

External Control of the *GAL* Network in *S. cerevisiae*: A View from Control Theory

Ruoting Yang¹, Scott C. Lenaghan¹, John P. Wiksw², Mingjun Zhang^{1*}

1 Department of Mechanical, Aerospace and Biomedical Engineering, The University of Tennessee, Knoxville, Tennessee, United States of America, **2** Vanderbilt Institute for Integrative Biosystems Research and Education, Departments of Biomedical Engineering, Molecular Physiology & Biophysics, and Physics & Astronomy, Vanderbilt University, Nashville, Tennessee, United States of America

Abstract

While there is a vast literature on the control systems that cells utilize to regulate their own state, there is little published work on the formal application of control theory to the external regulation of cellular functions. This paper chooses the *GAL* network in *S. cerevisiae* as a well understood benchmark example to demonstrate how control theory can be employed to regulate intracellular mRNA levels via extracellular galactose. Based on a mathematical model reduced from the *GAL* network, we have demonstrated that a galactose dose necessary to drive and maintain the desired *GAL* genes' mRNA levels can be calculated in an analytic form. And thus, a proportional feedback control can be designed to precisely regulate the level of mRNA. The benefits of the proposed feedback control are extensively investigated in terms of stability and parameter sensitivity. This paper demonstrates that feedback control can both significantly accelerate the process to precisely regulate mRNA levels and enhance the robustness of the overall cellular control system.

Citation: Yang R, Lenaghan SC, Wiksw JP, Zhang M (2011) External Control of the *GAL* Network in *S. cerevisiae*: A View from Control Theory. PLoS ONE 6(4): e19353. doi:10.1371/journal.pone.0019353

Editor: Diego Di Bernardo, Fondazione Telethon, Italy

Received: September 15, 2010; **Accepted:** March 31, 2011; **Published:** April 29, 2011

Copyright: © 2011 Yang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors would like to thank the National Academies Keck Futures Initiative for its support of this research, as well as the US Defense Threat Reduction Agency (HDTRA-09-1-0013), the NIH National Institute on Drug Abuse (RC2DA028981), and the Vanderbilt Institute for Integrative Biosystems Research and Education. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: mjzhang@utk.edu

Introduction

The complexity of nested feedback control loops associated with complex biological systems creates many challenges to understanding biological systems. At the same time, the need for precise control of gene expression is increasing in fields such as genetics, cell and molecular engineering, and in disease treatment via gene expression control. A straightforward application for control of gene expression might be to generate a maximum output in bioreactor-based systems. The outputs from this type of application would typically by specific proteins of interest for biopharmaceuticals. The approach can be further applied to create a system in which the control of a gene network may be coupled with the expression of a recombinant protein.

On the other hand, systems biologists have built numerous mathematical models for gene and protein networks that cells utilize to control themselves [1]. The applications of such models have gone beyond simple qualitative understanding of network dynamics. Some models have been used in synthetic biology to affect metabolism and eventually control the biological processes toward desired outputs [2]. The full advantages of control theory, however, have not been realized, and there is little work on external control of cellular functions such as gene expression.

From a control theory perspective, control techniques used for complex engineering systems should in principle be applicable to the regulation of cellular systems. There are in general two potential challenges. First, it is not easy to find simple mathematical models for cellular system control. Most existing cellular mathematical models are not formulated in a standard

control system framework. Typically, they lack global feedback loops from outputs to dynamically adjust the control dosage. Second, model parameters are usually very sensitive due to large uncertainties. The resulting control algorithm often becomes fragile, and the control system can remain stable only under small disturbances. Open-loop constant controls – for example, the gene knockdown or overexpression technique for yeast performed without adjusting dosages in real-time in response to the observed changes – are vulnerable to disturbances and individual heterogeneity.

It is desirable to have model-based, fine-tuned external control approaches to precisely regulate a reverse-engineered target. Unfortunately, achieving a reliable feedback control remains a challenge for cellular systems. Recently, Cantone *et al.* developed the first reverse-engineering benchmark system [3], which is based on the *GAL* network using real-time image feedback. Control engineers have begun to apply closed-loop control ideas based on this benchmark [4,5].

In this paper we have employed model reduction techniques to simplify the complex *GAL* network, and further demonstrate that the simplified model is favorable for control system design. We also present the benefit of closed-loop system design. Our goal is to demonstrate the usefulness of control theory for precise external regulation of cellular systems and to open discussions in the field about possible future benefits of more advanced theoretical control.

The *GAL* network in yeast, *Saccharomyces cerevisiae*, is one of the most studied gene networks, and hence is well suited to serve as a demonstration, benchmark system. Galactose utilization in the

network can be broken up into four distinct stages based on proteins present in the cytoplasm. In the first stage *GAL2p*, the galactose transporter protein, allows galactose to enter into the cytoplasm [6]. Once internalized, *GAL3p* binds to the cytoplasmic galactose and is activated. This causes the galactose to bind and inhibit the action of *GAL80p* [7,8]. With *GAL80p* bound, *GAL4p* is able to activate the transcription of *GAL2*, *GAL3*, and *GAL80* genes. In the presence of glucose, *GAL4p* is also inhibited, which leads to inhibition of *GAL1* expression [9]. Without accounting for the glucose network, the *GAL* system is composed of two positive feedback loops, *GAL3p* and *GAL2p*, and two negative feedback loops, *GAL80p* and *GAL1p*. The glucose network proteins bind upstream *GAL1*, *GAL3*, and *GAL4*, and suppress transcription through the action of Mig1. The glucose network also prevents the binding of *GAL2p* and external galactose and shuts down the galactose transportation. A mathematical model of the galactose network has previously been established by de Atauri *et al.* [10] and Ramsey *et al.* [11]. Bennett *et al.* combined a simplified glucose network with the galactose network model [12], and identified the model parameters using experimental data. A diagram of the simplified glucose and galactose networks used in this study is shown in **Figure 1**.

The key factor in designing a control system for a complex network that can be controlled externally is to design the system with measurable input(s) and output(s). For our study of the *GAL* network, the measurable inputs into the system are the concentrations of galactose and glucose. The most common mechanism currently available to monitor the output is a reporter to a gene of interest, normally green fluorescent protein (GFP), which allows researchers to measure the expression level of the gene. In the theoretical control model designed in this study, a GFP reporter coupled to *GAL1* can be used to measure the output from the system. In practice, a spectrofluorometer or a microscope with appropriate excitation and emission filters, possibly coupled to a microfluidic device, could be used for measuring the GFP concentration [3]. It would then be possible to develop a control algorithm for the expression of *GAL1* by tuning the concentration of galactose delivered externally to the cells. This model could be expanded to the control of other genes of interest (*GAL2*, *GAL3*, and *GAL80*) coupled to separate reporters.

Another challenge for control system design is the network scale and model complexity [13]. A gene network model is often constructed based on chemical kinetics with tens or hundreds of states and many parameters [12,14]. The most successful model for the yeast cell cycle alone has 61 coupled ordinary differential equations and 141 parameters [15]. In such models, many important system characteristics, such as steady states, cannot be solved analytically and have to be estimated using numerical simulation. Advanced control mechanisms directly applied to these systems may waste considerable time on calculation of complicated, yet inessential nonlinear reaction terms. This complexity may lead to fragility of the controlled system, causing the system to collapse due to noise and parameter perturbation. The reduction of model complexity thus becomes extremely important for complex control system design and implementation. One effective tool for model reduction is parameter sensitivity analysis, which elucidates the dependence of system dynamics on the parameters. A small sensitivity measure for a parameter implies that the value of this parameter can be substituted for a wide range of values without altering the system dynamics. Some reaction terms can thus be deemed negligible and result in a reduced model. The control design based on the reduced model can be expected to have a similar performance to that of the original model. This paper uses the *GAL1* mRNA level as a measurable output to be controlled (with the assumption that mRNA level is highly correlated with

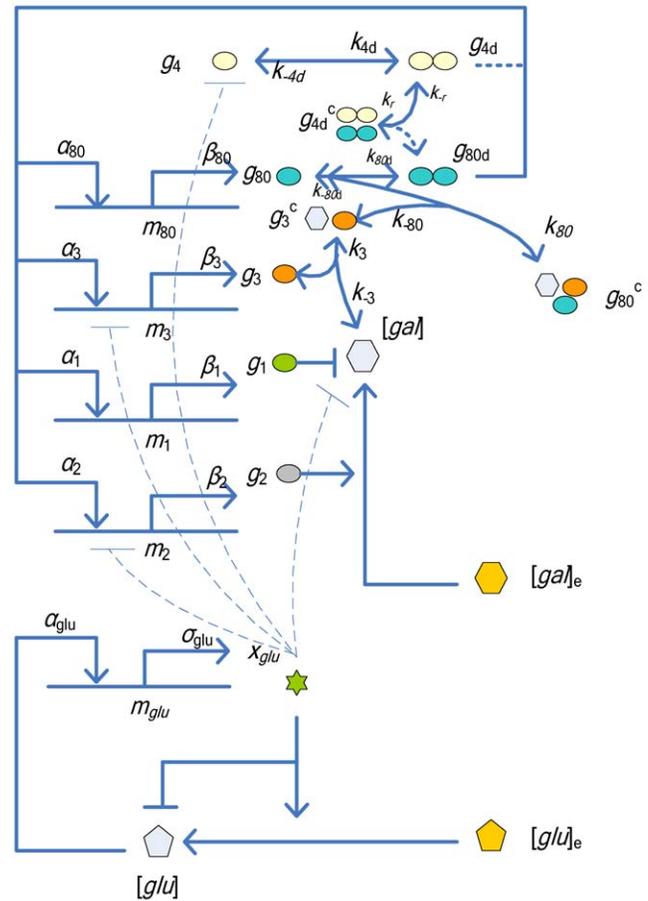


Figure 1. The gene regulatory network for galactose utilization (redrawn from Bennett *et al.* (2008) [12]). The extracellular inputs are $[gal]_e$ and $[glu]_e$. The natural output of the network is a group of proteins. The *GAL4* protein g_4 binds to upstream activation sites and activates the regulatory genes in the galactose network. The *GAL80* gene inhibits the inducing effects of *GAL4* and thereby provides negative feedback in the system. *GAL3* enhances expression of *GAL4* by binding with internal galactose (Gal), forming a *GAL3*-galactose complex, g_3^c , that inactivates g_{80} by binding to it and resulting in a complex g_{80}^c . In addition, the transporter *GAL2* increases the amount of internal galactose, which stimulates the galactose network. We use the *GAL1* mRNA level, m_1 , as the measurable output to control the system. According to sensitivity analysis, the interaction loops involving *GAL4p* dimerization (the dotted lines) can be eliminated from the original model to create a reduced model. doi:10.1371/journal.pone.0019353.g001

GAL1 translation) and investigates the sensitivity between the output as well as model parameters. The resulting reduced model based on sensitivity analysis successfully separates the galactose utilization network and the glucose network, and reduces the *GAL4p* subsystem and associated complexes. Furthermore, global sensitivity was also conducted to show how the system changes due to the simultaneous variation of all the parameters over a wide range of values. We have concluded that the global sensitivity analysis is consistent with the results of local sensitivity. In addition, the nonlinearity of the system is largely reduced, while maintaining a small deviation in the system dynamics. After system analysis, the original complex system was simplified to allow for the control system design. The goal of the control is to maintain one of the *GAL* mRNAs at a desired level.

