fig. 2.1C. At saturation (0.1% galactose), P_{GAL80}Venus was approximately 15% higher than P_{GAL3}Venus. The production rates, α_{G3}, α_{G80}, α_{oG80}, were chosen to have similar ratios to mirror the experimental measurements. Since Gal1p has been shown to bind to Gal80p with lower affinity than Gal3p, a scaling factor of ϵ was used to modify α_{gal} [26, 32].

Parameter	Description	Units	Parameter Set I*	Parameter set II†
k_{f81}	forward binding rate of Gal1p to Gal80p	$(\text{nM min})^{-1}$	100	100
k_{r81}	unbinding rate of Gal1p from Gal80p	min^{-1}	1500	2500
k_{f83}	forward binding rate of Gal3p to Gal80p	$(\text{nM min})^{-1}$	100	100
k_{r83}	unbinding rate of Gal3p from Gal80p	min^{-1}	1	462
k_{f84}	forward binding rate of Gal4p to Gal80p	$(\text{nM min})^{-1}$	100	100
k_{r84}	unbinding rate of Gal4p from Gal80p	min^{-1}	25	1300
α_{G1}	Gal1p production rate	nM min^{-1}	15	35
α_{G3}	Gal3p production rate	nM min^{-1}	0.9	8
α_{G4}	Gal4p production rate	nM min^{-1}	0.2	3.6
α_{oG80}	Basal Gal80p production rate	nM min^{-1}	0.6	5.9
α_{G80}	Gal80p production rate	nM min^{-1}	0.9	9
K_{G1}	*GAL1* transcriptional feedback threshold	nM	8	86.7
K_{G3}	*GAL3* transcriptional feedback threshold	nM	8	64.9
K_{G80}	*GAL80* transcriptional feedback threshold	nM	2	1.5
n_1	*GAL1* Hill coefficient	dimensionless	3	3
n_3	*GAL3* Hill coefficient	dimensionless	2	2
n_{80}	*GAL80* Hill coefficient	dimensionless	2	2
γ_{G1}	Gal1p degradation rate	min^{-1}	0.004	0.0263
γ_{G3}	Gal3p degradation rate	min^{-1}	0.004	0.004
γ_{G4}	Gal4p degradation rate	min^{-1}	0.004	0.0119
γ_{G80}	Gal80p degradation rate	min^{-1}	0.004	0.0073
γ_{C81}	Gal1p-Gal80p (C81) degradation rate	min^{-1}	0.004	0.0084
γ_{C83}	Gal3p-Gal80p (C83) degradation rate	min^{-1}	0.004	0.0527
γ_{C84}	Gal4p-Gal80p (C84) degradation rate	min^{-1}	0.004	0.0177
ϵ	scaling factor	dimensionless	0.1	1.02
α_{G1s}	constant Gal1p production rate	nM min^{-1}	0.1	0.1
α_{G3s}	constant Gal3p production rate	nM min^{-1}	0.1	6
α_{G80s}	constant Gal80p production rate	nM min^{-1}	1.5	13.5

Supplementary Table I: Description of model parameters used in this study. *Original parameter set estimated as described in Section S2.5 (Fig. 2.4 and Fig. 2.5). †New parameter set obtained by random parameter sampling (Latin hypercube method) that captures the dynamic interplay of Gal1p and Gal3p [34].

Forward binding rates (k_{f83}, k_{f81} and k_{f84}) were estimated using the limits of diffusion. The dissociation rates (k_{r83}, k_{r81}, k_{r84}) were free parameters with the requirement that $k_{r81} \gg k_{r83}$ [26, 32]. The protein concentrations of Gal1, Gal3, Gal4, Gal4 and the complexes C81, C83 and C84 were assumed to degrade linearly at approximately the rate of cell division.

S2.5.2 Parameter set II

To identify parameter sets that qualitatively matched the previously reported dynamic switch response of Gal1p and Gal3p [34], 10,000 parameter sets were sampled uniformly in linear scale in the 22-dimensional parameter space, using the Latin hypercube sampling method [45]. The following parameter ranges were used: 10-160 (nM min)$^{-1}$ for the forward binding constants (k_{f81}, k_{f83} and k_{f84}), 1-5000 min^{-1} for the dissociation constants (k_{f81}, k_{f83} and k_{f84}), 0.0035-0.06 min^{-1} for the degradation rates (γ_{G1}, γ_{G3}, γ_{G80}, γ_{G4}, γ_{C81}, γ_{C83} and γ_{C84}), 0.01-100 nM for the EC_{50} values in the Hill functions (K_{G1}, K_{G3} and K_{G80}), 0.1-40 nM min^{-1} for α_{G1}, 0.1-10 nM min^{-1} for α_{G3}, α_{oG80} and α_{G80}, 0.1-5 nM min^{-1} for α_{G4} and 0.01-2 for ϵ. The constitutive rates for the feedback knockouts were fixed at 0.1, 0.1 and 1.5 nM min^{-1} for α_{G1s}, α_{G3s} and α_{G80s}, respectively. D_H was computed for each parameter set and for each of the five models (WT, GAL1Δ fb, GAL3Δ fb, GAL80Δ fb and GAL1Δ fb GAL3Δ fb). First, these parameter sets were filtered based on the presence of bistability. Second, the subset of parameter sets that satisfied these constraints were simulated and the relative concentrations of C81 and C83 were calculated at an initial (10 min) and delayed (500 min) time point by simulation of the full WT model (7-state) before applying the quasi-steady-state assumption.

S2.6 Feedback loop deletion models

The individual *GAL1*, *GAL3*, *GAL80* and combined *GAL1* and *GAL3* feedback deletions were obtained by replacing the appropriate Hill functions representing transcriptional regulation by Gal4p with a constant or equivalently setting the appropriate thresholds in the Hill functions, K_{Gx}, $x = 1, 3, 80$ to zero. In the GAL80Δ fb model, the basal and constitutive production rate were lumped

into one parameter, α_{G80s}. The GAL1Δ GAL3Δ fb model was obtained from the GAL1Δ fb GAL3Δ fb model by setting $\alpha_{G1s} = 0$ and $\epsilon = 0$. The set of ordinary differential equations to model the five feedback loop knockout topologies are as follows.

S2.6.1 *GAL1* feedback deletion (GAL1Δ fb)

$$\frac{d[\mathrm{G1}]}{dt} = \epsilon\alpha_{gal} + \alpha_{G1s} + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],$$
$$\frac{d[\mathrm{G3}]}{dt} = \alpha_{gal} + \alpha_{G3}\left(\frac{[\mathrm{G4}]^2}{K_{G3}^2 + [\mathrm{G4}]^2}\right) + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],$$
$$\frac{d[\mathrm{G4}]}{dt} = \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],$$
$$\frac{d[\mathrm{G80}]}{dt} = \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^2}{K_{G80}^2 + [\mathrm{G4}]^2}\right) + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}],$$

S2.6.2 *GAL3* feedback deletion (GAL3Δ fb)

$$\frac{d[\mathrm{G1}]}{dt} = \epsilon\alpha_{gal} + \alpha_{G1}\left(\frac{[\mathrm{G4}]^3}{K_{G1}^3 + [\mathrm{G4}]^3}\right) + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],$$
$$\frac{d[\mathrm{G3}]}{dt} = \alpha_{gal} + \alpha_{G3s} + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],$$
$$\frac{d[\mathrm{G4}]}{dt} = \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],$$
$$\frac{d[\mathrm{G80}]}{dt} = \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^2}{K_{G80}^2 + [\mathrm{G4}]^2}\right) + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}],$$

S2.6.3 *GAL80* feedback deletion (GAL80Δ fb)

$$\frac{d[\mathrm{G1}]}{dt} = \epsilon\alpha_{gal} + \alpha_{G1}\left(\frac{[\mathrm{G4}]^3}{K_{G1}^3 + [\mathrm{G4}]^3}\right) + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],$$
$$\frac{d[\mathrm{G3}]}{dt} = \alpha_{gal} + \alpha_{G3}\left(\frac{[\mathrm{G4}]^2}{K_{G3}^2 + [\mathrm{G4}]^2}\right) + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],$$
$$\frac{d[\mathrm{G4}]}{dt} = \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],$$
$$\frac{d[\mathrm{G80}]}{dt} = \alpha_{G80s} + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}],$$

S2.6.4 *GAL1* and *GAL3* feedback deletions (GAL1Δ fb GAL3Δ fb)

$$\frac{d[\mathrm{G1}]}{dt} = \epsilon\alpha_{gal} + \alpha_{G1s} + \omega[\mathrm{G1}][\mathrm{G80}] - \gamma_{G1}[\mathrm{G1}],$$
$$\frac{d[\mathrm{G3}]}{dt} = \alpha_{gal} + \alpha_{G3s} + \delta[\mathrm{G3}][\mathrm{G80}] - \gamma_{G3}[\mathrm{G3}],$$
$$\frac{d[\mathrm{G4}]}{dt} = \alpha_{G4} + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G4}[\mathrm{G4}],$$
$$\frac{d[\mathrm{G80}]}{dt} = \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^2}{K_{G80}^2 + [\mathrm{G4}]^2}\right) + \omega[\mathrm{G1}][\mathrm{G80}] + \delta[\mathrm{G3}][\mathrm{G80}] + \beta[\mathrm{G80}][\mathrm{G4}] - \gamma_{G80}[\mathrm{G80}].$$

S2.7 General models of molecular sequestration with positive feedback

We constructed generalizable models of molecular sequestration and positive feedback to examine the relationship between the binding affinity of the activator-repressor pair(s) and the system's region of bistability. We first explored the parameter dependence of a simple model of an activator x that is regulated by a transcriptional repressor z with a Hill coefficient of 1 (noncooperative). In this model, x can sequester z to form an inactive heterodimer, hence generating a positive feedback loop.

Next, we analyzed the steady-state and dynamic properties of systems with two activators, x_1 and x_2 that are each regulated by the transcriptional repressor z and can sequester z into two inactive complexes (c_1 and c_2), thus forming one or two positive feedback loops. In these models, the mechanisms of sequestration and positive feedback are triggered by an input (u) that represents a basal production rate of x, x_1 and x_2.

S2.7.1 Model description for single noncooperative sequestration feedback loop

The three equations that implement a single noncooperative sequestration feedback loop (Supplementary Fig. 2.8A,B) are

$$\frac{dx}{dt} = u + \frac{\alpha K}{K+z} - k_f xz + k_r c - \gamma_x x,$$
$$\frac{dc}{dt} = k_f xz - k_r c - \gamma_c c,$$
$$\frac{dz}{dt} = \alpha_z - k_f xz + k_r c - \gamma_z z.$$

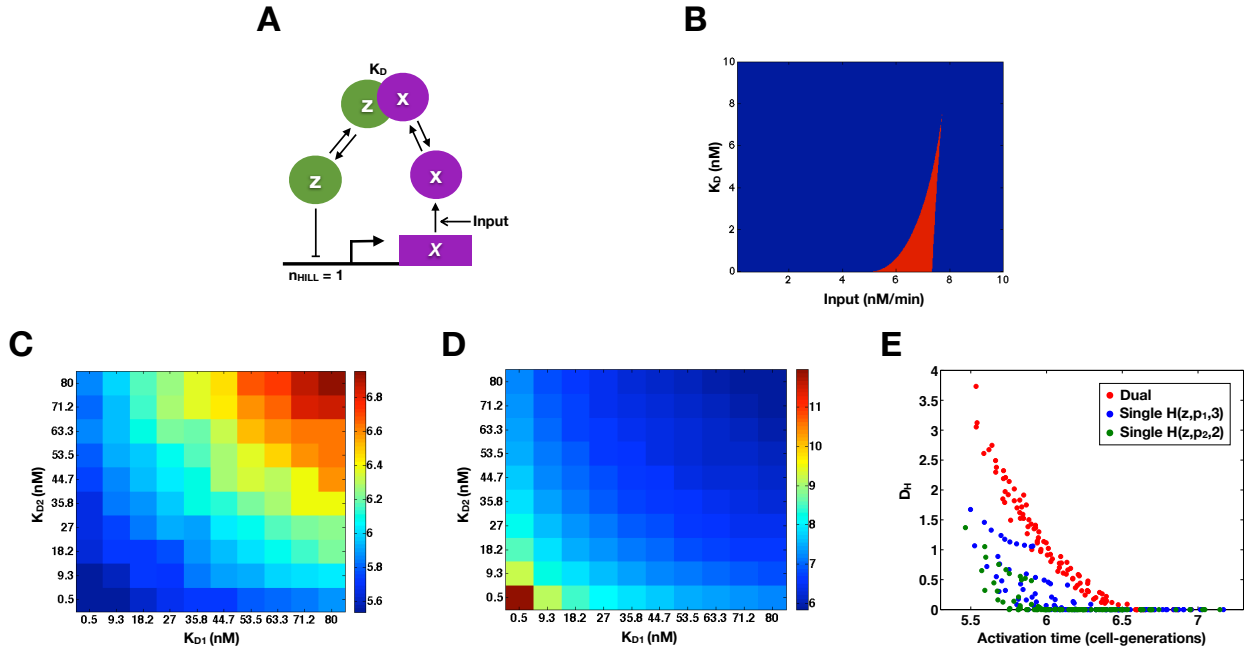
Here, u represents the input. Assuming that $\frac{dc}{dt} = 0$ (quasi-steady-state approximation), the model was reduced to

$$\frac{dx}{dt} = u + \frac{\alpha K}{K+z} + \beta xz - \gamma_x x,$$
$$\frac{dz}{dt} = \alpha_z + \beta xz - \gamma_z z,$$

where $\beta = k_f \left(\frac{k_r}{k_r+\gamma_c} - 1\right)$. The parameter values were set to $\alpha_x = 5$ nM min^{-1}, $\alpha_z = 10$ nM min^{-1}, $\gamma_x = \gamma_c = \gamma_z = 0.005$ min^{-1}, $K = 100$ nM, $k_f = 100$ (nM min)$^{-1}$. u and k_r varied within the range of 0.1-10 nM min^{-1} and 0.1-1000 min^{-1} (Supplementary Fig. 2.8B). Similar to the GAL model, a bifurcation analysis was performed by computing the roots of the cubic polynomial in z.

S2.7.2 Model description for double sequestration linked feedback loops

The ODE model that represents a double sequestration linked feedback loop system shown in Fig. 2.6A consists of the following equations



Supplementary Figure 2.8: Molecular sequestration can generate bistability without cooperativity. Relationships between the binding affinities and activation or deactivation response times for the dual feedback loop sequestration model. **(A)** Circuit topology consists of an activator x that can form inactive heterodimers with a transcriptional repressor, z. z transcriptionally represses x with a Hill coefficient of 1 (noncooperative). **(B)** Regions of bistability (red) and monostability (blue) for a set of input and K_D values (binding affinity of x to z). The region of bistability shrinks and eventually disappears as the binding affinity decreases. Model equations and parameter values are listed in Section S2.7. **(C)** Activation response times measured in cell-generations for the double feedback loop sequestration model for different values of K_{D1} and K_{D2}. **(D)** Deactivation response times measured in cell-generations for the double feedback loop sequestration model for different values of K_{D1} and K_{D2}. **(E)** Relationship between activation response times and range of bistability (D_H) for the double positive feedback loop sequestration model for a set of K_{D1} and K_{D2} values. For a fixed nonzero D_H, the double positive feedback loop system could exhibit a faster activation response compared to the single positive feedback loop models.

$$\frac{dx_1}{dt} = u + \frac{\alpha_1 K_1^3}{K_1^3 + z^3} - k_{f1} x_1 z + k_{r1} c_1 - \gamma_1 x_1,$$
$$\frac{dx_2}{dt} = u + \frac{\alpha_2 K_2^2}{K_2^2 + z^2} - k_{f2} x_2 z + k_{r2} c_2 - \gamma_2 x_2,$$
$$\frac{dc_1}{dt} = k_{f1} x_1 z - k_{r1} c_1 - \gamma_{c1} c_1,$$
$$\frac{dc_2}{dt} = k_{f2} x_2 z - k_{r2} c_2 - \gamma_{c2} c_2,$$
$$\frac{dz}{dt} = \alpha_z - k_{f1} x_1 z + k_{r1} c_1 - k_{f2} x_2 z + k_{r2} c_2 - \gamma_z z.$$

Here, u represents the input. Assuming the inactive complexes (c_1 and c_2) approach equilibrium significantly faster than the other species (quasi-steady-state approximation), the system of equations was reduced to

$$\frac{dx_1}{dt} = u + \frac{\alpha_1 K_1^3}{K_1^3 + z^3} + \beta_1 x_1 z - \gamma_1 x_1,$$
$$\frac{dx_2}{dt} = u + \frac{\alpha_2 K_2^2}{K_2^2 + z^2} + \beta_2 x_2 z - \gamma_2 x_2,$$
$$\frac{dz}{dt} = \alpha_z + \beta_1 x_1 z + \beta_2 x_2 z - \gamma_z z,$$

where $\beta_1 = k_{f1}\left(\frac{k_{r1}}{k_{r1}+\gamma_{c1}} - 1\right)$ and $\beta_2 = k_{f2}\left(\frac{k_{r2}}{k_{r2}+\gamma_{c2}} - 1\right)$.

The parameter values were set to $\alpha_1 = \alpha_2 = 5$ nM min^{-1}, $\alpha_z = 10$ nM min^{-1}, $\gamma_1 = \gamma_2 = \gamma_{c1} = \gamma_{c2} = \gamma_z = 0.005$ min^{-1}, $K_1 = K_2 = 100$ nM, $k_{f1} = k_{f2} = 100$ (nM min)$^{-1}$. u, k_{r1} and k_{r2} were each varied over a range of values. The single feedback loop models referred to as the "single $\mathrm{n_{Hill}} = 2$" and the "single $\mathrm{n_{Hill}} = 3$" were obtained by replacing the Hill functions by a constitutive production rate, α_{1s} or α_{2s}. These parameters were set to $\alpha_{1s} = 5$ nM min^{-1}, $\alpha_{2s} = 5$ nM min^{-1} or $\alpha_{1s} = 0.1$ nM min^{-1}, $\alpha_{2s} = 0.1$ nM min^{-1} for the activation or deactivation response time analysis (see below). Similar to the GAL model, a bifurcation analysis was performed by calculating the roots of polynomials in z. The dual, single $\mathrm{n_{Hill}} = 2$ and single $\mathrm{n_{Hill}} = 3$ models were simplified to seventh, fifth and sixth order polynomials in z.

Response time analysis

The activation and deactivation response times were computed by simulation of the full sequestration models described above before applying the quasi-steady-state approximation (six-state ODE model including an output species, y). Total simulation time was 5000 min. The equation for the output species was

$$\frac{dy}{dt} = \frac{\alpha_y K_y^3}{K_y^3 + z^3} - \gamma_y y,$$

and the parameters equaled $\alpha_y = 10$ nM min^{-1}, $K_y = 100$ nM, $\gamma_y = 0.005$ min^{-1}.

Activation response times

For the activation time simulations, $u = 0$ for $t \leq 500$ min and then $u = 10$ for $t > 500$ min. The initial conditions approximated the steady-state concentrations for the low state where $\mathrm{x}_{1o} = 0.005$ nM, $\mathrm{x}_{2o} = 0.01$ nM, $\mathrm{c}_{1o} = 0.12$ nM, $\mathrm{c}_{2o} = 2.4$ nM, $\mathrm{z}_o = 1998$ nM and $\mathrm{y}_o = 0.26$ nM. The time required for y (normalized between 0 and 1) to increase to half its maximum value was computed for each set of K_{D1} and K_{D2} values. In the single feedback loop models, the constitutive production rates of x_1 or x_2 (α_{1s} or α_{2s}) were set to 0.1 nM min^{-1} because this value approximated the Hill functions at equilibrium for $u = 0$.

Deactivation response times

For the deactivation time simulations, $u = 10$ for $t \leq 500$ min and then $u = 0$ for $t > 500$ min. The initial conditions approximated the steady-state concentrations for the high state where $\mathrm{x}_{1o} = 1943$ nM, $\mathrm{x}_{2o} = 1891$ nM, $\mathrm{c}_{1o} = 993$ nM, $\mathrm{c}_{2o} = 966$ nM, $\mathrm{z}_o = 40.9$ nM and $\mathrm{y}_o = 1872$ nM. The time required for y (normalized between 0 and 1) to decay to half its maximum value was computed for each set of K_{D1} and K_{D2} values. In the single feedback loop models, the constitutive production rates of x_1 or x_2 (α_{1s} or α_{2s}) were set to 5 nM min^{-1} because this value approximated the Hill

functions at equilibrium for $u = 10$.

Strain name	Genotype
Wild-type (WT) P_{GAL10}Venus	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3
GAL2Δ fb	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL2, ade2::ADE2, his3, GAL2Δ::KAN
GAL3Δ fb	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL3, ade2::ADE2, his3, GAL3Δ::KAN
GAL80Δ fb	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL80, ade2::ADE2, his3, GAL80Δ::HPH
GAL2Δ fb GAL3Δ fb	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL3, ade2::ADE2, his3::HIS3-P_{ADH1}GAL2, GAL3Δ::KAN, GAL2Δ::NAT
GAL2Δ fb GAL3Δ fb GAL80Δ fb	MATα ura3:URA3-P_{TET}GAL3, leu2::LEU2-P_{STE5}GAL80, ade2::ADE2-P_{GAL10}Venus, trp1::TRP1-P_{ADH1}GAL2, his3::HIS3-P_{TEFm4}rtTA-M2, GAL3Δ::KAN, GAL2Δ::NAT, GAL80Δ::HPH
GAL1Δ†	MATα leu2, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3, GAL1Δ
GAL1Δ† fb	MATα leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3, GAL1Δ
GAL1Δ† GAL2Δ fb	MATα leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL2, ade2::ADE2, his3, GAL2Δ::NAT, GAL1Δ
GAL1Δ† GAL3Δ fb	MATα leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL3, ade2::ADE2, his3, GAL3Δ::KAN, GAL1Δ
GAL1Δ† GAL80Δ fb	MATα leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL3, ade2::ADE2, his3, GAL80Δ::HPH, GAL1Δ
WT P_{GAL3}Venus	MATa leu2, trp1::TRP1-P_{GAL3}Venus, ura3, ade2::ADE2, his3
WT P_{GAL80}Venus	MATa leu2, trp1::TRP1-P_{GAL80}Venus, ura3, ade2::ADE2, his3
WT P_{TDH3}Venus	MATa leu2, trp1::TRP1-P_{TDH3}Venus, ura3, ade2::ADE2, his3
WT $P_{CYC1\text{-}G4BS}$Venus	MATa leu2, trp1::TRP1-$P_{CYC1\text{-}G4BS}$Venus, ura3, ade2::ADE2, his3
Wild-type diploid	MATa/α leu2/leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus/trp1, ura3/ura3, ade2::ADE2/ade2, his3/his3
MA0182*	MATa/α, ura3/ura3::URA3-P_{TETO2}GAL3, his3::HIS3/his3, ade2::ADE2-P_{MYO2}rtTA/ade2::ADE2-P_{GAL1}YFP, GAL3Δ::KAN/GAL3Δ::KAN

Supplementary Table II: Strains used in this study. All strains were W303. †Constructed using CSY53 background described in [28]. *Strain described in [20].

Chapter 3

Transient bistability generates anticipatory and deferred metabolic states

Introduction

To survive and reproduce, single microbial cells must integrate numerous extracellular signals to infer the state of their environment and respond by adjusting intracellular regulatory programs. These single-cell decisions have been shaped by evolution to weigh the costs and benefits of a particular response. Cellular decision-making relies on current information about the environment and may also reflect an anticipatory response to future changes in these conditions [78, 79]. Although suboptimal in the present environment, preemptive actions may provide crucial benefits to populations of cells upon a shift in the environmental state.

