Molecular systems analysis of a cis-encoded epigenetic switch

By

Leah M. Octavio

B.S. Chemical Engineering
University of Rhode Island, 2006

SUBMITTED TO THE PROGRAM IN COMPUTATIONAL SYSTEMS BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN COMPUTATIONAL AND SYSTEMS BIOLOGY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Signature of author ___
Program in Computational Systems Biology
August 11, 2011

Certified by
Gerald R. Fink, Ph.D.
Professor of Biology
Thesis co-advisor

Certified by
Narendra Maheshri, Ph.D.
Assistant Professor of Chemical Engineering
Thesis co-advisor

Accepted by
Christopher B. Burge, Ph.D.
Professor of Biology and Biological Engineering

Molecular Systems Analysis of a Cis-encoded Epigenetic Switch

By

Leah M. Octavio

Submitted to the Computational Systems Biology Program
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computational Systems Biology

Abstract An ability to control the degree of heterogeneity in cellular phenotypes may be important for cell populations to survive uncertain and ever-changing environments or make cell-fate decisions in response to external stimuli. Cells may control the degree of gene expression heterogeneity and ultimately levels of phenotypic heterogeneity by modulating promoter switching dynamics. In this thesis, I investigated various mechanisms by which heterogeneity in the expression of *FLO11* in *S. cerevisiae* could be generated and controlled. First, we show that two copies of the *FLO11* locus in *S. cerevisiae* switch between a silenced and competent promoter state in a random and independent fashion, implying that the molecular event leading to the transition occurs in *cis*. Through further quantification of the effect of *trans* regulators on both the slow epigenetic transitions between a silenced and competent promoter state and the fast promoter transitions associated with conventional regulation of *FLO11*, we found different classes of regulators affect epigenetic, conventional, or both forms of regulation. Distributing kinetic control of epigenetic silencing and conventional gene activation offers cells flexibility in shaping the distribution of gene expression and phenotype within a population. Next, we demonstrate how multiple molecular events occurring at a gene's promoter could lead to an overall slow step in *cis*. At the *FLO11* promoter, we show that at least two pathways that recruit histone deacetylases to the promoter and *in vivo* association between the region -1.2 kb from the ATG start site of the *FLO11* ORF and the core promoter region are all required for a stable silenced state. To generate bimodal gene expression, the activator Msn1p forms an alternate looped conformation, where the core promoter associates with the non-coding RNA *PWR1*'s promoter and terminator regions, located at -2.1 kb and -3.0 kb from the ATG start site of the *FLO11* ORF respectively. Formation of the active looped conformation is required for Msn1p's ability to stabilize the competent state without destabilizing the silenced state and generate a bimodal response. Our results support a model where multiple stochastic steps at the promoter are required to transition between the silenced and active states, leading to an overall slow step in *cis*. Finally, preliminary investigations of heterozygous diploids revealed possible transvection occurring at *FLO11*, where a silenced allele of *FLO11* appeared to transfer silencing factors to a desilenced *FLO11* allele on the homologous chromosome. These observations suggest a new mechanism through which heterogeneity in *FLO11* expression could be further controlled, in addition to the molecular events at the *FLO11* promoter we elucidated previously.

Thesis co-advisor: Gerald R. Fink, Ph.D.
Title: American Cancer Society Professor of Genetics, M.I.T.
Thesis co-advisor: Narendra Maheshri, Ph.D.
Title: Assistant Professor of Chemical Engineering, M.I.T.

ACKNOWLEDGEMENTS

I am grateful to my advisors, Prof. Gerald Fink and Prof. Narendra Maheshri, for unparalleled training in systems biology and yeast genetics, and for being my mentors. While most mentors go the extra mile for their students, Narendra and Gerry have gone an extra ten thousand miles for me. For this, I can never thank them enough.

Narendra is an incredible mentor, with endless energy, enthusiasm and passion for his work. Carrying out research in systems biology, a very young and interdisciplinary field, is uniquely challenging. Narendra not only performs research that pushes boundaries of systems biology but also possesses unshakeable faith in his work that has inspired me to believe in myself and to push the limits of what I can achieve.

Gerry is a brilliant biologist, who possesses a rare ability to conceive and pursue crazy ideas while employing good judgment and sense at the same time. He is also a very wise mentor. I have learned so much from his vast knowledge and experience in yeast biology, his impressive intuition and his rich, heterogeneous life experiences. I thank him for making me a stronger and wiser person. He is not only one of the most brilliant, but also one of the kindest, most patient and most open-minded professors I have known, and I would never have made it this far without what he has done for me.

I am also grateful to Prof. Chris Burge, for his assistance and tremendous support throughout my graduate career. I would not have gone far without his help. I also thank Prof. Kris Prather, Prof. Bruce Tidor, Prof. Arup Chakraborty and Prof. Alan Grossman, for being my informal mentors. I greatly appreciate their advice and the generosity of their time.

I thank my thesis committee members, Prof. Ernest Fraenkel and Prof. Alexander van Oudenaarden, for their valuable guidance and support for my thesis project and graduate career, and for allowing me to benefit from their expertise and experience in systems biology research.

I am very thankful to the Fink lab members. The Finklets are an amazingly talented group of people and have provided an extremely stimulating and supportive environment to do research in. I especially thank Stacie Bumgarner. Much of my thesis has built upon her graduate and post-graduate work on *FLO11*. Her comments and insights on my project were invaluable, and her curiosity and passion for unraveling the intricacies of the *FLO11* promoter was infectious. I also thank the Maheshri lab members, who have provided much technical assistance, feedback and reagents instrumental to my thesis project. In particular, I owe many thanks to Kamil Gedeon, who helped me carry out strain construction and experiments in Chapter 2.

Finally, I am enormously grateful to members of the giggle squad: Jaimie Lee, Tracy Washington and Mei Lyn Ong. I am forever thankful to them for their unconditional friendship and never-ending support, and, most of all, for believing me in those times when it seemed like nobody else did.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Gene silencing in yeast and generation of phenotypic heterogeneity in microbes

Chromatin states in eukaryotes can be generally divided into heterochromatin and euchromatin. Originally, heterochromatin was defined as the deeply condensed regions of the chromosomes observed throughout the cell cycle (Schultz, 1936). Heterochromatic regions are characterized by hypoacetylated H3 and H4 histones (Braunstein, Rose, et al, 1993; Suka, Suka, et al, 2001) (Schultz, 1936) that associate with silent information regulator (SIR) proteins (Sir2p, Sir3p, and Sir4p in yeast), forming a dense, compact structure with highly ordered nucleosomes known as "silenced" chromatin (Hecht, Strahl-Bolsinger, and Grunstein, 1996; Lieb, Liu, et al, 2001; Ravindra, Weiss, and Simpson, 1999; Rusche, Kirchmaier, and Rine, 2002; Strahl-Bolsinger, Hecht, et al, 1997; Weiss, and Simpson, 1998; Zhang, Hayashi, et al, 2002). In contrast, euchromatic regions have no densely packed, ordered structure, and are generally more accessible to transcriptional machinery than heterochromatin (Gottschling, 1992; Singh, and Klar, 1992).

In *S. cerevisiae*, classic heterochromatic or "silenced" regions comprise the telomeres, the mating type loci, and the ribosomal DNA genes (Rusche, Kirchmaier, and Rine, 2003). Telomeres and mating type loci both require SIR proteins for silencing, although the mechanisms to initiate silencing at these two regions are slightly different. At the mating type loci, Rap1p, Abf1p and the origin recognition complex (ORC) bind cooperatively to silencer sites and recruit Sir2p, (which deacetylates histones in the region) to nucleate silencing. Sir3p and Sir4p then bind to the deacetylated histones in a complex with Sir2p, stably maintaining the silenced state

(Hoppe, Tanny, et al, 2002; Luo, Vega-Palas, and Grunstein, 2002; Moretti, Freeman, et al, 1994; Moretti, and Shore, 2001; Rusche, Kirchmaier, and Rine, 2002; Zhang, Hayashi, et al, 2002; Rusche, Kirchmaier, and Rine, 2003). At the telomeres, a similar structure of Sir2p/Sir3p/Sir4p in a complex with hypoacetylated histones exists, but only Rap1p and another protein, yKu70p, are required to initiate the formation of the silenced state (Martin, Laroche, et al, 1999; Mishra, and Shore, 1999; Tsukamoto, Kato, and Ikeda, 1997).

Cells may utilize silenced states of genes as a strategy to exhibit heterogeneity or selectivity in gene expression, leading to an effective "switch" between particular phenotypes that could be beneficial to their fitness. In many cases, genes that are silenced and are expressed only in selected conditions are located in discrete regions of the chromosome. In the parasite *Trypanosoma bruceii,* for example, expression of variant surface glycoproteins (VSG) in the human bloodstream is allele exclusive, where among the hundreds of VSG alleles encoded in its genome, only a single allele is expressed and the rest are silenced (Horn, and Cross, 1997; Pays, 2005). Switching expression of one VSG to another enables *T. bruceii* to escape host immune surveillance; the process involves duplication of a previously silent allele and recombination into the active expression site at the telomeres (Barry, and McCulloch, 2001; Borst, Bitter, et al, 1998; Cross, Wirtz, and Navarro, 1998; Pays, 2005; Vanhamme, Lecordier, and Pays, 2001). Meanwhile, in the agent responsible for human malaria, *Plasmodium falciparum*, similar antigenic variation is exhibited. Here, the expression of membrane proteins encoded by the *var* gene family is also allele exclusive, although the mode of switching involves no DNA recombination or deletions (Chen, Fernandez, et al, 1998; Scherf, Hernandez-Rivas, et al, 1998). While the molecular mechanism to achieve exclusive allele expression is unknown, it is believed that an epigenetic mechanism may be involved as it has been observed that *var* promoter

silencing depends on its interactions with certain introns, and remodeling of *var* promoters by histone modifications could also silence expression (Dzikowski, Li, et al, 2007; Frank, Dzikowski, et al, 2006; Voss, Tonkin, et al, 2007). In the pathogenic yeast *C. glabrata*, the *EPA* gene family encodes adhesins critical to establishing a virulent phenotype and is silenced in a Sir2p-dependent manner (De Las Penas, Pan, et al, 2003). It has been demonstrated that in nicotinic acid (NA) limited environments, intracellular levels of NA in the NA auxotroph *C. glabrata* decrease, leading to lower nicotinamide adenine dinucleotide (NAD) levels and a subsequent reduction in the activity of the NAD-dependent histone deacetylase Sir2p (Domergue, Castano, et al, 2005). This suggests that during the course of infecting its host, *C. glabrata* may be able to sense NA-limited conditions in its host environment as a cue to initiate the switch to an adhesive phenotype.

In *S. cerevisiae*, a number of multigene families reside in subtelomeric regions, with most of these genes possessing functions required only in sub-optimal growth conditions. One such family comprises of the *FLO* genes (*FLO1*, *FLO5*, *FLO9*, *FLO10* and *FLO11*), which encode adhesins that localize to the cell surface (Teunissen, and Steensma, 1995). All the *FLO* genes are located in the subtelomeres, except *FLO11*, which is not in the telomeres or subtelomeres. Currently, mechanisms known to generate variation in cell-surface adhesins include high mutation frequency in upstream regulators (Halme, Bumgarner, et al, 2004), ploidy regulation (Braus, Grundmann, et al, 2003; Galitski, Saldanha, et al, 1999), recombination and mutations within the adhesins and promoter (Fidalgo, Barrales, et al, 2006) and epigenetic silencing in haploids (Halme, Bumgarner, et al, 2004). Sequence polymorphisms in the Ras GTPase activating proteins *IRA1* and *IRA2* lead to activation of the normally silent *FLO10* and altered cell-surface properties (Halme, Bumgarner, et al, 2004). In haploid and diploid cells of the

Σ1278b background, significant differences between *FLO11* induction were observed during nitrogen and amino acid starvation, where expression was generally higher in haploids (Braus, Grundmann, et al, 2003). Meanwhile, in a study dissecting the origin of a highly buoyant phenotype ("flor" phenotype) observed in yeast used to make sherry wine, adaptive mutations in both the *FLO11* promoter and open reading frame were found (Fidalgo, Barrales, et al, 2006). Two genetic changes: a 111-bp deletion in the *FLO11* promoter and an increase in the number of tandem repeats in the central domain of the *FLO11* ORF, increased the expression and hydrophobicity of Flo11p, conferring the floatable phenotype observed.

Recently, epigenetic silencing at *FLO11* was observed in haploid *S. cerevisiae* cells (Halme, Bumgarner, et al, 2004). Under expression-inducing conditions, a mixed population of cells with *FLO11* expression "ON" and "OFF" was observed, where the metastable silenced state was heritable for about ten generations. However, *FLO11* does not reside in telomeric or subtelomeric regions, and silencing at *FLO11* is SIR-independent, making the molecular mechanism of its silencing intriguing.

Timescale of promoter dynamics and effect on gene expression / phenotypic heterogeneity

Epigenetics is the inheritance of a gene expression state without changes to DNA sequence. Therefore, an epigenetic event may also be described in terms of the dynamics of gene expression. If a gene switches between an "OFF" and "ON" state at a rate slower than the cell's division rate, then the gene's expression state is inherited.

The steps leading to expression of a gene can be broken down as follows: the promoter switches from an inactive state to an active state, then mRNA is transcribed from the active state, and then the mRNA is translated into protein, which is then diluted by cell division and degraded by other

mechanisms. Transcription and translation of mRNA occur on a timescale faster than the cell division time, whereas the promoter switching dynamics may be slower or faster than the cell division time. "Conventional" gene regulation occurs when the promoter dynamics are faster than the rate of cell division, whereas "epigenetic" gene regulation refers to promoter dynamics slower than the rate of cell division. Essentially, when promoter dynamics are slower than the cell division rate, the cell population exhibits a distinct fraction of "OFF" cells (cells not expressing the gene) and "ON" cells (cells highly expressing the gene). This is because when a gene that was previously expressing high protein level switches "OFF" and remains "OFF" for at least several generations, the protein levels will be diluted away sufficiently so the cell will appear "OFF" in protein expression. The fraction of cells in the population that happen to have the promoter "OFF" at any given time will also appear "OFF" for protein expression. However, when a promoter switches faster than the cell division rate, the proteins transcribed during the time the promoter was "ON" will not be diluted away sufficiently during the short time that the promoter is "OFF". Hence, with fast promoter dynamics, all the cells in the population, with slight variations, generally express the same mean level of protein. Therefore, by controlling the dynamics of promoter switching in response to upstream signals, a cell population can choose to express the gene in a variegated (bimodal) fashion or a graded (unimodal) fashion, which may have important consequences in the ability of the cell population to thrive in their particular environment.

In Chapter 2, I explore the slow and fast promoter dynamics at the *FLO11* gene in the budding yeast *S. cerevisiae*. I will show that the slow promoter switching is encoded in *cis*, and that the various trans regulators of *FLO11* affect different combinations of fast and slow promoter switching rates, enabling the cell population to effectively tune their levels of *FLO11* expression

heterogeneity. Finally, in Chapter 3, I investigate how some of the *FLO11 trans* regulators orchestrate molecular events at the promoter to lead to an overall slow promoter switching rate.

Regulation of transcription in yeast by DNA looping

Long-range DNA interactions, or DNA looping, have been shown to regulate gene expression in several organisms, including yeast. In *S. cerevisiae*, gene looping was first demonstrated by O'Sullivan et al (O'Sullivan, Tan-Wong, et al, 2004), who showed that the promoter and terminator regions of the genes *FMP27* and *SEN1* associated *in vivo*, forming a looped conformation. More genes in *S. cerevisiae* were later shown to also exist in looped conformations, including genes with ORF's as short as 1 kb (Singh, and Hampsey, 2007). Recently, Rodley et al performed a global 3C experiment to assay for possible associations between all genomic loci in *S. cerevisiae*, and reported a large number of long-range inter- and intrachromosomal associations (Rodley, Bertels, et al, 2009).

The looped conformation of a gene can be dependent on its transcriptional state. O'Sullivan et al observed that gene looping was abolished in Kin28p mutants (O'Sullivan, Tan-Wong, et al, 2004). Coordination of transcriptional elongation is achieved by differential phosphorylation of heptad repeat at Ser2 and Ser5 at the C-terminal domain of RNA Pol II. Since Kin28p phosphorylates Ser5 early during elongation, they hypothesized that gene loops formed during early transcriptional activation. However, Ansari and Hampsey later found that the juxtaposition of the promoter and terminator regions also required interactions between TFIIB and Ssu72p, a component of the 3' end processing complex. Since the interactions with the 3' end processing machinery was necessary for gene looping, they argued that loops did not form during early

transcription activation, but rather, after at least a round of transcription had occurred (Ansari, and Hampsey, 2005).

Do these gene loops play any physiological role? Laine et al showed that gene looping at *GAL10* was associated with rapid reactivation kinetics, where the looped conformation is thought to facilitate rapid re-association of RNA Pol II with the *GAL 10* promoter upon reactivation (Laine, Singh, et al, 2009). In a study by Tan-Wong et al, gene loops formed at *HXK1* upon transcriptional activation were found to associate with the nuclear pore complex (NPC) (Tan-Wong, Wijayatilake, and Proudfoot, 2009). The gene loop structure at *HXK1* was found to persist for up to an hour after repression. If *HXK1* was reactivated during this 1-hour period, RNA Pol II was recruited to the promoter faster and therefore, transcription initiation was faster. The gene loop's association with the NPC was dependent on Mlp1p, a protein found in the intranuclear filaments of the NPC basket. Disruption of Mlp1p abolished tethering of the gene loop to the NPC and disrupted the looped gene structure, resulting in slower reactivation kinetics of *HXK1*.

In Chapter 3, I will show how the stable silenced and stable active states at *FLO11* are associated with different looped conformations of the promoter. The silenced and active looped conformation are dependent on *FLO11*-specific trans regulators and ncRNA transcription at the *FLO11* promoter, and I will argue that the looped conformations at *FLO11* are important for the stability of the silenced and active states, which give rise to the slow promoter dynamics observed.

Transvection in yeast

Transvection is a phenomenon where an allele on a chromosome interacts with its corresponding allele on the homologous chromosome (Lewis, 1954). A classic assay for transvection is to check whether a WT promoter of a mutant ORF, which in a homozygous background produces a mutant phenotype, could complement a mutant promoter of a WT ORF, which also results in a mutant phenotype in a homozygous background. If transvection occurred, a WT phenotype would be recovered in the heterozygous background. Transvection is dependent upon chromosomal pairing, and rearrangements which disrupt homologous pairing will prevent complementation (Pirrotta, 1999).

Complementation interactions observed during transvection often resulted in activation of the wild type coding region. For example, in Drosophila, the gene *yellow*, which is expressed in the body and the wing, is activated by two upstream enhancers. A mutant allele of yellow called y^2 contains a gypsy retrotransposon insertion between the enhancers and the *yellow* promoter that results in inactive *yellow*. The y^{3c3} allele was found to complement the y^2 allele and restore yellow expression. Interestingly, the y^{3c3} allele contained a mutation where the entire regulatory region and part of the coding region of *yellow* was deleted, so *yellow* was not expressed in a homozygous y^{3c3} background. In the heterozygous y^2/y^{3c3} background, however, the y^{3c3} allele was found to act in *trans* to induce chromatin conformation changes in the y^2 allele, allowing the distal enhancers to come into contact with the *yellow* promoter and activate expression (Geyer, Green, and Corces, 1990; Morris, Chen, et al, 1998).

While not as frequently observed, *trans* interactions between alleles which resulted into repression have also been reported. The bw^D allele in Drosophila, for instance, is a wild type *brown* coding region with a heterochromatic sequence inserted nearby, which silences expression of *brown*. When bw^D is homologously paired with wild type *brown*, the wild type allele becomes

silenced (Csink, and Henikoff, 1996; Dernburg, Broman, et al, 1996). Similarly, *trans* interaction between Polycomb response elements (PRE) have also been shown to occur. PRE (Polycomb response elements) are sites where Polycomb protein complexes form and silence nearby regions, and pairing of two PRE sequences on homologous chromosomes was found to enhance silencing (Chan, Rastelli, and Pirrotta, 1994; Fauvarque, and Dura, 1993; Kassis, 1994).

While most examples of transvection have been described in Drosophila, Aramayo and Metzenberg showed that transvection occurred in the fungus *Neurospora crassa* (Aramayo, and Metzenberg, 1996), a filamentous haploid ascomycete that exists in two non-switching mating types. In Neurospora, Asm-1 is a key regulator of sexual development. Deletion of Asm-1 results in an ascus-dominant phenotype, where nearly all the progeny are immature, colorless, and inviable spores. When a wild-type Asm-1 copy was inserted in another genomic locus in the *asm-1Δ* background and then mated with *asm-1Δ*, the ascus-dominant phenotype was observed. Further experiments showed that crosses with an Asm-1 copy in the normal location and the other copy in another chromosomal location resulted in infertility, whereas crosses with both copies in the second location were fertile. Aramayo and Metzenberg concluded that transvection must occur for normal regulation of Asm-1. The Asm-1 gene must be close to its homolog to function in diplophase, and transvection happened in a narrow window of time during which homologous chromosomes paired prior to crossover.

So far, transvection has been shown to occur in mammalian cells (Liu, Huang, et al, 2008; Rassoulzadegan, Magliano, and Cuzin, 2002; Sandhu, Shi, et al, 2009), Drosophila (Duncan, 2002; Muller, and Schaffner, 1990; Pirrotta, 1999) and Neurospora (Aramayo, and Metzenberg, 1996), but has never been reported to occur in *S. cerevisiae*. In Chapter 4, I will discuss experiments done to investigate the possibility of transvection at *FLO11*. Preliminary data were

consistent with a *trans* interaction occurring between a wild type silenced allele and a desilenced mutant allele in the heterozygous background. Current data support a model in which the silenced allele confers silencing to the desilenced copy, likely upon homologous chromosomal pairing.

The "systems biology" of the *FLO11* promoter

One of the major goals of the systems biology community is to understand how the many various components of a biological entity work together to achieve an output. For example, a gene regulatory network in a cell was often thought of as analogous to an electrical circuit, where a network of genes could process an input signal to produce an output. Depending on how the specific configurations in which the genes are "wired" together, the network may discriminate between signals of varying levels or frequencies, and it may produce an output either directly proportional or non-linearly dependent on the level of the input signal. Different dynamical behaviors may also be achieved by the gene network (for example, bistability, time delays or oscillations) when the appropriate feedback loops are incorporated into the system. Finally, also depending on the "wiring" of the genes in the system, the gene network may exhibit emergent "systems-level" properties, such as robustness or sensitivity to perturbations or fluctuations in levels of individual components and/or input signals.

While much attention has been focused on the systems properties and dynamical behavior of gene networks, I will argue that the same properties and behaviors can also be encoded at the level of a single gene promoter. My thesis focuses on the dynamics of *FLO11* and unravels some of the molecular mechanisms at the promoter which encode slow and fast timescales of promoter switching.

REFERENCES

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes Dev. *24,* 2969-2978.

Aramayo, R., and Metzenberg, R.L. (1996). Meiotic transvection in fungi. Cell *1,* 103-113.

Barry, J.D., and McCulloch, R. (2001). Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. Adv. Parasitol. 1-70.

Borst, P., Bitter, W., Blundell, P.A., Chaves, I., Cross, M., Gerrits, H., van Leeuwen, F., McCulloch, R., Taylor, M., and Rudenko, G. (1998). Control of VSG gene expression sites in Trypanosoma brucei. Mol. Biochem. Parasitol. *1,* 67-76.

Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D., and Broach, J.R. (1993). Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. Genes Dev. *4,* 592-604.

Braus, G.H., Grundmann, O., Bruckner, S., and Mosch, H.U. (2003). Amino acid starvation and Gcn4p regulate adhesive growth and FLO11 gene expression in Saccharomyces cerevisiae. Mol. Biol. Cell *10,* 4272-4284.

Chan, C.S., Rastelli, L., and Pirrotta, V. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. *11,* 2553-2564.

Chen, Q., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P., and Wahlgren, M. (1998). Developmental selection of var gene expression in Plasmodium falciparum. Nature *6691,* 392-395.

Cross, G.A., Wirtz, L.E., and Navarro, M. (1998). Regulation of vsg expression site transcription and switching in Trypanosoma brucei. Mol. Biochem. Parasitol. *1,* 77-91.

Csink, A.K., and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in Drosophila. Nature *6582,* 529-531.

De Las Penas, A., Pan, S.J., Castano, I., Alder, J., Cregg, R., and Cormack, B.P. (2003). Virulence-related surface glycoproteins in the yeast pathogen Candida glabrata are encoded in subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. Genes Dev. *18,* 2245-2258.

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard, D.A., and Sedat, J.W. (1996). Perturbation of nuclear architecture by long-distance chromosome interactions. Cell *5,* 745-759.

Domergue, R., Castano, I., De Las Penas, A., Zupancic, M., Lockatell, V., Hebel, J.R., Johnson, D., and Cormack, B.P. (2005). Nicotinic acid limitation regulates silencing of Candida adhesins during UTI. Science *5723,* 866-870.

Duncan, I.W. (2002). Transvection effects in Drosophila. Annu. Rev. Genet. 521-556.

Dzikowski, R., Li, F., Amulic, B., Eisberg, A., Frank, M., Patel, S., Wellems, T.E., and Deitsch, K.W. (2007). Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO Rep. *10,* 959-965.

Fauvarque, M.O., and Dura, J.M. (1993). polyhomeotic regulatory sequences induce developmental regulator-dependent variegation and targeted P-element insertions in Drosophila. Genes Dev. *8,* 1508-1520.

Fidalgo, M., Barrales, R.R., Ibeas, J.I., and Jimenez, J. (2006). Adaptive evolution by mutations in the FLO11 gene. Proc. Natl. Acad. Sci. U. S. A. *30,* 11228-11233.

Frank, M., Dzikowski, R., Costantini, D., Amulic, B., Berdougo, E., and Deitsch, K. (2006). Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. J. Biol. Chem. *15,* 9942-9952.

Galitski, T., Saldanha, A.J., Styles, C.A., Lander, E.S., and Fink, G.R. (1999). Ploidy regulation of gene expression. Science *5425,* 251-254.

Geyer, P.K., Green, M.M., and Corces, V.G. (1990). Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. EMBO J. *7,* 2247-2256.

Gottschling, D.E. (1992). Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo. Proc. Natl. Acad. Sci. U. S. A. *9,* 4062-4065.

Halme, A., Bumgarner, S., Styles, C., and Fink, G.R. (2004). Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. Cell *3,* 405-415.

Hecht, A., Strahl-Bolsinger, S., and Grunstein, M. (1996). Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. Nature *6595,* 92-96.

Hoppe, G.J., Tanny, J.C., Rudner, A.D., Gerber, S.A., Danaie, S., Gygi, S.P., and Moazed, D. (2002). Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. Mol. Cell. Biol. *12,* 4167-4180.

Horn, D., and Cross, G.A. (1997). Position-dependent and promoter-specific regulation of gene expression in Trypanosoma brucei. EMBO J. *24,* 7422-7431.

Kassis, J.A. (1994). Unusual properties of regulatory DNA from the Drosophila engrailed gene: three "pairing-sensitive" sites within a 1.6-kb region. Genetics *3,* 1025-1038.

Laine, J.P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. Genes Dev. *22,* 2604-2609.

Lewis EB (1954). "The theory and application of a new method of detecting chromosomal rearrangements in Drosophila melanogaster". *Am. Nat.* 88: 225–239.

Lieb, J.D., Liu, X., Botstein, D., and Brown, P.O. (2001). Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. Nat. Genet. *4,* 327-334.

Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B. *et al.* (2008). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J. Exp. Med. *8,* 1843-1858.

Luo, K., Vega-Palas, M.A., and Grunstein, M. (2002). Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. Genes Dev. *12,* 1528-1539.

Martin, S.G., Laroche, T., Suka, N., Grunstein, M., and Gasser, S.M. (1999). Relocalization of telomeric Ku and SIR proteins in response to DNA strand breaks in yeast. Cell *5,* 621-633.

Mishra, K., and Shore, D. (1999). Yeast Ku protein plays a direct role in telomeric silencing and counteracts inhibition by rif proteins. Curr. Biol. *19,* 1123-1126.

Moretti, P., Freeman, K., Coodly, L., and Shore, D. (1994). Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. Genes Dev. *19,* 2257-2269.

Moretti, P., and Shore, D. (2001). Multiple interactions in Sir protein recruitment by Rap1p at silencers and telomeres in yeast. Mol. Cell. Biol. *23,* 8082-8094.

Morris, J.R., Chen, J.L., Geyer, P.K., and Wu, C.T. (1998). Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. Proc. Natl. Acad. Sci. U. S. A. *18,* 10740-10745.

Muller, H.P., and Schaffner, W. (1990). Transcriptional enhancers can act in trans. Trends Genet. *9,* 300-304.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. Nat. Genet. *9,* 1014-1018.

Pays, E. (2005). Regulation of antigen gene expression in Trypanosoma brucei. Trends Parasitol. *11,* 517-520.

Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim. Biophys. Acta *1,* M1-8.

Rassoulzadegan, M., Magliano, M., and Cuzin, F. (2002). Transvection effects involving DNA methylation during meiosis in the mouse. EMBO J. *3,* 440-450.

Ravindra, A., Weiss, K., and Simpson, R.T. (1999). High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating-type locus HMRa. Mol. Cell. Biol. *12,* 7944-7950.

Rodley, C.D., Bertels, F., Jones, B., and O'Sullivan, J.M. (2009). Global identification of yeast chromosome interactions using Genome conformation capture. Fungal Genet. Biol. *11,* 879-886.

Rusche, L.N., Kirchmaier, A.L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. Annu. Rev. Biochem. 481-516.

Rusche, L.N., Kirchmaier, A.L., and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. Mol. Biol. Cell *7,* 2207-2222.

Sandhu, K.S., Shi, C., Sjolinder, M., Zhao, Z., Gondor, A., Liu, L., Tiwari, V.K., Guibert, S., Emilsson, L., Imreh, M.P., and Ohlsson, R. (2009). Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. Genes Dev. *22,* 2598-2603.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998). Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. EMBO J. *18,* 5418-5426.

Schultz, J. (1936). Variegation in Drosophila and the Inert Chromosome Regions. Proc. Natl. Acad. Sci. U. S. A. *1,* 27-33.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Mol. Cell *5,* 806-816.

Singh, J., and Klar, A.J. (1992). Active genes in budding yeast display enhanced in vivo accessibility to foreign DNA methylases: a novel in vivo probe for chromatin structure of yeast. Genes Dev. *2,* 186-196.

Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. (1997). SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. Genes Dev. *1,* 83-93.

Suka, N., Suka, Y., Carmen, A.A., Wu, J., and Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Mol. Cell *2,* 473-479.

Tan-Wong, S.M., Wijayatilake, H.D., and Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes Dev. *22,* 2610-2624.

Teunissen, A.W., and Steensma, H.Y. (1995). Review: the dominant flocculation genes of Saccharomyces cerevisiae constitute a new subtelomeric gene family. Yeast *11,* 1001-1013.

Tsukamoto, Y., Kato, J., and Ikeda, H. (1997). Silencing factors participate in DNA repair and recombination in Saccharomyces cerevisiae. Nature *6645,* 900-903.

Vanhamme, L., Lecordier, L., and Pays, E. (2001). Control and function of the bloodstream variant surface glycoprotein expression sites in Trypanosoma brucei. Int. J. Parasitol. *5-6,* 523-531.

Voss, T.S., Tonkin, C.J., Marty, A.J., Thompson, J.K., Healer, J., Crabb, B.S., and Cowman, A.F. (2007). Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. Mol. Microbiol. *1,* 139-150.

Weiss, K., and Simpson, R.T. (1998). High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating type locus HMLalpha. Mol. Cell. Biol. *9,* 5392-5403.

Zhang, Z., Hayashi, M.K., Merkel, O., Stillman, B., and Xu, R.M. (2002). Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing. EMBO J. *17,* 4600-4611.

Chapter 2

EPIGENETIC AND CONVENTIONAL REGULATION BY FLO11 ACTIVATORS ALLOW TUNING OF ITS EXPRESSION HETEROGENEITY

Author's Note: The material in this chapter was originally published as: Leah M. Octavio, Kamil Gedeon, and Narendra Maheshri (2009) "Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression." PLoS Genetics, Oct.5(10):e1000673.

I contributed to the work by constructing most of the strains, designing and carrying out all the experiments and analyzing all the data, with the assistance of my undergraduate research student, Kamil Gedeon.

ABSTRACT

Epigenetic switches encode their state information either locally, often via covalent modification of DNA or histones; or globally, usually in the level of a *trans*-regulatory factor. Here we examine how the regulation of *cis*-encoded epigenetic switches controls the extent of heterogeneity in gene expression, which is ultimately tied to phenotypic diversity in a population. We show that two copies of the *FLO11* locus in *S. cerevisiae* switch between a silenced and competent promoter state in a random and independent fashion, implying that the molecular event leading to the transition occurs locally at the promoter, in *cis*. We further quantify the effect of *trans* regulators on both the slow epigenetic transitions between a silenced and competent promoter state and the fast promoter transitions associated with conventional regulation of *FLO11*. We find different classes of regulators affect epigenetic, conventional, or both forms of regulation. Distributing kinetic control of epigenetic silencing and conventional gene activation offers cells flexibility in shaping the distribution of gene expression and phenotype within a population.

INTRODUCTION

Microbial cell populations employ a number of strategies to rapidly generate phenotypic diversity on relatively short time scales (Avery, 2006; Rando and Verstrepen, 2007). In some microbes, genes known as contingency loci contain tandem repeats of DNA whose recombination results in turning expression ON or OFF (Barry, Ginger, et al, 2003). Other genetic strategies include the directed recombination of silent alleles into a particular active locus, as is the case for mating type switching in yeasts and surface antigen expression in *T. brucei* (Pays, Vanhamme, and Perez-Morga, 2004), the causative agent of African sleeping sickness. Another widely used strategy that generates phenotypic heterogeneity in clonal microbial cell populations is epigenetic gene regulation. In contrast to genetic strategies, this refers to the heritable change in gene's expression that is not caused by changes in the underlying gene sequence. For example, the parasite *P. falciparum* (malaria) and the model organisms *S. cerevisiae* and *E. coli* use epigenetic mechanisms to variably express antigenic cell-surface proteins (Avery, 2006) and possibly escape immune surveillance and/or survive in a unpredictably changing environment.

Many epigenetically regulated genes can be considered switches as they have two heritable expression states, "ON" and "OFF." A stable epigenetic marker maintains each state and can be encoded in *cis* or in *trans*. The molecular basis of local, *cis* markers involve covalent modifications of DNA or DNA-associated proteins. These include DNA methylation (Low, Weyand, and Mahan, 2001) and histone modifications that define silenced heterochromatin or active euchromatin in eukaryotes (Wu, and Grunstein, 2000). Global, *trans* markers are often transcription factor activity; the mechanism for stable, slow switching of these levels is positive or double negative feedback loops that generate heritable bistable gene expression states

associated with high or low levels of transcription factor activity (Gardner, Cantor, and Collins, 2000; Kaufmann, and van Oudenaarden, 2007; Xiong, and Ferrell, 2003). Switches using either scheme respond to environmental factors, but heterogeneity is observed even with constant environmental conditions, suggesting that the switch can rarely and randomly be toggled due to fluctuations in the intracellular environment. The two schemes can be also combined. For example, in uropathogenic *E. coli* the expression of pyelonephritis-associated pili is regulated by an epigenetic switch that maintains its state through both DNA methylation and a positive feedback loop (Hernday, Braaten, et al, 2004).

The control of phenotypic heterogeneity is arguably as important as its rapid generation. Heterogeneity, or noise, in conventionally regulated gene expression has been well-studied in recent years. Single cell and single molecule studies have revealed that gene activation occurs in random, intermittent transcriptional bursts (Cai, Friedman, and Xie, 2006; Chubb, Trcek, et al, 2006; Golding, Paulsson, et al, 2005; Yu, Xiao, et al, 2006) due to fast promoter fluctuations (> once per cell cycle) between an inactive (but competent) and active promoter state. Mechanistically, this is an oversimplification as the promoter likely adopts a series of different states involving binding of various gene-specific and general transcriptional machinery that lead to productive transcription. Here, the active promoter state can be thought of as one where rapid initiation and reinitiation is possible. For example, for regulatable RNA Pol II–dependent promoters, transcriptional initiation is often rate-limiting and hence the active promoter state corresponds to pre-initiation complex formation. Expression heterogeneity caused by even these fast fluctuations can have consequences on phenotype and population-level fitness (Blake, Balazsi, et al, 2006).

Noise in gene expression can be partitioned depending on whether its source is intrinsic or extrinsic to the process of gene expression. Intrinsic noise is due to the random nature of chemical transformations, including transcription and translation events. However, the random bursts of transcription thought to be associated with fast promoter fluctuations occurring in *cis* appear to be the dominant source of intrinsic noise in eukaryotes (Kaern, Elston, et al, 2005; Maheshri, and O'Shea, 2007). Extrinsic noise is due to cell-to-cell variation in *trans* factors affecting gene expression: for example, general and gene-specific transcriptional machinery, ribosome number and tRNA availability, or even cell morphology. The two sources can be experimentally distinguished using a dual-reporter assay, where two copies of the same promoter are used to drive distinguishable fluorescent protein variants (Elowitz, Levine, et al, 2002). Extrinsic noise is variation in protein levels between different cells; intrinsic noise is variation in protein levels within the same cell.

How regulators control the kinetics of intrinsic promoter fluctuations dictates the resulting expression heterogeneity. Stochastic models can be used to directly quantify this relationship (Maheshri, and O'Shea, 2007). To date, most transcriptional regulators function by modulating the frequency of these bursts (Bar-Even, Paulsson, et al, 2006; Newman, Ghaemmaghami, et al, 2006), probably in large part by increasing the rate of transcriptional initiation. Therefore, regulators do not control expression heterogeneity independently of expression level. In fact, heterogeneity is under genetic control as noisy promoters tend to have particular characteristics: strong TATA boxes, highly regulable, and dependent on chromatin remodeling activities (Bar-Even, Paulsson, et al, 2006; Newman, Ghaemmaghami, et al, 2006; Raser, and O'Shea, 2004).

While conventional gene regulation involves fast fluctuations between inactive (competent) and active promoter states, epigenetic silencing of gene expression involves slow fluctuations (<

once per cell cycle) between a silenced and competent state. The kinetics of these fluctuations in *trans*-encoded switches involving feedback loops and associated with bistable gene expression have been studied in detail (Ingolia, and Murray, 2007; Kaufmann, Yang, et al, 2007; Mettetal, Muzzey, et al, 2006). Both theory and experiment suggest that extrinsic fluctuations in the *trans* factor that overcome the stability of the two epigenetic states lead to switching (Kaufmann, Yang, et al, 2007). However, much less is known of the precise role of regulators in modulating fluctuations of *cis*-encoded switches which must involve changes in the local promoter state. For example, activators could increase population-averaged expression by either stabilizing the competent state or destabilizing the silenced state. Again, the heterogeneity in expression is dictated by the specific kinetic role of the activator.

In a diploid organism, an epigenetically regulated gene might exhibit four different expression states if each copy switches independently. With global encoding, both copies respond to the same global factor and must switch in a correlated manner. However, with local encoding, each copy may respond independently if the fluctuation that trips the switch is a molecular event that occurs locally at one copy. In fact, a recent study demonstrated the random and independent switching of two copies of a reporter gene inserted within the canonically silenced mating type loci, *HMR* and *HML*, in *S. cerevisiae*. Four distinct expression states were observed in a *sir1* background, where SIR-protein dependent silencing of these loci is partially impaired (Xu, Zawadzki, and Broach, 2006).

Multiple *cis*-encoded epigenetic switches that toggle slowly and randomly could lead a combinatorial explosion of expression states and represent a powerful strategy to generate phenotypic diversity. Is independent switching employed in nature and how are slow fluctuations regulated? The *S. cerevisiae* Flo11p is a cell-wall adhesin protein and member of the *FLO* gene

family important in mediating cell-to-cell and hydrophobic cell-surface interactions (Verstrepen, and Klis, 2006). In addition to traditional regulation via the MAPK and PKA pathways (Pan, and Heitman, 2002; Rupp, Summers, et al, 1999), at least three mechanisms are known to generate variation in cell-surface adhesins: ploidy regulation (Galitski, Saldanha, et al, 1999), frequent recombination of tandem repeats within adhesin genes (Verstrepen, Jansen, et al, 2005) and epigenetic silencing (Halme, Bumgarner, et al, 2004). Silencing at *FLO11* occurs in a SIR-protein independent manner and is both promoter and position-specific (Halme, Bumgarner, et al, 2004). Given the importance of phenotypic diversity in the adhesive phenotype and the epigenetic silencing at *FLO11*, independent switching could represent a fourth mechanism for generating variation.

At 3.5 kb, the *FLO11* promoter is one of the largest in *S. cerevisiae* and regulated by many factors (Figure 1) whose kinetic roles are unknown. Silencing of *FLO11* is thought to occur through the recruitment of the histone deacetylase Hda1p via the repressor Sfl1p through a yet to be defined mechanism (Halme, Bumgarner, et al, 2004). The Sfl1p repressor binding site overlaps the Flo8p activator binding site (Pan, and Heitman, 2002). Activation of *FLO11* through the protein kinase A (PKA) pathway results in phosphorylation of both Sfl1p and Flo8p. While phosphorylation disables Sfl1p binding, it enables Flo8p binding (Pan, and Heitman, 2002; Rupp, Summers, et al, 1999). Additional transcription factors bind directly to this promoter (Borneman, Leigh-Bell, et al, 2006; Pan, and Heitman, 2002; Rupp, Summers, et al, 1999) including the MAPK regulated Ste12p/Tec1p and Phd1p. These three activators require Flo8p for activation and play a significant role in determining the overall level of expression (van Dyk, Pretorius, and Bauer, 2005). Two activators, Msn1p and Mss11p, do not require Flo8p for activation and operate through poorly understood mechanisms that do not seem to require DNA

binding (Gagiano, van Dyk, et al, 1999). Msn1p acts at longer distances to destabilize chromatin (Lorenz, and Heitman, 1998); Mss11p has glutamine rich activation domains and may associate weakly with Flo8p (van Dyk, Pretorius, and Bauer, 2005). All these activators modulate plasmid-borne *FLO11* expression, a context where silencing does not occur (Pan, and Heitman, 2002; Rupp, Summers, et al, 1999). However, their varied biochemical roles might imply distinct kinetic and functional roles in epigenetic regulation of *FLO11*.

Here, we provide evidence that *FLO11* is indeed a *cis*-encoded epigenetic switch and identify the kinetic roles of *trans* factors in the epigenetic and conventional regulation of *FLO11*. Within a diploid yeast, each locus switches in a slow, random, and independent manner, with switching rates dependent on environmental conditions. Using a stochastic kinetic model, we infer the kinetic role that different regulators have on the slow promoter fluctuations associated with epigenetic transitions between a silenced and competent promoter state and the fast promoter fluctuations associated with conventional gene activation. We find three classes of *FLO11* regulators: those that affect the stability of the competent state, affecting slow promoter fluctuations; those that regulate the burst frequency of transcription due to fast promoter fluctuations; and those that have both functions. Moreover, a single synthetic activator can mimic each of these three classes based on the location of its DNA binding site. Because the kinetic role of each regulator defines its impact on expression heterogeneity, this can be controlled by the choice of regulator class. Finally, ethanol controls the extent of gene silencing nearly independently of transcriptional activation through Flo8p, thereby dictating whether *FLO11* expression responds in a graded or heterogeneous manner to other signals.

MATERIALS AND METHODS

Yeast strains and media

To use the dual-reporter assay to study switching of the *FLO11* promoter, we replaced the *FLO11* ORF in a haploid Σ1278b from the Heitman laboratory (Lorenz, and Heitman, 1997) with *YFP-KanMX6* or *CFP-KanMX6* cassettes by PCR integration, and then mated to create diploids. All strains and plasmids used are provided in Supplemental Tables 1 and 2.

SD is synthetic defined media with 2% glucose. SD ura- media lacks uracil. SD ura- media with ethanol contains both 2% glucose and a specified amount of ethanol.

Single cell measurements and analysis

Cells from overnight cultures grown in SD ura- or SD ura- + ethanol were inoculated at an initial OD_{600} between 0.005 and 0.01, and grown for 15-20 hours in the same media. For the titration experiments, these cultures were treated with serial dilutions of doxycycline (0 to 10000 ng/ml) at 30°C in well-agitated deep well 96-well plates. Cells were harvested in mid-late log phase (OD_{600} between 0.5 and 1.5), and placed on ice while other samples were being processed. Expression was measured using a Zeiss AxioObserver microscope with filters optimized for yECitrine, mCherry, and Cerulean (Chroma). Metamorph software (Molecular Devices) was used to analyze images and quantify single cell YFP and CFP fluorescence from. Between 500 and 1500 cells were imaged for each sample. Fluorescence levels in the RFP channel was used to discard dead cells (usually <5% of population). Details of data preprocessing and estimation of λ and γ are given in the Supplemental Discussion.

In timelapse microscopy experiments, cells were loaded onto the ONIX Microfluidic Platform (CellASIC) with initially ~10 cells trapped in individual chambers. Media in the $\sim 10^{-5}$ μl chambers was

constantly replenished at a rate of 10 μl/hr. YFP, CFP, RFP fluorescence and bright field images were obtained every 15 minutes for 20 hours. Image stacks were segmented using custom Metamorph journals. Single cell tracking and fluorescence was determined using custom MATLAB routines.

