SPATIO TEMPORAL REGULATION OF CELL CYCLE TRANSCRIPTION

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF APPLIED PHYSICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Umut Eser

March 2013

ABSTRACT

The *Start* checkpoint regulates cell cycle commitment and associated transcription in the budding yeast, *Saccharomyces cerevisiae*. It was previously shown that commitment to cell division corresponds to activating the positive feedback loop of G1 cyclins controlled by the transcription factors SBF and MBF. Around this pivotal cell cycle event, over 300 genes (G1/S regulon) are expressed to facilitate the G1/S transition. Despite its importance, little was known about distinct temporal regulation within the G1/S regulon. We found that SBF and MBF target genes have a well-defined distribution of transcriptional activation times. We also showed that activation of G1 cyclins precedes the activation of the bulk of the G1/S regulon, which we named 'feedback-first' regulation. In budding yeast, feedback-first regulation ensures that commitment to cell division occurs before large-scale changes in transcription. Thus, the transition can be viewed as a two-step process whereby the decision to divide precedes synthesis of the cellular machinery required for division. Furthermore, we found that feedback-first regulation is conserved in the related yeast *S. bayanus* as well as human cells. This finding highlighted the importance of understanding the molecular mechanisms through which co-regulated genes can have distinct activation dynamics. We showed that timing is partially explained by the combinatorial use of SBF and MBF transcription factors, which implement a logical OR function for gene activation. In addition to combinatorial use of transcription factors, we analyzed genome-wide chromosome conformation capture data to examine the potential link between the timing of gene expression and 3-D genome architecture. The early-activated genes of the G1/S regulon are significantly enriched for the number of physical contacts to the rest of the genome. Further analysis revealed two main clusters, whose interactions co-vary and whose activation time distributions are distinct. Taken together, these our work explains a significant amount of

timing variation within cell cycle-dependent gene expression. Thus, we concluded that the cell utilizes both genome architecture and the combinatorial use of transcription factors to implement feedback-first regulation ensuring that commitment to cell division precedes genome-wide cell cycle-dependent transcription.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisor Dr. Jan Skotheim. This dissertation would not have been possible without his guidance and help. He is a great advisor in relation to both personal and academic life. His positive attitude, sincerity, and encouragement facilitated my engagement in life science as I was a physicist who had little experience in biology before I stated in his lab. I have learned many useful skills from him.

I am grateful to my committee members Dr. Steven Block and Dr. Daniel Fisher for their encouraging words, thoughtful criticism, and time and attention during busy academic quarters. I am grateful for the help from many lab members including: Melody Falleur-Fettig, who helped construct yeast strains and participate in the time-course microarray experiments shown in chapter 2; Dr. Andreas Doncic, in addition to providing friendly support, built an image analysis algorithm that was useful for my time-lapse single cell analysis shown in chapter 2; Amy Johnson helped me to identify the E2F transcription factor target list in the mammalian system, which I used in chapter 5; Devon Chandler-Brown who reviewed a portion of the thesis. In addition, I would like to thank the other lab members Amanda Amodeo, Jon Turner, Jenifer Ewald, Oguzhan Atay, Clayton Schwarz, Kurt Schmoler and David Jukam. I am grateful for the assistance of the administrators of the Departments of Biology and Applied Physics, especially, Lisa Pereira, Paula Perron, and Claire Nicholas, for their untiring effort in managing our paperwork.

Last but not the least, I would like to thank my family and the bay area Turkish community for all types of support. Without them, it would have been very difficult for me to adapt here.

DEDICATION

The milestone of my decision to become a scientist dates back to my highschool years where I have intensively studied physics olympiads. Therefore, I would like to dedicate my thesis to Samanyolu Highschools Science Olympiad group, especially to my olympiad coordinator Zafer Şimşek, who guided me to find my source of motivation and inspiration.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

MOTIVATION

Before explaining the details of my research, I would like to explain my motivation to study life science as a physicist. My physics training inculcated an urge to identify a unifying framework to explain widely varied phenomena. Physics has a long tradition of beginning with a purely empirical and descriptive view of the world and using these observations to build a theoretical framework that allows prediction of the properties (and even existence) of a wide variety of physical phenomena. In biosciences, however, due to technological limitations and the very complex nature of life, we have been gathering information at the empirical regime of knowledge with few theoretical frameworks that provide predictive power. However, many technical advances, such as such as fluorescence microscopy, data storage and processing, and deep sequencing, are rapidly lowering the cost of data acquisition and analysis. Therefore, I would like to utilize these technological advances and quantitative approaches to develop an predictive framework and identify the fundamental principles underlying biological phenomena. Let me briefly explain my approach to the biological sciences with an analogy:

Imagine that intelligent extraterrestrials on Mars receive the rover robot called Curiosity and they want to understand how that machine works (Figure 1). If we assume that they have no idea about our history of technology, we can say that the Martians should use multiple research stages to understand the weird object they just received. The first step will be to investigate the chemical composition of the robot's materials. Based on their chemical and electrical properties, they might proceed to characterize the components and their individual behavior. However, this will not suffice to understand how the robot works. Another layer of research will be conducted to conceptualize the relationship between individual components, which would reveal network diagrams for each robot function. Nevertheless, robots employ a large number of both elements and connections

to function. Therefore the resulting network diagram showing all the connections between the elements will look like a hairball. By looking at this map of interactions, the Martians might think that they understand the robot. But this delusion will only be valid until they see another robot (like that which will be sent in 2020 for the Mars rover mission, Figure 2). Once they notice that the wiring and circuit elements are very different, their claim that they understand robots will be reduced to merely understanding the Curiosity model robot. Hence, to understand robotics more generally, another layer of research revealing principles such as Kirchhoff laws and control theory will be necessary. Discovering the working principles or design principles of robots will allow the Martians to identify the robots' errors, fix them and even to make their own robots.

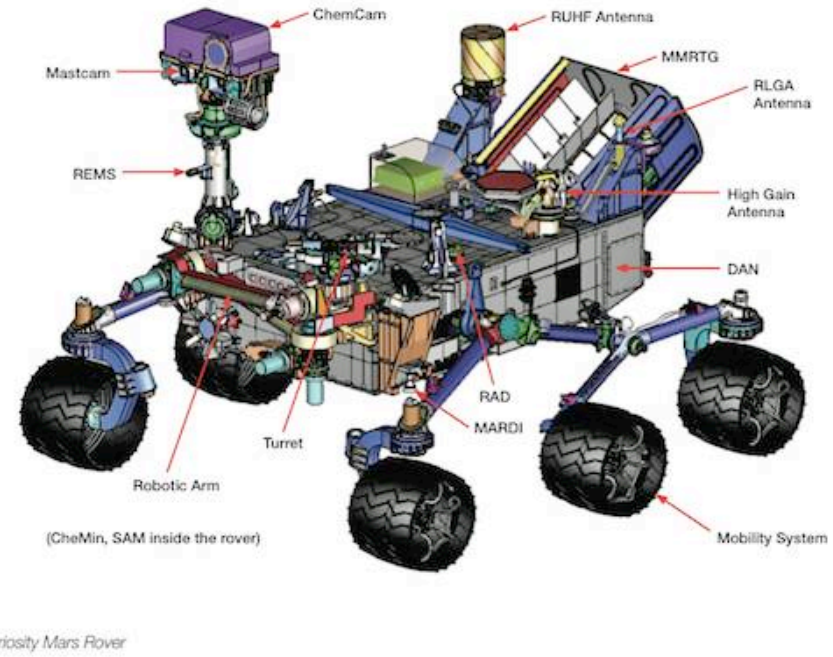


Figure 1**:** Schematic of the Curiosity rover identifying different components (adapted from http://www.theregister.co.uk/2012/08/08/mars_probe_cpu/)

As for this example of dissecting robotics, multiple layers of research are also required to understand living organisms. For instance, understanding how cells work necessitates identification and characterization of nucleic acids and proteins, finding their interactions and deciphering their signaling networks. Nevertheless, in addition to elucidating such schematics, we need to learn the general principles governing cellular signaling and information processing in order to deepen our understanding of nature, to increase the predictive power of our theories, and to improve our ability to design synthetic circuits for medical and engineering purposes.

Use of new quantitative approaches to understand the principles behind cellular phenomena is a growing area of biological research. For example, cells may undergo irreversible changes in their states. Such cellular transitions have been observed in such varied circumstances as bacterial sporulation and vertebrate stem cell differentiation. Another prominent example of an irreversible cellular transition is the point of commitment to the mitotic cell cycle in budding yeast, which forms the subject of this dissertation. Yeast commit to division within the G1 phase of the cell cycle; after cell division, but before initiation of DNA replication. The commitment point was apparently embedded within the activation of the expression of the largest cell cycle-dependent regulon. Indeed, 5-10% of all yeast genes are activated near the commitment point. Broadly speaking, my work examines the commitment transition at an unprecedented temporal resolution to reveal an exquisite pattern within the timing of transcriptional activation. My work reveals novel functions for fine-tuned transcriptional order and begins to address its molecular underpinnings.

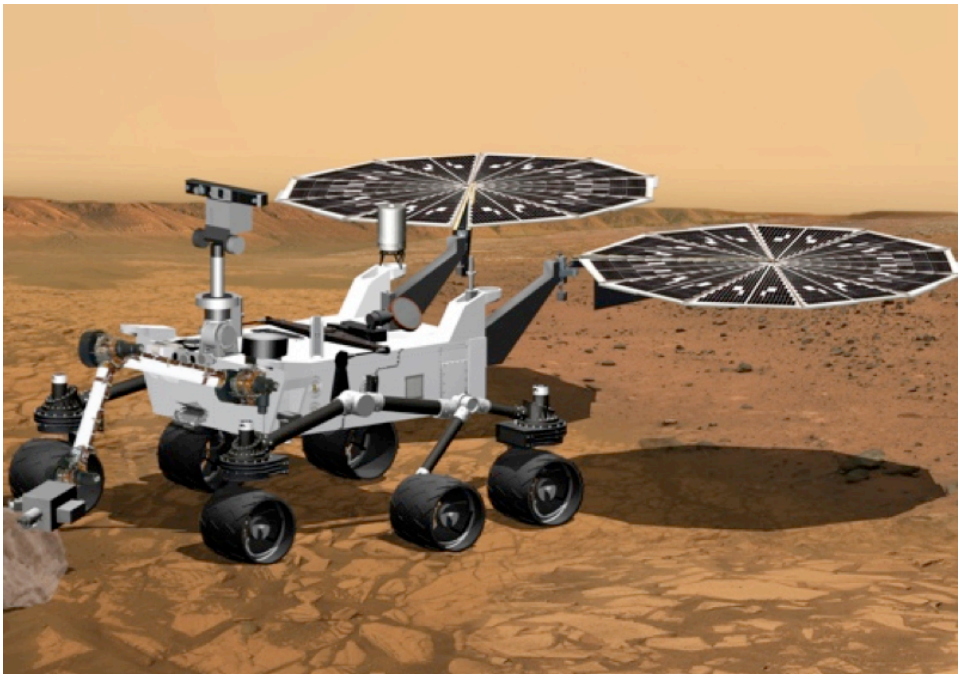


Figure 2: Image illustrating the concept of NASA's unnamed 2020 Mars mission rover (adapted from http://en.wikipedia.org/wiki/).

BACKGROUND

Order may be produced in a sequence of biochemical events through feedback control mechanisms or substrate-specific chemical kinetics. In the cell cycle, regulatory checkpoints ensure the proper order of many essential events through feedback control. DNA replication must be finished and damage repaired before mitosis, while anaphase is initiated only after complete spindle assembly (Morgan 2007). Checkpoints use designated regulatory molecules to restrain cell cycle progression until a set of criteria are satisfied (Hartwell and Weinert 1989). However, order without checkpoint control is observed in *Xenopus* embryos as cell cycle events are entrained by oscillations in cyclin dependent kinase (CDK) activity. Furthermore, addition of CDK substrates to *Xenopus* egg extracts in different stages of mitosis revealed that the order of substrate phosphorylation is independent of cell cycle phase (Georgi, Stukenberg et al. 2002). Thus, temporal order of phosphorylation in mitosis is likely the result of substrate-

specific kinetics. Here, we investigate the integration of chemical kinetics and feedback control at the *Start* transition in budding yeast.

The *Start* checkpoint, between cell division and DNA replication, prevents premature cell cycle entry (Hartwell, Culotti et al. 1974). Prior to *Start*, cells integrate internal (*e.g.*, cell size) and external (*e.g.*, mating pheromone) signals to make an all-or-none decision to divide. Beyond *Start*, cells are committed to divide regardless of changes in extracellular signals. Passage through *Start* corresponds precisely to the activation of the G1 cyclin positive feedback loop (Doncic, Fettig and Skotheim data not published). Thus, *Start* is a member of a growing list of cellular and developmental transitions driven by positive feedback (Pomerening, Sontag et al. 2003; Xiong and Ferrell 2003; Holt, Krutchinsky et al. 2008; Justman, Serber et al. 2009; Lopez-Aviles, Kapuy et al. 2009).

Positive feedback at *Start* is initiated by the G1 cyclin, Cln3 in complex with the cyclin dependent kinase Cdc28 (Figure 3). The primary target of Cln3 is the transcriptional inhibitor Whi5, whose inactivation is rate-limiting for the expression of the G1/S regulon (Costanzo, Nishikawa et al. 2004; de Bruin, McDonald et al. 2004). Cln3-Cdc28 phosphorylates and initiates Whi5 inactivation, which allows some transcription of two additional G1 cyclins, *CLN1* and *CLN2* (Tyers, Tokiwa et al. 1993). The downstream G1 cyclins then complete the positive feedback loop through the inactivation and nuclear exclusion of Whi5 and the full activation of the transcription factors SBF (Swi4-Swi6) and MBF (Mbp1-Swi6) (Andrews and Herskowitz 1989; Nasmyth and Dirick 1991; Koch, Moll et al. 1993; de Bruin, McDonald et al. 2004; Skotheim, Di Talia et al. 2008).

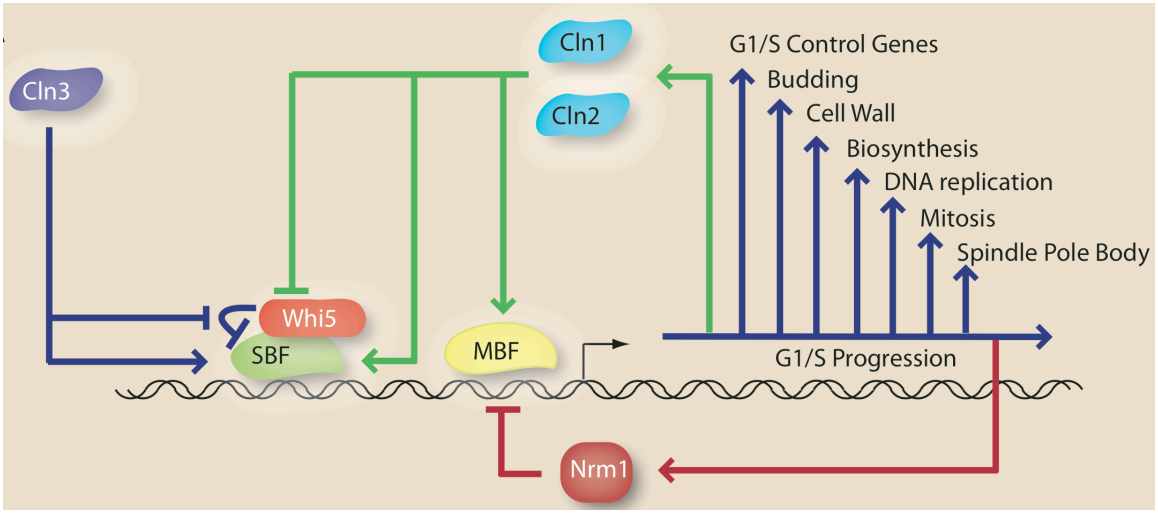


Figure 3: Schematic diagram of the G1/S transition.

Surprisingly, the transcription factors at the center of the positive feedback loop, SBF and MBF, are also responsible for the transcription of over 200 additional genes (Ferrezuelo, Colomina et al. 2010). Indeed, cell cycle commitment appears to coincide with the coordinated transcriptional activation of approximately 5% of all genes (Spellman, Sherlock et al. 1998). Although Whi5 phosphorylation is rate-limiting for activation of positive feedback, it is also likely to be rate limiting for the transcription of all SBF regulated genes due to the direct Whi5-SBF interaction (de Bruin, McDonald et al. 2004). The concurrent activation of the related heterodimeric transcription factor MBF also requires CDK activity, possibly through phosphorylation of the shared component Swi6 (Wijnen, Landman et al. 2002). Thus, given the integrated nature of the regulatory circuit and the ability of the upstream cyclin Cln3 to activated SBF- and MBF-dependent transcription in *cln1Δ cln2Δ* cells (Dirick, Bohm et al. 1995; Stuart and Wittenberg 1995), it is unclear if genome-wide changes in transcription occur after commitment to division.

Although G1/S transcription is largely regulated by SBF and MBF, single-cell studies have revealed significant differences in transcriptional activation of the 3 regulon

members *CLN2*, *RAD27* and *RFA1* (Skotheim, Di Talia et al. 2008). A rapid, feedback-driven increase in CDK activity drives the coherent and nearly simultaneous induction of these three genes in WT cells. However, significant differences in transcriptional activation timing are revealed in *cln1Δ cln2Δ* cells lacking positive feedback. *CLN2* is induced earlier than two other regulon members, which suggests a model in which full regulon expression would only occur after feedback loop activation to avoid detrimental transcription in cases where the cell does not commit to the mitotic cell cycle. Therefore, we hypothesized that the G1 cyclins *CLN1* and *CLN2*, involved in positive feedback, would be activated earlier than other genes in the G1/S regulon to ensure that commitment precedes the genome-wide change in transcription. This served as the entry point for my investigations.

OVERVIEW

The first chapter introduces the concepts and techniques that are used throughout the dissertation. Hence, it encompasses the basics of transcriptional dynamics and the technologies we use to examine it, *e.g.*, microarrays and computational approaches to detect the transcriptional activation timing. The second chapter reports the observation that the two G1 cyclins, namely *CLN1* and *CLN2,* are among the earliest activated genes of the G1/S regulon, which supports the hypothesis that genome-wide changes in transcription occur only after a cell is committed to division. By comparing sets of genes regulated by SBF, MBF, or by both factors together, we found that both transcriptional activation and inactivation can be approximated as logical OR functions (Chapter 3). Such combinatorial use of transcription factors partially explains the temporal variation of transcriptional activation. Given the heterogeneous composition of the nucleus, we analyzed a genome-wide chromosome conformation capture dataset (Duan et al. 2010) and showed that nuclear compartmentalization may affect the transcriptional activation timing (Chapter 4). Furthermore, *CLN1* and *CLN2* remain among the earliest activated cell cycle regulated genes in the related yeast, *S. bayanus*, which has significantly diverged gene expression (Tirosh, Weinberger et al. 2006; Guan, Dunham et al. 2010). A

similar analysis of human tissue culture cells revealed that functionally analogous feedback loop components E2F1, Skp2, and the cyclins E1 and E2 (Blagosklonny and Pardee 2002; Yung, Walker et al. 2007) are among the earliest activated cell cycle regulated targets of the E2F family of transcription factors (Chapter5). Taken together, our results demonstrate that feedback-first regulation, which places genome-wide changes in transcription downstream of positive feedback-dependent cell cycle commitment, is a common feature of G1/S control across eukaryotes. In addition, we have begun to unravel the molecular mechanisms underlying the temporal order of transcriptional activation.

CHAPTER1 : QUANTIFYING TRANSCRIPTIONAL DYNAMICS

MICROARRAYS

Monitoring the global-scale activity is essential for the quantitative studies of complex systems, such as cells. A microarray is an array of many spots that contain covalently bound single stranded DNA pieces, which revolutionized the molecular biology studies by introducing the genome-wide measurements (Schena, Shalon et al. 1995; DeRisi, Penland et al. 1996; Spellman, Sherlock et al. 1998; Hoheisel 2006). Microarray measurements are based on the hybridization of single stranded complementary DNA (cDNA) onto the probe oligo DNA. The higher the complementary base pairs, the stronger the strands form non-covalent bonds. A washing step is necessary to purify the array from the non-specifically bound sequences. Hence, measuring the intensity of fluorescence generated by the fluorescently labeled hybridized cDNAs quantifies the expression level relative to the intensity measurement of the same array under another condition (Figure 1-1).

One of the many widespread applications of microarrays is the quantification of the gene expression, *i.e.*, amount of mRNA. Gene expression is a crucial step in the central dogma as many of the cellular functions are regulated at the transcriptional level. High-throughput monitoring the transcriptional activity using microarrays includes computational analysis of the resulting genome-scale transcription data. The main challenge on the analysis side is often to distinguish between the real signal and the noise. DeRisi and coworkers defined an empirical criterion for a significant increase using concordance correlation analysis, *i.e.* reproducibility of the data (DeRisi, Penland et al. 1996). They reported that a 2-fold variation is usually sufficient to classify a data point as

a significant signal, which has been commonly misinterpreted as a universal threshold in microarray experiments (Hoheisel 2006).

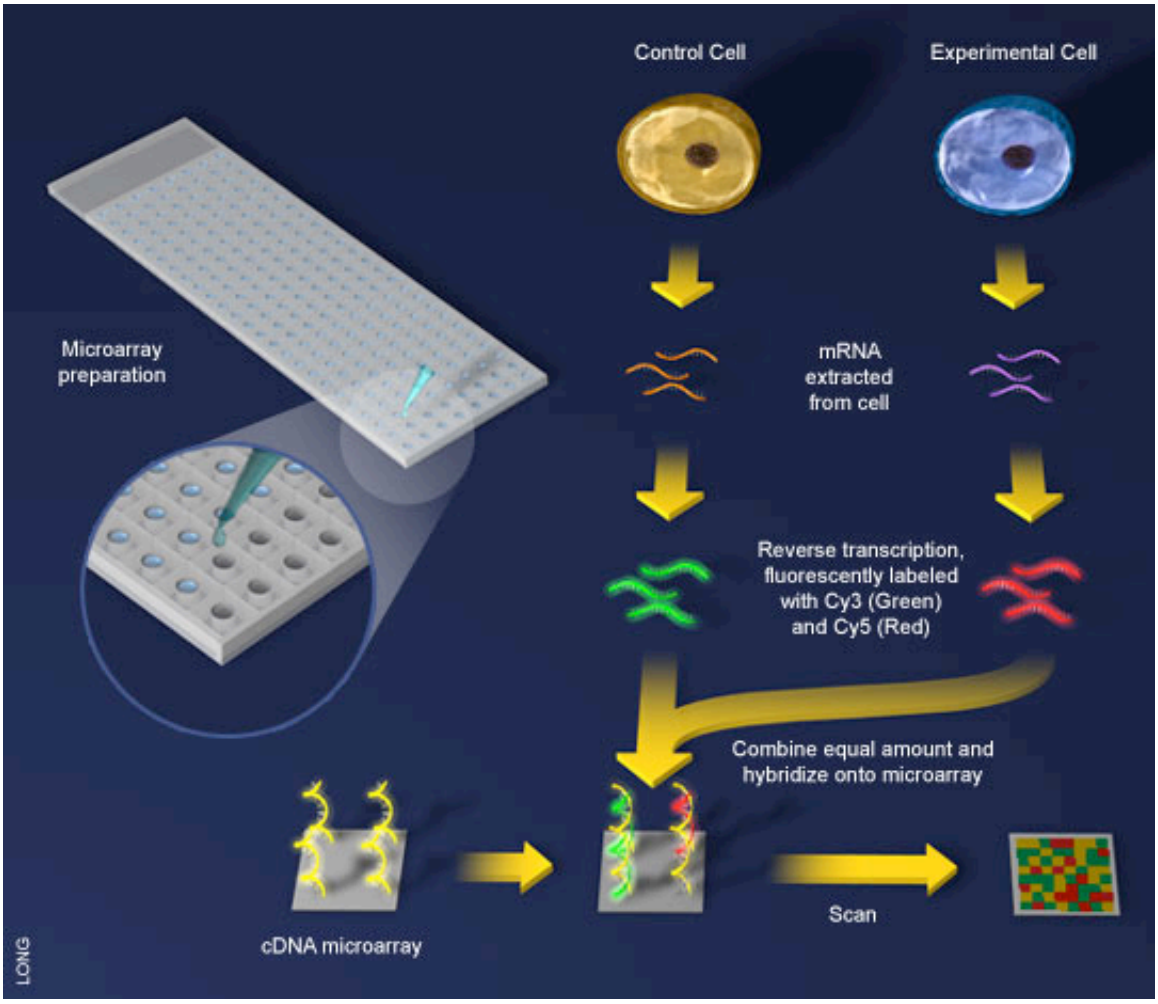


Figure 1-1: A simplified scheme showing an example microarray experiment procedure (adapted from http://www.scq.ubc.ca/spot-your-genes-an-overview-of-the-microarray/).

