The Interaction of Promoter Chromatin Architecture with the Cell Cycle Regulates Transcription Activation Kinetics

by

Christopher J. Zopf

B.S., Chemical Engineering

Columbia University (2006)

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemical Engineering

at the

Massachusetts Institute of Technology

February 2013

Signature of author: _____
Department of Chemical Engineering
January 31, 2013

Certified by: _____
Narendra Maheshri
Assistant Professor of Chemical Engineering
Thesis Supervisor

Accepted by: _____
Patrick S. Doyle
Professor of Chemical Engineering
Chairman, Committee of Graduate Students

The Interaction of Promoter Chromatin Architecture with the Cell Cycle Regulates Transcription Activation Kinetics

by

Christopher J. Zopf

Submitted to the Department of Chemical Engineering on January 31, 2013
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering

ABSTRACT

The relationship between regulatory *trans*-factors binding a gene's *cis*-regulatory sequence elements and the transcriptional output of that gene is fundamental to even the most complex network behaviors such as metabolism and differentiation. In eukaryotes, chromatin dynamics on gene promoter sequences is an integral part of regulation, and nucleosome remodeling is often required for transcription activation. Though the transient response of these regulated genes is often important in biological contexts, the role of promoter chromatin architecture in activation kinetics is still unclear. We sought to investigate this relationship as well as possible links to the cell cycle, over which chromatin state experiences dramatic changes. To study the activation kinetics of individual promoters, we develop a method to infer real-time transcription rates from protein expression in single *Saccharomyces cerevisiae* cells using time-lapse fluorescence microscopy. Comparison between the instantaneous transcription rate and cell-cycle phase in each cell demonstrates the majority of transcriptional variability is due to cell cycle-dependent effects with noisy expression restricted to S/G2/M. This is in stark contrast to current stochastic models of gene expression, most of which do not account for extrinsic effects, and reveals a permissive activation period beginning each S-phase. We then employ a switchable transactivator system to probe transient response kinetics as a function of promoter chromatin architecture at the *PHO5* promoter, a well-established model system for chromatin-regulated expression. While we show transactivator binding site affinity and location relative to nucleosomes influences promoter response kinetics, the effect is primarily through architecture-dependent reliance on a dominant, permissive activation period in S/G2. Together with similar observations at synthetic promoters using a chimerical, switchable transactivator, these results suggest the cell cycle has a general role in transcription activation. Based on the timing of the permissive period, DNA replication may play a direct role in transactivation. Thus, in network topologies involving noisy genes and positive feedback, the cell cycle-dependent transcription would lead to distinct predictions between frequently- and non-dividing cells. This work reveals an unappreciated yet dominant role for the cell cycle as a general regulator of transcription in eukaryotes with direct implications in better modeling and design of biological networks.

Thesis Supervisor: Narendra Maheshri
Title: Assistant Professor of Chemical Engineering

DEDICATION

To my parents, who cultivated my curiosity, and to my wife, who tolerates it every day.

ACKNOWLEDGEMENT

I wish to thank my advisor, Narendra Maheshri, who has continually served as a wellspring of ideas and enthusiasm. I am a better scientist owing to his guidance in devising, executing, interpreting, and communicating research. I appreciate all he has done to help me develop intellectually and professionally including many opportunities to speak at conferences and allowing me to take time away from my project for a clinical preceptorship. I will always be grateful for his genuine interest that I succeed and enjoy what I do.

I also thank the members of my thesis committee for all of their advice. J. Christopher Love provided welcome focus on many occasions, and Timothy Lu brought a new perspective to the project and to my own way of thinking. I also thank Douglas Lauffenburger for his recommendations at the beginning which helped shape my thesis.

I owe many years of shared thoughts, laughter, and commiseration to my labmates. I thank TL To, Tek Hyung Lee, Bradley Niesner, Shawn Finney-Manchester, Katie Quinn, Nicholas Wren, and Richard Joh for all of their suggestions and help in the lab. Our group was a wonderful family that made coming to work more interesting and helped me get through the longest days. I am also grateful for all of the help Katie and Nick have provided in continuing this work, and I appreciate all of the hard work of my UROPs Emily Jackson and Joshua Zeidman, who helped make great progress in developing the experimental system.

Fortuitously my Chemical Engineering classmates provided a wealth of friendship beyond the white board. From cookouts to ski trips, trivia, skits, and a variety of sports, I will always be grateful for our camaraderie throughout my time here. I especially appreciated the energy of David, Jerry, and Jordi who made sure to get me out of the lab as often as possible.

I am deeply indebted to my parents for providing me with so many opportunities, and their continued support made home feel that much closer. The support of my sisters and brother was always comforting, too.

Most of all, I thank my wife, Stephanie. Thank you for moving to Cambridge to be closer to me. Thank you for bringing me to the lab late at night so my next day would be productive. Thank you for being patient the dozens of times I came home late. Thank you for all that you've done to make the past few years less daunting than if I were facing it alone.

TABLE OF CONTENTS

Abstract 3

Dedication 4

Acknowledgement 5

CHAPTER 1. Introduction 11

1.1 A new perspective on gene regulation 11

1.2 Chromatin and transcription control 14

Promoter nucleosomes pose a barrier to transcription activation 14

1.3 An indirect role for the cell cycle in transcription 18

An S-phase window of opportunity 18

Transcription linked to the cell cycle indirectly through chromatin regulation 18

1.4 References 20

CHAPTER 2. Inferring an Instantaneous Transcription Rate in Single Cells from Fluorescence Time-lapse Microscopy Data 29

2.1 Introduction 29

2.2 Microfluidic culture and fluorescence microscopy to obtain movies of single yeast cells 30

2.3 Single-cell time series data extraction from microscopy movies using GRAFTS 32

Extracting single-cell volume and fluorescence time series from microscopy movies 32

2.4 A model to infer instantaneous transcription rate in single cells 37

Translational capacity is constant across the cell cycle 37

CFP and YFP mature rapidly, while RFP may mature more slowly 38

Spline-fitting to differentiate time series 41

Time accuracy of the calculated transcription transition times. 48

2.5 References 52

CHAPTER 3. Cell-cycle dependence of transcription dominates noise in gene expression .. 55

3.1 Introduction 55

3.2 Constitutive transcription increases in S/G2/M consistent with gene dosage 56

3.3 Random pulses of transcription during "bursty" expression only occur in S/G2/M and are correlated 61

Disaggregating static population distributions of mRNA level by cell-cycle phase confirms G1 inactivity at tetO promoters 66

3.4 Conclusion 67

3.5 References 68

CHAPTER 4. A combination of promoter architecture and cell-cycle phase influences delays in gene activation 73

4.1 Introduction 73

4.2 The *PHO5* promoter as a model for chromatin-regulated transcription 74

4.3 Reengineering the PHO pathway to accurately measure gene activation kinetics of PHO promoters 76

4.4 Activation kinetics is largely dictated by cell cycle phase 86

4.5 Promoter architecture affects cell cycle dependent delay 93

4.6 Memory of previous activation events can eliminate cell-cycle dependent delays 97

4.7 Activation kinetics at synthetic promoters is also cell cycle-dependent 100

4.8 Discussion 103

4.9 References 106

CHAPTER 5. Future Directions 113

5.1 Towards improved stochastic gene expression models 113

5.2 Investigating the permissive transcription period in S/G2 115

5.3 Promoter architecture-dependent activation kinetics: from mechanism to design 117

Towards better design of eukaryotic networks 119

5.4 References 120

Appendix 127

5.5 References 134

CHAPTER 1. INTRODUCTION

1.1 A new perspective on gene regulation

One of the most fascinating aspects of biology is how organisms are able to coordinate expression of multiple genes. Well-orchestrated, combinatorial expression programs enable a cell to metabolize raw materials and energy sources, to replicate, to mount stress responses, and to differentiate and form patterns in higher organisms. To accomplish these feats, organisms have evolved many strategies to regulate gene expression.

A predominant approach is through *trans*-acting transcription factors recognizing and binding their target DNA sequences, where they then act directly or indirectly, through more general cofactor recruitment, to establish a gene's transcription state. These transcription factors can be post-translationally regulated in turn by complex signaling networks, which monitor the cell state. However, DNA packaging onto chromatin in eukaryotes provides an additional layer of regulation. Even in prokaryotes, where the genome is thought to be readily accessible to protein binding, conformational changes due to DNA supercoiling and bacterial "chromatin" can influence gene regulation.

Because nucleosomes can occlude cognate sequences targeted by transcription factors in eukaryotes, the competition between nucleosomes and transcription factors to bind *cis*-regulatory elements in a gene's promoter or enhancer complicates the relationship between transcription factor activity and gene expression. Often, chromatin remodeling is required for induction, and much work (discussed in section 1.2) has been done to elucidate how the presence of chromatin regulates a promoter's response to a steady transcription factor stimulus. However, it is often the transient behavior of regulated genes that is crucial for a biological outcome, including proper stress responses, coordination for differentiation and patterning, and cell-cycle control. Yet the role of promoter chromatin architecture in regulating gene activation kinetics is less well understood.

The main goal of this thesis work is to investigate the role of chromatin dynamics during transcription activation. Specifically, how does promoter architecture, in terms of the relative

positioning of nucleosomes and various *cis*-regulatory elements, influence activation kinetics? Additionally, global processes in the cell may affect the promoter chromatin state. A conspicuous example would be the drastic changes to chromatin architecture throughout the cell cycle discussed in in section 1.3. This thesis therefore also addresses the question of what role, if any, does the cell cycle play in the relationship between promoter architecture and activation kinetics?

The tradeoff between data resolution and throughput makes the kinetic role of promoter architecture a challenging inquiry. Previous studies of activation at chromatin-bound promoters have been limited by population-average measurements, which obscure variability at the level of individual promoters, or by assays too onerous for a comparative survey of various architectures. Therefore, we have developed a moderate-throughput, single-cell approach to provide a promoter-level view of the rate and variability of transcription at a particular gene. The budding yeast *Saccharomyces cerevisiae* provides a model system for studying chromatin-regulated expression. Moreover, it offers straightforward genetic manipulation, and gene expression can be assayed readily using fluorescent protein reporters. We employed time-lapse fluorescence microscopy to obtain a full time-dependent trajectory of reporter expression in single yeast cells. In Chapter 2, we describe the experimental methods used to obtain fluorescence microscopy movies, and the analysis developed to use the single-cell expression time series to infer a corresponding instantaneous transcription rate. The analysis was facilitated by a custom-written graphical user interface for simplified movie analysis, visual data inspection, and automated time series calculations. By accounting for cell-cycle phase based on cell morphology throughout the bright field image series, any correlation to the cell cycle could also be examined.

We first applied these techniques to interrogate the effects of the cell cycle on transcription. As described in Chapter 3, observations of steady-state gene expression at high and low levels revealed an unappreciated dependence of transcription activity on cell-cycle phase. While constitutively active promoters display an instantaneous transcription rate pattern across the cell cycle that increased ~2-fold after replication in S-phase consistent with changes in gene dosage, a synthetic promoter known to have "noisy" expression at low levels was found to be transcriptionally active only after replication and before mitosis. These findings led to the conclusion that the majority of transcriptional variability across a yeast population under steady

conditions is due to cell cycle-dependent changes in transcriptional output. This is in stark contrast to current stochastic models of transcriptional bursting typically used to explain this variability. Furthermore, the dominant influence of cell-cycle phase in transcription pulse timing at steady state suggested a role during transient response activation as well.

To precisely probe the kinetics of promoter activation, we engineered a system in yeast to rapidly and observably switch transactivator activity based on the yeast phosphate starvation pathway. Phosphate-responsive promoters in general and the *PHO5* promoter in particular have served as a model system for studying the relationship between promoter chromatin and transcription regulation (discussed in Chapter 4). With each cell bearing a unique *PHO5* promoter variant driving expression of a fluorescent reporter, the activation response time to a step change in transactivator can be measured at the promoter level. As described in Chapter 4, the cell cycle plays a prominent role in dictating activation kinetics, with a clear permissive activation period that occurs post-budding and replication in S/G2. Conservative changes in promoter architecture revealed the effects of transactivator binding site affinity and location relative to well-positioned nucleosomes had varying effects on accelerating activation kinetics both in slower G1 periods, and in the faster S/G2 periods. Thus, promoter activation kinetics are linked to the cell-cycle in an architecture-dependent manner, which had the following unexpected consequences. Activation timing at two promoters in the same cell became correlated by the common dependence in *trans*. This result was surprising given the expectation that the nucleosome remodeling process in *cis* would dominate the variability in a promoter's response time. Also, transcriptional memory was observed at reactivated promoters, and manifested as a decreased reliance on the permissive S/G2 phase. Moreover, targeting the transactivator to synthetic promoters led to similar observations of an S/G2 permissive activation period, which suggests it is a general phenomenon.

Together these studies reshape our view of gene regulation in eukaryotes. As discussed in section 1.2, a regulatory role of chromatin in repressing or permitting transcription is well-established. The results presented extend our understanding of promoter architecture design rules to response kinetics. However, the dominant role of the cell cycle found here highlights the necessity of considering the global chromatin context in any model of transcription regulation. Given the changes to chromatin throughout the cell cycle and previous evidence of its effect on

coordinating development and viral expression programs as discussed in section 1.3, it is perhaps unsurprising to observe an effect of the cell cycle on transcription. Yet the crucial influence of the cell cycle in noisy expression and activation kinetics is remarkable in that it is almost wholly neglected in the study of genes not directly regulated by the cell cycle. Instead, this work suggests any gene in yeast whose expression is considered to be chromatin-regulated will be dependent on the cell cycle rather than solely the 15% of yeast genes previously shown to be cell cycle-regulated. This has profound implications for network design and function in eukaryotic synthetic biology or metabolic engineering applications. One result of a gene being reliant on a permissive S/G2 phase is that its expression dynamics will be largely dependent on growth rate, which determines the frequency of the S/G2 phase. Likewise, human health features examples of growth-linked, chromatin-regulated genes such as switched and reinforced epigenetic states leading to oncogene expression during tumorigenesis, which may require distinct intervention strategies in dividing and quiescent cells.

1.2 Chromatin and transcription control

Promoter nucleosomes pose a barrier to transcription activation

In eukaryotes, vast amounts of genetic information are condensed into the confines of the nucleus by loading the DNA onto chromatin. The basic packaging unit is the nucleosome. Each nucleosome consists of 147 base pairs (bp) of DNA wrapped around a histone octomer composed of two each H2A/H2B and H3/H4 heterodimers (Richmond and Davey 2003), and a stretch of linker DNA leading to the next nucleosome. At this level, chromatin resembles "beads on a string" and is in its least repressing, euchromatic state while addition of H1 linker histones and/or other scaffolding proteins further condenses the chromatin fiber into a heterochromatic state associated with gene silencing (Lu et al. 2009). However, even when in a more open, euchromatic state, important regulatory DNA sequences can still be bound by a core nucleosome.

Positioning of individual nucleosomes on the DNA is subject to several forces. DNA sequence influences its flexibility, and therefore the energetic favorability of wrapping around the histone core. Long poly(dA/dT) tracts tend to be inflexible, which inhibits nucleosome formation over

that sequence (Segal and Widom 2009). On the other hand, AA/TT/TA dinucleotides spaced at 10 bp intervals with GC dinucleotides repeating 5 bp out of phase facilitate bending and promoter nucleosome formation (Segal et al. 2006; Ioshikhes et al. 2006; Lowary and Widom 1998). Sequence alone cannot account for observed *in vivo* nucleosome positions, though. "Statistical positioning" was proposed to explain the degree of nucleosome organization across the yeast genome (Kornberg 1981; Kornberg and Stryer 1988). This theory proposes that regular nucleosome arrangements could arise on lengths of DNA bounded by barriers to nucleosome formation. Thus a well-positioned nucleosome abutting the barrier would set the pattern and restrict the possible locations of other nucleosomes along the stretch of DNA. Nucleosome positions observed within coding regions are consistent with this notion as positioning strength is strongest at the 5' end and decreases as the array progresses toward the 3' end (Yuan et al. 2005; Mavrich et al. 2008). This arrangement would arise from a barrier at the 5' end possibly in the form of an upstream poly(dA/dT) tract (Yuan et al. 2005) or a well-positioned nucleosome at the transcription start site (Mavrich et al. 2008). A recent model suggests the latter influences positioning across the gene while the former determines the upstream organization (Möbius and Gerland 2010). Additionally, there is a role for sequence-specific *trans*-factors establishing nucleosome positioning through remodeling processes as achieving proper *in vivo* positioning using yeast extracts *in vitro* was ATP-dependent (Zhang et al. 2011).

Combined, these three forces determine the chromatin architecture at a given gene's promoter or enhancer, and distinct regulatory strategies are achieved by the relative positioning of *cis*-regulatory elements and nucleosomes. In yeast, two large classes of promoter architecture are observed: "open" promoters where binding sites for transcription factors are present and accessible in a nucleosome-depleted region (NDR) proximal to the transcription start site, and "covered" promoters where nucleosomes occlude binding sites necessary to achieve transcription (Field et al. 2008; Tirosh and Barkai 2008). These two classes correspond to two different regulatory strategies. Open promoters favor constitutive expression since general transcription factors (GTFs) can be readily recruited, and thus this architecture is enriched for housekeeping genes. Conversely, covered promoters often require recruitment of a targeted transactivator to initiate chromatin remodeling to form a more open promoter state. The remodeling process can be tuned by the promoter architecture-specific competition between nucleosomes and transactivating factors to bind *cis*-regulatory elements at covered promoters. Befitting a class

enriched for stress response genes, remodeling requirements at covered promoters supply a regulation strategy with greater dynamic range, as well as more noise, than the open promoter class (Tirosh and Barkai 2008). Covered promoters not only provide a means to generate regulatory diversity between components in stress response pathways, their inherent noise can lead to phenotypic variation to spread the population across the fitness landscape. This promoter class provides a rich field of study in gene regulation and the basis of this thesis.

The initiating transcription factor may bind a cognate site in the DNA linker or it may have to compete for a nucleosome-occluded binding site. In the latter case, transcription factor binding sites are commonly located near the exit-entry portion of nucleosome-bound DNA (North et al. 2012), and transient, partial DNA unwinding from the histone core would provide occasional access (Anderson and Widom 2000; Polach et al. 2000). In vitro measurements indicate nucleosomes wholly and spontaneously unwrap several times per second (Li et al. 2005) making even deeply buried binding sites available for short periods of time. Rapid histone turnover at promoters (Dion et al. 2007; Rufiange et al. 2007) can also contribute to target site accessibility (Wang et al. 2011). Therefore, the nucleosome mainly represents a thermodynamic obstacle to transactivator binding rather than a kinetic barrier (Hayes and Hansen 2001).

However, at many covered promoters, the thermodynamic barrier posed by nucleosomes prevents immediate formation of a pre-initiation complex, and therefore poses a kinetic barrier to *transcription activation*. While the regulatory role of architecture has been studied in detail at steady state ((Lam et al. 2008), discussed below/in Chapter 4), remodeling has also been proposed as the rate-limiting step in transcription activation (Mao et al. 2010). In some cases, nucleosome eviction alone may be sufficient to allow spontaneous transcription (Han and Grunstein 1988; Ohsawa et al. 2009). Though transactivator binding to both nucleosome-free and occluded sites can occur through energy-independent nucleosome fluctuations, transactivators must precipitate promoter reconfiguration through energy-dependent processes to achieve an active transcription state.

Transactivator-driven remodeling can take place through many pathways, even at the same promoter (Barbaric et al. 2007), and the strength and location of nucleosomes in a given promoter architecture can determine the cofactors required to activate (Hertel et al. 2005;

Dhasarathy and Kladde 2005). Typically cofactor recruitment is done through the transactivator's activation domain, with protein binding abilities distinct from its DNA binding domain. A common conserved motif is an acidic activation domain, characterized by an enrichment of acidic residues, which stimulates transcription across all eukaryotes (Ptashne and Gann 1990). The activation domain directly binds to complexes with chromatin-remodeling and -modifying activities as well as transcription cofactors. Well-studied examples of transactivators in yeast with acidic domains such as Gcn4p, Gal4p, and Pho4p have been shown to directly interact with components of the histone acetyltransferase (HAT) complexes SAGA and NuA4 (Fishburn et al. 2005; Barbaric et al. 2003; Dhasarathy and Kladde 2005; Reeves and Hahn 2005), the chromatin remodeler SWI/SNF (Prochasson et al. 2003; Neely et al. 2002; Dhasarathy and Kladde 2005), and the cofactors Mediator and TFIID (Fishburn et al. 2005; Reeves and Hahn 2005; Herbig et al. 2010). Histone tail lysine acetylation by SAGA (H2B and H3 by Gcn5p) and by NuA4 (all but H3 by Esa1p) (Krebs 2007) may directly destabilize the charge interactions between DNA and histones as well as provide a targeting signal to factors containing bromodomains (Roth et al. 2003). The former mechanism has been called into question by the small effect of deleting histone tails on nucleosome unwrapping (Polach et al. 2000) so the main role of acetylation may be targeting and stabilizing binding of additional factors including ATPase-driven remodeling complexes. While chromatin remodelers can have various activities depending on the subunits associated with the core ATPase, SWI/SNF and INO80 are the main complexes associated with transcription activation in yeast (Clapier and Cairns 2009). These serve to slide and evict nucleosomes from the promoter region through energy-dependent mechanisms. Once the core promoter is finally cleared, the transactivator can also potentiate pre-initiation complex formation through its interactions with cofactors which bridge to components of the RNA Pol II holoenzyme.

The prominent role of energy-dependent processes in relieving chromatin repression at covered promoters suggests activation kinetics are crucial to understanding the overall regulatory dynamics. Since the requirement for various energy-dependent pathways depends on architecture (Dhasarathy and Kladde 2005), the details of a particular promoter architecture should determine activation kinetics as well as the steady state dynamics. As a representative of the covered promoter class, the *PHO5* promoter provides a model system in yeast for the study of chromatin-regulated transcription. Using a set of *PHO5* promoter variants (Lam et al. 2008), we show in

chapter 4 there is indeed a role for architecture in the transient induction kinetics. Surprisingly, activation kinetics are also strongly cell cycle-dependent in an architecture dependent manner.

1.3 An indirect role for the cell cycle in transcription

An S-phase window of opportunity

Global changes in chromatin structure occur throughout the cell cycle. Aside from direct regulation by cyclins, the cell cycle provides phase-specific changes in accessibility that may be important to chromatin-regulated genes. While chromosome condensation during mitosis is associated with gene repression (Johnson 1965), DNA replication during S-phase necessitates an open chromatin structure at least locally and forces changes to the epigenetic state (reviewed in (Alabert and Groth 2012)). As the replication fork progresses along the DNA, nucleosomes are removed by histone chaperones and randomly redistributed between the daughter strands (Annunziato 2012). How H3-H4 tetramers are recycled to the new DNA is not clear, possibly through the chaperones Asf1 and FACT (Alabert and Groth 2012), while the fate of the more dynamic H2A-H2B dimers is even less clear. Full nucleosome loading on both daughter strands, though, requires additional deposition of histones from the soluble pool. These new histones are more highly acetylated than those from the original locus (Annunziato and Seale 1983) generating a new epigenetic state. Histone methylation is also diluted through replication, but may not be reduced sufficiently to alleviate repression (Lanzuolo et al. 2011). Nascent chromatin undergoes a 10 to 20 minute maturation process whereby histone modifications from the original locus are reformed across the new loci (Annunziato 2012). The temporary, post-replication epigenetic state thus provides a "window of opportunity" for activation of many genes, and could be an unappreciated feature of chromatin-regulated genes in general.

Transcription linked to the cell cycle indirectly through chromatin regulation

The notion of a mechanistic link between transcription and DNA accessibility generated by replication has been around for some time (Wolffe 1991). Nascent chromatin provides a window of opportunity for both basal transcription (Almouzni and Wolffe 1993) as well as transactivator action (Kamakaka et al. 1993). In the latter case, opportunistic *trans*-factor binding likely supports formation of a competent promoter state rather than stimulating transcription initiation

itself (Wilson and Patient 1993). In other examples, replication has disrupted epigenetic states to the point of allowing transcription at heterochromatic centromeric repeats (Chen et al. 2008) and a transactivator to overcome subtelomeric silencing (Aparicio and Gottschling 1994).

This DNA replication-linked window of opportunity has been exploited by evolution. Timing of developmental genes is crucial to differentiation and patterning in higher eukaryotes. Replication can "license" an entire locus for coordinated expression (Fisher and Méchali 2003). It can also reactivate inherited silenced genes at a specific embryonic stage after a defined number of replication rounds (Forlani et al. 1998; Bancescu et al. 2004). Many developmental genes are later silenced and only become derepressed oncogenes during tumorigenesis. Unscheduled, accelerated replication can contribute to chromatin derepression both by diluting repressing factors at the oncogene locus and by enhancing activator accessibility (Crowe et al. 2000). In these examples, expression results from a pre-existing epigenetic program or a very rare, stochastic event from silenced chromatin. In either case, chromatin at these loci is tightly regulated.

It is currently unclear whether or not the window of opportunity for transactivators is a global phenomenon. While the several studies mentioned above show some generality in replication-linked transcription from various locations in the genome, the primary examples with phenotypic consequences in development or cancer occur at loci where the activation is part of a specific program or the result of rare misregulation. However, replication-dependent transcription has also been implicated in HIV gene expression (Williams et al. 1996), which integrates broadly into euchromatic regions of the host genome (Wang et al. 2007). In this case, activation through replication is not locus-specific, and must be efficient to effectively initiate infection. Therefore, susceptibility to replication-linked transcription may be a general feature of euchromatic genes. Since these genes are associated with being transcriptionally active, the phenotypic effect of a window of opportunity may not be as apparent as derepressing an oncogene, yet there are instances where replication-linked activation could matter. Proliferating and non-dividing cells would have entirely different expression patterns for a noisy gene regulated by a post-replication active period, and each cell type may activate a stress-response gene at a different rate if replication facilitated transcription. In the following chapters, we demonstrate a general, dominant role for the cell cycle in both noisy, steady state expression and in transcription

activation kinetics, which suggests the replication-linked window of opportunity for transactivators may be, in fact, global.

1.4 References

Adkins M, Howar S, Tyler J. 2004. Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Molecular Cell* **14**: 657–723.

Adkins M, Tyler J. 2006. Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Molecular Cell* **21**: 405–421.

Alabert C, Groth A. 2012. Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology* **13**: 153–67.

Almer A, Rudolph H, Hinnen A, Hörz W. 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. *The EMBO Journal* **5**: 2689–2785.

Almouzni G, Wolffe a P. 1993. Replication-coupled chromatin assembly is required for the repression of basal transcription in vivo. *Genes & Development* **7**: 2033–2047.

Anderson JD, Widom J. 2000. Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *Journal of Molecular Biology* **296**: 979–87.

Annunziato A, Seale R. 1983. Histone deacetylation is required for the maturation of newly replicated chromatin. *J Biol Chem* **258**: 12675–12684.

Annunziato AT. 2012. Assembling chromatin: The long and winding road. *Biochimica et Biophysica Acta* **1819**: 196–210.

Aparicio OM, Gottschling DE. 1994. Overcoming telomeric silencing: a trans-activator competes to establish gene expression in a cell cycle-dependent way. *Genes & Development* **8**: 1133–1146.

Bancescu DL, Glatt-Deeley H, Lalande M. 2004. Epigenetic activation of the 5-hydroxytryptamine (serotonin) receptor 2C in embryonal carcinoma cells is DNA replication-dependent. *Experimental Cell Research* **298**: 262–7.

Barbaric S, Luckenbach T, Schmid A, Blaschke D, Hörz W, Korber P. 2007. Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter in vivo. *The Journal of Biological Chemistry* **282**: 27610–21.

Barbaric S, Reinke H, Horz W. 2003. Multiple Mechanistically Distinct Functions of SAGA at the PHO5 Promoter. *Molecular and Cellular Biology* **23**: 3468–3476.

Barbaric S, Walker J, Schmid A, Svejstrup J, Hörz W. 2001. Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. *The EMBO Journal* **20**: 4944–4995.

Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SIS. 2008. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**: 734–7.

Clapier CR, Cairns BR. 2009. The biology of chromatin remodeling complexes. *Annual Review of Biochemistry* **78**: 273–304.

Crowe AJ, Piechan JL, Sang L, Michelle C, Barton MC. 2000. S-Phase Progression Mediates Activation of a Silenced Gene in Synthetic Nuclei. *Molecular and Cellular Biology* **20**.

Dhasarathy A, Kladde MP. 2005. Promoter Occupancy Is a Major Determinant of Chromatin Remodeling Enzyme Requirements. *Molecular and Cellular Biology* **25**: 2698–2707.

Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**: 1405–8.

Fascher K, Schmitz J, Hörz W. 1990. Role of trans-activating proteins in the generation of active chromatin at the PHO5 promoter in S. cerevisiae. *The EMBO Journal* **9**: 2523–2531.

Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E. 2008. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. ed. U. Ohler. *PLoS Computational Biology* **4**: e1000216.

Fishburn J, Mohibullah N, Hahn S. 2005. Function of a eukaryotic transcription activator during the transcription cycle. *Molecular Cell* **18**: 369–78.

Fisher D, Méchali M. 2003. Vertebrate HoxB gene expression requires DNA replication. *The EMBO Journal* **22**: 3737–48.

Forlani S, Bonnerot C, Capgras S, Nicolas JF. 1998. Relief of a repressed gene expression state in the mouse 1-cell embryo requires DNA replication. *Development (Cambridge, England)* **125**: 3153–66.

Han M, Grunstein M. 1988. Nucleosome loss activates yeast downstream promoters in vivo. *Cell* **55**: 1137–1145.

Hayes JJ, Hansen JC. 2001. Nucleosomes and the chromatin fiber. *Current Opinion in Genetics & Development* **11**: 124–9.

Herbig E, Warfield L, Fish L, Fishburn J, Knutson BA, Moorefield B, Pacheco D, Hahn S. 2010. Mechanism of Mediator recruitment by tandem Gcn4 activation domains and three Gal11 activator-binding domains. *Molecular and Cellular Biology* **30**: 2376–90.

Hertel CB, Langst G, Horz W, Korber P. 2005. Nucleosome Stability at the Yeast PHO5 and PHO8 Promoters Correlates with Differential Cofactor Requirements for Chromatin Opening. *Molecular and Cellular Biology* **25**: 10755–10767.

Ioshikhes IP, Albert I, Zanton SJ, Pugh BF. 2006. Nucleosome positions predicted through comparative genomics. *Nature Genetics* **38**: 1210–5.

Johnson TC. 1965. Ribonucleic acid and protein synthesis in mitotic HeLa Cells. *The Journal of Cell Biology* **27**: 565–574.

Kamakaka RT, Bulger M, Kadonaga JT. 1993. Potentiation of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly. *Genes & Development* **7**: 1779–1795.

Korber P, Barbaric S, Luckenbach T, Schmid A, Schermer UJ, Blaschke D, Hörz W. 2006. The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *The Journal of Biological Chemistry* **281**: 5539–45.

Korber P, Luckenbach T, Blaschke D, Hörz W. 2004. Evidence for histone eviction in trans upon induction of the yeast PHO5 promoter. *Molecular and Cellular Biology* **24**: 10965–74.

Kornberg R. 1981. The location of nucleosomes in chromatin: specific or statistical? *Nature* **292**: 579–580.

Kornberg RD, Stryer L. 1988. Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Research* **16**: 6677–6690.

Krebs JE. 2007. Moving marks: dynamic histone modifications in yeast. *Molecular BioSystems* **3**: 590–7.

Lam FH, Steger DJ, O'Shea EK. 2008. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**: 246–296.

Lanzuolo C, Lo Sardo F, Diamantini A, Orlando V. 2011. PcG complexes set the stage for epigenetic inheritance of gene silencing in early S phase before replication. ed. A. Akhtar. *PLoS Genetics* **7**: e1002370.

Li G, Levitus M, Bustamante C, Widom J. 2005. Rapid spontaneous accessibility of nucleosomal DNA. *Nature Structural & Molecular Biology* **12**: 46–53.

Lowary P., Widom J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of Molecular Biology* **276**: 19–42.

Lu X, Wontakal SN, Emelyanov A V, Morcillo P, Konev AY, Fyodorov D V, Skoultchi AI. 2009. Linker histone H1 is essential for Drosophila development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. *Genes & Development* **23**: 452–65.

Mao C, Brown CR, Falkovskaia E, Dong S, Hrabeta-Robinson E, Wenger L, Boeger H. 2010. Quantitative analysis of the transcription control mechanism. *Molecular Systems Biology* **6**: 431.

Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research* **18**: 1073–83.

Möbius W, Gerland U. 2010. Quantitative test of the barrier nucleosome model for statistical positioning of nucleosomes up- and downstream of transcription start sites. ed. E. Segal. *PLoS Computational Biology* **6**: 11.

Neef DW, Kladde MP. 2003. Polyphosphate Loss Promotes SNF/SWI- and Gcn5-Dependent Mitotic Induction of PHO5. *Molecular and Cellular Biology* **23**: 3788–3797.

Neely KE, Hassan AH, Brown CE, Howe L, Workman JL. 2002. Transcription Activator Interactions with Multiple SWI/SNF Subunits. *Molecular and Cellular Biology* **22**: 1615–1625.

North JA, Shimko JC, Javaid S, Mooney AM, Shoffner MA, Rose SD, Bundschuh R, Fishel R, Ottesen JJ, Poirier MG. 2012. Regulation of the nucleosome unwrapping rate controls DNA accessibility. *Nucleic Acids Research* **40**: 10215–27.

Nourani A, Utley R, Allard S, Côté J. 2004. Recruitment of the NuA4 complex poises the PHO5 promoter for chromatin remodeling and activation. *The EMBO Journal* **23**: 2597–3204.

Ohsawa R, Adkins M, Tyler J. 2009. Epigenetic inheritance of an inducibly nucleosome-depleted promoter and its associated transcriptional state in the apparent absence of transcriptional activators. *Epigenetics & Chromatin* **2**: 11.

O'Neill EM, Kaffman A, Jolly ER, O'Shea EK. 1996. Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* **271**: 209–212.

Polach KJ, Lowary PT, Widom J. 2000. Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes. *Journal of Molecular Biology* **298**: 211–23.

Pondugula S, Neef DW, Voth WP, Darst RP, Dhasarathy A, Reynolds MM, Takahata S, Stillman DJ, Kladde MP. 2009. Coupling Phosphate Homeostasis to Cell Cycle-Specific Transcription: Mitotic Activation of Saccharomyces cerevisiae PHO5 by Mcm1 and Forkhead Proteins. *Molecular and Cellular Biology* **29**: 4891–4905.

Prochasson P, Neely KE, Hassan AH, Li B, Workman JL. 2003. Targeting Activity Is Required for SWI/SNF Function In Vivo and Is Accomplished through Two Partially Redundant Activator-Interaction Domains. *Molecular Cell* **12**: 983–990.

Ptashne M, Gann AA. 1990. Activators and targets. *Nature* **346**: 329–31.