Chromatin Immunoprecipitation assay

Chromatin IP's were done based on the method of (Aparicio, Geisberg, and Struhl, 2004). Briefly, ~40 ml of cells were grown at 30°C in either SD complete or SD leu- to an OD_{600} = 0.8. Lysates from the fixed cells were sonicated to shear the chromatin to an average length of 500bp, and isolated chromatin was incubated with 2μl antibodies (Upstate/Millipore) against either histone H3 (Cat. No. 05-928), histone H4 (Cat. No. 05-858), acetylated histone H3 (Cat. No. 07-593) and acetylated histone H4 (Cat. No. 06-866). A sample with no antibodies was also prepared as a control. After reversal of cross-links, DNA from immunoprecipated chromatin was purified and analyzed using quantitative PCR (Applied Biosystems). Primers amplifying the -1.7 to -1.5 kb region of the *FLO11* promoter and primers amplifying a telomeric region in the right arm of chromosome VI (Yang, and Kirchmaier, 2006) as a control for hypoacetylated histone signals were used. Applied Biosystems 7300 software was used to obtain cycle threshold values. All signals from experimental samples were quantified relative to signal from a known amount of genomic DNA from an unmodified, cogenic Σ1287b strain (MLY43, Supplemental Table 1) that served as a positive control. The ratio of anti-histone / anti-acetylated histone signals was used as a measure of average H3 and H4 acetylation in the region.

Micrococcal nuclease assay

Micrococcal nuclease assays were performed as in (Lam, Steger, and O'Shea, 2008). All cells were grown at 30°C in either SD complete or SD leu- to OD_{600} = 0.5.

RESULTS

FLO11 *switches between silenced and competent states independently at each locus*

Under poor nutritional conditions, *FLO11* is partially silenced in haploid cells and heterogeneous in expression. Members of this population are capable of reversibly transitioning between the OFF (silenced) and ON (competent) state (Halme, Bumgarner, et al, 2004). To determine whether the transitions were due to a *cis* or *trans* fluctuation, we employed the dual-reporter assay, replacing the two copies of a *FLO11* ORF in diploid yeast with a distinct fluorescent protein variant (Venus YFP and Cerulean CFP) (Figure 2A). Importantly, we verified the independence and equivalence of the two reporters with respect to the presence of the other reporter (Figures S1 and S2). When grown in media with poor carbon sources, including ethanol, glycerol, galactose, and raffinose, we observed all four possible expression states (Figure 2B, data not shown). Because endogenous Flo11p is not present in the dual-reporter strain, we verified that Flo11p did not affect expression at the *FLO11* promoter in two ways. First, we added a plasmid constitutively expressing *FLO11* and found no significant effect on fluorescent protein expression (data not shown). Second, we compared fluorescent protein expression in the dual-reporter strain to strains where only one *FLO11* allele had been replaced with a fluorescent protein. There was no difference in expression levels (Figures S1 and S2).

If each allele switches independently, then at steady-state the proportion of cells in each expression state is given by p^2 (both ON), $(1-p)^2$ (both OFF), or $2p(1-p)$ (mixed ON/OFF and OFF/ON), where p is the proportion of cells with a particular allele ON. Note that p is identical for both YFP and CFP expression because the alleles are equivalent. We were able to verify the population's expression profile had reached steady-state (Figure 3A and S4). However, a naïve classification of expression state based on comparing a

cell's fluorescence level to background is incorrect because it does not consider the long lifetime of the fluorescent proteins which obscures the true expression state of the promoter. Therefore, we directly measured the eight transition rates by real-time monitoring of *FLO11* expression in single cells grown in a microfluidic chamber at constant conditions (Supplemental Discussion for details). All four ON to OFF and OFF to ON transition rates (Figures 2C, 2D and 2E) were found to be indistinguishable, demonstrating that each allele was switching independently. Furthermore, the fraction of cells turning ON or OFF are well-fit by a single exponential, confirming each transition is appropriately lumped as a pseudo-first order reaction.

Using a stochastic kinetic model, static distributions can reveal kinetic information

Time lapse microscopy provides an accurate determination of the slow epigenetic transition rates and proportion of each expression state, but it is experimentally challenging and low throughput. Therefore, after determining that transition rates were accurately described as first order, we devised a way to infer these rates directly from static snapshots, accounting for the long lifetimes of the fluorescent reporters. Two-state models have been widely employed to model faster promoter fluctuations associated with conventional gene regulation (Kaern, Elston, et al, 2005; Raj, Peskin, et al, 2006; Raser, and O'Shea, 2004). In such models, the promoter can transition between an inactive but competent state and an active state that leads to transcription. Many eukaryotic genes appear to reside in the competent state, with rare transitions to the short-lived active state that result in a "burst" of transcription.

The observed variation in the *FLO11* promoter can be divided into an intrinsic and extrinsic component. The *FLO11* promoter is subject to both fast intrinsic fluctuations and a slow epigenetic transition, as depicted by the augmented three-state model in Figure 3B. Extrinsic

noise also contributes to cell-to-cell variation in *FLO11* expression levels when the promoter is not silenced. However, when the promoter is (partially) silenced, the predominant source of variation in *FLO11* or reporter expression arises from the slow epigenetic transition between silenced to competent states because of (1) the smaller magnitude of the fast intrinsic and extrinsic fluctuations and the fact that (2) the faster fluctuations (< 1 cell generation) are more completely time-averaged by the long-lived reporter compared to the slow transition (> 1 cell generation). Therefore, we can lump the fast transition rates (λ', γ', μ') into an overall transcription rate μ and ignore extrinsic fluctuations. Gene expression can now be described using the commonly employed two-state model: $\frac{dx}{dt} = \mu f(t) - \delta x$, where x is the amount of reporter protein, μ is the (lumped) protein production rate, δ is a protein degradation rate (here the cell growth rate), and $f(t)$ is a "random telegraph process" that takes values of 0 or 1 corresponding to a silenced or active promoter state, with exponentially distributed times between switching events (Figure 3C). This stochastic equation has been solved analytically to yield a Beta distribution for protein number x at steady-state (Raj, Peskin, et al, 2006). The slow epigenetic transition rates, λ and γ, correspond to those measured in the time lapse experiment. To infer these rates we assume our measured distribution of protein x is steady (Figure S4) and fit it to the Beta distribution using a value of μ based on the expression level of the ON population in a bimodal condition (the parameter δ, the cell growth rate, is measured directly -- see Supplemental Discussion).

We tested this method in two different ways. First, we used the steady-state protein distribution of the time lapse experiments to estimate transition rates and found tight agreement between the inferred rates and those directly measured in time lapse (Figures 3A and 3C). Second, this model

allows proper estimation of the fraction of cells that appear ON in static distributions that are actually OFF because of the long lifetime of the fluorescent reporter (details in Supplemental Discussion). We applied this correction to static snapshots of cells grown in different conditions. Although the fraction of cells in each expression state varied, the overall statistics were always consistent with independent switching at each promoter (Figure 3D). Therefore, the upstream signaling network can map environmental inputs to a particular mixture of expression states through the modulation of transition rates.

A strategy for determining how regulators affect transition rates

Ultimately, environmental signals modulate epigenetic regulation of *FLO11* through downstream regulators. The effect of these regulators on both the mean level of expression and expression heterogeneity is succinctly and quantitatively described by their effect on the transition rates in the three-state model (Figure 3B). Therefore, we decided to titrate *trans* factors and measure the quantitative response of the *FLO11* promoter at the single cell level in hundreds of cells by fluorescence microscopy using the dual-reporter assay. For each condition and strain, we always grew cells for > 10 doublings, serially diluting them as needed to maintain low density and ensure a steady-state had been reached (further details in Supplemental Discussion).

To obtain transition rates, we fit the measured fluorescence distributions arising from each titration to the Beta distribution, the solution to the simplified two-state model. As described previously, a two-state model which lumps the fast transitions is only strictly applicable when slow epigenetic transitions related to silencing dominate. The four quadrant plot in Figure 4A summarizes the qualitative population-level response as given by the Beta distribution for various combinations of λ and γ. Each quadrant corresponds to regimes where λ and γ are slower

or faster than the division rate (δ). Epigenetic regulation occurs by definition in the lower left quadrant, when both λ and γ are slower than the division rate (λ/δ and $\gamma/\delta < 1$). Expression can turn completely OFF if the active state is destabilized (γ increases, shift to lower right quadrant), or the silenced state is stabilized (λ decreases, bimodal expression with vanishingly smaller percentage of cells ON). Opposite changes in λ and γ turn expression completely ON (and can lead to a shift to the upper right quadrant).

If epigenetic regulation is lost, expression levels can still change due to faster promoter fluctuations. A two-state model accurately describes intrinsic (but not extrinsic) fluctuations caused by transitions between the competent and active promoter states. For example, a conventionally regulated (but repressed) gene can be OFF and lie in the lower right quadrant. Activation leads to an increased burst frequency (λ' increases) and the graded, unimodal distribution of the upper right quadrant. Importantly, it is impossible to distinguish between conventional repression and epigenetic silencing in any population in the lower right quadrant where expression is completely OFF. Furthermore, fitting fluorescence expression distributions generated from a single *FLO11* promoter driven reporter that is conventionally regulated does not yield the fast promoter transition rates λ' and γ' because here extrinsic fluctuations are significant. The extrinsic noise is due to cell-to-cell variation in factors like morphology, ribosome number, and/or upstream components in the *FLO11* regulatory pathway and affects both promoters within the same cell in a correlated fashion. To properly measure the fast transition rates associated with conventional regulation, the intrinsic noise should be analyzed to determine the burst frequency (λ') and burst size (μ'/γ') (see Supplemental Discussion).

Three classes of regulators of FLO11 expression

To decouple complex upstream signaling events occurring at the promoter (Figure 1), individual *trans* factors were expressed heterologously under the control of a doxycycline-inducible promoter (Belli, Gari, et al, 1998). Because Sfl1p and Flo8p are post-translationally regulated, we needed a way to tune their relative strength. We chose ethanol, since the addition of ethanol activates *FLO11* expression in a Flo8p-dependent manner (see below). All the *trans* factor titrations were performed in either SD ura- or SD ura- with ethanol. Titrations of Sfl1p, Flo8p, and Mss11p were done in a *sfl1Δ*, *flo8Δ* or *mss11Δ* background, respectively.

For each titration point, we inferred the transition rates λ and γ using the two-state model. To summarize the effect of various activators and Sfl1p on the stability of the silenced and competent states, we plot the series of (γ, λ) values determined on the four quadrant plot of Figure 4A . In SD ura-, *FLO11* is OFF, corresponding to the lower right quadrant. Based on the response of the *FLO11* promoter (Figure 4B and 4C), we grouped the activators into 3 classes. Addition of three Class I activators, Tec1p, Ste12p, and Phd1p, appears to weakly stabilize the competent state, but expression remains extremely low. The Class II activators, Msn1p and Mss11p, stabilized the competent state by decreasing γ and entering the heterogeneous region where slow promoter fluctuations dominate. At some critical γ value, the silenced state is rapidly destabilized and the entire population turns ON. Flo8p constitutes a special class and was titrated in ethanol conditions where presumably some fraction of Flo8p is now phosphorylated and active. The population response is intermediate between Class I and Class II activators, but closer to Class II. Sfl1p has the exact opposite effect under the same ethanol conditions, consistent with the antagonistic role Sfl1p and Flo8p have through their overlapping binding sites and as being negatively and positively regulated by PKA, respectively.

To determine if activator class was correlated to binding site position or accessibility, we applied both an *in silico* nucleosomal occupancy model (Kaplan, Moore, et al, 2009) and performed micrococcal nuclease mapping (Figure 1 and S8) of the *FLO11* promoter. Both techniques suggested the -1200 region containing overlapping binding sites for Sfl1p and Flo8p is nucleosome free. In contrast, binding sites of Class I activators occur in nucleosomally occluded regions. Class II activators are not known to bind DNA but are potent activators, with Msn1p having a known ability to recruit chromatin remodeling machinery (Lorenz, and Heitman, 1998).

A synthetic activator mimics each activator class depending on its binding site position

Are binding site position, accessibility and/or competition with Sfl1p sufficient to determine activator class? If so, a synthetic activator could have qualitatively different regulatory profiles depending on binding site position. We engineered dual-reporter yeast strains with the 19 bp tetO sequence inserted at 3 different locations within the *FLO11* promoter. The first location was at -350 in a nucleosomally occluded region close to the TATA box and transcriptional start site. The second location was at -1470, on the outer-edge of the nucleosome upstream of the Sfl1p binding site and far from the core promoter. The third location was at -1160, within a nucleosome-free region directly overlapping the Sfl1p binding site. Sfl1p binds as a dimer at two sites (Conlan, and Tzamarias, 2001; Pan, and Heitman, 2002), so we replaced 19 bp of promoter sequence between the two sites with the tetO sequence to preserve the distance spanned by the two sites. The tetO is bound by rtTA, a synthetic activator that contains a strong acidic activation domain, VP16, (Belli, Gari, et al, 1998) known to recruit the SAGA complex in yeast (Berger, Pina, et al, 1992). We titrated rtTA in these three strains grown in SD ura- (all initially OFF). Each location functionally mimics the response of the respective class of activators (Figures 4B and 4D).

Importantly, the silenced state stability is reduced for the third tetO location compared to the second tetO location. This occurred even in the absence of ethanol, suggesting that the difference in silenced state stability between Class II activators and Flo8p is not due to an alternative ethanol-specific effect.

Two modes of Sfl1p repression correspond to a graded or heterogeneous response

When performing the Sfl1p titration above (Figure 4C and 5A), we did so in a *sfl1Δ* background in ethanol. *FLO11* was highly expressed (upper left quadrant) in the *sfl1Δ* background, as has been shown previously (Rupp, Summers, et al, 1999). Surprisingly, in the Sfl1p titration in SD ura- media without ethanol, both promoters turned off in a *graded* fashion (Figure 5B). This was in contrast to the heterogeneous population response observed for titrations of Class II activators (Figure 4) including Sfl1p titrations performed in ethanol (Figure 5A), a condition where Flo8p is presumably more active.

To explain this result, we hypothesized that a critical Sfl1p level is required to silence the promoter, and below this level Sfl1p still repressed transcription but in a conventional manner associated with faster promoter fluctuations. The model requires that Sfl1p is able to repress the *FLO11* promoter as a silencer or as a conventional repressor; evidence exists for both modes (Conlan, and Tzamarias, 2001; Song, and Carlson, 1998).

To test this model, we generated a dual-reporter strain in an *hda1Δ* background and added back Sfl1p heterologously. When we titrated Sfl1p in this background in SD ura- media, we observed a graded response (Figure 5C). This establishes that Sfl1p is capable of repressing expression in an Hda1p-independent manner. The graded response suggests Sfl1p is working as a conventional

repressor rather than affecting the slower fluctuations between the epigenetically silenced and competent promoter states. To further demonstrate that Hda1p is necessary to silence *FLO11*, we measured the average H3 and H4 histone acetylation state at the *FLO11* promoter by Chromatin-IP in the dual-reporter strain grown in SD ura- (Figure 5D). Only the wildtype *FLO11* promoter exhibited a hypoacetylated state, indicative of silenced chromatin and similar to a silenced telomeric region (Suka, Suka, et al, 2001). In both the *hda1Δ* and *sfl1Δ* backgrounds, the promoter was hyperacetylated, a chromatin state associated with lack of silencing. Together, this demonstrates that Sfl1p silences the promoter in an Hda1p-dependent manner and the silenced state at *FLO11* is correlated with hypoacetylation in at least one region (~ -1600 bp) of the promoter.

If silencing is eliminated in an *hda1Δ* background, titration of activators in the presence of high levels of Sfl1p will result in a graded response, since Sfl1p is now functioning as a conventional repressor. In addition, the threshold level of activator required to turn on *FLO11* will be lower. We performed these titrations (shown for Tec1p in Figure 5E and 5F and other activators in Figure S7) and confirmed this prediction. In addition, the intrinsic noise of these strains was proportional to the inverse square root of the mean expression level (Figure 5F inset and Figure S7), indicating that without slow promoter fluctuations, the activators regulate the burst frequency, λ' (Bar-Even, Paulsson, et al, 2006).

Ethanol controls the importance of silencing in a Flo8p-dependent manner

To further understand the role of ethanol in *FLO11* signaling, we grew the dual-reporter strain in SD ura- media in a range of ethanol concentrations. *FLO11* expression exhibited a graded response to increasing ethanol levels, but the average expression level remained low even at the

highest (3%) ethanol concentrations (Figure 6A). The graded response suggested a lack of silencing possibly due to increased Flo8p activity. This led to the hypothesis that at low levels of ethanol (<1%), Flo8p activity eliminates Sfl1p-mediated silencing, but has little effect on expression. Therefore, although *FLO11* expression is OFF (lower right quadrant), the promoter is actually in the competent state. To test this idea, we titrated Class I and II activators in 1% ethanol. Both classes were capable of increasing *FLO11* expression to high levels in a *graded* manner (Figure 6B), implying that silencing no longer occurred in these conditions. The corresponding synthetic activators had a similar effect. In contrast, titration of the synthetic activator mimicking Flo8p resulted in a response similar to ethanol (Figure 6C). Finally, Class I and II activator titrations in a *flo8Δ* in SD ura- with 1% ethanol (Figure 6D) reverted to the behavior seen in SD ura- conditions. Therefore, ethanol controls the extent of silencing at *FLO11* in a Flo8p-dependent manner. Both Flo8p and its synthetic analog affect slow promoter transitions, but neither is a strong conventional activator.

DISCUSSION

The main results of our work are perhaps best understood with reference to the 3-state model in Figure 7, a simple, but useful paradigm for describing the kinetics of epigenetic gene regulation. Regulators can affect either the slow transition rates associated with epigenetic silencing or the fast transition rates associated with conventional gene activation. While a good deal is known about how regulators affect the fast transition rates, less is known about how regulators affect the slow transition rates, and whether effects on slow and fast transition rates are coupled. Our work demonstrates that slow transitions at the two copies of the *FLO11* promoter in diploid yeast occur randomly and independently. Furthermore, we identify the role of various *trans* regulators of *FLO11* in controlling both slow and fast transition rates; it appears that this control is

distributed among various "classes" of regulators. Importantly, distributed control enables the cell to shape the diversity of *FLO11* expression within an isogenic population by utilizing different combinations of regulators.

Our demonstration that two copies of the *FLO11* promoter switch slowly and independently builds on previous work demonstrating (1) the *FLO11* promoter is epigenetically regulated (Halme, Bumgarner, et al, 2004), and (2) two copies of a partially silenced *URA3* promoter inserted at the mating type loci in *S. cerevisiae* switch independently in a *sir1* background (Xu, Zawadzki, and Broach, 2006). We now provide evidence that independent switching occurs in a natural gene whose epigenetic regulation is not SIR protein dependent. Given the rich diversity of the *FLO11* gene pool (Verstrepen, Jansen, et al, 2005) independent switching may be an additional mechanism for generating variation in adhesive phenotype.

To understand how expression heterogeneity at *FLO11* is controlled, we took a functional approach and determined the kinetic role of different regulators on both slow and fast promoter fluctuations. Class I regulates fast promoter fluctuations exclusively. Intrinsic noise measurements confirmed these regulators destabilize the competent state and increase the burst frequency, a common theme for regulators in yeast (Bar-Even, Paulsson, et al, 2006; Newman, Ghaemmaghami, et al, 2006). Interestingly, we find that the Class II and Flo8p regulate slow promoter fluctuations primarily by stabilizing the competent state. The previous study at the mating type loci found an activator, Ppr1p, could challenge the silenced state but affected both slow transition rates (Xu, Zawadzki, and Broach, 2006), but this was only measured at one level of Ppr1p. It is only by titrating regulators that we were able to clearly discern their functional roles. Whether activators generally affect one or both of the transition rates of silenced genes remains an open question, but likely depends on the mechanism of silencing.

While the regulator titrations indicate that ethanol desilences the promoter via Flo8p, the pathway(s) by which ethanol activates Flo8p is unknown. The simplest mechanism is that long term growth in ethanol activates PKA (specifically Tpk2p), which then activates Flo8p and deactivates Sfl1p. In fact, glucose is not required as the promoter response when Flo8p was titrated in synthetic media with ethanol or glycerol as the sole carbon source was similar to that in SD 1% ethanol (data not shown). However, activation of the PKA pathway via ethanol has not been reported. It is also possible that Flo8p is activated by ethanol via another pathway independent of PKA, although to our knowledge, it is not known to be the target of any other kinase. A third possibility is that ethanol may enable synergistic interactions between Flo8p and other regulators that leads to desilencing, although there is no evidence of Flo8p interacting with any Class I activators. These possibilities could be distinguished by monitoring cyclic AMP levels and the phosphorylation status of Flo8p and Sfl1p in ethanol.

Our results lend support to the idea that binding site position within the *FLO11* promoter can largely determine the kinetic role of the transcriptional regulator. However, the mechanistic description of how binding to particular sites affects slow epigenetic regulation and fast conventional regulation and the molecular nature of the silenced and competent promoter states is still unclear. A mechanistic explanation for the dual roles of Sfl1p is likely the clearest. Binding in the -1200 nucleosome-free region governs epigenetic silencing, possibly by recruiting Hda1p via Tup1p/Ssn6p corepressor (Conlan, and Tzamarias, 2001). Conventional regulation most likely occurs via Sfl1p binding to the -200 region which contains a putative Sfl1p binding site and has been shown to bind Sfl1p *in vitro*. Indeed, preliminary ChIP experiments suggest Sfl1p is bound to this region *in vivo* (Octavio and Maheshri, unpublished results).

Among the activators, the role of Flo8p and its synthetic analog is perhaps the clearest. Flo8p binding and Sfl1p binding at the -1200 region are likely mutually exclusive because of overlapping binding sites, and so Flo8p can prevent the Sfl1p-mediated establishment of silencing but probably not directly affect conventional Sfl1p repression. This would explain the ability of Flo8p and its corresponding synthetic analog to affect the slow epigenetic transition independently of the conventional activation. In fact, binding of tetR, which lacks the VP16 activation domain of rtTA, to the -1160 tetO site is sufficient to abrograte silencing (data not shown), implying steric hindrance is the major mode of action. Furthermore, any weak activation via Flo8p might be through its known role in binding to the promoters of other Class I activators (including the ones tested here) and presumably upregulating their expression. In this manner, Flo8p activation can put the *FLO11* promoter in a competent, "poised" state whose expression can be controlled by Class I –like regulators. There is evidence, though, that Flo8p can bind to other regions of the promoter (Borneman, Leigh-Bell, et al, 2006).

Several possibilities exist for the inability of the Class I activators to challenge the silenced state, yet still regulate the burst frequency of the competent *FLO11* promoter. For example, binding site proximity to the transcriptional start site could play a dominant role. Canonical yeast promoters are typically 150-400 bp with transcription factor binding sites are clustered 100-200 bp from the transcriptional start site (Pelechano, Garcia-Martinez, and Perez-Ortin, 2006); this proximity allows direct interaction with general transcriptional machinery. Therefore, it may be that Class I activators bind in the core region of the *FLO11* promoter, a region that may be inaccessible to transcription factors and/or transcriptional machinery in the silenced state. However, while the Class I synthetic analog binding site is at -350 in a nucleosome occluded

region, not all Class I activators have binding sites in this region (Borneman, Leigh-Bell, et al, 2006).

A second possibility not mutually exclusive with the first is that Class I activators need not bind in the downstream region but can influence transcription rates via long range (but fast) mechanisms including DNA looping, cryptic transcription, or long range chromatin remodeling. This would provide an explanation for the presence of Class I activator binding sites in these regions that are known to be bound *in vivo* in activating conditions (Borneman, Leigh-Bell, et al, 2006). In addition, it might explain why even at high levels of expression in the absence of silencing the intrinsic noise at the *FLO11* promoter is 10 times higher than that of a similarly highly expressed *PHO84* promoter (data not shown). However, with either explanation, the inability of Class I activators to challenge the silenced state is not clear. Altered chromatin structure or reduced binding site accessibility could be invoked as Class I activator sites tend to be under nucleosomes. However, other than some nucleosome depletion in the core promoter and the -1300 region, no gross nucleosomal rearrangements seem to occur upon silencing (Figure S8) although higher resolution mapping may reveal finer differences.

While we do not know the biochemical intermediates during the slow promoter transitions, the pseudo-first order rates suggest a single slow step, rather than a distributed control mechanism. This is similar to both the partially silenced mating type loci (Xu, Zawadzki, and Broach, 2006) and the epigenetically regulated *agn43* gene in *E. coli* (Lim, and van Oudenaarden, 2007). Possibilities for the slow epigenetic step governing ON to OFF might include Sfl1p binding or Sfl1p-mediated recruitment of silencing factors, among others. Both Class II activators Msn1p and Mss11p are capable of stabilizing the active state, but their localization and activity with respect to the *FLO11* promoter remains unclear (Lorenz, and Heitman, 1998; van Dyk, Pretorius,

and Bauer, 2005). The ability of the Class II synthetic analog to have a similar destabilizing effect on the OFF state by binding to the -1470 region as well as the differential acetylation state of that region strongly suggests chromatin remodeling in the upstream region affects accessibility of Class I activators and the transition to the competent state. High resolution mapping of the chromatin state of the entire *FLO11* promoter under various conditions should point toward the biochemical mechanism of the slow promoter transition and will be the focus of future work.

Our findings have implications for the regulation of various subtelomerically encoded gene families known to be epigenetically regulated. This includes the *FLO* gene family (Halme, Bumgarner, et al, 2004) and other closely related yeast adhesins (Verstrepen, and Klis, 2006) such as the *EPA* gene family in the pathogenic yeast *C. glabrata* silencing (De Las Penas, Pan, et al, 2003). Phenotypic variability in *EPA* gene expression might allow *C. glabrata* to rapidly colonize new host tissues and evade immune surveillance. Do such genes turn ON and OFF independently, does it depend on the mechanisms of their silencing (SIR-dependent, etc.), their relative chromosomal locations, or the presence of boundary elements? What promoter transitions do *trans* factors regulate? This understanding will allow the engineering of strains with well-defined levels of phenotypic heterogeneity. Such strains are a prerequisite to quantify what role if any phenotypic heterogeneity has on organismal fitness in natural environments.

SUPPLEMENTAL DISCUSSION

FLO11 signaling on solid media

We originally encountered the mixed expression states when analyzing colonies of the dual-reporter strain (Y45) grown on rich solid media (YPD, yeast extract, peptone, dextrose) for several days. Since colony growth on a complex solid media may be a more physiologically

relevant condition for the adhesion and filamentation response, we monitored reporter gene expression from the *FLO11* locus over the course of 10 days. To do so, we took an exponentially growing culture of the dual-reporter strain and various other control strains and spotted 10 μL containing 2 x 10^3 cells in 3 separate locations on YPD plates. These plates were left at room temperature for 10 days. Every two days, a mixture of cells originating from each spot was analyzed. In an effort to sample the "average" response of the colony rather than that of individual regions in the colonies, cells were collected by lightly tracing an X pattern in the spot using a toothpick. This insured that cells from both the periphery and edges of the spot were analyzed. We did this experiments with biological duplicates (different patches of the same strain) to verify that this technique yielded reproducible results (Figures S1, S2).

***FLO11* expression is not dependent on Flo11p and alleles are independent and equivalent.**

We created several different control strains to insure that 1) each fluorescent allele was equivalent, 2) each fluorescent allele was equivalent regardless of the haploid mating type from which it originated 3) each fluorescent allele's expression did not depend on the presence or absence of the other allele, 4) each fluorescent allele's expression was not affected by the presence of *FLO11*. For 1), compare Figure S1, row 2,3 with 4,5 Figure S2, any row. At 10 days, CFP and YFP levels are not quite equivalent on solid media, which we attribute to difficulties in detection and/or stability. This is never the case in liquid culture (for example, see Figure 2). For 2) compare Figure S1 row 2 with 3 or row 4 with 5. The variability in response between each pair is consistent with variability we see in biological duplicates (see Figure S2). For 3) and 4) compare Figure S1 rows 2,3 or 4,5 versus the marginal CFP or YFP expression in the diploids (Figure S2). The proportion of cells that turn ON for either YFP of CFP is similar with or without *FLO11*.

FLO11 expression in various environmental conditions

For the analysis in the main text of the paper, we were careful to always grow cells in SD media and harvest them in mid-log phase. This was to facilitate comparisons between samples and ensure that we were in a steady-state condition. We also analyzed expression from the dual-reporter strain in the following additional media conditions: YP + 2% glucose, 3% glycerol, 3% ethanol, 6% ethanol, YP 2% galactose, 2% raffinose; SLAD (synthetic low ammonium dextrose (2%)); SLARG (synthetic low ammonium galactose (2%)/raffinose(2%)); SD (2% glucose); YPD (2% glucose), SD (2% glucose), SLAD, SLARG + 100 μM tryptophol + 100 μM phenylethanol; YPD (2% glucose), SD (2% glucose), SLAD, SLARG + 100 μM tryptophol, YPD + 100 μM phenylethanol. The last two sets were performed due to a recent report that certain aromatic alcohols could influence *FLO11* expression (Chen, and Fink, 2006). All rich media (YP) conditions yielded a heterogeneous response that persisted as the culture approached saturation (data available upon request). However, in YPD *FLO11* became active near diauxic shift, presumably as glucose levels decreased. For SLAD and SLARG conditions, we saw very little *FLO11* expression. Because the low nitrogen condition in these two media are known to lead to filamentation (Guo, Styles, et al, 2000) on solid media, we also grew the dual-reporter strain on SLAD and SLARG plates, and monitored expression in the colony every two days for two weeks. We never observed significant *FLO11* expression (data not shown) suggesting that *FLO11* remained silenced. We are not certain why this was the case. Finally, the aromatic alcohols had very little if any effect on expression.

Fluorescence time trace analysis

For the timelapse experiment, cells (Y45) were grown in YP 1% ethanol 2% glycerol. This condition was selected because previous static snapshots revealed that approximately half the population was ON, so observing sufficient numbers of OFF and ON cells in the experiment would be easier. Prior to transfer to the microfluidic chamber, cells had been growing exponentially for >12 hrs. Upon transfer, the time evolution of the YFP expression is given in Figure 3A. At early times, only small numbers of cells are monitored, but within 10 hrs enough cells are present such that the steady-state expression distribution is indistinguishable to the pre-transfer distribution (two-sample KS test, $p = 0.19$ for CFP, $p = 0.21$ for YFP). We focused on microcolonies at the inlet only to ensure that the microenvironment was constant and uniform. Data was collected from three experiments, each run on a separate day.

Custom scripts in Metamorph were used to segment images to identify cells and generate masks. Fluorescence images from the experiment and the cell masks were loaded into MATLAB (Mathworks). Custom MATLAB routines were used to further process the masks and track individual cells. Processed tracks were manually reviewed and corrected if necessary (typically in frames with high cell density).

The observation length of each cell within a growing population of cells is not identical, and this observational bias *must* be considered for an accurate estimate of the switching rates. To determine waiting times for cells to switch from OFF to ON, we used a method similar to Kaufmann et al (Kaufmann, Yang, et al, 2007). We took the time between the birth of a cell that was OFF and (Kaufmann, Yang, et al, 2007)the time fluorescence signal sharply increases, and accounted for a 20 minute protein maturation for both CFP and YFP, which mature at similar rates (Kaufmann, Yang, et al, 2007). Because the sudden increase in fluorescence signal from the production of our reporter proteins was easily distinguishable from background

autofluorescence, we could clearly determine the moment a cell switches ON, within the 15 minute time resolution of our experiment. We therefore fit the distribution of OFF-ON waiting times using a bin size of 15 minutes.

We did not have as good a resolution for determining the time that a cell switches OFF. Because of the long-lived nature of our reporters, even after a cell switches OFF we only observe a sharp decrease in fluorescence signal when the cell divides. Prior to the division event, it is difficult to determine whether small fluctuations in the fluorescence are due to real changes in the protein production or other sources of noise such as photobleaching and tiny drifts in the z-direction during image capture. These fluctuations are negligible compared to the fluorescence change upon division. Therefore, we estimated the time a cell switched OFF by fitting the exponential decay of the fluorescence signal due to dilution from growth and extrapolating to determine the initial point in time when the signal began to decay. Because of the higher uncertainty in the exact moment the cell switched off in the period between division events, we fit the distribution of ON-OFF waiting times using a bin size of 158 minutes (the mean doubling time of the cells in this growth condition).

Calculation of switching rates from time traces

If the switching of the *FLO11* promoter is a Poisson process, the distribution of waiting times t between ON (OFF) and OFF (ON) states will be exponential, where λ is the switching rate (or, $1/\lambda$ is the mean waiting time between switching):

$$f(t) = \lambda e^{-\lambda t} \text{ (exponential PDF)} \quad (1)$$

Let $F(t)$ = probability that switch occurs before some time t. Then:

$$F(t) = 1 - e^{-\lambda t} \text{ (exponential CDF)} \quad (2)$$

Fitting the waiting times between ON (OFF) and OFF (ON) states to the above distribution would be valid only if we could observe every cell for an infinitely long time. However, we only observe cells for a finite time; moreover, in the limit of infinitely many cells, the observation time intervals are exponentially distributed, as cells are continuously born during the experiment.

To account for the uneven observation time intervals, we assume that the time intervals are exponentially distributed as follows, an assumption that was reasonably accurate for the number of cells in our experiment:

$$F(T) = \delta e^{-\delta T} \quad (3)$$

Where T = observation time, δ = mean growth rate of cells

The probability that a cell switches in some time *t'* is then sum of equations (4) and (5):

1) the probability that the switch occurs within *t'* and that we observed the cell for *t'*, which is given by:

$$\int_0^{t'} (1 - e^{-\lambda t}) \delta e^{-\delta t} dt \quad (4)$$

2) the probability that the switch occurs within *t'* and we observed the cell longer than *t'*, which is:

$$(1 - e^{-\lambda t'}) e^{-\delta t'} \quad (5)$$

The total probability is then:

$$\text{P (cell switches by time } t) = \frac{\lambda}{\lambda + \delta} e^{-(\lambda+\delta)t} \quad (6)$$

This is not a true cumulative distribution function as it will never go to 1 at long times, since cells are always being born and the observation time intervals of the newly born cells are short.

We fit the above probability density to the waiting times obtained from timelapse data using nonlinear least square optimization (fminsearch) in custom MATLAB routines.

Fitting static fluorescence distribution to a stochastic kinetic model

The solution to the stochastic kinetic model (Raj, Peskin, et al, 2006) is a Beta distribution, which can assume one of the four distinct shapes shown in Figure 4A, depending on the values of the parameters λ and γ (OFF-ON and ON-OFF transition rates). The model distribution accounts for variation in expression that arises solely from random promoter fluctuations between the ON and OFF states. In our experiments, additional factors such as cell-to-cell variation in the production rate μ between individual cells, intrinsic noise in mRNA transcription and translation, autofluorescence, and experimental error from the imaging equipment also contribute to the distribution we obtain, creating "noisy" ON and OFF peaks.

Intrinsic noise in mRNA production could be incorporated by convolving a Poisson distribution for mRNA levels with the Beta distribution. However, because of the separation in timescales, variation in protein levels from fast mRNA fluctuations is small compared to variation from the silenced promoter slowly switching between OFF and ON states. Similarly, extrinsic variation in production rates between different cells is smaller than the difference between ON and OFF cells, as indicated by our ability to observe "side populations". Error from the camera and imaging equipment are negligible; autofluorescence can be accounted for by kernel density

estimation. Because they are small, these extra sources of variation present in our raw experimental distribution do not change its qualitative shape; in almost all cases, we can easily distinguish bimodal distributions from unimodal ON or unimodal OFF distributions.

The noisy ON and OFF peaks, however, can complicate fitting of the raw distribution data to the model. Prior to fitting, we first applied a kernel density estimate to the raw data, where the kernel is the autofluorescence distribution. We observed the YFP fluorescence in OFF cells that were once ON was slightly above the autofluorescence measured in cells that did not express YFP. Therefore we used the YFP distribution from OFF cells that were once ON as the kernel. For cells that were once ON for CFP and have switched OFF, the level of CFP fluorescence was comparable to auto fluorescence. Then, we took the mean of the autofluorescence control as the OFF (zero) expression value. A portion of the measured fluorescence distribution mass falls below this value; we added this back to the first bin. Similarly, we took the mean of the ON peak as the value of μ/δ (maximum expression) and added back the portion of the fluorescence distribution mass above this value to the final bin.

Because of this operation, the resulting shape of the distribution is discontinuous between the first and second and penultimate and last bins. However, in most cases, this distribution could be fit directly to the beta distribution. Alternatively, the distribution could be further smoothed using a simple moving average prior to fitting to the model distribution. Both methods yield very similar fits.

We find that the values of the fit parameters λ and γ generally change 15-20% when the maximum expression value (μ/δ) selected is changed within a standard deviation about the mean of the ON peak. If we smooth the distribution using a moving average prior to fitting, the values

of λ and γ could also change depending on the size of the spanning window used. For a bimodal distribution, the effect of increasing the span size is generally to increase both λ and γ, as the mass is distributed along a wider range, creating less distinct ON/ OFF peaks. We therefore selected a small span sizes large enough to smooth jagged regions and always kept the span size uniform when comparing parameters between different regulator titrations.

In general, we have the best confidence in fitting bimodal distributions because we have an estimate of the maximum expression value from the ON peak. Unimodal ON distributions can be fit to a relatively wide range of high λ values and low to high γ values. For unimodal OFF distributions, it is difficult to estimate the maximum expression value. In titrations where the distributions were unimodal OFF (at low activator or high repressor levels) and then became bimodal at particular regulator levels, we set the value of μ/δ using the mean of the ON peak when in the bimodal regime. However, for regulator titrations of Tec1p, Ste12p or Phd1p in SD ura-, where distributions remained unimodal OFF at the highest activator levels, we chose the maximum fluorescence level as μ/δ. This choice of μ/δ represents a lower bound; the true maximum expression is likely higher. Fitting the same unimodal OFF distribution to a higher μ/δ will either decrease λ, increase γ, or both. Regardless, the estimated values of λ and γ for unimodal distributions no longer represent either fast or slow transition rates (see main text and below).

All fluorescence distributions were processed and the parameters λ and γ fit by maximum likelihood estimation using custom MATLAB routines.

Transition rates in epigenetic versus conventional regulation

When the two-state model (described by the Beta distribution) is used to derive OFF-ON transition rates, it is important to keep in mind that the underlying promoter states that we refer to as "OFF" or "ON" in the model are *different* depending on whether silencing occurs (Figure S3). As mentioned in the main text, in the case of the epigenetically regulated (silenced) promoter, the "OFF" to "ON" transition refers to the slow silenced to competent state transition. Because of the separation of timescales, the faster competent OFF to competent ACTIVE transition rates can all be lumped into a single, overall rate μ. Since the main source of variation in expression in an epigenetically regulated (partially silenced) promoter is the slow silenced to competent state transition, we can fit the static distributions and estimate λ and γ according to the procedure described in the previous section.

When the promoter is no longer silenced, only the fast promoter fluctuations between the competent OFF and competent ACTIVE states remain and cannot be neglected. This is now conventional regulation, where the variation in expression arises from faster extrinsic and intrinsic fluctuations. To measure the fast promoter transition rates λ' and γ', we cannot use the fitting procedure described above, as the fit rates are no longer accurate since extrinsic fluctuations are significant. We must instead analyze the intrinsic noise using the method of (Elowitz, Levine, et al, 2002). In many eukaryotes, including yeast, these fast promoter fluctuations occur as bursts. Promoters primarily stay in the competent state, and make infrequent, random transitions to the active state, which correspond to a burst of transcription (Maheshri, and O'Shea, 2007). In other words, the active state is highly unstable. Global studies in yeast indicate that *trans* regulators control the level of gene expression by modulating the

burst frequency, λ', and not the burst size μ'/γ' (Bar-Even, Paulsson, et al, 2006; Newman, Ghaemmaghami, et al, 2006).

Figure S3 shows the regimes where true transition rates in epigenetically regulated and conventionally regulated genes lie. For epigenetic regulation, these are the rates of switching between the silent to competent OFF states, whereas for conventional regulation, these are the rates of switching between the competent OFF to competent ACTIVE states. Global studies of noise in yeast gene expression estimate a burst size of ~1200 proteins (Bar-Even, Paulsson, et al, 2006; Newman, Ghaemmaghami, et al, 2006).

In addition, we observed that the intrinsic noise in a *sfl1Δ* strain was unusually high, about ten times the intrinsic noise measured at *PHO84* (data not shown), whose high expression is comparable to *FLO11* in a *sfl1Δ* background. The high intrinsic noise suggests the occurrence of a slow event (within a timescale of a cell division) prior to activation. Such an event might be chromatin remodeling to expose activator binding sites. If so, this could explain why much of the promoter remains nucleosome occluded in a *sfl1Δ* background (Figure S8). It remains an open question whether unusually high intrinsic noise might be a general property of heterochromatically silenced promoters when in a competent state, perhaps as a consequence of a particular nucleosomal structure that enables silencing.

Finally, we found no evidence that the slow transitions from silenced to competent states were correlated with replication (Figure S9).

Determining fraction of "side populations" in static distributions using estimated false positive rate

Cells expressing long-lived reporters at high levels require 3-5 divisions before fluorescence returns to autofluorescence levels and they actually appear OFF. The OFF state lifetime was ~3 cell division times on average in the time lapse experiments. Therefore, in a population of switching cells a fraction of cells that appear ON (above autofluorescence values) are actually OFF, resulting in a significant false positive rate. The false negative rate is negligible in comparison, because timelapse microscopy indicates that when OFF cells switch ON, fluorescence levels exceed autofluorescence significantly in much less than 1 cell division time.

We found good agreement with the values of λ and γ from fitting the static distributions to the stochastic kinetic model and the switching rates measured in the timelapse experiments. Therefore, we can apply this model to relate the true value of p , the fraction of cells ON for a particular reporter, to the measured value p' , the *apparent* fraction of ON cells. From our two-state model :

$$p = \frac{\lambda}{\lambda + \gamma} \tag{7}$$

The value p' is based on an arbitrary fluorescence threshold slightly above autofluorescence, chosen to be 3-5 standard deviations from the mean of the autofluorescence distribution, where the separation between the four states (both ON, both OFF, single CFP ON and single YFP ON) is well defined.

One can now define an estimate of the false positive rate, x, which is the fraction of cells that appear ON for a particular reporter but are really OFF. Then x can be simply calculated by

$$x = 1 - \frac{p}{p'} \tag{8}$$

where p' is obtained by direct measurement and p is estimated by fitting the fluorescence distribution to the stochastic model.

In the case of two reporters, the fraction of cells *apparently* ON for one reporter and not the other ("single ON") is:

$$\textit{apparent}\text{ "single ON" fraction (one particular reporter)} = p'(1-p') \tag{9}$$

The actual "single ON" fraction consists of those cells that both appear to be and are "single ON" plus those cells that appear to be "both ON" are really "single ON" for a particular reporter:

$$\textit{actual}\text{ single ON fraction} = (1-x)p'(1-p') + xp'(1-x)p' \tag{10}$$

If both reporters are independent, then the total "side populations" is just twice equation (10). It is easy to verify that this is equivalent to $2p(1-p)$, where p is the true fraction of ON cells, as shown in Figure 3D. We found that x was consistent across many conditions, and ranged between 0.30 and 0.36.

In all the regulator titrations and media conditions we have looked at, we have never observed well-separated subpopulations (both ON/OFF/ side populations) where the true fraction of ON cells $p > 0.5$. It appears that around this point (perhaps at the critical γ?), the OFF-ON switching rate (λ) becomes much faster than 2 cell divisions, blurring the separation between states and making the fractions of observed side populations extremely sensitive to the value of the threshold fluorescence assigned. Therefore, we analyzed only the well-separated populations in all our experiments, where the true fraction of ON cells p ranged from 0.03 to 0.4.

FLO11 activation model predictions

Our model of the kinetic roles of regulators makes several predictions about the nature of *FLO11* expression. Whereas a *sfl1Δ* expressed *FLO11* at high levels, an *sfl1Δflo8Δ* strain is reported to be OFF (Pan, and Heitman, 2002). We predict the promoter is in a competent state, and titration of either Class I or II activators will turn expression ON in a graded manner. Furthermore, the *FLO11* promoter should be hyperacetylated.

A second prediction is that the synthetic activator mimicking the Flo8p titration in SD ura- is capable of eliminating silencing at the promoter, even at low levels when expression is still OFF and the expression state is in the lower right quadrant (Figure 4C and 4D). Therefore constitutively expressing low levels of the synthetic activator in SD ura- media in the *absence* of ethanol should result in a competent promoter, which could be confirmed by ChIP for acetylation state and the titration of Class I activators.

Binding site locations at *FLO11* promoter

The location of binding sites shown in Figure 1 represent approximate positions only, as most sites were identified through a lacZ reporter assay where a fragment of the *FLO11* promoter was tested for its ability to activate expression of a lacZ reporter in various strain backgrounds. Therefore, the binding site(s) for each regulator lie within the range of the promoter fragment length used in the particular assay.

Ste12p and Tec1p binding sites were from (Rupp, Summers, et al, 1999), where 400 bp FLO11 promoter fragments were used in a lacZ assay comparing activation in a wildtype and *trans* regulator delete strain. Binding sites for Tec1p and Ste12p were also found at -699 to -704 and -

719 to -725 (Lo, and Dranginis, 1998); the consensus sequences (CATTC for Tec1p and TGAAAC for Ste12p) were originally identified using lacZ assays (Madhani, and Fink, 1997).

Using 250 bp promoter fragments, lacZ reporter assays and ChIP assays identified that Flo8p and Sfl1p both act upon the -1150 to -1400 region (Pan, and Heitman, 2002); other Flo8p binding sites shown were determined from the lacZ assay of (Rupp, Summers, et al, 1999), which used 400 bp fragments.

The consensus sequence (CATGCA) for Phd1p has been identified (Hughes, Badis, et al, 2008) using a microarray-based technique; this sequence is found at -1763 bp and -1775 bp.