In microarray time-course experiments, it is easier to overcome signal-to-noise problems because adjacent time points can provide similar information as biological replicates if one expects a smooth signal. Such a signal can be expected if the time scale of changes to gene expression is larger than the time scale of time-course sampling. Therefore,

applying an appropriate curve fitting to the expression profiles can identify noisy data points, which may subsequently be excluded from the analysis.

Microarray time-course experiments are usually used to observe the temporal behavior of gene expression. A common way to determine the expression timing of a gene is the clustering of similar expression profiles (Spellman, Sherlock et al. 1998; Peddada 2002; Kim and Kim 2007; Orlando, Lin et al. 2008). However, clustering loses information about the temporal order of gene activation. Another way to temporally locate gene activity is by detecting the peak time of an expression profile. However, this denotes the inactivation of gene expression rather than the time of activation. Therefore, a more rigorous approach is required to analyze the dynamics of a genome-wide transcriptional activation.

ALGORITHM TO DETECT TRANSCRIPTIONAL ACTIVATION TIME

In order to robustly identify the time of gene activation, we developed a novel algorithm to determine the time at which a specific gene is induced during the cell cycle. Although manually identifying activation points of cell cycle regulated genes is not difficult, we developed an automated algorithm to both avoid potential bias and increase throughput. Our algorithm is robust to noisy data, which can produce incorrect estimates for the activation time. We normalized all the time series and assumed that the time scale for changing transcript concentration is greater than 10 minutes. We therefore remove data points associated with large concentration changes on shorter timescales. Data points further than 20% of the dynamic range of the time series (maximum – minimum) from adjacent points are removed. We discarded time series with two or more removed data points. The mRNA level is then estimated using smoothing-splines. We selected the point where the 1st derivative first reaches 10% of its maximum. The smoothing parameter is optimized to minimize variation in biological replicates and the 1st derivative method is shown to be superior in estimating activation times relative to other methods (Figure 1-2).

The change in concentration of an mRNA species, $dX(t)/dt$, depends on a balance between the transcription rate, $r(t)$, and degradation rate, $\lambda X(t)$, so that

$$\frac{dX(t)}{dt} = r(t) - \lambda X(t)$$

Here, we want to find the time point where $r(t)$ increases rapidly, i.e., the activation time. For cell cycle dependent changes in r, the timescale of mRNA degradation (~20min)(Grigull 2004) is more rapid than the cell cycle timescale (>90 minutes) so that we may neglect degradation kinetics around the time of gene activation ($\lambda X(t) << r$) and consider the point where dX/dt increases above a threshold as the activation point (r > threshold). Hence, we define the activation time as the time point where the first derivative of the expression profile, dX/dt, reaches the defined threshold, which we take as 10% of the maximum value of the 1st derivative within a given time series. Alternatively, we can consider the value where the second derivative d^2X/dt^2 is maximum as the activation point, *i.e.*, where the increase in transcription rate is most rapid (2nd derivative method)(Skotheim, Di Talia et al. 2008). Since both the 1st derivative and 2nd derivative can be used to estimate activation, we must rationally choose between methods.

Since noisy data can produce incorrect estimates for the activation time, we require a systematic method to exclude outlying data points and discard low-quality time series. Because the time scale for changing transcript concentration is greater than 10 minutes, we remove data points associated with large concentration changes on shorter time scales. If a data point is more than 20% of the dynamic range of the time series (maximum – minimum), away from both the adjacent points, it is removed. We discarded time series where two or more points removed or missing from the original data.

To estimate *X(t),* we applied smoothing-spline function by using the Curve Fitting Toolbox in MATLAB.

CHOOSING THE SMOOTHING PARAMETER

For both methods of detecting the activation time, we optimized our smoothing spline. The smoothing spline, f, minimizes the following function:

$$p\sum_{j=1}^{n}\left|y(:,j)-f(x(j))\right|^2+(1-p)\int\left|D^2f(t)\right|^2\,dt$$

where x(j) is the time at the j^{th} time point, n is the number of data points, *p* is the smoothing parameter, y is the data value, $D^2f(t)$ is the second derivative of the smoothing fit function.

The smoothing parameter can be chosen between 0 and 1. For lower values, the fit over-smoothes and approaches a linear fit, which does not contain information about the activation time. For higher values, p~1, the fit just connects all the data points and is too sensitive to small experimental errors. This leads to inconsistent results among experimental replicates. Therefore, we can test the smoothing parameters by comparing the standard deviations for the distributions of the activation times from a set of experimental replicates.

To identify the best method and its associated optimal smoothing parameter, we produced a training set of data, which is composed of time series containing a step-wise change in transcription rate, r, at time t_0 and a small initial mRNA concentration, X_0. The analytical solution to equation 1 is as follows:

$$\frac{dX}{dt} = \begin{cases} -\lambda X, & for\ t < t_0 \\ r - \lambda X, & for\ t \geq t_0 \end{cases} \quad (3)$$

$$X = X_0 \ \ at \ \ t = 0 \quad (4)$$

$$X(t) = \begin{cases} X_0\, e^{-\lambda t}, & t < t_0 \\ \frac{r}{\lambda} + \left(X_0 - \frac{re^{\lambda t_0}}{\lambda}\right) e^{-\lambda t}, & t \geq t_0 \end{cases} \quad (5)$$

When we introduce a uniformly distributed noise function $\rho(t)$, multiplied by a tunable noise coefficient η, the expression profile becomes:

$$X(t) = \begin{cases} X_0\, e^{-\lambda t} + \eta\rho(t), & t < t_0 \\ \frac{r}{\lambda} + \left(X_0 - \frac{re^{\lambda t_0}}{\lambda}\right) e^{-\lambda t} + \eta\rho(t), & t \geq t_0 \end{cases} \quad (6)$$

Good activation time detection will yield little variation in activation times around t_0 in spite of the noise. Therefore, we tested the 1st and the 2nd derivative methods for a range of smoothing parameters (spanning between 0.001 and 0.87) and noise coefficients (from 0.05 to 5) (Figure 1-2B).

We set the activation time for the simulated data to t_0=13 min. with an initial transcript level $X_0 = 1$ (a.u.), and tested the training data for a range of noise levels ($\eta \in [0.05\text{-}5]$). The mean activation time of the 1st derivative method converged to the correct activation time (±1min) for the smoothing parameter values greater than 0.2 with a slowly increasing variance of ~5 min (Figure 1-2C). However, the mean activation time for the 2nd derivative method hits the defined activation time 13 min only for smoothing parameter values between 0.001 and 0.03meanwhile the variation changes from 4 to 8 min and is larger than for the 1st derivative method. For greater smoothing parameters, the 2nd derivative algorithms becomes more unreliable and mean activation time converges to 20 with the variation 9. Therefore, we choose 0.25 as a smoothing parameter and use the 1st derivative method in all our analysis.

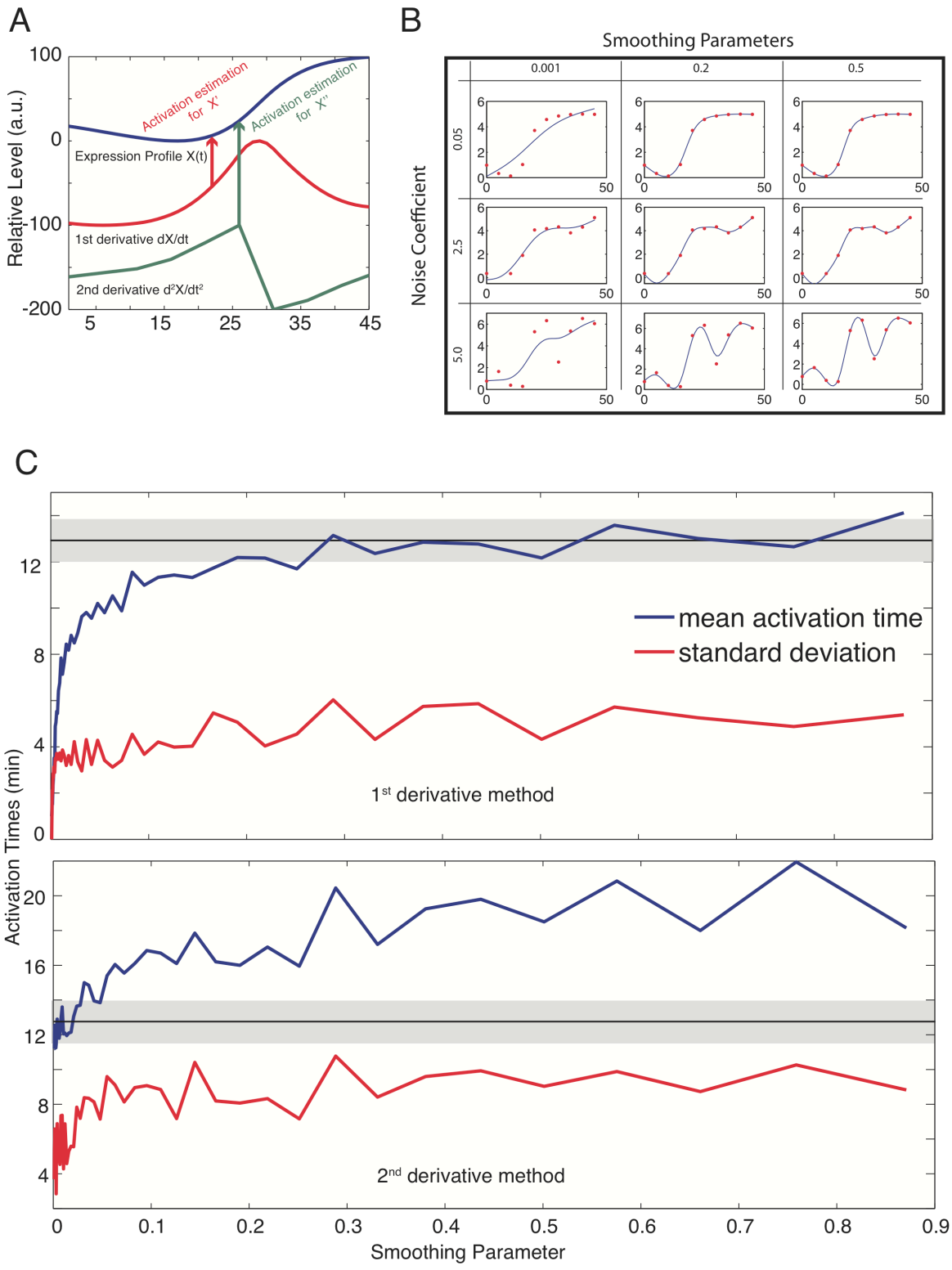

A
Relative Level (a.u.)
Activation estimation for X'
Activation estimation for X''
Expression Profile X(t)
1st derivative dX/dt
2nd derivative d²X/dt²
B
Smoothing Parameters
0.001
0.2
0.5
Noise Coefficient
0.05
2.5
5.0
C
mean activation time
standard deviation
1st derivative method
2nd derivative method
Activation Times (min)
Smoothing Parameter

Figure 1-2**:** A) Activation time of a gene can be calculated using the 1st derivative or the 2nd derivative. (B) Simulated data with three different noise levels (red dots) are plotted for three different smoothing parameters. Blue curves show the smoothing spline fit to the data, f(t). (C) The mean (blue) and the standard deviations (red) of the activation times of training set are shown for the 1^{st} derivative method (upper panel), and for the 2^{nd} derivative method (lower panel). We set the activation time to 13 min, shown as the horizontal line. The grey area denotes the activation time t_0 +\- 1min region.

CHAPTER2 : FEEDBACK-FIRST REGULATION

Single-cell studies have revealed significant differences in transcriptional activation of the 3 G1/S genes *CLN2*, *RAD27* and *RFA1* (Skotheim, Di Talia et al. 2008). A rapid, feedback-driven increase in CDK activity drives the coherent and nearly simultaneous induction of these three genes in WT cells. However, significant differences in transcriptional activation timing are revealed in *cln1Δ cln2Δ* cells lacking positive feedback. *CLN2* is induced earlier than two other regulon members, which suggests a model in which full regulon expression would only occur after feedback loop activation to avoid detrimental transcription in cases where the cell does not commit to the mitotic cell cycle. Therefore, we hypothesized that the G1 cyclins *CLN1* and *CLN2*, involved in positive feedback, would be activated earlier than other genes in the G1/S regulon to ensure that commitment precedes the genome-wide change in transcription.

DEFINING THE REGULON

To test our model that induction of positive feedback and concomitant cell cycle commitment precedes large-scale transcriptional change, we first need to accurately define the G1/S regulon. We are interested in the set of genes whose transcription is initiated due to increasing cyclin activity rather than upstream cyclin-independent processes (MacKay, Mai et al. 2001; Di Talia, Wang et al. 2009).

The set of cell cycle regulated genes was defined as the 800 genes with the largest amplitude mRNA concentration oscillation through the cell cycle (Spellman, Sherlock et al. 1998). To identify the set of G1 cyclin regulated genes, we relied on a second experiment by Spellman et al (1998), which identified a set of genes responding to exogenous Cln3 induction in G1 arrested *cln1Δ cln2Δ cln3Δ* cells. We took the top 413 as

the set of G1 cyclin inducible genes. The intersection of these two sets defines the 362-gene regulon (Figure 2-1B; Appendix D).

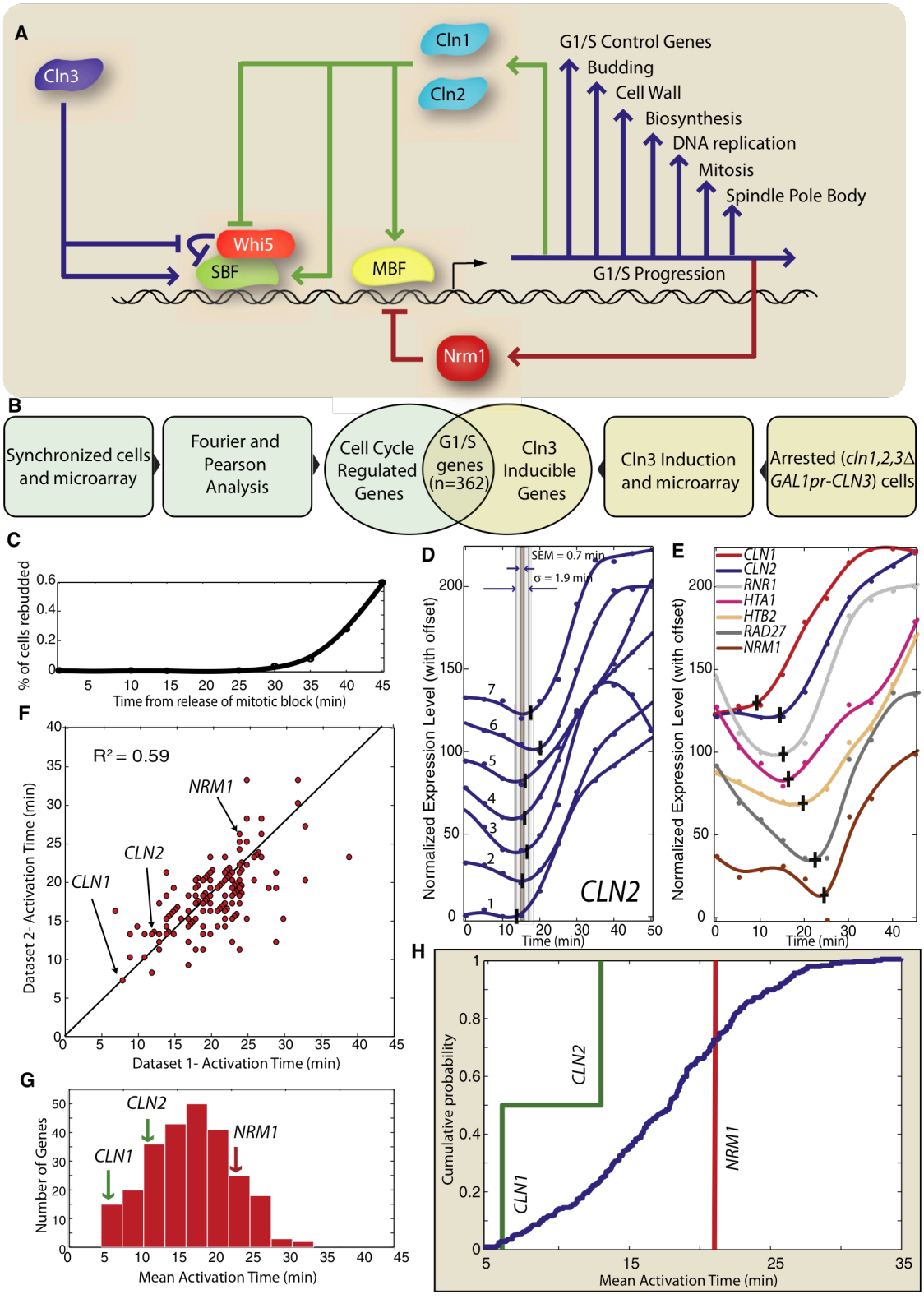

A
Cln3
Cln1
Cln2
G1/S Control Genes
Budding
Cell Wall
Biosynthesis
DNA replication
Mitosis
Spindle Pole Body
Whi5
SBF
MBF
G1/S Progression
Nrm1
B
Synchronized cells and microarray
Fourier and Pearson Analysis
Cell Cycle Regulated Genes
G1/S genes (n=362)
Cln3 Inducible Genes
Cln3 Induction and microarray
Arrested (cln1,2,3Δ GAL1pr-CLN3) cells
C
% of cells rebudded
Time from release of mitotic block (min)
D
SEM = 0.7 min
σ = 1.9 min
Normalized Expression Level (with offset)
CLN2
Time (min)
E
CLN1
CLN2
RNR1
HTA1
HTB2
RAD27
NRM1
Normalized Expression Level (with offset)
Time (min)
F
R² = 0.59
NRM1
CLN2
CLN1
Dataset 2- Activation Time (min)
Dataset 1- Activation Time (min)
G
CLN2
NRM1
CLN1
Number of Genes
Mean Activation Time (min)
H
Cumulative probability
CLN2
CLN1
NRM1
Mean Activation Time (min)

Figure 2-1: Positive feedback precedes genome-wide change in transcription at G1/S in *S. cerevisiae*. (A) Schematic diagram of the G1/S transition. (B) The G1/S regulon is defined as the intersection of the set of cell cycle regulated genes with the set of Cln3-inducible genes. (C) Synchrony of *cdc20Δ GALLpr-CDC20* metaphase block-release from Di Talia et al (2009). (D) A novel algorithm is applied to a smoothing-spline fit to detect activation of *CLN2* transcription in 7 mitotic block-release datasets. The standard deviation σ and the standard error of the mean (SEM) are calculated for each gene. (E) 7 genes in the G1/S regulon are activated at different times; data shown from a single dataset. The vertical and horizontal bars indicate the activation time and its SEM respectively. (F) Gene activation time correlation between two of the 7 datasets ($R^2 = 0.59$). Histogram (G) and corresponding cumulative distribution (H) of mean activation times for the 7 mitotic block-release datasets. *CLN1* and *CLN2*, two genes responsible for positive feedback, are among the earliest-activated genes. *NRM1*, a negative regulator of MBF, is activated later.

We analyzed 7 previously published microarray time-course datasets with 5-minute temporal resolution (Di Talia, Wang et al. 2009). All experiments were performed on *cdc20Δ GALLpr-CDC20* cells that were synchronized by mitotic arrest. Cells were released by switching the media containing galactose resulting in *CDC20* expression and a synchronous first cell cycle (Figure 2-1C).

Figure 2-1D shows the activation times for 7 independent *CLN2* expression profiles and their standard deviation and standard error of the mean. Because we have multiple time-courses, our error in estimating the activation time is low, *e.g.*, for *CLN2* we find the activation time to be 13 minutes after galactose addition with a standard deviation of 1.9 minutes and a standard error of the mean of 0.7 min. For genes within the G1/S regulon, we find that the average standard deviation is 4.7 min and the average standard error of the mean is 2.1 min. Despite regulation by the same transcription factors, the activation times of G1/S regulon members has a defined distribution (mean = 17.2 minutes; standard deviation = 5.9 minutes; Figure 2-1D-E, Figure 1-2D; Table 2-1: *Correlation coefficients (upper-right triangle, beige) and p-values (lower-left triangle, purple) for all pair-wise comparisons of the distribution of activation times for the 7 mitotic block-release experiments*.).

To test our model that feedback activation precedes regulon induction, we averaged the activation times from all 7 datasets for each gene (Figure 2-1G-H). These results were consistent with induction times measured in rtPCR time-courses (Figure 2-2). The positive feedback genes *CLN1* and *CLN2* are activated significantly earlier than the bulk

of the G1/S regulon. Indeed, within error, *CLN1* is the earliest activated gene, 5 minutes earlier than *CLN2*, suggesting a different temporal role even though these two genes are generally thought to be functionally redundant. However, it has been shown that *CLN1*, but not *CLN2*, transcription affects cell size (Flick, Chapman-Shimshoni et al. 1998), which our data suggests is due to timing. We note that for the feedback-first model to work it is sufficient to express either G1 cyclin, not necessarily both, prior to the majority of the regulon. Thus, we see that induction of the G1 cyclin positive feedback loop, which coincides with cell cycle commitment, precedes large-scale changes in the transcriptional program.

Interestingly, *NRM1*, the negative feedback element responsible for inactivating MBF regulated genes (de Bruin, Kalashnikova et al. 2006), is activated 15 min later than *CLN1* (Figure 2-1G-H) even though both genes are MBF targets (Ferrezuelo, Colomina et al. 2010). Thus, distinct temporal regulation allows positive feedback sufficient time for regulon transcription prior to *NRM1*-dependent inactivation.