Reeves WM, Hahn S. 2005. Targets of the Gal4 transcription activator in functional transcription complexes. *Molecular and Cellular Biology* **25**: 9092–102.

Reinke H, Hörz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Molecular Cell* **11**: 1599–2206.

Richmond TJ, Davey CA. 2003. The structure of DNA in the nucleosome core. *Nature* **423**: 145–50.

Roth SY, Denu JM, Allis CD. 2001. Histone acetyltransferases. *Annual Review of Biochemistry* **70**: 81-120.

Rudolph H. 1987. The Yeast PHO5 Promoter: Phosphate-Control Elements and Sequences Mediating mRNA Start-Site Selection. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 1340–1344.

Rufiange A, Jacques P-E, Bhat W, Robert F, Nourani A. 2007. Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Molecular Cell* **27**: 393–405.

Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, Wang J-PZ, Widom J. 2006. A genomic code for nucleosome positioning. *Nature* **442**: 772–8.

Segal E, Widom J. 2009. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Current Opinion in Structural Biology* **19**: 65–71.

Springer M, Wykoff DD, Miller N, O'Shea EK. 2003. Partially Phosphorylated Pho4 Activates Transcription of a Subset of Phosphate-Responsive Genes. *PLoS Biology* **1**: e28.

Steger D, Haswell E, Miller A, Wente S, O'Shea EK. 2003. Regulation of chromatin remodeling by inositol polyphosphates. *Science* **299**: 114-166.

Svaren J. 1997. Transcription factors vs nucleosomes: regulation of the PH05 promoter in yeast. *Trends in Biochemical Sciences* **22**: 93–97.

Tirosh I, Barkai N. 2008. Two strategies for gene regulation by promoter nucleosomes. *Genome Research* **18**: 1084–1175.

Vogel K, Horz W, Hinnen A. 1989. The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Molelcular and Cellular Biololgy* **9**: 2050–2057.

Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD. 2007. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Research* **17**: 1186–94.

Wang S-S, Zhou B, Zhou J-Q. 2011. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Molecular and Cellular Biology* **31**: 3171–3252.

Williams RD, Lee B a, Jackson SP, Proudfoot NJ. 1996. Activation domains of transcription factors mediate replication dependent transcription from a minimal HIV-1 promoter. *Nucleic Acids Research* **24**: 549–57.

Williams SK, Truong D, Tyler JK. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9000–5.

Wilson AC, Patient RK. 1993. DNA replication facilitates the action of transcriptional enhancers in transient expression assays. *Nucleic Acids Research* **21**: 4296–304.

Wolffe a P. 1991. Implications of DNA replication for eukaryotic gene expression. *Journal of Cell Science* **99 (Pt 2)**: 201–6.

Yuan G-C, Liu Y-J, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**: 626–30.

Zhang Z, Wippo CJ, Wal M, Ward E, Korber P, Pugh BF. 2011. A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332**: 977–80.

CHAPTER 2. INFERRING AN INSTANTANEOUS TRANSCRIPTION RATE IN SINGLE CELLS FROM FLUORESCENCE TIME-LAPSE MICROSCOPY DATA*

2.1 Introduction

Single-cell analysis of gene expression has furthered our understanding in many aspects of gene expression. Static snapshots of fluorescent reporter expression using flow cytometry or microscopy provide useful information on the distribution of single-cell expression, but lack the history and evolution of time series data required to directly inform gene expression dynamics. Fluorescence time-lapse microscopy presents a means to obtain both single-cell measurements and their history. Various experimental and analytical techniques have been developed to obtain and quantify movies of fluorescent reporter expression, thus imparting insights into gene regulation features (see (Locke and Elowitz 2009) for a review) such as cell-to-cell variation (Rosenfeld et al. 2005; Colman-Lerner et al. 2005), transcription initiation and elongation (Larson et al. 2011), transcriptional bursting (Golding et al. 2005; Suter et al. 2011), cell-cycle dependence (Cookson et al. 2010; Zopf et al.), and heritability (Kaufmann et al. 2007). However, obtaining quality single-cell fluorescence time series involves significant technical challenges in culturing a monolayer of cells in a controllable environment and regarding high-throughput quantification of the acquired fluorescence movies.

Here, we describe a procedure to obtain and analyze fluorescence movies of *S. cerevisiae* with no required experience in cell culture device manufacture or in software development. Instead, we employ a commercially available culture device and an easy-to-use, custom-written graphical user interface. The result is a time series for each single cell representing the time-dependent trajectory of a fluorescent reporter's expression. These time series rise and fall as the fluorescent protein is expressed or diluted/degraded, respectively, providing a kinetic view of transcriptional output. Using the rates of change in the expression level in each cell, we infer an instantaneous

* Portions of text and figures in this chapter were drawn from (Zopf and Maheshri) and (Zopf et al.).

transcription rate for each reporter through a lumped-parameter differential equation model. We also demonstrate the transcription rates calculated are reasonably accurate to draw both qualitative and quantitative conclusions in the following chapters.

2.2 Microfluidic culture and fluorescence microscopy to obtain movies of single yeast cells

The first step in obtaining quality, single-cell time series is generating high-quality fluorescence microscopy movies. Perfusion of cells trapped in a microfluidic chamber provides a means to image relatively stationary single cells and precise control over the extracellular environment. Though customized microfluidic culture chambers have been built and successfully employed previously (Cookson et al. 2005; Paliwal et al. 2007; Charvin et al. 2008), we use a commercially available microfluidic device. The system confines cells to monolayer growth and allows continual control of the perfusion environment. We detail here a basic protocol to generate fluorescence time series movies for budding yeast expressing one or more fluorescent reporters. In the following chapters, various changes are made to the culture media before and during microfluidic culture according to the requirements of the experiment, but all other procedures remain the same.

For each time series experiment, cells were picked from a single colony freshly grown on a synthetic solid media (containing no amino acids, or only those for which the strain is auxotrophic) with 2% glucose and agar. These cells were inoculated into synthetic complete (SC) media. The SC base comprised yeast nitrogen base without phosphate (MP Biomedicals, Santa Ana, CA; #4027-812) mixed with monobasic potassium phosphate solution (Sigma-Aldrich P8709) to set phosphate levels, and the pH was lowered to 4. Unless otherwise indicated, SC contained all amino acids, 2% glucose as the carbon source, and 5000 μM orthophosphate. The cells were then grown overnight on a roller drum at 30°C to OD_{600nm} ~0.1, diluted in fresh media, and grown again for 6-8 hrs to OD_{600nm}~0.1. These cells were loaded into a pre-washed Y04C microfluidic plate (CellAsic, Hayward, CA), which was primed with SC with appropriate carbon source and/or phosphate level according to the manufacturer's instructions and the particular experiment. Cells were perfused with SC at 6 psi throughout the experiment. Flow was

controlled using the programmable ONIX system (CellAsic) to rapidly switch between various media conditions discussed in the text.

Cell growth and expression was observed using a Zeiss Axio Observer.Z1 inverted microscope at 63x magnification (Zeiss Plan-Apochromat 63x/1.40 Oil DIC). Bright field (BF), BF out of focus (BFOOF, for segmentation) and fluorescence images were acquired with a Cascade II EMCCD camera (Photometrics, Tuscon, AZ) using MetaMorph software (Molecular Devices, Sunnyvale, CA), a Lumen 200 metal-halide arc lamp (PRIOR Scientific, Rockland, MA) for fluorescence excitation, appropriate filters for CFP, YFP, and RFP (Chroma Technology Corp, Bellows Falls, VT; set 89006), and acquisition settings optimized for rapid time points. Multiple positions in each culture chamber were imaged every 5 min. For each position at the start of each time point, a built-in MetaMorph autofocus module found the focal plane (f.p.) based on the BF channel. The software then acquired the BF image at +1 μm to the f.p. and a BFOOF image at -4 μm to the f.p. This switches the contrast rings surrounding each cell from light to dark, and comparison between these two images is used in segmentation. Finally the software acquired any fluorescence channels required by the experiment: one grayscale image for cyan fluorescent protein (CFP) at the f.p., and a 3-image Z-stack ±0.2 μm around the f.p. for both yellow (YFP) and red (RFP) fluorescent proteins. The Z-stack for the YFP and RFP channels was taken to provide a more accurate nuclear measurement for the switchable transactivator strain in Chapter 4, but this acquisition was preserved in all experiments for consistency in exposure.

Microscopy movies for each acquisition channel were compiled in MetaMorph for analysis in MATLAB (MathWorks, Natick, MA). A maximum projection was taken of each Z-stack at each time point, and was compiled into a movie for both YFP and RFP. In addition, a nuclear mask movie was created by thresholding the RFP movie when the switchable transactivator strain was used. The remaining steps in extracting single cell time series data from the movies were performed using our custom-written software.

2.3 Single-cell time series data extraction from microscopy movies using GRAFTS

We analyzed microscopy movies using a graphical user interface (GUI) –based software package in MATLAB (Mathworks, Natick, MA), dubbed the GUI for Rapid Analysis of Fluorescence Time Series (GRAFTS), to extract time series data for single cells. GRAFTS has similar features to the versatile, open-source software package Cell-ID (Gordon et al. 2007) in segmenting and tracking cells and in extracting fluorescence intensity and geometric information. However, GRAFTS provides important additional features. First, it offers easy interactive editing of segmentation and tracking results to verify data accuracy, rather than just statistical gating of outlier region traces after analysis. Moreover, it extends the analysis to automatically designate lineage and cell-cycle points of interest of budding yeast. Determining when mother and daughter divide to form two independent cell regions is crucial to determining whole cell (mother including any connected bud) measurements throughout the cell cycle (Cookson et al. 2010).

The suite consists of three modules to accomplish these tasks. The first segments cell regions based on the contrast between focused and unfocused bright field images, and allows the user to define and visually test segmentation parameters. The second tracks and measures cell regions through time; automatically assigns lineages; and enables visual inspection and error correction. A simple plotting GUI is included here to quickly query single-cell properties. The third module ascribes bud emergence and division times, and outputs whole cell time series data as well as their first and second time derivatives, which are used to infer transcription rate in the next section. As an example, we show the time series extracted for a single yeast cell expressing CFP constitutively from the *ADH1* promoter (Figure 2.1).

Extracting single-cell volume and fluorescence time series from microscopy movies

Pre-processing movies consisted of registering and segmenting cells. To account for imprecise return to each stage position at successive time points, movies were first registered using a 2D cross-correlation of BF images between time t and $t + 1$. Segmentation relied on the large differences in contrast of the cell wall between a BF (dark cell with light cell wall ring) and BFOOF (light cell with dark outer ring) images, and involved a series of morphological

operations and watershedding to identify single cell regions. A tracking and visualization GUI used a Matlab implementation of the IDL Particle tracking routine (Blair and Dufresne, available at: http://physics.georgetown.edu/matlab/) to assign cell IDs in each frame based on following region centroids. When a bud appears (new region and ID), the corresponding mother is assigned from the nearest neighbor regions. Lineage assignments for each bud are optimized by penalizing potential mothers based on distance between bud and mother perimeters, and based on potential matches to other buds in nearby times. After the initial automated analysis, the GUI then allows rapid visual inspection and manual curation of the accuracy of automated region segmentation, tracking IDs, and lineage assignments before finalizing cell measurement time series.

For each single cell region, we estimated raw cell volume similar to (Cookson et al. 2010): as an ellipsoid of cross-section corresponding to the region and constrained by the microfluidic trapping chamber, but with mothers and buds treated separately (Figure 2.1A & B). We interpreted the average pixel intensity in each region as a raw volumetric concentration (Figure 2.1C) because the depth of field (0.5 μm) was significantly less than the trapping chamber height (3.5 μm) and the fluorescence profile across the cell was flat, rather than the elliptical profile expected if light were captured from the entire cell volume. Both volume and concentration time series for each cell were conservatively smoothed to remove measurement noise using the spline method described in the next section; total protein was estimated as the product of the two (Figure 2.1D).

Though volume and total protein estimation of buds occurred separately from mothers, the data from the bud-mother pair were summed until cytokinesis to represent the whole cell (Figure 2.1A, B & E). We used morphological cues to define transitions between G1 and S/G2/M. Yeast cells pass through START to transition to S-phase, which occurs slightly before a bud appears (Cross 1995), so we estimated the time of bud appearance to approximate the G1/S transition. To do so, we linearly extrapolated the bud volume trace back to zero volume using the bud volume measured at the first five time points the bud was successfully segmented. Clear morphological markers for cytokinesis (the M/G1 transition) were the bud neck narrowing after nuclear division (easily observable when the nucleus is fluorescently labeled) and the formation of a dark line between the mother and daughter (bud) in a bright field image (Figure 2.2). These events strongly correlated with the end of a brief plateau in the bud volume, providing an easy way for

automated identification based on the slope in the bud volume trace (Figure 2.1B). By automatically assigning budding and division times, we have a high-throughput method to calculate volume and total protein for the contiguous whole cell throughout the cell cycle. Basing our model for inferring transcription rate on the total protein, we are able to avoid complications caused by growth-driven dilution and directly relate rates of change in total protein level to total mRNA level.

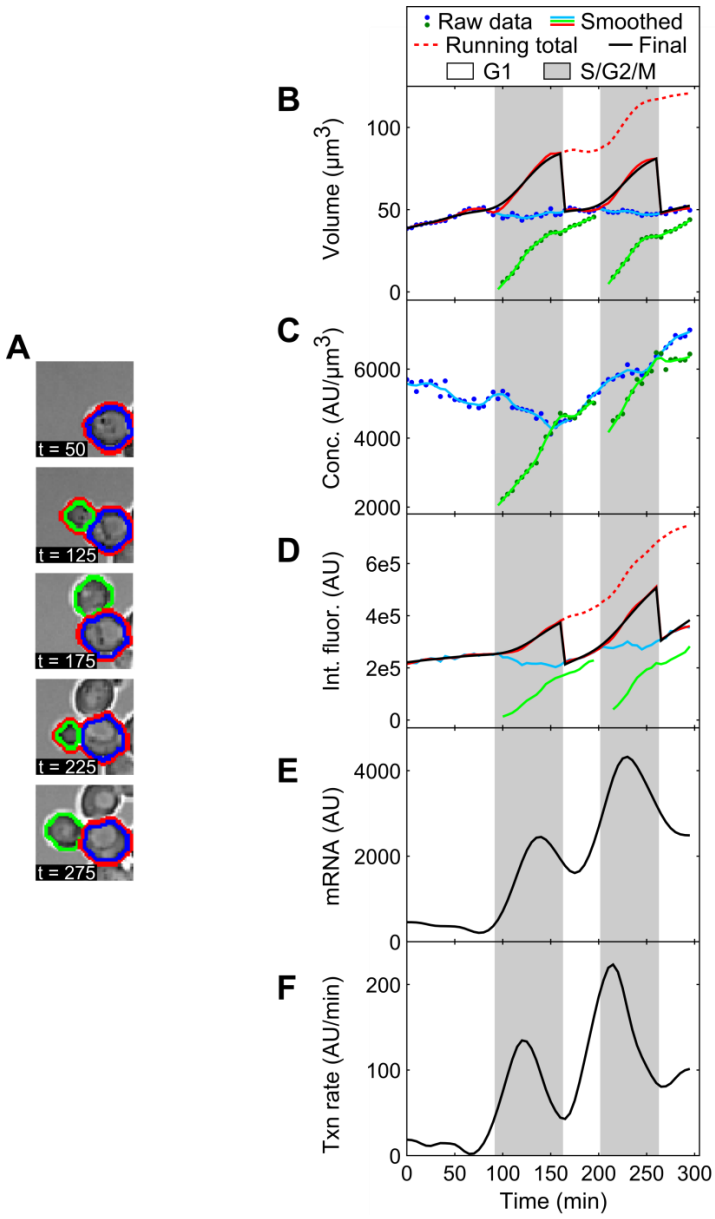


Figure 2.1 | Single-cell time series of a haploid yeast expressing PADH1-CFP over several generations.

(*A*) For the segmented mother cell (blue outline) and its buds (green outline), the contiguous whole cell trace is outlined in red. Raw mother and bud (*B*) volume and (*C*) protein concentration time series were smoothed to remove measurement noise, and (*D*) integrated CFP fluorescence was calculated as the product of volume and concentration. The whole cell (red) trace is extended past division to keep a running total that is easily fit to a differentiable smoothing spline across divisions (*B* & *D*). The (*E*) relative mRNA level and (*F*) instantaneous transcription rate are calculated using equations (2.1) and (2.2) and the spline fit in (*D*).

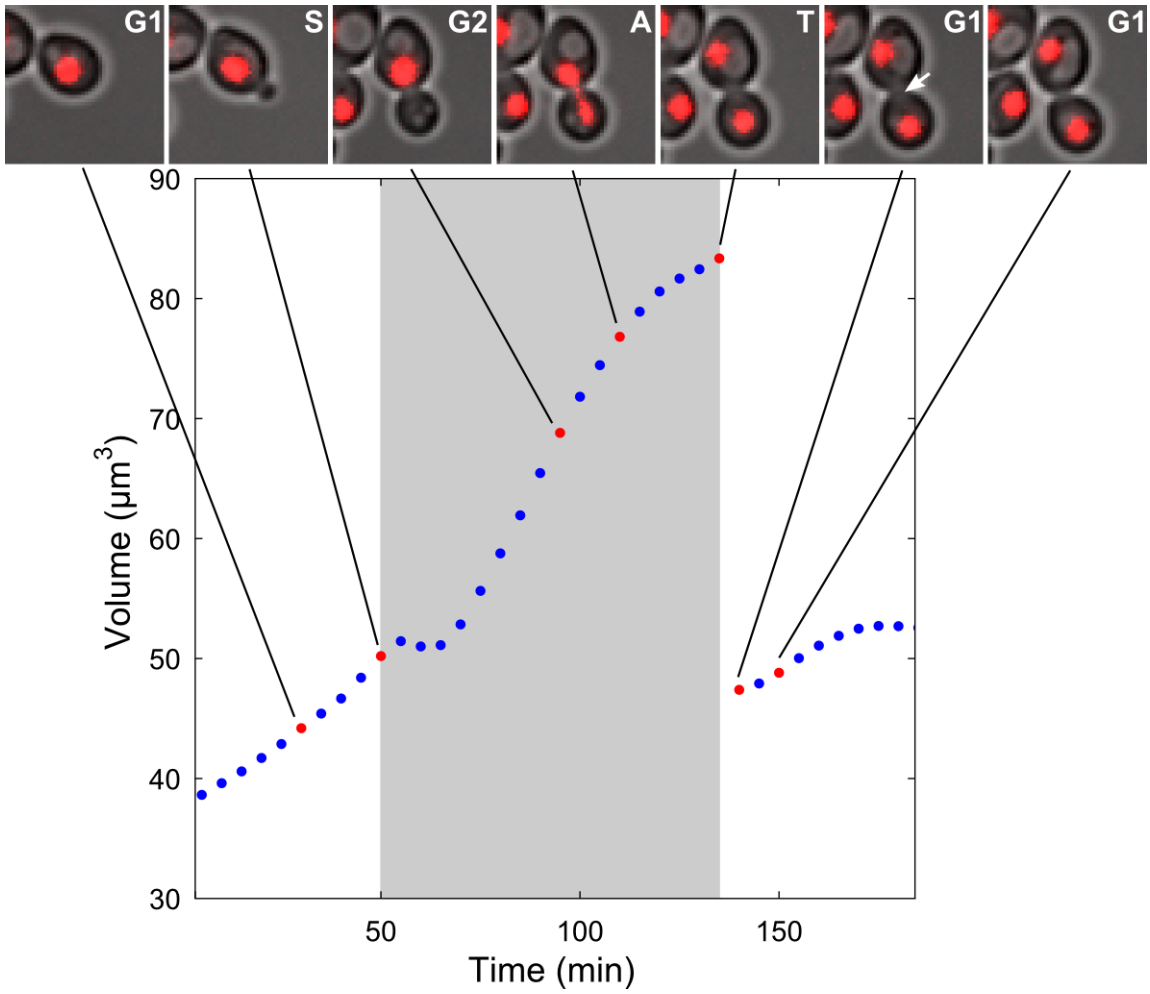


Figure 2.2 | Division time assignments correlate with nuclear division and cytokinesis time. At the top, bright field snapshots centered on a single cell are overlaid with RFP fluorescence marking the nucleus. Growth occurs in a microfluidic chamber, and the whole cell volume trace plotted corresponds to the contiguous volume of the central cell in the first image over time. The whole cell volume was obtained as in Figure 2.1B with automatically assigned budding and division times (beginning and end of gray shaded period, respectively). The cell begins in G1 and grows slowly until bud formation (characteristic of early S phase) at t = 50min. The nucleus migrates to the bud neck in G2 at t = 95min, and divides between mother and daughter during anaphase (A) at t = 110min. The automatically determined division time is t = 135min, at which point the nuclei have separated and the bud neck is narrowed in telophase (T). The subsequent G1 phase begins after the intersection of mother and daughter darkens (indicated by arrow) at the next time point, and by 150 min the dividing cell wall is distinct.

2.4 A model to infer instantaneous transcription rate in single cells

The total protein (fluorescence) spline $P(t)$ (Figure 2.1B) for each cell output by GRAFTS was used to calculate the protein production rate (proportional to mRNA per cell, $M(t)$) (Figure 2.1E), and the transcription rate, $A(t)$ (Figure 2.1F), using a simple continuous-time model of transcription and translation:

$$\frac{dP}{dt} = k_t M(t) \text{ or } M(t) = \frac{1}{k_t}\frac{dP}{dt} \tag{2.1}$$

$$\frac{dM}{dt} = A(t) - \gamma_M M(t) \text{ or } A(t) = \frac{1}{k_t}\left(\frac{d^2P}{dt^2} + \gamma_M \frac{dP}{dt}\right) \tag{2.2}$$

where γ_M is the mRNA degradation rate and k_t is the translation rate of mRNA to protein. The mRNA degradation rate for Venus *YFP* transcript in yeast was measured in our lab to be 0.04 min^{-1} (To and Maheshri 2010). The Cerulean *CFP* transcript differs by only a few base pairs, and it should have a similar degradation rate. While the transcript for tdTomato *RFP* is entirely different from that of *YFP*, we assume the same degradation rate in our *RFP* transcription calculations in Chapter 3. Then to use Eqn. (2.2) to represent transcription rate in single cells using only total protein level, three conditions must be met: (*i*) the translation rate k_t must not change so that inferred changes in $A(t)$ are due to changes in transcription; (*ii*) fluorophore maturation must not be slow relative to other processes; and (*iii*) we must be able to accurately calculate first and second time-derivatives of the total protein time series for accurate estimation of the transcription transition times. We address these points in the following sections. The result is that the quantity calculated by Eqn. (2.2) from each single-cell total protein time series represents a relative instantaneous reporter transcription rate delayed 10-15 minutes (due to fluorophore maturation) and smoothed over a 15-20 minute window (due to measurement and spline-fitting errors).

Translational capacity is constant across the cell cycle

We first contend translation rate k_t is nearly constant. In Chapter 3, we present evidence that the calculated transcription rate from Eqn. (2.2) fluctuates across the cell cycle, but increased protein production in S/G2 may be due to either increases in mRNA level or translational capacity (k_t in

Eqns. (2.1) and (2.2)); we argue for the former. First, while ribosomes numbers and activity are known to increase in yeast in S/G2 (Waldron et al. 1977; Elliott 1978), ribosome number is generally not considered rate-limiting for any particular gene as increasing gene dosage or mRNA number by transcriptional regulation leads to increased gene expression. Second, recent work in budding (Trcek et al. 2011) and fission (Zhurinsky et al. 2010) yeast suggests mRNA levels of constitutive genes increase during S/G2. Third, our lab found average protein to mRNA ratios of cells grouped by cell-cycle phase to show no discernible cell-cycle dependent trend (Zopf et al.). Therefore, we treat k_t as a proportionality constant.

CFP and YFP mature rapidly, while RFP may mature more slowly

The lumped-parameter model of gene expression represented by Eqns. (2.1) and (2.2) does not explicitly account for the maturation time of the reporter protein fluorophore. Each translated protein must undergo an autocatalytic oxidation reaction to form a fluorescent chromophore from peptide side-chains which creates a finite delay in the ability to observe new proteins (Reid and Flynn 1997). To be sure our estimations of transcription transitions were not artificially delayed by a lag in observation, we measured the maturation rate of CFP (Cerulean (Rizzo et al. 2009)), YFP (Venus (Nagai et al. 2002)), and RFP (tdTomato (Shaner et al. 2004)) similar to the method of (Gordon et al. 2007). We used a "3-color" diploid yeast strain with homologous 7xtetO promoters (P_{7xtetO}) driving either Cerulean (CFP) or Venus (YFP) and a constitutive *PGK1* promoter (P_{PGK1}) driving tdTomato (RFP). The tet-transactivator (tTA), which drives expression at the synthetic 7xtetO promoters, was also constitutively expressed by the *MYO2* promoter. Cells were cultured to mid log-phase growth, loaded into the microfluidics, and allowed to grow for 3 hrs. Then the flow was switched to media containing 30 μg/mL cycloheximide to block translation and images were acquired every 5 minutes. Initially the fluorescence increased, presumably due to maturation of the remaining immature protein pool. CFP and YFP fluorescence peaked quickly before a slow decline (Figure 2.3A). The loss of observable fluorophore occurs at longer time-scales and is most likely due to fluorophore bleaching because the fluorescent proteins are very stable. We fit the measured data to a simple model describing both maturation and loss of observable protein, *P*:

$$\frac{dI}{dt} = -k_m I \tag{2.3}$$

$$\frac{dP}{dt} = k_m I - \gamma_P P \tag{2.4}$$

Here, I is the immature protein, k_m is the maturation rate and γ_P is the loss rate, probably mostly photobleaching. The time-dependence of the observable protein can be easily determined after cycloheximide addition:

$$P(t) = \frac{k_m I_0}{\gamma_P - k_m}\left(e^{-k_m t} - e^{-\gamma_P t}\right) + P_0 e^{-\gamma_P t} \tag{2.5}$$

Here, I_0 and P_0 are the amounts of immature and mature protein at the time of cycloheximide addition. We fit Eqn. (2.5) to the CFP and YFP time series for each cell starting at the time of cycloheximide addition to obtain values for I_0, P_0, k_m, and γ_P for each trace. For RFP, the time-series never decreased, suggesting over the time of the experiment the maturation process was dominant over photobleaching (Figure 2.3A). Therefore, we fit the RFP time series assuming $\gamma_P = 0$ (reducing the fit to a single exponential). The median maturation half-lives ($\ln(2)/k_m$) for CFP, YFP, and RFP were found to be 10, 32, and 150 min, respectively (Figure 2.3B). The corresponding loss/bleaching half-lives ($\ln(2)/\gamma_P$) for CFP and YFP were 38 and 10 hrs, respectively. If some bleaching of RFP occurs, the 150 min maturation half-life represents a lower bound.

Unfortunately, we are not certain about the accuracy of these estimates for several reasons. First, on average CFP and YFP traces appear to change identically in response to cell-cycle dependent changes in transcription, within a multiplicative constant, in Chapter 3. This is true not only for situations of high expression with transcription occurring in both G1 and S/G2/M, but also with low expression where active transcription is only observed in S/G2/M. Given the difference in maturation, one would expect a larger delay in YFP expression. Second, in Chapter 4 we perform step tests in a diploid variant of the switchable transactivator strain expressing both CFP and YFP from a responsive promoter. The CFP and YFP responses are again identical within a multiplicative constant. Third, the fit value of the initial ratio of immature to mature protein (I_0 / P_0) is 0.10 and 0.33 for CFP and YFP, respectively. While this is qualitatively consistent with YFP having a slower maturation rate, it is quantitatively unreasonable. Cells are expressing CFP and YFP at steady-state levels before cycloheximide addition, so $I_0 / P_0 = \gamma_P / k_m$. Using the

inferred maturation and loss/bleaching rate yields an expected I_0 / P_0 of 0.005 and 0.05 for CFP and YFP. Fourth, the long maturation time of RFP should in principle dramatically smooth cell-cycle dependent transitions in the RFP trace, and yet we and others (Di Talia et al. 2007) observe these transitions.

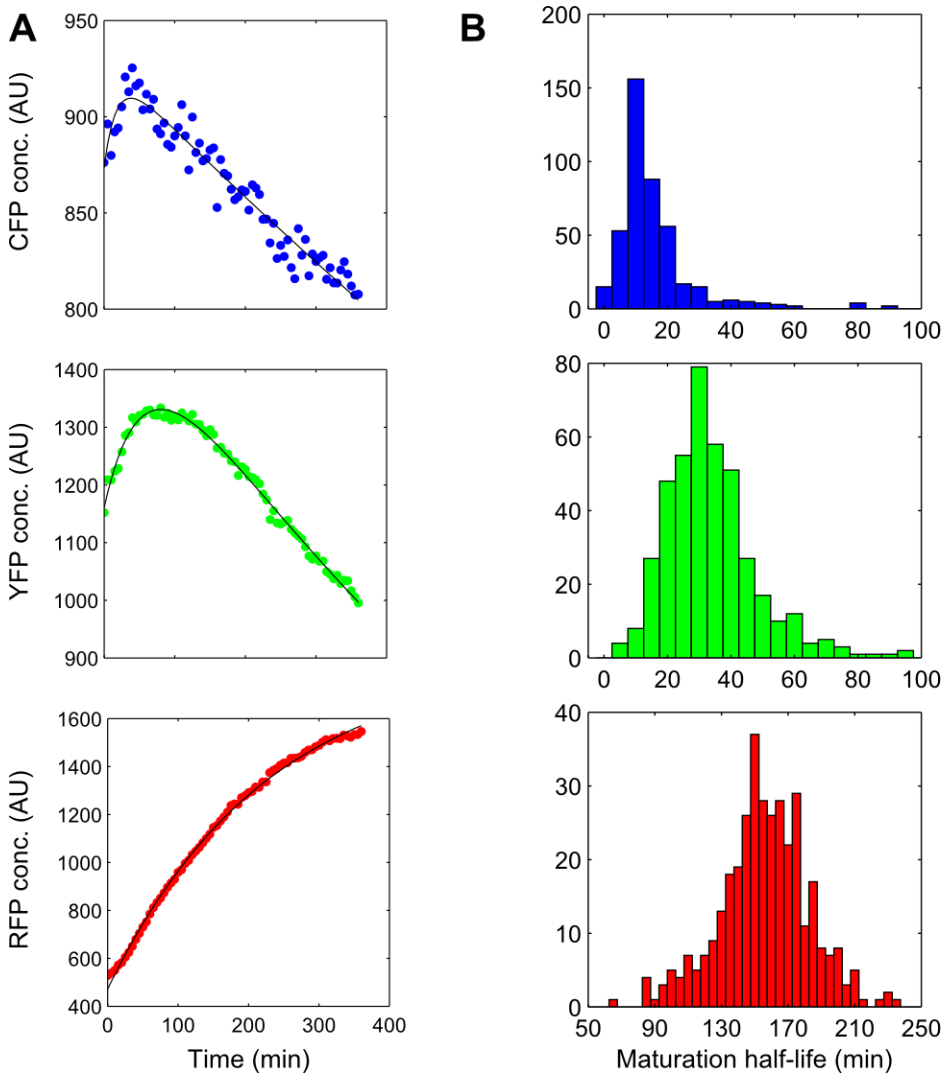


Figure 2.3 | Measurement of maturation rate for each fluorescent protein.
The 3-color strain was grown in microfluidics in synthetic media. After three hours, the chamber was perfused with the same media now containing 30 µg/mL cycloheximide to inhibit translation. The average intensity in each cell continued to increase for each fluorescent reporter, representing fluorophore maturation from the immature protein pool. (*A*) Single-cell, raw concentration time series were fit (black line) to Eqn. (2.5) (γ_P fixed at 0 for RFP) and (*B*) the histograms of the measured maturation half-lives ln(2)/k_m are shown. The medians are 10 min, 32 min, and 150 min for CFP (top panel), YFP (middle), and RFP (bottom), respectively.

Because of the uncertainties in maturation, we chose to analyze the time series data using Eqns. (2.1) and (2.2) which do not account for maturation. We discuss in detail how including maturation affects our results below. Because we have observed both YFP and CFP expression within 15 minutes of turning a gene on (using the kinetic strain), we suggest the maturation rate is similar and closer to the value measured for CFP (10-15 minute half-life). Importantly, the central findings in the following chapters that (*i*) transcription rate is greater in S/G2/M versus G1, (*ii*) at low expression levels transcription can be restricted to S/G2/M, and (*iii*) transcription activation delays are shorter in S/G2/M relative to G1 are qualitative and robust to changes in the maturation rate.

Spline-fitting to differentiate time series

In using Eqns. (2.1) and (2.2), we sought to estimate mRNA number and identify transitions from an inactive to active promoter state by examining rates of change in the protein time series. While such transitions might be more apparent from direct measurements of an unstable protein, its rapid degradation makes detection difficult at low levels of expression. Instead, we developed a method to estimate first and second time derivatives of the measured total fluorescence time series and thereby infer the transcription rate using Eqn. (2.2). In order to calculate an instantaneous growth rate, we consider the time-differentiation of the total volume time series as well. Because of noise in experimental data, estimating time derivatives is an ill-posed problem, and direct methods (e.g., finite differences) will amplify the noise with each application. Local splining techniques such as the Savitzky-Golay method (Savitzky and Golay 1964) can yield a smooth first derivative for noisy data, but will amplify noise for higher order derivatives required in our subsequent analysis. We instead fit each time series to a cubic smoothing spline following the method of de Boor (De Boor 1978). This algorithm calculates a single continuous spline, balancing accuracy and smoothness by minimizing a linear combination of the least square errors of the fit and the spline's second derivative (a measure of roughness). A single smoothing parameter, ρ, determines the relative weight in the minimization between the residuals and the curve roughness to allow an appropriate smoothing of noise in the data, y, while preserving real features. The resulting smoothing spline, f, then minimizes the expression

$$\rho \sum_{j=1}^{n} w(j)\left|y(j) - f(x(j))\right|^2 + (1-\rho)\int \left|D^2 f(t)\right|^2 dt \tag{2.6}$$

where *x* are the corresponding times for the data points *y*, with *w* the relative error weighting for each point.

We applied the MATLAB "csaps" implementation of this method to time series for each cell in two stages each with its own smoothing parameter. Due to the sensitivity of the spline fit to ρ, a suitable ρ is often near $1/(1+\Delta t^3/6*\beta)$ (as noted in the Matlab documentation for the "csaps" function), and so we approach spline smoothing by setting a suitable β. In the first stage, we fit a spline to both the volume and concentration time series to correct outlying data caused by large autofocusing and segmentation errors. Using a heuristically-chosen, conservative smoothing parameter (β = 3), the raw data was pre-treated to facilitate budding and division time assignments as well as the construction of integrated fluorescence time series, computed as the product of the volume and concentration splines (Figure 2.1D).