Single-cells deduce the state of their environment by interacting with extracellular signals. These molecular interactions are governed by the inherent stochastic nature of chemical reactions and diffusion, which can generate significant cell-to-cell variability across an isogenic population [36]. This diversification of phenotypes can lead to distinct functional consequences for individual cells [57]. For example, in response to an antibiotic stress, a fraction of *Escherichia coli* (*E. coli*) cells can survive in a persistence state without acquiring genetic resistance to the drug [8]. Non-genetic phenotypic heterogeneity such as persistence has been shown to confer crucial advantages for microorganisms by reducing the variance in population fitness in a fluctuating environment over time [80, 5].

In their natural environments, microbes grow on a variety of different carbon sources as opposed to a single carbon substrate [81]. In cultures with two carbon sources, microbial cells exhibit a diauxic or two-stage growth response in which the substrate that supports the highest growth rate (typically glucose) is consumed before a less preferred sugar [82, 83]. These two growth phases are separated by a delay of diminished growth referred to as the diauxic shift that involves induction of the regulatory and enzymatic pathway for the second carbon source.

As a model system, we used the *Saccharomyces cerevisiae* (*S. cerevisiae*) galactose gene-regulatory network (GAL) to explore how sugar deprived cells respond dynamically to a step input of glucose and galactose. The GAL pathway is a well-characterized eukaryotic gene-regulatory circuit that provides cells with the capability to metabolize galactose. This network includes a set of regulatory genes for sensing and controlling gene expression (Gal2p, Gal3p, Gal80p, Gal4p) and a set of enzymatic genes (Gal1p, Gal7p, Gal10p) for transforming galactose into glucose-6-phosphate as an entry point for glycolysis.

In the absence of galactose, the repressor Gal80p sequesters the transcriptional activator Gal4p from the general transcriptional machinery by binding to activation domain of Gal4p. Galactose enters the cell through the membrane-bound permease transporter Gal2p and activates the signal transducers Gal1p and Gal3p [23]. Upon activation by galactose and ATP, Gal1p and Gal3p can sequester Gal80p, which relieves the repression of Gal4p thus leading to GAL gene induction [24]. Gal1p is a bifunctional molecule and also performs the first step in galactose metabolism by phosphorylating galactose [25].

The regulatory proteins, Gal1p, Gal2p, Gal3p and Gal80p are induced in the presence of galactose by Gal4p, forming four feedback loops. These feedback loops have been shown to reduce gene expression noise, enhance response time, generate bistability and modulate the system's memory of previous environments [19, 20, 84]. The potential for bistability in the GAL network is established by the *GAL1* and *GAL3* feedback loops, which can produce two stable on and off-states in response to a range of galactose levels [19, 18].

Glucose repression is a complex regulatory network that modifies the activity of the GAL pathway

at both the transcriptional and post-transcriptional levels [83]. Transcriptional repression of the GAL genes is mediated by a set of DNA binding proteins that recruit the global transcriptional repression complex Cyc8-Tup1, which modifies the promoter state using multiple mechanisms such as chromatin remodeling [85, 86, 87, 88, 89]. In response to galactose, the GAL promoters can escape a repressed state by the liberation of Gal4p from Gal80p and inhibition of Cyc8-Tup1, which then initiates an ordered recruitment of the SAGA complex and TBP [90]. A PHD domain protein, Cti6p, has been shown to play an important role in overcoming the transcriptional repression of the *GAL1* promoter by Cyc8-Tup1 [91].

The bifunctional glucose kinase, Hxk2p, also has a critical role in glucose repression by regulating the nuclear to cytoplasmic ratio of Mig1p by blocking phosphorylation through the Snf1 kinase [92, 93]. In addition, Hxk2p has a glucose-dependent nuclear localization and physically interacts with Mig1p to regulate a set of target genes [94, 95, 96]. A previous study showed that the recruitment of Cyc8-Tup1 to the *GAL1* promoter, as opposed to the changes in the localization of Mig1, was the dominant regulatory interaction that triggers repression of the GAL system [97].

In this work, we investigated the dynamics of the GAL network in single-cells in response to a mixture of glucose and galactose. We found that similar concentrations of these sugars generated co-existing on and off-subpopulations that persisted for many cell-generations and eventually converged onto a single monomodal on-state. Even though a fraction of the cells are highly expressing the GAL genes, our data indicate that glucose is fully consumed before galactose. We identify the mapping between the population's metabolic state and the dynamics of gene expression, which indicates that the delayed activation precisely coincides with the metabolic shift from glucose to galactose. Using a computational model, we show that the observed transient bimodality originates from bistability that is eventually transformed into monostability when the concentration of glucose crosses a bifurcation point. By modifying the timing of gene expression relative to glucose consumption, we demonstrate that turning on the GAL genes many hours before they are required reduces the amount of time required to transition between carbon sources and provides a transient enhancement of fitness. However, our data demonstrates that constitutive GAL gene expression can reduce the

glucose consumption rate. Based on this tradeoff and evidence that the two subpopulations grow at different rates, we propose that the diversification of the population into two regulatory states provides a beneficial bet-hedging strategy for the cell population.

Results

The GAL system exhibits a transient bimodal response to a step input of similar concentrations of glucose and galactose

We explored the dynamic response of cells grown in rich media without any sugars to a mixture of glucose and galactose. To investigate the gene expression dynamics over a long time scale, we used automated flow cytometry to measure the single-cell fluorescence of a *GAL10* promoter driving Venus (YFP) approximately every 20 min for 14 hours (Fig. 3.1A) [Zuleta, I *et al.* 2013 *In preparation*]. For galactose concentrations significantly higher than glucose, the GAL genes turned on as a single monomodal distribution. By contrast, the GAL pathway did not induce over the course of the experiment for glucose concentrations significantly higher than galactose.

In response to similar concentrations of the two sugars, a fraction of the population turned on the GAL genes whereas the remaining fraction of cells persisted in an off-state, producing a time-dependent bimodality in the activity of the pathway across the population (Fig. 3.1A,B). Following variable time delays, the repressed subpopulation decisively converged in a narrow window of time to a monomodal on state. It should be noted that these synchronized single-cell dynamics are distinct from stochastic switching between two phenotypic states over time [98, 99]. The time-dependent bimodality was detected in over 35% of conditions tested, representing a significantly larger region of parameter space than observed previously in the absence of glucose [19, 20].

To understand the relationships between initial concentration of the two inputs and the dynamic features of the system, we quantified the response times, duration of bimodality and the fraction of cells in the on-state. In conditions that produced a detectable early activated subpopulation, glucose increased the response time of this cohort of cells whereas galactose decreased this response

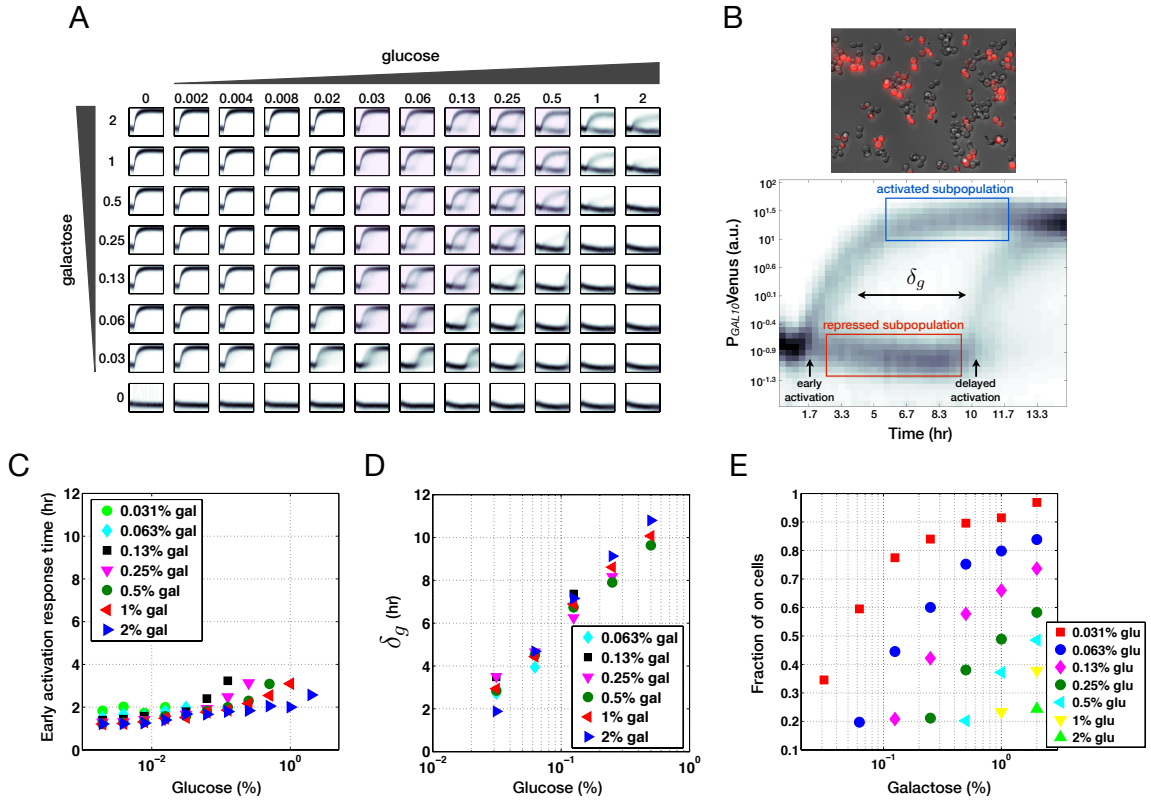


Figure 3.1: The galactose regulatory network in *S. cerevisiae* exhibited transient bimodality for similar concentrations of glucose and galactose. **(A)** Single-cell fluorescence distributions of Venus (YFP) driven by a *GAL10* promoter in WT *S. cerevisiae* obtained using dynamic automated flow cytometry [Zuleta, I *et al.* 2013 *In preparation*] in response to a step input of the two sugars. Dynamics of pathway activation in single-cells across many combinations of the two sugar inputs (right). Concentrations of galactose that significantly exceeded glucose produced monomodal induction and concentrations of glucose significantly higher than galactose generated monomodal repression. Cells exposed to similar concentrations of glucose and galactose displayed transient bimodality (diagonal). The delays between the first and second activation responses could be quantified for a subset of conditions (pink). **(B)** Microscopy image showing bimodality following a 6 hr induction with 0.25% glucose + 0.5% galactose (top). Representative data highlighting the transient bimodality phenotype (bottom). δ_g represents the time delay between the first activation and second activation responses (see Materials and Methods) for conditions highlighted in **A** (pink). **(C)** Response time of early activated subpopulation quantified using the half-max of the mean of high mode (see Materials and Methods). **(D)** δ_g increased with glucose and was not significantly modified by the initial galactose level (Supplementary Fig. 3.1A).**(E)** The fraction of cells in the early activated subpopulation at the midpoint of the transient bimodal region increased with galactose and decreased with glucose for a fixed concentration of galactose.

time (Fig. 3.1C). The duration of transient bimodality δ_g was computed as the difference between the half-max of the activated and repressed subpopulation using a Gaussian mixture model (GMM) (see Materials and Methods). Our data demonstrated that δ_g increased with the initial glucose concentration and was not significantly modified by the initial galactose concentration (Fig. 3.1D, Supplementary Fig. 3.1A). Whereas the response time of the early cohort of cells was fine-tuned by the initial sugar levels (Fig. 3.1C), δ_g was approximately linear with the initial glucose concentration and increased by up to 500%. The total amount of time the system exhibited bimodality decreased with galactose at low glucose levels by inducing the early activated subpopulation at earlier times (Supplementary Fig. 3.1B).

To determine if δ_g was modulated by the current level of glucose, we measured δ_g as a function of the initial cell density (N_o). By varying N_o, we could determine the relationship between the rate of sugar consumption and δ_g. Our results indicated that N_o was inversely related to δ_g, suggesting that the rate of glucose disappearance from the environment determines the timing of the delayed activation (Supplementary Fig. 3.1C). Next, we compared conditions that received one initial step of glucose to two step inputs separated by approximately 5 hours (Supplementary Fig. 3.1D). We found that all conditions that received two step inputs of glucose exhibited a larger δ_g across different glucose concentrations and initial cell densities (N_o) compared to conditions that received a single step input. It should be noted that the second pulse of glucose did not double the observed δ_g. This observation could be explained by a larger number of cells at the time of the second step of glucose that leads to a faster glucose consumption rate. Intriguingly, these data indicate that activation of the GAL pathway in the repressed subpopulation of cells is sensitive to the instantaneous concentration of glucose and hence δ_g can be adjusted in real time.

We analyzed the fraction of on cells in the early activated subpopulation at the mid-point of the transient bimodal region as a function of the sugars. The fraction of on cells significantly increased with the initial concentration of galactose and was reduced by the initial glucose concentration at a constant level of galactose (Fig. 3.1E). By contrast to δ_g, this quantity was not altered by a second step input of glucose and therefore the fraction of on cells distributed between the on and off states

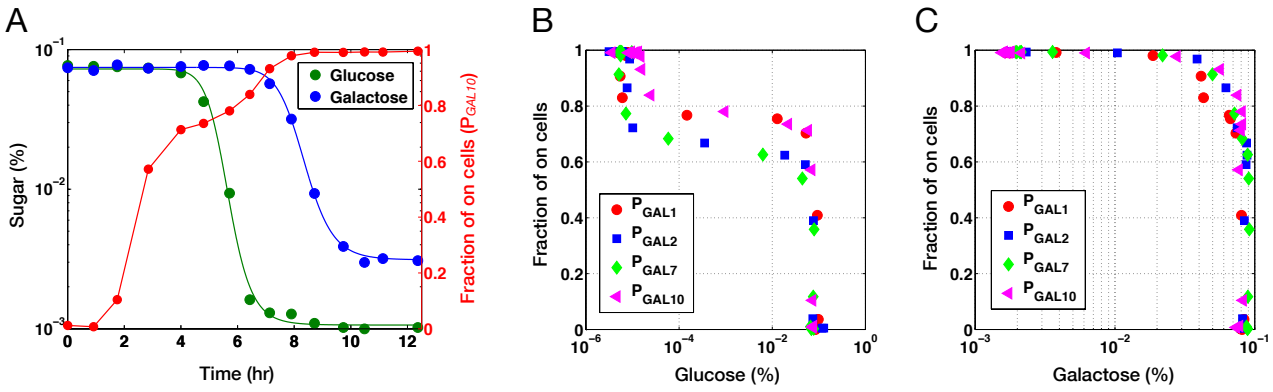


Figure 3.2: Inhibition of galactose metabolism did not require transcriptional repression of the GAL genes and the metabolic transition between carbon sources was tightly coupled to the dynamics of gene expression. **(A)** Representative glucose and galactose concentrations for the *GAL10* promoter fusion strain. Glucose was depleted first before galactose was consumed. Lines represent fitted Hill functions to the sugar data. The fraction of on cells increased immediately following the stimulus, plateaued as cells were consuming glucose and increased to one as the cells were switching from glucose to galactose. **(B)** Scatter plot of glucose concentrations and the fraction of cells in the on state for the four major GAL enzymes. The fraction of on cells increased initially at a constant glucose level, remained constant while glucose was being consumed and increased to one upon glucose depletion. **(C)** Scatter plot of galactose concentrations and the fraction of cells in the high population for the four GAL enzymes. The fraction of on cells increased at constant galactose level and galactose was consumed when approximately 85% cells were in the on-state.

is not modulated by the instantaneous glucose concentration (Supplementary Fig. 3.1E).

The gene expression dynamics and the metabolic transition between carbon sources are tightly coupled processes

To discover the mapping between the gene-expression dynamics and the concentration of the sugars, glucose and galactose were quantified (see Materials and Methods). For each sugar measurement, the single-cell gene expression was quantified for the *GAL1*, *GAL2*, *GAL7* and *GAL10* promoters fusions to Venus (Supplementary Fig. 3.2). These promoters exhibited similar dynamic profiles, indicating that the *GAL10* promoter was a reliable reporter of the network's activity in these conditions. As expected, our data demonstrated a sequential order of sugar utilization in which cells consumed glucose first before starting to deplete galactose from the culture (Fig. 3.2A) [100]. Our results showed that the initiation of galactose consumption was abrupt and switch-like following glucose exhaustion, as opposed to a gradual decrease in the level of galactose as glucose vanished from the culture.

The fraction of on-cells increased immediately following the two sugar stimulus and plateaued as the cells consumed glucose (Fig. 3.2A,B). The delayed activation response of the repressed subpopulation occurred at the time of glucose depletion, which resulted in an increase in the fraction of on-cells. We found that the consumption of galactose did not commence until approximately 85% of cells were in the on-state (Fig. 3.2C). Together, these results showed a precise relationship between the dynamics of gene expression and sugar metabolism. Indeed, both the regulation of gene expression and galactose metabolism were orchestrated together and exhibited switch-like dynamics.

According to the two-sugar Monod model, the expression of genes required for the less preferred substrate are repressed in the presence of the preferred carbon source, which blocks utilization of the secondary sugar. Counter to this model, our data indicates that the inhibition of galactose metabolism does not require transcriptional repression of the GAL genes. This absence of galactose consumption in spite of a subpopulation of cells highly expressing the GAL genes indicates that there is an unknown regulatory mechanism that blocks galactose metabolism. This observation is consistent with a rapid inhibition of galactose consumption in response to a glucose pulse at a timescale faster than can be explained by changes in the transcriptional state and sufficient degradation of the GAL proteins [101].

To assess the contribution of galactose to cellular fitness in the presence of glucose, we measured the sugars, growth and gene expression dynamics for a strain lacking the endogenous *GAL4* gene and expressing a DNA binding mutant of *GAL4* (C14Y) that is unable to activate transcription of the GAL genes [102] (Supplementary Fig. 3.3A). This mutant depleted glucose at a rate equivalent to WT but was not able to consume galactose (Supplementary Fig. 3.3B,C). These data demonstrate the quantitative enhancement of cellular fitness by galactose in the presence of glucose, hence revealing the importance of secondary sugar metabolism in competitive environments (Supplementary Fig. 3.3D).

Constitutive GAL gene expression accelerates the transition between sugars and produces a fitness cost

Our data indicated that a high expression level of the GAL genes was not sufficient to initiate galactose metabolism. To test whether the altering the gene expression state can modify the timing of galactose metabolism in the presence of glucose, we fully induced the system in advance of the wild-type by overexpressing Gal3p and monitored the dynamics of sugar consumption (Supplementary Fig. 3.4A). In this strain, the endogenous *GAL3* gene was deleted and Gal3p was expressed from an inducible *TET* promoter that could be regulated by an aTc-responsive transcription factor rtTA (GAL3Δ fb) [19, 43].

Constitutive GAL gene expression was not sufficient to abolish hierarchical sugar utilization. The GAL3Δ fb strain exhibited an approximately 10% diminished rate of glucose consumption compared to wild-type, indicating that constitutive GAL gene expression can have an inhibitory effect on glucose metabolism (Supplementary Fig. 3.4B). By contrast, galactose metabolism was accelerated in the mutant relative to WT (Supplementary Fig. 3.4C). These combined effects reduced the metabolic delay by approximately 50% compared to WT (Supplementary Fig. 3.4D).

GAL3Δ fb displayed a reduced growth rate on galactose by up to 20% relative to WT (Supplementary Fig. 3.4E, F). These data indicate that gratuitous expression of the GAL genes in environments where these genes are not required yields a fitness cost when growing on galactose. This diminished population fitness and rate of glucose consumption due to GAL gene overexpression provides insights into the benefits of tightly regulating the expression of the GAL genes in response to the availability of galactose in the environment.

Deterministic mathematical model of activation and repression in the GAL system can recapitulate features of dynamics

To understand how the structure of the GAL network could generate the observed dynamics, we constructed a mathematical model of this circuit that received an repression input from glucose and an activation signal from galactose (see Section S3.1 for a detailed description of the model). This

model included a signal transducer Gal1p (G1) that could inhibit the repressor Gal80p (G80) from sequestering the transcriptional activator Gal4p (G4) in the presence of galactose thus leading to GAL gene activation (Fig. 3.3A). A glucose-responsive repressor, R can be activated by the glucose signal, which enables this species to transcriptionally repress the promoters of *GAL1* and *GAL4*.

The model could recapitulate a region of bistability for similar concentrations of the two inputs, which corresponds to the region of bimodality in the experiments (Fig. 3.3B). Removing the *GAL80* feedback loop in the model augmented the percentage of conditions that produce bistability across different values of the two inputs (Supplementary Fig. 3.5A). The model's prediction qualitatively matched the experimental data since removing the *GAL80* feedback loop augmented the range of bimodality (Supplementary Fig. 3.5B-D).

To probe the observed transient nature of the bimodality, we explored the relationship between the hierarchical sugar consumption dynamics and the system's bistability. At time zero, a step input of the two sugars triggers bistability in the system (Fig. 3.3C). Over time, the glucose concentration decays to a critical threshold corresponding to a bifurcation point that produces an abrupt transformation of the system from bistable to monostable. Based on the model, the culture conditions synchronize the cell population dynamics due to the bistable character of the system, which generates a precipitous transition of the repressed subpopulation to the on-state. Precise mechanistic details of glucose repression of the GAL genes were not required to capture the transient bistability phenomenon, suggesting that a general combination of activation and repression of a bistable system can recapitulate the observed dynamics. Experimental characterization of δ_g in a set of mutants revealed that this phenotype is modulated by many factors (Section S3.2, Supplementary Fig. 3.6A-C). In particular, we found that δ_g can be significantly modified by perturbing the transcriptional state of the GAL genes, suggesting that transcriptional regulation has a fundamental role in the precise tuning of the duration of transient bimodality.

