

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

**Title**

Single cell growth and gene expression dynamics in model organisms

**Permalink**

<https://escholarship.org/uc/item/0cn311qj>

**Author**

Cookson, Natalie Anne

**Publication Date**

2009

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Single Cell Growth and Gene Expression Dynamics in Model Organisms**

A dissertation submitted in partial satisfaction of the requirements for the degree  
Doctor of Philosophy

in

Bioengineering

by

Natalie Anne Cookson

Committee in charge:

Jeff Hasty, Chair  
Michael Holst  
Tracy Johnson  
Shankar Subramaniam  
Lanping Amy Sung  
Lev Tsimring

2009

Copyright  
Natalie Anne Cookson, 2009  
All rights reserved.

The dissertation of Natalie Anne Cookson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

---

---

Chair

University of California, San Diego

2009



## DEDICATION

*Thanks to Scott for keeping me happy,  
the unconditional love of my family and friends for keeping me grounded,  
& the antics of Cougar and Kiwi for keeping me sane.  
In loving memory of Wiley.*

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Table of Contents . . . . .	v
List of Abbreviations . . . . .	viii
List of Figures . . . . .	x
List of Tables . . . . .	xii
Acknowledgements . . . . .	xiii
Vita . . . . .	xv
Abstract of the Dissertation . . . . .	xvii
Chapter 1 Introduction . . . . .	1
1.1 Synthetic Biology Overview . . . . .	1
1.2 Synthetic Biology in Practice . . . . .	3
1.2.1 Molecular Biology . . . . .	3
1.2.2 Microfluidic Technology . . . . .	5
1.2.3 Computational Modeling . . . . .	6
1.2.4 Successes in Synthetic Biology . . . . .	8
Chapter 2 Monitoring Dynamics of Single-Cell Gene Expression over Multiple Cell Cycles . . . . .	10
2.1 Introduction . . . . .	10
2.2 Design and Fabrication . . . . .	11
2.3 Experimental Results . . . . .	13
2.4 Materials and Methods . . . . .	17
2.4.1 Fabrication Procedure . . . . .	17
2.4.2 Device Characterization . . . . .	18
2.4.3 Strain and Cell Culture . . . . .	25
2.4.4 Loading Procedure . . . . .	27
2.4.5 Microscopy . . . . .	27
2.4.6 Image Processing . . . . .	28
2.4.7 Quantitative Controls and Error Compensation . . . . .	30
2.5 Acknowledgements . . . . .	33

Chapter 3	Origins of extrinsic variability in eukaryotic gene expression . . . . .	35
3.1	Introduction . . . . .	35
3.2	Results and Discussion . . . . .	36
3.3	Computational Model . . . . .	43
3.4	Materials and Methods . . . . .	46
3.4.1	Plasmids, Yeast Strains and Growth Conditions . . . . .	46
3.4.2	Gene Expression Experiments . . . . .	48
3.4.3	Variability of gated subpopulations . . . . .	48
3.4.4	Modeling studies . . . . .	48
3.4.5	GAL4p regulation of other promoters . . . . .	50
3.4.6	Fusion Strains . . . . .	51
3.4.7	Gene Expression Experiments . . . . .	52
3.5	Acknowledgements . . . . .	52
Chapter 4	Metabolic gene regulation in a dynamically changing environment . .	53
4.1	Introduction . . . . .	53
4.2	Results and Discussion . . . . .	54
4.3	Materials and Methods . . . . .	62
4.3.1	Dynamic environment experiments . . . . .	62
4.3.2	Microfluidic chips and waveform generation . . . . .	63
4.3.3	mRNA degradation experiments . . . . .	63
4.4	Acknowledgements . . . . .	63
Chapter 5	Cell cycle dependent variations in protein concentration . . . . .	65
5.1	Introduction . . . . .	65
5.2	Results and Discussion . . . . .	67
5.3	Materials and Methods . . . . .	75
5.3.1	Strain and cell culture . . . . .	75
5.3.2	Data acquisition . . . . .	76
5.3.3	Flow cytometry data acquisition . . . . .	76
5.3.4	Forward scatter size calibration . . . . .	76
5.3.5	Fluorescence Integration . . . . .	78
5.3.6	Computational model . . . . .	79
5.3.7	Volume Growth Characteristics . . . . .	86
5.4	Acknowledgements . . . . .	88
Chapter 6	Queuing up for enzymatic processing: Correlated signaling through coupled degradation . . . . .	89
6.1	Introduction . . . . .	89
6.2	Results and Discussion . . . . .	90
6.3	Materials and Methods . . . . .	97
6.3.1	Plasmid Construction . . . . .	97
6.3.2	Flow Cytometry . . . . .	98

6.3.3 Microscopy . . . . .	98
6.4 Acknowledgements . . . . .	99
Chapter 7 Summary . . . . .	100
References . . . . .	104

## LIST OF ABBREVIATIONS

AU	arbitrary units
bp	base pair
CFP	cyan fluorescent protein
div.	division
DNA	deoxyribonucleic acid
<i>E. coli</i>	<i>Escherichia coli</i>
eqn.	equation
fig.	figure
FL	fluorescence
FP	fluorescent protein
FSC	forward scatter
GFP	green fluorescent protein
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
LB	lysogeny broth
mRNA	messenger ribonucleic acid
MW	molecular weight
no.	number
OD	optical density
ODE	ordinary differential equation
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDMS	poly(dimethylsiloxane)
PIV	particle image velocimetry
RBITC	Rhodamine B isothiocyanate
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SD	synthetic drop-out
T $\mu$ C	Tesla microchemostat

UV .....	ultraviolet
yeGFP .....	yeast-enhanced cyan fluorescent protein
yeGFP .....	yeast-enhanced green fluorescent protein
yeVFP .....	yeast-enhanced Venus fluorescent protein
YFP .....	yellow fluorescent protein
YG .....	yellow-green

## LIST OF FIGURES

Figure 1.1: Three essential tools for engineering gene circuits. . . . .	2
Figure 2.1: The T $\mu$ C design was optimized to allow for long-term growth of cells in a monolayer. . . . .	12
Figure 2.2: Analysis of long-term expression data acquired using the T $\mu$ C. . . . .	15
Figure 2.3: Single-cell yeVFP dynamics for 119 cells. . . . .	16
Figure 2.4: Flow vector field produced by implementing a PIV method to process images of fluorescent beads used as flow tracers. . . . .	19
Figure 2.5: Schematics used in simulating nutrient delivery in the T $\mu$ C. . . . .	21
Figure 2.6: Comparison of experimental data and analytical solution for nutrient transport into the trapping region using 10 kDa Dextran-conjugated RBITC. . . . .	24
Figure 2.7: Analytical analysis of the time evolution of 1D concentration profiles for different advective velocities. . . . .	24
Figure 2.8: Comparison of experimental data and analytical solution for nutrient transport into the trapping region using Sulforhodamine 101. . . . .	25
Figure 2.9: Map of the pRS41-yv vector used to integrate three copies of the <i>gal1</i> promoter driving expression of <i>yeVFP</i> into the yeast genome. . . . .	26
Figure 2.10: Major steps of image segmentation. . . . .	29
Figure 2.11: Quantitative controls performed to ensure accuracy of the fluorescence data. . . . .	31
Figure 3.1: Experimental results for GFP expression for different copy numbers and galactose concentrations. . . . .	37
Figure 3.2: Comparison of Models A & B with experimental findings. . . . .	41
Figure 3.3: Illustrations of cell growth and galactose regulatory network. . . . .	44
Figure 3.4: Cell size distributions and coefficients of variation for GFP expression measured within narrow gates. . . . .	49

Figure 3.5: Coefficient of variation as a function of galactose for three promoters regulated by GAL4p. . . . .	50
Figure 4.1: Design and implementation of the microfluidic platform developed for our study. . . . .	55
Figure 4.2: Regulation in the galactose utilization network. . . . .	58
Figure 4.3: Experimental and computational results for cells expressing a <i>GALI-yECFP</i> fusion gene in response to alternating glucose and galactose media for strains K699 and YPH499. . . . .	60
Figure 4.4: Experimental and computational comparison of two yeast strains, one of which (YPH499) is known to have a deficiency in the galactose utilization network. . . . .	61
Figure 5.1: Single cell volume trajectories for a population of <i>S. cerevisiae</i> . . . . .	69
Figure 5.2: Volume growth and division statistics. . . . .	71
Figure 5.3: Comparison of experimental volume trajectories to model simulations and flow cytometry data. . . . .	73
Figure 5.4: Size data for flow calibration beads. . . . .	77
Figure 5.5: Correlation between the fluorescence profile of a cell and the corresponding elliptical shape. . . . .	79
Figure 5.6: Growth rate calculated by measuring the OD <sub>600</sub> over time. . . . .	88
Figure 6.1: Rate-limited processing can couple the queue lengths of different job types. . . . .	92
Figure 6.2: Steady-state behavior of a synthetic signaling network. . . . .	95
Figure 6.3: Dynamic behavior of a synthetic signaling network. . . . .	96



## LIST OF TABLES

Table 3.1: Summary of integration plasmids. . . . .	47
Table 3.2: Summary of yeast strains. . . . .	47
Table 3.3: Summary of yeast genes fused to <i>yEGFP</i> . . . . .	51
Table 5.1: Statistics of growth characteristics as they change over the generations.	87

## ACKNOWLEDGEMENTS

I cannot fully express my gratitude to Dr. Jeff Hasty, for warmly welcoming me into his lab and for always taking the time to provide guidance, advice, and friendship over the years. The scientific knowledge that I have acquired from him is priceless, and if I have absorbed a fraction of his instinct and intuition in research, I will consider myself lucky. I am equally lucky to have a second advisor, Dr. Lev Tsimring, one of the most talented scientists and photographers I will ever know. I also warmly acknowledge my fellow researchers in the Systems Biodynamics lab, whose help, advice, and collaboration made my time as a graduate student productive, rewarding, and fun. In particular, I am indebted to Jennifer Marciniak, for introducing me to bench work, Lee Pang, for his microfluidics expertise, Scott Cookson, for his dedication to perfection, William Mather, Matt Bennett and Dmitri Volfson for their computational genius and patience, Jesse Stricker and Bridget Baumgartner for their microbial expertise, and Michael Ferry, for always knowing the answers to my questions. Finally, I thank the United States Department of Energy (DOE) Computational Science Graduate Fellowship for their funding and support of my graduate work.

Chapter 2 contains material originally published as Cookson, S.\*, Ostroff, N.\*, Pang, W. L.\*, Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032-E1-6. (\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

Chapter 3 contains material originally published as Volfson, D.\*, Marciniak, J.\*, Ostroff, N., Blake, W., Tsimring, L., and Hasty, J., 2006: Origins of extrinsic variability in eukaryotic gene expression. *Nature*, 439: 861-864. (\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

Chapter 4 contains material originally published as Bennett, M.\*, Pang, W.\*, Ostroff, N., Baumgartner, B., Nayak, S., Tsimring, L., and Hasty, J., 2008: Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119-22.

(\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

Chapter 5 contains material originally published as Cookson, N., Cookson, S., Tsimring, L., and Hasty, J., 2009: Cell cycle dependent variations in protein concentration. *Nucl. Acids Res.*, in press. Copyright permission to republish here was granted by Oxford University Press.

Chapter 6 contains material originally submitted for publication as Cookson, N.\*, Mather, W.\*, Mondragon-Palomino, O., Williams, R., Tsimring, L. and Hasty, J., 2009: Queuing up for enzymatic processing: Correlated signaling through coupled degradation. *Science*, submitted. (\*equal contribution). Copyright permission to republish here was granted by AAAS.

## VITA

2002	Bachelor of Arts in Physics University of Pennsylvania
2007	Master of Science in Bioengineering University of California, San Diego
2009	Doctor of Philosophy in Bioengineering University of California, San Diego

## PUBLICATIONS

### Peer Reviewed Journal Articles

Cookson, S.\*, **Ostroff, N.\***, Pang, W. L.\*, Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032-E1-6. (\*equal contribution).

Volfson, D.\*, Marciniak, J.\*, **Ostroff, N.**, Blake, W., Tsimring, L., and Hasty, J., 2006: Origins of extrinsic variability in eukaryotic gene expression. *Nature*, 439: 861-864.

Bennett, M.\*, Pang, W.\*, **Ostroff, N.**, Baumgartner, B., Nayak, S., Tsimring, L., and Hasty, J., 2008: Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119-22. (\*equal contribution).

**Cookson, N.**, Tsimring, L., and Hasty, J., 2009: The pedestrian watchmaker: Genetic clocks from engineered oscillators. *FEBS Lett.*, in press.

**Cookson, N.**, Cookson, S., Tsimring, L., and Hasty, J., 2009: Cell cycle dependent variations in protein concentration. *Nucl. Acids Res.*, in press.

**Cookson, N.\***, Mather, W.\*, Mondragon-Palomino, O., Williams, R., Tsimring, L. and Hasty, J., 2009: Queuing up for enzymatic processing: Correlated signaling through coupled degradation. *Science*, submitted. (\*equal contribution)

### Book Chapter

**Ostroff, N.**, Ferry, M., Cookson, S., Johnson, T., and Hasty, J., 2008: Synthetic Biology: Bioengineering at the Genomic Level. *An Introductory Text to Bioengineering*, World Scientific Publishing Co., 427-451.

## FIELDS OF STUDY

Major Field: Bioengineering (Synthetic Biology)

Studies in Biological Dynamics

Professor Jeff Hasty and Dr. Lev S. Tsimring

ABSTRACT OF THE DISSERTATION

**Single Cell Growth and Gene Expression Dynamics in Model Organisms**

by

Natalie Anne Cookson

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2009

Jeff Hasty, Chair

The developing discipline of synthetic biology attempts to recreate in artificial systems the emergent properties found in natural biology. Progress in this field requires a thorough understanding of the basic cellular functions that underly complex biological networks. Here, we present several studies that use existing and novel methods to probe the dynamic behavior of the model organisms *Saccharomyces cerevisiae* and *Escherichia coli* at the single cell level. First, we develop a microfluidic chemostat for monitoring single-cell gene expression within large populations of *S. cerevisiae* over many cellular generations. Second, we investigate the sources of extrinsic variability in eukaryotic gene expression using a combination of computational modeling and fluorescence data generated from multiple promoter-gene inserts in *S. cerevisiae*. Third, we use an enhanced version of the microfluidic chemostat to subject a large population of *S. cerevisiae* to a periodically varying carbon source, uncovering a novel regulatory property of a well-characterized metabolic network. Fourth, we use fluorescence mi-

croscopy to acquire long-term volume trajectories for a large population of *S. cerevisiae* cells and reveal cell cycle dependent variations in protein concentration. Finally, we design and construct a synthetic signaling network in *E. coli* to investigate the coupling effect of “waiting lines” for enzymatic processing and discover correlated signaling through coupled protein degradation. Together, these studies illustrate the need for new approaches to studying fundamental cellular processes, in order to ultimately advance the goals of synthetic biology.

# Chapter 1

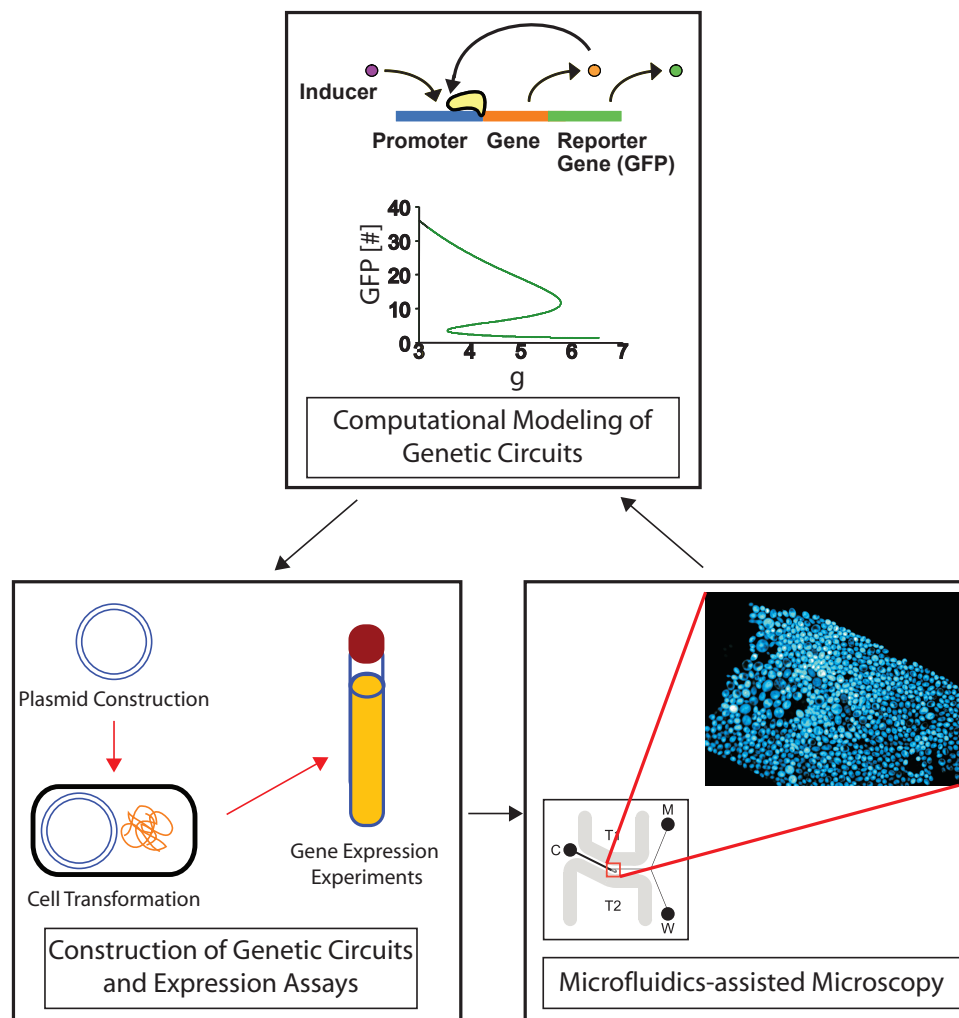
## Introduction

### 1.1 Synthetic Biology Overview

Systems biology has grown rapidly in the wake of the human genome project, as it has become clear that an integration of experimental and computational research will be required to quantitatively describe complex biological systems. The utilization of high-throughput technologies has led to the successful reconstruction of gene regulatory networks in many organisms (Tavazoie et al., 1999; Ibarra et al., 2002; Ideker et al., 2001b; Gardner et al., 2003), along with the development of quantitative models for many complex and fundamental cellular processes (Breedon, 2003; Simon et al., 2001; Vogelstein et al., 2000; Kohn and Pommier, 2005; Begley and Samson, 2004; Bartek et al., 2004). To complement the progress of genome-scale measurement technologies, the developing discipline of synthetic biology aims to design and build novel biological systems that reproduce the behavior of natural systems and contribute to our understanding of how complex biological functions arise from the connectivity of gene regulatory networks (Hasty et al., 2002; Basu et al., 2005; Mangan et al., 2003; Rosenfeld et al., 2005a; Pedraza and van Oudenaarden, 2005a).

Within the broad field of synthetic biology, our research focuses on the design, construction, and modeling of engineered gene circuits. Because the genetic networks





**Figure 1.1:** Three essential tools for engineering gene circuits. Molecular biology is used to construct synthetic DNA constructs and integrate them into an organism of interest. Single-cell measurement technologies, such as microfluidics-assisted microscopy, provide time-lapse quantitative data describing the dynamic behavior of a synthetic genetic circuit. Computational modeling provides a framework for understanding and predicting cellular function.

found in cells are often highly integrated and quite complex, redesigning simpler synthetic systems for study is a valuable approach to understanding gene regulation. The emergence and growing prominence of synthetic biology is largely due to vast improvements in molecular biology techniques, advances in the microfabrication of microscopy devices, and the increased application of mathematical modeling to biological systems over the past few decades (Fig. 1.1). Therefore, it is important to understand the significant contribution that each of these aspects makes to the overall goal of synthetic biology. First, the tools of molecular biology allow scientists to easily manipulate and clone DNA, enabling the rapid design and construction of simple genetic circuits. Second, microfabricated microfluidic devices, tailored to the acquisition of long-term data from large populations of individual cells, provide scientists with a means to monitor the dynamic behavior of a novel biological circuit. And finally, the application of engineering-based modeling techniques to the study of biological systems provides insight into the design of new circuits, as well-formulated models can predict the behavior of a circuit before it is even constructed. The growing excitement and recent successes in synthetic biology illustrate the great impact that the union of biology, engineering, and technology can have on our understanding of biological systems.

## **1.2 Synthetic Biology in Practice**

### **1.2.1 Molecular Biology**

The increasing ease with which scientists are able to modify DNA and genetically engineer organisms has led to a recent burst of interest in synthetic biology. Over the past few decades, several breakthroughs in molecular biology techniques have led to a more detailed understanding of basic biological functions and a greater ability to mimic these functions for the purpose of modifying natural biological systems. Recombinant DNA technology allows for the artificial rearrangement and cloning of DNA, and many tasks that were previously tedious or even impossible have become efficient and

cost-effective over the past few decades. There are several techniques that are essential to the field of synthetic biology. First is the ability to cleave DNA, using enzymes called restriction nucleases that cut DNA at a specific target sequence. Second is DNA cloning, which produces billions of identical copies of a DNA molecule of interest. And third is rapid and reliable sequencing a purified DNA fragment. These days, a scientist can isolate a specific gene, make unlimited copies of it, and determine its exact sequence all in a day's work. Not only can we copy a gene, but we can also alter it and transfer it back into a cell to be incorporated into the organism's genome.

The ease with which we can genetically engineer organisms has already served many purposes, such as aiding in the discovery of the function of new genes, advancing the understanding of evolutionary links between organisms, and enabling the mass production of proteins for use as hormones or vaccines. In addition, now that we have access to the genomic sequences of many organisms, we can begin to unravel complex regulatory networks, develop a strong understanding of native biological systems, and even piece network components back together to create new functionality. Ultimately, we can envision the ability to construct simple gene modules that could be integrated into a diseased cell to perform the function of a damaged or missing cellular component. With the technology for analyzing proteins, DNA and RNA advancing rapidly, we have the increasing potential to obtain a complete understanding of how a cell's genetic network regulates its behavior and its response to the extracellular environment.

The tools of recombinant DNA technology, including the techniques to "cut and paste," clone, and rapidly sequence DNA, have become indispensable to bioscientists interested in probing natural systems. They have also made possible the birth of a new era of synthetic biology, in which we will use these techniques to genetically engineer novel genetic circuits as we strive not only to understand native systems but also to design and build novel biological functions and systems. The output, or "read-out," of synthetic gene networks is typically a fluorescent protein, such as GFP. There are various techniques for measuring the fluorescence output of a cell, such as flow cytometry, which allows the rapid measurement of size and fluorescence for a large population of cells.

However, this only provides a snap-shot in time, and does not allow for the tracking of individual cells over time. Alternatively, traditional microscopy assays yield useful and very detailed information about the dynamical behavior of a small number of cells but typically run only for a short time period and sacrifice the ability to generate good statistics over a population. Microfluidic devices offer a nice solution, as they allow us to tightly control the cellular environment and measure the production of fluorescent reporters in many cells simultaneously over long periods of time.

### 1.2.2 Microfluidic Technology

Future successes in synthetic biology depend upon our ability to accurately extract quantitative measurements of dynamic gene regulation and to integrate this data into complete quantitative models of biological systems. This will require novel experimental techniques that enable the generation of time series data for long durations in a large number of individual living cells. An ideal data acquisition system would allow for the growth of a large population of cells in a defined environment which can be monitored by high resolution microscopy. With such a setup, the gene expression state of each cell could be monitored for the length of the experiment, providing accurate data about the temporal progression of each individual cell in the population. To this end, researchers have increasingly employed devices with fluid channels on the micron scale known as microfluidic devices.

Recently, the use of microfluidic technology has become widely popular in the field of synthetic biology and beyond. Microfluidics facilitates the study of cellular behavior because it provides the necessary tools for recreating *in vivo*-like cellular microenvironments. Microfluidics involves the handling and manipulation of very small fluid volumes, enabling the creation and control of microliter-volume reactors. The flow within microfluidic devices is laminar, ensuring that the system does not include turbulent flows which could disrupt cellular growth and observation. Over the past several years, microfluidic devices have been used to observe cellular development within dy-

dynamic microenvironments. Examples include devices designed to generate thermal or chemical gradients (Dertinger et al., 2001; Mao et al., 2002; Lin et al., 2004) and others designed to incorporate large-scale networks of fluidic channels for high-throughput cellular analysis (Hong et al., 2004; Fu et al., 2002; Balaban et al., 2004).

However, while there has been much success in the design and implementation of microfluidic devices for some purposes, it has been a challenge to design successful platforms for the long-term monitoring of single cells within a large population. Therefore, much recent research has focused on this goal using various design strategies. To monitor long-term gene expression dynamics within a larger microbial population, several researchers have developed microfluidic chemostats (Groisman et al., 2005a; Balagaddé et al., 2005). In continually providing fresh nutrients and removing cellular waste to support exponential growth, the microfluidic chemostat presents a nearly constant environment that is ideal for investigating single-cell biodynamics. In Chapter 2, we discuss the implementation of a microfluidic platform designed for the acquisition of long-term single-cell trajectories of gene expression in *S. cerevisiae*. The ability to generate single-cell expression profiles for a large number of cells is essential to understanding the roles of regulatory motifs within native and synthetic gene networks. Through such an understanding, we aim to develop an engineering-based approach to building gene-regulatory circuits, where design specifications generated from computational modeling drive the construction of regulatory networks with desired properties.

### 1.2.3 Computational Modeling

The final aspect of synthetic biology involves the application of computational modeling to understand and predict the behavior of gene networks (Hasty et al., 2002; Sprinzak and Elowitz, 2005; Gardner et al., 2000; Becskei and Serrano, 2000; Ozbudak et al., 2002; Basu et al., 2005; Blake et al., 2006). This involves the development of a set of “rules” that govern network behavior based on the connections and interactions between its components. First, genetic wiring diagrams are translated into equations

that can be analyzed. This step is analogous to the physicist's determination of Ohm's Law for the components and Kirchhoff's laws for the circuits. Next, tools from applied math and computer science are used to analyze the model in order to extract the "design criteria" for a desired output. Then modern recombinant DNA techniques are used to construct gene-regulatory networks in living cells according to the design specifications.

Our ability to synthesize and manipulate gene networks, study their behavior in living organisms, and develop a thorough quantitative understanding of their behavior has led to significant discoveries regarding some of the most fundamental cellular processes (Kepler and Elston, 2001; Ozbudak et al., 2002; Gardner et al., 2003; Carrera et al., 2009). Essential to this approach is the ability to develop computational models that can simulate and predict the behavior of cellular networks in growing and proliferating cells. For example, models of gene regulation have been developed to elucidate sources of noise in gene expression, to study the role of feedback in cellular networks, and have led to discoveries of novel network structure (Elowitz et al., 2002; Volfson et al., 2006; Ozbudak et al., 2002; Becskei and Serrano, 2000; Bennett et al., 2008; Basu et al., 2005; Simpson et al., 2003; Blake et al., 2003; Wang et al., 2006).

In addition, the construction of synthetic networks according to the specifications of quantitative models has led to the refinement of our understanding of the principles of cellular regulation (Gardner et al., 2000; Guido et al., 2006; Hasty et al., 2001a; McMillen et al., 2002). The successes in applying computational techniques to further our understanding of cellular behavior give promise to the field of synthetic biology. By combining the tools of molecular biology, microfluidic technology, and computational modeling, we can begin to probe and even mimic native network architecture and highlight key design principles that, while buried deep in the intricate regulatory web, are actually the driving force for fundamental cellular function.

### 1.2.4 Successes in Synthetic Biology

Recently, there has been significant activity directed towards designing synthetic gene networks that mimic the functionality of natural systems. Various architectures have been explored, such as toggle switches, binary logical circuits, and negative and positive autoregulatory networks (Gardner et al., 2000; Isaacs et al., 2003; Tyson et al., 2003; Becskei and Serrano, 2000; Guet et al., 2002). One particularly interesting genre of synthetic circuits is the genetic oscillator. The crucial role of time-keeping has required organisms to develop sophisticated regulatory networks to ensure the reliable propagation of periodic behavior. While many of these clock networks have been studied in great detail and their fundamental regulatory mechanisms are well-understood, it remains a challenge to engineer a synthetic system that is capable of mimicking their behavior.

The difficulty of emulating native behavior highlights the importance of information that is still unknown. The parameters that determine network dynamics are difficult to measure and impossible to ascertain from static network diagrams. Therefore, “forward engineering” can be a highly informative approach to studying the dynamics that arise from such complex network topologies, as step-by-step reconstruction can contribute invaluable information about key network properties.

Over the past several years, there have been several successful attempts at developing a synthetic oscillatory network controlled at the gene regulation level (Elowitz and Leibler, 2000; Atkinson et al., 2003; Tigges et al., 2009; Stricker, J.\* et al., 2008). These networks involved only two or three components, and mathematical modeling was instrumental in the process of designing and analyzing the network structure and demonstrating the theory behind their ability to drive periodic behavior. In *E. coli*, the *repressilator* (Elowitz and Leibler, 2000) consisted of a ring architecture of cyclic repression that was capable of generating sustained oscillations in a subset of the cells that were examined, while a two-component feedback-based circuit (Atkinson et al., 2003) was shown to generate damped oscillations. A synthetic mammalian oscillator based

on an autoregulated sense-antisense transcription control circuit yielded self-sustained and tunable oscillatory gene expression in a fraction of the cells observed (Tigges et al., 2009).

These examples represent progress in implementing an engineering-based approach to the study of gene networks, in which computational modeling is used to guide the design of novel networks and accurately predict their dynamic behavior. However, the lack of robustness in each of these networks demonstrates the need to focus on a network architecture that more closely mimics native networks. Recently, a novel and very robust two-component oscillator was created, based on principles observed to be critical for the core of a circadian clock network (Stricker, J.\* et al., 2008). The foundation of the design of the oscillator was coupled positive and negative feedback loops, and computational modeling was used to develop design criteria for robust oscillations in this system.

The success in building a network from the ground up that is capable of complex dynamic behavior validates the synthetic biology approach and demonstrates the power of computational modeling for understanding and predicting biological behavior. Using these tools, we can mimic native network architecture and highlight key features that, while buried deep in the intricate regulatory web, are actually the driving force for fundamental cellular function. In the following chapters, we describe various studies aimed at probing the fundamental functions of cells that underlay the complex behavior that sustains life, such as cellular growth and gene expression dynamics, noise and variability in gene regulation, and the cellular response to a changing environment. By understanding these basic functions, we hope to reveal design principles that are key to driving and regulating dynamic behavior.