In the second stage, the total fluorescence and total volume time series (the sum of mother and bud series) were again fit to a spline using a second smoothing parameter chosen to provide reliable time-derivatives for each. An ideal β would result in a spline *f* that minimizes fluctuations in the derivatives due to noise in the measurement but preserves real features in the derivative time series. This β depends on the characteristics of both noise and real features in a particular data series, but it is difficult to distinguish these *a priori* for any one condition. We therefore chose to fit all fluorescence time series to a single, pre-determined β. By treating all data consistently, we maintain clarity in interpretation across all experiments with the understanding that series with fewer and smaller real features will yield slightly noisier splines. A computational strategy allowed the determination of a single, suitable β to spline total fluorescence for the different experimental conditions (also suitable for total volume, which was much less sensitive to the choice of smoothing parameter).

First, we postulate typical features in a transcription rate time series, and then computationally generate representative mRNA and protein time series from that transcription rate. The prominent features we include in the simulated transcription rate are informed by experimental observations in Chapters 3 and 4: cell cycle-dependent oscillations, different expression levels,

and sharp transitions. Total protein accumulation accelerates after budding (Figure 2.1D), so we simulate the transcription rate in an oscillatory manner with a period of the cell cycle and an arbitrary mean of 1:

$$A_{sim}(t) = u(t)\left[\alpha \sin\left(2\pi\left(\frac{t}{t_{cyc}} + \varphi\right)\right) + 1\right] \tag{2.7}$$

where α is a scaling factor to set the peak to trough ratio (P:T) and t_{cyc} is the doubling time. We vary P:T from 1:1 to 3:1 when setting α to simulate behavior from flat transcription to a cycle-dependence greater than gene dosage. The cycle length t_{cyc} is either 100 or 200 min, the approximate bounds for growth in glucose or raffinose. The inclusion of $u(t)$, the Heaviside step function, models an ideal switch in response to a step change in input. This is included in all simulations to combine transient and steady-state transcription behavior in each fit. Finally, to generate multiple asynchronous time series akin to the cell population we measure, we include φ as a random number from 0 to 1 to randomize the initial cell cycle position. Solving Eqn. (2.2) for $M(t)$ yields:

$$M_{sim}(t) = \mu \;\; u(t)e^{-\gamma_M t}\int e^{\gamma_M t}A_{sim}(t)dt \tag{2.8}$$

with μ a transcription rate scaling factor that depends on promoter strength, and varies from 1 to 100 to match observed expression levels. γ_M is the mRNA degradation rate, set to 0.04 min^{-1} measured in (To and Maheshri 2010). The total protein can likewise be computed by solving Eqn. (2.1) for $P(t)$:

$$P_{sim}(t) = k_t \int M_{sim}(t)dt + \xi(t) \tag{2.9}$$

where k_t is the translation rate (a redundant scaling factor kept at 1 here), and $\xi(t)$ adds high-frequency measurement noise in total protein so that we may determine the effect of measurement errors on choosing β.

The noise term $\xi(t)$ models the noise observed in the total fluorescence measurement after the first splining stage when there is no transcription – i.e., a baseline level (Figure 2.4A). Baseline fluorescence time series could be experimentally obtained using the switchable transactivator

strain under repressing conditions when transcription does not occur, and the fluorescence level corresponds to cellular autofluorescence (Chapter 4). We describe this baseline noise– the product of smoothing raw, high-frequency noise in the volume and concentration time series – using an autoregressive-moving average (ARMA) model. This parametric approach allows us to generate stochastic baseline noise reflecting the statistics of experimental observations. To each of 86 representative baseline traces, we fit a 4^{th} order ARMA model:

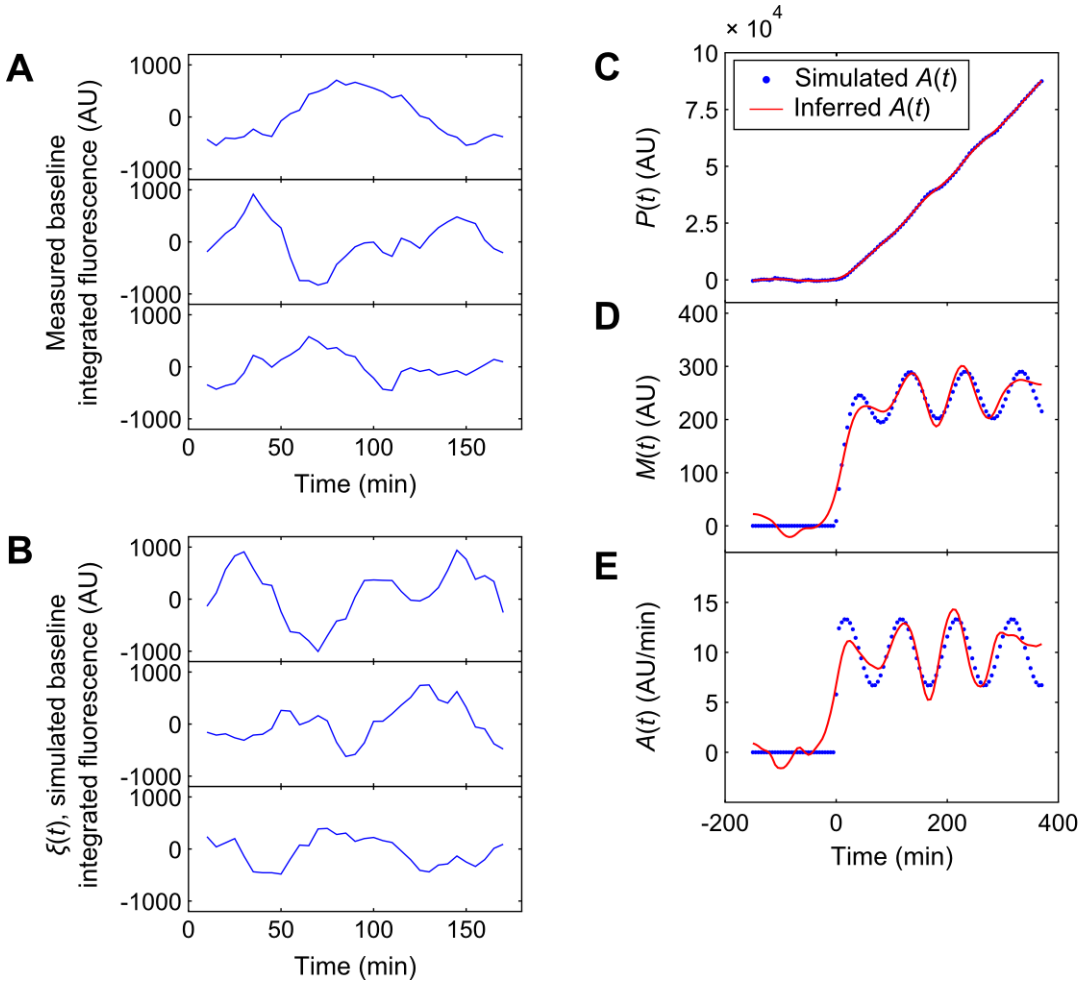
$$\xi(t) + a_1\xi(t-1) + ... + a_4\xi(t-4) = e(t) + c_1e(t-1) + ... + c_4e(t-4) \quad (2.10)$$

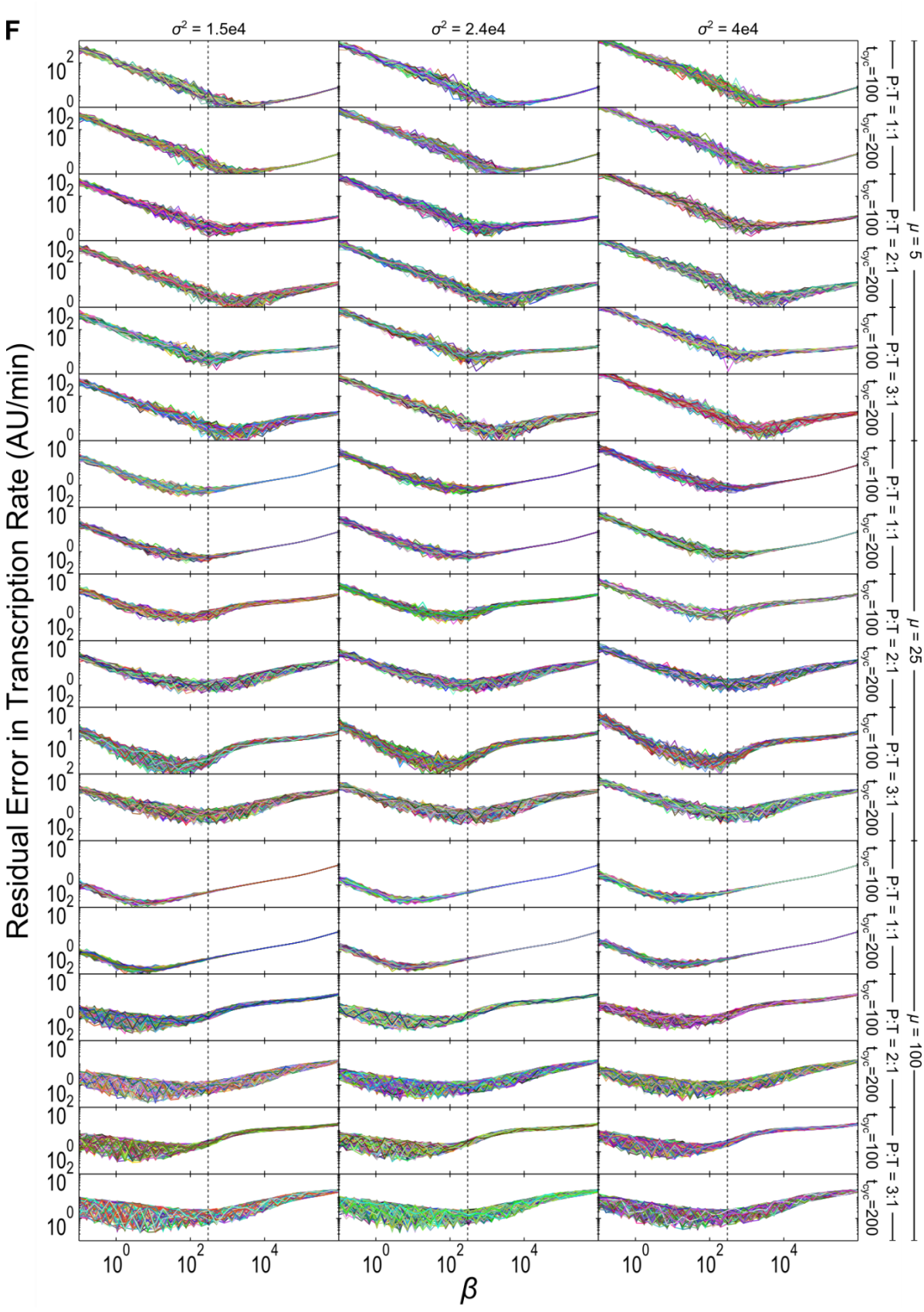
where the coefficients a_i and c_i define the autoregressive and moving average polynomials, respectively, and $e(t)$ is Gaussian white noise time series with variance σ^2. To define our noise term $\xi(t)$ as a representative ARMA(4,4) model, we take the median of all fits for each coefficient. This model then allows us to generate unique baseline noise traces with representative statistics simply by simulating a unique $e(t)$ (Figure 2.4B). To ensure we capture a realistic range of noise strength in the simulations, we vary σ^2 by setting it to either the 25^{th}, 50^{th}, or 75^{th} percentile of the fit values.

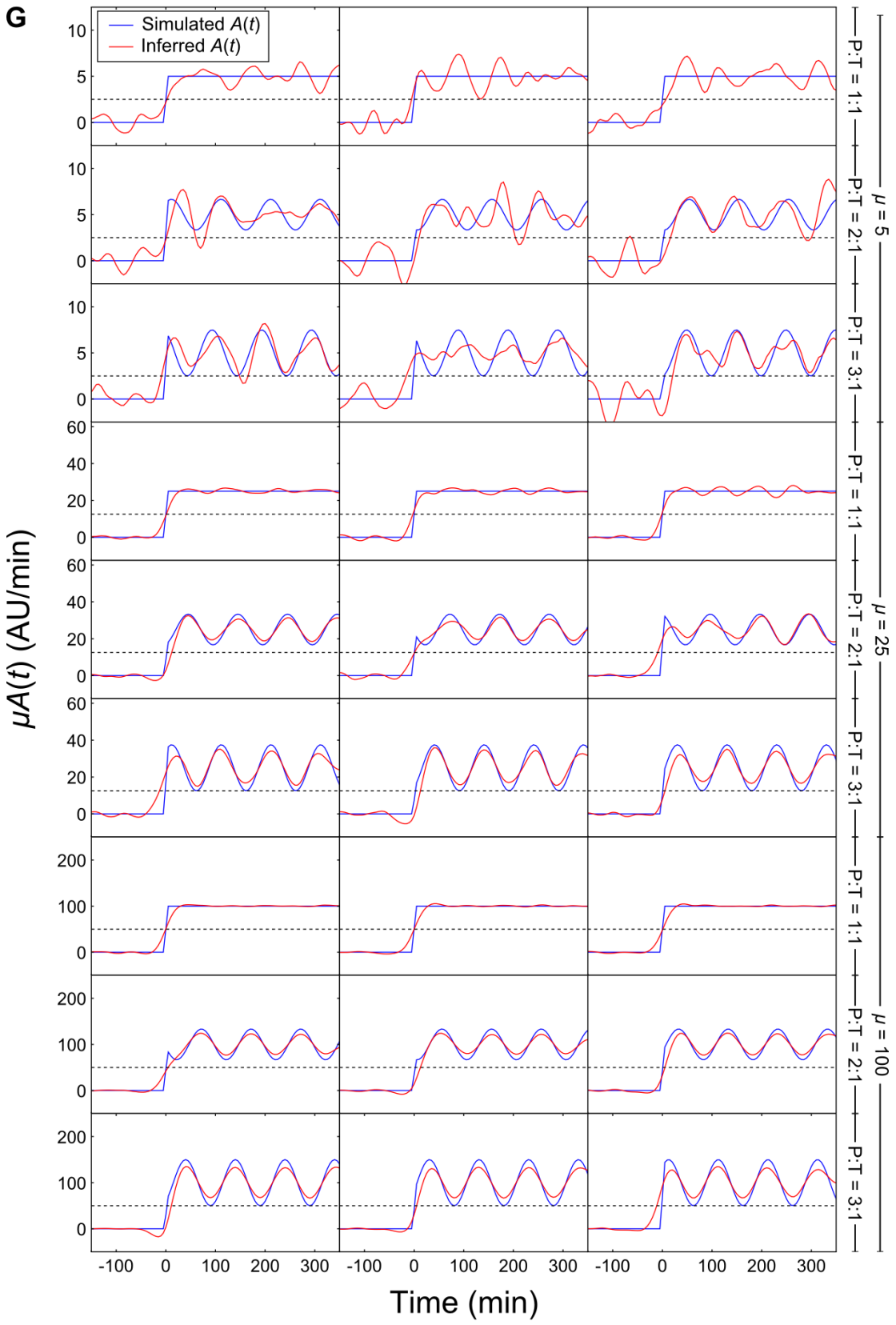
Next, we fit a spline to each simulated protein time series over a range of values for β to infer the underlying transcription rate (using Eqns. (2.1) and (2.2); Figure 2.4C-E). We compared this inferred rate for each β to the initially postulated $A_{sim}(t)$ and determined the β corresponding to the minimum residual error. To be sure our chosen β is suitable for splining all the data, the parameters α, t_{cyc}, μ, and σ^2 are varied within the ranges given above to generate noisy protein time series reflecting the gamut of experimental data. For each combination of the parameters shown in Figure 2.4F, 100 protein time series with a random instance of φ and $\xi(t)$ were generated, and the inferred, underlying transcription rate error was calculated for the range of β shown. The residual-minimizing β is insensitive to the parameters t_{cyc}, and α over the ranges encountered in the experiments; however, the magnitude of the noise term (σ^2) relative to the expression level (μ) strongly influences the range of β that provides an accurate inference of transcription rate. Predictably, relatively noisier protein traces require a higher β for satisfactory smoothing, while stronger transcription (i.e., stronger features) can be more accurately inferred at lower β. Therefore, there is no single β that will perfectly infer transcription at the noisiest

Figure 2.4 | Simulation of realistic total protein time series to determine optimal smoothing parameter for spline fit.

Noise characteristics of experimental integrated fluorescence data before splining (*A*, representative single cell traces from the 1xtetO and 7xtetO step tests, Chapter 4) were used to simulate baseline protein time series (*B*). (*C-E*) $A(t)$ was simulated for a step up at $t = 0$ from zero to an oscillating steady state with the period of a typical cell cycle (*E*, t_{cyc} = 100 min, P:T = 2:1). A simulated mRNA (*D*, $\mu = 10$) and total protein trace (*C*) were calculated from $A(t)$ using Eqns. (2.7)-(2.10), with noise added to the protein time series $\xi(t)$ similar to the high-frequency, measurement noise observed in data (σ^2 = 2.4e4). A smoothing spline was fit to the simulated protein time series and the corresponding mRNA and transcription rate time series were inferred using Eqns. (2.1) and (2.2) in the Methods. (*F*) The simulation was repeated for all combinations of parameters described in the text that span observed characteristics of protein time series across all experiments. For each parameter set (panels), 100 noisy, asynchronous traces were generated and the residual error between the simulated and the inferred transcription was found across a range of smoothing parameters β (colored lines). The chosen β = 300 (vertical, dashed line) minimizes the residual error across all parameter sets and was used for all data sets. (*G*) Sample comparisons of simulated and inferred transcription rate $\mu A(t)$ (using β = 300) for each parameter set with t_{cyc} = 100min. Black dotted line represents half the steady state average for $t > 0$ (the threshold for determining activation in a step test).




F
σ² = 1.5e4
σ² = 2.4e4
σ² = 4e4
Residual Error in Transcription Rate (AU/min)
β
μ = 5
μ = 25
μ = 100
P:T = 1:1
P:T = 2:1
P:T = 3:1
t_cyc=100
t_cyc=200


G
Simulated A(t)
Inferred A(t)
μA(t) (AU/min)
Time (min)
μ = 5
μ = 25
μ = 100
P:T = 1:1
P:T = 2:1
P:T = 3:1

baseline levels (upper right panels) and at the clear, strongest levels (lower left panels), but these simulations inform our decision in making this inherent tradeoff.

The important biological observation of cell cycle-dependent transcription relies not on quantitative accuracy in the inferred transcription rate but on correct inference of the *timing of transitions* in transcription rate. Undersmoothing preserves the sharpness of transitions, but noise is amplified in derivatives, possibly indicating false starts and stops in transcription. On the other hand, oversmoothing preserves transition times on average, but individual transitions are degraded in a trace-dependent manner. For this reason, a β of 300 was heuristically chosen to most accurately infer transcription rate for the range of scenarios in Figure 2.4F. This β best preserves transition times at the expense of tracking some noisy fluctuations during periods of gradual change and eroding the edges of the sharpest features (Figure 2.4G).

Two additional properties of experimental time series data must be taken into account when using spline fits to estimate derivatives. First, there is limited data at the beginning and end of a time series. To address this, we tapered the first and last three weights in *w* to diminish edge effects on nearby derivatives (all interior points were weighted equally). Second, sharp jumps in data across cell divisions lead to undefined second order derivatives. Therefore, the time series across a mother/daughter division was extended by maintaining a continuous whole cell "running total" of volume and fluorescence of the mother even after division (Figure 2.1C&E). This allowed a single spline with well-defined first and second derivatives to describe the entire, multi-generation time series for a cell. As with any smoothing approach for noisy time series, potentially real, sharp transitions in the derivative across division will be lost. Still, we capture the main features in the time derivatives of these single-cell time series using a cubic smoothing spline thus providing the basis for an analysis of gene activity over time.

Time accuracy of the calculated transcription transition times.

Inferring transcription rates from measured protein levels limits the time resolution of transcription events. We are particularly interested in the ability to infer large *transitions* in transcription rate and do so by applying Eqns. (2.1) and (2.2) to splines fitting total observable protein time series. There are three main sources of error which can limit our accuracy in determining these transitions. (*i*) Measurement noise at the protein level degrades feature

boundaries and can obscure the times at which the true transcription rate changes. (*ii*) Cell-to-cell variability in the mRNA degradation rate used in Eqn. (2.2) introduces uncertainty when inferring transcription using the average mRNA degradation rate. Also, for inference we use a model where mRNA degradation rate is deterministic, but there are stochastic fluctuations in mRNA numbers during degradation. These effects should not decrease the time resolution of our inference at low expression levels by more than 17 min, the mean lifetime of a single mRNA (To and Maheshri 2010). (*iii*) Maturation of the protein fluorophore creates an observational lag in the protein time series tightly distributed around an ~10 min time-scale (Figure 2.3A, as discussed above). We neglect contributions from cell-to-cell variability in maturation (mostly due to errors in fitting short, noisy time series) as well as translation rate, which is relatively faster and likely also tightly distributed (as discussed above).

We then used simulated data to estimate the time resolution considering the three main limitations. Transcription rate and mRNA time series were generated as in Figure 2.4 but excluding cell cycle-dependent oscillations for clarity. Immature protein $I(t)$ was simulated as an intermediate between mRNA and observable protein:

$$I(t) = k_t u(t) e^{-k_m t} \int e^{k_m t} M_{sim}(t) dt \tag{2.11}$$

where k_m is the maturation rate. The observable, mature protein $P_{sim}(t)$ was then simulated similar to Eqn. (2.7) including the measurement noise term:

$$P_{sim}(t) = k_m \int I(t) dt + \xi(t) \tag{2.12}$$

To simulate idealized transcription rate transitions, we model the transcription rate time series input as a square pulse (Figure 2.5A). This approximation is relevant as transcription rate transitions during step tests in Chapter 4 appear step-like within the time-resolution found below. For the step change from 0 to 1 (and vice versa), the time of transition between the two plateaus is determined when the transcription rate reaches a midpoint value of 0.5. One thousand mature protein data series were simulated, and a transcription rate series was inferred for each by splining and using only Eqns. (2.1) and (2.2) (not accounting for a maturation step, as in Figure 2.4). Transition times for activation and deactivation in each simulated trace were estimated as

when the inferred transcription time series rose above or fell below the 0.5 threshold, respectively. Compared to the input transition times of both activation at t = 0 min and deactivation at t = 200 min, the inferred times were delayed by ~15 min on average and had a mean absolute deviation (MAD) of ~1 min (Figure 2.5B). Next, we allowed the mRNA degradation rate to vary during simulation of Eqn. (2.8). For each of the 1000 traces generated, the degradation rate was randomly chosen from a normal distribution of mean 0.04 min^{-1} (~17 min half-life) with a standard deviation of 0.009 min^{-1} equal to the error measured in (To and Maheshri 2010). Transcription time series were inferred from these samples using the average degradation rate, which lead to shifts in the active steady state and, thus, the appropriate threshold. Calculating transition times using the original midpoint threshold for these transcription time series does not influence the precision of the inferred activation time, but inferred deactivation times now have ~8.4 min MAD Figure 2.5C). The transition times estimated from each inferred transcription rate time series are therefore accurate within a ~17 min window, comparable to the time-scale of mRNA degradation.

All results reported in the following chapters are based on inferences using only Eqns. (2.1) and (2.2), and maturation has not been accounted for in any calculation. We performed simulations combining protein maturation and measurement noise to probe the effect of maturation on inferring the transcription transition time. We find an observational lag of ~15 min, which is approximately one 5 min time point longer than the expected ~10 min delay based on the maturation half-life. As transcription is only evaluated at 5 min intervals, a threshold-crossing occurring mid-interval is treated as occurring at the next 5 min time point. The observational delay can therefore be interpreted as being 10-15 min (2-3 time points). Additionally, modeling mRNA degradation deterministically using a single, average rate for each cell limits the time accuracy to ~17 min. All observed transcription rate transition events must then be interpreted as having occurred 10-15 min in the past with a resolution of no worse than ~17 min (between 3 and 4 time points). This is a conservative estimate for activation transition times, which are not strongly influenced by mRNA degradation.

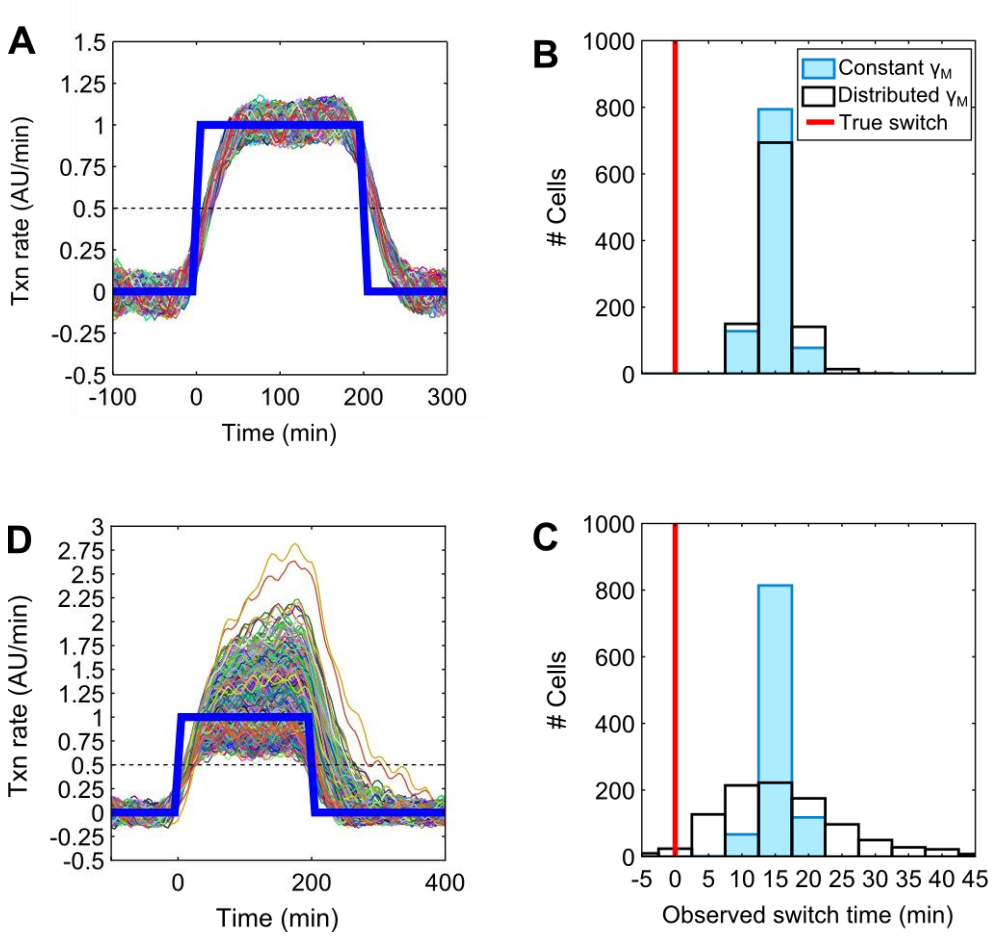


Figure 2.5 | Simulations to estimate the time accuracy of the transcription rate transition time assignments.

(*A*) Transcription rate time series (thin colored lines) are inferred for noisy protein data simulated from the square pulse transcription rate generator function (thick, blue line) as in Figure 2.4 (n = 1000). Here a protein maturation step is included in simulating the protein data, but not in the transcription rate inference. (*B*) Activation and (*D*) deactivation times "observed" for each inferred transcription trace in *A* are delayed 10-15 min relative to the true times (blue histogram compared to vertical, red line) with a mean absolute deviation of ~2 min. (*C*) Allowing the mRNA degradation rate γ_M in Eqn. (2.8) to vary during simulation, but inferring transcription rate using the single, average γ_M value, results in greater dispersion in the sample traces. The "observed" activation times did not vary much more, but deactivation time mean absolute deviation increased to ~8.4 min (black histograms in *B* and *D*).

2.5 References

De Boor C. 1978. *A Practical guide to splines*. Springer-Verlag, Berlin.

Charvin G, Cross FR, Siggia ED. 2008. A microfluidic device for temporally controlled gene expression and long-term fluorescent imaging in unperturbed dividing yeast cells. *PloS One* **3**: e1468.

Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437**: 699–706.

Cookson NA, Cookson SW, Tsimring LS, Hasty J. 2010. Cell cycle-dependent variations in protein concentration. *Nucleic Acids Research* **38**: 2676–2681.

Cookson S, Ostroff N, Pang WL, Volfson D, Hasty J. 2005. Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Molecular Systems Biology* **1**: 2005.0024.

Cross FR. 1995. Starting the cell cycle: what's the point? *Current Opinion in Cell Biology* **7**: 790–797.

Elliott SG. 1978. Rate of Macromolecular Synthesis through the Cell Cycle of the Yeast Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* **75**: 4384–4388.

Golding I, Paulsson J, Zawilski S, Cox E. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**: 1025–1061.

Gordon A, Colman-Lerner A, Chin TE, Benjamin KR, Yu RC, Brent R. 2007. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nature Methods* **4**: 175–181.

Kaufmann BB, Yang Q, Mettetal JT, Van Oudenaarden A. 2007. Heritable stochastic switching revealed by single-cell genealogy. ed. M.S. Gelfand. *PLoS Biology* **5**: e239.

Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. 2011. Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332**: 475–478.

Locke JCW, Elowitz MB. 2009. Using movies to analyse gene circuit dynamics in single cells. *Nature Reviews Microbiology* **7**: 383–92.

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

Paliwal S, Iglesias PA, Campbell K, Hilioti Z, Groisman A, Levchenko A. 2007. MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* **446**: 46–51.

Reid BG, Flynn GC. 1997. Chromophore formation in green fluorescent protein. *Biochemistry* **36**: 6786–91.

Rizzo MA, Davidson MW, Piston DW. 2009. Fluorescent protein tracking and detection: fluorescent protein structure and color variants. *Cold Spring Harbor Protocols* doi:10.1101/pdb.top63.

Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene Regulation at the Single-Cell Level. *Science* **307**: 1962–1965.

Savitzky A, Golay MJE. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**: 1627–1639.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* **22**: 1567–72.

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* **332**: 472–474.

Di Talia S, Skotheim JM, Bean JM, Siggia ED, Cross FR. 2007. The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* **448**: 947–51.

To T-L, Maheshri N. 2010. Noise Can Induce Bimodality in Positive Transcriptional Feedback Loops Without Bistability. *Science* **327**: 1142–1145.

Trcek T, Larson D, Moldón A, Query C, Singer R. 2011. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* **147**: 1484–1581.

Waldron C, Jund R, Lacroute F. 1977. Evidence for a high proportion of inactive ribosomes in slow-growing yeast cells. *The Biochemical Journal* **168**: 409–15.

Zhurinsky J, Leonhard K, Watt S, Marguerat S, Bähler J, Nurse P. 2010. A coordinated global control over cellular transcription. *Current Biology : CB* **20**: 2010–5.

Zopf CJ, Maheshri N. 2013. Acquiring fluorescence time-lapse movies of budding yeast and analyzing single-cell dynamics using GRAFTS. *Journal of Visualized Experiments (in press).*

Zopf CJ, Quinn K, Zeidman J, Maheshri N. 2013. Cell-cycle dependence of stochastic gene expression. *PLoS Computational Biology (submitted).*

CHAPTER 3. CELL-CYCLE DEPENDENCE OF TRANSCRIPTION DOMINATES NOISE IN GENE EXPRESSION[†]

3.1 Introduction

At the single-cell level, mRNA and protein levels of regulable genes are often found to be highly variable (Newman et al. 2006; Raj and van Oudenaarden 2009; Taniguchi et al. 2010). The resulting long-tailed mRNA and protein distributions are well-described by stochastic models (Peccoud and Ycart 1995; Raj et al. 2006; Shahrezaei and Swain 2008; Taniguchi et al. 2010) of transcriptional bursting, where a promoter undergoes random and intermittent periods of highly active transcription. Real-time observations of transcription in multiple organisms appear consistent with this behavior (Golding et al. 2005; Maiuri et al. 2011; Chubb et al. 2006; Muramoto et al. 2012; Larson et al. 2011; Choi et al. 2008; Taniguchi et al. 2010; Suter et al. 2011). Thus, both static and dynamic views attribute much of the observed mRNA variability to the stochastic nature of reactions intrinsic to transcription. Consequently, the standard stochastic model of gene expression has been widely used to infer steady-state dynamics (Raj et al. 2006; Mao et al. 2010; To and Maheshri 2010; Tan and van Oudenaarden 2010; Munsky et al. 2012). However, earlier studies examining the origin of variability in protein expression found such variability is not solely due to stochasticity in reactions intrinsic to gene expression, but also extrinsic factors. This was done by looking for correlations in expression between identical copies of one promoter (Elowitz et al. 2002; Raser and O'Shea 2004; Volfson et al. 2006) and/or between that promoter and a global or pathway-specific gene (Pedraza and Van Oudenaarden 2005; Colman-Lerner et al. 2005). Not only is the importance of extrinsic factors clear, without time-series measurements the intrinsic noise measured by these techniques may not completely be ascribed to stochastic reactions in gene expression (Hilfinger and Paulsson 2011). While global extrinsic factors have been suggested to largely impact translation (Raj et al. 2006), their influence on transcription and transcriptional bursting is unclear.

[†] Text and figures are largely drawn from (Zopf et al.).

The cell cycle has global effects on total protein and RNA synthesis that should play a role in transcription (Volfson et al. 2006; Larson et al. 2011; Trcek et al. 2011). With few exceptions (Volfson et al. 2006), most deterministic and stochastic models of gene regulation do not account for cell cycle variability. Using dynamic real-time protein measurements in single cell, and supported by static single molecule mRNA measurements in our lab, we show that much of the variability in a synthetic tetO promoter typical of noisy genes in yeast is driven by differences in transcription rate between G1 and S/G2/M.

3.2 Constitutive transcription increases in S/G2/M consistent with gene dosage

We examined cell-cycle dependent effects by microscopically monitoring fluorescent protein expression every 5 minutes in growing monolayers of yeast within a microfluidic chamber (as described in Chapter 2). We used a "3-color" diploid yeast strain with homologous 7xtetO promoters (P_{7xtetO}) driving either Cerulean (CFP, (Rizzo et al. 2009)) or Venus (YFP, (Nagai et al. 2002)) and a constitutive *PGK1* promoter (P_{PGK1}) driving tdTomato (RFP, (Shaner et al. 2004)) (Figure 3.1). A constitutive *MYO2* promoter drives expression of the tet-Trans-Activator (tTA), a fusion of the tetR DNA binding protein and the VP16 acidic activation domain which together activate transcription from the synthetic tetO binding site-*CYC1* minimal promoters. Addition of doxycycline inhibits tTA activity to control expression level from the tetO promoters, which is essentially constitutive without doxycycline present. Correlations in transcriptional activity between different promoters allowed distinction between different sources of fluctuations (Colman-Lerner et al. 2005; Volfson et al. 2006). Using the methods developed in Chapter 2, we infer the instantaneous transcription, protein production, and growth rates in single cells from the microscopy movies.