The model predicted that the duration of bimodality was inversely related to the current glucose concentration because the system would be closer to the bifurcation point that transforms bistability into a monostable on-state (Fig. 3.3C). We designed an experiment to test this trend predicted by

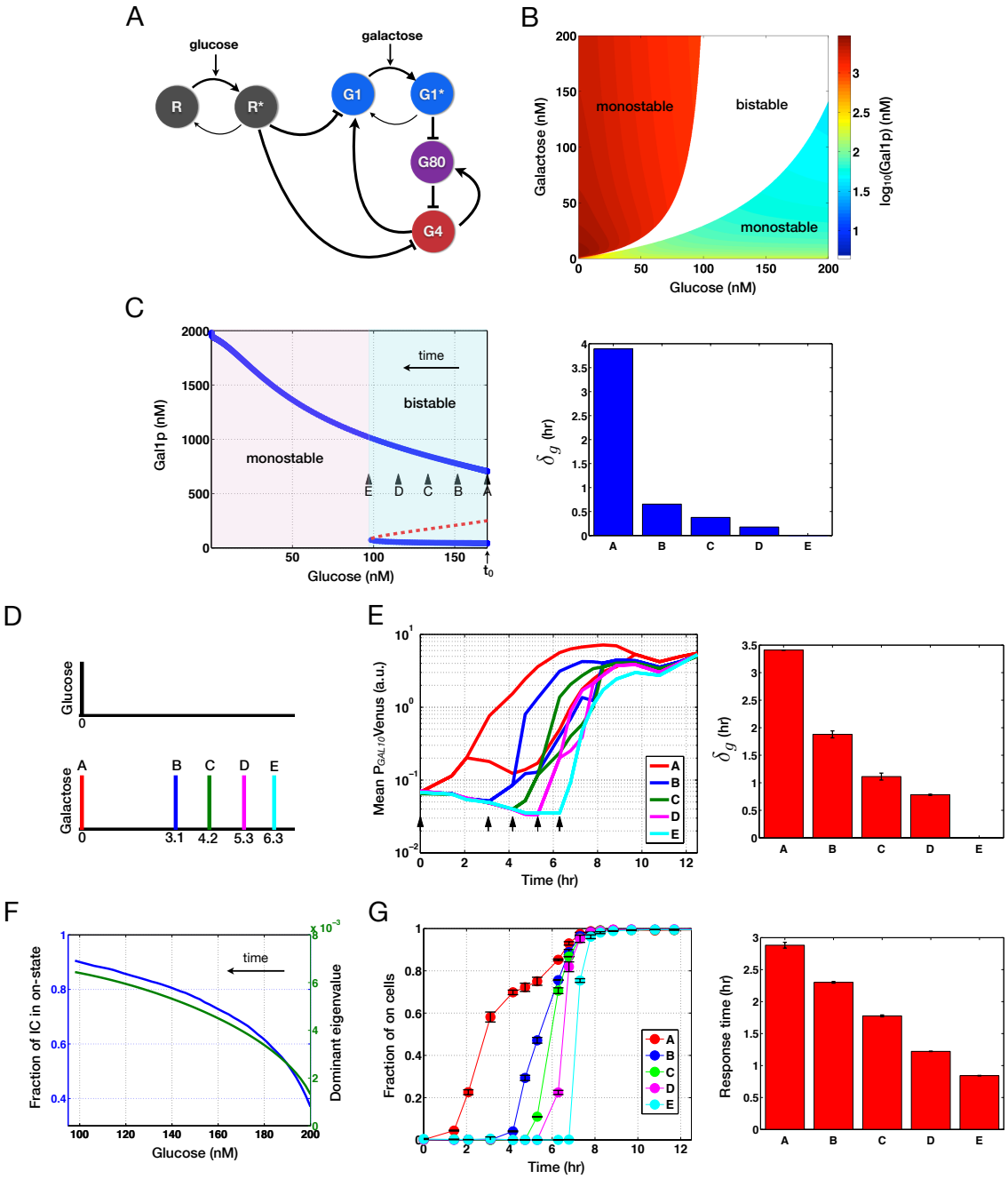

A
glucose
galactose
R
R*
G1
G1*
G80
G4
B
monostable
bistable
monostable
Galactose (nM)
Glucose (nM)
log10(Gal1p) (nM)
C
Gal1p (nM)
time
bistable
monostable
E D C B A
Glucose (nM)
t0
δg (hr)
A B C D E
D
Glucose
Galactose
A B C D E
0 3.1 4.2 5.3 6.3
E
Mean PGAL10Venus (a.u.)
A B C D E
Time (hr)
δg (hr)
A B C D E
F
Fraction of IC in on-state
time
Dominant eigenvalue
Glucose (nM)
G
Fraction of on cells
A B C D E
Time (hr)
Response time (hr)
A B C D E

Figure 3.3

Figure 3.3: Computational model of repression by glucose and activation by galactose recapitulates transient bistable gene expression dynamics and predicts dynamic trends observed in the experimental data. **(A)** Schematic diagram showing the topology of the circuit in the model. Glucose activates a repressor, R to R* that can transcriptionally repress the GAL genes by inhibiting the transcription factor, Gal4p (G4) and the signal transducer, Gal1p (G1). Galactose activates G1 to G1*, which enables G1* to sequester Gal80p (G80). This inhibition of G80 liberates G4 to transcriptionally induce G1 and G80, forming a positive and negative feedback loop. **(B)** The model produces bistability for range of similar concentrations of glucose and galactose (white), a monostable on-state for high galactose concentrations (red) and a monostable off-state for high levels of glucose (blue). **(C)** The transient bimodal response can be explained by an initial bistability in the system triggered by a specific combination of the two inputs. Over time (right to left), glucose decays before galactose is consumed. As a consequence, the system passes through a bistable regime for a range of glucose levels and crosses a bifurcation point at a threshold concentration of glucose. For glucose concentrations below this value, the system exhibits a monostable on-state. In the model, delayed step inputs of galactose to a system that is consuming glucose (highlighted by the arrows) exhibits a shorter duration of bistability (right). **(D)** Experimental design of delayed galactose pulse experiment to test the prediction of the model. 0.1% glucose was added at time zero to conditions A-E and 0.1% galactose was added over a range of times. **(E)** Mean expression level of a *GAL10* promoter fusion to Venus for the galactose pulse experiment (left). Arrows indicate the time when galactose was added to the culture. The duration of bimodality was inversely related to the delay in the galactose input (right). **(F)** The fraction of initial conditions (IC) and dominant eigenvalue decreases as a function of glucose. Here, galactose = 150 nM. **(G)** Fraction of on-cells in the galactose pulse experiment over time for condition A-E. Corroborating the model's prediction, the response time of the fraction of on-cells was inversely related to the delay in the galactose stimulus (right). Arrow bars indicate 1 s.d. from the mean of two replicates.

the model by applying a step input of 0.1% galactose at different times to a set of conditions that had received 0.1% glucose from time zero. In condition A, both sugars were added simultaneously at time zero (Fig. 3.3D). In conditions B-E, glucose was present from time zero and galactose was added 3.1, 4.2, 5.3 and 6.3 hours following the glucose stimulus. As demonstrated in the model, the transient bimodal region contracted with the increased delay in the galactose stimulus and was graded for a sufficiently long delay in condition E (Fig. 3.3E, Supplementary Fig. 3.7).

We analyzed how the domain of attraction and the local stability of the high equilibrium points change as function of the two inputs. The domain of attraction was determined by random sampling of 5000 initial conditions using the Latin hypercube method and computing the fraction of these initial conditions that are absorbed by the high equilibrium state [45]. In the model, the concentration of glucose was inversely related to the domain of attraction and the absolute value of the dominant eigenvalue of the linearization of the high equilibrium state (Fig. 3.3F). Here, the dominant eigenvalue is defined as the eigenvalue with the smallest absolute value. By contrast, galactose increased

the fraction of initial conditions that were assimilated by the high equilibrium point, reflecting the experimental data (Supplementary Fig. 3.8A). Assuming that individual cells in a population have distinct initial conditions that can map to disparate equilibrium states, these results predict a decrease in the response time of the fraction of on cells as glucose decays. To experimentally assess this trend, we analyzed the response time of the fraction of cells in the on-state for conditions A-E in the delayed galactose pulse experiment. Corroborating the results of the model, we found that the time required for the cell population to reach 50% of cells in the on-state decreased with the delay in the galactose stimulus (Fig. 3.3G).

To explore how a single cell could enter the repressed or activated states immediately following the stimulus, we examined the model's dynamics for time varying glucose and galactose concentrations. In these simulations, the concentration of glucose is depleted first before galactose begins to disappear. In response to a combination of the two inputs that generates bistability, our model demonstrated that varying the initial ratio of repressors (R and G80) to activators (G1 and G4) could generate an early or delayed activation of the system (Supplementary Fig. 3.8B). The decision of a single cell to occupy the high or low state immediately following the two input stimulus was investigated experimentally in a panel of mutants (Section S3.1, Supplementary Fig. 3.6D). The majority of these perturbations to the system modified the fraction of on-cells as a function of glucose, indicating that transcriptional regulation has a critical role and that this phenotype is controlled by a complex interplay of the regulatory machinery.

Anticipatory gene expression accelerates galactose metabolism and provides a transient increase in fitness

To investigate whether induction of the GAL genes many cell generations prior to galactose metabolism provides a functional role in the mixed sugar environment, we measured the glucose and galactose concentrations for the delayed galactose pulse experiment. Since the cells consume glucose first before galactose, the presence of galactose should not benefit cells until glucose is fully consumed. In this experiment, we could test whether the timing of GAL gene induction alters the profile of

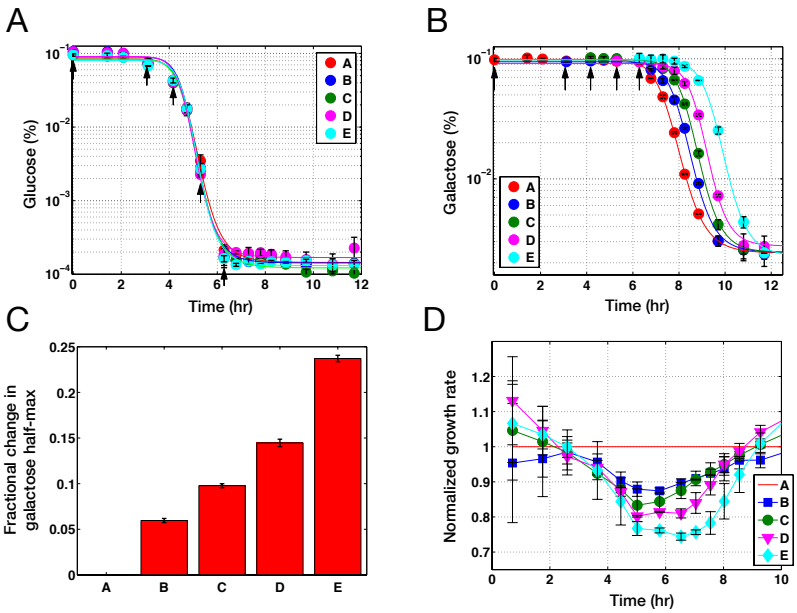


Figure 3.4: Delayed step inputs of 0.1% galactose following an initial step input of 0.1% glucose produces delays in galactose metabolism and a transient decrease in growth rate. Arrows indicate the time when galactose was added to the culture. **(A)** Glucose concentrations for each condition. Lines represent fitted Hill functions to the sugar data. **(B)** Galactose concentrations for each condition. **(C)** Fractional change in the half-max of the galactose concentrations for each condition relative to A. **(D)** Normalized growth rates of conditions B-F compared to A (red line). Error bars represent 1 s.d. from the mean of two replicates.

sugar consumption and the cell population's growth rate.

Our data demonstrated that glucose decayed at a similar rate across all conditions (Fig. 3.4A). However, we observed increasing delays in the consumption of galactose with the delay in the galactose stimulus (Fig. 3.4B,C). During the shift between carbon sources, we observed a transient growth rate advantage of up to 25% between the condition that received galactose at time zero (A) and following a 6.3 hour delay (E) (Fig. 3.4D, Supplementary Fig. 3.9). Taken together, these data indicate turning on the GAL pathway in a fraction of cells in the population before these proteins are used to metabolize galactose, leads to a faster consumption rate of galactose and a transient growth rate advantage during the metabolic transition.

To control GAL gene expression independently of galactose, we tested a synthetic estradiol-responsive Gal4 chimera to induce GAL gene expression at specific times in a strain lacking the endogenous *GAL4* gene [103]. In this experiment, each condition was induced with 0.1% glucose and 0.1% galactose from time zero and the GAL genes were turned on by applying a step input in estradiol

at different times before the cells had depleted glucose from their environment (Supplementary Fig. 3.10A). Since the synthetic inducible system is not connected to the feedback structure of the natural circuit, the gene expression was graded as opposed to bimodal.

Interestingly, our data showed that the delay in the activation of gene expression correlated with a faster glucose consumption rate of up to approximately 15%, indicating that GAL gene expression can inhibit glucose consumption (Supplementary Fig. 3.10B). Consistent with the results of the galactose pulse experiment, we found that turning on the GAL genes at later times delayed galactose metabolism (Supplementary Fig. 3.10C). These combined effects resulted in a significant increase in the delay of the transition between carbon sources of up to approximately 22% comparing condition E to A (Supplementary Fig. 3.10C). Together, these data highlight a potential tradeoff involved in inducing the GAL genes before they are required that arises from the cost of a reduced rate of glucose decay and the benefit of faster galactose consumption, which facilitates the transition between carbon substrates. In addition, this experiment showed that the uncharacterized control mechanism that enforces hierarchical sugar utilization does not arise from the natural regulation of the circuit including the feedback loops, suggesting that this regulatory checkpoint may be upstream at the level of the enzymes.

These two experiments revealed a delay on the timescale of hours for cells to start consuming galactose from the time that the system was turned on by either galactose or estradiol. We found that the duration of the delay in galactose consumption was augmented by turning the system on closer to the point of glucose exhaustion and thus the time when these genes were required (Fig. 3.4B,C and Supplementary Fig. 3.10C). Presumably, this stalled metabolic state is partially due to the time required for sufficient accumulation of the metabolic machinery including transporters and enzymes for galactose utilization. However, these data cannot be entirely explained by the timing of the buildup of necessary proteins since we observed a reverse trend in the gene expression response time that decreased with the increase in the galactose stimulus (Fig. 3.3G).

Cells in high subpopulation have a lower growth rate than cells in low subpopulation

In the transient bimodal regime, the low and high subpopulations were clearly distinguishable for several hours. We assumed that the potential transitions between the two states were insignificant since switching between states would yield intermediate fluorescence levels due to the gradual accumulation or dilution of fluorescent protein. During this period, we computed the growth rate of the individual subpopulations by quantifying the number of cells that accumulated in the low and high modes over time using a thresholding method (see Materials and Methods). Our results showed that the low subpopulation grew on average 15% faster than the high mode across a range of sugar levels (Supplementary Fig. 3.11). This diminished growth rate could be explained a reduced glucose consumption rate in the high mode compared to the low mode.

Discussion

Faced with uncertain and fluctuating environmental stimuli, single microbial cells sense and respond to their current surroundings but can also anticipate a probable future shift in the environmental state. These decision-making strategies can be a deciding factor in competitive environments by providing crucial fitness advantages for microorganisms. We examined how single cells infer the state of their environment and make decisions using a galactose metabolic gene-regulatory network in response to a mixture of carbon sources.

Our dynamic measurements of the GAL system revealed that similar concentrations of glucose and galactose produced two coexisting on and off gene expression states that persisted for a period of time. Following variable delays, cells in the repressed subpopulation abruptly turned on the GAL genes and therefore all cells eventually occupied the same monomodal steady-state. This transient behavior is reminiscent of adaptive cellular stress responses that show significant disparities between the system's initial and long-term behavior [104].

In conditions that produced transient bimodality, the off population was responsive to the current glucose level, but these cells were unaware of the presence of galactose and deferred a commitment

to GAL gene activation over many cell-generations. This delayed decision-making strategy is similar to the timing of developmental decisions such as the postponement of lineage selection by stem-cells until appropriate signals are received and the delayed commitment to sporulation in *B. subtilis* [105, 106].

In homeostatic cellular responses, the mean expression level of specific genes expressed uniformly across the cell population are adjusted to be proportional to the strength of the input [107]. Here, the composition of the sugar mixture adjusts the percentage of cells that induce the GAL genes many hours before galactose is consumed. Although glucose is abundant in some conditions that generated bimodality, a fraction of the population highly expresses the GAL genes many hours prior to the forthcoming transition to galactose metabolism. Advanced preparation to changing conditions may be a prevalent strategy in biological networks [78]. For example, anticipation of future environmental changes has been suggested to determine growth rate in response to specific environmental cues and may have provided critical fitness advantages for microorganisms during evolution [108].

Our systematic characterization of the GAL system in the mixed carbon source environment revealed that there are two layers of regulation that repress galactose utilization in the presence of glucose. The first is a set of mechanisms to repress GAL gene expression and the second is an unknown inhibition of galactose metabolism that depends on the availability of glucose. Interestingly, the release of metabolic inhibition displayed a similar switch-like dynamic response to the derepression of transcriptional regulation even though gene expression is decoupled from metabolism. Constitutive GAL gene expression did not abolish the galactose metabolic repression but reduced the time required to shift between carbon substrates. These data suggest that there is a feedback interaction between the expression state of the pathway and the initiation of metabolism.

We identified a mapping between the single-cell gene expression state, concentrations of sugars and the population's growth rates. Our results revealed that glucose depletion, the onset of galactose metabolism and the delayed gene activation response are tightly coupled dynamic processes. We hypothesize that this coordinated timing is a consequence of a regulatory interconnection between these mechanisms. The ortholog of Gal80p (KlGal80p) in *Kluyveromyces lactis* and Gal80p

in *S. cerevisiae* (ScGal80p) have been shown to inhibit galactokinase activity by interacting with KlGal1p and ScGal1p [29, 109]. It would be interesting to explore whether the dual roles of Gal80p in repressing the transcriptional regulation of the GAL genes and the production of galactose-1-phosphate contributes to the tightly coupling in the observed dynamics of metabolism and gene expression. A previous study showed that a deletion of *HXK2* disrupts the hierarchical order of glucose and galactose metabolism [110]. Our data demonstrated that cells deleted for this gene were not able to induce the repressed subpopulation. Taken together, we hypothesize that Hxk2p is a crucial player in linking the availability of glucose to the repression of galactose consumption and timing of gene induction.

Turning on the GAL genes can generate an energetic cost and the intermediate galactose-1-phosphate is toxic to cells [42, 111]. In addition, we found that the expression of GAL genes can reduce the rate of glucose consumption. On the other hand, turning on the GAL genes just in time for galactose metabolism can significantly augment the metabolic delay of the diauxic shift and transiently decrease cellular fitness. Suboptimal strategies, such as enduring the costs of expressing a pathway when its not required, could be explained by the selection pressure of competitive environments that reward fast cellular responses. For example, there is a fundamental tradeoff due to thermodynamics between the total yield and the rate of production of ATP [112]. In competitive conditions, a rapid rate of ATP production is advantageous even though the total efficiency is reduced. In fact, *S. cerevisiae* uses fermentation as opposed to respiration even in the presence of oxygen, suggesting that competition may have shaped the metabolic strategy of this organism [113].

The GAL bimodal population consists of a fraction of cells that endure the costs of highly expressing the GAL genes in the absence of galactose utilization and cells in the repressed subpopulation that induce GAL gene expression just in time to transition to the secondary carbon source. We propose that the transient bimodality balances the costs of activating the pathway with the benefits of preparing in advance for a future change in glucose availability. Similar to the tradeoff between efficiency and rate of ATP production, this preemptive strategy enabled a faster rate of galactose consumption and a transient enhancement of fitness that may have provided evolutionary advantages

for this organism.

Methods

Growth conditions and flow cytometry

Cells were grown in yeast peptone media for approximately 12-hours prior to induction with glucose and galactose. Single-cell fluorescence was measured on an LSRII analyzer (BD Biosciences). A blue (488 nm) laser was used to excite YFP and emission was detected using a 530/30 nm filter. The number of cells quantified for each dynamic measurement ranged from approximately 1000 to 20,000. Strains used in this study are described in Supplementary Table .

Automated flow-cytometry measurements

A 500 μl culture volume was used in 96-well plate format for the automated flow cytometry measurements. For each time point, 30μl was removed from the culture for measurement on the cytometer and 30μl of fresh media containing the 1X concentration of glucose and galactose was used to maintain a constant culture volume.

Flask measurements

A 60 ml culture volume was used for the experiments in which the sugar concentrations were quantified. Less than 5% of the total volume was removed over the course of the experiment to quantify the single-cell fluorescence, sugar concentrations and OD. Cell mass (OD_{600}) was measured on a Nanodrop 2000c spectrophotometer (Thermo Scientific).

Quantitative analysis of gene expression dynamics

Flow cytometry distributions were analyzed using a Gaussian mixture model algorithm (GMM, MATLAB) and classified as bimodal as described in [19]. The response time was defined as the time to reach the half-max of the fraction of on-cells. The delay time δ_g was computed by subtracting the half-max of the mean of the activated and repressed subpopulations. The fraction of on-cells were quantified at the midpoint of the transient bimodal region.

Sugar measurements

Glucose and galactose were measured using the Amplex Red glucose oxidase and galactose oxidase kits (Molecular Probes, Life Technologies). A Tecan Safire plate reader (Tecan) was used to quantify fluorescence and absorbance. The metabolic lag was determined by computing the difference between the half-max of the galactose decay and the half-max of the glucose decay. In the galactose pulse and estradiol experiments, the delay in galactose metabolism was computed by subtracting the half-max of the galactose decay from the time in which glucose was fully consumed (the time in which the derivative of glucose first crossed zero).

Computational modeling

Code for mathematical modeling was written in MATLAB (Mathworks) and Mathematica (Wolfram Research). The domain of attraction of the equilibrium points were determined by randomly sampling 5000 initial conditions using the Latin Hypercube Method. We computed the fraction of these initial conditions that were absorbed by the high equilibrium point.

Acknowledgements

The authors would like to thank Dr. Adam Rosenthal and Dr. Michael Chevalier for insightful discussions.

Supplementary Information

1. Two sugar mathematical model

2. Characterization of δ_g and fraction of on cells in a set of mutants

S3.1 Two sugar mathematical model

An ODE model of the GAL gene-regulatory circuit was constructed based on the GAL system. This model was able to provide explanations for experimental data and insights about dynamics of the system. We assumed the following:

- No distinction was made between Gal1p and Gal3p since these regulators have the same mechanism for activating the GAL genes by sequestration of Gal80p [25]. Therefore, we assumed in the model that these two species could be represented by one protein, Gal1p.
- Intracellular transport of glucose and galactose through the glucose hexose transporters and Gal2p was not modeled for simplicity.
- Glucose repression of the GAL system could be modeled as a single repressor, R, that is activated by glucose and then transcriptionally represses Gal4p and Gal1p.
- Dimerization of Gal4p and Gal80p was not modeled for simplicity [73, 74].
- We did not differentiate between nuclear and cytoplasmic partitioning of the GAL proteins because this is a subject of debate [70, 71, 72].

The Hill coefficients for G1 (n_{G1}, n_{R1}), G4 (n_{R4}) and G80 (n_{80}) were estimated as 3, 2, 1 and 2.