FIGURES AND TABLES

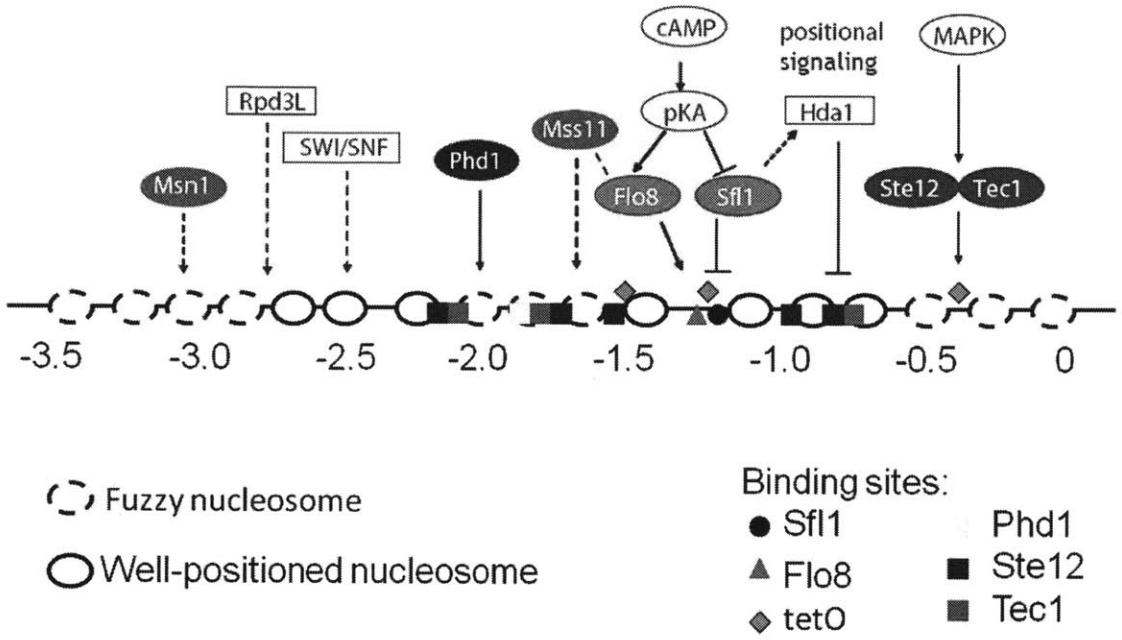

Figure 1: Signals from many *trans* factors converge at the complex *FLO11* promoter Regulators of *FLO11* transcription. Nucleosomal positions are based on a thermodynamic model for nucleosomal occupancy (Kaplan, Moore, et al, 2009). Binding sites are approximate and based on literature but most sites have not been confirmed directly. The three locations where a tetO sequence was inserted are also shown. See text and Supplemental Discussion for further details.

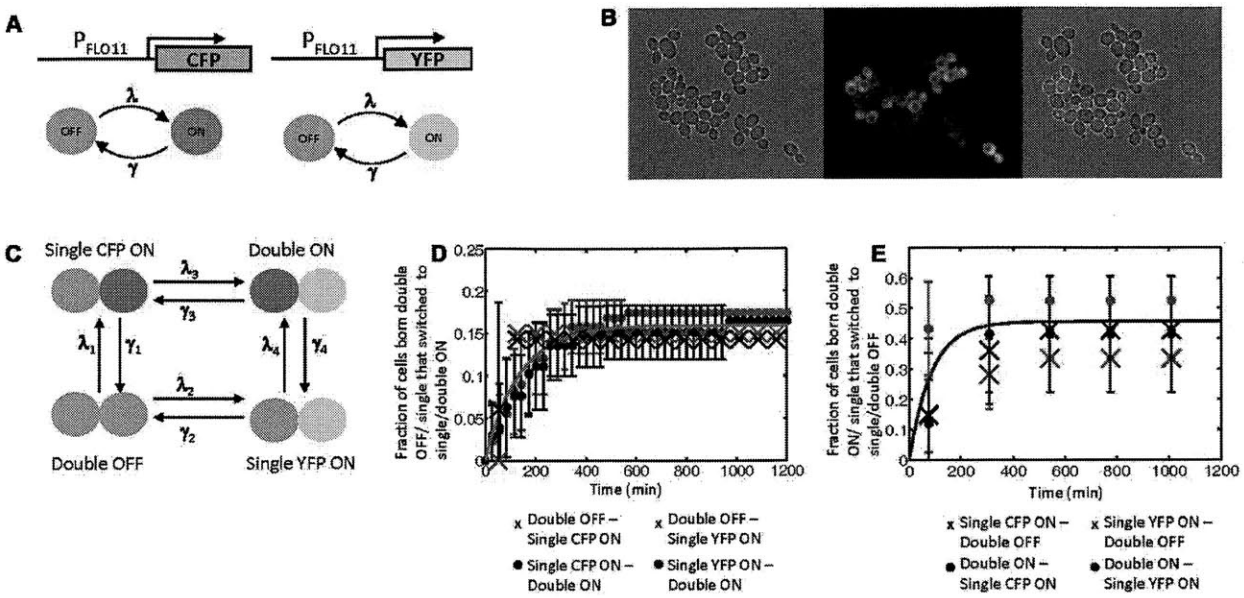

Figure 2: Mixed expression states and independent switching at the *FLO11* locus. A) Dual reporter assay. Each *FLO11* allele turns ON and OFF slowly with identical rates λ and γ because the two reporters are equivalent. B) Mixed expression states. A dual reporter strain grown in rich media (no glucose) supplemented with 1% ethanol and 2% glycerol (false color overlay CFP = red, YFP = Green). All four possible expression states are seen.C) Transition rates. Equivalence of reporters implies $\lambda_1 = \lambda_2, \lambda_3 = \lambda_4, \gamma_1 = \gamma_2, \gamma_3 = \gamma_4$. Independent switching implies $\lambda_1 = \lambda_3$ and $\gamma_1 = \gamma_3$. D) OFF to ON transition rates of different expression states: λ_1 (X), λ_2 (X), λ_3 (•), λ_4 (•). Each marker represents the fraction of cells observed to switch at the particular time, and the pink curve is the same as the fit curve in Figure 3A. E) As in D but for ON to OFF transition rates: γ_1 (X), γ_2 (X), γ_3 (•), γ_4 (•). The blue curve is the same as the fit curve in Figure 3A. D and E demonstrate that transition rates at one allele are independent of the state of the other allele. Even the null hypothesis that γ_2 and γ_4 are equivalent cannot be rejected at the 5% significance level (two-way T-test, p = 0.28) nor can the null hypothesis that their distributions are identical (two-way KS test, p = 0.47).

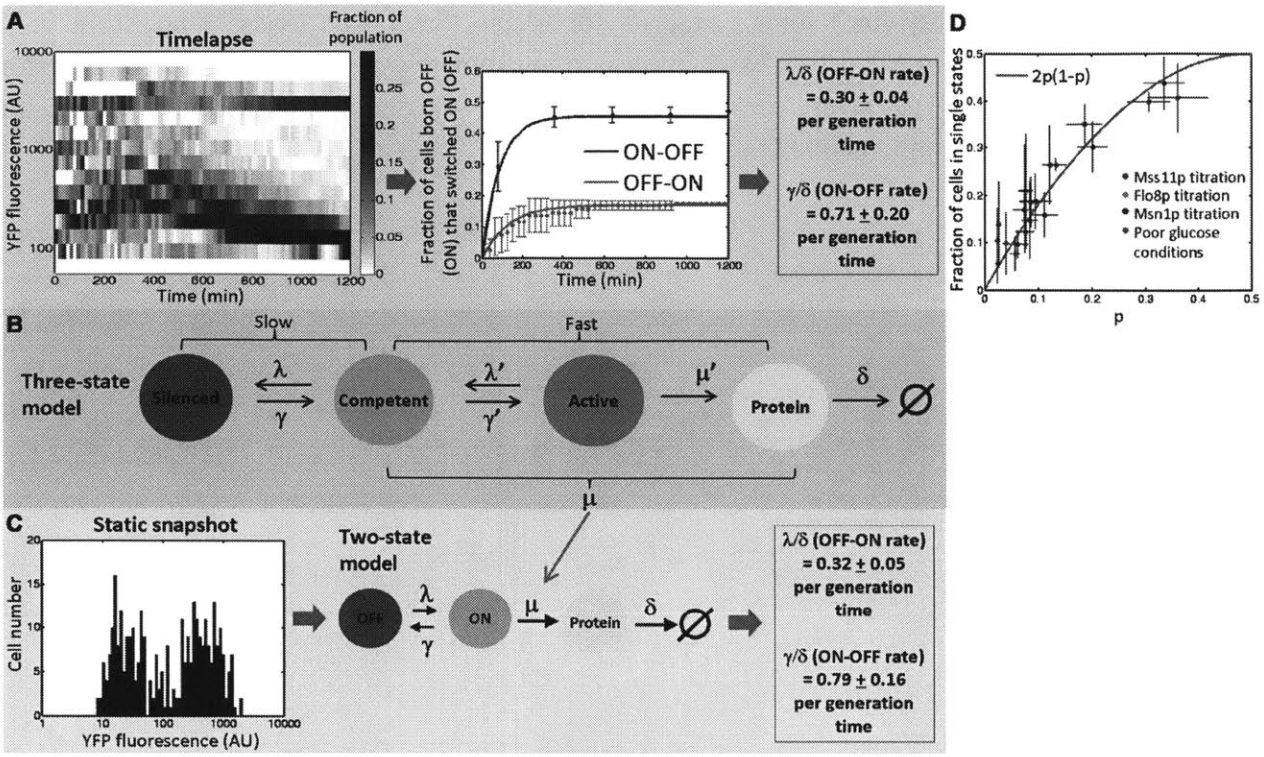

Figure 3: A method for inferring kinetics of switching from static steady state distribution. A) (Left) Time evolution of the population distribution of YFP expression from a dual-reporter strain growing in YP 1% ethanol, 2% glycerol within a microfluidic chamber over 20 hours. Colorbar indicates fraction of cells (n = 230 over time course). This strain had been growing in identical conditions in liquid culture prior to transfer to the microfluidic chamber. The distribution changes early on because of the small initial sample size (n = 10). (Right) Marginal transition rates between ON and OFF states. Blue/pink dots indicate fraction of cells ON/OFF at birth and observed to switch OFF/ON. Corresponding curves are fits of the model for exponentially distributed switching times from ON to OFF and OFF to ON, with adjusted rates shown next to the plot. Error bars correspond to 3 s.d. from the mean calculated by a bootstrap analysis. B) Three-state model of *FLO11* activation showing separation of timescales between epigenetic (silencing) and conventional regulation. When slow transitions associated with silencing are present, the fast transitions of transcriptional bursting can be lumped into a single rate μ. The model then collapses into the two-state model in C. C) (Left) Static distribution of YFP fluorescence of dual-reporter strain grown in identical media conditions as A but in deep well plates rather than the microfluidic device. Transition rates inferred from this snapshot using a stochastic kinetic model (right) agree closely with those obtained by timelapse microscopy. D) Modulation of switching rates. The stochastic kinetic model's prediction of the fraction of cells in the mixed expression state corresponds to independent switching (given by $2p(1-p)$,

corresponding to the gray line) for a range of conditions. Error bars (x-axis) are from 95% confidence intervals from MLE fit of switching rate to estimate true fraction of ON cells; error bars (y-axis) are due to errors in the estimation of threshold fluorescence value for autofluorescence (see Supplemental Discussion).

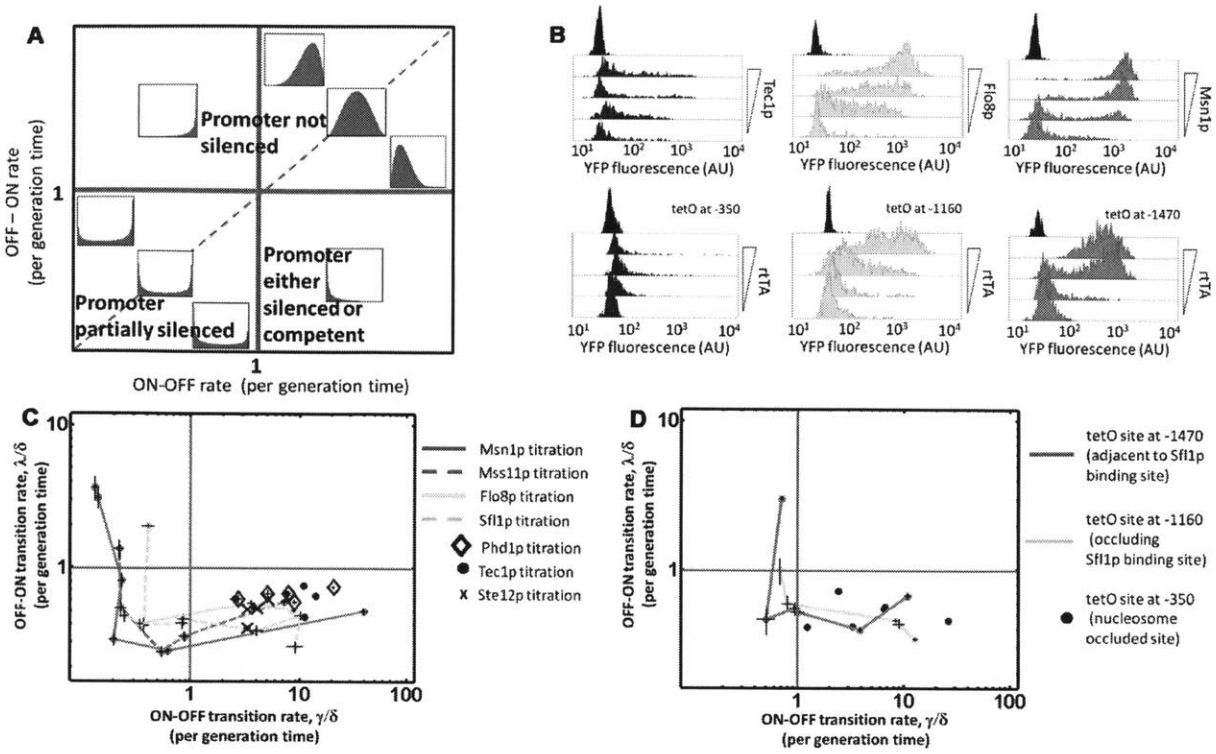

Figure 4: Three different kinetic roles for regulators of *FLO11*. A) The qualitative shape of the Beta distribution for various values of OFF → ON (λ/δ) and ON → OFF (γ/δ) transition rates (normalized with respect to the growth rate δ). When both rates are slower than growth (lower left quadrant) they characterize slow epigenetic transitions between the silenced and competent states. The expression distribution is bimodal, representing stable ON and OFF populations. These rates can be inferred by measuring the expression distribution by fluorescence microscopy and fitting to the Beta distribution. For unimodal ON distributions, epigenetic silencing no longer occurs. If only fast intrinsic fluctuations between the competent and active promoter state were present, the same two-state model would apply, but now predict the fast transition rates and unimodal distributions (upper half of plot). However, because extrinsic fluctuations also matter, direct fitting of measured unimodal distributions does *not* yield the fast transition rates (see text and Supplemental Discussion for details). B) Representative fluorescence histograms of the three activator classes. (Top) Tec1p titrated in wildtype background in SD ura-; Flo8p titrated in *flo8Δ* background in SD ura- + 1% ethanol; Msn1p titrated in wildtype background in SD ura-. (Bottom) rtTA titrated in strain with tetO at -350 (nucleosome occluded site), at -1160 (site occludes Sfl1p binding site), and at -1470 (site directly upstream of the -1200 nucleosome free region). Histograms are derived from fluorescence microscopy (cell number > 300). The fluorescence distribution of an OFF strain (Y92) used to measure autofluorescence is shown at the top of each plot. C) Kinetic roles of regulators. Increasing levels of various activators of *FLO11* decrease γ, stabilizing the active state without significantly changing λ. Class I activators

cannot decrease γ significantly (blue). Class II activators can shift the transition rates into the lower left quadrant which corresponds to partially silenced, bimodal expression (pink). Flo8p has a less stable silenced state compared to the class II activators. It appears that at a critical value of γ the regulators abolish silencing, and the response enters the upper left quadrant. D) Synthetic activator titration. Titrations of synthetic activators mimic the three classes of activators, depending on the location of the binding site. All titrations (B, C and D) were in SD ura- except Sfl1p and Flo8p where 1% ethanol was added. Error bars represent 95% confidence intervals (to fits of experimental data to the Beta distribution for a single value of μ - details in Supplemental Discussion).

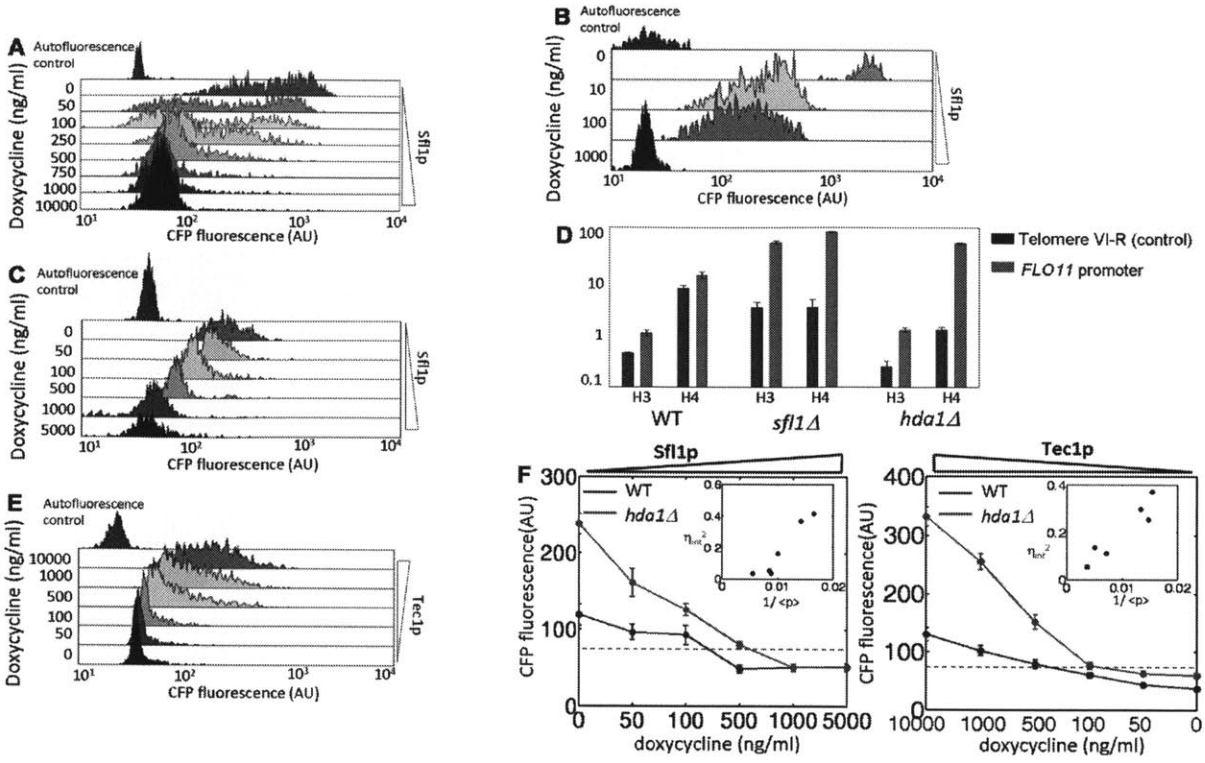

Figure 5: Hda1p is necessary for silencing and a heterogeneous response. A) Sfl1p titrated in *sfl1Δ* background in SD ura- + 1% ethanol leads to a heterogeneous response. B) Sfl1p titrated in *sfl1Δ* background in SD ura- leads to a graded response. C) Sfl1p titrated in *hda1Δ* background in SD ura- + 0.5% ethanol reverts to a graded response. Expression is lower with no doxycycline because of endogenous Sfl1p expression. D) ChIP assay probing for acetylated H3 and H4 histones at *FLO11* promoter (strains grown in SD complete or SD leu-). Probes amplified the -1.7 to -1.5 kb region of the promoter. Signal (y-axis) represents anti-acetylated histone/anti-histone ratio, or an effective average acetylation per histone in the region. Deletion of both *sfl1* and *hda1* result in hyperacetylation of the *FLO11* promoter which is associated with the abrogation of silencing. Therefore *SFL1*-dependent silencing at *FLO11* requires *HDA1*. Error bars are standard error of triplicate quantitative PCR samples. E) When the activator Tec1p is titrated in an *hda1Δ* background, the response is also graded (SD ura-). F) Mean levels of expression during Sfl1p and Tec1p titrations in wildtype (blue curve) and *hda1Δ* (pink curve) backgrounds. Elimination of silencing because of lack of Hda1p lowers the threshold level at which activators function. Furthermore, the population response is graded (panel C and E, see Figure S7 for other activator titrations). Error bars represent 3 s.d. around mean calculated from bootstrap analysis. Inset: The square of intrinsic noise of Sfl1p (left) and Tec1p (right) titrated in *hda1Δ* is proportional to the reciprocal of protein abundance (here shown as the mean fluorescence level). This indicates that without silencing, regulators modulate expression by controlling burst frequency (λ').

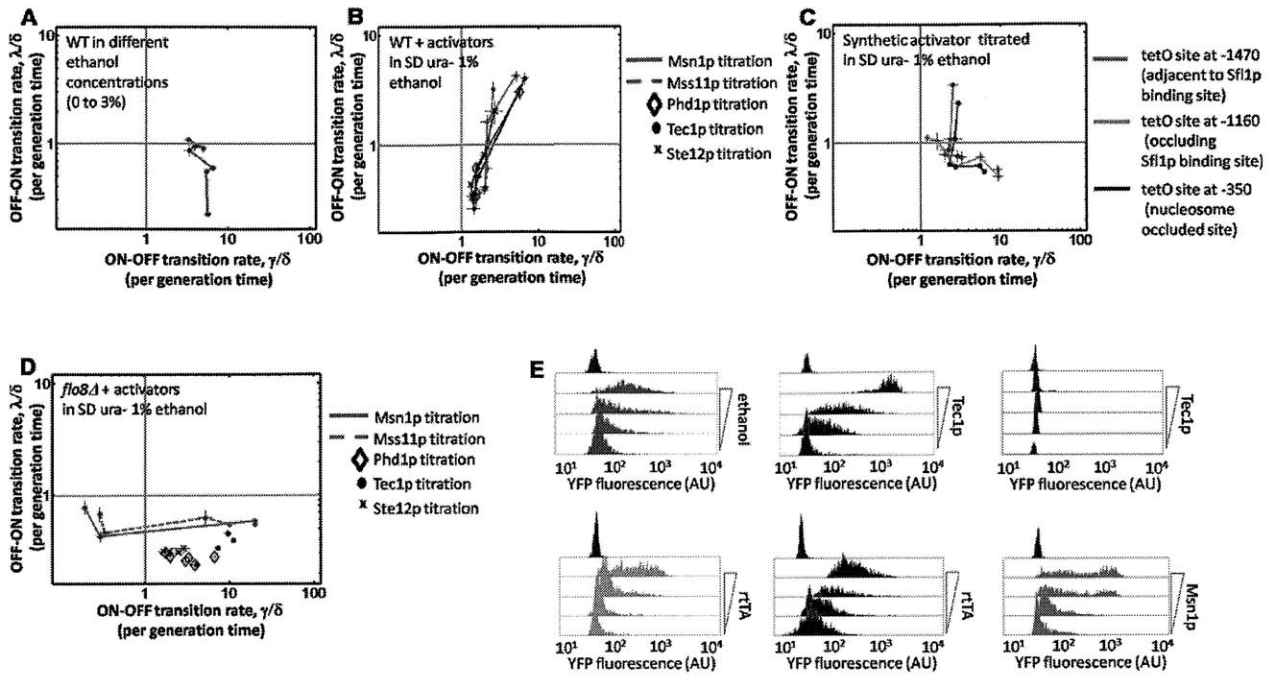

Figure 6: Ethanol modulates silencing at the promoter via Flo8p. A) Wildtype grown in SD complete with ethanol added to final concentration ranging between 0 and 3%. B) Activators titrated in wildtype background in SD ura- + 1% ethanol. All responses are graded, suggesting loss of silencing at the promoter. C) Synthetic activator (rtTA) titration in SD ura- + 1% ethanol. As in panel B, responses are also graded. D) Activators titrated in *flo8Δ* in SD ura- + 1% ethanol. The response is closer to that in Figure 4A rather than Figure 6B, indicating that ethanol abolishes silencing at the promoter through Flo8p. E) Representative fluorescence histograms of titrations shown in A, B, C and D. (Top) wildtype titrated as in panel A, Tec1p titrated as in panel B, and as in panel D. (Bottom) rtTA titrated in strain with tetO site at -1160 (occluding Sfl1p binding site), and at -350 (nucleosome occluded site) as in panel C; Msn1p titrated as in panel D.

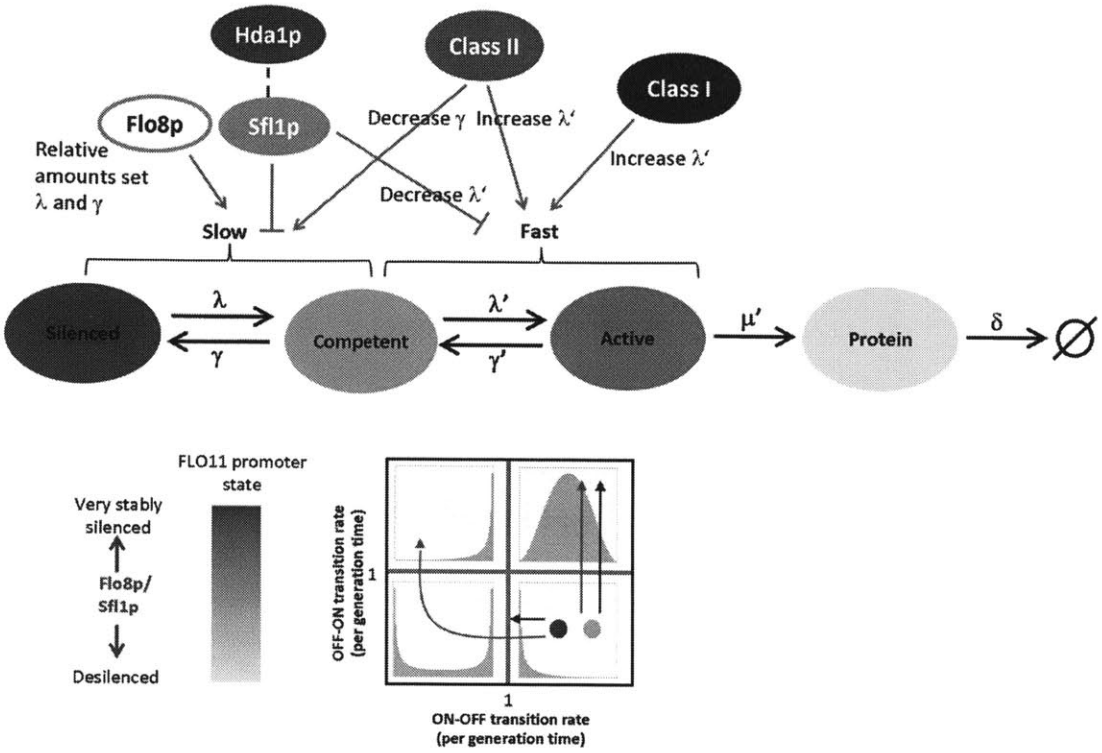

Figure 7: Functional roles of regulators of *FLO11* promoter activation shape the population response. Silencing at the promoter is established by binding of Sfl1p and recruitment of Hda1p. Relative activities of Flo8p and Sfl1p determine the chromatin state of the promoter. The underlying promoter state of an OFF population can be revealed by addition of Class I and II activators (bottom), as Class I activators cannot effectively activate transcription at a silenced promoter, whereas Class II activators can activate expression by sufficiently stabilizing the competent state. At very high levels, Class II activators disrupt silencing; at this point, all cells are also expressing highly. In contrast, an intermediate level of Flo8p activity can "open" the promoter converting the silenced state to a stable competent state, while expression remains low. This opening might be related to chromatin modifications. The combination of Flo8p activation and Class I activators allows the decoupling of chromatin state and expression level, whereas activation by Class II activators alone would not.

SUPPLEMENTAL TABLE 1: Yeast strains used in study

Strains constructed from Σ1278 ML *ura3-52, leu2Δ::hisG* (Lorenz, and Heitman, 1997)

Strain	***Genotype* (Source)**
MLY42	**Σ1278b** *MAT* α *ura3-52 leu2Δ::hisG* (Lorenz and Heitman, 1997)
MLY43	**Σ1278b** *MAT* **a** *ura3-52 leu2Δ::hisG*
Y35	*MAT* α *flo11Δ*::CFP-KanMX6 *ura3-52 leu2Δ::hisG* (GAL HO switched Y37)**
Y36	*MAT* α *flo11Δ*::YFP-KanMX6 *ura3-52 leu2Δ::hisG* (GAL HO switched Y38)**
Y37	*MAT* a *flo11Δ*::CFP-KanMX6 *ura3-52 leu2Δ::hisG*
Y38	*MAT* a *flo11Δ*::YFP-KanMX6 *ura3-52 leu2Δ::hisG*
Y39	*MAT* a/α *ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y40	*MAT* a/α *FLO11/flo11Δ::CFP-KanMX6 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y41	*MAT* a/α *flo11Δ::CFP-KanMX6/FLO11 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y42	*MAT* a/α *FLO11/flo11Δ::YFP-KanMX6 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y43	*MAT* a/α *flo11Δ::YFP-KanMX6/FLO11 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y44	*MAT* a/α *flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y45	*MAT* α/**a** *flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 ura3-52/ura3-52 leu2::hisG leu2Δ::hisG*
Y92	*MAT* α/a *flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 flo8Δ::LEU2/flo8Δ::LEU2 ura3-52/ura3-52 leu2 Δ::hisG/leu2Δ::hisG*
Y93	*MAT* α/a *flo11Δ::YFP-KanMX6/ flo11Δ::CFP-KanMX6 sfl1Δ::LEU2/sfl1 Δ::LEU2 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y171	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 hda1Δ::LEU2/hda1Δ::LEU2 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*
Y197	*MAT* α/a *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 mss11Δ::LEU2 /mss11Δ::LEU2 ura3-52/ura3-52 leu2 Δ::hisG/ leu2Δ::hisG*
Y253	*MAT* α/a *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -350 region of FLO11 promoter::tetO site/-350 region of FLO11 promoter::tetO site ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*
Y254	*MAT* α/a *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -1400 region of FLO11 promoter::tetO site/-1400 region of FLO11 promoter::tetO site ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*
Y255	*MAT* α/a *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -1470 region of FLO11 promoter::tetO site/-1470 region of FLO11 promoter::tetO site ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*

*** All strains constructed are in the Σ1278b background and cogenic to MLY43.**

**** Haploid strains Y35, Y36 were constructed by HO-mediated mating type switching of strains Y37, Y38, to ensure all diploids were created from isogenic haploid strains.**

Strains with selected *FLO11* regulators knocked out were created by PCR integration to delete the regulator ORF in the YFP and CFP haploids and then mated to obtain diploids. Yeast were transformed by the PEG/Lithium Acetate method (Gietz, and Woods, 2002) and all integrations were verified through colony PCR. Strains with plasmids transformed were kept under selection media.

SUPPLEMENTAL TABLE 2: Plasmids used in study

Plasmid	*Description*
pYC	*CEN URA3* ADH1 promoter-rtTA
pYCFlo8	7xtetO site (XhoI/BamHI), *FLO8 ORF* (BamHI/NotI) in pYC
pYCSfl1	7xtetO site (XhoI/BamHI), *SFL1 ORF* (BamHI/NotI) in pYC
pYCPhd1	7xtetO site (XhoI/BamHI), *PHD1 ORF* (BamHI/NotI) in pYC
pYCTec1	7xtetO site (XhoI/BamHI), *TEC1 ORF* (BamHI/NotI) in pYC
pYCSte12	7xtetO site (XhoI/BamHI), *STE12 ORF* (BamHI/NotI) in pYC
pYCMsn1	7xtetO site (XhoI/BamHI), *MSN1 ORF* (BamHI/NotI) in pYC
pYCMss11	7xtetO site (XhoI/BamHI), *MSS11 ORF* (BamHI/NotI) in pYC

SUPPLEMENTAL FIGURES

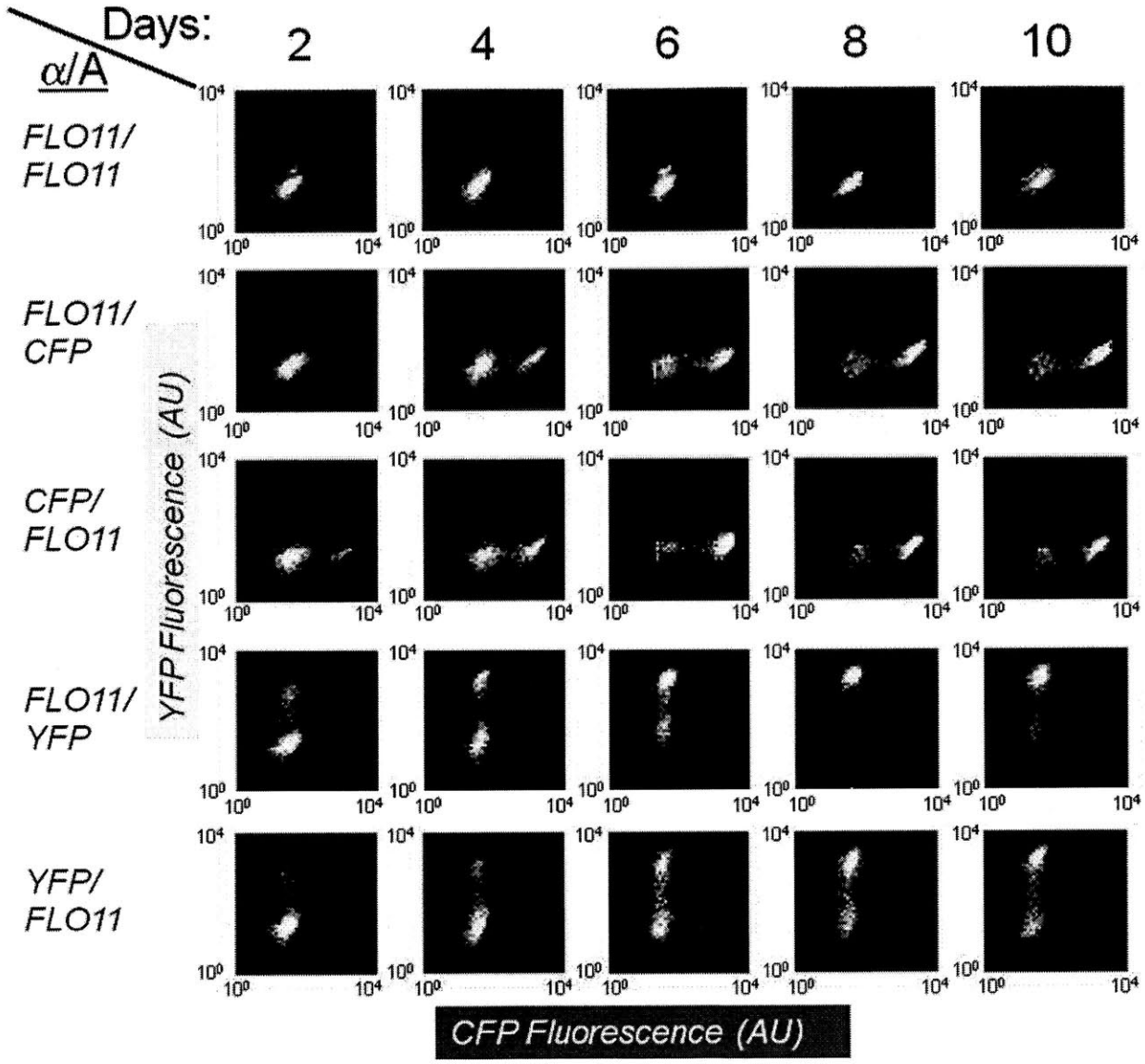

FIGURE S1: *FLO11* expression on solid media – single reporter strains. A) 10 μL of a mid-log phase culture of various single and dual reporter strain was spotted on a fresh YPD plate. Labels on the left indicate whether and at which locus the *FLO11* ORF was replaced with a particular fluorescent protein variant. Plates were left at room temperature. Cells from all regions of the spot were sampled (see supplemental text) and CFP and YFP expression of these samples was monitored every 2 days by fluorescence microscopy. Density plots for each sample are given, where the x-axis is log CFP fluorescence levels and the y-axis is log YFP fluorescence levels. Cellular autofluorescence can be estimated based on the (first) control strain). Both fluorescence reporters respond equivalently, whether integrated at the **a** or α locus.

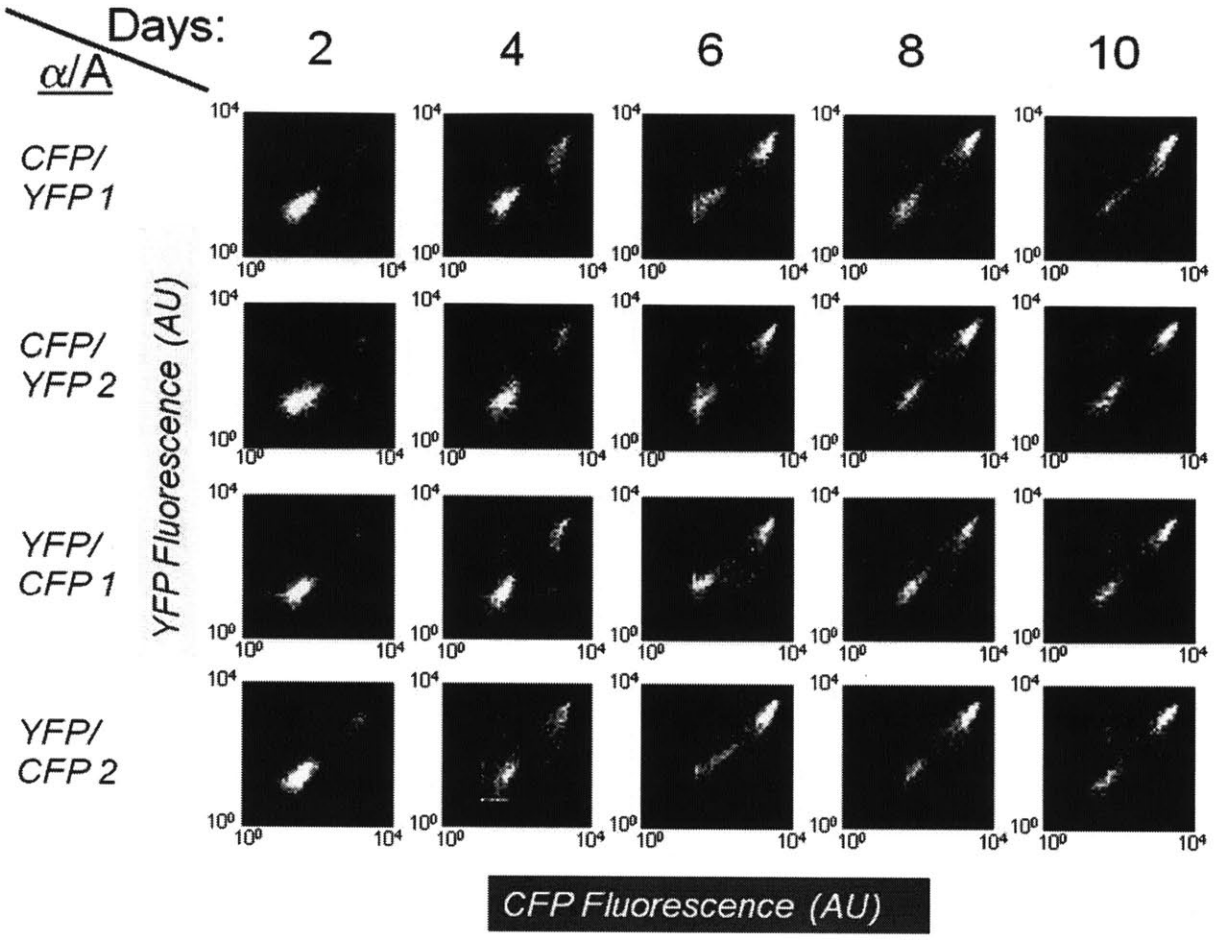

FIGURE S2: *FLO11* expression on solid media – dual reporter strains. As in panel Figure S1. Four different dual reporter strains were constructed, two with CFP at the α locus and YFP and the **a** locus, and two in the opposite configuration. Their response is similar, verifying that both reporters are equivalent. Furthermore, the expression distribution of each individual fluorescent reporter is equivalent to the corresponding single reporter strain in panel A, verifying independence and the fact that *FLO11* expression doesn't feedback and affect its own expression.

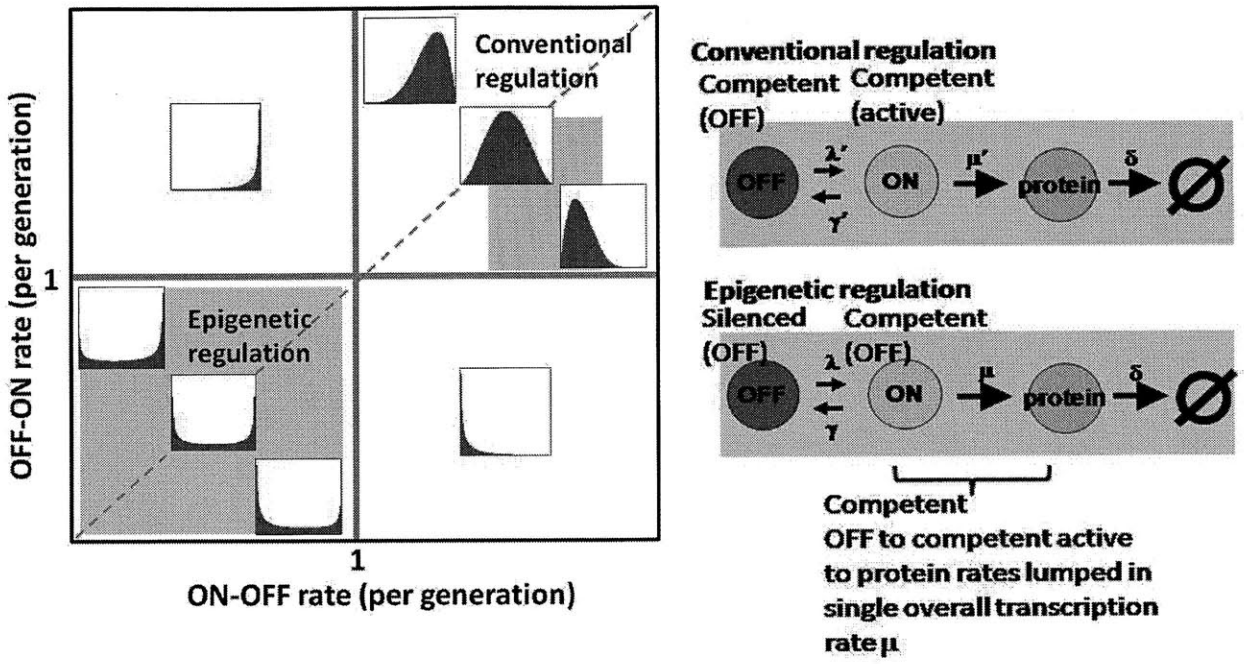

FIGURE S3: Switching times and fluorescence distribution. The qualitative behavior of the Beta distribution, as described in Figure 4A in the main text, is characterized by a bimodal distribution at low transition rates (< once per cell generation). As the rates increase to greater than once per generation, the two peaks blend, and a gamma-like distribution is observed (upper right quadrant). In epigenetic regulation (silencing present), the rate limiting step is the silenced to competent OFF transition. The regime in which these rates lie is indicated by the shaded green region; by definition, these rates should lie in the region where the transitions occur less than once per cell generation. Meanwhile, in conventional regulation (silencing absent), the silenced to competent OFF transition is gone, and the fast competent OFF to competent ACTIVE transition now dominate. These rates correspond to the burst frequency (λ') and burst size (μ'/γ'). Previous studies that have measured these rates indicate that they lie approximately in the purple shaded region shown above. Rare bursting and measured burst sizes indicate that this region lies below the diagonal, although the region's width is only a guess and could vary.

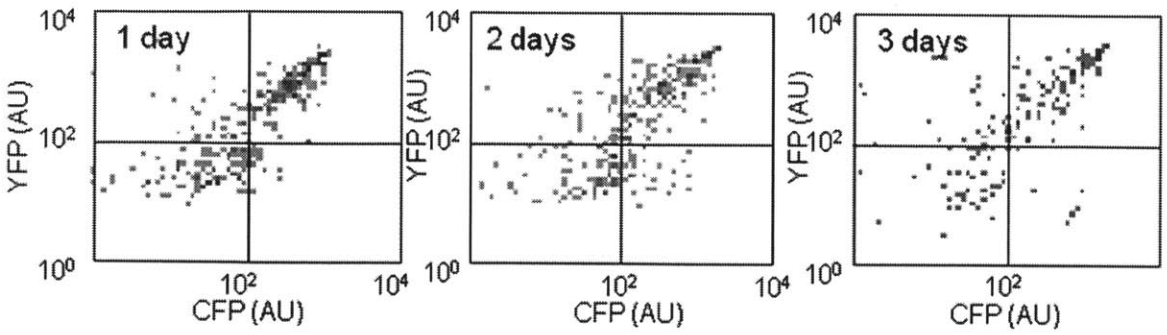

FIGURE S4: Static snapshots of Y45 cells in YP 1% Ethanol, 2% Glycerol maintained in exponential phase by dilution. The null hypothesis that distributions of YFP fluorescence at each time point are equivalent cannot be rejected (two-way Kolmogorov-Smirnov test, $p = 0.60$, 0.25, 0.85 for day 1 vs. day 2, day 2 vs. day 3 and day 1 vs. day 3 respectively). Similarly, the null hypothesis that CFP fluorescence distributions at each time point are equivalent cannot be rejected (two-way Kolmogorov-Smirnov test, $p = 0.66$, 0.36, 0.88 for day 1 vs. day 2, day 2 vs. day 3 and day 1 vs. day 3 respectively).