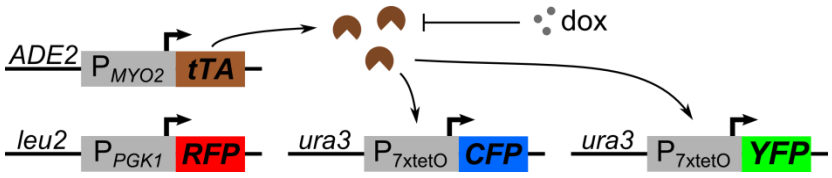


Figure 3.1 | The "3-color" diploid strain expresses three fluorescent reporters.

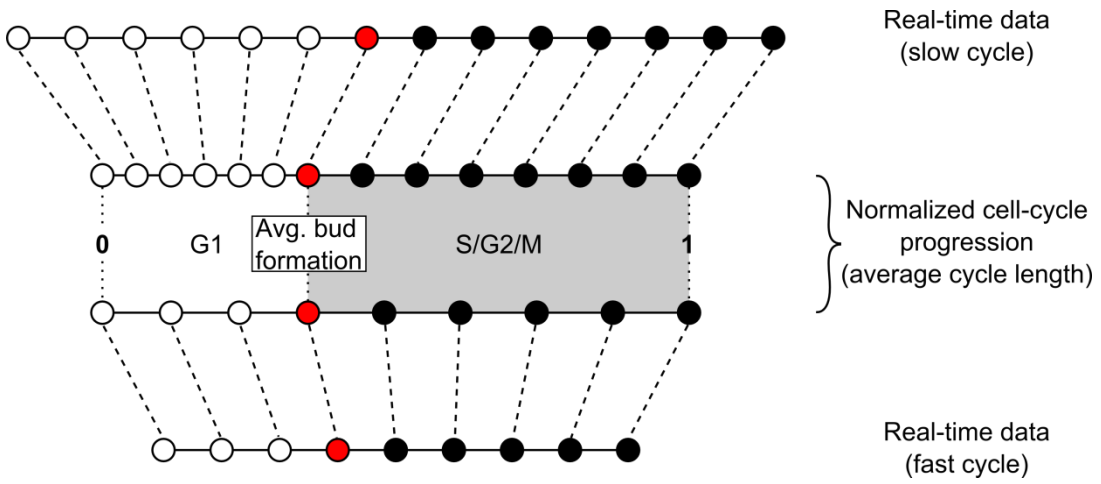


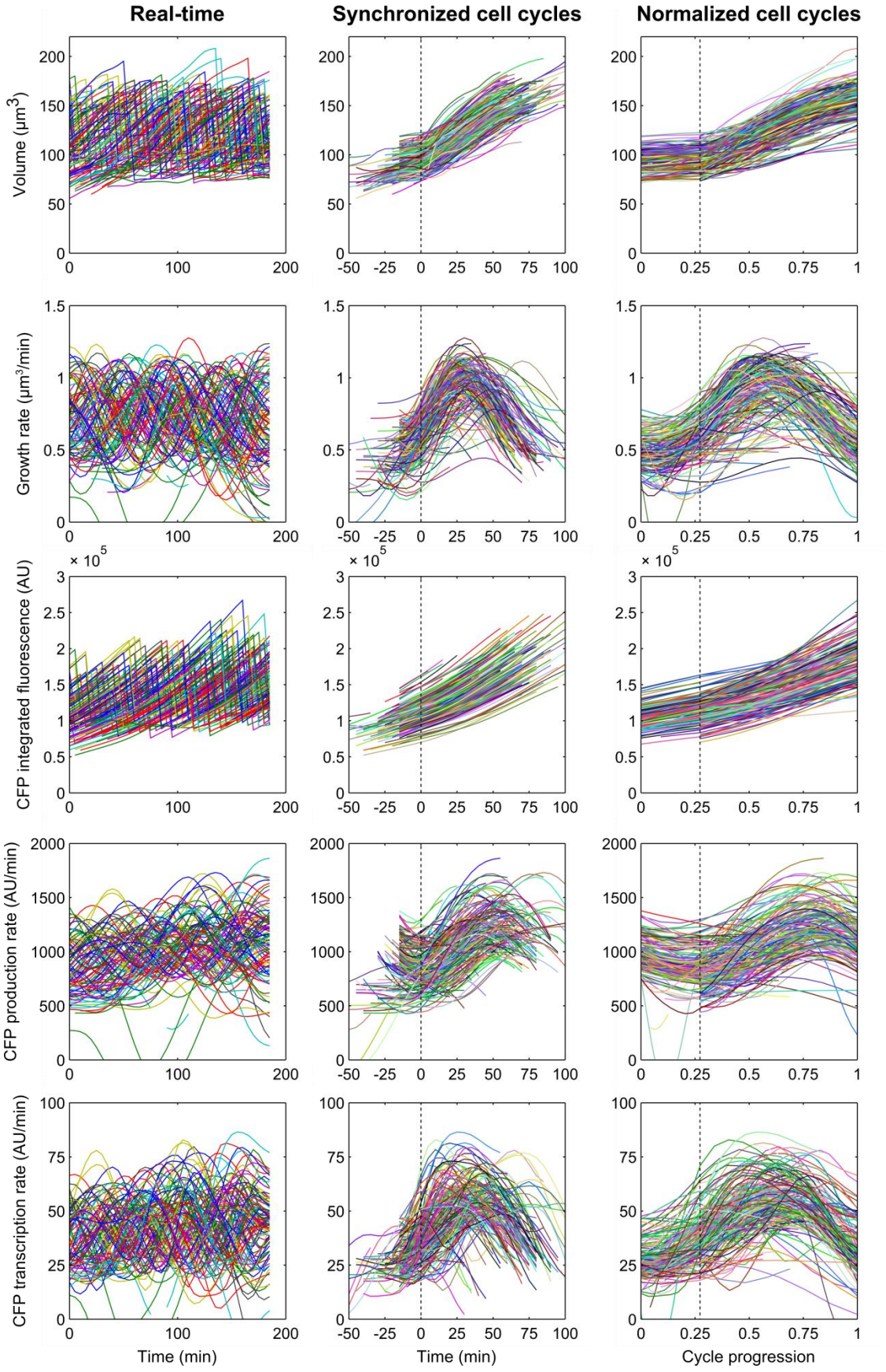
Figure 3.2 | Mapping data from each cell cycle to a normalized cell cycle progression representing the average cycle.
In real-time, the data points (circles) are evenly spaced. The endpoints of each series represent the first post-division data point to the division at the end of the cell cycle depicted. Each cell cycle consists of a G1 phase (white circles), bud formation (red circles), and S/G2/M phase (black circles). The data from each cell cycle in real-time (top and bottom series) are mapped to the average cell cycle (middle two series) by aligning bud formation at the average time, and linearly compressing or expanding intervals within each phase. The resulting normalized cell cycles can then be compared in terms of progression.

To examine the effects of cell-cycle phase on expression, we aligned growth and expression data with respect to cell-cycle progression. We subdivided single-cell time series data between division events, synchronized the data by bud formation time, and rescaled time such that the pre- and post-bud phases mapped to the population-average time in those phases (Figure 3.2). Thus, division occurs at 0 and 1, and all traces for each measurement bud at the same cell-cycle progression (Figure 3.3). From these plots it is apparent there are qualitatively similar cell cycle-dependent changes in growth and expression at the single-cell level so the average behavior will be representative. We then take the mean of all traces binned across the normalized cell-cycle progression. Cells exhibit a slow growth phase up to bud formation corresponding to G1 and very early S, followed by faster growth in S/G2/M (Figure 3.4A), consistent with (Goranov et al. 2009; Cookson et al. 2010). The instantaneous protein production rate similarly has two modes, but lags the instantaneous growth rate (Figure 3.4B). In contrast, the instantaneous transcription and growth rate correlate, approximately doubling in S/G2 relative to G1 (Figure 3.4C).

To confirm the relationship between transcription rate and cell cycle phase, we next changed the average growth rate by growing cells in media with alternate carbon sources (galactose or raffinose) or depleted phosphate (100 μM). All conditions lengthened the cell cycle predominately through extending G1. In each case, the above findings in glucose are robust to the changes in the cell's average growth rate (Figure 3.5). The gradual rise in transcription rate over ~30 min may mask a sharper change because of smoothing and reporter maturation (Figure 3.6). Thus, the ~2-fold change in transcription rates (Figure 3.4 and Figure 3.5) could be entirely consistent with the change in gene dosage at replication, which occurs during S-phase around the time of bud formation.

Figure 3.3 | *In silico* synchronization of single-cell time series mapped to average cell cycle (next page).

The first column contains the real-time series from each cell growing at steady-state. In the second column, each cell cycle is plotted between division events and synchronized so that budding occurs at *t* = 0 (vertical dashed lines). In the third column, the cell-cycle progression for each cell has been normalized to the average cycle, extending from 0 to 1 with the bud appearing at the average budding point of 0.27.


Real-time
Synchronized cell cycles
Normalized cell cycles
Volume (μm3)
Growth rate (μm3/min)
CFP integrated fluorescence (AU)
× 10^5
CFP production rate (AU/min)
CFP transcription rate (AU/min)
Time (min)
Time (min)
Cycle progression

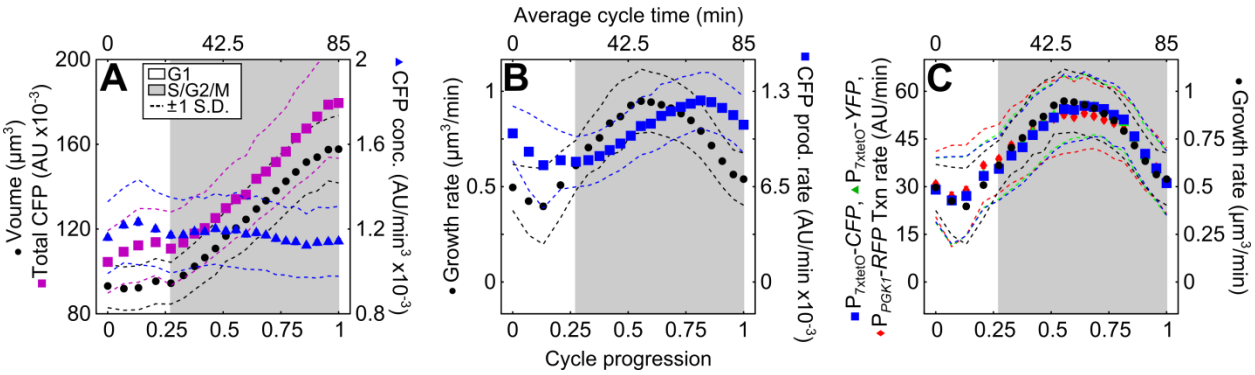


Figure 3.4 | Instantaneous transcription rate in single yeast cells correlates with growth across the cell cycle in glucose media.
(*A*) Both volume and total CFP rise rapidly post-bud formation but at slightly different rates because of the small (<10%) decrease in CFP concentration, similar to (Cookson et al. 2010). (*B*) Mean instantaneous growth and protein production rates are lower in G1 and peak in S/G2/M. (*C*) Instantaneous transcription rate of P_{7xtetO} and P_{PGK1} correlates with instantaneous growth rate in glucose. YFP and RFP transcription averages were normalized to CFP. Error bars represent the bin SEM from bootstrapping over 171 data points in each bin. Dotted lines indicate the bin S.D.

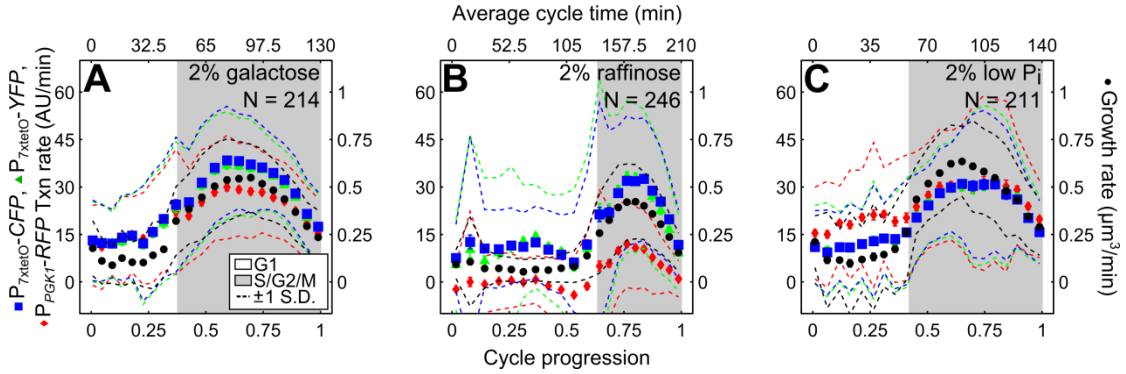


Figure 3.5 | Correlation between instantaneous transcription rate and growth rate is robust to changes in cell-cycle length.
Instantaneous transcription rate of P_{7xtetO} and P_{PGK1} correlates with instantaneous growth rate in galactose (*A*), raffinose (*B*), and low phosphate (*C*) media. YFP and RFP transcription averages were normalized by the ratio between the CFP average in the condition specified here and in glucose (Figure 3.4C). Error bars represent the bin SEM from bootstrapping over N data points in each bin. Dotted lines indicate the bin S.D.

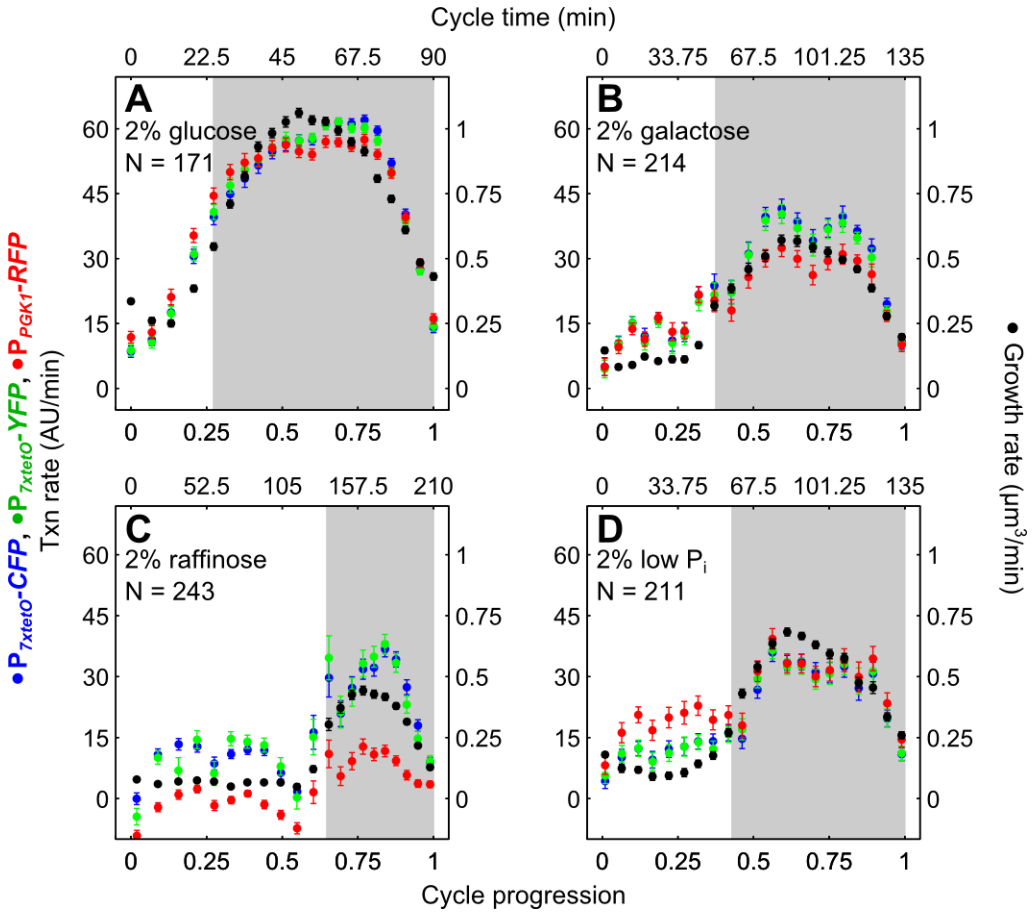


Figure 3.6 | Less smoothing results in sharper transitions in average instantaneous growth and transcription rates across cell cycle phases but general trends are robust to choice of smoothing parameter.

The same raw data in as in Figure 3.4 and Figure 3.5 are processed using a more conservative smoothing parameter β = 3 (compared to β = 3, see Chapter 2) that does not smooth noisy features, which results in slightly different cell cycle lengths compared to the previous figures. After synchronizing and averaging single-cell traces as previously, the general trends of the transcription and growth rates are similar. Notably, less smoothing leads to sharper transitions at budding and division with roughly constant rates within each cycle phase. Gray areas again show the average S/G2/M phase, and white areas show the average G1 phase. Error bars represent the bin SEM from bootstrapping over N data points in each bin.

3.3 Random pulses of transcription during "bursty" expression only occur in S/G2/M and are correlated

We next added 50 ng/mL dox to reduce P_{7xtetO} expression in the 3-color diploid to levels where transcription is thought to occur in infrequent, independent bursts at each locus (Raj et al. 2006;

To and Maheshri 2010). At low enough frequencies, these bursts are presumably resolvable by the real-time transcription rate analysis. Inspection of single-cell traces of transcription rate for CFP and YFP showed occasional pulses of active expression from the 7xtetO promoters relative to the baseline fluctuations (Figure 3.7). Comparatively, RFP transcription from the constitutive *PGK1* promoter follows the cell-cycle pattern observed above (Figure 3.4).

Due to the limited time resolution of spline-fitting the data, determining an activation time for each pulse is not straightforward. In the case of an underlying instantaneous switch between two transcription rates, we demonstrated in Chapter 2 that the activation time can be accurately inferred as the time at which the transcription rate reaches the midpoint between the two rates even when the observed transcription time series is noisy and spline-smoothed. Unlike for step changes, though, an appropriate threshold cannot be determined simply at 50% of the pulse maximum for several reasons. First, the pulse maximum may not reflect the true steady-state transcription rate as the "on" periods are short; smoothing may degrade the peak height. Second, each pulse has a different height and midpoint and weaker pulses are more susceptible to feature degradation by spline-smoothing. Third, some periods have irregular shapes and cannot be modeled as a sharp step change.

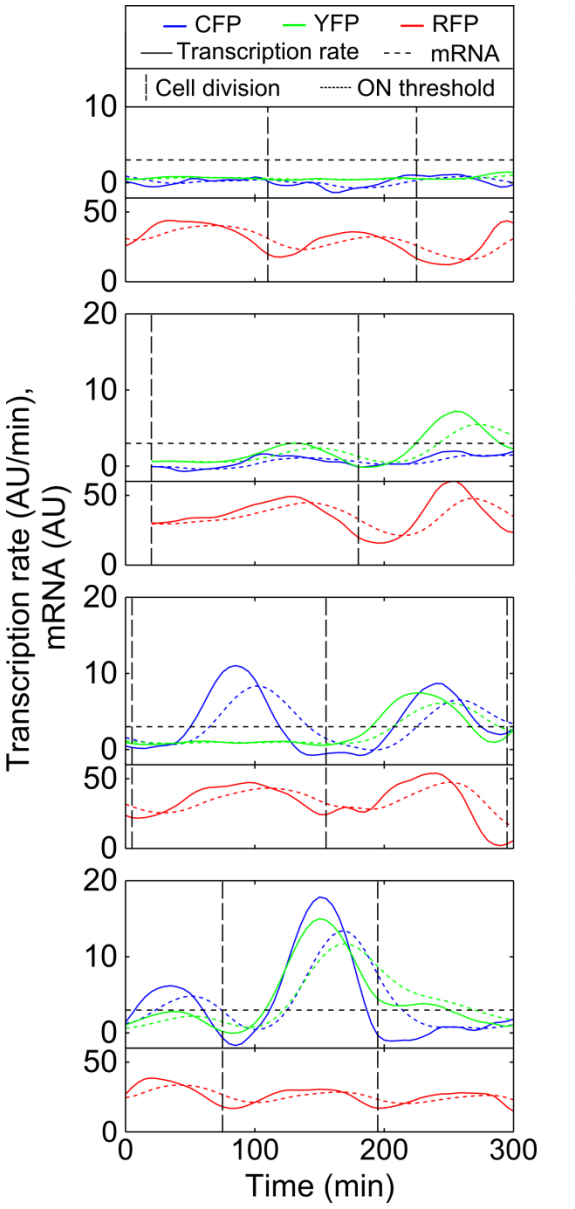


Figure 3.7 | Sporadic transcriptional activation in real-time at two copies of P_{7xtetO}.
Transcription rate time series (solid lines) for four representative cells, one cell in each panel, showing transcription rates from P_{7xtetO} driving either *CFP* or *YFP* (blue and green, respectively, top of each panel). Both promoters are off in G1, may or may not turn ON in S/G2 (with a transcription rate that exceeds the dotted black line denoting the threshold). In contrast, a constitutively expressed RFP (red) from P_{PGK1} (bottom of each panel) is always ON and exhibits the usual cell-cycle dependent fluctuations. Colored, dashed lines show the corresponding relative mRNA level resulting from transcriptional activity. Vertical dashed lines demarcate cell divisions.

To distinguish between "on" and "off" periods, then, we defined an absolute transcription rate threshold between real, detectable transcription and baseline noise in measurement. We expect

transcription probabilities to be the same for the YFP and CFP reporters in Figure 3.7 so we required the threshold to yield similar statistics for each reporter despite CFP traces having greater baseline noise. Beginning in a range of transcription rate values known to be above baseline noise (based on the baseline data used in Chapter 2), the threshold was decreased as low as possible (3 AU/min) without significant qualitative differences arising between the transcription probabilities for the two reporters. We then confirmed this threshold was appropriate by inspecting mRNA traces after an "on" period in transcription. For strong transcription "on" periods, the following "off" period should correspond to an exponential decay in the mRNA time series reflecting pure degradation of mRNA with the measured half-life of ~17 min (To and Maheshri 2010). This is indeed the case for strong transcription "on" periods at a threshold of 3 AU/min (Figure 3.8), but the exponential mRNA decay after weak "on" transcription periods is both smaller and more susceptible to smoothing making it harder to detect. An absolute threshold does introduce systematic errors in transition time assignment based on "on" period height (weak periods are detected later and for less time, stronger periods are detected earlier and longer), but the transcription rate transitions are still sharp enough that assignment errors are within the ~15-20 min time resolution discussed in Chapter 2. We therefore used a threshold of 3 AU/min to designate "on" period boundaries for all transcription features (Figure 3.7).

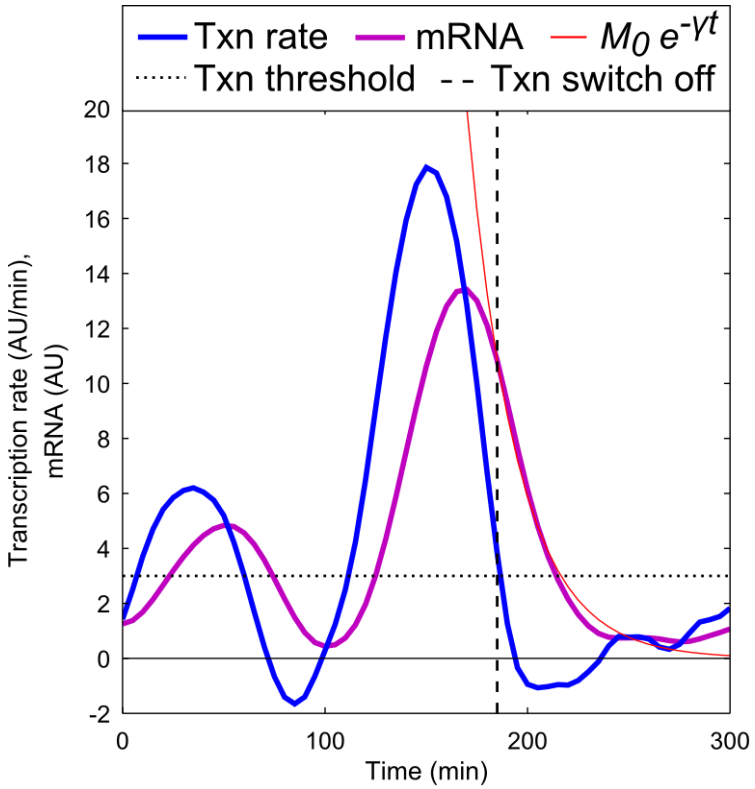


Figure 3.8 | mRNA levels fall exponentially with the expected degradation rate after periods of strong transcription.
For the single cell in the bottom panel of Figure 3.7, the CFP transcription rate (blue) and mRNA level (purple) show sporadic transcription at low expression levels. At the end of a strong "on" period (vertical, dashed line), where transcription rate drops below the threshold (horizontal, dotted line), the inferred mRNA level drop is consistent with the expected first-order degradation process $M(t) = M_0 e^{-\gamma t}$ (red line), where γ is the mRNA degradation rate and M_0 is a constant.

After assigning each transcription rate data point as "on" or "off" based on the threshold and mapping time to normalized cell cycle progression as above, we find that "on" pulses are restricted to S/G2, generally begin within 20 minutes of bud formation, and last until division (Figure 3.9A). Because of the tendency to activate around the time of budding, the timings of pulses from the two homologous P_{7xtetO} promoters are correlated. If both P_{7xtetO} copies turn on, >70% of the time they do so within 15 minutes of each other (ρ = 0.42, Figure 3.9B). In general, the "on" periods are not independent ($p < 10^{-5}$, χ^2 test; ρ = 0.42) at each locus (Figure 3.9C). These results are in striking contrast to the view of transcriptional bursting as intrinsically driven with exponential interarrival times (Raj et al. 2006; Raj and van Oudenaarden 2009; Larson et al. 2009; Golding et al. 2005).

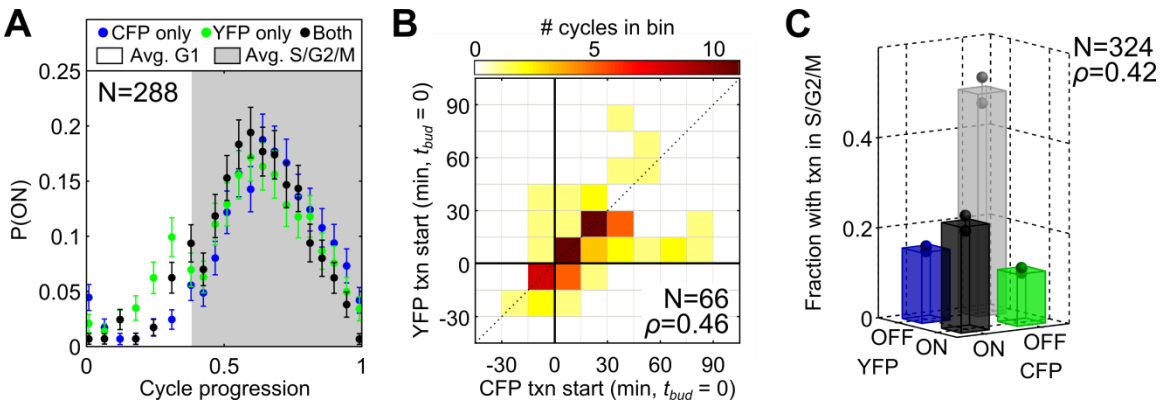


Figure 3.9 | Transcriptional bursts from homologous loci are cell-cycle dependent and partially correlated.
(*A*) The probability that each P_{7xtetO}'s transcription rate is above background, computed by averaging individual cell responses at different cell-cycle progression, increases after G1. (*B*) A 2D histogram of activation time for each promoter when both activate (t = 0 at budding). Most activation occurs near budding and is correlated. (*C*) Classifying single-cell S/G2/M periods from (*A*) by whether each P_{7xtetO} activates reveals correlations in sporadic expression. Error bars represent SEM from bootstrapping.

Disaggregating static population distributions of mRNA level by cell-cycle phase confirms G1 inactivity at tetO promoters

Further support for these real-time observations was obtained in our lab by using single-molecule mRNA fluorescence in situ hybridization (FISH) to probe how mRNA numbers in single cells varied with cell-cycle phase, classified based on the presence and size of a bud (Zopf et al.). Figure 3.10A&B describe mRNA distributions from cells with P_{1xtetO} and P_{7xtetO}, but no activator present (basal expression). The G1 distributions are zero-peaked with a long-tail that disappears by early S, suggesting transcription does not occur in G1, consistent with Figure 3.9A. Progression through S/G2/M leads to a unimodal non-zero-peaked distribution in G2/M consistent also with real-time observations (Figure 3.9B) indicating the time when an inactive promoter turns ON in S/G2 is variable. With intermediate expression, there is also increased activity in S/G2/M, but the G1 distribution is qualitatively different: a non-zero peak indicates transcription now occurs in G1 (Figure 3.10C&D). However, low transcription activity does not imply G1 inactivity. The weak but constitutive *DOA1* promoter (P_{DOA1}) has a non-zero G1 peak even with low mRNA levels (Figure 3.10E). While the mRNA distributions for the total

population in Figure 3.10 exhibit the long-tailed shape predicted by the standard model (Paulsson et al. 2000; Raj et al. 2006), partitioning the data by cell-cycle phase reveals it is incorrect.

3.4 Conclusion

Our results indicate the G1 to S/G2 transition has strong effects on transcriptional activity beyond differences in gene dosage for the tetO promoters, which have characteristics (strong TATA box, regulable) of "noisy" promoters identified in genome-wide studies (Bar-Even et al. 2006; Newman et al. 2006). Temporary disruption of a repressed promoter's chromatin architecture during DNA replication could explain the pulse timing in Figure 3.9B. Whatever the event, it does not occur independently at homologous loci. Our data alter the interpretation of studies where static mRNA or protein distributions are fit to stochastic models of gene expression to infer steady-state dynamics (Raj et al. 2006; Mao et al. 2010; To and Maheshri 2010; Munsky et al. 2012). This difficulty of using static data to pinpoint origins of variability has been anticipated (Taniguchi et al. 2010; Hilfinger and Paulsson 2011), although even static mRNA FISH data can reveal additional dynamic information (Wyart et al. 2010) as demonstrated by disaggregating mRNA distributions by cell-cycle phase (Figure 3.10) (Zopf et al.). New models incorporating cell-cycle linked pulses of transcription should alter predictions of gene network behavior.

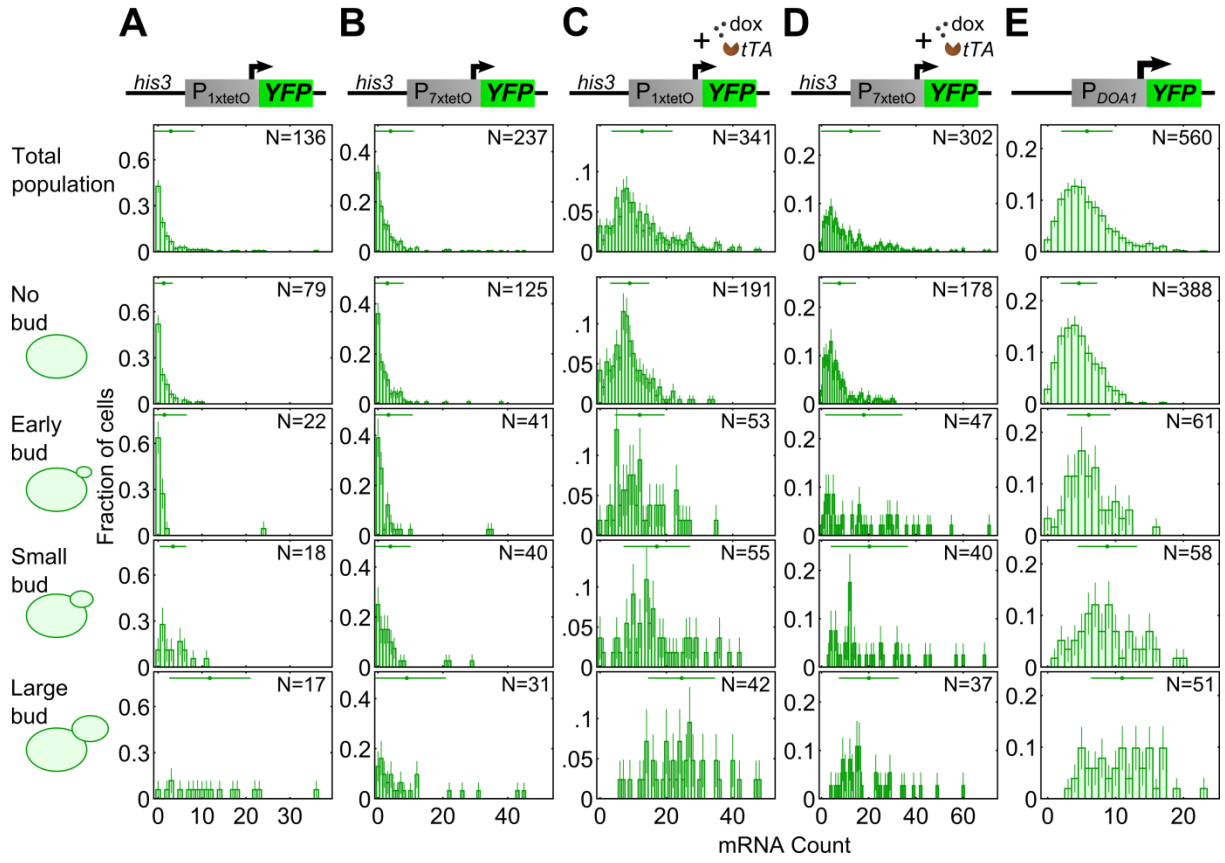


Figure 3.10 | Large differences in transcriptional activity between S/G2/M and G1 depend on the promoter.
(*A*) YFP mRNA distributions in a haploid yeast with integrated P_{1xtetO}-*YFP* and no tTA are shown in a column as a function of cell-cycle phase. Points and horizontal lines above each distribution are the experimental mean and standard deviation, respectively. (*B*) As in (*A*) but with P_{7xtetO}. (*C* & *D*) As in (*A* & *B*) but with tTA and 100 or 500 ng/mL dox added for P_{1xtetO} and P_{7xtetO}, respectively. (*E*) Integrated P_{DOA1}-*YFP* with native *DOA1* expressed from a plasmid. (Figure courtesy of Katie Quinn.)

3.5 References

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

Choi P, Cai L, Frieda K, Xie X. 2008. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**: 442–448.

Chubb JR, Trcek T, Shenoy SM, Singer RH. 2006. Transcriptional pulsing of a developmental gene. *Current Biology* **16**: 1018–25.

Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437**: 699–706.

Cookson NA, Cookson SW, Tsimring LS, Hasty J. 2010. Cell cycle-dependent variations in protein concentration. *Nucleic Acids Research* **38**: 2676–2681.