## Chapter 2

# Monitoring Dynamics of Single-Cell Gene Expression over Multiple Cell Cycles

### 2.1 Introduction

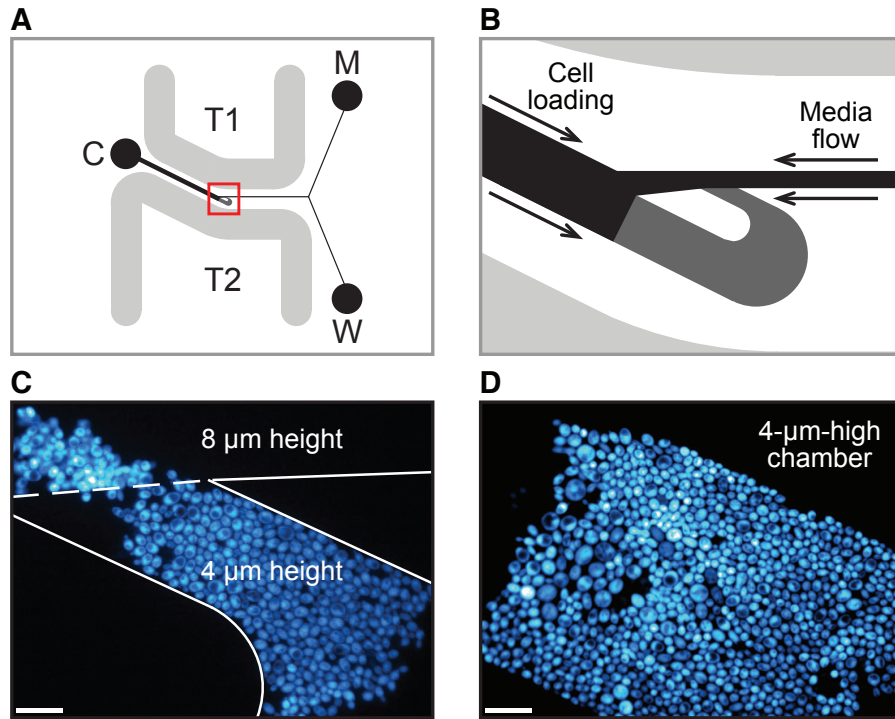
The utility of single-cell measurements with high temporal resolution has been demonstrated by recent bacterial studies, which used optical microscopy to observe *E. coli* over long time periods and reveal interesting temporal fluctuations and cell-cell variability that would otherwise be masked by population-wide measurements (Pedraza and van Oudenaarden, 2005a; Rosenfeld et al., 2005a). However, traditional microscopy methods, which typically involve viewing cells on a microscope slide or an agar plate, limit both the length and quality of an experimental run, as large groups of cells tend to grow out of the focal plane. Here, we present the development and application of a novel “free-running” microfluidic platform tailored to the generation and analysis of single-cell fluorescence data over many cellular generations. We demonstrate the utility of the platform with experimental data acquired from *S. cerevisiae*, a workhorse for eukaryotic cell biology. In addition, we present software designed for the extraction of

time-series data from a sequence of fluorescence images. The availability of this device should greatly aid quantitative modelers of both native and synthetic genetic circuits by facilitating the long-term observance of dynamical properties of gene regulation in *S. cerevisiae* and other model organisms.

## 2.2 Design and Fabrication

The T $\mu$ C is based on an implementation of the classic Tesla diode loop (Tesla, 1920; Duffy et al., 1999; Bendib and François, 2001), modified for the imaging of a monolayer culture of cells growing in exponential phase for many generations. The construction is such that the side-arm of the diode forms a shallow trapping region that constrains a population of cells to the same focal plane (Fig. 2.1). Fluid flow is utilized to continuously purge cells that grow beyond the trapping region boundaries so that the device can function as a standard chemostat. Coupled with optical autofocus, this design feature allows for the T $\mu$ C to operate in “free-run” mode over long time periods without the need for external adjustment.

In order to achieve free-running long experimental runs, a critical design objective was to avoid clogging between the media port and the trapping region. We developed a three-port chip design in which the main channel extending from the cell port splits into both a media channel and a waste channel downstream of the trapping region, which prohibits the collection of cells in the media port during the loading process (Fig. 2.1A). By constructing the height of the bypass channel to be two or three times the height of the trapping region, substantial flow can be maintained throughout this channel while flow through the trapping region remains minimal. Once cells are loaded, they receive nutrients via a combination of diffusion and advection. As the colony grows, fluidic resistance increases through the trapping region, and diffusion dominates the transport process (see Device Characterization section of Materials and Methods). Supplied with abundant nutrients, the cells are able to grow exponentially to fill the trapping region in a monolayer (Fig. 2.1C, D). The open walls of the trapping region allow for



**Figure 2.1:** The T $\mu$ C design was optimized to allow for long-term growth of cells in a monolayer. (A) Three separate ports for cell loading (C), media supply (M), and waste (W) minimize potential clogging of media supply lines. With this layout, we are able to generate optimal fluid flow both for the loading of cells into the trapping region and for the delivery of nutrients over many generations. Strong flow toward the trapping region provides the momentum necessary to carry cells into the region against high resistance. Cells that do not enter the trapping region are forced into the waste port by strong flow from the media port. Once cells are loaded, flow is reversed to run from the media port to both the waste and cell ports. Running temperature-regulated water through thermal lines T1 and T2 maintains the device at an optimal temperature. (B) A magnified view of the diode loop (red boxed region in A). The height of the trapping region (dark gray) is customized based on species. The flow channels (black) are 2–3 times higher than the trapping region. An open trapping region (black/gray interface) allows for peripheral cells to be pushed from the observation region as the colony grows. (C) Shallow trapping regions confine cells to a monolayer. Cells residing at the trapping region entrance highlight the benefit of a height-constrained growth environment. Scale bar, 20  $\mu$ m. (D) The 4- $\mu$ m-high yeast T $\mu$ C full of cells after 24 h of growth. Scale bar, 20  $\mu$ m.

peripheral cells to escape when they are pushed into the high flow of the main channel, thus permitting continuous exponential growth long after the trapping region becomes full.

Based on these design considerations, the T $\mu$ C is constructed using standard microfabrication techniques (see Fabrication Procedure section of Materials and Methods). Briefly, a PDMS chip is created with the desired channels and microstructures from a patterned silicon wafer template. Device fabrication is relatively straightforward, as it requires only a single PDMS layer. Furthermore, flow control is maintained by passive gravitational forces alone, eliminating the need for complex on-chip actuators, flow circuitry, and run-time software. The minimal three-port design allows for a rapid experimental setup that has been optimized for convenience and requires approximately 1 h of bench time. The simplicity of device construction and experimental setup makes the platform accessible to experimentalists with minimal experience in microfabrication.

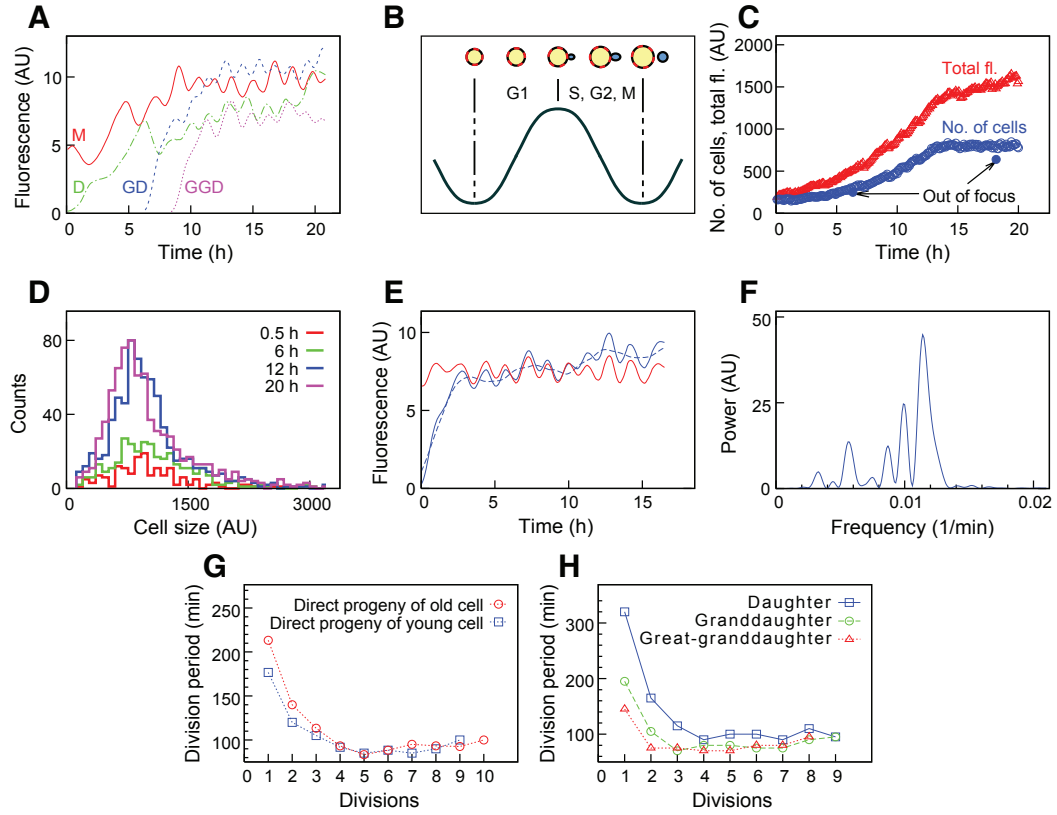
## 2.3 Experimental Results

We illustrate the utility of the device using the budding yeast *S. cerevisiae*, an excellent test for the T $\mu$ C device owing to the tendency of cells to flocculate, or aggregate into clumps. This property presents a major obstacle in the attainment of long-term single-cell temporal data in this important model organism (Hartwell et al., 1997; Simon et al., 2001; Simon and Yen, 2003; Chen et al., 2004; McMurray and Gottschling, 2004; Scheibel et al., 2004). To address this problem, we designed the trapping region height to be approximately equal to the diameter of a single yeast cell. The significant advantage of monolayer growth in a height-constrained chamber is demonstrated by visualizing the group of cells residing at the trapping region boundary (Fig. 2.1C). The cells growing inside the 4- $\mu$ m-high cell chamber are collectively in focus, and individual cells can easily be distinguished. In contrast, cells just outside of the chamber, where the height is 8  $\mu$ m, grow in multiple layers, producing blurry aggregates from which quantitative single-cell data are very difficult to extract using wide-field microscopy. Although this

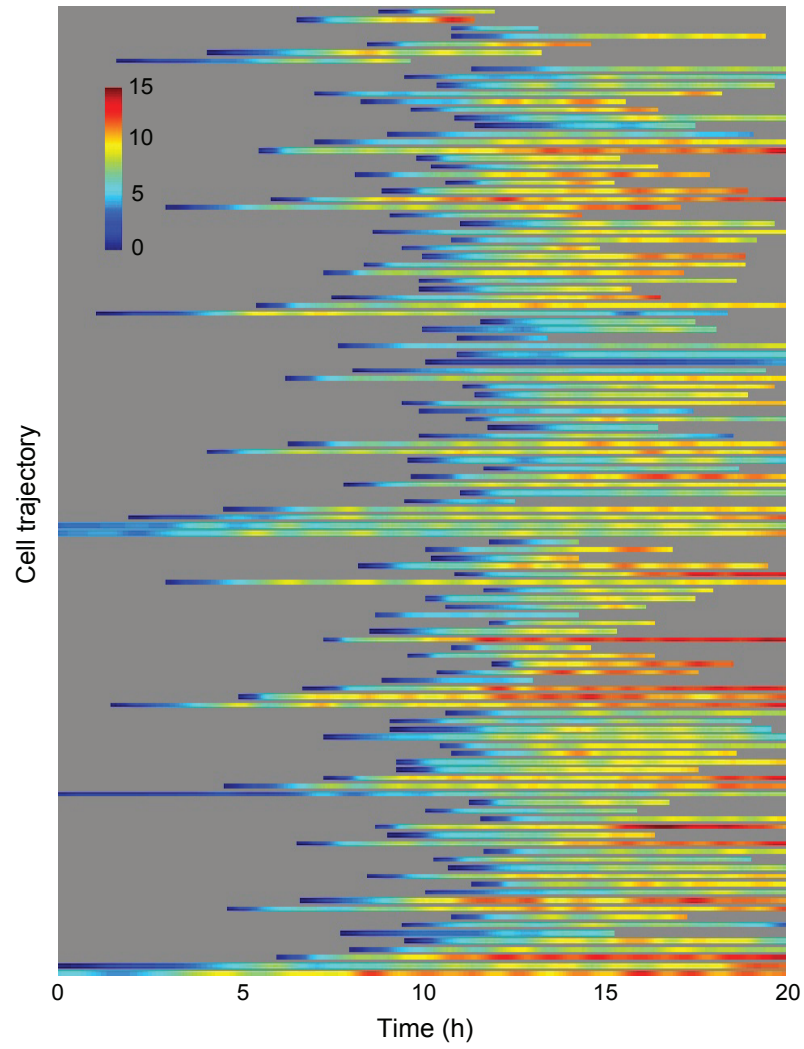
may not impede the observation of a few cells over a short period of time, as these cells begin to divide it becomes increasingly difficult to resolve individual cells and quantify their behavior.

Using the T $\mu$ C device to grow a monolayer colony of cells, we can obtain a long sequence of consistently focused fluorescence images. The extraction of single-cell expression dynamics from a sequence of images involves two major steps: (i) we segment each image into individual cells and (ii) we resolve the temporal evolution from the segmented images (see Image Processing section of Materials and Methods). We utilized cells that exhibit galactose-dependent production of yeast-enhanced Venus fluorescent protein (yeVFP) from three integrated copies of the *gal1* promoter driving *yeVFP* expression. Fully induced cells are tracked through a series of images, yielding fluorescence trajectories that rise and fall with each cell cycle (Figs. 2.2A and 2.3). Since segmentation involves tracking a daughter cell from the moment the bud begins to emerge, the fluorescence signal of the mother is observed to decrease as yeVFP freely diffuses to the bud throughout the S, G2, and M phases. The fluorescence signal rises again at the beginning of the next G1 phase, when the mother and daughter have fully divided and the mother can resume the accumulation of yeVFP (Fig. 2.2B). The cell count grows exponentially until the chamber fills, at which point the chemostat enables extended run times as the population can continue growing by pushing peripheral cells out of the trapping region (Fig. 2.2C). Histograms of cell sizes at different time points throughout the experiment retain a constant distribution, indicating that growth conditions remain optimal for the duration of the run (Fig. 2.2D).

The most novel utility of the T $\mu$ C is the ability to observe single-cell dynamics over long time periods. Quantitative accuracy of the fluorescence trajectories can be ensured by compensating for errors that may be introduced by experimental conditions and image analysis (see Quantitative Controls and Error Compensation section of Materials and Methods). These trajectories can be used to extract many types of information about individual cells, such as the average cell cycle time for any individual cell (Fig. 2.2E, F). We can also use fluorescence to monitor how the division rate of an



**Figure 2.2:** Analysis of long-term expression data acquired using the T $\mu$ C. (A) Four representative, directly related (mother, daughter, granddaughter, great-granddaughter) trajectories showing fluorescence of each cell as a function of time. (B) Cartoon illustrating how the cell tracking leads to oscillations in a gene expression time series. (C) Total number of segmented cells as a function of time and the sum of the fluorescent signals of all cells as a function of time. (D) Histograms of cell sizes at various times throughout the experiment. (E) Data processing: raw data smoothed with an 8-point Savitsky-Golay filter (solid blue line); long-term trend obtained with a 45-point Savitsky-Golay filter (dashed blue); detrended data input to Lomb-Scargle transform (solid red). (F) Sample frequency spectrum of a typical trajectory, where the  $0.0116 \text{ min}^{-1}$  frequency peak corresponds to a cell division period of 86.1 min. (G) The time per division versus division number for progeny of a nearly senescent cell and progeny of a younger cell. Each curve is averaged over three cells. (H) The time per division versus division number for the daughter, granddaughter, and great-granddaughter of a nearly senescent cell.



**Figure 2.3:** Single-cell yeVFP dynamics for 119 cells. Each row corresponds to a density plot that depicts the evolution of the amount of yeVFP in a single cell over a period of many hours (the scale bar denotes arbitrary units (AU) consistent with Fig. 2.2). Variations in cycle times and long-term trends are clearly discerned.

individual cell evolves as it ages. Given that the local minima of a trajectory mark the beginning of each G1 phase, we can calculate the duration of each cell division as the time between each G1 start. As expected, the first cell cycle of a new bud is unusually long, as the bud first has to grow to a certain size before it can begin producing buds of its own. However, our fluorescence data suggest that it often takes a young cell two or three cycles to recover to a normal division rate (Fig. 2.2G). This phenomenon is particularly pronounced in daughters of old, nearly senescent cells, as reported in various studies of aging in yeast (Egilmez and Jazwinski, 1989). Fig. 2.2G compares the evolution of division times for direct progeny of an old cell and direct progeny of a young cell, averaged over three examples of each. For progeny of the young cell, we see that the first division time is long, as expected, and then the cells quickly recover down to a steady division rate. In contrast, progeny of older cells take longer to reach this steady state, as indicated by the longer division times for the first few cell cycles. This phenomenon is highlighted in Fig. 2.2H, where the division times for the daughter, granddaughter, and great-granddaughter of a nearly senescent cell are plotted. We observe that the daughter takes three cycles to recover to steady state, the granddaughter takes two cycles to recover, and the great-granddaughter recovers immediately. These cycle-time results are consistent with previous studies that utilized different assays (Egilmez and Jazwinski, 1989), and the ability to simultaneously track gene expression over long periods highlights the utility of the device.

## **2.4 Materials and Methods**

### **2.4.1 Fabrication Procedure**

Microfluidic devices were constructed using well-established techniques (Duffy et al., 1999; Hansen and Quake, 2003) and the UCSD Integrated Technology Laboratory (ITL). Briefly, photolithographic photomasks were drawn using FreeHand MX (Macromedia Inc.; San Francisco, CA), printed onto transparency film at high resolution (Out-

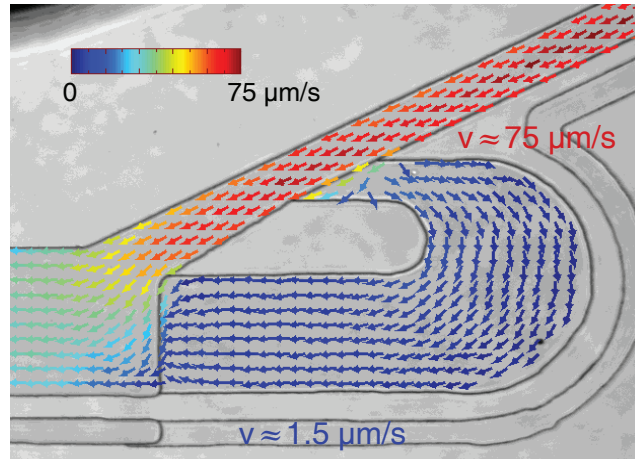


put City; Poway, CA), and mounted to clean borosilicate glass plates (McMaster-Carr; Los Angeles, CA). To make master molds, SU-8 2000 (MicroChem Corp.; Newton, MA) was spin-coated to appropriate depths using a Headway PWM32 programmable spinner (Headway Research Inc.; Garland, TX) and patterned by UV exposure via appropriate photomasks using a contact mask aligner (HTG; San Jose, CA). After all photolithographic steps were completed, SU-8 feature heights were verified using a DEK-TAK 3030ST profilometer (Sloan Technology Corp.; Santa Barbara, CA), and treated with vaporous chlorotrimethylsilane (CTMS) for 5–10 min. PDMS (Sylgard 184; Dow Corning; Midland, MI), was mixed in a 10:1 ratio with the supplied crosslinking agent and degassed in a vacuum desiccator at  $\sim$ -1 atm for 30 min to 1 h. The degassed PDMS was then poured over the silicon/SU-8 master to a depth of approximately 0.5 cm and cured at 80°C for 1 h. After curing, the hardened PDMS monolith was carefully released from the master. Fluidic ports for media/cell loading and heated water lines were bored with 20- and 16-gauge Luer stub adapters, respectively, and flushed with 0.2- $\mu$ m-filtered dH<sub>2</sub>O. Individual chips were sectioned from the PDMS monolith, sonicated in a 0.1% v/v TWEEN 80 solution for 15 min, and rinsed with dH<sub>2</sub>O. Scotch 810 office tape was used to remove any remaining particles from the PDMS surface. Finally, each chip was exposed to O<sub>2</sub> plasma at 30 W for 30 s in a 500-II Plasma Asher (Technics Plasma; Danville, CA) and brought into contact with plasma-cleaned 24×40 mm<sup>2</sup> number (no.) 1  $\frac{1}{2}$  coverslips (Corning Inc.; Corning, NY), which forms a strong irreversible bond between the two surfaces (Wu et al., 2002). Our photomask files are freely available for academic use and may be downloaded from our website (<http://biodynamics.ucsd.edu/download.htm>).

## 2.4.2 Device Characterization

We used two methods to characterize fluid flow through the T $\mu$ C under typical run-time conditions. First, we used 0.5  $\mu$ m yellow-green (YG) beads as tracers to map the flow field through the trapping region under the pressures used to deliver nutrients.

We acquired a movie of bead flow at  $10 \text{ frames s}^{-1}$  and input this to a MATLAB implementation of the particle image velocimetry (PIV) method (Sveen and Cowen, 2004). PIV provides a vector field that accurately describes the magnitude and direction of flow, and the results affirmed our prediction of creeping flow through the trapping region (Fig. 2.4). From this vector field, the maximum velocity through the feeder channel was found to be  $\sim 75 \mu\text{m s}^{-1}$ , and the mean velocity through the trapping region was  $\sim 1.5 \mu\text{m s}^{-1}$ .



**Figure 2.4:** Flow vector field produced by implementing a PIV method to process images of fluorescent beads used as flow tracers. Unit velocity vectors show the direction of flow, and colors (from blue to red) correspond to the velocity magnitude (from low to high) on a log scale.

Second, in order to ensure that a culture growing in the trapping region would receive sufficient nutrients, we performed a more detailed flow characterization by simulating nutrient delivery using two red fluorescent dyes. Nutrient transport throughout the trapping region of the T $\mu$ C is driven by both advective and diffusive processes, and the key parameters involved are the diffusion coefficient of each nutrient species and the advective velocity through the trapping region. We modeled this transport process by deriving a closed-form solution of the time-dependent chemical concentration profile throughout the trapping region. The analytical solution was fit to experimental data, yielding an estimate of the advective velocity.

The chemical species mass balance for unsteady diffusion with advection is

$$\frac{\partial c}{\partial t} + \vec{v} \cdot \vec{\nabla} c = \mathcal{D} \nabla^2 c. \quad (2.1)$$

Because flow within the microfluidic device is restricted to the laminar regime, analysis can be performed on the central streamline of flow through the trapping region (Fig. 2.5A). It is assumed that components of the velocity and concentration gradient normal to the direction of flow are negligible relative to the tangent components. Using this assumption, the physical process can be reduced to a 1D system (Fig. 2.5B), simplifying Equation (Eqn.) 2.1 to

$$\frac{\partial c}{\partial t} + V_0 \frac{\partial c}{\partial x} = \mathcal{D} \frac{\partial^2 c}{\partial x^2} \quad (2.2)$$

with the following boundary and initial conditions:

$$c(0, t) = c_0, \quad t > 0 \quad (2.3)$$

$$c(L, t) = c_0, \quad t > 0 \quad (2.4)$$

$$c(x, 0) = 0, \quad 0 < x < L. \quad (2.5)$$

Here,  $V_0$  is the average channel velocity,  $\mathcal{D}$  is the molecular diffusivity of the chemical species, and  $c_0$  is both the steady-state and boundary-supported concentration.

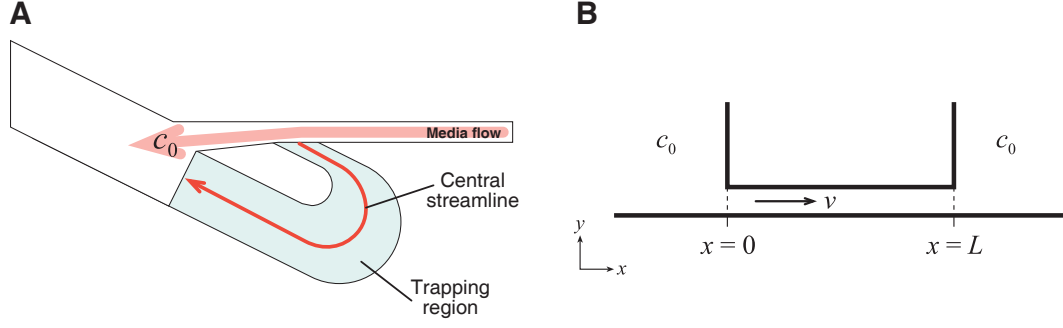
This system of equations can be further simplified by introducing the following dimensionless parameters:

$$C = \frac{c - c_0}{c_0} \quad (2.6)$$

$$\xi = \frac{x}{L} \quad (2.7)$$

$$\tau = \frac{\mathcal{D}}{L^2} t. \quad (2.8)$$

Hence, Eqns. 2.2–2.5 become



**Figure 2.5:** Schematics used in simulating nutrient delivery in the T $\mu$ C. **(A)** Analysis is performed on the central streamline of flow through the trapping region. **(B)** 1D reduction of the system used to characterize diffusive and advective transport.

$$\frac{\partial C}{\partial \tau} + N_{\text{Pe}} \frac{\partial C}{\partial \xi} = \frac{\partial^2 C}{\partial \xi^2} \quad (2.9)$$

and

$$C(0, \tau) = 0, \quad \tau > 0 \quad (2.10)$$

$$C(1, \tau) = 0, \quad \tau > 0 \quad (2.11)$$

$$C(\xi, 0) = -1, \quad 0 < \xi < 1, \quad (2.12)$$

where

$$N_{\text{Pe}} = \frac{V_0 L}{\mathcal{D}} \quad (2.13)$$

is a dimensionless ratio of advective to diffusive transport that commonly arises in heat and mass transfer problems and is known as the Peclet number. Because this system is linear with respect to  $C$  and has homogeneous boundary conditions, it can easily be solved using the separation of variables method (Haberman, 1998). This yields the closed-form solution,

$$C(\xi, \tau) = \exp\left(\frac{N_{\text{Pe}}}{2}\xi\right) \sum_{n=1}^{\infty} A_n \sin(n\pi\xi) \exp(-\lambda_n \tau) \quad (2.14)$$

or

$$c(x, t) = c_0 \left[ \exp\left(\frac{N_{\text{Pe}}}{2L}x\right) \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}x\right) \exp\left(-\lambda_n \frac{\mathcal{D}}{L^2}t\right) + 1 \right], \quad (2.15)$$

where

$$\lambda_n = \frac{1}{4} [(2n\pi)^2 + N_{\text{Pe}}^2] \quad (2.16)$$

and

$$A_n = 4 \exp\left(-\frac{N_{\text{Pe}}}{2}\right) \left[ \frac{2n\pi \cos(n\pi) + N_{\text{Pe}} \sin(n\pi) - 2n\pi \exp\left(\frac{N_{\text{Pe}}}{2}\right)}{N_{\text{Pe}}^2 + (2n\pi)^2} \right]. \quad (2.17)$$

Experimental concentration profiles were visualized using 10 kDa conjugated Rhodamine B isothiocyanate (RBITC). We estimated the diffusion constant of this dye using the Stokes-Einstein equation for large diffusing particles (Bird et al., 2002) as

$$\mathcal{D}_{AB} = \frac{k_b T}{6\pi\mu_B s_A}, \quad (2.18)$$

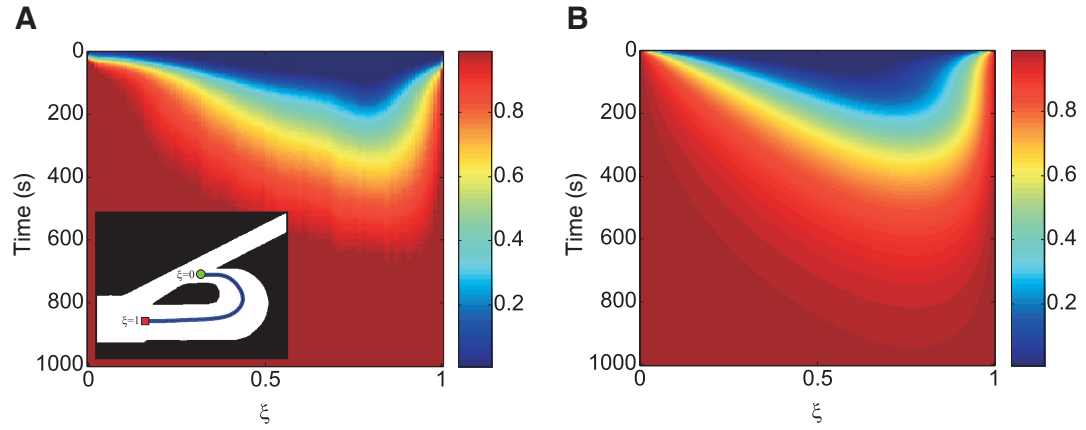
where  $\mathcal{D}_{AB}$  is the diffusion coefficient of solute  $A$  in solvent  $B$ ,  $k_b$  is Boltzmann's constant,  $T$  is the ambient temperature,  $\mu_B$  is the dynamic viscosity of solvent  $B$ , and  $s_A$  is the molecular Stokes radius of solute  $A$ . The value of  $s_A$  was determined using a previously-reported correlation for dextrans based on molecular weight (MW) (Venturoli and Rippe, 2005):

$$s_A = 0.488 (MW)^{0.437}. \quad (2.19)$$

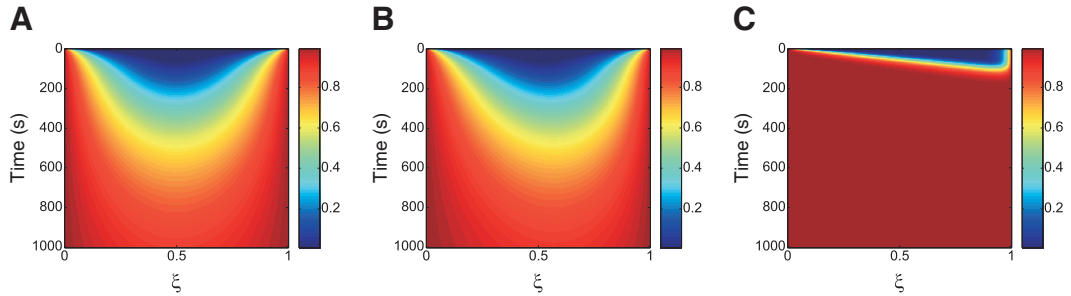
Using this approximation,  $\mathcal{D}_{AB}$  for a 10 kDa dextran was found to be  $7.90 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$ . The calculated diffusivity of the dye was input to our analytical model of the transport process, which yielded a value of  $1 \text{ } \mu\text{m s}^{-1}$  for the advection velocity,  $V_0$ , and a path length,  $L$ , of  $570 \text{ } \mu\text{m}$  (approximately the path length analyzed in the experimental data), altogether providing a  $N_{\text{Pe}}$  of 7.125.

