

Divergence of a conserved elongation factor and transcription regulation in budding and fission yeast

Gregory T. Booth,¹ Isabel X. Wang,² Vivian G. Cheung,² and John T. Lis¹

¹Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853-2703, USA; ²Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, USA

Complex regulation of gene expression in mammals has evolved from simpler eukaryotic systems, yet the mechanistic features of this evolution remain elusive. Here, we compared the transcriptional landscapes of the distantly related budding and fission yeast. We adapted the Precision Run-On sequencing (PRO-seq) approach to map the positions of RNA polymerase active sites genome-wide in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*. Additionally, we mapped preferred sites of transcription initiation in each organism using PRO-cap. Unexpectedly, we identify a pause in early elongation, specific to *S. pombe*, that requires the conserved elongation factor subunit Spt4 and resembles promoter-proximal pausing in metazoans. PRO-seq profiles in strains lacking Spt4 reveal globally elevated levels of transcribing RNA Polymerase II (Pol II) within genes in both species. Messenger RNA abundance, however, does not reflect the increases in Pol II density, indicating a global reduction in elongation rate. Together, our results provide the first base-pair resolution map of transcription elongation in *S. pombe* and identify divergent roles for Spt4 in controlling elongation in budding and fission yeast.

[Supplemental material is available for this article.]

Budding and fission yeast are both widely used model organisms for the study of transcription, yet their genomes are as distantly related as metazoans and budding yeast (Dujon 2010). Details of transcription and its regulation have been characterized through a variety of high-resolution and genome-wide studies for *Saccharomyces cerevisiae* (Hahn and Young 2011; Pelechano and Steinmetz 2013). However, although location and production of genomic RNA has been investigated in fission yeast (Wilhelm et al. 2008; DeGennaro et al. 2013), various aspects of the transcriptional landscape of *Schizosaccharomyces pombe* remain more vague. Through our investigation of elongating RNA polymerase distributions in *S. pombe*, we unexpectedly found evidence of promoter-proximal pause-like distributions of RNA polymerase across the *S. pombe* genome. These findings prompted our inquiry into the mechanistic differences between budding and fission yeast at the level of transcription.

Unlike most well-studied metazoans, budding yeast displays a relatively uniform distribution of elongating Pol II across transcription units (Steinmetz et al. 2006). The lack of significant barriers to elongation in *S. cerevisiae* supports a model in which Pol II recruitment and initiation are the predominant steps regulating gene activity (Stargell and Struhl 1996; Ptashne and Gann 1997). In contrast, the majority of genes in metazoans, including mammals and flies, are intricately regulated at a promoter-proximal pausing step during elongation (Fuda et al. 2009; Adelman and Lis 2012). Characteristics of transcription elongation also differ between metazoans and *S. cerevisiae* beyond the cleavage and polyadenylation signal (CPS) (Porrua and Libri 2015). In contrast to budding yeast, elongating Pol II in mammals experiences post-CPS slowing or pausing while continuing to transcribe for several kilobases prior to termination (Proudfoot 1989; Gromak et al. 2006; Core et al. 2008; Laitem et al. 2015). These differences in Pol II dynamics reflect distinct mechanisms for controlling gene transcription

that likely reflect the vast evolutionary distance separating these organisms.

How promoter-proximal pausing evolved into a ubiquitous mechanism regulating transcription in mammals is unknown. Metazoans with pervasive pausing possess a four-subunit complex termed negative elongation factor (NELF) with no known orthologs in yeast, *C. elegans*, or plants (Hartzog and Fu 2013). NELF cooperates with DRB-sensitivity inducing factor (DSIF), composed of the highly conserved Spt4 and Spt5 subunits, to induce stable pausing of RNA polymerase shortly after initiation (Wada et al. 1998; Wu et al. 2003). However, studies in which NELF was highly depleted *in vivo* have shown that cells still retain substantial promoter-proximal pausing (Brannan et al. 2012; Core et al. 2012), consistent with the idea that DSIF is sufficient to induce promoter-proximal pausing (Wada et al. 1998). In fact, recent studies found that, despite lacking NELF homologs, some *C. elegans* genes are paused under nutrient-limiting conditions (Kruesi et al. 2013; Maxwell et al. 2014). Nonetheless, it remains unknown whether such pausing *in vivo* is dependent on Spt4-Spt5.