Elowitz M, Levine A, Siggia E, Swain P. 2002. Stochastic gene expression in a single cell. *Science* **297**: 1183–1189.

Golding I, Paulsson J, Zawilski S, Cox E. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**: 1025–1061.

Goranov AI, Cook M, Ricicova M, Ben-Ari G, Gonzalez C, Hansen C, Tyers M, Amon A. 2009. The rate of cell growth is governed by cell cycle stage. *Genes & Development* **23**: 1408–22.

Hilfinger A, Paulsson J. 2011. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12167–72.

Larson DR, Singer RH, Zenklusen D. 2009. A single molecule view of gene expression. *Trends in Cell Biology* **19**: 630–7.

Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. 2011. Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332**: 475–478.

Maiuri P, Knezevich A, De Marco A, Mazza D, Kula A, McNally JG, Marcello A. 2011. Fast transcription rates of RNA polymerase II in human cells. *EMBO Reports* **12**: 1280–5.

Mao C, Brown CR, Falkovskaia E, Dong S, Hrabeta-Robinson E, Wenger L, Boeger H. 2010. Quantitative analysis of the transcription control mechanism. *Molecular Systems Biology* **6**: 431.

Munsky B, Neuert G, Van Oudenaarden A. 2012. Using gene expression noise to understand gene regulation. *Science* **336**: 183–7.

Muramoto T, Cannon D, Gierlinski M, Corrigan A, Barton GJ, Chubb JR. 2012. Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 7350–5.

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

Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. 2006. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**: 840–6.

Paulsson J, Berg OG, Ehrenberg M. 2000. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7148–53.

Peccoud J, Ycart B. 1995. Markovian Modeling of Gene-Product Synthesis. *Theoretical Population Biology* **48**: 222–234.

Pedraza J, Van Oudenaarden A. 2005. Noise propagation in gene networks. *Science* **307**: 1965–1974.

Raj A, Van Oudenaarden A. 2009. Single-molecule approaches to stochastic gene expression. *Annual Review of Biophysics* **38**: 255–70.

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA synthesis in mammalian cells. ed. U. Schibler. *PLoS Biology* **4**: e309.

Raser JM, O'Shea EK. 2004. Control of stochasticity in eukaryotic gene expression. *Science* **304**: 1811–1814.

Rizzo MA, Davidson MW, Piston DW. 2009. Fluorescent protein tracking and detection: fluorescent protein structure and color variants. *Cold Spring Harbor Protocols* **2009**: pdb.top63.

Shahrezaei V, Swain P. 2008. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 17256–17317.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* **22**: 1567–72.

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* **332**: 472–474.

Tan RZ, Van Oudenaarden A. 2010. Transcript counting in single cells reveals dynamics of rDNA transcription. *Molecular Systems Biology* **6**: 358.

Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A, Xie XS. 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**: 533–8.

To T-L, Maheshri N. 2010. Noise Can Induce Bimodality in Positive Transcriptional Feedback Loops Without Bistability. *Science* **327**: 1142–1145.

Trcek T, Larson D, Moldón A, Query C, Singer R. 2011. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* **147**: 1484–1581.

Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J. 2006. Origins of extrinsic variability in eukaryotic gene expression. *Nature* **439**: 861–4.

Wyart M, Botstein D, Wingreen NS. 2010. Evaluating gene expression dynamics using pairwise RNA FISH data. *PLoS Computational Biology* **6**: e1000979.

Zopf CJ, Quinn K, Zeidman J, Maheshri N. 2013. Cell-cycle dependence of stochastic gene expression. *PLoS Computational Biology (submitted)*.

CHAPTER 4. A COMBINATION OF PROMOTER ARCHITECTURE AND CELL-CYCLE PHASE INFLUENCES DELAYS IN GENE ACTIVATION[‡]

4.1 Introduction

Promoter chromatin contributes substantially to a gene's transcriptional regulation. As discussed in Chapter 1, the positioning of nucleosomes relative to *cis*-regulatory elements in yeast distinguishes two classes of promoters: "open" promoters where binding sites for transcription factors (TFs) are present and accessible in a nucleosome-depleted region (NDR), and "covered" promoters where nucleosomes flanking an NDR occlude recognition sequences necessary to achieve transcription (Tirosh and Barkai 2008; Field et al. 2008). The covered promoter class is enriched for regulated genes and has greater dynamic range and expression noise (Tirosh and Barkai 2008) owing to the competition between nucleosomes, transactivors, and the general transcription machinery to bind the same sequences. Promoter-specific studies at inducible *GAL* and *PHO* promoters highlight this competition by demonstrating the requirement of chromatin disruption for transcription (Lohr 1997). In the next section, we discuss how the *PHO5* promoter has served as a model for chromatin regulated-transcription, and how this has improved our understanding of the relationship between promoter architecture and its steady-state expression dynamics.

However, the role of promoter chromatin architecture in regulating gene activation kinetics is less well understood despite the transient behavior of regulated genes being crucial to proper stress response, pathway coordination, and cell-cycle control. A comparison of Msn2p target promoters revealed slower response kinetics for nucleosome-occluded architectures, but this and other kinetic studies of *PHO5* promoter activation thus far have been limited by population-average measurements (Barbaric et al. 2001; Jessen et al. 2006; Lam et al. 2008; Zhou and Zhou 2011). Single-cell approaches to signaling and observing gene induction have likewise been susceptible to the confounding possibility of unobservable delays in signal transduction (Charvin

[‡] Portions of the text and figures in this chapter are drawn from (Zopf et al. 2013a, 2013b).

et al. 2008). To address these issues, we developed a switchable transactivator in *S. cerevisiae* based on the phosphate starvation pathway. Here we apply this system to a library of *PHO5* promoter variants (Lam et al. 2008) to study the relationship between promoter architecture and kinetics of transcription activation, and we find that cell-cycle phase plays a major role in determining the activation delay at individual promoters.

4.2 The *PHO5* promoter as a model for chromatin-regulated transcription

Phosphate-responsive promoters in general and the *PHO5* promoter in particular have served as a model system for studying the relationship between promoter chromatin and transcription regulation. As a member of the phosphate starvation pathway, *PHO5* is regulated by the TF Pho4p (Fascher et al. 1990). Pho4p is the master regulator of the yeast phosphate starvation pathway, and translocates to the nucleus in response to low intracellular phosphate (O'Neill et al. 1996) in turn activating promoters of "*PHO"* genes (Springer et al. 2003). These promoters for the *PHO* genes lie in the "covered" class with a canonical architecture of at least 3 well-positioned nucleosomes. The -1 nucleosome occludes the TATA box, and Pho4p binding sites of varying affinity can be present either "under" or in a NDR between the -2 and/or -3 nucleosome. The *PHO5* promoter in particular has a low-affinity Pho4p binding site in the NDR (traditionally referred to as UASp1, but here referred to as a nucleosome-free site or NFS), and a high-affinity Pho4p binding site under the -2 nucleosome (traditionally referred to as UASp2, but here referred to as a nucleosome-occluded site or NOS) (Almer et al. 1986; Rudolph 1987; Vogel et al. 1989) (Figure 4.1). Despite the possibility of stochastic DNA unwinding from the -2 nucleosome, Pho4p binding at the UASp2 is not observable until it has bound the UASp1 and instigated chromatin remodeling (Svaren 1997). Various cofactor knockouts have demonstrated a degree of redundancy in the activation of the *PHO5* promoter (Barbaric et al. 2007; Korber et al. 2006), but transcription activation does require chromatin disassembly (Adkins et al. 2004).

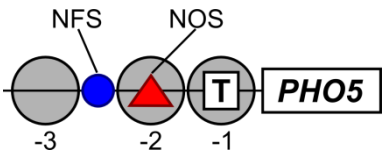


Figure 4.1 | The *PHO5* promoter.

The roles of several HATs, remodelers, and chaperones have been implicated in efficient *PHO5* activation. Initially, the cofactor Pho2p recruits NuA4 to poise the promoter through its H4 acetylation activity (Nourani et al. 2004). This may be to balance the repressing effect of Rpd3L histone deacetylation recruited to the promoter by Set1-dependent H3K4 methylation, which serves to prevent histone turnover at the promoter (Wang et al. 2011). Acting before Pho4p arrival, promoter poising is crucial to efficient induction, and elimination of these acetylation or deacetylation activities slows activation or increases basal expression, respectively. Upon phosphate depletion, Pho4p binds UASp1 and recruits the SAGA HAT and SWI/SNF chromatin remodeling complexes (Barbaric et al. 2003; Neely et al. 2002; Dhasarathy and Kladde 2005) with additional roles observed for the HAT Rtt109 (Williams et al. 2008) and the remodeler INO80 (Steger et al. 2003). The combined activities of these complexes drive accessibility of UASp2, more stable binding of Pho4p (Barbaric et al. 2001; Reinke and Hörz 2003), and eventual removal of nucleosomes across the promoter (Korber et al. 2004), which liberates the TATA box from the -1 nucleosome. Efficient eviction also requires the histone chaperone Asf1 (Adkins et al. 2004; Korber et al. 2006). Pho4p finally stabilizes pre-initiation complex formation (Mao et al. 2010) possibly mediated by the association between SAGA and the TATA binding protein (Barbaric et al. 2003). Continued transcription requires Pho4p to maintain an open promoter by antagonizing histone deposition by the chaperone Spt6, but not necessarily to maintain the initiation complex (Adkins and Tyler 2006). The chromatin remodeling function recruited by Pho4p is therefore the crucial purpose in its binding to the *PHO5* promoter. An additional factor that binds the *PHO5* promoter, Mcm1p, has been implicated in mitotic induction of the promoter (Pondugula et al. 2009), but this phenomenon is related to cycle-dependent phosphate depletion. Appreciable activation only occurs in the presence of Pho4p (Neef and Kladde 2003), which suggests only to a supporting role for Mcm1p possibly contributing to nucleosome disruption.

Other *PHO* promoters possess different numbers and affinity of Pho4p binding sites within these two regions. A detailed study using a library of promoter mutants based on *PHO5* showed that *PHO* promoter chromatin architecture serves to disentangle the regulatory parameters in the steady-state response (Lam et al. 2008). Specifically, the NFS site affinity governs the threshold for induction by Pho4p while the NOS dominates the expression level. This strategy provides the cell a way to activate subsets of the *PHO* pathway in response to intermediate stimuli

independently of the output strength. Thus, at intermediate levels of phosphate when nuclear Pho4p activity is moderated, the cell can first activate the high-affinity exposed site genes *PHM4* and *PHO84* to mobilize internal phosphate stores and take up phosphate with higher affinity membrane transporters. Only if phosphate conditions become more dire will Pho4p activity be sufficient to activate the low-affinity exposed site promoters like that of *PHO5* and other phosphate scavenging genes thus conserving energy and materials for the most severe stresses (Lam et al. 2008; Steger et al. 2003). Since promoter architecture clearly plays a significant role in determining the response of a gene to a transactivator input, and because nucleosome remodeling is essential to efficient activation kinetics, a role for architecture in the transient induction kinetics is likely.

4.3 Reengineering the PHO pathway to accurately measure gene activation kinetics of PHO promoters

We sought to measure to kinetics of activation at a series of *PHO5* promoter variants with precise changes in promoter architecture that mimic differences in natural *PHO* promoters. In high phosphate conditions, the master transcription factor Pho4p, responsible for the phosphate starvation response by activating *PHO* genes, is preferentially exported to the cytoplasm via phosphorylation of multiple serine residues (Kaffman et al. 1994; Kaffman et al. 1998) Transfer of yeast cells to low external phosphate conditions leads to eventual depletion of intracellular phosphate levels (Shirahama et al. 1996), inhibition of Pho4p phosphorylation, and translocation of the transcription factor from the cytoplasm to the nucleus (Kaffman et al. 1994; O'Neill et al. 1996) where it activates *PHO* promoters. Previous measures of *PHO5* activation kinetics in single cells usually entail transferring yeast cells from high to low phosphate conditions, and then measuring gene expression (Thomas and O'Shea 2005) or chromatin state (Jessen et al. 2006) of the *PHO5* promoter. While such measurements have indicated cell-to-cell variability in remodeling and expression, they are unable to decouple the kinetics of the upstream phosphate signaling pathway from gene activation. In fact, in one of these studies activation delays were found to be due to the time required to deplete vacuolar stores of polyphosphate (Thomas and O'Shea 2005). Additionally, the single-cell measurements employed were static population snapshots. Though these studies provided timecourse data for the evolution of the population response, single promoter activation trajectories cannot be inferred.

To obviate these problems and increase the time resolution of the kinetic measurement, we developed a switchable transactivator system (1) to rapidly control and observe the nuclear translocation of Pho4p, and (2) to measure subsequent expression from a *PHO5* promoter variant driving a CFP reporter, all in the same single cell (Figure 4.2). This required multiple changes in the endogenous pathway. To ensure rapid translocation in response to a step change in external phosphate levels, we eliminated major feedback pathways and vacuolar storage of polyphosphate. Normally, cells employ a dual positive and negative feedback loop architecture to respond to changes in external phosphate concentration. This results in switching the complement of high and low affinity phosphate transporters present on the membrane and can even result in a bimodal population response at intermediate levels of phosphate (Wykoff et al. 2007). Our strategy was to eliminate these feedback loops and retune the system to meet our criteria. We started with strain EY2210 (Appendix, Table A.1, (Wykoff et al. 2007)). The *spl2Δ* prevents down regulation of low affinity phosphate transporters in low phosphate conditions and hence eliminates the positive feedback loop. Also *PHO84*, encoding the high affinity transporter, is replaced with CFP. With no high affinity phosphate transporter, which normally is upregulated in low phosphate conditions (Bun-Ya et al. 1991), there is no negative feedback loop. This strain also expresses an Nhp2p-mCherry protein fusion to fluorescently label the nucleus. Because we sought to look at CFP expression from tetO promoters, we replaced *CFP* at the *PHO84* locus with a *KlURA3* (using a PCR fragment from B108, Appendix, Table A.2) and regenerated uracil auxotrophy by eliminating a portion of *KlURA3* (using a short PCR fragment from B464, Appendix, Table A.2). Utilization of vacuolar stores of phosphate can delay the response to changes in external levels (Thomas and O'Shea 2005) so we eliminated a key vacuolar transporter, Phm4p, by integrating a *phm4Δ::HIS3MX6* PCR fragment. In this background, growth in the microfluidic chamber in low external phosphate (5 μM) slows relative to growth in high phosphate (5000 μM) (Figure 4.3); however, this occurs gradually and does not significantly impact either changes in Pho4p's nuclear localization or the transcription rate until late times. (While 5 μM external phosphate leads to a severe growth defect in culture and is reportedly below the K_m of the low affinity transporters (Wykoff and O'Shea 2001), we have found both nutrients and drugs are more potent in the microfluidic chamber. This is likely due to both constant perfusion and a reduction in mass transfer limitations present in culture.) These genetic modifications thus enabled toggling the nuclear state of Pho4p using an external

phosphate concentration that has minimal effects on growth rate on the time-scale of reporter gene activation.

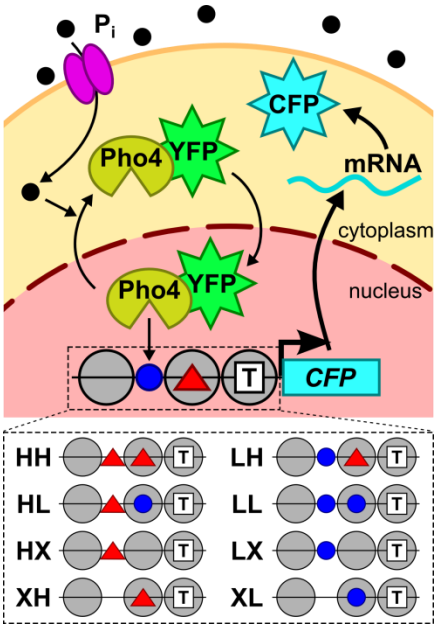


Figure 4.2 | A switchable transactivator system to probe promoter activation kinetics in yeast.
Extracellular phosphate concentration controls the nucleo-cytoplasmic localization of the Pho4-YFP TF fusion through an engineered phosphate starvation pathway. When localized to the RFP-marked nucleus, the TF drives expression of CFP from the *PHO5* promoter or a variant differing only in the affinities of the open and nucleosome-occluded Pho4p binding sites (high: "H", Δ; low: "L", ○; ablated: "X", no symbol).

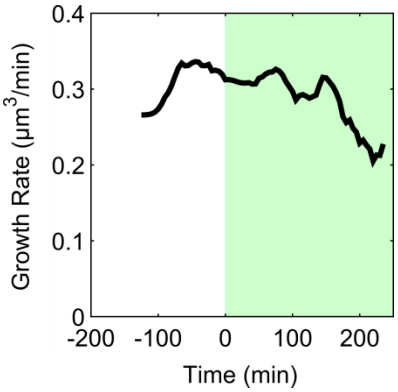


Figure 4.3 | Mean growth rate decreases only gradually in low phosphate.
The switchable transactivator strain was grown in the microfluidic device in high phosphate (white area) before switching to low phosphate media (light green area). The mean instantaneous growth rate varies

over time due to cell-cycle synchronization imposed by trapping similarly-sized cells in the device and cell cycle-dependent changes in growth rate (discussed in Chapter 3).

To observe the translocation, we expressed a Pho4-YFP (Citrine (Griesbeck et al. 2001)) fusion and compared its signal colocalization with the RFP-marked nucleus. We chose to constitutively overexpress Pho4-YFP to rapidly attain and exceed a phosphate-starved nuclear Pho4p level, rather than rely on slow positive feedback through Pho4p-dependent increases in expression of *PHO81* (Yoshida et al. 1989; Creasy et al. 1996), the inhibitor of the cyclin-CDK which drives Pho4p nuclear export. This ensures the downstream steady-state promoter response is quickly saturated with respect to Pho4p and not sensitive to small fluctuations in its concentration. Finally, the nuclear YFP signal we observe represents a pool of Pho4p phosphoforms that may have varying transcriptional activity based on phosphorylation of S223, referred to as site 6 (O'Neill et al. 1996; Springer et al. 2003). Therefore we utilized a mutant $Pho4^{SD6}$ (Komeili and O'Shea 1999) with the site 6 serine converted aspartate, which fixes site 6 in a "phosphorylated" state and prevents binding to the Pho2p cofactor. Thus $Pho4^{SD6}$ has reduced but still potent transcriptional activity (Springer et al. 2003), and fluorescence observed from the nuclear Pho4-YFP pool represents a uniform activity. This reengineered strain was then grown in a microfluidic device that allows rapid (< 1 min) changes in external phosphate concentrations. By monitoring large step changes in Pho4-YFP nuclear translocation in single cells using real-time fluorescence microscopy at 5 min intervals, we determined the response delays of individual target promoter activation using the automated computational tools described in Chapter 2.

Figure 4.4 describes the kinetics of Pho4-YFP localization (nuclear YFP signal – cytoplasmic YFP signal) and activation of the *PHO5* promoter driving *CFP*, respectively. In Figure 4.4B, we report the resulting instantaneous *CFP* transcription rate, which we infer using the measured CFP total fluorescence time series and the continuous-time transcription model in Chapter 2. While the average transcription rate over the population reaches a constant steady-state after 100 mins in low phosphate, in single cells the transcription rate exhibits cell cycle-dependent fluctuations (Figure 4.5) consistent with our previous observations at constitutively-active endogenous or induced synthetic promoters in Chapter 3. When individual cells activate transcription, they exhibit a smooth rise in the transcription rate. While the absolute value to which this rate rises depends on the particular cell, all cells show a similar characteristic shape and duration of the

rise (Figure 4.6). Given the measurement frequency, noise, and the spline fits used to smooth and differentiate the CFP time series (see section 2.4), this characteristic rise time (~40 min) in the inferred transcription rate is consistent with a sharp transition in the actual transcription rate smoothed over the ~20 min time resolution window. (The inferred, smoothed transcription rate will begin to rise ~20 min before the actual transition, and will not fully reach the "on" state until ~20 min after the switch.)

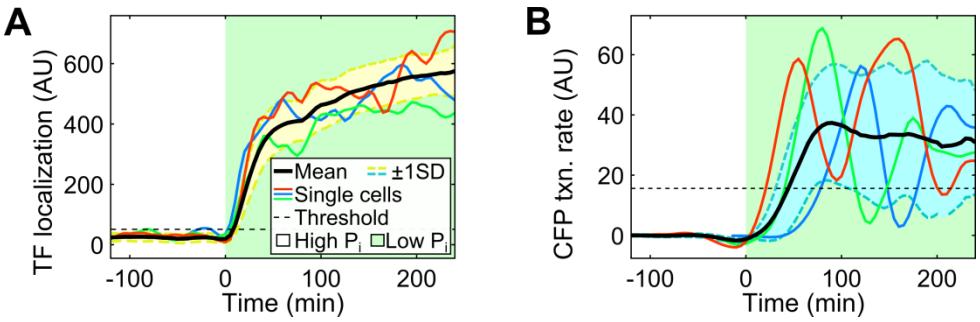


Figure 4.4 | A step change in nuclear Pho4-YFP leads to *CFP* transcription activation in single cells.
(*A*) After a step to low phosphate growth conditions, the TF localization (nuclear – cytoplasmic YFP intensity) in single cells quickly rises above a threshold level sufficient for promoter activation. (*B*) In response to the TF localization in each cell in *A*, an instantaneous *CFP* transcription rate is calculated that crosses the "on" threshold at distributed times. Data shown for the wild-type *PHO5* promoter. (Symbols as in *A*.)

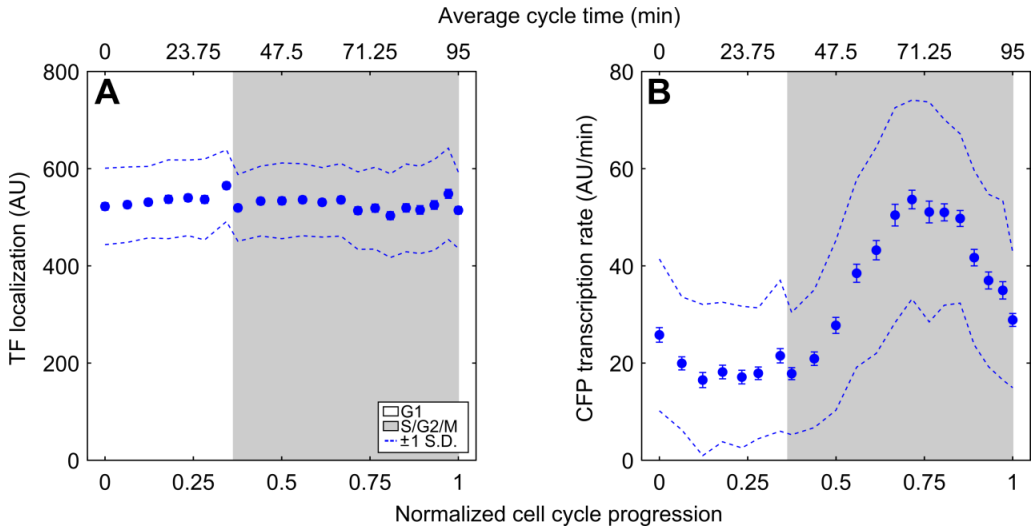


Figure 4.5 | Transcription from the *PHO5* promoter oscillates across the cell cycle at steady state.
Data for each cell cycle at steady state (>100 minutes after the phosphate switch) is mapped to a normalized cell cycle progression as in Chapter 3 (Figure 3.2) and the mean is taken for bins across the cell cycle. (*A*) Mean Pho4-YFP localization is not a function of cell cycle progression while (*B*) mean *CFP* transcription rate exhibits strong cell cycle oscillations. Dashed lines represent 1 S.D. from the mean (points), and error bars represent the S.E.M. by bootstrapping.

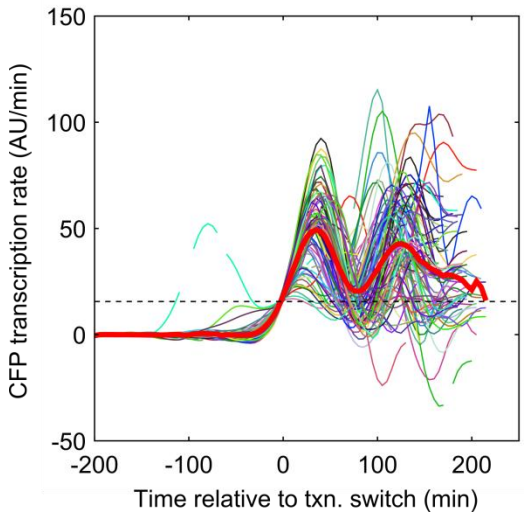
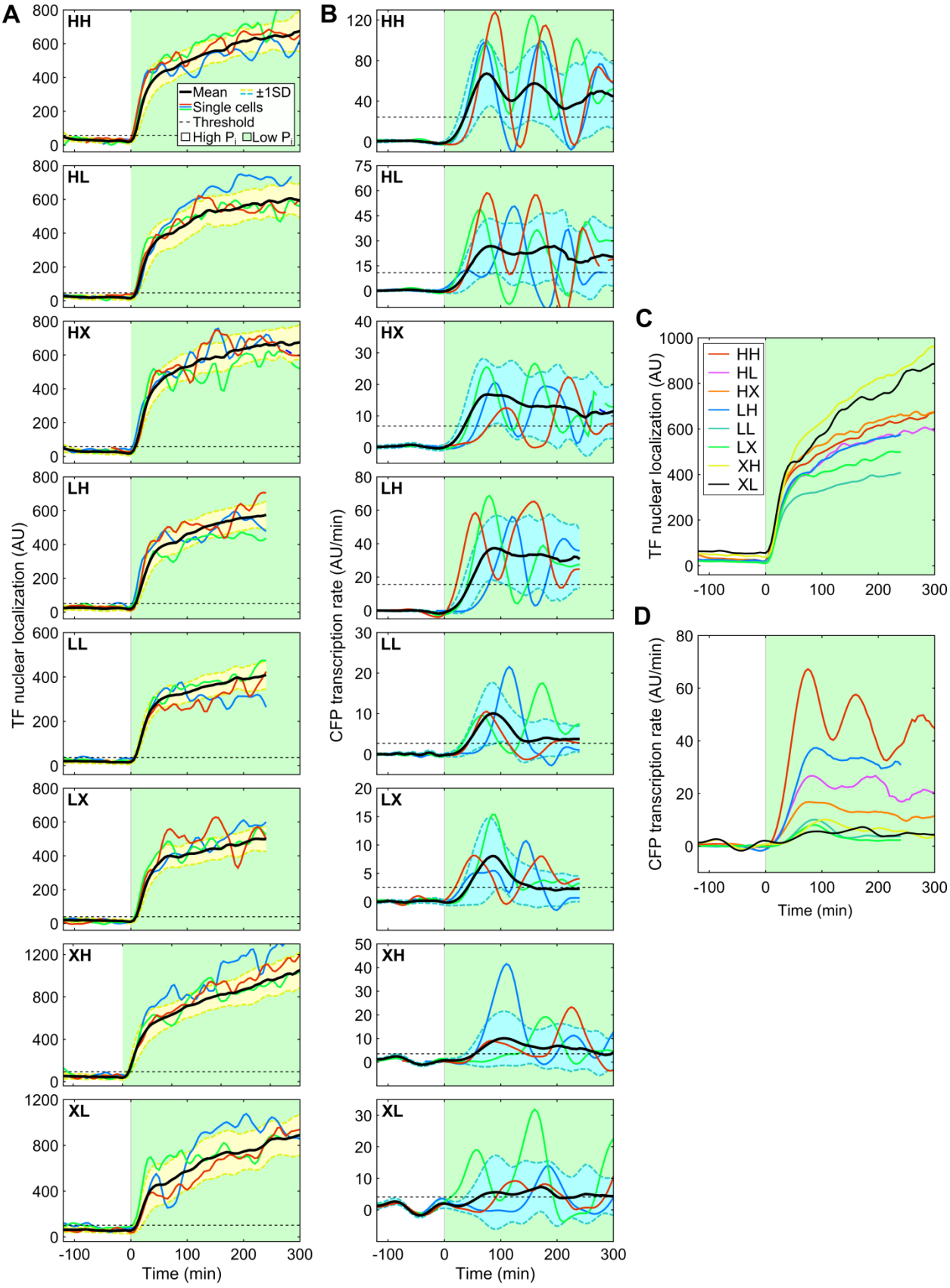


Figure 4.6 | Single cell traces exhibit qualitatively similar rises to steady-state transcription.
Single cell transcription rate time series (thin, multicolored lines) were synchronized by the time at which each crosses the switch "on" transcription rate threshold (dashed, black line), and the mean is plotted at each time interval (thick, red line).

We then need an unbiased method to assign transcription transition times for precise comparisons of promoter variant activation kinetics in the case of a switch on to an oscillating steady state. If we model the transcription rate response to a step change in phosphate as a Heaviside function, a spline fit to simulated data yields a value at the activation time that is midway between the baseline and active transcription rates (see section 2.4). Even with cell cycle-dependent oscillations, activation would occur on average at the time transcription rate reaches the midway point to the new, time-averaged steady state (Figure 2.4). The switch time is then estimated at the midpoint between the "off" and "on" levels. The "off" level is ~0 AU/min for each promoter, but a step test of each promoter variant depicted in Figure 4.2 results in a range of active transcription strengths (Figure 4.7). Therefore, the "on" transcription rate resulting from the step change must be calculated for each promoter variant. In each case, the steady-state transcription rate varies with the cell cycle, which yields an oscillating, dynamical steady state. However, the mean transcription rate trends down at later times for the weaker LL and LX promoters (Figure 4.7B) most likely due to a strong link to the decreasing growth rate (section 3.2, Figure 4.3) while the mean for XH and XL hovers around the detection limit. Since the long-time mean is not necessarily representative of the "on" state at early times, we approximate the initial "on" level based on the first peak to trough in transcription rate after activation. We take the mean of the midpoint between the first peak >3 AU/min and the following trough in transcription rate in all cells as the new mean "on" level for transcription rate immediately after activation. We are then able to calculate automatically a transcription rate threshold for each promoter architecture as half of the estimated mean "on" level.

Figure 4.7 | Step responses for each promoter in the *PHO5* variant library (next page).
(A) Single-cell and mean Pho4-YFP nuclear localization in response to a step change in phosphate for each promoter variant quickly reach the localization threshold determined sufficient to activate transcription (Figure 4.8). (B) Single-cell and mean CFP transcription rate driven by the TF localization in A reach the promoter-specific switch "on" threshold at variable times. (C) Mean TF localization and (D) mean CFP transcription rate overlay for all promoter variants.


A
B
C
D
HH
HL
HX
LH
LL
LX
XH
XL
Mean
±1SD
Single cells
Threshold
High P_i
Low P_i
TF nuclear localization (AU)
CFP transcription rate (AU/min)
Time (min)

To find the delay in activation, we determine when nuclear Pho4-YFP levels have crossed a threshold sufficient to activate expression. We can estimate this threshold by ordering single cells by the localization level at the time of activation (Figure 4.8). We take as the threshold the localization level at which 5% of single cells have activated. Due to the low variability in single-cell localization profiles (Figure 4.7A), this threshold is robust; raising it merely shifts the localization time for all cells equally. The activation delay for each cell is then estimated as the time between localization and transcriptional activation. This value represents delays in remodeling, initiation, elongation, translation, and maturation of the CFP reporter (~ 11 min half-life, section 2.4). Because the latter three processes are identical for every promoter variant, any differences in delays we see are related to forming an active promoter state.

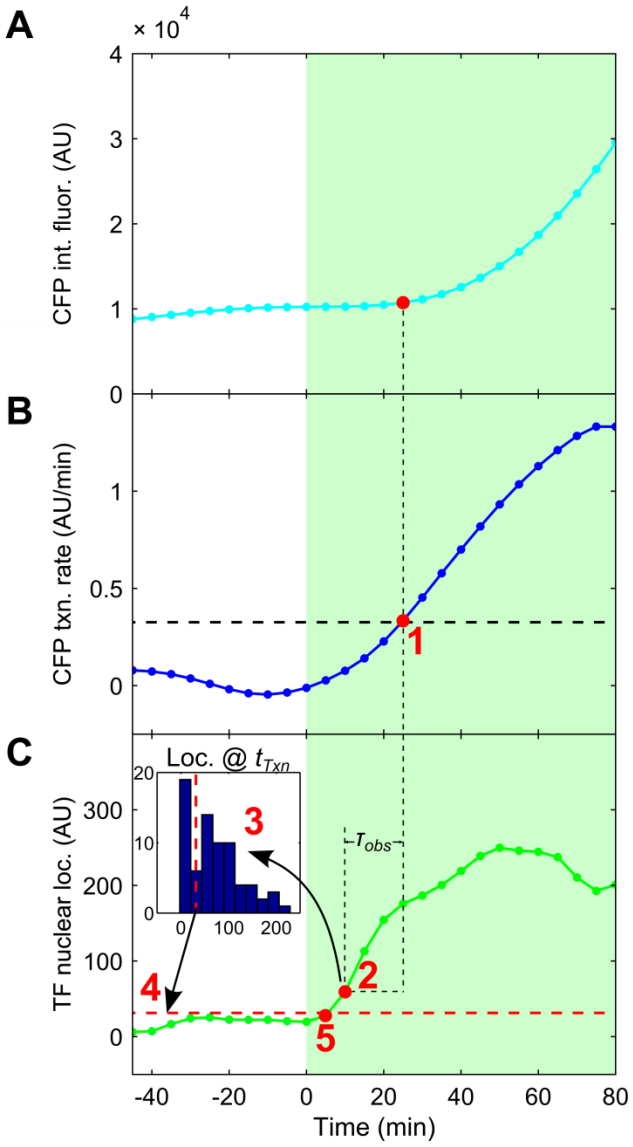


Figure 4.8 | Determining transcription activation and effective TF nuclear localization times in single cells during step tests (next page).
After the switch to low phosphate at t = 0 (green shaded area), a sample cell begins to express CFP (*A*, cyan) in response to the TF localizing to the nucleus (*C*, green). The inferred CFP transcription rate (*B*, blue) rises after the TF localizes to an effective level. Transcription and localization transition times for each cell are then found in 5 calculation steps (corresponding to the red numbers on plots). (1) The transcription threshold (horizontal black dashed line, *B*) is set at 50% of the "on" transcription rate in the population as discussed in the text. This threshold point corresponds well to the beginning of a rise in observable protein (*A*). (2) Because of the 10-15 min lag in observing transcription (τ_{obs}) due to protein maturation, the TF localization level 15 min prior to the observed time of transcription activation is noted. (3) Steps 1 and 2 were repeated for all cell traces creating a distribution of localization values (inset in *C*). (4) The effective TF localization threshold is set at the 5th localization percentile (vertical, dashed line in inset). This corresponds to a TF level where many cells do activate transcription

suggesting that nuclear TF levels are no longer limiting. Any resulting delays in transcription activation are interpreted as due to the transcription process and not due to a lack of nuclear TF. (5) Each cell's localization time is then determined when nuclear TF levels cross the effective TF localization threshold (horizontal red dashed line, C).