Based on these assumptions, the model that captures a set of molecular interactions for bistability in the wild-type (WT) GAL network in the presence of glucose and galactose is

$$\frac{d[\mathrm{G1}]}{dt} = \alpha_o + \alpha_{G1}\left(\frac{[\mathrm{G4}]^{n_1}}{K_{G1}^{n_1} + [\mathrm{G4}]^{n_1}}\right)\left(\frac{K_{R1}^{n_{R1}}}{K_{R1}^{n_{R1}} + [\mathrm{Rs}]^{n_{R1}}}\right) - k_{fg}[\mathrm{gal}][\mathrm{G1}] + k_{rg}[\mathrm{G1s}] - \gamma_{G1}[\mathrm{G1}],$$
$$\frac{d[\mathrm{G1s}]}{dt} = k_{fg}[\mathrm{gal}][\mathrm{G1}] - k_{rg}[\mathrm{G1s}] - k_{f81}[\mathrm{G1s}][\mathrm{G80}] + k_{r81}[\mathrm{C81}] - \gamma_{G1s}[\mathrm{G1s}],$$
$$\frac{d[\mathrm{R}]}{dt} = \alpha_R - k_{fR}[\mathrm{glu}][\mathrm{R}] + k_{rR}[\mathrm{Rs}] - \gamma_R[\mathrm{R}],$$
$$\frac{d[\mathrm{Rs}]}{dt} = k_{fR}[\mathrm{glu}][\mathrm{R}] - k_{rR}[\mathrm{Rs}] - \gamma_{Rs}[\mathrm{Rs}],$$
$$\frac{d[\mathrm{G4}]}{dt} = \alpha_{G4}\left(\frac{K_{R4}^{n_{R4}}}{K_{R4}^{n_{R4}} + [\mathrm{Rs}]^{n_{R4}}}\right) - k_{f84}[\mathrm{G4}][\mathrm{G80}] + k_{r84}[\mathrm{C84}] - \gamma_{G4}[\mathrm{G4}],$$
$$\frac{d[\mathrm{G80}]}{dt} = \alpha_{oG80} + \alpha_{G80}\left(\frac{[\mathrm{G4}]^{n_{80}}}{K_{G80}^{n_{80}} + [\mathrm{G4}]^{n_{80}}}\right) - k_{f81}[\mathrm{G1s}][\mathrm{G80}] + k_{r81}[\mathrm{C81}] - k_{f84}[\mathrm{G4}][\mathrm{G80}] + k_{r84}[\mathrm{C84}] - \gamma_{G80}[\mathrm{G80}],$$
$$\frac{d[\mathrm{C81}]}{dt} = k_{f81}[\mathrm{G1s}][\mathrm{G80}] - k_{r81}[\mathrm{C81}] - \gamma_{C81}[\mathrm{C81}],$$
$$\frac{d[\mathrm{C84}]}{dt} = k_{f84}[\mathrm{G4}][\mathrm{G80}] - k_{r84}[\mathrm{C84}] - \gamma_{C84}[\mathrm{C84}].$$

Using the quasi-steady-state assumption, the concentrations of the complexes, G1-gal (G1s), R-glu (Rs), Gal1p-Gal80p (C81) and Gal4p-Gal80p (C84) reached their respective equilibria significantly faster the dynamics of G1, R, G4 and G80. This assumption was used to simplify the system of equations to the following four ODEs

$$\frac{d\mathrm{G1}}{dt} = \alpha_o - \gamma_{G1}\mathrm{G1} - k_{fg}\mathrm{G1gal} + \frac{\mathrm{G1}k_{rg}p_3}{p_4 + p_5\mathrm{R}} + \frac{\alpha_{G1}K_{R1}^2\mathrm{G4}^3}{\left(K_{R1}^2 + \mathrm{R}^2p_2^2\right)\left(K_{G1}^3 + \mathrm{G4}^3\right)},$$
$$\frac{d\mathrm{R}}{dt} = \alpha_R - \gamma_R\mathrm{R} - k_{fR}\mathrm{gluR} + k_{rR}p_2\mathrm{R},$$
$$\frac{d\mathrm{G80}}{dt} = \alpha_{oG80} - \gamma_{G80}\mathrm{G80} - \frac{p_6\mathrm{G1G80}}{p_4 + p_5\mathrm{G80}} + p_7\mathrm{G80G4} + \frac{\alpha_{G80}\mathrm{G4}^2}{K_{R4}^2 + \mathrm{G4}^2},$$
$$\frac{d\mathrm{G4}}{dt} = \frac{\alpha_{G4}K_{R4}}{K_{R4} + p_2\mathrm{R}} - \gamma_{G4}\mathrm{G4} + p_8\mathrm{G80G4}.$$

where $p_0 = \frac{k_{fs}}{\gamma_{C81}+k_{rs}}$, $p_1 = \frac{k_{f84}}{\gamma_{C84}+k_{r84}}$, $p_3 = \frac{\mathrm{glu}k_{fR}}{\gamma_{Rs}+k_{rR}}$, $p_3 = k_{fg}\mathrm{gal}$, $p_4 = \gamma_{G1s}+k_{rg}$, $p_5 = k_{f81}-k_{r81}p_0$, $p_6 = k_{f81}p_3 + k_{r81}p_0p_3$, $p_7 = p_1k_{r84} - k_{f84}$ and $p_8 = p_1k_{r84} - k_{f84}$.

At steady-state, $\frac{d[\mathrm{G1}]}{dt} = \frac{d[\mathrm{R}]}{dt} = \frac{d[\mathrm{G4}]}{dt} = \frac{d[\mathrm{G80}]}{dt} = 0$ and we could solve for the equilibrium concentrations of G1, R and G4 in terms of G80. The equilibrium value of G80, $\mathrm{G80}_e$ was determined by

computing the roots of a ninth order polynomial

$$a_0 + a_1 \mathrm{G80}_e + \cdots + a_9 \mathrm{G80}_e^9 = 0,$$

where the coefficients, a_i, are functions of the model parameters. The *GAL80* feedback deletion model was solved by applying the same procedure. In this model, $K_{R4} = 0$ and $\mathrm{G80}_e$ was determined by computing the roots of a sixth order polynomial. The stability of the equilibrium points was determined by computing the eigenvalues of the Jacobian matrix of the system of equations [14].

S3.2 Characterization of δ_g and fraction of on cells in a set of mutants

To gain mechanistic insight about the observed gene expression dynamics, we characterized a set of mutants that modify the transcriptional state of the GAL promoters. For this experiment, we started with similar cell densities for each mutant ($\mathrm{OD}_{600} \approx 0.3$). Our results showed that a deletion of *CTI6* significantly extended δ_g. Indeed, there were cells in the repressed subpopulation that never turned on the GAL genes over a period of approximately 14 hours (Supplementary Fig. 3.6A-1). These data suggests that the timing of the repressed state's activation depends on the interplay of CTI6, SAGA and Cyc8-Tup1 at the promoter-level. We also found that the percent of cells in the off-state was significantly higher in the *CTI6*Δ strain compared to WT (Supplementary Fig. 3.6D).

We explored the dynamics of the system in a strain lacking the *HXK2* gene. Interestingly, deletion of the dominant glucose kinase, Hxk2, significantly extended δ_g compared to WT (Supplementary Fig. 3.6A-2). Similar to the *CTI6*Δ cells, the repressed subpopulation did not switch to the on-state during the course of the experiment. This mutant was particularly intriguing since it was the only modification to the system that decoupled the fraction of on-cells and delay dynamics. Indeed, there were a higher number of cells in the on-state but δ_g was augmented (Supplementary Fig. 3.6C, D). Previous results have shown that cells deleted for Hxk2p consume glucose and galactose simultaneously [110]. Taken together, these results suggest that the glucose-dependent metabolic inhibition of galactose metabolism and the transcriptional derepression are tightly coupled dynamic processes.

We next characterized the dynamics of a strain in which Gal80p was regulated by an inducible *TET* promoter as opposed to its endogenous Gal4p-dependent promoter, therefore abolishing the feedback regulation of this protein (GAL80Δ fb) [19]. In the presence of 19.7 ng/ml aTc, the repressed subpopulation did not induce the GAL genes over the course of the experiment (Supplementary Fig. 3.6A-3). These data demonstrate that a high concentration of Gal80p relative to WT can block the repressed population from inducing the GAL genes at the appropriate time, highlighting the sensitivity to disruptions in the regulation of this repressor.

Three mutants that displayed accelerated δ_g phenotypes compared to WT were identified. The *MIG1* binding site is degenerate and multiple different proteins can bind to the same sequence [114]. A gene deletion of *MIG1* could therefore increase the probability that these alternative transcriptional regulators bind to the putative *MIG1* sites, thus potentially altering the expression levels of target genes. To directly reduce the promoter-level repression without significantly perturbing promoter occupancy, four point mutations were introduced into Mig1 (Mig1_{4m}) that have been shown to reduce the affinity of Mig1 and the Cyc8-Tup1 complex [115, 97, 116]. This mutant exhibited a reduced δ_g and a higher fraction of on-cells compared to WT (Supplementary Fig. 3.6B-1, C, D).

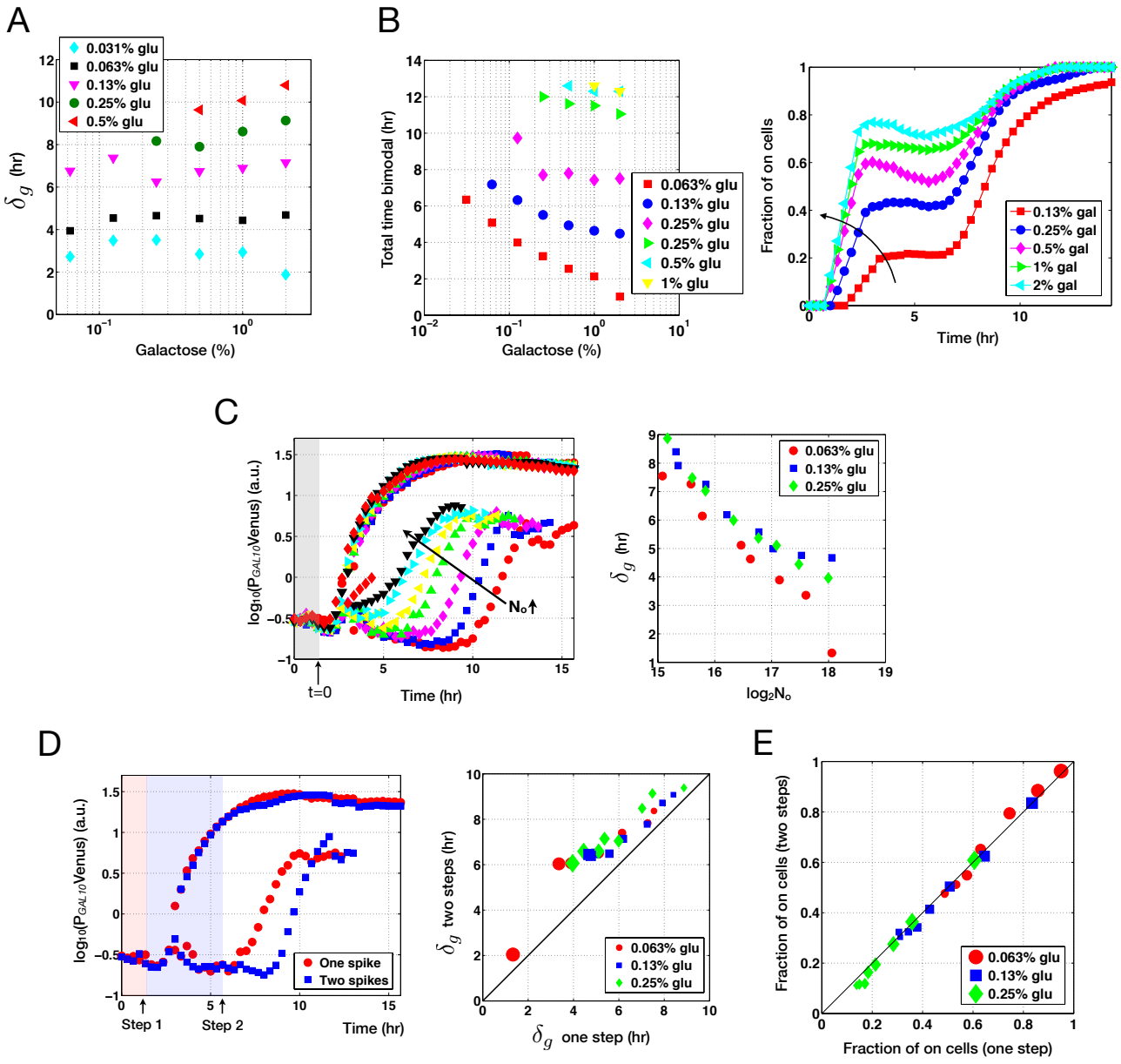
We examined the role of the Cyc8-Tup1 complex in the regulation of the transient bimodal dynamics. Since a deletion the Cyc8-Tup1 complex has severe pleiotrophic effects, including flocculation, diploids were constructed that contain only one allele of *CYC8* and *TUP1* and therefore half the dosage of this complex. We compared the dynamics of this mutant to a WT diploid and observed a significantly smaller δ_g and larger fraction of on-cells compared to WT (Supplementary Fig. 3.6B-2, C, D).

To measure the contribution of Gal80p-Gal4p sequestration on the dynamics of the system, the affinity of Gal80p to Gal4p was reduced by mutating F856C and M861C in *GAL4* [117]. Mutation F856C reduces the affinity of Gal80 to Gal4 more significantly than mutation M861C. These two mutants exhibited a significantly smaller δ_g than WT (Supplementary Fig. 3.6B-3,C). F856C had a smaller δ_g compared to M861C, suggesting that the binding affinity of the Gal80p-Gal4p complex is a crucial variable that controls the duration of bimodality. In addition, these two mutants exhibited

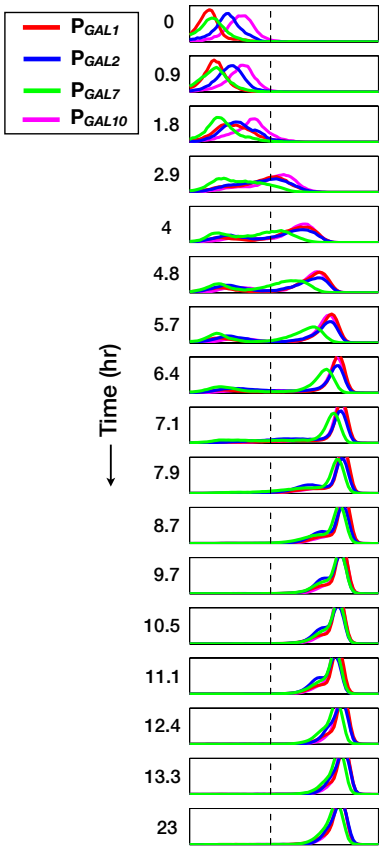
a larger fraction of on-cells compared to WT (Supplementary Fig. 3.6D). Taken together, these results suggest that the timing of the delay dynamics and the single-cell decision to occupy the low or high state is strongly regulated at the promoter level.

We measured the growth rates of these strains to determine if the observed changes in the system's dynamics could be explained by alterations in the growth rates and therefore perhaps the consumption rate of glucose. However, the changes in δ_g and the fraction of on-cells were significantly greater than the variation in growth rates and were therefore dominated by the modifications to the system's regulation (Supplementary Fig. 3.12).

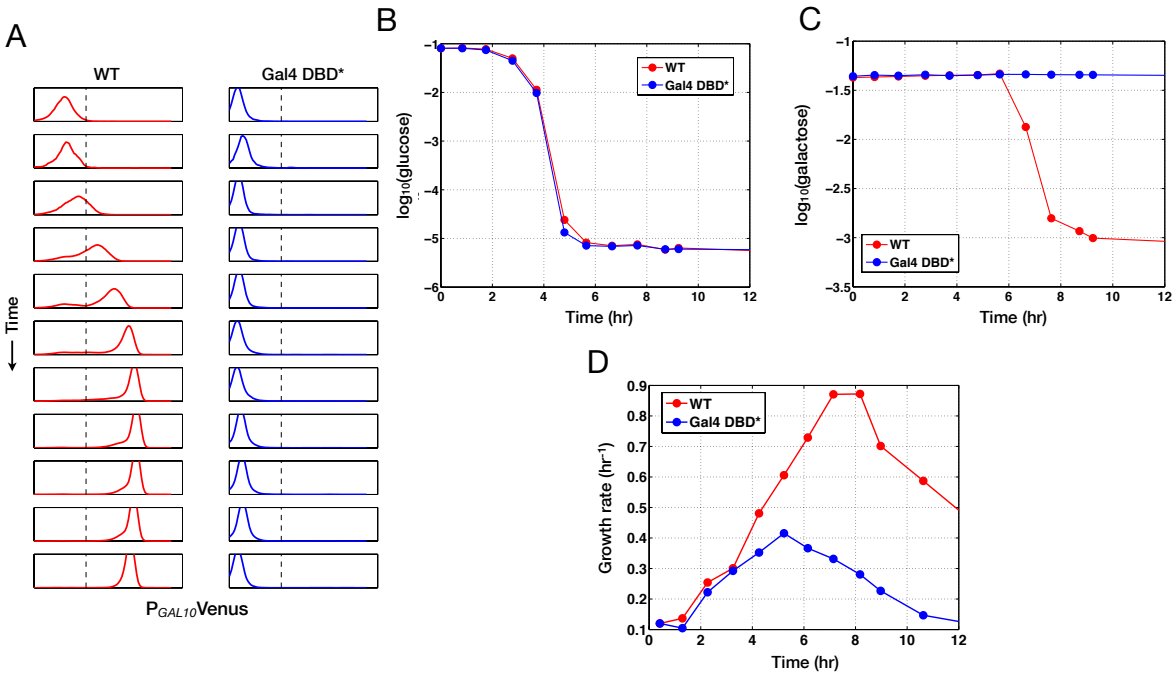
Supplementary Figures



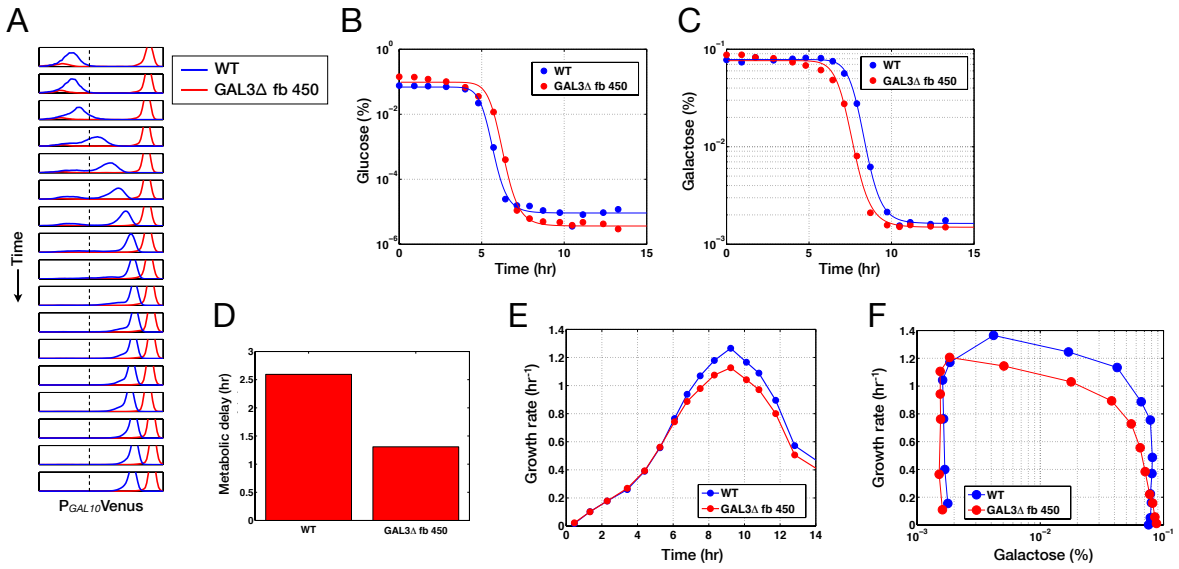
Supplementary Figure 3.1: Duration of transient bimodality as a function of the two sugars. **(A)** The amount of time the system exhibited bimodality was quantified by subtracting the half-max of the activated and repressed subpopulations (δ_g) as a function of galactose for a set of glucose levels. **(B)** Total amount of time the cell populations exhibited bimodality as a function of galactose for a range of glucose levels. Representative dynamics of the fraction of on-cells for 0.125% glucose over a range of galactose concentrations (right). The amount of time the system was bimodal increased as a function of galactose for low concentrations of glucose by inducing a fraction of the population at earlier times (arrow). **(C)** Means of the subpopulations over time quantified using a Gaussian mixture model (GMM) for a combination of glucose and galactose for a range of initial cell densities (N_o). N_o and the delay time are inversely related (right). **(D)** Two steps of glucose produces larger δ_g. Representative data showing the means of the subpopulations that either received single step of glucose (red circles) or two steps (blue squares). Comparison of δ_g across a range of initial cell densities (N_o) for three glucose concentrations in conditions with a single or two steps of glucose (right). Data point size is proportional to N_o. **(E)** Comparison of the fraction of on cells across a range of N_o for three glucose concentrations for conditions with a single compared to two steps of glucose. Data point size is proportional to N_o.



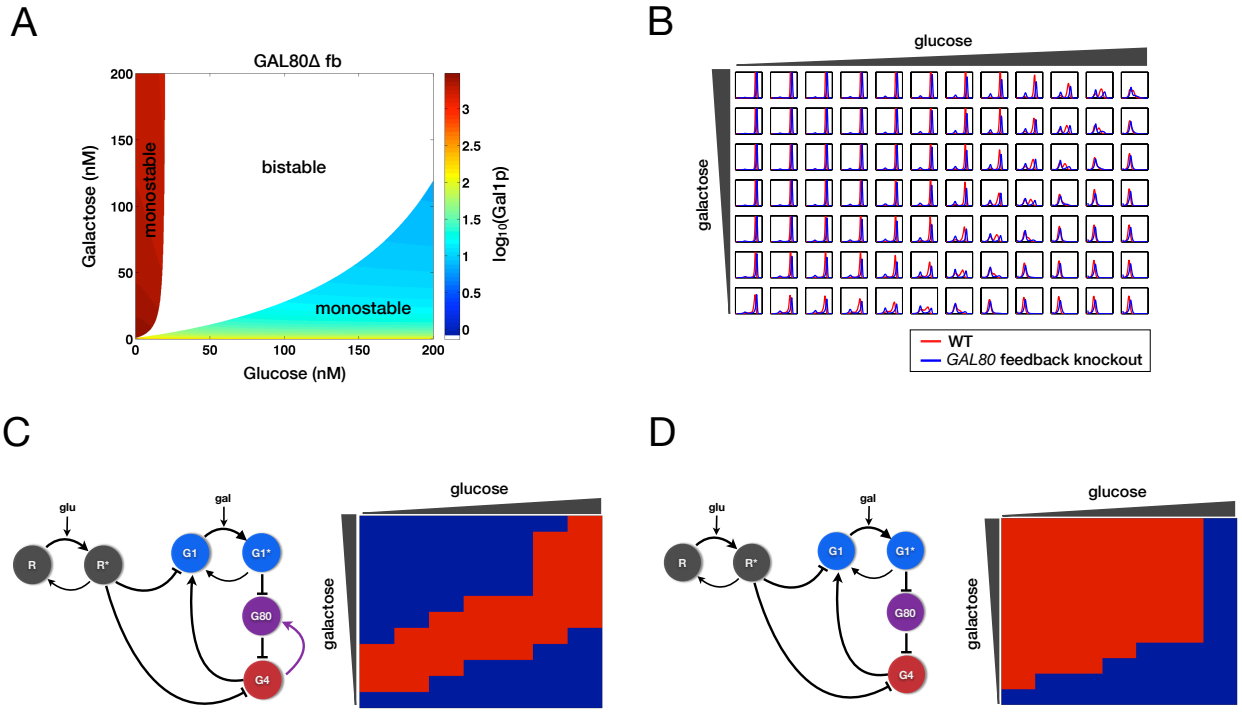
Supplementary Figure 3.2: Flow-cytometry distributions of fluorescent protein (Venus) fusions to the *GAL1*, *GAL2*, *GAL7* and *GAL10* promoters over time.