Experimental data was acquired by imaging fluorescent dye entering the trapping region and recording intensity values along the central streamline. Fluorescence images were acquired every 5 s for 1.4 h and post-processed by background subtraction, flat field correction, and smoothing using a 5-pixel Gaussian kernel. A quadratic interpolating spline was fit to 6–10 points manually chosen to define the central streamline and then subdivided into 100 points. For each image, fluorescence values at each point along the path were calculated as the mean pixel intensity over a  $10\text{-}\mu\text{m}$  line normal to the spline. Experimental concentration was assumed to be linearly proportional to fluorescence signal and was normalized by the steady-state value. The temporal evolution profiles of the experimental data and the analytical model for the empty device are shown in Fig. 2.6. Analytical data is calculated from a 100-term Fourier series and plotted as  $C'(\xi, t) = C(\xi, t) + 1$ . A comparison of the two profiles shows good agreement, supporting the existence of creeping flow on the order of  $1\text{ }\mu\text{m s}^{-1}$  within the trapping region of the device under normal operating conditions. In addition, the calculated velocity of the creeping flow closely matches the value obtained from the PIV analysis ( $1.5$  versus  $1.0\text{ }\mu\text{m s}^{-1}$ ). The asymmetry in the concentration profile over  $\xi$  is due primarily to advective transport. Closer inspection of the analytical solution reveals that as  $N_{\text{Pe}} \rightarrow 0$ , Eqns. 2.14–2.17 degenerate to the solution for simple 1D unsteady diffusion and the physical process becomes spatially symmetric about  $\xi = 0.5$ . Moreover, the speed of the transport process is highly dependent on the value of  $N_{\text{Pe}}$  (Fig. 2.7).

Our transport analysis based on the RBITC dye provides an estimate for the Peclet number in the case of a heavy molecule perfusing into the empty trapping region. To better approximate the diffusive transport of nutrients through the trapping region, we repeated this analysis with the red dye Sulforhodamine 101 (Acid-Free Texas Red), which has a diffusion constant on the same order of magnitude as a typical nutrient molecule (calculated to be  $5.72 \times 10^{-6}\text{ cm}^2\text{ s}^{-1}$ ). With the diffusion constant an order of magnitude larger than RBITC, the Peclet number was an order of magnitude smaller, and we observed increased domination of diffusive over advective transport in both experimental and analytical results (Fig. 2.8). In experiments involving cells growing within

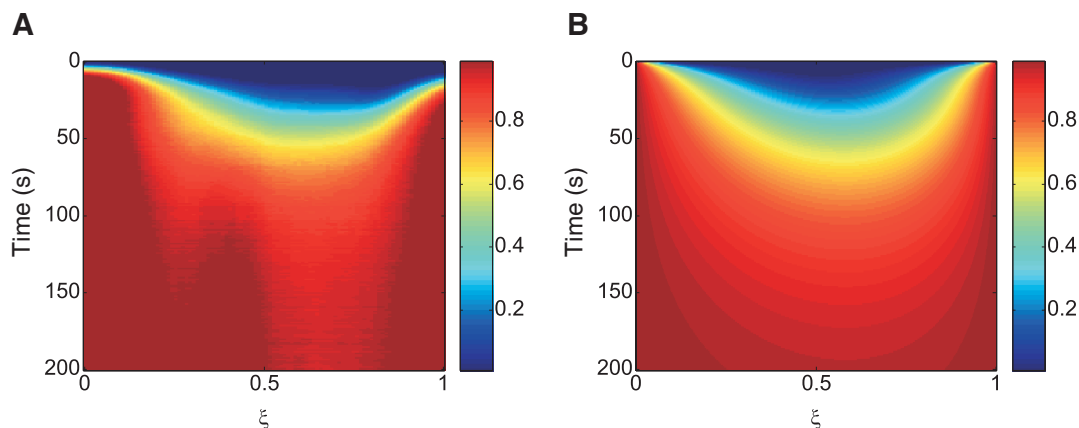


**Figure 2.6:** Comparison of experimental data and analytical solution for nutrient transport into the trapping region using 10 kDa Dextran-conjugated RBITC. **(A)** Experimental data collected along the central streamline of the trapping region. The inset shows the path analyzed (blue line), starting from the green circle and ending at the red square. **(B)** Analytical solution of 1D diffusion with advection for large molecule transport.



**Figure 2.7:** Analytical analysis of the time evolution of 1D concentration profiles for different advective velocities. **(A)**  $V_0 = 0 \mu\text{m s}^{-1}$ ,  $N_{\text{Pe}} = 0$ . **(B)**  $V_0 = 0.2 \mu\text{m s}^{-1}$ ,  $N_{\text{Pe}} = 1.425$ . **(C)**  $V_0 = 5 \mu\text{m s}^{-1}$ ,  $N_{\text{Pe}} = 36.625$ .

the trapping region, we expected to see the Peclet number decrease even further, as advective velocity should be slowed by the increased resistance provided by colony expansion. Indeed, upon monitoring Sulforhodamine 101 dye penetration with the chamber both empty and full of cells, we found that diffusion dominates the transport process.



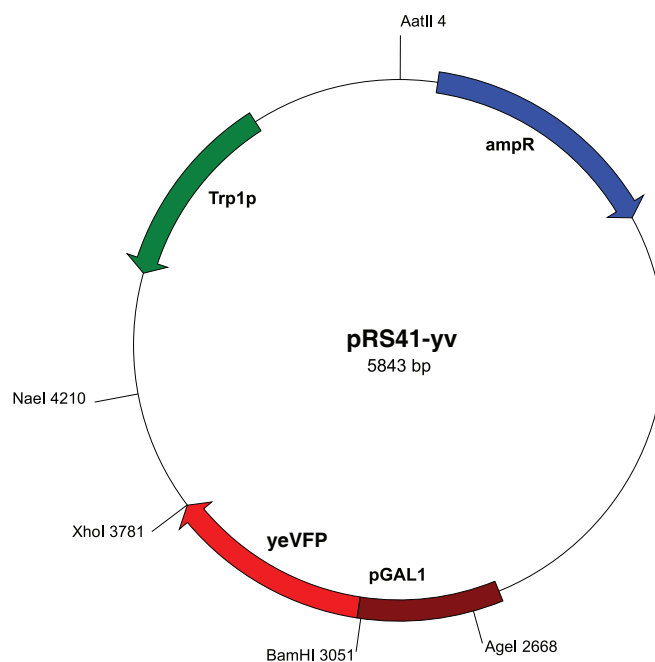
**Figure 2.8:** Comparison of experimental data and analytical solution for nutrient transport into the trapping region using Sulforhodamine 101. (A) Experimental data collected along the central streamline of the trapping region. (B) Analytical solution of 1D diffusion with advection for small molecule transport.

### 2.4.3 Strain and Cell Culture

The yeast strain was created by targeted chromosomal integration of the pRS41-yv vector (Fig. 2.9) at the *gal1-10* promoter locus of *S. cerevisiae* strain K699 (a, ADE2, ura3, his3, trp1, leu2). The vector was constructed using standard recombination techniques from two commercially available vectors. The tryptophan marker from pRS404 was inserted into the pESC-His vector containing the *gal1-10* promoter locus of *S. cerevisiae* by cutting both vectors at the AatII and NaeI restriction sites. The gene encoding a YFP variant, yeVFP, was inserted into the BamHI and XhoI restriction sites downstream of the *gal1* promoter for fluorescence production inducible by the addition of galactose to the medium. In preparation for transformation, the resulting vector, pRS41-yv, was cut at the AgeI restriction site within the *gal1-10* promoter to ensure sufficient homol-



ogy for recombination into the yeast genome. Selection of integrants was performed by first growing cultures in synthetic drop-out (SD) medium containing 2% w/v glucose and supplemented with all amino acids except tryptophan. Subsequently, flow cytometry was used to identify a triple integrant in order to maximize yeVFP production. For experiments, cultures were grown in SD supplemented with all amino acids and containing 2% w/v galactose for full induction of the *galI* promoter. To minimize flocculation of yeast while growing in the incubator shaker, a single colony was initially inoculated into a microcentrifuge tube containing 1 ml of medium and vortexed on high for 2 min before being transferred to a culture tube containing 4 ml total medium. Cultures were then grown at 30°C for 18–24 h to an optical density at 600 nm (OD<sub>600</sub>) of  $1.0 \pm 0.25$ . In preparation for loading, four samples from a single culture were diluted to an OD<sub>600</sub> of 0.05 in microcentrifuge tubes containing 1 ml of medium and vortexed for 5–10 min. Cultures were then recombined for a total loading volume of 4 ml.



**Figure 2.9:** Map of the pRS41-yv vector used to integrate three copies of the *galI* promoter driving expression of yeVFP into the yeast genome.

#### 2.4.4 Loading Procedure

In preparation for cell observation, devices were mounted to the microscope stage and connected to thermal water baths to maintain the optimal growth temperature of 30°C. Thermal connections were made using Tygon microbore tubing (0.050 in inner diameter, 0.090 in outer diameter; Cole Parmer; Vernon Hills, IL) connected to chip thermal ports with 16-gauge dispensing needles (McMaster-Carr; Los Angeles, CA). Water temperatures flowing into and out of the device were monitored using in-line thermocouples. Following device priming with filtered dH<sub>2</sub>O, an open 10 ml syringe serving as a waste reservoir was filled with medium and suspended 5 in above the device. A similar connection for medium was made at the device media port at a height of 25 in. Cells and 2.5- $\mu$ m-diameter YG fluorescent beads were loaded into the device at the cell port, with the beads serving as a fluorescence intensity standard, until several of each entered the trapping region. During loading, all flows were directed toward the waste port to prevent contamination of the media line. To minimize clogging, strong flow from the media port was used to flush clear all cells residing in the shared fluid channels. Finally, the three reservoirs were brought to their final run-time heights, with the cell reservoir fixed at a height 1 in above the waste reservoir and the medium reservoir fixed at a height 1 in above the cell reservoir. These differential heights provided for both gentle flow of medium through the trapping region and strong flow of medium into the waste port, thereby feeding the cells to be monitored while preventing discarded cells from re-entering the system.

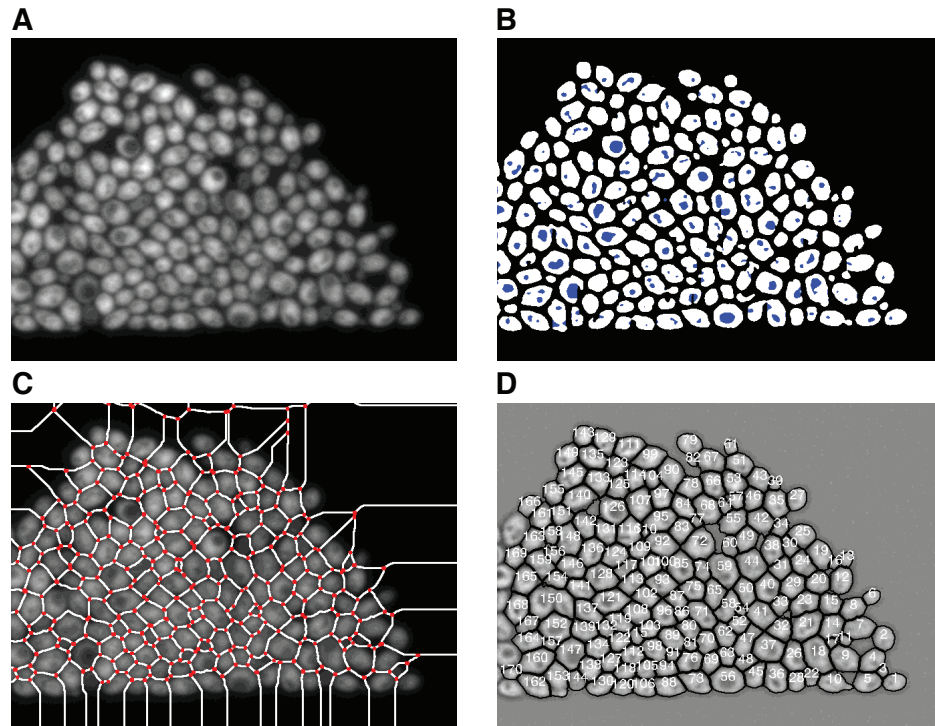
#### 2.4.5 Microscopy

Image acquisition was performed on a Diaphot TMD epifluorescent inverted microscope (Nikon Instruments Inc.; Tokyo, Japan) outfitted with fluorescence excitation and emission filter wheels, a ProScan II XY-motorized stage with fine focus control and a hardware-based autofocus controller (Prior Scientific Inc.; Rockland, MA), and Uniblitz VS35 high-speed shutters (Vincent Associates; Rochester, NY). Images were

acquired using an Orca-ER cooled CCD camera (Hamamatsu Photonics; Hamamatsu, Japan) and a custom positioning and multispectral acquisition application written in LabVIEW (National Instruments; Austin, TX). Imaging for the autofocus controller was performed using a Cohu 4900 Series CCD camera (Cohu Inc.; San Diego, CA) mounted to one of the microscope eyepieces. Fluorescence visualization of yeVFP and red fluorescent beads was performed with narrow band-pass excitation and emission filters (filter set #86006 for CFP/YFP/DsRed; Chroma Technology Corp.; Rockingham, VT).

### 2.4.6 Image Processing

The extraction of single-cell expression dynamics from a sequence of images involves two major steps: (i) we segment each image into individual cells, and (ii) we resolve the temporal evolution from the segmented images. To begin the segmentation process, each image is flat field corrected and pre-processed with a  $3 \times 3$ -pixel<sup>2</sup> Gaussian filter to remove high-frequency noise (Fig. 2.10A). Approximate locations of the cells are determined using a seeding technique (Adams and Bischof, 1994). To form seeds, we use the differential of the image averaged with boxcar filters of small and large sizes ( $5 \times 5$  pixel<sup>2</sup> and  $15 \times 15$  pixel<sup>2</sup>, respectively) to support the robust location of local intensity maxima independent of the mean level of the signal. The resulting binary mask is over-segmented in most cases due to the presence of vacuoles which appear as islands of the background inside the foreground cell seeds (Fig. 2.10B). These islands are eliminated by the transferring of closely-connected regions of the background into the foreground. After this correction, the resulting binary image is processed with a morphological distance transform and a watershed segmentation (Vincent, 1993). These operations produce a set of dams which separate cell seeds and are typically located at troughs of the grayscale intensity image (Fig. 2.10C). In the final step of segmentation, the seeds are used for growing the grayscale region in an operation which involves expanding the area of the cells from the peaks within the seeds down to a specified level



**Figure 2.10:** Major steps of image segmentation. For visual clarity, only one fourth of the entire image is shown. (A) Original fluorescence image after flat field correction and the removal of high-frequency noise. (B) Results of initial seeding (white) along with filled “holes” (blue) form the resulting binary mask. (C) Results of the watershed segmentation, where segments (white lines) joining nodes (red dots) may be eliminated in the interactive mode to allow the merging of neighboring cells. (D) To complete segmentation, recognized cells are enumerated and their boundaries are displayed in black. Cells crossing the boundaries of the field of view are abandoned.

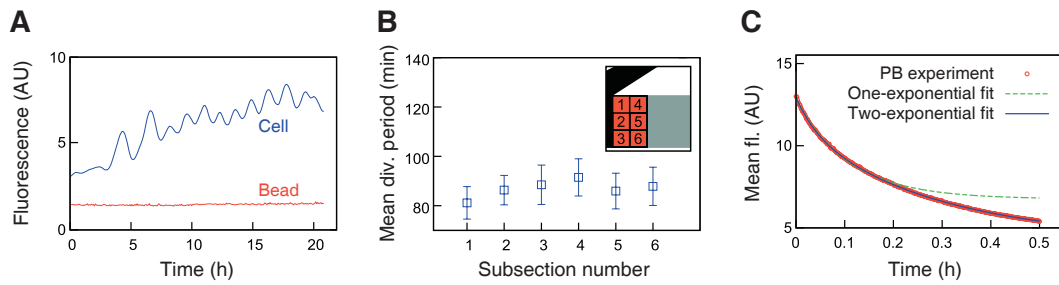
above the background (Fig. 2.10D). Here, the watershed dams serve to preclude the joining of neighboring cells during the growth process. In addition, the same skeleton of dams is used to allow for an effective means of manual correction, since the removal of a dam segment between touching cells usually results in their joining during the growing of the grayscale region. The resulting algorithm is implemented using the Interactive Data Language (IDL) programming language (ITT Visual Information Solutions; Boulder, CO). In automated mode, which is used in most cases, the segmentation of a single image with 800 cells takes about 30 s on a standard personal workstation. In this mode, we determined that the fraction of errors in segmentation typically does not exceed 5% of the total number of recognized cells. Segmentation errors arise primarily when dark regions inside a cell take on values close to the background level, causing the software to draw a watershed line through the cell, or when the boundary between two cells is difficult to observe, causing them to be considered a single cell.

For the second step of image analysis, in which cells are tracked through a set of segmented images, we adopted a set of tools developed in the context of colloidal suspensions and granular materials (Crocker and Grier, 1996). These tools implement cluster analysis, and we used the center of mass coordinates and the total fluorescence integrated over the cellular area as our cluster variables. Errors in this step arise primarily when a clump of cells near the trapping region boundary breaks off and leaves the chamber, causing a shift in the positions of many cells in the chamber. A manual correction option during each step of the image processing makes it possible to correct for errors, and future implementations of the software will include improvements upon these methods. Our image analysis software is freely available for academic use and may be downloaded from our website (<http://biodynamics.ucsd.edu/download.htm>).

## **2.4.7 Quantitative Controls and Error Compensation**

Quantitative accuracy of the fluorescence trajectories can be ensured by compensating for errors that may be introduced by experimental conditions and image analysis.

As a quantitative control in our cell experiments, we simultaneously imaged calibration-standard fluorescent beads of similar brightness. Because these beads are resistant to photobleaching, any variation in intensity from frame to frame will be a measure of deviations due to either focal drift or fluorescent bulb fluctuations. The trajectory of a bead is shown in Fig. 2.11A alongside a typical cell trajectory. The coefficient of variation of the bead signal was found to be 3.0%. While this low noise level is negligible when analyzing general trends in data, the bead trajectory can be used to compensate for experimental error in studies for which these fluctuations would be more detrimental.



**Figure 2.11:** Quantitative controls performed to ensure accuracy of the fluorescence data. (A) The fluorescence trajectory of a bead (red) alongside a sample cell trajectory (blue). (B) Mean division period in each of six subsections of the field of view. Error bars represent standard deviation, and the inset shows a map of subsections. (C) Results of a photobleaching experiment to determine the decay rate of yeVFP in response to fluorescent light exposure. Experimental data is shown as red circles. The best fit is given by a two-exponential function (solid blue curve). However, in the analysis we used a one-exponential fit (dashed green curve) for simplicity.

A concern regarding cells being constrained to a shallow chamber is that nutrients may be unevenly distributed, leading to varying growth rates throughout the trapping region. To address this question, we divided the field of view into subsections and calculated division period statistics for all cells residing in each area. We extracted the average cell cycle time for each trajectory by analyzing a power spectrum of the time-series data. Spectra were calculated using the Lomb-Scargle transform for unevenly-spaced data to account for occasional gaps due to tracking errors. In calculating the average cell division period in each of six regions, we found a mean and standard deviation of 86.88 and 3.43 min respectively (Fig. 2.11B). This low variation reflects the

constant micro-environment throughout the observed region.

Another typical concern in fluorescence microscopy is that the intermittent exposure of fluorescent molecules to high-intensity light will bleach them over time, adversely affecting the data. To address this, we acquired a photobleaching curve and modeled the effects of exposure on the quantitative integrity of our fluorescence trajectories. In a typical experiment, fluorescence images were acquired every 5 min with exposure times of 50–250 ms. With these acquisition settings, effects of photobleaching of yeVFP molecules were determined to be negligible. A photobleaching curve was obtained by subjecting a full field of view of cells expressing *yeVFP* to constant exposure of fluorescent light. Images were captured every 1.6 s, and mean fluorescence over a region of interest was plotted against time (Fig. 2.11C). We fit this data using both one- and two-exponential functions and for simplicity used a one-exponential fit to obtain the decay rate,  $\gamma_p$ , for yeVFP due to exposure from fluorescent light.

To model the effect of photobleaching on our acquisition of a single-cell fluorescence trajectory, we assume that the fluorescence data obtained from a cell,  $\dot{x}$ , is a combination of the true fluorescence signal,  $f(x, t)$ , plus a degradation term dependent on the amount of yeVFP present in the cell,  $x$ :

$$\dot{x} = f(x, t) - \gamma_p x. \quad (2.20)$$

In order to extract the true signal from the observed signal, we use an explicit first-order discrete model of the ordinary differential equation (ODE) as follows:

$$(x_{i+1} - x_i) \frac{1}{\Delta t} = f_i(x, t) - \gamma_{eff} x_i. \quad (2.21)$$

Here,  $\gamma_{eff}$  is the effective decay rate representing the effect of photobleaching due to a single exposure of duration  $\Delta t$ . This equation can be rearranged to give an expression for  $f_i(x, t)$ :

$$f_i(x, t) = (x_{i+1} - x_i) \frac{1}{\Delta t} + \gamma_{eff} x_i. \quad (2.22)$$

We now want to approximate the actual value of the fluorescence data as it would be recorded if there was no degradation due to photobleaching. The corrected fluorescence value,  $\tilde{x}_{i+1}$ , at each time point can be evaluated based on the current and previous data points as follows:

$$\tilde{x}_{i+1} = x_i + \Delta t f_i(x, t) = x_i + \Delta t \left[ (x_{i+1} - x_i) \frac{1}{\Delta t} + \gamma_{eff} x_i \right] = x_{i+1} + \Delta t \gamma_{eff} x_i. \quad (2.23)$$

The effective decay rate,  $\gamma_{eff}$ , should have the same effect on fluorescent intensity over  $\Delta t$  that the calculated decay rate,  $\gamma_p$ , has over the duration of exposure,  $t_{exp}$ . That is,  $\gamma_{eff} \Delta t = \gamma_p t_{exp}$ . Given a maximum exposure time of 250 ms and the value of  $\gamma_p$  calculated from the photobleaching curve of 0.256, we find that the value of  $\gamma_{eff} \Delta t$  yields a negligible correction term of 0.1% for each fluorescence value,  $x_{i+1}$ . Correcting a given trajectory for this effect has a negligible effect on the data.

An additional concern regarding the extraction of quantitative information from fluorescence microscopy data is that an individual cell, which often moves throughout the field of view during an experiment due to colony expansion, may not experience spatially-constant illumination. To address this concern, we applied a flat field correction to each image to compensate for nonlinearities in our microscope optics. The correction matrix was determined by filling the field of view with a fluorescent dye of even thickness and acquiring an image at the run-time settings. Since the ideal signal should be constant over all pixels, this image could be used as a weighting template to remove optical artifacts from our data due to variations in either the pixel-to-pixel sensitivity of the camera or distortions in the optical path.

## 2.5 Acknowledgements

Chapter 2 contains material originally published as Cookson, S.\*, Ostroff, N.\*, Pang, W. L.\*, Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene



expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032-E1-6. (\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

# Chapter 3

## Origins of extrinsic variability in eukaryotic gene expression

### 3.1 Introduction

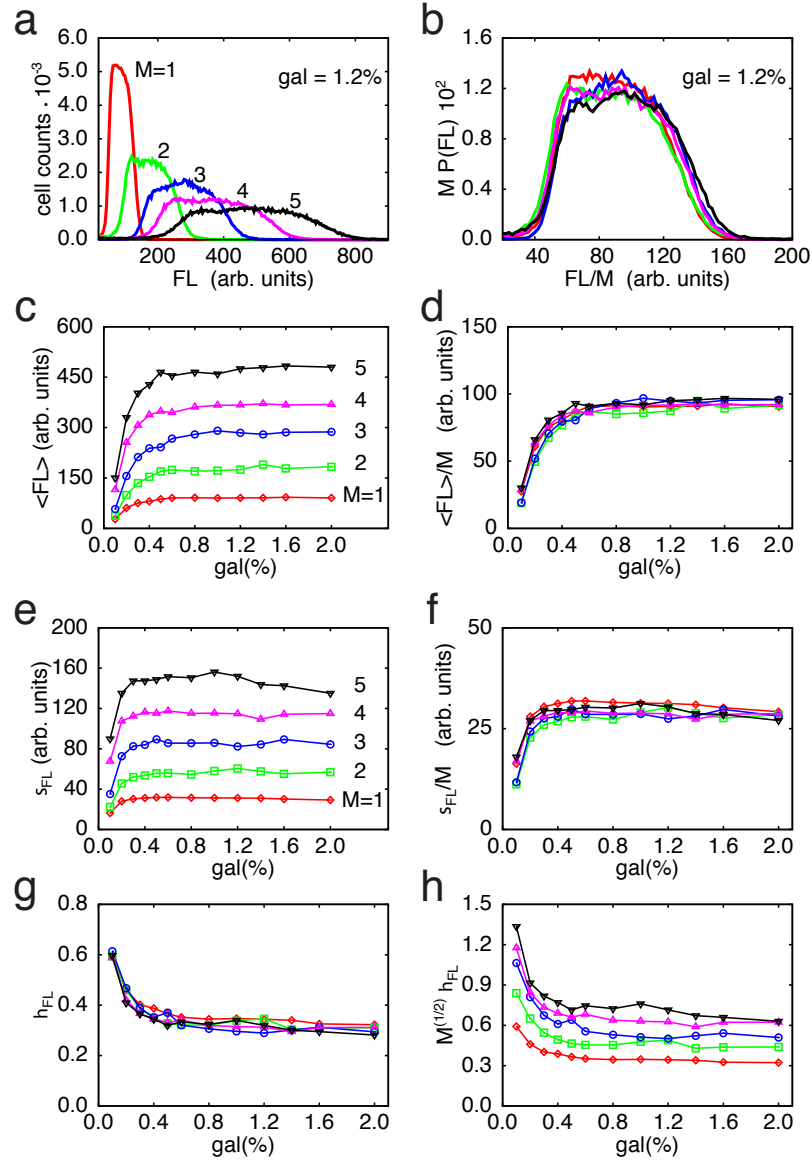
Variable gene expression within a clonal population of cells has been implicated in a number of important processes including mutation and evolution (Rao et al., 2002; Reanne, 2002), determination of cell fates (Sternberg and Felix, 1997; Grossman, 1995) and the development of genetic disease (Kemkemer et al., 2002; Cook et al., 1998). Recent studies have demonstrated that a significant component of expression variability arises from extrinsic factors thought to influence multiple genes in concert (Elowitz et al., 2002; Raser and O'Shea, 2004; Pedraza and van Oudenaarden, 2005b; Rosenfeld et al., 2005b), yet the biological origins of this extrinsic variability have received little attention. Here we combine computational modeling (Arkin et al., 1998; Ozbudak et al., 2002; Kepler and Elston, 2001; Hasty et al., 2002; Blake et al., 2003; Isaacs et al., 2003; Kaern et al., 2005; Simpson et al., 2003) with fluorescence data generated from multiple promoter-gene inserts in *Saccharomyces cerevisiae* to identify two major sources of extrinsic variability. One unavoidable source arising from the coupling of gene expression with population dynamics leads to a ubiquitous noise floor in

expression variability. A second source which is modeled as originating from a common upstream transcription factor exemplifies how regulatory networks can convert noise in upstream regulator expression into extrinsic noise at the output of a target gene (Pedraza and van Oudenaarden, 2005b). Our results highlight the importance of the interplay of gene regulatory networks with population heterogeneity for understanding the origins of cellular diversity.

## 3.2 Results and Discussion

To investigate variability in eukaryotic gene expression, we used the native *GALI* promoter of the yeast *Saccharomyces cerevisiae* with the *yEGFP* (yeast-enhanced Green Fluorescent Protein) as a quantifiable marker. We constructed five yeast strains by varying the number of copies integrated into the *GALI-10* locus on chromosome II. These multiple-copy constructs can be used to determine if the variability in gene expression is due to intrinsic or extrinsic sources. Adopting the standard nomenclature, *intrinsic* noise originates from the small number of regulatory molecules participating in the inherently noisy biochemical reactions leading to expression, while *extrinsic* variability arises from sources such as the variation of the elements of the transcriptional machinery common to all genes or fluctuating environmental variables. This extrinsic variability can have stochastic or deterministic origins, and we reserve the use of the term *noise* to denote a random, as opposed to a deterministic, origin for the variations.

In the absence of glucose, the *GALI* promoter is activated in response to galactose, and our cell strains were induced to produce *yEGFP* using varying amounts of galactose ranging from 0.1% to 2%. Single-cell fluorescence data was collected using flow cytometry, and representative distributions are depicted in Fig. 3.1a for fixed galactose concentration. Information regarding the underlying transcriptional process can be extracted by analyzing the scaling properties of the data sets with respect to the copy number. For example, the mean fluorescence increases as a function of galactose concentration and the copy number (Fig. 3.1c), and when fluorescence is divided by the



**Figure 3.1:** Experimental results for GFP expression for different copy numbers and galactose concentrations. **a**, Histograms of GFP measurements for copy numbers from  $M = 1$  to  $M = 5$  above the saturation ( $gal = 1.2\%$ ). **b**, The collapse of GFP distributions under the transformation  $F \rightarrow F/M$ ,  $P(F) \rightarrow MP(F/M)$  implies an extrinsic source of variability **c**, Induction curves for copy numbers from  $M = 1$  to  $M = 5$ . **d**, Collapse of the induction curves implies that transcription from each promoter is independent. **e**, Standard deviations of GFP corresponding to induction curves. **f**, The collapse of the standard deviation implies an extrinsic source of variability **e**. **g**, The collapse of the CV for different copy number implies an extrinsic source of variability. **h**, Lack of collapse implies the variability is not of intrinsic origin.

corresponding copy number, the induction curves for all of the strains “collapse” to a single curve (Figs. 3.1c,d). This indicates that each promoter-gene pair transcribes at the same mean rate upon insertion. We can utilize scaling in an analogous manner to deduce information about the origins of variability in the system. Depending on the relative magnitude of intrinsic noise versus extrinsic variability, we expect to observe different dependencies (scalings) on the copy number  $M$ .

In order to illustrate how scaling can be used to determine whether the observed variability is intrinsic or extrinsic, suppose there are  $M$  promoter-gene pairs producing GFP molecules, and let the level of production of the  $i$ th copy be the sum of two terms representing the mean and fluctuations about the mean:  $g_i = \langle g_i \rangle + \tilde{g}_i$ . The time-average of the fluctuating term is zero ( $\langle \tilde{g}_i \rangle = 0$ ) and the variance is given by  $v_i = \langle g_i^2 \rangle - \langle g_i \rangle^2$ . Then the mean and variance for the total amount of GFP produced by  $M$  copies are  $\langle G_M \rangle = \sum_{i=1}^M \langle g_i \rangle$  and  $V_M = \sum_{i=1}^M \left[ v_i + \sum_{j \neq i} \langle \tilde{g}_i \tilde{g}_j \rangle \right]$ , respectively. The first term in the expression for  $V_M$  represents the sum of the variances of the individual copies, and the second term is the contribution of cross-correlations between them. For identical copies,  $\langle g_i \rangle = \langle g \rangle$ ,  $v_i = v$ ,  $\langle \tilde{g}_i \tilde{g}_j \rangle = c$ , and the total mean is proportional to the copy number,  $\langle G_M \rangle = M \langle g \rangle$ , while the total variance is given by  $V_M = Mv + M(M-1)c$ . When  $c = 0$ , the fluctuations are completely uncorrelated and the variance is proportional to the copy number. This is the intrinsic noise limit, where expression events from individual copies are completely independent. When  $c = v$ , the variance is proportional to the square of the copy number. In this limit the fluctuations are completely correlated, and it is often caused by extrinsic factors. Using the results for the variance and mean, the coefficient of variation (CV)  $\eta$  (standard deviation divided by the mean), is expected to scale as  $\eta \propto M^{-1/2}$  for the case of purely intrinsic noise and be independent of the copy number in the case of extrinsic variability. Thus, if the system is dominated by intrinsic factors, the quantity  $M^{1/2}\eta$  should collapse to the same galactose-dependent curve for all copy numbers. On the other hand, if the system is dominated by extrinsic sources, then  $\eta$  itself should collapse for all copy numbers.