4.4 Activation kinetics is largely dictated by cell cycle phase

We measured activation kinetics of the eight *PHO5* promoter variants (Figure 4.2) subject to a drop in phosphate when grown in glucose. Upon activation, single cells from all variants exhibit the same characteristic rise to an oscillating steady-state transcription rate (Figure 4.7). As discussed in the previous section, this implies activation occurs as a fast (<20 min) switch to an active transcription state at the single promoter level, rather than a slow progression through a series of more and more active transcription states. The XH and XL variants which lack a NDR site exhibit the longest delay, whereas the HH variant exhibits the shortest delay (Figure 4.9). The delay may correlate roughly with steady-state expression levels, but certainly not in a linear manner, as significantly large differences in delays between HX, XH, and XL correspond to small differences in steady-state expression. Moreover, LH may be an exception as it has a longer delay than the lower expressing HX and HL variants, although this difference is not significant within error. At the very least, we can conclude that delay times *are* related to promoter architecture.

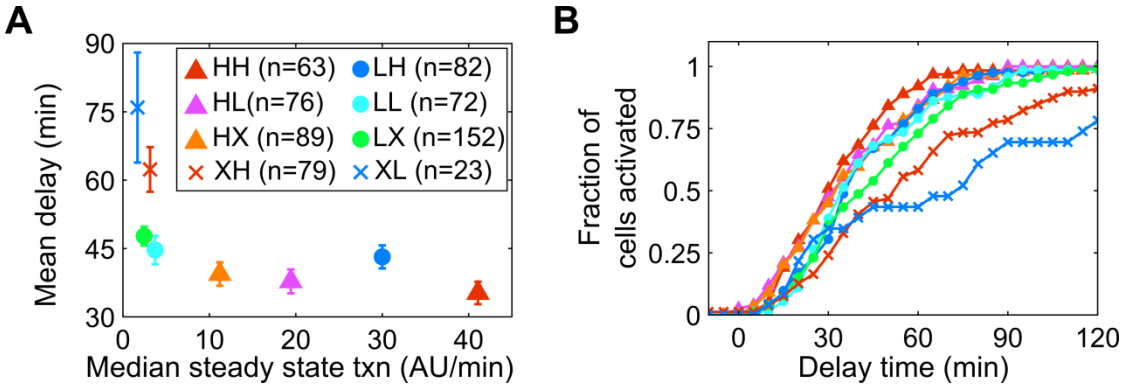


Figure 4.9 | Transcription activation delay time is related to promoter architecture.
(*A*) Each point represents the mean delay time calculated for the step tests of each promoter variant in Figure 4.7 plotted against the mean steady state transcription level for times >100 min after the phosphate switch. Error bars represent S.E.M. by bootstrapping, and are negligible for the mean steady state transcription. (*B*) The distribution of single-cell delay times from each promoter.

Because we measured delays at the single cell level, we could go beyond mean delay times and compare delay time distributions (Figure 4.9B). Especially with the XL and XH variants, there appear to be two groups of cells with shorter and longer delay times. In Chapter 3, we found that cell-cycle phase can have large effects on *steady-state* transcription rate of low expressed noisy genes that goes beyond gene dosage. Therefore we classified cells by cell-cycle phase in which localization occurred using the absence (G1) or presence (S/G2/M) of a bud. In comparing the mean delay times of these two groups (Figure 4.10), localization in G1 as compared to S/G2/M leads to a longer activation delay across promoter architectures despite no difference in localization strength between the two phases initially or over time (Figure 4.11). However, the activation delay differences are not always significant.

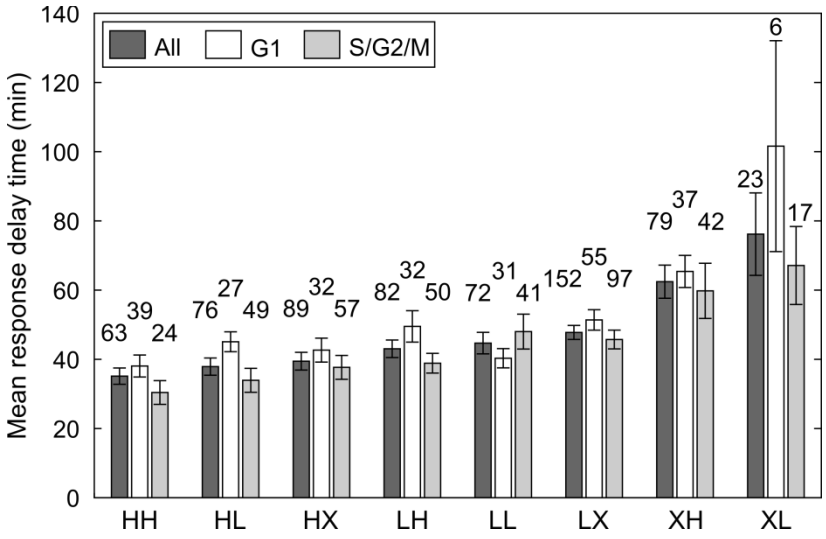


Figure 4.10 | Mean promoter response delay by cell cycle phase for step tests of the *PHO5* promoter variants in glucose media.

Bars represent the mean for all cells in which a delay was measured (dark gray), or the mean of those cells in which TF localization occurred during G1 (white) or S/G2/M (light gray). Numbers above each bar represent the number of cells averaged. Error bars represent S.E.M. by bootstrapping.

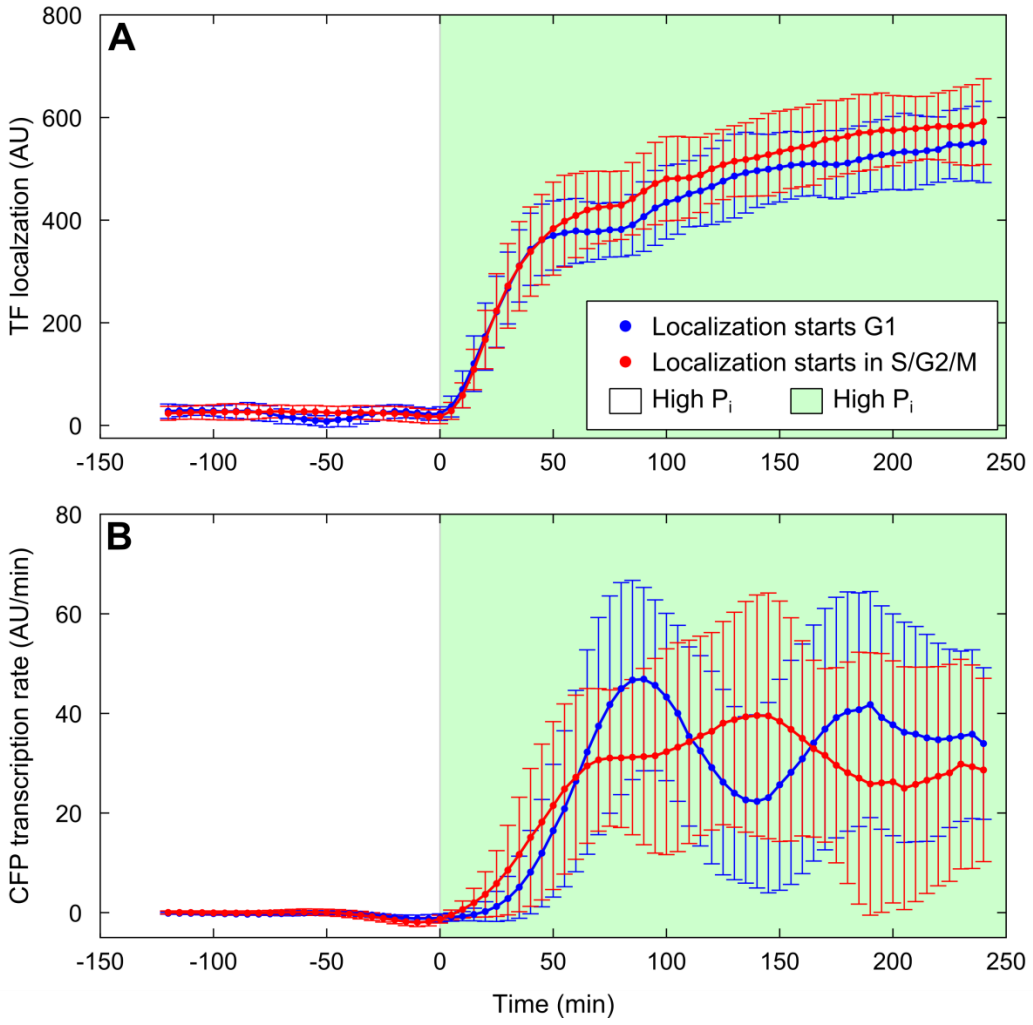


Figure 4.11 | TF localization level does not correlate to cell-cycle phase at the start of localization.
(*A*) From the wild-type *PHO5* step test in Figure 4.4, mean Pho4-YFP localization for the two classes of cells where TF localization starts in G1 versus S/G2/M. (*B*) The corresponding mean *CFP* transcription rate for each class. Error bars represent ±1 standard deviation at each time point. There is no difference in localization at early times. The small systematic difference in localization at later times is within one standard deviation, and has no effect on mean transcription rate.

We reasoned that the short time cells spend in G1 when grown in glucose (~30 min with ~100 min doubling time, Figure 4.12) might confound these results in two ways. First, some fraction of cells is likely misclassified in G1 or S/G2/M because localization is only estimated accurately within one time point (±5 minute sampling frequency). Second, cells that localize Pho4-YFP in G1 may actually activate quickly because they quickly reach the transition to S phase (<30 min).

We addressed these concerns by extending the time cells spent in G1 by repeating the step test experiments using 2% galactose as the carbon source (Figure 4.12). We chose to focus on the HH, LH, XH, and HX variants, which span different delay times and steady-state expression levels and in which the NFS is systematically varied. With the extended G1, the differences in delay times are now significant depending on whether localization occurs in G1 versus S/G2/M (Figure 4.13). To see if the widening disparity in delay times was due to the extended G1 in galactose, we visualized single cell traces by when localization and activation occur in relation to normalized cell-cycle progression (as in section 3.2). Analyzing single cell traces in this manner reveals (Figure 4.14): if localization occurs in early S/G2, there is a short delay to activation (magenta); if localization occurs in G1, there is a longer delay with the majority of events not activating until the G1/S transition (red) and a minority of events activating in G1 (blue); and if localization occurs in late G2/M (arbitrarily defined with a cell cycle progression >0.7) activation usually does not occur until the next cell cycle, and frequently not until the subsequent G1/S transition. While steady-state transcription in G1 is lower versus S/G2/M (Figure 4.15, and Figure 4.5, Figure 3.4, Figure 3.5) and activation times are determined based on a fixed threshold, the large delays are due to the initial absence of transcription in G1 and late G2/M (Figure 4.16). This analysis not only confirmed activation in G1 is slower than post-budding, it also shows that differences between the activation delay during G1 versus S/G2 are even larger because we had lumped the early S/G2 (short delay) and late G2/M events (long delay) events (Figure 4.13). Taken together, it appears that early S/G2 is a permissive period for activation of *PHO* promoters.

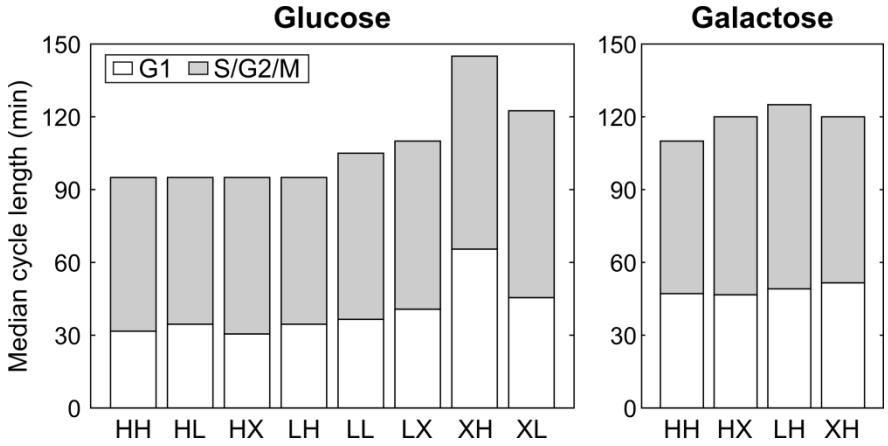


Figure 4.12 | The median cell cycle length for cells throughout step tests in glucose or galactose media.
For each observed cell cycle, the full cell cycle length (full bar height) was calculated as the time between divisions, and the G1→S transition was calculated as the median bud formation time as a fraction of the total cell cycle. The G1 length (white portion) is the product of the full cell cycle length and the median cycle progression at bud formation, and the remaining cycle time is attributed to S/G2/M (gray portion).

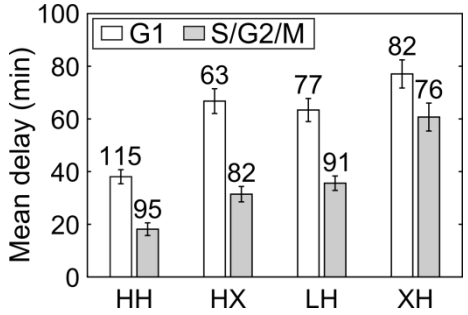


Figure 4.13 | Transcription activation is slow in G1.
For a step test in media containing galactose as the carbon source, the mean promoter response delay times categorizing each cell as G1 (white) or S/G2/M (gray) based on the point in the cell cycle at which TF localization occurred. The number of cells in each category appears above each bar, and error bars represent S.E.M by bootstrapping.

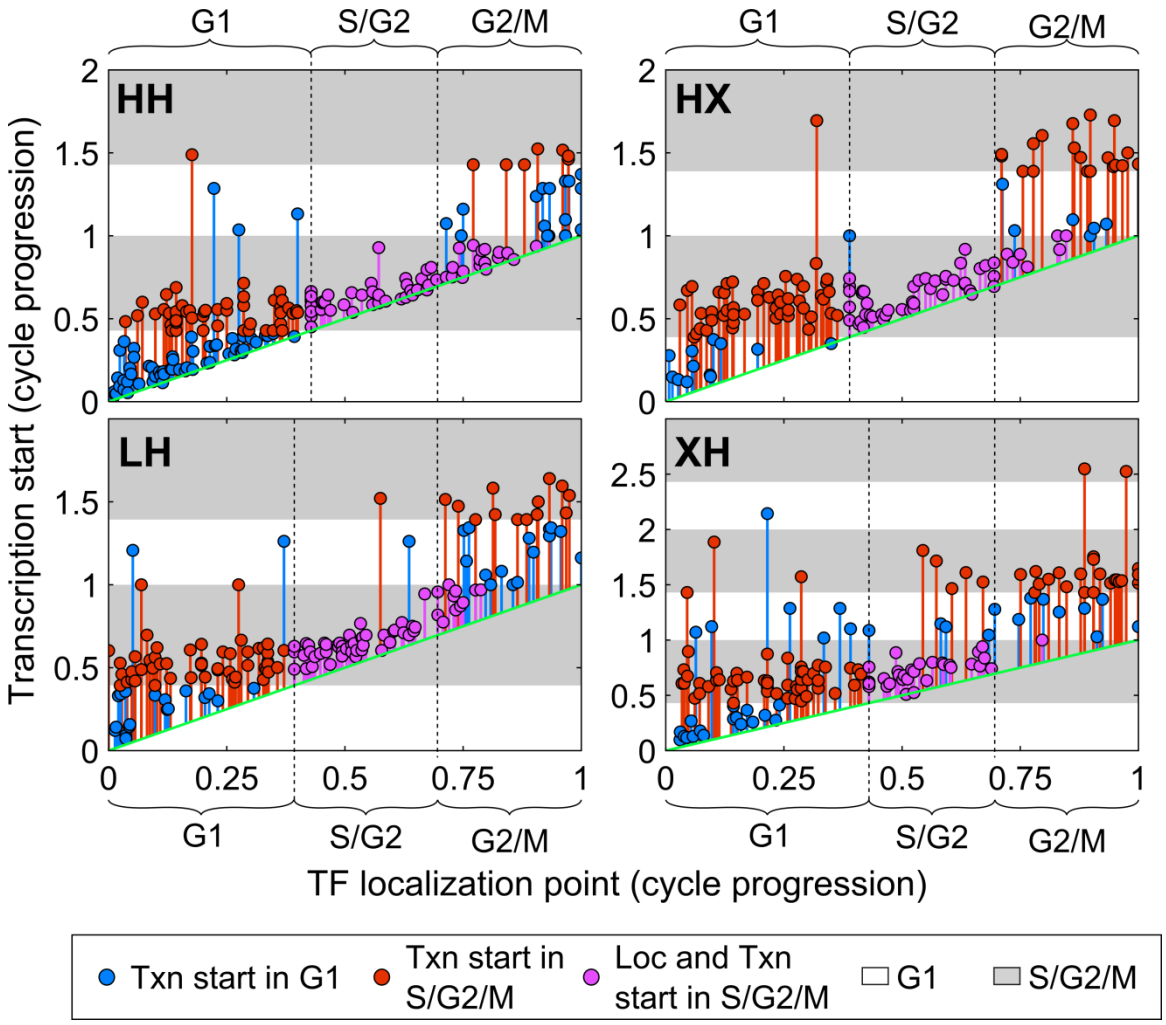


Figure 4.14 | Transcription activation is dependent on a permissive period in S/G2.
For the step tests in Figure 4.13, single-cell time of transcription start (circles) against the localization time as normalized cell cycle progression. The green diagonal represents the point of TF localization. Vertical distance from the green diagonal to the circle represents the promoter response delay time. Average cell-cycle G1 length is demarcated on the localization axis. A boundary is arbitrarily defined between S/G2 and G2/M at 0.7, since there appears to be a difference in the rate of transcription activation between early and late S/G2/M.

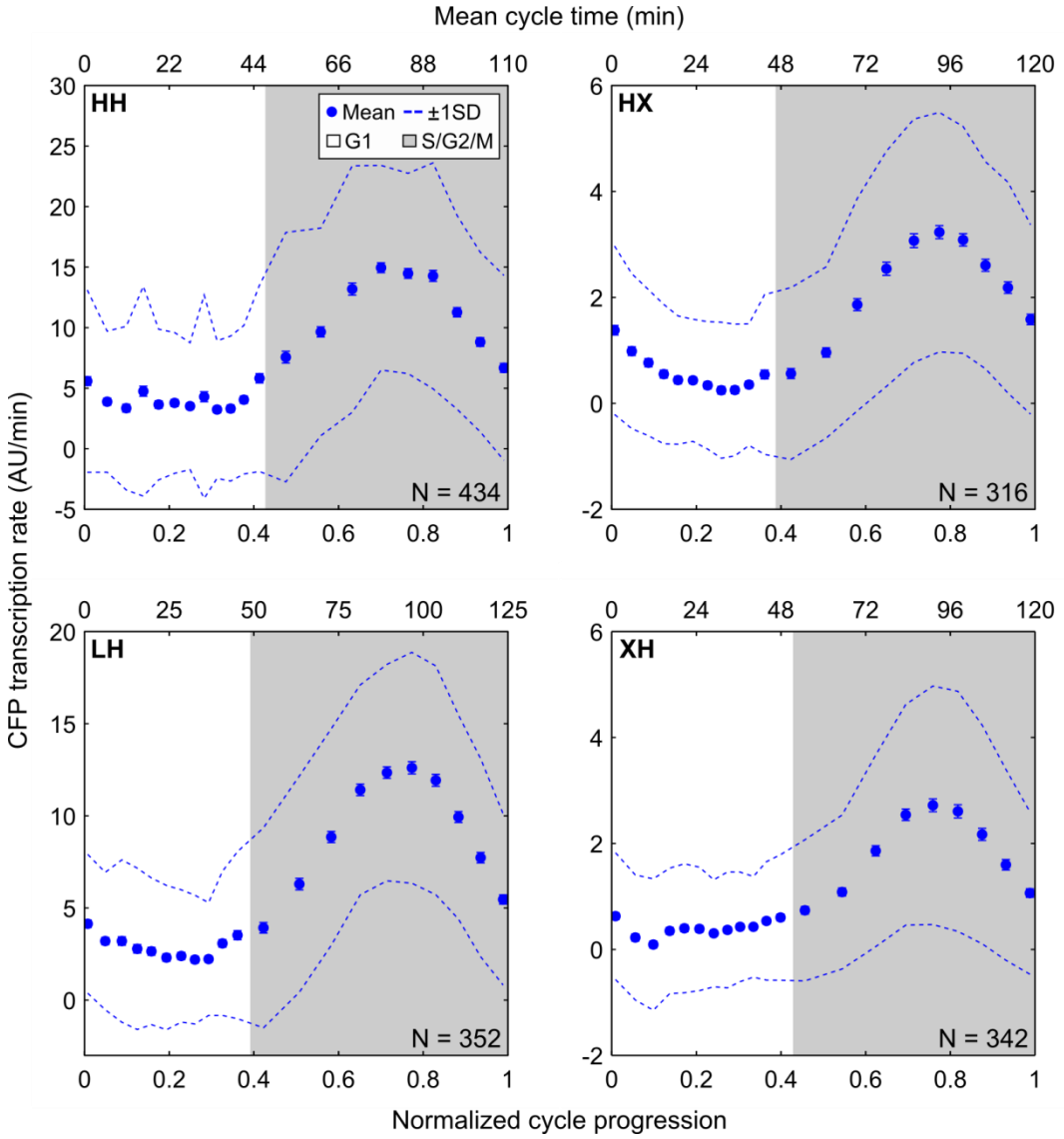


Figure 4.15 | Steady-state transcription varies across the cell cycle at *PHO5* promoter variants
CFP transcription rate data from each galactose step test at long times (>100 minutes in low phosphate, when at least 80% cells have activated transcription for XH and at least 90% for the other variants), was binned according to the corresponding cell cycle progress at each time point. The bootstrapped means (points) are plotted for each normalized cycle progression bin along with one standard deviation above and below (dashed lines). Error bars represent S.E.M. of the mean by bootstrapping, and "N" is the number of data points in each bin.

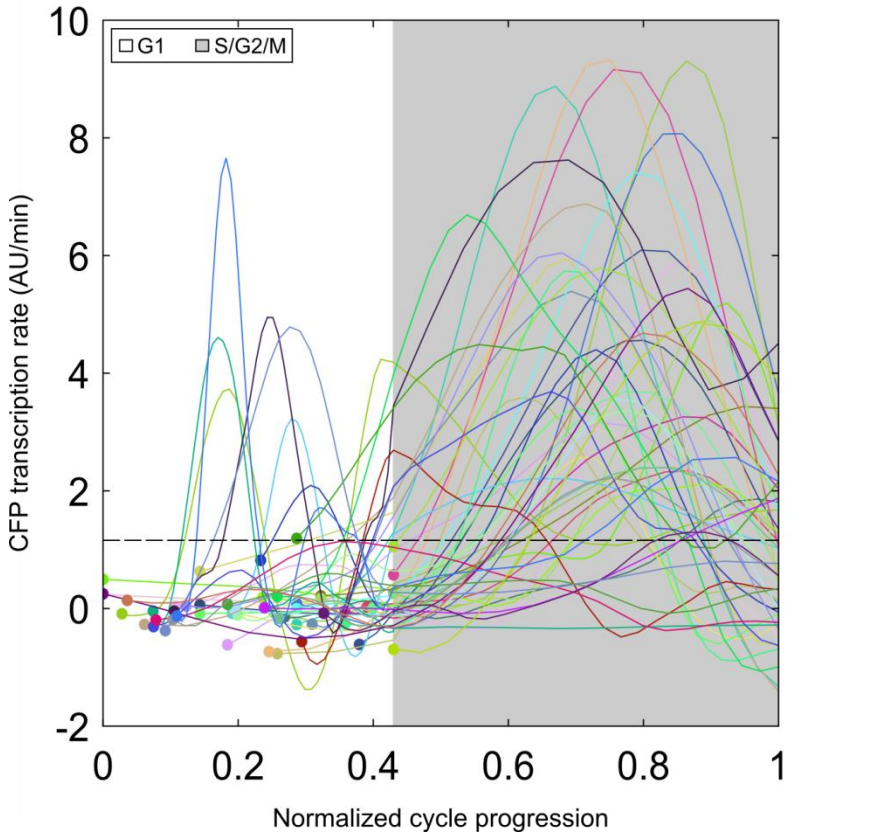


Figure 4.16 | Lack of *CFP* transcription activation in G1 is due to long times spent waiting until S/G2.
Transcription rate is plotted from the time of TF localization (points) for 50 cells in which localization occurred during G1. The plurality of cells that do not activate *CFP* transcription until S/G2/M do not transcribe at all during G1 rather than exhibiting detectable transcription that does not reach the threshold (dashed line).

4.5 Promoter architecture affects cell cycle dependent delay

Cell-cycle phase appears to be an event in *trans* that strongly influences promoter transitions in *cis* that are required for activation. We next asked whether and how promoter architecture affects the two distinct delays occurring in G1 versus early S/G2. Comparing the G1 delay with an aggregate S/G2/M delay between HH, LH, XH, and HX shows that the absence of a NFS site strongly affects the delay in S/G2/M (Figure 4.13). In addition, both LH and HX exhibit indistinguishable delay times even though they have significantly different steady-state expression levels (Figure 4.17). To incorporate the impact of localizing in late G2/M, we defined

and calculated probability of three events: activation in S/G2 given S/G2 localization, activation in G1 given G1 localization, and activation in G1 given localization without activation in the prior G2/M. The first event is a measure of the S/G2 delay whereas the second and third events are measures for the G1 delay. In Table 1, we list these probabilities for different promoter architectures and group them based on statistically significant differences (pairwise Fisher's exact test comparisons – see table legend). These data also reveal that the *absence* of the NFS extends the S/G2 delay (by decreasing the probability of activation in S/G2 given localization in S/G2). Moreover, the aggregate affinity of both sites is important to the G1 delay. HH and (to a lesser extent) LH variants are more likely to activate transcription in G1 compared to the HX and XH variants. While LH and HX have very similar delay times in Figure 4.13, the pairwise comparisons show LH is more likely to activate transcription in G1 following TF localization in G2/M. This could be due to the discrete cell-cycle phase categorization used in the statistical test compared to the continuous measurement of delay time. In Figure 4.14, about half of the LH cells that localize in G2/M and activate transcription in G1 activate towards the end of G1. The delay time for these cells is not much different than the HX cells that activate in early S/G2, but the classification is distinct. Further extending G1 by growing cells in media with raffinose as the carbon source should clarify whether, given enough time, LH can activate in G1 faster than HX. At the very least, we can conclude that both a NFS *and* a NOS are necessary for activation in G1 by comparing HH to both XH and HX, and we consider it likely that aggregate affinity dictates G1 delay.

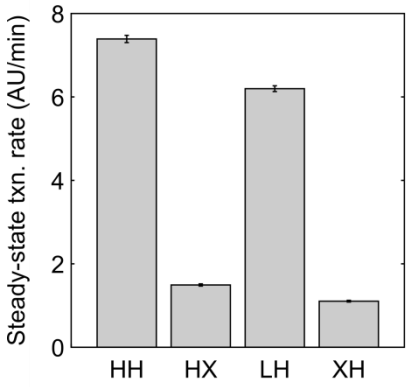


Figure 4.17 | Mean steady-state *CFP* transcription rate from *PHO5* promoter variants in galactose step tests.
Bars represent the bootstrapped mean of all data in each plot of Figure 4.15, and error bars represent the S.E.M.

Table 1 | Reliance on the S/G2 permissive period for activation is dependent on promoter architecture

Localization starts in:	G1	S/G2	G2/M
Probability transcription starts in:	G1 vs. Later	S/G2 vs. Later	Next G1 vs. Next S/G2/M
HH	0.52 +	0.78 +	0.68 +
LH	0.25 –	0.76 +	0.54 +
HX	0.20 –	0.64 +	0.24 –
XH	0.22 –	0.33 –	0.26 –

Promoters in the "+" group are significantly more likely to activate transcription during the first indicated period compared to promoters in the "–" group in each column (pairwise Fisher's exact test, $p < 0.05$).

Our data predict that being in or transitioning to early S/G2, an event in *trans*, can accelerate promoter transitions in *cis* to shorten the activation delay. Moreover, given a NFS site, lower aggregate affinity of Pho4p binding sites should extend the G1 delay. We tested this prediction in a different way, by expressing two copies of the LH, LL, or LX variants in diploid cells in glucose. A short activation delay in S/G2 would lead activation of both copies of these promoter variants to appear strongly correlated. However, the strength of the correlation would decrease

with aggregate affinity if the slower G1 delay times are due to slow promoter transitions in *cis*. These predictions were roughly borne out as two copies of LH and LL variants showed highly correlated ($\rho \sim 0.75$) activation whereas the LX variant correlation was less ($\rho \sim 0.5$) but still strong (Figure 4.18A). Indeed, as predicted when localization occurs in S/G2 the two promoter copies are almost perfectly correlated because activation is largely restricted to S/G2 (Figure 4.18B). Importantly, this trend in correlated activation was independent of the steady-state transcription rate. The XH variant has an equivalent steady-state transcription rate to LL and LX, but the absence of a NFS site leads to longer S/G2 delays and, as predicted, a lower correlation ($\rho \sim 0.4$) (Figure 4.18A & B).

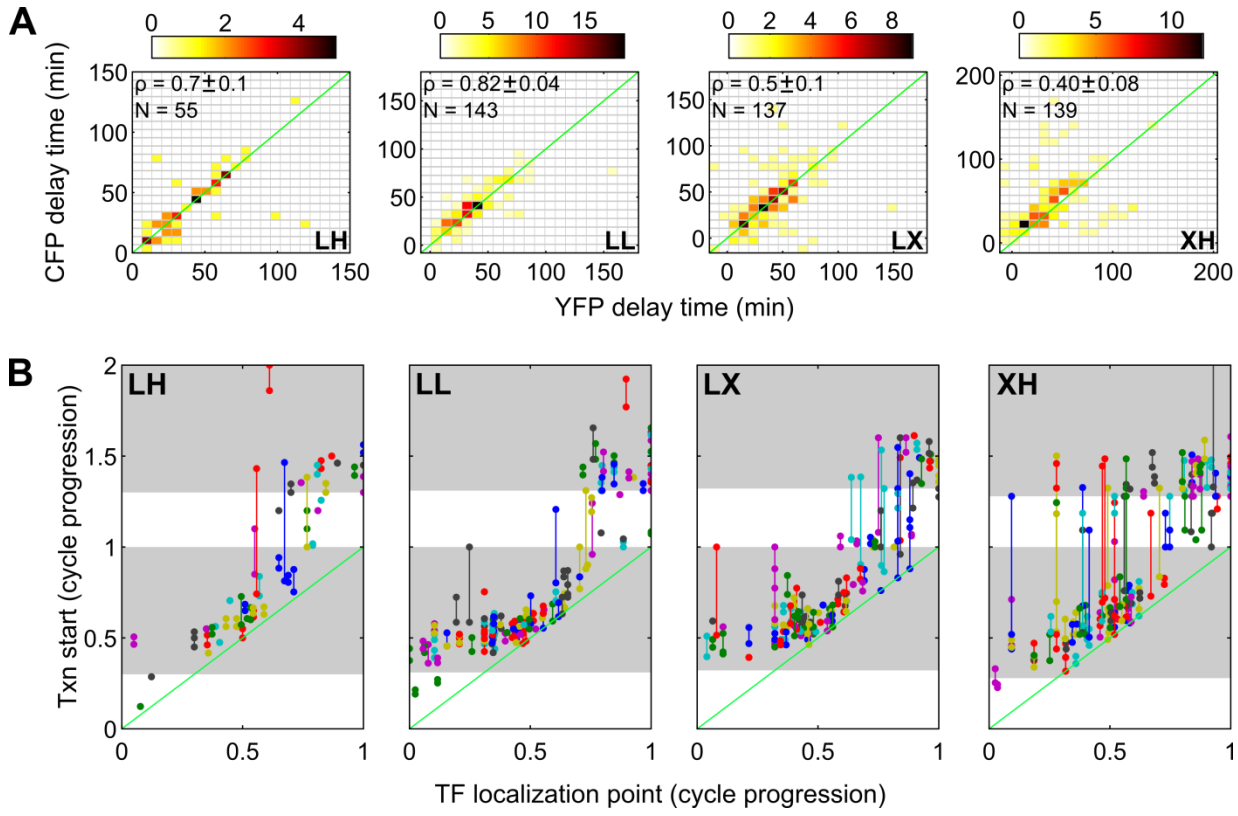


Figure 4.18| The cell cycle correlates transcription activation at two homologous promoter copies.

(*A*) Single-cell delay times for step tests of a diploid strain with two homologous copies of the indicated promoter driving expression of either *CFP* or *YFP*. (*B*) For the step tests in *A*, single-cell time of transcription start against the localization time as normalized cycle progression. For each cell (colors), the transcription start times for the two promoter copies (circles) are connected to show the difference in activation times. Vertical distance from the green diagonal to each point represents each promoter's response delay time.

4.6 Memory of previous activation events can eliminate cell-cycle dependent delays

A number of yeast promoters, including *INO1* and *GAL1*, have been shown to reactivate faster after having been recently transcribed as compared to a naïve reporter (Brickner et al. 2007; Kundu et al. 2007). In these cases, the transcriptional memory has lasted multiple cell cycles and corresponded to movement of the promoter to the nuclear periphery. We sought to determine if transcriptional memory exists at *PHO5* promoter variants and whether it accelerates reactivation in a cell cycle-dependent way. We chose to focus on *PHO5* (LH) and the HX variant that exhibit similar delay times. Cells with these variants were grown in galactose media and subject to a first step test, leaving cells in low phosphate for 200 minutes. This was followed by a period in high phosphate to delocalize the Pho4-YFP to the cytoplasm, and then a second step test to low phosphate (Figure 4.19A). The high phosphate period intervening between the two step tests had a duration of either 75 minutes or 150 minutes, shorter or longer than the average cell cycle length, respectively (Figure 4.20). Transcription quickly ceased after removal of Pho4p from the nucleus, consistent with rapid chromatin reassembly (Schermer et al. 2005; Adkins and Tyler 2006), and no transcriptional activity was observed during the high phosphate "off" period (Figure 4.19B). Upon reactivation, the mean response delay decreased in both cases (Figure 4.21).