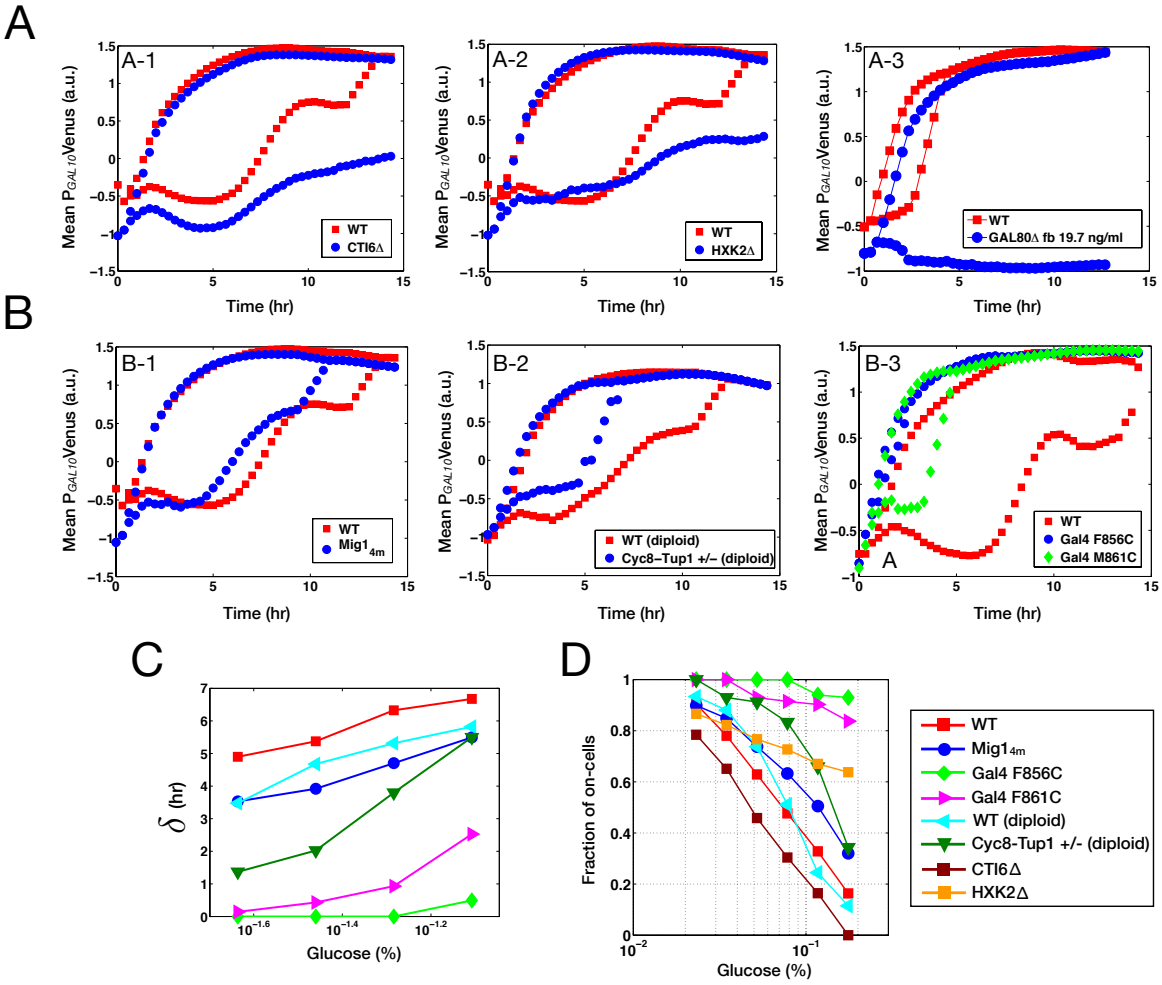
Supplementary Figure 3.3: Consumption of galactose significantly enhances the growth rate of the cell population. **(A)** Single-cell fluorescence distributions of a *GAL10* promoter fusion to Venus in WT and a strain expressing a Gal4 DNA binding mutant that is not able to induce the GAL genes (Gal4 DBD*). **(B)** Glucose concentrations over time. Lines represent fitted Hill functions to the sugar data. **(C)** Galactose concentrations over time. **(D)** Growth rates over time.



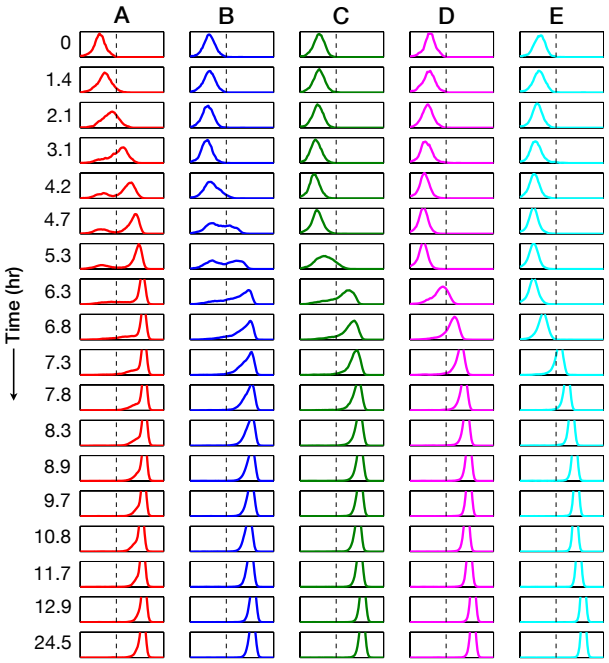
Supplementary Figure 3.4: Constitutive expression of the GAL genes approximately 12 hours prior to induction with glucose and galactose decreased the delay of the metabolic shift by 50% compared to WT and generated a diminished growth rate while consuming galactose. **(A)** Single-cell fluorescence of the *GAL10* promoter in WT and the *GAL3* feedback deletion strain (GAL3Δ fb) induced with 450 ng/ml aTc. **(B)** Glucose concentrations over time. Lines represent fitted Hill functions to the sugar data. **(C)** Galactose concentrations over time. **(D)** Metabolic delay computed by subtracting the half-max of the glucose and galactose concentrations. **(E)** Growth rates over time. **(F)** Relationship between the concentration of galactose and the growth rate.



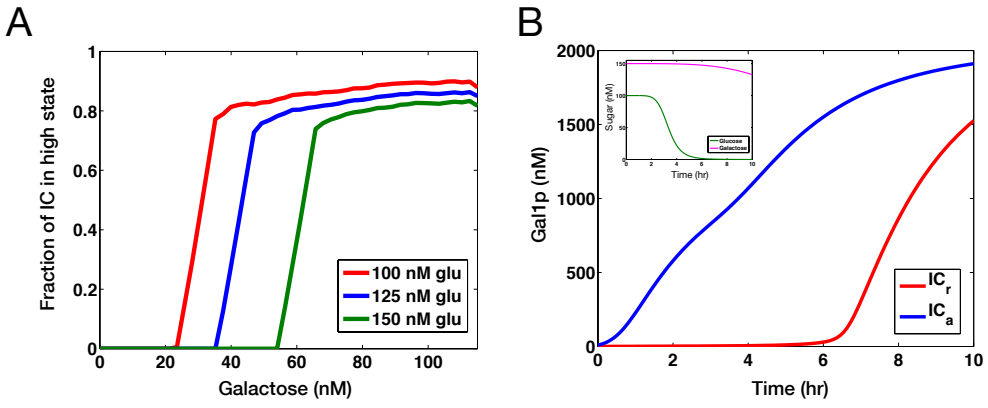
Supplementary Figure 3.5: Removing the *GAL80* feedback loop expands the region of bistability across different concentrations of glucose and galactose in the model and experiments. **(A)** Removing the *GAL80* feedback loop in the model expands the set of conditions that produce bistability. **(B)** Single-cell fluorescence distributions a *GAL10* promoter fusion to Venus in the WT and the *GAL80* feedback loop knockout (GAL80Δ fb). **(C)** Topology of the WT circuit in the computational model (left). Gaussian mixture model (GMM) classification of bimodality for the single-cell fluorescence distributions shown in **B** (right). Red represents a bimodal distribution and blue denotes a monomodal distribution. **(B)** Topology of the GAL80Δ fb model (left). GMM classification of bimodality for this mutant across the range of sugar levels (right).



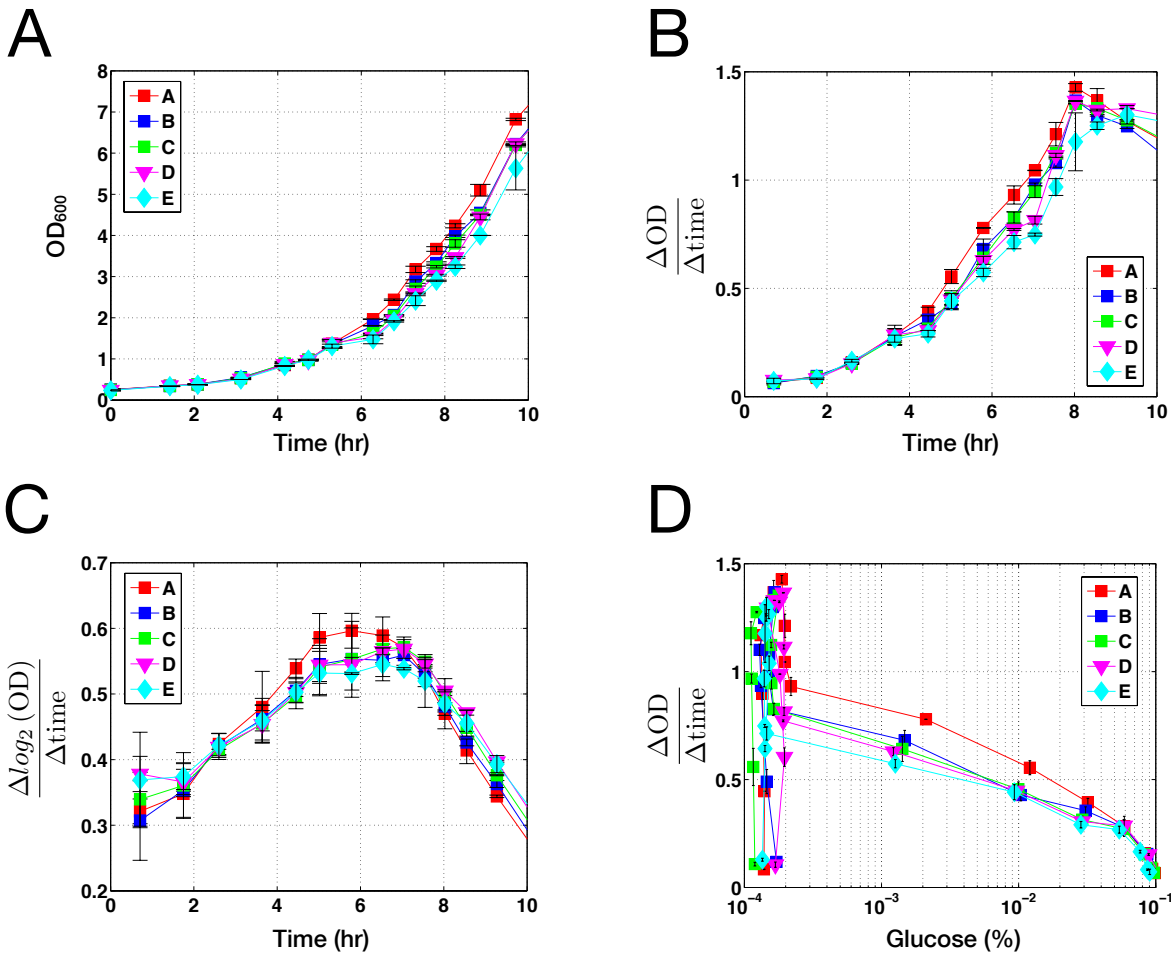
Supplementary Figure 3.6: Modulation of δ_g and the fraction of on cells in a series of mutants. **(A)** Representative data of mutants that significantly extend δ_g. Gene deletions of *CTI6* (A-1) and *HXK2* (A-2) that significantly increase δ_g in response to a step of 0.05% glucose and 0.13% galactose. Feedback deletion of *GAL80* disrupts delayed activation in response to a step of 0.06% glucose and 0.5% galactose (A-3). **(B)** Representative data of mutants that significantly reduce δ_g. Quadruple mutant of Mig1 ($Mig1_{4m}$) (B-1) (0.05% glucose + 0.13% galactose), halved dosage of diploid Cyc8p-Tup1p compared to WT diploid (B-2) (0.05% glucose + 0.13% galactose) and point mutants of Gal4p with reduced affinity to Gal80p significantly reduces δ_g (D-3) (0.08% glucose + 0.13% galactose). **(C)** Comparison of δ_g in response to a step of 0.13% galactose and a range of glucose levels. Mutants with extended δ_g in **A** could not be quantified. **(D)** Fraction of on-cells for the set of mutants at the midpoint of the transient bimodal region.



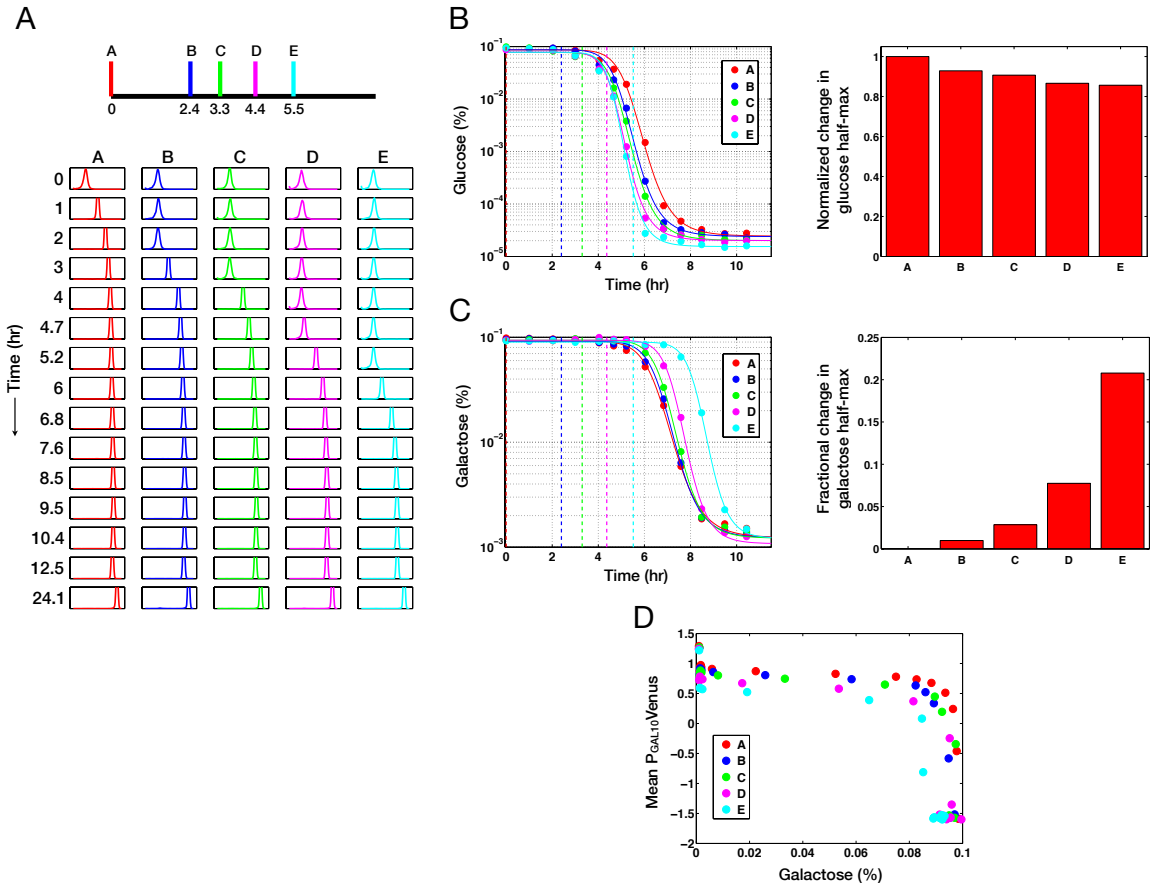
Supplementary Figure 3.7: Single-cell fluorescence distributions of conditions A-E in the delayed galactose pulse experiment. The dotted line denotes the threshold used to compute the fraction of on-cells.



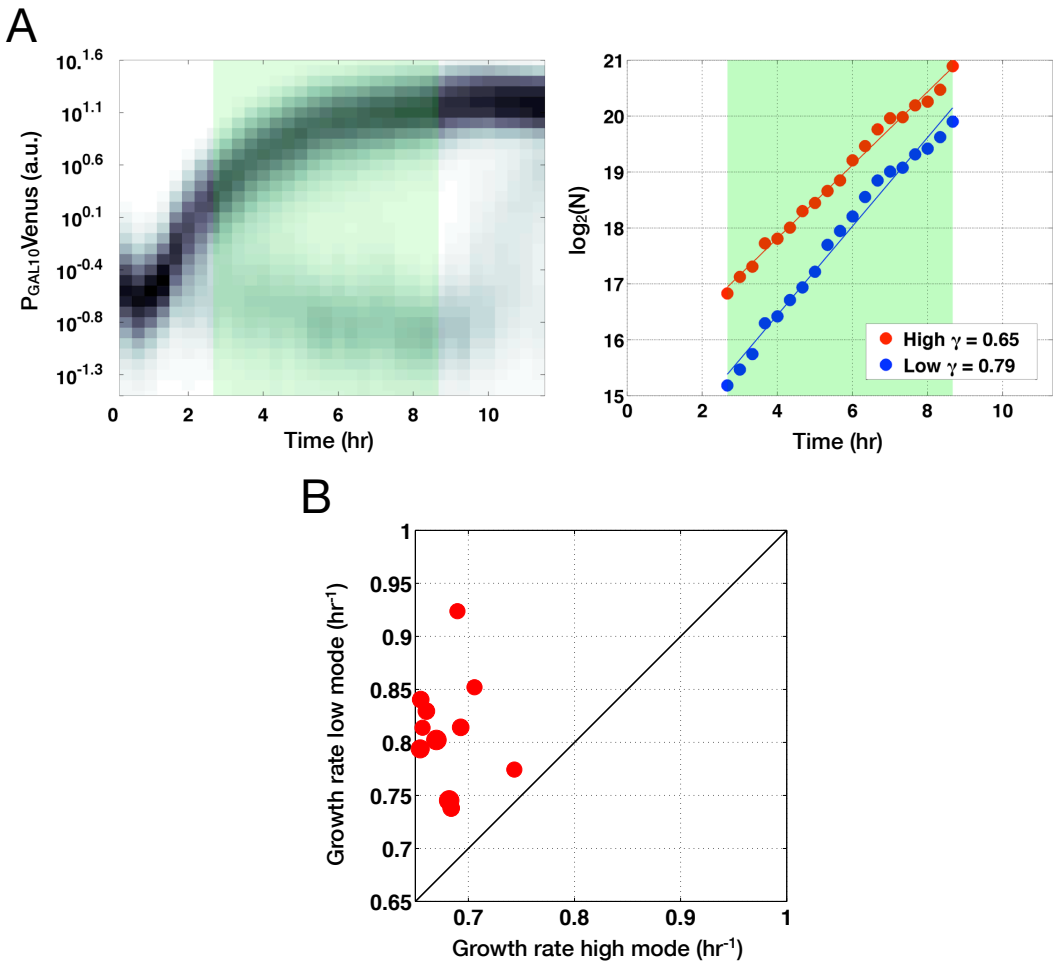
Supplementary Figure 3.8: Model can recapitulate early and delayed activation responses and the modulation of the fraction of on-cells in response to initial sugar concentrations. **(A)** For a concentration of the two inputs that generates bistability (glucose = 100 nM, galactose = 150 nM), the model can recapitulate the early and delayed activation response by starting the system from distinct initial conditions (IC). IC_a has a low initial concentration of repressors (R and G80) whereas IC_r has a high initial concentration of repressors. Here, glucose disappears before galactose (inset). **(B)** Galactose expands the basin of attraction of the high equilibrium state, qualitatively reflecting the experimentally observed increase in the fraction of on-cells as a function of galactose.



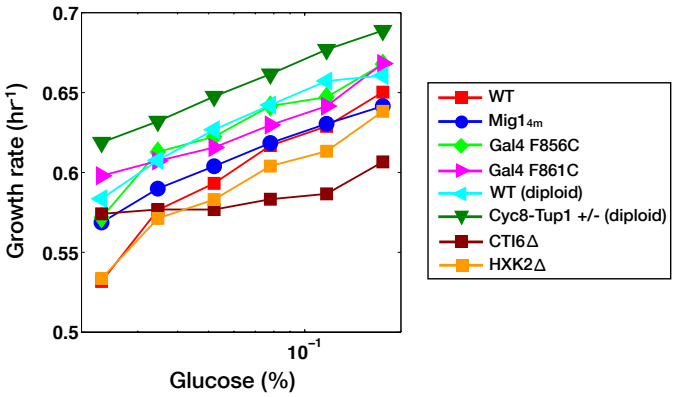
Supplementary Figure 3.9: Growth profiles of conditions A-E in the delayed galactose pulse experiment. **(A)** Optical density 600 nm (OD_{600}) over time. **(B)** Change in OD_{600} per unit time. **(C)** Change in the log of OD_{600} per unit time. **(D)** Relationship between the glucose concentration and the change in OD_{600} per unit time.



Supplementary Figure 3.10: Early activation of the GAL genes using an estradiol-inducible Gal4 chimera prior to glucose depletion reduced the rate of glucose consumption and the amount of time to consume galactose. **(A)** Experimental design in which conditions A-E received a step input of 0.1% glucose + 0.1% galactose from time zero and the GAL genes were induced with 400 nM estradiol at different times over the course of the experiment (top). Single-cell fluorescence distributions of a *GAL10* promoter fusion to Venus over time (bottom). **(B)** Glucose concentrations over time (left). Normalized change in the half-max of glucose consumption relative to condition A (right). Lines represent fitted Hill functions to the data. Dashed lines indicate the time when estradiol was added to each culture. **(C)** Galactose concentrations over time. Fractional change in the half-max of the galactose concentrations relative to condition A. **(E)** Relationship between galactose concentrations and the mean expression levels for each condition.



Supplementary Figure 3.11: Growth rates of the high and low modes quantified from the automated flow cytometry experiment in Fig. 3.1. The growth rate was determined by a linear regression on the number of cells collected for each mode during the transient bimodal region. The diameter of each data point is proportional to the initial concentration of glucose.



Supplementary Figure 3.12: Growth rates of a set of strains computed by a regression on the number of cells obtained at each time point over a period of 12 hours.

Strain name	Genotype
Wild-type (WT) P_{GAL1}Venus	MATa leu2::LEU2, trp1::TRP1-P_{GAL1}Venus, ura3::URA3, ade2::ADE2, his3::HIS3
WT P_{GAL2}Venus	MATa leu2::LEU2, trp1::TRP1-P_{GAL1}Venus, ura3::URA3, ade2::ADE2, his3::HIS3
WT P_{GAL7}Venus	MATa leu2::LEU2, trp1::TRP1-P_{GAL7}Venus, ura3::URA3, ade2::ADE2, his3::HIS3
WT P_{GAL10}Venus	MATa leu2::LEU2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3, ade2::ADE2, his3::HIS3
Gal4-C14Y P_{GAL10}Venus	MATa leu2::LEU2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{GAL4}Gal4-C14Y, ade2::ADE2, his3::HIS3, GAL4::HPH
Gal4-ER P_{GAL10}Venus	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{ADH1}GAL4DBD-ER-MSN2AD, ade2::ADE2, his3, GAL4::HPH
*HXK2*Δ	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3, HXK2::NAT
*CTI6*Δ	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3, CTI6::NAT
$Mig1_{4m}$	MATa leu2::LEU2-P_{MIG1}MIG1-L490A, L493A, L496A, L498A, trp1::TRP1-P_{GAL10}Venus, ura3, ade2::ADE2, his3, MIG1::NAT
Gal4 F856C	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{GAL4}GAL4-F856C, ade2::ADE2, his3, GAL4::HPH
Gal4 F861C	MATa leu2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{GAL4}GAL4-F861C, ade2::ADE2, his3, GAL4::HPH
WT diploid	MATa/α leu2/leu2, trp1::TRP1-P_{GAL1}Venus/trp1, ura3/ura3::URA3, ade2::ADE2/ade2, his3/his3
Cyc8-Tup1 (+/-) diploid	MATa/α leu2/leu2, trp1::TRP1-P_{GAL1}Venus/trp1, ura3/ura3::URA3, ade2::ADE2/ade2, his3/his3, CYC8::NAT/CYC8, TUP1/TUP1::KAN
GAL3Δ fb*	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL3, ade2::ADE2, his3, GAL3Δ::KAN
GAL80Δ fb*	MATa leu2::LEU2-P_{TEFm4}rtTA-M2, trp1::TRP1-P_{GAL10}Venus, ura3::URA3-P_{TET}GAL80, ade2::ADE2, his3, GAL80Δ::HPH

Supplementary Table I: Strains used in this study. All strains were W303. *Strain described in [19].

Chapter 4

Inference of regulatory regions of a promoter library using statistical analysis

Introduction

Genetic regulation is a fundamental and ubiquitous mechanism for generating adaptive cellular responses to changes in environment conditions. Genetic programs rely on promoter sequences to map the activity states of a upstream regulators to the rate of gene expression of downstream targets. Recently, the location of transcription factor binding sites have been extensively characterized across many genomes [118, 119]. However, we do not yet quantitatively understand how multiple *cis* DNA sequences combine to produce gene expression profiles. Due to the combinatorial complexity of the upstream inputs and the potential time and context dependence of the response, a detailed understanding of how promoters integrate these signals poses many challenges.

