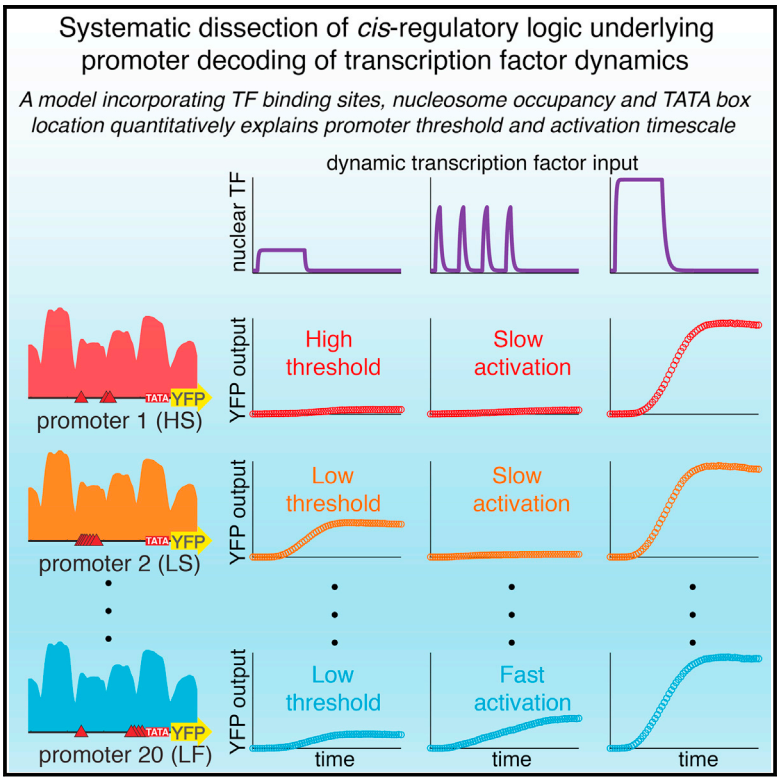
Report

Cell Reports

cis Determinants of Promoter Threshold and Activation Timescale

Graphical Abstract



Highlights

- *cis*-regulatory logic underlying threshold and activation timescale is elucidated
- Via modulation of TF binding sites, the promoter class can be switched
- A simple model quantitatively accounts for dynamic promoter behavior
- The promoter amplitude threshold can be decoupled from the activation timescale

Authors

Anders S. Hansen, Erin K. O'Shea

Correspondence

erin_oshea@harvard.edu

In Brief

Cells control gene expression in part by regulating the dynamics of transcription-factor activity. Hansen and O'Shea investigate the *cis*-regulatory logic underlying promoter decoding of transcription-factor dynamics for 20 promoter variants and reveal that two properties characterizing the gene expression response—threshold and activation timescale—can be decoupled.

Hansen & O'Shea, 2015, Cell Reports *12*, 1226–1233
August 25, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2015.07.035

OPEN ACCESS
CellPress

Cell Reports
Report

cis Determinants of Promoter Threshold and Activation Timescale

Anders S. Hansen[1,2,3,5] and Erin K. O'Shea[1,2,3,4,*]
[1]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA
[2]Howard Hughes Medical Institute
[3]Faculty of Arts and Sciences Center for Systems Biology
[4]Department of Molecular and Cellular Biology
Harvard University, Northwest Laboratory, 52 Oxford Street, Cambridge, MA 02138, USA
[5]Present address: Department of Molecular and Cell Biology, Howard Hughes Medical Institute, 475D Li Ka Shing Center, University of California, Berkeley, Berkeley, CA 94720, USA
*Correspondence: erin_oshea@harvard.edu
http://dx.doi.org/10.1016/j.celrep.2015.07.035

SUMMARY

Although the relationship between DNA *cis*-regulatory sequences and gene expression has been extensively studied at steady state, how *cis*-regulatory sequences affect the dynamics of gene induction is not known. The dynamics of gene induction can be described by the promoter activation timescale (AcTime) and amplitude threshold (AmpThr). Combining high-throughput microfluidics with quantitative time-lapse microscopy, we control the activation dynamics of the budding yeast transcription factor, Msn2, and reveal how *cis*-regulatory motifs in 20 promoter variants of the Msn2-target-gene *SIP18* affect AcTime and AmpThr. By modulating Msn2 binding sites, we can decouple AmpThr from AcTime and switch the *SIP18* promoter class from high AmpThr and slow AcTime to low AmpThr and either fast or slow AcTime. We present a model that quantitatively explains gene-induction dynamics on the basis of the Msn2-binding-site number, TATA box location, and promoter nucleosome organization. Overall, we elucidate the *cis*-regulatory logic underlying promoter decoding of TF dynamics.

INTRODUCTION

Regulation of gene expression is important for the ability of cells to maintain homeostasis and survive stress. The expression level of a gene depends on *cis*-regulatory motifs present in promoters that are interpreted by transcription factors (TFs), which control the rate of transcription (Levo and Segal, 2014; Rando and Winston, 2012). A major goal is to quantitatively understand and predict gene expression from knowledge of *cis*-regulatory DNA sequence and TF activity. Accordingly, how the promoter input-output relationship depends on the number and location of TF binding sites, nucleosome stability, and positioning in the promoter, the affinity of TF binding sites, and the presence of other *cis*-regulatory motifs have been extensively studied at steady state (Gertz et al., 2009; Lam et al., 2008; Mogno et al., 2013; Rajkumar et al., 2013; Raveh-Sadka et al., 2012; Sharon et al., 2012). Moreover, several studies have explored the relationship between promoter architecture and cell-to-cell variability in expression (noise) at steady state (Hornung et al., 2012; Sharon et al., 2014). However, a significant aspect of gene regulation occurs out of steady state: the kinetics of gene induction crucially determine how cells respond dynamically to signals and stresses, but how gene induction kinetics are influenced by regulatory DNA sequences is poorly understood. Along these lines, recent studies demonstrate that cells transmit gene expression information about external signals and stresses by regulating TF activation dynamics (Behar and Hoffmann, 2010; Castillo-Hair et al., 2015; Levine et al., 2013; Purvis and Lahav, 2013). Yet, despite this, the relationship between promoter *cis* elements and how the promoter responds to dynamical TF inputs has not been studied.