In this paper, both a constant open-loop control and a proportional-output feedback closed-loop control are designed based on the reduced model. First, an analytic formula for the precise dosage of the external

galactose is given, instead of manually tuning galactose by trial and error. Feedback control is introduced to improve the robustness of the controlled system and enhance the convergence rates. The simulation results show that both controls can achieve the control objective, *i.e.*, maintaining *GAL1* mRNA at a desired level. Similar analytic galactose dosages can be achieved for all the other measurable outputs, such as *GAL2*, *GAL3*, and *GAL80*. The difference between the open-loop control of the system with constant control input, and feedback control with time-dependent parameters, is that the feedback control significantly shortens the time required to achieve the steady state. The feedback loop indirectly increases the degradation rate of the internal galactose so that the system more rapidly reaches balance. The feedback control also significantly reduces parameter sensitivity. Feedback control takes advantage of information about the system state to regulate output, and in doing so significantly decreases the local sensitivity of measurements to the system parameters. Hence feedback control can resist much larger parameter perturbations/uncertainties as compared to the constant open-loop control; that is, an open-loop control with constant control input.

Methods

Mathematical Model for Yeast GAL Network

The *GAL* network is a good starting point to demonstrate feedback and feed forward external controls for cellular systems. Though simple, it is well understood and is easily manageable to interpret experimental results. On the other hand, it is complex enough to test sophisticated control algorithms. As illustrated in Figure 1, a mathematical model for the *GAL* network has been proposed based on the interactions between proteins and internalized galactose. The figure shows a gene regulatory network for galactose utilization (redrawn from Bennett *et al.* (2008) [12]). The extracellular inputs are $[gal]_e$ and $[glu]_e$. The natural output of the network is a group of proteins. The *GAL4* protein g_4 binds to upstream activation sites and activates the regulatory genes in the galactose network. The *GAL80* gene inhibits the inducing effects of *GAL4* and thereby provides negative feedback in the system. *GAL3* enhances expression of *GAL4* by binding with internal galactose (Gal), forming a *GAL3*-galactose complex, g_3^c , that inactivates g_{80} by binding to it and resulting in a complex g_{80}^c . In addition, the transporter *GAL2* increases the amount of internal galactose, which stimulates the galactose network. We use the *GAL1* mRNA level, m_1 , as the measurable output to control the system. According to sensitivity analysis, the interaction loops involving *GAL4*p dimerization (the dotted lines) can be eliminated from the original model to create a reduced model. In the figure, g_i is the number of the galactose network protein monomers ($i = 1, 2, 3, 4, 80$), m_i is amount of mRNA ($i = 1, 2, 3, 80$), g_{id} is the number of protein dimers ($i = 4, 80$), g_3^* is the number of *GAL3*p proteins bound to galactose, g_{4d}^* is the number of *GAL4*p dimers bound to Gal80p dimers, and g_{80}^* is the number of *GAL80*p proteins bound to the Gal3p-galactose complex.

A complete mathematical model of the above gene network, which we term the Original model, can be described with 22 mathematical equations [12]. The first five represent the mass-action kinetics of galactose protein monomers, including dimerization, and the interaction with the internal galactose:

$$\dot{g}_1 = \beta_1 m_1 - \delta_1 g_1 \quad (1)$$

$$\dot{g}_2 = \beta_2 m_2 - \delta_2 g_2 \quad (2)$$

$$\dot{g}_3 = \beta_3 m_3 - \delta_1 g_3 - k_3 x_{gal} g_3 + k_{-3} g_3^c \quad (3)$$

$$\dot{g}_4 = \beta_4 R(x_{glu}) - \delta_1 g_4 - 2k_{4d} g_4^2 + 2k_{-4d} g_{4d} \quad (4)$$

$$\begin{aligned} \dot{g}_{80} = & \beta_{80} R(x_{glu}) - \delta_1 g_{80} - k_{80} g_3^c g_{80} + \\ & k_{-80} g_{80}^c + 2k_{-80d} g_{80d} - 2k_{80d} g_{80}^2. \end{aligned} \quad (5)$$

Equations (6)–(9) describe mRNA kinetics accounting for the transcription and translation of the galactose genes as well as the degradation of mRNA

$$\dot{m}_1 = \alpha_1 R(x_{glu}) F(g_{4d}, g_{80d}, 4) - \gamma_1 m_1 - \frac{v_{sd} m_1 x_{glu}}{k_{sd} + m_1} \quad (6)$$

$$\dot{m}_2 = \alpha_2 F(g_{4d}, g_{80d}, 2) - \gamma_2 m_2 \quad (7)$$

$$\dot{m}_3 = \alpha_3 R(x_{glu}) F(g_{4d}, g_{80d}, 1) - \gamma_3 m_3 - \frac{v_{sd} m_3 x_{glu}}{k_{sd} + m_3} \quad (8)$$

$$\dot{m}_{80} = \alpha_{80} F(g_{4d}, g_{80d}, 1) - \gamma_{80} m_{80}. \quad (9)$$

Equations (10) – (14) represent the kinetics of protein dimers and the associated complexes:

$$\dot{g}_{4d} = k_{4d} g_4^2 - k_{-4d} g_{4d} - k_r g_{80d} g_{4d} + k_{-r} g_{4d}^c - \delta_1 g_{4d} \quad (10)$$

$$\dot{g}_{80d} = k_{80d} g_{80}^2 - k_{-80d} g_{80d} - k_r g_{80d} g_{4d} + k_{-r} g_{4d}^c - \delta_1 g_{80d} \quad (11)$$

$$\dot{g}_{4d}^c = k_r g_{80d} g_{4d} - k_{-r} g_{4d}^c - \delta_1 g_{4d}^c \quad (12)$$

$$\dot{g}_3^c = k_3 g_3 [gal] - k_{-3} g_3^c - k_{80} g_{80} g_3^c + k_{-80} g_{80}^c - \delta_1 g_3^c \quad (13)$$

$$\dot{g}_{80}^c = k_{80} g_{80} g_3^c - k_{-80} g_{80}^c - \delta_1 g_{80}^c. \quad (14)$$

The metabolic reactions and transport of galactose are represented by a single equation:

$$\begin{aligned} \dot{[gal]} = & R(x_{glu}) T_1([gal], [gal]_e) - \frac{k_{cat} g_1 [gal]}{k_{mgk} + [gal]} - \\ & k_3 g_3 [gal] + k_{-3} g_3^c - \delta_1 [gal]. \end{aligned} \quad (15)$$

Equations (16) – (18) describe a simplified glucose network, including the glucose-mediated enzymatic decay of *GALI* and *GAL3* mRNA:

$$\dot{m}_{glu} = \frac{\alpha_{glu} + \varepsilon_{glu} ([glu]/c_{glu})^b}{1 + ([glu]/c_{glu})^b} - \gamma_{glu} m_{glu} \quad (16)$$

$$\dot{x}_{glu} = \sigma_{glu} m_{glu} - \delta_{glu} x_{glu} \quad (17)$$

$$[\dot{glu}] = ([glu], [glu]_e) - \frac{\mu_{glu} [glu] x_{glu}}{k_{glu} + [glu]} - \delta_d [glu], \quad (18)$$

where m_{glu} , x_{glu} , and $[glu]$ are the amount of glucose network mRNA, associated protein, and cellular internal glucose, respectively. They compose a simplified glucose network. The inhibitory effect due to products of the glucose network that act on various processes of the galactose network is represented as follows,

$$R(x_{glu}) = \frac{\rho^q}{\rho^q + x_{glu}^q}, \quad (19)$$

where

$$T_1([gal], [gal]_e) = \frac{k_{tr} g_2}{k_{mtr} + [gal]_e + [gal] + a_{tr} [gal]_e [gal] / k_{mtr}} [gal]_e - [gal] \quad (20)$$

and

$$T_2([glu], [glu]_e) = \frac{k_{tr2} x_{glu}}{k_{mtr2} + [glu]_e + [glu] + a_{tr2} [glu]_e [glu] / k_{mtr2}} [glu]_e - [glu] \quad (21)$$

are the transport rates of external galactose and external glucose into the cell, respectively. The cooperative fractional saturation function describing the number of upstream activation sites occupied on a promoter, assuming that N sites exist [10], is given by

$$F(g_{4d}, g_{80d}, N) = \frac{\sum_{i=0}^N \binom{N}{i} (K_q K_p g_{4d} g_{80d})^i \sum_{h=1}^{N-i} \binom{N-i}{h} C_P^{h+i-1} C_Q^{i-1} (K_p g_{4d})^h}{\sum_{i=0}^N \binom{N}{i} (K_q K_p g_{4d} g_{80d})^i \sum_{h=0}^{N-i} \binom{N-i}{h} C_P^{h+i-1} C_Q^{i-1} (K_p g_{4d})^h} \quad (22)$$

The above mathematical model is mainly based on the biochemical reactions occurring throughout the network. This approach of modeling the intermediates in a pathway has been widely used in both chemistry and biology. The variables and parameters used for the model are defined and summarized in **Tables 1** and **2**.

Our goal for galactose network control is to regulate the *GAL* family mRNA level by manipulating galactose and glucose concentrations. In this paper, we select *GAL1* as the gene of interest. The mRNA level of this target gene can be measured experimentally using a GFP conjugated to *GAL1*, which can be measured for intact cells, or by means of a microarray or other assay that requires lysing a small fraction of the cells under

observation. In practice, other reporter genes may be used to monitor *GAL1* expression, for example, the red fluorescent protein from the gene dsRed [16]. In addition, it is notable that glucose suppresses all *GAL* genes [17]. External glucose must be kept at a very low level, *i.e.*, $[glu]_e \approx 0$, or the *GAL* network will be inhibited. Once no external glucose is supplied, the glucose network will soon reach its steady state, which means one can focus on the galactose network.

From a control engineering perspective, one key step is to identify inputs (control) and outputs (preferably measurable), so that the system can be put into a control framework for discussion. $[gal]_e$ and $[glu]_e$, external galactose and glucose levels, are the control variables in this study, while the mRNA levels of *GAL1*, *GAL2*, *GAL3*, and *GAL80*, m_1 , m_2 , m_3 , m_{80} , are measurable variables. The goal of this research was to design a time-course of $[gal]_e$ such that one of the mRNAs m_1 , m_2 , m_3 , m_{80} is maintained at a desired level.

While there are several methods that might be suitable for culturing yeast under external control, microfluidic devices would be ideal for experimental implementation of the cellular control system [18]. Fluorescence microscopy could be used to quantify, in real-time, the levels of GFP-labeled *Gal1p* from a small, synchronized population of *S. cerevisiae*, with the output of a camera or photomultiplier tube serving as the sensor measurement to the control system. Ion mobility-mass spectrometry could also be used to monitor the secreted metabolites and signaling molecules in real-time and intracellular species after cell lysis [19]. Microfluidic valves could control the concentrations of various chemicals and serve as the control system outputs.

In the next section, we will conduct a sensitivity analysis to simplify this subsystem. As shown in **Figure 1**, the interaction loops involving *GAL4p* dimerization can be removed. The reduced

Table 1. Model Variables.

Variable	Description	Initial value (molec.) [11]
g_1 (molec./cell)	Gal1p	132.3
g_2 (molec./cell)	Gal2p	1156.7
g_3 (molec./cell)	Gal3p	4341.2
g_4 (molec./cell)	Gal4p	0.1563
g_{80} (molec./cell)	Gal80p	0.1138
m_1 (molec./cell)	<i>GAL1</i> mRNA	0.2647
m_2 (molec./cell)	<i>GAL2</i> mRNA	0.3305
m_3 (molec./cell)	<i>GAL3</i> mRNA	0.9044
m_{80} (molec./cell)	<i>GAL80</i> mRNA	1.1871
g_{4d} (molec./cell)	Gal4p dimer	308.92
g_{80d} (molec./cell)	Gal80p dimer	157.229
g_{4d}^c (molec./cell)	Gal4p dimer - Gal80p dimer complex	0
g_3^c (molec./cell)	Gal3p - gal. complex	0
g_{80}^c (molec./cell)	Gal80p - Gal3p - Gal. complex	0
$[gal]$ (molec./cell)	Internal galactose	0
m_{glu} (molec./cell)	Glucose network mRNA	4000
x_{glu} (molec./cell)	Glucose network proteins	15000
$[glu]$ (molec./cell)	Internal glucose	150003
$[gal]_e$ (molec.)	External galactose outside the cell	$0-2.366 \times 10^8$
$[glu]_e$ (molec.)	External glucose outside the cell	$0-2.957 \times 10^8$

doi:10.1371/journal.pone.0019353.t001

Table 2. Model Parameters.