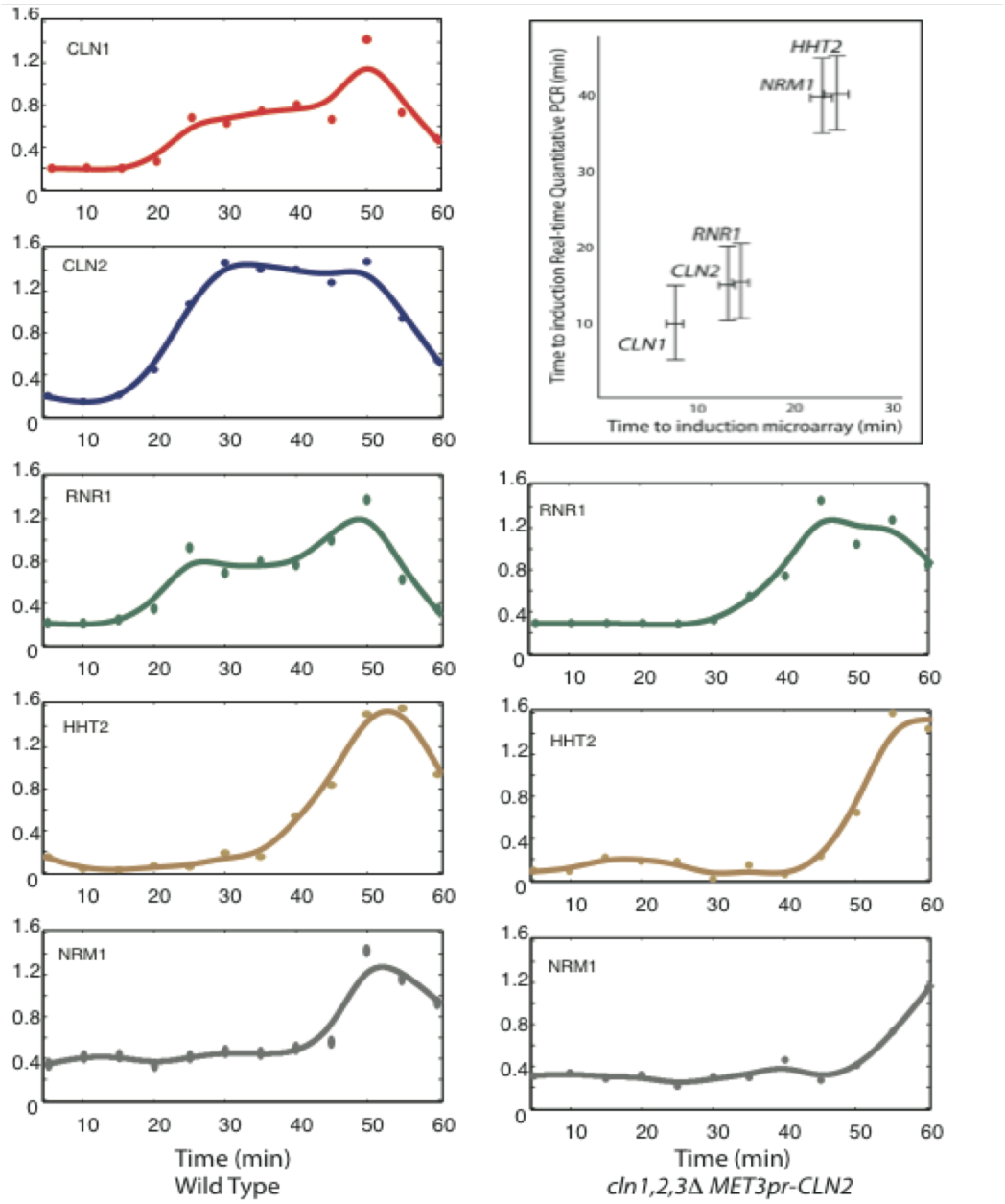


Figure 2-2: Pheromone and G1-cyclin block-release experiments. Gene expression measured using real time PCR corroborates our conclusions about gene activation timing from microarray analysis. The 3 curves on the RHS are from a cln1Δcln2Δcln3Δ MET3pr-CLN2 G1 block-release time course. Plot in upper RHS corresponds to the pheromone block experiment.

ACTIVATION TIMES OF G1/S REGULON ARE REPRODUCIBLE

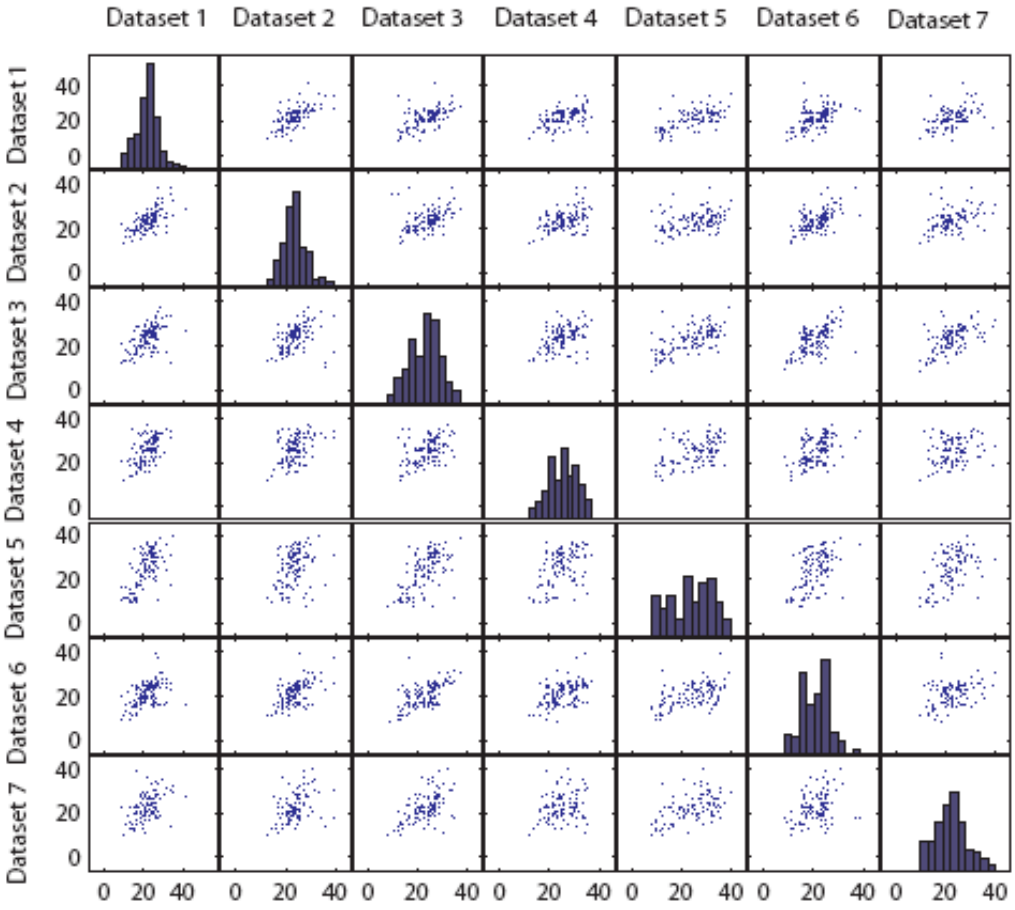


Figure 2-3: *The scatter plots of each dataset pair are shown in a matrix. Histograms of activation times of each dataset are plotted diagonal. Dataset1: WT, Dataset2: ash1Δ,Dataset3: ace2Δ, Dataset4: ash1Δ ace2Δ, Dataset5: ASH1*, Dataset6: ACE2-G128E, Dataset7: ASH1* ACE2-G128E*

We applied our algorithm to the 7 time-course microarray data published by Di Talia et al. (2009). Correlation coefficients, $\boldsymbol{R}$, are calculated for all pair-wise comparisons by using the following equation:

$$R(X,Y) = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

where $\boldsymbol{E}$ is the expected value operator and $\boldsymbol{\mu}$ is the mean value, i.e. $\boldsymbol{\mu_X = EX}$, and $\boldsymbol{\sigma}$ is the standard deviation. The correlation coefficients are shown in the upper right triangle in the Table 2-1. The p-values, calculated using *kstest*, indicate that the activation times of the G1/S regulon members are consistent among all datasets (Table 2-1).

Table 2-1: *Correlation coefficients (upper-right triangle, beige) and p-values (lower-left triangle, purple) for all pair-wise comparisons of the distribution of activation times for the 7 mitotic block-release experiments.*

	Dataset 1	**Dataset 2**	**Dataset 3**	**Dataset 4**	**Dataset 5**	**Dataset 6**	**Dataset 7**
Dataset 1		0.72	0.6	0.59	0.62	0.69	0.56
Dataset 2	2.90E-19		0.72	0.57	0.48	0.59	0.51
Dataset 3	2.30E-11	6.64E-19		0.46	0.55	0.76	0.61
Dataset 4	1.50E-10	1.69E-10	9.62E-07		0.40	0.45	0.21
Dataset 5	5.03E-10	3.77E-06	8.26E-08	4.00E-04		0.57	0.35
Dataset 6	5.89E-16	9.22E-12	1.00E-20	2.13E-06	2.08E-08		0.64
Dataset 7	4.02E-09	5.86E-08	2.78E-11	3.00E-02	1.80E-03	4.62E-12	

DELAYED POSITIVE FEEDBACK DOES NOT RESCUE *CLN1Δ CLN2Δ* CELLS

To examine the functional consequences of feedback timing, we integrated a CLN2 allele regulated by the NRM1 promoter into a cln1Δ cln2Δ cell containing MET3pr-CLN2, CLN2pr-GFPpest and RAD27-mCherry. Cells were grown overnight on media lacking methionine (MET3pr-CLN2 on) prior to switching to media containing methionine (MET3pr-CLN2 off) for single-cell analysis of one cell cycle (Skotheim, Di Talia et al. 2008). Cells exhibited similarly incoherent gene expression (time between CLN2pr and RAD27pr induction) and cell size defect as cln1Δ cln2Δ cells (Figure 2-4A-B; Figure

2-5). However, the fitness defect was partially reduced (Figure 2-4C). This indicates the importance of running the positive feedback loop from an early activated promoter.

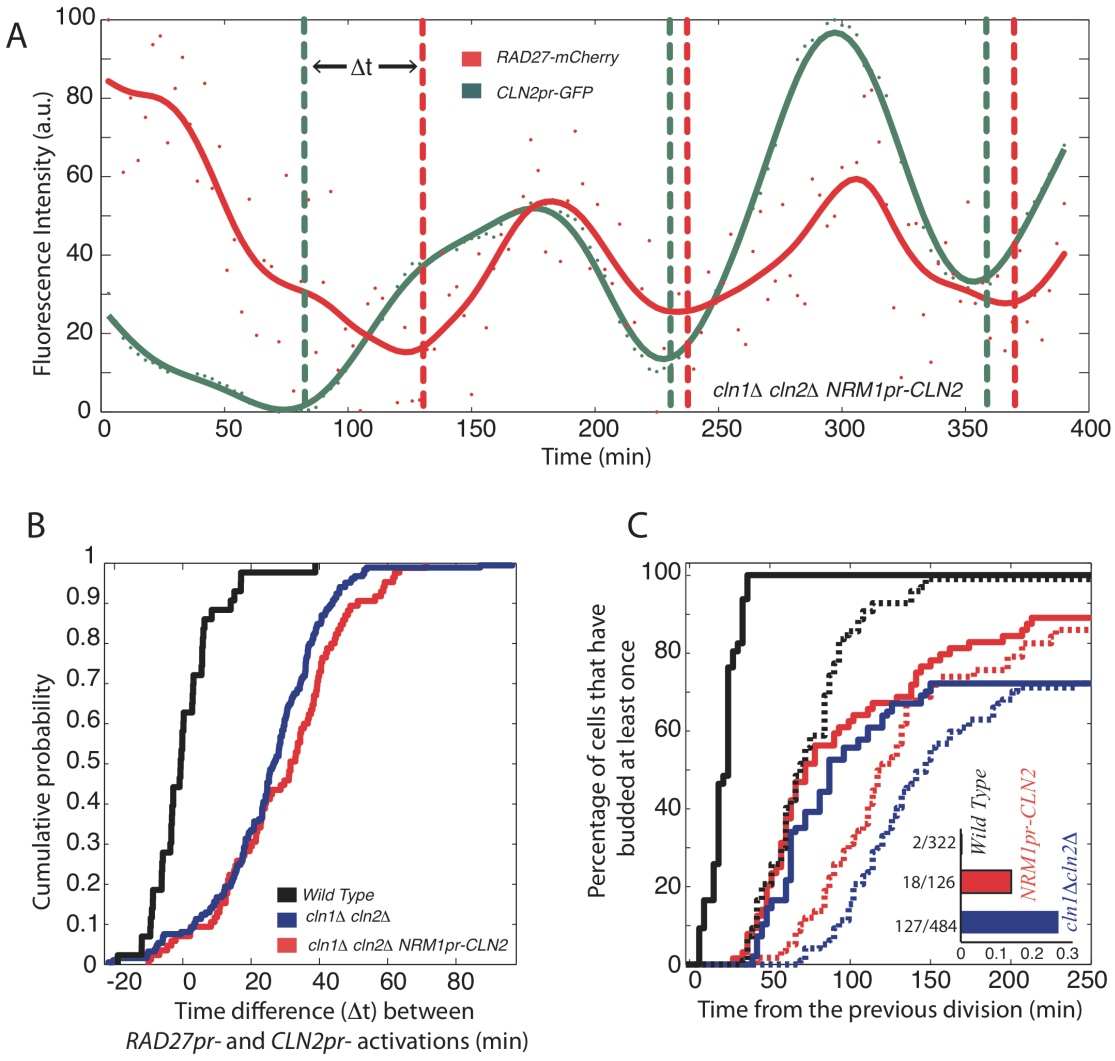


Figure 2-4: Phenotypic consequences of delayed positive feedback. (A) Time course of incoherent *RAD27-mCherry* and *CLN2pr-GFP* expression in a single *cln1Δ cln2Δ NRM1pr-CLN2* cell. (B) Time difference between *CLN2pr* and *RAD27pr* induction measured as in Skotheim et al (2008); cells not showing significant induction of either promoter were omitted from the analysis. (C) A cumulative plot for the first bud emergence measured from cell division. Solid and dashed lines correspond to mother and daughter cells respectively. Inset shows fraction of G1-arrested cells.

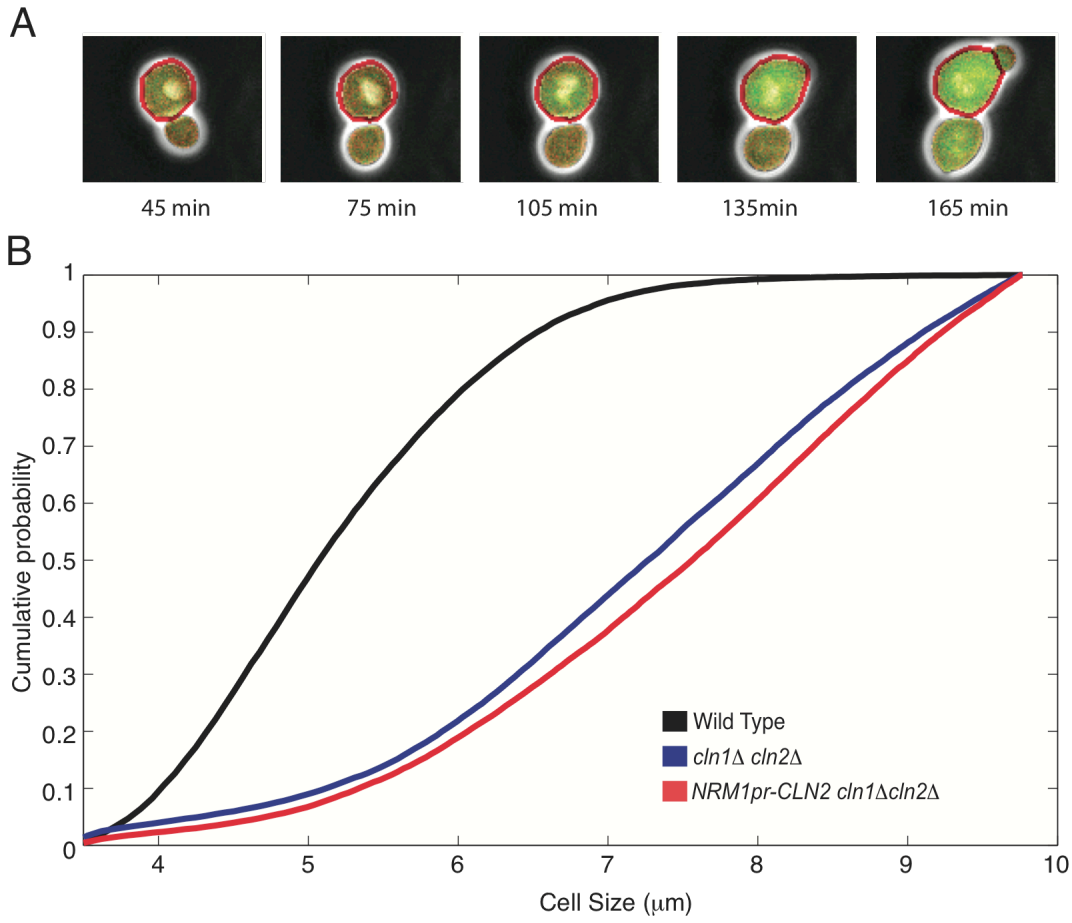


Figure 2-5: A) Time-lapse microscopy images of the cells having genotype *NRM1pr-CLN2 cln1Δ cln2Δ* are shown. By using an image segmentation program, the cells are automatically detected (red contour). The promoter activations are monitored by fluorescence intensities (GFP for *CLN2* and mCherry for RAD27). B) Delayed positive-feedback cells (*cln1Δ cln2Δ NRM1pr-CLN2*) have a similar cell size phenotype to cln1Δ cln2Δ cells. Cells are grown in log phase and cell size is measured using a Coulter Counter.

FEEDBACK-FIRST REGULATION IS ROBUST TO CHANGES IN CARBON SOURCE AND SYNCHRONIZATION METHOD

To further test our feedback-first model, we examined the effects of varying carbon source and synchronization method, which are both known to affect gene expression (Flick, Chapman-Shimshoni et al. 1998; Levy, Ihmels et al. 2007; Brauer, Huttenhower et

al. 2008). We performed a micro-array time course after synchronizing cells with mating pheromone in media with either glucose or galactose. Carbon source does not have a large effect as differences in activation times were similar to experimental replicates (Figure 2-6A).

To analyze the effect of synchronization method, we examined cells lacking endogenous G1 cyclins (*cln1Δcln2Δcln3Δ*) but containing an integrated *MET3pr-CLN2* construct (see methods). Cells were arrested in G1 before being transferred to media with a low level of methionine to activate exogenously controlled *CLN2* transcription at physiological levels. We then compared activation times between the cyclin blocked and the pheromone blocked cells (Figure 2-6B). Our three G1 block-release experiments varying carbon source and synchronization method produced similar timing profiles.

We examined the distribution of activation times pooled from the 3 separate G1 block experiments (Figure 2-6C). Although transcriptional order is affected by the arrest phase (Figure 2-6D, Figure 2-7), *CLN1* is activated at the first possible time-point (5 minutes after release) in agreement with the feedback-first model.

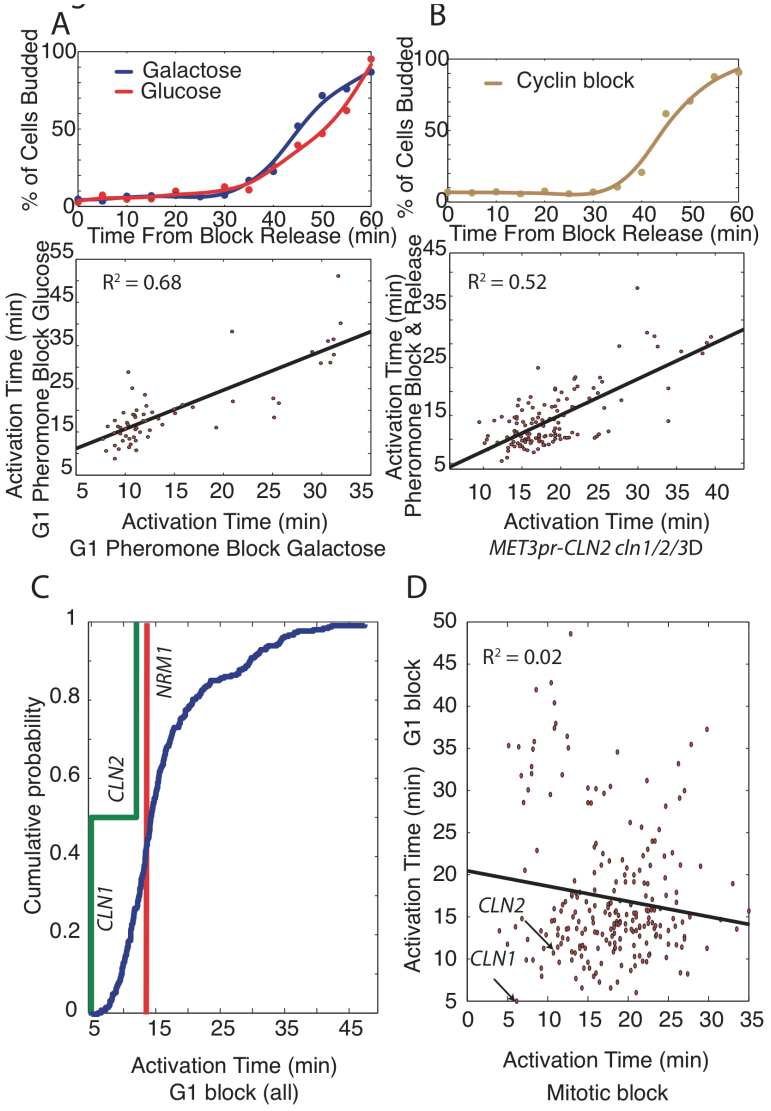


Figure 2-6: Synchronization phase, but not carbon source or synchronization method, affects gene activation timing. (A) Bud-index measurements and gene activation time correlation for G1 pheromone block-release time-course microarray experiments with glucose or galactose carbon sources. (B) Bud index for G1 block-release using *cln1Δ cln2Δ cln3Δ MET3pr-CLN2* cells and correlation of gene activation times for pheromone and G1 cyclin block-release experiments. (C) Significant correlation between the 3 G1 block-release datasets allows them to be pooled together to produce a histogram of activation times for the G1/S regulon again demonstrating feedback-first regulation. (D) Activation times from G1 and mitotic block-release experiments are not correlated.