To investigate the relationship between the architecture of a promoter and how it decodes TF dynamics, we study the *SIP18* promoter, which is activated by the budding yeast TF, Msn2. During stress exposure, Msn2 encodes information about stress identity in its nuclear translocation dynamics—for example, Msn2 exhibits brief nuclear pulses with dose-dependent frequency under glucose limitation but a sustained pulse with dose-dependent amplitude under oxidative stress (Hao et al., 2013; Hao and O'Shea, 2012; Petrenko et al., 2013). Msn2 target genes can differentially decode Msn2 dynamics such that stress-relevant target genes are predominantly expressed under the relevant stress (Hansen and O'Shea, 2013, 2015; Hao and O'Shea, 2012). However, at a mechanistic level, we do not currently understand how promoters decode TF dynamics differently.

Conceptually, we can characterize the gene expression response with two parameters: (1) the amplitude threshold (AmpThr), which quantifies how sensitive the promoter is to the nuclear TF concentration, and (2) the activation timescale (AcTime), which quantifies gene induction kinetics, i.e., how long it takes the TF to activate the gene (Hansen and O'Shea, 2013). Based on these two variables, four extreme promoter

classes exist (high or low AmpThr with slow or fast AcTime, corresponding to the HS, HF, LS, and LF classes). To understand gene induction dynamics and why natural promoters decode the same dynamical TF signals differently, we must understand the mechanistic and *cis*-regulatory logic that determines AmpThr and AcTime.

Here, we combine high-throughput microfluidics and quantitative time-lapse microscopy to pharmacologically control Msn2 translocation dynamics and measure how 20 *SIP18* promoter variants decode Msn2 dynamics in single cells. We find that just three variables—the number of Msn2 binding sites, nucleosome occupancy over Msn2 binding sites, and their distance to the TATA box—are sufficient to quantitatively account for AmpThr and AcTime. Furthermore, we find that AmpThr and AcTime can be decoupled. Although the wild-type *SIP18* promoter (WT *pSIP18*) belongs to the HS class, by modulating the number and location of Msn2 binding sites, we can switch it to the LF or LS class. Additionally, we show that AcTime, but not AmpThr, determines the gene expression noise level. Finally, we propose a mechanistic model that plausibly explains promoter class from promoter *cis* elements and chromatin organization.

RESULTS

Models for Inferring Promoter AmpThr and AcTime

To systematically investigate how AmpThr and AcTime depend on promoter *cis* elements, we sought an efficient way of determining AmpThr and AcTime for a large number of promoter variants. Previously, to determine AmpThr and AcTime, we exposed an Msn2 target promoter of interest driving dual YFP/CFP reporter expression in diploid cells to a panel of 30 different Msn2 inputs, fit a mathematical model, and then calculated AmpThr and AcTime based on the fit (Supplemental Information; Hansen and O'Shea, 2013). Given that our microfluidic platform enables us to perform five experiments in parallel (Hansen et al., 2015), we sought inference models to quantitatively estimate AmpThr and AcTime from just five experiments in haploid cells. To obtain a training data set for the models, we exposed seven wild-type Msn2 target promoters for which we already know AmpThr and AcTime (Hansen and O'Shea, 2013) to five dynamical Msn2 inputs (the five *x*-conditions in Figure 1A; see Figure S1 for full data set). To understand why we chose these five experiments, consider the behavior of *SIP18*, an oxidative stress response gene (Rodríguez-Porrata et al., 2012), and of *HXK1*, a glucose limitation response gene (Herrero et al., 1995). *SIP18* has a high AmpThr and slow AcTime (HS class), whereas *HXK1* has a low AmpThr and fast AcTime (Hansen and O'Shea, 2013). Because the *SIP18* promoter activates slowly, it filters out brief nuclear Msn2 pulses (*x*(FM4) and *x*(FM8) in Figure 1A) observed during glucose limitation (Hao and O'Shea, 2012), whereas *HXK1* strongly induces under these conditions. Similarly, having a high AmpThr allows *SIP18* to filter out low-amplitude pulses (*x*(175 nM) in Figure 1A) and only induce during sustained Msn2 activation (*x*(3 μM) in Figure 1A) observed during oxidative stress (Hao and O'Shea, 2012). *HXK1*, on the other hand, has a low AmpThr and induces significantly during low-amplitude input (*x*(175 nM) in Figure 1A). Thus, YFP expression under *x*(175 nM), *x*(690 nM), and *x*(3 μM) allows us to infer AmpThr, whereas we can infer AcTime from YFP expression under *x*(690 nM), *x*(FM4), and *x*(FM8). Building on this intuition, we developed simple inference models to estimate AmpThr and AcTime with two and three fitted parameters, respectively (Figure 1B; Experimental Procedures). These simple models could account for >98% of the variance in AmpThr and AcTime (Figure 1B). However, a model with a sufficient number of fitted parameters can fit any data set. Overfitting is a particular concern because it reduces the predictive power of the model. To test whether the inference models were overfit, we use leave-one-out cross-validation (LOOCV): we leave out one promoter, fit the model to the remaining six, and use these fit parameters to predict AmpThr and AcTime for the promoter that was left out. We repeat this for all seven promoters. Even when we correct for overfitting using LOOCV, the inference models still account for >96% and >93% of the AmpThr and AcTime variances, respectively (Figure 1B). Thus, the inference models are not overfit and we can use them to calculate AmpThr and AcTime from just five experiments.

Systematic Dissection of How Msn2 STRE Location in the *SIP18* Promoter Affects AmpThr, AcTime, and Strength

Msn2 binds the stress-response element (STRE) (5′-CCCCT-3′) with sub-micromolar affinity and with limited flanking base preference (Siggers et al., 2014; Stewart-Ornstein et al., 2013). To systematically investigate how AmpThr and AcTime depend on Msn2 STRE number and location, we focus on the *SIP18* promoter, which contains three STREs. The "null mutant" without the two STREs at −386 and −367 bp between the −2 and −3 promoter nucleosomes cannot activate expression (Figure 1C), and we therefore consider the STRE at −524 bp to be non-functional on its own. We next developed a combinatorial promoter DNA synthesis method (Figure S2A) and divided the *SIP18* promoter into four regions: A, B, C, and D. We added 2, 3, or 4 STREs (A2-4, B2-4, C2-4, and D2-4 in Figure 1C) to each region, mimicking natural Msn2 target genes, which also contain STRE clusters: *HXK1*, for example, contains five clustered STREs (Figure S2C). We also made "scattered" mutants (S4; SA-SD). Whereas region A and C are in the accessible linker regions between nucleosomes, region B and D are located in sequences within strongly positioned nucleosomes (Figure 1C). We exposed strains containing each of these 16 promoter variants (chromosomally integrated at the *SIP18* locus and driving *sip18*::YFP) to the five dynamical Msn2 inputs (Figure 1A) and measured YFP expression (Figure S1; raw single-cell time-trace data are available as Supplemental Source Data). We used the models (Figure 1B) to infer AmpThr and AcTime for each promoter variant (Figure 1D) from the YFP measurements. Whereas WT *pSIP18* has a very high AmpThr and very slow AcTime, most promoter variants have lower AmpThr and faster AcTime (Figure 1D). We find that most binding-site changes cause incremental effects and the variants generally fall along the AmpThr/AcTime diagonal. Thus, AmpThr and AcTime appear to be coupled for these 16 mutants. We observe minor discrepancies: D3, for example, appears to be slightly faster than both D2 and D4 (Figure 1C). We attribute this to slight experimental or measurement error or to a minor effect on nucleosome organization.