Parameter	Description	Value [12]
β_1 (min^{-1})	Translation rate of Gal1p g_1	9.92
β_2 (min^{-1})	Translation rate of Gal2p g_2	6.94
β_3 (min^{-1})	Translation rate of Gal3p g_3	18.0
β_4 ((molec./cell)/min)	Max translation rate of Gal4p g_4	0.86
β_{80} (min^{-1})	Translation rate of Gal80p g_{80}	4.00
α_1 ((molec./cell)/min)	Max translation rate of GAL1	1.09
α_2 ((molec./cell)/min)	Max translation rate of GAL2	1.20
α_3 ((molec./cell)/min)	Max translation rate of GAL3	36.0
α_{80} ((molec./cell)/min)	Max translation rate of GAL80	3.00
δ_1 (min^{-1})	Deg. rate of gal. proteins and complexes	0.0033
γ_1 (min^{-1})	Deg. rate of GAL1 mRNA	0.036
γ_2 (min^{-1})	Deg. rate of GAL2 mRNA	0.026
γ_3 (min^{-1})	Deg. rate of GAL3 mRNA	0.036
γ_{80} (min^{-1})	Deg. rate of GAL80 mRNA	0.036
k_3 ((molec./cell)/min)	Binding rate of g_3 to x_{gal}	5.0×10^{-8}
k_{-3} (min^{-1})	Dissociation rate of g_3^c	890
k_{4d} ((molec./cell)/min)	Binding rate of g_4 to g_4	0.1
k_{-4d} (min^{-1})	Dissociation rate of g_{4d}	1.0
k_{80d} ((molec./cell)/min)	Binding rate of g_{80} to g_{80}	0.10
k_{-80d} (min^{-1})	Dissociation rate of g_{80d}	170
k_{80} ((molec./cell)/min)	Binding rate of g_3^c to g_{80}	0.10
k_{-80} (min^{-1})	Dissociation rate of g_{80}^c	0.03
k_r ((molec./cell)/min)	Binding rate of g_{4d} to g_{80d}	0.10
k_{-r} (min^{-1})	Dissociation rate of g_{4d}^c	1.80
k_{cat} (min^{-1})	Galactose metabolism rate	3350
k_{mgk} (molec./cell)	Galactose metabolism constant	1.29×10^7
k_{tr} (min^{-1})	Galactose transport rate	4350
k_{mir} (molec./cell)	Galactose transport constant	2.15×10^8
a_{tr} (unitless)	Galactose interactive constant	10.0
K_p (cell/molec.)	Equilibrium constant of g_{4d} binding to UAS	0.091
K_q (cell/molec.)	Equilibrium constant of g_{80d} binding to g_{4d}	0.0556
C_p (unitless)	Cooperative binding constant of g_{4d} to UAS	1
C_q (unitless)	Cooperative binding constant of g_{80d} to g_{4d} -UAS complexes	30
α_{glu} ((molec./cell)/min)	Basal transcription rate of glucose DNA	215
ϵ_{glu} ((molec./cell)/min)	Activated transcription rate of glucose DNA	6.452×10^5
σ_{glu} (min^{-1})	Translation rate of glucose proteins	0.4
δ_{glu} (min^{-1})	Degradation rate of glucose proteins	0.1
γ_{glu} (min^{-1})	Degradation rate of glucose mRNA	0.0633
δ_d (min^{-1})	Dilution rate of glucose	0.0033
c_{glu} (molec./cell)	Hill constant for glucose induction	1.075×10^7
b (unitless)	Hill coefficient for glucose induction	1.8
k_{tr2} (min^{-1})	Glucose transport rate	4350
k_{mir2} (molec./cell)	Glucose transport constant	6.022×10^8
a_{tr2} (unitless)	Glucose interactive constant	1.0
μ_{glu} (min^{-1})	Glucose transport rate	5350
K_{glu} (molec./cell)	Glucose transport constant	1.29×10^7
ρ (molec./cell)	Hill constant for gal. repression	1.29×10^7
q (unitless)	Hill coefficient for gal. repression	0.8

Table 2. Cont.

Parameter	Description	Value [12]
γ_{sd} (min^{-1})	Glucose induced mRNA degradation rate	9.30×10^{-6}
k_{sd} (molec./cell)	Glucose induced mRNA degradation constant	30

doi:10.1371/journal.pone.0019353.t002

model lowers the complexity to design the associated controller, allowing analytical implementation of a feasible design. We will show later that the reduced model may cause bounded deviation from the original model in transient response; however, it results in insignificant deviation from the steady states.

Model Reduction

Given the steady states of the glucose utilization network, the original model still has 35 parameters. The large number of parameters and significant nonlinearities make it difficult to design an effective controller. As with any biological system, the number of parameters involved in the induction of a specific pathway could be large, with physical, chemical, and biological parameters all affecting the system. In terms of a modeling-based representation of biological systems, the goal of our effort is to determine the parameters that most affect the induction of the system.

A simple example of this concept can be observed in glycolysis. Through the complex intermediates and interactions within this pathway, three regulatory steps are often considered, biologically, the most important rate-limiting steps in the pathway. These irreversible steps, phosphorylation of glucose, phosphorylation of fructose-6-phosphate, and transfer of phosphate to phosphoenolpyruvate and then to ADP, would be the most important control points. The proteins involved in these steps, as well as the mRNA that generates these proteins, are crucial to the overall process. Investigating the influence of the parameters on the control target is an effective tool to reduce the model complexity and assist control system design. In order to reduce the model, we should know whether the output significantly changes under a small change of the parameters, as some parameters may not be as effective as others. By understanding how small changes in parameters affect the overall system, we can determine the most important parameters involved in generating the desired output. In the case of *GAL1* mRNA, a small change in the value of the *GAL4p* parameter will have a large effect on *GAL1* mRNA expression, indicating that this parameter significantly influences the uncertainties of *GAL1* mRNA expression. Thus, this parameter is crucial to control the system. Small changes in a parameter that was not as closely linked to *GAL1* mRNA expression specifically may not have a large effect on the expression. In fact, most biological systems may be sensitive to certain parameters over a large range. If the sensitivity of a parameter is small compared to other parameters, the parameter could be arbitrarily chosen from a wide range of values without altering the dynamic performance of the system. This practice has been used successfully in control systems engineering, including chemical process control, robotics control, aerospace control, etc.

For the above model, we will first conduct local sensitivity analysis that will quantify the change of the system states due to constant perturbation of a parameter at time zero. Consider a general form of nonlinear ordinary differential equations:

$$\frac{dX}{dt} = f(X, \phi, t), \quad X(t_0) = X_0, \quad (23)$$

where X is the state vector with n components, and ϕ is the parameter vector with m components. The initial condition of the above equations is set to x_0 . Define the sensitivity coefficients of the state x_i with respect to parameter $\phi_j, j = 1, \dots, m$, by

$$s(x_i; \phi_j, t) = \frac{\partial x_i}{\partial \phi_j}, \quad (24)$$

and then the state's sensitivity trajectory can be described by the following:

$$\begin{aligned} \frac{ds(X; \phi_j, t)}{dt} &= \frac{d(\partial X / \partial \phi_j)}{dt} = \frac{\partial f}{\partial X} \frac{\partial X}{\partial \phi_j} + \frac{\partial f}{\partial \phi_j} \\ &= \frac{\partial f}{\partial X} s(X; \phi_j, t) + \frac{\partial f}{\partial \phi_j}, \end{aligned} \quad (25)$$

where the state vector is denoted by $X = [x_1 \ \dots \ x_n]^T$. The initial values of the sensitivity equations can be obtained by

$$s(x_i; \phi_j, t_0) = \frac{\partial x_i(t_0)}{\partial \phi_j} = \begin{cases} 0, & x_i \neq \phi_j \\ 1, & x_i = \phi_j \end{cases} \quad (26)$$

In order to avoid scale differences in the parameters, the sensitivity coefficients are normalized as follows,

$$ns(x_i; \phi_j, t) = \frac{\partial x_i / x_i}{\partial \phi_j / \phi_j}. \quad (27)$$

The normalized sensitivity coefficient $ns(x_i; \phi_j, t)$ means that a 1 percent change of state x_i results in a 1 percent change of the parameter ϕ_j over the time course t . The sensitivity trajectory $s(x; \phi_j, t)$ can be calculated numerically by equations (25–26). The importance of the parameters can thus be distinguished by the magnitude of the normalized sensitivity coefficients in the time-course of the system following the constant perturbation at time zero.

The normalized sensitivity coefficients for the original model are shown in **Figure 2**, where each small block describes a normalized sensitivity for a certain parameter at a certain time. In **Figure 2A**, the galactose input is kept at a constant,

$[gal]_e = 4 \times 10^7$ (molec.) with the measured system output represented by m_1 . It is easy to observe that the external galactose input is positively correlated to the output of m_1 . The glucose network remains in the steady state when the internal glucose is emptied and there is no external glucose supply. The steady state of $[glu] = 13586$ (molec./cell) and $m_{glu} = 3397$ (molec./cell) can be determined by solving $\dot{m}_{glu} = \dot{g}_{glu} = [glu] = 0$. The function $R([glu])$ in equation (19) was found to equal 0.9959. Thus, the glucose network can be separated from the galactose network. This is in agreement with experimental data, where a lack of glucose as a carbon source and the presence of galactose will induce the *GAL* network. Based on the sensitivity analysis of m_1 with respect to all the parameters (Table 2), it becomes clear which parameters can be omitted from the system. The heat map in Figure 2 shows the normalized sensitivity of *GAL1* mRNA concentration (m_1) after changes in all 35 parameters. The Figure 2A shows the sensitivities when $[gal]_e = 4 \times 10^7$ molec. The Figure 2B represents the sensitivities when $[gal]_e = 4 \times 10^2$ molec. The sensitivities of the parameters δ_1 , a_{tr} , v_{sd} , γ_1 , β_4 , k_{4d} , k_{-4d} , k_{-r} , k_r , K_p , C_p and k_{sd} are much smaller than those of the other parameters (near zero, green color). In general, when the external galactose $[gal]_e$ is higher, the parameters are more sensitive, with the exception of α_1 and γ_1 .