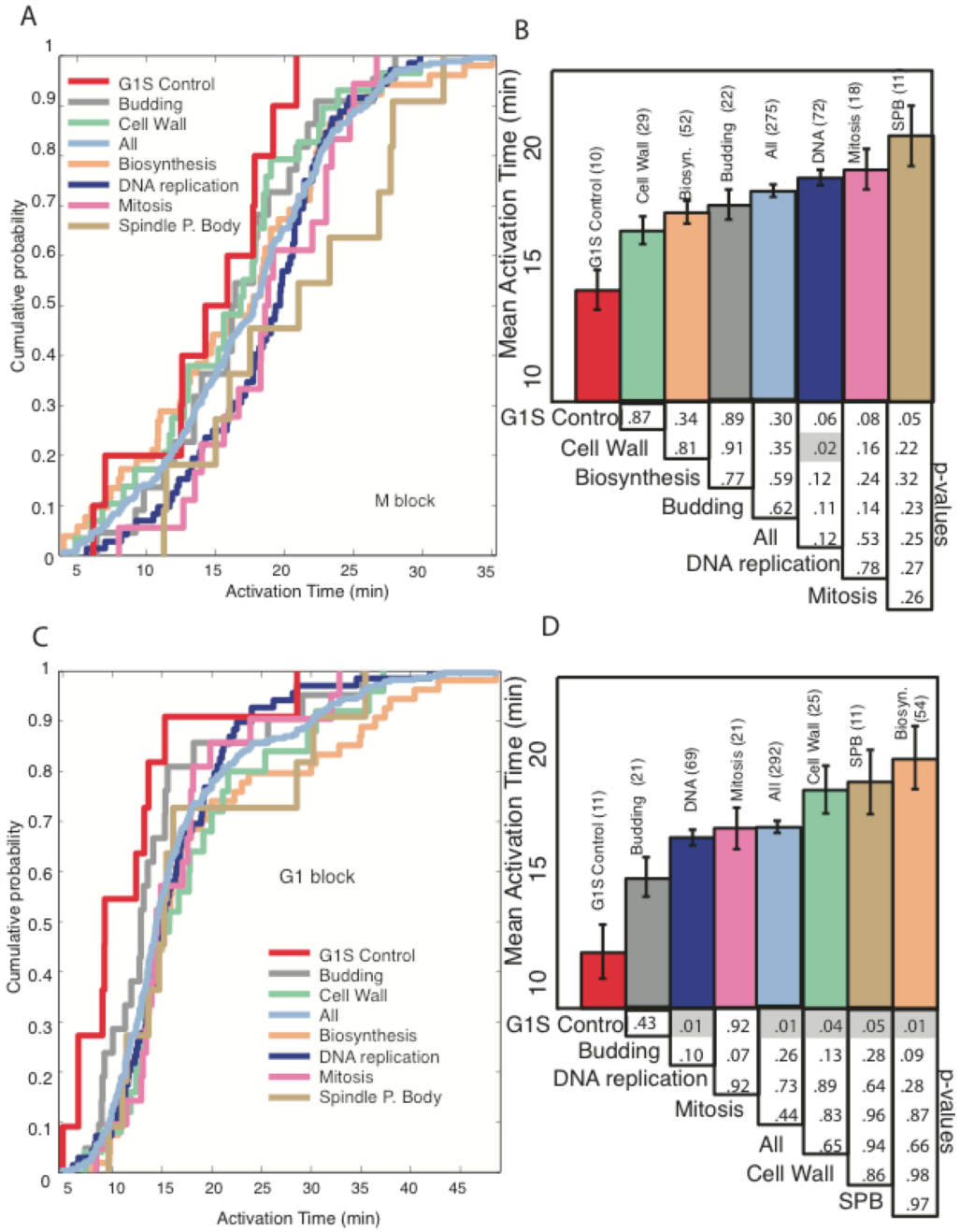


Figure 2-7: Activation time analysis of functional categories of G1/S regulon in Mitotic and G1 block-release microarray experiments, respectively. Cumulative probability of activation (A), (C) and mean activation time (B), (D) of each functional subgroup with standard error of the mean and the p- values of pair-wise comparisons are shown.

GENE ACTIVATION IS CORRELATED IN FREELY CYCLING CELLS AND MITOTIC BLOCK-RELEASE EXPERIMENTS

Since transcriptional order changes with the arrest phase, we decided to investigate which block is more similar to the free-running cell cycle using time-lapse fluorescence microscopy (Skotheim, Di Talia et al. 2008). We analyzed protein accumulation in 11 strains expressing C-terminal GFP fusion proteins from the endogenous loci (Ghaemmaghami, Huh et al. 2003), and two strains containing an integrated *CLN1* or *CLN2* promoter driving the expression of a destabilized $Venus_{PEST}$(Mateus and Avery 2000). We selected this group of strains to span the distribution of activation times. Automated cell segmentation allows us to analyze the fluorescent intensity change in single-cells through the cell cycle (Figure 2-8A). We detected activation timing relative to bud emergence and analyzed more than 50 cells of each strain (Figure 2-8B-C; Table 2-2). We found that the mean single-cell activation times in the unperturbed cell cycle correlated more with the mitotic block experiments ($R^2 = 0.72$; Figure 2-8D) than the G1 block experiments ($R^2 = 0.21$; Figure 4E). This result also implies that the order of mRNA transcription is largely reflected in protein accumulation. Thus, the mitotic block experiments are more representative of freely cycling cells.

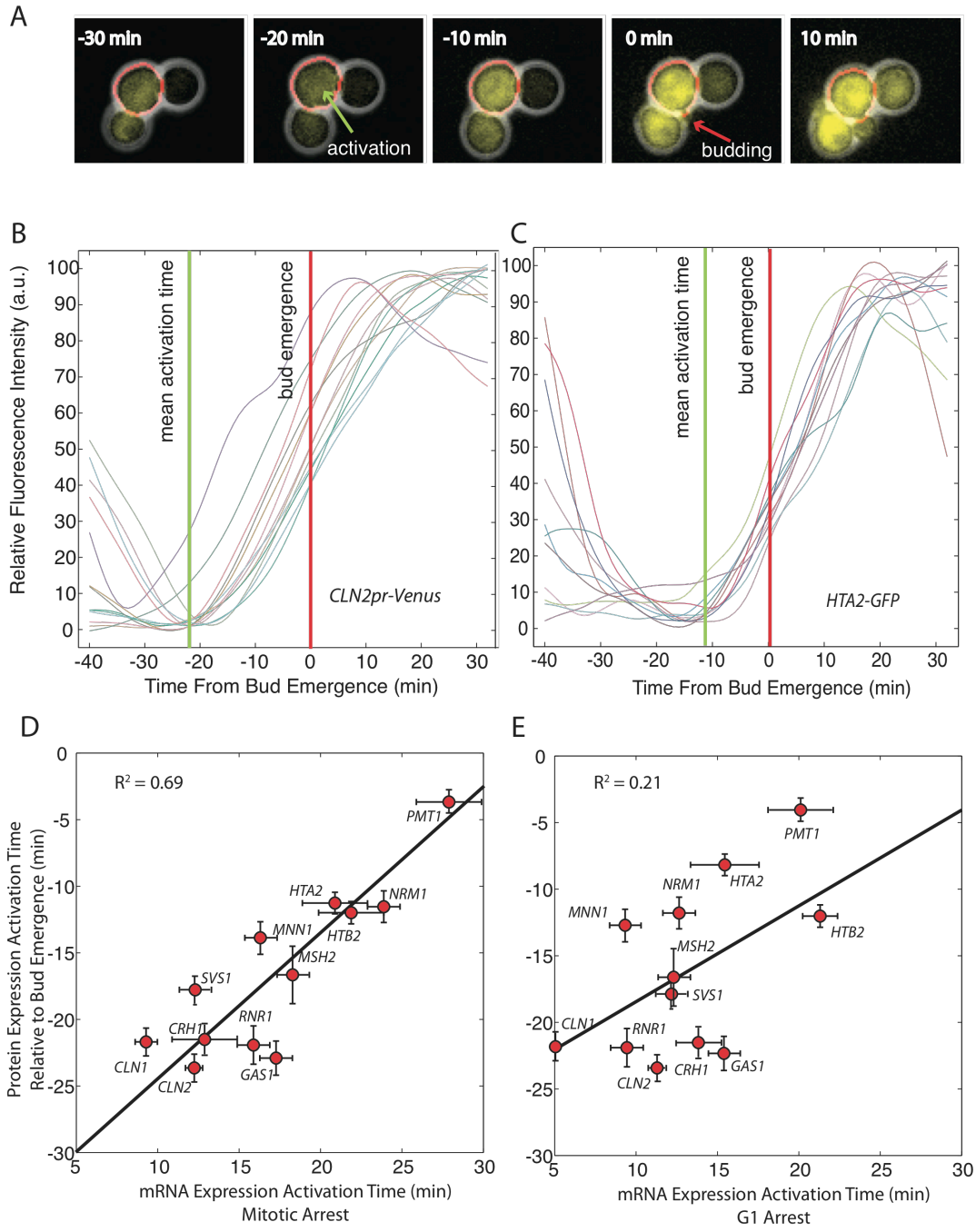


Figure 2-8: Gene activation is correlated in the free running cell cycle and mitotic block-release experiments. (A) Composite phase and fluorescence images of *CLN2pr-Venus*$_{PEST}$ cells. Venus yellow fluorescent protein contains a destabilizing PEST sequence. The red contour denotes the cell boundary detected by automatic segmentation. Gene activation time calculated from fluorescence intensity time courses aligned at bud emergence for (B) *CLN2pr-Venus*$_{PEST}$ and (C) *HTA2-GFP* cells. Gene activation

times ± SEM for 10 strains containing GFP-fused proteins and 2 strains containing promoter-Venus constructs expressed at the endogenous locus correlated with mean activation times from microarray time-courses for cells synchronized at mitosis (D) or G1 (E).

Table 2-2: Activation times (minutes) for protein accumulation are measured relative to bud emergence

Marker	Mean (min)	Standard error (min)	Standard deviation (min)	N
CRH1-GFP	-22.4	1.1	7.6	46
HTB2-GFP	-12.0	0.8	5.1	38
NRM1-GFP	-11.6	1.2	6.8	33
RNR1-GFP	-22.0	2.7	8.1	32
PMT1-GFP	-3.7	0.9	3.8	18
GAS1-GFP	-22.8	1.3	7.0	30
SVS1-GFP	-17.6	1.5	6.4	18
HTA2-GFP	-8.3	0.8	5.8	52
CLN1pr-VenusPEST	-22.5	1.1	7.0	37
CLN2pr-VenusPEST	-23.4	1.1	6.9	20
MNN1-GFP	-12.7	1.6	8.1	25

Since transcription activation times change with the phase of the block used, we decided to analyze previously published cell cycle synchronized microarray time courses (Spellman, Sherlock et al. 1998; Pramila, Wu et al. 2006; Orlando, Lin et al. 2008). Although quantitative comparisons of individual genes are difficult due to either poor temporal resolution or lack of experimental replicates, we are able to detect correlations of genes within the G1/S regulon. We found that G1 blocks, including elutriation, correlate with our G1 block data (Table 2-3). Interestingly, the $cdc15^{ts}$ data from Spellman et al (1998) correlates with our G1 block experiments rather than the mitotic block experiments even though this is an anaphase block indicating that an event occurring in cells blocked downstream of Cdc20 may be responsible for differences in gene activation timing. We note that release from G1 arrest and free cycling are both likely to be physiologically relevant.

Table 2-3: Comparison of activation time distributions are calculated using previously published datasets *(Spellman, Sherlock et al. 1998; Pramila, Wu et al. 2006; Orlando, Lin et al. 2008; Di Talia, Wang et al. 2009)*

			Distribution of gene activation times			
			G1 block-release		Mitotic block-release	
Data Source	Year	Synchronization	Correlation Coefficient	p-values	Correlation Coefficient	p-values
This work	2010	G1 block	1	0	-0.13	0.15
Di Talia et al.	2009	M block	-0.13	0.15	1	0
Pramila et al.	2006	a factor	0.41	8.44E-11	0.23	0.01
Orlando et al.	2008	Elutriation(WT)	0.48	1.47E-08	-0.02	0.87
Orlando et al.	2008	Elutriation(clbsΔ)	0.65	2.42E-12	-0.17	0.24
Spellman et al	1998	a factor	0.08	0.53	-0.4	0.02
Spellman et al.	1998	cdc15	0.36	4.6E-03	0.04	0.84
Spellman et al.	1998	cdc28	0.26	0.08	-0.39	0.11

CHAPTER3 : TRANSCRIPTIONAL "OR" GATE

SBF- AND MBF-DEPENDENT ACTIVATION IS A LOGICAL OR GATE

In the previous chapter, we showed that the activation order of G1/S genes changes in different arrest phases. We hypothesized that the observed differences in gene activation time in different blocks might be due to differential regulation of specific transcription factors. The majority of genes in what we defined as the G1/S regulon are regulated by the transcription factors SBF and MBF (Ferrezuelo, Colomina et al. 2010). For our analysis, we divided the activation times of the G1/S genes into three categories: 136 SBF-only targets, 63 MBF-only targets, and 36 dual-regulated SBF and MBF targets.

Since combinatorial use of transcription factors may yield differential activation timing, we analyzed the activation times of the SBF only, MBF only, and dual-regulated genes. For our G1 arrest data, we find that MBF-only targets are activated earlier than SBF-only targets ($p < 0.01$). Furthermore, the distribution of the dual regulated targets is more similar to the earlier-activated MBF-only targets ($p = 0.90$) than the more tardy SBF-only targets ($p = 0.01$; Figure 3-1A).

In the mitotic block-release, the SBF-only targets are activated earlier than the MBF-only targets ($p = 0.08$). This is the opposite order than in the G1-block experiments and consistent with the lack of correlation between activation times of individual G1/S regulon members (Figure 2-6D). Furthermore, we find that the common targets are much more likely to follow the SBF-only distribution ($p = 0.79$) than the MBF-only distribution ($p = 0.06$; Figure 3-1B). We note that the SBF distribution is broader so that the late-activated SBF genes are activated later than the late-activated MBF genes. However, the late-activated dual-regulated genes now appear to follow MBF.

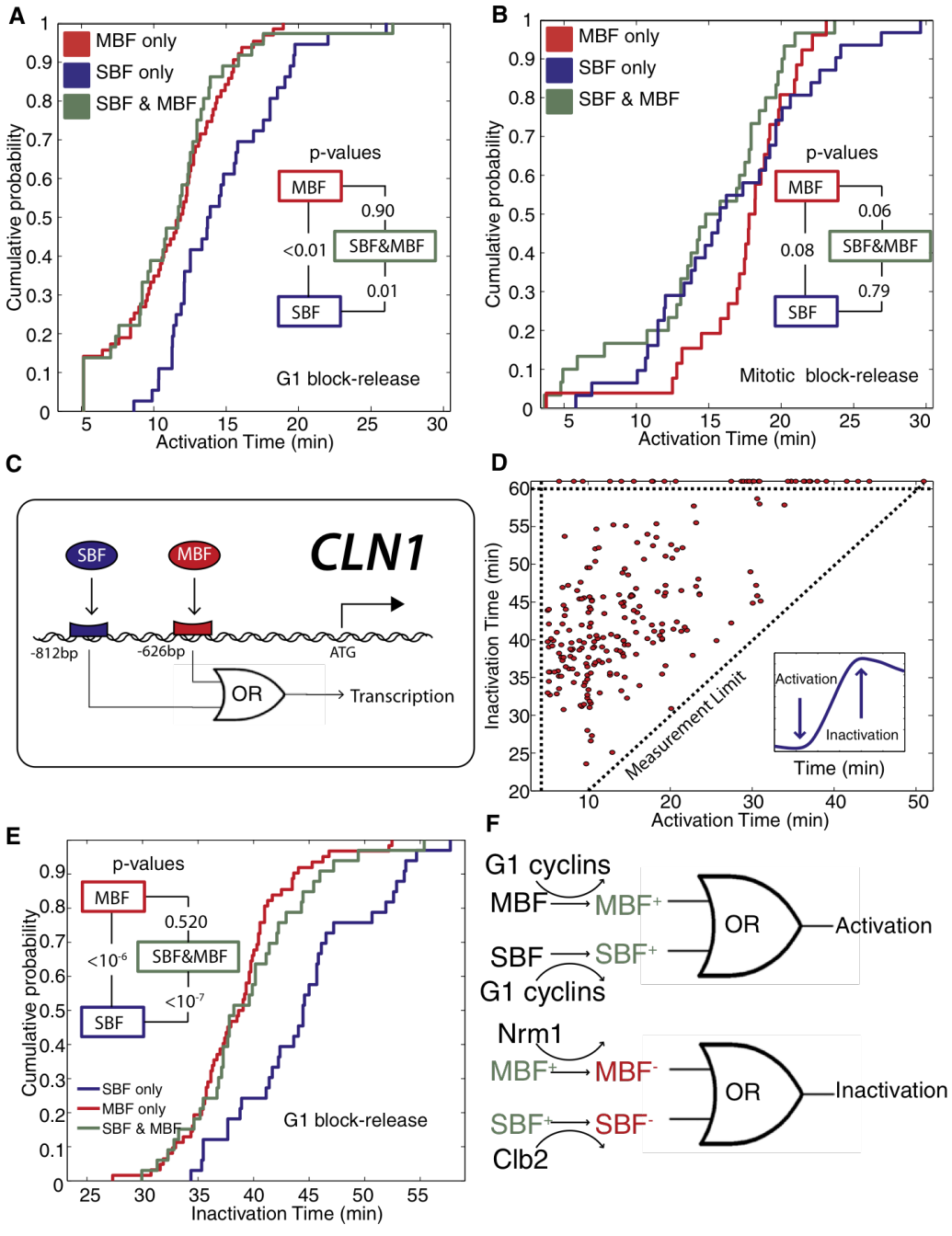


Figure 3-1: SBF and MBF dual-regulated promoters act as logical OR gates in response to activation and inactivation signals. Cumulative probability of activation times for SBF-only, MBF-only and SBF/MBF dual-regulated targets are plotted for (A) G1 block-release and (B) mitotic block-release experiments. Inset

shows p-values comparing each pair of distributions. (C) Schematic showing logical regulation of the early-activated *CLN1* promoter denoting SBF and MBF consensus binding sites. (D) Inactivation time for each gene, where the 1st derivative is zero and the 2nd derivative is negative (inset), is uncorrelated with activation for G1 block-release experiments. Points above the horizontal dotted line represent genes peaking later than 60 min. (E) Cumulative probability of inactivation for SBF-only, MBF-only and SBF/MBF dual targets for G1 block-release experiments. Inset shows p-values comparing each pair of distributions. (F) The transcriptional activation and inactivation can be modeled as a logical OR gate. For dual-regulated genes, activating either SBF or MBF suffices for activation, while inactivating MBF suffices for inactivation. Different colors denote different possible states of a transcription factor.

Taken together, our results from the two different types of experiments suggest that the dual-regulated targets are activated by the earliest active transcription factor. In the G1 block experiments, the co-regulated genes are activated by MBF, while in the mitotic block experiments the co-regulated genes are activated by SBF. This implies that transcriptional activation is functioning as a logical OR gate, where either an active SBF or an active MBF is sufficient to activate transcription.

LOGICAL INACTIVATION

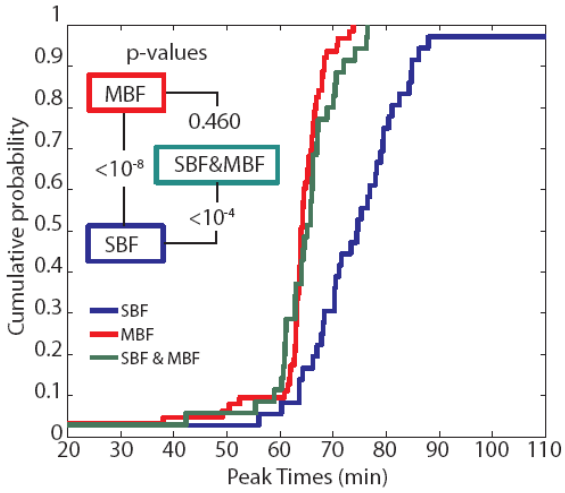
Our results analyzing transcriptional activation encouraged us to perform a similar analysis on transcriptional inactivation, which we estimate as the time of the peak transcript level (Figure 3-1C). The peak time is defined to be the point where the 1st derivative of the smoothed data is zero and the 2nd derivative is negative. We then implemented an algorithm for unbiased peak detection and analyzed our G1 block-release data. Inactivation is not well correlated with activation (Figure 3-1D).

Next, we decided to analyze inactivation in light of our SBF-only, MBF-only and dual regulated gene lists. Whereas mitotic cyclins are responsible for SBF inactivation (Amon, Tyers et al. 1993), MBF inactivation is performed by Nrm1 possibly through a direct interaction (de Bruin, Kalashnikova et al. 2006). In *nrm1Δ* cells, mitotic cyclins are capable of inactivating MBF-regulated genes; however, inactivation is delayed about 10 minutes relative to WT (de Bruin, Kalashnikova et al. 2006). This suggests that mitotic cyclin-dependent inactivation occurs later than Nrm1-dependent inactivation and that we

should expect to see MBF-only targets inactivated earlier than SBF-only targets. Consistent with previous results (Ferrezuelo, Colomina et al. 2010), we find that MBF-only targets are inactivated earlier than SBF-only targets ($p < 10^{-7}$; Figure 3-1E, Figure 3-2). The distribution of inactivation times for the dual regulated genes was much more similar to the MBF-only genes ($p=0.52$) than the SBF-only genes ($p < 10^{-7}$). Inactivation of MBF is sufficient to turn off gene expression regardless of the presence of an active SBF transcription factor. Thus, both activation and inactivation may be represented by logical OR gates (Figure 3-1F).

Inactivation Times of SBF and MBF Specific Targets

To further investigate the logic of transcriptional inactivation, we analyzed the previously published time-course microarray data where the cells were synchronized via elutriation (Orlando, Lin et al. 2008). We found the result consistent with the analysis in the main text. The SBF and MBF dual regulated targets have the same inactivation time distribution as the MBF-only targets, supporting the Boolean OR-gate model of



transcriptional inactivation.

Figure 3-2: *Cumulative distribution of inactivation for SBF-only, MBF-only, and SBF/MBF dual regulated targets. Microarray data from Orlando et al. (2008).*

CHAPTER4 : EFFECT OF SPATIAL GENOME ORGANIZATION ON TRANSCRIPTION

Feedback-first regulation ensures that commitment to cell division occurs before large-scale changes in transcription and is conserved in the related yeast *S. bayanus* and human cells. In the previous chapter, we showed that timing differences within the G1/S regulon, comprised of both SBF and MBF targets, are partially explained by the combinatorial use of SBF and MBF transcription factors, which implement a logical OR function for gene activation. However, the molecular mechanisms underlying the precise temporal order at the basis of feedback-first regulation remained unclear. More specifically, we do not know what specifies the order of activation within the specific set of genes that are targets of SBF, but not MBF, and vice versa (Figure 3-1A,B). Here in this chapter, we analyze genome-wide chromosome conformation capture data to uncover a potential link between the timing of cell cycle regulated gene expression and 3-D genome architecture.

NUCLEAR COMPARTMENTALIZATION

A eukaryotic nucleus is not a homogenous bag of nucleic acids and proteins. Transcription in eukaryotes reflects this fact and comprises many layers of regulation, from *cis*-elements to higher-level chromosome organization (Misteli 2004; Fraser and Bickmore 2007; Meaburn and Misteli 2007; Babu, Janga *et al*. 2008; Takizawa, Gudla et al. 2008; Lieberman-Aiden, van Berkum et al. 2009). Such organization includes centromere clustering, telomere pairing, nuclear pore complex attachments, and nucleolus formation (Gotta, Laroche et al. 1996; Andrulis, Neiman et al. 1998; Jin, Fuchs et al. 2000; Dekker, Rippe et al. 2002; Bystricky, Laroche et al. 2005; Taddei, Van Houwe et al. 2009). All these features have been described for the budding yeast *S. cerevisiae*, the subject of our study (Figure 4-1).