In order to ascertain if the GAL system is dominated by intrinsic noise or ex-

trinsic variability, we plot the unscaled ( $\eta$ ) and scaled ( $M^{1/2}\eta$ ) coefficient of variation as a function of galactose (Figs. 3.1g,h). We observe that even though different strains exhibit small differences, the collapse of the data in the unscaled case indicates that the coefficient of variation is independent of the copy number. An even more striking collapse is presented in Fig. 3.1b, where we rescale fluorescence distributions  $P(F)$  for a representative value of galactose concentration above saturation (gal=1.2%) and for different strains. In obtaining this collapse, we first normalized the distributions and rescale by the copy number as follows:  $F \rightarrow F/M$ ,  $P(F) \rightarrow MP(F/M)$ . This collapse implies that not only the standard deviation is proportional to the copy number, but generally the  $n$ th moment of the fluorescence distribution is proportional to  $M^n$ . These results corroborate findings by Raser and O'Shea (Raser and O'Shea, 2004), who recently used a “two-color” approach (Elowitz et al., 2002) to demonstrate that extrinsic sources dominate the variations observed in the GAL system. In the Materials and Methods section below, we demonstrate that our findings are unchanged when the genes are moved to different chromosomes, and describe how our data analysis can be generalized for intermediate cases where there is mixture of scaling arising from the contribution of both intrinsic and extrinsic sources.

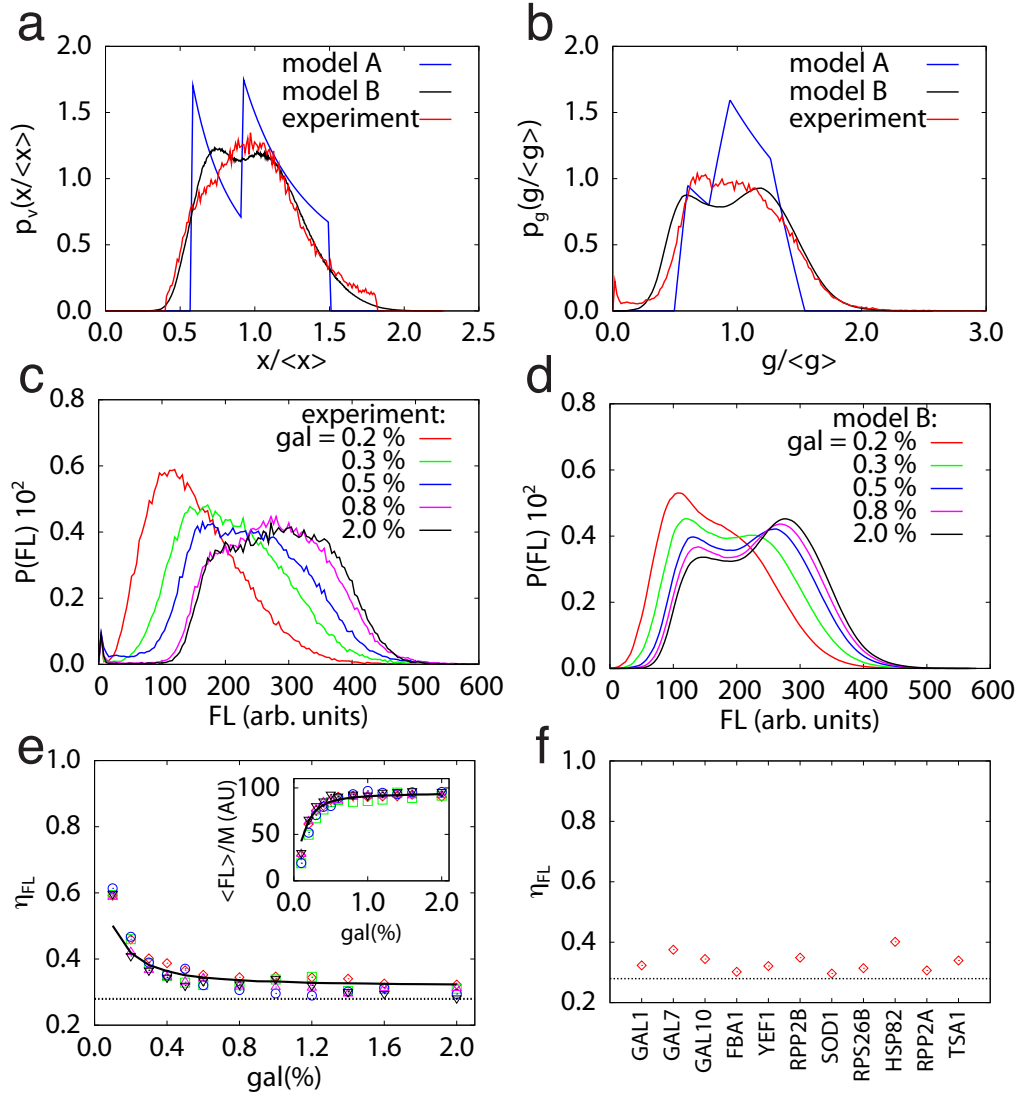
Our analysis of the experimental data provides an important constraint when attempting to deduce the biological origin of the variability, since a nontrivial test of any model description is that it must exhibit correct copy number scaling. We first construct a fully deterministic model which couples a mesoscopic description of cell growth and division with a microscopic description of gene expression (Modeling Box: Model A). For colony growth, we generalize the classical model introduced by Hartwell and Unger (Hartwell and Unger, 1977), which is known to capture the major experimental observations regarding the cell size distributions. We assume that each cell grows exponentially at a rate  $\alpha$  and divides after reaching the same size  $x_0$ . The division is asymmetric so that the size of a mother cell after division is  $\xi/\kappa$  larger than that of a daughter cell ( $\xi + \kappa = 1$ ). After a transient, such a colony reaches a stationary distribution of sizes with two characteristic peaks corresponding to subpopulations of daughter

and mother cells (see Fig. 3.2a where the theoretical distribution is compared with experimental data obtained from the flow cytometry forward scatter data).

In addition to the process of growth and division, Model A assumes that each cell is producing GFP at a constant rate  $\gamma$ , and GFP is distributed at division in the same proportion as the volume. The joint probability of finding a cell of a given size  $x$ , and with GFP content  $g$ , has the functional form  $\mathcal{P}_{vg}(x, g) = \mathcal{P}_v(x)\mathcal{P}_g(g/a, \kappa)$ , where  $a = \gamma/\alpha$  is the ratio of the GFP production rate and volume growth rate. Using this result, we then derive an explicit expression for the distribution  $\mathcal{P}_g(g/a, \kappa)$  (see Fig. 3.2b), and from there generate expressions for the mean and standard deviation of the stationary distribution of GFP,  $\langle g \rangle = a$  and  $\sigma_g = a\eta_g(\kappa)$ , where  $\eta_g(\kappa) = \left[1 - \frac{2}{3\kappa\xi} (\kappa^2 \ln(\kappa)^2 + \kappa\xi \ln(\kappa) \ln(\xi) + \xi^2 \ln(\xi)^2)\right]^{1/2}$ .

The model results for the mean and standard deviation have nontrivial implications consistent with our experimental findings. Since the rate of GFP production  $\gamma = a\alpha$  is proportional to the copy number  $M$ , the functional form  $\mathcal{P}_g(g/a, \kappa)$  implies that distributions of GFP for different copy numbers  $M$  should collapse after the transformation  $g \rightarrow g/M$ ,  $\mathcal{P}_g \rightarrow M\mathcal{P}_g$ . In particular, since the mean value of the GFP distribution is proportional to the copy number and the standard deviation is proportional to the mean, the CV is independent of  $M$  as observed in the experiments. In addition, the model yields an explicit dependence of the CV on the partition ratio  $\kappa$ , and using a previously reported value for this parameter (Woldringh et al., 1993) we obtain a CV which is in close quantitative agreement with our experiments for high concentrations of galactose (Fig. 3.2e, straight line). Importantly, this result was obtained without any additional fitting (in this model, the CV is shown to be a function of  $\kappa$  only). As an additional confirmation of the role of the population dynamics in genetic variability we calculated the coefficient of variation in gated subpopulations of cells within narrow windows of sizes and found that gating reduces the CV by as much as 50%.

Since the variability in the model arises from the coupling of population effects and deterministic gene expression, an interesting model prediction is that there is a “noise floor” that provides the lower limit for expression variability. If there are no other



**Figure 3.2:** Comparison of Models A & B with experimental findings. **a**, Distributions of cells sizes. **b**, Distributions of GFP at fixed galactose concentration and  $M = 1$ . **c**, Distributions of GFP for  $M = 3$  and for a number of galactose concentrations. **d**, The same as **c** but from Model B. **e**, Collapse of the induction curves (symbols) along with the results of simulations of Model B (black solid line). **f**, Validation of model predictions. The model noise floor (dashed black line) along with experiments utilizing different fusions (symbols).



sources of variability, this implies that for various highly expressing genes the observed variability should be similar. We tested this prediction by constructing GFP fusions to 10 highly expressed *Saccharomyces cerevisiae* genes. As predicted, we observed a near-constant level of variability for all of the constructed fusions (Fig. 3.2f). This level is consistent with the lower noise floor observed in the GAL system (Fig. 3.2e), and can be taken as the lower limit for the variability in a growing population of cells. These results suggest that for high expression levels population effects dominate and gene expression within a single cell is mostly deterministic.

While our fully deterministic Model A accounts for the majority of the observed variations in the inducible GAL1-GFP system, there are still certain limitations of this simple model. In particular, the theoretical GFP distribution shows a large peak that is not present in the experimental data (Fig. 3.2b). It occurs since Model A ignores the stochastic fluctuations in the cell size at division, as well as the binomial nature of division of GFP among mother and daughter cells, and thus the GFP distribution is narrow. Furthermore, Model A does not account for the increase of variability at small galactose values. We next generalize the model to describe the variability at small galactose values (Model B). In this regime, the variations increase as the number of GFP molecules decreases. While this dependence is consistent with an intrinsic noise source, the scaling of the CV with copy number implies an extrinsic source and thus fluctuations in expression from uncorrelated *GALI* promoters cannot be the origin. We therefore look upstream of the *GALI* promoter. The *GALI* promoter is activated by GAL4p protein dimers, so a common stochastic component of gene expression may originate from fluctuations in the level of the GAL4p activators. We tested the idea of activator mediated fluctuations with a quantitative model incorporating the galactose-dependent degradation of GAL4p (Figs. 3.2a-c). The generalized model shows excellent agreement for the distributions and scaling of the CV as a function of galactose. In addition, Model B predicts that other GAL4p-regulated promoters should exhibit extrinsic scaling at low galactose, and we present experimental validation of this prediction in the Materials and Methods section below.

We have shown how experimental classification of variability can be used to constrain the space of possible models, and lead to the development of a model incorporating variations arising from the coupling of population growth and gene expression. While this dominant source of variability is likely to exist in all systems, the generalization of our model to include upstream activators of the *GAL1* promoter is system specific. It would be interesting to explore other inducible eukaryotic systems to see if the scaling of the variability also implies an extrinsic source for low induction levels. In the context of the reported correlations observed in the transcripts of proximally located genes (Cohen et al., 2000), our results suggest that the biological origin of these correlations could be a common upstream transcriptional regulator.

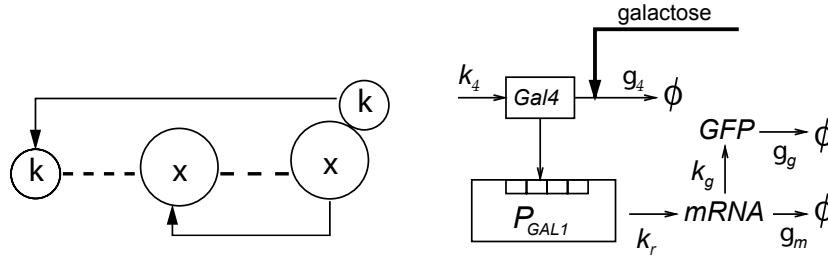
### 3.3 Computational Model

Our two models are constructed by coupling a subsystem for gene expression from the *GAL1* promoter to a subsystem representing the growth and division of individual cells in the population. For the gene expression subsystem, we introduce a simple scheme analogous to the previously reported model (Blake et al., 2003). In our model each of the  $M$  identical promoters may be in either inactive state ( $O_1$ ) or in active state ( $O_2$ ). The latter represents preinitiation complex with all the components necessary for transcription in place. The production of *mRNA* proceeds from state  $O_2$  with rate  $k_m$ . After *mRNA* has been produced, translation ensues with rate  $k_g$ . Both *mRNA* and *GFP* are allowed to decay at rates  $\gamma_m$  and  $\gamma_g$ , respectively. The effect of the inducer is modeled by taking into account the presence of GAL4p proteins which are responsible for the activation of  $P_{GAL1}$  promoter. These proteins dimerize and bind to four binding sites of  $P_{GAL1}$  activating the process of transcription. When multiple copies of the promoter-gene pair are inserted, they share the pools of *GAL4p*, *mRNA*, and *GFP*. We assume that GAL4p monomers are constitutively produced with rate  $\alpha_4$ . Another protein (GAL80p) binds to GAL4p and renders it unable to activate the promoter. We model this by taking the degradation rate of GAL4p ( $\gamma_4$ ) to be inversely proportional

to the galactose concentration. Combining these assumptions, the differential equations describing the deterministic model for gene expression may be deduced from the chemical reactions. The rate equations for  $GAL4p$ ,  $mRNA$ , and  $GFP$ , respectively, read

$$\dot{g}_4 = \alpha_4 - \gamma_4 g_4, \quad \dot{r} = M k_r H(g_4) - \gamma_r r, \quad \dot{g} = k_g r - \gamma_g g$$

where  $H(g_4) = K g_4^z / (1 + K g_4^z)$  is the Hill function for the activation process and we use the cooperativity  $z = 8$  which stems from the binding of four  $GAL4p$  dimers;  $K$  is a lumped parameter which describes the net effect of dimerization/dissociation and binding/unbinding; the dependence on the galactose concentration arises from  $\gamma_4 \propto gal^{-1}$ .



**Figure 3.3:** Illustrations of cell growth and galactose regulatory network.

**Model A:** Our primary motivation is to elicit the extrinsic variability arising from cell growth and division, so we first develop a deterministic model for the gene expression component. The characteristic timescale here is set by the growth rate of the cells. The stability of GFP implies a negligible degradation rate  $\gamma_g$ , and since the rates of degradation of  $GAL4p$  and  $mRNA$  are fast compared with the cellular growth rate, we can eliminate the dynamical equations for  $g_4$  and  $m$ . Utilizing these considerations, the dynamics of GFP is then reduced to production at a constant rate,  $\dot{g} = \Gamma = M [k_g k_r H(\alpha_4 / \gamma_4) / \gamma_m]$ .

We next couple a deterministic model for cellular growth and division with the deterministic production of GFP. To account for population dynamics we consider an asynchronous population of cells, each growing exponentially with rate  $\alpha$  and dividing

in an asymmetric fashion (Hartwell and Unger, 1977) after reaching a certain maximum size. We let the size of a mother cell after division be  $\xi/\kappa$  times larger than that of the daughter cell, where  $\kappa < 1/2$  ( $\xi + \kappa = 1$ ). The process of growth and division repeats indefinitely, producing a growing structured population of cells of multiple generations where each cell is described by a position within the cell cycle and genealogical age. The coupling between this growth model and gene expression arises through the partitioning of GFP content at division. The system of evolution equations describing the size ( $x$ ) and number of GFP molecules ( $g$ ) is given by

$$\dot{x} = \alpha x, \quad \dot{g} = \Gamma,$$

which should be supplemented by the renewal equations (division rules), according to  $x_0 \rightarrow \kappa x_0 + \xi x_0$ ,  $g(x_0) \rightarrow \kappa g(x_0) + \xi g(x_0)$  where  $x_0, g(x_0)$  are the cell size and GFP content at the moment of division respectively. The additional assumptions behind this model may be summarized as follows. We distinguish only two generations of cells, daughters and mothers, thus neglecting the difference between mothers of second generation and higher. We assume that cells are growing with the same rate regardless of their age, and that both daughters and mothers divide after reaching the same size and in the same proportion. Finally, we assume GFP is distributed among daughter cells in the same proportion as the size (fast diffusion of GFP molecules). Overall, the statistics of the population is modeled in terms of the equation for the joint probability  $\mathcal{P}_{vg}(x, g)$  to find a cell of a given size  $x$  and with a certain amount of GFP  $g$ , which is solved analytically.

**Model B** generalizes Model A by including several realistic features of both gene expression and population dynamics. For gene expression we utilize a hybrid approach whereby production and degradation of *GAL4p* proteins is modeled stochastically using the Gillespie algorithm (Gillespie, 1977), while production of *mRNA* and *GFP* is simulated with the set of ordinary differential equations introduced above. Consistent with our experimental observations, our hybrid approach eliminates all intrinsic fluctuations except for those induced by *GAL4p* monomers. This intrinsic noise is then

transferred to the level of  $mRNA$  production through fluctuations in the production rate of  $mRNA$ , which is proportional to the Hill function  $H(g_4)$ . Here the intrinsic noise manifests itself as extrinsic fluctuations common to all  $M$  copies. For the population dynamics we introduce genealogical age by utilizing different values of  $\kappa$  depending on the generation (Woldringh et al., 1993), and introduce a stochastic component in division by drawing a cell size at division from a narrow gaussian distribution centered at  $x_0$ . In addition, upon division the amounts of  $GAL4p$ ,  $mRNA$  and  $GFP$  are distributed among the offspring in a binomial fashion.

## 3.4 Materials and Methods

### 3.4.1 Plasmids, Yeast Strains and Growth Conditions

All plasmid backbones were derived from pRS403, pRS404, pRS405, or pRS406 shuttle vectors (Stratagene). The *GALI-10* promoter region and the *yEGFP* was obtained from pESC1-yG. The first plasmid was constructed by combining the integrative feature of pRS405 (Stratagene), and the *GALI-yEGFP* portion of pESC1-yG. Each plasmid was digested using the restriction enzyme PvuII and the pertinent pieces were ligated together to make pRS51-yG. Plasmids pRS31-yG, pRS41-yG, and pRS61-yG differ from pRS51-yG in their yeast selective markers (see Table 3.1). All plasmids contain an ampicillin resistance marker and the ColEI replication origin. All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. Plasmids were introduced into *Escherichia coli* XL10-Gold (Stratagene) by using a standard heat shock transformation protocol. All bacterial cells were grown in LB (Fisher) with 100  $\mu\text{g/mL}$  of ampicillin (Sigma). All yeast strains (see Table 3.2) in this study are derived from the *Saccharomyces cerevisiae* parent strain YPH500 ( $\alpha$ , *ura3-52*, *lys2-801*, *ade2-101*, *trp1 $\Delta$ 63*, *his3 $\Delta$ 200*, *leu2 $\Delta$ 1*) (Stratagene). The strains were created by targeted chromosomal integration of shuttle vector constructs at either the *GALI-10* locus on chromosome II or the *TRP1* locus on chromosome IV. All fusion strains were con-

structed as described by Sheff et al. (Sheff and Thorn, 2004a). The yeast strains were grown in synthetic drop-out (SD) medium supplemented for selection of correct integrands containing 2% glucose at 30° C. The strains were selected for the proper number of promoter-gene pairs by growing the cells on 2% galactose and assaying the cells for average fluorescence. All cloning steps were performed in *Escherichia coli* XL10-Gold (Stratagene).

**Table 3.1:** Summary of integration plasmids.

Plasmid Name	Description
pRS31-yG	integrating vector with <i>GALI</i> promoter driving <i>yEGFP</i> ; his3 marker
pRS41-yG	integrating vector with <i>GALI</i> promoter driving <i>yEGFP</i> ; trp1 marker
pRS51-YG	integrating vector with <i>GALI</i> promoter driving <i>yEGFP</i> ; leu2 marker
pRS61-yG	integrating vector with <i>GALI</i> promoter driving <i>yEGFP</i> ; ura3 marker

**Table 3.2:** Summary of yeast strains.

Strain	Description
YPH500	WT strain
JYM101	pRS31-yG integrated at <i>GALI-10</i> locus
JYM113	pRS31-yG and pRS41-yG integrated at <i>GALI-10</i> locus
JYM154	pRS31-yG, pRS51-yG, and pRS61-yG integrated at <i>GALI-10</i> locus
JYM198	pRS31-yG, pRS41-yG, pRS51-yG, and pRS61-yG integrated at <i>GALI-10</i> locus
JYM199	pRS31-yG, pRS41-yG, pRS51-yG, and pRS61-yG integrated at <i>GALI-10</i> locus
JYM203	pRS41-yG integrated at TRP1 locus
JYM213	pRS31-yG integrated at <i>GALI-10</i> locus; pRS41-yG integrated at TRP1 locus
FUS001	<i>yEGFP</i> fused to <i>FBA1</i>
FUS002	<i>yEGFP</i> fused to <i>YEF3</i>
FUS003	<i>yEGFP</i> fused to <i>RPP2B</i>
FUS004	<i>yEGFP</i> fused to <i>SOD1</i>
FUS005	<i>yEGFP</i> fused to <i>RPS26B</i>
FUS006	<i>yEGFP</i> fused to <i>HSP82</i>
FUS007	<i>yEGFP</i> fused to <i>RPP2A</i>
FUS008	<i>yEGFP</i> fused to <i>TSA1</i>
FUS009	<i>yEGFP</i> fused to <i>GALI</i>
FUS010	<i>yEGFP</i> fused to <i>GAL7</i>
FUS011	<i>yEGFP</i> fused to <i>GAL10</i>

### 3.4.2 Gene Expression Experiments

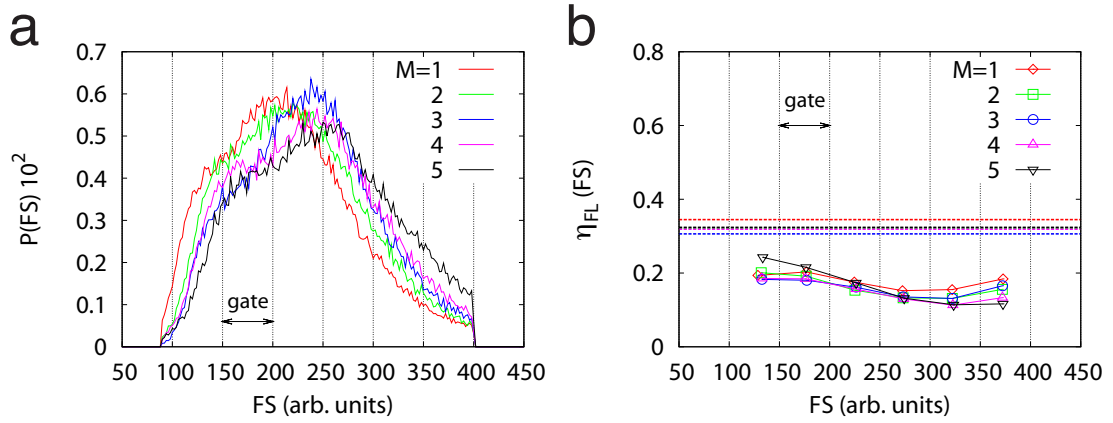
Exponentially growing yeast cells were diluted from 1:40 to 1:70 into medium containing 2% raffinose and a range of galactose concentrations (0.1-2.0% galactose) as inducer. After 14-18 hours, the cells were assayed at mid-log phase at an  $OD_{600}$  of  $0.55 \pm 0.15$ . Expression data were collected using a Becton-Dickinson FACSCalibur flow cytometer with a 488-nm argon excitation laser and a 515- to 545-nm emission filter (FL1) at a low flow rate. Forward scatter values and fluorescence values were collected for 100,000 cells. The standard list-mode files obtained from the flow cytometer were converted to ASCII format with MFI (E. Martz, University of Massachusetts, Amherst; <http://www.umass.edu/microbio/mfi>) and analyzed using Matlab (The MathWorks, Inc., Natick, Massachusetts).

### 3.4.3 Variability of gated subpopulations

We calculated coefficient of variation for subpopulations of cells within narrow windows of cell sizes separated in Fig. 3.4a by thin vertical lines. As seen in Fig. 3.4b, the coefficient of variation for subpopulations with small distribution of cell sizes (as measured by the forward scatter in flow cytometer) is up to 50% smaller than that for the whole population. This observation implied that cell growth and division should play an important role in the observed variability of gene expression.

### 3.4.4 Modeling studies

Model A was solved analytically. For the comparison with the experiment we used an effective ratio of the volumes of daughter after division to mother before the division,  $\kappa = 0.34$ , representing the weighted average of  $\kappa$  among several generations (Woldringh et al., 1993) (Fig. 3.1g). Model B was solved numerically using a hybrid technique (Adalsteinsson et al., 2004) whereby the stochastic dynamics of GAL4p protein have been modeled with the Direct Gillespie algorithm (Gillespie, 1977), the in-



**Figure 3.4:** Cell size distributions (a) and coefficients of variation for gfp expression measured within narrow gates separated by thin vertical lines, for several values of the copy number  $M$ .

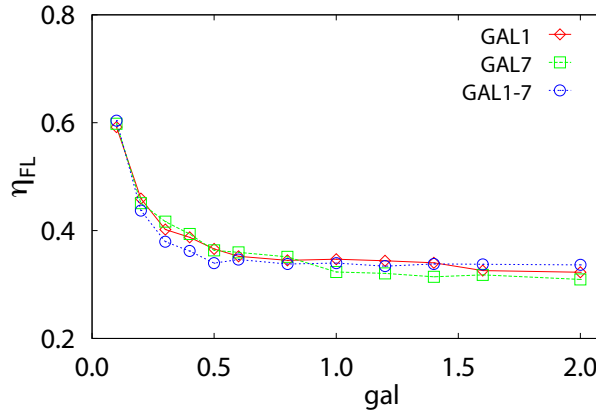
intermediate reactions of dimerization and binding of GAL4p dimers were eliminated employing quasi-stationary assumptions for the sake of simplicity and the dynamics of *mRNA* and *GFP* were simulated using the rate equations presented in Box 1. For the population dynamics we used previously reported measurements regarding the structure of the population of *Saccharomyces cerevisiae* (Woldringh et al., 1993). Namely, for the first generation, we used  $\kappa = 0.4$  and for subsequent generations  $\kappa = 0.3$ , the mean size of the cells at divisions was chosen to be 1.089, 1.179, 1.268, for genealogical ages from 2 to 4 respectively and 1.357 for all older generations; and the size at division for a particular cell was drawn from a narrow gaussian distribution near these mean values with coefficient of variation 0.15. The parameters of the simulations in units of average growth rate are,  $k_r = 80$ ,  $\gamma_m = 4$ ,  $k_g = 4$ ,  $\gamma_g = 0$ ,  $k_4 = 8$ ,  $\gamma_4 = 0.5 gal^{-1}$ ,  $K = 0.05$ . The protocol of the simulations closely resembled the experimental procedure. We started from a small collection of cells (typically  $10^3$ ), with Gaussian distributions of sizes and quantities of GAL4p, mRNA, and GFP near the expected mean values. This small population was “grown” until it reached 100,000 cells. The state of the population was recorded within a narrow time window right after that. We checked that this measurement was consistent with the time-ensemble average over a longer period of time



indicating that the evolution of the population reached the stationary state.

### 3.4.5 GAL4p regulation of other promoters

If the extrinsic variability observed from the *GAL1* promoter at low galactose is due to the GAL4p regulator, an interesting model prediction is that other GAL4p-regulated promoters should also exhibit extrinsic noise at low galactose. We tested this prediction with the construction of two additional strains. In the first strain, we fused GFP to GAL7p, whose promoter is activated by GAL4p. In a second strain, we fused to both GAL7p and to GAL1p. As predicted by the model, we observed scaling which indicates that extrinsic noise dominates the system at low galactose levels (Fig. 3.5). While it is still certainly possible that other upstream factors are responsible for the observed downstream variations, these results provide additional support for our hypothesis that variability in the GAL4p activator is a dominant factor.



**Figure 3.5:** Coefficient of variation as a function of galactose for three promoters regulated by GAL4p. Two strains contain GAL1p and GAL7p fusions, respectively, while a third strain contains both a GAL1p and GAL7p fusion. The collapse is predicted by the model.

### 3.4.6 Fusion Strains

Yeast-enhanced Green Fluorescence Protein (yEGFP) was fused to several yeast genes that were found to be highly expressed during normal growth conditions and three genes of the galactose utilization pathway (see Table 3.3). Fusion strains were created by amplifying pKT128 Sheff and Thorn (2004a) using the Phusion<sup>TM</sup> High-Fidelity DNA Polymerase (Finnzymes) PCR system following the manufacturer's instructions. Each primer consists of 40 base pairs that overlap with the yeast gene of interest in order to encourage homologous recombination. The portion of the plasmid that is amplified contains *yEGFP* and the *SpHIS5* selective marker. The PCR product was introduced into YPH500 using heat shock.

**Table 3.3:** Summary of yeast genes fused to *yEGFP*

Gene	Function
FBA1	Fructose 1,6-bisphosphate aldolase: catalyses the reversible formation of fructose-1,6 bisphosphate from dihydroxyacetone phosphate and glyceraldehyde-3-phosphate.
YEF3	Translational elongation factor 3: stimulates the binding of aminoacyl-tRNA to ribosomes.
RPP2B	Ribosomal protein P2 beta: regulates the accumulation of P1 in the cytoplasm
SOD1	Cu, Zn superoxide dismutase: catalyzes the dismutation of superoxide anion
RPS26B	Ribosomal protein S26B: protein component of the small (40S) ribosomal subunit
HSP82	Heat shock protein 90: a cytoplasmic chaperone that is required for pheromone signaling and negative regulation of Hsf1p
RPP2A	Ribosomal protein P2 alpha: regulates the accumulation of P1 in the cytoplasm
TSA1	Thioredoxin-peroxidase: reduces H <sub>2</sub> O <sub>2</sub> and alkyl hydroperoxides
<i>GAL1</i>	Galactokinase: phosphorylates alpha-D-galactose to alpha-D-galactose-1-phosphate
<i>GAL7</i>	Galactose-1-phosphate uridyl transferase: synthesizes glucose-1-phosphate and UDP-galactose from UDP-D-glucose and alpha-D-galactose-1-phosphate
<i>GAL10</i>	UDP-glucose-4-epimerase: catalyzes the interconversion of UDP-galactose and UDP-D-glucose in galactose metabolism; also catalyzes the conversion of alpha-D-glucose or alpha-D-galactose to their beta-anomers

### **3.4.7 Gene Expression Experiments**

Yeast strains are plated from frozen stock on agar plates with synthetic drop-out (SD) medium supplemented for selection of transformants containing 2% glucose and grown for 40-48 hours. One colony is chosen and grown in liquid SD medium with appropriate supplements and 2% raffinose. Raffinose is used as the primary carbon source to avoid the inhibitory effects of glucose on the galactose-utilization pathway. To prepare cells for the flow cytometer, they are pelleted, resuspended in filtered PBS (Sigma), and vortexed to reduce clumps.

