



A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome

Bryan J. Venters and B. Franklin Pugh

Genome Res. 2009 19: 360-371 originally published online January 5, 2009

Access the most recent version at doi:[10.1101/gr.084970.108](https://doi.org/10.1101/gr.084970.108)

References This article cites 87 articles, 35 of which can be accessed free at:
<http://genome.cshlp.org/content/19/3/360.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." in white. In the center, there is a white-bordered box containing the text "LEARN MORE". On the right, there is a photograph of a woman wearing a red and white superhero cape and mask, with a green molecular structure logo above the word "CELLECTA" in white capital letters.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © 2009 by Cold Spring Harbor Laboratory Press

A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome

Bryan J. Venters and B. Franklin Pugh¹

Center for Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The predominant organizational theme by which the transcription machinery and chromatin regulators are positioned within promoter regions or throughout genes in a genome is largely unknown. We mapped the genomic location of diverse representative components of the gene regulatory machinery in *Saccharomyces cerevisiae* to an experimental resolution of <40 bp. Sequence-specific gene regulators, chromatin regulators, mediator, and RNA polymerase (Pol) II were found primarily near the downstream border from the “–1” nucleosome, which abuts against the ~140-bp nucleosome-free promoter region (NFR). General transcription factors TFIIA, -B, -D, -E, -F, -H were located near the downstream edge from the NFR. The –1 nucleosome dissociated upon Pol II recruitment, but not upon recruitment of only TBP and TFIIB. The position of many sequence-specific regulators in promoter regions correlated with the position of specific remodeling complexes, potentially reflecting functional interactions. Taken together the findings suggest that the combined action of activators and chromatin remodeling complexes remove the –1 nucleosome after the preinitiation complex (PIC) has partially assembled, but before or concomitant with Pol II recruitment. We find PIC assembly, which includes Pol II recruitment, to be a significant rate-limiting step during transcription, but that additional gene-specific rate-limiting steps associated with Pol II occur after recruitment.

[Supplemental material is available online at www.genome.org. The array data from this study have been submitted to ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) under accession nos. E-MEXP-1676 and E-MEXP-1677.]

The eukaryotic transcription cycle is regulated by a large number of proteins, ranging from sequence-specific DNA binding factors to chromatin regulators to the general transcription machinery and their regulators (Struhl et al. 1998; Berger 2000; Pugh 2000; Orphanides and Reinberg 2002; Li et al. 2007). Their collective function is to express a subset of genes as dictated by a complex interplay of environmental signals that is only partly understood. Sequence-specific factors orchestrate the transcription cycle because their activities are controlled by cellular signals. DNA sequences directly and selectively dictate their recruitment to promoters. To better understand the transcription cycle and its applicability to all genes, we sought a clearer picture of how the transcription machinery and its regulators are organized at genes. From this organization, functional linkages among the many regulatory proteins may be deduced. While the location of the transcription machinery and its regulators have been largely pegged to their target genes throughout the *Saccharomyces* genome (Lee et al. 2002; Ng et al. 2002; Harbison et al. 2004; Steinmetz et al. 2006; Zanton and Pugh 2006), higher resolution maps that delineate the arrangement of proteins relative to each other are largely absent. Such maps establish the predominant *in vivo* organization of the transcription machinery in the context of positioned nucleosomes, and provide an important reference state for attempts to biochemically reconstitute and probe physiologically relevant mechanisms of transcription complex assembly.

Recent high-resolution maps of nucleosome organization have revealed a predominant chromatin architecture where a “–1”

H2A.Z-containing nucleosome resides immediately upstream of a nucleosome-free promoter region (NFR), and a “+1” H2A.Z-containing nucleosome resides immediately downstream where in yeast it occludes the transcriptional start site (TSS) (Guillemette et al. 2005; Li et al. 2005; Raisner et al. 2005; Yuan et al. 2005; Zhang et al. 2005; Albert et al. 2007; Lee et al. 2007; Shivaswamy et al. 2008). The –1, NFR, +1 arrangement provides the stage upon which the transcription preinitiation complex (PIC) assembles, and raises the intriguing question as to which nucleosome (–1 vs. +1), if any, controls PIC formation, and at what point during assembly or initiation it is displaced.

Many transcriptional regulators can be classified as components of a smaller number of protein complexes that regulate specific stages of the transcription cycle. Therefore, we sought to generate high-resolution location maps of representative factors from these different stages with particular focus on PIC formation. The factors examined include eight sequence-specific transcription factors (Rap1, Ifh1, Cin5, Swi4, Reb1, Rfx1, Xbp1, and Yap6), seven chromatin regulators (ISW1a, ISW1b, ISW2, SWI/SNF, RSC, INO80, and SWR-C), seven components of the general transcription machinery (TBP, TFIIA, -B, -D, -E, -F, and -H), two components of RNA polymerase (Pol) II (Rpb3, Rpo21), and three components of Mediator (Rgr1, Nut1, and Srb5).

These location maps allowed us to (1) assess PIC assembly characteristics of regions giving rise to noncoding RNA; (2) determine the relative placement of sequence-specific regulators, chromatin regulators, and PIC components at promoter regions, and whether such placement can pre-exist without recruitment of Pol II; (3) determine whether Pol II recruitment to promoters is associated with displacement of the –1 and/or +1 nucleosome; (4) establish where along individual genes Pol II transcription is

¹Corresponding author.

E-mail bfp2@psu.edu; fax (814) 863-8595.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.084970.108>.

rate-limited; and (5) assess whether the local position of sequence-specific regulators is linked to the local position of chromatin remodeling complexes, which might provide an indicator of functional interactions.

Results

Antisense transcripts originate from a canonical promoter architecture

Our initial studies employed high-density tiled (5-bp probe spacing) Affymetrix microarrays to conduct genome-wide location analysis on *Saccharomyces* growing exponentially in rich media. We tested a select subset of factors involved in PIC assembly, including a sequence-specific regulator (Reb1), chromatin remodeling complexes RSC (Rsc9) and SWR-C (Swr1), two general transcription factors (TBP and TFIIB/Sua7), and two RNA polymerase II subunits (Rpb3 and Rpo21). Examples of the distribution of these proteins at individual genes are shown in Figure 1. As expected, these factors were present at highly and moderately expressed genes and relatively depleted, but not necessarily absent, from lowly expressed genes (panel A). Repressed genes, such as sporulation-specific *SPS100*, were devoid of factors.

Genes that possessed confirmed antisense transcripts (Perocchi et al. 2007), including those originating internal to a gene, also had PICs or partial PICs (GTFs but not Pol II) (Zanton and Pugh 2006) assembled near the antisense TSS (Fig. 1B). Interestingly, the well-studied and quiescent *GAL10* gene contains an antisense transcript that originates internal to *GAL10* in a region that bears a standard “-1, NFR, +1” configuration (panel D). The -1, NFR, +1 configurations are defined in part by the presence of Reb1 (Raisner et al. 2005), which was present internal to *GAL10*, along with TBP and TFIIB, and low levels of Pol II. In general, confirmed antisense promoter regions generally adopted the canonical arrangement of nucleosomes and PIC components (Supplemental Fig. S1), indicating that many are likely to arise through normal PIC formation, rather than through undirected Pol II entry at potentially permissive NFR sites. This represents the first comprehensive assessment of PIC assembly at putative noncoding RNA promoters.

At divergently transcribed genes, whose promoter regions about over a rather narrow window (<400 bp), separate PICs were identifiable (Fig. 1C). This indicates that two spatially separated PICs can be accommodated at divergently transcribed genes.

Mapping accuracy

Local spatial resolution afforded by chromatin immunoprecipitation (ChIP)-chip has two limiting constraints: tiling distance