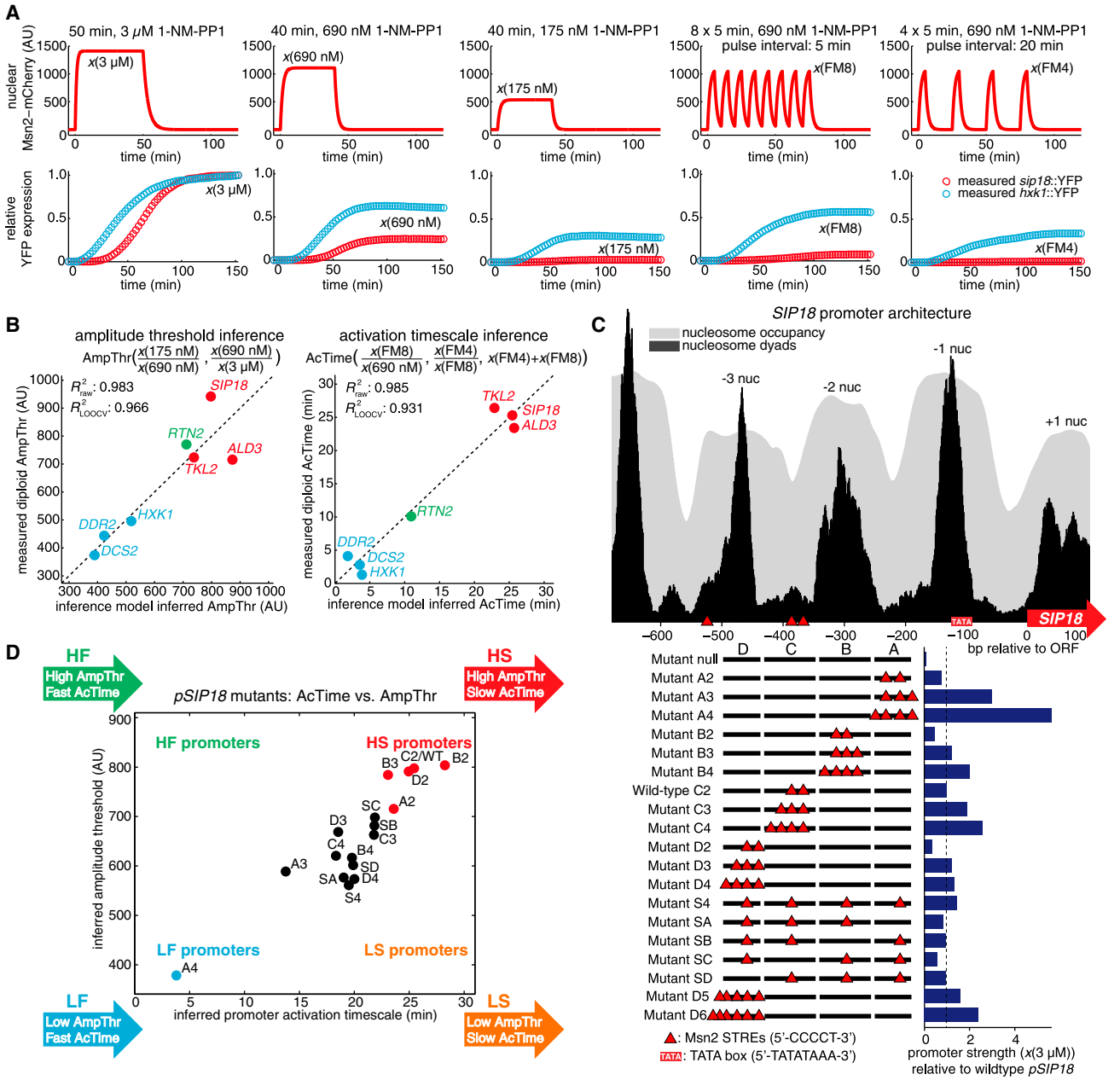


Figure 1. Inferring AmpThr and AcTime for *pSIP18* Promoter Variants

(A) Inferring promoter class from five experiments. Top panels: Msn2-mCherry nuclear translocation dynamics in five different dynamical treatments of 1-NM-PP1. 1-NM-PP1 inhibits PKA^{as} and causes Msn2 to translocate to the nucleus (Hansen et al., 2015). Bottom panels: corresponding *sip18*::YFP (red) and *hxk1*::YFP (blue) gene expression measured using time-lapse microscopy for each 1-NM-PP1 treatment are shown. From the YFP expression ratios, it is possible to estimate the promoter amplitude threshold (AmpThr) and activation timescale (AcTime).

(B) Models for inferring AmpThr and AcTime. For the seven wild-type promoters for which we know AmpThr and AcTime from measurements in diploid cells (Hansen and O'Shea, 2013), we performed the five experiments shown in (A) in haploid cells and fit to models defined in Experimental Procedures. These models account for most of the variance in AmpThr and AcTime even when corrected for overfitting using leave-one-out cross-validation (LOOCV).

(C) Overview of *pSIP18* promoter architecture. Nucleosome occupancy (gray) and nucleosome centers (dyads; black) are plotted using MNase-seq data from Hansen and O'Shea (2013). The wild-type promoter contains three Msn2 binding sites (STREs: 5′-CCCCT-3′) and a TATA box (Basehoar et al., 2004). The promoter was divided into four regions: A; B; C; and D. Promoter variants containing two, three, or four STREs in each region or "scattered" among the regions were constructed (Figure S2A). For each mutant, the experiments shown in (A) were performed and the promoter strength (response to 50 min; 3 μM 1-NM-PP1) relative to the wild-type *pSIP18* promoter is shown in blue bars.