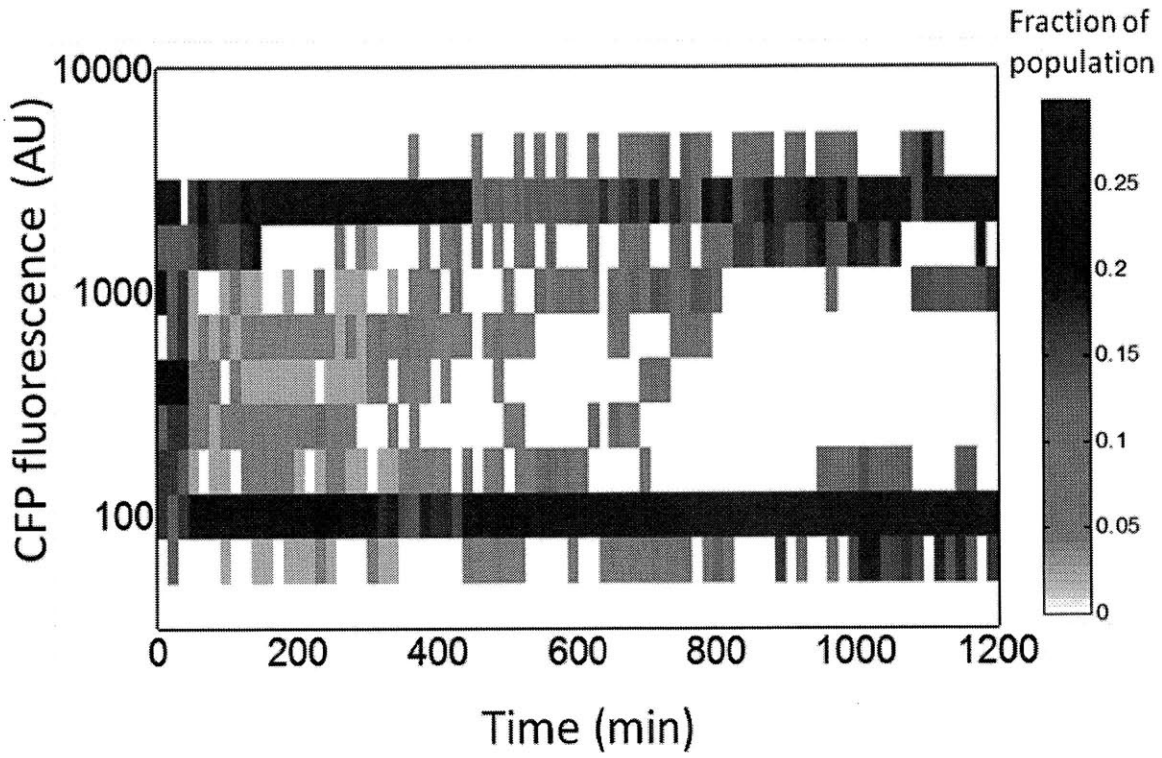

FIGURE S5: CFP expression distribution during timelapse. As in Figure 3A, except for CFP rather than YFP.

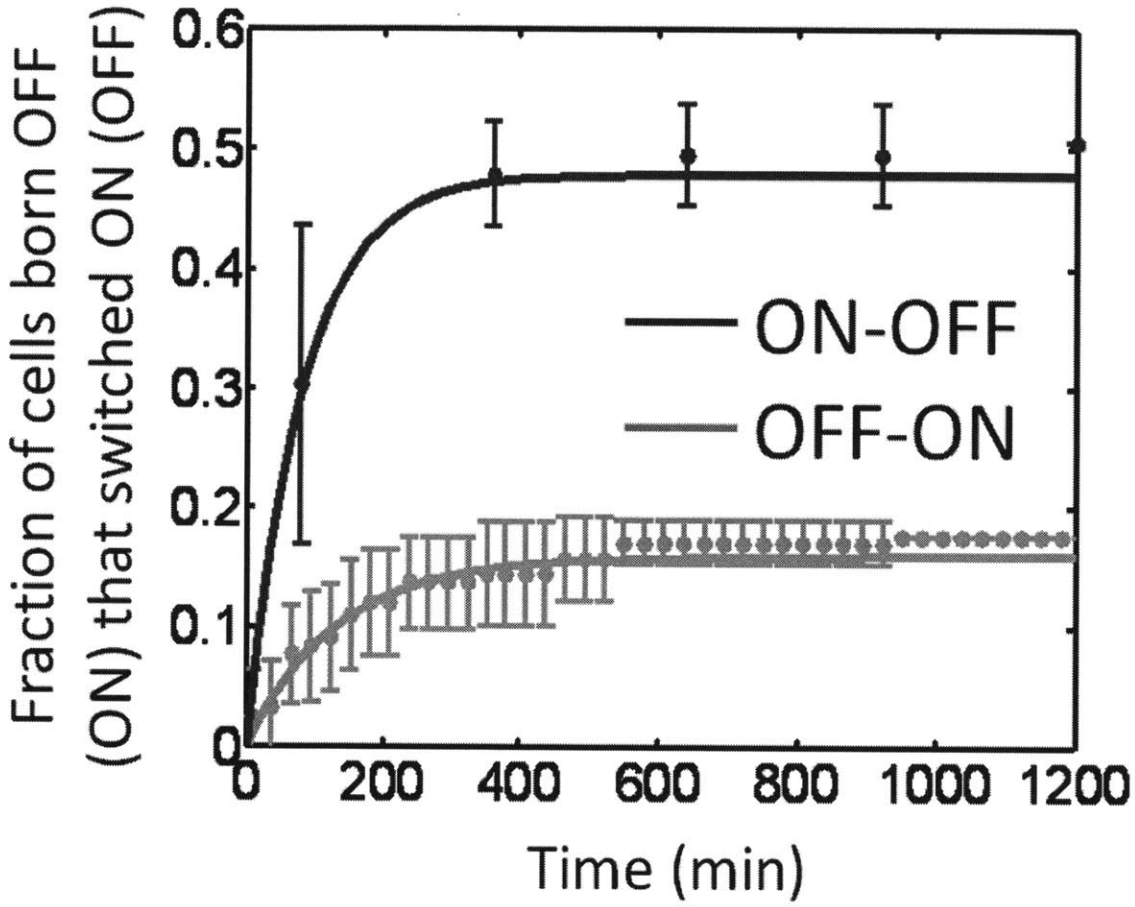

Figure S6: **Switching rates of CFP reporter.** As in Figure 3A in the main text, but for CFP rather than YFP. The fit yields switching rates for CFP were λ/δ (OFF-ON) = 0.25 $\pm$ 0.03 generation^{-1} (pink), γ/δ (ON-OFF) =0.90 $\pm$ 0.17 generation^{-1} (blue).

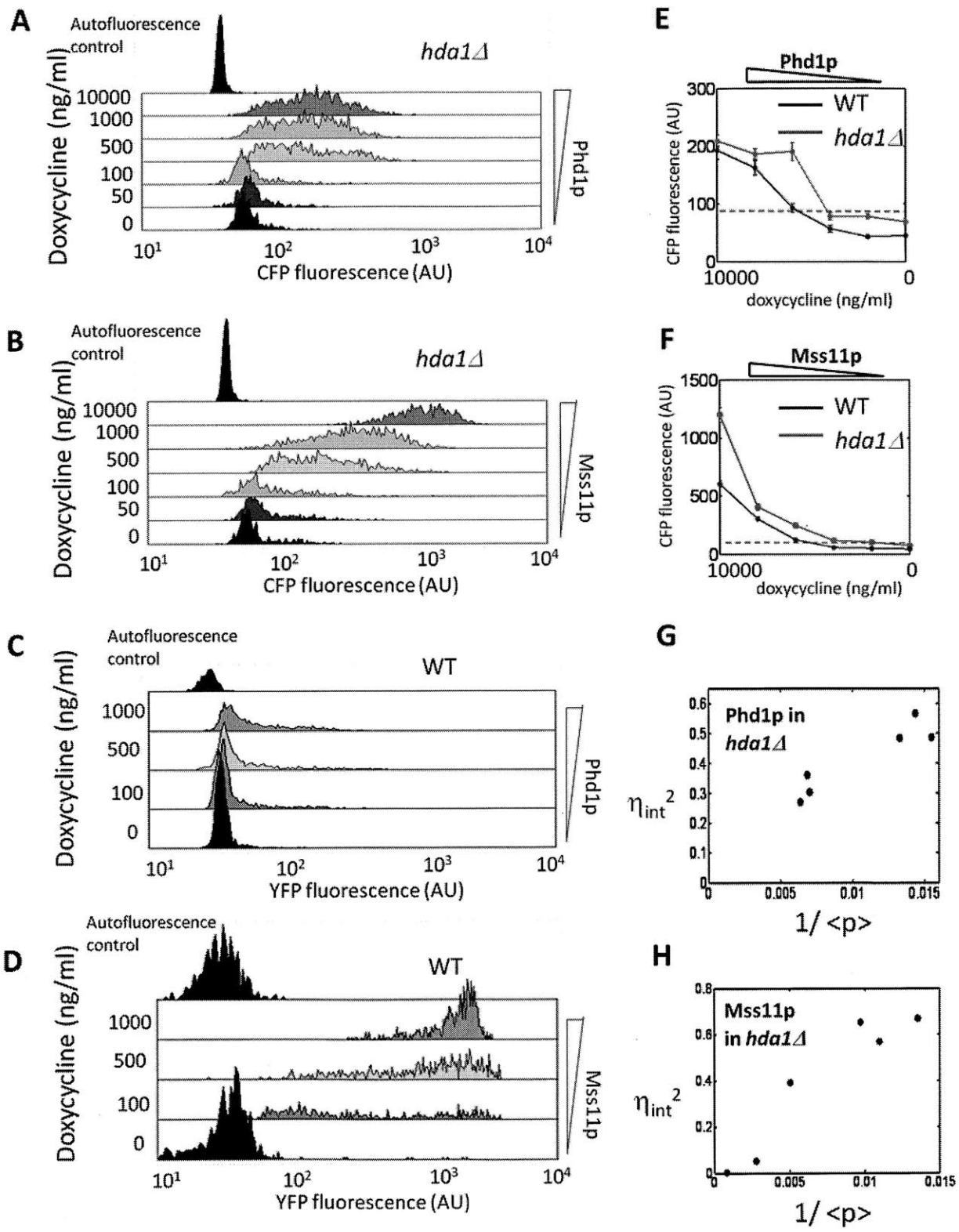

FIGURE S7: Loss of Hda1p converts activators Mss11p and Phd1p response from heterogeneous to graded. As in Figure 4B and Figure 5E in the main text where Tec1p was

titrated in a wildtype (Y45) and *hda1Δ* background respectively, the other activators Phd1p and Mss11p also exhibit a graded response in an *hda1Δ* background (panels A and B). In the wildtype background, Phd1p (panel C) like Tec1p is unable to stabilize the ON state enough to enter the bimodal regime, whereas Mss11p (panel D), like Msn1p (shown in Figure 4B), is able to do so. As in Figure 5F, elimination of silencing in the *hda1Δ* background lowers the threshold level at which Mss11p and Phd1p function (panels E and F). Error bars represent 3 standard deviations around the mean from bootstrap analysis. Like Tec1p in *hda1Δ* in Figure 5F, both Mss11p and Phd1p control burst frequency (λ') in the absence of silencing, as the square of the intrinsic noise of Phd1p titrated in *hda1Δ* (panel G) and Mss11p titrated in *hda1Δ* (panel H) scale with the reciprocal of protein abundance. All titrations were done in SD ura-.

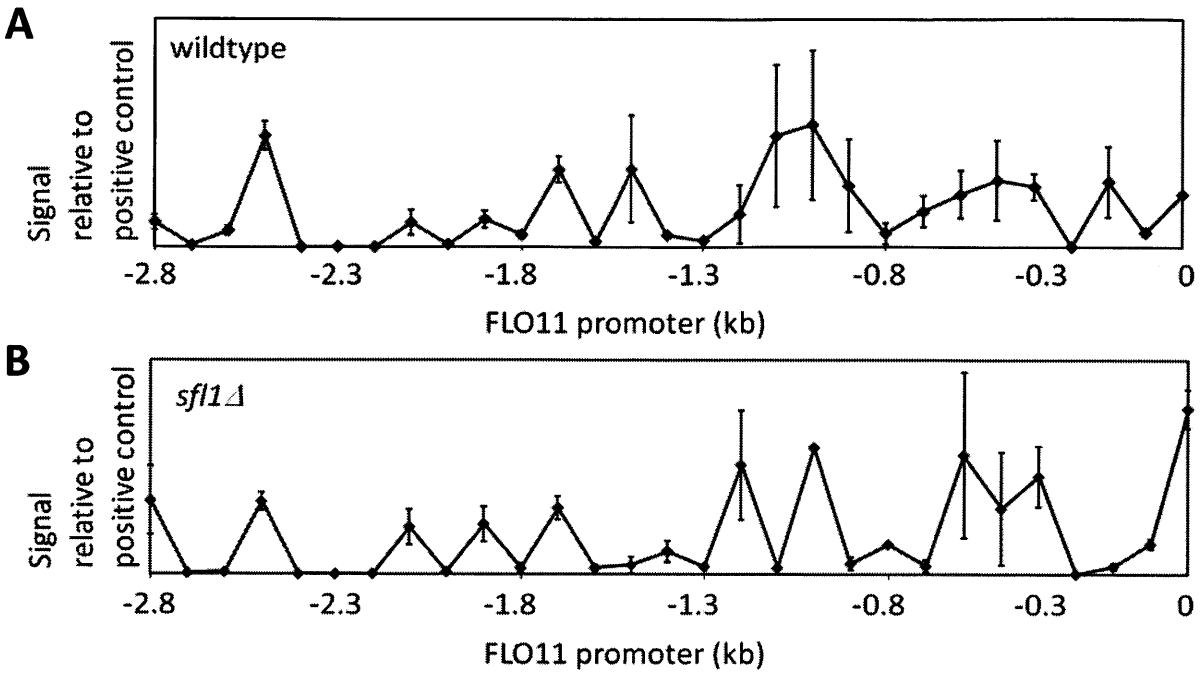

Figure S8: Micrococcal nuclease mapping of *FLO11* was performed on cells grown in conditions where the promoter was completely silenced (growth of wildtype Y45 in SD complete) in panel A or completely active (growth of an *sfl1Δ* strain in SD complete plus 2% glucose) in panel B. The overall structure agrees well with the *in silico* predictions in Figure 1. Although nucleosomal occupancy in the -1400 bp region and the -150 bp region appears to be further depleted in the active state, there is surprisingly no gross rearrangement of nucleosomal structure between the silenced and active state. Error bars are standard error from triplicate quantitative PCR samples.

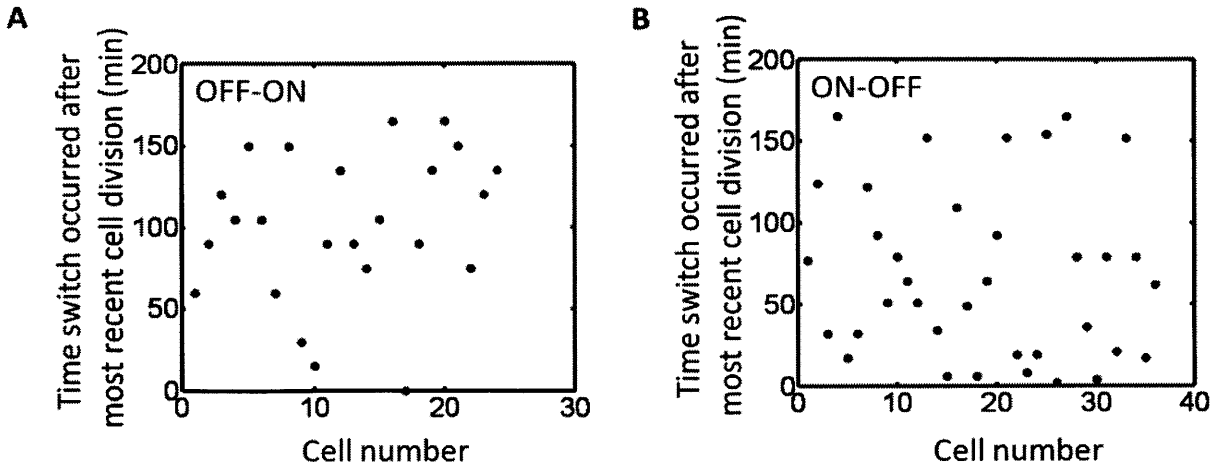

Figure S9: OFF(ON)-ON(OFF) switch at *FLO11* is not correlated with cell cycle stage. The time at which a cell born OFF switches ON (panel A) or the time when a cell born ON switches OFF (panel B) after its most recent division event occurred is shown for each cell observed to switch during the timelapse experiment in Figure 3A. Points at which the switch occurs do not appear to cluster at any particular position during the cell cycle.

REFERENCES

Aparicio, O., Geisberg, J.V., and Struhl, K. (2004). Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. Curr. Protoc. Cell. Biol. Unit 17.7.

Avery, S.V. (2006). Microbial cell individuality and the underlying sources of heterogeneity. Nat. Rev. Microbiol. *8,* 577-587.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. Nat. Genet. *6,* 636-643.

Barry, J.D., Ginger, M.L., Burton, P., and McCulloch, R. (2003). Why are parasite contingency genes often associated with telomeres? Int. J. Parasitol. *1,* 29-45.

Belli, G., Gari, E., Aldea, M., and Herrero, E. (1998). Functional analysis of yeast essential genes using a promoter-substitution cassette and the tetracycline-regulatable dual expression system. Yeast *12,* 1127-1138.

Berger, S.L., Pina, B., Silverman, N., Marcus, G.A., Agapite, J., Regier, J.L., Triezenberg, S.J., and Guarente, L. (1992). Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. Cell *2,* 251-265.

Blake, W.J., Balazsi, G., Kohanski, M.A., Isaacs, F.J., Murphy, K.F., Kuang, Y., Cantor, C.R., Walt, D.R., and Collins, J.J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. Mol. Cell *6,* 853-865.

Borneman, A.R., Leigh-Bell, J.A., Yu, H., Bertone, P., Gerstein, M., and Snyder, M. (2006). Target hub proteins serve as master regulators of development in yeast. Genes Dev. *4,* 435-448.

Cai, L., Friedman, N., and Xie, X.S. (2006). Stochastic protein expression in individual cells at the single molecule level. Nature *7082,* 358-362.

Chen, H., and Fink, G.R. (2006). Feedback control of morphogenesis in fungi by aromatic alcohols. Genes Dev. *9,* 1150-1161.

Chubb, J.R., Trcek, T., Shenoy, S.M., and Singer, R.H. (2006). Transcriptional pulsing of a developmental gene. Curr. Biol. *10,* 1018-1025.

Conlan, R.S., and Tzamarias, D. (2001). Sfl1 functions via the co-repressor Ssn6-Tup1 and the cAMP-dependent protein kinase Tpk2. J. Mol. Biol. *5,* 1007-1015.

De Las Penas, A., Pan, S.J., Castano, I., Alder, J., Cregg, R., and Cormack, B.P. (2003). Virulence-related surface glycoproteins in the yeast pathogen Candida glabrata are encoded in

subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. Genes Dev. *18,* 2245-2258.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *5584,* 1183-1186.

Gagiano, M., van Dyk, D., Bauer, F.F., Lambrechts, M.G., and Pretorius, I.S. (1999). Msn1p/Mss10p, Mss11p and Muc1p/Flo11p are part of a signal transduction pathway downstream of Mep2p regulating invasive growth and pseudohyphal differentiation in Saccharomyces cerevisiae. Mol. Microbiol. *1,* 103-116.

Galitski, T., Saldanha, A.J., Styles, C.A., Lander, E.S., and Fink, G.R. (1999). Ploidy regulation of gene expression. Science *5425,* 251-254.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *6767,* 339-342.

Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 87-96.

Golding, I., Paulsson, J., Zawilski, S.M., and Cox, E.C. (2005). Real-time kinetics of gene activity in individual bacteria. Cell *6,* 1025-1036.

Guo, B., Styles, C.A., Feng, Q., and Fink, G.R. (2000). A Saccharomyces gene family involved in invasive growth, cell-cell adhesion, and mating. Proc. Natl. Acad. Sci. U. S. A. *22,* 12158-12163.

Halme, A., Bumgarner, S., Styles, C., and Fink, G.R. (2004). Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. Cell *3,* 405-415.

Hernday, A.D., Braaten, B.A., Broitman-Maduro, G., Engelberts, P., and Low, D.A. (2004). Regulation of the pap epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. Mol. Cell *4,* 537-547.

Hughes, T., Badis, G., Pena-Castillo, L., Chan, E., Talukder, S., Coburn, D., Terterov, D., Yang, A., Min, R., Gosset, A. *et al.* (2008) Paper presented at Yeast Genetics and Molecular Biology Meeting. (Toronto: Toronto).

Ingolia, N.T., and Murray, A.W. (2007). Positive-feedback loops as a flexible biological module. Curr. Biol. *8,* 668-677.

Kaern, M., Elston, T.C., Blake, W.J., and Collins, J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. *6,* 451-464.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., and Segal, E. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *7236,* 362-366.

Kaufmann, B.B., and van Oudenaarden, A. (2007). Stochastic gene expression: from single molecules to the proteome. Curr. Opin. Genet. Dev. *2,* 107-112.

Kaufmann, B.B., Yang, Q., Mettetal, J.T., and van Oudenaarden, A. (2007). Heritable stochastic switching revealed by single-cell genealogy. PLoS Biol. *9,* e239.

Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. Nature *7192,* 246-250.

Laroche, T., Martin, S.G., Gotta, M., Gorham, H.C., Pryde, F.E., Louis, E.J., and Gasser, S.M. (1998). Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. Curr. Biol. *11,* 653-656.

Lim, H.N., and van Oudenaarden, A. (2007). A multistep epigenetic switch enables the stable inheritance of DNA methylation states. Nat. Genet. *2,* 269-275.

Lo, W.S., and Dranginis, A.M. (1998). The cell surface flocculin Flo11 is required for pseudohyphae formation and invasion by Saccharomyces cerevisiae. Mol. Biol. Cell *1,* 161-171.

Lorenz, M.C., and Heitman, J. (1998). Regulators of pseudohyphal differentiation in Saccharomyces cerevisiae identified through multicopy suppressor analysis in ammonium permease mutant strains. Genetics *4,* 1443-1457.

Lorenz, M.C., and Heitman, J. (1997). Yeast pseudohyphal growth is regulated by GPA2, a G protein alpha homolog. EMBO J. *23,* 7008-7018.

Low, D.A., Weyand, N.J., and Mahan, M.J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect. Immun. *12,* 7197-7204.

Madhani, H.D., and Fink, G.R. (1997). Combinatorial control required for the specificity of yeast MAPK signaling. Science *5304,* 1314-1317.

Maheshri, N., and O'Shea, E.K. (2007). Living with noisy genes: how cells function reliably with inherent variability in gene expression. Annu. Rev. Biophys. Biomol. Struct. 413-434.

Mettetal, J.T., Muzzey, D., Pedraza, J.M., Ozbudak, E.M., and van Oudenaarden, A. (2006). Predicting stochastic gene expression dynamics in single cells. Proc. Natl. Acad. Sci. U. S. A. *19,* 7304-7309.

Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature *7095,* 840-846.

Pan, X., and Heitman, J. (2002). Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. Mol. Cell. Biol. *12,* 3981-3993.

Pays, E., Vanhamme, L., and Perez-Morga, D. (2004). Antigenic variation in Trypanosoma brucei: facts, challenges and mysteries. Curr. Opin. Microbiol. *4,* 369-374.

Pelechano, V., Garcia-Martinez, J., and Perez-Ortin, J.E. (2006). A genomic study of the inter-ORF distances in Saccharomyces cerevisiae. Yeast *9,* 689-699.

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS Biol. *10,* e309.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of genetic and epigenetic inheritance. Cell *4,* 655-668.

Raser, J.M., and O'Shea, E.K. (2004). Control of stochasticity in eukaryotic gene expression. Science *5678,* 1811-1814.

Rupp, S., Summers, E., Lo, H.J., Madhani, H., and Fink, G. (1999). MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast FLO11 gene. EMBO J. *5,* 1257-1269.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *1,* 19-27.

Song, W., and Carlson, M. (1998). Srb/mediator proteins interact functionally and physically with transcriptional repressor Sfl1. EMBO J. *19,* 5757-5765.

Suka, N., Suka, Y., Carmen, A.A., Wu, J., and Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Mol. Cell *2,* 473-479.

van Dyk, D., Pretorius, I.S., and Bauer, F.F. (2005). Mss11p is a central element of the regulatory network that controls FLO11 expression and invasive growth in Saccharomyces cerevisiae. Genetics *1,* 91-106.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. Nat. Genet. *9,* 986-990.

Verstrepen, K.J., and Klis, F.M. (2006). Flocculation, adhesion and biofilm formation in yeasts. Mol. Microbiol. *1,* 5-15.

Wu, J., and Grunstein, M. (2000). 25 Years After the Nucleosome Model: Chromatin Modifications. Trends Biochem. Sci. *12,* 619-623.

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

Xu, E.Y., Zawadzki, K.A., and Broach, J.R. (2006). Single-cell observations reveal intermediate transcriptional silencing states. Mol. Cell *2,* 219-229.

Yang, B., and Kirchmaier, A.L. (2006). Bypassing the catalytic activity of SIR2 for SIR protein spreading in Saccharomyces cerevisiae. Mol. Biol. Cell *12,* 5287-5297.

Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X.S. (2006). Probing gene expression in live cells, one protein molecule at a time. Science *5767,* 1600-1603.

Chapter 3

DNA LOOP FORMATION STABILIZES SILENCED AND ACTIVE STATES OF A YEAST PROMOTER

Author's Note: The material presented in this chapter are in a manuscript to be submitted for publication.

I contributed to this work by conceiving most of the questions asked, constructing most of the strains needed, designing and carrying out the experiments, and analyzing the data.

ABSTRACT

The ability to control the degree of heterogeneity in cellular phenotypes may be important for cell populations to survive uncertain and ever-changing environments or make cell-fate decisions in response to external stimuli. Cells may control the degree of gene expression heterogeneity and ultimately levels of phenotypic heterogeneity by modulating promoter switching dynamics. In particular, a cell population can have well-defined fractions of "off" and "on" cells if the promoter switching rate is slower than the cell division rate. Well-studied mechanisms to enable slow promoter dynamics are those which encode the slow step in *trans*, where slow promoter switching arises from rare fluctuations in levels of a *trans*-factor. However, mechanisms to encode the slow step in *cis* are poorly understood. Here we demonstrate how multiple molecular events occurring at a gene's promoter might lead to an overall slow step in *cis*. At the *FLO11* promoter in *S. cerevisiae*, we show at least two pathways that recruit histone deacetylases to the promoter and *in vivo* association between the region -1.2 kb from the ATG start site of the *FLO11* ORF and the core promoter region are all required for a stable silenced state. To generate bimodal gene expression, the activator Msn1p forms an alternate looped conformation, where the core promoter associates with the non-coding RNA *PWR1*'s promoter and terminator regions, located at -2.1 kb and -3.0 kb from the ATG start site of the *FLO11* ORF respectively. Formation of the active looped conformation is required for Msn1p's ability to stabilize the competent state

without destabilizing the silenced state and generate a bimodal response. Our results support a model where multiple stochastic steps at the promoter are required to transition between the silenced and active states, leading to an overall slow step in *cis*.

INTRODUCTION

Events such as DNA looping and non-coding RNA transcription have been increasingly implicated in regulation of gene expression. Although the prevalence of DNA looping is recognized in various organisms, including yeast, mice and humans (Di, Wang, et al, 2008; Fullwood, Liu, et al, 2009; Kagey, Newman, et al, 2010; Rodley, Bertels, et al, 2009; Zhao, Tavoosidana, et al, 2006), its functional roles are only beginning to be elucidated. In human cells, looping has been frequently correlated with transcriptional activation or repression, suggesting that long-range DNA interactions play an important role in gene regulation (Sexton, Bantignies, and Cavalli, 2009). In *S. cerevisiae*, looping has been shown to confer faster gene reactivation kinetics, where association between a gene's promoter and terminator regions is thought to facilitate rapid re-association of RNA Pol II and more efficient recycling of transcriptional machinery components (Hampsey, Singh, et al, 2010; Laine, Singh, et al, 2009; Tan-Wong, Wijayatilake, and Proudfoot, 2009).

Non-coding RNA (ncRNA) transcription has been found to occur across virtually all euchromatic regions of the genome, in both mammalian cells and yeast (Bertone, Stolc, et al, 2004; Carninci, Kasukawa, et al, 2005; Dutrow, Nix, et al, 2008; Kapranov, Cawley, et al, 2002; Rinn, Euskirchen, et al, 2003). In yeast, ncRNA transcription generally represses expression of a gene in *cis* either by transcriptional interference and/or recruitment of chromatin-modifying enzymes which promote repressive chromatin. In transcriptional interference, an antisense transcript can suppress transcription of a sense transcript (or vice-versa), likely through collision

of transcriptional machinery. Such a mechanism was shown to regulate expression of *IME4*, a key regulator of meiosis in *S. cerevisiae* (Hongay, Grisafi, et al, 2006). Transcriptional interference may also involve ncRNA transcription across a promoter region that possibly occludes binding of transcription factors; at the *SER3* locus, intergenic ncRNA transcription was required to repress of *SER3* expression (Martens, Laprade, and Winston, 2004). Transcription of ncRNA across a region has also been found to recruit histone deacetylases such as Hda1p (Camblong, Iglesias, et al, 2007) and Rpd3p (Houseley, Rubbi, et al, 2008), which locally deacetylate histones at a gene promoter and create a chromatin environment that is prohibitive to transcription. While most studies have reported ncRNA transcription resulting in gene repression, ncRNA transcription has also been shown to result in gene activation. In *S. pombe*, a cascade of ncRNA transcription at the *fbp1* locus results in chromatin remodeling, which enables transcription factors to access the promoter DNA and activate *fbp1* expression (Hirota, Miyoshi, et al, 2008).

Recently, S. Bumgarner et al discovered that non-coding RNA transcripts were transcribed at the *FLO11* promoter in *S. cerevisiae*: the sense transcript, *ICR1*, which is transcribed from the ~ -3 kb region and terminates close to the *FLO11* ORF ATG start site, and the antisense transcript *PWR1,* which is transcribed from the ~ -2.1 kb region and terminates near the -3 kb region (Bumgarner, Dowell, et al, 2009). At ~ 3.5 kb, the *FLO11* promoter is one of the largest in *S. cerevisiae* and its regulation is complex. *FLO11* is epigenetically regulated (Halme, Bumgarner, et al, 2004); under nutrient-limited conditions, its expression is variegated; the variegation results from the gene switching between a silenced and active state once every several cell generations (Halme, Bumgarner, et al, 2004). Many factors, including Sfl1p, Flo8p, Tec1p, Ste12p, Phd1p, Msn1p and Mss1p, have been shown to regulate *FLO11* expression (Borneman, Leigh-Bell, et al,

2006; Gagiano, van Dyk, et al, 1999; Pan, and Heitman, 2002; Rupp, Summers, et al, 1999). Moreover, ncRNA transcription at the promoter was also shown to regulate silencing and activation of *FLO11* (Bumgarner, Dowell, et al, 2009). Because of the promoter's large size and the multiple factors known to act upon it, we decided to further investigate the possible roles DNA looping might play in regulation of *FLO11*'s gene expression dynamics.

Our previous analysis of *FLO11*'s gene expression dynamics revealed that a 3-state model best described the promoter. In this model, the promoter switches slowly (< once per cell generation) between a silenced and competent state, and it switches at a fast rate (> once per cell generation) between a competent and active state (Octavio, Gedeon, and Maheshri, 2009). Silencing of *FLO11* requires the repressor Sfl1p (Halme, Bumgarner, et al, 2004). When the *FLO11* promoter is silenced, addition of the activators Tec1p, Ste12p, and Phd1p are unable to challenge the silenced state, so expression of *FLO11* is low (Octavio, Gedeon, and Maheshri, 2009). The activator Flo8p, which binds to the same promoter region as Sfl1p, can effectively destabilize the silenced state, and activation by Flo8p leads to a fairly graded response (Octavio, Gedeon, and Maheshri, 2009). Meanwhile, the activator Msn1p can stabilize the active state at *FLO11* without destabilizing the silenced state, effectively yielding a bimodal response (Octavio, Gedeon, and Maheshri, 2009). Msn1p's mechanism of activation is poorly understood, although it has been shown to act at longer distances to destabilize chromatin (Lorenz, and Heitman, 1998).

By correlating the dynamics of the *FLO11* promoter with its underlying molecular phenotype, we show that the stable silenced and stable active states at *FLO11* are associated with different looped conformations of the promoter. Furthermore, the silenced and active looped conformations are dependent on *FLO11*-specific trans regulators and ncRNA transcription at the

FLO11 promoter. Our results demonstrate that DNA looping at the *FLO11* promoter plays an important role in stabilizing the silenced and active states, giving rise to the overall slow promoter dynamics and the resulting variegation in *FLO11* expression.

MATERIALS AND METHODS

Yeast strains and media

All yeast strains were derived from the haploid Σ1278b from the Heitman laboratory (Lorenz, and Heitman, 1997), in which the *FLO11* ORF was replaced with the *YFP-KanMX6* or *CFP-KanMX6* cassettes by PCR integration. The haploid strains were subsequently mated to create diploids. All strains and plasmids used are provided in Supplemental Tables 1 and 2.

SC refers to synthetic defined media with 2% glucose and complete amino acid supplement. SD refers to synthetic defined media with 2% glucose and no amino acid supplements, and SC ura- refers to synthetic defined media with 2% glucose and amino acid supplement lacking uracil. Unless otherwise noted, cells were grown in SC, SD, or SC ura- media for all experiments.

Activator titration and promoter response analysis

Single-cell measurement and analysis of YFP and CFP fluorescence in individual cells were performed as described in (Octavio, Gedeon, and Maheshri, 2009). Analysis of the expression distributions to determine promoter kinetics is also described in (Octavio, Gedeon, and Maheshri, 2009).

As explained in (Octavio, Gedeon, and Maheshri, 2009), the resulting YFP or CFP expression distribution upon titration of an activator can reveal whether the underlying promoter state is silenced or desilenced. For *FLO11*, when the promoter is silenced, titration of a Class I activator (Tec1p, Ste12p or Phd1p) will yield no response and titration of a Class II activator (Msn1p and

Mss11p) will yield a bimodal response. When the promoter is desilenced, titration of either Class I or II activator will yield a graded response.

Chromatin conformation capture assay

Chromatin conformation capture (3C) assays were done based on the method of Singh, Ansari, and Hampsey 2009. Briefly, ~ 40 ml SC or SC ura- were inoculated with overnight cell cultures to initial OD_{600} = 0.005 to 0.1, and then grown to between OD_{600} = 0.8 and 1. Chromatin isolated from fixed cells was digested at 37°C with the restriction enzyme AluI for 5 hours. Complete digestion was confirmed by PCR of digested template (Figure S1). After digestion, restriction enzymes were heat deactivated, and the digested chromatin was diluted and ligated with Quick Ligase for 2 hours at room temperature. After reversal of crosslinks by incubation in 65°C overnight, DNA was isolated and analyzed by PCR using divergent primer pairs designed to yield a PCR signal only when successful ligation of corresponding *FLO11* promoter fragments occurred (Figure S1). Between 200 and 300 ng of template DNA was used in each PCR reaction. PCR signal using a primer pair amplifying the ~1.5kb region of the promoter, which is not digested by the restriction enzyme AluI, was used as loading control.

Chromatin Immunoprecipitation assay

Chromatin IP's to measure histone acetylation state were performed as in (Octavio, Gedeon, and Maheshri, 2009), except for the following modifications. Cells for harvesting were grown exactly as cells grown for 3C assay described above. Isolated chromatin was incubated with antibodies overnight at 4°C, and Dynal Protein G magnetic beads (Invitrogen Cat. # 10007D) were used instead of Sepharose beads. ChIP signals reported correspond to the ratio of anti-acetylated histone / anti-histone signal at the *FLO11* promoter region normalized by the ratio of the anti-acetylated histone/ anti-histone signal at a region in telomere VI-R (Yang, and

Kirchmaier, 2006) in each IP sample. Therefore, the resulting signal is equivalent to the "average histone acetylation at the promoter region divided by the average acetylation at a telomeric region."

To measure occupancy of HA-tagged regulators bound to the *FLO11* promoter, Anti-HA High Affinity antibodies were used (Roche Cat. # 11 867 423 001). Fold enrichment signals were calculated as the ratio of the signal at the promoter / signal at an unbound region(telomere VI-R) in the input vs. the signal at the promoter / signal at an unbound region in the IP. Control signals correspond to a fold enrichment signal calculated using chromatin isolated from a strain with no HA-tagged alleles incubated with anti-HA, and serves as a background signal.

RT-PCR assay

To measure *ICR1* transcript levels, total RNA was isolated from cells grown as in 3C and ChIP assays using RNeasy Mini Kit (Qiagen Cat. # 74104). Reverse transcriptase quantitative PCR was performed using OneStep RT-PCR Kit (Qiagen Cat. # 210212), with primers specific to the -800 bp region of the *FLO11* promoter and an internal control primer specific to the housekeeping non-coding RNA *SCR1* to normalize signals. To measure *PWR1* transcript levels, reverse transcriptase quantitative PCR was performed in two steps: first, total RNA isolated from cells was transcribed into cDNA using Transcriptor First Strand cDNA Synthesis Kit (Roche Cat. # 04 379 012 001) and primers specific to the *PWR1* strand and the housekeeping non-coding RNA *SCR1*. The cDNA was then used as template in quantitative PCR reactions performed as in (Octavio, Gedeon, and Maheshri, 2009) using primers specific to the -2500 bp region of the *FLO11* promoter. Signal was normalized to *SCR1*.

RESULTS

The -1.2 kb and -0.2 kb promoter regions associate to form a DNA loop in the silenced state.

We performed a chromatin conformation capture (3C) assay (Singh, Ansari, and Hampsey, 2009) to probe for *in vivo* associations between the core promoter region -- ~200 bp upstream of ATG of the *YFP/CFP* open reading frame (ORF) -- and upstream promoter regions in the wild-type (WT) dual-reporter strain grown in synthetic media (SC) with 2% glucose, where the promoter is silenced (Octavio, Gedeon, and Maheshri, 2009). Our 3C strategy is depicted in Figure 1A. As controls, we monitored extent of digestion of cross-linked chromatin to ensure complete digestion of the DNA prior to ligation (Figure S1). We also checked that every primer pair (P1-A, P1-B, etc.) used in the 3C analysis did not yield any PCR products prior to ligation of the DNA (Figure S1). Therefore, any PCR product detected after ligation of the DNA should result only from successful amplification of ligation products resulting from the DNA fragments in close proximity during ligation in the dilute solution, indicating *in vivo* association between these DNA regions. Our 3C analysis revealed that in this silenced background, the +328 bp to -328 bp of ATG of the *YFP/CFP* ORF and the -859 bp to -1380 bp regions associate *in vivo* to form a loop (Figure 1B). These regions correspond to the core *FLO11* promoter region and the ~ -1.2kb region, respectively.

To determine whether the loop between the core promoter region and the -1.2 kb region was correlated with the silenced promoter state, we performed 3C in mutant backgrounds where the promoter was desilenced. Previously, Sfl1p, Hda1p and Set1p were shown to be necessary for silencing at *FLO11* (Bumgarner, Dowell, et al, 2009; Halme, Bumgarner, et al, 2004). In the desilenced *sfl1Δ*, *hda1Δ* and *set1Δ* backgrounds, 3C results indicated that the loop between the -

1.2 kb region and core promoter region was abolished and no *in vivo* associations occurred between the core promoter region and further upstream regions (Figures 1C, 1D and 1E).

The *FLO11* core promoter and ncRNA *PWR1* promoter and terminator regions associate to form a second DNA loop in the Msn1p-activated state.

We then performed 3C on our diploid dual-reporter strain with the *FLO11* activator Msn1p overexpressed, where *FLO11* is expressed highly (Octavio, Gedeon, and Maheshri, 2009). Our 3C analysis revealed *in vivo* association between the core promoter, a region spanning the -2.2 to -2.4 kb region encompassing the *PWR1* start site, and a region spanning the -3.1 to -3.3 kb region coinciding with a potential *PWR1* terminator (Bumgarner, Dowell, et al, 2009) (Figure 2B). We will refer to this loop conformation as the "active loop", to distinguish from the "silenced loop" observed between ~ -1.2 kb and the core promoter in the silenced state.

Because loops formed between a yeast gene's promoter and terminator regions are dependent on TFIIB (Singh, and Hampsey, 2007) and enable faster re-induction of transcription (Laine, Singh, et al, 2009), we overexpressed Msn1p in a WT background with the *sua7-1* allele expressed on a plasmid. The *sua7-1* allele encodes TFIIB with an E62K mutation that shown to impair looping but not mRNA levels at several *S. cerevisiae* genes (Pinto, Ware, and Hampsey, 1992; Pinto, Wu, et al, 1994; Singh, and Hampsey, 2007). 3C results showed that the active loop was abolished in the background carrying the *sua7-1* allele (Figure 2C). To further confirm the occurrence of the active loop, we performed chromatin immunoprecipitation (ChIP) to probe for TFIIB occupancy at the promoter in the WT and WT + *sua7-1* backgrounds. When Msn1p was overexpressed, TFIIB was present at the *FLO11* and *PWR1* core promoters in both WT and WT + *sua7-1* backgrounds (Figure 2D). However, the presence of TFIIB at the *PWR1* terminator was greatly diminished in the WT + *sua7-1* background where the active loop is abolished (Figure

2D). Finally, we overexpressed HA-tagged Msn1p in a WT background and assayed Msn1p localization by ChIP. Msn1p localized [to the *FLO11* core promoter and the -1.7 kb to -2.1 kb region encompassing the *PWR1* promoter] at the -1.7 kb to -2.1 kb region, corresponding to the *PWR1* promoter, and weakly localized at the -2.8 kb and -0.2 kb regions, which correspond to the *PWR1* terminator and *FLO11* core promoter respectively (Figure 2E). This localization profile suggests that Msn1p might be directly, rather than indirectly, involved in active loop formation and/or maintenance at the promoter. Together, these results are consistent with the Msn1p-mediated active loop forming via TFIIB at the core promoter and at the *PWR1* promoter interacting with *PWR1* termination machinery.

The experiments with the strains carrying the *sua7-1* allele on a plasmid will be repeated in a background where *sua7-1* is integrated in the native *SUA7* locus; please refer to the Appendix for more details on strain construction and experiments to be performed.

Noncoding RNAs *ICR1* and *PWR1* are required for silenced and active loop formation respectively.

Since transcription of two non-coding RNA species at the promoter, *ICR1* and *PWR1*, is known to correlate with the silenced and active states of *FLO11,* respectively (Bumgarner, Dowell, et al, 2009), we asked whether ncRNA transcription might be important in forming the loops observed in the silenced and active states. Transcription of the ncRNA *ICR1* is required for silencing of *FLO11* in the haploid Σ1278b strain background (Bumgarner, Dowell, et al, 2009). To investigate possible roles of *ICR1* transcription in loop formation at the *FLO11* promoter in our dual reporter diploid strain, we abolished *ICR1* transcription by inserting a terminator construct ~ 3 kb upstream of the ATG start site of the *YFP/CFP* ORF, which was confirmed by RT-PCR (Figure 5F). We will refer to this strain background as "*icr1-term*". We then performed 3C on

icr1-term to probe for *in vivo* associations between the core *FLO11* promoter region and upstream regions. We found no evidence for *in vivo* associations between the core promoter and far upstream regions, including the silenced loop formed between the -1.2 kb and core promoter regions (compare Figure 3A with Figure 1A). The results indicate that along with Sfl1, Hda1p, and Set1p (Figures 1C, 1D and 1E), *ICR1* transcription is necessary for forming and/or maintaining the silenced loop.

Next, because *PWR1* transcription is correlated with active *FLO11* expression in the haploid Σ1278b strain background (Bumgarner, Dowell, et al, 2009), we investigated its possible role in forming the loop observed in the Msn1p-activated *FLO11* promoter. RT-PCR analysis showed that *PWR1* transcription was elevated upon overexpression of Msn1p in our dual-reporter diploid (Figure 3B). To determine whether transcription of *PWR1* was required to form the active loop, we deleted the -2100 to -2350 bp region of the promoter corresponding to *PWR1* transcription start sites (Bumgarner, Dowell, et al, 2009) and confirmed that *PWR1* transcription was abolished by RT-PCR (Figure S3). We will refer to this strain background as "*pwr1Δ*." When we overexpressed Msn1p in *pwr1Δ*, 3C revealed that the active loop was abolished (Figure 3C), with only a very weak association between the *FLO11* core and *PWR1* promoter remaining, indicating that *PWR1* transcription was required for forming and/or maintaining the active loop at *FLO11*.

In silenced looped state, Sfl1p and *ICR1* transcription recruit HDACs to the promoter, which keeps the promoter fully hypoacetylated.