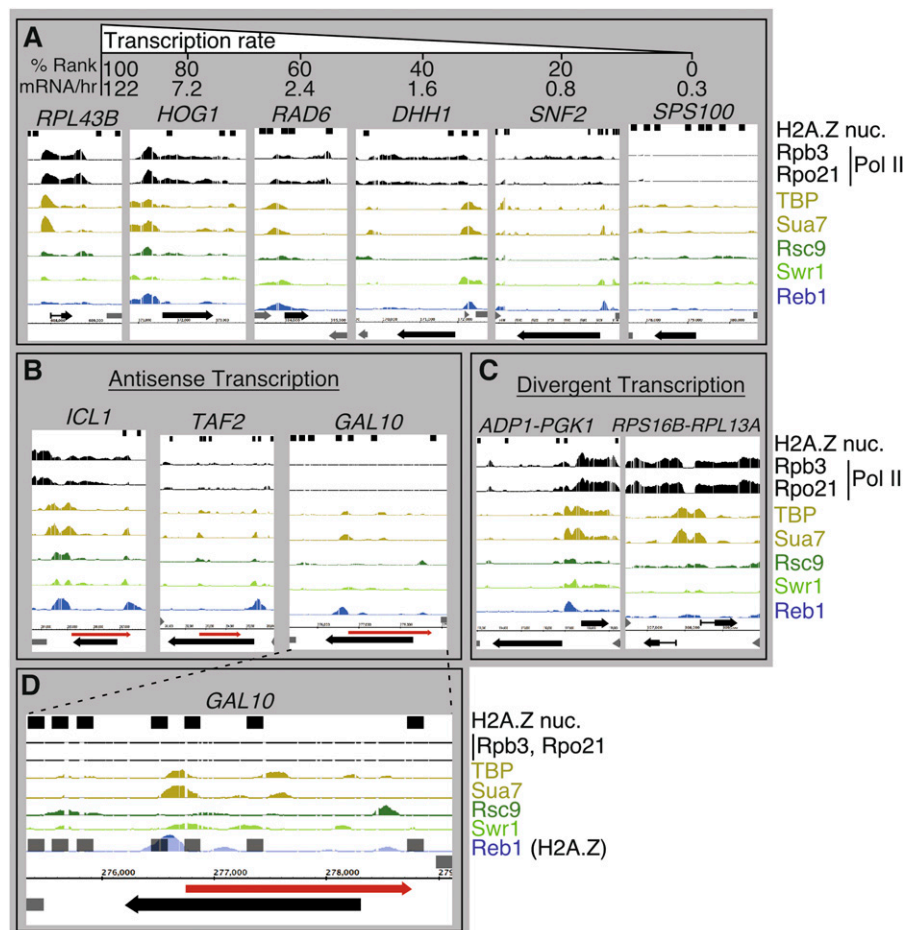


Figure 1. Distribution of transcription factor occupancy at selected genes. ChIP-chip on Affymetrix high-density tiling arrays was performed on representative transcription factors. Each individual bar in the graph represents the normalized signal from a 150-bp sliding window in 5-bp steps for the bound factor indicated to the right of each panel. The locations of H2A.Z-containing nucleosomes are also shown as 147-bp black boxes (Albert et al. 2007). The distribution of factors is anecdotal and should not be interpreted as the consensus location with respect to specific genomic landmarks. (A) Anecdotal example of genes covering a range of transcription frequencies (Holstege et al. 1998), showing a corresponding level of PIC components. (B) Anecdotal example of factor occupancy in regions producing confirmed antisense transcripts (red arrow) (David et al. 2006; Perocchi et al. 2007). (C) Anecdotal example of factor occupancy at divergent, highly transcribed genes. (D) An enlarged view of *GAL10*. (*GAL10* sense transcription is expected to be silent under the glucose-containing media in which this experiment was performed.)

(base-pair genomic distance between probes) and median DNA length of the ChIP DNA sample. In principle, shorter lengths of ChIP DNA provide greater accuracy and may be the limiting consideration in spatial resolution, rather than tiling distance. Consequently, we sought greater accuracy by shearing ChIP DNA to a median size of 200 bp, rather than the standard 500–1000 bp (Supplemental Fig. S2).

Mapping accuracy was evaluated by comparing the distance between the measured location of sequence-specific factors and their cognate DNA binding sites. For TBP, analyses were conducted only for TBP-occupied promoters in which a TATA-box is present (Basehoar et al. 2004). For TFIIB, which lacks a well-defined DNA binding site in yeast, we used the location of TBP to which TFIIB binds. As shown in Figure 2, most measured locations were within 35 bp of their target locations. The accuracy of the sequence-specific factors might be higher in that computational

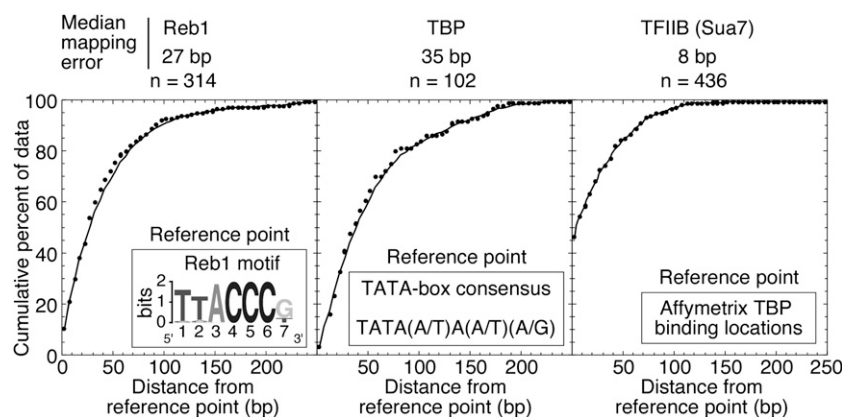


Figure 2. Error assessment of factor mapping. The cumulative error plots for Reb1, TBP, and Sua7 are shown. Cumulative error is the maximum allowable bp distance between predicted and measured locations. Only data that were within 250 bp of a reference point were considered. The median mapping error and number of data points (n) are indicated above each plot.

identification of cognate DNA binding sites does not guarantee that the sites are actually used. The median error between the measured TBP and TFIIB locations was substantially less (8 bp), indicating an overall greater precision. This represents the highest resolution mapping of factors reported to date.

TBP and other transcriptional regulators have a canonical organization in the NFR regardless of the presence of a TATA box

We examined whether the transcription machinery and chromatin regulators had a consensus location within the -1 , NFR, $+1$ canonical chromatin architecture (Fig. 3A–E). The frequency distribution of bound factors was determined relative to the TSS of all genes, and superimposed on the known canonical distribution of H2A.Z nucleosomes and TATA boxes. Reb1 is a sequence-specific transcription factor that functions in part to maintain an open chromatin structure (Fourel et al. 2002; Angermayr and Bandlow 2003; Yu et al. 2003; Raisner et al. 2005). Reb1 was narrowly distributed, peaking 120 bp upstream of the TSS (Fig. 3A). Reb1 was located in the NFR ~ 30 bp downstream from the -1 -nucleosome border, which is consistent with the proposed role of Reb1 in maintaining an NFR (Raisner et al. 2005).

As expected, TBP and TFIIB/Sua7 were tightly distributed over where the TATA box typically resides, ~ 60 bp upstream of the TSS (Fig. 3B,C). However, most promoters lack a TATA box, which leaves open the question as to where TBP assembles when a TATA box is not present. TBP (and TFIIB) were tightly distributed at the same relative distance from the TSS at promoters containing and lacking a TATA box (Supplemental Fig. S3). Essentially, the same result was obtained for the other mapped factors (data not shown), suggesting that a PIC is established at a fixed distance from nucleosomes and the TSS, regardless of the presence of a TATA box.

RSC1 and 2 are compositionally related major chromatin remodeling complexes in *Saccharomyces* (Cairns et al. 1996), and thus might be expected to be spread equally amongst all nucleosomes at active genes. However, the Rsc9 subunit, which is common to both RSC complexes, cross-linked primarily in the vicinity of the -1 nucleosome (Fig. 3D; Supplemental Fig. S4A), suggesting that its predominant target is the -1 nucleosome. Our Rsc9 target genes overlap significantly with RSC target genes reported else-

where ($P = 10^{-54}$ and 10^{-17} , respectively) (Damelin et al. 2002; Ng et al. 2002). Our observations suggest a mechanism in which RSC, as one of many remodeling complexes, facilitates removal of -1 nucleosomes to allow Pol II to access promoters.

In contrast to RSC, the SWR-C/SWR1 complex, which is responsible for ATP-dependent deposition of H2A.Z into nucleosomes (Krogan et al. 2003; Kobor et al. 2004; Mizuguchi et al. 2004), was found predominantly at both the -1 and $+1$ nucleosomes, which is where most of H2A.Z resides (Fig. 3D; Supplemental Fig. S4B). The ratio of SWR-C to H2A.Z at both locations was similar (data not shown), indicating that SWR-C largely resides at its site of action and in proportion to the amount of H2A.Z deposited.

Pol II was largely enriched in the promoter region, peaking at -100 bp from the TSS, which is ~ 40 bp upstream of where TBP and TFIIB assemble and ~ 60 bp downstream from the -1 nucleosome border (Fig. 3E). Similar results were obtained for TATA-containing and TATA-less promoters (data not shown). Our findings appear to contrast with recent reports of Pol II being more or less evenly distributed throughout genes, one using lower density arrays (Steinmetz et al. 2006) and another that used PCR tiling of an individual gene (Mason and Struhl 2005). We do find Pol II distributed throughout genes (see below), but this is largely restricted to highly expressed genes, which the previous studies focused on. However, for most genes there was an overall enrichment at promoters, which suggests that Pol II on average spends more time at promoter regions than at other locations in the gene. For yeast, this represents a change from the recent view that Pol II is spread uniformly throughout genes (Mason and Struhl 2005), and indicates that its distribution is more similar to that of metazoans than previously recognized. This distinction is important because it suggests that once Pol II is recruited to yeast genes it is subjected to additional rate-limiting regulatory steps during early elongation that would have little basis in the recent view.