However, it would be careless to omit these parameters without further investigation. For example, the degradation rate δ_1 is important for achieving the steady states of the original model, so it must be retained. K_p and C_p come from the cooperative fractional saturation function $F(g_{4d}, g_{80d}, N)$. Both g_{4d} and g_{80d} in the function contribute to the activation of mRNAs m_1 , m_2 , m_3 , and m_{80} . However, the contributions may not be equal. If we define $P = C_p K_p g_{4d}$; $Q = C_Q K_q g_{80d}$, then we have

$$F(g_{4d}, g_{80d}, 1) = \frac{P}{1 + P + PQ} = \frac{1}{1/P + 1 + Q} \quad (28)$$

When $1 + Q \gg 1/P$, the above equation (28) can be approximated by

$$F(g_{4d}, g_{80d}, 1) \approx \frac{1}{1 + Q} = 1 - \left(\frac{Q}{1 + Q} \right) \quad (29)$$

From (29), we observed that the output m_1 is positively correlated to the external galactose input. The sensitivity analysis has told us that K_p and C_p can be set to be large enough to satisfy $1 + Q \gg 1/P$. Thus, only g_{80d} contributes to the cooperative

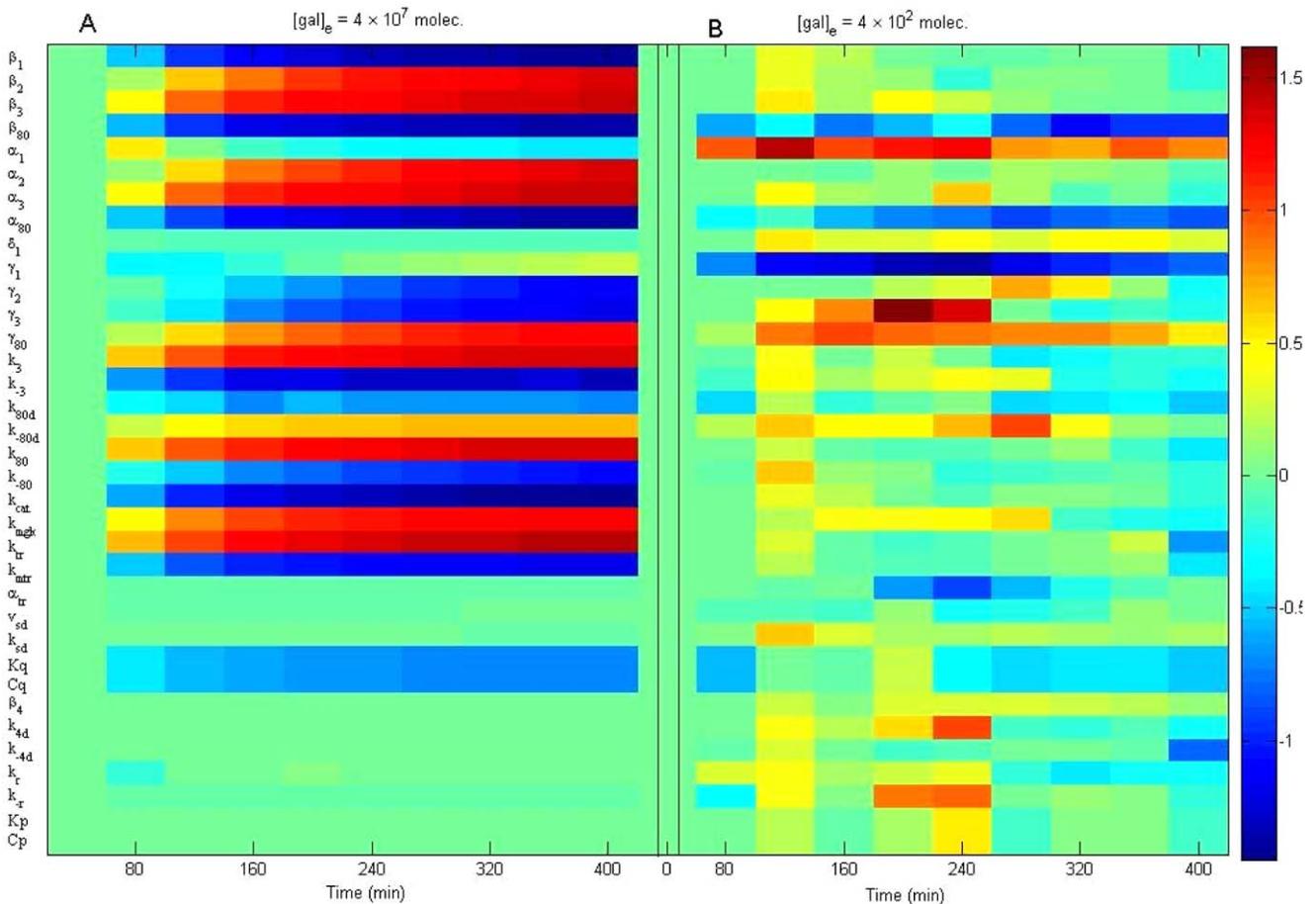


Figure 2. Normalized parameter sensitivity trajectory with respect to time. The heat map shows the normalized sensitivity of *GAL1* mRNA concentration (m_1) towards change in all 35 parameters. Panel (A): Sensitivities when $[gal]_e = 4 \times 10^7$ molec. Panel (B): Sensitivities when $[gal]_e = 4 \times 10^2$ molec. The sensitivities of the parameters δ_1 , a_{tr} , v_{sd} , γ_1 , β_4 , k_{4d} , k_{-4d} , k_{-r} , k_r , K_p , C_p and k_{sd} are much smaller than those of the other parameters (approximately zero, green color). In general, when the external galactose $[gal]_e$ is higher, the parameters are more sensitive, besides α_1 and γ_1 .

doi:10.1371/journal.pone.0019353.g002

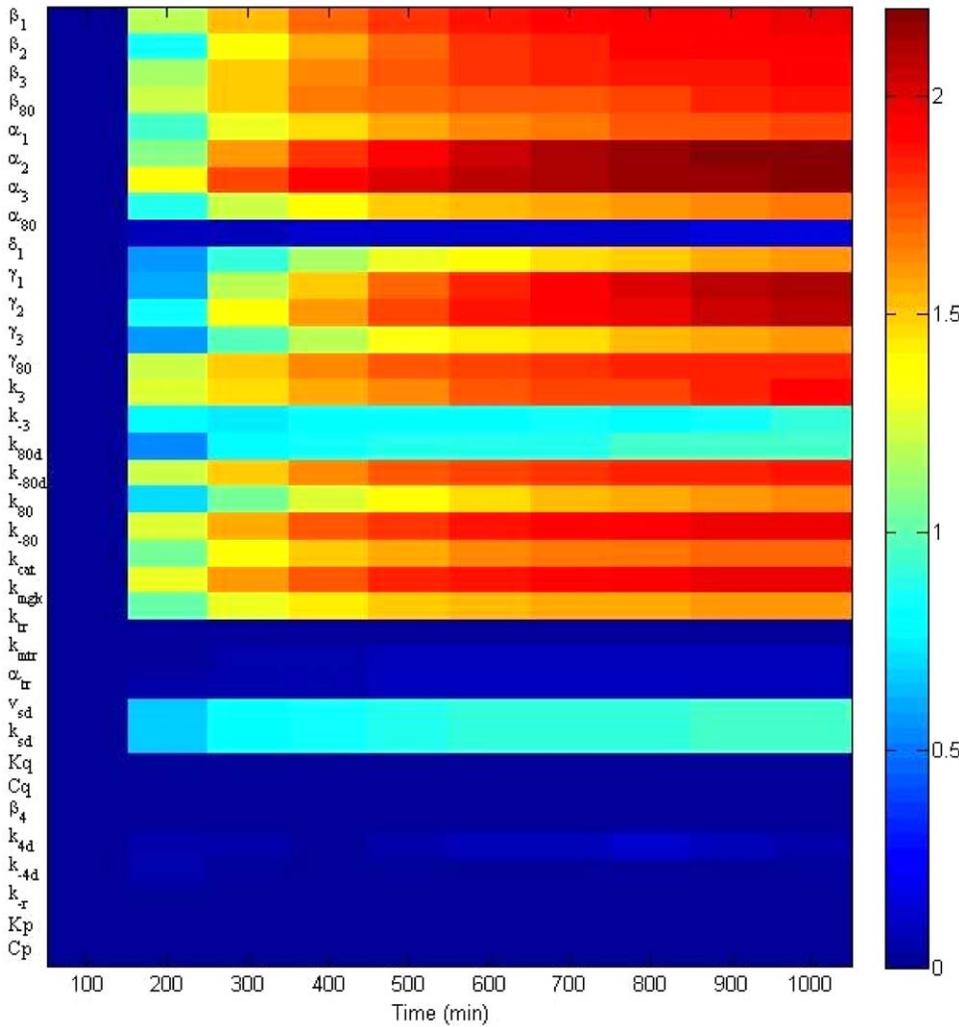


Figure 3. Local sensitivity analysis at the steady states of four GAL mRNAs with respect to all model parameters. The heat map denotes the average absolute value of the normalized sensitivity coefficients of four GAL mRNA across 1000 minutes. The parameters $\delta_1, a_{tr}, v_{sd}, k_{sd}, \beta_4, k_{4d}, k_{-4d}, k_{-r}, k_r, K_p,$ and C_p are smaller than 0.25, while the others are larger than 1. This implies the reaction terms associated with insensitive parameters can be omitted no matter which mRNA is chosen as the control target. One can conclude not only that the control model targeting GAL1 mRNA need not include the terms associated with the above parameters, but also that control strategies targeting all four GAL mRNAs can suppress the same terms. The reduced model can thus be used for the control design targeting any of the four GAL mRNAs. doi:10.1371/journal.pone.0019353.g003

fractional saturation function. Similarly, the other two cooperative functions can be simplified as follows,

$$F(g_{4d}, g_{80d}, 2) = \frac{2P + P^2 + 2P^2Q}{1 + 2P + P^2 + 2P^2Q + 2PQ + P^2Q^2} \approx 1 - \left(\frac{Q}{1+Q}\right)^2, \text{ if } 1+Q \gg \frac{1}{P}, \tag{30}$$

and

$$F(g_{4d}, g_{80d}, 4) \approx 1 - \left(\frac{Q}{1+Q}\right)^4, \text{ if } 1+Q \gg \frac{1}{P}. \tag{31}$$

Therefore, the cooperative fractional saturation effect of g_{4d} and g_{80d} can be approximated by the contribution from g_{80d} alone. This can also be explained by the sensitivity analysis in **Figure 1**,

3, and **4**, which shows that k_{4d}, k_{-4d}, k_{-r} as well as k_r are less “essential” to the system output. Thus, we can set $k_{4d}, k_{-4d}, k_r,$ and k_{-r} to be zero, since they are small numbers and have small sensitivity coefficients. This results in a removal of the loops involving the GAL4p dimer, g_{4d} , and the GAL4p/GAL80p dimer complex, g_{4d}^c . As shown in **Figure 1**, the subsystem (g_4, g_{4d}, g_{4d}^c and g_{80d}) can be simplified as g_{80d} only, because the rest of the subsystems will not affect the steady state of the output. The goal of the proposed control is to maintain a level of mRNA expression at some predetermined level. The model reduction will thus have little influence on the control performance in terms of output deviation after a long enough time. The contributions from the glucose network to m_1 , and m_3 , are small when no glucose exists, which is necessary to induce the GAL gene network, as explained earlier. This is consistent with the small sensitivity coefficients with respect to v_{sd} and k_{sd} . We thus set v_{sd} equal to zero and k_{sd} equal to a large number. Also, small sensitivity of the coefficient with respect to a_{tr} leads to a slight model reduction on $T_1([gal], [gal]_e)$ by setting $a_{tr} = 0$.