## **3.5 Acknowledgements**

Chapter 3 contains material originally published as Volfson, D.\*, Marciniak, J.\*, Ostroff, N., Blake, W., Tsimring, L., and Hasty, J., 2006: Origins of extrinsic variability in eukaryotic gene expression. *Nature*, 439: 861-864. (\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

# Chapter 4

## Metabolic gene regulation in a dynamically changing environment

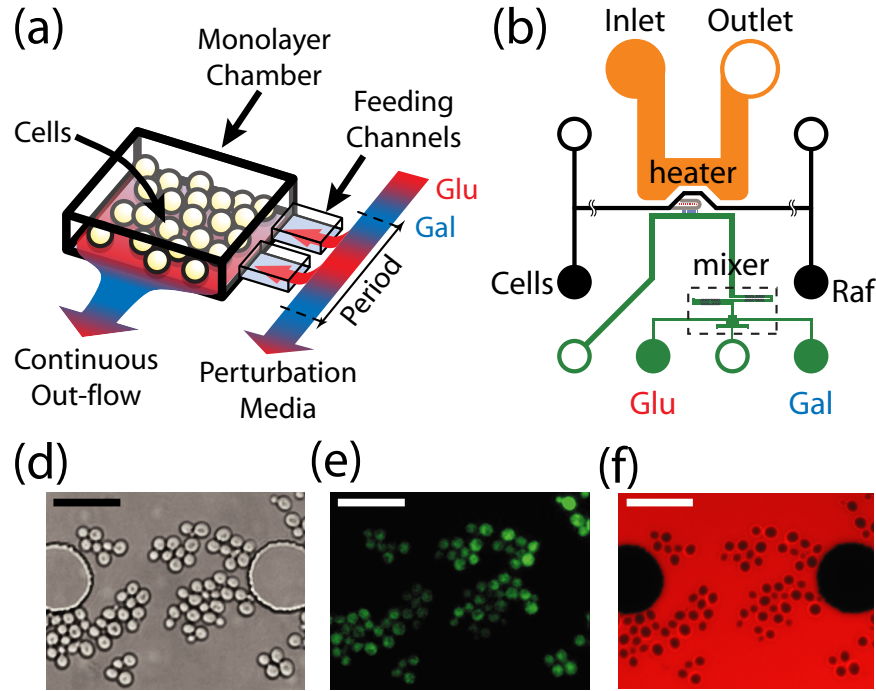
### 4.1 Introduction

Natural selection dictates that cells constantly adapt to dynamically changing environments in a context-dependent manner. Gene-regulatory networks often mediate the cellular response to perturbation (Beadle and Tatum, 1941; Jacob and Monod, 1961; Douglas and Hawthorne, 1966), and an understanding of cellular adaptation will require experimental approaches aimed at subjecting cells to a dynamic environment that mimics their natural habitat (Thattai and Shraiman, 2003; Lipan and Wong, 2005; Kussell and Leibler, 2005; Kruse and Julicher, 2005; Ronen and Botstein, 2006; Thattai and van Oudenaarden, 2004). Here, we monitor the response of *S. cerevisiae* metabolic gene regulation to periodic changes in the external carbon source by utilizing a microfluidic platform that allows precise, dynamic control over environmental conditions. We find that the metabolic system acts as a low-pass filter that reliably responds to a slowly changing environment, while effectively ignoring fluctuations that are too fast for the cell to mount an efficient response. We use computational modeling calibrated with experimental data to determine how frequency selection in the system is controlled by

the interaction of coupled regulatory networks governing the signal transduction of alternative carbon sources. Experimental verification of model predictions leads to the discovery of two novel properties of the regulatory network. First, we reveal a previously unknown mechanism for post-transcriptional control, by demonstrating that two key transcripts are degraded at a rate that depends on the carbon source. Second, we compare two *S. cerevisiae* strains and find that they exhibit the same frequency response despite having markedly different induction characteristics. Our results suggest that while certain characteristics of the complex networks may differ when probed in a static environment, the system has been optimized for a robust response to a dynamically changing environment. Importantly, the integration of a novel experimental platform with numerical simulations revealed previously masked network properties, and the approach establishes a framework for dynamically probing organisms in order to reveal mechanisms that have evolved to mediate cellular responses to unpredictable environments.

## 4.2 Results and Discussion

In order to probe the response of a metabolic gene network to a fluctuating environment, we developed a microfluidic platform which can subject a population of cells to a continuously varying media supply (Fig. 4.1). The device is designed to generate a fluctuating media signal by dynamically combining two media reservoirs according to a time-dependent function. Feeding channels deliver the media downstream to a customizable growth chamber, which for this study was constructed to constrain a population of yeast cells to grow in a monolayer, allowing for long-term data acquisition (Cookson et al., 2005). The composition of the media is dynamically controlled by a fluidic switch (Groisman et al., 2005b), such that changes in the upstream source may be detected almost immediately by the cells. The fluidic switch was optimized to generate a linear range of mixing ratios from the two media inputs, allowing a variety of periodic waveforms or random signals to be generated.



**Figure 4.1:** Design and implementation of the microfluidic platform developed for our study. **(a)** Conceptual design of the imaging chamber. The chamber is coupled to the switch output channel via multiple  $1\ \mu\text{m}$  tall “feeding” channels. The feeding channels are fed by a controllable wave-form generator that creates sinusoidal perturbations in the glucose concentration while maintaining constant background levels of galactose. **(b)** An overview of the design shows the layout of the device. The device makes use of three flow networks for (1) loading cells (middle, black), (2) generating microenvironmental waveforms (bottom, green), (3) and controlling on-chip temperature (top, orange). The imaging chamber (center, gray region) is designed to be about  $4\ \mu\text{m}$  tall in order to constrain a population of yeast cells to grow in a monolayer. **(c)** Representative brightfield image of cells growing in the imaging chamber. These images are used to measure the total size of the colony. Scale bar is  $25\ \mu\text{m}$  in length, and large circles are support posts in the chamber. **(d)** Green fluorescence image of the same cells as in (c). These images allow us to measure the amount of Gal1p in each cell. **(e)** Red fluorescence image of the chamber. The glucose media also contains a red fluorescent dye, and therefore the intensity of the red fluorescence is proportional to the amount of glucose in the chamber at any given time.

As a quantifiable reporter of the cellular response to environmental fluctuations, we fused the native Gal1p protein of *S. cerevisiae* to the yeast-optimized enhanced cyan fluorescent protein (yECFP) (Sheff and Thorn, 2004b; Raser and O'Shea, 2004). The enzymes for galactose utilization, including Gal1p, are among the most tightly regulated proteins in yeast. Because glucose requires much less energy to metabolize, cells will only consume galactose if glucose is not available. Therefore, *S. cerevisiae* has evolved a highly complex regulatory network to ensure that the galactose enzymes will be strongly activated when they are needed, but tightly repressed if glucose is present in the environment (Fig. 4.2(a)). Because the network is well studied and involves regulatory motifs common to many higher organisms, galactose utilization is a paradigm for gene regulation. In order to build on the current understanding of its robust regulatory mechanisms, we employed our microfluidic platform to monitor the dynamics of network activation and repression in response to sinusoidal perturbations of glucose over a galactose background.

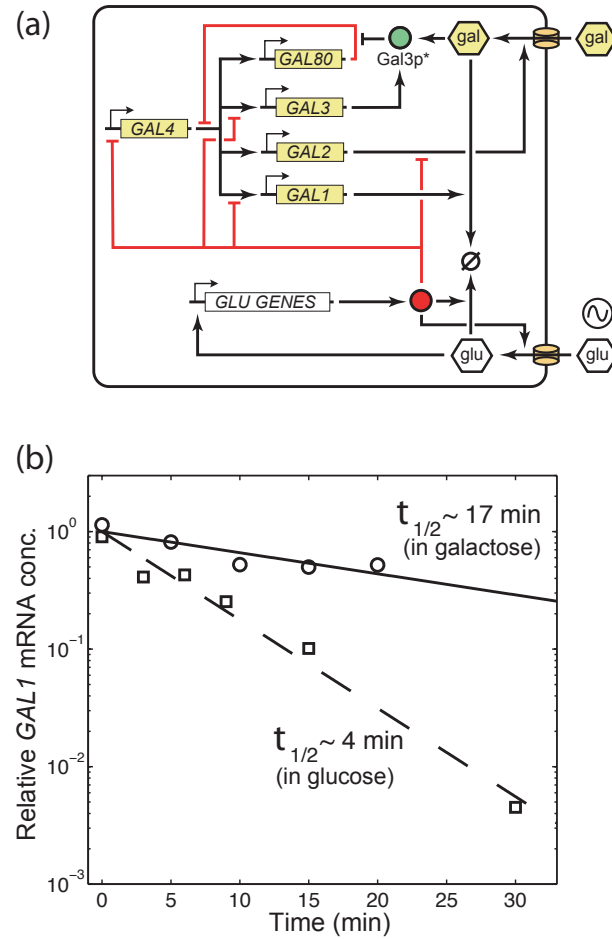
A population of yeast cells was subjected to sinusoidal glucose waves over a 0.2% (w/v) galactose background, with varying glucose concentration from 0.0% (no repression of *GAL1* transcription) to 0.25%. For each run we changed the frequency of the glucose signal, varying the period from 0.75 to 4.5 hr, and we imaged the population for a minimum of four full cycles. Time-lapse fluorescence imaging of the cell population in the growth chamber was used to calculate the amplitude ratio and phase shift of the cellular response relative to glucose signal. The results show a maximum response frequency of about  $5.6 \text{ rads hr}^{-1}$  (1.125 hr period). At this frequency, the response trace was indistinguishable from a normal step function response, whereas at the lower frequencies the temporal fluorescence trajectories clearly oscillated in response to the signal. In this sense, the galactose system appears to function as a low-pass filter that reliably responds to a slowly changing environment, while effectively ignoring fluctuations that are too fast for the cell to mount an efficient response.

Since the sinusoidal driving of the galactose utilization network leads to complex cellular behavior, we used computational modeling to simulate the response and uncover

key aspects of the network architecture that give rise to the observed behavior (Hasty et al., 2001b). In particular, we were interested in how the interplay of the galactose and glucose utilization networks gives rise to the observed frequency response to carbon source fluctuations. By itself, the turnover of Gal1p-yECFP, either due to dilution or active degradation (or both), leads to low pass filtering of periodic signals. However, feedback loops inherent to gene regulatory networks can alter the response of proteins to stimuli (Savageau, 1974). Therefore, in order to simulate the effects of galactose activation and glucose repression on our experimental data, we adapted a comprehensive model of the galactose network originally described by deAtauri et al. (de Atauri et al., 2004). This model includes the transcription and translation of the *GAL1*, *GAL2*, *GAL3*, *GAL4* and *GAL80* genes as well as the interactions of their respective proteins with each other and galactose (such as dimerization, transport and metabolism). Whenever possible we used parameter values either at or close to the values reported by deAtauri et al. In addition to this galactose network model, it was necessary to model the dynamics of the glucose network. The glucose network is much more complex than that of galactose (Demir and Kurnaz, 2006; Kaniak et al., 2004; Verma et al., 2005) and models for it are much less well established. Therefore, we chose to model the glucose network with a simplified module describing a basic transport regulatory system. In it, protein products of the glucose network are responsible for transporting external glucose into the cell while internalized glucose acts to induce transcription in the network, giving rise to a positive feedback loop (see Fig. 4.2(a)).

Calibration of the computational model to the experimental data led to several important observations that would not have arisen from an analysis of steady-state batch culture data. The large amplitude ratios observed at low frequencies suggested that when glucose was added to the system the degradation rates of some galactose network components were greater than in the absence of glucose. Previous studies have suggested that components of the glucose network can actively degrade mRNA produced by genes involved in the galactose/glucose switch (Ronen and Botstein, 2006), and such a phenomenon has also been shown to exist for the mRNA of other genes (Scheffler et al.,

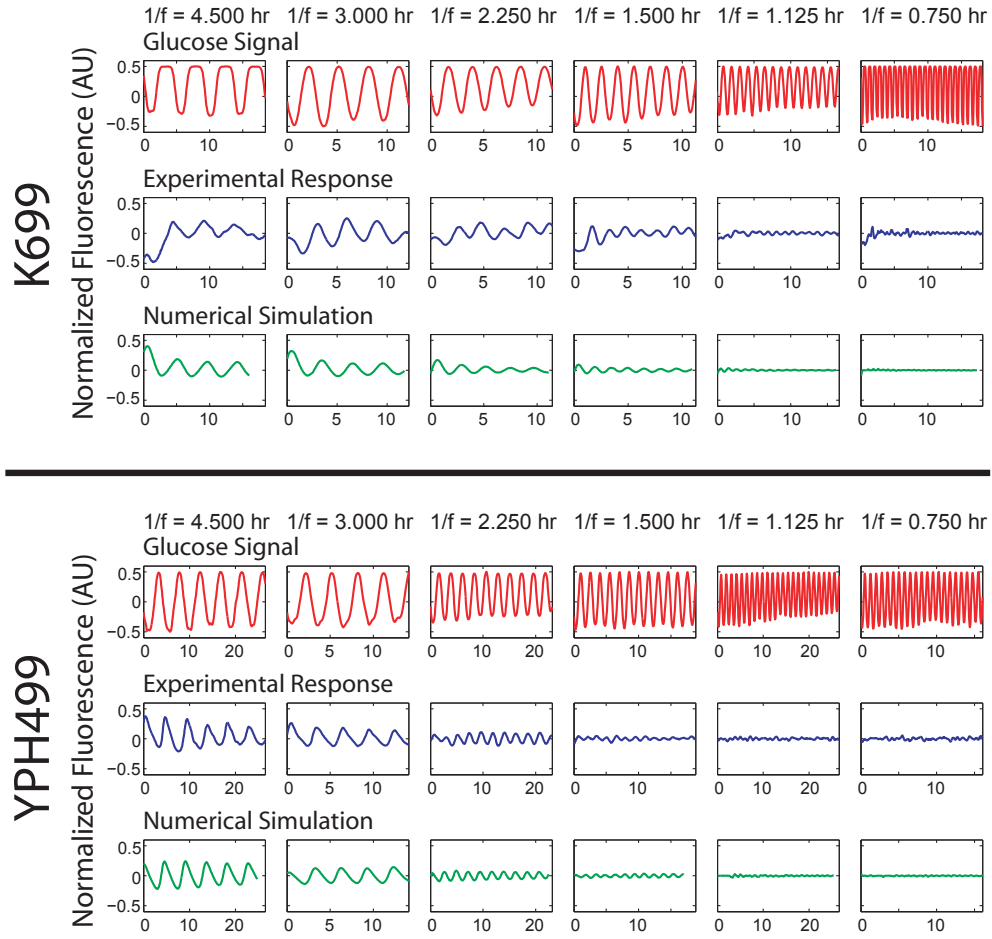




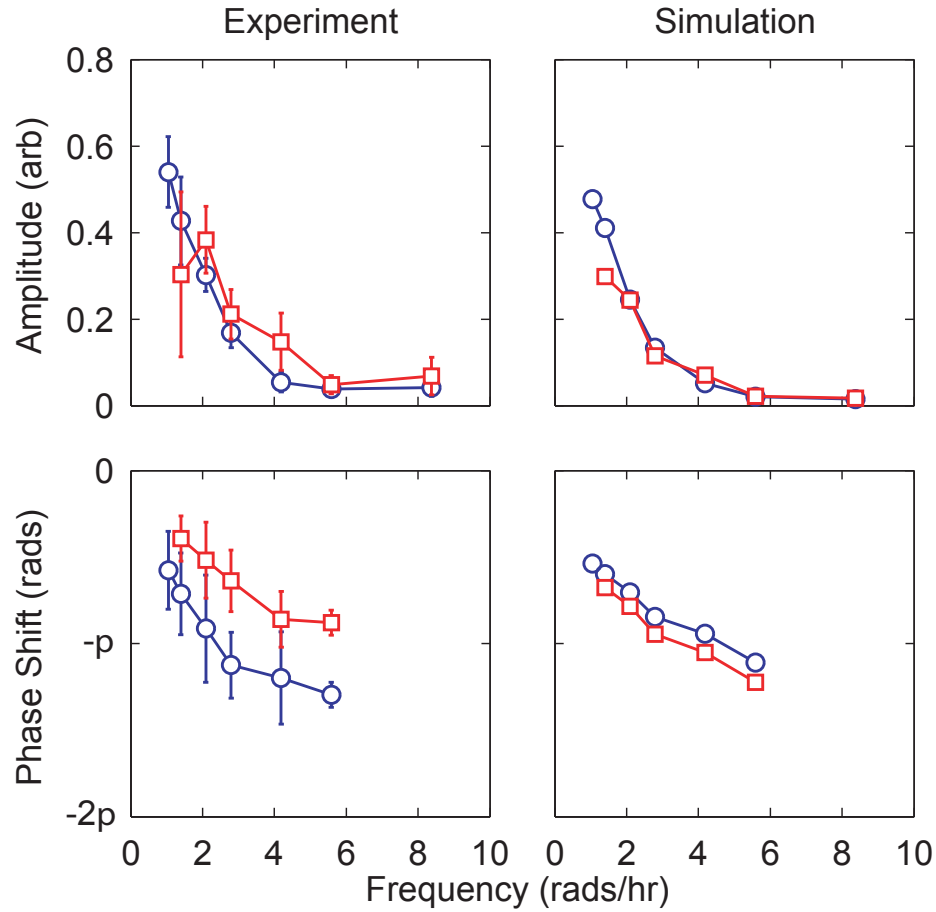
**Figure 4.2:** Regulation in the galactose utilization network (a) Schematic of the gene regulatory networks involved. The regulatory genes in the galactose network are activated by the Gal4p protein which binds to upstream activation sites. The *GAL80* gene provides negative feedback in the system by prohibiting the inducing affects of Gal4p. Positive feedback is provided by both *GAL2* and *GAL3*. Internalized galactose can bind to Gal3p and the resulting complex binds to Gal80p. Gal80p bound to the Gal3p-galactose complex is incapable of repressing Gal4p. Also, the transporter Gal2p increases the amount of internal galactose which stimulates the galactose network. The glucose network inhibits the transport of galactose and represses transcription of the galactose network in the presence of glucose through the action of Mig1p, which can bind to upstream regulatory sites of *GAL1*, *GAL3* and *GAL4* (Verma et al., 2005). The glucose network also regulates the hexose transporter genes (*HXT*) which are responsible for transporting glucose into the cell (Boles and Hollenberg, 1997), which then activates the glucose network. (b) Experimentally measured decay of *GAL1* transcripts in galactose (circles) and glucose (squares). Also shown are the best-fit lines corresponding to half-lives of around 17 min in galactose (solid line) and 4 min in glucose (dashed line), similar to the values predicted by the numerical model. Data is normalized to the initial concentration of mRNA predicted by the best-fit lines.

1998; de la Cruz et al., 2002; Andrade et al., 2005). Therefore, we added enzymatic decay terms (governed by Michaelis-Menten dynamics) to the equations describing the dynamics of the *GAL1* and *GAL3* mRNA and found that it greatly increased the accuracy of the model. These two genes are among those in the galactose network that are targeted by the glucose induced Mig1p that represses transcription by binding to upstream regulatory sites (Verma et al., 2005). Thus, if proteins from the glucose network do actively degrade galactose network transcripts, *GAL1* and *GAL3* are likely targets. To test this prediction, we measured the degradation rates of *GAL1* and *GAL3* in both galactose and glucose. Both transcripts showed a 2-8 fold increase in their decay rates when in the presence of glucose (see Fig. 4.2(b)), consistent with the values predicted by the computational model. This form of post-transcriptional regulation, in which glucose acts to down-regulate *GAL* protein synthesis, is a previously unknown source of regulation in the galactose utilization network. Furthermore, the inclusion of glucose mediated mRNA decay results in a model that accurately reproduces the dynamic response of a population of cells to sinusoidal repression over a large range of frequencies (Fig. 4.3).

Batch-culture induction characteristics for metabolic genes can vary from strain to strain or depend sensitively on the growth state of the culture. Therefore, we were also interested in using the model to determine how galactose induction differences would affect the response to the glucose fluctuations. The model demonstrated that significantly different galactose induction does not necessarily lead to significant differences in the response characteristics (data not shown). In other words, the model led to the hypothesis that deficiencies in network induction capabilities might not hinder a cell's ability to adapt and thrive in a changing environment. The yeast strain used to collect this data, K699, is sensitive to external galactose concentrations, with full induction of the galactose network occurring around 0.05% (w/v) galactose. In order to test our hypothesis, we turned to a strain (YPH499) that is known to have a deficiency in the galactose utilization network, which causes it to require more galactose than "normal" to induce production of the galactose enzymes (Rohde et al., 2000). YPH499 is a derivative of a *GAL2* mutant strain, and while the mutations were reportedly repaired, the *GAL2* alleles



**Figure 4.3:** Experimental and computational results for cells expressing a *GALI-yECFP* fusion gene in response to alternating glucose and galactose media for strains K699 and YPH499. The top row of each strain depicts the input glucose signal measured during each experimental run and also used to simulate the responses. The mean fluorescence of a red tracer dye, representing the glucose concentration in the media, is normalized and subtracted from 1 to represent the “induction” signal used in the experimental and computational runs above. The middle rows show normalized and detrended fluorescence trajectories for a population of cells as they respond to glucose waves of various frequencies over a galactose background. In the absence of glucose, galactose induces transcription of *GALI-yECFP* causing an increase in cellular fluorescence. However, as glucose is introduced into the extracellular environment, transcription of the galactose enzymes is shut off, causing a decrease in fluorescence signal as the Gal1p-yECFP protein is degraded. Oscillation periods shown from left to right are 4.5, 3.0, 2.25, 1.5, 1.125, and 0.75 hr. For input waves with a period shorter than 1.125 hr, cells no longer respond to sinusoidal repression in a periodic fashion, demonstrating their ability to “filter” out high frequency environmental fluctuations. The bottom rows show simulation results for the same frequencies as above. The model, calibrated to experimental induction and repression data, accurately reproduces the cellular responses over a large range of frequencies.



**Figure 4.4:** Experimental and computational comparison of two yeast strains, one of which (YPH499) is known to have a deficiency in the galactose utilization network. Amplitude (top row) and phase shift (bottom row) of the response of cells to sinusoidal repression at various frequencies are shown for both K699 (red) and YPH499 (blue) strains. For the highest frequency trial, reliable phases could not be calculated at all due to the noise, and have been omitted from the graphs. The experimental data (left column) demonstrate that the amplitude responses of the two strains are strikingly similar, especially considering their significantly different induction curves. This phenomenon was predicted by model simulations, as slight modifications to the model parameters that affected induction and repression curves did not affect the cell population's robust response to a dynamic environment. This suggests that the complex structure of the glucose and galactose networks may confer robustness to cells even if faced with seemingly detrimental network deficiencies. The phase responses (bottom row) of the two strains did show a marked difference, with YPH499 cells exhibiting a greater phase lag than K699.

in many of the derivative strains have been shown to cause significantly impaired galactose uptake (Rohde et al., 2000). The Gal2p protein is responsible for the transport of extracellular galactose into the cell and its activity is markedly different in YPH499 than in K699. Our flow cytometry population data demonstrated that YPH499 cells require about ten times more galactose to reach full induction than do K699 cells.

Despite the difference in induction sensitivity between K699 and YPH499, our model predicted that inefficient Gal2p transport does not translate into a less robust response to a fluctuating environment. This suggests that the complex interplay of the glucose and galactose networks may confer robustness to cells even if faced with deficiencies in the induction characteristics. To validate this finding, we repeated the microfluidic runs at each frequency, this time using the YPH499 strain with a Gal1p-yECFP fusion. As predicted, the amplitude responses of the two strains are strikingly similar (Fig. 4.4), especially considering the significant difference in their galactose sensitivity. We do not at present know the underlying mechanistic property of the regulatory network that leads to the robust response of the two strains. Future studies might endeavor to deduce this mechanism through the systematic deconstruction of the regulatory elements in a single strain. While the present study demonstrates how robustness *can* occur despite large differences in induction characteristics, one could further investigate the generality of this phenomenon by comparing the responses of many different strains to different types of temporal perturbation.

## 4.3 Materials and Methods

### 4.3.1 Dynamic environment experiments

Cells containing a *GAL1-yECFP* fusion were imaged every 5 min for up to 24 hr using time lapse fluorescent microscopy to estimate the concentration of GAL1p as a function of time. Cells were constrained to grow in a custom designed microfluidic platform which allows for dynamically controlled mixing of two growth media. Here,

our inducing media contained 2% raffinose + 0.2% galactose while the repressing media contained 2% raffinose + 0.2% galactose + 0.25% glucose. The resulting images were processed with cell segmentation and tracking software, and the population averaged fluorescence concentrations were measured. To ensure correct waveform generation, glucose concentrations were monitored by introducing a red fluorescent tracer dye (0.01 mg/mL, sulforhodamine 101) to the repressing media.

### 4.3.2 Microfluidic chips and waveform generation

The PDMS (polydimethylsiloxane) microfluidic devices were designed to allow for monolayer growth of yeast cells in the imaging chamber and were fabricated using standard replica molding techniques (Whitesides et al., 2001, 2004; Xia and Whitesides, 1998). An upstream fluidic switch controlled the media input into the chamber by mixing the flows of the inducing and repressing media. The mixing ratio of the two media was governed by a software controlled, custom-designed pressurization system able to consistently produce time-varying waveforms.

### 4.3.3 mRNA degradation experiments

The degradation rates of *GAL1* and *GAL3* transcripts were measured using standard rt-qPCR techniques. Knockout strains for both genes were first created, and then ectopic *GAL1* and *GAL3* were placed back into the cell under the control of a doxycycline repressible promoter. mRNA half-lives were measured from cells grown in the presence or absence of glucose.

## 4.4 Acknowledgements

Chapter 4 contains material originally published as Bennett, M.\*, Pang, W.\*, Ostroff, N., Baumgartner, B., Nayak, S., Tsimring, L., and Hasty, J., 2008: Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119-22.

(\*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

# Chapter 5

## Cell cycle dependent variations in protein concentration

### 5.1 Introduction

Computational modeling of biological systems has become an effective tool for analyzing cellular behavior and for elucidating key properties of the intricate networks that underlie experimental observations. While most modeling techniques rely heavily on the concentrations of intracellular molecules, little attention has been paid to tracking and simulating the significant volume fluctuations that occur over each cell division cycle. Here, we use fluorescence microscopy to acquire single cell volume trajectories for a large population of *S. cerevisiae* cells. Using this data, we generate a comprehensive set of statistics that govern the growth and division of these cells over many generations, and we discover several interesting trends in their size, growth, and protein production characteristics. We use these statistics to develop an accurate model of cell cycle volume dynamics, starting at cell birth. Finally, we demonstrate the importance of tracking volume fluctuations by combining cell division dynamics with a minimal gene expression model for a constitutively expressed fluorescent protein. The significant oscillations in the cellular concentration of a stable, highly expressed protein mimic the observed ex-



perimental trajectories and demonstrate the fundamental impact that the cell cycle has on cellular functions.

Synthetic biology has emerged as an important field in the effort to quantitatively understand biological systems (Hasty et al., 2002; Sprinzak and Elowitz, 2005). Bridging the gap between engineering and biology, this broad field includes a wide range of disciplines ranging from synthetic biochemistry to the recreation of life through artificial reconstruction of entire genomes (Iafolla and McMillen, 2006; Gibson et al., 2008). An important aspect of synthetic biology involves the design and construction of engineered gene circuits. Combining the powerful tools of molecular biology and computational modeling, synthetic gene networks can be designed to perform a specific biological function, and experimental data can be used to refine our quantitative understanding of the predicted behavior.

Our ability to synthesize and manipulate gene networks and study their behavior in living organisms has led to significant discoveries regarding some of the most fundamental cellular processes (Kepler and Elston, 2001; Ozbudak et al., 2002; Gardner et al., 2003; Carrera et al., 2009). In addition, the construction of synthetic networks according to the specifications of quantitative models has led to the refinement of our understanding of the principles of cellular regulation (Gardner et al., 2000; Guido et al., 2006; Hasty et al., 2001a; McMillen et al., 2002). Essential to this approach is the ability to develop computational models that can simulate and predict the behavior of cellular networks in growing and proliferating cells. In particular, the yeast *S. cerevisiae* has served as an important eukaryotic model for cellular functions as fundamental as gene regulation and as complex as cell cycle orchestration (Raser and O'Shea, 2004; Acar et al., 2005; Breeden, 2003; Chen et al., 2004). Models of gene regulation have been developed to elucidate sources of noise in gene expression and the effect of noise on fitness, to study the role of feedback in cellular networks, and have led to discoveries of novel network structure (Elowitz et al., 2002; Volfson et al., 2006; Ozbudak et al., 2002; Becskei and Serrano, 2000; Bennett et al., 2008; Basu et al., 2005; Simpson et al., 2003; Blake et al., 2003; Wang et al., 2006; Blake et al., 2006; Smith et al., 2007; Bayer et al.,

2009; Nevozhay et al., 2009; Austin et al., 2006).

As we continue to develop these models as tools to refine our quantitative understanding of basic biological functions, an important and often overlooked contribution to network dynamics is the effect of volume fluctuations associated with the cell growth and division cycle. The majority of gene regulatory models rely on rate constants that are concentration dependent, yet typically the cellular volume is assumed to be constant. By ignoring volume fluctuations that affect concentrations in a quasi-periodic manner, the protein concentrations of “constitutively expressed” genes are often considered to be constant, which can lead to incomplete conclusions regarding gene network behavior.

To address this issue, we measured volume dynamics along with gene expression in a population of yeast cells. We analyzed the growth characteristics of budding cells, starting at birth, and generated a set of descriptive statistics governing the growth and division process. We used this information to develop an accurate cell division model, which takes into account the two distinct linear growth rates observed in the G1 and S phases. When combined with constitutive production of a fluorescent protein, the analysis reveals an oscillatory trend in the protein concentration over time. This effect of cellular growth and division is commonly overlooked, but it can play an important role in the behavior of both native and synthetic gene networks.

## 5.2 Results and Discussion

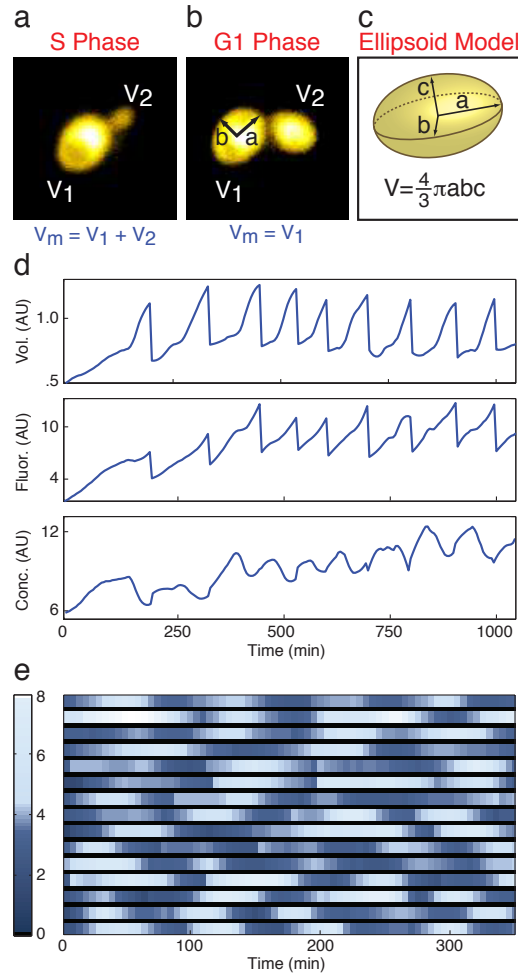
In order to obtain a quantitative set of cell cycle growth characteristics, we tracked a large population of cells growing in a monolayer inside a microfluidic chemostat. The cells expressed a stable Venus YFP from the constitutive *gal1* promoter. A segmentation algorithm was applied to each image, providing an accurate area and position measurement for each cell at each time point (Raser and O’Shea, 2004; Cookson, S. et al., 2005). Volume for each cell was measured by calculating the major and minor axes for each cell area and fitting an ellipsoid to these values (Figure 5.1a, b, and c). The height of the cell was assumed to be equal to the minor axis, unless the minor axis was

larger than the height of the chamber. In that case, the height of the cell was set equal to the height of the chamber.