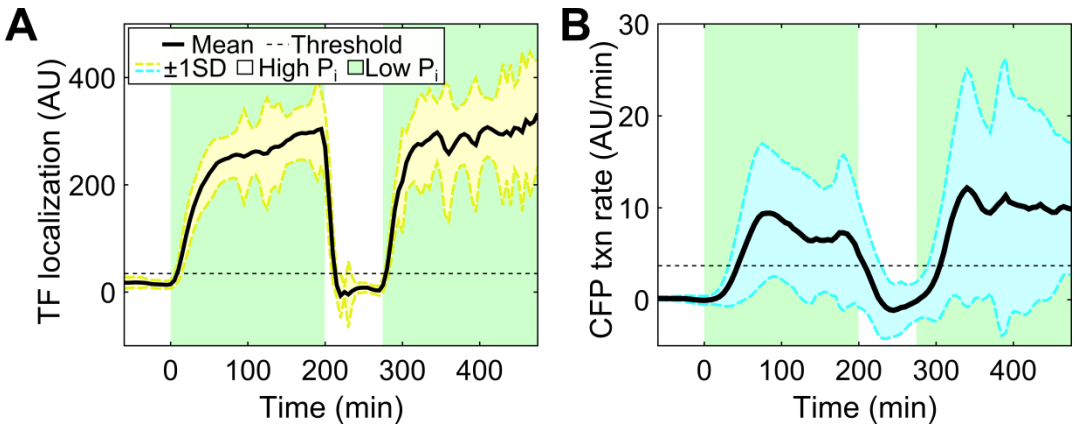


Figure 4.19 | Reactivation of the *PHO5* promoter reveals memory of previous transcription. (*A*) TF localization in response to a 75 min pulse of high phosphate ("off" pulse) between periods of low phosphate. Data shown for strain bearing the wild-type *PHO5* promoter (LH). (*B*) CFP transcription rate of LH in response to TF localization in *A* reactivates faster. (Symbols as in *A*.) (*C*) Mean promoter

activation delay times in all or cell-cycle categorized single-cells on the first and second step for the indicated promoters and “off” pulse lengths. Error bars represent S.E.M. by bootstrapping.

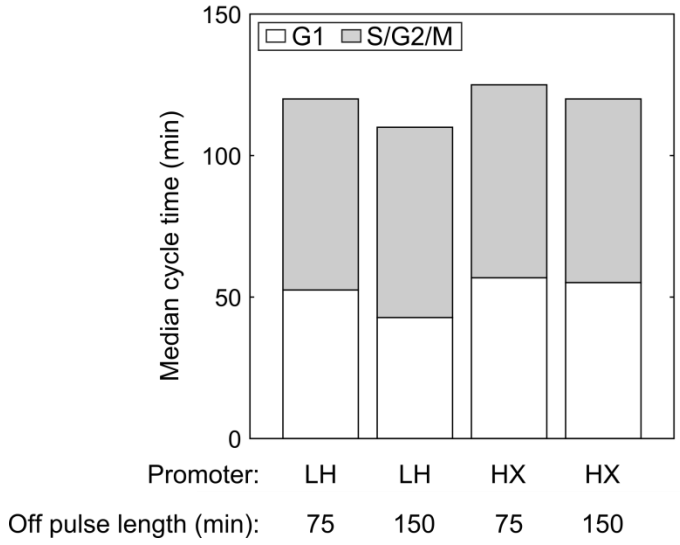


Figure 4.20 | Median cell cycle length for the full time course of the transcriptional memory reactivation step tests.
Cell cycle (full bar), G1phase (white portion), and S/G2/M phase (gray portion) lengths calculated as in Figure 4.12.

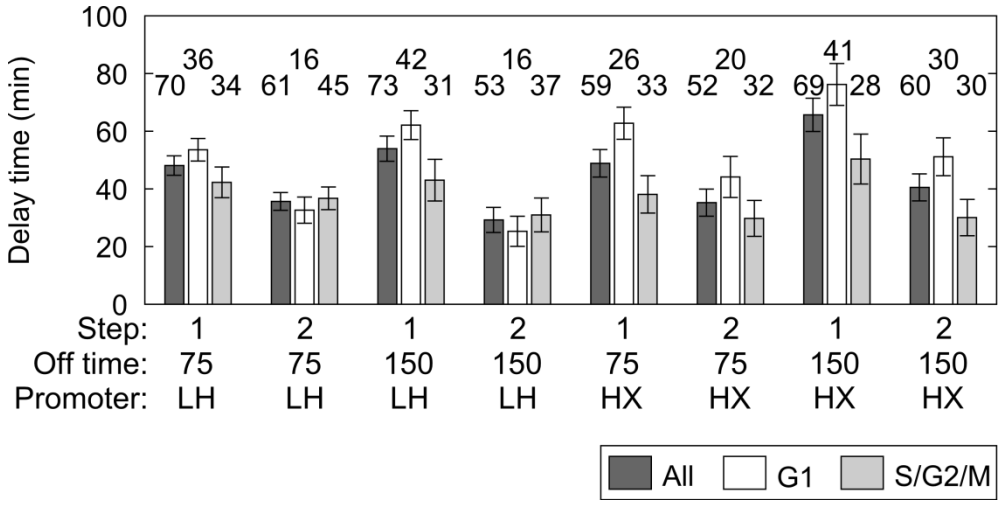


Figure 4.21 | Memory at the promoter decreases cell cycle-dependence.
Mean promoter activation delay times in all or cell-cycle categorized single-cells on the first and second step for the indicated promoters and “off” pulse lengths. Error bars represent S.E.M. by bootstrapping.

Since direct observation of the nuclear Pho4-YFP input to the promoter precludes pathway feedback as the source of memory, the accelerated kinetics reflects transcriptional memory at the *PHO5* promoter itself. Furthermore, the memory survives over the length of the cell cycle. Repeating the tests on the HX architecture, which had a similar response delay to LH (Figure 4.13) despite a stark difference in architectures, we again found a faster reactivation response (Figure 4.21). This suggests the memory mechanism is not strictly architecture-dependent.

Categorizing each cell as before based on the cell cycle phase at the time of TF localization reveals a link between the transcriptional memory and the cell-cycle dependence of gene activation delays. In each test, reactivation delays for G1 cells are shorter than for the initial response while the change in response delay for S/G2/M localized cells tends not to change significantly (Figure 4.21). For each promoter, the shorter G1 delay upon reactivation is due to the significantly lower likelihood of passing through G1 without activation if localization occurs in G1 or the previous G2/M period (Table 2, pairwise Fischer exact test, variable p-value). In contrast to the accelerated reactivation for cells during G1, those in which the TF localized in S/G2 do not reactivate transcription any faster on the second stimulation, which contributes to the lack of a significant change in delay time for S/G2/M cells in Figure 4.21. Thus, the memory creates an accelerated response by increasing the rate of gene activation in G1, thereby diminishing the reliance on the S/G2 phase. Interestingly, upon reactivation the G1 response delay time decreases for HX but disappears for LH (Figure 4.21), which intimates a role for architecture in exploiting the memory mechanism. This establishes a supporting role for the NOS in more permissive reactivation similar to the dependence in G1 on aggregate affinity when activating a naïve promoter (Table 1). Taken together, these results indicate a long-lived, post-transcriptional state of the *PHO5* promoter decreases the G1 delay, which practically eliminates the cell-cycle dependence. A similar memory test of the XH promoter, with a much longer S/G2 delay, may reveal significantly faster reactivation also occurs during S/G2, which would suggest the post-transcriptional promoter state accelerates activation kinetics in a more general, cell-cycle independent way.

Table 2 | Transcriptional memory increases the probability of transcription activation in G1.

Localization starts in:		G1		S/G2		G2/M	
Probability transcription starts in:		G1 vs. Later		S/G2 vs. Later		Next G1 vs. Next S/G2/M	
LH 75 min	Pulse 1	0.09		0.75		0.31	
	Pulse 2	0.43	+++	0.67	–	0.62	+
LH 150 min	Pulse1	0.19		0.65		0.50	
	Pulse 2	0.57	+++	0.74	–	0.67	–
HX 75 min	Pulse 1	0.08		0.56		0.20	
	Pulse 2	0.32	++	0.79	–	0.67	+++
HX 150 min	Pulse 1	0.15		0.65		0.11	
	Pulse 2	0.46	+++	0.71	–	0.75	+++

Pairwise Fisher's exact tests were performed between the first and second pulse of low phosphate in each memory test. Probabilities corresponding to a significant increase in the likelihood of activating transcription during the first indicated period in each column are denoted as (+++: $p \leq 0.05$; ++: $p < 0.1$; +: $p < 0.2$) while no significant change is labeled as "–".

4.7 Activation kinetics at synthetic promoters is also cell cycle-dependent

Finally, we asked if cell-cycle phase affects the kinetics of gene activation more generally by measuring the time to activate at synthetic 1xtetO and 7xtetO promoters (consisting of 1 and 7 tetO binding sites fused to the *CYC1* minimal promoter, respectively) in response to a step change in a TF input. Rather than inducing the promoters through the tet-Trans-Activator and its repressor doxycycline, of which concentration changes may affect the TF slowly and at variable times in single cells (Charvin et al. 2008), we developed a chimerical, Pho4p-based switchable transactivator to target the tetO promoters (Figure 4.22). Instead of integrating Pho4-YFP into the switchable transactivator strain background, we introduced a Pho4tetR-YFP at the *leu2* locus. The C-terminal Pho4p DNA binding domain is well-defined (Ogawa and Oshima 1990) and easily eliminated (McAndrew et al. 1998), so we replaced it with a tetR C-terminally fused to Citrine. YFP was fused to tetR using the peptide linker (GDGAGLIN) reported in (Sheff and

Thorn 2004), but modified for more common codon usage and to convert the PacI restriction site to a SpeI site by exchanging valine for isoleucine (GGT GA**T** GGT GCT **GGA CTA GTT** AA**T**, translated as GDGAGL**V**N). The modified linker was then used to fuse tetR to the *PHO4* shuttling domain (base pairs 1-600). Because Pho4p is known to dimerize with the Pho2p cofactor we also deleted the Pho2p binding domain in *PHO5* (base pairs 601-741) {{967 Hirst 1994;}}, further insulating the pathway from native phosphate regulation. Into the resulting strain (Y947, Appendix, Table A.1), we integrated either the 1xtetO or 7xtetO promoter driving *CFP* at the *URA3* locus using integrating plasmids (B579 or B720, Appendix, Table A.2).

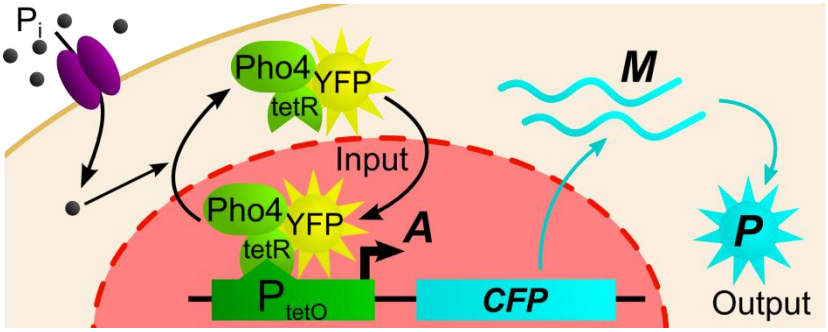


Figure 4.22 | A chimerical, switchable transactivator system to probe activation kinetics at tetO promoters.
In response to a step change in phosphate concentration, both TF localization and subsequent gene activation from the 1xtetO and 7xtetO promoter are identical (Figure 4.23A&B). The activation delay time distribution (Figure 4.23C), with a median 17 min delay, is likely dominated by fluorophore maturation (~10 min delay, section 2.4). Therefore, on average both promoters respond quickly to the TF. However, when we separate cells by the cell-cycle phase when TF localization occurs, the post-budding delay distribution is significantly different from the pre-budding delay distribution (Figure 4.23D, 2-sample K-S test: 1xtetO, $p = 0.05$; 7xtetO, $p < 0.001$), with median response times in post-budded cells 7 (1xtetO) or 10 (7xtetO) minutes earlier. We repeated the 7xtetO promoter step test in 2% raffinose to extend G1 and better sample G1 cells. Both the median 20 min activation delay and the 10 min gap between pre- and post-bud cells' delay (K-S test: $p = 0.005$) are similar to results in glucose, with prolonged delays in activation restricted to G1 (Figure 4.24). These results suggest the cell cycle may play a more general role in transcription activation.

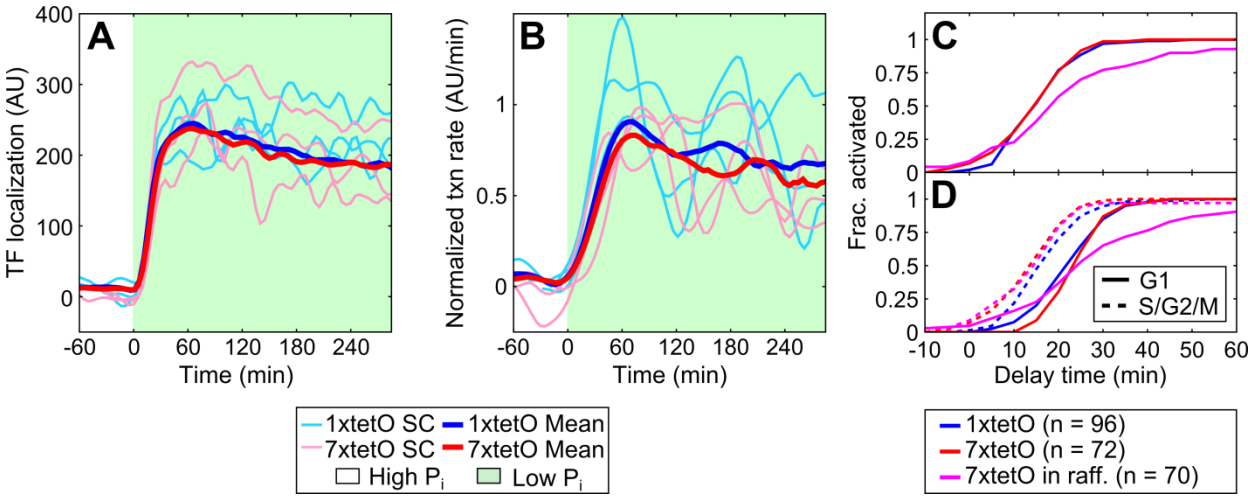


Figure 4.23 | Synthetic tetO promoter activation depends on cell-cycle phase.
(*A*) Localization of the chimerical TF in response to low phosphate was identical in strains carrying either 1xtetO- or 7xtetO-*CFP*. Thick lines represent the mean, and thin lines represent single cell traces (SC). (*B*) The resulting *CFP* transcription from each tetO promoter was normalized to the average first peak height for comparison. (*C*) The distribution of delay times is the same for both tetO promoters in glucose, but growth in raffinose slows the response. (*D*) Disaggregating the distributions by cell-cycle phase reveals faster activation post-budding.

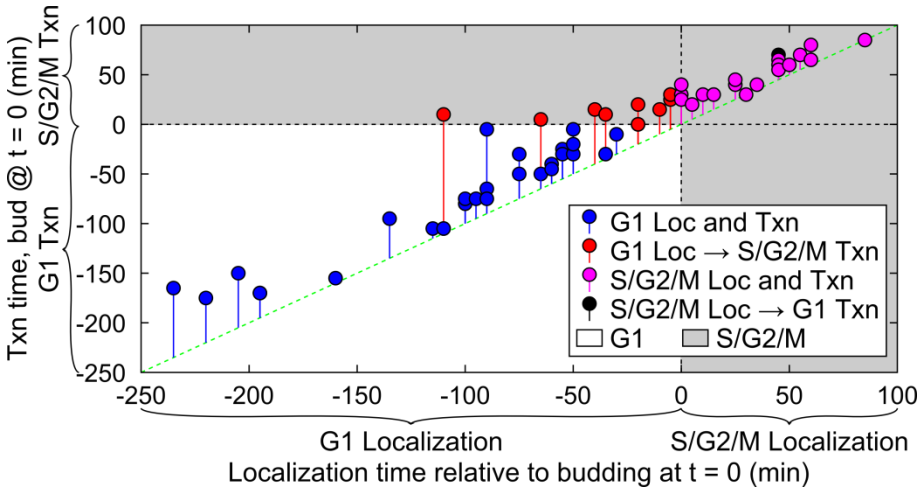


Figure 4.24 | Synthetic tetO promoter activation is slow in G1.
The time of localization and transcription start with respect to bud formation are plotted for single cells from the 7xtetO step test in raffinose. Vertical distance from the green diagonal represents the delay in transcription activation after TF localization.

4.8 Discussion

While a wealth of studies have probed the dynamics and mechanism of relieving chromatin repression at the *PHO5* promoter (reviewed in (Rando and Winston 2012)), the role of promoter architecture in gene activation kinetics remained unclear. To isolate the effect of transactivator binding site affinity and location from upstream signaling dynamics, we performed step tests on a library of previously described *PHO5* promoter variants (Lam et al. 2008) in a yeast strain background engineered for rapid and observable switches in Pho4p activity. We find that promoter response delay is strongly linked to the cell-cycle phase, with varying binding site requirements in different phases. The dominance of the cell cycle in *trans* leads to surprisingly high correlation in response timing between homologous copies of even the slowest, most variable architectures. Moreover, accelerated reactivation revealed transcriptional memory manifests as a decreased cell cycle-dependence. The prominent role of the cell cycle is an unexpected influence on the regulatory capacity of promoter chromatin.

The presence of chromatin at the promoter provides a tunable activation delay in response to a stimulus, but how architecture determines that delay is closely related to transitions through cell-cycle phases. The prohibitive effect of promoter chromatin is most prominent in the extended G1 during growth in galactose. Here, the NFS affinity correlated with likelihood of activating during G1, but in the absence of a NOS, the HX promoter could not efficiently activate (Table 1). This suggests a step-wise model for activation under repressive chromatin conditions in G1 consistent with previous work demonstrating the inability of Pho4p to bind the NOS initially at the repressed promoter without Pho4p transcriptional activity first recruited to the NFS (Venter et al. 1994; Svaren et al. 1994; Korber et al. 2006). NFS-bound Pho4p recruits the histone acetyltransferase activity of the SAGA complex (Barbaric et al. 2003) to acetylate the -2 nucleosome possibly destabilizing DNA-histone contacts and providing NOS access even without remodeling (Adkins et al. 2004). From there, nucleosome loss spreads from the NDR until activation (Jessen et al. 2006). Under this model, aggregate promoter binding site affinity and dispersed Pho4p binding provide an advantage in efficiently driving nucleosome eviction. However, activation during the permissive period in S/G2 occurs through a faster pathway unavailable in G1. Through some event in *trans* during S/G2, the NOS is no longer necessary for rapid activation, and Pho4p-NFS binding can efficiently drive nucleosome remodeling, which is

required to recruit the general transcription machinery (Adkins et al. 2007). We argue that although the promoter chromatin is not completely accessible in S/G2, the remodeling process is expedited by a looser nucleosome structure. This is evidenced by the accelerated S/G2/M response for the XH variant without a NFS, which can only be activated by Pho4p competitively binding the hidden NOS to drive remodeling. While our strain may be overexpressing Pho4p relative to the endogenous TF under no stress, the nuclear Pho4p concentration remains constant across the cell cycle (Figure 4.5A). Therefore, the strong probability of activating once S/G2 is reached can only be explained by a Pho4p-independent increase in the NOS accessibility. Our observation of an S/G2 permissive period for faster activation likely reflects alternating periods of promoter chromatin accessibility throughout the cell cycle. One possible *PHO5*-specific cause of cell cycle-dependent chromatin disruption is Mcm1p binding between the -1 and -2 nucleosomes. Though it is has been associated with mitotic induction of *PHO5*, its activity is shifted earlier in the cell cycle if phosphate storage is eliminated (Neef and Kladde 2003; Pondugula et al. 2009) as in our switchable transactivator strain. However, our observation of accelerated activation post-budding even at synthetic tetO promoters (Figure 4.23) argues in favor of a global effect of the cell cycle on transcription activation, which relegates Mcm1p binding to a supporting role in *PHO5* activation. The relative lengths of the cell-cycle phases then determine the relevance of architecture to the activation delay. In the case of exponential growth in glucose, G1 is much shorter than S/G2/M, which leads to frequent permissive periods. Thus, the population average follows the NFS-dependence of S/G2 permissive period activation. Slow growth through a long G1 phase decreases the frequency of permissive periods, and creates more reliance on the overall promoter architecture.

As suggested by the onset of the permissive period in S-phase, the increase in chromatin accessibility could be caused by DNA replication. Nucleosomes are quickly redeposited after replication fork passage, making opportunistic Pho4p binding on a bare promoter unlikely. However, the nucleosomes from the template DNA are decomposed and the constituent histones are distributed between the two copies in combination with histones in *trans* (Annunziato 2012; Alabert and Groth 2012). Reforming nucleosomes with cytosolic histones may reduce Set1-generated H3K4 methylation at the promoter which recruits the histone deacetylase (HDAC) activity of the Rpd3L complex (Wang et al. 2011) thereby disrupting the balance of NuA4 HAT and Rpd3L HDAC activities which poise the *PHO5* promoter (Nourani et al. 2004; Wang et al.

2011). Indeed, loss of Rdp3L activity at the *PHO5* promoter was recently shown to increase nucleosome turnover, basal transcription, and activation kinetics (Wang et al. 2011). While DNA replication is not required for *PHO5* promoter activation (Schmid et al. 1992), our results indicate it may serve to accelerate successful Pho4p-driven remodeling. DNA replication could also loosen chromatin structure through a non-promoter-specific pathway as we also observed a similar S/G2 permissive period in activation for synthetic tetO-based promoters (Figure 4.24).

The transcriptional memory elicits an accelerated response by making transcription activation generally more permissible in G1. Though a previous study found little change in reactivation kinetics at the *PHO5* promoter (Zhou and Zhou 2011), our method controls for delays in the phosphate signaling pathway and measures kinetics at the single-promoter level and at higher resolution. There are several possible mechanisms by which epigenetic memory at the promoter could lead to greater chromatin accessibility across the cell cycle. It is possible that removal of the Pho4p stimulus to deactivate a transcribing *PHO5* promoter may lead to a period of greater lability while repressive chromatin marks are being reestablished. However, nucleosomes are rapidly redeposited from a separate histone source (Schermer et al. 2005) making preservation of accessibility-enhancing histone modifications unlikely, and Set1-dependent repression is reset in less than one hour (Zhou and Zhou 2011). Attractive, histone modification-independent alternatives include maintained localization at the nuclear periphery as observed at the promoters of *INO1* and *GAL1* over several generations (Brickner et al. 2007). Another possibility is prolonged, inherited activity of the SWI/SNF chromatin remodeler perpetuating weaker nucleosome positioning as at the *GAL1* promoter (Kundu et al. 2007) since SWI/SNF is also recruited to induce the *PHO5* promoter (Dhasarathy and Kladde 2005). Neither of these pathways would depend directly on the affinity of the NFS or NOS but instead require only that the complement of required cofactors have been successfully recruited. Whereas the reactivated LH promoter is completely cell-cycle independent, the remaining discrepancy between G1 and S/G2/M activation for HX demonstrates the memory mechanism must somehow account for promoter architecture. We therefore favor a model which includes a long-lived labile chromatin state at the promoter.

In the context of the complete phosphate signaling pathway, the interplay between promoter cell-cycle- and architecture-dependence may accentuate the delay for slower variants. Nutrient

depletion tends to arrest yeast cells in G1 (Figure 4.12 and Figure 3.5) and therefore extends the prohibitive period. This may provide additional time for the conservative response genes, which have high affinity NFS, to alleviate the stress by mobilizing intracellular stores (*PHO8*, *PHM2*, *PHM4*) or increasing the phosphate affinity of cell-surface transporters (*PHO84*) before resorting to secretion of valuable materials to scavenge phosphate from the extracellular environment (*PHO5*) (Lam et al. 2008; Springer et al. 2003). The cell cycle thus adds complexity to a coordinated stress response. Additionally, transcriptional memory at promoters would enable a rapid switch to the radical emergency response expression program in the face of repeated stress. Since regulable promoters in yeast typically contain a nucleosome-free region enriched for transcription factor binding sites (Tirosh and Barkai 2008) and this canonical nucleosome pattern is preserved in metazoans (Barski et al. 2007; Tirosh and Barkai 2008), an activation time-scale separation strategy linked to the cell-cycle may be a general scheme to selectively activate transcription at subsets of promoters in stress-response networks.

4.9 References

Adkins M, Howar S, Tyler J. 2004. Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Molecular Cell* **14**: 657–723.

Adkins M, Tyler J. 2006. Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Molecular Cell* **21**: 405–421.

Adkins MW, Williams SK, Linger J, Tyler JK. 2007. Chromatin Disassembly from the PHO5 Promoter Is Essential for the Recruitment of the General Transcription Machinery and Coactivators. *Molecular and Cellular Biology* **27**: 6372–6382.

Alabert C, Groth A. 2012. Chromatin replication and epigenome maintenance. *Nature reviews Molecular Cell Biology* **13**: 153–67.

Almer A, Rudolph H, Hinnen A, Hörz W. 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. *The EMBO Journal* **5**: 2689–2785.

Annunziato AT. 2012. Assembling chromatin: The long and winding road. *Biochimica et Biophysica Acta* **1819**: 196–210.

Barbaric S, Luckenbach T, Schmid A, Blaschke D, Hörz W, Korber P. 2007. Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter in vivo. *The Journal of Biological Chemistry* **282**: 27610–21.

Barbaric S, Reinke H, Horz W. 2003. Multiple Mechanistically Distinct Functions of SAGA at the PHO5 Promoter. *Molecular and Cellular Biology* **23**: 3468–3476.

Barbaric S, Walker J, Schmid A, Svejstrup J, Hörz W. 2001. Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. *The EMBO Journal* **20**: 4944–4995.

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–37.

Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee P-C, Widom J, Brickner JH. 2007. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. ed. T. Mistelli. *PLoS Biology* **5**: e81.

Bun-Ya M, Nishimura M, Harashima S, Oshima Y. 1991. The PHO84 gene of Saccharomyces cerevisiae encodes an inorganic phosphate transporter. *Molecular and Cellular Biology* **11**: 3229–3238.

Charvin G, Cross FR, Siggia ED. 2008. A microfluidic device for temporally controlled gene expression and long-term fluorescent imaging in unperturbed dividing yeast cells. *PloS One* **3**: e1468.

Creasy C, Shao D, Begman L. 1996. Negative transcriptional regulation of PH081 expression in Saccharomyces cerevisiae. *Gene* **168**: 23–32.

Dhasarathy A, Kladde MP. 2005. Promoter Occupancy Is a Major Determinant of Chromatin Remodeling Enzyme Requirements. *Molecular and Cellular Biology* **25**: 2698–2707.

Fascher K, Schmitz J, Hörz W. 1990. Role of trans-activating proteins in the generation of active chromatin at the PHO5 promoter in S. cerevisiae. *The EMBO Journal* **9**: 2523–2531.

Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E. 2008. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. ed. U. Ohler. *PLoS Computational Biology* **4**: e1000216.

Griesbeck O, Baird GS, Campbell RE, Zacharias DA, Tsien RY. 2001. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *The Journal of Biological Chemistry* **276**: 29188–94.

Jessen W, Hoose S, Kilgore J, Kladde M. 2006. Active PHO5 chromatin encompasses variable numbers of nucleosomes at individual promoters. *Nature Structural & Molecular Biology* **13**: 256–319.

Kaffman A, Herskowitz I, Tjian R, O'Shea EK. 1994. Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science* **263**: 1153–1156.

Kaffman A, Rank NM, O'Shea EK. 1998. Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes & Development* **12**: 2673-2683.

Komeili A, O'Shea EK. 1999. Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science* **284**: 977–980.

Korber P, Barbaric S, Luckenbach T, Schmid A, Schermer UJ, Blaschke D, Hörz W. 2006. The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *The Journal of Biological Chemistry* **281**: 5539–45.

Korber P, Luckenbach T, Blaschke D, Hörz W. 2004. Evidence for histone eviction in trans upon induction of the yeast PHO5 promoter. *Molecular and Cellular Biology* **24**: 10965–74.

Kundu S, Horn PJ, Peterson CL. 2007. SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. *Genes & Development* **21**: 997–1004.

Lam FH, Steger DJ, O'Shea EK. 2008. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**: 246–296.

Lohr D. 1997. Nucleosome Transactions on the Promoters of the Yeast GAL and PHO Genes. *Journal of Biological Chemistry* **272**: 26795–26798.

Mao C, Brown CR, Falkovskaia E, Dong S, Hrabeta-Robinson E, Wenger L, Boeger H. 2010. Quantitative analysis of the transcription control mechanism. *Molecular Systems Biology* **6**: 431.

McAndrew PC, Svaren J, Martin SR, Horz W, Goding CR. 1998. Requirements for chromatin modulation and transcription activation by the Pho4 acidic activation domain. *Molecular and Cellular Biology* **18**: 5818–5827.

Neef DW, Kladde MP. 2003. Polyphosphate Loss Promotes SNF/SWI- and Gcn5-Dependent Mitotic Induction of PHO5. *Molecular and Cellular Biology* **23**: 3788–3797.

Neely KE, Hassan AH, Brown CE, Howe L, Workman JL. 2002. Transcription Activator Interactions with Multiple SWI/SNF Subunits. *Molecular and Cellular Biology* **22**: 1615–1625.

Nourani A, Utley R, Allard S, Côté J. 2004. Recruitment of the NuA4 complex poises the PHO5 promoter for chromatin remodeling and activation. *The EMBO Journal* **23**: 2597–3204.

Ogawa N, Oshima Y. 1990. Functional domains of a positive regulatory protein, PHO4, for transcriptional control of the phosphatase regulon in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **10**: 2224–2236.

O'Neill EM, Kaffman a, Jolly ER, O'Shea EK. 1996. Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* **271**: 209–12.

Pondugula S, Neef DW, Voth WP, Darst RP, Dhasarathy A, Reynolds MM, Takahata S, Stillman DJ, Kladde MP. 2009. Coupling Phosphate Homeostasis to Cell Cycle-Specific Transcription: Mitotic Activation of Saccharomyces cerevisiae PHO5 by Mcm1 and Forkhead Proteins. *Molecular and Cellular Biology* **29**: 4891–4905.

Rando OJ, Winston F. 2012. Chromatin and transcription in yeast. *Genetics* **190**: 351–87.

Reinke H, Hörz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Molecular Cell* **11**: 1599–2206.

Rudolph H. 1987. The Yeast PHO5 Promoter: Phosphate-Control Elements and Sequences Mediating mRNA Start-Site Selection. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 1340–1344.

Schermer UJ, Korber P, Hörz W. 2005. Histones Are Incorporated in trans during Reassembly of the Yeast PHO5 Promoter. *Molecular Cell* **19**: 279–285.

Schmid A, Fascher K-D, Hörz W. 1992. Nucleosome disruption at the yeast PHO5 promoter upon PHO5 induction occurs in the absence of DNA replication. *Cell* **71**: 853–864.

Sheff MA, Thorn KS. 2004. Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. *Yeast* **21**: 661–670.

Shirahama K, Yazaki Y, Sakano K, Wada Y, Ohsumi Y. 1996. Vacuolar function in the phosphate homeostasis of the yeast Saccharomyces cerevisiae. *Plant & Cell Physiology* **37**: 1090–3.

Springer M, Wykoff DD, Miller N, O'Shea EK. 2003. Partially Phosphorylated Pho4 Activates Transcription of a Subset of Phosphate-Responsive Genes. *PLoS Biology* **1**: e28.

Steger D, Haswell E, Miller A, Wente S, O'Shea EK. 2003. Regulation of chromatin remodeling by inositol polyphosphates. *Science* **299**: 114-166.

Svaren J. 1997. Transcription factors vs nucleosomes: regulation of the PH05 promoter in yeast. *Trends in Biochemical Sciences* **22**: 93–97.

Svaren J, Schmitz J, Horz W. 1994. The transactivation domain of Pho4 is required for nucleosome disruption at the PHO5 promoter. *The EMBO Journal* **13**: 4856–4862.

Thomas MR, O'Shea EK. 2005. An intracellular phosphate buffer filters transient fluctuations in extracellular phosphate levels. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 9565–9570.

Tirosh I, Barkai N. 2008. Two strategies for gene regulation by promoter nucleosomes. *Genome Research* **18**: 1084–1175.

Venter U, Svaren J, Schmitz J, Schmid A, Hörz W. 1994. A nucleosome precludes binding of the transcription factor Pho4 in vivo to a critical target site in the PHO5 promoter. *The EMBO Journal* **13**: 4848–4903.

Vogel K, Horz W, Hinnen A. 1989. The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Molecular and Cellular Biology* **9**: 2050–2057.

Wang S-S, Zhou B, Zhou J-Q. 2011. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Molecular and Cellular Biology* **31**: 3171–3252.

Williams SK, Truong D, Tyler JK. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9000–5.

Wykoff D, O'Shea E. 2001. Phosphate transport and sensing in Saccharomyces cerevisiae. *Genetics* **159**: 1491–1500.

Wykoff DD, Rizvi AH, Raser JM, Margolin B, O'Shea EK. 2007. Positive feedback regulates switching of phosphate transporters in S. cerevisiae. *Molecular Cell* **27**: 1005–1013.

Yoshida K, Ogawa N, Oshima Y. 1989. Function of the PHO regulatory genes for repressible acid phosphatase synthesis in Saccharomyces cerevisiae. *Molecular & General Genetics* **217**: 40–46.

Zhou BO, Zhou J-Q. 2011. Recent transcription-induced histone H3 lysine 4 (H3K4) methylation inhibits gene reactivation. *The Journal of Biological Chemistry* **286**: 34770–6.

Zopf CJ, Quinn K, Zeidman J, Maheshri N. 2013a. Cell-cycle dependence of stochastic gene expression. *PLoS Computational Biology (submitted)*.

Zopf CJ, Wren N, Maheshri N. 2013b. A combination of promoter architecture and cell cycle phase influences delays in gene activation. *Nucleic Acids Research (submitted)*.

CHAPTER 5. FUTURE DIRECTIONS

The single-cell real-time analysis techniques developed in this thesis have led to a new view of the cell cycle as a general regulator of transcription activation, rather than being restricted to influencing classically identified cell cycle-regulated genes. Yet a full picture is by no means complete, and many questions remain regarding the universality, mechanism, and phenotypic impact of this phenomenon. Given the dominance of the cell cycle in transcriptional variability, how does this change predictions of stochastic models of gene regulatory networks? What is the molecular mechanism through which the cell cycle exerts its influence? Given the timing of the permissive period onset in S-phase, does DNA replication directly enhance transcription activation? What are the required factors that mediate the permissive period? The same factors likely also play a role in the S/G2 permissive period observed during the kinetic stress response at *PHO5* promoters. However, promoter variants do not equally depend on or utilize the permissive period to activate transcription. This begs the question, what are the molecular activation pathways available to some *PHO5* promoter architectures but not others, and how do they depend on the cell cycle? Promoter reactivation may take place through pathways newly accessible by prior activation. What is the memory mechanism? How general is it, or is it entirely gene-specific? These results would also come to bear on synthetic biology applications. In addition to providing foundational design rules for eukaryotic promoters with desired kinetics, this unappreciated role of the cell cycle should play an important role in modeling and implementing robust networks. In the remaining sections, we suggest potential research strategies to address the above questions.