(legend continued on next page)

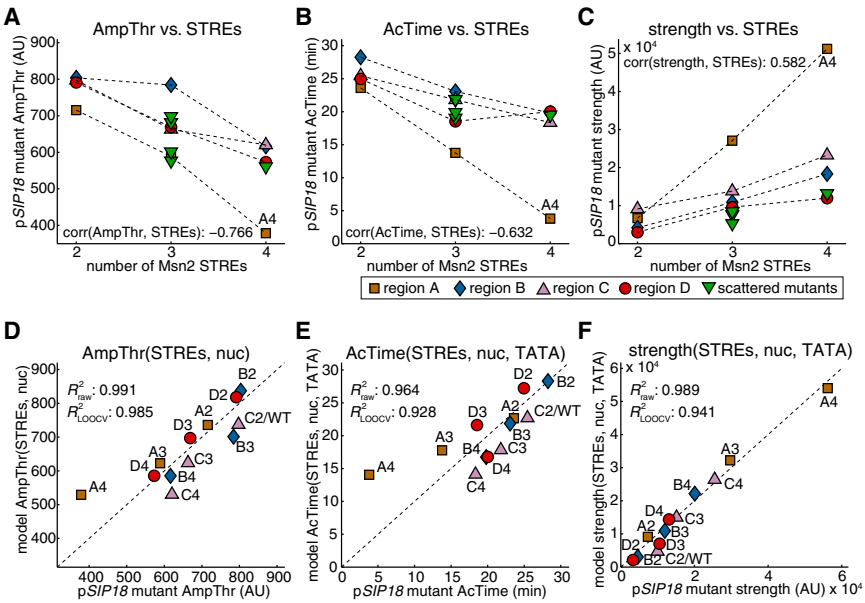


Figure 2. A Quantitative Relationship between Promoter *cis* Elements and Promoter Class

(A–C) The number of STREs alone cannot explain promoter class. For each *pSIP18* mutant, the AmpThr (A), AcTime (B), and expression strength (C) is plotted as a function of the number of STREs and the Pearson correlation coefficient (ρ) is shown. Each mutant is colored according to its STRE region (see legend).

(D–F) Simple models quantitatively account for variation in measured AmpThr, AcTime, and strength. Simple non-linear phenomenological models with four fitted parameters and three variables—the number of STREs (STREs), the nucleosome occupancy over the STREs (nuc) (Figure S2B), and the distance from the STREs to the TATA box (TATA)—can account for more than 90% of the variance in AmpThr, AcTime, and strength. Full details on the models and variables are given in Supplemental Information.

See also Figure S2B.

Although previous studies have shown that changing the location of the TF binding site does not measurably affect promoter nucleosome positioning (Lam et al., 2008) and that neither does replacing the *SIP18* ORF with a YFP reporter (Hansen and O'Shea, 2013), we cannot fully exclude that minor differences in nucleosome positioning may account for some of the differences observed. Mutant A4 shows the biggest change: A4 entirely switches to the LF class and has a very low AmpThr and a fast AcTime (Figure 1D). Furthermore, A4 shows ∼5-fold higher strength than WT *pSIP18* (defined as absolute YFP level under *x*(3 μM); Figure 1C). Thus, both AmpThr and AcTime are tunable in *cis*.

Three Variables—STRE Number, STRE Distance from TATA Box, and Nucleosome Occupancy of STREs—Suffice to Quantitatively Account for AmpThr, AcTime, and Strength

Next, we sought to mechanistically understand how AmpThr, AcTime, and promoter strength are determined. In the simplest scenario, the number of STREs could simply determine these variables. However, although AmpThr (Figure 2A), AcTime (Figure 2B), and promoter strength (Figure 2C) generally show a monotonic relationship with the number of STREs, the number of STREs alone cannot fully explain this relationship. For example, among the mutants with four STREs, B4, C4, D4, and S4 have similar AcTime, but A4 stands out with a much-lower AcTime (Figure 2A).

There could be several reasons why STRE number alone fails to account for promoter behavior. For example, nucleosomes restrict TF binding and nucleosome remodeling may be required for initiation of transcription (Lam et al., 2008; Rando and Winston, 2012; Raveh-Sadka et al., 2012; Weiner et al., 2012). Furthermore, for some promoters, removal of the nucleosome occupying the TATA box can be sufficient to activate transcription (Adkins and Tyler, 2006; Zhang and Reese, 2007). We therefore hypothesized that, together, the number of STREs (STREs), the level of nucleosome occupancy over the STREs (nuc; Figure S2B), and the average distance from the STREs to the TATA box (TATA) might account for the observed spread in AmpThr, AcTime, and strength. To test this, we developed simple non-linear phenomenological models relating AmpThr (Figure 2D), AcTime (Figure 2E), and strength (Figure 2F) to these three variables (Experimental Procedures). Even when correcting for overfitting using LOOCV, the simple models could account for >98%, >92%, and >94% of the variance in AmpThr, AcTime, and strength, respectively (Figures 2D–2F). In fact, just two variables—STREs and nuc—were sufficient to largely account for AmpThr. Although slight discrepancies were observed (e.g., the AmpThr model overestimates the A4 AmpThr; Figure 2D), when taken together, these results demonstrate that three mechanistic variables suffice to quantitatively account for >90% of the variance in AmpThr, AcTime, and strength.

To assess whether the mechanistic insight obtained for the *pSIP18* mutant promoters generalizes to other promoter backgrounds, we applied the AmpThr and AcTime models to the six other wild-type promoters for which we already know AmpThr and AcTime (Figure 1B). The models could accurately predict AmpThr (R^2 = 0.945), but the prediction of AcTime was associated with higher error (R^2 = 0.710; Figure S2D)—this may be

(D) AmpThr versus AcTime for *pSIP18* mutants. For each mutant in (C), YFP expression was measured for each Msn2 input in (A) and the models in (B) applied to estimate the AmpThr and AcTime for each mutant. Mutants have been colored based on the promoter classification, and intermediate promoters are shown in black.

See also Figures S1 and S2A. Raw time-trace data for all strains in Figure 1 (66,088 single cells) are available as Supplemental Source Data.