Cells were tracked from birth, and the volumes of a mother and bud cell were added together throughout each S phase (Figure 5.1a). The first appearance of the bud marks the beginning of S phase, so we used this visual cue to split each cell cycle into two growth phases. Using this time point for each cell cycle, we were able to measure the time spent in each phase and the growth rate of the cell in each phase (approximated by a linear fit to the slope of the volume). At the time of mitosis, the two volume trajectories were split and their volumes were considered separate starting at the beginning of the next G1 phase (Figure 5.1b). For the mother cell, this yields a discontinuity at the point of mitosis, with the volume dropping as the daughter receives a fraction of the volume and the mother retains the remaining fraction (Figure 5.1d, top).

YFP gene expression trajectories were generated by integrating the fluorescence signal over the entire cell at each time point (Figure 5.1d, middle), and concentration was calculated as the total fluorescence divided by the total volume (Figure 5.1d, bottom). The volume trajectories clearly show two distinct linear growth phases: a period of slow growth during the G1 phase of the cell cycle, and a period of faster growth as the cell begins to bud and progress through S phase towards mitosis. The production of YFP, on the other hand, remains roughly constant throughout each cell cycle. Because the tracking algorithm involves measuring the total fluorescence of a stable protein over each cell as a whole (including mother and bud, when appropriate), the production of YFP is calculated as the slope of the fluorescence trajectory between two budding events. The combination of these two phenomena leads to clear oscillations of the concentration trajectories. A density plot of fifteen concentration trajectories demonstrates both the oscillations in individual trajectories as well as the large degree of cell-to-cell variability (Figure 5.1e).

For each cell, we tracked various growth characteristics to obtain distributions and determine trends in these characteristics (Figure 5.2a, See Materials and Methods below). Each cell included in the analysis was tracked from birth, in order to be able



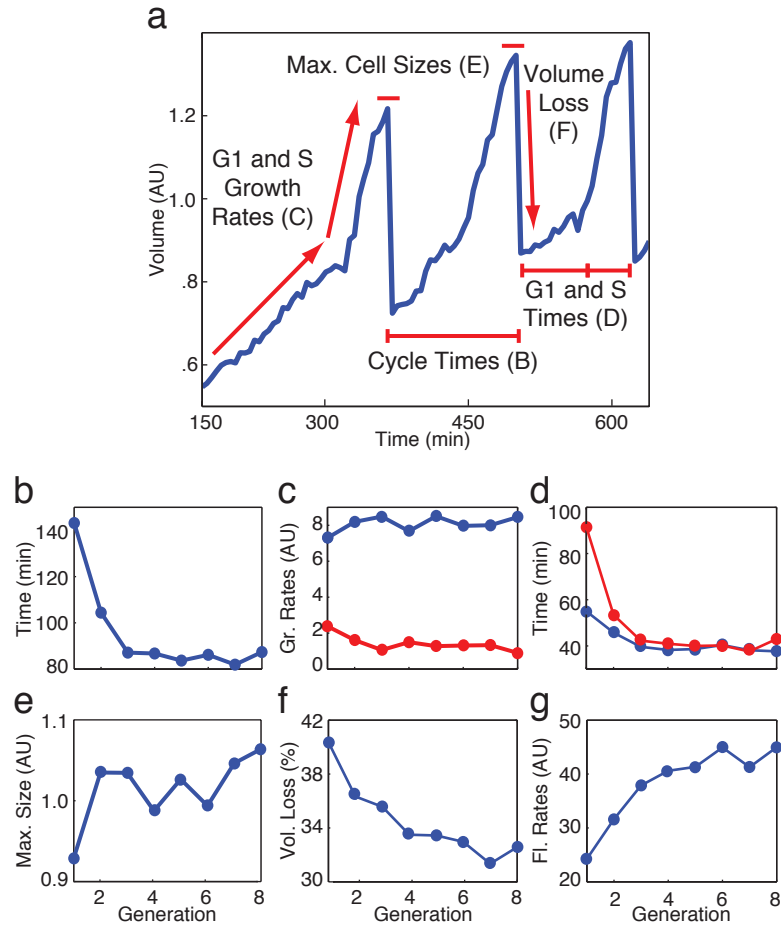
**Figure 5.1:** Single cell volume trajectories for a population of *S. cerevisiae*. **(a)** Asymmetrically budding yeast cells were tracked through many generations of growth and division. During S phase, a mother cell begins to produce a small daughter cell. Until mitosis, the mother and daughter are connected, so their volumes are added together. **(b)** After mitosis, the mother and daughter cells have split, and their volumes are independent. During G1 phase, they grow very slowly as they prepare to produce their next buds. **(c)** Each cell was approximated to be an ellipsoid, with a major and minor axis ( $a$  and  $b$ , respectively). For larger cells with a minor axis greater than the height of the imaging chamber, the height of the cell was set equal to the height of the chamber; otherwise, the height of the cell was set equal to the minor axis. **(d)** Trajectories were generated for each cell, based on these volume measurements. The volume (top) clearly displays two different growth rates for G1 and S phase. Fluorescence (middle) was integrated over the entire area of each cell, and concentration (bottom) was calculated as total fluorescence divided by total volume. The combination of constant fluorescence production and bilinear volume growth yields oscillatory concentration trajectories. **(e)** Concentration trajectories for fifteen cells show clear oscillations and a high degree of variability.

to confidently discern generation-dependent behaviors. That is, some statistics were observed to change over the first few generations of life, as the cell gradually reached a healthy adult size and growth rate. For example, the time required to complete each cell division cycle decreases over the first few generations (Figure 5.2b). This is a well known phenomenon, which has been observed using various techniques over the years (Egilmez and Jazwinski, 1989; Cookson, S. et al., 2005).

Other traits that we observed and characterized include growth rates during both G1 and S phases, growth phase durations, fraction of volume loss at the end of each mitosis occurrence, and cellular volume at the beginning and end of each cell cycle. While some characteristics are fairly constant, others vary greatly from cell to cell or even for an individual cell throughout its trajectory. Figure 5.2 shows all of the characteristics that we quantified and how they vary by generation. For each characteristic, we quantified the mean and standard deviation across all cells as well as across each individual cell's multiple generations.

Our volume trajectories demonstrate two clear growth phases, which could each be approximated by a linear growth rate. As expected, the G1 phase growth rate is significantly slower than the S phase growth rate (Figure 5.2c, S in blue, G1 in red). Interestingly, the growth rates in each phase remain constant throughout the generations observed. However, the time spent in each phase changes over the first three cycles, following a similar trend as the overall cell cycle time (Figure 5.2d). The time spent in G1 has a more extreme trend, with the time almost dropping in half over the first three generations. The S phase growth rate on the other hand speeds up a little as well, but our data demonstrates that the extra time required to complete the cell cycle early in life is spent mostly in the G1 phase. This is consistent with other studies that have demonstrated that changes in growth rate in different media conditions are mainly reflections of altered lengths of the G1 phase, with the duration of the S phase remaining fairly constant (Slater et al., 1977).

We also looked at cell size trends. For each cell, we examined how the maximum size reached at the end of each cell cycle varies over time, as well as how the mean of



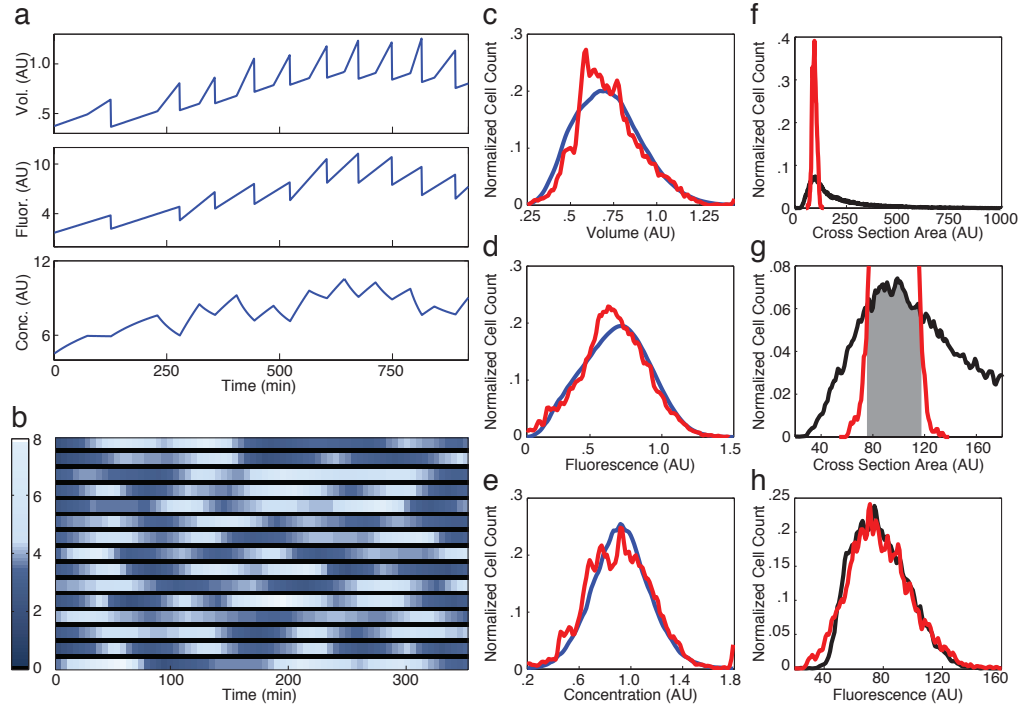
**Figure 5.2:** Volume growth and division statistics. **(a)** For each trajectory, we determined the mean and standard deviation of certain cell growth characteristics, starting at birth. These features were compared across each cell's various generations as well as across all trajectories. **(b)** The time to complete each cell cycle shows a downward trend over the first few generations. **(c)** The G1 (red) and S (blue) growth rates remain constant over time. **(d)** The time spent in each phase decreases over the first few generations, although the majority of time during the early generations is spent in G1. **(e)** The size of cells increases linearly with generation, although a larger jump in size of about 5% is observed between the first and second generations. **(f)** The percent of volume lost by a mother cell to its bud decreases on average for about five generations. **(g)** Finally, the fluorescence production rate is observed to increase over the first five or six generations, on average.

the maximum size compares to that of other cells (Figure 5.2e). As consistent with previous findings, our trajectories show a slight linear increase of cellular volume with age (Egilmez et al., 1990). However, by tracking cells from birth, we were also able to detect a more marked increase in size between the first and second generations of about 5% on average. We also tracked the percentage of volume lost by a mother to its new bud, which also had an interesting trend (Figure 5.2f). We found that it takes over five cycles for the value to level off to about 32%. This indicates that cells in their early generations bud daughter cells that are larger relative to their own size than they do throughout most of their lives.

These statistics and trends come together to form a comprehensive picture of how the volume changes in a quasi-periodic manner over many cellular generations, starting at birth. To determine how these volume fluctuations can affect the concentrations of intracellular proteins, we also measured fluorescence in the same cells. The fluorescence trajectories revealed a constant rate of protein production during each cell cycle (Figure 5.1d, middle). While fairly constant and linear, we found that the protein production rate for this constitutive promoter increases over the first five cycles before leveling off (Figure 5.2g). This could be due to many factors, such as a gradual building up of protein synthesis machinery, freeing up of resources originally dedicated to maturing into a healthy adult cell, or a priority shift to metabolic gene expression as the cell matures.

We compiled all of this information into a model that combines growth and division of budding yeast cells with constitutive YFP production. The model generates volume and fluorescence trajectories for newly budded cells for an arbitrary number of generations, which closely match the experimental results (Figure 5.3a and b). Histograms of volume, fluorescence, and concentration distributions provide a good metric for comparing the experimental and simulated trajectories, and our simulations match the means and standard deviations of the experiments well (Figure 5.3c, d, and e).

Additionally, histograms from the microscopy data can be used as a standard for understanding flow cytometry data, which is commonly used to measure gene ex-



**Figure 5.3:** Comparison of experimental volume trajectories to model simulations and flow cytometry data. **(a)** Simulated volume (top), fluorescence (middle), and concentration (bottom) trajectories have similar trends as the experimental data (Figure 1). **(b)** Simulated concentration trajectories for fifteen cells. **(c, d, e)** Comparisons of histograms of volume, fluorescence, and concentration from the simulations (blue) and microscopy (red) data show well matched means and standard deviations. **(f)** Comparison of size data from the flow cytometry (black) and microscopy (red) data demonstrates that the flow cytometer generates a deceptively wide distribution. **(g)** We can use the microscopy data to determine how to gate the flow cytometry data in order to sample just individual yeast cells. This metric yields a gate of 20% on either side of the forward scatter peak. **(h)** Using these gated cells, we can compare the fluorescence histograms from the flow cytometry and microscopy data, which yields very well matched means and standard deviations.



pression noise but is often difficult to interpret. The forward scatter (FSC) measurement provided by a flow cytometer is believed to be proportional to the size of the object being measured, but it is unclear exactly what this value represents. In addition, the tendency of yeast cells to clump and their unusual shape due to asymmetric budding result in misleading size distributions. In order to compare our microscopy size data to the FSC histogram provided by the flow cytometer, we had to first calibrate the FSC measurement to some physical characteristic of a yeast cell. To do this, we used spherical flow calibration beads of various diameters ranging from 2 to 15  $\mu\text{m}$ . We ran these beads through the flow cytometer on the same settings we use to collect our yeast data, and we calculated the mean FSC measurement. In order to determine if the FSC was proportional to the length, area, or volume of a sphere, we compared the diameter, diameter<sup>2</sup>, and diameter<sup>3</sup> of the beads to the mean FSC value. It was clear from our data that the FSC measurement is proportional to the diameter of the bead (See Materials and Methods below).

Using the scope data as a reference for the true distribution of sizes for individual budding cells, we show that the forward scatter data provided by the flow cytometer yields a much wider distribution (Figure 5.3f), which translates to much higher fluorescence variability than is actually present in a healthy population of single cells. The presence of clumps of two or more cells as well as smaller particles (dead cells or media debris) creates a much wider size distribution than a homogeneous population of single cells would yield. We can use our microscopy size histogram to determine a metric for gating the flow cytometry size histogram about the mode of the distribution (Figure 5.3g). The fluorescence distribution of only the gated cells accurately reproduces the microscopy data (Figure 5.3h) and yields a coefficient of variation that is three times lower than that of the entire ungated data set.

As synthetic biology moves steadily towards the goal of developing accurate and predictive models of cellular behavior, it is clear that computational modeling will rely heavily on quantitative measurements of cellular phenomena. The cell division cycle is a highly dynamic and noisy process that affects a cell's behavior on many levels, and

the fluctuations in cellular volume will affect the function of both native and synthetic gene networks. It will be important to account for these fluctuations in the construction of synthetic circuits designed to perform a specific function, such as to generate oscillations or toggle between two states, as they will inevitably affect network dynamics. In addition, determining the sources and effects of noise in biological circuits is becoming increasingly critical, as we try to develop more detailed and robust models of genetic networks. These quasi-periodic fluctuations will have an impact on the study of noise biology, which relies heavily on precise terms for the dilution of cellular components due to cell growth. Our results demonstrate the importance of accurately representing the complex growth trends of a population of cells in order to account for the effect that the resulting volume fluctuations have on the concentrations of intracellular components.

## 5.3 Materials and Methods

### 5.3.1 Strain and cell culture

The yeast strain was created by targeted chromosomal integration of the pRS61-yv vector at the *gal1-10* locus of *S. cerevisiae* strain K699 (a, ADE2, ura3, his3, trp1, leu2). This vector was constructed using standard recombination techniques and contains the *gal1* promoter locus of *S. cerevisiae* driving production of the yeast enhanced Venus fluorescent protein (yEVFP), a YFP variant (Raser and O'Shea, 2004). Cultures were grown in synthetic drop-out (SD) medium supplemented with all amino acids except uracil for selection of correct integration and containing 2% glucose. After selection, cultures were grown in SD supplemented with all amino acids and containing 2% galactose for full induction of the production of the yEVFP protein. Cultures were grown at 30 °C for 12-18 hours to an OD<sub>600</sub> of  $1.0 \pm .25$ . In preparation for loading into the microfluidic device, the sample was passed back to an OD<sub>600</sub> of 0.1 and allowed to grow for about 3 hours to re-enter exponential growth.

### 5.3.2 Data acquisition

Image acquisition was performed on a Nikon Diaphot TMD epifluorescent inverted microscope with a hardware based autofocus controller (Prior Scientific, Rockland, Massachusetts). Images were acquired using a Hamamatsu ORCA-ERG cooled CCD camera, and fluorescence visualization was performed with narrow bandpass excitation and emission filters for YFP (Chroma, Inc., Rockingham, Vermont). The cells were imaged inside a microfluidic chemostat as previously described (Cookson, S. et al., 2005).

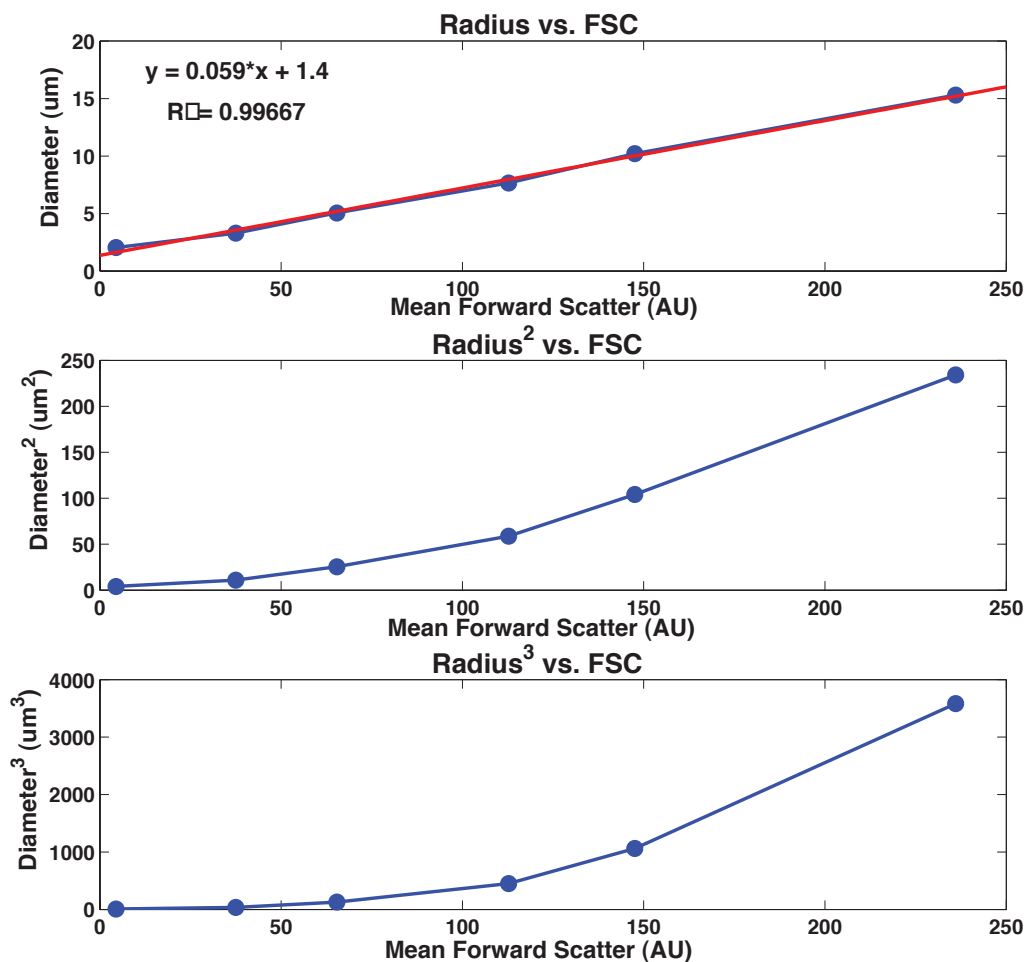
### 5.3.3 Flow cytometry data acquisition

Flow cytometry data was acquired with a Becton-Dickinson FACSCalibur flow cytometer. To prepare the cells to be assayed, we grew them overnight in medium containing 2% galactose. In the morning, cells were passed to an  $OD_{600}$  of .05 and allowed to grow for 8 hours. At this time, 1 mL of the culture was spun down at 8000 RPM for 1 minute, and the cells were resuspended in sterile phosphate buffered saline (PBS). The cells were run through the flow cytometer, and 50,000 cells were sampled. Again, MATLAB was used to analyze the flow cytometry data.

### 5.3.4 Forward scatter size calibration

In order to compare our microscopy size data to the FSC histogram provided by the flow cytometer, we had to first determine what size dimension the FSC data represents. To do this, we used spherical flow calibration beads of various diameters ranging from 2 to 15  $\mu\text{m}$  (SPHERO Flow Cytometry Size Standard Kit). Using a Becton-Dickinson FACSCalibur flow cytometer, we ran these beads through on the same settings we use to collect our yeast data. For each size bead, we collected 50,000 samples, and we used MATLAB (The MathWorks, Inc.) to calculate the mean FSC measurement. In order to determine if the FSC was proportional to the length, area, or

volume of a sphere, we compared the diameter, diameter<sup>2</sup>, and diameter<sup>3</sup> of the beads to the mean FSC value. It is clear from our data that the FSC measurement is proportional to the diameter of the bead (Figure 5.4). Also, a linear fit to the diameter vs. FSC data provides a nice equation to convert FSC to length.



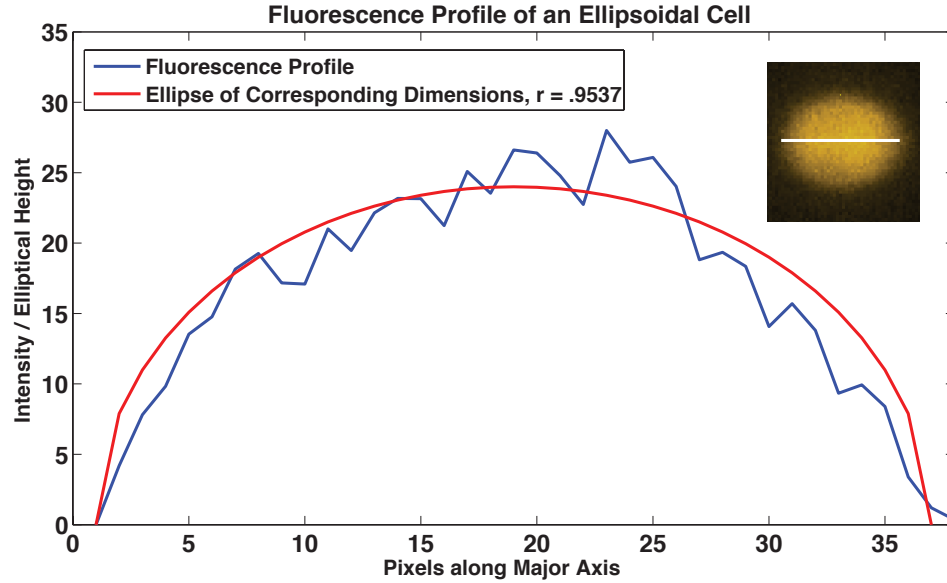
**Figure 5.4:** Size data for flow calibration beads. The top panel compares the diameter of the beads to the mean forward scatter reported by the flow cytometer. The bottom panel compares the diameter squared (area) of the beads, and the bottom compares the diameter cubed (volume) of the beads to the forward scatter.

### 5.3.5 Fluorescence Integration

In order to measure total fluorescence emitted by a single cell, we must be able to assume that our image acquisition method is able to capture the fluorescence signal from the entire cell. Wide-field microscopy is ideal for this purpose, as a wide-field microscope collects the light from all points in the focal plane at once plus all the light from illuminated regions of the sample that are above and below the focal plane. In addition, a 40x objective was chosen in order to strike a good balance between fluorescence signal, depth of field, and field of view.

The depth of field (the axial range through which an objective can be focused without any appreciable change in image sharpness) of the 40x oil objective used is about  $1.0\text{ }\mu\text{m}$ . While light outside of this region will be collected, this signal may not be integrated linearly. However, as the imaging region is only  $4\text{ }\mu\text{m}$  high, and the cells are confined to a monolayer within this region, the loss of light due to out-of-focus regions of a cell's volume is likely insignificant and should affect all cells in a similar manner. As a check to ensure that images were composed of light emitted from all regions of the cell, we examined the fluorescence profile extracted along the major axis of a healthy adult cell. Assuming an ellipsoidal cell shape and nonspecific YFP localization, we would expect this fluorescence profile to fit an elliptical curve created based on the major and minor axis dimensions of the cell.

Below, this verification is demonstrated, with a high degree of correlation (Figure 5.5). Slight imperfections in the correlation may be due to deviations from a purely ellipsoidal cell shape as well as the imperfect distribution of YFP throughout the cellular volume. Fluorescent proteins should not be localized to any particular area and should be robust to pH differences, so we expect the signal to be more or less constant throughout the cellular volume. However, since we are measuring the fluorescence of the cell as a whole, the accurate measurement of total fluorescence does not rely on the homogeneity of the fluorescent protein.



**Figure 5.5:** Correlation between the fluorescence profile of a cell taken along its major axis (inset) and the corresponding elliptical shape.

### 5.3.6 Computational model

MATLAB was used to model the growth and division of a population of yeast cells. The data from the trajectory analysis was used to generate histograms for each characteristic. These histograms were fit to Gaussian distributions, and the mean and standard deviation was calculated for each. These statistics were fed into the MATLAB model. Each cell was initialized using values drawn from a distribution based on this data. As cells were modeled from birth, the trends of the characteristics as a function of generation were also taken into account. That is, cells were initialized to a size drawn from the distribution of cells sizes at birth. Similarly, the fluorescence production rate, growth rate, and growth phase characteristics were chosen from histograms of first generation measurements. After initialization, a cell was grown over time according to the current value of the growth rate, and fluorescence was produced based on the current value of fluorescence production.

The cell size was grown at the current value of the G1 growth rate until a thresh-

old time was reached, set by the product of the cycle time and the fraction of time to be spent in G1. After this time, the cell size was grown at the current value of the S growth rate. The fluorescence was grown at a constant rate for the entire duration of the current cell cycle. When cell cycle end time was reached, the size and fluorescence of a cell was multiplied by the current value of  $\epsilon$ , the fraction of volume to be retained by the mother at the time of division. For the next cycle, the values above were reset, again drawn from the distributions extracted from experimental data. The cells were grown for 836 minutes, the average length of the experimental trajectories. One thousand cells were simulated and compared to the experimental data.