Pause release in metazoans is modulated by the kinase activity of cyclin dependent kinase 9 (CDK9), a complex referred to as positive transcription elongation factor b (pTEFb) (Wada 1998). CDK9 phosphorylates NELF, DSIF, and Pol II to permit pause release and elongation into the gene body (Peterlin and Price 2006). This transition coincides with the release of NELF from the complex, at which point phosphorylated DSIF is converted into a positive elongation factor (Ping 2000). In all eukaryotes, DSIF stably associates with transcribing Pol II throughout the duration of transcription. Mutations in either DSIF subunit (Spt4-Spt5) in *S. cerevisiae* have been found to result in elongation defects (Swanson et al. 1991; Keogh et al. 2003), but how Spt4-Spt5 evolved an additional role in promoter-proximal pausing remains unknown.

Corresponding author: johnlis@cornell.edu

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.204578.116>.

© 2016 Booth et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Our understanding of pausing as well as numerous other transcription-coupled phenomena has been greatly enhanced through the use of nuclear-run on approaches (Love et al. 1985). More recently, Global and Precision Run-On sequencing (GRO/PRO-seq) have provided genome-wide views of the distribution of engaged RNA Polymerases with strand specificity in metazoan organisms (Core et al. 2008; Kwak et al. 2013). Despite the use of various run-on-based approaches in *S. cerevisiae* (Pelechano et al. 2010; Rodríguez-Gil et al. 2010; McKinlay et al. 2011; Jordán-Pla et al. 2014), neither GRO-seq nor PRO-seq has been fully demonstrated in yeast. Here, we use a yeast-optimized, Precision Run-On sequencing approach to assay elongating RNA polymerases genome-wide in distantly related budding and fission yeast, providing the first base-pair resolution transcription analysis in *S. pombe*. We further apply this technique to investigate the importance of the Spt4-Spt5 complex in global transcription elongation. Comparing the effect of Spt4 deletion in both *S. cerevisiae* and *S. pombe* provides insight into the evolution of its role in regulating Pol II transcription elongation in eukaryotes.

Results

PRO-cap and PRO-seq capture sites of transcription initiation and elongation in *S. cerevisiae* and *S. pombe*

We first set out to map positions of all elongation competent RNA polymerases in both budding and fission yeast using the Precision Run-On sequencing (PRO-seq) protocol, which we adapted for use in yeast. Furthermore, we used Precision Run-On 5' cap sequencing (PRO-cap), a modification to the PRO-seq protocol which includes selection for short, capped nascent RNAs, to precisely identify the initiation sites for transcription units in both species (Supplemental Table S2, *S. cerevisiae*; Supplemental Table S3, *S. pombe*; Kwak et al. 2013). PRO-seq signal in *S. pombe* is observed across the transcription unit of active genes, providing the direction, relative amount and distribution of elongating polymerases (Fig. 1A, top track). In contrast, PRO-cap signal is largely confined to the 5' regions near annotated TSSs, indicating a successful enrichment of nascent RNA 5' cap sites (Fig. 1A, lower track).

To more precisely define the initiation sites of transcription units, we identified the base with the most PRO-cap reads above background within 250 bases upstream of or downstream from the annotated TSS (Fig. 1A; Supplemental Fig. S1A, green models).