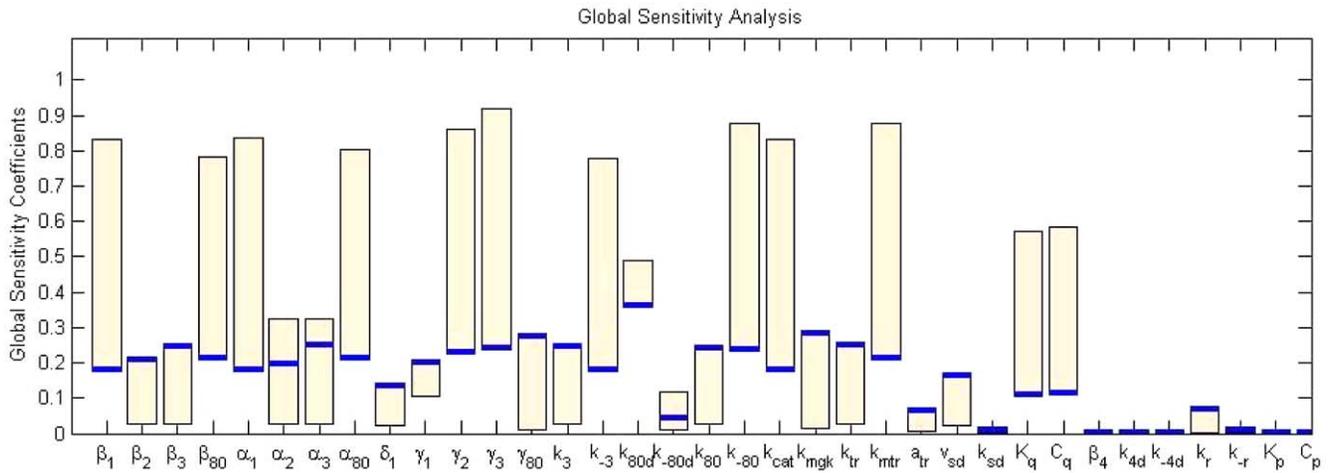


Figure 4. Global sensitivity analysis at the steady states of four *GAL* mRNAs with respect to all model parameters under constant control $[gal]_e = 4 \times 10^7$ molec. The bars denote the range of the sensitivity coefficients of four *GAL* mRNAs with respect to each parameter. The blue solid line stands for the global sensitivity coefficients for *GAL1* mRNA. doi:10.1371/journal.pone.0019353.g004

Thus, a Reduced Order Model can be achieved based on the analysis above. The Reduced Order Model in a state space form can thus be described as follows:

$$\dot{g}_1 = \beta_1 m_1 - \delta_1 g_1 \tag{32}$$

$$\dot{g}_2 = \beta_2 m_2 - \delta_1 g_2 \tag{33}$$

$$\dot{g}_3 = -\delta_1 g_3 - k_3 x_{gal} g_3 + k_{-3} g_3^c + \beta_3 m_3 \tag{34}$$

$$\dot{g}_{80} = \beta_{80} m_{80} - \delta_1 g_{80} - k_{80} g_3^c g_{80} + k_{-80} g_{80}^c + 2k_{-80d} g_{80d} - 2k_{80d} g_{80}^2 \tag{35}$$

$$\dot{m}_1 = \alpha_1 F(g_{80d}, 4) - \gamma_1 m_1 \tag{36}$$

$$\dot{m}_2 = \alpha_2 F(g_{80d}, 2) - \gamma_2 m_2 \tag{37}$$

$$\dot{m}_3 = \alpha_3 F(g_{80d}, 1) - \gamma_3 m_3 \tag{38}$$

$$\dot{m}_{80} = \alpha_{80} F(g_{80d}, 1) - \gamma_{80} m_{80} \tag{39}$$

$$\dot{g}_{80d} = k_{80d} g_{80}^2 - k_{-80d} g_{80d} - \delta_1 g_{80d} \tag{40}$$

$$\dot{g}_3^c = k_3 [gal] g_3 - k_{-3} g_3^c - k_{80} g_3^c g_{80} + k_{-80} g_{80}^c - \delta_1 g_3^c \tag{41}$$

$$\dot{g}_{80}^c = k_{80} g_3^c g_{80} - k_{-80} g_{80}^c - \delta_1 g_{80}^c \tag{42}$$

$$[gal] = k_{tr} g_2 \frac{[gal]_e - [gal]}{k_{mtr} + [gal]_e + [gal]} - \frac{k_{cat} g_1 [gal]}{k_{mgk} + [gal]} - k_3 g_3 [gal] + k_{-3} g_3^c - \delta_1 [gal] \tag{43}$$

where

$$F(g_{80d}, N) = 1 - \left(\frac{C_Q K_Q g_{80d}}{1 + C_Q K_Q g_{80d}} \right)^N \tag{44}$$

All parameters are summarized in **Table 2**.

The comparison of the outputs of the Original Model and the Reduced Model is shown in **Figure 5**. The differences between the two models occur in the transition process before the first 200 min, while the difference decreases significantly in the steady state. According to the experimental results from [11,12], the galactose network takes 4–7 hours (240–420 minutes) to reach a steady state. Our simulation results in **Figure 3** are consistent with the experimental data. When the external galactose level is set at $[gal]_e = 4 \times 10^7$ molec. or $[gal]_e = 4 \times 10^2$ molec., the steady-state deviation of *GAL1* mRNA, m_1 (molec./cell) is similar between the full and reduced models. However, the deviation of the transition process, the time before achieving the steady state, between the two models with the smaller external galactose level is significant. This result is consistent with the sensitivity measure in **Figure 3**.

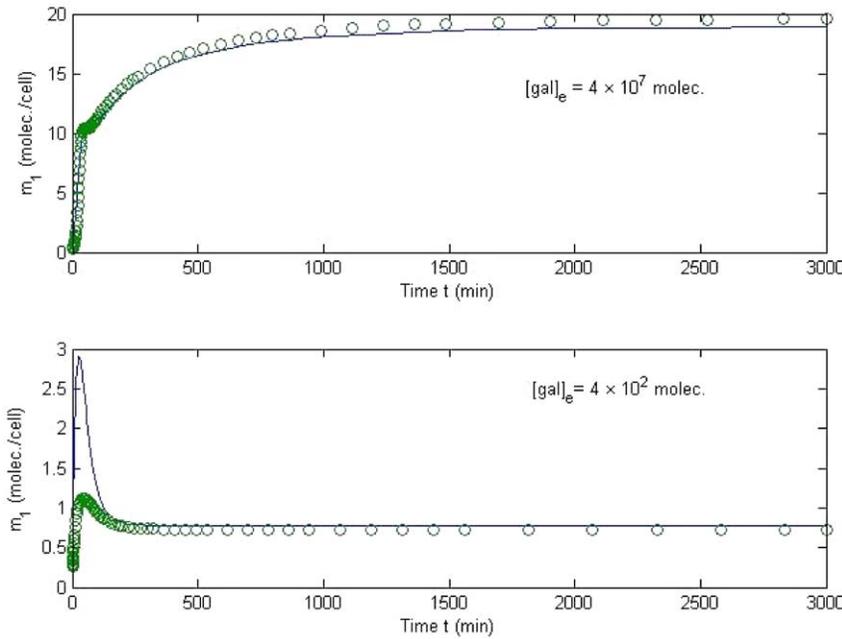


Figure 5. Comparison of original model and reduced model. The difference between the two models occurs in the transition process before the first 200 min, while the difference decreases significantly in the steady state. When $[gal]_e = 4 \times 10^7$ molec. and $[gal]_e = 4 \times 10^2$ molec. are applied to both models, the steady-state deviation of the output m_1 (molec./cell) is less than 1 percent. However, the deviation of the transition process between the two models with the smaller external galactose is significant. This result is consistent with the sensitivity measure in Figure 3. The solid line is for m_1 concentration in the original model, while the open symbol is for the reduced model.
doi:10.1371/journal.pone.0019353.g005

Biologically, the reduced model varies from the complex model only within the first 200 minutes of expression for relatively low galactose levels, 400 molecules. The steady-state levels of *GAL1* mRNA expression after the first 200 min for both galactose levels tested were the same in both models (Figure 5). In this way, the reduced model can be used to explain the expression of *GAL1* mRNA, once the expression level has reached a steady state. Before the steady state is reached, only a complex model can completely model the effects of *GAL1* mRNA expression from the introduction of low levels of galactose. A similar technique can be applied if the control target changes to any of the other mRNAs or proteins. This paper exemplifies *GAL1* mRNA as a primary control target, but the applications of the method are not limited to a certain mRNA. Choosing *GAL2* or *GAL3*, as well as *GAL80*, will produce very similar reduced models; however, using *GAL4* as a control target will obviously yield a different reduced model because of the simplification in the *GAL1* output system obtained by the removal of the *GAL4p* subsystem. A reduced model for *GAL4* would have to include the *GAL4p* subsystem, and thus a much different reduced model would be generated.

Figure 3 describes the local sensitivity analysis at the steady states of four *GAL* mRNAs with respect to all model parameters. The bars denote the range of the sensitivity coefficients of four *GAL* mRNAs with respect to each parameter. The parameters δ_1 , a_{tr} , v_{sd} , k_{sd} , β_4 , k_{4d} , k_{-4d} , k_{-r} , k_r , K_p , and C_p are smaller than 0.25, while the others are greater than one. This implies that the reaction terms associated with the insensitive parameters can be omitted no matter which mRNA is chosen as the control target. The design of a control model to control steady-state expression of *GAL1*, *GAL2*, *GAL3*, and *GAL80* mRNA (m_1 , m_2 , m_3 and m_{80}) is described in the following section.

The local sensitivity analysis describes the change of system dynamics at time t after a step-wise perturbation to a parameter

occurs at time zero. One might wonder what the effect would be when all parameters are perturbed simultaneously. Global sensitivity analysis investigates the effect of simultaneous large variations of all parameters on the states. The common methods to test global sensitivity include the Fourier amplitude sensitivity test (FAST) [20], extended FAST [21], Sobol's method [22], the partial rank correlation coefficient (PRCC) [23], as well as the weighted average local sensitivity method [24].

FAST is one of the classical methods to determine global sensitivity. Assuming that the time dependence of each uncertain parameter ϕ_i is associated with a frequency ω_j , we can describe that parameter by a transformation function of a sinusoid, i.e.,

$$\phi_j(s) = G_j[\sin(\omega_j s)], \quad j = 1, 2, \dots, m \quad (45)$$

for all m parameters, where s is the time.

The first order global sensitivity of x_i with respect to the variation of the parameter ϕ_i can be defined as the ratio of the variance of ϕ_i to the total variance

$$GS_i(\phi_j) = \sigma_{i,\omega_j}^2 / \sigma_i^2, \quad (46)$$

where the total variance is

$$\sigma_i^2 = 2 \sum_{k=1}^{\infty} [A_i^2(k) + B_i^2(k)] \quad (47)$$

and the variance of ϕ_i becomes

$$\sigma_{i,\omega_j}^2 = 2 \sum_{k=1}^{\infty} [A_i^2(k\omega_j) + B_i^2(k\omega_j)], \quad (48)$$

with

$$A_i(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x_i(\phi, s) \cos(\omega s) ds \quad (49)$$

$$B_i(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x_i(\phi, s) \sin(\omega s) ds. \quad (50)$$

Although the above local sensitivity analysis pointed out some candidate parameters and their associated terms that may be removed, it only guarantees the scenario when the parameters are tuned one by one. When the parameters are simultaneously adjusted, some may become sensitive. Thus, we must also apply global sensitivity analysis to rule out corresponding parameters.

As shown in Figure 4, the bars denote the range of the sensitivity coefficients of four *GAL* mRNAs with respect to each parameter under constant control $[gal]_e = 4 \times 10^7$ molec. The blue solid line stands for the global sensitivity coefficients for *GAL1* mRNA. In particular, a_{tr} , k_{sd} , β_4 , k_{4d} , k_{-4d} , k_{-r} , K_p , and C_p are nearly zero. The maximum of the coefficients δ_1 , v_{sd} and k_{-80d} is smaller than 0.2, while that of the others is larger than 0.2, consistent with the local sensitivity coefficients in Figure 3. This implies that the reaction terms associated with insensitive parameters can be omitted no matter which mRNA is chosen as the control target. Thus, both local and global sensitivity analyses indicate that the reduced model, equations (32–43), can be used for regulating all four *GAL* mRNAs.