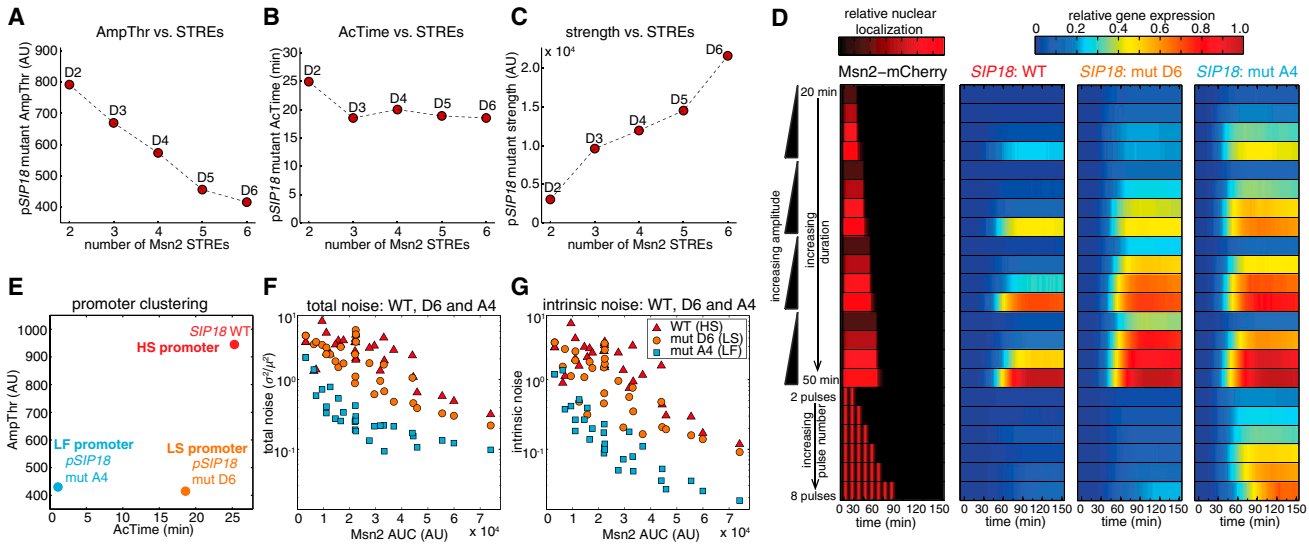


Figure 3. Decoupling Promoter AmpThr from AcTime: Mutant D6 Is an LS Promoter

(A–C) Decoupling AmpThr from AcTime in region D. AmpThr (A), AcTime (B), and strength (C) are plotted as a function of the number of STREs in region D. Mutant D6 appears to belong to the LS class.

(D) Systematic experimental dissection of how WT *pSIP18*, D6, and A4 decode Msn2 dynamics. Each row corresponds to a specific Msn2-mCherry input (left, in red), and the corresponding gene expression response for each of the three promoters is shown in the corresponding rows on the right. The gene expression responses for each promoter are internally normalized to their maximal expression level. Each row is the per-cell average of a few hundred cells from at least two replicates. WT data are from Hansen and O'Shea (2013). Data (15,875 single cells in total) for all 30 experiments for A4 and D6 are shown in Figure S3.

(E) Clustering of promoters. The full 30-experiment data set for A4 and D6 was fit to a previously described differential equations model (Hansen and O'Shea, 2013) and best-fit parameters inferred. Numerically, AmpThr is defined as the nuclear Msn2-mCherry level (AU) required to reach the half-maximal level of promoter activity obtained at 3 μM 1-NM-PP1 (the maximal nuclear Msn2-mCherry level). AcTime is defined as the time (min) it takes to reach the half-maximal level of promoter activity reached at steady state at 690 nM 1-NM-PP1. Both AmpThr and AcTime were obtained from model simulations. Full details are given in Supplemental Information.

(F and G) Total noise (σ^2/μ^2; F) and intrinsic (G) noise are plotted against the Msn2 AUC ($\int_0^\infty [\mathrm{Msn2}(t)]\mathrm{d}t$), and the experiments are colored based on promoter class: WT (HS, red); D6 (LS, orange); and A4 (LF, blue). Each dot corresponds to the noise (mean across time points after gene expression has reached a plateau) for each of the 30 experiments performed in (D).

See also Figure S3 and Supplemental Source Data for all raw A4 and D6 data.

a result of differences in chromatin background (see Supplemental Information for a detailed discussion). Overall, this analysis indicates that the *cis*-regulatory logic underlying AmpThr and AcTime for *pSIP18* may generalize to other eukaryotic promoters.

Decoupling Promoter AmpThr from AcTime

Even though we investigated a series of quite different mutant and wild-type promoters, all promoters roughly fall along the AmpThr/AcTime diagonal (Figures 1B and 1D). This would seem to suggest that AmpThr and AcTime cannot be decoupled: when AmpThr is lowered, a corresponding decrease in AcTime is also observed and vice versa. Nevertheless, for the D region mutants, we noticed that increasing the number of STREs lowers AmpThr without having a strong effect on AcTime (Figures 2A and 2B). We therefore hypothesized that adding even more than four STREs should yield mutants with a low AmpThr but relatively slow AcTime (LS class). To test this, we made mutants D5 and D6 (Figure 1C), repeated the experiments, and calculated AmpThr and AcTime using the inference models (Figure 1B). Indeed, mutant D5 and D6 show low AmpThr (Figure 3A) while remaining slow (Figure 3B). We see a monotonic relationship between STRE number and strength (Figure 3C). Overall, the D5 and D6 results suggest that it is possible to decouple AmpThr from AcTime and furthermore indicate that the LS promoter class also exists.

Mutants A4 and D6 Switch Promoter Class

Based on the five haploid experiments and the inference model estimate of AmpThr and AcTime (Figures 1D, 3A, and 3B), mutants A4 and D6 switch promoter class to LF and LS, respectively. To verify our inference approach and confirm these results, we made diploid strains with dual *sip18*::CFP and *sip18*::YFP reporters on the homologous chromosomes driven by the A4 and D6 promoters. Having dual CFP/YFP reporters allows us to also study gene-expression noise (Elowitz et al., 2002). We then exposed each diploid mutant to 30 experiments systematically varying Msn2 pulse duration, amplitude, pulse number, and interval (Figures 3D and S3), fit a previously described model (Hansen and O'Shea, 2013), and calculated the actual AmpThr and AcTime (Supplemental Information). Indeed, these results confirm that A4 is an LF promoter and D6 an LS promoter (Figure 3E). Having a slow AcTime, both WT *pSIP18* and D6 filter out oscillatory and short-duration Msn2 input, whereas A4 with a fast AcTime responds strongly (Figure 3D). Similarly, both A4 and D6 have a low AmpThr and

therefore activate strongly to low levels of Msn2 activation, whereas WT *pSIP18* largely filters out low-amplitude input (Figure 3D). Taken together, these results confirm that A4 and D6 completely switch the promoter class.