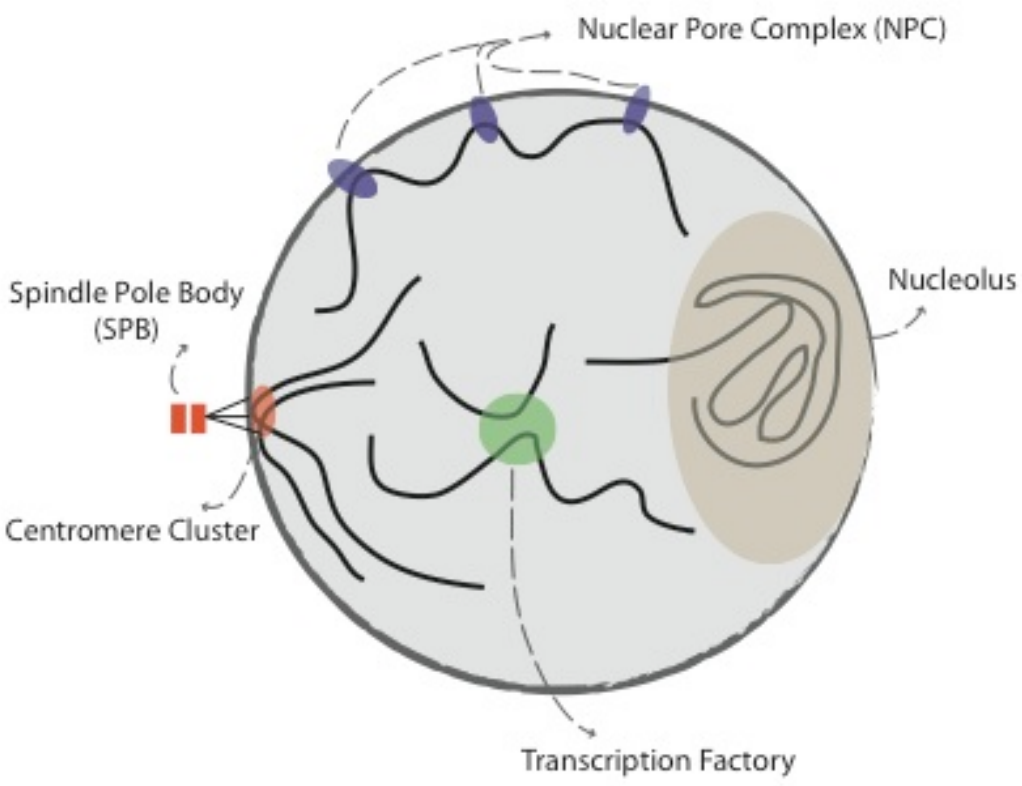


Figure 4-1: Illustration of distinct compartments within the yeast nucleus.

Despite the significant evidence for yeast nuclear architecture, the model of chromosome territories, which states that the chromosomes occupy distinct regions in the nucleus, is not commonly accepted (Haber and Leung 1996; Meaburn and Misteli 2007; Berger, Cabal et al. 2008). Recently, the identification of nuclear architecture has been facilitated by high-throughput experimental methods based on Chromosome Conformation Capture (3C), called Hi-C (Duan, Andronescu et al. ; Dekker, Rippe et al. 2002; Simonis, Klous et al. 2006; Zhao, Tavoosidana et al. 2006; Lieberman-Aiden, van Berkum et al. 2009). The common procedure of these experiments includes cross-linking the DNA fragments, which happen to be near one another, digesting with a restriction enzyme, circularizing,

processing and sequencing. If a circularized piece of DNA contains fragments from two different parts of the genome (Figure 4-2). This technique thus gives something akin to a frequency distribution of the physical interactions within the genome.

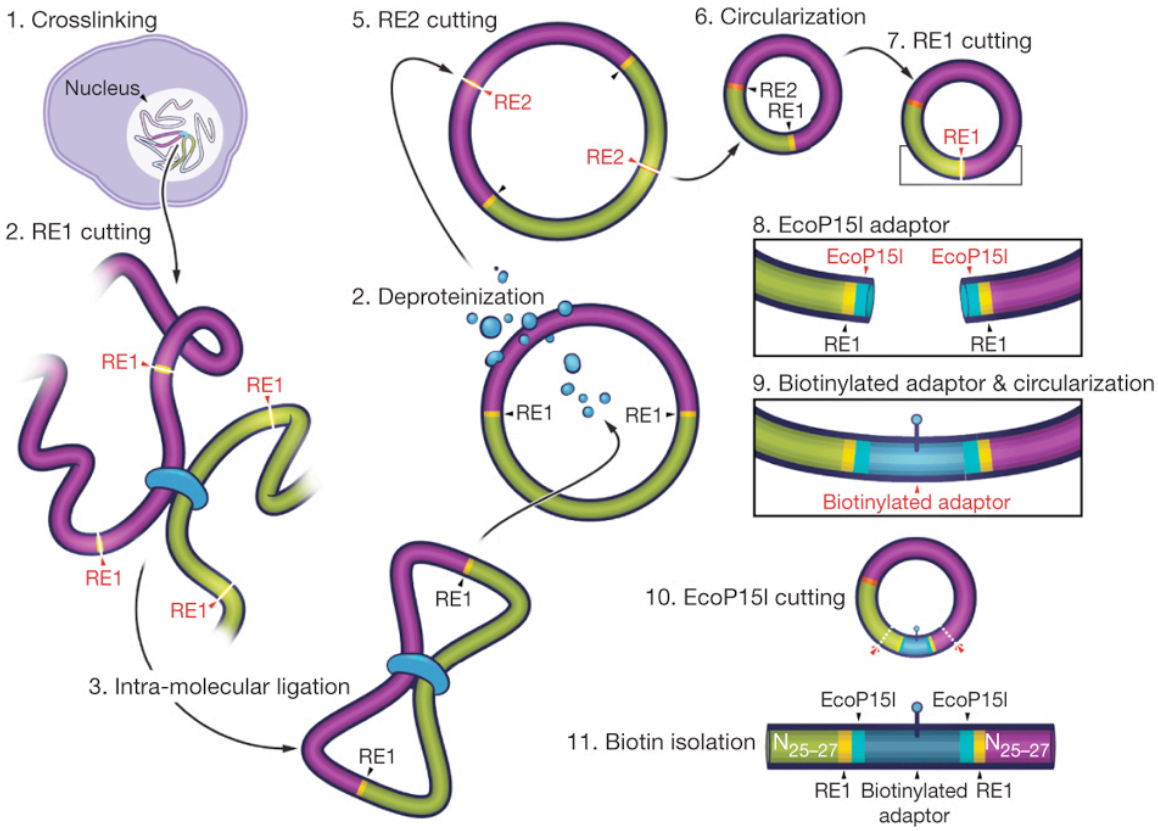


Figure 4-2: Schematic of genome-wide chromatin conformation capture methods (adapted from Duan et al. 2010).

Here, in this chapter we analyzed Hi-C data published by Duan et al. to investigate the possible relationship between sublocalization of genes and transcription. We were driven to test the hypothesis that if compartmentalization affects the kinetics of cell cycle-dependent transcriptional activation, then we should observe genes activated with similar timing to be colocalized in space, which predicts a higher frequency of interactions in a Hi-C data set.

GENES WITH SIMILAR ACTIVATION TIMING COLOCALIZE

The Hi-C data contain interactions among pairs of fragment reads that are unevenly distributed along the genome with ~1 kbp resolution. We consider two genes as interacting if there is an interaction between at least one pair of fragments, which lies within 2.5 kbp at TSS of these genes as suggested previously (Dai and Dai 2012). The number of all possible pair-wise interactions among N genes is given by $M = \binom{N}{2}$. If X is the number of all distinct interactions observed, the ratio X/M gives us the interaction enrichment or the colocalization score for a given set of N genes. To check if the interaction enrichment of this gene set is significantly higher than that of randomly chosen N genes, we bootstrapped our data by re-sampling N randomly picked genes 1000 times to produce a smooth distribution of colocalization scores to comprise a null model. The percentile of the selected set of genes denotes the significance level of colocalization. Therefore, the p-value is calculated by the percentage of random cases, which have higher colocalization score than the genes of interest we are examining (Figure 4-3).

To test the hypothesis that genes, which are activated at similar times share a distinct region of the nucleus, we first sorted the G1/S genes based on their activation times. Then we calculated the colocalization score and the p-value for a group of N adjacent genes sliding over all genes one-by-one. In other words, a window with a size N slides through the gene list and produces the colocalization score for each group. We calculated the results for groups of N = 5, 10, 15, 20 and 25 genes. We defined the colocalization coverage as the number of counts that a gene is considered as colocalized with the adjacent genes in the list by a p-value<0.05. We see significant colocalization of genes that are activated early and late in the G1/S regulon consistent with the model that genes of similar activation timing reside in similar regions of the nucleus (Figure 4-4).

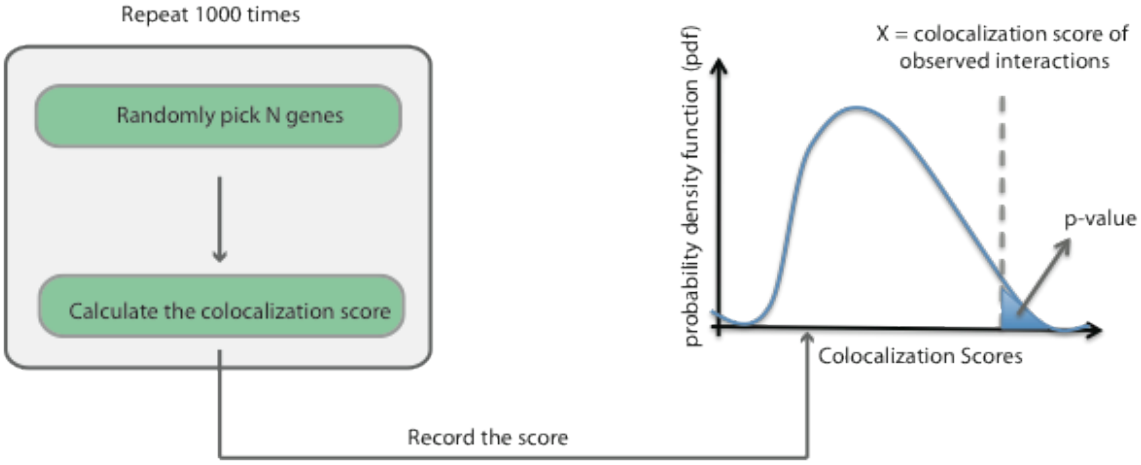


Figure 4-3: The bootstrap procedure to calculate the significance of colocalization of a selected group of genes.

Interestingly, spatial clustering of these genes does not obviously correlate with the type of TF (SBF or MBF) suggesting that the nuclear architecture is common to both factors. It has previously been shown in metazoan cells that transcription enzymes cluster spatially into regions denoted as transcription factories, which may regulate the transcription of genes differentially (Casolari, Brown et al. 2004; Janga, Collado-Vides et al. 2008). Our work here is consistent with the model of transcription factories, which have not, to our knowledge, been shown previously to apply to yeast cells.

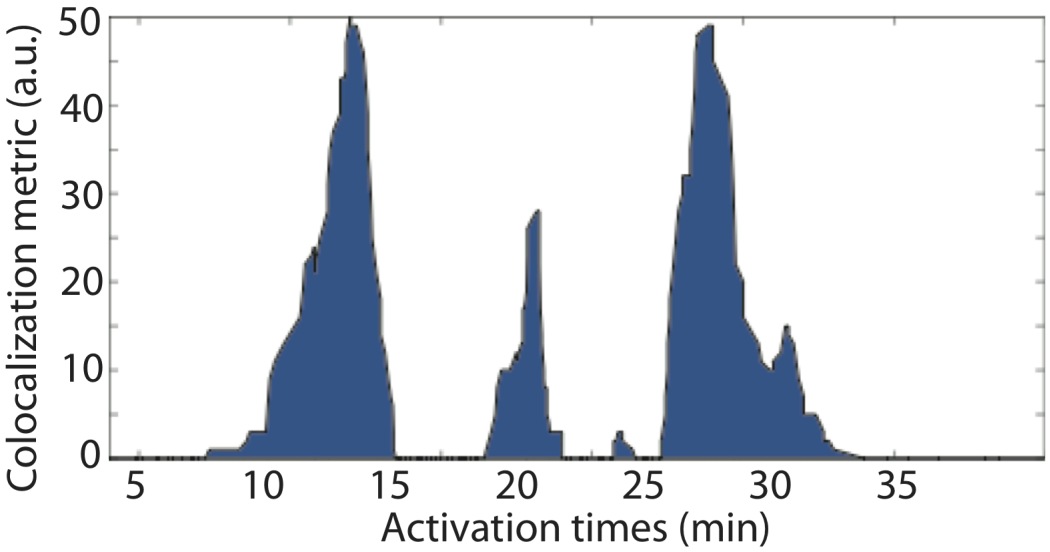


Figure 4-4: Colocalization of the genes activated at similar times. Genes are arranged by activation time. We examine the degree of colocalization of a sliding window of varying number of genes (see text for details) indicating that early- and late-activated G1/S regulon members cluster in space.

DISTINCT ACTIVATION TIME DISTRIBUTIONS BETWEEN GENE CLUSTERS

So far, we have seen that the early and the late activated genes have enriched interactions among their groups suggesting a correlation between localization and transcriptional activation. Another way to examine this relationship is to compare the activation time distributions of the genes in different clusters in the nucleus. To determine the gene clusters in the nucleus, distance information between any two genes is needed. Therefore, similar to Duan et al (2010), we defined a distance metric by assigning the nearest upstream and downstream fragments to the start codon of each gene, then calculated the total interaction between all possible pair of genes in the G1/S regulon (Figure 4-5A). Hence, the metric is defined by subtracting the total interactions between four fragments assigned to two genes from the maximum possible interactions, *i.e.*, 4. In other words, the distance metric has the score 0 for the closest pair of genes, which have 4 interactions, and 4 for the most distant ones, which have no interaction. Figure 4-5B shows the

heatmap representation of distance metric of the genes in G1/S regulon, which are sorted alphabetically. Next, we used an agglomerative hierarchical clustering algorithm with average linkage clustering criterion and applied a threshold (cophenetic distance>20) to limit the dendrogram to 6 clusters (Figure 4-6).

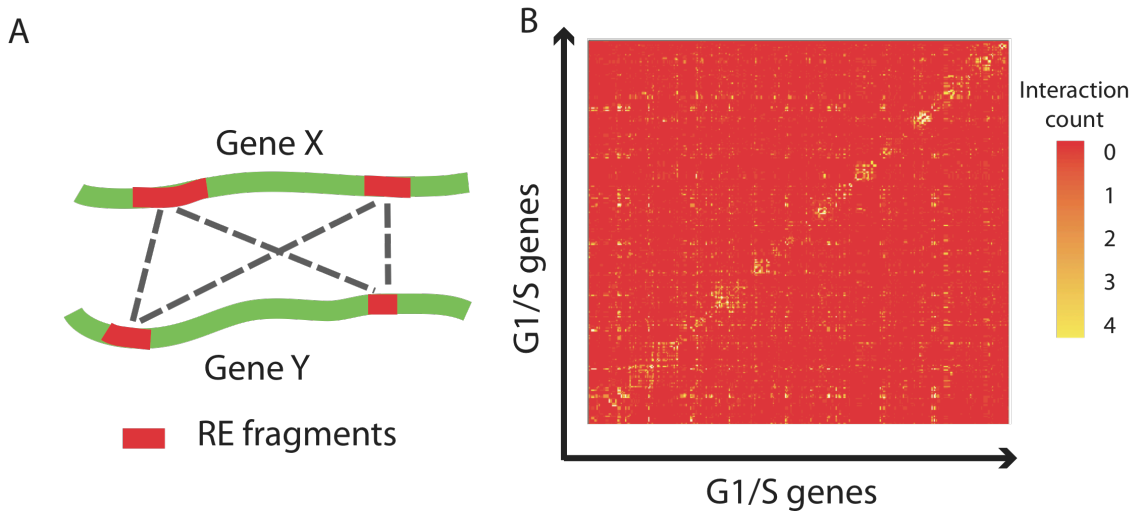


Figure 4-5: (A) Possible interactions between the assigned fragments of any two genes (Gene X and Gene Y). (B) Heatmap of a matrix showing the distance between gene pairs. Rows and columns denote the G1/S genes and they are alphabetically sorted.

Interestingly, the genes in two clusters (clusters 1 and 2) show significantly different activation time distributions with the mean 12.2 (± 1.7) min. and 16.6 (± 1.4) min., respectively ($p<0.01$; Figure 4-7). Different clusters with different activation time distributions underlines the heterogeneity effect of nuclear organization. But, what, physically, do these functional clusters indicate? We hypothesize, that distinct clusters correspond to distinct regions of the nucleus that may be composed of different chromatin distributions.

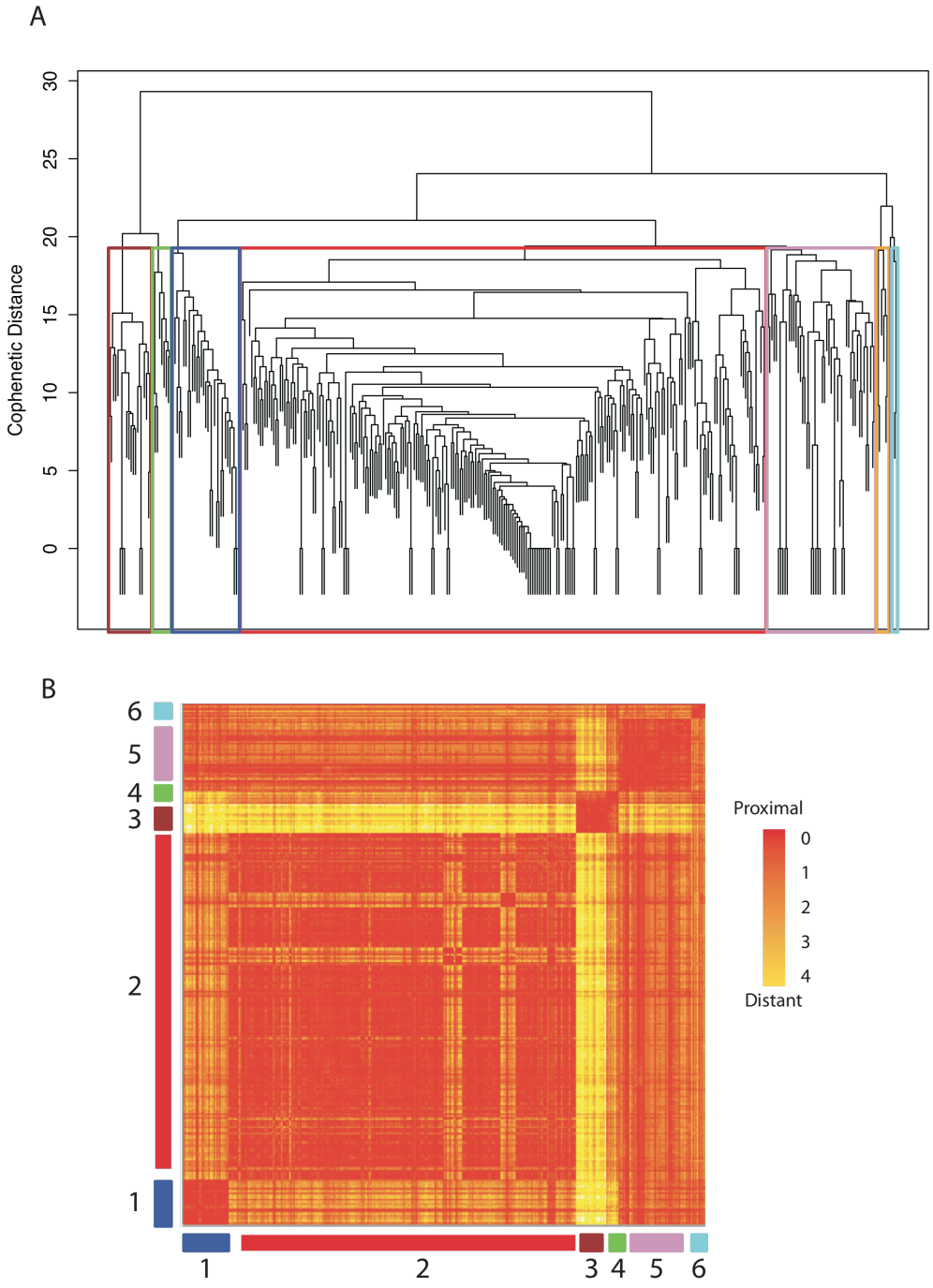

A
Cophenetic Distance
0
5
10
15
20
25
30
B
6
5
4
3
2
1
1
2
3
4
5
6
Proximal
0
1
2
3
4
Distant

Figure 4-6: Average-linkage hierarchical clustering of G1/S genes. Color-code indicating clusters remains same for the clusters in both (A) the denrogram, and (B) the heatmap.

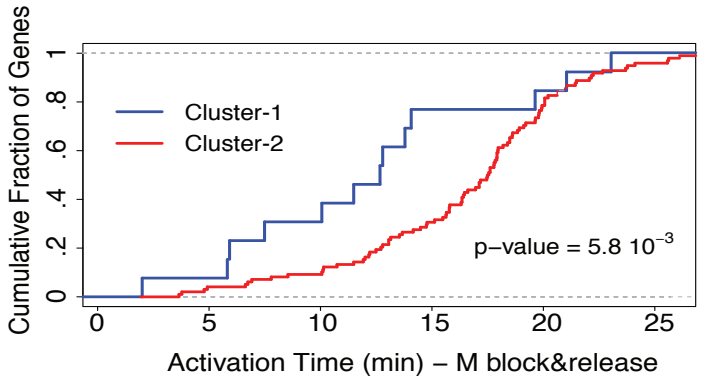


Figure 4-7: Activation time distributions of clusters 1 and 2 as indicated in Fig. 4-6 above.

But, what is the correspondence between a Hi-C frequency map and the spatial structure of the genome? Clearly, the Hi-C data we analyze represent the statistical ensemble of many genome configurations. Because some parts of the genome are constrained by protein complex interactions, such as within the nucleolus, nuclear pore complex, spindle pole body and centromeres, the amount of different configurations that are represented in the ensemble could be considered a local parameter. Each cell can maximally contribute only a single interaction for each fragment to the ensemble. To calculate the total amount of interactions per gene, we assign the closest fragment to each gene. Hence, we obtain the total number of interaction distributions of the clusters, which can be considered a measure of their number of different configurations. Significantly, we observe distinct distributions of total interaction amount for the early and the late clusters, *i.e.*, cluster 1 and cluster 2 (Figure 4-8A). This result suggests that the more interactions correlate with the earlier activation. Indeed, if we stratify all the G1/S genes based on a threshold of the interaction amounts (Appendix C, Figure C-1); we find a statistically more significant distinction between the activation timing distributions (Figure 4-8B).