The -1 nucleosome is removed upon Pol II recruitment, but not when only TBP and TFIIB are recruited

The location of Pol II toward the upstream region of promoters is consistent with a number of models placing upstream DNA near Pol II via looping or through wrapping around Pol II that is contained in a PIC (Coulombe et al. 1994; Robert et al. 1998). Given the dimensions of Pol II (Gnatt et al. 2001), this placement should allow Pol II to occupy the core promoter without dissociation of the $+1$ nucleosome, located downstream. Displacement of the upstream -1 nucleosome, however, may be necessary for Pol II occupancy. Displacement could involve lateral movement of the nucleosome to a more upstream location or dismantling of the nucleosome. To address the status of the -1 nucleosome when Pol II is present, we examined the distribution of nucleosomes at genes that are enriched with Pol II in the promoter region (Fig. 3F). Indeed, we find that the -1 nucleosome is depleted rather than laterally displaced when Pol II is present, suggesting that removal of the -1 nucleosome may be necessary for productive docking of

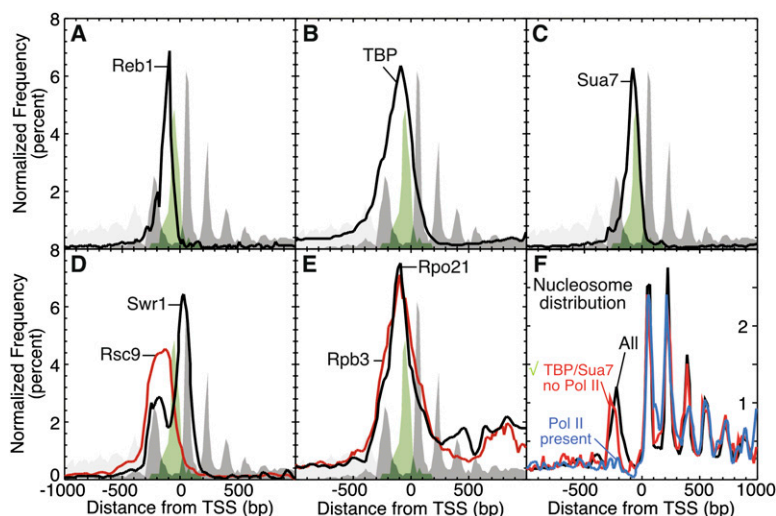


Figure 3. High-resolution map of the organization of the transcription machinery at promoter regions. ChIP-chip on Affymetrix high-density tiling arrays was used to determine the binding location for representative transcription initiation factors. To provide reference landmarks, the genome-wide frequency distributions for H2A.Z nucleosomes determined by ChIP-seq (Albert et al. 2007) and for conserved TATA consensus sites (Basehoar et al. 2004) are shown in gray and green, respectively. (A–E) The composite frequency distributions for the binding locations (peak calls from normalized signal with an FDR <5%) for representative components of the transcription machinery are shown relative to the TSS of protein coding genes. (F) Shown are the composite frequency distributions for the nucleosome sequence reads (Mavrich et al. 2008) for all genes (black trace), for genes containing high levels of Pol II in the promoter region (–500 bp to 100 bp relative to TSS, blue trace), and for genes containing both TBP and TFIIB, but not Pol II, in the promoter region (red trace). The bin counts are expressed as a percentage of the total for each trace over the indicated range so as to compare relative peak heights within traces. Genes that contained high levels of Pol II generally had lower overall levels of nucleosomes.

Pol II at promoters. This might explain the prevalence of remodeling complexes in the vicinity of the –1 nucleosome.

Strikingly, when the distribution of nucleosomes was examined at promoters that contained TBP and TFIIB, but no Pol II (i.e., partial PICs), the –1 nucleosome was present (Fig. 3F). This suggests that the –1 nucleosome is removed during PIC assembly, after TBP and TFIIB have been recruited, but before or concurrent with Pol II recruitment.

Pol II is rate-limited at distinct stages of the transcription cycle

The transcription cycle involves recruitment of Pol II to promoters, then transcription initiation, elongation, and termination (Wade and Struhl 2008). Any of these steps can potentially be rate-limiting. Recent studies find that an early elongation step is rate-limiting in the expression of ~10% of all *Drosophila* genes (Muse et al. 2007; Zeitlinger et al. 2007; Lee et al. 2008). In mammals, most genes are also enriched with Pol II at their 5' ends, but with little or no detected mRNA product (Guenther et al. 2007), and this has been interpreted to reflect widespread post-recruitment regulation of Pol II that occupies the 5' end of most genes. In *Saccharomyces*, it has been suggested that the transition from initiation to elongation is rapid, with little post-recruitment regulation except in isolated examples (Wade and Struhl 2008). That is, stationary phase cells and select genes may have preassembled PICs (Martens et al. 2001; Sekinger and Gross 2001; Radonjic et al. 2005). Pol II is also reported to be present at nearly all *Saccharomyces* genes (Steinmetz et al. 2006). Taken together, these findings have led to speculation that genes may be largely regulated after Pol II has been recruited, rather than before (Core and Lis 2008; Margaritis and Holstge 2008; Price 2008; Wade and Struhl

2008). The high-resolution mapping of Pol II, presented here, provides further insight into this concept.

To examine where Pol II is rate-limited during transcription, we collectively examined Pol II density along individual genes (Fig. 4A). Genes having similar distributions of Pol II were clustered. At >20× the resolution of previous data sets, this analysis is able to identify subgenomic locations of Pol II abundance, which we interpret to reflect rate-limiting steps during transcription (including initiation, elongation, and termination). We identified three groups in which Pol II was either enriched at promoters, enriched immediately downstream from promoters, or spread throughout the genes. Such groupings suggest that rate-limiting steps for transcription are tailored toward specific genes, and include promoter clearance, early elongation, and elongation/termination. This limited ability to clear promoters and/or conduct early elongation does not distinguish between stably bound poised or paused polymerases and polymerases that actively enter the region but terminate prematurely (Arigo et al. 2006; Thiebaut et al. 2006; Camblong et al. 2007; Perocchi et al. 2007; Proudfoot and Gullerova

2007; Uhler et al. 2007; Milligan et al. 2008).

We examined each group of genes for a wide range of properties so as to gain insight into why Pol II distributes differently at these collections of genes. In group 1, there were 1007 genes (17% of the genome) in which Pol II was enriched at promoters relative to the body of the gene (Fig. 4A, upper panels). Group 1 genes were remarkable in that the binding of more than 87 sequence-specific transcription factors was statistically enriched at these genes, using a *P*-value cut-off of <10^{–5}. This represents ~70% of the known factors in this class. Of these factors, those that were particularly enriched include Reb1, Rfx1, Mig1, and Rcs1 (*P* ~ 10^{–17} to ~10^{–12}, Supplemental Table S1). Sequence-specific factors are thought to orchestrate PIC assembly, and so their presence may be linked not only to recruitment of Pol II but also to its retention at promoters in a poised state.

In group 2, 1070 genes (18% of the genome) were enriched with Pol II at their 5' ends, but Pol II was generally located downstream from the promoter (Fig. 4A, middle panels). These genes were transcriptionally more active than group 1, but less active than group 3, when normalized to the amount of Pol II present. Thus, early transcription elongation appears to be a slow step in the transcription cycle of group 2 genes. Conceivably, Pol II may be paused or arrested or may be in some other rate-limited state at these genes. A potentially similar phenomenon occurs in multicellular eukaryotes where Pol II pauses soon after initiation (Muse et al. 2007; Lee et al. 2008), but this remains to be tested.

There were 964 genes (16% of the genome) in group 3 (Fig. 4A, lower panels). They displayed a comparatively even distribution of Pol II across their genes or an enrichment toward the 3' end. In comparison with other groups, group 3 genes had a high transcription frequency relative to Pol II density, indicating that

events on the DNA (promoter clearance, elongation, and termination) are much quicker than for group 1 and 2 genes. These genes tended to be the Rap1-regulated ribosomal protein genes ($P < 10^{-57}$; Supplemental Table S1) and were transcribed at the relatively high average rate of 14.3 mRNA/hr (four times higher than group 1). For these genes, the rate-limiting step may be the intrinsic maximal rate at which Pol II traverses the gene and terminates. As such, these genes might contain high levels of transcription-related chromatin regulators because of the strong coupling between Pol II and transcription-coupled chromatin modification. Indeed, group 3 genes were enriched with SWR-C ($P < 10^{-22}$) and the transcription-coupled histone methylase complex COMPASS ($P < 10^{-13}$).

Peaks of Pol II were observed at the 3' end of genes, which is where enrichment of Pol II often ended. These peaks mapped to recently identified 3' NFRs (Fig. 4B; Mavrich et al. 2008; Shivaswamy et al. 2008) and were ~60 bp downstream from the cleavage and polyadenylation sites and ~60 bp upstream of a nucleosome border. The distribution of Pol II was unchanged if only those regions that were downstream from two convergent genes was examined (T-T in Fig. 4B), which rules out "bleeding" effects from nearby promoters. These observations place the terminating Pol II at a very confined location between two nucleosomes. Conceivably, nucleosomal collision soon after transcript cleavage might hamper the progress of Pol II, allowing termination factors to act.

Most genes have relatively little Pol II bound

We next sought a quantitative measure of Pol II levels at individual genes to address the perception that all genes are actively transcribed. The first step was to define maximal Pol II occupancy. Across the entire genome, *RPS3* had the highest average density of Pol II throughout its gene (involving 366 Rpo21 ChIP measurements that were confirmed by 366 Rpb3 ChIP measurements), and thus represented a measured maximum of Pol II occupancy. Consistent with this, the set of more than 100 highly transcribed ribosomal protein genes achieved on average >60% of this maximal level, whereas the generally repressed sporulation-specific genes or the genes in the lowest fifth percentile of transcription frequency (61 and 311 genes, respectively) had on average ~1%–2% of the maximal level of Pol II occupancy (Fig. 4C). In addition, >50% of the genome had <5% of the maximum Pol II density (Fig. 4D), indicating that in contrast to current views most genes contain relatively little Pol II and are infrequently transcribed.

On the whole, highly expressed genes had a high Pol II density that was evenly distributed across genes, reflecting rapid recruitment and a rapid transition into elongation (Fig. 4E). In contrast, the lower density of Pol II at lowly expressed genes was concentrated at the 5' end of the genes. This reflects both low recruitment and slow transitions into productive elongation. Taken together, these findings suggest that most yeast genes are subjected to both recruitment and post-recruitment regulation of Pol II.