Gene-Expression Noise Level Is Affected by AcTime Not AmpThr

The relationship between gene-expression noise level (σ^2/μ^2) and wild-type promoter variants (Bar-Even et al., 2006; Newman et al., 2006) or synthetic promoter variants (Hornung et al., 2012; Sharon et al., 2014) has been extensively studied at steady state, but it is not clear how noise depends on AmpThr and AcTime. Previously, we observed a high negative correlation between noise and AcTime. Based on this, we argued that noise should strongly depend on AcTime—such that slow promoters show dramatically higher noise in gene expression—but that AmpThr should not strongly affect noise (Hansen and O'Shea, 2013). To causally test this, it is necessary to compare noise levels for promoters where only either AmpThr or AcTime are changed. Comparing WT *pSIP18* and D6, AmpThr changes from low to high without much of a change in AcTime (Figure 3E). Likewise, comparing D6 and A4, AcTime changes from slow to fast without much of a change in AmpThr. For each experiment in Figures 3D and S3, we calculate the total (Figure 3F) and intrinsic (Figure 3G) noise (Elowitz et al., 2002) and plot the noise as a function of the Msn2 area under the curve (Msn2 AUC) (time-integrated Msn2 activation) such that each dot in Figures 3F and 3G corresponds to a single Msn2 input for a single promoter. We find that, whereas WT *pSIP18* and mutant D6 exhibit high total and intrinsic noise, mutant A4 shows lower noise (Figures 3F and 3G). Because mutants A4 and D6 differ only by AcTime, this experimentally demonstrates that the noise level is strongly affected by AcTime, but not much affected by AmpThr.

DISCUSSION

Recently, it has become clear that cells transmit information and control cell fate by regulating the dynamics of master TFs (Levine et al., 2013; Purvis and Lahav, 2013). For example, in murine neural progenitor cells, control of TF Asc1 dynamics is sufficient to control cell fate: oscillatory Asc1 activity leads to cell proliferation, whereas sustained Asc1 activity causes neuronal differentiation (Imayoshi et al., 2013). The mechanism is believed to be due to different Asc1 dynamics inducing different downstream gene-expression programs, which requires that target genes show different induction kinetics. The same dynamical signaling logic appears to hold for the budding yeast TF, Msn2 (Hao and O'Shea, 2012). However, how promoter *cis* elements affect promoter decoding of TF dynamics was not understood.

Here, we systematically investigate how STRE number and location in the *SIP18* promoter affects gene-induction dynamics. We find that AmpThr and AcTime can be decoupled and that just three variables—number of STREs, their accessibility, and their distance from the TATA box—suffice to quantitatively explain more than 90% of the variance in AmpThr and AcTime. Whereas the strong dependence of AmpThr on the number of STREs and nucleosome occupancy over the STREs could perhaps be expected from previous steady-state studies (Sharon et al., 2012), the strong dependence of AcTime on the distance from the STREs to the TATA box is surprising.

What is the cause of this relationship? And, mechanistically, how can we explain why A4 and D6 fall into the LF and LS classes? If we assume that (1) remodeling of the nucleosome occupying the TATA box is required for transcription, which is well supported (Adkins and Tyler, 2006; Zhang and Reese, 2007) and (2) remodeling of nucleosomes adjacent to Msn2 binding is fast, but remodeling of nucleosomes distal to Msn2 binding is slow, which is consistent with local recruitment of chromatin remodelers by TFs (Larschan and Winston, 2001; Weake and Workman, 2010), a mechanistic model emerges that explains promoter class from promoter architecture (Figure 4). Wild-type SIP18 promoter (HS class) is slow because, although remodeling of the −2 and −3 nucleosomes adjacent to where Msn2 binds is fast, remodeling of the distal −1 nucleosome occupying the TATA box is slow and remodeling of the −1 nucleosome is required for activation (Figure 4A). Similarly, the additional STREs in mutant D6 (LS class) greatly lowers the AmpThr, but D6 remains slow because Msn2 binds too far upstream of the −1 nucleosome to rapidly remodel it (Figure 4B). Conversely, mutant A4 has both a low AmpThr due to its clustered STREs in a nucleosome-free region and a fast AcTime because Msn2 binds adjacent to the TATA box (Figure 4C). Thus, remodeling the −1 nucleosome is rapid for A4, which therefore belongs to the LF class.

Although this mechanistic model (Figure 4) plausibly explains promoter class and induction dynamics from promoter architecture, it is a simplification of the complex sequence of events taking place during gene activation (Hager et al., 2009). Nevertheless, a direct prediction of the model is that the −2 and −3 nucleosomes should be remodeled faster than the −1 nucleosome occupying the TATA box for the wild-type *SIP18* promoter, and this is supported by nucleosome remodeling time course data (Figure S4).

Recent advances in DNA synthesis now make it possible through massively parallel approaches such as FACS-seq to study thousands of promoters (Noderer et al., 2014; Sharon et al., 2012). Through such approaches, it will be interesting to investigate the extent to which the *cis*-regulatory logic underlying promoter decoding of TF dynamics we uncovered for the *SIP18* promoter generalizes to other Msn2-regulated promoters and to promoters regulated by other TFs.

EXPERIMENTAL PROCEDURES

Strains

All *Saccharomyces cerevisiae* strains used in this work are in the W303 background. The combinatorial promoter synthesis method is illustrated in Figure S2A, a list of strains is given in Table S1, and further information about how they were constructed is given in Supplemental Information.