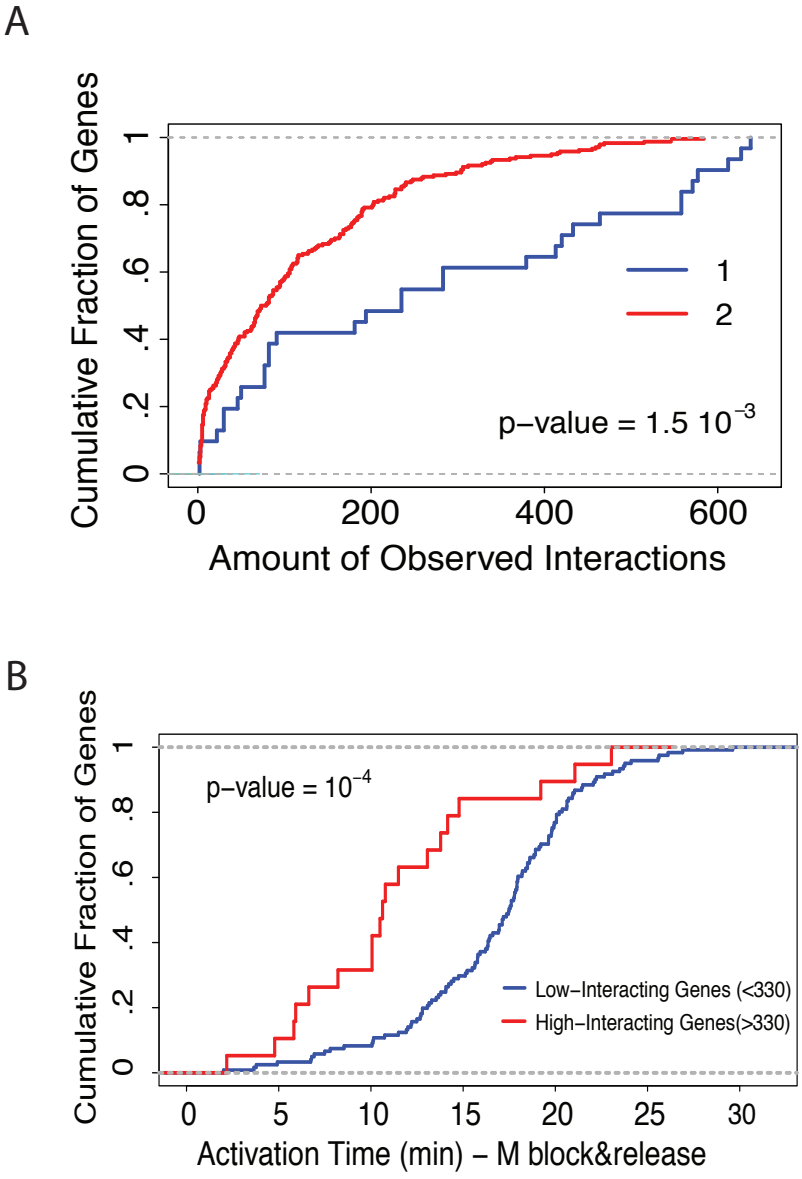


Figure 4-8: Amount of interactions observed in the data is one of the determinants of spatial clustering. (A) Distributions of interaction amounts of the genes in cluster-1 (blue) and in cluster-2 (red) are significantly different. (B) Genes interacting with more than 330 fragments have earlier activation times than that are interacting less than 330 fragments, a result robust to the choice of cutoff.

Thus, in this chapter, we have tested the hypothesis that co-localization of genes within distinct compartments within the nucleus using Hi-C data. Our analysis suggests that

nuclear compartmentalization can have a significant affect on how quickly genes within particular regions within the yeast genome can be activated. This defines a second axis, distinct from the combinatorial use of transcription factors, the cell can tune to generate the differential activation timing of cell cycle-dependent transcription that underpins feedback-first regulation. Future work will test this hypothesis by integrating a subset of SBF-regulated genes into a large set of locations within the genome and measuring activation time.

CHAPTER5 : FEEDBACK-FIRST MOTIF IN OTHER EUKARYOTES

We found that *S. cerevisiae* activates positive feedback and commits to another round of cell division before making large-scale changes to its transcriptional program. This temporal organization of the G1/S regulon may be an efficient way to ensure that cell cycle associated genes are only transcribed after a cell has decided to divide. If feedback-first regulation increases fitness then we should expect to see it conserved in divergent evolutionary lineages.

FEEDBACK-FIRST REGULATION IN THE BUDDING YEAST *S. BAYANUS*

To examine the conservation of feedback-first regulation, we analyzed a closely related yeast *Saccharomyces bayanus*, for which cell cycle synchronized microarray data was available. Compared to *S. cerevisiae*, *S. bayanus* has 67% local similarity of intergenic regions indicating significant divergence of gene regulation (Cliften, Sudarsanam et al. 2003). Gene orthologs are easily identified by sequence and the *S. bayanus* genes are conveniently annotated using the *S. cerevisiae* nomenclature (Cliften, Sudarsanam et al. 2003). Indeed, studies on the evolution of gene expression among *sensu stricto* yeast species revealed substantial differences (Tirosh, Weinberger et al. 2006; Guan, Dunham et al. 2010).

We analyzed the *S. bayanus* time-course microarray dataset from the GEO database (GSE16544). Cells were synchronized in G1 using mating pheromone and samples were taken every 10 minutes for 300 minutes following release (Guan, Dunham et al. 2010). To define a set of genes that are cell cycle regulated, we calculated the cross-correlation coefficients with two known cell cycle regulated genes, *CLN2* and the G2 gene *KIN2*. We sorted the genes based on their cross-correlation scores and selected the 714 genes that

were in the top 1000 of both cross-correlations. To eliminate spurious profiles, we considered only genes showing multiple well-defined oscillations.

We analyzed the correlation of cell cycle regulated gene expression in the two budding yeasts. Of the 800 and the 223 well-defined cell cycle regulated genes in *S. cerevisiae* and *S. bayanus* respectively, only 79 were cell cycle regulated in both species (Figure 5-1A). Furthermore, the activation times of the common cell cycle regulated genes is weakly correlated (R^2 = 0.22; Figure 5-1B). Our observation of significant changes in transcriptional activation timing through the cell cycle is consistent with the emerging picture of significantly diverged transcription across the *sensu stricto* (Tirosh, Weinberger et al. 2006).

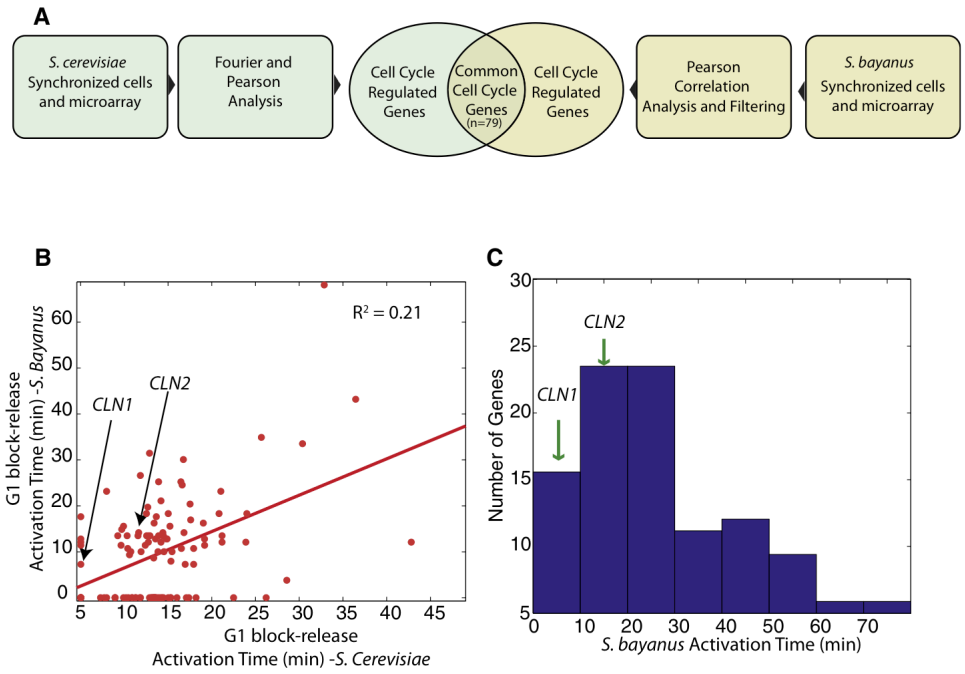


Figure 5-1: Feedback-first regulation is conserved in the budding yeast *S. bayanus*. Activation times are analyzed for all cell cycle regulated genes in a *S. bayanus* pheromone block-release microarray time-course (Guan, Dunham et al. 2010). (A) Intersection of cell cycle regulated genes in both budding yeasts. (B) Weak correlation between gene activation times in *S. bayanus* and *S. cerevisiae* for G1 block-release experiments. (C) Histogram of activation times of the cell cycle regulated genes in *S. bayanus* indicates that the G1 cyclins responsible for positive feedback, *CLN1* and *CLN2*, are among the early-activated genes

To test for the conservation of feedback-first regulation, we analyzed the distribution of first activation times (<80 min). The activation times for *CLN1* and *CLN2* was calculated to be 6 and 15 minutes respectively. Thus, the G1 cyclins are among the earliest activated genes in the *S. bayanus* cell cycle, which indicates conservation of feedback-first regulation (Figure 5-1C).

TEMPORAL ANALYSIS OF E2F-DEPENDENT TRANSCRIPTION IN HUMAN CELLS

Our finding that two yeasts engage positive feedback prior to full regulon activation suggests that this regulatory motif is widespread. Thus, we chose to examine a mammalian system. Although many of the components of the genetic network regulating the G1/S transition in mammals do not have well-defined orthologs in yeast, both networks contain multiple positive feedback elements indicating similar network topology (Figure 5-2A). There is a functional analogy between the cyclin D-E2F-Rb-cyclinE and the Cln3-SBF/MBF-Whi5-Cln1/2 pathways. Furthermore, both budding yeasts and mammals regulate commitment to cell division in response to multiple internal and external signals at the G1/S transition (Planas-Silva and Weinberg 1997; Blagosklonny and Pardee 2002; Yao, Lee et al. 2008).

Mammalian G1 progression is initiated by mitogen-dependent activation of cyclin D–CDK4/6 complexes, which phosphorylate and partially inactivate the transcriptional inhibitor Rb (Blagosklonny and Pardee 2002). This allows for the initiation of transcription by the E2F family (E2F1-3) of transcription factors. Included in this set of genes are cyclin E1 and cyclin E2, which complex with CDK2 to phosphorylate Rb and

thereby complete a positive feedback loop (Bracken, Ciro et al. 2004). Additionally, at the G1/S transition E2F initiates transcription of E2F1, which may form a second transcriptional positive feedback loop(Johnson, Ohtani et al. 1994). The SCF component Skp2, responsible for the specific degradation of the CDK inhibitor p27, is also an E2F target (Yung, Walker et al. 2007). Therefore, multiple potential positive feedback loops may act during the mammalian G1/S transition. If our feedback-first model applies to mammalian cell cycle control, we expect to see feedback loop components transcribed before other E2F targets.

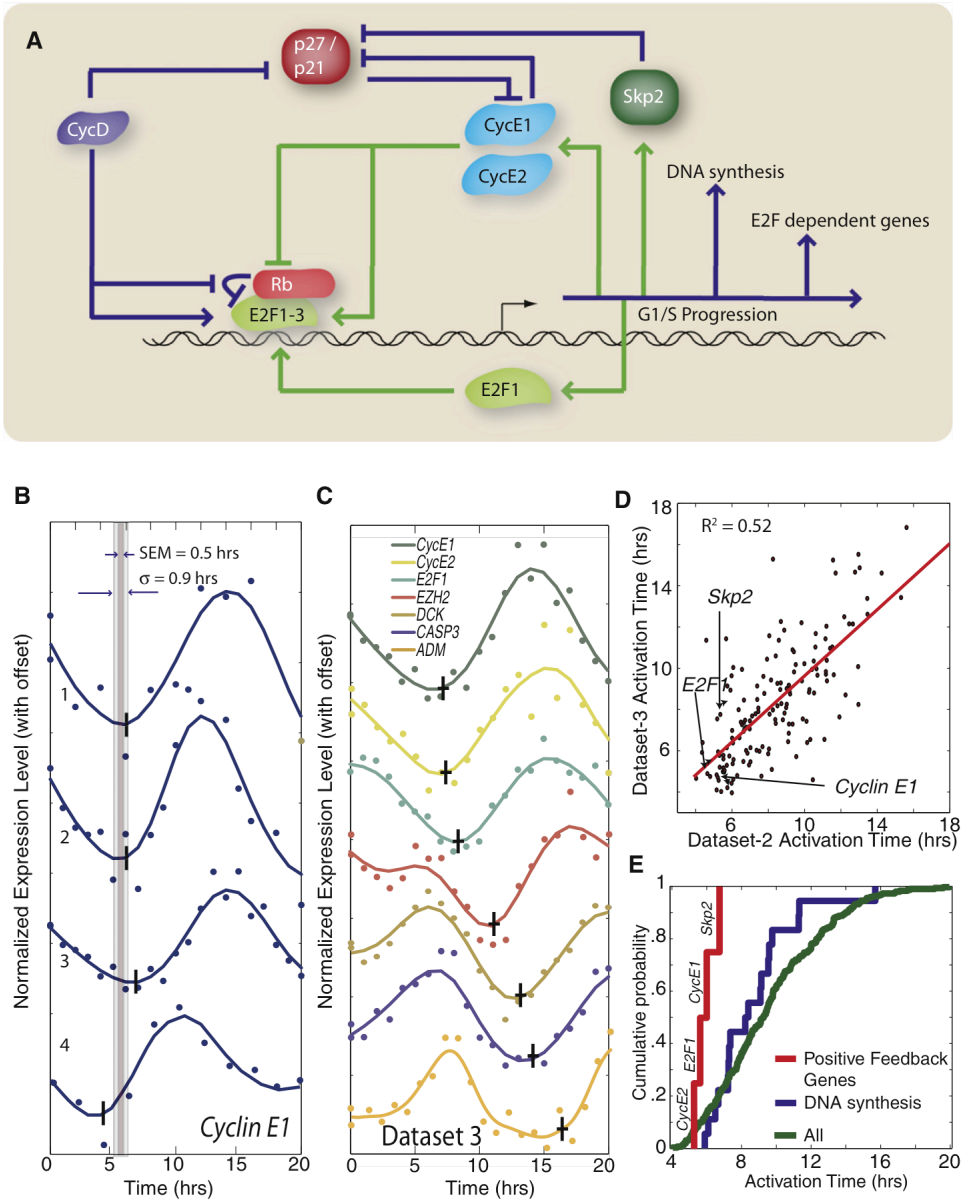


Figure 5-2: Feedback-first regulation is conserved in human cells. (A) Schematic diagram of G1/S regulation in human cells. (B) Cyclin E1 activation is consistent in 4 different cell cycle synchronized microarray experiments from Whitfield et al. (2002). The standard deviation σ and the standard error of the mean (SEM) is calculated for each gene. (C) 7 genes regulated by the E2F family of transcription factors are activated at different times; data shown from a single dataset. The vertical and horizontal bars indicate the mean activation time and the SEM respectively. (D) Gene activation time correlation between 2 datasets (R^2 = 0.52). (E) Cumulative distribution of mean activation times for cell cycle regulated E2F-targets. Genes responsible for positive feedback at the G1/S transition, including the cyclins E1 and E2 the transcription factor E2F1, and the SCF component Skp2, are transcribed earlier than other E2F-targets

($p<0.01$) and earlier than the set of E2F targets specifically involved in DNA replication ($p=0.03$). This demonstrates the conservation of feedback-first regulation in eukaryotes.

To test our hypothesis that the positive feedback elements are transcribed early, we first need to define a set of cell cycle regulated E2F targets (Markey, Angus et al. 2002; Ren, Cam et al. 2002). Therefore, we compiled a list of 311 cell cycle regulated E2F targets from two previous studies (Table S5) (Muller, Bracken et al. 2001; Whitfield, Sherlock et al. 2002; Xu, Bieda et al. 2007).

We analyzed gene activation timing in human HeLa cells for 4 cell cycle synchronized microarray time courses (Whitfield, Sherlock et al. 2002). We see consistent activation of individual genes across the datasets. For example, cyclin E1 is activated at 5.4 ± 0.6 hours on average with a standard deviation of 1.1 hours (Figure 5-2B). Our analysis identifies distinct activation times for E2F regulated genes (Figure 5-2C). Three experiments synchronized cell with a double thymidine block and one experiment used a thymidine block followed by a nocodazole block (Whitfield, Sherlock et al. 2002). However, we see no difference in relative activation timing due to the two different synchronization methods as all four data sets yield comparable results (Figure 5-2D). As for the two yeast analyses, we observe that genes responsible for positive feedback (cyclin E1, cyclin E2, Skp2 and E2F1) are among the first transcribed at the G1/S transition consistent with our feedback-first model (Figure 5-2E).

CHAPTER6 : DISCUSSION

We showed that genome-wide transcription is restricted until positive feedback commits a cell to division. This regulatory organization was previously unclear because transcription of the G1 cyclins and the rest of the G1/S regulon are both dependent on the same transcription factors and appear concurrent when analyzed with clustering-based algorithms. In contrast, both our activation detection algorithm and parametric algorithms preserve dynamic information (Chechik, Oh et al. 2008).

TEMPORAL SEPARATION OF POSITIVE AND NEGATIVE FEEDBACK LOOPS

An interesting feature of the G1/S regulon is that both positive (*CLN1,2*) and negative (*NRM1*) feedback elements are regulated under the same CDK-dependent transcription factors. Activating both feedbacks at the same time would be much like stepping on the brake and gas pedals simultaneously to detrimental effect (Figure 2). To avoid this outcome, promoter specific kinetics may allow temporal separation of positive and negative feedback loops. A similar process was found to regulate mitotic entry in *Xenopus* egg extracts(Georgi, Stukenberg et al. 2002). Wee1 and Cdc25 phosphorylation by CDK1, which is associated with positive feedback at mitotic entry, occurs before the phosphorylation of other CDK1 targets including the APC component Cdc27. Thus, both feedback-first regulation and the temporal separation of positive and negative feedback loops may be enacted through the evolution of differential rate constants.

FEEDBACK-FIRST REGULATION ENSURES COMMITMENT TO CELL DIVISION PRIOR TO LARGE-SCALE CHANGES IN GENE EXPRESSION

Transcribing genes when they are needed may increase efficiency by avoiding unnecessary protein synthesis. The subunits of the *E. coli* flagella were found to be synthesized in the order that they are needed for assembly (Kalir, McClure et al. 2001). Fine temporal control of transcription during amino acid synthesis ensured that enzymes were made in the order they were needed (Zaslaver, Mayo et al. 2004). In the cell cycle context, ribonucleotide reductase is transcribed just before S phase (Elledge and Davis 1990), and histones are transcribed during S phase to be assembled with newly replicated DNA (Borun, Gabrielli et al. 1975; Hereford, Osley et al. 1981). Taken together, these studies indicate that fine temporal order of events may provide a fitness advantage.

A transcriptional oscillation with specific temporal order occurs through the cell division cycle in both prokaryotes and eukaryotes. This extensive oscillation entails ~10-20% of all *Caulobacter* and budding yeast genes (Cho, Campbell et al. 1998; Spellman, Sherlock et al. 1998; Laub, McAdams et al. 2000). However, a comparative analysis of the yeasts *S. pombe* and *S. cerevisiae* revealed that temporal regulation of most orthologous genes is not well conserved (Rustici, Mata et al. 2004; Oliva, Rosebrock et al. 2005; Peng, Karuturi et al. 2005). Indeed, the order of gene activation at the G1/S transition in *S. cerevisiae* depends on the synchronization phase (Figure 2-6D). Further comparison of the two yeasts with human cell lines and the plant *Arabidopsis* revealed that only 5 orthologs are cell cycle regulated in all 4 species (Jensen, Jensen et al. 2006). However, different protein subunits of the same complex were often found to have cell cycle regulated transcription in different species, suggesting conserved transcriptional control of the complex rather than the individual subunits (Jensen, Jensen et al. 2006). Thus, although periodic transcription of individual genes varies, there may still be conserved regulatory features.

TOWARDS THE MECHANISTIC BASIS OF TRANSCRIPTION ORDER

We observed considerable variation in gene activation timing among genes regulated by a specific transcription factor. In budding yeast, we showed that a significant amount of

this variation is due to combinatorial use of the transcription factors SBF and MBF resulting in logical OR gates for both transcriptional activation and inactivation. Thus, the genes regulated by both SBF and MBF transcription factors are activated early in mitotic block-release experiments, where SBF is activated before MBF, and in G1 block-release experiments, where MBF is activated before SBF. This may be functionally relevant as the earliest activated G1 cyclin *CLN1* is regulated by both factors (Flick, Chapman-Shimshoni et al. 1998; Ferrezuelo, Colomina et al. 2010), which may ensure feedback-first regulation in a variety of physiological contexts.

Cells utilize the digital nature of the genetic code but many of the complex biological processes require analog tuning of the signals. In other words, regulating the expression levels and the timing of many genes only with a single transcription factor makes it difficult to optimize the cost for adaptation. Compartmentalization of transcription factories is an elegant way of adjusting the *trans*-signal by a transcription factor. Furthermore, our analysis in chapter 4 emphasized the effect of nuclear organization on the transcriptional activation timing.

The chromosomes are tethered at specific sites to distinct loci in the nucleus, constraining the conformation for certain chromosome regions. We showed that the earlier activated genes have more interactions, which means that they are less confined than the genes activated later that have fewer interactions. The relationship between the transcriptional activation and the interaction amount can possibly be explained by the chromatin state. Certain regions have closed structure, called heterochromatin, and known as transcriptionally more silent; whereas other regions are more flexible, called euchromatin, and have been identified as transcriptionally more active (Kurdistani, Tavazoie et al. 2004; Pokholok, Harbison et al. 2005). Thus, the amount of total interactions of a gene might correlate with the openness of the chromatin, which is already reported as in relation to transcription. Future work will aim at explaining the relationship between the activation timing and the chromatin state.

There are other possible mechanisms that can also contribute to the molecular basis for the significant temporal variation in G1/S transcription unexplained by the combinatorial use of SBF and MBF. One possibility is that differential transcription timing may arise through the combinatorial use of additional transcription factors (Kato, Hata et al. 2004). In such a model, intermediate times might be produced by regulating a promoter with a late-activated and an early-activated transcription factor. An example of this type of regulation is that the Fkh2-regulated genes show different activation times at G2/M depending on Yox1-coregulation (Darieva, Clancy et al. 2010). This model suggests that the late activated SBF targets might also be regulated by a late-activated transcription factor such as Fkh2. A large number of transcription factors might therefore account for the variation in gene activation time.

A second possibility is that promoter-specific rate-constants underlie gene activation kinetics. This could arise through promoter-specific transcription factor and nucleosome arrangements or TATA-box sequence (Lam, Steger et al. 2008; Chechik and Koller 2009; Bai, Charvin et al. 2010; Mogno, Vallania et al. 2010). Thus, in response to a single input such as CDK activity, the organization of kinetic parameters can result in differential activation timing (Shen-Orr, Milo et al. 2002). We note that all these classes of mechanisms are not mutually exclusive and likely cooperate to tune gene expression.

Here, we identify such a conserved regulatory feature of the eukaryotic cell cycle. We find that commitment via positive feedback precedes large-scale transcriptional activation at the G1/S transition. Our study was able to identify feedback-first regulation because we employ a novel algorithm to analyze activation and inactivation separately. We revealed feedback-first regulation in the yeasts, *S. cerevisiae* and *S. bayanus* as well as in human cells. The conservation of feedback-first regulation leads us to anticipate its widespread use in cellular and developmental transitions.

APPENDIX A: STRAINS

YEAST STRAINS

Table A-1: *Yeast strains used in this study*

Name	Genotype	Source
JS38-1a	*MAT**a** cln1Δ::HIS3 cln2Δ cln3Δ::LEU2 TRP1::MET3pr-CLN2 HTB2-mCherry-spHIS5 WHI5-GFP-kanMX ADE2 ura3-1 can1-1* (W303 background)	(Skotheim, Di Talia et al. 2008)
JS163-8d	*MAT**a** bar1Δ::URA3 ADE2 leu2-3,100 his3-11,15 trp1-1 can1-1* (W303 background)	(Eser, et al. 2011)
JS209	*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 CLN2::CLN2pr-Venus$_{PEST}$* (BY4741 background)	(Eser, et al. 2011)
JS210	*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 CLN1::CLN1pr-Venus$_{PEST}$* (BY4741 background)	(Eser, et al. 2011)

C-terminal GFP fusion strains in the BY4741 background used for live-cell imaging were from the UCSF collection(Ghaemmaghami, Huh et al. 2003). We created JS209 by integrating pJS19 at the CLN2 locus after EcoNI digestion. Similarly, we created JS201 by integrating pJS25 at the CLN1 locus after AgeI digestion.