Organizational theme of transcriptional regulatory proteins in promoter regions

The data presented in Figure 1 suggest that regulatory proteins, whether they bind a specific DNA sequence or not, have their

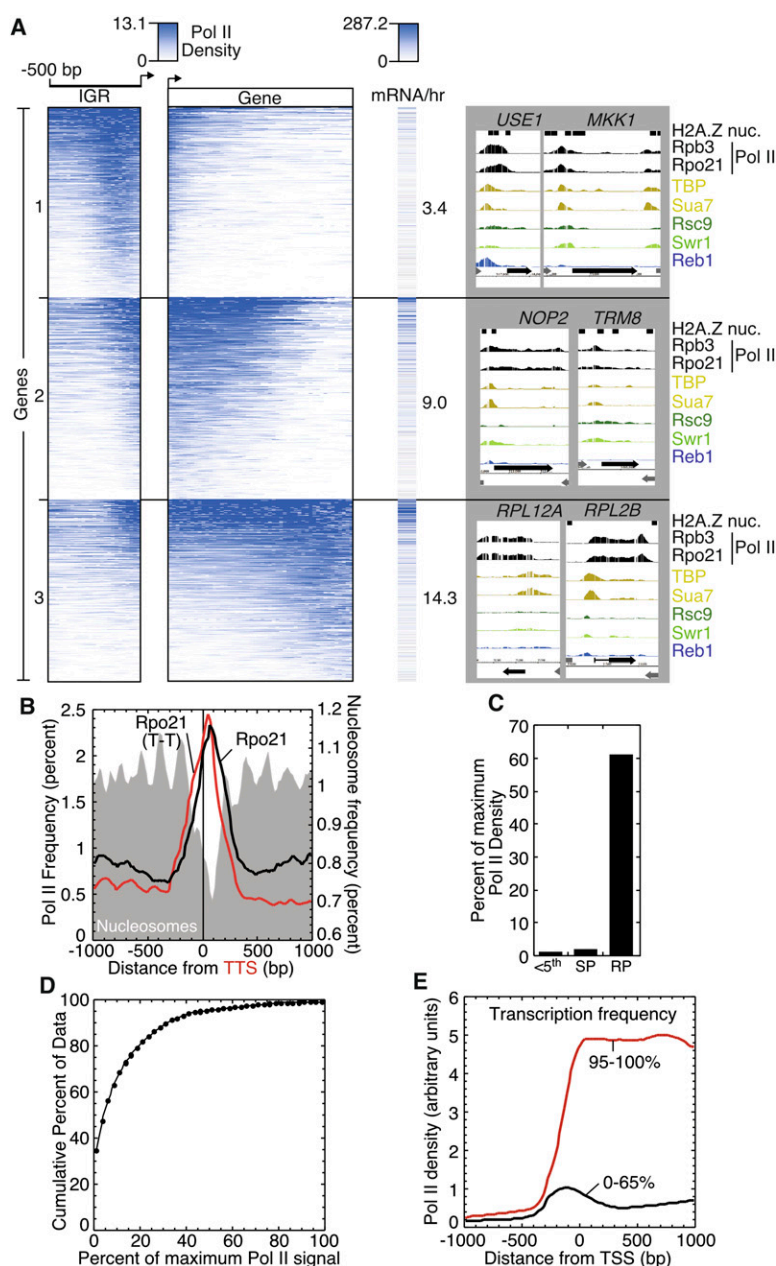


Figure 4. (Legend on next page)

consensus location in promoter regions. We expanded the study to look at the positions of a larger variety of regulatory proteins. We sought to increase cost efficiency without compromising resolution. To do this we designed microarrays with one probe at each end of every yeast promoter region (near the UAS and TSS, respectively), and an ORF probe toward the 3' end of each gene. By assessing the ratio of binding between the two promoter probes, we interpolated the location of the bound factor in the promoter region.

To validate the interpolation approach, we compared measured genome-wide locations to those obtained using the Affymetrix high density tiling arrays. As shown in Supplemental Figure S5, the precision of both platforms was very similar (<40-bp median error). We mapped the spatial location of 27 transcriptional regulatory proteins (eight sequence-specific factors, seven general transcription factors, five co-activators, and seven chromatin remodeling complexes) across promoter regions (Fig. 5). Since each factor has different ChIP efficiencies, absolute occupancy levels cannot be inferred by the area under each plot. These factors were chosen so that we could examine the genome-wide distribution of a broad range of factors that regulate different phases of transcription.

Sequence-specific regulators were chosen to represent three diverse sets of transcriptional programs (Nasmyth and Dirick 1991; Shore 1994; Mai and Breeden 1997; Huang et al. 1998; Furuchi et al. 2001): activation of housekeeping and ribosomal proteins genes (Rap1, Ifh1, and Reb1), activation of genes specific to a pathway (Swi4, Cin5, and Yap6), and repression (Rfx1 and Xbp1). All were found in the vicinity of the -1 nucleosome (Fig. 5A-C), which is in agreement with the location of their cognate binding site (Supplemental Fig. S6). Most factors peaked near the downstream edge from the -1 nucleosome, placing both the nucleosome and the factors in position to regulate each other's occupancy or position. Rap1 was more central to the nucleosome position, where it potentially binds or excludes nucleosomes (Yu and Morse 1999). In addition to a nucleosomal location, Rfx1 and Xbp1 repressor

proteins were also enriched in the NFR, which places them in position to regulate the PIC.

Of the chromatin remodeling complexes, nearly all (RSC, SWI/SNF, ISW1a, ISW1b, ISW2, and INO80, but not SWR-C) were predominantly located at the -1 nucleosome compared with the +1 nucleosome (Fig. 5D-F). The detection of SWR-C at the -1 and +1 nucleosomes confirms that these microarrays detect factor occupancy at both nucleosome sites. Relative occupancy internal to and toward the 3' end of genes was relatively low compared with promoter regions (Fig. 5L). However, some remodelers such as ISW1, SWI/SNF, and INO80 displayed significant occupancy in the genic regions. When the occupancy of all factors was compared on a gene-by-gene basis, most remodeling complexes (except SWR-C) occupied a similar set of genes, which only partially overlapped with a common set of genes occupied by the transcription machinery (Supplemental Fig. S7). Thus, not all remodelers are located at active genes, which might be expected if some are repressing gene expression.

Our finding that Isw2 was predominantly located at the -1 nucleosome appears to contrast with a report of Isw2 functioning within genic regions (Whitehouse et al. 2007). However, our reanalysis of the previous data also showed Isw2 predominantly bound in the vicinity of the -1 nucleosome (Supplemental Fig. S8). Interestingly, our reanalysis of a catalytically inactive Isw2(K215R) mutant places it in the vicinity of both the -1 and +1 nucleosomes, suggesting that the catalytic activity of Isw2 (and perhaps other remodeling complexes) results in a relatively transient presence at the +1 nucleosome.

SAGA functions to acetylate nucleosomes and deliver TBP to promoters (Grant et al. 1997; Dudley et al. 1999; Bhaumik and Green 2002). The Spt3 subunit of SAGA mapped to the -1 nucleosome region (Fig. 5G), as had been observed at the *GAL1* promoter (Larschan and Winston 2001). SAGA at this location might be interacting with the -1 nucleosome via bromodomains contained in its Gcn5 and/or Spt7 subunits as demonstrated in vitro (Hassan et al. 2002). Alternatively, SAGA might be interacting with activators bound in this region, potentially through the Tra1 subunit (Brown et al. 2001).

TFIID is a large multisubunit complex that functions to link activators and chromatin to TBP delivery at most promoters. The Taf1 subunit of TFIID was broadly distributed at promoter regions (Fig. 5G), which is consistent with TFIID interactions with multiple factors, including (1) activators (Garbett et al. 2007), (2) nucleosomes via bromodomains contained in the TFIID-interacting subunit Bdf1 (Matangkasombut and Buratowski 2003), and (3) core promoter interactions via TBP and TAFs (Buratowski et al. 1989).