— Begin MATLAB Code —

```
function conc_array = YeastGrowth()
numcells = 1000; % total # cells to simulate
totaltime = 836; % total # time points - the average length of the experimental trajectories
volume_array = zeros(numcells,totaltime); % array to store volume trajectories
fluor_array = zeros(numcells,totaltime); % array to store the fluorescence trajectories
for c = 1:numcells % iterate through cells
initialSizeMean = .4352; % mean initial volume of cells
initialSizeStd = .0772; % std. dev. of initial volume of cells
maxSizeMean = 1.00; % mean max volume reached by a cell at the end of each cell cycle
maxSizeStd = .1199; % std. dev. of max volume
maxSizeStdInd = .0842; % once a max volume is chosen for a particular cell, it will vary by generation about this mean
meanCycleGrowth = 0.3577; % mean volume increase over one cell cycle
stdCycleGrowth = 0.0948; % std. dev. of volume increase over one cell cycle
cycletimes = [143.20 104.42 86.64]; % mean cycle times for the 1st, 2nd, and all following generations
cycletimes_stds = [31.00 26.32 11.50]; % std. dev. of cycle times
```

```

epsilon_means = [0.5963 0.6367 0.6443 0.6650 0.6656 0.6703 0.6760]; % fraction of
volume the mother retains at bud event for 1st, 2nd, 3rd, ... and all following generations
epsilon_stds = [0.0479 0.0522 0.0483 0.0489 0.0470 0.0474 0.0513]; % std. dev. of ep-
silons
g1_fraction_means = [.611 .523 .511]; % mean fraction of cell cycle time spent in G1
phase
g1_fraction_stds = [.097 .112 .094]; % std. dev. of fractions
G1_rate = .0014; % mean growth rate for G1 phase
G1_std = .000099; % std. dev. growth rate for G1 phase
fluor_rate_means = [2426 3163 3789 4061 4313]; % mean rate of YFP production for
the first 4 cycles and all cycles after
fluor_rate_stds = [1039 882 908 1083 1004]; % std. dev. of rate of YFP production for
the first 4 cycles and all cycles after
initialConcMean = 5.8353e+05; % mean initial YFP concentration in a new cell
initialConcStd = 1.5686e+05; % std. dev. of initial YFP concentration in a new cell
cellarr = zeros(totalltime,12); % trajectory information of single cell: # rows = # time
points
cellarr(:,1) = (1:totalltime); % column #1 stores the current time
initial_size = randn(1)*initialSizeStd + initialSizeMean;
cellarr(1,2) = initial_size; % column #2 stores the current volume - starting size chosen
from distribution
cellarr(1,12) = initial_size; % column #12 stores the initial volume of the cell for the
current cell cycle
cellarr(1,3) = 1; % column #3 stores the current generation - various parameters are de-
pendent on generation #
cellGrowth = randn(1)*stdCycleGrowth + meanCycleGrowth; % the amount the cell
will grow over the current cell cycle
maxSizeNew = initial_size + cellGrowth;
cellarr(1,4) = maxSizeNew; % column #4 stores the maximum volume that a cell will

```



```

reach during the current generation
cycleTimeCell = randn(1)*cycletimes_stds(1) + cycletimes(1);
cellarr(1,5) = cycleTimeCell; % column #5 stores the time to complete current cell divi-
sion cycle
alpha0 = rand(1)*G1_std + G1_rate;
cellarr(1,6) = alpha0; % column #6 stores the growth rate during the G1 phase
eps = randn(1)*epsilon_stds(1) + epsilon_means(1);
cellarr(1,8) = eps; % column #8 stores the current epsilon value - the fraction of volume
that the mother maintains after division
g1_fraction = randn(1)*g1_fraction_stds(1) + g1_fraction_means(1);
cellarr(1,9) = g1_fraction; % column #9 stores the fraction of the cell cycle duration
spent in G1 phase
halfdouble = cycleTimeCell*g1_fraction; % S phase growth rate is calculated based on
other factors - G1 fraction, G1 growth rate, cell cycle time, and maximum size
maxSize = cellarr(1,4);
d1 = alpha0*halfdouble;
d2 = maxSize - d1 - initial_size;
alpha1 = d2/(cycleTimeCell-halfdouble);
cellarr(1,7) = alpha1; % column #7 stores the growth rate during S phase
fluor_rate = randn(1)*fluor_rate_stds(1) + fluor_rate_means(1);
cellarr(1,10) = fluor_rate; % column #10 stores the rate of YFP production in the cell
initialConc = randn(1)*(initialConcStd) + initialConcMean; % YFP amount is calcu-
lated based on the initial YFP concentration and the initial size of the cell
initialFluor = initialConc*initial_size;
cellarr(1,11) = initialFluor; % column #11 stores the total amount of YFP in the cell
tau = 1.0; % time step in minutes
tct = 0; % initialize time counter - helps keep track of cell cycle phase
for tpt = 2:totaltime % step through time and generate a volume trajectory
alpha0 = cellarr(tpt-1,6); % get current G1 growth rate

```

```

alpha1 = cellarr(tpt-1,7); % get current S growth rate
cycletime_curr = cellarr(tpt-1,5); % get current cell cycle duration
g1_fraction_curr = cellarr(tpt-1,9); % get current G1 fraction
fluor_rate_curr = cellarr(tpt-1,10); % get current YFP production rate
halfdouble = cycletime_curr*g1_fraction_curr; % calculate the time to end the G1 phase
for this generation
if (tct*tau < halfdouble) % check cell cycle phase to determine growth rate
cellarr(tpt,2) = cellarr(tpt-1,2) + alpha0; % linear size growth - slower in G1 phase than
S phase else
cellarr(tpt,2) = cellarr(tpt-1,2) + alpha1; % after G1 phase, grow at faster rate
end
cellarr(tpt,11) = cellarr(tpt-1,11) + fluor_rate_curr; % linear YFP production
tct = tct+1; % increment tct
cellarr(tpt,3) = cellarr(tpt-1,3); % update cell array with previous values - these will only
change if it is time to divide
cellarr(tpt,4) = cellarr(tpt-1,4);
cellarr(tpt,5) = cellarr(tpt-1,5);
cellarr(tpt,6) = cellarr(tpt-1,6);
cellarr(tpt,7) = cellarr(tpt-1,7);
cellarr(tpt,8) = cellarr(tpt-1,8);
cellarr(tpt,9) = cellarr(tpt-1,9);
cellarr(tpt,10) = cellarr(tpt-1,10);
cellarr(tpt,12) = cellarr(tpt-1,12);
if(cellarr(tpt-1,5) < tct) % check if cell needs to divide; if so, proceed with division
size = cellarr(tpt,2); % get current volume
eps = cellarr(tpt,8); % get current epsilon
yfp = cellarr(tpt,11); % get current YFP level
cellarr(tpt,11) = yfp*eps; % divide YFP and volume according to epsilon value
initial_size = size*eps;

```

```

cellarr(tpt,2) = initial_size; % update current volume
cellarr(tpt,12) = initial_size; % update initial volume
cellarr(tpt,3) = gen; % update generation
cellGrowth = randn(1)*stdCycleGrowth + meanCycleGrowth;
maxSizeNew = initial_size + cellGrowth; % generate new max volume
if gen == 2 % increase next size appropriately between generations 1 and 2 - this is a
larger increase than in later gens
maxSizeNew = maxSizeNew*1.06;
end
maxSizeNew = maxSizeNew*1.026; % increase next size appropriately after generation
2
cellarr(tpt,4) = maxSizeNew; % update max volume for next cycle
alpha0 = rand(1)*G1_std + G1_rate;
cellarr(tpt,6) = alpha0; % update G1 growth rate for next cell division cycle
ind = gen; % create an index based on the new generation - 1, 2, 3 or 4 and beyond
if (ind > 3)
ind = 3;
end
cycleTimeCell = randn(1)*cycletimes_stds(ind) + cycletimes(ind); % use appropriate
value for cycle time and G1 fraction based on generation
cellarr(tpt,5) = cycleTimeCell; % update time to complete next cell division cycle
g1_fraction = randn(1)*g1_fraction_stds(ind) + g1_fraction_means(ind);
cellarr(tpt,9) = g1_fraction; % update G1 fraction for next cell division cycle
halfdouble = cycleTimeCell*g1_fraction;
d1 = alpha0*halfdouble;
d2 = maxSizeNew - d1 - initial_size;
alpha1 = d2/(cycleTimeCell-halfdouble);
cellarr(tpt,7) = alpha1; % update S growth rate for next cell division cycle
ind = gen; % create an index based on the new generation - 1, 2, 3...6, or 7 and beyond

```

```

if (ind > 7)
ind = 7;
end
eps = randn(1)*epsilon_stds(ind) + epsilon_means(ind); % use appropriate value for epsilon based on generation
cellarr(tpt,8) = eps; % update epsilon for next cell division cycle
ind = gen; % create an index based on the generation - 1, 2, 3, 4 or 5 and beyond
if (ind > 5)
ind = 5;
end
fluor_rate = randn(1)*fluor_rate_stds(ind) + fluor_rate_means(ind); % use appropriate value for YFP production rate based on generation
cellarr(tpt,10) = fluor_rate; % update YFP production rate
tct = 0; % reset time counter for next cell division cycle
end
end
volume_array(c,:) = cellarr(:,2)'; % fill next row of volume array with current cell's volume trajectory
fluor_array(c,:) = cellarr(:,11)'; % fill next row of fluorescence array with current cell's fluorescence trajectory
end
mean(volume_array(:)) % display the mean of the volume array
std(volume_array(:)) % display the std. dev. of the volume array
mean(fluor_array(:)) % display the mean of the fluor array
std(fluor_array(:)) % display the std. dev. of the fluor array
vbins = .24:.01:1.46; % plot the distributions for volume and fluorescence
figure(1);
[vhist,vbins] = hist(volume_array(:),vbins);
hist(volume_array(:),100)

```

```
fbins = 60000:11800:1500000;
figure(2);
[fhist,fbins] = hist(fluor_array(:),fbins);
hist(fluor_array(:),100)
output = [vhist; vbins; fhist; fbins];
```

— End MATLAB Code —

### 5.3.7 Volume Growth Characteristics

The code is provided in the preceding section. This code details the values used for each of the characteristics that we measured and integrated into the cell volume model. All size measurements were normalized so that the average adult cell had a mean maximum volume (volume attained at the end of each cell cycle) of 1.0. The characteristics measured were the following:

- Initial Size: The volume of a newly budded daughter cell.
- Maximum Size: The maximum volume reached by a cell at the end of each cell division cycle.
- Cycle Growth: The volume increase over a single cell division cycle.
- Cycle Time: Time to complete the cell division cycle.
- Epsilon: Fraction of volume the mother retains at bud event.
- G1 Fraction: Fraction of cell cycle time spent in G1 phase.
- G1 Rate: Growth rate for G1 phase.
- S Rate: Growth rate for S phase.
- Fluor. Rate: Rate of YFP production.

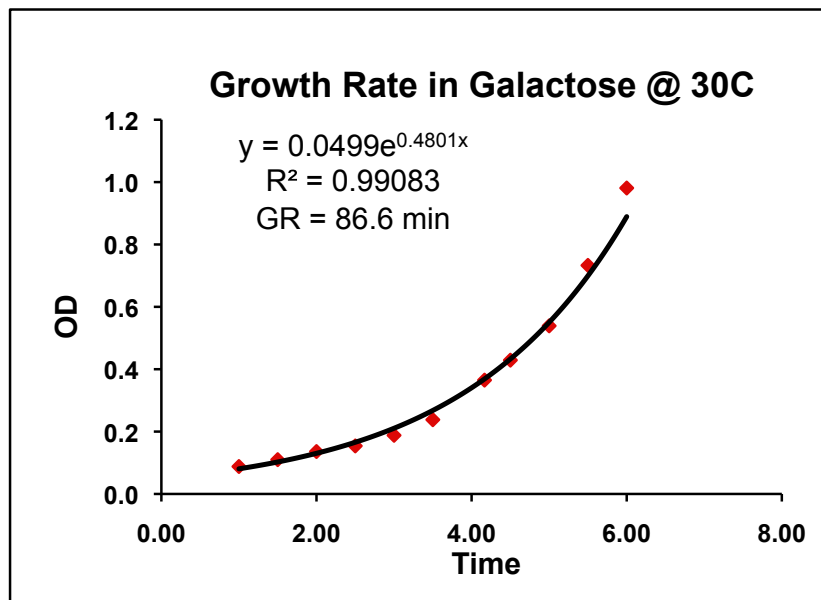
- Initial Fluor.: Initial YFP level in a new cell.
- Initial Conc.: Initial YFP concentration in a new cell.

In the table below, we detail the mean and standard deviation of each of these characteristics. If a characteristic changed throughout the early generations (gen #), we document the specific value for each cycle, until the value leveled off (at which point we leave the table blank).

**Table 5.1:** Statistics of growth characteristics as they change over the generations.

Characteristic	Gen 1	Gen 2	Gen 3	Gen 4	Gen 5	Gen 6	Gen 7
Initial Size Mean	.4532						
Initial Size Std.	.0772						
Max. Size Mean	1.0						
Max. Size Std.	.1199						
Cyc. Growth Mean	.3577						
Cyc. Growth Std.	.0948						
Cyc. Time Mean	143.20	104.42	86.64				
Cyc. Time Std.	31.00	26.32	11.50				
Epsilon Mean	.5963	.6367	.6443	.6650	.6656	.6703	.6760
Epsilon Std.	.0479	.0522	.0483	.0489	.0470	.0474	.0513
G1 Fraction Mean	.611	.523	.511				
G1 Fraction Std.	.097	.112	.094				
G1 Rate Mean	.0014						
G1 Rate Std.	.0001						
S Rate Mean	.0082						
S Rate Std.	.003						
Fluor. Rate Mean	2,426	3,163	3,789	4,061	4,313		
Fluor. Rate Std.	1,039	882	908	1,083	1,004		
Init. Fluor. Mean	267,800						
Init. Fluor. Std.	100,500						
Init. Conc. Mean	583,530						
Init. Conc. Std.	156,860						

We also measured the growth rate in batch culture of this strain, using the standard OD<sub>600</sub> technique and found a doubling time of 86.6 minutes, which corresponds with the mean cycle time for a healthy adult cell (Figure 5.6).



**Figure 5.6:** Growth rate calculated by measuring the OD<sub>600</sub> over time.

## 5.4 Acknowledgements

Chapter 5 contains material originally published as Cookson, N., Cookson, S., Tsimring, L., and Hasty, J., 2009: Cell cycle dependent variations in protein concentration. *Nucl. Acids Res.*, in press.. Copyright permission to republish here was granted by Oxford University Press.

## Chapter 6

# Queuing up for enzymatic processing: Correlated signaling through coupled degradation

### 6.1 Introduction

Genomic technologies have led to significant progress in the deduction of network connections arising from transcriptional and post-transcriptional regulation, yet few studies have considered global mechanisms of indirect coupling that effectively increase network connectivity. Such indirect coupling may arise from an abundance of target molecules relative to a limited number of processing components, resulting in “waiting lines” that occur in the context of queuing theory. Using the *E. coli* clpXP degradation machine as a model processing system, we induce a “master” signaling network and monitor a “slaved” constitutive system that is indirectly coupled to the master through the common clpXP degradation processor. We observe significant correlations between master and slave in their steady-state dose-response curves and dynamic signaling responses. Our results demonstrate how global processing queues can provide a mechanistic origin for intracellular correlations between many molecular species.



Genetic wiring diagrams have arisen as a post-sequencing paradigm for depicting the interactions between vast and diverse cellular species (Ideker et al., 2001a; Sauer, 2004; Li et al., 2004). These diagrams typically arise from high-throughput experiments that probe cellular networks for interactions between genes either at the level of transcriptional regulation or protein-protein interactions, enabling the connection of components in a large-scale network “map.” An alternative approach to drawing connections between cellular components involves subjecting cells to a particular stimulus, such as DNA damage or starvation, observing the response of a large number of cellular components, and looking for correlated behavior (Alon et al., 1999; Golub et al., 1999). This approach often leads to much more highly interconnected networks than are predicted by standard gene and protein interaction maps, suggesting the existence of alternative, indirect sources of coupling.

## 6.2 Results and Discussion

Here, we use notions from queuing theory to investigate the effect that “waiting lines” can have on the coupling of proteins, resulting from enzymatic processes involving many target molecules and a relatively small number of processing machines (see Fig. 6.1a). Suppose that two proteins  $x_1$  and  $x_2$  involved in different signaling pathways do not directly interact with each other but are being processed downstream with the same enzyme  $E$ . If  $E$  is abundant, its interaction with  $x_1$  and  $x_2$  is not limited by the amount of  $E$ , and thus  $x_1$  and  $x_2$  remain uncoupled. If however the amount of  $E$  is limited,  $x_1$  and  $x_2$  “compete” for  $E$  and this competition introduces correlations between  $x_1$  and  $x_2$ , which may be hard to distinguish from direct crosstalk that arises from regulatory mechanisms or bimolecular protein-protein interactions. A common example of such shared enzymatic machinery is the superfamily of AAA+ proteases which target for degradation multiple types of proteins which are either damaged or no longer required (Maurizi et al., 1990; Wojtkowiak et al., 1993; Sauer et al., 2004; Levchenko et al., 2000). These “degradation factories” are a natural object of study in the context of

*queuing theory*, which deals with delays in processing that occur when multiple types of “customers” wait for service from one or more “servers” (Kelly, 1979; Bramson, 1998; Williams, 1998; Bramson and Dai, 2001). When servers are busy, customers of different types compete for service, and that introduces correlations in their waiting times: increasing the number of customers of one type will increase the waiting time of other types as well.

The coupling effect of a common enzyme can be understood in the framework of a deterministic model based on a Michaelis-Menten kinetics approximation for proteins processed in dividing cells (Michaelis and Menten, 1913). The concentrations  $x_1$  and  $x_2$  for two distinguishable proteins are assumed to obey the dynamics

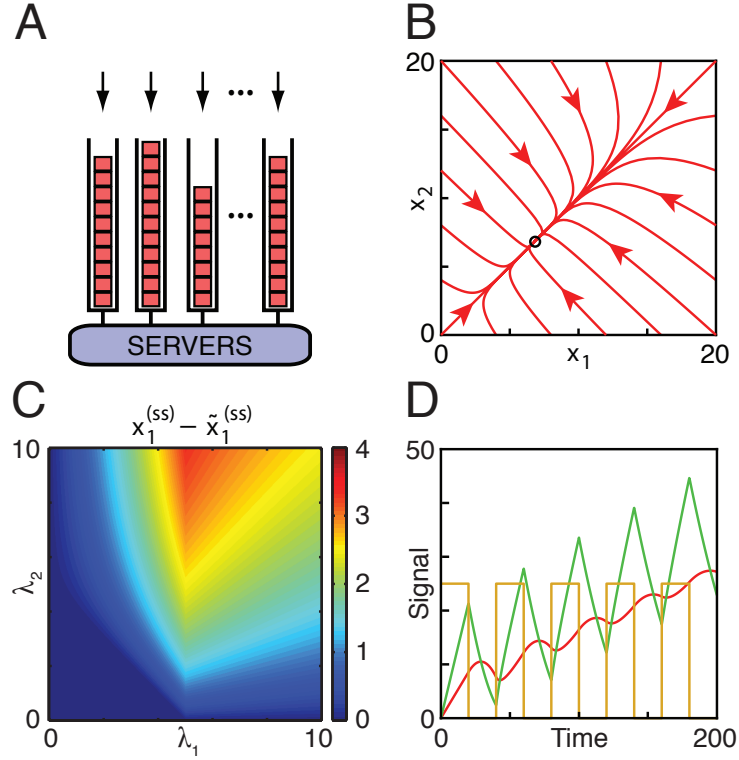
$$\frac{dx_1}{dt} = \lambda_1 - \gamma x_1 - \mu x_1 / (K + x_1 + x_2), \quad \frac{dx_2}{dt} = \lambda_2 - \gamma x_2 - \mu x_2 / (K + x_1 + x_2) \quad (6.1)$$

where  $\lambda_1, \lambda_2$  are production rates,  $\gamma$  is the dilution rate due to cell growth,  $K$  is a Michaelis-Menten molar constant, and  $\mu$  is the maximum enzymatic processing rate. Near so-called balance, i.e. when  $\lambda_1 + \lambda_2 \approx \mu$ ,  $x_1$  and  $x_2$  strongly contract onto a slow manifold (see Fig. 6.1b). The steady state solutions  $x_1^{(ss)}$  and  $x_2^{(ss)}$  in the limit  $K \rightarrow 0$  are (see SI for the more general result)

$$x_1^{(ss)} = \theta(\eta) \eta \lambda_1 / \gamma (\lambda_1 + \lambda_2), \quad x_2^{(ss)} = \theta(\eta) \eta \lambda_2 / \gamma (\lambda_1 + \lambda_2) \quad (6.2)$$

where  $\eta = \lambda_1 + \lambda_2 - \mu$ , and  $\theta(\cdot)$  is the Heaviside step function:  $\theta(\eta) = 1$  if  $\eta \geq 0$  and  $\theta(\eta) = 0$  if  $\eta < 0$ . From Equations 6.2, coupling between non-zero steady state concentrations arises when  $\lambda_1$  and  $\lambda_2$  are simultaneously nonzero and  $\lambda_1 + \lambda_2 > \mu$  (see Fig. 6.1c). This coupling can also be seen for a driven system (see Fig. 6.1d), where a signal in  $\lambda_1$  can propagate into  $x_2$  signal.

As a model enzymatic queuing process, we explored the degradation of ssrA-tagged proteins by the native *E. coli* ClpXP machine. ClpXP is a protease composed of two multimeric subunits, ClpX and ClpP. ClpX is a hexameric ATP-ase which binds and denatures proteins that are targeted for rapid removal from the system; these proteins are then translocated into the degradation chamber within ClpP, where they are



**Figure 6.1:** (A) Rate-limited processing can couple the queue lengths of different job types. Here, jobs (red squares) are flowing into their respective queues (open black boxes). The set of servers removes these jobs from their queues and processes each job with some characteristic time. Queue lengths become coupled due to different jobs competing for the attention of the servers. (B) Flow of the variables  $x_1$  and  $x_2$  with the dynamics in Equations 6.1, using parameters  $\lambda_1 = \lambda_2 = 100$ ,  $\mu = 200$ ,  $\gamma = 1$ , and  $K = 1$ . Coupling due to common degradation tends to contract  $x_1$  and  $x_2$  to a slow manifold, defined approximately by  $x_1/x_2 = \lambda_1/\lambda_2$ . A steady state (black circle) is eventually reached. (C) Coupling is apparent at the level of steady state concentrations. Using Equations 6.2 for the steady state concentrations  $x_1^{(ss)}$  and  $\tilde{x}_1^{(ss)}$  with ( $\lambda_2 \neq 0$ ) and without ( $\lambda_2 = 0$ ) the effect of degradation coupling, respectively, the difference  $x_1^{(ss)} - \tilde{x}_1^{(ss)}$  is plotted as a measure of coupling. This coupling arises when  $\lambda_1$  and  $\lambda_2$  are simultaneously nonzero and  $\lambda_1 + \lambda_2$  is sufficiently large. (D) Degradation coupling in a driven system. Here,  $\lambda_2 = 1$  is constant, while  $\lambda_1$  varies between 0 and 2.5 (indicated by yellow line, scaled 10x). Oscillating signal in  $x_1$  (green) is propagated into  $x_2$  (red) by the degradation coupling. Other parameters are  $\mu = 2$ ,  $\gamma = 0$ ,  $K \rightarrow 0$ . Note that the drive alternatively places the system above and below balance.

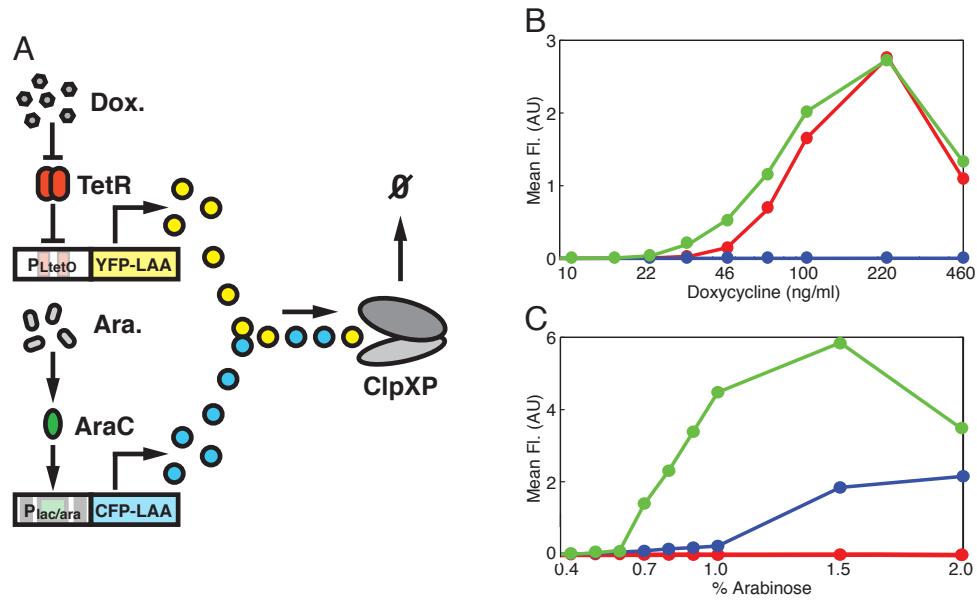
destroyed (Maurizi et al., 1990; Wojtkowiak et al., 1993; Sauer et al., 2004; Levchenko et al., 2000). An important role of ClpXP is to target proteins that have stalled in the ribosome; these incomplete proteins are marked with a well-characterized *ssrA* tag (Keiler et al., 1996). Because ClpXP is involved in regulating the activity of metabolic enzymes as well as increasing degradation rates in stress response, thousands of cellular substrates (Damerau and St John, 1993; Schweder et al., 1996; Flynn et al., 2003; Neher et al., 2006) are often competing to be processed by only hundreds of degradation machines (Farrell et al., 2005). This results in a “waiting line” for degradation and can produce a coupling of cellular components that is unexpected and unaccounted for in the context of traditional gene and protein interaction maps.

To mimic a cellular environment in which ClpXP machinery is saturated, causing target molecules to enter a processing queue, we constructed a synthetic system to over-express two different tagged proteins from separate and uncorrelated promoters (Fig. 6.2A). The  $P_{LtetO-1}$  promoter, used to drive expression of YFP (yeast-enhanced venus fluorescent protein, (Raser and O’Shea, 2004)), is tightly repressible by the *Tet* repressor (TetR) and can be regulated over an up to 5000-fold range by supplying doxycycline to the culture (Lutz and Bujard, 1997). The hybrid promoter,  $P_{lac/ara-1}$ , used to drive expression of CFP (yeast-enhanced cerulean fluorescent protein, (Raser and O’Shea, 2004)), is tightly repressed by the *Lac* repressor (LacR) and activated by AraC and can be regulated over an up to 1800-fold range in the presence of IPTG and arabinose in the culture. For maximum expression off of the  $P_{lac/ara-1}$  promoter, we used 1mM of IPTG in all samples and used various levels of arabinose to tune the induction level of CFP. Both YFP and CFP were tagged on their C terminus with the well-characterized 11-residue “LAA” tag (AANDENYALAA), marking them as targets for rapid degradation by ClpXP (Keiler et al., 1996). In order to ensure stable synthesis and maintenance of the regulatory proteins, the synthetic system was transformed into an *E. coli* DH5 $\alpha$ Z1 host that produces constitutive levels of TetR, LacR, and AraC off of the chromosome (Lutz and Bujard, 1997). Our analysis of the process of enzymatic decay of highly expressed proteins predicts that at certain levels of expression, the mean level

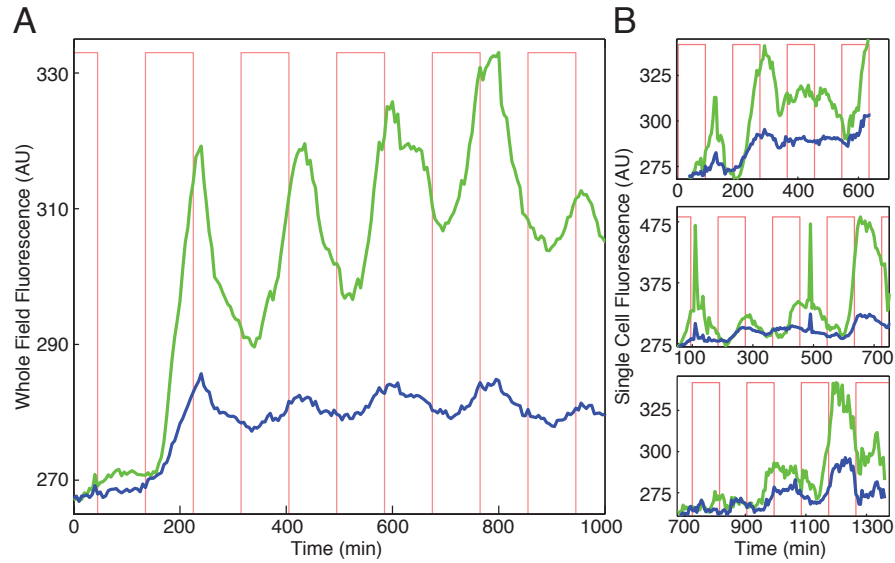
of one protein will be significantly coupled to the mean of the other.

We used two-color flow cytometry to generate induction curves for the two inducible systems. As expected, the mean CFP expression level was not induced by the presence of doxycycline alone (Fig. 6.2B, blue curve), but the mean YFP expression showed typical induction characteristics in response to doxycycline (Fig. 6.2B, red curve). Interestingly, the addition of both arabinose and doxycycline to the medium increased the mean expression level of YFP for most doxycycline levels (Fig. 6.2B, green curve). That is, by inducing expression of CFP at the same time that YFP expression is induced, the ClpXP degradation machines are faced with increasingly overloaded levels of tagged proteins, resulting in a decreased overall degradation rate and increased mean levels of both proteins. A similar result is observed in response to arabinose induction (Fig. 6.2C). In this case, YFP expression is observed to show no response to arabinose alone, as expected (Fig. 6.2C, red curve), while CFP mean expression does increase in response to arabinose alone (Fig. 6.2C, blue curve). Again, we can clearly observe the effect of inducing both promoters simultaneously, as the addition of both arabinose and doxycycline dramatically increases the induction characteristics of CFP (Fig. 6.2C, green curve).

In order to further investigate the implications for enzymatic queueing, we designed a novel microfluidic platform to drive and monitor the signaling responses of a “master” and “slave” system at the single-cell level. We drove production of the master YFP circuit with a square-wave signal of doxycycline (see Materials and Methods below) and used two-color microscopy to observe the response of both the master and the slave CFP systems. Whole-field fluorescence of a population of *E. coli* cells demonstrates how the coupling of the two proteins through the shared degradation pathway serves to drive the slave system in response to the behavior of the master system, as both the YFP (Fig. 6.3A, green) and CFP (Fig. 6.3A, blue) signals oscillate with the driving signal (Fig. 6.3A, red). Similar trends can be observed in the fluorescence trajectories of individual cells (Fig. 6.3B). Correlated behavior is observed in response to the driving signal as well as in long-term expression trends.



**Figure 6.2:** Steady-state behavior of a synthetic signaling network. (A) Two independently produced fluorescent proteins are coupled only by degradation. LAA-tagged YFP is produced by the  $P_{LtetO}$  promoter, which is repressed by *tetR* in the absence of doxycycline. LAA-tagged CFP is produced by the  $P_{lac/ara}$  promoter, which is activated by *araC* in the presence of arabinose. (B) Mean YFP production of a large population *E. coli*. The synthetic system showed no YFP response to the addition of arabinose (blue curve), a typical induction response to the addition of doxycycline (red), and an increased response to the presence of both inducers, indicating the overloaded burden on the clpXP machinery as both CFP and YFP were being produced simultaneously. (C) Mean CFP production of a large population *E. coli*. The synthetic system showed no CFP a typical induction response to the addition of arabinose (blue curve), no response to the addition of doxycycline (red), and an increased response to the presence of both inducers.



**Figure 6.3:** Dynamic behavior of a synthetic signaling network. (A) Using a microfluidic platform capable of generating a time-dependent induction signal, a large population of *E. coli* expressing the synthetic network were subject to a series of doxycycline pulses (red). The YFP and CFP production, integrated over the entire colony, demonstrates the direct response of the  $P_{LtetO}$  promoter, producing YFP (green) as *tetR* is periodically deactivated, as well as the indirect response of the CFP signal (blue) due to the time-dependent saturation of the clpXP machinery. (B) Trajectories for several individual cells demonstrate the response at a single-cell level.

The steady-state induction characteristics and dynamic signaling properties of the two synthetic networks provide an unambiguous demonstration of how queueing for enzymatic processing can induce correlations when there is a significant abundance of proteins relative to the number of functional proteases. If such behavior were observed in a native or uncharacterized system, it would likely be assumed that these two proteins were correlated in a direct way, such as by coordinated gene expression or a protein-protein interaction. However, our results demonstrate the strong correlated behavior that can be observed between two components that are only indirectly coupled via an overloaded enzymatic process. The observed behavior suggests the importance of considering this type of coupling as a form of post-transcriptional regulation when compiling and interpreting genetic wiring diagrams, as this could provide new connections in existing networks or help to understand existing but currently inexplicable links.

## 6.3 Materials and Methods

### 6.3.1 Plasmid Construction

The plasmid pNO-2CLAA was constructed using the pZE24-mcs2a cloning plasmid (Lutz and Bujard, 1997), which has a kanamycin resistance marker and the hybrid  $P_{ara/lac-1}$  promoter upstream of a multiple cloning site (mcs). The sequence for CFP was tagged by PCR with a carboxy-terminal ssrA tag (AANDENYALAA) (Keiler et al., 1996) and inserted between the KpnI and HindIII sites of the mcs, creating pZE42-CFP-LAA. The YFP fragment was similarly tagged and inserted onto the pZS31-mcs2a cloning plasmid, which contains a chloramphenicol resistance marker and the  $P_{LtetO-1}$  promoter upstream of a multiple cloning site. The fragment of this plasmid containing the marker gene and  $P_{LtetO-1}$  driving YFP was copied by PCR and inserted into the SacI site of pZE24-CFP-LAA, creating the final plasmid pNO-2CLAA, containing the independently controlled fluorescent proteins.



### 6.3.2 Flow Cytometry

Flow cytometry data was taken with a Becton-Dickinson LSR II Cell Analyzer, fitted with 405nm and 488nm lasers. The cells were grown overnight in non-inducing medium: LB plus kanamycin for plasmid selection. The cells were passed in the morning into LB plus kanamycin plus various levels of inducer, either doxycycline, arabinose/IPTG, or both. The cells were grown in a 37°C shaker at 300rpm. After 3 hours, the cells were harvested by centrifugation, resuspended in sterile PBS, and put on ice until they were ready to be sampled. Using the LSR II, 100,000 cells were assayed, and the mean CFP and YFP fluorescence was calculated using MATLAB (The MathWorks, Inc.).