Control Design

As discussed earlier, any of the four *GAL* mRNAs can be chosen as the control target. For demonstration purposes, we have chosen to control *GAL1* mRNA (m_1), *i.e.*, essentially maintain the *GAL1* mRNA at a desired level m_1^* , *i.e.*,

$$m_1 = m_1^* \text{ and } \dot{m}_1 = \ddot{m}_1 = \dots = 0. \quad (51)$$

In order to keep equation (51) true, the steady state of the system becomes

$$g_{80d}^* = \frac{1/C_q K_q}{1 - [1 - \gamma_1 m_1^*/\alpha_1]^{1/4}} - \frac{1}{C_q K_q}. \quad (52)$$

As a test case for this model, we will demonstrate the control of *GAL1* mRNA expression by controlling the external level of galactose. It is important to note that any gene or protein

represented in the reduced model can be controlled in a similar manner. Following this same pattern, recombinant genes regulated by *GAL1* could be controlled in the same manner.

Similarly, we can find the steady states for m_2, m_3, m_{80} , and g_1 ,

$$m_2^* = \alpha_2 F(g_{80d}^*, 2) / \gamma_2 \quad (53)$$

$$m_3^* = \alpha_3 F(g_{80d}^*, 1) / \gamma_3, \quad (54)$$

$$m_{80}^* = \alpha_{80} F(g_{80d}^*, 1) / \gamma_{80} \quad (55)$$

$$g_1^* = \beta m_1^* / \delta_1 \quad (56)$$

$$g_2^* = \beta m_2^* / \delta_1 \quad (57)$$

$$g_{80}^* = \sqrt{(k_{-80d} + \delta_1) g_{80d}^* / k_{80d}} \quad (58)$$

$$(g_3^c)^* = (k_{-80} + \delta_1) (g_{80}^c)^* / (k_{80} g_{80}^*). \quad (59)$$

Observing the galactose network, we can obtain two auxiliary equations:

$$\frac{d(g_3 + (g_3^c)^* + (g_{80}^c)^*)}{dt} = \beta_3 m_3 - \delta_1 (g_3 + (g_3^c)^* + (g_{80}^c)^*) \quad (60)$$

$$\frac{d(g_{80} + 2g_{80d} + (g_{80}^c)^*)}{dt} = \beta_{80} m_{80} - \delta_1 (g_{80} + 2g_{80d} + (g_{80}^c)^*). \quad (61)$$

Equation (60) indicates that the total mass of $(g_3, (g_3^c)^*, (g_{80}^c)^*)$, is up-regulated by m_3 . Biologically, this is a simple concept to understand, as expression of *GAL3* mRNA, m_3 , is necessary for the generation of *GAL3p*, g_3 , and the Gal3p dimer, $(g_3^c)^*$, in a dose-dependent manner. Similarly, the *GAL80p/GAL3p* dimer complex is dependent on the expression of *GAL3p*. Equation (61) indicates

that the total mass of $(g_{80}, g_{80d}, (g_{80}^c)^*)$, is up-regulated by m_{80} , in a manner similar to that described above. Both of these equations show the interplay between the biological mechanisms responsible for their creation.

Through equations (60) and (61) we can determine the remaining steady states g_3^* , $(g_{80}^c)^*$ and $[gal]^*$ using the auxiliary equations

$$g_3^* + (g_3^c)^* + (g_{80}^c)^* = \beta_3 m_3^* / \delta_1 \tag{62}$$

$$g_{80}^* + 2g_{80d}^* + (g_{80}^c)^* = \beta_{80} m_{80}^* / \delta_1. \tag{63}$$

Thus,

$$(g_{80}^c)^* = \beta_{80} m_{80}^* / \delta_1 - m_{80}^* - 2g_{80d}^* \tag{64}$$

$$g_3^* = \beta_3 m_3^* / \delta_1 - (g_3^c)^* - (g_{80}^c)^* \tag{65}$$

$$[gal]^* = [(k_{-3} + \delta_1)(g_3^c)^* + \delta_1 (g_{80}^c)^*] / (k_3 g_3^*). \tag{66}$$

Given our desire to enforce the control condition of $m_1 = m_1^*$, we need to control the external galactose, $[gal]_e^*$, which at steady state can be calculated as follows,

$$[gal]_e^* = \frac{k_{mir} \Omega + [gal]^* (1 + \Omega)}{1 - \Omega}, \tag{67}$$

where $\Omega = \left(\frac{k_{cat} g_1^* [gal]^*}{k_{mgk} + [gal]^*} + k_3 g_3^* [gal]^* - k_{-3} (g_3^c)^* + \delta_1 [gal]^* \right) / (k_{ir} g_2^*)$.

Similarly, if the control objective is to regulate *GAL2* mRNA, $m_2 = m_2^*$, then

$$g_{80d}^* = \frac{1/C_q K_q}{1 - \sqrt{1 - \gamma_2 / \alpha_2 m_2^*}} - \frac{1}{C_q K_q} \tag{68}$$

$$m_1^* = \alpha_3 F(g_{80d}^*, 4) / \gamma_1. \tag{69}$$

The rest of the terms can be achieved using the same equations (54)–(67). Because this system has only a single steady

state, one can always find a nominal control value for external galactose $[gal]_e^*$ to maintain any control objective at a desired level. In other words, we can always compute the necessary amount of galactose offline in order to maintain *GAL1* mRNA or any other factor, such as *GAL2* mRNA or *GAL3p*, at a level of interest.

Thus, offline, we can calculate an open-loop constant control, which is the simplest type of controller that does not take into consideration state information that is fed back into the system for control purposes.

$$\text{Open-loop Constant Control: } [gal]_e = [gal]_e^* \tag{70}$$

for any of the targeting mRNA, m_1 , m_2 , m_3 , and m_{80} , if these mRNAs are not measurable in real-time. However, if they are measurable, one can improve the control algorithm by introducing a closed-loop feedback control. A simple proportional output feedback control for m_1 can be described as

$$\text{Closed-loop Feedback Control: } [gal]_e = [gal]_e^* + k(m_1^* - m_1) \tag{71}$$

where the feedback gain $k = 10^8$. The gain k directly affects the convergence rate of the controlled GAL network. However, the external galactose amount should not exceed the maximum concentration of galactose. Thus, the control dosage should be bounded by

$$[gal]_e = \min([gal]_e, [gal]_e^{\max}) \tag{72}$$

Of course, one can use a proportional-integral-derivative (PID) control in practice. Integral control can adjust the static error, but too much integral gain can reduce the stabilization of the system. Derivative control can adjust the convergence rates, however, it is too sensitive to high frequency noise. For pursuing different requirements, many more complicated control methods can be applied, such as optimal control for minimizing the galactose dosage and robust control for reducing the uncertainty, among others. Because the control objective of regulating the mRNA concentration to a desired value, is a relatively simple task, without strict requirements for reaction time and galactose dosage, a simple proportional control is adequate to accomplish the control goal.

Results

Control Stability Analysis

Next, we will prove the stability of this equilibrium because the equilibrium is a unique steady state. If it is stable, then the state will tend to this steady state given the nominal control. The control objective m_1 will be maintained at the desired level m_1^* . Biologically, this system is stable based on experimental results [12]. The following is a theoretical proof based on the mathematical model.

The Jacobian matrix at this equilibrium is

$$J = \begin{bmatrix} -\delta_1 & 0 & 0 & 0 & \beta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\delta_1 & 0 & 0 & 0 & \beta_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\delta_1 - k_3[gal]^* & 0 & 0 & 0 & \beta_3 & 0 & 0 & k_{-3} & 0 & -k_3g_3^* \\ 0 & 0 & 0 & -\delta_1 - 4k_{80d}g_{80}^* - k_{80}(g_3^*)^* & 0 & 0 & 0 & \beta_{80} & 2k_{-80d} & -k_{80}g_{80}^* & k_{-80} & 0 \\ 0 & 0 & 0 & 0 & -\gamma_1 & 0 & 0 & 0 & F_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\gamma_2 & 0 & 0 & F_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_3 & 0 & F_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_{80} & F_{80} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2k_{80d}g_{80}^* & 0 & 0 & 0 & 0 & -\delta_1 - k_{-80d} & 0 & 0 & 0 \\ 0 & 0 & k_3[gal]^* & -k_{80}(g_3^*)^* & 0 & 0 & 0 & 0 & 0 & -\delta_1 - k_{-3} - k_{80}g_{80}^* & k_{-80} & k_3g_3^* \\ 0 & 0 & 0 & k_{80}(g_3^*)^* & 0 & 0 & 0 & 0 & 0 & k_{80}g_{80}^* & -\delta_1 - k_{-80} & 0 \\ G_1 & G_2 & -k_3[gal]^* & 0 & 0 & 0 & 0 & 0 & 0 & k_{-3} & 0 & -\delta_1 - G_3 \end{bmatrix} \quad (73)$$

where

$$F_1 = -4\alpha_1(K_q C_q)^4 (g_{80d}^*)^3 / (1 + K_q C_q g_{80d}^*)^5, \quad (74)$$

$$F_2 = -2\alpha_2(K_q C_q)^2 g_{80d}^* / (1 + K_q C_q g_{80d}^*)^3, \quad (75)$$

$$F_3 = -\alpha_3 K_q C_q / (1 + K_q C_q g_{80d}^*)^2, \quad (76)$$

$$F_{80} = -\alpha_{80} K_q C_q / (1 + K_q C_q g_{80d}^*)^2, \quad (77)$$

$$G_1 = \frac{-k_{cat}[gal]^*}{k_{mgk} + [gal]^*}, \quad (78)$$

$$G_2 = k_{tr} \frac{[gal]_e^* - [gal]^*}{k_{mtr} + [gal]_e^* + [gal]^*}, \text{ and} \quad (79)$$

$$G_3 = k_3 g_3^* + \frac{k_{cat} k_{mgk} g_1^*}{(k_{mgk} + [gal]^*)^2} + \frac{(2[gal]_e^* + k_{mtr}) g_2^* k_{tr}}{(k_{mtr} + [gal]_e^* + [gal]^*)^2}. \quad (80)$$

If $m_1^* = 20$, one finds the eigenvalues of the Jacobian matrix are $[-895.99, -194.68, -12.54, -12.04, -0.0295, -0.0387 + 0.0024i, -0.0387 - 0.0024i, -0.0360, -0.0015, -0.0033, -0.0033, -0.0033]$. A standard stability analysis of the dynamic systems finds that the equilibrium is asymptotically stable. Thus the system will tend to the steady state asymptotically.

Because most of the eigenvalues are small in magnitude, it implies that the convergence of this controlled system will take a

long time. Biologically, after we set the external galactose level to a constant value, the yeast galactose utilization network will slowly achieve balance (a steady state) and the *GAL1* mRNA level will gradually reach and maintain the desired level. At this steady-state level, there will not be oscillations in the mRNA or protein levels associated with this system.

In order to change the eigenvalues to negative, we follow a common practice in control systems engineering and introduce a linear output feedback control

$$[gal]_e = [gal]_e^* + k(m_1^* - m_1), \quad (81)$$

where $k = 10^8$. $J(12,5) = -\frac{(2[gal]_e^* + k_{mtr})g_2^*k_{tr}k}{(k_{mtr} + [gal]_e^* + [gal]^*)^2}$. The eigenvalues of the Jacobian matrix indicate how fast the control can reach a steady state. In general, the more negative, the faster. In the above control, the eigenvalues of the Jacobian matrix (73) move far away from the y axis: $[-895.99, -194.67, -16.40, -5.33, -2.91, -0.026, -0.0361, -0.0360, -0.0033, -0.0033, -0.0033]$. This implies that the system will converge faster under the linear feedback control. Increasing the value of k will lead to faster convergence. However, the growth of the convergence rate is limited due to the maximum concentration of galactose. From **Figure 1**, we can find that the internalized galactose level increases with rising external galactose levels, and the internalized galactose level enhances *GAL 4*, which in turn promotes *GAL1* mRNA expression. Increased *GAL1* mRNA expression will also increase the degradation rate of the internalized galactose, thus the internalized galactose will reach a steady state faster. In essence, the introduction of the feedback control loop actually increases the external galactose level, but dynamically adjusts the amount such that the *GAL1* mRNA can still reach the desired level. In this process, the degradation rate of the internalized galactose is largely enhanced, consistent with a significant increase of the magnitude of several negative eigenvalues.