General transcription factors are thought to function at most genes and are required for transcription initiation (Orphanides and Reinberg 2002). In addition to TFIID, these include TFIIA, TFIIB, TFIIE, TFIIH, and TFIIF. TFIIA peaked just upstream of TBP, and TFIIB just downstream (Fig. 5H), which is in agreement with their location as defined by DNase I footprinting and by X-ray crystallography (Buratowski et al. 1989;

Figure 4. Pol II density across genes. (A) Cluster plot of Pol II density across individual genes. Normalized Pol II (Rpo21) ChIP-chip signal was binned in 20-bp increments for intergenic regions (denoted "IGR") from the TSS to 500 bp upstream. The same was done for genic regions (denoted Gene) except that the data were parsed into 50 equal sized bins (and thus represents a percentage of the gene length, rather than absolute distance). The normalized signal within each bin was averaged, *K*-means clustered ($k = 4$, the fourth group of ~3000 genes is not shown because very little Pol II was observed in the vicinity of the genes), and displayed using Treview (Eisen et al. 1998). Pol II enrichment and transcription frequency are shown in blue. The transcription frequency, denoted as mRNA/hr (Holstege et al. 1998), for each gene was aligned to the cluster order and displayed, and the average value for each cluster was reported. Also shown are representative screenshots of the normalized signal distribution for Pol II and other factors at individual genes from groups 1-3. The identity of the genes in groups 1-3 can be found in Supplemental Table S5 (column J). (B) The composite frequency distribution of Pol II (Rpo21) intensity peaks relative to the transcript termination site (TTS) (i.e., the site of polyA addition). All TTS are shown in black, and only those that are downstream from two converging genes are shown in red (designated T-T). The distribution of nucleosomes is shown as a gray fill (Mavrich et al. 2008). (C) Pol II is mostly absent from lowly transcribed genes and abundant at highly transcribed genes. The gene having the highest average Pol II density was identified (*RPS3*), and set to the maximum level of Pol II occupancy (100 arbitrary units). The average Pol II density for three classes of genes, relative to this maximum was assessed: 311 genes in the lowest fifth percentile of transcription frequency (Holstege et al. 1998), 61 mid-late sporulation genes (SP) (Gasch et al. 2000), and 132 ribosomal protein genes (RPG). For each group of genes, the median percentage value is plotted. (D) Most genes have very little bound Pol II. The percentage of all 5228 protein coding genes (y -axis) that has less than the indicated Pol II level (as a percentage of the level found at *RPS3*) is plotted. For example, 50% of all genes have <5% of the maximum level of Pol II. The region examined for each gene was from 100 bp upstream of the TSS (or 160 bp upstream of the ORF start, if no TSS was present) to the TTS (or ORF endpoint, if no TTS was present). (E) Highly and lowly expressed genes display distinct Pol II distributions across genes. The average Pol II density throughout genes is plotted for the 311 genes in the upper fifth percentile (denoted 95-100%) and 3208 genes in the lower 65th percentile (denoted 0-65%) of transcription frequency (Holstege et al. 1998).

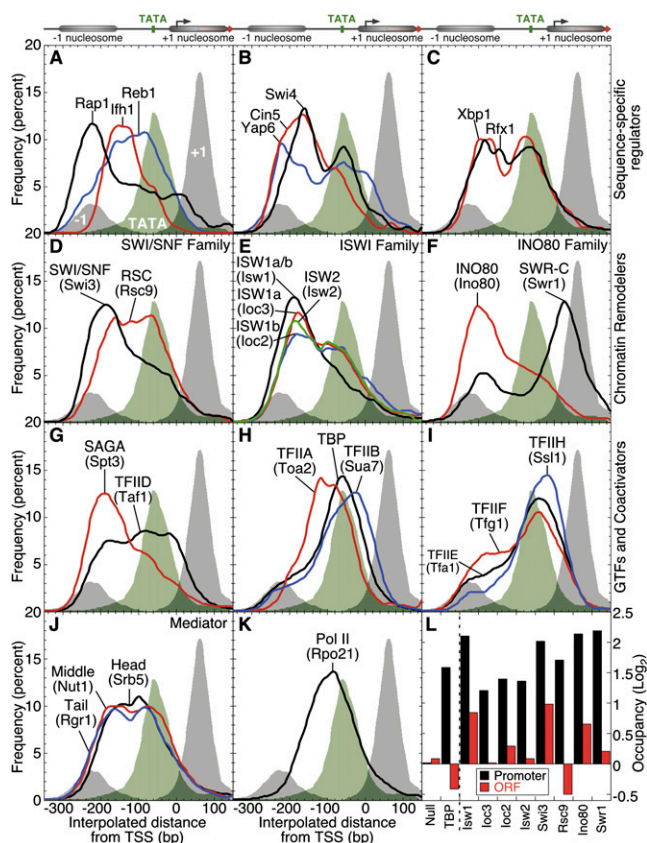


Figure 5. Genome-wide spatial assembly of the transcription machinery in the context of chromatin. (A–K) The composite frequency distributions for the binding locations for components of the transcription machinery are shown for all protein coding genes relative to the TSS (false discovery rates for each factor are reported in Supplemental Table S1). The distribution of TATA boxes when present (green fill) (Basehoar et al. 2004) and H2A.Z promoter nucleosomes (gray fill) (Albert et al. 2007), which are shown in each panel. (L) Occupancy detected toward the 3' end of each gene (designated ORF) is compared to occupancy in the promoter region. Shown in the bar graph are the median occupancy values for all genes in which occupancy of at least one region (ORF or Promoter) met the 5% FDR cut-off.

Nikolov et al. 1995; Geiger et al. 1996). TFIIE, TFIIF, and TFIIF were located just downstream from TBP and Pol II, and near the upstream edge of the +1 nucleosome (Fig. 5I). It is remarkable that the *in vivo* consensus location of many of the general transcription factors recapitulates well-defined biochemically reconstituted locations.

Mediator is a large multisubunit complex that links activators to Pol II (Kornberg 2005), and functions at most genes (Holstege et al. 1998). Mediator is composed of three distinct regions: the head (Srb5), middle (Nut1), and tail (Rgr1) (Dotson et al. 2000; Kang et al. 2001). The distribution of these proteins were virtually identical (Fig. 5J), and were broadly placed across promoters between the –1 nucleosome region and the TATA box region, much like the distribution of Pol II (Fig. 5K).

Sequence-specific regulators are location-linked to specific chromatin remodeling complexes

Chromatin remodeling complexes may be targeted to specific nucleosomes via interactions with sequence-specific DNA binding proteins (Cosma et al. 1999; Natarajan et al. 1999; Yudkovsky et al.

1999; Goldmark et al. 2000; Neely et al. 2002). Retention might be further enhanced through direct interactions with acetylated nucleosomes, in that many chromatin remodeling complexes contain bromodomains. Little is known about which sequence-specific DNA binding proteins specify which remodeling complex, if any, to recruit to promoter regions. We investigated this issue through the following two-point rationale: (1) A sequence-specific factor that recruits a remodeling complex should have both factors statistically enriched at the same promoters (co-occupancy in Fig. 6A and Supplemental Table S3); and (2) a sequence-specific factor that varies in its location (relative to the TSS) when all of its target promoters are examined, might mirror a similar location pattern for chromatin remodeling complexes that it interacts with (location linkage in Fig. 6A and Supplemental Table S4). Thus, two factors are “location-linked” if their positions covary at genes in which they co-occupy. To this end we performed a location-linkage analysis of each of the sequence-specific factors with each chromatin remodeling complex.

Some sequence-specific factors such as Rap1, Iff1, and Cin5 did not exhibit significant co-occupancy with any of the remodeling complexes tested, and thus lacked evidence for interactions with chromatin remodeling complexes (Fig. 6B, blue). Rfx1, Xbp1, and Yap6 displayed strong co-occupancy with all tested remodeling complexes except SWI/SNF (Fig. 6B, red). *P*-values ranged from 10^{-170} to 10^{-31} . Furthermore, Xbp1 and Rfx1 displayed location-linkage with all tested chromatin remodeling complexes, except SWI/SNF (Fig. 6C). Affinity capture experiments confirm an interaction between Rfx1 and ISW1a/b, ISW2, and RSC (Gavin et al. 2002).

Swi4 positively regulates G1-specific cyclin genes (Nasmyth and Dirick 1991). Swi4 co-occupied promoters with SWI/SNF ($P \sim 10^{-8}$), and the two were location-linked. As expected, these genes include cell cycle-regulated genes, such as *PCL1* and *SIM1*. While the location-linkage is consistent with their genetic interaction (Stern et al. 1984), there is no prior evidence that Swi4 recruits SWI/SNF to promoters. We examined the genome-wide recruitment of SWI/SNF to Swi4 target sites in a *swi4Δ* strain. Many of the 34 genes identified by location linkage had reduced levels of SWI/SNF in the absence of Swi4 (Fig. 6D), suggesting that Swi4 directs SWI/SNF to these genes. These findings were verified by standard gene-specific ChIP (Fig. 6E).

Discussion

Regulators and components of the transcription machinery occupy canonical positions in promoter regions

The Pol II transcription machinery assembles and operates in the context of nucleosomes, which have well-defined canonical positions at the beginning and end of protein-coding genes. The transcription machinery also assembles in full or in part at many places in the genome that are not associated with the promoter region of known genes. Many of these locations may be associated with the start of antisense transcripts, functional noncoding RNAs, or cryptic unstable transcripts (CUTs) (Arigo et al. 2006; Thiebaut et al. 2006; Camblong et al. 2007; Perocchi et al. 2007; Proudfoot and Gullerova 2007; Uhler et al. 2007; Milligan et al. 2008). Whether at the promoters of coding genes or antisense transcripts, there is a predominant pattern by which the transcription machinery is arranged (Fig. 7).