Our 3C results showed that the stable silenced state of the promoter was highly correlated with the presence of a loop between the core promoter and the ~ -1.2 kb region. In a previous study, we suggested that Sfl1p directly bound the core promoter because of its ability to function as a conventional repressor in the absence of silencing (Octavio, Gedeon, and Maheshri, 2009) and

the presence of a Sfl1p binding site in the core promoter (Conlan, and Tzamarias, 2001). Furthermore, Sfl1p is thought to recruit Sfl1p at *FLO11* (Halme, Bumgarner, et al, 2004), and the silenced state appeared hypoacetylated in at least the -1.7 kb region (Octavio, Gedeon, and Maheshri, 2009). Therefore, we hypothesized that this loop generated a stable, silenced state by localizing histone deacetylases at the promoter and maintaining a repressive chromatin environment. To test this, we looked at the chromatin state and the occupancy of not only Sfl1p and Hda1p at the promoter in silenced (looped) and desilenced (no loop) backgrounds, but also *ICR1* and the histone methyltransferase Set1p, as all four are necessary for silencing at *FLO11* (Bumgarner, Dowell, et al, 2009; Halme, Bumgarner, et al, 2004; Bumgarner, 2008) and silenced loop formation (Figures 1, 3A).

We measured the acetylation state of H3 and H4 histones across the *FLO11* promoter in the WT silenced strain using ChIP and found histones across the entire promoter region were hypoacetylated (Figure 4A). Next, we performed ChIP for HA-tagged Sfl1p and Hda1p in the WT silenced background revealed that Sfl1p localized to the ~ -1.2kb region and the ~ -0.2 kb region, consistent with the loop formed between these two regions (Figure 4B). The Hda1p signal also peaked at the -1.2 kb and -0.2 kb regions, colocalizing with Sfl1p; but in contrast to Sfl1p, it was localized significantly over the control across the entire promoter (Figure 4C). To test whether Sfl1p was necessary for Hda1p's localization at the *FLO11* promoter, we performed ChIP for HA-tagged Hda1p in *sfl1Δ* and found that Hda1p no longer occupied the promoter in this background (compare Figure 4D with Figure 4C). In both *sfl1Δ* and *hda1Δ* strains where the promoter is desilenced, ChIP results revealed that H3 and H4 histones were hyperacetylated (compare Figures 4E and 4F with Figure 4A), indicating that Sfl1p and Hda1p were necessary for hypoacetylation of the entire *FLO11* promoter region.

Next, we looked at *icr1-term*. We first confirmed that the promoter was desilenced using our activator titration and promoter response assay (see Methods), in our diploid *icr1-term* background. Titration of either Tec1p (a Class I activator of *FLO11*) (Figures 5A and 5B) or Msn1p (a Class II activator) (Figure S2) in these backgrounds yields a graded response, indicating that the underlying promoter state was desilenced.

We then focused on defining causal relationships between *ICR1* transcription and Sfl1p and Hda1p activity. One reason *ICR1* transcription may be necessary for silencing would be if it were required to recruit Sfl1p and/or Hda1p to the promoter. To test this idea, we performed ChIP for HA-tagged Sfl1p and HA-tagged Hda1p in *icr1-term* and found Sfl1p and Hda1p still localized to the -1.2 kb and -0.2 kb regions, but Hda1p no longer occupied the rest of the promoter region as in WT (compare Figure 5D with Figure 4B, and Figure 5E with Figure 4C). This result indicated that *ICR1* transcription was not required for Sfl1p and Hda1p to localize to the promoter; however, *ICR1* transcription allowed Hda1p to localize across the entire region. We next tested whether Sfl1p and/or Hda1p is required for *ICR1* transcription by quantifying *ICR1* transcript levels relative to the housekeeping ncRNA *SCR1* in the WT, *sfl1Δ* and *hda1Δ* backgrounds (Figure 5F). *ICR1* transcript levels were only reduced in the *sfl1Δ* background, indicating that *ICR1* transcription was independent of Hda1p but promoted by Sfl1p.

Because Hda1p localization is restricted in *icr1-term*, hypoacetylation might also be restricted. Indeed, ChIP analysis of the acetylation state of the *FLO11* promoter in the *icr1-term* background revealed that H3 and H4 histones across the region were primarily hyperacetylated, with local regions of deacetylation at the -1.2 kb and -0.2 kb regions where Sfl1p and Hda1p localized (compare Figure 5G with Figures 5D and 5E). Therefore *ICR1* transcription is required for hypoacetylation of both H3 and H4 throughout the *FLO11* promoter.

Either the process of transcription and/or the presence of the *ICR1* ncRNA acting in *cis* could be the necessary intermediate that leads to hypoacetylation, the diffuse Hda1p localization across the entire promoter, and silenced loop formation. Because Hda1p is restricted to H3 (Carmen, Rundlett, and Grunstein, 1996), at least one additional histone deacetylase with activity for H4 must be recruited to the *FLO11* promoter. We hypothesized that *ICR1* transcription ultimately recruits such a deacetylase via Set1p and Set3p. Set1p is a histone methyltransferase which as a member of the COMPASS complex can mono-, di- or trimethylate H3K4 (Briggs, Bryk, et al, 2001; Miller, Krogan, et al, 2001; Nagy, Griesenbeck, et al, 2002; Roguev, Schaft, et al, 2001), and has been shown to be recruited to transcribed regions via its interaction with the C-terminal domain of RNA Pol II (Ng, Robert, et al, 2003). Set1p is required for silenced loop formation at the *FLO11* promoter (Figure 1E). Meanwhile, Set3p has been shown to recognize the H3K4 dimethylation mark deposited by Set1p (Kim, and Buratowski, 2009) and is a member of a complex that contains a number of histone deacetylases with activity for H4, including Hos2p (Pijnappel, Schaft, et al, 2001).

To test the hypothesis that *ICR1* transcription contributes to silencing of the promoter by recruitment of additional histone deacetylases via Set1p and Set3p, first we checked the stability of the silenced state in the *set3Δ* mutant background, as we would expect this background to be desilenced like *icr1-term* and *set1Δ*. We verified the graded response upon titration of Tec1p (Class I activator) or Msn1p (Class II activator) in *set1Δ* (Figure 5B and S2) and found the identical response in *set3Δ* (Figure 5C and S2), indicating that the *FLO11* promoter is desilenced. Next, using ChIP we found the acetylation profile of the promoter in the *set1Δ* background mirrored the profile in the *icr1-term* background (compare Figure 5H with Figure 5G), with the -1.2 kb and -0.2 kb regions deacetylated locally. RT-PCR analysis showed that

ICR1 transcript levels in *set1Δ* were similar to WT levels (Figure 5F). Together, these data suggested that Set1p catalyzed events downstream of *ICR1* transcription that is required for silencing. Finally, to test whether deacetylation of the entire promoter required recruitment of the Set3 complex to the *FLO11* promoter, we examined a strain where wild-type *SET3* was replaced with the *set3W140A* allele. The W140A mutation in Set3p has been shown to abrogate Set3p's ability to recognize the H3K4 dimethylation mark (Shi, Kachirskaia, et al, 2007). Therefore, if Set3p was recruited to the *FLO11* promoter only via H3K4 dimethylation marks deposited by Set1p, Set3p^{W140A} should not be able to localize to the *FLO11* promoter and deacetylation would be restricted to the -1.2 and -0.2 kb regions. ChIP analysis in the *set3W140A* strain revealed an acetylation profile that mirrored the *icr1-term* and *set1Δ* backgrounds, confirming this prediction (Figure 5I).

Taken together, these results indicate that Sfl1p binding to the -1.2 and -0.2 kb regions mediates Hda1p localization to these regions, and *ICR1* transcription enables Hda1p localization across the entire promoter which is necessary for a stably silenced, hypoacetylated promoter state. In addition, the silencing function of *ICR1* requires Set1p, and the ability of Set3p to bind H3K4Me2 marks associated with Set1p activity.

Active loop enables Msn1p to stabilize the active state and generate a bimodal *FLO11* expression profile.

Our 3C results indicated that the association between the *FLO11* promoter and the promoter and terminator regions of the ncRNA *PWR1* (the active loop) was highly correlated with the Msn1p-activated state. However, this active loop conformation was not required for full activation of *FLO11* expression, as evidenced by the 3C results in *sfl1Δ* where no loops were observed (Figure 1C); *sfl1Δ* is a desilenced background where the *FLO11* promoter is highly active in all cells in

the population (Halme, Bumgarner, et al, 2004; Octavio, Gedeon, and Maheshri, 2009). In addition, no loops were observed in in WT + *sua7-1* or *pwr1Δ* (Figures 2C and 3C) even when Msn1p was expressed at extremely high levels (10000 ng/mL doxycycline) that resulted in complete activation of *FLO11* (Figure 6 and Figure S5). Therefore, we decided to investigate whether the active loop formation might instead play a role in the particular mechanism by which Msn1p regulates the slow dynamics of *FLO11* activation that results in bimodal gene expression at steady state.

In previous work, we found that different activators of *FLO11* affected different combinations of the fast and slow promoter switching rates, resulting into either graded or bimodal expression profiles upon activation of *FLO11* (Octavio, Gedeon, and Maheshri, 2009). In particular, addition of intermediate levels of Msn1p resulted in a bimodal steady-state expression profile. This implied Msn1p was able to stabilize the active state without destabilizing the silenced state. Since the active loop was not necessary for *FLO11* activation, we hypothesized it was instead required to stabilize the active state in conditions where the silenced state remained a stable alternative. To test this, we titrated Msn1p in the *pwr1Δ* and WT + *sua7-1* backgrounds, where active looping at the *FLO11* promoter no longer occurred, and looked at the promoter response. The bimodal response we observed when titrating Msn1p in a WT background was abolished in both these mutant backgrounds (Figure 6 and Figure S5). At intermediate Msn1p levels, the promoter ON-OFF rates are about 7 to 10-fold faster (see fit rates in Figures S6 and S7) in the mutant backgrounds with the active loop abolished than in the WT background, confirming that the active loop was necessary for Msn1p's ability to stabilize the active state at *FLO11* without destabilizing the silenced state.

The activator titrations in the *sua7-1* background will be repeated in a background where *sua7-1* is integrated in the native *SUA7* locus; please refer to the Appendix for more details on strain construction and experiments to be performed.

DISCUSSION

Our study reveals that the molecular signature of the two epigenetic states of *FLO11* observed to coexist within a population (Halme, Bumgarner, et al, 2004; Octavio, Gedeon, and Maheshri, 2009) are two (mutually exclusive) looped conformations of the *FLO11* promoter. Transitions between the two states require multiple transformations, leading to their overall stability. This silenced state is characterized by *ICR1* transcription, hypoacetylation of H3 and H4, localization of Hda1p across the entire promoter, and high Sfl1p occupancy at the -0.2 and -1.2 kb region containing previously identified Sfl1p binding sites (Pan, and Heitman, 2002), and the DNA loop formed between these regions. Sfl1p has been shown to form multimers *in vivo* (Conlan, and Tzamarias, 2001) and the DNA loop may in part be formed or stabilized by multimerized Sfl1p bridging its two binding sites. The Msn1p-dependent active state is characterized by *PWR1* transcription, hyperacetylation of H3 and H4, localization of Msn1p across the promoter, the absence of Sfl1p binding, and an active DNA loop formed between the *PWR1* and *FLO11* core promoter and the *PWR1* terminator. Msn1p occupancy peaks at the *PWR1* and *FLO11* core promoter, and likely does so also at the *PWR1* terminator located near the *ICR1* promoter, although we did not probe in this region.

Our analysis of various mutants which are not silenced, but are still repressed, indicate a probable transition state between the two stable epigenetic states and the molecular events involved in switching. In the absence of *ICR1* transcription (an in conditions where Sfl1p is

active), Sfl1p binds to both the -0.2 and -1.2 regions and recruits Hda1p, possibly via the Tup1/Ssn6 corepressor transcription (Conlan, and Tzamarias, 2001; Davie, Edmondson, et al, 2003). This desilenced, repressed intermediate state is the transition state where Hda1p localization and hypoacetylation is restricted to where Sfl1p can bind, DNA loops are not present, and Class I activators are capable of activating *FLO11* expression. We propose transition to the silenced state begins with *ICR1* transcription, which recruits the Set1p histone methyltransferase, the catalytic member of the COMPASS complex recruited by RNA Polymerase II during (Ng, Robert, et al, 2003). Set1p-mediate deposition of tri, di, and mono, methylation marks on histones in the transcribed region allows Set3p binding to H3K4Me2 as part of the Set3C complex that includes the histone deacetylases Hos2 and Hst1 (Briggs, Bryk, et al, 2001; Miller, Krogan, et al, 2001; Nagy, Griesenbeck, et al, 2002; Roguev, Schaft, et al, 2001). Because Hda1p activity is restricted to H3 and Hos2 and Hst1 have known activity on H4 (Wu, Suka, et al, 2001), Set3C is a strong candidate for the observed H4 hypoacetylation (Pijnappel, Schaft, et al, 2001), although other HDACS associated with *FLO11* such as Rpd3p (Bumgarner, Dowell, et al, 2009) may play some role. Still, Hda1p is required for the silenced state, so the HDAC activity of Set3C is not sufficient for full hypoacetylation of the promoter and/or loop formation. Hypoacetylation is required for the stable formation of the silenced DNA loop, through an unknown mechanism. Once formed, the DNA loop could further stabilize the silenced state by localizing repressive factors at the core promoter.

Activating a silenced *FLO11* promoter either requires destabilizing the silenced state or stabilizing the active state, the choice of which dictates the level of heterogeneity in the population response. Activating *FLO11* by destabilizing the silenced state and activating transcription at the desilenced promoter results in a *graded* response, where the cell population

exhibits a unimodal expression profile. With the exception of *sfl1Δ*, *FLO11* expression is repressed in our desilenced promoter mutants, representing the transition state where Sfl1p is still bound to the core promoter and conventionally represses expression. Addition of activators in this background increases expression levels in a graded manner. Because of their global importance, it is unlikely cells would abolish Hda1p or Set1p activity to desilence the promoter and activate expression in a graded fashion in this manner. More likely, cells achieve graded activation by displacement of Sfl1p binding at the -1.2kb region, which would abolish the association between this region and the core promoter. In support of this, we have shown previously that titration of the activator Flo8p, which also binds to the -1.2 kb region (Pan, and Heitman, 2002), destabilized the silenced state and stabilized the active state at the promoter in a *coupled* manner, thereby yielding a fairly graded response that can be further tuned by modulating other Class I activators (Octavio, Gedeon, and Maheshri, 2009). Moreover, when we prevented Sfl1p binding to the -1.2 kb region by having tetR bind to a tetO site inserted at this region, titration of activators in this background yielded a graded response (Figure S4).

When the silenced state is stable relative to the transition state, the activators Msn1p and Mss11p encompass a second class of activators that stabilize the activate state *without* destabilizing the silenced state when present at intermediate levels (Octavio, Gedeon, and Maheshri, 2009). We envision the silenced *FLO11* promoter dynamically interchanging with the transition state where the loop comes apart and histone deacetylases can no longer maintain hypoacetylation. A round of *ICR1* transcription can reset the chromatin state and reestablish the loop. However, intermediate levels of Msn1p can drive promoter to the active state. We propose this begins with Msn1p-mediated *PWR1* transcription antagonizing *ICR1* transcription and providing a platform on which to form the active loop. Msn1p must also play a role in displacing

Sfl1p in the -0.2 kb region and perhaps from the -1.2 kb region as well. While Msn1p has recently been shown to directly bind to DNA (Fordyce, Gerber, et al, 2010), and we cannot exclude the possibility that Msn1p's role is indirect in any or all of these steps. Nevertheless, Msn1p is present within the active loop.

The stability of this epigenetic switch is predicated on the ability to form DNA loops (Figure 4, 6). We suspect both loops serve to exclude transcription factor DNA binding, acetylation status, and ncRNA transcription associated with the other state, while promoting its own state. The silent loop prevents binding of activators, and transcription of *PWR1*, while bringing histone deacetylases in close proximity promoter and possibly promoting their "spreading" across the entire promoter. In the active loop, we envision an activator-mediated "scaffold" at the juxtaposition of the *PWR1* terminator and the *PWR1* and *FLO11* core promoter that (i) excludes Sfl1p binding to the -0.2 kb (and probably the -1.2 kb) region, (ii) promoters *PWR1* transcription and (iii) recruits histone acetylases that can function across the promoter. Such a scaffold has been suggested to explain why TFIIB-mediated loops at other yeast genes can accelerate the kinetics of reinitiation (Hampsey, Singh, et al, 2010; Laine, Singh, et al, 2009; Tan-Wong, Wijayatilake, and Proudfoot, 2009). We envision an activator-mediated "scaffold" at the juxtaposition of the *PWR1* terminator and the *PWR1* and *FLO11* core promoter that excludes Sfl1p binding to the -0.2 kb region, and probably even the -1.2 kb region.

Multiple properties of the *FLO11* epigenetic switch we describe hear are likely employed in a variety of organisms. First, the use of multiple DNA loops in a single genomic region to encode stable epigenetic states has been observed at the b1 locus in maize (Louwers, Bader, et al, 2009) and the Igf2 cluster (Murrell, Heeson, and Reik, 2004). Second, TFIIB-mediated DNA looping is present at many yeast genes and the "illicit" looping we observe between a coding and non-

coding promoter may be common because of the widespread cryptic transcription yeast (Neil, Malabat, et al, 2009; Xu, Wei, et al, 2009). This illicit interaction requires factors yet to be identified. Third, our proposed mechanism of ncRNA-mediated recruitment of SET3C histone deacetylases via Set1p-mediated H3K4me2 builds upon previous work demonstrating the ncRNA-mediate recruitment of RPD3S /Ead3 via Set2p-mediated H3K36me3 (Houseley, Rubbi, et al, 2008). What determines whether a particular pathway is employed remains to be uncovered (at *FLO11*, the Set2p pathway does not function (Bumgarner, Dowell, et al, 2009)). While we favor the view that repression in both cases is mediated solely by the transcriptional mark deposited and not due to DNA or protein interactions with the ncRNA, direct interactions remain possible. Recently the lincRNA HOTTIP transcribed from the 5' end of the *Hox* cluster in mammals was reported to interact with WDR5, which through recruitment of MLL family proteins leads to H3K4 methylation (Wang, Yang, et al, 2011) .

Finally our work demonstrates molecular mechanisms of how a complex promoter might integrate different *trans* signals to modulate slow and fast promoter dynamics encoded in *cis* and ultimately, levels of phenotypic heterogeneity. While the details may differ, events similar to those at *FLO11* (looping, ncRNA transcription, recruitment of histone-modifying enzymes etc.) are common in epigenetically-regulated developmental genes in eukaryotes. A better integrated study of these events is needed to decipher their intricate interplay and ultimate effect on the regulation (or mis-regulation) of phenotypic switching and cell fate decisions.

FIGURES AND TABLES

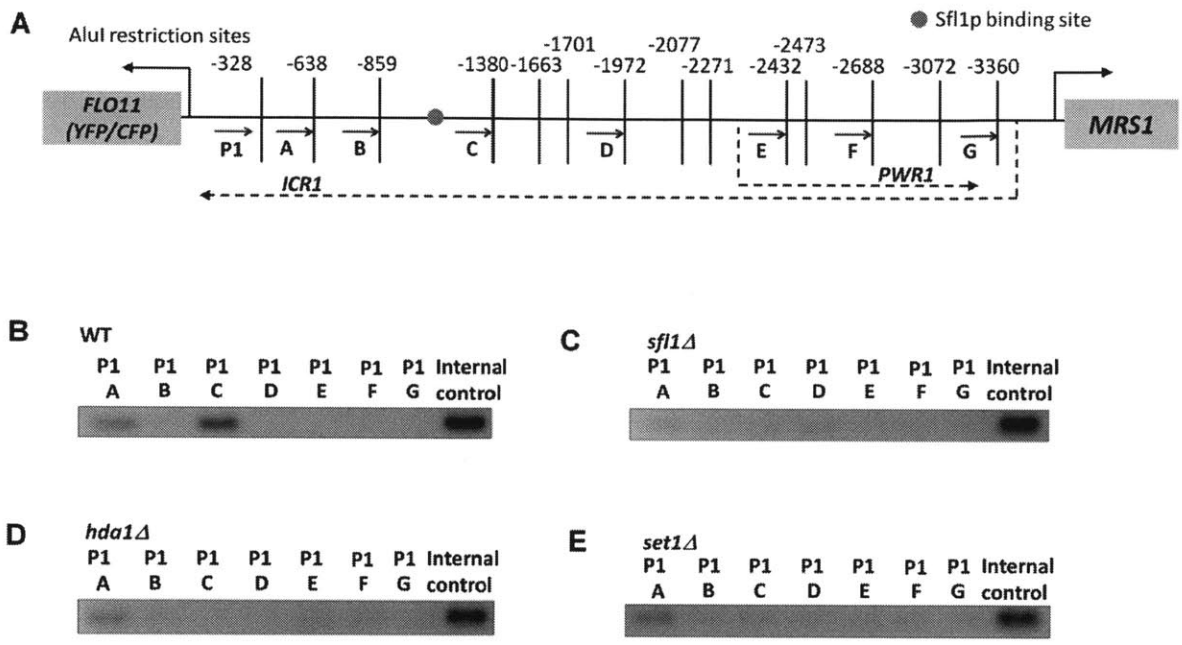

Figure 1: Silenced state of *FLO11* promoter is correlated with *in vivo* association between -1.2 kb and -0.2 kb regions. (A) *FLO11* promoter regions probed in chromatin conformation capture (3C) assays. In 3C assays, AluI was used to digest chromatin. AluI restriction sites across the ~3600 bp intergenic region between the *FLO11* (or *YFP/CFP*) ORF and the *MRS1* ORF are mapped; positions indicated are relative to the ATG start site of the *YFP/CFP* ORF. Divergent primer pairs were designed to probe for possible interactions of the -328 bp to + 328 bp region (primer 'P1') with each of the following regions: -328 bp to -638 bp (primer 'A'); -638 bp to -859 bp (primer 'B'); -859 bp to -1380 bp (primer 'C'); -1701 bp to -1972 bp (primer 'D'); -2271 bp to -2432 bp (primer 'E'); -2473 bp to -2688 bp (primer 'F'); and, -3072 bp to -3360 bp (primer 'G'). (B-E) 3C assay probing for *in vivo* associations between the *FLO11* core promoter region and upstream promoter regions in the various strains. PCR signal using primers probing the ~ -1500 bp undigested region serves as an internal loading control. A visible band from PCR with primers 'P1' and 'A' indicates association between the +328 bp to -328 bp region and the -328 bp to -638 bp region; because these regions are contiguous, the signal is likely from non-specific contacts between immediately neighboring nucleosomes. (B) In WT, PCR product observed from the reaction using primers 'P1' and 'C' indicates *in vivo* association between the core promoter region (+328 bp to -328 bp region) and the -859 bp to -1380 bp region, which are the regions Sfl1p is bound to. This association is lost and no other associations are seen in the (C) *sfl1Δ*, (D) *hda1Δ*, and (E) *set1Δ* backgrounds. All strains were grown in SC.

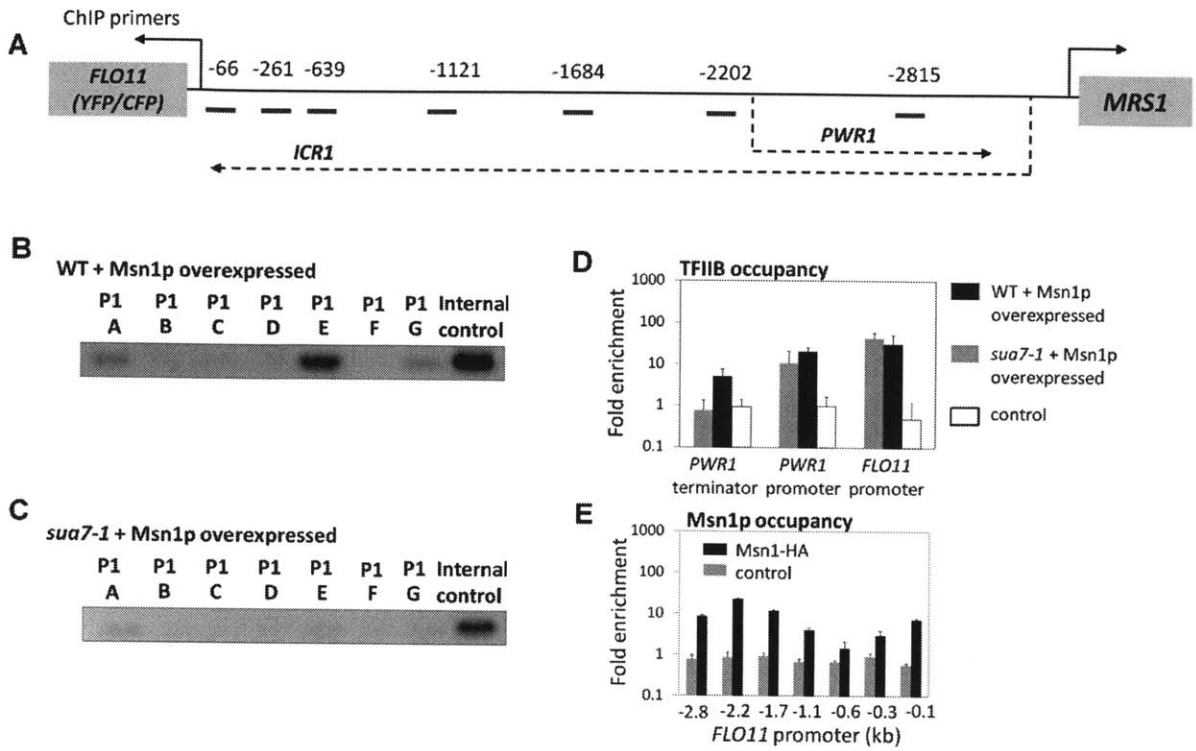

Figure 2: *FLO11* core promoter associates with *PWR1* promoter and *PWR1* terminator regions in Msn1p-activated *FLO11* promoter. (A) *FLO11* promoter regions probed in chromatin immunoprecipitation (ChIP) assays and in chromatin conformation capture (3C) assays. In ChIP assays, quantitative PCR was performed on isolated DNA from IP samples using primer pairs which amplified regions 170-240 bp in length, centered at 66, 261, 639, 1121, 1684, 2202, and 2815 bases upstream from the ATG site of the *YFP/CFP* ORF. Chromatin was sheared to fragments of ~500 bp average length. (B-C) 3C assay probing for *in vivo* associations between the *FLO11* core promoter region and upstream promoter regions in (B) the WT and (C) WT + *sua7-1* background with Msn1p overexpression. Strains were grown on SC ura- with 10000 ng/ml doxycycline to overexpress Msn1p. (B) Signals from the PCR reactions with primer pairs 'P1' + 'E' and 'P1' + 'G' indicated an association between the core promoter region (+328 bp to -328 bp) and the -2271 bp to -2432 bp region and association between the core promoter region and the -3072 bp to -3360 bp region in WT. These upstream regions correspond to the *PWR1* promoter and *PWR1* terminator regions, respectively. (C) These signals were greatly diminished in WT + *sua7-1* indicating that associations between the core promoter region and the *PWR1* promoter region and *PWR1* terminator region are abrogated. No association between the core promoter region and the -859 bp to -1380 bp (Sfl1p bound) region or any other region was observed for either background.(D) ChIP assay to measure TFIIB occupancy at the *FLO11* promoter in WT + Msn1p overexpressed and WT + *sua7-1* + Msn1p overexpressed

backgrounds. Signal (y-axis) represents fold enrichment in the IP over input after appropriate normalization (see Methods); error bars represent ± s.e.m. Control signal is from a WT strain carrying no HA-tagged alleles treated with anti-HA; this serves as a background signal. In the WT background with Msn1p overexpressed (10000 ng/ml doxycycline), signals greater than 10-fold over background at the *FLO11* core promoter region (~ -0.1 kb), the *PWR1* promoter (~ -2.2 kb) and the *PWR1* terminator (~ -2.8kb) indicated that TFIIB was bound to these regions. In the WT + *sua7-1* background with Msn1p overexpressed (10000 ng/ml doxycycline), the ChIP signal remained at least 10-fold over background at the *FLO11* core promoter region and the *PWR1* promoter, indicating that TFIIB still occupied these regions; however, the signal at the *PWR1* terminator was indistinguishable from background signal, indicating TFIIB no longer occupied the *PWR1* terminator. (E) ChIP assay to measure Msn1p occupancy at the *FLO11* promoter in WT + HA-tagged Msn1p overexpressed (10000 ng/ml doxycycline) background. As in (D), the control signal is fold enrichment from a WT strain with no HA-tagged alleles and serves as a background signal. Because no enrichment should occur, control signal values should be ~1. Signals at -2.8 kb, -2.4kb, -2.2kb, -1.7kb and -0.1 kb are at least several-fold over background, indicating Msn1p association with these promoter regions. Error bars represent ± s.e.m.

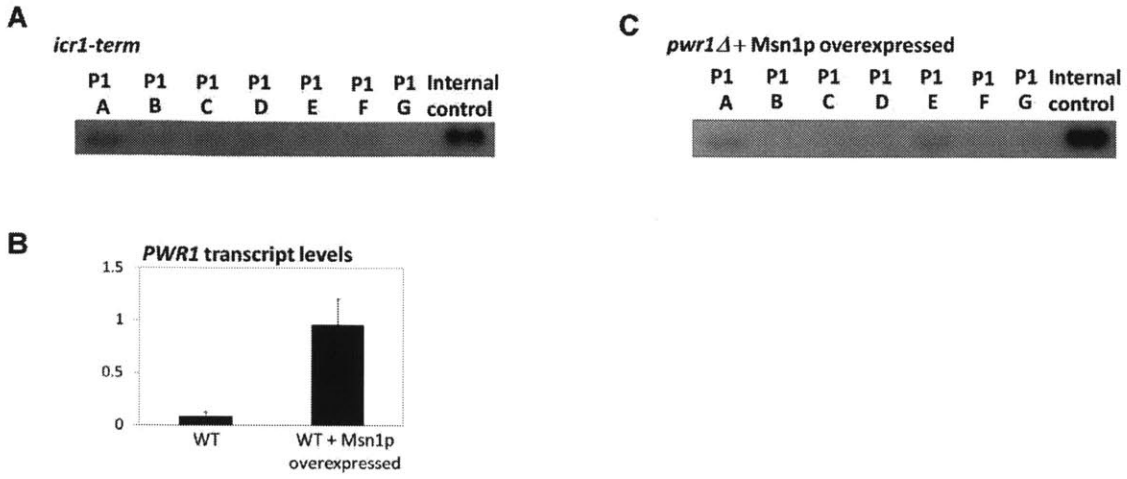

Figure 3: Looping in silenced and active promoter states is dependent on non-coding RNA *ICR1* and *PWR1* transcription respectively. (A) 3C assay probing for *in vivo* associations between the *FLO11* core promoter and upstream promoter regions in *icr1-term* grown in SC. No association between the core promoter region and any far upstream promoter region was observed. (B) RT-PCR assay to measure *PWR1* transcript level in several strain backgrounds grown in SC or SC ura-. Signal (y-axis) is *PWR1* transcript levels relative to internal control ncRNA *SCR1* transcript levels ± s.e.m. *PWR1* transcript levels were ~10-fold higher in the WT + Msn1p overexpressed (10000 ng/ml doxycyline) background than in WT. (C) 3C assay as in (A), but in the *pwr1Δ* + Msn1p overexpressed strain grown on SC ura- 10000 ng/ml doxycycline. Signals from the PCR reactions with primer pairs 'P1' + 'E' and 'P1' + 'G' were greatly diminished indicating that associations between the core promoter region and the *PWR1* promoter region and *PWR1* terminator region were abolished in the *pwr1Δ* background.

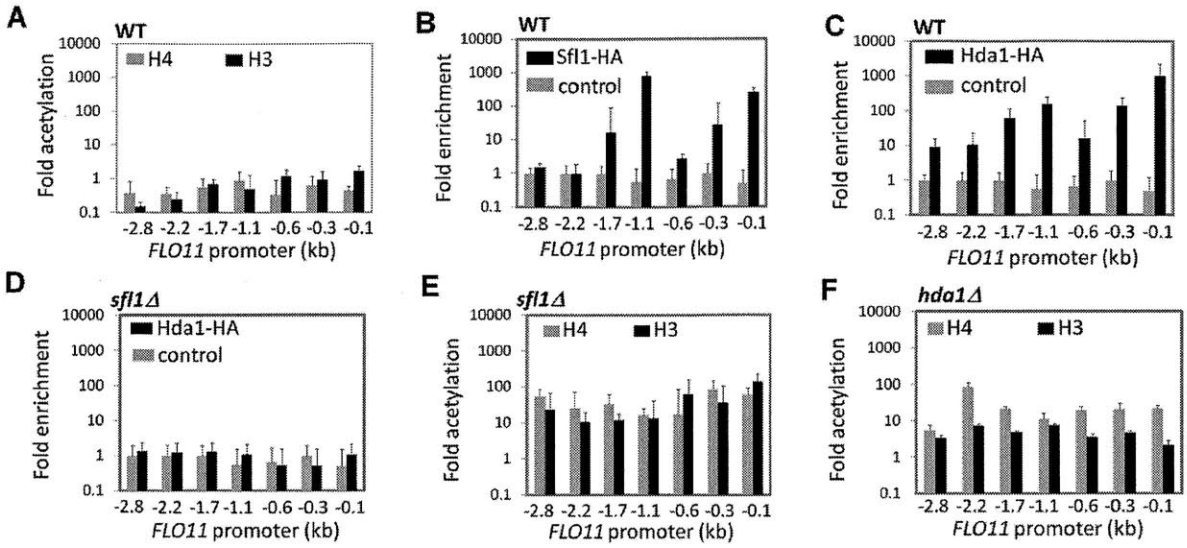

Figure 4: Hypoacetylation of histones at promoter in silenced state requires Sfl1p and Hda1p. A) ChIP assay probing for acetylated H3 and H4 histones across *FLO11* promoter region in WT two-color diploid grown in SC using primer pairs as in Figure 2A. Previous activator titration/ promoter response analysis revealed that the promoter is silenced in this background (Octavio, Gedeon, and Maheshri, 2009). Signal (y-axis) represents degree of acetylation at promoter relative to histone acetylation at a telomeric region (see Methods) $\pm$ s.e.m. Acetylation levels of H3 and H4 histones across the entire promoter were similar to acetylation at the telomeric region, indicating that the promoter region was hypoacetylated in this background.(B-D) ChIP assays probing for occupancy of Sfl1p and Hda1p in various strains grown in SC. Sample and control signals (y-axis) represent fold enrichment + s.e.m. as described in Figure 2D. Control signals were always from the same strain treated with no HA antibody. (B) ChIP for Sfl1p gives signals a strong signal at ~ -1100 bp and ~ -1700 bp upstream of the ATG of the *YFP/CFP* ORF, consistent with Sfl1p binding to the -1150 to -1400 bp region of the promoter as reported by Pan and Heitman (Pan, and Heitman, 2002). Signals greater than 10-fold over background at -300 and -100 bp indicated that Sfl1p was also localized to the core *FLO11* promoter region. (C) ChIP for Hda1p in the WT (silenced promoter background) yielded signals at least several-fold over background in all regions probed, suggesting that Hda1p was spread across the entire promoter region. Signal peaks at the ~ -1100 bp and ~ 0 bp regions coincide with Sfl1p's localization to these promoter regions. (D) In an *sfl1Δ*, Hda1p was no longer localized to the promoter as ChIP signals for Hda1p were similar to background across the entire promoter. (E-F) H3 and H4 histone acetylation monitored by ChIP as in (A) but in an *sfl1Δ* (E) and (F) *hda1Δ*, both grown in SC. In both strain backgrounds, the promoter is desilenced (Octavio, Gedeon, and Maheshri, 2009). Acetylation of histones across the *FLO11* promoter was 10 to 100-fold higher in the *sfl1Δ* and at least several-fold higher in the *hda1Δ*, as compared to telomeric histone acetylation levels.

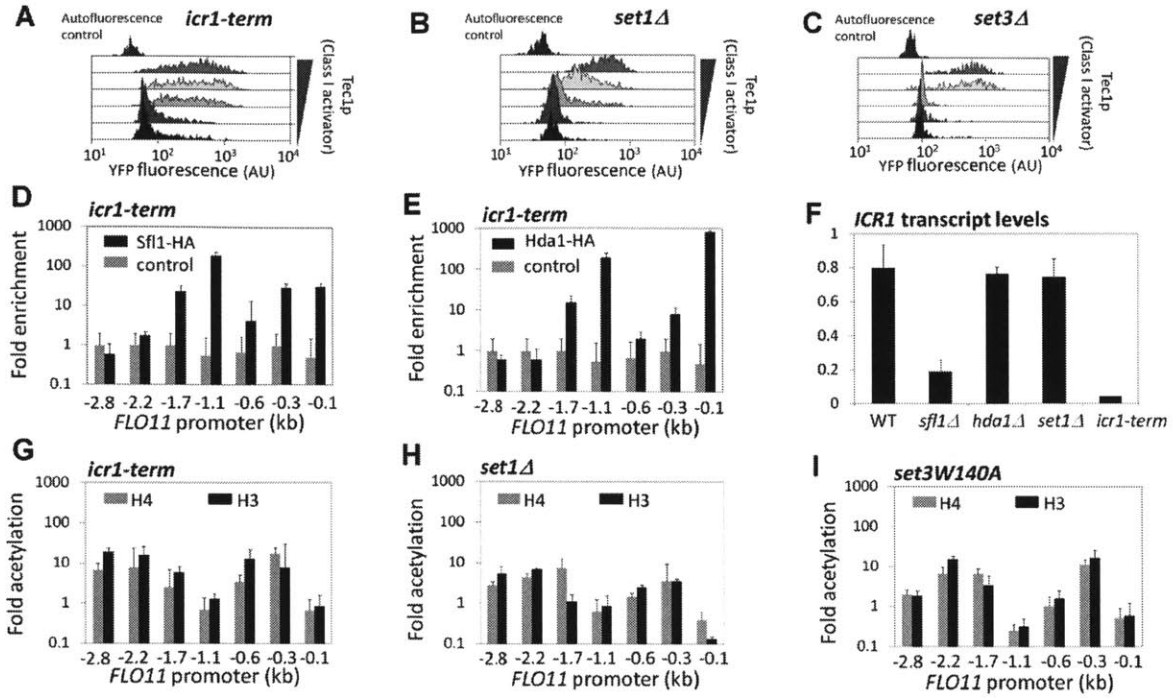

Figure 5: ICR1 transcription recruits additional histone deacetylases to the promoter via Set1p and Set3p. (A-C) Activator titration / promoter response analyses in the (A) *icr1-term*, (B) *set1Δ* and (C) *set3Δ* strain backgrounds, all grown in SC. Titration of Tec1p (a class I activator) (Octavio, Gedeon, and Maheshri, 2009) in all three backgrounds yielded a graded response, indicating that the underlying promoter state was desilenced. (D-E) (D) Sfl1p and (E) Hda1p occupancy at the promoter in *icr1-term* grown in SC. (D) Signals for Sfl1p at the -1700, -1100, -300 and -100 bp regions were at least ten-fold over background, indicating that Sfl1p was localized to the same promoter regions in the *icr1-term* background as in WT.(E) Only signals at the -1700, -1100 and -100 bp regions were at least ten-fold over background for Hda1p, indicating that it was no longer spread across the entire promoter region but localized only to regions where Sfl1p was present. F) RT-PCR analysis measures *ICR1* transcript levels in several strain backgrounds. RT-PCR analysis was performed using primers probing the region ~ 800 bp upstream of the ATG of the *YFP/CFP* ORF and primers for the housekeeping non-coding RNA *SCR1* as internal control. *ICR1* transcript levels are similar in the WT, *hda1Δ* and *set1Δ* backgrounds; in *sfl1Δ*, *ICR1* transcript level is about a quarter of the WT level. (G-H) H3 and H4 histone acetylation levels across *FLO11* promoter in (G) *icr1-term*, (H) *set1Δ*, and (I) *set3W140A*, all grown in SC. For all backgrounds, histones at all regions probed, except at the -1100 bp and -100 bp regions, are at least several fold higher than telomere acetylation levels, whereas the -1100 bp and -100 bp regions have about acetylation levels similar to or less than the telomeric region.

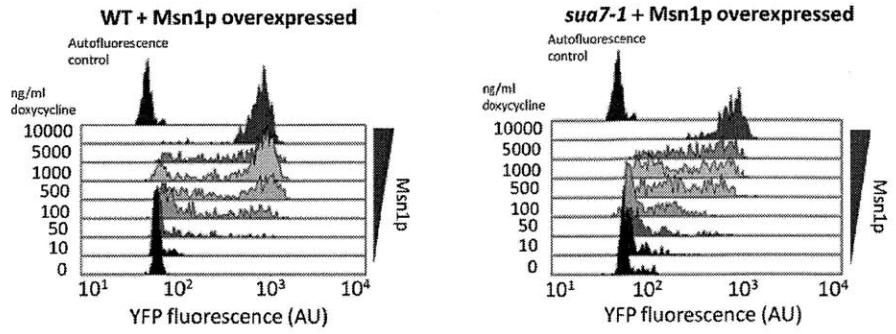

Figure 6: Msn1p-dependent active loop is required for Msn1p's ability to stabilize active promoter state without destabilizing silenced state and yield bimodal response. Msn1p titration in WT and WT + *sua7-1* backgrounds. Msn1p was titrated (0 to 10000 ng/ml doxycycline) in WT with empty pRS315 vector and WT carrying the *sua7-1* allele in a pRS315 vector. Titration of Msn1p in both strain backgrounds was carried out so that each strain was grown in media with identical doxycycline concentrations, and fluorescence microscopy to measure expression distributions for all samples were performed on the same day. At intermediate concentrations of Msn1p, where bimodal expression was observed in the WT background (left), bimodal expression was not observed in the WT + *sua7-1* at the same intermediate concentrations of Msn1p (right).

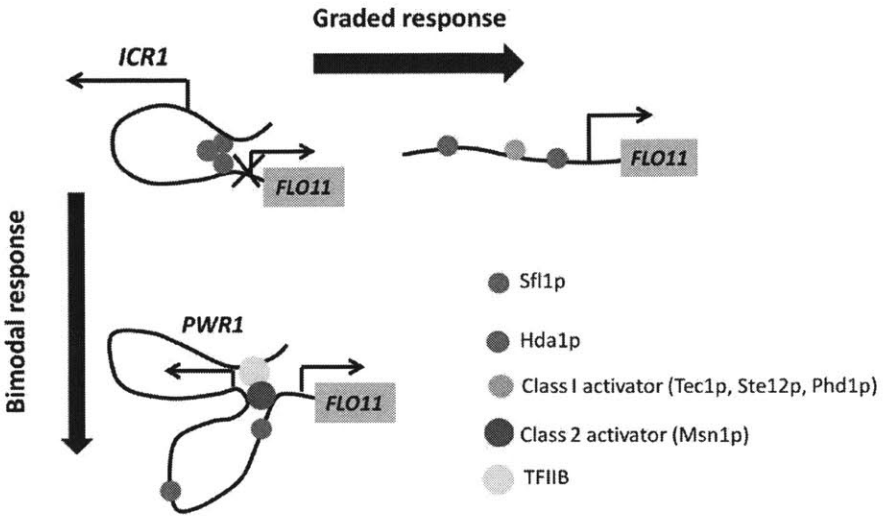

Figure 7: Model for cis-encoded bimodal gene expression at *FLO11* encoded by Sfl1p-mediated repressor looping and Msn1p-mediated activator looping The silenced *FLO11* promoter state is stabilized by several events. First, *ICR1* transcription deposits histone methylation marks via Set1p, which then recruits additional histone deacetylases to the promoter via Set3p. Sfl1p recruits Hda1p to the core promoter and -1.2 kb region, and *ICR1* transcription enables effective spreading of Hda1p across the entire promoter region. Finally, the silenced promoter is also looped between the core promoter and -1.2 kb regions. When the promoter is desilenced by inhibition of any of the steps that lead to silencing (for example, deletion of *HDA1* or *ICR1* transcription), activation of the promoter will yield a graded response. If the promoter is silenced, increasing levels of the activator Msn1p generates a bimodal response. Msn1p stabilizes the active state by forming a loop between the *FLO11* core promoter region and the *PWR1* promoter and terminator regions, which requires active *PWR1* transcription and TFIIB-mediated interactions. Bimodal gene expression is achieved at *FLO11* as switching between the silenced promoter state and Msn1-dependent active states at *FLO11* is slow (i.e. occurs on a timescale of once per several generations). Although individual reactions at the promoter might occur on a timescale faster than once per cell generation, the overall transition rate between the silenced and Msn1-dependent active state is slow because many individual stochastic steps are required to transition

SUPPLEMENTAL TABLE 1: Yeast strains used in study

Strains constructed from Σ1278 ML *ura3-52, leu2Δ::hisG* (Lorenz, and Heitman, 1997)

Strain	***Genotype***	***Reference / Source***

MLY42	**Σ1278b** *MAT* α *ura3-52 leu2Δ::hisG*	Lorenz and Heitman, 1997
MLY43	**Σ1278b** *MAT* **a** *ura3-52 leu2Δ::hisG*	Lorenz and Heitman, 1997
Y35	*MAT* α *flo11Δ*::CFP-KanMX6 *ura3-52 leu2Δ::hisG* (GAL HO switched Y37)	Maheshri laboratory collection
Y36	*MAT* α *flo11Δ*::YFP-KanMX6 *ura3-52 leu2Δ::hisG* (GAL HO switched Y38)	Maheshri laboratory collection
Y37	*MAT* a *flo11Δ*::CFP-KanMX6 *ura3-52 leu2Δ::hisG*	Maheshri laboratory collection
Y38	*MAT* a *flo11Δ*::YFP-KanMX6 *ura3-52 leu2Δ::hisG*	Maheshri laboratory collection
Y45	*MAT* α/**a** *flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 ura3-52/ura3-52 leu2::hisG leu2Δ::hisG*	Maheshri laboratory collection
Y93	*MAT* α/a *flo11Δ::YFP-KanMX6/ flo11Δ::CFP-KanMX6 sfl1Δ::CgLEU2/sfl1 Δ :Cg:LEU2 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*	Maheshri laboratory collection
Y171	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 hda1Δ::CgLEU2/hda1Δ::CgLEU2 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*	Maheshri laboratory collection
LOY001	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 set1Δ ::CgLEU2/set1Δ :: CgLEU2 ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY004	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 set3Δ ::CgLEU2/set3Δ ::CgLEU2 ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY006	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -3000 bp FLO11 promoter :: TEF terminator/-3000bp FLO11 promoter ::TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY009	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -2100 to -2350 FLO11 promoter:: loxP/-2100 to -2350bp FLO11 promoter::loxP ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study

All strains constructed are in the Σ1278b background and cogenic to MLY43.