If the transcription factors are known, there are two rational approaches to characterize the relationship between promoter sequence and transcription rate. Varying the composition, number and spacing of binding sites in synthetic promoters has provided many insights about how these variables can shape a transcriptional response [120, 121, 122]. Hypothesis-driven mutations or deletions to specific regions of a natural promoter, referred to as "promoter bashing", have revealed the location and contributions of different regulatory elements [123, 124].

However, the rules identified in the context of synthetic promoters may not apply to the regulation

of the genomic promoter due to numerous uncharacterized inputs to the promoter that do not function in isolation. In addition, the comprehensive set of transcriptional regulators for a promoter are frequently not known, especially in less studied model organisms. To address these challenges, a random mutagenesis strategy can be used to analyze the quantitative contributions of regions of a promoter sequence without prior information about the location and identity of *cis* regulatory sequences. Recently, libraries of random promoter sequence variants have been used to elucidate the relationship between the mean expression level and noise, and the arrangements of nucleosomes [125, 126].

To quantitatively probe the relationship between DNA sequence and phenotype, we used the bidirectional *GAL1-10* promoter in *Saccharomyces cerevisiae* (*S. cerevisiae*) as a model system. The galactose regulatory network in *S. cerevisiae* is a paradigm for eukaryotic gene regulation and has provided many insights about how cells control a genetic program in response to environmental signals. In *S. cerevisiae*, *GAL1*, *GAL10* and *GAL7*, which perform the first three enzymatic steps in galactose metabolism, constitute a gene cluster that has been evolutionarily conserved across several related yeast species [127]. A previous study showed that this gene cluster does not provide a fitness advantage for cells in specific environments tested thus far. However, the strong correlation between GAL gene loss and gene clustering across many yeast species suggests that tightly coupled genes may be easier to eliminate over evolution [128]. A bioinformatic comparison of the *GAL1-10* promoter in *S. cerevisiae* to three related yeast species (*S. paradoxus*, *S. mikatae* and *S. bayanus*), revealed a very high conservation of the *GAL4* binding sites and TATA boxes [129]. Interestingly, bidirectional gene expression has been shown to be a ubiquitous phenomenon in yeast and it is not yet understood what features of promoters are important for coordination of gene expression in two directions [128, 130, 131].

The transcriptional activator Gal4p remains bound to the *GAL1-10* promoter in the presence and absence of galactose. The *GAL1-10* promoter is repressed in the absence of galactose due to inhibition of the activation domain of Gal4p by the repressor Gal80p. The Cyc8-Tup1 complex is also a dominant source of repression of this promoter and collaborates with a set of transcriptional

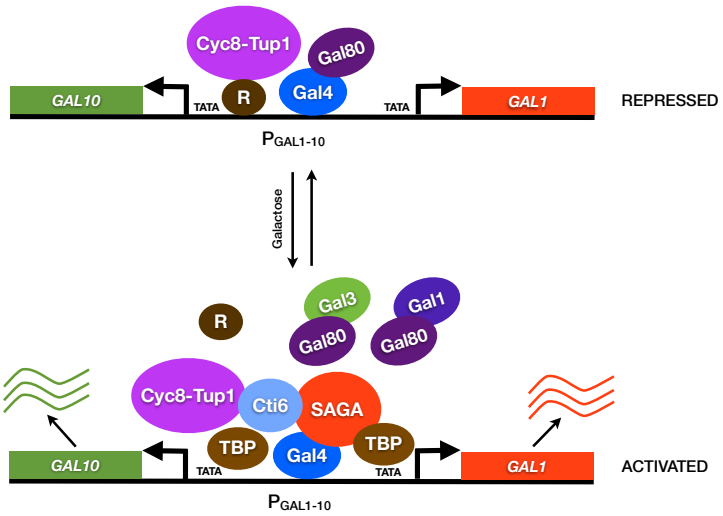


Figure 4.1: The bidirectional *GAL1-GAL10* promoter ($P_{GAL1-10}$) switches from a repressed to an activated state in the presence of galactose. The general repression complex Cyc8-Tup1 is a dominant source of active repression and remains tethered to the promoter by directly interacting with a set of DNA binding repressors (R) including, for example, Mig1/2 and Nrg1/2. The repressor Gal80p occludes the potent activation domain of the transcriptional activator, Gal4p in the absence of galactose. The addition of galactose triggers the activation of the signal transducers, Gal1p and Gal3p, which enables these proteins to sequester Gal80p. As a consequence, Gal4p is able to recruit the pre-initiation complex including the Spt-Ada-Gcn5 acetyltransferase (SAGA) and the TATA binding protein (TBP) leading to gene induction. The PHD domain protein, Cti6p, facilitates the off → on transition by physically interacting with Cyc8-Tup1 and SAGA.

repressors including Mig1, Mig2, Nrg1 and Nrg2 to block gene expression [116, 85, 86]. The addition of galactose to cells triggers the activation of the signal transducers Gal1p and Gal3p, which enables these proteins to sequester Gal80p. The liberation of the Gal4p activation domain leads to the ordered assembly of the transcriptional pre-initation complex including Spt-Ada-Gcn5 acetyltransferase (SAGA) and the TATA binding protein (TBP) (Fig. 4.1). The Plant Homeo Domain (PHD) protein Cti6p plays a critical role in facilitating the off to on transition of this promoter by physically interacting with Cyc8-Tup1 and SAGA [91].

In *S. cerevisiae*, Gal4p induces transcription of the galactose transporter Gal2p, signal transducers Gal1p, Gal3p and the repressor Gal80p, thus forming four feedback loops. The two positive feedback loops established by Gal1p and Gal3p have been shown to generate bistability, which can lead to two coexisting high and low expression levels in a homogenous environment across a population of cells [19]. This bimodality in gene expression has been observed for range of intermediate galactose levels and for similar concentrations of glucose and galactose [18, 20].

In this work, we propose a method to quantitatively analyze the contributions of individual nucleotides to a phenotype using a combination of random mutagenesis and statistical modeling. Statistical models have been used to identify the relationships between protein sequence, function and stability, and have aided the field of protein engineering [132, 133]. Here, we use a regression framework to map promoter sequence to phenotype. We explore the context dependence of single nucleotides in non-repressive (galactose) and repressive bimodal (glucose + galactose) conditions. In addition, we identify the sequence determinants of coordinated expression by comparing the expression of *GAL1* and *GAL10* using two fluorescent proteins in the same cell. Our models make predictions about the relative importance of nucleotides in the sequence and suggest several unexplored regulatory regions of the promoter. This technique can be applied to any promoter to identify *de novo* binding sites *in vivo* without a priori knowledge of the system. Furthermore, our method can be used to quantitatively dissect the relative importance of individual nucleotides in a promoter and does not require a large library size.

Results

Characterization of the library of sequence variants

To study the bidirectional expression of the *GAL1-10* promoter, Venus (YFP) was expressed from *GAL10* and mCherry (RFP) was expressed from *GAL1* (see Materials and Methods). We first characterized the promoter strengths of *GAL1* and *GAL10* by separately expressing mCherry from *GAL10* and Venus from *GAL1*. By comparing the fluorescence of Venus and mCherry between the two constructs, we determined that the *GAL1* promoter is approximately 15% stronger than *GAL10* at steady-state and has a ≈5 min faster response time (Fig. 4.2A,B). From these dynamic data, we identified a delay of 41.8 min in the half-max of mCherry compared to the half-max of Venus and a 30% reduced brightness of mCherry compared to Venus in these conditions.

We used random mutagenesis to generate a library of 343 sequence variants of the bidirectional promoter (see Materials and Methods). After sequencing, we determined the mean, minimum and

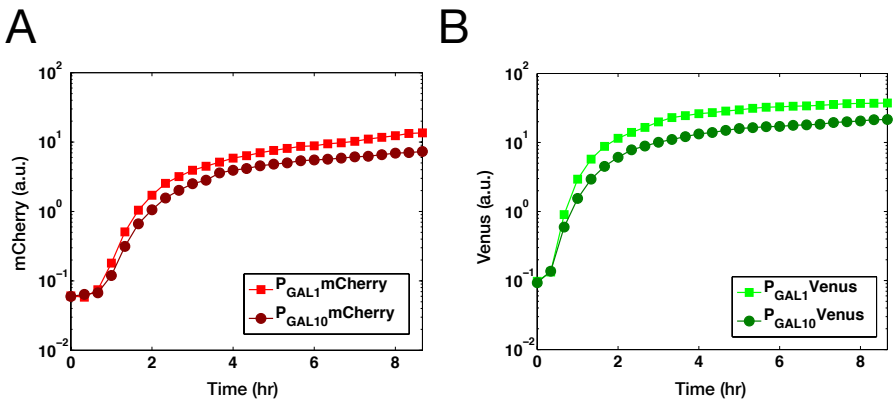


Figure 4.2: The *GAL1* promoter is stronger than the *GAL10* promoter. Dynamic measurements of the bidirectional $P_{GAL1\text{-}10}$ promoter in the presence of 2% galactose. The two fluorescent proteins (mCherry and Venus) were expressed separately from the *GAL1* and *GAL10* promoters for comparison.

maximum number of mutations to be 10.4, 0 and 19. Mutations were distributed uniformly over the promoter sequence with the exception of two regions of diminished mutation rates close to the transcription start sites of *GAL1* and *GAL10* (Fig. 4.3).

The phenotypes of these mutants were measured over time in non-repressive conditions (0.25% galactose) and following 5 hours of induction in repressive conditions in the presence of glucose (0.25% glucose + 1% galactose) (Fig. 4.4). Most of the mutants exhibited wild-type expression levels in the absence of glucose, suggesting that in these conditions mutations were nearly neutral (Fig. 4.4A,B). In the presence glucose and galactose, we observed a bimodal distribution of gene expression across the population of cells for the majority of mutants (Fig. 4.4C,D). The mean of the high mode in repressive conditions displayed larger phenotypic variability compared to the mean in the absence of glucose, indicating that mutations have a more significant effect on the promoter activity in this condition (Fig. 4.4E,F).

The relationship between the number of mutants for each variant and the mean of the bimodal distribution of gene expression in the presence of 0.25% glucose + 1% galactose was examined for both directions (Fig. 4.5A,B). The majority of mutations reduced the mean expression level, suggesting that mutations that produced a detectable change in phenotype either reduced activation or increased repression. There were many sequences that displayed phenotypes similar to or higher than wild type but contained a large number of mutations, indicating that this promoter could

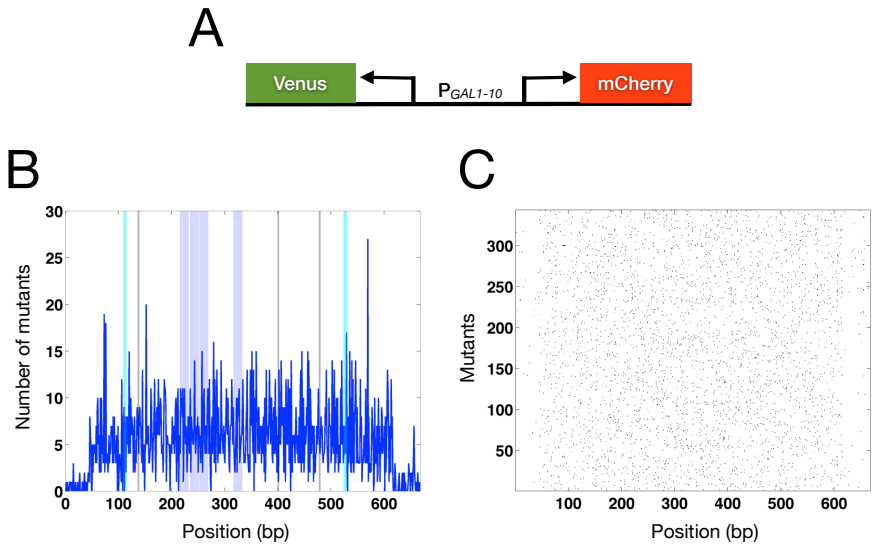


Figure 4.3: Sequence characterization of the library of 343 $P_{GAL1\text{-}10}$ mutants. **(A)** Venus and mCherry were expressed from the *GAL10* and *GAL1* promoters, respectively. **(B)** The number of sequences with a specific mutations. Light blue denotes TATA boxes, gray indicates putative *MIG1* binding sites and dark blue represents *GAL4* binding sites. **(C)** Heat map of the library highlighting the distribution of mutations. Black indicates a mutation at the specific position.

tolerate many mutations and maintain wild-type expression capabilities. Overall, increasing the number of mutations produced phenotypes with lower mean expression levels.

Previous studies have identified a scaling relationship between the mean abundance and the noise across many genes in yeast [134]. In our data, we determined how mutations in the promoter sequence modify the relationship between the noise (coefficient of variation, CV) and the mean of each mutant induced with 0.25% galactose at steady-state for the two directions (Fig. 4.6A,B). For the majority of mutants, the CV and the mean were uncorrelated suggesting that some mutations can independently adjust the noise and the mean. By contrast, in some mutants the CV was inversely proportional to the mean for both *GAL1* and *GAL10*, indicating that specific mutations can simultaneously alter the noise and mean properties of this promoter.

Lasso regression provides insight about cis regulatory regions

To investigate how individual base pairs contribute to the observed variation in phenotypes, we implemented least absolute shrinkage and selection operator regression (Lasso) [135]. Lasso regression constrains the L^1 norm of the regression coefficients to be less than a specified value (see

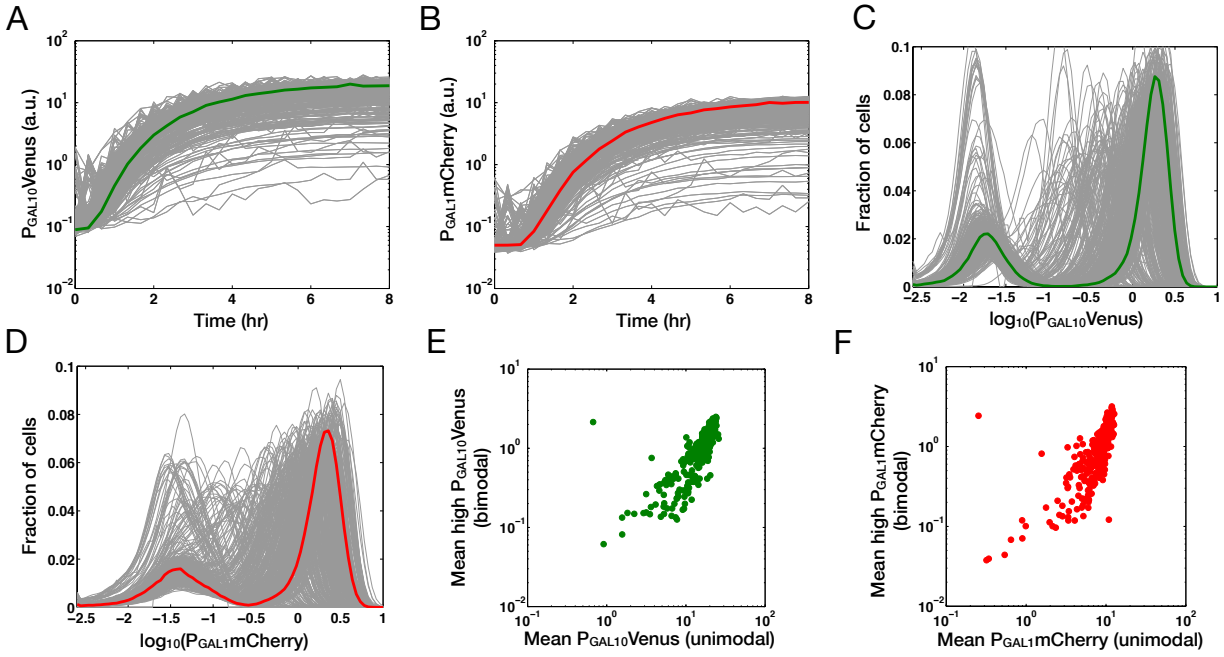


Figure 4.4: Characterization of the $P_{GAL1\text{-}10}$ mutant phenotypes. **(A)** Dynamic measurements of the *GAL10* promoter driving Venus induced with 0.25% galactose. For all subplots, the bold colored line indicates the WT sequence. **(B)** Dynamic measurements of the the *GAL10* promoter driving mCherry induced with 0.25% galactose. **(C)** Single-cell distributions of $P_{GAL1\text{-}10}$Venus induced with 0.25% glucose + 1% galactose. **(D)** Single-cell distributions of $P_{GAL1\text{-}10}$mCherry induced with 0.25% glucose + 1% galactose. **(E)** Relationship between the mean induced with 0.25% galactose at steady-state in **A** and the mean of the high mode for the bimodal distributions in **C**. **(F)** Relationship between the mean induced with 0.25% galactose at steady-state in **B** and the mean of the high mode for the bimodal distributions in **D**. The mean of the high mode was extracted using a Gaussian mixture model (GMM).

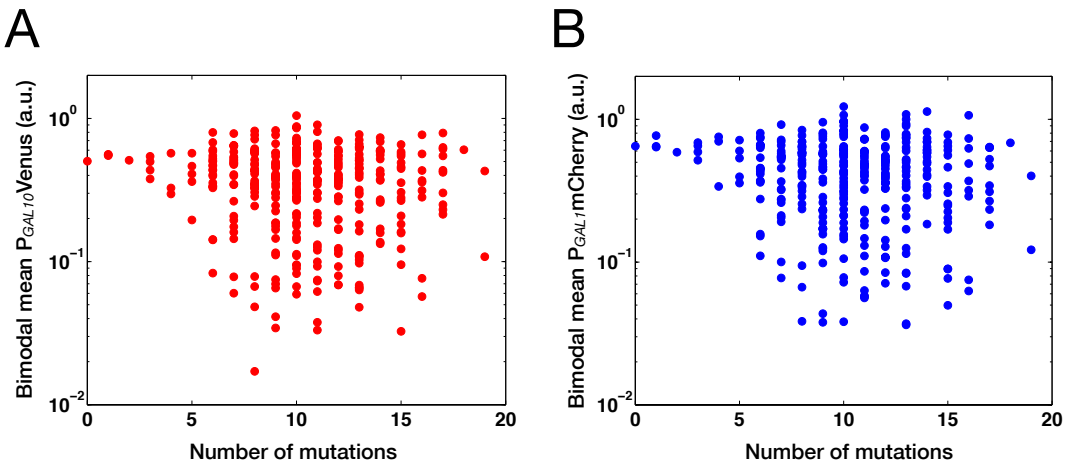


Figure 4.5: Relationship between the number of mutations and the mean of the bimodal distribution of gene expression (0.25% glucose + 1% galactose). **(A)** Relationship between the number of mutations and the mean of P_{GAL10}Venus. **(B)** Relationship between the number of mutations and the mean of P_{GAL1}mCherry.

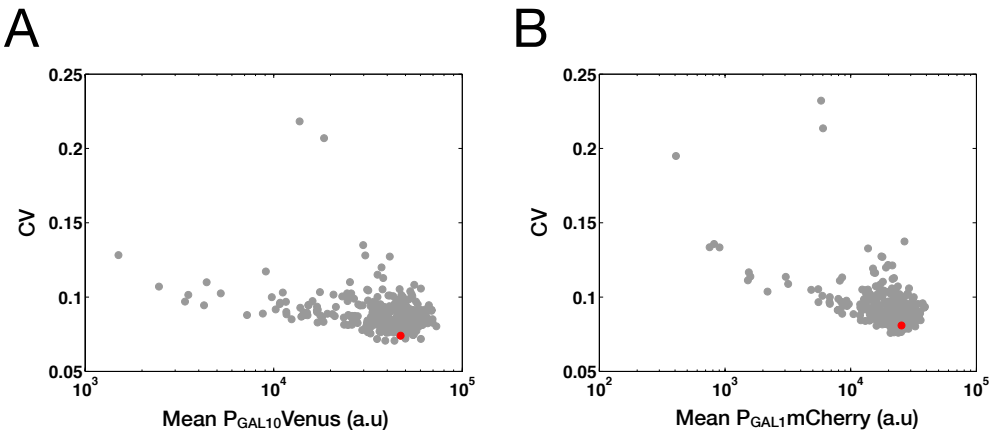


Figure 4.6: Relationship between the mean fluorescence at steady-state in the presence of 0.25% galactose and the coefficient of variation (CV). **(A)** Scatter plot of the mean of P_{GAL10}Venus and CV. **(B)** Scatter plot of the mean of P_{GAL1}mCherry and CV. Red data points indicate wild type.

Materials and Methods). This method is particularly useful for solving underdetermined problems that involve many variables. The solutions of Lasso tend to be sparse and hence may produce a more interpretable mapping between the dependent and independent variables. The regularization coefficient was selected by identifying the value that produced the largest correlation between the predicted and measured phenotypes using cross-validation. A representative correlation for the two directions of the *GAL1-10* promoter is shown in Fig. 4.7A,B.

We applied this method to the mean steady-state expression level of the mutant library in non-repressive conditions. The largest regression coefficient for *GAL10* (Venus) was at the precise position of the TATA box (Fig. 4.7A). The weights on the edges of the four *GAL4* sites were also significant and suggested that the last and first base pairs of the third and fourth *GAL4* binding sites are the most important for this phenotype compared to the other regions of the *GAL4* binding sites. Gal4p recognizes the first and last three base pairs of the binding site CGG-$(N)_{11}$-CCG, corroborating this pattern observed in the regression coefficients [136]. The TATA box of *GAL1* (mCherry) exhibited a moderately high weight and a similar pattern for the *GAL4* binding sites, indicating that the third and fourth sites are the most critical.

In repressive conditions (0.25% glucose + 1% galactose), we computed the regression coefficients using the mean of the bimodal population and the mean of the high mode (Fig. 4.8B,C). The mean of the bimodal population took into account the fraction of cells distributed between the two expression

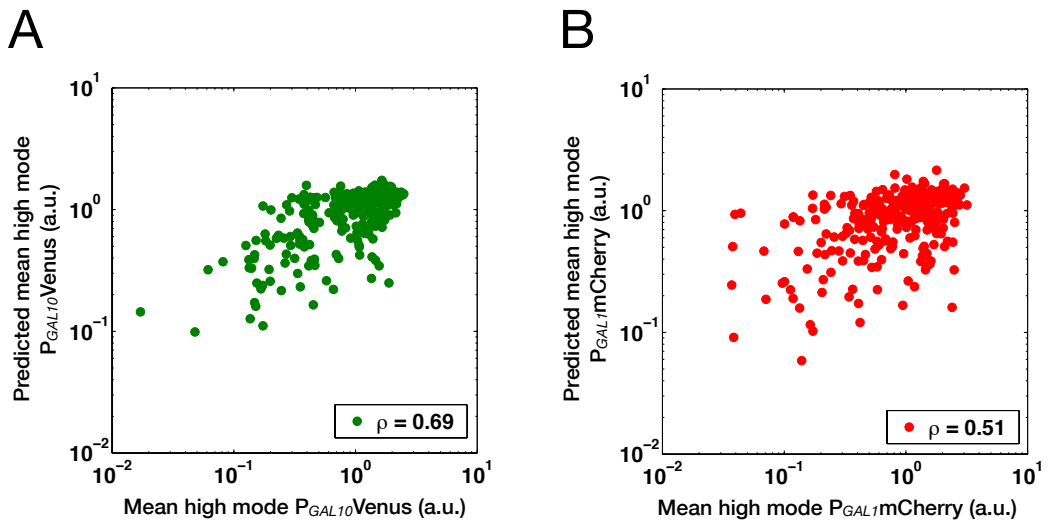


Figure 4.7: Relationship between model predictions and measured phenotypes for the library of sequence variants. **(A)** Scatter plot of model predictions and experimentally measured mean of the high mode in the presence of 0.25% glucose + 1% galactose for P$_{GAL10}$Venus. **(B)** Scatter plot of model predictions and actual mean of the high mode for P$_{GAL1}$mCherry. ρ represents the correlation coefficient.

states, which can be modulated by the initial cell density (Chapter 3). Similar to the non-repressive conditions, the weight of the *GAL10* TATA box was large for Venus. In this condition, the second, third and fourth *GAL4* binding sites were the most significant. P$_{GAL1}$mCherry exhibited small weight for the *GAL1* TATA box and similar relative weights in the *GAL4* binding sites compared to the *GAL10* direction. Interestingly, the second putative *MIG1* binding site displayed a negative regression coefficient for Venus and for the high mode of mCherry indicating that this base pair contributes to activation as opposed to repression in this environment.