Importantly, the –1 nucleosome, which marks the upstream border of the promoter region, is a likely target of substantial regulation, assuming that occupancy is linked to function. In

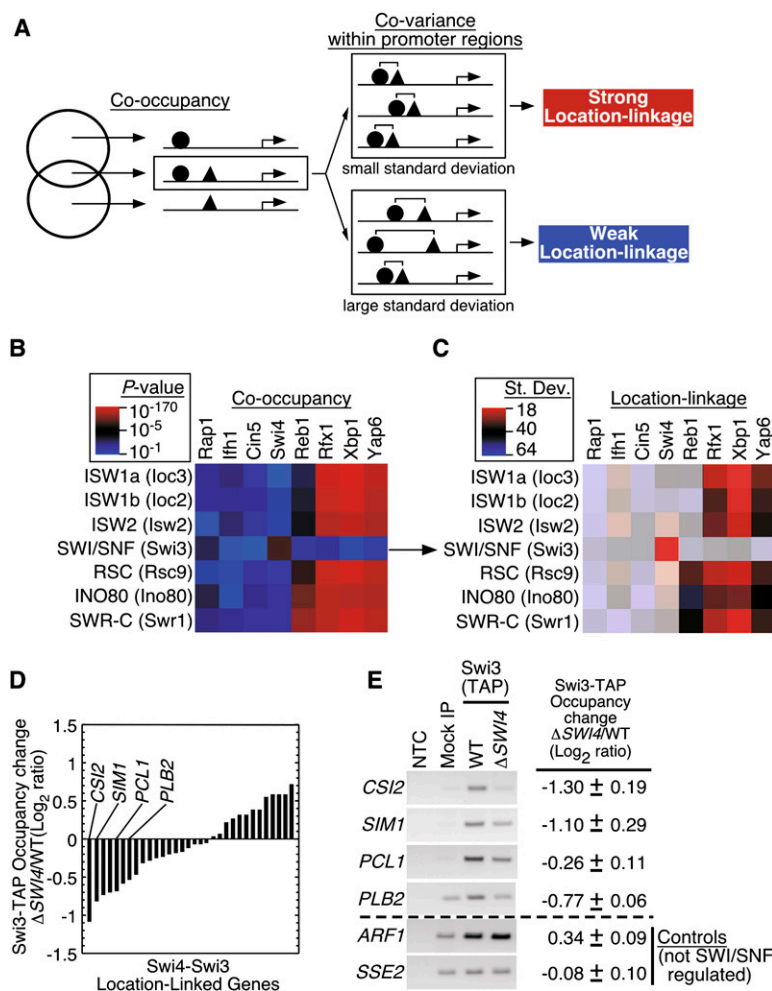


Figure 6. Sequence-specific regulators are location-linked to chromatin remodelers. (A) Schematic overview of the location-linkage analysis. Statistically significant promoter co-occupancy between a sequence-specific regulator and chromatin remodeler was computed using CHITEST function in Excel. Two factors are location-linked if their positions covary at genes in which they co-occupy. The extent to which two factors covary across all co-occupied promoters was determined by computing the standard deviation for the distances between their co-occupied binding locations. (B) Promoter co-occupancy between sequence-specific regulators and chromatin remodelers. Red reflects greater co-occupancy significance; blue, less significance relative to $P = 10^{-5}$. (C) Location linkage of sequence-specific regulators and chromatin remodelers. Brighter red indicates far less than the average standard deviation of pairwise distances (greater location-linkage). Darker blue indicates greater than average standard deviation (less location-linkage). Co-occupancy pairs that did not pass criterion 1 are shown with lighter shading. (D) ChIP-chip was used to determine the changes in occupancy for Swi3-TAP (SWI/SNF) in a *swi4Δ* strain. The Swi3-TAP occupancy changes for the genes that are Swi4-Swi3 location linked (panels B,C) are displayed as a bar graph. (E) Standard ChIP was performed on four of the genes most dependent on Swi4 for SWI/SNF recruitment (*CSI2*, *SIM1*, *PCL1*, and *PLB2*). A no template control (NTC) and a mock IP using an untagged BY4741 strain are shown. The PCR products from four biological replicates were quantified, and the log₂ ratios (*swi4Δ/WT*) ± SD for the Swi3-TAP changes in occupancy are displayed to the right of the PCR product for each gene. Loss of *SWI4* did not diminish Swi3-TAP expression or its global distribution (Supplemental Fig. S9).

particular, all tested chromatin remodeling complexes except SWR-C selectively occupy the -1 nucleosomal region compared to the +1 region and the 3' genic region. SWR-C occupies regions in proportion to the level of H2A.Z deposited. Second, sequence-specific regulators bind to regions of the promoter that overlap with the -1 nucleosome. Some might stabilize and position the -1 nucleosome, while others might exclude or be excluded by the -1 nucleosome. Within the rather narrowly defined -1 nucleosomal region, sequence-specific regulators vary in their location

relative to the TSS. For certain sequence-specific regulators, such as Swi4, Reb1, Rfx1, Xbp1, and Yap6, this gene-specific positioning is reflected in a linked positioning of specific chromatin remodeling complexes. This location-linkage is consistent with specific regulators directing specific remodeling complexes to the -1 nucleosome. However, this needs to be tested directly, in that location-linkage may be an indirect consequence of interactions with other linked proteins. Third, removal of the -1 nucleosome appears to be necessary for loading of Pol II, after TBP and TFIIB have assembled.

PIC assembly is at least partially rate-limiting in the expression of most genes, but post-recruitment regulation is widespread

The transcription cycle in yeast, as in other organisms, involves an interwoven series of highly regulated events, which can be rate-limited by Pol II recruitment or essentially any stage of the initiation/elongation/termination phase. Historically, unexpressed genes have been thought to be rate-limited by steps leading up to Pol II recruitment (Ptashne and Gann 1997). However, this notion has recently been challenged by the finding of Pol II at most genes, from yeast to humans (Steinmetz et al. 2006; Guenther et al. 2007), which has led to the suggestion that the expression of many or most genes is rate-limited after Pol II recruitment (Struhl 2007; Margaritis and Holstege 2008; Price 2008; Wade and Struhl 2008). However, this cannot be entirely correct. The ability to detect Pol II occupancy is a statistical metric (fold enrichment above background) and does not necessarily reflect 100% occupancy. The principle of mass action ensures that Pol II occupies every genomic locus at some nonzero level, but this level is less than a few percent of its maximum potential. How one parses observed levels of Pol II occupancy between experimental noise (DNA contamination), biological noise (genomic binding of Pol II, without functional consequence), and real transcription (genomic binding of Pol II, with functional consequences) can lead to different conclusions. Indeed, it has been suggested that as much as 90% of Pol II initiation is biological noise (Struhl 2007). But this conclusion depends upon assumptions inherent in distinguishing experimental vs. biological noise.

Our studies indicate that lowly transcribed genes in general have relative levels of Pol II occupancy that commensurate with the relative level of mRNA output, and that most genes have <5%

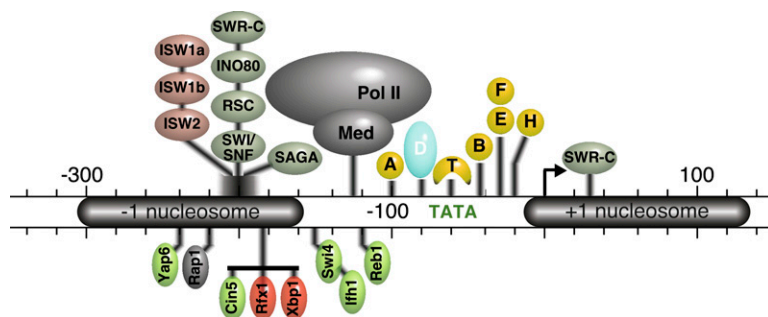


Figure 7. Illustration of the spatial assembly of the transcription machinery and chromatin regulators at promoters bracketed by the -1 and $+1$ nucleosomes. The coordinate system is based upon the location of the TSS (arrow). Locations are more approximate ($\pm \sim 20$ bp) than indicated by the lines attached to the coordinate. This view is a composite, and thus not all factors occupy the same set of genes.

of the maximal level of Pol II occupancy. Indeed such relationships are evident but not discussed in other genome-wide studies of Pol II occupancy (Steinmetz et al. 2006; Guenther et al. 2007). Thus, steps leading up to and including stable Pol II recruitment are likely to be limiting the expression of most genes from yeast to humans. Nonetheless, it is clear that recruitment of Pol II is not always sufficient to achieve maximal mRNA output, in that multiple post-recruitment barriers to elongation modulate this output.

Our data identify three groups of genes with different post-recruitment barriers that modulate mRNA output. These barriers include promoter clearance, early elongation, and termination. The data do not distinguish whether the barrier is caused by a polymerase translocation block or enhanced dissociation of Pol II immediately downstream from where Pol II accumulates. It is striking that genes that are highly enriched with a variety of sequence-specific regulators are rate-limited during transcription by promoter clearance. Some of these regulators might be retaining Pol II at promoters until signaling events trigger its release (Radonjic et al. 2005). Other genes may be able to clear their promoters more quickly but are slow in transiting into the early elongation phase, resulting in less than maximal expression. Such regulation is reminiscent of Pol II promoter proximal pausing in metazoans (Muse et al. 2007; Zeitlinger et al. 2007; Lee et al. 2008). The biochemical state of this early elongation complex is known to be quite different from the late elongation complex (Core and Lis 2008; Margaritis and Holstege 2008; Price 2008; Wade and Struhl 2008). Both involve post-translational modification and compositional alterations of the Pol II and its associated factors, which might be involved in the slowdown. The accumulation of Pol II at the end of genes might represent a kinetically rate-limited dissociation of Pol II during termination, or might represent cryptic unstable antisense transcription arising from Pol II activity in the 3' NFR.

Methods

Yeast strains, growth conditions, and epitope tagging

Strains used for ChIP-chip were obtained from the Yeast TAP-Fusion Library (Open Biosystems), in which tagged proteins were immunoprecipitated with IgG antibodies. Strain BY4741 was used for TBP ChIP-chip, and TBP was immunoprecipitated using anti-TBP rabbit serum. In both high- and low-density tiling array ChIP-chip experiments, BY4741 was used as an untagged Mock IP control for occupancy normalization.