We set the desired level of *GAL1* mRNA as $m_1 = 20$. From equation (67), the nominal control values for $[gal]_e^*$ can be obtained as 4.0552×10^7 molec./cell. Application of the nominal control should cause the system to slowly reach a stable equilibrium. As shown in **Figure 6**, it takes more than 1500 min for m_1 to reach the desired level. From **Figure 5**, we can see that achieving the desired high *GAL1* mRNA level usually demands a much longer time than is required for low *GAL1* mRNA levels. The metabolic reactions involved in promoting mRNA expression and subsequent translation into proteins require a lengthy time. Multiple positive and negative feedback loops inside the network actually prevent mRNAs and the associated proteins from rapid changes due to external stimuli.

The effects of protein concentration profiles by feedback control on *GAL1* mRNA expression on the *GAL* network are illustrated in **Figure 7**. All states (unit: molec./cell) converge to a stable steady state, while g_2 , g_3^c and g_{80}^c have slower convergence rates than the other states. It is not surprising to see that all concentrations tend to a steady state, since the system has a unique stable steady state. But the transient response of the galactose system is equally important, because a large fluctuation is not desired in a reliable bioreactor. The figure shows that the system does not have large overshoots and is thus safe for implementation in practice.

Figure 8 compares the set point regulation performance of the feedback control between the original model and the reduced model using feedback control. Setting three desired levels of m_1 as 10, 15, and 20, feedback control based on both the original model (solid) and the reduced model (open) can be maintained at the desired levels within a similar time frame. The reduced model-based feedback control, however, leads to slightly larger magnitudes of oscillations at the transient response. The difference is small enough not to result in any severe fluctuation in the biological system. Thus, the control design based on the reduced model can be applied to the original model directly.

Sensitivity Analysis

The steady-state sensitivity describes the change of the steady state given a small perturbation of a certain parameter [25]. It can be given by the first-order partial derivative of a state with respect to the parameter. Using the finite difference method, the normalized steady-state sensitivity (NSSS) coefficient for the steady state of state x_i with respect to the parameter ϕ_j can be approximated by

$$SS_{ij} = \frac{\partial Z_{ss_i}(\phi_j) / Z_{ss_i}(\phi_j)}{\partial \phi_j / \phi_j} \approx \frac{Z_{ss_i}(\phi_j + \Delta \phi_j) - Z_{ss_i}(\phi_j)}{\Delta \phi_j} \frac{\phi_j}{Z_{ss_i}(\phi_j)}, \quad (82)$$

where Z_{ss_i} is the steady state of the state x_i .

A small magnitude of NSSS of state x_i with respect to the parameter ϕ_j implies that the parameter uncertainty of ϕ_j has no significant influence on the state x_i locally. In contrast, a large NSSS implies serious influence by the parameter uncertainty. It is necessary to emphasize “locally” because the steady-state sensitivity is a local sensitivity measurement that perturbs one parameter and fixes the other parameters. As shown in **Figure 9**, we calculated the NSSS of all states with respect to the output m_1 for both constant control and the feedback control algorithm. The NSSS of the feedback control case has significantly smaller magnitude in terms of average maximum sensitivities for all states (0.93 v.s. 3.42) and output state (0.02 v.s. 1.66). Especially for the output state, the magnitude of NSSS was reduced 100-fold. This implies that the parameter uncertainty does not affect the steady state of the output. The feedback control law (81) can always drive the output to the reference level subject to uncertainties and perturbation in the parameters locally.

Biologically, we can conclude from the above data that the feedback control is less influenced by disturbances within the system, as shown in **Figure 9**. For mRNA expression, this means that a feedback control prevents the system from being overly sensitive to any specific parameter. The perturbations or parameter uncertainty in any parameter will not cause a significant deviation to the steady mRNA level. This is necessary for complex biological systems, where parameter uncertainty and perturbations widely exist. If the system was overly sensitive over some parameters, then it would be difficult to maintain the mRNA at a desired level. A series of checks and balances are always present to prevent the system from fluctuating out of control.

Discussion

This paper employs control theory and a systems engineering approach to demonstrate the regulation *in silico* of the budding

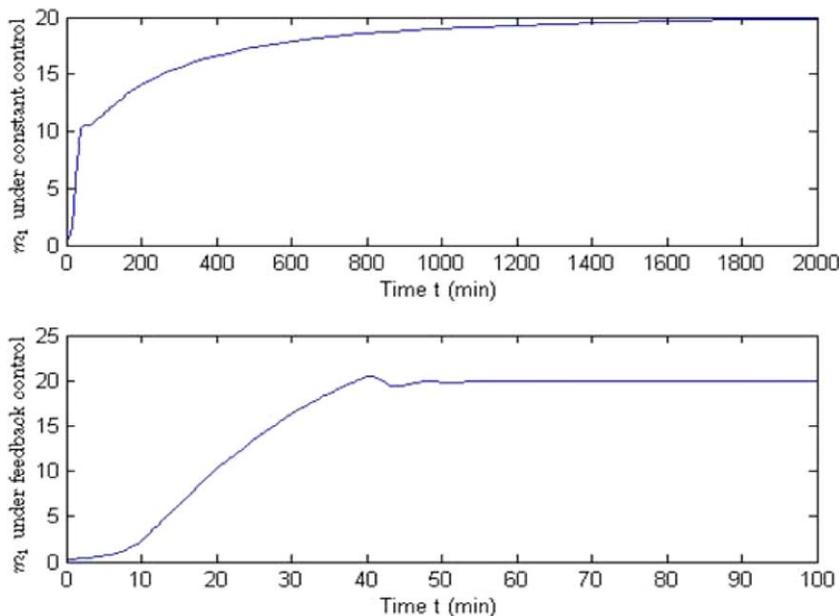


Figure 6. Comparison between constant and feedback control for the Reduced Order Model. It takes more than 1500 min for *GAL1* mRNA, m_1 to reach the desired level of 20 (molec./cell) using the constant control (top); however, only 40 min are required for the feedback control (bottom). According to the simulation, feedback control is much more effective at bringing the yeast system into a steady-state level of mRNA expression.

doi:10.1371/journal.pone.0019353.g006

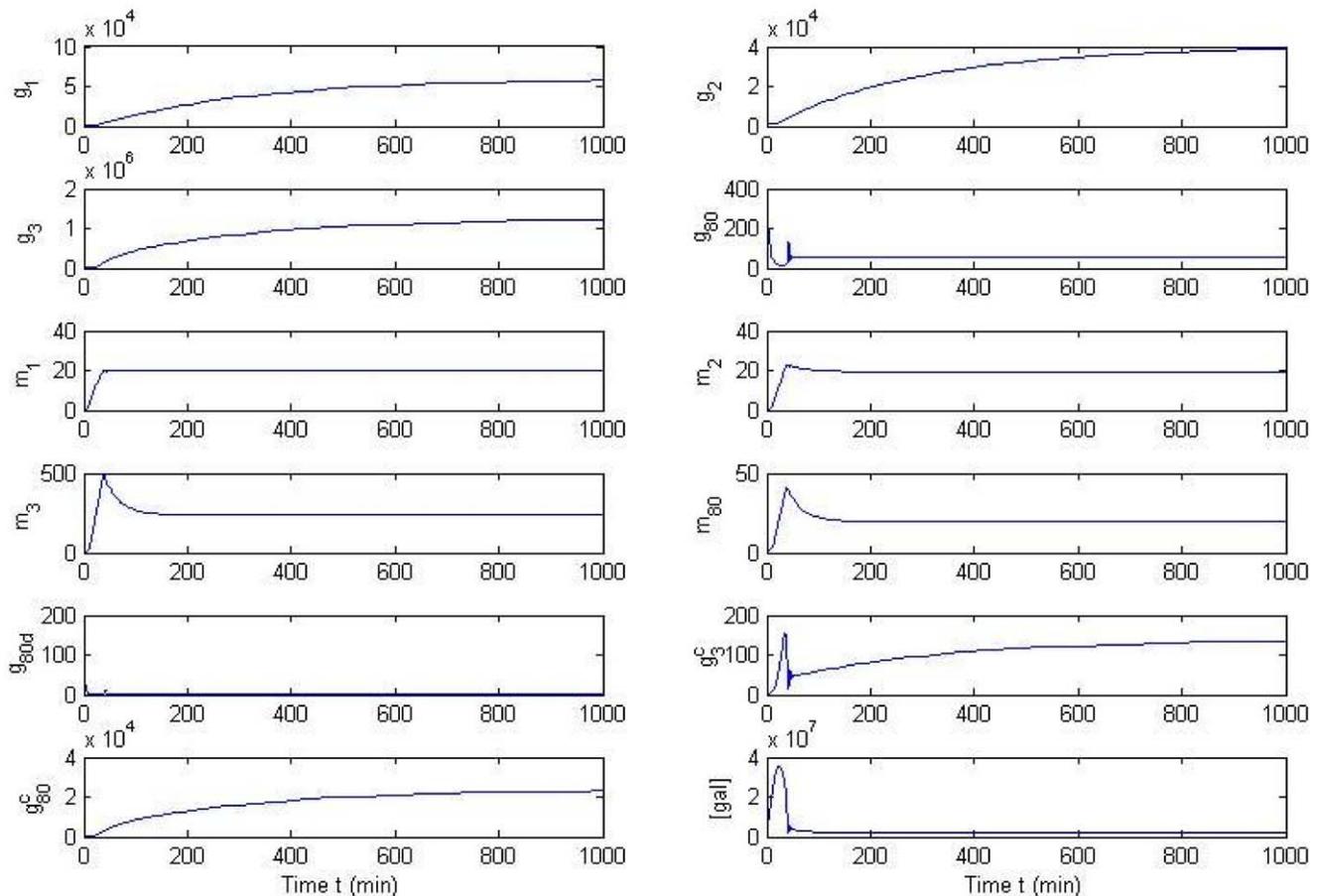


Figure 7. Effect of protein concentration on mRNA expression by feedback control on the GAL network. All states (unit: molec./cell) converge to a stable steady state, while g_2 , g_3^c and g_{80}^c have slower convergence rates than the other states. The transient process of states g_{80} and g_3^c is listed in the small windows above. The overshoot of the states is kept in a reasonable range. No severe fluctuations or oscillations are found in the GAL network under feedback control.
doi:10.1371/journal.pone.0019353.g007

yeast, *Saccharomyces cerevisiae*, into a desired gene expression and/or metabolic activity pattern. The GAL network is used as an example to demonstrate the effectiveness of a systems approach for complex cellular system control. We demonstrate that although the structure and computational models are complicated, it is possible to use a “simple” control algorithm to manipulate the system. According to the local parameter sensitivity analysis, some terms can be eliminated from the model without adversely affecting its long-term behavior, and the control design based on the reduced model can still be applied to the original model. Both constant control and feedback control laws have been designed for the galactose network. To evaluate the control performance, we conducted local steady-state sensitivity and global sensitivity analysis. The results accordingly imply that feedback control significantly suppresses the parameter uncertainties. This sheds light on the potential to provide novel control components to eukaryotic transcriptional regulation networks, as well as other gene regulatory networks.