The results of the regression indicated that the TATA boxes for *GAL1* and *GAL10* are specific to each gene and mutations in these regions do not significantly modify the more distant gene. Whereas the *GAL4* binding sites were significant for both Venus and mCherry, there were a number of base pairs that were distinct between the two directions, suggesting that mutations in this bidirectional promoter can yield distinct consequences for the two genes. In the presence of glucose, the dominant weights were primarily clustered in the *GAL4* binding sites and there were only a few nucleotides with positive weights suggesting a potential repression site. Indeed, only the third putative *MIG1* binding site had a small positive weight for mCherry using the mean of the high mode. Since

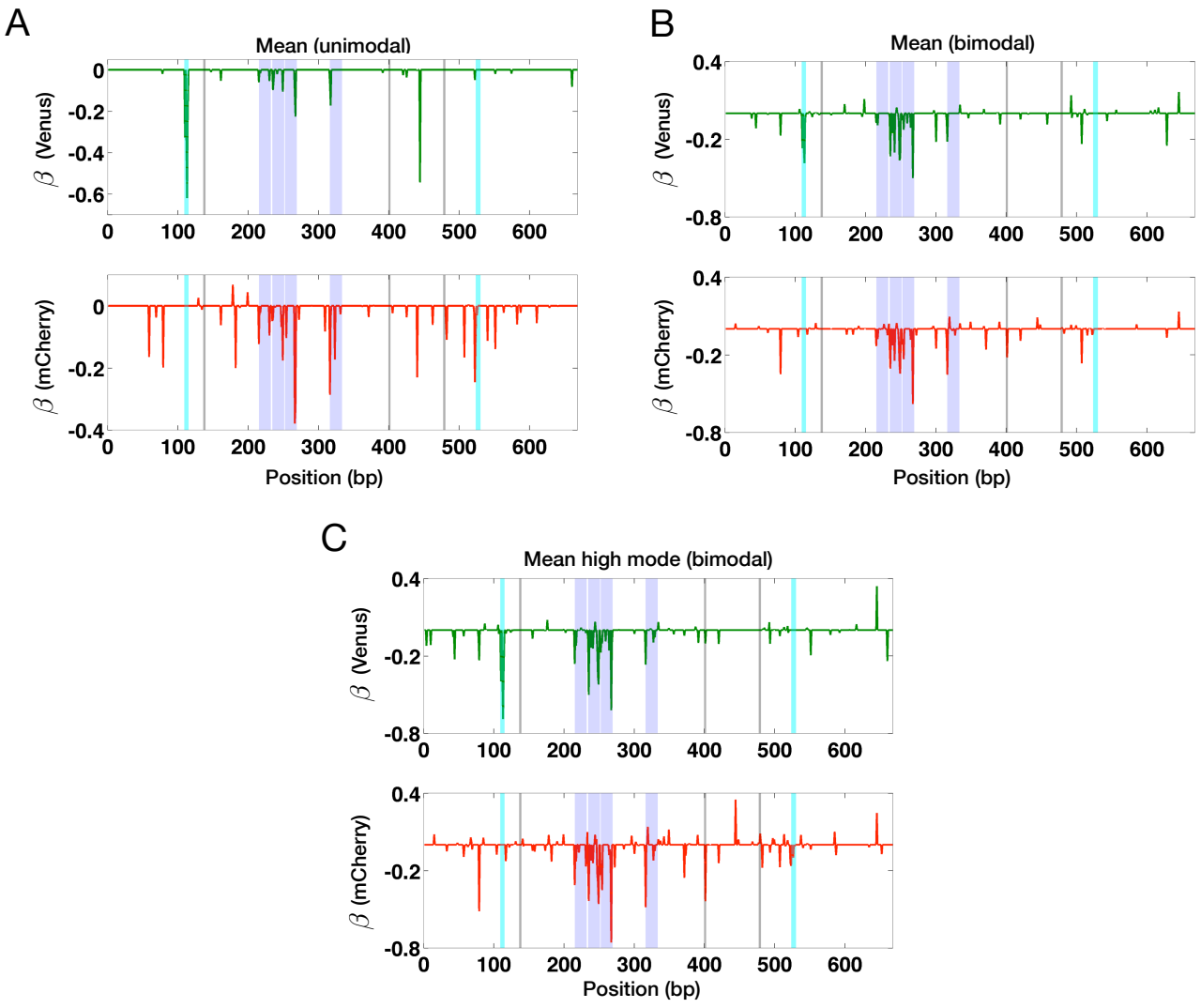


Figure 4.8: Lasso regression coefficients using three different phenotypes. Light blue denotes TATA boxes, gray indicates putative *MIG1* binding sites and dark blue represents *GAL4* binding sites. **(A)** Regression coefficients (β) using the mean of P_{GAL10}Venus (top) and P_{GAL1}mCherry (bottom) in the presence of 0.25% galactose at steady-state. **(B)** Regression coefficients using the mean of the bimodal distribution of gene expression. To induce bimodality, cells were exposed to 0.25% glucose + 1% galactose.**(C)** Regression coefficients using the mean of the high mode for a bimodal distribution of gene expression. The mean of the high mode was computed using a Gaussian mixture model (GMM).

the repressive glucose signal did not yield dominant positive regression coefficients compared to the condition lacking glucose, we hypothesize that repression of this promoter is implemented by blocking *GAL4* activation as opposed to harnessing active repression mechanisms implemented by transcriptional repressors.

Experimental characterization of model predictions demonstrates strong correlation

To test the model, site-directed mutagenesis was used to mutate specific nucleotides predicted to alter the expression level of either *GAL1* or *GAL10*. For each position, the wild-type sequence was changed to the most frequently represented nucleotide in the library. These positions included the *GAL10* TATA box (T110C and T113C), *GAL4* binding sites (G235A, C267T, C248T and C316A), *MIG1* binding site (G478T) and the putative *GAL1* TATA box (A526C) [86].

As predicted, mutations to the *GAL10* TATA box significantly reduced the expression level of *GAL10* by 95-100%. Interestingly, these mutations increased the expression level of*GAL1* by up to 14%, suggesting that a disruption of the TATA box of the opposing promoter direction can moderately enhance the rate of transcription (Fig. 4.9A,B). One potential explanation for these data is that the two directions of the promoter compete for binding to TBP.

The mean of both *GAL1* and *GAL10* were significantly reduced for mutations in the *GAL4* binding sites. Whereas positions 235 and 248 produced similar quantitative effects on mCherry and Venus, nucleotides 267 and 316 reduced the mean of *GAL1* by 60% and 69% compared to 47% and 38% for *GAL10*. These data demonstrate that the two directions of the bidirectional promoter can be differentially modulated by mutations in the regulatory regions. Our results showed that the mutation to the experimentally characterized *MIG1* binding site (478) increased the expression level of mCherry by 24% and Venus by 12%. Finally, the mutation to the *GAL1* TATA box (526) decreased the expression level of mCherry by 66%.

We computed a strong correlation coefficient of 0.76 and 0.89 between the model predictions of the mutant phenotypes and the experimental measurements for *GAL1* and *GAL10* (Fig. 4.9C,D).

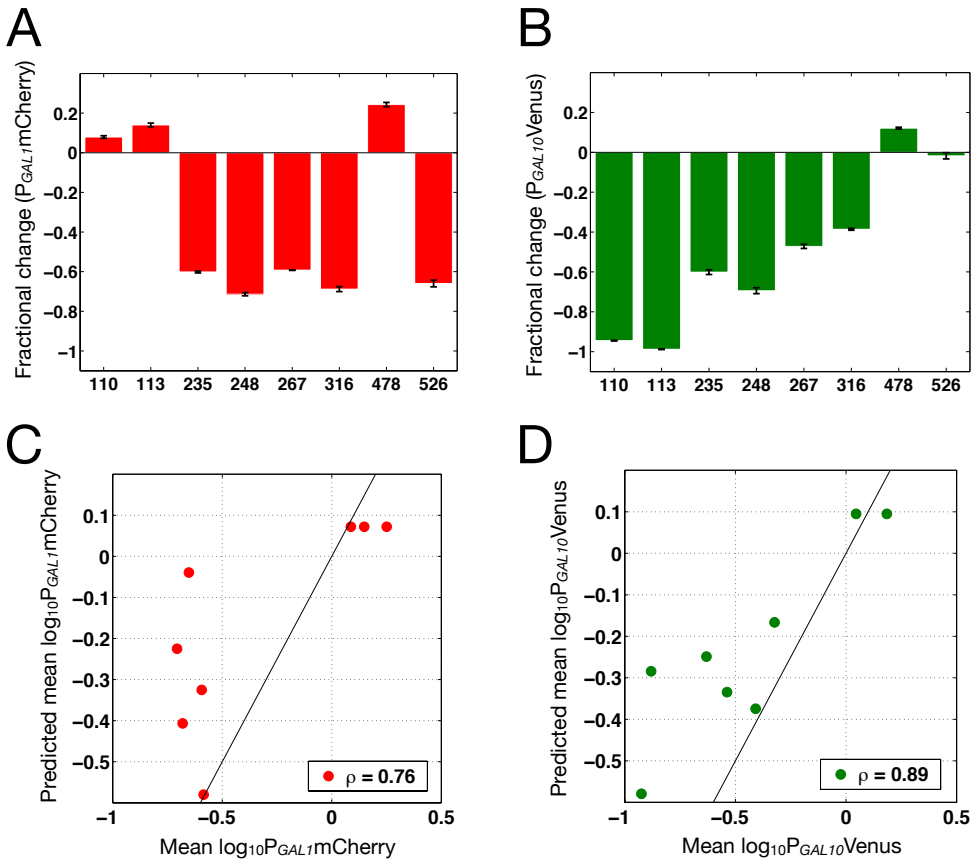


Figure 4.9: Experimental characterization of specific nucleotides predicted by the regression model to modify the expression level. **(A)** Fractional change in the mean of the high mode induced with 1% galactose + 0.25% glucose for P$_{GAL1}$mCherry for a set of mutations. **(B)** Fractional change in the mean of the high mode induced with 1% galactose + 0.25% glucose for P$_{GAL10}$Venus for a set of mutations. **(C)** Relationship between the model predictions of the mutation and the experimentally measured mean of P$_{GAL1}$mCherry. **(D)** Relationship between the model predictions of the mutation and the experimentally measured mean of P$_{GAL10}$Venus. ρ denotes the correlation coefficient. Error bars represent the mean of two independent replicates.

Together these data demonstrate that the regression model can accurately reveal the location of dominant activation and repression regulatory regions in a promoter and predict the quantitative effects of mutations to these *cis* regulatory sequences.

Rational methods to explore Gal4-dependent promoter regulation

To investigate the rules of *GAL4* dependent gene regulation, we constructed a set of synthetic promoters. These synthetic promoters were designed to test the effects of varying the distance between binding sites, number of binding sites on the effective Hill coefficient of the dose response

and to measure the *in vivo* strengths of the individual *GAL4* binding sites in the bidirectional *GAL1-10* promoter for comparison with the random mutagenesis results.

Using the weak constitutive *EAF1* promoter as a backbone, we positioned two strong *GAL4* binding sites 400 base pairs (bp) upstream of the transcriptional start site and increased the number of nucleotides separating these two sites from 0 to 16 in increments of two (Fig. 4.10A). The *EAF1* promoter was used because its constitutive activity was not significantly altered in the presence of glucose or galactose. Interestingly, the mean expression levels exhibited a sinusoidal trend with respect to the distance between binding sites with peaks at 2 and 12 bp in both the absence and presence of 2% galactose. Given that the there are approximately 10-10.5 bp per turn in the DNA helix, two Gal4p dimers are on the same side of the double helix for the observed optimal spacing of 2 and 14 bp. We hypothesize that a distance less than 2 bp produces steric hindrance leading to destabilization of Gal4p from its binding site. Corroborating previous studies, these data suggest that Gal4p molecules can interact synergistically to enhance gene expression [137, 31].

We next varied the number of *GAL4* binding sites from 1-3 with a distance of 16 bp separating these sites and measured the corresponding galactose dose responses. Increasing the number of binding sites produced an increase in the effective cooperativity of the dose response from approximately 1.25 to 2.2 (Fig. 4.10B), indicating that promoter level cooperativity can be augmented by additional binding sites. Finally, we measured the *in vivo* strengths of the individual binding sites from the *GAL1-10* promoter as a comparison with the results of the random mutagenesis regression analysis. Galactose induction of the the first and fourth binding sites did not produce a detectable change in fluorescence whereas the second and third binding binding sites produced a large fold change (Fig. 4.10C). These data demonstrate the ordering of strengths of these binding sites is $1 < 4 < 2 < 3$ in an isolated synthetic promoter context.

Discussion

In this work, we demonstrated that the combination of random mutagenesis and regression analysis can be used to build a detailed mapping of promoter sequence to phenotype. Regression analysis

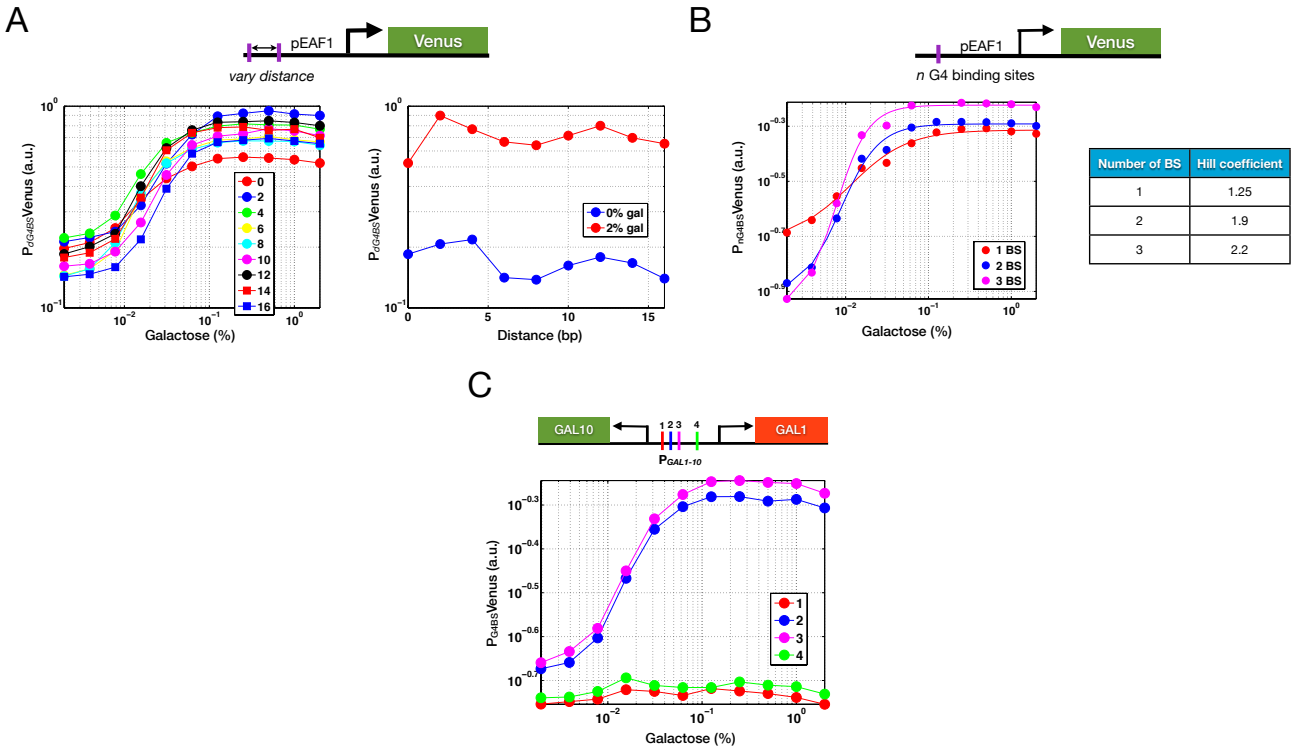


Figure 4.10: Rational methods to probe Gal4-dependent gene regulation. **(A)** Helical spacing of two *GAL4* binding sites in a synthetic promoter. The distance between two *GAL4* sites was varied in 2 base pair (bp) steps from 0 to 16 bp upstream of the *EAF1* promoter (top). Galactose dose responses (left) and the relationship between the spacing of *GAL4* binding sites and the mean expression levels (right). **(B)** The cooperativity of the galactose dose response increases with the number of *GAL4* binding sites. Lines represent fitted Hill functions to the data. **(C)** The four *GAL4* binding sites in $P_{GAL1\text{-}10}$ vary significantly in strength in the context of the *EAF1* promoter. Measurements were taken after 6 hours.

using regularization can solve underdetermined problems and thus circumvents the need for large libraries of sequence variants. These methods can be used to identify sensitive and robust regions of a promoter sequence, novel regulatory sites and does not require a priori knowledge of the regulatory inputs. As long as the promoter exhibits sufficient phenotypic variation, this approach can be used to discover cis sequence determinants for any phenotypic response.

Our model indicated that the *GAL4* binding sites and TATA boxes were the dominant contributing nucleotides in conditions with only an activation signal (galactose) and combinations of activation and repression (glucose + galactose). Based on the similarity in the weights between these two conditions, we propose that the main mechanism of repression involves inhibition of Gal4p as opposed to active repression. Instead of functioning as a potent source of repression, we hypothesize that repressors such as Mig1p fine tune the activity of the *GAL1-10* promoter.

Mutation 316 in the fourth *GAL4* binding site significantly reduced promoter activity. By contrast, this binding site did not produce a detectable change in fluorescence in the synthetic promoter context. These results indicate that the role of a binding site in the natural promoter can be significantly different than in a synthetic promoter context. Previous studies have shown that Gal4p dimers can interact cooperatively which provides an explanation for the observed difference in *in vivo* activity of site four in the genomic and synthetic promoter [137].

Methods

Strains and constructs

The plasmid used in this study was derived from a set of yeast single integration vectors constructed in the lab of Wendell Lim (UCSF). This vector contains a marker and targeting sequence for *TRP1* locus and was linearized for transformation by digesting with PmeI. Promoters were cloned between the PspOMI and XhoI restriction sites, Venus (YFP) between KpnI and PspOMI and mCherry between XhoI and BamHI. These plasmids contained an *ADH1* terminator downstream of BamHI site. The strain background used in this study was W303a.

Growth conditions and flow cytometry

Cells were grown in appropriate dropout media supplemented with 2% filter-sterilized raffinose at 30°C. Flow cytometry measurements were made an LSRII analyzer (BD Biosciences) with a High Throughput Sampler extension (HTS). A blue (488 nm) laser was used to excite Venus and the emission was detected on the LSRII using a 530/30 nm filter. At least 10,000 cells were collected for single measurement and between 1000-20,000 cells were collected for dynamics measurements.

Mutant library construction

The intergenic region between the *GAL10* and *GAL1* genes (668 bp) was cloned from the yeast genome. We used the GeneMorph II Random Mutagenesis Kit (Stratagene) to generate mutations in this promoter. Five successive rounds of mutagenesis were performed to increase the mutation rate by transferring 1μl of the PCR reaction into a fresh PCR reaction.

Lasso regression

Given a set of input measurements x_1,$x_2, \ldots x_p$ and an output measurement y, the lasso fits a linear model $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots b_p x_p$. The parameters are identified by minimizing $\sum_{i=1}^{N} (y_i - \hat{y})^2$ using the constraint $\sum_j |b_j| \leq \alpha$. The parameter α can be tuned. For small α, some of the b_j parameters can be forced to zero. Cross-validation was used to identify the best value of α.

Bibliography

[1] M Nakajima, K Imai, H Ito, T Nishiwaki, Y Murayama, H Iwasaki, T Oyama, and T Kondo. Reconstitution of circadian oscillation of cyanobacterial kaiC phosphorylation in vitro. *Science*, 308(5720):414–5, 2005.

[2] GM Suel, J Garcia-Ojalvo, LM Liberman, and MB Elowitz. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, 440(7083):545–50, 2006.

[3] SM Block, JE Segall, and HC Berg. Adaptation kinetics in bacterial chemotaxis. *Journal of bacteriology*, 154(1):312–23, 1983.

[4] A Novick and M Weiner. Enzyme induction as an all-or-none phenomenon. *Proceedings of the National Academy of Sciences*, 43(7):553–66, 1957.

[5] IG de Jong, P Haccou, and OP Kuipers. Bet hedging or not? A guide to proper classification of microbial survival strategies. *Bioessays*, 33(3):215–23, 2011.

[6] H Maamar, A Raj, and D Dubnau. Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science*, 317(5837):526–9, 2007.

[7] JW Veening, EJ Stewart, TW Berngruber, F Taddei, OP Kuipers, and LW Hamoen. Bet-hedging and epigenetic inheritance in bacterial cell development. *Proceedings of the National Academy of Sciences*, 105(11):4393–8, 2008.

[8] NQ Balaban, J Merrin, R Chait, L Kowalik, and S Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–5, 2004.

[9] SF Levy, N Ziv, and ML Siegal. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biology*, 10(5):e1001325, 2012.

[10] N Avraham, I Soifer, M Carmi, and N Barkai. Increasing population growth by asymmetric segregation of a limiting resource during cell division. *Molecular Systems Biology*, 9:656, 2013.

[11] MA Strauch, JJ Wu, RH Jonas, and JA Hoch. A positive feedback loop controls transcription of the spoOF gene, a component of the sporulation phosphorelay in *Bacillus subtilis*. *Molecular Microbiology*, 7(6):967–74, 1993.

[12] H Maamar and D Dubnau. Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Molecular Microbiology*, 56(3):615–24, 2005.

[13] O Brandman and T Meyer. Feedback loops shape cellular signals in space and time. *Science*, 322(5900):390–395, 2008.

[14] KJ Aström and RM Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.

[15] N Rosenfeld, MB Elowitz, and U Alon. Negative autoregulation speeds the response time of transcription networks. *Journal of Molecular Biology*, 323(5):785–793, 2002.

[16] A Becskei and L Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.

[17] L Ashall, CA Horton, DE Nelson, P Paszek, CV Harper, K Sillitoe, S Ryan, DG Spiller, JF Unitt, DS Broomhead, DB Kell, DA Rand, V See, and MR White. Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription. *Science*, 324(5924):242–6, 2009.

[18] SR Biggar and GR Crabtree. Cell signaling can direct either binary or graded transcriptional responses. *The EMBO Journal*, 20(12):3167–76, 2001.

[19] OS Venturelli, H El-Samad, and RM Murray. Synergistic dual positive feedback loops established by molecular sequestration generate robust bimodal response. *Proceedings of the National Academy of Sciences*, 109(48), 2012.

[20] M Acar, A Becskei, and A Oudenaarden. Enhancement of cellular memory by reducing stochastic transitions. *Nature*, 435:228–232, 2005.