Chromatin immunoprecipitation

ChIP assays were performed as described previously (Zanton and Pugh 2004). Briefly, each strain (500 mL) was grown at 25°C to mid-log phase ($A_{600} = 0.8$) in 500 mL YPD. Cells were incubated for 2 h with 1% formaldehyde and quenched for 5 min with glycine. The harvested cells were lysed with glass beads, and the chromatin pellet was washed and sonicated. Sheared chromatin was immunoprecipitated with IgG-Sepharose, or with an immobilized protein-specific antibody. This ChIP-enriched DNA was amplified by ligation-mediated PCR as described elsewhere (Harbison et al. 2004), and 75–300 bp LM-PCR-amplified fragments were gel purified according to the manufacturer's protocol (Qiagen). The gel-purified DNA yield was measured using the Nanodrop ND-1000 spectrophotometer.

DNA labeling and microarray hybridization

DNA labeling and hybridization to DNA microarrays were performed as described previously (Zanton and Pugh 2004). For the custom tiling microarrays, 100 ng of gel-purified LM-PCR ChIP-enriched DNA was amplified by 15 PCR cycles, coupled to Cy3 or Cy5 fluorescent dyes, and subsequently cohybridized to the arrays. For each ChIP sample, at least two biological replicates performed employing a dye swap for each replicate.

For high-density tiling arrays, 100 ng of gel-purified LM-PCR ChIP-enriched DNA was amplified according to the manufacturer's protocol (Affymetrix), and 7.5 μ g of biotin-labeled DNA was hybridized to Affymetrix *S. cerevisiae* Tiling 1.0R Array.

Array design and analysis

Low-density tiling microarrays

Oligonucleotide probes (60 bp) were designed and synthesized by Operon Biotechnologies, and were primarily against regions -320 to -260 and -90 to -30 from the start of each ORF in the *S. cerevisiae* genome, with positional adjustments made to keep a uniform T_m . Additional genic oligos comprised the Yeast Genome Oligo Set Version 1.1 (Operon Biotechnologies). An additional 1205 probes were synthesized against nonpromoter intergenic regions, and in snRNA and tRNA promoters. Each microarray consisted of $>21,000$ oligonucleotide probes spotted onto aminosilane treated glass slides. Data were filtered, analyzed, and occupancy levels determined as previously described (Zanton and Pugh 2006). Raw data are available at ArrayExpress under accession numbers E-MEXP-1676 and E-MEXP-1677. Normalized occupancy levels are available in Supplemental Table S5.

Fractional distribution analysis of normalized 25°C occupancy values was used to determine the transcription factor binding location at each promoter. Assuming a normal signal value distribution for the binding site of a transcription factor, fractional distribution analysis interpolates the binding site by using the fractional occupancy of adjacent upstream (-320 to -260 bp) and downstream (-90 to -30 bp) probes tiled across the promoter. Briefly, the binding site is determined by computing the ratio of the signal from the upstream probe over the sum of the signal from the upstream and downstream probes, resulting in a value between 1 and 0. This ratio is multiplied by the distance between adjacent probes, and added/subtracted from the upstream

probe coordinate depending on whether the probe is derived from the Watson or Crick strand. For example, a factor would map to the middle of the region between the upstream and downstream probes (–145 bp) if the fractional distribution ratio was 0.5. Therefore, the closer the ratio value is to 1, the closer a factor maps to the upstream probe, and vice versa. The binding locations for each transcription factor relative to the TSS and the absolute coordinates are reported in Supplemental Table S6.

False discovery rates were determined using a method described previously (Li et al. 2008) and are reported in Supplemental Table S1. Briefly, based on an untagged control ChIP-chip occupancy data set, we generated a frequency distribution plot of the ChIP values. Plotting the data on a \log_2 scale generally produces a “normal” (symmetrical) distribution for untagged control (background) data. Based on the assumption that the untagged control produces no true positives, at a given threshold the false discovery rate (FDR) can be estimated by computing the ratio of data points in the untagged control to the data points in the factor ChIP. The FDRs for each factor on low-density arrays is reported in Supplemental Table S2, and were used to generate Figure 5.

High-density tiling microarrays

Affymetrix *S. cerevisiae* Tiling 1.0R Arrays were used on select factors where noted. The *S. cerevisiae* Tiling 1.0R Array contains 3.2 million perfect match and mismatch probe pairs tiled through the complete yeast genome at a 5-bp tiling resolution. ChIP-enriched DNA was labeled and hybridized to *S. cerevisiae* Tiling 1.0R Array according to the manufacturer's protocol. Two independent biological replicates for TBP, Sua7, Reb1, Rsc9, Swr1, Rpo21, Rpb3, and Mock IP (BY4741) were analyzed. Raw data (i.e., cel files) are available at ArrayExpress under accession numbers E-MEXP-1676 and E-MEXP-1677. Normalized occupancy levels are available as wig files online at www.genome.org.

Signal analysis, interval analysis, and peak calling were performed using Model-based Analysis of Tiling-arrays (MAT) software (Johnson et al. 2006). Replicates for each factor were normalized to two replicates of the Mock IP (BY4741). A bandwidth and maximum gap value of 75 bp was used for the interval analysis. Peaks were called using a 5% FDR significance threshold (a 1% FDR was used for Reb1).

Location-linkage analysis

Location linkage involved pairwise comparisons between a sequence-specific regulator and a chromatin remodeling complex, and required two criteria to be met: (1) statistically significant promoter co-occupancy across the genome ($P < 10^{-5}$), and (2) less than the average standard deviation (40 bp) of pairwise distances among co-occupied promoter regions. For the purpose of displaying co-occupancy P -values as a colored display table (Fig. 6B), the values were transformed with the following formula: $\log_2(-\log_{10}(P\text{-value}) / -\log_{10}(5))$. As a result, $P < 10^{-5}$ is displayed by shades of red (more significant), whereas $P > 10^{-5}$ is displayed by shades of blue (less significant).

For the second criterion, the standard deviation (σ) of pairwise distances was determined. For the purpose of displaying the distribution of pairwise standard deviations as a colored display table (Fig. 6C), the standard deviation values were transformed with the following formula: $\log_2(\sigma/\sigma_{\text{ave}})$, where σ_{ave} is the averaged standard deviation of all pairwise combinations (equal to 40 bp) that passed criterion 1. As a result, positive transformed values (red) reflect a greater than average local spatial coupling at promoters. Co-occupancy pairs that did not pass criterion 1 are shown with lighter shading in Figure 6C.

Standard ChIP

Samples were prepared for ChIP-chip as described above except that gene-specific primers were used to PCR-amplify the promoter regions of the indicated genes. Each ChIP band in Figure 6E corresponds to 3.3 mL of cell culture ($OD_{600} = 0.8$). The PCR products were resolved on a 2% agarose gel and quantified using Image J software. The average value and standard deviation from four independent replicates are reported in Figure 6E.

Acknowledgments

We thank Travis N. Mavrich for designing and testing the oligo arrays and members of the Pugh laboratory and the Center for Gene Regulation for many helpful discussions. We thank Istvan Albert and Cizhong Jiang for bioinformatics support. This work was supported by NIH grant ES013768. This project is funded, in part, under a grant with the Pennsylvania Department of Health using Tobacco Settlement Funds. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations, or conclusions.

References

- Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. 2007. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**: 572–576.
- Angermayr, M. and Bandlow, W. 2003. Permanent nucleosome exclusion from the Gal4p-inducible yeast GCY1 promoter. *J. Biol. Chem.* **278**: 11026–11031.
- Arigo, J.T., Eyler, D.E., Carroll, K.L., and Corden, J.L. 2006. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell* **23**: 841–851.
- 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.
- Berger, S.L. 2000. Gene regulation. Local or global? *Nature* **408**: 412–413, 415.
- Bhaumik, S.R. and Green, M.R. 2002. Differential requirement of SAGA components for recruitment of TATA-box-binding protein to promoters in vivo. *Mol. Cell Biol.* **22**: 7365–7371.
- Brown, C.E., Howe, L., Sousa, K., Alley, S.C., Carrozza, M.J., Tan, S., and Workman, J.L. 2001. Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* **292**: 2333–2337.
- Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549–561.
- Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R.D. 1996. RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87**: 1249–1260.
- Camblong, J., Iglesias, N., Fickentscher, C., Dieppo, G., and Stutz, F. 2007. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**: 706–717.
- Core, L.J. and Lis, J.T. 2008. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**: 1791–1792.
- Cosma, M.P., Tanaka, T., and Nasmyth, K. 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **97**: 299–311.
- Coulombe, B., Li, J., and Greenblatt, J. 1994. Topological localization of the human transcription factors IIA, IIB, TATA box-binding protein, and RNA polymerase II-associated protein 30 on a class II promoter. *J. Biol. Chem.* **269**: 19962–19967.
- Damelin, M., Simon, I., Moy, T.I., Wilson, B., Komili, S., Tempst, P., Roth, F.P., Young, R.A., Cairns, B.R., and Silver, P.A. 2002. The genome-wide localization of Rsc9, a component of the RSC chromatin-remodeling complex, changes in response to stress. *Mol. Cell* **9**: 563–573.
- David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W., and Steinmetz, L.M. 2006. A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci.* **103**: 5320–5325.
- Dotson, M.R., Yuan, C.X., Roeder, R.G., Myers, L.C., Gustafsson, C.M., Jiang, Y.W., Li, Y., Kornberg, R.D., and Asturias, F.J. 2000. Structural

- organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci.* **97**: 14307–14310.
- Dudley, A.M., Rougeulle, C., and Winston, F. 1999. The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo. *Genes & Dev.* **13**: 2940–2945.
- Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* **95**: 14863–14868.
- Fourrel, G., Miyake, T., Defossez, P.A., Li, R., and Gilson, E. 2002. General regulatory factors (GRFs) as genome partitioners. *J. Biol. Chem.* **277**: 41736–41743.
- Furuchi, T., Ishikawa, H., Miura, N., Ishizuka, M., Kajiyama, K., Kuge, S., and Naganuma, A. 2001. Two nuclear proteins, Cin5 and Ydr259c, confer resistance to cisplatin in *Saccharomyces cerevisiae*. *Mol. Pharmacol.* **59**: 470–474.
- Garbett, K.A., Tripathi, M.K., Cencki, B., Layer, J.H., and Weil, P.A. 2007. Yeast TFIID serves as a coactivator for Rap1p by direct protein–protein interaction. *Mol. Cell. Biol.* **27**: 297–311.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Cell. Biol.* **11**: 4241–4257.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141–147.
- Geiger, J.H., Hahn, S., Lee, S., and Sigler, P.B. 1996. Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272**: 830–836.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Komberg, R.D. 2001. Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876–1882.
- Goldmark, J.P., Fazzio, T.G., Estep, P.W., Church, G.M., and Tsukiyama, T. 2000. The Isw2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p. *Cell* **103**: 423–433.
- Grant, P.A., Duggan, L., Cote, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., et al. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: Characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & Dev.* **11**: 1640–1650.
- Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**: 77–88.
- Guillemette, B., Bataille, A.R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. 2005. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol.* **3**: e384. doi: 10.1371/journal.pbio.0030384.
- Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. 2004. Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**: 99–104.
- Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111**: 369–379.
- Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717–728.
- Huang, M., Zhou, Z., and Elledge, S.J. 1998. The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell* **94**: 595–605.
- Johnson, W.E., Li, W., Meyer, C.A., Gottardo, R., Carroll, J.S., Brown, M., and Liu, X.S. 2006. Model-based analysis of tiling-arrays for ChIP-chip. *Proc. Natl. Acad. Sci.* **103**: 12457–12462.
- Kang, J.S., Kim, S.H., Hwang, M.S., Han, S.J., Lee, Y.C., and Kim, Y.J. 2001. The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* **276**: 42003–42010.
- Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D., and Rine, J. 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol.* **2**: e131. doi: 10.1371/journal.pbio.0020131.
- Kornberg, R.D. 2005. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**: 235–239.
- Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadian, V., et al. 2003. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol. Cell* **12**: 1565–1576.
- 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.
- Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**: 799–804.
- Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R., and Nislow, C. 2007. A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **39**: 1235–1244.
- Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol. Cell. Biol.* **28**: 3290–3300.
- Li, B., Pattenden, S.G., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J., and Workman, J.L. 2005. Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. *Proc. Natl. Acad. Sci.* **102**: 18385–18390.
- Li, B., Carey, M., and Workman, J.L. 2007. The role of chromatin during transcription. *Cell* **128**: 707–719.
- Li, X.Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D.A., Iyer, V.N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C.L., et al. 2008. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**: e27. doi: 10.1371/journal.pbio.0060027.
- Mai, B. and Breeden, L. 1997. Xbp1, a stress-induced transcriptional repressor of the *Saccharomyces cerevisiae* Swi4/Mbp1 family. *Mol. Cell. Biol.* **17**: 6491–6501.
- Margaritis, T. and Holstege, F.C. 2008. Poised RNA polymerase II gives pause for thought. *Cell* **133**: 581–584.
- Martens, C., Krett, B., and Laybourn, P.J. 2001. RNA polymerase II and TBP occupy the repressed CYC1 promoter. *Mol. Microbiol.* **40**: 1009–1019.
- Mason, P.B. and Struhl, K. 2005. Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol. Cell* **17**: 831–840.
- Matangkasombut, O. and Buratowski, S. 2003. Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol. Cell* **11**: 353–363.
- Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**: 1073–1083.
- Milligan, L., Decourty, L., Saveanu, C., Rappsilber, J., Ceulemans, H., Jacquier, A., and Tollervy, D. 2008. A yeast exosome cofactor Mpp6 functions in RNA surveillance and in the degradation of ncRNA transcripts. *Mol. Cell. Biol.* **28**: 5446–5457.
- Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**: 343–348.
- Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. 2007. RNA polymerase is poised for activation across the genome. *Nat. Genet.* **39**: 1507–1511.
- Nasmyth, K. and Dirick, L. 1991. The role of SWI4 and SWI6 in the activity of G1 cyclins in yeast. *Cell* **66**: 995–1013.
- Natarajan, K., Jackson, B.M., Zhou, H., Winston, F., and Hinnebusch, A.G. 1999. Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator. *Mol. Cell* **4**: 657–664.
- Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. 2002. Transcription activator interactions with multiple SWI/SNF subunits. *Mol. Cell. Biol.* **22**: 1615–1625.
- Ng, H.H., Robert, F., Young, R.A., and Struhl, K. 2002. Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes & Dev.* **16**: 806–819.
- Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**: 119–128.
- Orphanides, G. and Reinberg, D. 2002. A unified theory of gene expression. *Cell* **108**: 439–451.
- Perocchi, F., Xu, Z., Clauder-Munster, S., and Steinmetz, L.M. 2007. Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Res.* **35**: e128. doi: 10.1093/nar/gkm683.
- Price, D.H. 2008. Poised polymerases: On your mark... get set... go!. *Mol. Cell* **30**: 7–10.
- Proudfoot, N. and Gullerova, M. 2007. Gene silencing CUTs both ways. *Cell* **131**: 649–651.
- Ptashne, M. and Gann, A. 1997. Transcriptional activation by recruitment. *Nature* **386**: 569–577.
- Pugh, B.F. 2000. Control of gene expression through regulation of the TATA-binding protein. *Gene* **255**: 1–14.
- Radonjic, M., Andrau, J.C., Lijnzaad, P., Kemmeren, P., Kockelkorn, T.T., van Leenen, D., van Berkum, N.L., and Holstege, F.C. 2005. Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for

- rapid response upon *S. cerevisiae* stationary phase exit. *Mol. Cell* **18**: 171–183.
- Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. 2005. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**: 233–248.
- Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F., and Coulombe, B. 1998. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell* **2**: 341–351.
- Sekinger, E.A. and Gross, D.S. 2001. Silenced chromatin is permissive to activator binding and PIC recruitment. *Cell* **105**: 403–414.
- Shivaswamy, S., Bhingre, A., Zhao, Y., Jones, S., Hirst, M., and Iyer, V.R. 2008. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* **6**: e65. doi: 10.1371/journal.pbio.0060065.
- Shore, D. 1994. RAP1: A protean regulator in yeast. *Trends Genet.* **10**: 408–412.
- Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* **24**: 735–746.
- Stern, M., Jensen, R., and Herskowitz, I. 1984. Five SWI genes are required for expression of the HO gene in yeast. *J. Mol. Biol.* **178**: 853–868.
- Struhl, K. 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* **14**: 103–105.
- Struhl, K., Kadosh, D., Keaveney, M., Kuras, L., and Moqtaderi, Z. 1998. Activation and repression mechanisms in yeast. *Cold Spring Harb. Symp. Quant. Biol.* **63**: 413–421.
- Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. 2006. Transcription termination and nuclear degradation of cryptic unstable transcripts: A role for the nrd1–nab3 pathway in genome surveillance. *Mol. Cell* **23**: 853–864.
- Uhler, J.P., Hertel, C., and Svejstrup, J.Q. 2007. A role for noncoding transcription in activation of the yeast *PHO5* gene. *Proc. Natl. Acad. Sci.* **104**: 8011–8016.
- Wade, J.T. and Struhl, K. 2008. The transition from transcriptional initiation to elongation. *Curr. Opin. Genet. Dev.* **18**: 130–136.
- Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. 2007. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**: 1031–1035.
- Yu, L. and Morse, R.H. 1999. Chromatin opening and transactivator potentiation by RAP1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**: 5279–5288.
- Yu, Q., Qiu, R., Foland, T.B., Griesen, D., Galloway, C.S., Chiu, Y.H., Sandmeier, J., Broach, J.R., and Bi, X. 2003. Rap1p and other transcriptional regulators can function in defining distinct domains of gene expression. *Nucleic Acids Res.* **31**: 1224–1233.
- Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**: 626–630.
- Yudkovsky, N., Logie, C., Hahn, S., and Peterson, C.L. 1999. Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes & Dev.* **13**: 2369–2374.
- Zanton, S.J. and Pugh, B.F. 2004. Changes in genomewide occupancy of core transcriptional regulators during heat stress. *Proc. Natl. Acad. Sci.* **101**: 16843–16848.
- Zanton, S.J. and Pugh, B.F. 2006. Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock. *Genes & Dev.* **20**: 2250–2265.
- Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. 2007. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* **39**: 1512–1516.
- Zhang, H., Roberts, D.N., and Cairns, B.R. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**: 219–231.

Received August 15, 2008; accepted in revised form December 18, 2008.