PLASMIDS

Table A-2*: Plasmids used in this study*.

Name	Description	Source
pJS19	pRS406-*CLN2pr-Venus-PEST*	(Eser, et al. 2011)
pJS25	pRS406-*CLN1pr-Venus-PEST*	(Eser, et al. 2011)
pGC08D	pRS404-*CLN2pr-Venus-PEST*	G. Charvin

pGC08D was a kind gift from G. Charvin. To construct pJS19, the *CLN2pr-Venus-PEST* insert from pGC08D was ligated to the pRS406 vector following digestion with BglII. A 1kb *CLN1* promoter fragment with terminal PacI and BamHI restriction sites was

obtained by PCR of genomic DNA replace the 1kb *CLN2* promoter in pJS19 to create pJS25.

APPENDIX B: EXPERIMENTAL PROCEDURES

MICROARRAY EXPERIMENTS AND ANALYSIS

Mitotic block release analysis was based on data collected in (Di Talia, Wang et al. 2009). The sequential order of activations is highly consistent between datasets indicating a defined temporal regulation even though the genotypes for the 7 time-courses were not identical (Figure 2-1F, Figure 2-3; Table 2-1).

G1 block-release experiments were performed at 30°C. Cells were harvested immediately after inoculation and then every 5 minutes thereafter. Microarray hybridization was performed at the Stony Brook Microarray Facility. For pheromone block experiments, *bar1Δ* cells were grown in log-phase in either SCD (%2) or SCG (3%) before being arrested for 135 minutes in 95nM α-factor. Cells were then washed and inoculated into pheromone-free media. The cln-block experiment was performed using *cln1Δ cln2Δ cln3Δ MET3pr-CLN2* cells grown to early log-phase in SCD - met (media without methionine; exogenous *CLN2* on), then 0.2g/L met was added for 120 minutes to arrest cells in early G1 (*CLN2* off). Cells were then washed and inoculated into 4mg/L met (*CLN2* partially on) to provide the amount of *CLN2* expression resulting in budding kinetics similar to WT cells released from a pheromone block.

There was some ambiguity in identifying the gene activation time for *CLN1* in the *S. bayanus* data set because either the 2nd or 3rd data point for *CLN1* was likely an outlier. Therefore, we averaged the activation times for the dataset after having removed either the 2nd or 3rd data point.

TIME-LAPSE FLUORESCENCE MICROSCOPY

Wide-field fluorescence and phase-contrast images were captured every 3 minutes for 6 hours from cells prepared as previously described (Bean, Siggia et al. 2006). Cells were

automatically segmented and the mean fluorescence intensity was measured. Bud emergence was identified manually using phase images.

APPENDIX C: ANALYZING THE HI-C DATA

We separated the genes into two groups, high-interacting and low interacting, based on their interaction amount. To see whether the distributions of activation times of high and low interacting genes are significant, we applied Kolmogorov-Smirnov test (KS-test). By changing the threshold from 1 to 600, we found that for a wide range of threshold, p-value is less than 0.05. This result confirms that the significance of the statistics does not sensitively depend on the choice of threshold. Therefore, in chapter 4, we used 330 as a threshold for grouping the high and low interacting genes.

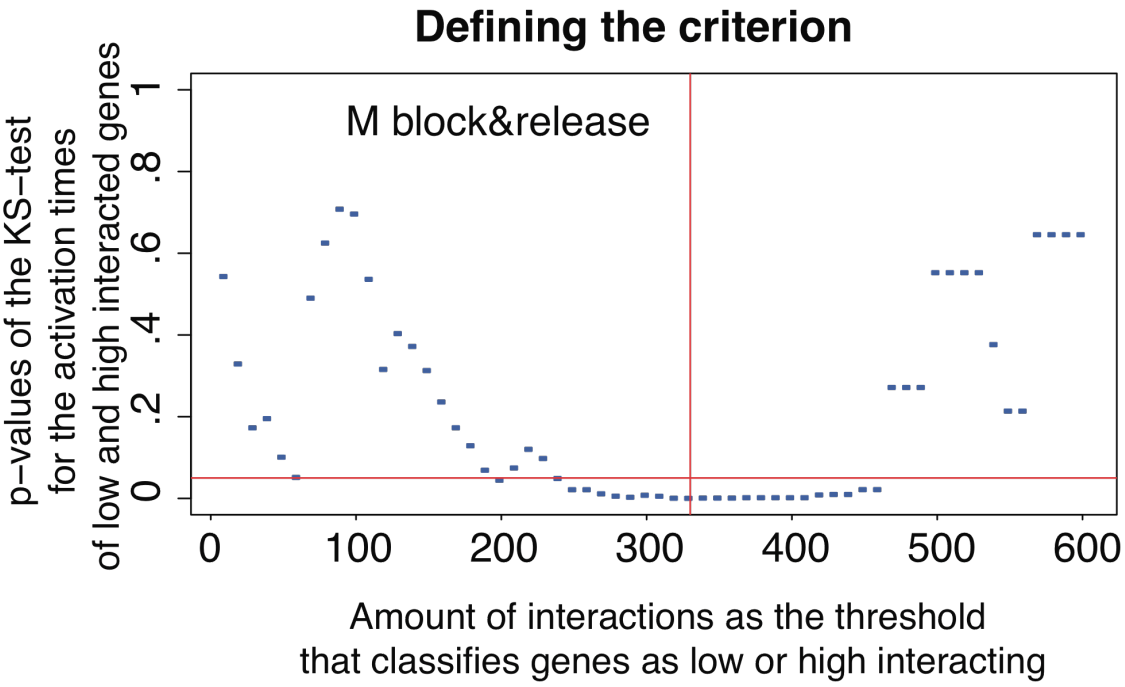


Figure C-1: p-value of the KS-test applied to the activation time distributions of high and low interacting genes. x-axis is the threshold values that classify the high and low interaction groups.

APPENDIX D: G1/S GENES ACTIVATION TIMES

Table D-1: List of G1/S genes with their activation times in mitotic and G1 block-release experiments

YORF	ORF	function	TF	Activation Time- M block	STD	SEM	Activation Time- G1 block
YGR041W	BUD9	biosynthesis		4	NaN	NaN	13
YHR067W	HTD2	biosynthesis		4	NaN	NaN	NaN
YBR071W	YBR071W	biosynthesis	SBF&MBF	5	2.1	0.9	12
YAL067C	SEO1	uncharacterized		5	0.9	0.3	35
YGL028C	SCW11	cell wall		5	3.0	1.1	NaN
YOR075W	UFE1	transport		5	2.6	1.5	NaN
YPL054W	LEE1	uncharacterized		5	NaN	NaN	NaN
YMR199W	CLN1	G1/S control	SBF&MBF	6	2.3	0.9	5
YLR049C	YLR049C	uncharacterized	MBF	6	NaN	NaN	14
YIR017C	MET28	bud		6	2.1	0.9	35
YDR528W	YDR528W	biosynthesis		6	NaN	NaN	NaN
YKR077W	YKR077W	biosynthesis	MBF	6	1.6	0.6	NaN
YLR286C	CTS1	DNA		6	2.3	0.9	NaN
YLR300W	EXG1	cell wall	SBF&MBF	6	2.9	1.1	NaN
YML027W	YOX1	TF transcription at M/G1	SBF&MBF	7	5.7	2.3	10
YGR189C	CRH1	cell wall	SBF&MBF	7	3.9	1.7	15
YOL019W	YOL019W	uncharacterized	SBF	7	2.9	1.2	21
YPR106W	ISR1	G1/S control / DNA		7	NaN	NaN	29
YER124C	DSE1	uncharacterized		7	3.6	1.4	32
YDR353W	TRR1	biosynthesis		7	3.1	1.4	NaN
YHR123W	EPT1	biosynthesis		7	2.3	1.0	NaN
YER070W	RNR1	DNA	SBF&MBF	8	5.3	2.2	10
YIL076W	SEC28	transport		8	3.5	2.5	12
YGR055W	MUP1	transcription		8	2.4	0.9	30
YPL269W	KAR9	mitosis		8	4.2	3.0	32
YGL184C	STR3	biosynthesis		8	2.5	1.0	33
YKR069W	MET1	biosynthesis		8	2.5	1.0	35
YHR143W	DSE2	cell wall		8	3.1	1.3	36
YNL208W	YNL208W	uncharacterized		8	2.3	1.3	NaN
YDL101C	DUN1	DNA	MBF	9	7.9	3.9	8
YLR313C	SPH1	bud	MBF	9	3.4	1.5	9

YDL156W	YDL156W	cell wall	MBF	9	2.7	1.1	13
YEL068C	YEL068C	uncharacterized		9	7.5	4.3	23
YLL062C	YLL062C	DNA		9	2.1	0.8	42
YAR003W	SWD1	transcription		9	2.3	1.3	NaN
YER091C	MET6	biosynthesis		9	2.8	1.0	NaN
YPL232W	SSO1	transport		9	5.3	2.0	NaN
YDL227C	HO	endonuclease	SBF	10	3.4	1.4	11
YIL140W	SRO4	bud	SBF&MBF	10	4.0	1.8	13
YOR308C	YOR308C	RNA processing		10	4.0	2.3	30
YKL001C	MET14	biosynthesis		10	2.4	0.9	43
YKL042W	SPC42	SPB	MBF	11	4.7	1.8	10
YDR356W	NUF1	SPB	MBF	11	5.7	2.8	11
YOL007C	YOL007C	uncharacterized	SBF	11	4.6	1.8	11
YEL064C	YEL064C	bud		11	1.5	0.8	13
YAL053W	YAL053W	cell wall		11	2.7	1.0	14
YJR006W	HYS2	DNA		11	6.8	3.4	15
YKR091W	SRL3	DNA		11	4.0	1.6	15
YLL012W	YEH1	biosynthesis		11	2.9	1.3	16
YFR030W	MET10	biosynthesis		11	1.3	0.5	37
YJR137C	ECM17	biosynthesis		11	2.3	0.9	38
YGL125W	MET11	biosynthesis		11	6.2	2.5	40
YIL104C	SHQ1	RNA processing		11	2.9	1.7	NaN
YDL103C	QRI1	cell wall		12	13.3	7.7	12
YPL163C	SVS1	cell wall	SBF	12	3.8	1.4	13
YNL225C	CNM1	DNA		12	6.3	2.6	14
YDL127W	PCL2	DNA	SBF	12	13.8	5.2	16
YPR018W	RLF2	DNA	MBF	12	5.6	2.3	16
YJR155W	YJR155W	DNA		12	4.0	1.8	17
YKR010C	TOF2	transcription silencing		12	7.0	3.1	32
YLL061W	MMP1	biosynthesis		12	2.8	1.0	36
YCL024W	KCC4	bud		12	3.1	1.2	NaN
YKL065C	YET1	uncharacterized		12	NaN	NaN	NaN
YGL207W	SPT16	DNA		13	NaN	NaN	10
YHR030C	SLT2	cell wall		13	2.8	1.1	11
YOR115C	YOR115C	transport		13	6.2	2.8	11
YPL256C	CLN2	Cdh1 inhibitor / G1/S ?	SBF	13	1.9	0.7	12
YPR120C	CLB5	G1/S control	MBF	13	3.7	1.4	14
YKL101W	HSL1	biosynthesis	SBF&MBF	13	4.3	1.6	15
YDR309C	GIC2	bud	SBF&MBF	13	1.9	0.7	16
YPR174C	YPR174C	uncharacterized	MBF	13	7.2	3.2	17

YLR121C	YPS3	cell wall	SBF	13	2.4	0.9	18
YIL066C	RNR3	DNA		13	2.9	1.7	19
YPL274W	SAM3	biosynthesis		13	2.1	0.8	35
YIL074C	SER33	biosynthesis		13	1.3	0.5	49
YBR073W	RDH54	DNA		13	3.9	1.9	NaN
YDL055C	PSA1	cell wall	SBF	13	2.3	0.9	NaN
YER149C	PEA2	mitosis		13	4.6	2.7	NaN
YHR149C	SKG6	bud	SBF&MBF	13	2.7	1.0	NaN
YNL309W	STB1	G1/S control	MBF	14	7.1	3.5	7
YMR179W	SPT21	transcription silencing	SBF&MBF	14	5.2	2.1	8
YGR042W	YGR042W	uncharacterized		14	2.6	1.2	10
YNL300W	YNL300W	bud	SBF	14	8.5	6.0	11
YDL211C	YDL211C	mating type switching		14	8.1	4.7	12
YML012W	ERV25	DNA	MBF	14	5.4	2.7	12
YDR113C	PDS1	biosynthesis	SBF&MBF	14	7.6	2.9	13
YBR088C	POL30	DNA	SBF&MBF	14	2.9	1.1	14
YJL181W	YJL181W	uncharacterized	MBF	14	5.4	2.4	14
YOR144C	YOR144C	biosynthesis	MBF	14	5.5	2.3	17
YJR054W	YJR054W	mitosis	SBF	14	3.3	1.7	18
YKL089W	MIF2	mitosis	MBF	14	5.9	2.4	18
YJL078C	PRY3	uncharacterized		14	2.4	1.2	NaN
YOL034W	SMC5	DNA	MBF	14	5.3	2.6	NaN
YLR032W	RAD5	telomere silencing	MBF	15	3.5	2.5	9
YNL082W	PMS1	cell wall	MBF	15	3.6	1.6	9
YLR467W	YRF1-5	telomeres		15	4.5	2.3	10
YOL017W	YOL017W	MBF corepressor	MBF	15	5.5	2.3	11
YGL061C	YGL061C	biosynthesis		15	5.3	2.2	14
YBR007C	DSF2	uncharacterized		15	10.2	4.5	16
YAR008W	SEN34	DNA	MBF	15	3.9	1.7	17
YKL008C	YKL008C	vacuole		15	5.7	2.6	29
YPL253C	VIK1	SPB		15	11.3	8.0	29
YPL250C	YPL250C	uncharacterized		15	2.4	1.0	30
YBR108W	AIM3	uncharacterized		15	4.7	2.1	NaN
YOR195W	SLK19	biosynthesis	MBF	15	3.9	1.7	NaN
YDR507C	GIN4	bud	SBF&MBF	16	6.0	2.3	7
YER001W	MNN1	biosynthesis	SBF&MBF	16	3.7	1.4	8
YGR099W	TEL2	DNA		16	9.9	7.0	8
YGR296W	YRF1	DNA		16	5.0	2.5	11
YGR014W	MSB2	bud	SBF	16	6.1	2.5	13
YPL267W	ACM1	G1/S control	SBF&MBF	16	2.7	1.0	13

YPR075C	OPY2	stress		16	4.6	2.7	14
YIL144W	TID3	mitosis		16	2.5	1.1	15
YKR013W	PRY2	DNA	SBF	16	4.6	1.7	16
YCL061C	YCL061C	bud	MBF	16	2.6	1.1	18
YDR440W	PCH1	DNA		16	5.7	2.6	20
YDR448W	ADA2	transcription		16	10.3	4.6	23
YEL047C	FRD1	biosynthesis		16	4.5	2.0	24
YPL032C	SVL3	SPB	SBF	16	3.5	1.4	30
YLR343W	GAS2	cell wall		16	10.5	4.7	NaN
YNL273W	TOF1	bud	MBF	16	5.4	2.2	NaN
YPL153C	RAD53	cell wall	MBF	16	2.1	0.9	NaN
YGR188C	BUB1	mitosis		17	6.0	3.5	9
YJL115W	ASF1	DNA	MBF	17	6.0	2.3	12
YHR061C	GIC1	cell wall		17	4.5	2.0	13
YBL003C	HTA2	DNA	SBF	17	3.9	1.5	16
YOR073W	YOR073W	transcription silencing		17	3.8	1.6	21
YDR225W	HTA1	DNA	SBF	17	3.9	1.5	23
YDR261C	EXG2	cell wall		17	3.5	1.4	30
YPR202W	YPR202W	uncharacterized		17	4.4	2.2	NaN
YGR221C	YGR221C	G1/S control / DNA		18	4.4	1.8	7
YJL118W	YJL118W	uncharacterized		18	12.7	7.3	11
YLR103C	CDC45	DNA	MBF	18	3.3	1.4	11
YOR176W	HEM15	biosynthesis		18	6.3	2.8	11
YML133C	YML133C	uncharacterized	MBF	18	2.5	1.5	12
YNL262W	POL2	DNA		18	4.1	1.5	12
YMR076C	PDS5	mitosis	MBF	18	4.3	1.6	13
YOL090W	MSH2	bud	MBF	18	4.6	1.7	13
YDR451C	YDR451C	biosynthesis		18	3.9	1.9	14
YER095W	RAD51	biosynthesis	MBF	18	1.5	0.8	14
YNL166C	BNI5	bud		18	5.0	1.9	14
YMR307W	GAS1	G1/Scontrol / bud	SBF	18	4.5	1.7	15
YIL117C	PRM5	uncharacterized		18	3.8	1.7	16
YLR212C	TUB4	SPB	MBF	18	3.3	1.7	16
YPR135W	CTF4	biosynthesis	MBF	18	3.5	2.5	16
YNL030W	HHF2	DNA	SBF&MBF	18	3.2	1.2	17
YOR247W	SRL1	cell wall	SBF	18	11.6	4.4	17
YPR141C	KAR3	mitosis		18	3.9	1.9	18
YDR144C	MKC7	cell wall	MBF	18	3.6	1.5	22
YLR342W	FKS1	cell wall		18	1.3	0.6	30
YCR065W	HCM1	DNA	SBF&MBF	18	6.0	2.3	NaN
YDL102W	CDC2	biosynthesis	MBF	18	3.4	1.7	NaN

YOL094C	RFC4	DNA		18	8.9	3.6	NaN
YNL339C	YNL339C	cell wall	MBF	19	2.9	1.7	7
YDR279W	YDR279W	G1/Scontrol	MBF	19	6.9	2.8	9
YGL027C	CWH41	biosynthesis		19	NaN	NaN	11
YER003C	PMI40	biosynthesis	SBF	19	11.8	4.5	12
YPL127C	HHO1	DNA	SBF	19	2.6	1.5	12
YDL003W	MCD1	mitosis	SBF&MBF	19	5.3	2.0	14
YGR140W	CBF2	bud	SBF&MBF	19	3.8	1.7	14
YJL201W	ECM25	uncharacterized		19	4.4	1.8	14
YNL263C	YIF1	transport	MBF	19	10.2	4.1	14
YPL208W	YPL208W	biosynthesis / translation	MBF	19	4.4	2.0	14
YDL157C	YDL157C	uncharacterized	MBF	19	6.5	2.9	15
YMR238W	DFG5	cell wall		19	3.5	1.4	16
YFR027W	ECO1	DNA		19	3.6	2.1	17
YPL221W	YPL221W	cell wall		19	8.6	3.5	19
YBL002W	HTB2	DNA	SBF	19	4.4	1.7	22
YDR247W	YDR247W	DNA		19	2.9	1.7	35
YJL074C	SMC3	mitosis	MBF	19	3.9	1.7	NaN
YJL187C	SWE1	mitosis	SBF&MBF	19	3.5	1.4	NaN
YNL102W	POL1	DNA	SBF&MBF	19	5.2	2.1	NaN
YNR066C	YNR066C	cell wall		19	NaN	NaN	NaN
YPR159W	KRE6	biosynthesis		19	4.7	1.9	NaN
YLL022C	YLL022C	bud / polarization		20	2.5	1.5	9
YPL124W	NIP29	DNA	SBF	20	3.7	1.4	13
YDR097C	MSH6	DNA	MBF	20	3.5	1.4	14
YML061C	PIF1	DNA		20	7.1	3.2	14
YDL164C	CDC9	DNA	MBF	20	6.8	3.1	15
YEL042W	GDA1	biosynthesis		20	4.5	2.6	16
YJL073W	JEM1	chaperone		20	11.4	4.7	17
YNL233W	BNI4	DNA	MBF	20	4.3	1.7	19
YDR297W	SUR2	biosynthesis	SBF&MBF	20	3.3	1.2	20
YML060W	OGG1	DNA	MBF	20	3.5	1.4	21
YBL063W	KIP1	DNA		20	8.7	4.3	26
YML102W	CAC2	DNA		21	3.8	1.7	6
YPL255W	BBP1	G1/S control / bud	SBF	21	5.8	2.2	9
YHR153C	SPO16	sporulation	MBF	21	3.2	1.4	10
YKL113C	RAD27	DNA	SBF&MBF	21	7.3	2.8	10
YHR154W	RTT107	DNA	SBF&MBF	21	6.5	2.5	12
YGL093W	YGL093W	cell wall	SBF	21	2.4	1.2	13
YJR043C	POL32	DNA	MBF	21	2.7	1.2	13
YNR009W	NRM1	negative regulator	SBF	21	4.3	1.6	13

		of MBF					
YGL225W	GOG5	DNA	SBF	21	5.7	2.2	14
YMR144W	YMR144W	uncharacterized	SBF&MBF	21	2.9	1.1	14
YLR457C	NBP1	SPB		21	NaN	NaN	15
YMR305C	YMR305C	transcription / telomere	SBF&MBF	21	5.8	2.2	15
YOR074C	CDC21	DNA	MBF	21	11.1	4.2	15
YOR321W	PMT3	biosynthesis		21	NaN	NaN	15
YBR009C	HHF1	DNA	SBF	21	3.7	1.4	19
YPR175W	DPB2	DNA	MBF	21	5.8	2.2	20
YDR224C	HTB1	DNA	SBF	21	5.3	2.0	21
YDL105W	QRI2	DNA	MBF	21	7.6	3.1	NaN
YMR078C	CTF18	DNA		21	5.2	2.1	NaN
YLR383W	RHC18	biosynthesis	SBF&MBF	22	5.6	2.5	10
YNL231C	PDR16	biosynthesis	SBF&MBF	22	7.2	2.7	10
YGR152C	RSR1	bud	SBF&MBF	22	5.4	2.0	12
YGR109C	CLB6	biosynthesis	SBF&MBF	22	NaN	NaN	14
YER118C	SSU81	biosynthesis	MBF	22	6.3	2.4	16
YAR007C	RFA1	DNA	MBF	22	7.8	3.5	17
YJL173C	RFA3	DNA	SBF&MBF	22	3.5	1.5	19
YOR084W	YOR084W	DNA		22	4.0	1.5	19
YLR372W	SUR4	biosynthesis		22	4.2	1.6	22
YOR188W	MSB1	biosynthesis		22	5.0	1.9	23
YKL165C	MCD4	bud		22	3.4	1.4	26
YPR034W	ARP7	transcription		22	4.7	2.7	31
YFL008W	SMC1	DNA / mitosis	MBF	22	8.4	3.2	NaN
YIL141W	YIL141W	uncharacterized		22	17.2	8.6	NaN
YKL108W	YKL108W	DNA	MBF	23	5.8	2.2	10
YKL045W	PRI2	biosynthesis	MBF	23	7.7	3.2	11
YML109W	ZDS2	DNA	MBF	23	5.8	2.2	13
YEL061C	CIN8	mitosis		23	5.3	2.1	14
YER016W	BIM1	mitosis	MBF	23	3.6	1.6	14
YJL019W	MPS3	SPB		23	8.1	4.0	14
YNL072W	RNH35	cell wall		23	6.4	3.2	14
YLR183C	TOS4	uncharacterized	SBF	23	5.0	1.9	15
YER019W	ISC1	biosynthesis		23	4.8	2.4	16
YGR153W	YGR153W	membrane		23	3.5	2.5	16
YLL002W	KIM2	Dna		23	4.7	1.8	16
YIL026C	IRR1	mitosis	MBF	23	4.8	2.1	17
YDL095W	PMT1	cell wall		23	8.7	4.3	21
YBR010W	HHT1	DNA		23	7.2	3.0	24
YNL165W	YNL165W	DNA		23	NaN	NaN	24