Microfluidics, Time-Lapse Microscopy, and Image Analysis

Briefly, yeast cells were grown overnight at 30°C with shaking at 180 rpm to an $OD_{600\ nm}$ of ~0.1 in low fluorescence medium (which exhibits minimal autofluorescence), quickly collected by suction filtration, loaded into the five channels of a microfluidic device pretreated with concanavalin A to

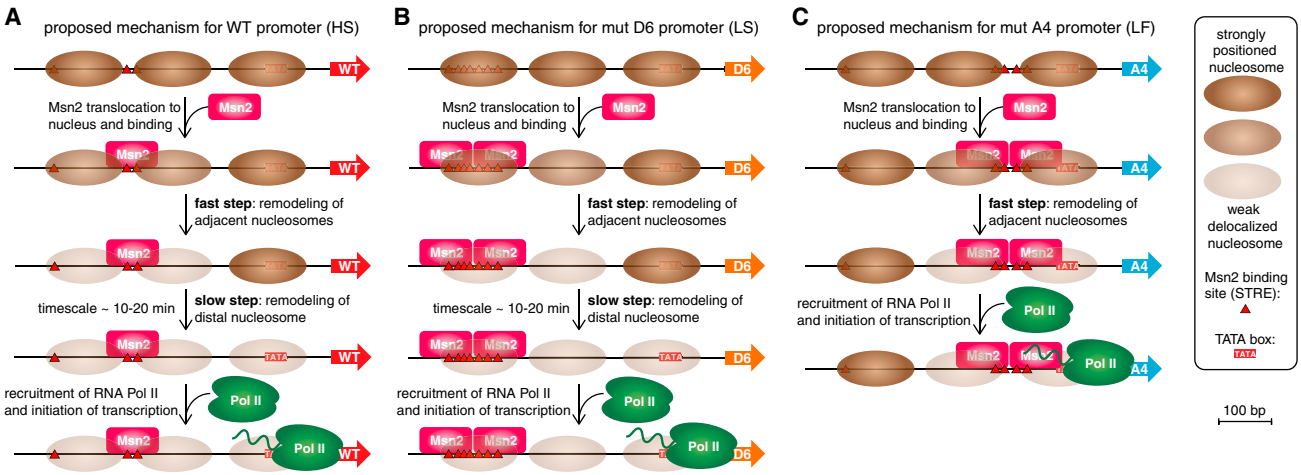


Figure 4. A Mechanistic Model Can Explain Promoter Class from *cis* Elements

A simplified, mechanistic model that assumes that (1) remodeling of the nucleosome occupying the TATA box is required for gene induction and (2) that nucleosome remodeling adjacent to where Msn2 binds is fast but nucleosome remodeling distal to Msn2 binding is slow can explain observed difference in promoter class. Promoter architecture and nucleosome sizes are drawn to scale.

(A) Model for wild-type promoter (HS). Msn2 binds to non-clustered STREs with low affinity and rapidly remodels adjacent nucleosomes. After slow remodeling of distal nucleosome, the TATA box is available and initiation of transcription occurs.

(B) Model for D6 promoter (LS). Msn2 binds to clustered STREs with high affinity and rapidly remodels adjacent nucleosomes. After slow remodeling of distal nucleosome, the TATA box is available and initiation of transcription occurs.

(C) Model for A4 promoter (LF). Msn2 binds to clustered STREs with high affinity and rapidly remodels adjacent nucleosomes. Because the TATA box is now available, initiation of transcription immediately occurs.

See also Figure S4 for MNase-seq time course data.

retain cells, and the setup mounted on an inverted fluorescence microscope kept at 30°C. Programmable solenoid valves deliver the 1-NM-PP1 pulse treatments shown in Figure 1A to each microfluidic channel. The microscope automatically maintains focus and acquires phase-contrast, YFP, CFP, RFP, and iRFP images with a 63×/1.4 NA objective from each of five microfluidic channels for 64 frames with a 2.5-min time resolution. Approximately 100–200 cells were imaged in each field of view. The resulting movies were analyzed using custom written segmentation and tracking software (MATLAB), and gene expression is quantified as the average YFP concentration per cell. A full protocol describing microfluidics, microscopy, image analysis, and control software (MATLAB) for solenoid valves is given elsewhere (Hansen et al., 2015; Hansen and O'Shea, 2013). Raw time-trace data for all single cells (Figures 1 and 3) are available as Supplemental Source Data.

Inference Models for Inferring AmpThr and AcTime

The AmpThr and AcTime inference models are given by

$$\mathrm{AmpThr}(\mathrm{STREs},\mathrm{nuc}) = c_1 \cdot e^{c_2(\mathrm{STREs}-1)} \cdot e^{c_3 \cdot \mathrm{nuc}}$$

$$\mathrm{AcTime}(\mathrm{STREs},\mathrm{TATA},\mathrm{nuc}) = c_1 \cdot e^{c_2(\mathrm{STREs}-1)} \cdot e^{c_3 \cdot \mathrm{TATA}} \cdot e^{c_4 \cdot \mathrm{nuc}}$$

$$\mathrm{Strength}(\mathrm{STREs},\mathrm{TATA},\mathrm{nuc}) = c_1 \left(\frac{(\mathrm{STREs}-1)}{\mathrm{TATA}^{c_2} \cdot \mathrm{nuc}^{c_3}} \right)^{c_4}$$

Further details on how the variables are quantitatively defined, values for best-fit parameters, and the models are given in Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, four figures, one table, and supplemental source data and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.07.035.

$$\mathrm{AmpThr}\left(\frac{x(175\ \mathrm{nM})}{x(690\ \mathrm{nM})}, \frac{x(690\ \mathrm{nM})}{x(3\ \mu\mathrm{M})}\right) = k_1 f\left(\frac{x(175\ \mathrm{nM})}{x(690\ \mathrm{nM})}\right) + k_2 f\left(\frac{x(690\ \mathrm{nM})}{x(3\ \mu\mathrm{M})}\right)$$

$$\mathrm{AcTime}\left(\frac{x(\mathrm{FM8})}{x(690\ \mathrm{nM})}, \frac{x(\mathrm{FM4})}{x(\mathrm{FM8})}, x(\mathrm{FM4}) + x(\mathrm{FM8})\right) = k_1 f\left(\frac{x(\mathrm{FM8})}{x(690\ \mathrm{nM})}\right) + k_2 f\left(\frac{x(\mathrm{FM4})}{x(\mathrm{FM8})}\right) + k_3 f(x(\mathrm{FM4}) + x(\mathrm{FM8}))$$

The x variables are defined in Figure 1A, and *f* is a variant of the Gumbel CDF. Full details and best-fit parameters are given in Supplemental Information.

Non-linear Phenomenological Models for AmpThr, AcTime, and Promoter Strength

The phenomenological models for AmpThr, AcTime, and strength take the variables STREs (number of STREs), nuc (defined in Figure S2B), and TATA (mean distance from STREs to TATA box) as input and are given by

ACKNOWLEDGMENTS

We thank Nan Hao and members of the E.K.O. lab for discussions and critically reading the manuscript. This work was performed in part at the Center for Nanoscale Systems at Harvard University, a member of the National Nanotechnology Infrastructure Network (NNIN), which is supported by the National Science Foundation under NSF award no. ECS-0335765. Image analysis and model simulations were run on the Odyssey cluster supported by the FAS

Division of Science, Research Computing Group at Harvard University. The Howard Hughes Medical Institute supported this work.