Strains with selected *FLO11* regulators knocked out were created by PCR integration to delete the regulator ORF in the YFP and CFP haploids and then mated to obtain diploids. Yeast were transformed by the PEG/Lithium Acetate method (Gietz, and Woods, 2002) and all integrations were verified through colony PCR. Strains with plasmids transformed were kept under selection media.

SUPPLEMENTAL TABLE 2: Plasmids used in study

Plasmid	***Genotype***	***Reference / Source***

pYC	*CEN URA3* ADH1 promoter-rtTA	Maheshri laboratory collection
pRS315	*CEN LEU2*	Sikorski and Hieter 1989
pRS316	*CEN URA3*	Sikorski and Hieter 1989
pYCTec1	7xtetO site (XhoI/BamHI), *TEC1* ORF (BamHI/NotI) in pYC	Maheshri laboratory collection
pYCMsn1	7xtetO site (XhoI/BamHI), *MSN1* ORF (BamHI/NotI) in pYC	Maheshri laboratory collection
pRS316Sfl1-HA	P_{SFL1}-*SFL1*-HA ORF (C-terminus HA tag) (ClaI/XhoI) in pRS316	This study
pRS316Hda1-HA	P_{HDA1}-*HDA1*-HA ORF (C-terminus HA tag) (ClaI/KpnI) in pRS316	This study
pYCMsn1-HA	7xtetO site (XhoI/BamHI), *MSN1*-HA ORF (C-terminus HA tag) (BamHI/NotI) in pYC	This study
pRS316Set3W140A	P_{SET3}- *set3W140A* ORF (XhoI/BamHI) in pRS316	This study
pRS315Sua7-1	P_{SUA7}-*sua7-1* ORF (ApaI/NotI) in pRS315	This study
pRS315Sua7-HA	P_{SUA7}-*SUA7*-HA ORF (C-terminus HA tag) (ClaI/XhoI) in pRS315	This study
pRS315Sua7-1-HA	P_{SUA7}-*sua7-1*-HA ORF (C-terminus HA tag) (ApaI/NotI) in pRS315	This study

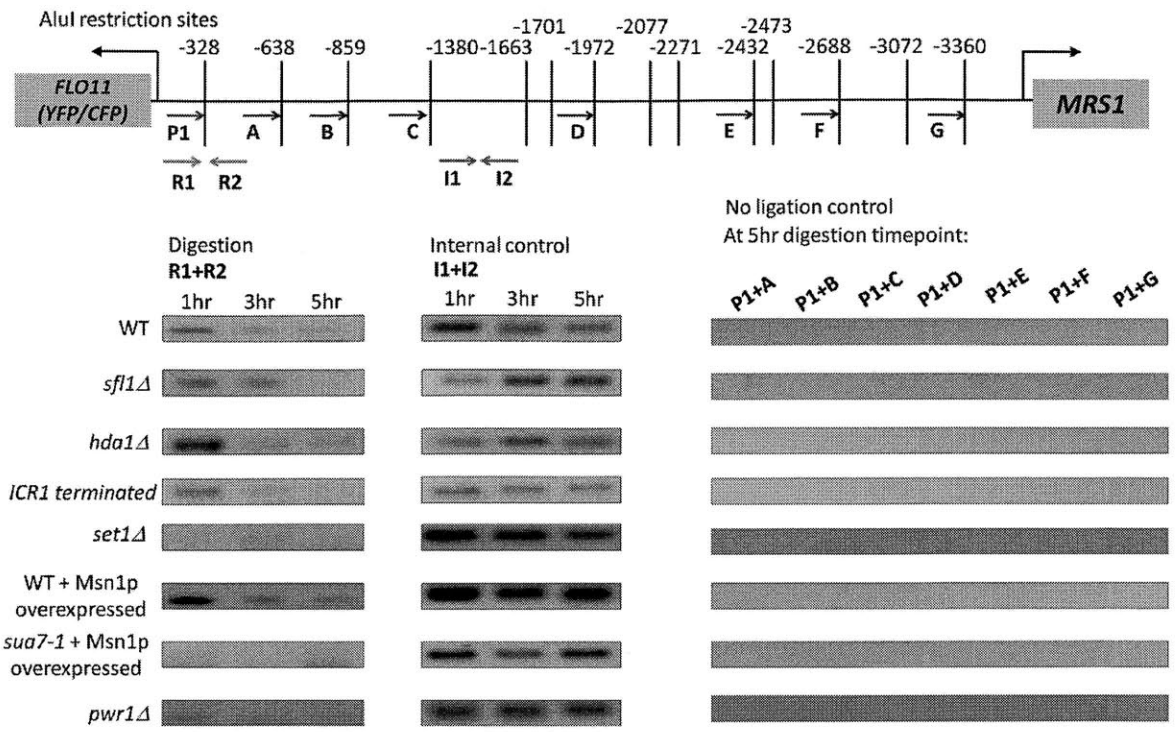

Figure S1: Chromatin conformation capture (3C) controls. At 1 hour, 3 hours, and 5 hours after digestion with AluI, ~0.5ul crude digested chromatin template was used in 50 ul PCR reactions using primers 'R1' + 'R2' to check for extent of digestion. These convergent primers amplify the region around -328 bp which is cleaved by AluI. Primers 'I1' + 'I2' amplify the region around -1.5 kb that is not digested by AluI and serve as internal loading controls. After 5 hours, chromatin in all samples reached complete digestion. PCR using the fully digested template was then performed using primer pairs 'P1' + 'A', 'P1' + 'B', 'P1' + 'C', 'P1' + 'D', 'P1' + 'E', 'P1'+ 'F' and 'P1' + 'G' which are divergent primers that should not amplify any product prior to addition of ligase to the digested chromatin and should amplify only promoter fragments that were successfully ligated in dilute solution , indicative of crosslinking between the fragments from *in vivo* association.

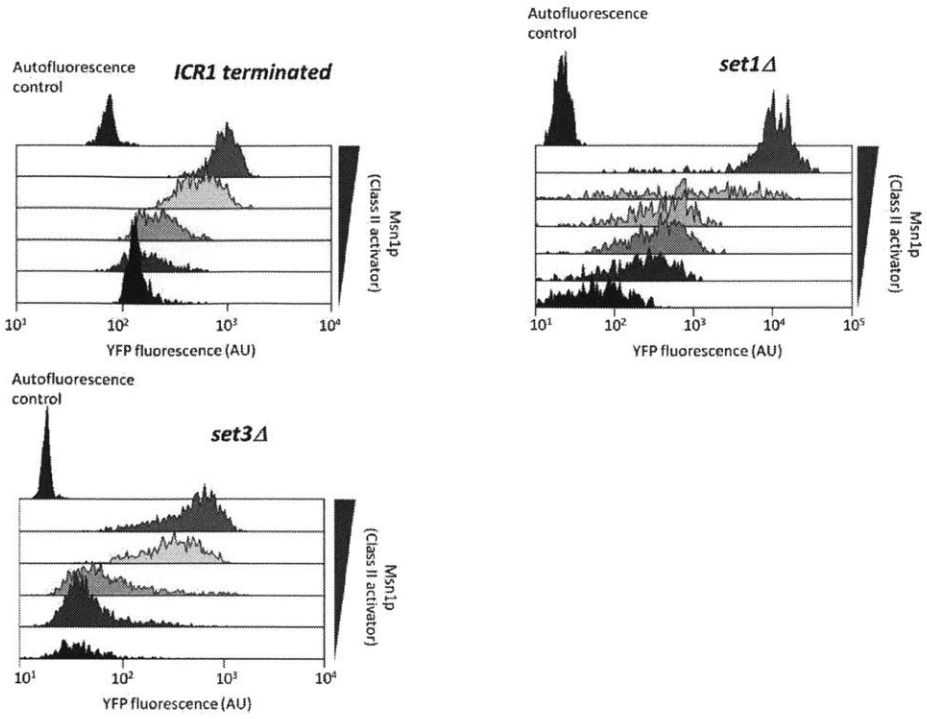

Figure S2: Titration of Msn1p (Class II activator) in *icr1-term*, *set1Δ*, and *set3Δ* backgrounds. To determine whether the *FLO11* promoter state was silenced or desilenced, titration of both Tec1p (Class I activator) (Figures 3A, 3B and 3C) and Msn1p (Class II activator) were performed. If the promoter was silenced, titration of a Class I activator would yield no response whereas titration of a Class II activator would yield a bimodal response. If the promoter was desilenced, then titration of either class of activator would generate a graded response. Like titration of the Class I activator Tec1p (Figures 3A, 3B and 3C), titration of the Class II activator Msn1p in the *icr1-term*, *set1Δ* and *set3Δ* backgrounds yields a graded response, indicating that the promoter is desilenced.

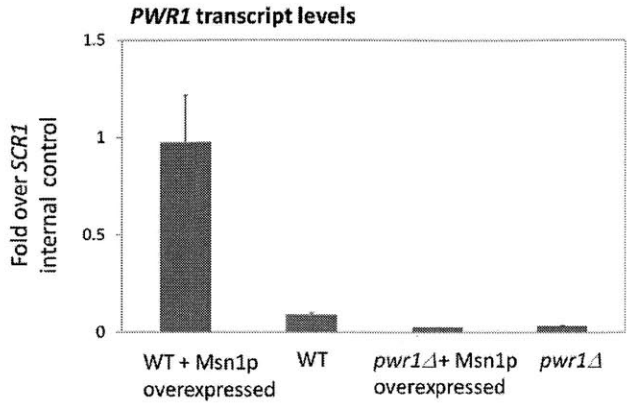

Figure S3: *PWR1* transcript levels in *pwr1Δ* with in WT background with endogenous and overexpressed levels of Msn1p. *PWR1* transcript levels in several strain backgrounds grown in SC or SC ura- were measured using RT-PCR (see Methods) to check *PWR1* transcript levels in *pwr1Δ* (strain background where the -2100 to -2350 bp region of the promoter corresponding to *PWR1* transcription start sites were deleted and replaced by a loxP sequence). Signal (y-axis) is *PWR1* transcript levels relative to internal control ncRNA *SCR1* transcript levels $\pm$ s.e.m. Because *PWR1* expression in WT with endogenous Msn1p levels is low and comparable to levels in *pwr1Δ* with endogenous Msn1p levels, we looked at both backgrounds with Msn1p overexpressed to the same level (10000 ng/ml doxycyline). In the Msn1p overexpressed backgrounds, *PWR1* transcript levels are greatly diminished in the *pwr1Δ* compared to WT, confirming that *PWR1* transcription is highly reduced in *pwr1Δ*

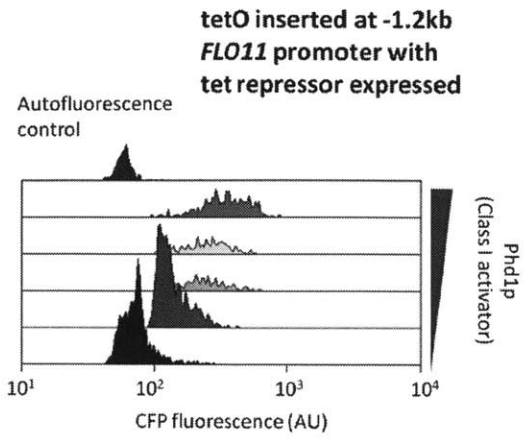

Figure S4: Titration of Phd1p (Class I activator) in background with tetO inserted at -1.2 kb of *FLO11* promoter and the tet repressor expressed. We expressed the tet repressor in the strain background with tetO inserted at -1.2 kb to occlude Sfl1p binding to this region. A graded response upon titration of a Class I activator (Phd1p) was observed, suggesting that preventing Sfl1p from binding to the -1.2 kb region is sufficient to desilence the promoter.

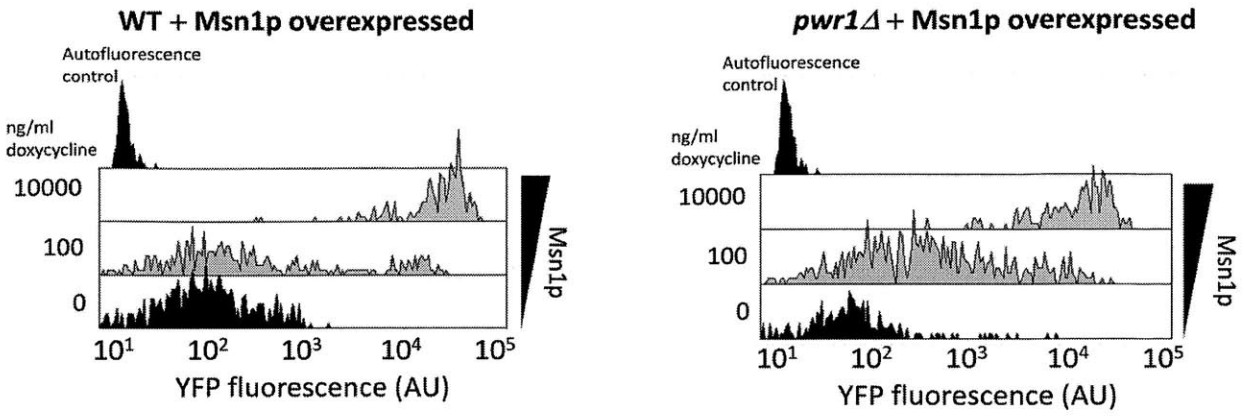

Figure S5: Msn1p titration in WT and *pwr1Δ* backgrounds. Msn1p was titrated (0 to 10000 ng/ml doxycycline) in the WT and *pwr1Δ* backgrounds. Titration of Msn1p in both strain backgrounds was carried out as described in Figure 6. At the intermediate concentration of Msn1p (100 ng/ml doxycycline), where bimodal expression is observed in the WT background (left), bimodal expression is not observed in the *pwr1Δ* background at the same intermediate concentration of Msn1p (right).

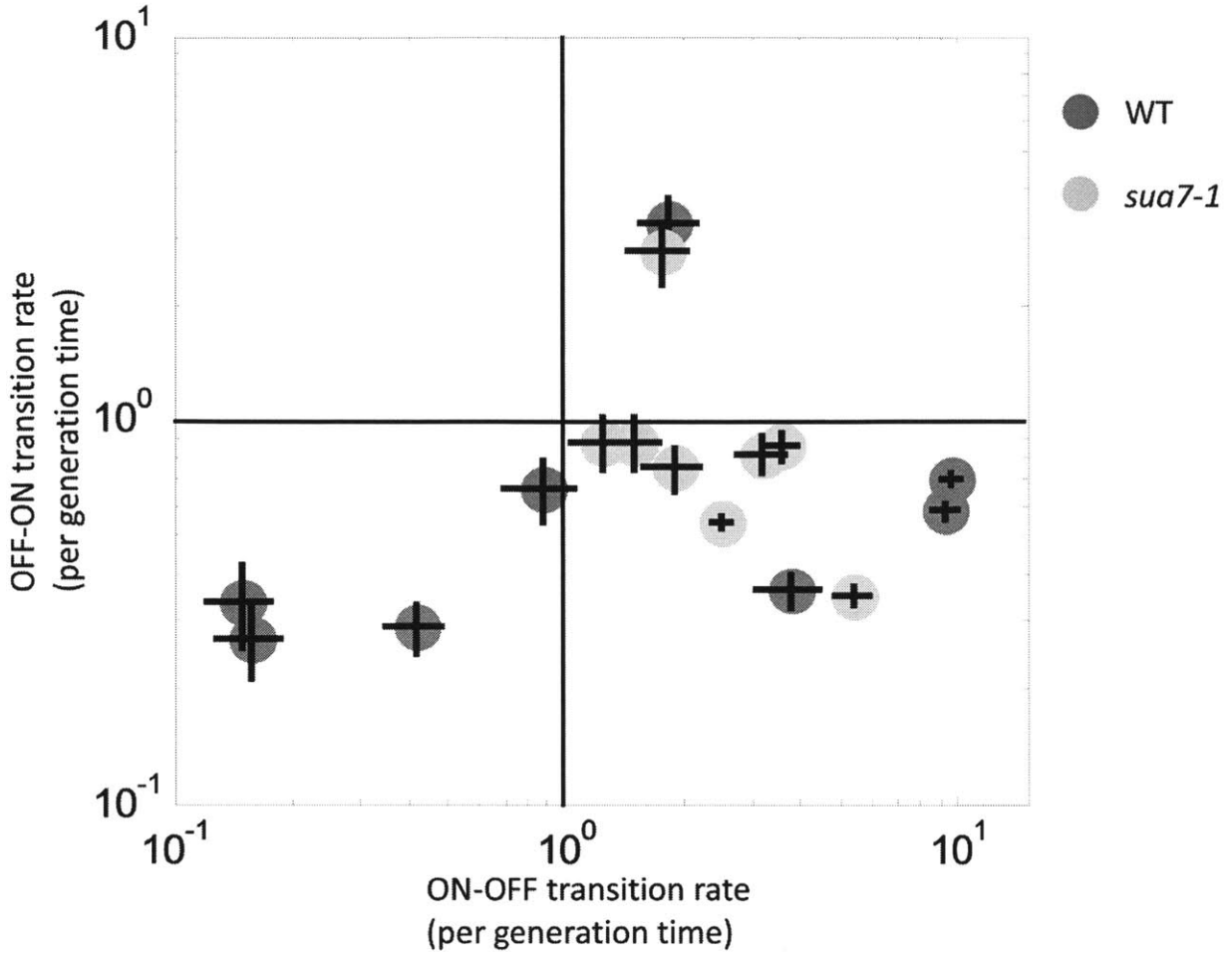

Figure S6: Promoter OFF-ON and ON-OFF rates fit to expression distributions of Msn1p titration in WT and *sua7-1* backgrounds in Figure 6. Error bars represent 95% confidence intervals to fits of experimental data to the Beta distribution for a single value of μ, the lumped mRNA/protein production rate - details in Supplemental Discussion of (Octavio, Gedeon, and Maheshri, 2009). At intermediate Msn1p levels in *sua7-1*, ON-OFF rates (green) are ~ 10-fold faster than at same intermediate Msn1p levels in WT (pink).

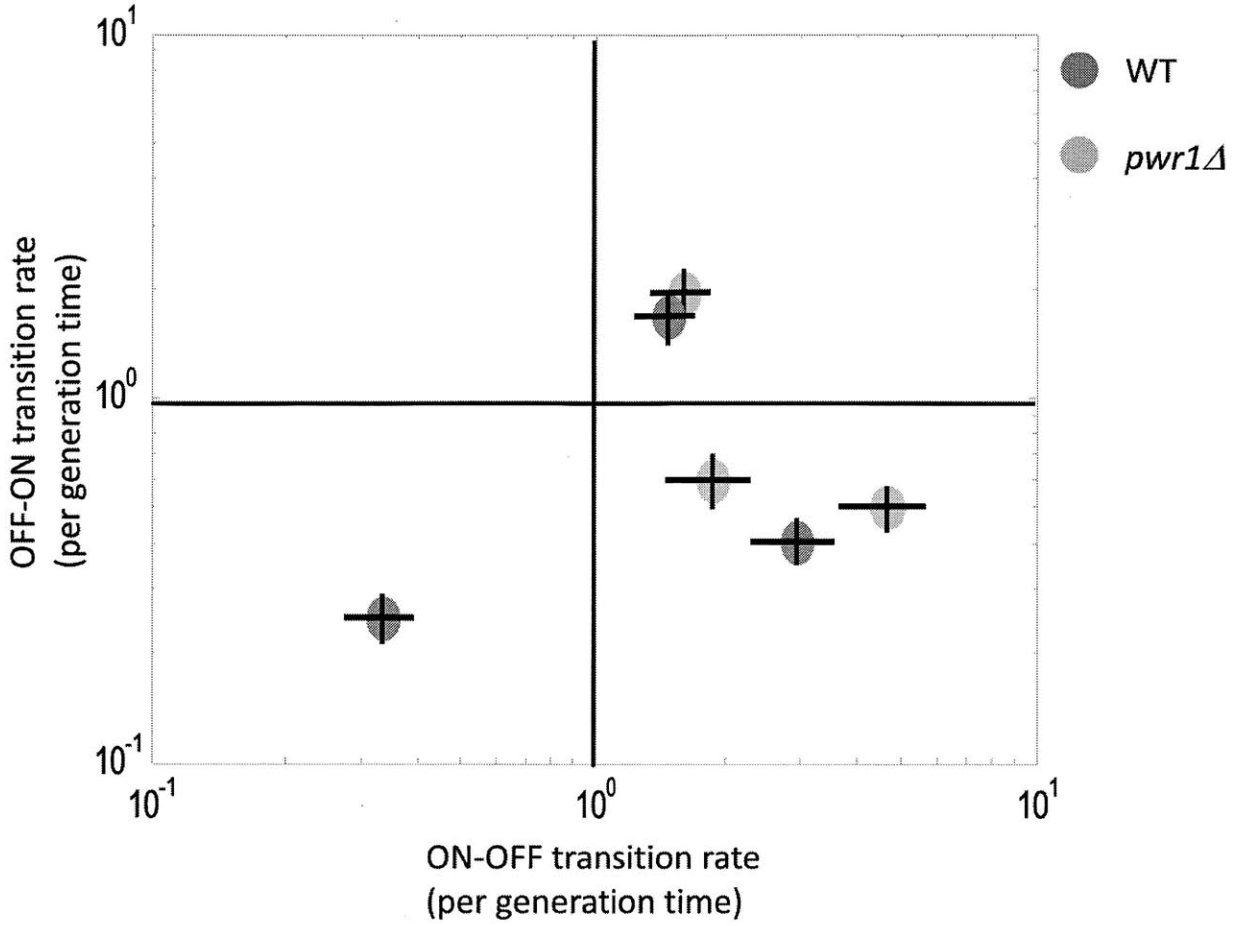

Figure S7: Promoter OFF-ON and ON-OFF rates fit to expression distributions of Msn1p titration in WT and *pwr1Δ* backgrounds in Figure S5. Error bars represent 95% confidence intervals to fits of experimental data to the Beta distribution for a single value of m, the lumped mRNA/protein production rate - details in Supplemental Discussion of (Octavio, Gedeon, and Maheshri, 2009) . At intermediate Msn1p level (100ng/ml doxycycline) in *pwr1Δ*, ON-OFF rates (green) are ~ 7-fold faster than at same intermediate Msn1p level in WT (pink).

REFERENCES

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M., and Snyder, M. (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *5705,* 2242-2246.

Borneman, A.R., Leigh-Bell, J.A., Yu, H., Bertone, P., Gerstein, M., and Snyder, M. (2006). Target hub proteins serve as master regulators of development in yeast. Genes Dev. *4,* 435-448.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev. *24,* 3286-3295.

Bumgarner, S.L., Dowell, R.D., Grisafi, P., Gifford, D.K., and Fink, G.R. (2009). Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast. Proc. Natl. Acad. Sci. U. S. A. *43,* 18321-18326.

Bumgarner, S.L. (2008). Mechanisms underlying cell-to-cell diversity in clonal populations of yeast (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., and Stutz, F. (2007). Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell *4,* 706-717.

Carmen, A.A., Rundlett, S.E., and Grunstein, M. (1996). HDA1 and HDA3 are components of a yeast histone deacetylase (HDA) complex. J. Biol. Chem. *26,* 15837-15844.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C. *et al.* (2005). The transcriptional landscape of the mammalian genome. Science *5740,* 1559-1563.

Conlan, R.S., and Tzamarias, D. (2001). Sfl1 functions via the co-repressor Ssn6-Tup1 and the cAMP-dependent protein kinase Tpk2. J. Mol. Biol. *5,* 1007-1015.

Davie, J.K., Edmondson, D.G., Coco, C.B., and Dent, S.Y. (2003). Tup1-Ssn6 interacts with multiple class I histone deacetylases in vivo. J. Biol. Chem. *50,* 50158-50162.

Di, L.J., Wang, L., Zhou, G.L., Wu, X.S., Guo, Z.C., Ke, X.S., Liu, D.P., and Liang, C.C. (2008). Identification of long range regulatory elements of mouse alpha-globin gene cluster by quantitative associated chromatin trap (QACT). J. Cell. Biochem. *1,* 301-312.

Dutrow, N., Nix, D.A., Holt, D., Milash, B., Dalley, B., Westbroek, E., Parnell, T.J., and Cairns, B.R. (2008). Dynamic transcriptome of Schizosaccharomyces pombe shown by RNA-DNA hybrid mapping. Nat. Genet. *8,* 977-986.

Fordyce, P.M., Gerber, D., Tran, D., Zheng, J., Li, H., DeRisi, J.L., and Quake, S.R. (2010). De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. Nat. Biotechnol. *9,* 970-975.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H. *et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *7269,* 58-64.

Gagiano, M., van Dyk, D., Bauer, F.F., Lambrechts, M.G., and Pretorius, I.S. (1999). Msn1p/Mss10p, Mss11p and Muc1p/Flo11p are part of a signal transduction pathway downstream of Mep2p regulating invasive growth and pseudohyphal differentiation in Saccharomyces cerevisiae. Mol. Microbiol. *1,* 103-116.

Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 87-96.

Halme, A., Bumgarner, S., Styles, C., and Fink, G.R. (2004). Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. Cell *3,* 405-415.

Hampsey, M., Singh, B.N., Ansari, A., Laine, J.P., and Krishnamurthy, S. (2010). Control of eukaryotic gene expression: Gene loops and transcriptional memory. Adv. Enzyme Regul.

Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C.S., Shibata, T., and Ohta, K. (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature *7218,* 130-134.

Hongay, C.F., Grisafi, P.L., Galitski, T., and Fink, G.R. (2006). Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell *4,* 735-745.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol. Cell *5,* 685-695.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S. *et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *7314,* 430-435.

Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P., and Gingeras, T.R. (2002). Large-scale transcriptional activity in chromosomes 21 and 22. Science *5569,* 916-919.

Kim, T., and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. Cell *2,* 259-272.

Laine, J.P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. Genes Dev. *22,* 2604-2609.

Lorenz, M.C., and Heitman, J. (1998). Regulators of pseudohyphal differentiation in Saccharomyces cerevisiae identified through multicopy suppressor analysis in ammonium permease mutant strains. Genetics *4,* 1443-1457.

Lorenz, M.C., and Heitman, J. (1997). Yeast pseudohyphal growth is regulated by GPA2, a G protein alpha homolog. EMBO J. *23,* 7008-7018.

Louwers, M., Bader, R., Haring, M., van Driel, R., de Laat, W., and Stam, M. (2009). Tissue- and expression level-specific chromatin looping at maize b1 epialleles. Plant Cell *3,* 832-842.

Martens, J.A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *6991,* 571-574.

Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl. Acad. Sci. U. S. A. *23,* 12902-12907.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat. Genet. *8,* 889-893.

Nagy, P.L., Griesenbeck, J., Kornberg, R.D., and Cleary, M.L. (2002). A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proc. Natl. Acad. Sci. U. S. A. *1,* 90-94.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *7232,* 1038-1042.

Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature *7095,* 840-846.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell *3,* 709-719.

Octavio, L.M., Gedeon, K., and Maheshri, N. (2009). Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. PLoS Genet. *10,* e1000673.

Pan, X., and Heitman, J. (2002). Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. Mol. Cell. Biol. *12,* 3981-3993.

Pijnappel, W.W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R., and Stewart, A.F. (2001). The S. cerevisiae SET3 complex includes

two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev. *22,* 2991-3004.

Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell *5,* 977-988.

Pinto, I., Wu, W.H., Na, J.G., and Hampsey, M. (1994). Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J. Biol. Chem. *48,* 30569-30573.

Rinn, J.L., Euskirchen, G., Bertone, P., Martone, R., Luscombe, N.M., Hartman, S., Harrison, P.M., Nelson, F.K., Miller, P., Gerstein, M., Weissman, S., and Snyder, M. (2003). The transcriptional activity of human Chromosome 22. Genes Dev. *4,* 529-540.

Rodley, C.D., Bertels, F., Jones, B., and O'Sullivan, J.M. (2009). Global identification of yeast chromosome interactions using Genome conformation capture. Fungal Genet. Biol. *11,* 879-886.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W.W., Wilm, M., Aasland, R., and Stewart, A.F. (2001). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. EMBO J. *24,* 7137-7148.

Rupp, S., Summers, E., Lo, H.J., Madhani, H., and Fink, G. (1999). MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast FLO11 gene. EMBO J. *5,* 1257-1269.

Sexton, T., Bantignies, F., and Cavalli, G. (2009). Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. Semin. Cell Dev. Biol. *7,* 849-855.

Shi, X., Kachirskaia, I., Walter, K.L., Kuo, J.H., Lake, A., Davrazou, F., Chan, S.M., Martin, D.G., Fingerman, I.M., Briggs, S.D. *et al.* (2007). Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. J. Biol. Chem. *4,* 2450-2455.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *1,* 19-27.

Singh, B.N., Ansari, A., and Hampsey, M. (2009). Detection of gene loops by 3C in yeast. Methods *4,* 361-367.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Mol. Cell *5,* 806-816.

Tan-Wong, S.M., Wijayatilake, H.D., and Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes Dev. *22,* 2610-2624.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A. *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *7341,* 120-124.

Wu, J., Suka, N., Carlson, M., and Grunstein, M. (2001). TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. Mol. Cell *1,* 117-126.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *7232,* 1033-1037.

Yang, B., and Kirchmaier, A.L. (2006). Bypassing the catalytic activity of SIR2 for SIR protein spreading in Saccharomyces cerevisiae. Mol. Biol. Cell *12,* 5287-5297.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U. *et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat. Genet. *11,* 1341-1347.

Chapter 4

TRANSVECTION IN YEAST

ABSTRACT

Transvection is a phenomenon in which an allele interacts with its corresponding allele on the homologous chromosome, which results in the regulation of one allele being transferred to the other allele. Typically, transvection is identified by complementation assays designed to check whether a mutant regulatory region driving a wild-type coding region of an allele, which results in a mutant phenotype, can complement a wild-type regulatory region driving expression of a mutant coding region, which also results in a mutant phenotype. Transvection occurs if the complementation interaction can rescue the wild-type phenotype in the heterozygous background. While transvection has been observed in Drosophila (Duncan, 2002; Muller, and Schaffner, 1990; Pirrotta, 1999), mammalian cells (Liu, Huang, et al, 2008a; Rassoulzadegan, Magliano, and Cuzin, 2002; Sandhu, Shi, et al, 2009), and the fungus Neurospora (Aramayo, and Metzenberg, 1996), it has never been found in Saccharomyces. Here preliminary experiments investigating *FLO11* expression profiles in heterozygous backgrounds revealed a possible trans interaction between the *FLO11* alleles, where the silenced wild type promoter copy is able to silence the mutant desilenced promoter on the homologous chromosome. In a homozygous diploid background, wild-type *FLO11* is partially silenced and exhibits a bimodal expression profile. Meanwhile, in a homozygous diploid background, a mutant *FLO11* allele with a *K. lactis* URA3-terminator insertion at the promoter is desilenced and exhibits a unimodal ON expression profile. When we looked at the expression profile of a wild-type *FLO11* x desilenced mutant *FLO11* heterozygous strain, we expected to observe a bimodal expression profile for the wild-type allele (indicative of partial silencing) and a unimodal ON expression profile for the

desilenced mutant allele. Instead, we observed that the desilenced mutant allele exhibited a *bimodal* expression profile, indicating that it had now become partially silenced in the heterozygous background. Further experiments using a different insertion construct to desilence the *FLO11* promoter revealed similar expression profiles in the heterozygous background, suggesting that the phenomenon observed was not construct-specific and may be an indication of transvection truly occurring at *FLO11*.

INTRODUCTION

The term "transvection" was first coined by Ed Lewis in 1954 to describe a phenomenon where pairing between two alleles allowed one allele to provide factors necessary for the function of the homologous allele (Lewis, 1954). Lewis observed that the phenotypes of mutant alleles of the *Ubx* genes in Drosophila were greatly enhanced when chromosomal pairing between the mutant and wild-type alleles was disrupted, which implied that expression of *Ubx* was increased by chromosomal pairing. Later, he also showed that the chromosomal pairing could allow different combinations of mutant alleles to complement each other; these mutations were usually combinations of mutations in the regulatory region of one allele and the coding region of the homologous allele.

While transvection was first thought to be a quirky phenomenon restricted only to a few genes in Drosophila, it was later found to be more commonplace than previously thought. In Drosophila, besides the *Ubx* genes, transvection between the homologous alleles of *yellow*, *white* and *brown* genes have also been reported (Pirrotta, 1999). Transvection between *yellow* alleles was shown to result in chromatin modifications that led to expression activation (Geyer, Green, and Corces, 1990; Morris, Chen, et al, 1998), whereas transvection between *brown* alleles led to silencing of

the homologous allele (Csink, and Henikoff, 1996; Dernburg, Broman, et al, 1996). Meanwhile, in some malignant human B cell lines (mantle cell lymphoma (MCL) and multiple myeloma (MM) cell lines), Liu et al (Liu, Huang, et al, 2008b) showed that translocation of regulatory elements was required for the DNA hypomethylation of the cyclin D1 gene promoter on the same chromosome, as well as trans-allelic hypomethylation of the cyclin D1 allele on the homologous chromosome. In the fungus *Neurospora crassa*, Aramayo and Meztzenberg (Aramayo, and Metzenberg, 1996) found that transvection between alleles of a key regulator of sexual development, Asm-1, was necessary in diplophase for the Asm-1 alleles to be properly regulated. Disruption of transvection between the alleles resulted in improper Asm-1 regulation, which led to generation of immature, inviable spores.

Transvection has been demonstrated in *Drosophila*, mammalian cells and *Neurospora* (Aramayo, and Metzenberg, 1996; Duncan, 2002; Liu, Huang, et al, 2008a; Muller, and Schaffner, 1990; Pirrotta, 1999; Rassoulzadegan, Magliano, and Cuzin, 2002; Sandhu, Shi, et al, 2009); it has never been shown to occur in *Saccharomyces*. Here we describe preliminary data suggesting the occurrence of transvection at the *FLO11* loci in *S. cerevisiae*. So far, data suggest that when it is in the silenced state, a wild-type partially silenced *FLO11* allele is able to silence a desilenced *FLO11* allele on the homologous chromosome. Future experiments will be needed to determine whether the trans-effect is dependent upon homologous chromosome pairing, and to further characterize the mechanism of transvection between the *FLO11* alleles.

MATERIALS AND METHODS

Yeast strains and media

We replaced the *FLO11* ORF in a haploid Σ1278b from the Heitman laboratory (Lorenz, and Heitman, 1997) with *YFP-KanMX6* or *CFP-KanMX6* cassettes by PCR integration. The haploid with *YFP-KanMX6* is mating type α and the haploid with *CFP-KanMX6* is mating type **a**. In each haploid background, we inserted either *K. lactis* URA3-TEF terminator or *C. glabrata* LEU2-TEF terminator cassettes by PCR integration at various locations within the *FLO11* promoter to create desilenced promoter strains. Combinations of these desilenced promoter and wild-type silenced promoter haploid strains were then mated to create heterozygous and homozygous diploids. All strains and plasmids used are listed in Supplemental Tables 1 and 2.

Single cell measurements and analysis

For microscopy, cells were either picked from 2-day old YPD plates and streaked directly onto a drop of water on a slide, or harvested from overnight liquid cultures which were first grown in YPD, SC ura- or YP 1% ethanol 2% glycerol, then inoculated at an initial OD_{600} between 0.005 and 0.01, and then grown for 15-20 hours in the same media. For the titration experiments, these cultures were treated with serial dilutions of doxycycline (0 to 10000 ng/ml) at 30°C in culture tubes placed in a rolling drum. Cells were harvested in mid-late log phase (OD_{600} between 0.5 and 1.5), and placed on ice while other samples were being processed. Expression was measured using a Zeiss AxioObserver microscope with filters optimized for yECitrine, mCherry, and Cerulean (Chroma). Metamorph software (Molecular Devices) was used to analyze images and quantify single cell YFP and CFP fluorescence from. Between 500 and 1000 cells were imaged

for each sample. Fluorescence levels in the RFP channel was used to discard dead cells (usually <5% of population).

RESULTS

Expression profiles reveal trans interaction between *FLO11* alleles in heterozygous background, where the wild type silenced allele can silence a mutant desilenced allele on the homologous chromosome.

To explore whether transvection at *FLO11* occurred, we started constructing various homozygous and heterozygous strains based on our original haploid Σ1278b strain with the *FLO11* ORF replaced by YFP in the α haploid and CFP in the **a** haploid. We then mated haploids with the wild-type *FLO11* promoter with haploids bearing a *K. lactis* URA3-TEF terminator construct insertion -1.5 kb upstream of the ATG start site of the YFP/CFP ORF (I will refer to this mutant allele as P'_{FLO11}) to create heterozygous strains. To create the homozygous strains, we mated the wild-type *FLO11* promoter haploids of opposite mating types and the mutant P'_{FLO11} haploids of opposite mating types (Figure 1).

We grew the four diploid strains (P_{FLO11}-YFP x P_{FLO11}-CFP, P_{FLO11}-YFP x P'_{FLO11}-CFP, P'_{FLO11}-YFP x P_{FLO11}-CFP and P'_{FLO11}-YFP x P'_{FLO11}-CFP) on YPD plates for 2 days, picked cells directly from the plate and visualized YFP and CFP expression by microscopy. Figure 4 shows the CFP and YFP expression profiles of the diploids. The homozygous wild-type *FLO11* promoter strain (P_{FLO11}-YFP x P_{FLO11}-CFP) distinctly exhibited the four possible expression states (both alleles OFF, both alleles ON, only YFP allele ON and only CFP allele ON), indicative of partial silencing of *FLO11*. Meanwhile, all the cells in the homozygous mutant *FLO11* promoter strain (P_{FLO11}-YFP x P_{FLO11}-CFP) population were ON for both CFP and YFP, indicating that the *FLO11* promoter is desilenced. In the heterozygous diploids (P_{FLO11}-YFP x P'

$_{FLO11}$-CFP and P'$_{FLO11}$-YFP x P$_{FLO11}$-CFP), the expression profiles showed subpopulations of cells with the wild-type allele ON, cells with the wild-type allele OFF, cells with the mutant desilenced promoter ON and *cells with the mutant desilenced promoter OFF*. For example, the expression profile of P$_{FLO11}$-YFP x P'$_{FLO11}$-CFP shows cells in the YFP-ON / CFP-ON quadrant, YFP-OFF / CFP-ON quadrant and the YFP-OFF / CFP-OFF quadrant, which correspond respectively to cells with the wild-type allele ON / mutant desilenced allele ON, cells with the wild-type allele OFF / mutant desilenced allele ON, and cells with the wild-type allele OFF / *mutant desilenced allele OFF*. The fact that we observe cells with the desilenced mutant promoter stably OFF in the heterozygous background, where it was previously desilenced in the homozygous background, suggests that pairing of the desilenced mutant allele with the wild-type partially silenced copy allows silencing of the desilenced mutant allele. Furthermore, we do not observe cells in the quadrant where the wild-type allele is ON and the desilenced mutant allele is OFF. This suggests that the silencing of the desilenced mutant allele is dependent on the state of the wild-type allele, where only the silent wild-type allele is able to transfer silencing to the desilenced allele. When the wild-type allele switches ON, the active wild-type allele is no longer able to silence the desilenced allele.

We then grew the diploid strains in liquid media instead of YPD plates and looked at their expression profiles. We grew the cells in YP 1% ethanol 2% glycerol, a liquid condition we previously used where we observed the wild-type diploid exhibit clearly bimodal expression, which indicated that the promoter was partially silenced in this condition. Figure 5 shows CFP and YFP expression profiles of cell populations grown in this condition. While the profiles are similar to those in Figure 4, the separation between the subpopulations in the different quadrants

is less distinct, especially in the heterozygous strains, compared to the separation observed in the cells grown on YPD plates.

Next, to eliminate the possibility of the phenomenon we observed being construct-specific, we made several other promoter insertion mutants using constructs other than *K. lactis* URA3-TEF terminator. We constructed other desilenced promoter mutants by inserting the *C. glabrata* LEU2-TEF terminator construct at -3 kb from the ATG start site of the YFP / CFP ORF (Figure 2), and at -2 kb from the ATG start site of YFP / CFP ORF; this construct also replaces the native *FLO11* promoter sequence between -2 kb and -3 kb (Figure 3). Like previously, we mated the appropriate wild-type and mutant haploid strains to create the homozygous and heterozygous diploid (Figure 2 and Figure 3).

When we grew the homozygous diploid with the *C. glabrata* LEU2-TEF terminator insertion at the -3 kb region of the promoter in YP 1% ethanol 2% glycerol, we found that essentially all the cells were ON for CFP and YFP, consistent with the promoter being desilenced (Figure 6). Like earlier, the homozygous wild-type promoter diploid is partially silenced in YP 1% ethanol 2% glycerol, and its expression profile exhibits cells in the four possible gene expression states (Figure 6). The heterozygous diploids P''_{FLO11}-YFP x P_{FLO11}-CPF and P_{FLO11}-YFP x P''_{FLO11}-CFP (where P''_{FLO11} is the promoter mutant with the *C. glabrata* LEU2-TEF terminator inserted at the -3 kb region) exhibited expression profiles similar to the profiles in Figures 4 and 5. Again, cells with the desilenced promoter copy stably OFF were observed. Like earlier, no cells with the desilenced promoter copy OFF and the wild-type copy ON were observed, consistent with the hypothesis that the wild-type silenced copy can only transfer silencing factors to the desilenced copy when the wild-type promoter is in the silenced state, and not when it is in the active state. This indicated that the transvection-like effect observed previously was not an oddity particular

to the *K. lactis* URA3-TEF terminator insertion strain originally constructed, but appeared to be a real phenomenon.

Trans interaction does not depend on ncRNA transcripts *ICR1* and *PWR1* complementary sequence pairing

Previously, S. Bumgarner et al discovered that two ncRNA transcripts, *ICR1* and *PWR1*, were transcribed from the *FLO11* promoter region (Bumgarner, Dowell, et al, 2009). The sense transcript *ICR1* was transcribed from the -3 kb region, and terminated close to the *FLO11* ORF ATG start site, whereas the antisense *PWR1* was transcribed from the -2.1 kb region and terminated near the -3 kb region (Bumgarner, Dowell, et al, 2009). To determine whether the transvection between the *FLO11* promoter alleles required the ncRNA transcripts *ICR1* and *PWR1* to pair via their complementary sequence, we looked at the diploids with the *C. glabrata* LEU2-TEF terminator construct replacing the native promoter sequence between -2 kb and -3 kb, which would disrupt any sequence pairing between the ncRNA transcripts emanating from the wild-type and mutant promoter copies (Figure 3). Here we grew the diploids in another condition in which we had previously observed a strongly bimodal *FLO11* expression profile: cells with intermediate Msn1p levels grown in SC liquid media. To increase Msn1p levels in the homozygous and heterozygous diploid backgrounds, we transformed a centromeric plasmid with the *MSN1* ORF driven by a tetO promoter and the *ADH1* promoter driving the reverse tet-transactivator (rtTa) so that we could induce Msn1p expression by addition of doxycycline.

We grew the diploids at 100 ng/ml doxycycline in SC -ura. At this level of doxycycline, *FLO11* expression in the wild-type homozygous diploid is bimodal, indicating that the promoter is partially silenced (Figure 7). The expression profile of the homozygous mutant promoter diploid

P'''$_{FLO11}$-YFP x P''' $_{FLO11}$-CFP (where P'''$_{FLO11}$ is the *FLO11* promoter is the promoter with the *C. glabrata* LEU2-TEF terminator construct replacing the native promoter sequence between -2 kb and -3 kb) showed that all cells were ON for CFP and YFP, consistent with the promoter being desilenced in this background (Figure 7). In the heterozygous backgrounds, the expression profiles exhibited the characteristic "triangular" profile, similar to previous profiles observed (Figure 7). The triangular profile observed in this experiment was consistent with trans allelic silencing occurring between *FLO11* alleles, but with the silenced state being much less stable, likely due to the elevated activator levels. Therefore, as transvection still appeared to occur in these heterozygous backgrounds, the trans interaction between the alleles do not seem to depend on pairing between the ncRNA transcripts via their complementary sequences.

DISCUSSION

Results so far suggest that transvection occurred between *FLO11* alleles in *S. cerevisiae*, and that this trans interaction results in the wild-type promoter copy silencing the desilenced promoter on the homologous chromosome. Interestingly, we did not detect this phenomenon in our early analyses of switching kinetics at *FLO11* (Octavio, Gedeon, and Maheshri, 2009); our previous characterization of promoter switching in the homozygous diploids had suggested that switching occurred independently at each locus, given that the rates of switching OFF to ON and ON to OFF at an allele appeared to be independent of the current state of the allele on the homologous chromosome.