### 6.3.3 Microscopy

Image acquisition was performed on a Nikon Eclipse TI epifluorescent inverted microscope outfitted with fluorescence filter cubes optimized for YFP and CFP imaging and a phase-contrast based autofocus algorithm. Images were acquired using a Photometrics CoolSNAP HQ2 cooled CCD camera, controlled by Nikon Elements software. Images were acquired every 1 minute in phase contrast, in order to provide the optimal temporal coverage to suit the automated tracking program. Fluorescent images were acquired every five minutes. The cells were imaged inside a microfluidic with an upstream switch, with the ability to mix or switch between two different media sources. A custom application written in LabVIEW (National Instruments, Austin, Texas) controlled linear actuators, to which two reservoirs containing inducing and non-inducing medium were attached. Using this algorithm, a square wave of 3 hour period was generated, subjecting the cells to alternating 90 minute terms of inducer or non-inducer.

## 6.4 Acknowledgements

Chapter 6 contains material originally submitted for publication as Cookson, N., Mather, W., Mondragon-Palomino, O., Williams, R., Tsimring, L. and Hasty, J., 2009: Queuing up for enzymatic processing: Correlated signaling through coupled degradation. *Science*, submitted. Copyright permission to republish here was granted by AAAS.

# Chapter 7

## Summary

The field of synthetic biology is growing rapidly as there is increasing success in the application of computational techniques to describe and predict dynamic cellular behavior. This progress is aided by the continuing development of powerful tools in molecular biology, microfluidic, and computational techniques. The ability to construct networks from simple biological components, based on computational guidelines and predictions will require a thorough understanding of the basic cellular functions that underly complex biological networks. Here, we have presented several studies aimed at probing the dynamic behavior of the model organisms *Saccharomyces cerevisiae* and *Escherichia coli* at the single cell level.

In Chapter 2, we demonstrated the utility of a novel microchemostat designed for monitoring long-term gene expression dynamics in *S. cerevisiae* (Cookson\*, Ostroff\*, Pang\*, Volfson, and Hasty, 2005, \*equal contribution). This device provides many benefits over traditional methods for measuring cellular fluorescence, such as flow cytometry and traditional microscopy. Flow cytometry provides a snapshot of single-cell data but does not offer the ability to track any given cell. These population measurements provide very useful information about the steady-state behavior of biological networks, but they inherently sacrifice details at the single-cell dynamical level. Traditional microscopy assays yield useful and very detailed information on the dynamical behavior

of a small number of cells but typically run only for a short time period and sacrifice the ability to generate good statistics over a population. In contrast, the T $\mu$ C allows for long-term single-cell dynamical measurements over a large population. Within the context of systems biology, the ability to generate such data at the single-cell level will aid in the development of predictive dynamical modeling (Chen et al., 2004) and facilitate the application of novel approaches such as the use of frequency-space analysis to quantify the variability inherent in gene expression (Simpson et al., 2003). The unique capability of the T $\mu$ C for studying both the dynamics and the variability of biological processes within a population of living cells represents an important step toward bringing quantitative single-cell data to the field of systems biology.

In Chapter 3, we studied the biological origins of variable gene expression within a clonal population of yeast cells (Volfson\*, Marciniak\*, Ostroff, Blake, Tsimring, and Hasty, 2006, \*equal contribution). We combined computational modeling with fluorescence data generated from multiple promoter-gene inserts in *S. cerevisiae* to identify two major sources of extrinsic variability. One unavoidable source arising from the coupling of gene expression with population dynamics leads to a ubiquitous noise floor in expression variability. A second source which is modeled as originating from a common upstream transcription factor exemplifies how regulatory networks can convert noise in upstream regulator expression into extrinsic noise at the output of a target gene (Pedraza and van Oudenaarden, 2005b). Our results highlight the importance of the interplay of gene regulatory networks with population heterogeneity for understanding the origins of cellular diversity.

In Chapter 4, we monitored the response of *S. cerevisiae* metabolic gene regulation to periodic changes in the external carbon source by utilizing a microfluidic platform that allows precise, dynamic control over environmental conditions (Bennett\*, Pang\*, Ostroff, Baumgartner, Nayak, Tsimring, and Hasty, 2008, \*equal contribution). We found that the metabolic system acts as a low-pass filter that reliably responds to a slowly changing environment, while effectively ignoring fluctuations that are too fast for the cell to mount an efficient response. We used computational modeling calibrated

with experimental data to determine how frequency selection in the system is controlled by the interaction of coupled regulatory networks governing the signal transduction of alternative carbon sources. Experimental verification of model predictions led to the discovery of two novel properties of the regulatory network. First, we revealed a previously unknown mechanism for post-transcriptional control, by demonstrating that two key transcripts are degraded at a rate that depends on the carbon source. Second, we compared two *S. cerevisiae* strains and find that they exhibit the same frequency response despite having markedly different induction characteristics. Our results suggest that while certain characteristics of the complex networks may differ when probed in a static environment, the system has been optimized for a robust response to a dynamically changing environment.

In Chapter 5 we used fluorescence microscopy to acquire single cell volume trajectories for a large population of *S. cerevisiae* cells (Cookson, N., Cookson, S., Tsimring, and Hasty, 2009). Using this data, we generated a comprehensive set of statistics that govern the growth and division of these cells over many generations, and we discovered several interesting trends in their size, growth, and protein production characteristics. We used these statistics to develop an accurate model of cell cycle volume dynamics, starting at cell birth. Finally, we demonstrated the importance of tracking volume fluctuations by combining cell division dynamics with a minimal gene expression model for a constitutively expressed fluorescent protein. The significant oscillations in the cellular concentration of a stable, highly expressed protein mimicked the observed experimental trajectories and demonstrated the fundamental impact that the cell cycle has on cellular functions.

In Chapter 6 we investigated a mechanism of indirect protein coupling that effectively increases the connectivity of biological networks (Cookson\*, Mather\*, Mondragon-Palomino, Williams, Tsimring, and Hasty, 2009, \*equal contribution). The indirect coupling we studied arises from an abundance of target molecules relative to a limited number of processing components, resulting in “waiting lines” that occur in the context of queuing theory. Using the *E. coli* clpXP degradation machine as a model system, we

constructed a synthetic signaling network to demonstrate the unexpected coupling of two otherwise independent components, resulting from being subject to a degradation queue. Our results highlight the importance of considering indirect sources of coupling between biological components, as they can yield observed correlated behavior that is strikingly similar to that of more directly connected components.

Together, these studies highlight the importance of studying the dynamics of biological systems, using methods that allow the tracking of long-term behavior in single cells. The ability to take high-resolution data over long time periods for comparison with computational modeling allows us to elucidate novel properties of biological systems. The synthetic biology approach will become increasingly effective as these technologies continue to improve and we gain more information about fundamental cellular behavior at the single-cell level. By understanding these basic functions, we hope to reveal design principles that are key to driving and regulating dynamic behavior. This will allow the advanced development of synthetic networks from the ground up, that are capable of dynamic behavior that can be controlled and accurately predicted. Ultimately, we can envision applications in the development of “personalized” medicine, with the ability to construct simple gene modules that could be integrated into a diseased cell to perform the function of a damaged or missing cellular component, returning the system to a healthy state.

# References

- Acar, M., Becskei, A., and van Oudenaarden, A., 2005: Enhancement of cellular memory by reducing stochastic transitions. *Nature*, **435**(7039), 228–232.
- Adalsteinsson, D., McMillen, D., and Elston, T. C., 2004: Biochemical Network Stochastic Simulator (BioNetS): software for stochastic modeling of biochemical networks. *BMC Bioinformatics*, **5**(1), 24.
- Adams, R., and Bischof, L., 1994: Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, **16**(6), 641–7.
- Alon, U., Barkai, N., Notterman, D., Gish, K., Ybarra, S., Mack, D., and Levine, A., 1999: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences*, **96**(12), 6745.
- Andrade, R. P., Kötter, P., Entian, K. D., and Casal, M., 2005: Multiple transcripts regulate glucose-triggered mRNA decay of the lactate transporter *JEN1* from *Saccharomyces cerevisiae*. *Biochem. Bioph. Res. Co.*, **332**, 254–262.
- Arkin, A., Ross, J., and McAdams, H. H., 1998: Stochastic kinetic analysis of developmental pathway bifurcation in phage  $\lambda$ -infected *Escherichia coli* cells. *Genetics*, **149**(4), 1633–48.
- Atkinson, M. R., Savageau, M. A., Myers, J. T., and Ninfa, A. J., 2003: Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, **113**(5), 597–607.
- Austin, D., Allen, M., McCollum, J., Dar, R., Wilgus, J., Sayler, G., Samatova, N., Cox, C., and Simpson, M., 2006: Gene network shaping of inherent noise spectra. *Nature*, **439**(7076), 608–611.
- Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S., 2004: Bacterial persistence as a phenotypic switch. *Science*, **305**(5690), 1622–5.

- Balagaddé, F. K., You, L., Hansen, C. L., Arnold, F. H., and Quake, S. R., 2005: Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science*, **309**(5731), 137–40.
- Bartek, J., Lukas, C., and Lukas, J., 2004: Checking on DNA damage in S phase. *Nat. Rev. Mol. Cell Biol.*, **5**(10), 792–804.
- Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R., 2005: A synthetic multicellular system for programmed pattern formation. *Nature*, **434**(7037), 1130–4.
- Bayer, T., Hoff, K., Beisel, C., Lee, J., and Smolke, C., 2009: Synthetic control of a fitness tradeoff in yeast nitrogen metabolism. *Journal of biological engineering*, **3**(1), 1.
- Beadle, G. W., and Tatum, E. L., 1941: Genetic control of biochemical reactions in *Neurospora*. *Proc. Natl. Acad. Sci. USA*, **27**, 499–506.
- Becskei, A., and Serrano, L., 2000: Engineering stability in gene networks by autoregulation. *Nature*, **405**(6786), 590–3.
- Begley, T. J., and Samson, L. D., 2004: Network responses to DNA damaging agents. *DNA Repair (Amst)*, **3**(8–9), 1123–32.
- Bendib, S., and Français, O., 2001: Analytical study of microchannel and passive microvalve “application to micropump simulator”. In *Design, Characterisation, and Packaging for MEMS and Microelectronics 2001*, 283–91. Adelaide, Australia.
- Bennett, M. R., Pang, W. L., Ostroff, N. A., Baumgartner, B. L., Nayak, S., Tsimring, L. S., and Hasty, J., 2008: Metabolic gene regulation in a dynamically changing environment. *Nature*, **454**(7208), 1119–22.
- Bird, R. B., Stewart, W. E., and Lightfoot, E. N., 2002: *Transport Phenomena*. J. Wiley, New York, 2nd edition.
- Blake, W., Balázsi, G., Kohanski, M., Isaacs, F., Murphy, K., Kuang, Y., Cantor, C., Walt, D., and Collins, J., 2006: Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular cell*, **24**(6), 853–865.
- Blake, W. J., Kærn, M., Cantor, C. R., and Collins, J. J., 2003: Noise in eukaryotic gene expression. *Nature*, **422**(6932), 633–7.
- Boles, E., and Hollenberg, C. P., 1997: The molecular genetics of hexose transport in yeasts. *FEMS Microbiol. Rev.*, **21**, 85–111.
- Bramson, M., 1998: State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems*, **30**(1-2), 89–148.



- Bramson, M., and Dai, J. G., 2001: Heavy traffic limits for some queueing networks. *Annals of Applied Probability*, **11**(1), 49–90.
- Breeden, L. L., 2003: Periodic transcription: a cycle within a cycle. *Curr. Biol.*, **13**(1), R31–8.
- Carrera, J., Rodrigo, G., and Jaramillo, A., 2009: Towards the automated engineering of a synthetic genome. *Molecular BioSystems*, **5**(7), 733–743.
- Chen, K. C., Calzone, L., Csikasz-Nagy, A., Cross, F. R., Novak, B., and Tyson, J. J., 2004: Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell*, **15**(8), 3841–62.
- Cohen, B. A., Mitra, R. D., Hughes, J. D., and Church, G. M., 2000: A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat. Genet.*, **26**(2), 183–6.
- Cook, D. L., Gerber, A. N., and Tapscott, S. J., 1998: Modeling stochastic gene expression: implications for haploinsufficiency. *Proc. Natl. Acad. Sci. U. S. A.*, **95**(26), 15641–6.
- Cookson, S., Ostroff, N., Pang, W. L., Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Sys. Biol.*, **1**(1), msb4100032:E1–E6.
- Cookson, S., Ostroff, N., Pang, W. L., Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032–E1–6. (\*equal contribution).
- Crocker, J. C., and Grier, D. G., 1996: Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.*, **179**(1), 298–310. See also “Particle tracking using IDL” online tutorial developed by E. Weeks, <http://www.physics.emory.edu/weeks/idl/>.
- Damerau, K., and St John, A., 1993: Role of clp protease subunits in degradation of carbon starvation proteins in escherichia coli. *Journal of bacteriology*, **175**(1), 53.
- de Atauri, P., Orrell, D., Ramsey, S., and Bolouri, H., 2004: Evolution of ‘design’ principles in biochemical networks. *Syst. Biol.*, **1**(1), 28–40.
- de la Cruz, B. J., Prieto, S., and Scheffler, I. E., 2002: The role of the 5’ untranslated region (UTR) in glucose-dependent mRNA decay. *Yeast*, **19**, 887–902.
- Demir, O., and Kurnaz, I. A., 2006: An integrated model of glucose and galactose metabolism regulated by the GAL genetic switch. *Comp. Biol. Chem.*, **30**, 179–192.

- Dertinger, S. K. W., Chiu, D. T., Jeon, N. L., and Whitesides, G. M., 2001: Generation of gradients having complex shapes using microfluidic networks. *Anal. Chem.*, **73**(6), 1240–6.
- Douglas, H. C., and Hawthorne, D. C., 1966: Regulation of genes controlling synthesis of the galactose pathway enzymes in yeast. *Genetics*, **54**, 911–916.
- Duffy, D. C., Schueller, O. J. A., Brittain, S. T., and Whitesides, G. M., 1999: Rapid prototyping of microfluidic switches in poly(dimethyl siloxane) and their actuation by electro-osmotic flow. *J. Micromech. Microeng.*, **9**(3), 211–7.
- Egilmez, N., Chen, J., and Jazwinski, S., 1990: Preparation and partial characterization of old yeast cells. *Journal of Gerontology*, **45**(1), B9.
- Egilmez, N. K., and Jazwinski, S. M., 1989: Evidence for the involvement of a cytoplasmic factor in the aging of the yeast *Saccharomyces cerevisiae*. *J. Bacteriol.*, **171**(1), 37–42.
- Elowitz, M. B., and Leibler, S., 2000: A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**(6767), 335–8.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S., 2002: Stochastic gene expression in a single cell. *Science*, **297**(5584), 1183–6.
- Farrell, C., Grossman, A., and Sauer, R., 2005: Cytoplasmic degradation of ssra-tagged proteins. *Molecular Microbiology*, **57**(6), 1750–1761.
- Flynn, J., Neher, S., Kim, Y., Sauer, R., and Baker, T., 2003: Proteomic discovery of cellular substrates of the clpx protease reveals five classes of clpx-recognition signals. *Molecular cell*, **11**(3), 671–683.
- Fu, A. Y., Chou, H. P., Spence, C., Arnold, F. H., and Quake, S. R., 2002: An integrated microfabricated cell sorter. *Anal. Chem.*, **74**(11), 2451–7.
- Gardner, T. S., Cantor, C. R., and Collins, J. J., 2000: Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**(6767), 339–42.
- Gardner, T. S., di Bernardo, D., Lorenz, D., and Collins, J. J., 2003: Inferring genetic networks and identifying compound mode of action via expression profiling. *Science*, **301**(5629), 102–5.
- Gibson, D., Benders, G., Andrews-Pfannkoch, C., Denisova, E., Baden-Tillson, H., Zaveri, J., Stockwell, T., Brownley, A., Thomas, D., Algire, M., Merryman, C., Young, L., Noskov, V., Glass, J., Venter, J., Hutchison, C. I., and Smith, H., 2008: Complete chemical synthesis, assembly, and cloning of a mycoplasma genitalium genome. *Science's STKE*, **319**(5867), 1215.

- Gillespie, D. T., 1977: Exact stochastic simulation of coupled chemical-reactions. *J. Phys. Chem.*, **81**(25), 2340–61.
- Golub, T., Slonim, D., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J., Coller, H., Loh, M., Downing, J., Caligiuri, M., et al., 1999: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**(5439), 531.
- Groisman, A., Lobo, C., Cho, H., Campbell, J. K., Dufour, Y. S., Stevens, A. M., and Levchenko, A., 2005a: A microfluidic chemostat for experiments with bacterial and yeast cells. *Nat. Methods*, **2**(9), 685–9.
- Groisman, A., Lobo, C., Cho, H., Campbell, J. K., Dufour, Y. S., Stevens, A. M., and Levchenko, A., 2005b: A microfluidic chemostat for experiments with bacterial and yeast cells. *Nature Methods*, **2**, 685–689.
- Grossman, A. D., 1995: Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet*, **29**, 477–508.
- Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S., 2002: Combinatorial synthesis of genetic networks. *Science*, **296**(5572), 1466–70.
- Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., Elston, T. C., and Collins, J. J., 2006: A bottom-up approach to gene regulation. *Nature*, **439**(7078), 856–60.
- Haberman, R., 1998: *Elementary Applied Partial Differential Equations: with Fourier Series and Boundary Value Problems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition.
- Hansen, C., and Quake, S. R., 2003: Microfluidics in structural biology: smaller, faster ... better. *Curr. Opin. Struct. Biol.*, **13**(5), 538–44.
- Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W., and Friend, S. H., 1997: Integrating genetic approaches into the discovery of anticancer drugs. *Science*, **278**(5340), 1064–8.
- Hartwell, L. H., and Unger, M. W., 1977: Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. *J. Cell Biol.*, **75**(2 Pt 1), 422–35.
- Hasty, J., Isaacs, F., Dolnik, M., McMillen, D., and Collins, J. J., 2001a: Designer gene networks: towards fundamental cellular control. *Chaos*, **11**(1), 207–20.
- Hasty, J., McMillen, D., and Collins, J. J., 2002: Engineered gene circuits. *Nature*, **420**(6912), 224–30.

- Hasty, J., McMillen, D., Isaacs, F., and Collins, J. J., 2001b: Computational studies of gene regulatory networks: *in numero* molecular biology. *Nat. Rev. Genet.*, **2**(4), 268–79.
- Hong, J. W., Studer, V., Hang, G., Anderson, W. F., and Quake, S. R., 2004: A nanoliter-scale nucleic acid processor with parallel architecture. *Nat. Biotechnol.*, **22**(4), 435–9.
- Iafolla, M., and McMillen, D., 2006: Extracting biochemical parameters for cellular modeling: A mean-field approach. *J. Phys. Chem. B*, **110**(43), 22019–22028.
- Ibarra, R. U., Edwards, J. S., and Palsson, B. O., 2002: *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature*, **420**(6912), 186–9.
- Ideker, T., Thorsson, V., Ranish, J., Christmas, R., Buhler, J., Eng, J., Bumgarner, R., Goodlett, D., Aebersold, R., and Hood, L., 2001a: Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, **292**(5518), 929.
- Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L., 2001b: Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, **292**(5518), 929–34.
- Isaacs, F. J., Hasty, J., Cantor, C. R., and Collins, J. J., 2003: Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. U. S. A.*, **100**(13), 7714–9.
- Jacob, F., and Monod, J., 1961: Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, **3**, 318–56.
- Kaern, M., Elston, T., Blake, W., and Collins, J., 2005: Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, **6**, 451–464.
- Kaniak, A., Xue, Z., Macool, D., Kim, J., and Johnston, M., 2004: Regulatory network connecting two glucose signal transduction pathways in *Saccharomyces cerevisiae*. *Euk. Cell*, **3**, 221–231.
- Keiler, K. C., Waller, P. R. H., and Sauer, R. T., 1996: Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger rna. *Science*, **271**(5251), 990–993.
- Kelly, F. P., 1979: *Reversibility and Stochastic Networks*. John Wiley and Sons, Chichester.
- Kemkemer, R., Schrank, S., Vogel, W., Gruler, H., and Kaufmann, D., 2002: Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 *in vitro*. *Proc. Natl. Acad. Sci. USA*, **99**, 13783–8.

- Kepler, T. B., and Elston, T. C., 2001: Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.*, **81**(6), 3116–36.
- Kohn, K. W., and Pommier, Y., 2005: Molecular interaction map of the p53 and Mdm2 logic elements, which control the Off-On switch of p53 in response to DNA damage. *Biochem. Biophys. Res. Commun.*, **331**(3), 816–27.
- Kruse, K., and Julicher, F., 2005: Oscillations in cell biology. *Curr. Opin. Cell Biol.*, **17**(1), 20–26.
- Kussell, E., and Leibler, S., 2005: Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, **309**(5743), 2075–2078. doi: 10.1126/science.1114383.
- Levchenko, I., Seidel, M., Sauer, R., and Baker, T., 2000: A specificity-enhancing factor for the ClpXP degradation machine. *Science*, **289**(5488), 2354.
- Li, S., Armstrong, C., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P., Han, J., Chesneau, A., Hao, T., et al., 2004: A map of the interactome network of the metazoan *C. elegans*. *Science*, **303**(5657), 540.
- Lin, F., Saadi, W., Rhee, S. W., Wang, S. J., Mittal, S., and Jeon, N. L., 2004: Generation of dynamic temporal and spatial concentration gradients using microfluidic devices. *Lab Chip*, **4**(3), 164–7.
- Lipan, O., and Wong, W. H., 2005: The use of oscillatory signals in the study of genetic networks. *Proc. Natl. Acad. Sci. USA*, **102**(20), 7063–8.
- Lutz, R., and Bujard, H., 1997: Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I<sub>1</sub>-I<sub>2</sub> regulatory elements. *Nucleic Acids Res.*, **25**(6), 1203–10.
- Mangan, S., Zaslaver, A., and Alon, U., 2003: The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.*, **334**(2), 197–204.
- Mao, H., Yang, T., and Cremer, P. S., 2002: A microfluidic device with a linear temperature gradient for parallel and combinatorial measurements. *J. Am. Chem. Soc.*, **124**(16), 4432–5.
- Maurizi, M., Clark, W., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., and Gottesman, S., 1990: Sequence and structure of clp p, the proteolytic component of the atp-dependent clp protease of *Escherichia coli*. *Journal of Biological Chemistry*, **265**(21), 12536.

- McMillen, D., Kopell, N., Hasty, J., and Collins, J. J., 2002: Synchronizing genetic relaxation oscillators by intercell signaling. *Proc. Natl. Acad. Sci. U. S. A.*, **99**(2), 679–84.
- McMurray, M. A., and Gottschling, D. E., 2004: Aging and genetic instability in yeast. *Curr. Opin. Microbiol.*, **7**(6), 673–9.
- Michaelis, L., and Menten, M. L., 1913: The kinetics of the inversion effect. *Biochemische Zeitschrift*, **49**, 333–369.
- Neher, S., Villén, J., Oakes, E., Bakalarski, C., Sauer, R., Gygi, S., and Baker, T., 2006: Proteomic profiling of clpxp substrates after dna damage reveals extensive instability within sos regulon. *Molecular cell*, **22**(2), 193–204.
- Nevozhay, D., Adams, R., Murphy, K., Josić, K., and Balázsi, G., 2009: Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences*, **106**(13), 5123.
- Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A., 2002: Regulation of noise in the expression of a single gene. *Nat. Genet.*, **31**(1), 69–73.
- Pedraza, J. M., and van Oudenaarden, A., 2005a: Noise propagation in gene networks. *Science*, **307**(5717), 1965–9.
- Pedraza, J. M., and van Oudenaarden, A., 2005b: Noise Propagation in Gene Networks. *Science*, **307**(5717), 1965–1969.
- Rao, C. V., Wolf, D. M., and Arkin, A. P., 2002: Control, exploitation and tolerance of intracellular noise. *Nature*, **420**(6912), 231–7.
- Raser, J. M., and O’Shea, E. K., 2004: Control of stochasticity in eukaryotic gene expression. *Science*, **304**(5678), 1811–4.
- Reaney, D., 2002: Genetic noise in evolution? *Nature*, **307**, 318–319.
- Rohde, J. R., Trinh, J., and Sadowski, I., 2000: Multiple signals regulate gal transcription in yeast. *Mol. Cell Biol.*, **20**(11), 3880–3886.
- Ronen, M., and Botstein, D., 2006: Transcriptional response of steady-state yeast cultures to transient perturbations in carbon source. *Proc. Natl. Acad. Sci. USA*, **103**, 389–394.
- Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B., 2005a: Gene regulation at the single-cell level. *Science*, **307**(5717), 1962–5.
- Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B., 2005b: Gene Regulation at the Single-Cell Level. *Science*, **307**(5717), 1962–1965.

- Sauer, R., Bolon, D., Burton, B., Burton, R., Flynn, J., Grant, R., Hersch, G., Joshi, S., Kenniston, J., Levchenko, I., et al., 2004: Sculpting the proteome with AAA+ proteases and disassembly machines. *Cell*, **119**(1), 9–18.
- Sauer, U., 2004: High-throughput phenomics: experimental methods for mapping fluxomes. *Current Opinion in Biotechnology*, **15**(1), 58–63.
- Savageau, M. A., 1974: Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, **252**(5484), 546–9.
- Scheffler, I. E., de la Cruz, B. J., and Preito, S., 1998: Control of mRNA turnover as a mechanism of glucose repression in *Saccharomyces cerevisiae*. *Biochem. Cell Biol.*, **30**, 1175–1193.
- Scheibel, T., Bloom, J., and Lindquist, S. L., 2004: The elongation of yeast prion fibers involves separable steps of association and conversion. *Proc. Natl. Acad. Sci. U. S. A.*, **101**(8), 2287–92.
- Schweder, T., Lee, K., Lomovskaya, O., and Matin, A., 1996: Regulation of escherichia coli starvation sigma factor (sigma s) by clp<sub>XP</sub> protease. *Journal of bacteriology*, **178**(2), 470.
- Sheff, M. A., and Thorn, K. S., 2004a: Optimized cassettes for fluorescent protein tagging in *saccharomyces cerevisiae*. *Yeast*, **21**(8), 661–670.
- Sheff, M. A., and Thorn, K. S., 2004b: Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast*, **21**(8), 661–70.
- Simon, I., Barnett, J., Hannett, N., Harbison, C. T., Rinaldi, N. J., Volkert, T. L., Wyrick, J. J., Zeitlinger, J., Gifford, D. K., Jaakkola, T. S., and Young, R. A., 2001: Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell*, **106**(6), 697–708.
- Simon, J. A., and Yen, T. J., 2003: Novel approaches to screen for anticancer drugs using *Saccharomyces cerevisiae*. *Methods Mol. Biol.*, **223**, 555–76.
- Simpson, M. L., Cox, C. D., and Sayler, G. S., 2003: Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci. U. S. A.*, **100**(8), 4551–6.
- Slater, M., Sharrow, S., and Gart, J., 1977: Cell cycle of *saccharomyces cerevisiae* in populations growing at different rates. *Proceedings of the National Academy of Sciences*, **74**(9), 3850–3854.
- Smith, M., Sumner, E., and Avery, S., 2007: Glutathione and gts1p drive beneficial variability in the cadmium resistances of individual yeast cells. *Molecular Microbiology*, **66**(3), 699.

- Sprinzak, D., and Elowitz, M. B., 2005: Reconstruction of genetic circuits. *Nature*, **438**(7067), 443–8.
- Sternberg, P. W., and Felix, M., 1997: Evolution of cell lineage. *Curr Opin Genet Dev*, **7**, 543–50.
- Stricker, J.\*, Cookson, S.\*, Bennett, M. R.\*, Mather, W. H., Tsimring, L. S., and Hasty, J., 2008: A fast, robust and tunable synthetic gene oscillator. *Nature*. doi:10.1038/nature07389. (\*equal contribution).
- Sveen, J. K., and Cowen, E. A., 2004: *PIV and Water Waves*, volume 9 of *Advances in Coastal and Ocean Engineering*, chapter Quantitative imaging techniques and their application to wavy flows, 1–49. World Scientific, Singapore.
- Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J., and Church, G. M., 1999: Systematic determination of genetic network architecture. *Nat. Genet.*, **22**(3), 281–5.
- Tesla, N., 1920: U. S. Patent No. 1,329,559.
- Thattai, M., and Shraiman, B. I., 2003: Metabolic switching in the sugar phosphotransferase system of escherichia coli. *Biophys. J.*, **85**(2), 744–754.
- Thattai, M., and van Oudenaarden, A., 2004: Stochastic gene expression in fluctuating environments. *Genetics*, **167**, 523–530.
- Tigges, M., Marquez-Lago, T., Stelling, J., and Fussenegger, M., 2009: A tunable synthetic mammalian oscillator. *Nature*, **457**(7227), 309–312.
- Tyson, J. J., Chen, K. C., and Novak, B., 2003: Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.*, **15**(2), 221–31.
- Venturoli, D., and Rippe, B., 2005: Ficoll and dextran vs. globular proteins as probes for testing glomerular permselectivity: effects of molecular size, shape, charge, and deformability. *Am. J. Physiol. Renal Physiol.*, **288**(4), F605–13.
- Verma, M., Bhat, O. J., and Venkatesh, K. V., 2005: Steady-state analysis of glucose repression reveals hierarchical expression of proteins under Mig1p control in *Saccharomyces cerevisiae*. *Biochem. J.*, **388**, 843–849.
- Vincent, L., 1993: Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Trans. Image Process.*, **2**(2), 176–201.
- Vogelstein, B., Lane, D., and Levine, A. J., 2000: Surfing the p53 network. *Nature*, **408**(6810), 307–10.



- Volfson, D., Marciniak, J., Blake, W. J., Ostroff, N., Tsimring, L. S., and Hasty, J., 2006: Origins of extrinsic variability in eukaryotic gene expression. *Nature*, **439**(7078), 861–4.
- Wang, X., Hao, N., Dohlman, H. G., and Elston, T. C., 2006: Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade. *Biophys. J.*, **90**(6), 1961–78.
- Whitesides, G. M., Jiang, J., Sia, S., Linder, V., Parviz, B., Sigel, A., and Lee, J., 2004: Soft lithography and bioanalysis. *Abstr. Paper Am. Chem. Soc.*, **227**, U113–U113.
- Whitesides, G. M., Ostuni, E., Takayama, S., Jiang, X. Y., and Ingber, D. E., 2001: Soft lithography in biology and biochemistry. *Ann. Rev. Biomed. Eng.*, **3**, 335–373.
- Williams, R. J., 1998: Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems*, **30**(1-2), 27–88.
- Wojtkowiak, D., Georgopoulos, C., and Zylicz, M., 1993: Isolation and characterization of clpx, a new atp-dependent specificity component of the clp protease of escherichia coli. *Journal of Biological Chemistry*, **268**(30), 22609.
- Woldringh, C. L., Huls, P. G., and Vischer, N. . E., 1993: Volume growth of daughter and parent cells during the cell cycle of *saccharomyces cerevisiae* a/ $\alpha$  as determined by image cytometry. *J. Bact.*, **175**(10), 3174–81.
- Wu, Z. Y., Xanthopoulos, N., Raymond, F., Rossier, J. S., and Girault, H. H., 2002: Polymer microchips bonded by O<sub>2</sub>-plasma activation. *Electrophoresis*, **23**(5), 782–90.
- Xia, Y. N., and Whitesides, G. M., 1998: Soft lithography. *Angew. Chem. Int. Ed. Engl.*, **37**(5), 550–75.