The approach we propose can not only be used to control the yeast GAL network, but the fundamental principles can be applied to a wide range of biological network control. Complex networks involving cell-cycle controls, DNA repair, and other genes of interest can be controlled in a similar manner, given an effective

means of monitoring the output of the system. The most attractive application is the control of bioreactors. If done properly, the ability to yield a large amount of the target in a relatively short period of time, while minimizing the effects of toxic levels of gene expression products can be achieved. If these systems are tightly controlled at the cellular level, proteins of interest, such as immunogenic proteins used in the production of vaccines, restriction enzymes, biopharmaceuticals, biochemicals can be tightly regulated. This would allow for a cost-effective means of producing a large variety of products for research and commercial use. Applications to other systems include bacterial expression systems and even mammalian cell systems, based on the effectiveness of the bioreactor and the access to the requisite control variables. Of course, other physical and chemical parameters would be necessary to completely control a bioreactor-based system, including pH, temperature, and nutrient conditions. The major difference between these and the above approach is that the proposed approach focuses on cellular system control, which is still a widely open field. Our future research will implement this control system with a fully established microfluidic device to monitor the real-time expression of GAL-GFP fusion proteins. By validating the above control in an experimental environment we will demonstrate the usefulness and validity of our approach.

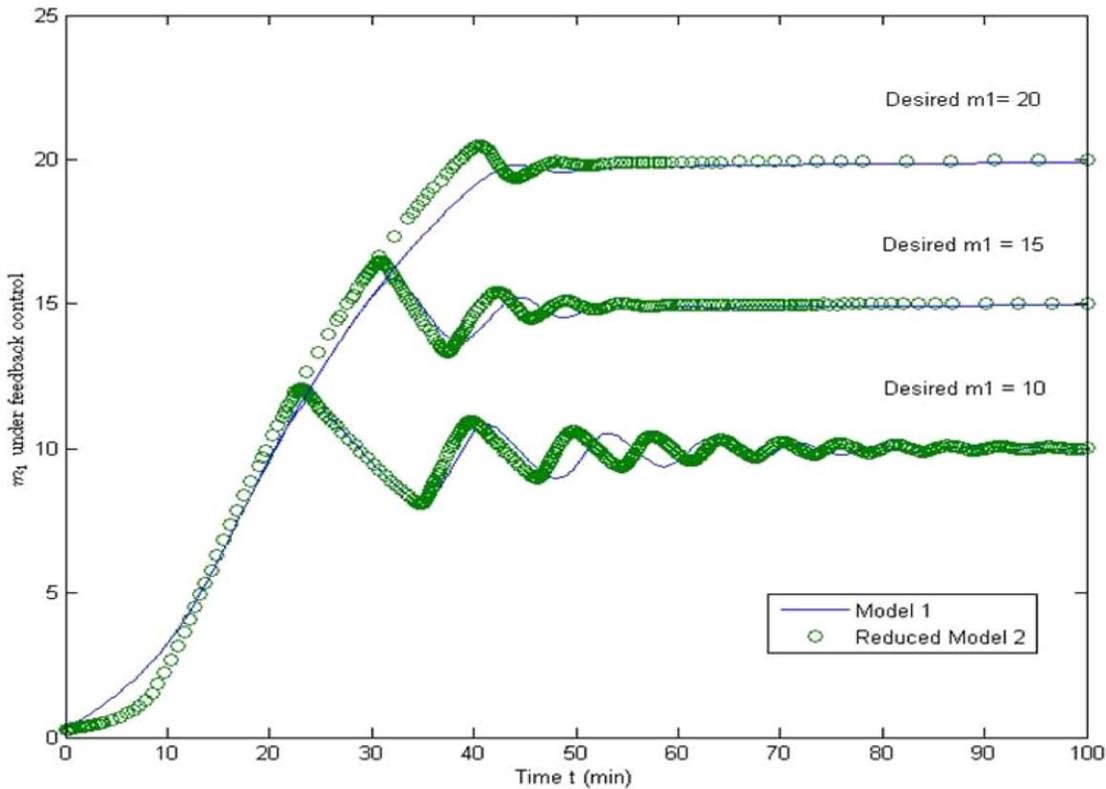


Figure 8. Comparison of referenced tracking between the original model and the reduced model using feedback control. Under the feedback control, the tracking performance of the original model and the reduced model is similar. Thus, the control design based on the reduced model can be applied to the original model directly.
doi:10.1371/journal.pone.0019353.g008

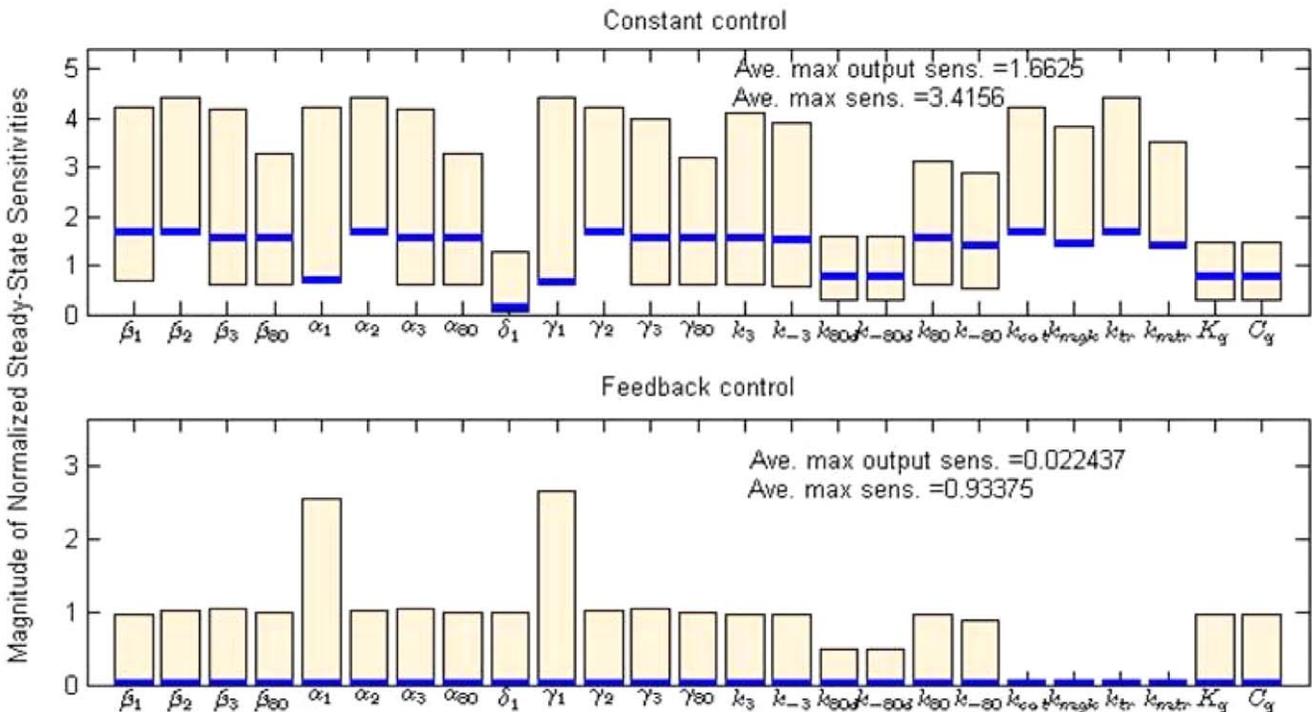


Figure 9. Comparison of normalized steady-state sensitivities (NSSS) between constant and feedback control. The NSSS of the feedback control case has significantly smaller magnitude in terms of average maximal sensitivities for all states (0.93 v.s. 3.42) and the output state (0.02 v.s. 1.66). The bars describe the maximum and minimum values of NSSS of the states with respect to each parameter. The solid lines denote the output NSSS on each parameter.
doi:10.1371/journal.pone.0019353.g009

Acknowledgments

We are indebted to Allison Price for her editorial assistance.

References

- Bhalla US, Iyengar R (1999) Emergent properties of networks of biological signaling pathways. *Science* 283: 381–387.
- Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403: 339–342.
- Cantone I, Marucci L, Iorio F, Ricci MA, Belcastro V, et al. (2009) A yeast synthetic network for *in vivo* assessment of reverse-engineering and modeling approaches. *Cell* 137: 172–181.
- Menolascina F, di Bernardo M, di Bernardo D (2010) Design and implementation of a feedback control strategy for IRMA, a novel synthetic gene regulatory network. Atlanta, GA, USA: Proceedings of the 49th IEEE Conference on Decision and Control.
- Menolascina F, di Bernardo M, di Bernardo D Analysis, design and implementation of a novel scheme for in-vivo control of synthetic gene regulatory networks. *Automatica*, (In press).
- Horak J, Wolf D (1997) Catabolite inactivation of the galactose transporter in the yeast *Saccharomyces cerevisiae*: ubiquitination, endocytosis, and degradation in the vacuole. *J Bacteriol* 179: 1541–1549.
- Yano K, Fukasawa T (1997) Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 94: 1721–1726.
- Diep CQ, Tao X, Pilauri V, Losiewicz M, Blank TE, et al. (2008) Genetic evidence for sites of interaction between the Gal3 and Gal80 proteins of the *Saccharomyces cerevisiae* GAL gene switch. *Genetics* 178: 725–736.
- Johnston M, Flick JS, Pexon T (1994) Multiple mechanisms provide rapid and stringent glucose repression of GAL gene expression in *Saccharomyces cerevisiae*. *Mol Cell Biol* 14: 3834–3841.
- de Atauri P, Orrell D, Ramsey S, Bolouri H (2004) Evolution of 'design' principles in biochemical networks. *Systems Biology, IEE Proceedings* 1: 28–40.
- Ramsey SA, Smith JJ, Orrell D, Marelli M, Petersen TW, et al. (2006) Dual feedback loops in the GAL regulon suppress cellular heterogeneity in yeast. *Nat Genet* 38: 1082–1087.
- Bennett MR, Pang WL, Ostroff NA, Baumgartner BL, Nayak S, et al. (2008) Metabolic gene regulation in a dynamically changing environment. *Nature*.
- Bornholdt S (2005) SYSTEMS BIOLOGY: Less is more in modeling large genetic networks. *Science* 310: 449–451.
- de Atauri P, Orrell D, Ramsey S, Bolouri H (2004) Evolution of 'design' principles in biochemical networks. *IET Sys Bio* 1: 28–40.
- Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, et al. (2004) Integrative Analysis of Cell Cycle Control in Budding Yeast. *Mol Biol Cell* 15: 3841–3862.
- Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, et al. (2002) A monomeric red fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* 99: 7877–7882.
- Trumbly RJ (1992) Glucose repression in the yeast *Saccharomyces cerevisiae*. *Molecular Microbiology* 6: 15–21.
- Wikswa JP, Prokop A, Baudenbacher F, Cliffl D, Csukas B, et al. (2006) Engineering challenges of bioNEMS: the integration of microfluidics, micro- and nanodevices, models and external control for systems biology. *IEE Proceedings Nanobiotechnology*, 153: 81–101.
- McLean JA, Schultz JA, Woods AS (2010) Ion Mobility-Mass Spectrometry for biological and Structural Mass Spectrometry. In: Cole RB, ed. *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicabilities, and Biological Applications*: John Wiley and Sons. pp 411–439.
- Saltelli A, Tarantola S, Chan KPS (1999) A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* 41: 39–56.
- Saltelli A, Tarantola S, Campolongo F (2000) Sensitivity analysis as an ingredient of modeling. *Statistical Science* 15: 377–395.
- Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comp Simul* 55: 271–280.
- Draper N, Smith H (1981) *Applied regression analysis*. New York: Wiley.
- Bentele M, Lavrik I, Ulrich M, Stosser S, Heermann DW, et al. (2004) Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* 166: 839–851.
- Varma A, Morbidelli M, Wu H (1999) *Parameteric sensitivity in chemical systems*. Cambridge, UK: Cambridge University Press.

Author Contributions

Conceived and designed the experiments: RY SL JW MJZ. Performed the experiments: RY SL JW MJZ. Analyzed the data: RY SL JW MJZ. Contributed reagents/materials/analysis tools: RY SL JW MJZ. Wrote the paper: RY SL JW MJZ.