To confirm the occurrence of transvection at *FLO11*, the next question to answer is whether the transvection at *FLO11* is dependent on homologous chromosomal pairing. This can be answered by looking at whether transvection still occurred in a diploid with one of the promoter alleles moved to a different genomic location, or in a haploid with a native *FLO11* promoter copy in the

genome, and another *FLO11* promoter copy on a plasmid. If transvection were occurring, then we would expect the alleles to behave independently in strain backgrounds where homologous chromosomal pairing of the *FLO11* promoter alleles was disrupted.

Currently, our model for possible transvection at *FLO11* is as follows. The wild type promoter copy switches slowly between a silenced and active state, and when it is in the silenced state and interacts in trans with the desilenced mutant promoter allele, it may transfer silencing to the other copy. This is consistent with our model of silencing at *FLO11*, which is described in Chapter 3. Since silencing at *FLO11* appears to require the establishment and maintenance of a looped, hypoacetylated state in which histone deacetylases are localized strongly, it is possible that localization of a desilenced *FLO11* promoter to a silenced allele could allow the silencing factors at the silent copy to transfer to the desilenced allele and establish the silenced state. Once the hypoacetylated, looped state is established, the silenced "structure" is stable enough to last at least several generations, allowing us to observe the stable "OFF" desilenced alleles in the heterozygous background. Again, this is consistent with our model for silencing at *FLO11*, where the looped configuration enables effective localization of silencing factors to the *FLO11* promoter, which stabilizes the silenced state.

In the homozygous wild type background, the trans interaction between the two alleles may still occur although the trans effect was not detected in our previous timelapse experiments. Nonetheless, the experiments with the heterozygous strains clearly showed the transvection effect. Future experiments will be needed to determine whether the trans-allelic effect observed is dependent on homologous chromosomal pairing and what molecular mechanism(s) underlie this phenomenon.

FIGURES AND TABLES

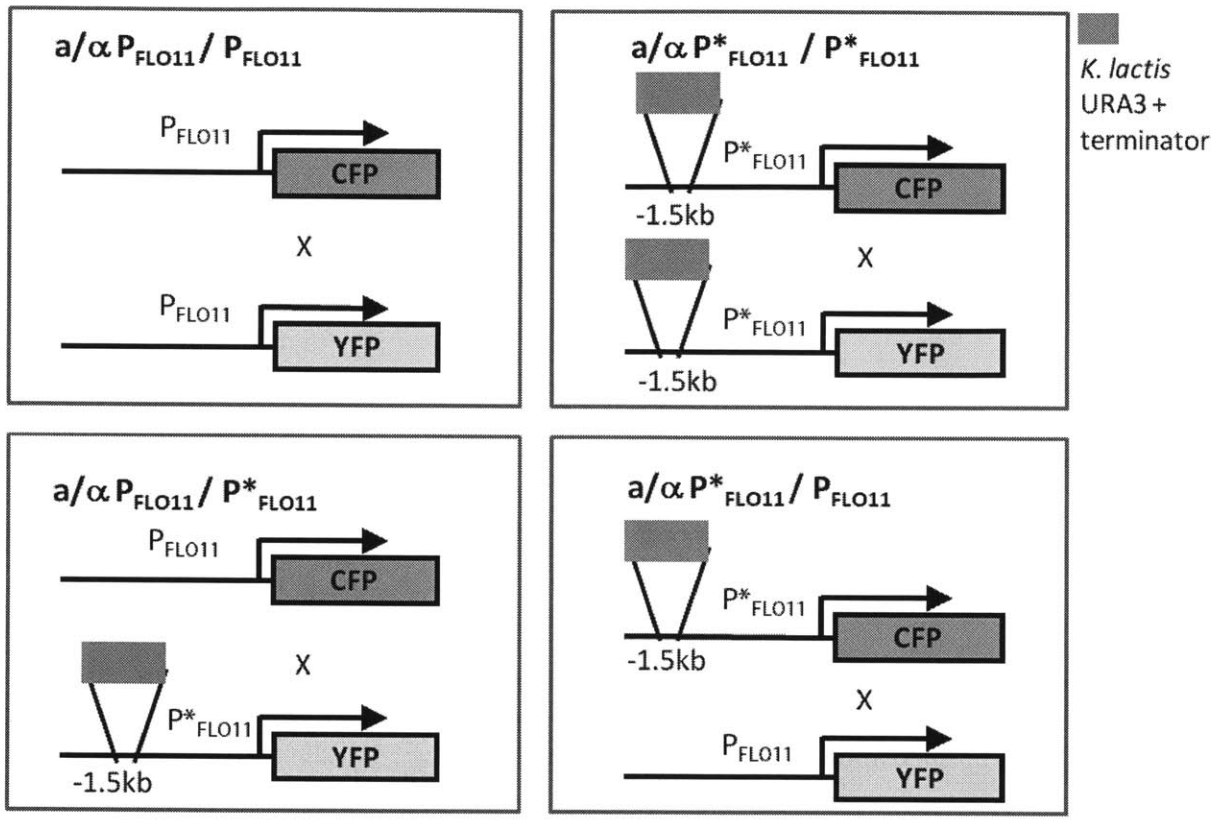

Figure 1: Strains used to assay for transvection at *FLO11* (K. lactis URA3-terminator at -1.5 kb insertion). In the WT homozygous background (top left panel), the *FLO11* promoter is partially silenced. In the mutant homozygous background (top right panel), a *K. lactis* URA3-terminator construct was inserted in the *FLO11* promoter at -1.5 kb; the promoter is no longer silenced in this background. Heterozygous strains were created by mating a haploid WT with a haploid mutant (with *K. lactis* URA3-terminator construct inserted at -1.5 kb of the *FLO11* promoter) of the opposite mating type (bottom left and right panels).

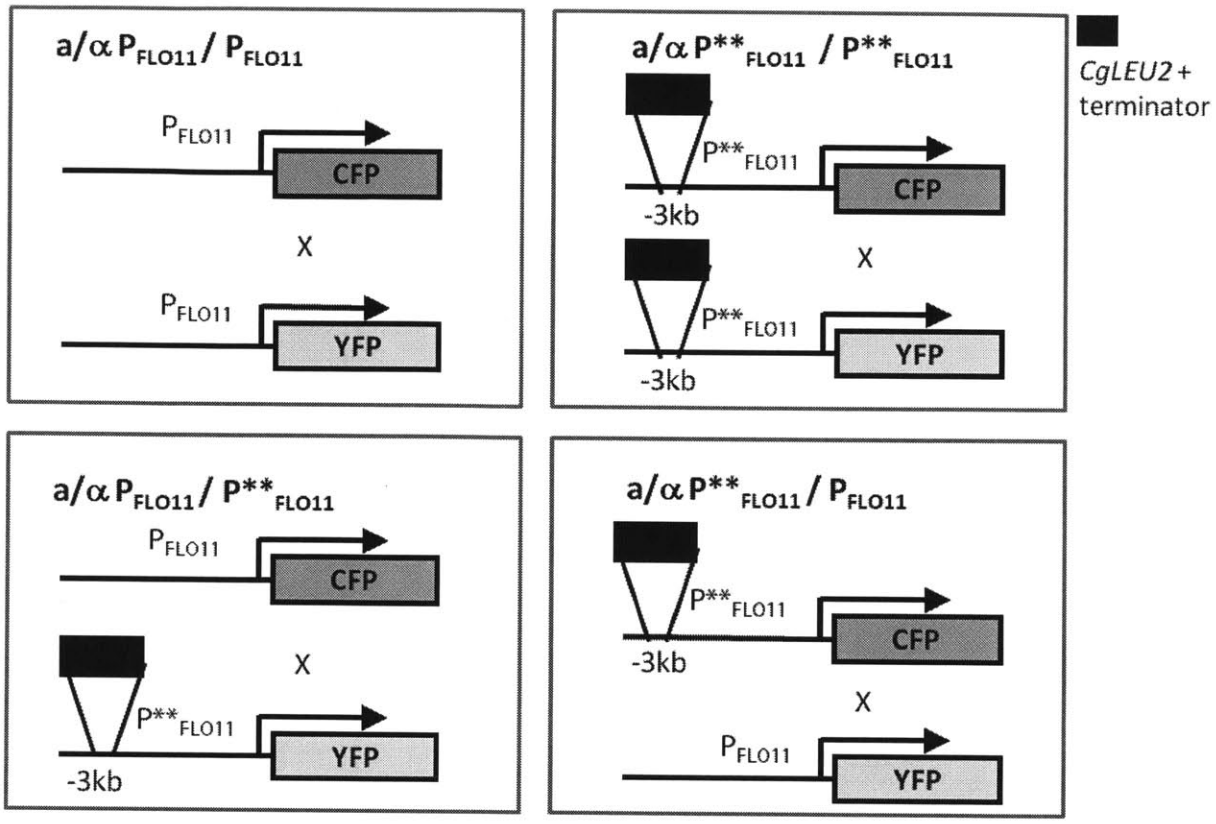

Figure 2: Strains used to assay for transvection at *FLO11* (CgLEU2-terminator insertion at -3 kb). In the WT homozygous background (top left panel), the *FLO11* promoter is partially silenced. In the mutant homozygous background (top right panel), a *CgLEU2*-terminator construct was inserted in the *FLO11* promoter at -3 kb; the promoter is no longer silenced in this background. Heterozygous strains were created by mating a haploid WT with a haploid mutant (with *Cg*LEU2-terminator construct inserted at -3 kb of the *FLO11* promoter) of the opposite mating type (bottom left and right panels).

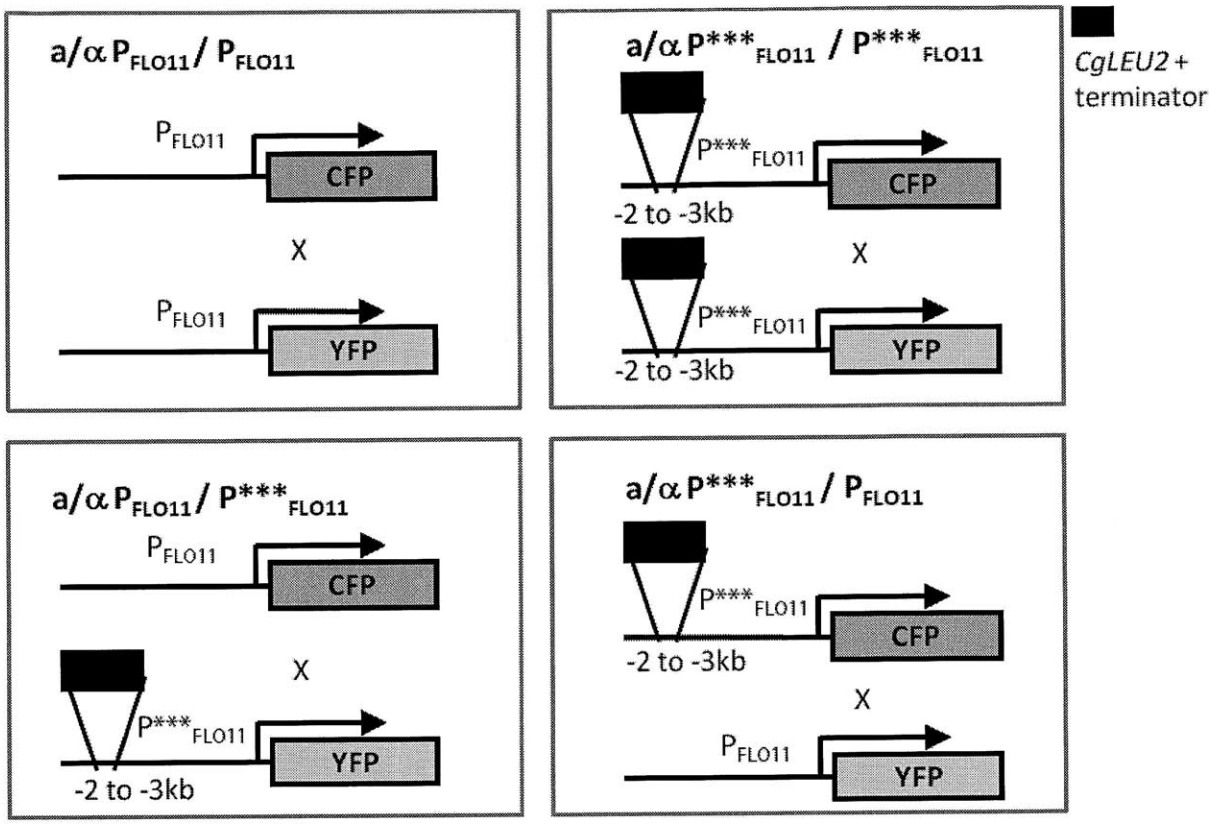

Figure 3: Strains used to assay for transvection at *FLO11* (CgLEU2-terminator insertion at -2 kb, replacing native sequence between -2 and -3 kb). In the WT homozygous background (top left panel), the *FLO11* promoter is partially silenced. In the mutant homozygous background (top right panel), a *CgLEU2*-terminator construct was inserted in the *FLO11* promoter to replace the -2 to -3 kb native sequence. The promoter is no longer silenced in this background. Heterozygous strains were created by mating a haploid WT with a haploid mutant (with *CgLEU2* -terminator construct inserted to replace -2 to -3 kb of native *FLO11* promoter sequence) of the opposite mating type (bottom left and right panels).

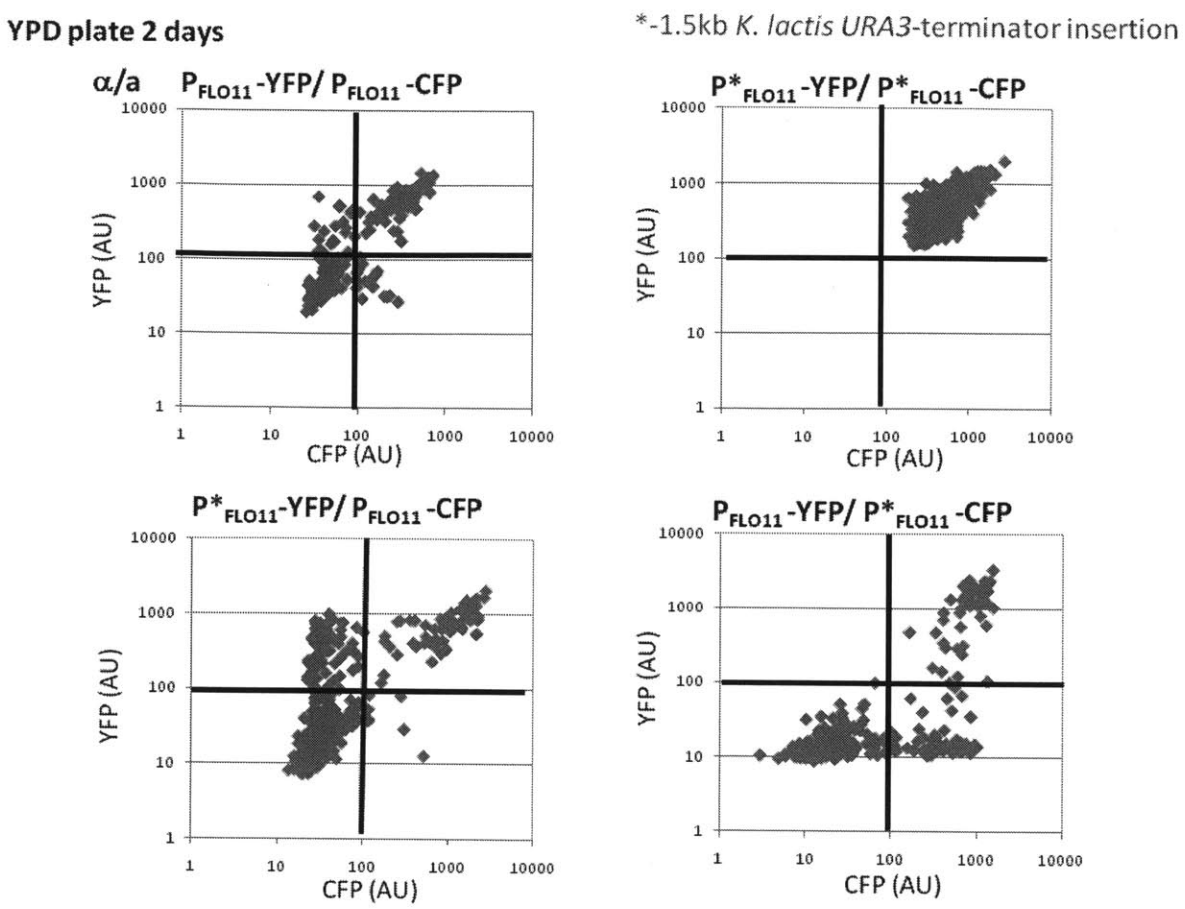

Figure 4: CFP and YFP expression profiles in diploid strains constructed for transvection assay (YPD plate). In the WT homozygous background (top left panel), where the *FLO11* promoter is partially silenced, four possible expression states are observed (cells with both alleles ON, both OFF, and cells with either YFP ON only or CFP ON only). In the mutant homozygous background (top right panel), with *K. lactis* URA3-terminator construct inserted in the *FLO11* promoter at -1.5 kb, all the cells are ON for both alleles, indicating the promoter is no longer silenced in this background. In the heterozygous strains, cells with both alleles OFF were observed (bottom left and right panels), indicating that the mutant desilenced promoter allele was silenced in the heterozygous background.

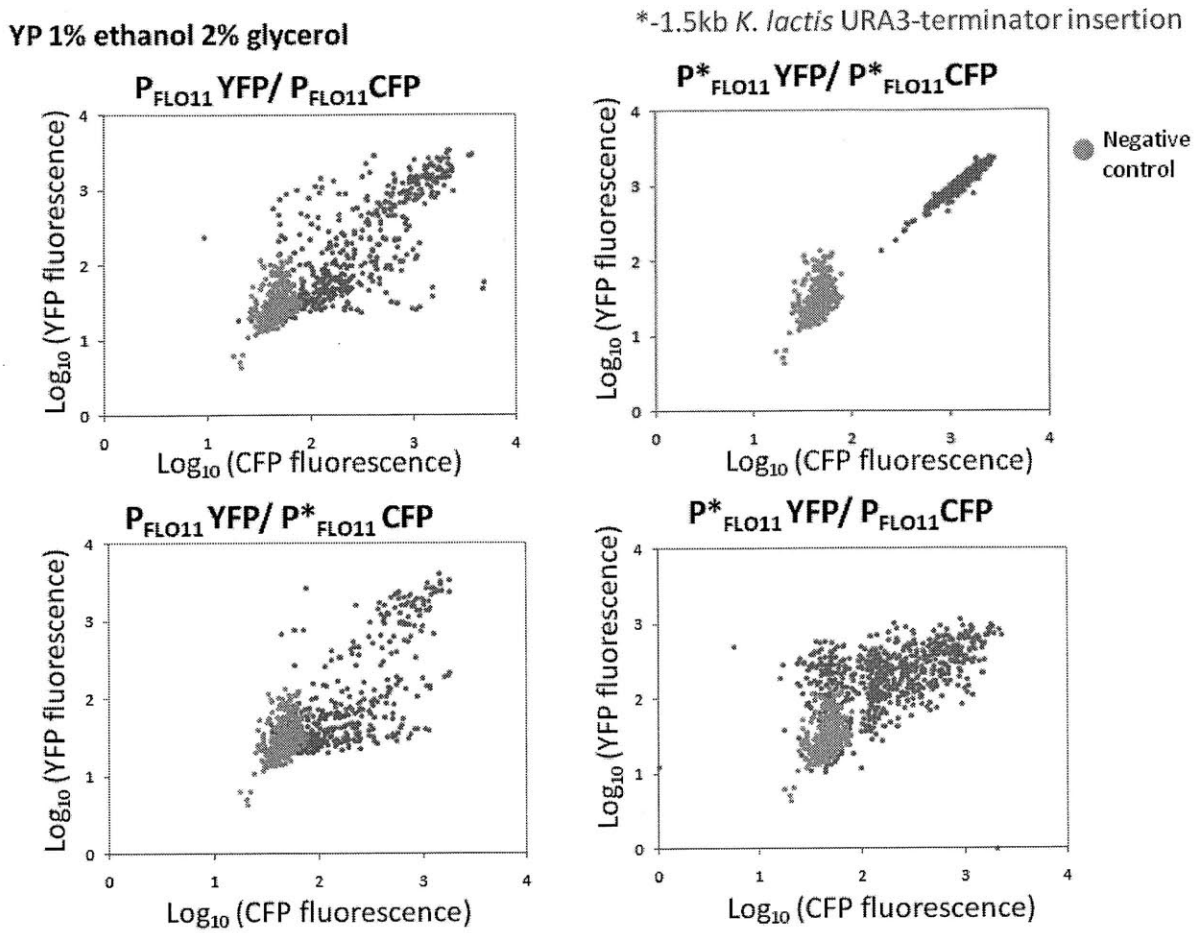

Figure 5: CFP and YFP expression profiles in diploid strains constructed for transvection assay (YP 1% ethanol 2% glycerol liquid media). Negative control strain is Y92 (*flo8Δ*), which does not express *FLO11*. Like in Figure 4, but with the cells grown in liquid YP 1% ethanol 2% glycerol instead of YPD plates. Since cells with both alleles OFF were still observed (bottom left and right panels), this suggests that transvection still occurred and the phenomenon observed previously was not specific to the media used.

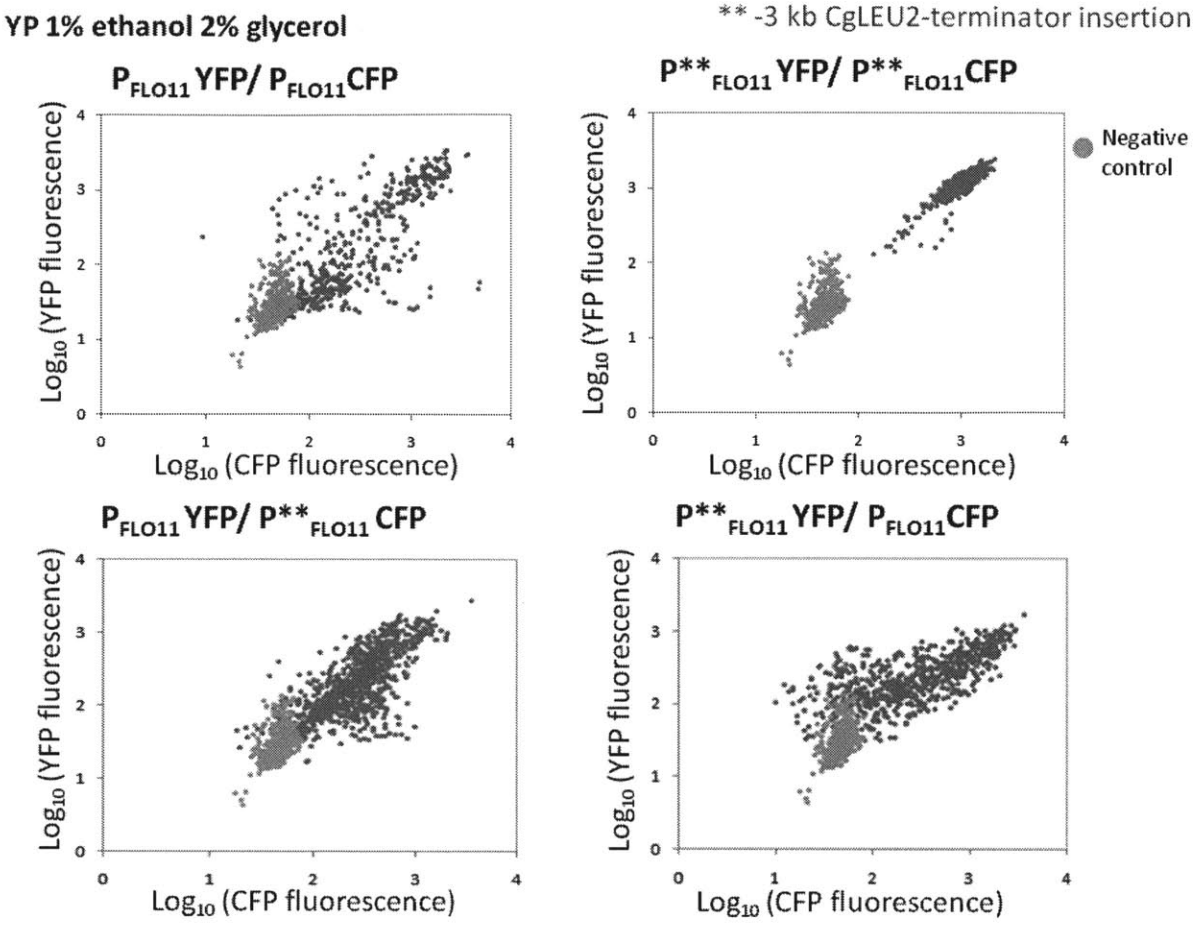

Figure 6: CFP and YFP expression profiles in diploid strains constructed for transvection assay (YP 1% ethanol 2% glycerol liquid media). Like in Figure 5, but the desilenced mutant strain constructs in Figure 2 was used to construct the heterozygotes. In the heterozygous strains, cells with both alleles OFF were observed (bottom left and right panels), indicating that the mutant desilenced promoter allele was silenced in the heterozygous background, suggesting transvection still occurred and was not specific to the original mutant strain construct used.

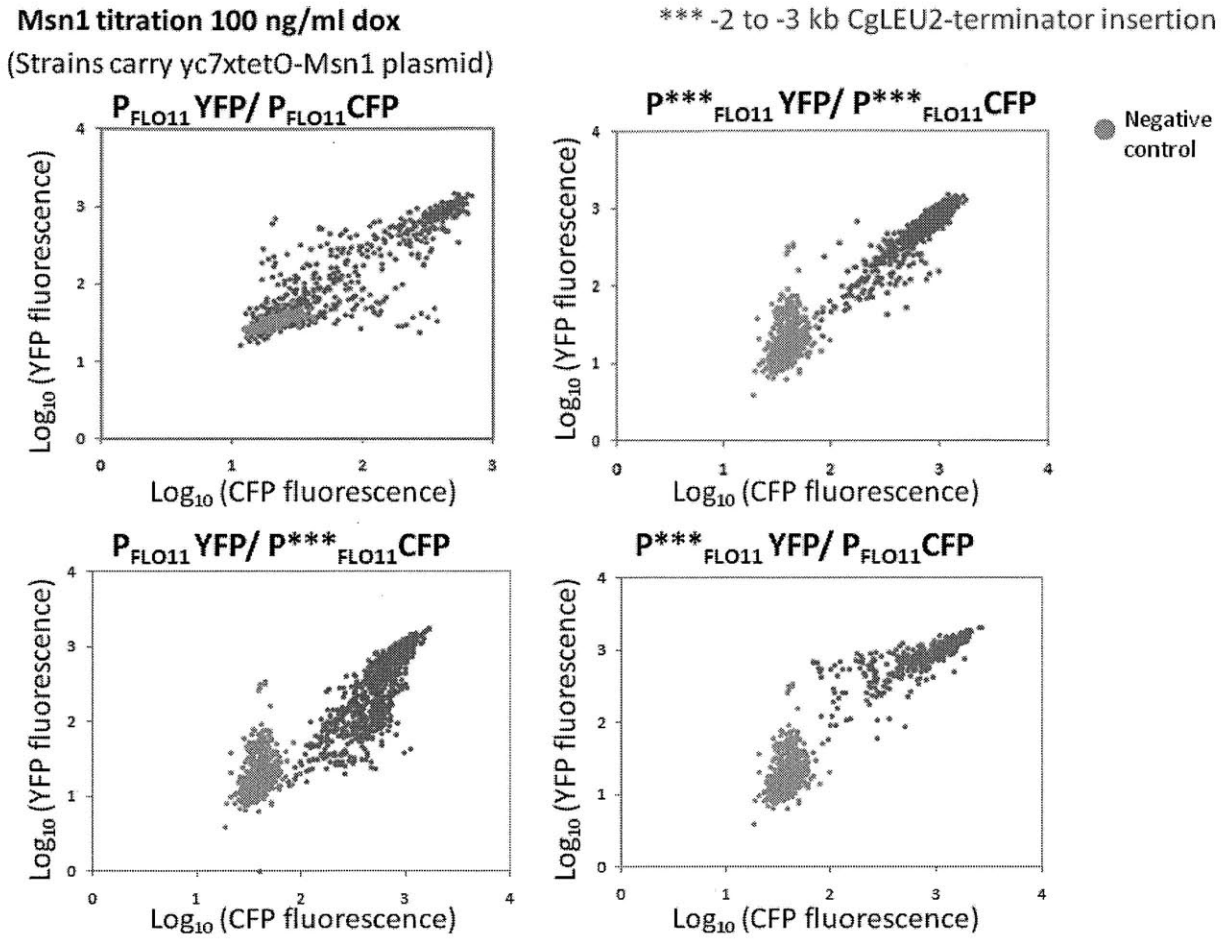

Figure 7: CFP and YFP expression profiles in diploid strains constructed for transvection assay (SC 2% glucose Msn1p titration, 100 ng/ml doxycycline). The desilenced mutant strain constructs in Figure 3 were used to construct the heterozygotes, and strains were grown in another condition known to induce bimodal *FLO11* expression (SC 2% glucose with intermediate levels of the activator Msn1p). The desilenced mutant strains in Figure 3 were constructed so that any pairing of the ncRNAs *ICR1* and *PWR1* via their sequence would be abolished. In the heterozygous strains (bottom left and right panels), the "triangular" expression profiles suggest transvection still occurred in this background, indicating that complementary sequence pairing between *ICR1* and *PWR1* are likely not necessary for transvection to occur. The "triangular" profiles of the heterozygotes are consistent with our model for transvection at *FLO11*, but with the silenced state being much less stable.

SUPPLEMENTAL TABLE 1: Yeast strains used in study

Strains constructed from Σ1278 ML *ura3-52, leu2Δ::hisG* (Lorenz, and Heitman, 1997)

Strain	***Genotype***	***Reference / Source***
MLY42	**Σ1278b** *MAT* α *ura3-52 leu2Δ::hisG*	Lorenz and Heitman, 1997
MLY43	**Σ1278b** *MAT* **a** *ura3-52 leu2Δ::hisG*	Lorenz and Heitman, 1997
Y36	*MAT* α *flo11Δ*::YFP-KanMX6 *ura3-52 leu2Δ::hisG*	Maheshri laboratory collection
Y37	*MAT* a *flo11Δ*::CFP-KanMX6 *ura3-52 leu2Δ::hisG*	Maheshri laboratory collection
Y45	*MAT* α/**a** *flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 ura3-52/ura3-52 leu2::hisG leu2Δ::hisG*	Maheshri laboratory collection
Y92	*MAT* α/a *flo11Δ::YFP-KanMX6/ flo11Δ::CFP-KanMX6 flo8Δ::CgLEU2/flo8 Δ ::CgLEU2 ura3-52/ura3-52 leu2Δ::hisG/leu2Δ::hisG*	Maheshri laboratory collection
LOY006	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -3000 bp FLO11 promoter ::CgLEU2-TEF terminator/-3000bp FLO11 promoter ::CgLEU2-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY011	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -1500bp FLO11 promoter:: K. lactis URA3-TEF terminator /-1500bp FLO11 promoter:: K. lactis URA3-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY012	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -1500bp FLO11 promoter:: K. lactis URA3-TEF terminator /+ ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY013	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 +/-1500bp FLO11 promoter:: K. lactis URA3-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY014	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -3000 bp FLO11 promoter ::CgLEU2-TEF terminator /+ ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY015	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 +/-3000 bp FLO11 promoter ::CgLEU2-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY016	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -2000 to -3000 bp FLO11 promoter ::CgLEU2-TEF terminator/ -2000 to -3000 bp FLO11 promoter ::CgLEU2-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study

SUPPLEMENTAL TABLE 1: Yeast strains used in study (continued)

Strain	***Genotype***	***Reference / Source***
LOY017	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 -2000 to -3000 bp FLO11 promoter ::CgLEU2-TEF terminator/+ ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study
LOY018	*MAT* α/**a** *flo11Δ::YFP-KanMX6 / flo11Δ::CFP-KanMX6 +/ -2000 to -3000 bp FLO11 promoter ::CgLEU2-TEF terminator ura3-52/ura3-52 leu2 Δ::hisG/ leu2 Δ::hisG*	This study

All strains constructed are in the Σ1278b background and cogenic to MLY43.

Strains with selected *FLO11* regulators knocked out were created by PCR integration to delete the regulator ORF in the YFP and CFP haploids and then mated to obtain diploids. Yeast were transformed by the PEG/Lithium Acetate method (Gietz, and Woods, 2002) and all integrations were verified through colony PCR. Strains with plasmids transformed were kept under selection media.

SUPPLEMENTAL TABLE 2: Plasmids used in study

Plasmid	***Genotype***	***Reference / Source***
pYC	*CEN URA3* ADH1 promoter-rtTA	Maheshri laboratory collection
pYCMsn1	7xtetO site (XhoI/BamHI), *MSN1* ORF (BamHI/NotI) in pYC	Maheshri laboratory collection

REFERENCES

Aramayo, R., and Metzenberg, R.L. (1996). Meiotic transvection in fungi. Cell *1,* 103-113.

Bumgarner, S.L., Dowell, R.D., Grisafi, P., Gifford, D.K., and Fink, G.R. (2009). Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast. Proc. Natl. Acad. Sci. U. S. A. *43,* 18321-18326.

Csink, A.K., and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in Drosophila. Nature *6582,* 529-531.

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard, D.A., and Sedat, J.W. (1996). Perturbation of nuclear architecture by long-distance chromosome interactions. Cell *5,* 745-759.

Duncan, I.W. (2002). Transvection effects in Drosophila. Annu. Rev. Genet. 521-556.

Geyer, P.K., Green, M.M., and Corces, V.G. (1990). Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. EMBO J. *7,* 2247-2256.

Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 87-96.

Lewis EB (1954). "The theory and application of a new method of detecting chromosomal rearrangements in Drosophila melanogaster". *Am. Nat.* 88: 225–239.

Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B. *et al.* (2008a). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J. Exp. Med. *8,* 1843-1858.

Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B. *et al.* (2008b). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J. Exp. Med. *8,* 1843-1858.

Lorenz, M.C., and Heitman, J. (1997). Yeast pseudohyphal growth is regulated by GPA2, a G protein alpha homolog. EMBO J. *23,* 7008-7018.

Morris, J.R., Chen, J.L., Geyer, P.K., and Wu, C.T. (1998). Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. Proc. Natl. Acad. Sci. U. S. A. *18,* 10740-10745.

Muller, H.P., and Schaffner, W. (1990). Transcriptional enhancers can act in trans. Trends Genet. *9,* 300-304.

Octavio, L.M., Gedeon, K., and Maheshri, N. (2009). Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. PLoS Genet. *10,* e1000673.

Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim. Biophys. Acta *1,* M1-8.

Rassoulzadegan, M., Magliano, M., and Cuzin, F. (2002). Transvection effects involving DNA methylation during meiosis in the mouse. EMBO J. *3,* 440-450.

Sandhu, K.S., Shi, C., Sjolinder, M., Zhao, Z., Gondor, A., Liu, L., Tiwari, V.K., Guibert, S., Emilsson, L., Imreh, M.P., and Ohlsson, R. (2009). Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. Genes Dev. *22,* 2598-2603.

Chapter 5

SUMMARY OF RESULTS, DISCUSSION AND CONCLUDING REMARKS

SUMMARY AND DISCUSSION

In *S. cerevisiae*, *FLO11* is a member of the FLO gene family which encodes adhesins that enable cells to adhere to solid substrates and to other cells in liquid media (Guo, Styles, et al, 2000). Regulation of *FLO11* expression is complex, where its promoter is one of the largest known in yeast, and many trans factors are known to act upon it (Borneman, Leigh-Bell, et al, 2006; Gagiano, van Dyk, et al, 1999; Pan, and Heitman, 2002; Rupp, Summers, et al, 1999). Recently, *FLO11* expression was shown to be epigenetically regulated (Halme, Bumgarner, et al, 2004), where the gene switched between OFF and ON states at a slow rate (~ once every several cell generations). Here I further characterized the slow switching dynamics at *FLO11*, determined how different regulators affected the slow and fast promoter switching dynamics, and finally, investigated the molecular mechanisms that allowed the trans regulators to generate bimodal gene expression in *cis*.

In Chapter Two, we used the dual-reporter assay to determine whether the slow switching at *FLO11* occurred in *cis* or in *trans*. In this assay, we replaced the native *FLO11* ORFs in diploid yeast with YFP and CFP, so that each copy of the *FLO11* promoter drove expression of a different fluorescent reporter. When we grew the two-color diploid in a poor carbon media to steady-state protein expression, we observed the four possible expression states: cells with both alleles ON, both alleles OFF, cells with YFP ON only, and cells with only CFP ON. Because we observed cells with one allele ON and the other allele OFF, this indicated that slow switching at *FLO11* occurred in *cis*. Next, using timelapse microscopy, we measured the rates of switching

between OFF and ON states at each *FLO11* locus. Within statistical error, the switching rates were identical at each allele, indicating the switching was independent at each locus. Finally, we investigated how different trans regulators of *FLO11* affected promoter switching dynamics. We titrated individual regulators in the two-color reporter background, and estimated switching kinetics using a gene expression model. Based on the promoter response upon activator titration, we found that the activators of *FLO11* could be classified into three types, based on their kinetic roles. Class I activators, which includes Tec1p, Ste12p and Phd1p, affected only the fast promoter switching rates, and did not affect slow switching rates. Next, Class II activators Msn1p and Mss11p affected both fast and slow switching rates; in particular, they stabilized the competent state without destabilizing the silenced state, enabling bimodal *FLO11* expression. Finally, the activator Flo8p stabilized the competent state and destabilized the silenced state in a coupled manner, leading to a fairly graded response. The repressor, Sfl1p, meanwhile, had dual modes of repression: it could silence (i.e. affect slow switching rates) the promoter, and also, in the absence of silencing, it could conventionally repress expression (i.e. affect fast promoter switching rates).

In Chapter Three, we investigated the molecular mechanisms that led to Sfl1p's ability to stably silence the *FLO11* promoter, and the mechanism that enabled Msn1p to stabilize the competent state without destabilizing the silenced state, to yield a bimodal response. We found that when the promoter was silenced, Sfl1p was bound to the -1.2 kb region and the core promoter region, and these two regions also associated *in vivo*. Sfl1p was required to recruit the histone deacetylase Hda1p to the promoter, which then spread across the entire promoter region. We also found that the ncRNA transcript *ICR1*, which was previously found to be necessary for silencing of the promoter (Bumgarner, Dowell, et al, 2009), was required to recruit additional

histone deacetylases to the promoter via Set1p and Set3p, which led to full hypoacetylation and stable silencing. Next, upon overexpression of the activator Msn1p, a new looped configuration at the promoter was observed, where the core *FLO11* promoter region associated with the -2.1 kb (*PWR1* promoter) and -3 kb (*PWR1* terminator) regions. Formation and/or maintenance of this loop required *PWR1* transcription and TFIIB-mediated interactions between transcription initiation and termination machinery. The loop observed in the Msn1p-activated promoter state was also required for Msn1p's ability to stabilize the competent state without destabilizing the silenced state, which enables bimodal *FLO11* expression.

In Chapter Four, we explored the possible occurrence of transvection at *FLO11*. We constructed heterozygous two-color diploids with one mutant desilenced copy of the *FLO11* promoter and a wild-type partially silenced copy on the homologous chromosome. We then grew these heterozygous diploids, the homozygous wild-type two-color diploids and homozygous desilenced mutant diploids on YPD plates for two days, and then observed expression profiles. In the homozygous wild-type diploid, *FLO11* expression was bimodal, indicating the promoter was partially silenced, consistent with what we had observed before. *FLO11* expression in the homozygous diploid with the mutant desilenced promoter was unimodal ON, indicating absence of silencing, as we expected. However, in the heterozygous diploids, we observed three distinct expression states in the cell population: cells with both alleles ON, cells with the desilenced mutant allele ON and the wild-type allele OFF, and cells with both alleles OFF. This indicated that the desilenced mutant allele, which should be ON, was now somehow stably silenced since we observed a distinct population with both alleles OFF. Further experiments with other heterozygous strains using different desilenced *FLO11* promoter mutant alleles paired with a wild-type allele on the homologous chromosome revealed similar expression profiles with a

distinct population of cells with both alleles OFF, suggesting that the phenomenon may be real and not an artifact of the original mutant constructs used. Future experiments will be needed to determine whether the trans-allelic silencing effect is dependent on homologous chromosome pairing and to elucidate the mechanism of this phenomenon.

In summary, I have characterized the kinetics of slow switching at *FLO11* and determined that the slow promoter fluctuation occurs in *cis* and *FLO11*'s trans regulators affected different combinations of slow and fast switching rates at the promoter, enabling the cell population to effectively modulate not only levels of *FLO11* expression but also the levels of expression heterogeneity. The remarkably wide timescale range of the promoter state fluctuations at *FLO11* allows the gene to be activated in either a bimodal manner or graded manner, depending on the upstream signals the promoter receives.

One way of activating *FLO11* is by destabilizing the silenced state altogether, and activating transcription at the desilenced promoter. This mode of activation results in a *graded* response, where the cell population exhibits a unimodal expression profile. In our desilenced promoter mutants, the promoter is not looped and generally hyperacetylated. However, with the exception of *sfl1Δ*, *FLO11* expression in these backgrounds was low, most likely because Sfl1p is still bound and conventionally represses expression. Addition of activators in this background increases expression levels in a graded manner. It is unlikely that cells would abolish Hda1p or Set1p activity to desilence the promoter and activate expression in a graded fashion in this manner. More likely, cells achieve graded activation by displacement of Sfl1p binding at the -1.2 kb region, which would abolish the association between this region and the core promoter. In support of this, we have shown previously that titration of the activator Flo8p, which also binds to the -1.2 kb region (Pan, and Heitman, 2002), destabilized the silenced state and stabilized the

active state at the promoter in a *coupled* manner, thereby yielding a fairly graded response. Moreover, when we prevented Sfl1p binding to the -1.2 kb region by having tetR bind to a tetO site inserted at this region, titration of activators in this background yielded a graded response.

How can cells achieve *bimodal* activation of *FLO11*? At intermediate levels, the activator Msn1p stabilizes the active state *without* destabilizing the silenced state, allowing bimodal *FLO11* expression. We believe that unlike the activator Flo8p which can easily displace Sfl1p from the -1.2 kb site by competing for the same binding, Msn1p is unable to effectively destabilize the silenced state as it may not easily access regions of the *FLO11* promoter when the promoter is silenced; furthermore, at every step of attempting to establish an active state, it must always "compete" with multiple repressive events. To establish the active state, Msn1p must "wait" until conditions at the promoter are favorable. Because molecular events at the promoter are stochastic, Msn1p might access the promoter at one instance but may get kicked off the promoter if *ICR1* happened to be transcribed shortly after. Then Msn1p must "wait" again to be able to access the promoter. But, it may also happen that when Msn1p accessed the promoter, it got a "lucky" break and *PWR1* was successfully transcribed. If Msn1p was once more "lucky," the establishment of the active loop conformation and localization of transcriptional machinery at the *FLO11* core promoter would be successful, allowing *FLO11* to be transcribed highly. This looped active state conformation is highly stable, as Sfl1p and the other repressive factors that silence the promoter must now "compete" with multiple activating events at various steps to establish the silenced state. Again, because the transition between the silenced and Msn1p-activated state involves many steps, where each step occurs stochastically and in competition with other events, the overall rate switching rate between the silenced and active states is slow. In other words, it will take many attempts for either Sfl1p or Msn1p to be successful in

establishing either the silenced or active state because the chances of having a "lucky" streak is rare; success happens only about once every several cell generations.

This work demonstrates molecular mechanisms of how a complex promoter might integrate different signals to modulate slow and fast promoter dynamics and ultimately, levels of phenotypic heterogeneity. Essentially, different *trans* regulators of *FLO11* are able to orchestrate molecular events that lead to promoter conformations that determine the overall promoter switching rates. Importantly, the slow promoter dynamics at the *FLO11* promoter is encoded in *cis*, where the bistability in gene expression is due to the multiple events in *cis* (looping, ncRNA transcription, recruitment of histone-modifying enzymes etc.) that reinforce each state. We also uncovered novel physiological roles for DNA looping, where the silenced and active looped conformations at *FLO11* help stabilize each respective state, forming the basis for *cis*-encoded epigenetic switching. Given the recent findings demonstrating widespread occurrence of DNA looping and ncRNA transcription across the genome in other eukaryotes, including human cells, at least some of the epigenetic phenomenon in human cells are likely to involve long-range DNA interactions and ncRNA transcription, in addition to histone modifications and *trans* regulators. A better integrated study of these events is needed to decipher their intricate interplay and ultimate effect on the regulation (or mis-regulation) of phenotypic switching and cell fate decisions.

FUTURE DIRECTIONS

First, work presented in Chapter 3 requires a few additional experiments before submitting for publication. Prof. Narendra Maheshri will be continuing this work in his lab. The experiments involving *sua7-1* will be repeated in a different strain background, with the endogenous sua7 allele replaced with *sua7-1* allele in its native locus (more details in the Appendix).

Next, possible future directions of this work can be divided into the following categories:

1. Further explore mechanisms of regulation of *FLO11* expression dynamics by other *FLO11* regulators.
2. Investigate mechanism of promoter loop formation.
3. Further study transvection at *FLO11*.

Further studies of molecular mechanisms of gene expression dynamics at *FLO11* can be divided into 1) studies characterizing the detailed molecular mechanisms by which other trans regulators, Mss11p, Tec1p, Ste12p, Phd1p and other activators found in more recent genetic screens control *FLO11* expression and 2) studies combining biophysical theories of nucleosome modifications, trans regulator binding events, ncRNA transcription, and DNA looping to more generally understand how these ubiquitous *cis* events impact the dynamics of gene expression.