Received: May 24, 2015
Revised: July 8, 2015
Accepted: July 15, 2015
Published: August 13, 2015

REFERENCES

Adkins, M.W., and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. Mol. Cell *21*, 405–416.

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. *38*, 636–643.

Basehoar, A.D., Zanton, S.J., and Pugh, B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. Cell *116*, 699–709.

Behar, M., and Hoffmann, A. (2010). Understanding the temporal codes of intra-cellular signals. Curr. Opin. Genet. Dev. *20*, 684–693.

Castillo-Hair, S.M., Igoshin, O.A., and Tabor, J.J. (2015). How to train your microbe: methods for dynamically characterizing gene networks. Curr. Opin. Microbiol. *24*, 113–123.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.

Gertz, J., Siggia, E.D., and Cohen, B.A. (2009). Analysis of combinatorial cis-regulation in synthetic and genomic promoters. Nature *457*, 215–218.

Hager, G.L., McNally, J.G., and Misteli, T. (2009). Transcription dynamics. Mol. Cell *35*, 741–753.

Hansen, A.S., and O'Shea, E.K. (2013). Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. Mol. Syst. Biol. *9*, 704. http://dx.doi.org/10.1038/msb.2013.56.

Hansen, A.S., and O'Shea, E.K. (2015). Limits on information transduction through amplitude and frequency regulation of transcription factor activity. eLife *4*.

Hansen, A.S., Hao, N., and O'Shea, E.K. (2015). High-throughput microfluidics to control and measure signaling dynamics in single yeast cells. Nat. Protoc. *10*, 1181–1197. http://dx.doi.org/10.7554/eLife.06559.

Hao, N., and O'Shea, E.K. (2012). Signal-dependent dynamics of transcription factor translocation controls gene expression. Nat. Struct. Mol. Biol. *19*, 31–39.

Hao, N., Budnik, B.A., Gunawardena, J., and O'Shea, E.K. (2013). Tunable signal processing through modular control of transcription factor translocation. Science *339*, 460–464.

Herrero, P., Galíndez, J., Ruiz, N., Martínez-Campa, C., and Moreno, F. (1995). Transcriptional regulation of the Saccharomyces cerevisiae HXK1, HXK2 and GLK1 genes. Yeast *11*, 137–144.

Hornung, G., Bar-Ziv, R., Rosin, D., Tokuriki, N., Tawfik, D.S., Oren, M., and Barkai, N. (2012). Noise-mean relationship in mutated promoters. Genome Res. *22*, 2409–2417.

Imayoshi, I., Isomura, A., Harima, Y., Kawaguchi, K., Kori, H., Miyachi, H., Fujiwara, T., Ishidate, F., and Kageyama, R. (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science *342*, 1203–1208.

Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. Nature *453*, 246–250.

Larschan, E., and Winston, F. (2001). The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. Genes Dev. *15*, 1946–1956.

Levine, J.H., Lin, Y., and Elowitz, M.B. (2013). Functional roles of pulsing in genetic circuits. Science *342*, 1193–1200.

Levo, M., and Segal, E. (2014). In pursuit of design principles of regulatory sequences. Nat. Rev. Genet. *15*, 453–468.

Mogno, I., Kwasnieski, J.C., and Cohen, B.A. (2013). Massively parallel synthetic promoter assays reveal the in vivo effects of binding site variants. Genome Res. *23*, 1908–1915.

Newman, J.R.S., 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 *441*, 840–846.

Noderer, W.L., Flockhart, R.J., Bhaduri, A., Diaz de Arce, A.J., Zhang, J., Khavari, P.A., and Wang, C.L. (2014). Quantitative analysis of mammalian translation initiation sites by FACS-seq. Mol. Syst. Biol. *10*, 748. http://dx.doi.org/10.15252/msb.20145136.

Petrenko, N., Chereji, R.V., McClean, M.N., Morozov, A.V., and Broach, J.R. (2013). Noise and interlocking signaling pathways promote distinct transcription factor dynamics in response to different stresses. Mol. Biol. Cell *24*, 2045–2057.

Purvis, J.E., and Lahav, G. (2013). Encoding and decoding cellular information through signaling dynamics. Cell *152*, 945–956.

Rajkumar, A.S., Dénervaud, N., and Maerkl, S.J. (2013). Mapping the fine structure of a eukaryotic promoter input-output function. Nat. Genet. *45*, 1207–1215.

Rando, O.J., and Winston, F. (2012). Chromatin and transcription in yeast. Genetics *190*, 351–387.

Raveh-Sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., Zeevi, D., Sharon, E., Weinberger, A., and Segal, E. (2012). Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. Nat. Genet. *44*, 743–750.

Rodríguez-Porrata, B., Carmona-Gutierrez, D., Reisenbichler, A., Bauer, M., Lopez, G., Escoté, X., Mas, A., Madeo, F., and Cordero-Otero, R. (2012). Sip18 hydrophilin prevents yeast cell death during desiccation stress. J. Appl. Microbiol. *112*, 512–525.

Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and Segal, E. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. Nat. Biotechnol. *30*, 521–530.

Sharon, E., van Dijk, D., Kalma, Y., Keren, L., Manor, O., Yakhini, Z., and Segal, E. (2014). Probing the effect of promoters on noise in gene expression using thousands of designed sequences. Genome Res. *24*, 1698–1706.

Siggers, T., Reddy, J., Barron, B., and Bulyk, M.L. (2014). Diversification of transcription factor paralogs via noncanonical modularity in C2H2 zinc finger DNA binding. Mol. Cell *55*, 640–648.

Stewart-Ornstein, J., Nelson, C., DeRisi, J., Weissman, J.S., and El-Samad, H. (2013). Msn2 coordinates a stoichiometric gene expression program. Curr. Biol. *23*, 2336–2345.

Weake, V.M., and Workman, J.L. (2010). Inducible gene expression: diverse regulatory mechanisms. Nat. Rev. Genet. *11*, 426–437.

Weiner, A., Chen, H.V., Liu, C.L., Rahat, A., Klien, A., Soares, L., Gudipati, M., Pfeffner, J., Regev, A., Buratowski, S., et al. (2012). Systematic dissection of roles for chromatin regulators in a yeast stress response. PLoS Biol. *10*, e1001369.

Zhang, H., and Reese, J.C. (2007). Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3. Proc. Natl. Acad. Sci. USA *104*, 8833–8838.