5.1 Towards improved stochastic gene expression models

We used real-time inferences of transcription rate in single yeast cells to investigate the effect of the cell cycle on instantaneous expression. By comparing the transcription rate to cell cycle progression under steady growth conditions in Chapter 3, we were able to show that the cell cycle is a surprisingly dominant source of variability in transcription. Extrinsic factors are known to play a significant role in expression variation both between cells and within the same cell over

time (Elowitz et al. 2002; Raser and O'Shea 2004; Volfson et al. 2006; Pedraza and van Oudenaarden 2005; Colman-Lerner et al. 2005), but the restriction of transcription to the period post-budding highlights the cell cycle as a driving force in noisy gene expression. The combination of no expression in G1 and the S-phase permissive period leads to a large (>> 2) fold-change in transcription when expression pulses do occur, which contributes significantly to the variability across an asynchronously dividing population as well as correlates expression between reporters in single cells. Thus, this finding calls into question the inference of steady-state dynamics using the standard stochastic model of gene expression (To and Maheshri 2010; Raj et al. 2006; Mao et al. 2010; Munsky et al. 2012; Tan and van Oudenaarden 2010), which attributes all population variability to the intrinsic stochasticity of transcription activation at a promoter.

Explicitly including the effects of the cell cycle on transcription rate will lead to more accurate stochastic models of gene expression. Future work should investigate under what circumstances predictions are altered by inclusion of cell cycle-dependent effects. At the very least, gene dosage creates a baseline 2-fold change in the *observed* transcription rate at an active reporter after replication in any dividing system. Moreover, in systems where progression through the cell cycle may influence the *actual activation probability* of a reporter as we observed at low levels of expression from the noisy 7xtetO synthetic promoter (Figure 3.9), subsets of the population in different cell-cycle phases follow entirely different expression dynamics. Indeed, for the single-mRNA measurements of noisy and constitutive expression in Chapter 3 (Figure 3.10), a simple model where transcription switched between low and high rates in G1 and S/G2/M, respectively, could reproduce the main features of observed mRNA distributions across the population (Zopf et al.). This model could also explain the shift of the mRNA distribution through the cell cycle (Figure 3.10) whereas the standard model fails to capture the periodic generation of the long tail. Thus, considering the cell cycle is crucial to meaningful interpretation of noisy gene expression in dividing cells. This is not to say all transcriptional bursts are cell cycle-linked nor that there is no intrinsic stochasticity at the promoter; clear bursts of expression appear in quiescent cells (Suter et al. 2011) and not every S-phase we observed lead to a transcription pulse (Figure 3.7). However, models featuring cell-cycle phase-specific expression regimes would account for a large extrinsic source of variability that is mostly ignored at present. Theoretical investigations should lead the way in predicting the effect on regulatory networks. In particular, we anticipate

the effect of the cell cycle to be important in network topologies where noisy promoters experience periods of low expression. One example is a positive feedback loop, which has been associated with bimodal and bistable expression in populations (To and Maheshri 2010). Restriction of transcription to S/G2 should alter the stability of the low expression state.

Adaptation of current experimental techniques will provide support for or counterevidence against theoretical predictions. In this thesis we used bud morphology to assign cells to the G1 or S/G2/M phases of the cell cycle, but more accurate phase assignments are possible. More accurate phase assignments could be coupled to FISH through co-observing nuclear localization of a Whi5p-fluorescent protein fusion, which delocalizes to the cytoplasm between the START checkpoint in G1 until just before mitosis, and by marking the bud-neck with a Cdc10p-fluorescent protein fusion as in (Charvin et al. 2008). Cell-cycle phase in other organisms could be similar monitored with phase-specific fluorescent reporter fusions. Likewise, incorporating cell-cycle phase data into other real-time methods of observing transcription (Taniguchi et al. 2010; Golding et al. 2005; Maiuri et al. 2011; Chubb et al. 2006; Muramoto et al. 2012; Larson et al. 2011; Choi et al. 2008; Suter et al. 2011) will facilitate more mechanistically relevant models. These efforts would be invaluable to determining the contribution of the S-phase permissive period to transcription variability in various systems.

5.2 Investigating the permissive transcription period in S/G2

Our working hypothesis is that the process of DNA replication leads to the observed permissive period for activation during early S/G2. Passage of the replication fork leads to expulsion and reassembly of nucleosomes along the DNA (Alabert and Groth 2012). This may provide a window of opportunity for transactivator access which coupled with some level of transcriptional feedback would maintain an active transcriptional state throughout S/G2/M. To test the importance of replication we could observe the transcription rate of a non-replicating reporter. To achieve such a reporter, a 7xtetO synthetic promoter-*YFP* gene fusion could be expressed from a centromeric plasmid in which the replication origin is flanked by parallel loxP sequences. Expression of an inducible *CRE* recombinase during microfluidic culture would then recombine the loxP sequences (Sternberg and Hamilton 1981), looping out the origin, to generate a non-replicating reporter plasmid that is stable over multiple generations. The reporter's tet-

transactivator (tTA)-driven transcription can then be observed in real-time in the noisy expression regime. Cells in which the plasmid maintains the capacity to replicate can be discarded during analysis *in silico* if expression is observed in both lineages following recombinase expression. If no transcriptional bursts occur, or if they are infrequent and uncorrelated to cycle phase, replication directly generates the permissive period. If S-phase is still permissive for transcription activation even in the absence of replication, the increase in activation probability is due instead to an S/G2 phase-specific *trans*-factor or -event.

Whether or not replication directly causes the increased activation probability, the above experiment would not pinpoint the factors involved. Global changes to chromatin during S-phase are a possible *trans*-event contributing to the permissive period even in the absence of replication. As a model system for chromatin study, many chromatin-modifying or –remodeling complexes and histone chaperones in yeast are conserved or have homologues in metazoans. Therefore, factors found to be important to a general S/G2 permissive period in yeast may have a role in the replication-linked window of opportunity discussed in section 1.3, which is crucial to biologically-relevant processes such as development and tumorigenesis. Identified factors could offer potential drug targets to slow cancer progression or prevent metastases.

We next propose a strategy to identify proteins and complexes required to generate the S-phase permissive period. First, candidate proteins likely to have a role will be selected. Promising targets include those associated with the post-replication window of chromatin maturation. The chaperones Asf1 (Groth et al. 2007), CAF-1 (Kaufman 1996), and Rtt106 (Fazly et al. 2012) are responsible for deposition of H3/H4 tetramers after fork passage; deletion mutants targeting these may slow chromatin maturation and extend the permissive S/G2 window. The process of H2A/B deposition after replication is less clear, but may involve Nap1, a chaperone of H2A/B histones (Ito et al. 1996), which may have a similar effect upon deletion. Additionally, post-translational modifications of histones associated with the chromatin maturation process can be targeted by deleting the histone acetyltransferase Rtt109 (H3K56ac, (Recht et al. 2006; Williams et al. 2008)), or by using histone mutants (H4K5R, H4K12R (Megee et al. 1990)).

Next, in a strain with constitutive expression of tTA, a single candidate protein will be deleted to create a library of deletion mutants. Then each mutant will be "color-coded". A set of plasmids

where each bears a combination of 1, 2, or 3 fluorescent reporters (*RFP*, *YFP*, and *CFP*) can be transformed into the deletion mutant strains to provide 7 unique color-profiles (7 combinations of 1 or more colors). On each of these plasmids, one fluorescent reporter will be driven by the 7xtetO promoter to allow observation of noisy transcription pulses while any remaining reporters will be driven by constitutive promoters. This creates 7 sub-libraries in the deletion mutant library. We can then co-culture a strain from each sub-library in the same microfluidic chamber to observe expression from the 7xtetO reporter in real-time and use the color code to distinguish between deletions. This strategy would allow rapid time-lapse screening of the role of 84 candidate proteins with only 12 microfluidic chambers (3 experiment-days with 4 channels per microfluidic plate). This strategy could be extended to study potentially varying roles of factors in noisy expression from subtelomeric and heterochromatic regions as well, both of which are strongly repressed by chromatin. Combining the library screen with the test of replication-dependence outlined above, a mechanistic understanding of the S-phase permissive period for noisy gene expression in yeast, and possibly higher eukaryotes, will be realized soon.

5.3 Promoter architecture-dependent activation kinetics: from mechanism to design

We also employed a novel transactivator switch to probe the kinetics of transcription activation as a function of promoter architecture. While activation kinetics are certainly influenced by promoter architecture, cell-cycle progression analysis in Chapter 4 again revealed a strong dependence on an S/G2 permissive period to achieve active transcription in response to a step change in transactivator. By eliminating delays in upstream signaling, we expected to abate the extrinsic processes which dominate *PHO5* expression (Raser and O'Shea 2004) and the mitotic activation of *PHO5* caused by phase-specific fluctuations in intracellular phosphate levels (Neef and Kladde 2003; Pondugula et al. 2009). Thus, we anticipated variability in nucleosome remodeling at the promoter would dictate variability in the activation delay. Instead, the time of transcription factor localization to the nucleus relative to the onset of S/G2 dominated the delay due to slow kinetics in G1. The importance of this cell-cycle effect, though, is tuned to different degrees based on promoter architecture. While correlated by the cell cycle in *trans*, the response times from two homologous promoters in a single cell varied dependent on architecture. Taken together, these results point to an architecture-dependent ability to activate through multiple

pathways (Barbaric et al. 2007) which apparently have varying degrees of dependence on the cell cycle.

An important question to answer next is what are the molecular pathways important for efficient activation from the different promoter architectures? It has been shown that architecture and transactivator occupancy can determine differential cofactor requirements (Hertel et al. 2005; Dhasarathy and Kladde 2005), but how these relate to the cell cycle is also an open question. DNA replication could again play a direct role in enhancing promoter accessibility. With some additional engineering, the switchable transactivator strain background could be adapted for the non-replicating reporter strategy described above to test this hypothesis. Another likely contributor is binding of the late cell-cycle factor Mcm1p. Though Mcm1p binds the *PHO5* promoter and facilitates mitotic activation in coordination with Pho4p following cell cycle-linked depletion of intracellular phosphate (Neef and Kladde 2003; Pondugula et al. 2009), our engineered strains lack intracellular phosphate stores to buffer fluctuations. The cells then may experience phosphate shortages more quickly and shift mitotic activation earlier in the cell cycle (Neef and Kladde 2003). The possibility that Mcm1p binding creates a more accessible chromatin state for Pho4p-driven activation can be easily tested by eliminating Mcm1p binding sites from the *PHO5* promoter and looking for a decrease in S/G2 permissiveness for activation. Though *PHO5* activation may have a direct link to the cell cycle mediated by Mcm1p, the fact that we observe the S/G2 permissive period even when activating a synthetic promoter in Chapter 4 suggests additional, more general pathways are also involved.

While it may be difficult to adapt the deletion mutant library strategy to our highly engineered switchable transactivator strain, proteins and complexes identified as important to the S/G2 permissive period in noisy expression may have a role in the cell cycle-dependence of *PHO5* activation as well. Targeted experiments using the switchable transactivator would elucidate the kinetic role of these proteins and complexes, in addition to those known to contribute to efficient activation, promoter poising, and reassembly such as SAGA (Barbaric et al. 2001, 2003; Reinke and Hörz 2003; Gregory et al. 1998; Barbaric et al. 2007; Dhasarathy and Kladde 2005), NuA4 (Nourani et al. 2004), COMPASS (Wang et al. 2011; Carvin and Kladde 2004), Rpd3L (Wang et al. 2011), SWI/SNF (Neely et al. 2002; Dhasarathy and Kladde 2005; Barbaric et al. 2001), INO80 (Steger et al. 2003), Asf1 (Adkins et al. 2004; Korber et al. 2006), and Spt6 (Adkins and

Tyler 2006). Though many are not required for viability, the dynamic balance between these factors can be important to determining the initial promoter state (Gregory et al. 1998; Wang et al. 2011; Nourani et al. 2004) so deletion mutants would not provide a view of activation pathways at the endogenous *PHO5* promoter. Instead, we propose replacing catalytic and targeting subunits of the complexes with temperature sensitive alleles (either available through a large collection at the Boone Laboratory, University of Toronto or developed using one of several techniques (Tan et al. 2009b; Ben-Aroya et al. 2008)). During microfluidic culture, an objective heater can be used to shift the chamber to the non-permissive temperature prior to a step change in the transactivator. This would allow a test of the promoter response kinetics in the absence of specific factor activities to determine which pathways are utilized most readily by a particular promoter architecture. This technique can be extended to study the molecular mechanism of memory as well by shifting the temperature before, during, and after the first active transcription period in a reactivation test. By detailing how promoter architectures require or access different activation pathways, a more fundamental understanding of eukaryotic gene regulation will be gained.

Towards better design of eukaryotic networks

The grand goal of synthetic biology to build circuits to predictably perform complicated tasks using a toolbox of individual genetic components has had some successes, but has fallen far short of expectations to date. Often the components do not function or interact as expected (Kelly et al. 2009). Robust networks have also been difficult to achieve partially due to unexpected extrinsic effects. Unanticipated interactions with the host cell physiology can affect expression of genes in the engineered network (Tan et al. 2009a), and growth rate and the cell cycle may also play a role in undermining predicted behaviors.

As an example, one particular function of synthetic gene circuits that has elicited great interest is the design of a clock or oscillator (reviewed in (Purcell et al. 2010)). Though many published oscillator networks appear decoupled from the cell cycle, robust function remains an issue. In the case of the classic repressilator, transition to stationary phase halted the oscillatory nature of this gene network (Elowitz and Leibler 2000). The strong cell cycle-dependence we observed for activation of regulated genes in yeast may contribute to difficulty implementing synthetic

oscillators in eukaryotes. Not only does the presence of chromatin at eukaryotic promoters complicate gene regulation compared to that in prokaryotes, but large, periodic changes in its degree of repression would add an unexpected forcing function, which must be accounted for but could also be exploited.

Our studies at the *PHO5* promoter provide the foundation for a set of rules to design eukaryotic promoters with desired kinetic properties, and the switchable transactivator system provides a means to test novel promoters. We found that a nucleosome-free transactivator binding site allowed efficient activation during the S/G2 permissive period while aggregate promoter affinity was important to the ability to activate in G1. These principles were established for relatively conservative promoter changes to isolate the effect of each binding site, but other promoter features could provide additional regulatory behaviors. Longer and less variable delays could be achieved through stronger nucleosomes positioning, more nucleosomes between the nucleosome-depleted region and the TATA box, and additional occluded binding sites. The degree of cell cycle-dependence is also likely to be tunable and probably dominated by nucleosome positioning strength. After the investigation outlined above, architecture-dependent activation pathways can also be selected through promoter design. Different promoter architectures can be created through DNA sequence using a nucleosome positioning algorithm such as (Kaplan et al. 2009), or selected from other *PHO* or endogenous yeast promoters. The kinetics of each can then be probed using an appropriate chimerical, switchable transactivator. Detailed characterization of promoter kinetics with regard to the cell cycle will then allow more robust circuit construction through more accurate network models. While effects of the cell cycle are now being considered in modeling synthetic networks in bacteria (Chen et al. 2004; Yoda et al. 2007; Tuttle et al. 2005), they should prove crucial to successfully predicting network behavior of regulated genes in eukaryotes.

5.4 References

Adkins M, Howar S, Tyler J. 2004. Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Molecular Cell* **14**: 657–723.

Adkins M, Tyler J. 2006. Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Molecular Cell* **21**: 405–421.

Barbaric S, Luckenbach T, Schmid A, Blaschke D, Hörz W, Korber P. 2007. Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter in vivo. *The Journal of Biological Chemistry* **282**: 27610–21.

Barbaric S, Reinke H, Horz W. 2003. Multiple Mechanistically Distinct Functions of SAGA at the PHO5 Promoter. *Molecular and Cellular Biology* **23**: 3468–3476.

Barbaric S, Walker J, Schmid A, Svejstrup J, Hörz W. 2001. Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. *The EMBO Journal* **20**: 4944–4995.

Ben-Aroya S, Coombes C, Kwok T, O'Donnell KA, Boeke JD, Hieter P. 2008. Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae. *Molecular Cell* **30**: 248–58.

Carvin C, Kladde M. 2004. Effectors of lysine 4 methylation of histone H3 in Saccharomyces cerevisiae are negative regulators of PHO5 and GAL1-10. *The Journal of Biological Chemistry* **279**: 33057–33119.

Charvin G, Cross FR, Siggia ED. 2008. A microfluidic device for temporally controlled gene expression and long-term fluorescent imaging in unperturbed dividing yeast cells. *PloS One* **3**: e1468.

Chen L, Wang R, Kobayashi T, Aihara K. 2004. Dynamics of gene regulatory networks with cell division cycle. *Physical Review E* **70**: 011909.

Choi P, Cai L, Frieda K, Xie X. 2008. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**: 442–448.

Chubb JR, Trcek T, Shenoy SM, Singer RH. 2006. Transcriptional pulsing of a developmental gene. *Current Biology* **16**: 1018–25.

Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437**: 699–706.

Dhasarathy A, Kladde MP. 2005. Promoter Occupancy Is a Major Determinant of Chromatin Remodeling Enzyme Requirements. *Molecular and Cellular Biology* **25**: 2698–2707.

Elowitz M, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**.

Elowitz M, Levine A, Siggia E, Swain P. 2002. Stochastic gene expression in a single cell. *Science* **297**: 1183–1189.

Fazly A, Li Q, Hu Q, Mer G, Horazdovsky B, Zhang Z. 2012. Histone chaperone Rtt106 promotes nucleosome formation using (H3-H4)2 tetramers. *The Journal of Biological Chemistry* **287**: 10753–60.

Golding I, Paulsson J, Zawilski S, Cox E. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**: 1025–1061.

Gregory PD, Schmid A, Zavari M, Lui L, Berger SL, Hörz W. 1998. Absence of Gcn5 HAT Activity Defines a Novel State in the Opening of Chromatin at the PHO5 Promoter in Yeast. *Molecular Cell* **1**: 495–505.

Groth A, Corpet A, Cook AJL, Roche D, Bartek J, Lukas J, Almouzni G. 2007. Regulation of replication fork progression through histone supply and demand. *Science* **318**: 1928–31.

Hertel CB, Langst G, Horz W, Korber P. 2005. Nucleosome Stability at the Yeast PHO5 and PHO8 Promoters Correlates with Differential Cofactor Requirements for Chromatin Opening. *Molecular and Cellular Biology* **25**: 10755–10767.

Ito T, Bulger M, Kobayashi R, Kadonaga JT. 1996. Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. *Molecular and Cellular Biology* **16**: 3112–24.

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**: 362–6.

Kaufman PD. 1996. Nucleosome assembly: the CAF and the HAT. *Current Opinion in Cell Biology* **8**: 369–373.

Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, De Mora K, Glieberman AL, Monie DD, Endy D. 2009. Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of Biological Engineering* **3**: 4.

Korber P, Barbaric S, Luckenbach T, Schmid A, Schermer UJ, Blaschke D, Hörz W. 2006. The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *The Journal of Biological Chemistry* **281**: 5539–45.

Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. 2011. Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332**: 475–478.

Maiuri P, Knezevich A, De Marco A, Mazza D, Kula A, McNally JG, Marcello A. 2011. Fast transcription rates of RNA polymerase II in human cells. *EMBO Reports* **12**: 1280–5.

Mao C, Brown CR, Falkovskaia E, Dong S, Hrabeta-Robinson E, Wenger L, Boeger H. 2010. Quantitative analysis of the transcription control mechanism. *Molecular Systems Biology* **6**: 431.

Megee P, Morgan B, Mittman B, Smith M. 1990. Genetic analysis of histone H4: essential role of lysines subject to reversible acetylation. *Science* **247**: 841–845.

Munsky B, Neuert G, Van Oudenaarden A. 2012. Using gene expression noise to understand gene regulation. *Science* **336**: 183–7.

Muramoto T, Cannon D, Gierlinski M, Corrigan A, Barton GJ, Chubb JR. 2012. Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 7350–5.

Neef DW, Kladde MP. 2003. Polyphosphate Loss Promotes SNF/SWI- and Gcn5-Dependent Mitotic Induction of PHO5. *Molecular and Cellular Biology* **23**: 3788–3797.

Neely KE, Hassan AH, Brown CE, Howe L, Workman JL. 2002. Transcription Activator Interactions with Multiple SWI/SNF Subunits. *Molecular and Cellular Biology* **22**: 1615–1625.

Nourani A, Utley R, Allard S, Côté J. 2004. Recruitment of the NuA4 complex poises the PHO5 promoter for chromatin remodeling and activation. *The EMBO Journal* **23**: 2597–3204.

Pedraza J, Van Oudenaarden A. 2005. Noise propagation in gene networks. *Science* **307**: 1965–1974.

Pondugula S, Neef DW, Voth WP, Darst RP, Dhasarathy A, Reynolds MM, Takahata S, Stillman DJ, Kladde MP. 2009. Coupling Phosphate Homeostasis to Cell Cycle-Specific Transcription: Mitotic Activation of Saccharomyces cerevisiae PHO5 by Mcm1 and Forkhead Proteins. *Molecular and Cellular Biology* **29**: 4891–4905.

Purcell O, Savery NJ, Grierson CS, Di Bernardo M. 2010. A comparative analysis of synthetic genetic oscillators. *Journal of the Royal Society, Interface / the Royal Society* **7**: 1503–24.

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA synthesis in mammalian cells. ed. U. Schibler. *PLoS Biology* **4**: e309.

Raser JM, O'Shea EK. 2004. Control of stochasticity in eukaryotic gene expression. *Science* **304**: 1811–1814.

Recht J, Tsubota T, Tanny JC, Diaz RL, Berger JM, Zhang X, Garcia BA, Shabanowitz J, Burlingame AL, Hunt DF, et al. 2006. Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 6988–93.

Reinke H, Hörz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Molecular Cell* **11**: 1599–2206.

Steger D, Haswell E, Miller A, Wente S, O'Shea EK. 2003. Regulation of chromatin remodeling by inositol polyphosphates. *Science* **299**: 114-166.

Sternberg N, Hamilton D. 1981. Bacteriophage P1 site-specific recombination. *Journal of Molecular Biology* **150**: 467–486.

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* **332**: 472–474.

Tan C, Marguet P, You L. 2009a. Emergent bistability by a growth-modulating positive feedback circuit. *Nature Chemical Biology* **5**: 842–8.

Tan G, Chen M, Foote C, Tan C. 2009b. Temperature-sensitive mutations made easy: generating conditional mutations by using temperature-sensitive inteins that function within different temperature ranges. *Genetics* **183**: 13–22.

Tan RZ, Van Oudenaarden A. 2010. Transcript counting in single cells reveals dynamics of rDNA transcription. *Molecular Systems Biology* **6**: 358.

Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A, Xie XS. 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**: 533–8.

To T-L, Maheshri N. 2010. Noise Can Induce Bimodality in Positive Transcriptional Feedback Loops Without Bistability. *Science* **327**: 1142–1145.

Tuttle LM, Salis H, Tomshine J, Kaznessis YN. 2005. Model-driven designs of an oscillating gene network. *Biophysical Journal* **89**: 3873–83.

Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J. 2006. Origins of extrinsic variability in eukaryotic gene expression. *Nature* **439**: 861–4.

Wang S-S, Zhou B, Zhou J-Q. 2011. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Molecular and Cellular Biology* **31**: 3171–3252.

Williams SK, Truong D, Tyler JK. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9000–5.

Yoda M, Ushikubo T, Inoue W, Sasai M. 2007. Roles of noise in single and coupled multiple genetic oscillators. *The Journal of Chemical Physics* **126**: 115101.

Zopf CJ, Quinn K, Zeidman J, Maheshri N. 2013. Cell-cycle dependence of stochastic gene expression. *PLoS Computational Biology (submitted).*

APPENDIX

Table A.1 | Yeast strains used in this study

Strain	Relevant Genotype	Parent Strain	Reference
EY131	MATa pho4Δ::TRP1	W303	Gift from E. O'Shea
EY2150	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho84Δ::CFP-kanR PHO4::cYFP-HIS3	W303	Gift from E. O'Shea
EY2210	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::PHO4(1-741bp)-cYFP-HIS3 pho84Δ::CFP-kanR	W303	Gift from E. O'Shea
EY2436	pho5Δ::CFP-KANMX6		Gift from E. O'Shea
EY2439	P_{PHO5}-pho5Δ::P_{mutHL}-CFP-KANMX6		Gift from E. O'Shea
EY2441	P_{PHO5}-pho5Δ::P_{mutLX}-CFP-KANMX6		Gift from E. O'Shea
EY2442	P_{PHO5}-pho5Δ::P_{mutHX}-CFP-KANMX6		Gift from E. O'Shea
EY2470	P_{PHO5}-pho5Δ::P_{mutHH}-CFP-KANMX6		Gift from E. O'Shea
EY2472	P_{PHO5}-pho5Δ::P_{mutLL}-CFP-KANMX6		Gift from E. O'Shea
EY2474	P_{PHO5}-pho5Δ::P_{mutHXH}-CFP-KANMX6		Gift from E. O'Shea
EY2478	P_{PHO5}-pho5Δ::P_{mutHXL}-CFP-KANMX6		Gift from E. O'Shea
Y1	MATa trp1-1 can1-100 leu2-3,112 his 3-11,5 ura3 GAL+ ADE+	W303	Laboratory collection
Y3	MATa ade2-1 trp1-1 can1-100 leu2-3,112 his 3-11,5 ura3 GAL+	W303	Laboratory collection

Strain	Relevant Genotype	Parent Strain	Reference
Y6	MATα ade2-1 trp1-1 can1-100 leu2-3,112 his 3-11,5 ura3 GAL+	W303	Laboratory collection
Y47	MATa leu2::P_{ADH1}-CFP-hisG::URA3::kanR::hisG	Y1	Gift from E. O'Shea (Raser and O'Shea 2004)
Y114	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho84Δ::CFP-kanR pho4Δ::TRP1	EY2210 (EY131 PCR template)	This study
Y152	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::CFP-kanR phm4Δ::P_{GPD}-PHO81-HIS3MX6 leu2Δ::P_{TEF1m7}-PHO4-cYFP	Y114 (+ B133 PCR product, + B141)	Laboratory collection
Y216	MATα ade2-1 ura3::P_{7xtetO}-vYFP-kanR	Y6	Laboratory collection (Lee and Maheshri 2012)
Y231	MATa leu2::P_{PGK1}-RFP ADE2::P_{MYO2}-tTA ura3::P_{7xtetO}-CFP-kanR	Y3	Laboratory collection (Lee and Maheshri 2012)
Y236	MATa/α leu2::P_{PGK1}-RFP ADE2::P_{MYO2}-tTA ura3/ura3::P_{7xtetO}-CFP-kanR/ P_{7xtetO}-vYFP-kanR	Y231xY216	This study
Y320	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::KlURA3	Y114	This study
Y532	MATa doa1::vYFP + pRS316-DOA1	Y1 + B858	This study
Y784	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3	Y320	This study

Strain	Relevant Genotype	Parent Strain	Reference
Y785	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6	Y784	This study ("switchable transactivator" base strain)
Y789	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 pho5Δ::CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2436	This study
Y790	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHL}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2439	This study
Y791	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutLX}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2441	This study
Y792	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHX}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2442	This study
Y793	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHH}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2470	This study
Y794	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutLL}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2472	This study
Y795	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6	Y785 + *PHO5* locus PCR product from	This study

Strain	Relevant Genotype	Parent Strain	Reference
	P_{PHO5}-pho5Δ::P_{mutXH}-CFP-KANMX6	EY2474	
Y796	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutXL}-CFP-KANMX6	Y785 + *PHO5* locus PCR product from EY2478	This study
Y833	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 pho5Δ:: CFP-KANMX6 leu2Δ::P_{TEF1m7}-PHO4^{SD6-}cYFP-hisG::URA3::kanR::hisG	Y789 + B631	This study
Y837	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHL}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-PHO4^{SD6}-cYFP-hisG::URA3::kanR::hisG	Y790 + B631	This study
Y840	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutLX}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-PHO4^{SD6}-cYFP-hisG::URA3::kanR::hisG	Y791 + B631	This study
Y844	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHX}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-PHO4^{SD6}-cYFP-hisG::URA3::kanR::hisG	Y792 + B631	This study
Y848	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutHH}-CFP-KANMX6	Y793 + B631	This study

Strain	Relevant Genotype	Parent Strain	Reference
	leu2Δ::P_{TEF1m7}-$PHO4^{SD6}$-cYFP-hisG::URA3::kanR::hisG		
Y852	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutLL}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-$PHO4^{SD6}$-cYFP-hisG::URA3::kanR::hisG	Y794 + B631	This study
Y855	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutXH}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-$PHO4^{SD6}$-cYFP-hisG::URA3::kanR::hisG	Y795 + B631	This study
Y858	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutXL}-CFP-KANMX6 leu2Δ::P_{TEF1m7}-$PHO4^{SD6}$-cYFP-hisG::URA3::kanR::hisG	Y796 + B631	This study
Y868	MATα ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6	Y785 mating-type switch (using B278)	This study
Y869	MATα ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 pho5Δ::vYFP-KANMX6	Y868 + *YFP-KANMX6* PCR product (B182)	This study
Y871	MATa/α pho5Δ::CFP /pho5Δ::vYFP diploid	Y833 x Y869	This study
Y933	MATα ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6	Y869 + EB1425 pop-in/pop-out as	This study

Strain	Relevant Genotype	Parent Strain	Reference
	P_{PHO5}-pho5Δ::P_{mutLX}-vYFP-KANMX6	in (Lam et al. 2008)	
Y936	MATα ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutLL}-vYFP-KANMX6	B89	pJRL2
Y937	MATa/α P_{PHO5}-pho5Δ::P_{mutLX}-CFP / P_{PHO5}-pho5Δ::P_{mutLX}-vYFP diploid	Y840 x Y933	This study
Y940	MATa/α P_{PHO5}-pho5Δ::P_{mutLL}-CFP / P_{PHO5}-pho5Δ::P_{mutLL}-vYFP diploid	Y852 x Y936	This study
Y942	MATα ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 P_{PHO5}-pho5Δ::P_{mutXH}-vYFP-KANMX6	B89	pJRL2
Y944	MATa/α P_{PHO5}-pho5Δ::P_{mutXH}-CFP / P_{PHO5}-pho5Δ::P_{mutXH}-vYFP diploid	Y855 x Y942	This study
Y947	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 leu2Δ::P_{TEF1m7}-PHO4ΔP2-tetR-cYFP	Y785 (+B798)	This study
Y960	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4Δ::TRP1 pho84Δ::klura3 phm4Δ::HIS3MX6 leu2Δ::P_{TEF1m7}-PHO4ΔP2-tetR-cYFP URA3::P_{1xtetO}-CFP	Y947 (+B579)	This study
Y962	MATa ADE+ NHP2::RFP-NAT spl2Δ::LEU2 pho4::TRP1 pho84Δ::klura3 phm4::HIS3MX6 leu2Δ::P_{TEF1m7}-PHO4ΔP2-tetR-cYFP URA3::P_{7xtetO}-CFP	Y947 (+B720)	This study

Table A.2 | Plasmids used in strain construction

Plasmid	Base vector	Relevant gene	Construction information
B48	Yeast centromeric	P_{1xtetO}-CFP	Laboratory stock
B89	pJRL2	P_{PGK1}-vYFP	Laboratory stock, *leu2* integrating vector (Raser and O'Shea 2004)
B108	pBluescript	KlURA3	Laboratory stock, contains *URA3* gene from *K. lactis*; PCR template for *pho84Δ* in Y320
B129	pRS303	HIS3MX6	*HIS3MX6* from pFA6a-*HIS3MX6* (Longtine et al. 1998)(BglII/EcoRI digested) ligated into pRS303 (BamHI/EcoRI); PCR template for *phm4*Δ in Y785
B133	pRS303	P_{GPD}-PHO84-HIS3MX6	Into pRS303, cloned P_{GPD}-*PHO84* (SalI/PvuII into SalI/EcoRV) and *HIS3MX6* (BglII/EcoRI into BamHI/EcoRI)
B141	pJRL2	P_{TEF1m7}-PHO4 - cYFP	PCRed *PHO4-Citrine* from EY2150 genome and clone P_{TEF1m7} (SalI/EcoRI) – *PHO4 – Citrine* (EcoRI/NotI) into SalI/NotI digested pJRL2 (B89). P_{TEF1m7} from (Nevoigt et al. 2006).
B163	pRS303	P_{1xtetO}-vYFP	Laboratory stock
B165	pRS303	P_{7xtetO}-vYFP	Laboratory stock
B198	pRS303	P_{7xtetO}-tTA	Laboratory stock
B228	pCM189	P_{ADH1}-tTA	Laboratory stock
B229	pRS303	P_{1xtetO}-tdTomato	Laboratory stock
B278	pRS316	P_{GAL1}-HO	Laboratory stock
B464	pBluescript	klura3	Digested B108 with EcoRV and self-ligated to remove portion of ORF and make *KlURA3* non-

Plasmid	Base vector	Relevant gene	Construction information
			functional; PCR template for *klura3* Y784
B579	pRS306	P_{1xtetO}-CFP	XhoI/NotI digested P_{1xtetO}-*CFP* from B48 ligated in front of T_{ACT1} in pRS306 (XhoI/NotI)
B720	pRS306	P_{7xtetO}-CFP	XhoI/BamHI digested P_{7xtetO}-*CFP* from B198 ligated in front of T_{ACT1} in pRS306 (XhoI/NotI)
B798	pJRL2	P_{TEF1m7}-PHO4ΔP2-tetR-cYFP	P_{TEF1m7} (SalI/EcoRI) – *PHO4*(1-600bp) (EcoRI/SpeI) – linker – *tetR* (SpeI/PacI) – linker – *Citrine* (PacI/NotI) ligated into SalI/NotI digested pJRL2 (B89)
B798	pJRL2 hisG::URA3::kanR::hisG	P_{TEF1m7}-$PHO4^{SD6}$-cYFP	*PHO4* gene PCRed from the MluI-restriction site 58bp 5' to S223 (using a primer carrying the S→D mutation TCT→GAT) to the PacI-restrcition site between *PHO4* and *cYFP*, and cloned into B141 (MluI/PacI)
EB1424	pRS306	P_{mutXH}-*PHO5*	Gift from F. Lam (Lam et al. 2008)
EB1425	pRS306	P_{mutLX}-*PHO5*	Gift from F. Lam (Lam et al. 2008)
EB1734	pRS306	P_{mutLL}-*PHO5*	Gift from F. Lam (Lam et al. 2008)

5.5 References

Lam FH, Steger DJ, O'Shea EK. 2008. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**: 246–296.

Lee T-H, Maheshri N. 2012. A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Molecular Systems Biology* **8**: 576.

Longtine MS, McKenzie 3rd A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**: 953–961.

Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. 2006. Engineering of Promoter Replacement Cassettes for Fine-Tuning of Gene Expression in Saccharomyces cerevisiae. *Applied and Environmental Microbiology* **72**: 5266–5273.

Raser JM, O'Shea EK. 2004. Control of stochasticity in eukaryotic gene expression. *Science* **304**: 1811–1814.