To determine the detailed molecular mechanisms by which other regulators of *FLO11* control *FLO11* expression, the occupancy of the promoter by regulators under different conditions and/or mutant backgrounds could be determined from additional ChIP experiments. Correlating this data with promoter histone acetylation levels, nucleosomal occupancy, chromatin conformation and single-cell *FLO11* expression data would yield a better picture of

how the pathways in which regulators activated or repressed expression, and whether the pathways involved mechanisms that resulted in slow or fast promoter dynamics.

While the approach of elucidating individual activation and repression mechanisms by each of the many different regulators of *FLO11* would be fairly straightforward and would yield useful and interesting information, the ultimate "story" from this approach would be *FLO11*-specific, and may likely be of interest only to the *FLO11* and / or transcription dynamics communities. A more difficult but potentially more rewarding approach to studying regulation of *FLO11* transcription dynamics would be to develop solid theoretical / biophysical models to explain the data and reconcile the looping, ncRNA transcription, transcription factor binding, nucleosome / chromatin remodeling and modification events with resulting gene expression dynamics. If such models can successfully capture the important, general features of gene expression dynamics as a function of *cis* events at the promoter, these models could be "scaled up" to describe gene expression at a genomewide level. Many important cell-fate decision processes (e.g. stem cell to differentiated cell, or normal to cancerous cell transitions) involve changes in expression patterns genomewide, yet the mechanisms by which these transitions occur and how the resulting gene expression patterns are kept stable for many generations are poorly understood. A general model which encapsulates the events relevant to influencing and stabilizing gene expression genomewide may help in accelerating our understanding of how underlying genomewide expression patterns transition and give rise to various cellular identities / states.

Next, another interesting direction to pursue would be investigation of the mechanisms behind loop formation. First, the proteins involved in the looping interaction should be identified; this could be determined by genetic screens for candidate proteins and then

biochemical experiments to pull out the proteins in complex with the candidate protein. Then, the proteins in the loop complex could be further characterized with *in vitro* experiments. If any of the proteins involved bind or localize to many other regions in the genome, or are involved in other general cell process(es), then the protein(s)' activity and function in the other genomic regions and/or cell process can be further investigated. It would be of interest to determine if there are general regulatory pathways that control long-range DNA interactions and higher order chromatin structures genomewide, and if so, how these pathways operate and how these pathways interact with transcription regulation and possibly other cell processes.

Finally, the next step in the transvection project is to carry out experiments to determine whether the trans-allelic effect at *FLO11* observed in the heterozygous strain backgrounds is still observed in backgrounds where homologous pairing between the *FLO11* alleles is disrupted. One quick way to determine this is to look at the expression profile in a haploid with one genomic copy of the *FLO11* promoter driving a fluorescent reporter, and another copy of the *FLO11* promoter driving a fluorescent reporter (of a different color) on a plasmid. If the silencing of *FLO11* "in trans" requires homologous chromosomal pairing, then we should not observe any transvection in this experiment. Another way to carry out this assay is to observe *FLO11* expression in a diploid heterozygous background, with one copy of the *FLO11* promoter driving expression of a fluorescent reporter at the native *FLO11* locus, and the other copy of the *FLO11* promoter driving a fluorescent reporter transplanted to a non-native genomic location. Again, if the phenomenon observed previously was truly transvection at *FLO11*, then we would expect the trans-allelic effect to be abolished in this experiment, where homologous chromosomal pairing is disrupted.

REFERENCES

Borneman, A.R., Leigh-Bell, J.A., Yu, H., Bertone, P., Gerstein, M., and Snyder, M. (2006). Target hub proteins serve as master regulators of development in yeast. Genes Dev. *4,* 435-448.

Bumgarner, S.L., Dowell, R.D., Grisafi, P., Gifford, D.K., and Fink, G.R. (2009). Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast. Proc. Natl. Acad. Sci. U. S. A. *43,* 18321-18326.

Gagiano, M., van Dyk, D., Bauer, F.F., Lambrechts, M.G., and Pretorius, I.S. (1999). Msn1p/Mss10p, Mss11p and Muc1p/Flo11p are part of a signal transduction pathway downstream of Mep2p regulating invasive growth and pseudohyphal differentiation in Saccharomyces cerevisiae. Mol. Microbiol. *1,* 103-116.

Guo, B., Styles, C.A., Feng, Q., and Fink, G.R. (2000). A Saccharomyces gene family involved in invasive growth, cell-cell adhesion, and mating. Proc. Natl. Acad. Sci. U. S. A. *22,* 12158-12163.

Halme, A., Bumgarner, S., Styles, C., and Fink, G.R. (2004). Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. Cell *3,* 405-415.

Pan, X., and Heitman, J. (2002). Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. Mol. Cell. Biol. *12,* 3981-3993.

Rupp, S., Summers, E., Lo, H.J., Madhani, H., and Fink, G. (1999). MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast FLO11 gene. EMBO J. *5,* 1257-1269.

APPENDIX

STRAIN CONSTRUCTION / EXPERIMENTS TO ADDRESS ISSUES IN CHAPTER 3

Prof. Narendra Maheshri will be continuing work presented in Chapter 3 in his lab. The following strain construction and experiments will be carried out:

Experiments investigating role of TFIIB in active looping:

Strains to be constructed:

1. HA-tagged *sua7-1* allele integrated at the *SUA7* locus to replace native *SUA7* ORF in the diploid two-color strain background

   Genotype: *MAT α/a flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 sua7Δ :: HA-tagged sua7-1 CgLEU2/ sua7Δ :: HA-tagged sua7-1 CgLEU2 ura3-52/ura3-52 leu2 Δ::hisG/leu2Δ::hisG*

2. HA-tagged *SUA7* allele integrated at *SUA7* locus to replace native *SUA7* ORF in the diploid two-color strain background

   Genotype: *MAT α/a flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 sua7Δ :: HA-tagged sua7 CgLEU2/ sua7Δ :: HA-tagged sua7 CgLEU2 ura3-52/ura3-52 leu2 Δ::hisG/leu2Δ::hisG*

To make the first strain listed above, the integration construct will be made by PCR fusion of HA-tagged *sua7-1* allele and a CgLEU2 marker. The ends of the construct will have ~ 50 bp homology to the terminator of *SUA7* and the interior portion of *SUA7* where the *sua7-1* (E62K) mutation is located. Here, the single base mutation is encoded in the primer used to make the

PCR fusion construct. The construct is designed so that the CgLEU2 marker is downstream of the HA-tagged *sua7-1* allele when it is integrated.

To make the second strain, the integration construct will be made by PCR using primers that amplify the CgLEU2 marker; the primers will also contain sequences for the HA tag, stop codon and 50 bp homology to the *SUA7* terminator region. The integration construct will be designed so that the HA tag is fused to the C-terminus of Sua7p.

Each construct will be integrated in the haploid YFP and CFP (Y36 and Y37) strains, which are of opposite mating types. The haploids will then be mated to obtain the diploid strain, and the diploid will be transformed with a *URA3*-marked centromeric plasmid carrying the *ADH1* promoter driving rtTa expression and a 7xtetO promoter driving *MSN1* expression.

Experiments to perform:

1. Msn1p titration in the *sua7-1* background.
2. ChIP for HA-tagged TFIIB in the *sua7-1* and WT background with Msn1p overexpressed.
3. 3C probing for interactions between *FLO11* core promoter and upstream regions in the *sua7-1* background with Msn1p overexpressed.

The Msn1p titration in the *sua7-1* background will be performed according to the procedure described in page 30. ChIP for HA-tagged TFIIB in the *sua7-1* and WT backgrounds will be performed according to the procedure in page 92. The same primers probing the *FLO11* promoter region used in Chapter 3 will be used, and the same anti-HA antibodies (Roche Cat. # 11 867 423 001) will be used. The negative control will be a strain not carrying an HA-tagged allele treated with anti-HA. Occupancy levels of TFIIB at the *FLO11* promoter will be measured

relative to occupancy of TFIIB at a telomeric region, using primers probing the *FLO11* promoter and a region in tel-VIR. Finally, 3C will be performed to probe for in vivo association between the *FLO11* core promoter and upstream promoter regions using the procedure in page 92. Like all 3C experiments carried out in Chapter 3, the same restriction enzyme (AluI) will be used, and the same primers to probe for the promoter region associations will be used.

Experiments investigating mechanism of role of *PWR1* transcription in active looping

Strain to be constructed:

1. tetO site inserted at -2.1 kb (*PWR1* promoter) in the diploid two-color strain

   Genotype: *MAT α/a flo11Δ::YFP-KanMX6/flo11Δ::CFP-KanMX6 -2.1 kb FLO11 promoter:: tetO site loxP / -2.1 kb FLO11 promoter:: tetO site loxP ura3-52/ura3-52 leu2 Δ::hisG/leu2Δ::hisG*

Currently, the YFP and CFP haploids carrying the tetO site insertion at -2.1 kb of the *FLO11* promoter are available, but the haploids are of the same mating type. The mating type of one of the haploids will be switched, and once the opposite mating type is obtained, the haploids of opposite mating types will be mated to get the diploid. Finally, the diploid will be transformed with a *URA3*-marked plasmid carrying the *ADH1* promoter driving rtTa and 7xtetO driving *FLO8* expression.

Experiment to be performed:

The strain will be grown in SC media with various concentrations of ethanol (0.5 to 2%) to modulate Flo8p activity and various concentrations of doxycycline (0 to 1000 ng/ml) to control expression of *PWR1*, and the fluorescence microscopy of the cells will be performed according

to the procedure in page 30. The experiment is designed to mimic displacement of Sfl1p at -1.2 kb of the promoter and simultaneous *PWR1* transcription, and this experiment will test whether transcription of *PWR1* and displacement of Sfl1p at -1.2 kb may be sufficient to give the bimodal gene expression observed in the Msn1p titrations.

FLO11 BIOPHYSICAL MODEL AND SIMULATION RESULTS

Model description

To gain biophysical insight into how loops at the *FLO11* promoter may stabilize the silenced and active chromatin states, we used simulations of a model relating epigenetic memory to nucleosome modifications (Figure 1) to explore how binding of activators, repressors and various looped configurations at the *FLO11* promoter influenced the stability of the macroscopic chromatin state.

The model was developed by Dodd et al and is described in Dodd, I.B., Micheelsen, M.A., Sneppen, K., and Thon, G. (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell, 129 (4), 813-822. Essentially, in the model, each nucleosome in an array of N nucleosomes can be in one of three possible states: M (methylated), U (unmodified) and A (acetylated) (Figure 1). M nucleosomes are nucleosomes with modifications associated with a "repressive" or "silenced" chromatin state, while A nucleosomes have modifications associated with "active" chromatin states. At each timestep during the simulation of the model, either a "feedback" interaction takes place, or a "noise" move is made; the probability of a feedback interaction occurring is defined by a parameter representing "feedback strength". If a "feedback" interaction occurs, a nucleosome in the array is randomly picked and is able to interact with any other nucleosome in the array. The nucleosomal interaction results in one nucleosome

influencing the modification state of the other nucleosome by having the other nucleosome's modification state transition towards the original nucleosome's current modification state. For example, an M nucleosome would make an A nucleosome become a U nucleosome; or, an M nucleosome would make a U nucleosome into an M nucleosome. Modification states of nucleosomes transition between M to U to A, and vice-versa; no direct M to A (or vice-versa) transition is allowed. If a "noise" move is made instead, a random nucleosome is selected in the array, and its modification state will transition either towards M or A, with equal probability.

One change we made in our model implementation compared to the model implementation of Dodd et al is that we did not allow nucleosomes in the array to interact with any other nucleosome in the array at random. Instead, nucleosomes in the array had a very high probability of interacting with nearest neighboring nucleosomes and a much lower probability of interacting with nucleosomes farther away. The probability of interaction was defined by a power-law model, where the probability of interaction decreased sharply with increasing distance between the nucleosomes.

To simulate the effect of activator / repressor binding and looping on the chromatin state of the *FLO11* (which is represented by an array of 16 nucleosomes), we added extra parameters that represented the frequency of activator / repressor binding and looping. Binding of an activator resulted into influencing the nucleosomes at the binding region to transition towards the A state, whereas binding of a repressor resulted into transition towards the M state. Looping between specified regions of the array (which were originally regions distant from each other in the absence of looping) increased the probability in which the nucleosomes in these regions interacted with each other.

Below is a quick summary of our implementation of the model:

Define parameter α (feedback strength).

Iterate following for many timesteps:

Step 1: Select nucleosome n1 randomly from array of nucleosomes

Step 2: Select nucleosome n2 with whatever defined probability (Dodd et al used both a uniform distribution and power law; here we used a power law where nucleosomes interact most frequently with their nearest neighbors and much less frequently with far away nucleosomes).

Step 3: Select random number x [0,1]. If $x < \alpha$ attempt a "feedback" move.

If $x > \alpha$ attempt a "noise" move.

Feedback move:

Basically, make n2 more like n1: if n2 is M, n1 becomes U if it's A, or M if it's U.

if n2 is A, n1 becomes U if it's M, or A if it's U.

Unmodified histones cannot make modifications; modifications occur one step at a time (no direct M → A or A → M moves).

Noise move:

If n1 is U, change to M or A with probability ½

If n1 is A or M, change it to U with probability ½

To the basic model implementation, **supplement each step** with the following:

To add repressor(s):

1. Define repressor binding sites (specify the nucleosome(s) at site)

We did not implement a more complicated probability law for selecting which nucleosome gets modified when the repressor is bound. If repressor is bound, nucleosome(s) at site selected will be modified with probability =1.

2. Define parameter s_r and select random number x [0,1]. If $x < s_r$ then repressor is "bound."

Change nucleosome(s) at repressor site to step one step towards M

To add activator(s):

Same formulation as above for repressor, except define s_a (frequency of activator binding) and change nucleosome(s) at activator step to step one step towards A if activator bound

To add loop(s):

Define s_{lr} (frequency of repressor looping) and sites that contact each other in loop

Repressor loop:

Select random number x [0,1]. If repressor bound and $x < s_{lr}$, then select n1 according to power law in the neighborhood of the site that repressor contacts via looping. Modify n1 to take one step towards M.

Activator loop:

Same formulation as repressor loop, except define s_{la} (frequency of activator looping) instead, and modify n1 (nucleosome selected according to power law in the neighborhood of the site that activator contacts via looping) to take one step towards A.

Simulation parameters

In this model simulation, the feedback strength $\alpha = 0.7$ and a power law distribution was used to define nucleosomal interactions, where nucleosomes affect nearest neighbors only up to 3 nucleosomes away, with decreasing probability of affecting farther neighbors. The *FLO11* promoter was represented by an array of 16 nucleosomes (roughly 3.2 kb if 1 nucleosome occupied about ~ 0.2 kb of DNA).

Figures of model and simulation results

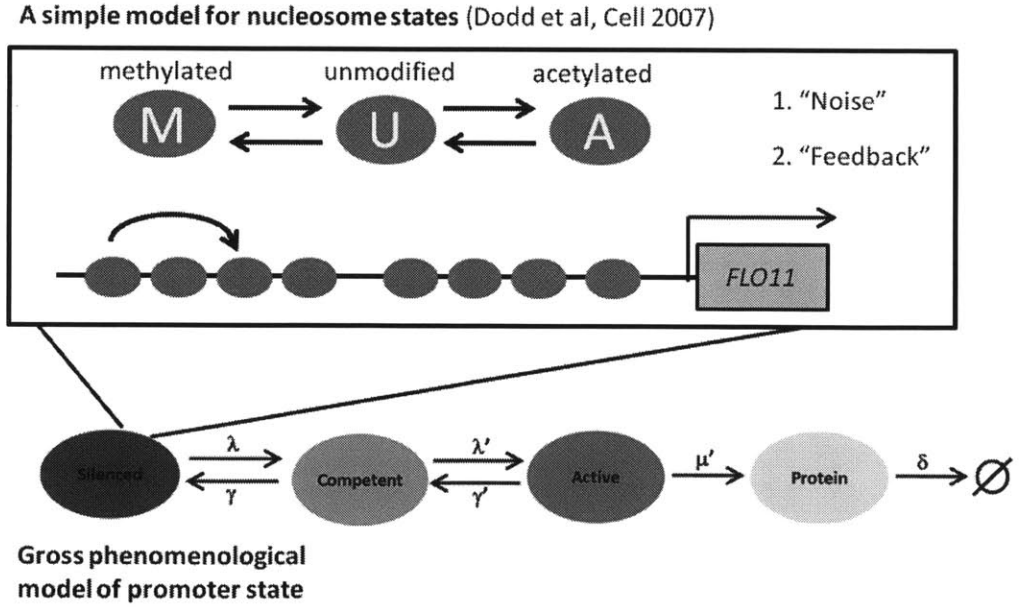

Figure 1: In the model for epigenetic memory developed by Dodd et al., each nucleosome in the region transitions between three different states. During a simulation of this model, a random

nucleosome from the region is selected, and with some probability, its state will either be influenced by the state of a neighboring nucleosome or will transition to another at random. Simulations of this model with activators, repressors and looping added were used to computationally investigate the effects of these molecular events on the "gross" promoter state (the "silenced" or "competent" states).

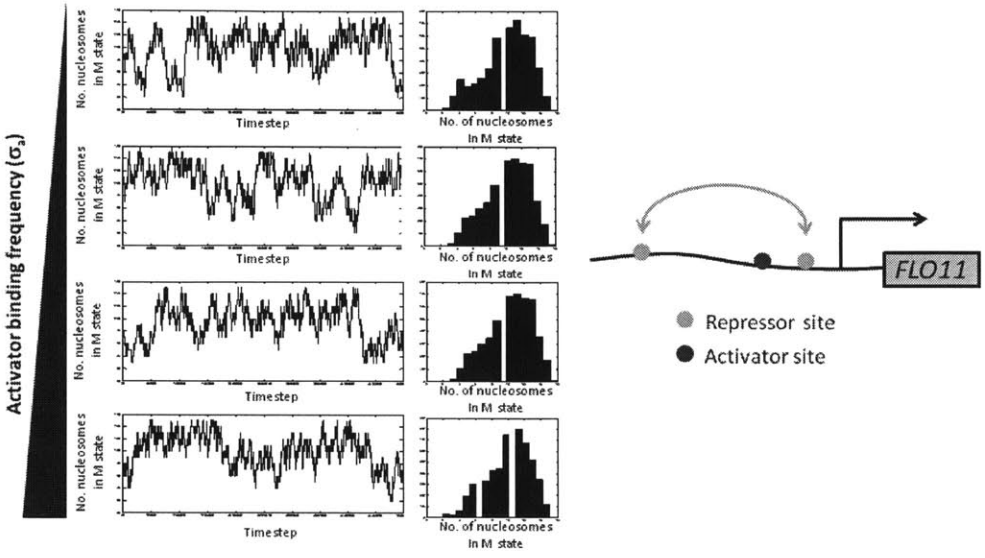

Figure 2: Model simulation of a promoter region with two repressor sites that can form a loop and one activator site between the repressor sites. Results show that upon increasing activator levels in this background, the entire region remains in a high "M" state (corresponding to a repressive chromatin state).

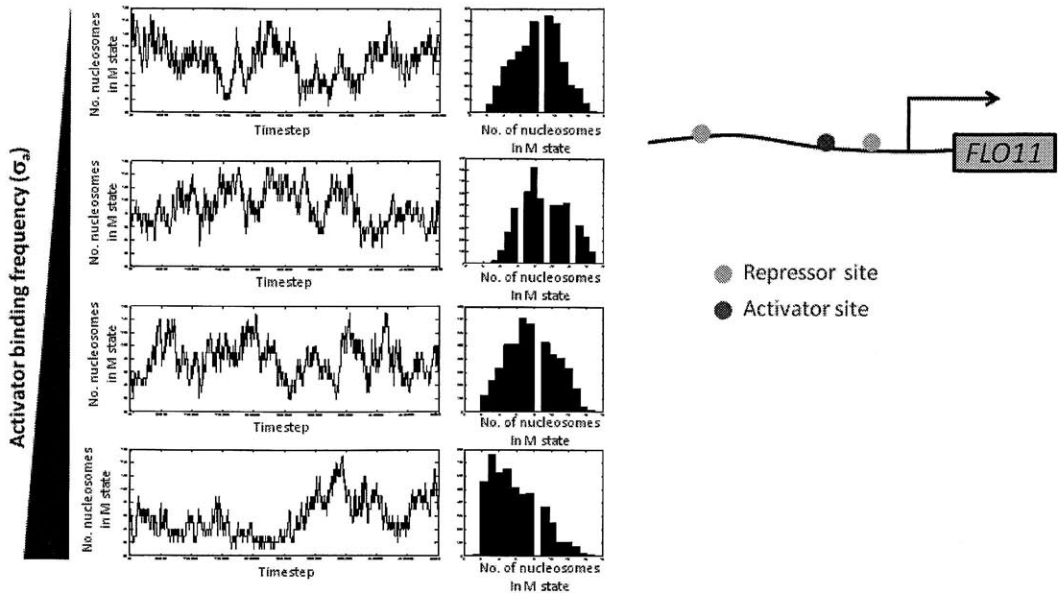

Figure 3: Model simulation of a promoter region with two repressor sites that do not form a loop and one activator site between the repressor sites. Results show that upon increasing activator levels in this background, the average number of "M" nucleosomes in the region decreases, and the entire region eventually remains at a low "M" state at high activator levels (corresponding to an active chromatin state).

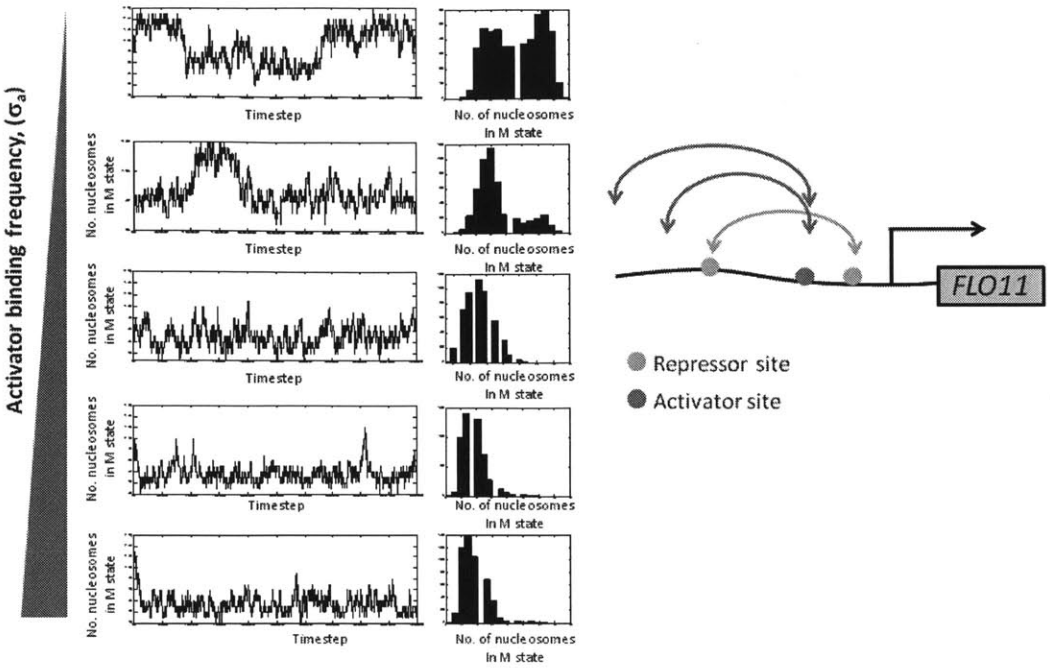

Figure 4: Model simulation with two repressor sites able to form a loop and one activator site able to form loops with upstream regions (configuration designed to mimic Msn1 activation in a Sfl1-silenced promoter background). Results show that at intermediate activator levels, the region can exhibit bistable chromatin states (stable repressive chromatin and stable active chromatin states).

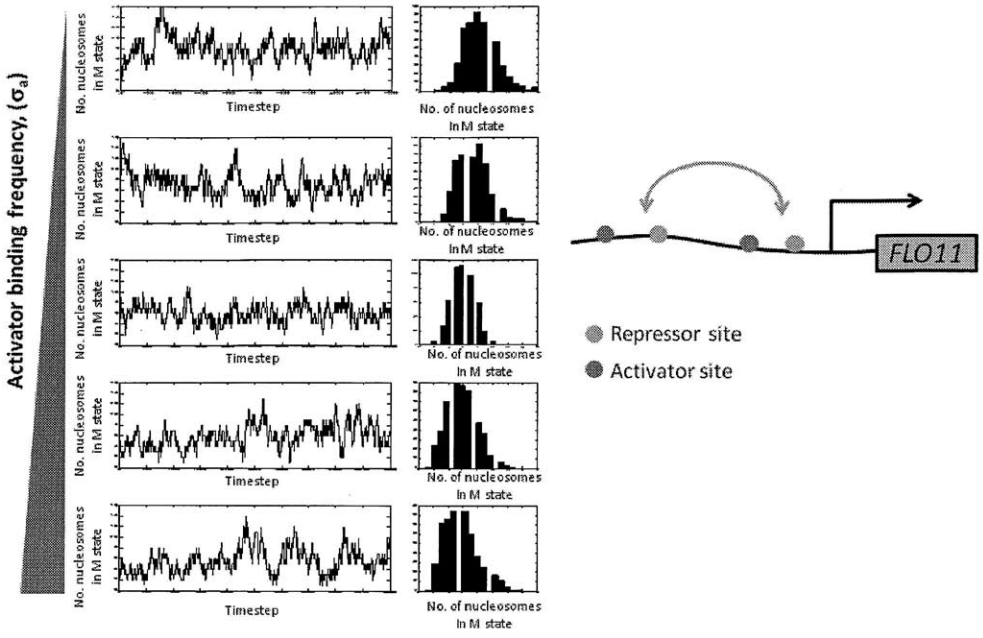

Figure 5: Model simulation with two repressor sites able to form a loop and two activator sites that do not form a loop. Results show that even if the activator binds at two sites, but does not form a loop, the region does not exhibit bistable chromatin states.

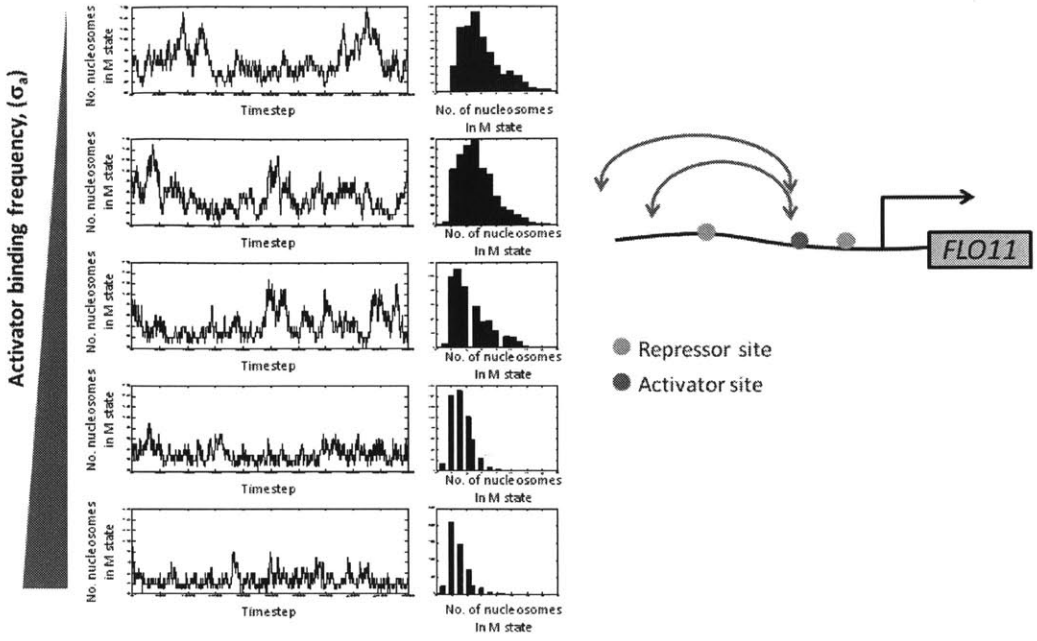

Figure 6: Model simulation with activator able to form loop with upstream regions and two repressor sites but no repressor loop is formed. Results indicate that if only activator looping occurred and repressor looping did not, the region does not exhibit bistable chromatin states.

ACTIVATOR TITRATIONS in *tup1Δ* and *ssn6Δ*

We have also created homozygous *tup1Δ* and *ssn6Δ* in our diploid two-color strain, and titrated various activators in these deletion backgrounds. We performed these activator titrations to determine whether Tup1p and Ssn6p might also be required for silencing at the *FLO11* promoter. The results were unusual; the expression profiles, which were heterogeneous, showed highly correlated CFP and YFP expression, compared to the expression profiles of the activator titrations in the WT background, which showed uncorrelated CFP and YFP expression. All titrations were done according to the procedure described in page 30.

Figures

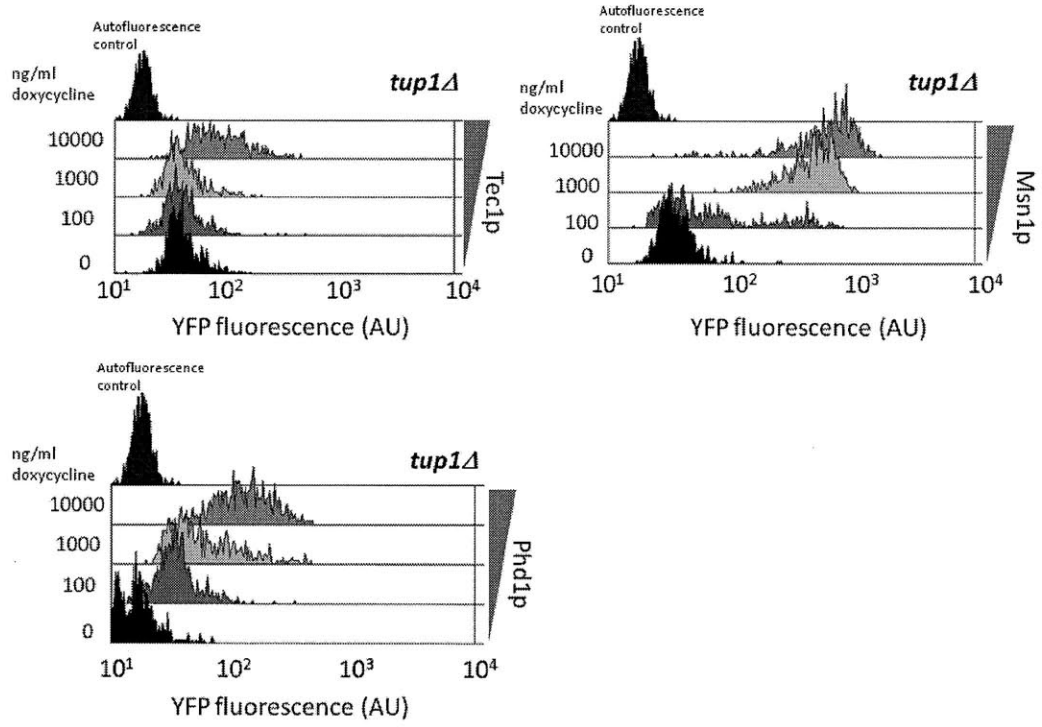

Figure 1: Titration of activators Tec1p, Phd1p and Msn1p in *tup1Δ*. All titrations were performed in SC –ura 2% glucose.

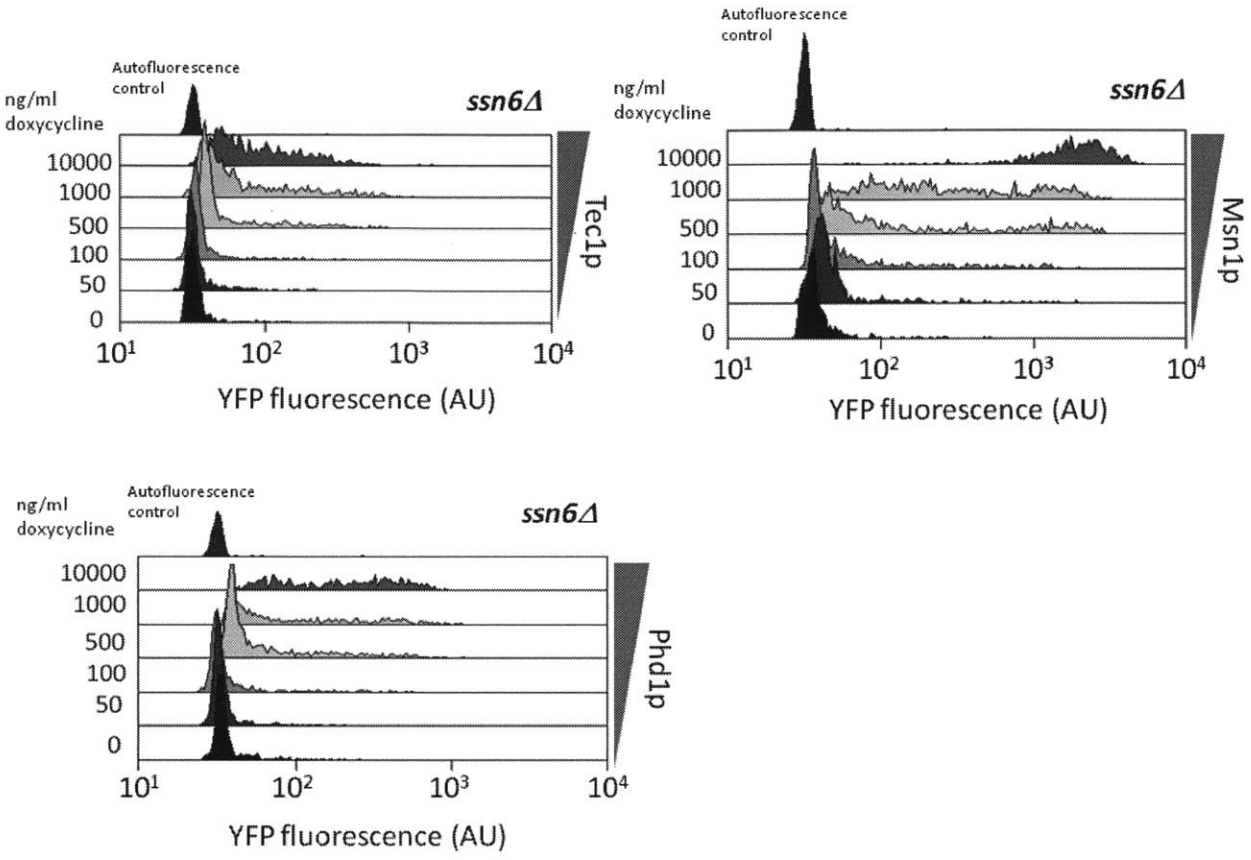

Figure 2: Titration of activators Tec1p, Phd1p and Msn1p in *ssn6Δ*. All titrations were performed in SC –ura 2% glucose.

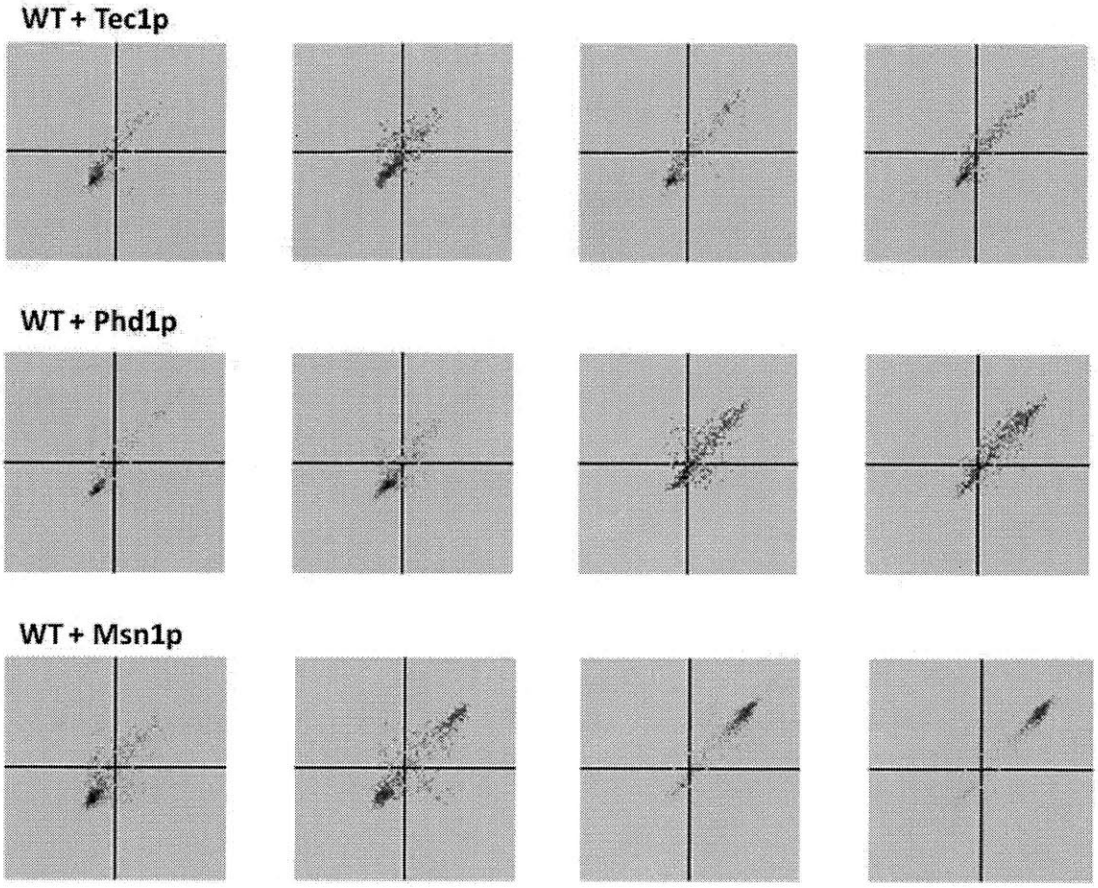

Figure 3: 2D histogram of YFP/CFP expression profiles of activator titrations. Activators Tec1p, Phd1p and Msn1p were titrated in the WT two-color diploid grown in SC –ura 2% glucose. The y-axis is YFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. The x-axis is CFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. CFP and YFP expression are generally uncorrelated.

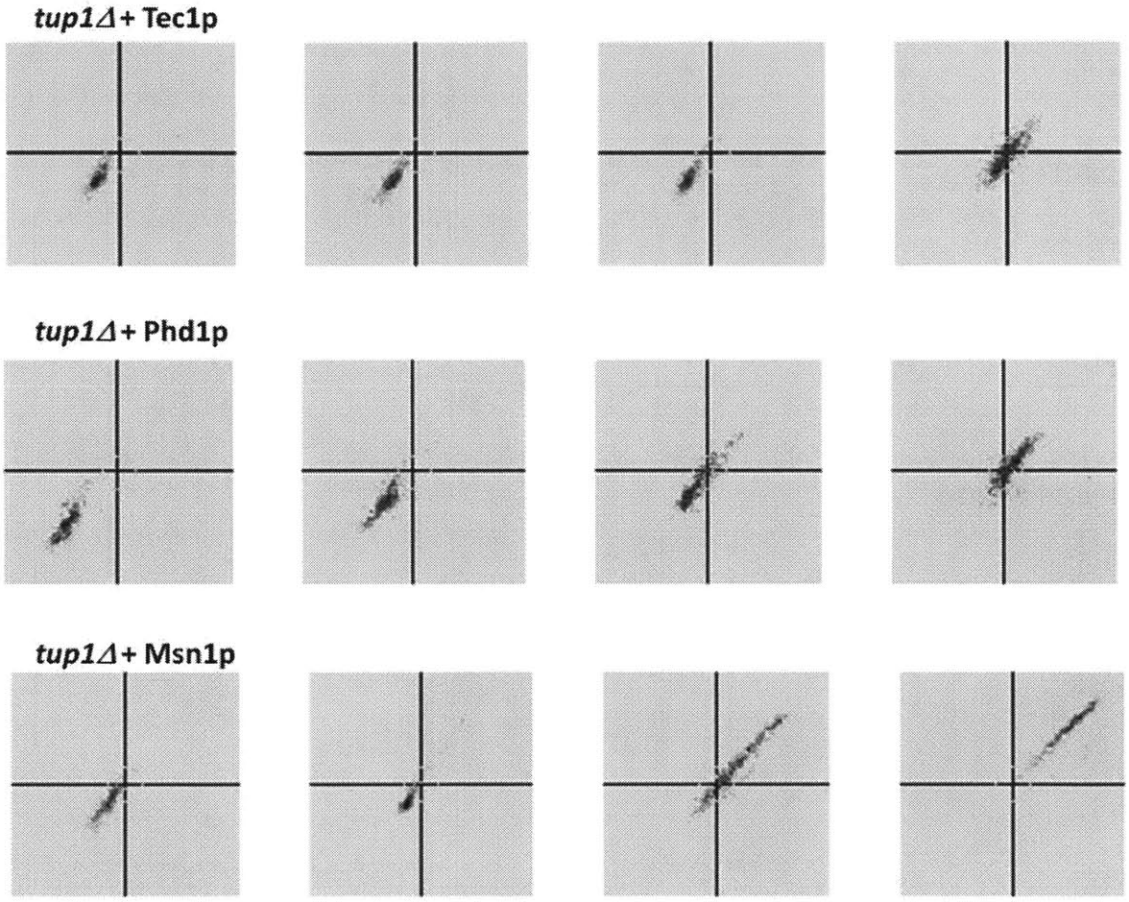

Figure 4: 2D histogram of YFP/CFP expression profiles of activator titrations. Activators Tec1p, Phd1p and Msn1p were titrated in the *tup1Δ* two-color diploid grown in SC –ura 2% glucose. The y-axis is YFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. The x-axis is CFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. CFP and YFP expression are highly correlated, compared to expression in WT background (Figure 3).

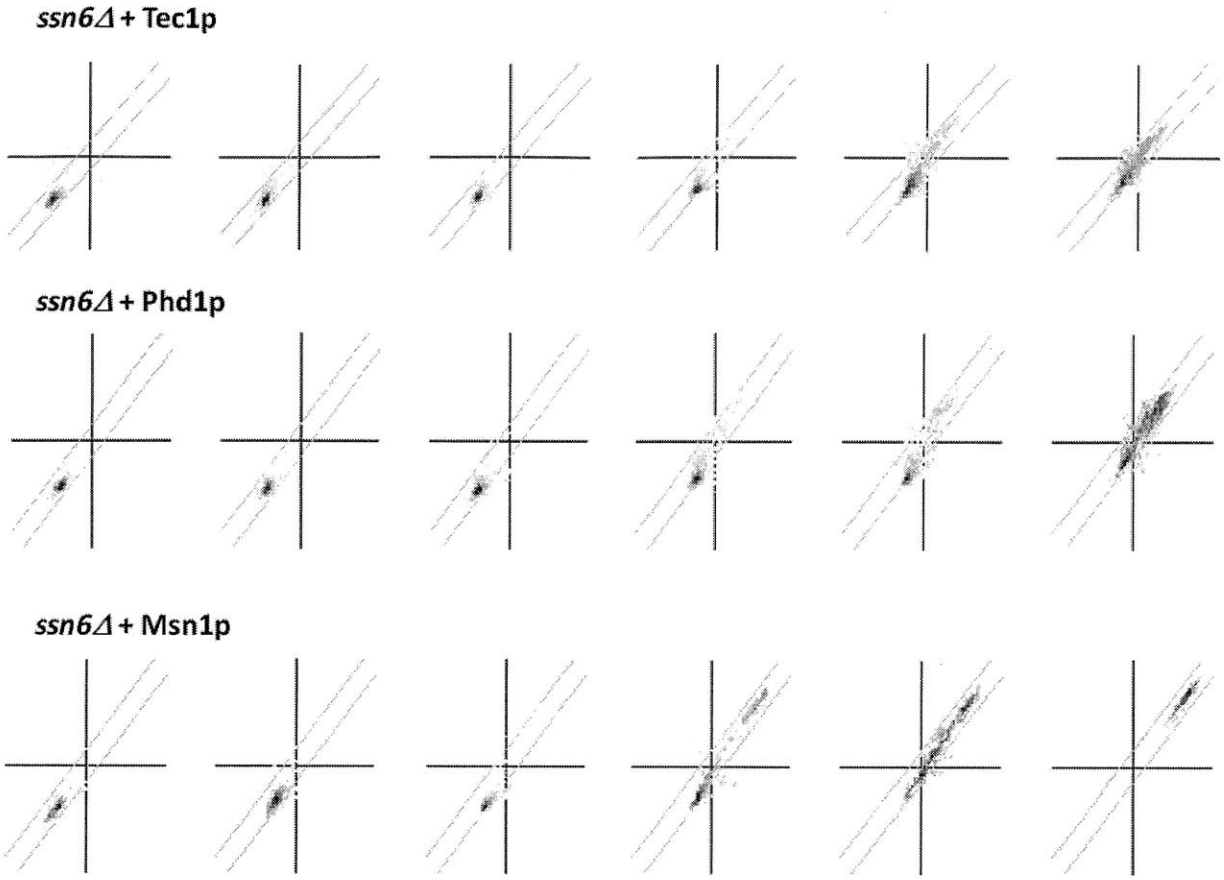

Figure 5: 2D histogram of YFP/CFP expression profiles of activator titrations. Activators Tec1p, Phd1p and Msn1p were titrated in the *ssn6Δ* two-color diploid grown in SC –ura 2% glucose. The y-axis is YFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. The x-axis is CFP expression (AU) in $\log_{10}$ scale, with the range from 10^0 to 10^4. CFP and YFP expression are highly correlated, compared to expression in WT background (Figure 3).