[21] FJ Isaacs, J Hasty, CR Cantor, and JJ Collins. Prediction and measurement of an autoregulatory genetic module. *Proceedings of the National Academy of Sciences*, 100(13):7714–7719, 2003.

[22] TE Torchia, RW Hamilton, CL Cano, and JE Hopper. Disruption of regulatory gene GAL80 in *Saccharomyces cerevisiae*: effects on carbon-controlled regulation of the galactose/melibiose pathway genes. *Molecular Cellular Biology*, 4(8):1521–7, 1984.

[23] JF Tschopp, SD Emr, C Field, and R Schekman. GAL2 codes for a membrane-bound subunit of the galactose permease in *S*accharomyces cerevisiae.

[24] TE Torchia and JE Hopper. Genetic and molecular analysis of the GAL3 gene in the expression of the galactose/melibiose regulon of *Saccharomyces cerevisiae*. *Genetics*, 113(2):229–46, 1986.

[25] P Bhat and J Hopper. Overproduction of the GAL1 or GAL3 protein causes galactose-independent activation of the GAL4 protein: evidence for a new model of induction for the yeast GAL/MEL regulon. *Molecular Cell Biology*, 12(6):2701–2707, 1992.

[26] DJ Timson, HC Ross, and RJ Reece. Gal3p and Gal1p interact with the transcriptional repressor Gal80p to form a complex of 1:1 stoichiometry. *Biochemistry Journal*, 363(Pt 3):515–20, 2002.

[27] W Bajwa, TE Torchia, and JE Hopper. Yeast regulatory gene GAL3: carbon regulation; UASGal elements in common with GAL1, GAL2, GAL7, GAL10, GAL80, and MEL1; encoded protein strikingly similar to yeast and *Escherichia coli galactokinases*, journal = Molecular Cellular Biology, volume = 8, number = 8, pages = 3439-47, year = 1988.

[28] MK Hawkins and CD Smolke. The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in *Saccharomyces cerevisae*. *Journal of Biological Chemistry*, 281(19):13485–13492, 2006.

[29] CA Sellick, TA Jowitt, and RJ Reece. The effect of ligand binding on the galactokinase activity of yeast Gal1p and its ability to activate transcription. *Journal of Biological Chemistry*, 284(1):229–36, 2009.

[30] JB Thoden, CA Sellick, DJ Timson, RJ Reece, and HM Holden. Molecular structure of *Saccharomyces cerevisiae* Gal1p, a bifunctional galactokinase and transcriptional inducer. *Journal of Biological Chemistry*, 280(44):36905–11, 2005.

[31] TC Hittinger and SB Carroll. Gene duplication and the adaptive evolution of a classic genetic switch. *Nature*, 449:677–681, 2007.

[32] M Johnston. A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. *Microbiological Reviews*, 51(4):458–476, 1987.

[33] I Zacharloudakis, T Gligoris, and D Tzamarias. A yeast catabolic enzyme controls transcriptional memory. *Current Biology*, 17(23):2041–2046, 2007.

[34] D Abramczyk, S Holden, CJ Page, and RJ Reece. Interplay of a ligand sensor and an enzyme in controlling expression of the *Saccharomyces cerevisiae* GAL genes. *Eukaryotic Cell*, 11(3):334–42, 2011.

[35] MN Artyomov, J Das, M Kardar, and A Chakraborty. Purely stochastic binary decisions in cell signaling models without underlying deterministic bistabilities. *Proceedings of the National Academy of Sciences*, 104(48):18958–18963, 2007.

[36] M Samoilov, S Plyasunov, and A Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences*, 102(7):2310–2315, 2005.

[37] TB Kepler and TC Elston. Stochasticity in transcriptional regulation: origins, consequences and mathematical representations. *Biophysical Journal*, 81(6):3116–3136, 2001.

[38] I Lestas, J Paulsson, NE Ross, and G Vinnicombe. Noise in gene regulatory networks. *Automatic Control, IEEE Transactions*, 53(Special issue on systems biology):189–200, 2008.

[39] GM Guidi and A Goldbeter. Bistability without hysteresis in chemical reaction systems: A theoretical analysis of irreversible transitions between multiple steady states. *Journal of Physical Chemistry*, 101:9367–9376, 1997.

[40] T Nagai, K Ibata, ES Park, M Kubota, K Mikoshiba, and A Miyawaki. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology*, 20(1):87–90, 2002.

[41] HC Douglas and DC Hawthorne. Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*. *Genetics*, 49:837–44, 1964.

[42] MI Riley and RC Dickson. Genetic and biochemical characterization of the galactose gene cluster in *Kluyveromyces lactis*. *Journal of bacteriology*, 158(2):705–12, 1984.

[43] S Urlinger, U Baron, M Thellmann, MT Hasan, H Bujard, and W Hillen. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, 97(14):7963–8, 2000.

[44] D Angeli, JE Ferrell, and ED Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America*, 101(7):1822–7, 2004.

[45] RL Iman, JM Davenport, and DK Zeigler. Latin hypercube sampling program users guide. *Technical Report Sandia Labs*, 1980.

[46] Jr. Ferrell, JE. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochemical Sciences*, 21(12):460–6, 1996.

[47] NE Buchler and M Louis. Molecular titration and ultrasensitivity in regulatory networks. *Journal of Molecular Biology*, 384(5):1106–19, 2008.

[48] NE Buchler and FR Cross. Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Molecular Systems Biology*, 5:272, 2009.

[49] SY Kim and Jr Ferrell, JE. Substrate competition as a source of ultrasensitivity in the inactivation of Wee1. *Cell*, 128(6):1133–45, 2007.

[50] Jr Ferrell, JE. Feedback regulation of opposing enzymes generates robust, all-or-none bistable responses. *Current Biology*, 18(6):R244–5, 2008.

[51] MJ Dunlop, RS Cox, JH Levine, RM Murray, and MB Elowitz. Regulatory activity revealed by dynamic correlations in gene expression noise. *Nature Genetics*, 40(12):1493–8, 2008.

[52] CD Thron. A model for a bistable biochemical trigger of mitosis. *Biophysical Chemistry*, 57(2-3):239–51, 1996.

[53] Jr Ferrell, JE. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology*, 14(2):140–8, 2002.

[54] CV Cabrera, MC Alonso, and H Huikeshoven. Regulation of scute function by extramacrochaete in vitro and in vivo. *Development*, 120(12):3595–603, 1994.

[55] A Bhattacharya and NE Baker. A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell*, 147(4):881–92, 2011.

[56] M Van Doren, HM Ellis, and JW Posakony. The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development*, 113(1):245–55, 1991.

[57] JW Veening, WK Smits, and OP Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Reviews Microbiology*, 62:193–210, 2008.

[58] T Pfeiffer and S Bonhoeffer. Evolution of cross-feeding in microbial populations. *The American Naturalist*, 163(6), 2004.

[59] E Nevoigt, J Kohnke, CR Fischer, H Alper, U Stahl, and G Stephanopoulos. Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Applied Environmental Microbiology*, 72(8):5266–73, 2006.

[60] R Lutz and H Bujard. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Research*, 25(6):1203–10, 1997.

[61] CM Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[62] KJ Livak and TD Schmittgen. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. *Methods*, 25:402–408, 2001.

[63] MA Teste, M Duquenne, JM Francois, and JL Parrou. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Molecular Biology*, 10:99, 2009.

[64] A Dickenstein and IZ Emiris. *Solving Polynomial Equations: Foundations, Algorithms, and Applications (Algorithms and Computation in Mathematics)*. Springer, 2005.

[65] C Mateus and SV Avery. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast*, 16(14):1313–23, 2000.

[66] WK Huh, JV Falvo, LC Gerke, AS Carroll, RW Howson, JS Weissman, and EK O'Shea. Global analysis of protein localization in budding yeast. *Nature*, 425(6959):686–91, 2003.

[67] CB Tyson, PG Lord, and AE Wheals. Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. *Journal of bacteriology*, 138(1):92–8, 1979.

[68] JR Newman, S Ghaemmaghami, J Ihmels, DK Breslow, M Noble, JL DeRisi, and JS Weissman. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*, 441(7095):840–6, 2006.

[69] MA Savageau. *Biochemical Systems Analysis*. Addison-Wesley Publishing Company, 1976.

[70] G Peng and JE Hopper. Evidence for Gal3p's cytoplasmic location and Gal80p's dual cytoplasmic-nuclear location implicates new mechanisms for controlling Gal4p activity in *Saccharomyces cerevisiae*. *Molecular Cellular Biology*, 20(14):5140–8, 2000.

[71] R Wightman, R Bell, and RJ Reece. Localization and interaction of the proteins constituting the GAL genetic switch in *Saccharomyces cerevisiae*. *Eukaryotic Cell*, 7(12):2061–8, 2008.

[72] O Egriboz, F Jiang, and JE Hopper. Rapid GAL gene switch of *Saccharomyces cerevisiae* depends on nuclear Gal3, not nucleocytoplasmic trafficking of Gal3 and Gal80. *Genetics*, 189(3):825–36, 2011.

[73] M Hong, MX Fitzgerald, S Harper, C Luo, DW Speicher, and R Marmorstein. Structural basis for dimerization in DNA recognition by Gal4. *Structure*, 16(7):1019–26, 2008.

[74] V Pilauri, M Bewley, C Diep, and J Hopper. Gal80 dimerization and the yeast GAL gene switch. *Genetics*, 169(4):1903–14, 2005.

[75] K Melcher and E Xu. Gal80-Gal80 interaction on adjacent Gal4p binding sites is required for complete GAL gene repression. *The EMBO Journal*, 20:841–851, 2001.

[76] E Giniger and M Ptashne. Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proceedings of the National Academy of Sciences*, 85(2):382–6, 1988.

[77] DW Griggs and M Johnston. Promoter elements determining weak expression of the GAL4 regulatory gene of *Saccharomyces cerevisiae*. *Molecular Cell Biology*, 13(8):4999–5009, 1993.

[78] A Mitchell, GH Romano, B Groisman, A Yona, E Dekel, M Kupiec, O Dahan, and Y Pilpel. Adaptive prediction of environmental changes by microorganisms. *Nature*, 460(7252):220–4, 2009.

[79] TJ Perkins and PS Swain. Strategies for cellular decision-making. *Molecular Systems Biology*, 5:326, 2009.

[80] AL Bishop, FA Rab, ER Sumner, and SV Avery. Phenotypic heterogeneity can enhance rare-cell survival in 'stress-sensitive' yeast populations. *Molecular Microbiology*, 63(2):507–20, 2007.

[81] W Harder and L Dijkhuizen. Strategies of mixed substrate utilization in microorganisms. *Philosophical Transactions of the Royal Society*, 297(1088):459–80, 1982.

[82] J. Monod. Recherches sur la croissance des cultures bacteriennes (studies on the growth of bacterial cultures). *Actualites Scientifique et Industrielles*, 911:1–215, 1942.

[83] JM Gancedo. Yeast carbon catabolite repression. *Microbiology Molecular Biology Review*, 62(2):334–61, 1998.

[84] SA Ramsey, JJ Smith, D Orrell, M Marelli, TW Petersen, P de Atauri, H Bolouri, and JD Aitchison. Dual feedback loops in the GAL regulon suppress cellular heterogeneity in yeast. *Nature Genetics*, 38(9):1082–7, 2006.

[85] H Zhou and F. Winston. NRG1 is required for glucose repression of the SUC2 and GAL genes of *Saccharomyces cerevisiae*. *BMC Genetics*, 2:5, 2001.

[86] JO Nehlin, M Carlberg, and H Ronne. Control of yeast GAL genes by MIG1 repressor: a transcriptional cascade in the glucose response. *EMBO J*, 10(11):3373–7, 1991.

[87] KH Wong and K Struhl. The Cyc8-Tup1 complex inhibits transcription primarily by masking the activation domain of the recruiting protein. *Genes and Development*, 25(23):2525–39, 2011.

[88] RL Smith and AD Johnson. Turning genes off by Ssn6-Tup1: a conserved system of transcriptional repression in eukaryotes. *Trends in Biochemical Sciences*, 25(7):325–30, 2000.

[89] CA Keleher, MJ Redd, J Schultz, M Carlson, and AD Johnson. Ssn6-Tup1 is a general repressor of transcription in yeast. *Cell*, 68(4):709–19, 1992.

[90] E Larschan and F Winston. The *S. cerevisiae* SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes and Development*, 15(15):1946–56, 2001.

[91] M Papamichos-Chronakis, T Petrakis, E Ktistaki, I Topalidou, and D Tzamarias. Cti6, a PHD domain protein, bridges the Cyc8-Tup1 corepressor and the SAGA coactivator to overcome repression at GAL1. *Molecular Cell*, 9(6):1297–305, 2002.

[92] D Ahuatzi, A Riera, R Pelaez, P Herrero, and F Moreno. Hxk2 regulates the phosphorylation state of Mig1 and therefore its nucleocytoplasmic distribution. *Journal of Biological Chemistry*, 282(7):4485–93, 2007.

[93] H Ma and D Botstein. Effects of null mutations in the hexokinase genes of *Saccharomyces cerevisiae* on catabolite repression. *Molecular Cellular Biology*, 6(11):4046–52, 1986.

[94] F Randez-Gil, P Herrero, P Sanz, JA Prieto, and F Moreno. Hexokinase pii has a double cytosolic-nuclear localisation in *Saccharomyces cerevisiae*. *FEBS Letters*, 425(3):475–8, 1998.

[95] D Ahuatzi, P Herrero, T de la Cera, and F Moreno. The glucose-regulated nuclear localization of hexokinase 2 in *Saccharomyces cerevisiae* is Mig1-dependent. *Journal of Biological Chemistry*, 279(14):14440–6, 2004.

[96] P Fernandez-Garcia, R Pelaez, P Herrero, and F Moreno. Phosphorylation of yeast hexokinase 2 regulates its nucleocytoplasmic shuttling. *Journal of Biological Chemistry*, 287(50):42151–64, 2012.

[97] M Papamichos-Chronakis, T Gligoris, and D Tzamarias. The Snf1 kinase controls glucose repression in yeast by modulating interactions between the Mig1 repressor and the Cyc8-Tup1 co-repressor. *EMBO Reports*, 5(4):368–72, 2004.

[98] A Arkin, J Ross, and HH McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics*, 149(4):1633–48, 1998.

[99] A Hernday, M Krabbe, B Braaten, and D Low. Self-perpetuating epigenetic pili switches in bacteria. *Proceedings of the National Academy of Sciences*, 99:16470–6, 2002.

[100] S Ostergaard, KO Walloe, SG Gomes, L Olsson, and J Nielsen. The impact of GAL6, GAL80, and MIG1 on glucose control of the GAL system in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 1(1):47–55, 2001.

[101] LN Sierkstra, NP Nouwen, JM Verbakel, and CT Verrips. Analysis of glucose repression in *Saccharomyces cerevisiae* by pulsing glucose to a galactose-limited continuous culture. *Yeast*, 8(12):1077–87, 1992.

[102] M Johnston and J Dover. Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA binding domain. *Proceedings of the National Academy of Sciences*, 84(8):2401–5, 1987.

[103] JF Louvion, B Havaux-Copf, and D Picard. Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast. *Gene*, 131(1):129–34, 1993.

[104] AP Gasch, PT Spellman, CM Kao, O Carmel-Harel, MB Eisen, G Storz, D Botstein, and PO Brown. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell*, 11(12):4241–57, 2000.

[105] AM Pietersen and M van Lohuizen. Stem cell regulation by polycomb repressors: postponing commitment. *Current Opinion Cellular Biology*, 20(2):201–7, 2008.

[106] JH Levine, ME Fontes, J Dworkin, and MB Elowitz. Pulsed feedback defers cellular differentiation. *PLoS Biology*, 10(1), 2012.

[107] D Pincus, MW Chevalier, T Aragon, E van Anken, SE Vidal, H El-Samad, and P Walter. Bip binding to the er-stress sensor ire1 tunes the homeostatic behavior of the unfolded protein response. *PLoS Biology*, 8(7):e1000415, 2010.

[108] S Levy and N Barkai. Coordination of gene expression with growth rate: a feedback or a feed-forward strategy? *FEBS Letters*, 583(24):3974–8, 2009.

[109] A Anders, H Lilie, K Franke, L Kapp, J Stelling, ED Gilles, and KD Breunig. The galactose switch in *Kluyveromyces lactis* depends on nuclear competition between gal4 and Gal1 for Gal80 binding. *Journal of Biological Chemistry*, 281(39):29337–48, 2006.

[110] LM Raamsdonk, JA Diderich, A Kuiper, M van Gaalen, AL Kruckeberg, JA Berden, and K Van Dam. Co-consumption of sugars or ethanol and glucose in a *Saccharomyces cerevisiae*strain deleted in the HXK2 gene. *Yeast*, 18(11):1023–33, 2001.

[111] J van den Brink, M Akeroyd, R van der Hoeven, JT Pronk, JH de Winde, and P Daran-Lapujade. Energetic limits to metabolic flexibility: responses of *Saccharomyces cerevisiae* to glucose-galactose transitions. *Microbiology*, 155(Pt 4):1340–50, 2009.

[112] T Pfeiffer, S Schuster, and S Bonhoeffer. Cooperation and competition in the evolution of atp-producing pathways. *Science*, 292(5516):504–7, 2001.

[113] JP van Dijken, RA Weusthuis, and JT Pronk. Kinetics of growth and sugar consumption in yeasts. *Antonie Van Leeuwenhoek*, 63(3-4):343–52, 1993.

[114] J Wu and RJ Trumbly. Multiple regulatory proteins mediate repression and activation by interaction with the yeast Mig1 binding site. *Yeast*, 14(11):985–1000, 1998.

[115] J Ostling, JP Cassart, J Vandenhaute, and H Ronne. Four hydrophobic amino acid residues in the c-terminal effector domain of the yeast mig1p repressor are important for its in vivo activity. *Molecular Genetics and Genomics*, 260(2-3):269–79, 1998.

[116] MA Treitel and M Carlson. Repression by SSN6-TUP1 is directed by MIG1, a repressor/activator protein. *Proceedings of the National Academy of Sciences*, 92(8):3132–6, 1995.

[117] AZ Ansari, RJ Reece, and M Ptashne. A transcriptional activating region with two contrasting modes of protein interaction. *Proceedings of the National Academy of Sciences*, 95(23):13543–8, 1998.

[118] CT Harbison, DB Gordon, TI Lee, NJ Rinaldi, KD Macisaac, TW Danford, NM Hannett, JB Tagne, DB Reynolds, J Yoo, EG Jennings, J Zeitlinger, DK Pokholok, M Kellis, PA Rolfe,

KT Takusagawa, ES Lander, DK Gifford, E Fraenkel, and RA Young. Transcriptional regulatory code of a eukaryotic genome. *Nature*, 431(7004):99–104, 2004.

[119] Y Shen, F Yue, DF McCleary, Z Ye, L Edsall, S Kuan, U Wagner, J Dixon, L Lee, VV Lobanenkov, and B Ren. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409):116–20, 2012.

[120] J Gertz, ED Siggia, and BA Cohen. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, 457(7226):215–8, 2009.

[121] J Gertz and BA Cohen. Environment-specific combinatorial cis-regulation in synthetic promoters. *Molecular Systems Biology*, 5:244, 2009.

[122] E Sharon, Y Kalma, A Sharp, T Raveh-Sadka, M Levo, D Zeevi, L Keren, Z Yakhini, A Weinberger, and E Segal. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat Biotechnology*, 30(6):521–30, 2012.

[123] M Johnston and RW Davis. Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Molecular Cellular Biology*, 4(8):1440–8, 1984.

[124] AM Stathopoulos and MS Cyert. Calcineurin acts through the CRZ1/TCN1-encoded transcription factor to regulate gene expression in yeast. *Genes and Development*, 11(24):3432–44, 1997.

[125] G Hornung, M Oren, and N Barkai. Nucleosome organization affects the sensitivity of gene expression to promoter mutations. *Molecular Cell*, 46(3):362–8, 2012.

[126] G Hornung, R Bar-Ziv, D Rosin, N Tokuriki, DS Tawfik, M Oren, and N Barkai. Noise-mean relationship in mutated promoters. *Genome Research*, 22(12):2409–17, 2012.

[127] JC Slot and A Rokas. Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi. *Proceedings of the National Academy of Sciences*, 107(22):10136–41, 2010.

[128] GI Lang and D Botstein. A test of the coordinated expression hypothesis for the origin and maintenance of the GAL cluster in yeast. *PLoS One*, 6(9):e25290, 2011.

[129] M Kellis, N Patterson, M Endrizzi, B Birren, and ES Lander. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, 423(6937):241–54, 2003.

[130] Z Xu, W Wei, J Gagneur, F Perocchi, S Clauder-Munster, J Camblong, E Guffanti, F Stutz, W Huber, and LM Steinmetz. Bidirectional promoters generate pervasive transcription in yeast. *Nature*, 457(7232):1033–7, 2009.

[131] H Neil, C Malabat, Y d'Aubenton Carafa, Z Xu, LM Steinmetz, and A Jacquier. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*, 457(7232):1038–42, 2009.

[132] PA Romero, E Stone, C Lamb, L Chantranupong, A Krause, A Miklos, RA Hughes, B Fechtel, AD Ellington, FH Arnold, and G Georgiou. Schema designed variants of human arginase i & ii reveal sequence elements important to stability and catalysis. *ACS Synthetic Biology*, 1(6):221–228, 2012.

[133] Y Li, DA Drummond, AM Sawayama, CD Snow, JD Bloom, and FH Arnold. A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments. *Nature Biotechnology*, 25(9):1051–6, 2007.

[134] A Bar-Even, J Paulsson, N Maheshri, M Carmi, E O'Shea, Y Pilpel, and N Barkai. Noise in protein expression scales with natural protein abundance. *Nature Genetics*, 38(6):636–43, 2006.

[135] R Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58(1):267–288, 1996.

[136] S Vashee, H Xu, SA Johnston, and T Kodadek. How do Zn2 cys6 proteins distinguish between similar upstream activation sites? comparison of the DNA-binding specificity of the GAL4 protein *in vitro* and *in vivo*. *Journal of Biological Chemistry*, 268(33):24699–706, 1993.

[137] T Kang, T Martins, and I Sadowski. Wild type GAL4 binds cooperatively to the GAL1-10 UASG in vitro. *Journal of Biological Chemistry*, 268(13):9629–35, 1993.

[138] J Stelling, ED Gilles, and 3rd Doyle, FJ. Robustness properties of circadian clock architectures. *Proceedings of the National Academy of Sciences*, 101(36):13210–5, 2004.

[139] N Barkai and S Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–7, 1997.

[140] G von Dassow, E Meir, EM Munro, and GM Odell. The segment polarity network is a robust developmental module. *Nature*, 406(6792):188–92, 2000.

[141] A Eldar, R Dorfman, D Weiss, H Ashe, BZ Shilo, and N Barkai. Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. *Nature*, 419(6904):304–8, 2002.

[142] H Kitano. Biological robustness. *Nature Review Genetics*, 5(11):826–37, 2004.

[143] ME Csete and JC Doyle. Reverse engineering of biological complexity. *Science*, 295(5560):1664–9, 2002.

[144] N Bluthgen and H Herzel. How robust are switches in intracellular signaling cascades? *Journal of Theoretical Biology*, 225(3):293–300, 2003.

[145] H El-Samad, H Kurata, JC Doyle, CA Gross, and M Khammash. Surviving heat shock: control strategies for robustness and performance. *Proceedings of the National Academy of Sciences*, 102(8):2736–41, 2005.

[146] KL Brown and KT Hughes. The role of anti-sigma factors in gene regulation. *Molecular Microbiology*, 16(3):397–404, 1995.

[147] JD Norton, RW Deed, G Craggs, and F Sablitzky. Id helix-loop-helix proteins in cell growth and differentiation. *Trends Cellular Biology*, 8(2):58–65, 1998.