YOR114W	YOR114W	uncharacterized	SBF&MBF	24	3.9	1.6	13
YOR320C	GNT1	biosynthesis		24	4.7	1.8	13
YNL304W	YNL304W	DNA		24	3.6	1.8	14
YDL010W	YDL010W	TF for S phase expression	MBF	24	6.6	2.7	16
YNL283C	WSC2	cell wall / stress	SBF	24	3.0	1.1	25
YDL197C	ASF2	transcription silencing	MBF	25	2.2	1.1	10
YOR083W	WHI5	negative regulator of SBF		25	NaN	NaN	12
YLR045C	STU2	mitosis	SBF	25	1.9	0.8	13
YOR372C	NDD1	mitosis	SBF	25	3.3	1.2	18
YDL093W	PMT5	biosynthesis		25	3.4	1.5	20
YNL031C	HHT2	DNA	SBF	25	3.8	1.6	22
YBR008C	FLR1	DNA		25	1.2	0.7	28
YNL312W	RFA2	membrane / vesicles	MBF	26	8.2	3.4	12
YBL009W	ALK2	uncharacterized	SBF	26	5.1	1.9	14
YDR501W	PLM2	uncharacterized		26	3.4	1.5	15
YGL200C	EMP24	transport		26	4.6	2.1	17
YBR070C	ALG14	biosynthesis	SBF&MBF	26	4.9	2.4	18
YHR173C	YHR173C	sporulation	SBF	26	3.5	1.4	23
YHR098C	YHR098C	bud		26	4.0	1.8	29
YOR355W	GDS1	uncharacterized		26	6.9	3.4	33
YGL163C	RAD54	DNA		27	9.9	4.4	8
YML020W	YML020W	uncharacterized		27	3.5	2.5	8
YGR238C	KEL2	mitosis		27	2.9	1.7	11
YDR307W	YDR307W	cell wall		27	NaN	NaN	12
YNL126W	SPC98	vacuole		27	NaN	NaN	13
YML021C	UNG1	DNA		27	3.3	1.7	14
YOL012C	HTA3	DNA	SBF	27	5.8	2.4	21
YOR307C	SLY41	biosynthesis		27	2.6	1.5	30
YHR172W	SPC97	SPB		27	3.7	1.5	NaN
YBR087W	RFC5	DNA		28	4.0	1.6	15
YJL092W	HPR5	SPB		28	3.5	1.4	15
YGL195W	GCN1	SPB		28	3.2	1.3	35
YDR368W	YPR1	bud		28	NaN	NaN	NaN
YMR295C	YMR295C	uncharacterized		28	NaN	NaN	NaN
YBR161W	YBR161W	DNA		29	NaN	NaN	21
YNL181W	YNL181W	DNA		30	6.4	2.9	11
YLR380W	YLR380W	cell wall		30	4.6	2.1	37
YKL127W	PGM1	biosynthesis	SBF	31	3.5	2.5	NaN
YKR037C	SPC34	SPB		32	3.9	1.6	15

YPR107C	YTH1	DNA		32	NaN	NaN	NaN
YLR154C	YLR154C	biosynthesis		33	NaN	NaN	19
YDL018C	ERP3	uncharacterized	MBF	34	3.5	2.5	14
YDR481C	PHO8	biosynthesis		35	NaN	NaN	16
YEL077C	YEL077C	uncharacterized		NaN	NaN	NaN	8
YJL119C	YJL119C	uncharacterized		NaN	NaN	NaN	8
YLL067C	YLL067C	uncharacterized		NaN	NaN	NaN	8
YPL209C	IPL1	mitosis		NaN	NaN	NaN	8
YBL113C	YBL113C	uncharacterized		NaN	NaN	NaN	9
YDL163W	YDL163W	G1/S control / bud		NaN	NaN	NaN	9
YHL049C	YHL049C	uncharacterized		NaN	NaN	NaN	9
YBL111C	YBL111C	uncharacterized		NaN	NaN	NaN	10
YBL112C	YBL112C	uncharacterized		NaN	NaN	NaN	10
YEL076C	YEL076C	uncharacterized		NaN	NaN	NaN	10
YEL076C-A	YEL076C-A	biosynthesis		NaN	NaN	NaN	10
YHL050C	YHL050C	bud		NaN	NaN	NaN	10
YHR218W	YHR218W	uncharacterized		NaN	NaN	NaN	10
YOR283W	YOR283W	uncharacterized		NaN	NaN	NaN	10
YBR089W	YBR089W	uncharacterized		NaN	NaN	NaN	11
YEL075C	YEL075C	biosynthesis		NaN	NaN	NaN	11
YER189W	YER189W	biosynthesis		NaN	NaN	NaN	11
YFL064C	YFL064C	osmosensor		NaN	NaN	NaN	11
YLL066C	YLL066C	biosynthesis		NaN	NaN	NaN	11
YLR135W	YLR135W	SPB / mitosis		NaN	NaN	NaN	11
YLR462W	YLR462W	uncharacterized		NaN	NaN	NaN	11
YNL289W	PCL1	biosynthesis		NaN	NaN	NaN	11
YCL022C	YCL022C	uncharacterized		NaN	NaN	NaN	12
YLR235C	YLR235C	uncharacterized		NaN	NaN	NaN	12
YGR151C	YGR151C	uncharacterized		NaN	NaN	NaN	13
YHR110W	YHR110W	DNA	MBF	NaN	NaN	NaN	13
YHR127W	HSN1	uncharacterized		NaN	NaN	NaN	13
YIL177C	YIL177C	uncharacterized		NaN	NaN	NaN	13
YLR233C	EST1	telomeres		NaN	NaN	NaN	13
YPL241C	CIN2	mitosis		NaN	NaN	NaN	13
YDL011C	YDL011C	mitosis		NaN	NaN	NaN	14
YDR053W	YDR053W	mitosis / DNA		NaN	NaN	NaN	14
YDR400W	URH1	biosynthesis		NaN	NaN	NaN	14
YER170W	ADK2	DNA		NaN	NaN	NaN	14
YGL038C	OCH1	biosynthesis	SBF&MBF	NaN	NaN	NaN	14
YIL132C	YIL132C	DNA		NaN	NaN	NaN	14
YLR050C	YLR050C	biosynthesis		NaN	NaN	NaN	14

YBL035C	POL12	DNA	MBF	NaN	NaN	NaN	15
YJL018W	MPS3	uncharacterized		NaN	NaN	NaN	15
YKL066W	YKL066W	uncharacterized		NaN	NaN	NaN	15
YKR090W	PXL1	bud	SBF	NaN	NaN	NaN	15
YLR151C	PCD1	biosynthesis		NaN	NaN	NaN	15
YDR488C	PAC11	transport		NaN	NaN	NaN	16
YGL185C	YGL185C	uncharacterized		NaN	NaN	NaN	16
YDL096C	YDL096C	biosynthesis		NaN	NaN	NaN	17
YDR503C	LPP1	TF / M/G1 transcription		NaN	NaN	NaN	17
YKL067W	YNK1	biosynthesis		NaN	NaN	NaN	17
YKL103C	LAP4	protein degradation		NaN	NaN	NaN	17
YFL060C	SNO3	DNA		NaN	NaN	NaN	18
YOR248W	YOR248W	cell wall		NaN	NaN	NaN	18
YPR076W	YPR076W	uncharacterized		NaN	NaN	NaN	18
YGL060W	YBP2	mitosis		NaN	NaN	NaN	20
YJL217W	YJL217W	uncharacterized		NaN	NaN	NaN	20
YLR326W	YLR326W	cell wall		NaN	NaN	NaN	20
YNL015W	PBI2	protein degradation		NaN	NaN	NaN	22
YGR113W	DAM1	mitosis		NaN	NaN	NaN	24
YCL023C	YCL023C	uncharacterized		NaN	NaN	NaN	27
YML100W	TSL1	biosynthesis		NaN	NaN	NaN	30
YLR455W	YLR455W	uncharacterized		NaN	NaN	NaN	31
YOR270C	VPH1	vacuole		NaN	NaN	NaN	31
YML117W	YML117W	telomere silencing / Mitosis		NaN	NaN	NaN	33
YKR012C	YKR012C	uncharacterized		NaN	NaN	NaN	43

BIBLIOGRAPHY

Amon, A., M. Tyers, et al. (1993). "Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins." Cell **74**(6): 993-1007.

Andrews, B. J. and I. Herskowitz (1989). "Identification of a DNA binding factor involved in cell-cycle control of the yeast HO gene." Cell **57**(1): 21-29.

Andrulis, E. D., A. M. Neiman, et al. (1998). "Perinuclear localization of chromatin facilitates transcriptional silencing." Nature **394**(6693): 592-595.

Babu, M. M., S. C. Janga, et al. (2008). "Eukaryotic gene regulation in three dimensions and its impact on genome evolution." Curr Opin Genet Dev **18**(6): 571-582.

Bai, L., G. Charvin, et al. (2010). "Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle." Dev Cell **18**(4): 544-555.

Bean, J. M., E. D. Siggia, et al. (2006). "Coherence and timing of cell cycle Start examined at single-cell resolution. ." Mol Cell **21**(1): 3-14.

Berger, A. B., G. G. Cabal, et al. (2008). "High-resolution statistical mapping reveals gene territories in live yeast." Nat Methods **5**(12): 1031-1037.

Blagosklonny, M. V. and A. B. Pardee (2002). "The restriction point of the cell cycle." Cell Cycle **1**(2): 103-110.

Borun, T. W., F. Gabrielli, et al. (1975). "Further evidence of transcriptional and translational control of histone messenger RNA during the HeLa S3 cycle." Cell **4**(1): 59-67.

Bracken, A. P., M. Ciro, et al. (2004). "E2F target genes: unraveling the biology." Trends Biochem Sci **29**(8): 409-417.

Brauer, M. J., C. Huttenhower, et al. (2008). "Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast." Mol Biol Cell **19**(1): 352-367.

Bystricky, K., T. Laroche, et al. (2005). "Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization." J Cell Biol **168**(3): 375-387.

Casolari, J. M., C. R. Brown, et al. (2004). "Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization." Cell **117**(4): 427-439.

Chechik, G. and D. Koller (2009). "Timing of gene expression responses to environmental changes." Journal of Computational Biology **16**(2): 279-290.

Chechik, G., E. Oh, et al. (2008). "Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network." Nature biotechnology **26**(11): 1251-1259.

Cho, R. J., M. J. Campbell, et al. (1998). "A Genome-Wide Transcriptional Analysis of the Mitotic Cell Cycle." Molecular cell **2**(1): 65-73.

Cliften, P., P. Sudarsanam, et al. (2003). "Finding functional features in Saccharomyces genomes by phylogenetic footprinting." Science **301**(5629): 71-76.
Costanzo, M., J. L. Nishikawa, et al. (2004). "CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast." Cell **117**(7): 899-913.
Dai, Z. and X. Dai (2012). "Nuclear colocalization of transcription factor target genes strengthens coregulation in yeast." Nucleic Acids Res **40**(1): 27-36.
Darieva, Z., A. Clancy, et al. (2010). "A competitive transcription factor binding mechanism determines the timing of late cell cycle-dependent gene expression." Mol Cell **38**(1): 29-40.
de Bruin, R. A., T. I. Kalashnikova, et al. (2006). "Constraining G1-specific transcription to late G1 phase: the MBF-associated corepressor Nrm1 acts via negative feedback." Mol Cell **23**(4): 483-496.
de Bruin, R. A., W. H. McDonald, et al. (2004). "Cln3 activates G1-specific transcription via phosphorylation of the SBF bound repressor Whi5." Cell **117**(7): 887-898.
Dekker, J., K. Rippe, et al. (2002). "Capturing chromosome conformation." Science **295**(5558): 1306-1311.
DeRisi, J., L. Penland, et al. (1996). "Use of a cDNA microarray to analyse gene expression patterns in human cancer." Nat Genet **14**(4): 457-460.
Di Talia, S., H. Wang, et al. (2009). "Daughter-specific transcription factors regulate cell size control in budding yeast." PLoS Biol **7**(10): e1000221.
Dirick, L., T. Bohm, et al. (1995). "Roles and regulation of Cln-Cdc28 kinases at the start of the cell cycle of Saccharomyces cerevisiae." Embo J **14**(19): 4803-4813.
Duan, Z., M. Andronescu, et al. "A three-dimensional model of the yeast genome." Nature **465**(7296): 363-367.
Elledge, S. J. and R. W. Davis (1990). "Two genes differentially regulated in the cell cycle and by DNA-damaging agents encode alternative regulatory subunits of ribonucleotide reductase." Genes Dev **4**(5): 740-751.
Eser, U., M. Falleur-Fettig, et al. (2011). "Commitment to a cellular transition precedes genome-wide transcriptional change." Mol Cell **43**(4): 515-527.
Ferrezuelo, F., N. Colomina, et al. (2010). "The transcriptional network activated by Cln3 cyclin at the G1-to-S transition of the yeast cell cycle." Genome Biology **11**(6).
Ferrezuelo, F., N. Colomina, et al. (2010). "The transcriptional network activated by Cln3 cyclin at the G1-to-S transition of the yeast cell cycle." Genome Biol **11**(6): R67.
Flick, K., D. Chapman-Shimshoni, et al. (1998). "Regulation of cell size by glucose is exerted via repression of the CLN1 promoter." Mol Cell Biol **18**(5): 2492-2501.
Fraser, P. and W. Bickmore (2007). "Nuclear organization of the genome and the potential for gene regulation." Nature **447**(7143): 413-417.
Georgi, A. B., P. T. Stukenberg, et al. (2002). "Timing of events in mitosis." Curr Biol **12**(2): 105-114.
Ghaemmaghami, S., W. K. Huh, et al. (2003). "Global analysis of protein expression in yeast." Nature **425**(6959): 737-741.

Gotta, M., T. Laroche, et al. (1996). "The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae." J Cell Biol **134**(6): 1349-1363.
Grigull, J. (2004). "Genome-Wide Analysis of mRNA Stability Using Transcription Inhibitors and Microarrays Reveals Posttranscriptional Control of Ribosome Biogenesis Factors." Molecular and Cellular biology **24**(12): 14.
Guan, Y., M. Dunham, et al. (2010). "Systematic planning of genome-scale experiments in poorly studied species." PLoS Comput Biol **6**(3): e1000698.
Haber, J. E. and W. Y. Leung (1996). "Lack of chromosome territoriality in yeast: promiscuous rejoining of broken chromosome ends." Proc Natl Acad Sci U S A **93**(24): 13949-13954.
Hartwell, L. H., J. Culotti, et al. (1974). "Genetic control of the cell division cycle in yeast." Science **183**(120): 46-51.
Hartwell, L. H. and T. A. Weinert (1989). "Checkpoints: controls that ensure the order of cell cycle events." Science **246**(4930): 629-634.
Hereford, L. M., M. A. Osley, et al. (1981). "Cell-cycle regulation of yeast histone mRNA." Cell **24**(2): 367-375.
Hoheisel, J. D. (2006). "Microarray technology: beyond transcript profiling and genotype analysis." Nat Rev Genet **7**(3): 200-210.
Holt, L. J., A. N. Krutchinsky, et al. (2008). "Positive feedback sharpens the anaphase switch." Nature **454**(7202): 353-357.
Janga, S. C., J. Collado-Vides, et al. (2008). "Transcriptional regulation constrains the organization of genes on eukaryotic chromosomes." Proc Natl Acad Sci U S A **105**(41): 15761-15766.
Jensen, L. J., T. S. Jensen, et al. (2006). "Co-evolution of transcriptional and post-translational cell-cycle regulation." Nature **443**(7111): 594-597.
Jin, Q. W., J. Fuchs, et al. (2000). "Centromere clustering is a major determinant of yeast interphase nuclear organization." J Cell Sci **113 (Pt 11)**: 1903-1912.
Johnson, D. G., K. Ohtani, et al. (1994). "Autoregulatory control of E2F1 expression in response to positive and negative regulators of cell cycle progression." Genes Dev **8**(13): 1514-1525.
Justman, Q. A., Z. Serber, et al. (2009). "Tuning the activation threshold of a kinase network by nested feedback loops." Science **324**(5926): 509-512.
Kalir, S., J. McClure, et al. (2001). "Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria." Science **292**(5524): 2080-2083.
Kato, M., N. Hata, et al. (2004). "Identifying combinatorial regulation of transcription factors and binding motifs." Genome Biol **5**(8): R56.
Kim, J. and J. H. Kim (2007). "Difference-based clustering of short time-course microarray data with replicates." BMC Bioinformatics **8**: 253.
Koch, C., T. Moll, et al. (1993). "A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase." Science **261**(5128): 1551-1557.
Kurdistani, S. K., S. Tavazoie, et al. (2004). "Mapping global histone acetylation patterns to gene expression." Cell **117**(6): 721-733.

Lam, F. H., D. J. Steger, et al. (2008). "Chromatin decouples promoter threshold from dynamic range." Nature **453**(7192): 246-250.
Laub, M. T., H. H. McAdams, et al. (2000). "Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle." Science **290**(5499): 2144-2148.
Levy, S., J. Ihmels, et al. (2007). "Strategy of Transcription Regulation in the Budding Yeast." PLoS ONE **2**(2): e250.
Lieberman-Aiden, E., N. L. van Berkum, et al. (2009). "Comprehensive mapping of long-range interactions reveals folding principles of the human genome." Science **326**(5950): 289-293.
Lopez-Aviles, S., O. Kapuy, et al. (2009). "Irreversibility of mitotic exit is the consequence of systems-level feedback." Nature **459**(7246): 592-595.
MacKay, V. L., B. Mai, et al. (2001). "Early cell cycle box-mediated transcription of CLN3 and SWI4 contributes to the proper timing of the G(1)-to-S transition in budding yeast." Mol Cell Biol **21**(13): 4140-4148.
Markey, M. P., S. P. Angus, et al. (2002). "Unbiased analysis of RB-mediated transcriptional repression identifies novel targets and distinctions from E2F action." Cancer Res **62**(22): 6587-6597.
Mateus, C. and S. V. Avery (2000). "Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry." Yeast **16**(14): 1313-1323.
Meaburn, K. J. and T. Misteli (2007). "Cell biology: chromosome territories." Nature **445**(7126): 379-781.
Misteli, T. (2004). "Spatial positioning; a new dimension in genome function." Cell **119**(2): 153-156.
Mogno, I., F. Vallania, et al. (2010). "TATA is a modular component of synthetic promoters." Genome Res **20**(10): 1391-1397.
Morgan, D. O. (2007). The cell cycle : principles of control. London
Sunderland, MA, New Science Press ;
Sinauer Associates.
Muller, H., A. P. Bracken, et al. (2001). "E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis." Genes Dev **15**(3): 267-285.
Nasmyth, K. and L. Dirick (1991). "The role of SWI4 and SWI6 in the activity of G1 cyclins in yeast." Cell **66**(5): 995-1013.
Oliva, A., A. Rosebrock, et al. (2005). "The cell cycle-regulated genes of Schizosaccharomyces pombe." PLoS Biol **3**(7): e225.
Orlando, D. A., C. Y. Lin, et al. (2008). "Global control of cell-cycle transcription by coupled CDK and network oscillators." Nature **453**(7197): 944-947.
Peddada, S. D. (2002). "Gene selection and clustering for time-course and dose–response microarray experiments using order-restricted inference." Bioinformatics **19**(7): 834-841.
Peng, X., R. K. Karuturi, et al. (2005). "Identification of cell cycle-regulated genes in fission yeast." Mol Biol Cell **16**(3): 1026-1042.

Planas-Silva, M. D. and R. A. Weinberg (1997). "The restriction point and control of cell proliferation." Curr Opin Cell Biol **9**(6): 768-772.
Pokholok, D. K., C. T. Harbison, et al. (2005). "Genome-wide map of nucleosome acetylation and methylation in yeast." Cell **122**(4): 517-527.
Pomerening, J. R., E. D. Sontag, et al. (2003). "Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2." Nat Cell Biol **5**(4): 346-351.
Pramila, T., W. Wu, et al. (2006). "The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle." Genes Dev **20**(16): 2266-2278.
Ren, B., H. Cam, et al. (2002). "E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints." Genes Dev **16**(2): 245-256.
Rustici, G., J. Mata, et al. (2004). "Periodic gene expression program of the fission yeast cell cycle." Nat Genet **36**(8): 809-817.
Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science **270**(5235): 467-470.
Shen-Orr, S. S., R. Milo, et al. (2002). "Network motifs in the transcriptional regulation network of Escherichia coli." Nature genetics **31**(1): 64-68.
Simonis, M., P. Klous, et al. (2006). "Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C)." Nat Genet **38**(11): 1348-1354.
Skotheim, J. M., S. Di Talia, et al. (2008). "Positive feedback of G1 cyclins ensures coherent cell cycle entry." Nature **454**(7202): 291-296.
Spellman, P. T., G. Sherlock, et al. (1998). "Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization." Mol Biol Cell **9**(12): 3273-3297.
Stuart, D. and C. Wittenberg (1995). "CLN3, not positive feedback, determines the timing of CLN2 transcription in cycling cells." Genes Dev **9**(22): 2780-2794.
Taddei, A., G. Van Houwe, et al. (2009). "The functional importance of telomere clustering: global changes in gene expression result from SIR factor dispersion." Genome Res **19**(4): 611-625.
Takizawa, T., P. R. Gudla, et al. (2008). "Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP." Genes Dev **22**(4): 489-498.
Tirosh, I., A. Weinberger, et al. (2006). "A genetic signature of interspecies variations in gene expression." Nat Genet **38**(7): 830-834.
Tyers, M., G. Tokiwa, et al. (1993). "Comparison of the Saccharomyces cerevisiae G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins." Embo J **12**(5): 1955-1968.
Whitfield, M. L., G. Sherlock, et al. (2002). "Identification of genes periodically expressed in the human cell cycle and their expression in tumors." Molecular Biology of the Cell **13**(6): 1977.
Whitfield, M. L., G. Sherlock, et al. (2002). "Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors." Molecular Biology of the Cell **13**(6): 1977-2000.

Wijnen, H., A. Landman, et al. (2002). "The G(1) cyclin Cln3 promotes cell cycle entry via the transcription factor Swi6." Mol Cell Biol **22**(12): 4402-4418.

Xiong, W. and J. E. Ferrell, Jr. (2003). "A positive-feedback-based bistable 'memory module' that governs a cell fate decision." Nature **426**(6965): 460-465.

Xu, X., M. Bieda, et al. (2007). "A comprehensive ChIP-chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members." Genome Res **17**(11): 1550-1561.

Yao, G., T. J. Lee, et al. (2008). "A bistable Rb-E2F switch underlies the restriction point." Nat Cell Biol **10**(4): 476-482.

Yung, Y., J. Walker, et al. (2007). "A Skp2 autoinduction loop and restriction point control." The Journal of cell biology **178**(5): 741.

Yung, Y., J. L. Walker, et al. (2007). "A Skp2 autoinduction loop and restriction point control." J Cell Biol **178**(5): 741-747.

Zaslaver, A., A. E. Mayo, et al. (2004). "Just-in-time transcription program in metabolic pathways." Nat Genet **36**(5): 486-491.

Zhao, Z., G. Tavoosidana, et al. (2006). "Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions." Nat Genet **38**(11): 1341-1347.