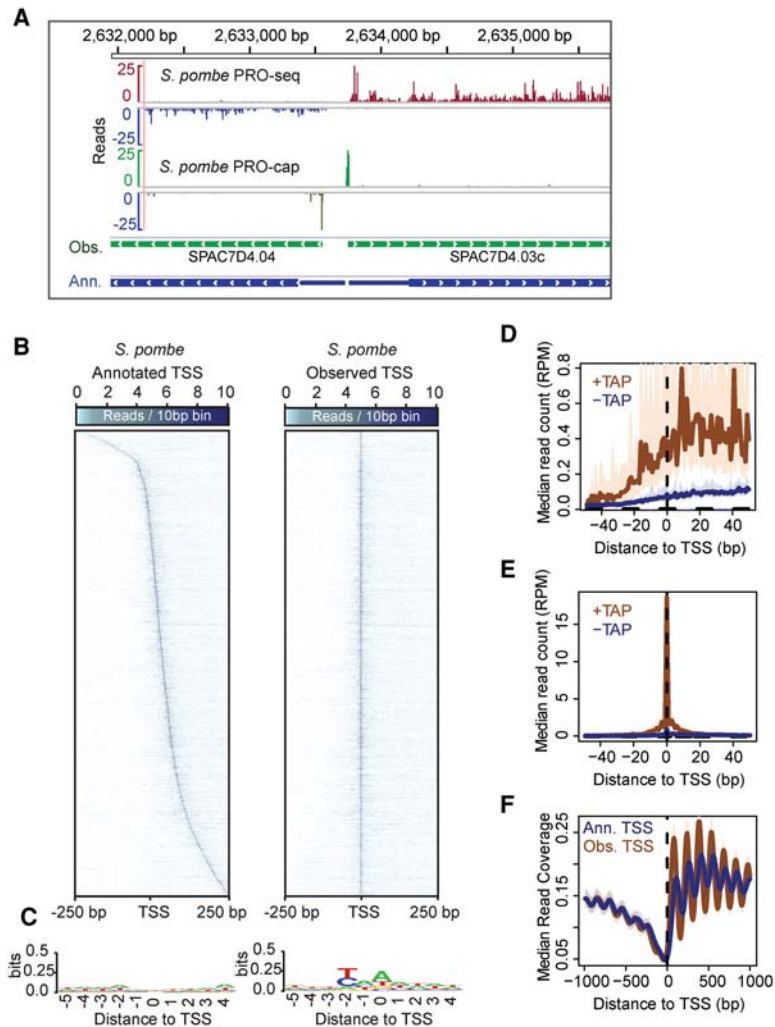


Figure 1. PRO-seq and PRO-cap capture transcription elongation and initiation genome-wide in *S. pombe*. (A) Browser tracks of PRO-seq (plus strand: red; minus strand: blue) and PRO-cap data (plus strand: green; minus strand: olive) derived from *S. pombe*. Green gene models below the data tracks show the reannotated, “observed” transcription start-sites based on PRO-cap data. Blue gene models correspond to the available annotations of genes. (B) Heatmaps of *S. pombe* PRO-cap signal for each base within ± 250 bp around the annotated TSS (left) and PRO-cap observed TSS (right) for all active and filtered genes ($n = 3214$). Genes within heatmaps are sorted by increasing downstream distance of observed TSS relative to annotated TSS. (C) Sequence logos of 10-bp sequence centered on either annotated TSS (left) or observed TSS (right) generated using WebLogo (Crooks et al. 2004). (D) PRO-cap signal from samples prepared either with or without Tobacco Acid Pyrophosphatase (TAP) treatment centered on annotated TSS. The TAP-minus samples represent empirical levels of background for each genomic position. (E) Median PRO-cap signal from samples prepared either with or without TAP treatment centered on observed TSS. (F) Median MNase-seq coverage centered on annotated TSSs (blue) or observed TSSs (brown). For the metagene plots, the y-axis shows the median read counts for each base-pair (D, E), or median read coverage within 10-bp bins (F). In D, E, and F, the 12.5% and 87.5% quantiles are shown in lightly shaded regions.

We refer to the selected base as the “observed TSS,” and to previous annotations as the “annotated TSS.” Heatmaps of PRO-cap read depth surrounding previously “annotated TSSs” in *S. pombe* show that initiation is distributed around the annotated TSS, but that the annotation rarely matches the base-pair resolution measurement of initiation that we make with PRO-cap (Fig. 1B, left panel). In contrast, the distribution of PRO-cap reads is centered and highly focused at our “observed TSSs” in *S. pombe*, indicating a narrow window within which start-site selection occurs (Fig. 1B, right panel). Moreover, a moderate sequence preference for

initiating at an adenosine, two bases downstream from a thiamine, was revealed under the observed TSS, whereas no base preferences underlie the PomBase annotations for the same genes (Fig. 1C). More recently, groups have used other approaches to reannotate transcription units across the fission yeast genome (Li et al. 2015; Eser et al. 2016). Comparing our PRO-cap-defined annotations in *S. pombe* with TSSs determined by Eser et al. (2016) revealed a much closer agreement between these two independent approaches than with PomBase annotations, although some differences still remained (Supplemental Fig. S2A,B).

For *S. cerevisiae*, we compared our PRO-cap observed TSSs with the 5' ends of the longest major transcript isoforms identified by Pelechano et al. (2013) (Supplemental Fig. S1). Although a striking enrichment of PRO-cap signal occurred directly over expected 5' ends of these transcript isoforms (Supplemental Fig. S1B,D), we found many preferred TSSs to be more downstream, consistent with the existence of many shorter major transcript isoforms in *S. cerevisiae* (Supplemental Fig. S1B; Pelechano et al. 2013).

To assess the amount of signal attributable to background in our PRO-cap experiments, we simultaneously prepared PRO-cap libraries from the same starting material in which nascent RNAs were not treated with Tobacco Acid Pyrophosphatase (TAP) prior to 5' adaptor ligations. These sequenced RNAs represent molecules lacking a 5' cap structure and therefore are unlikely to represent true transcription start sites. Looking at a 100-bp window around annotated or observed TSSs in *S. pombe*, we observed a highly focused enrichment of reads over background centered on the observed TSS (Fig. 1E), a trend that does not exist when centering previous annotations (Fig. 1D). The same analysis in *S. cerevisiae* revealed PRO-cap signal highly consistent with the transcript isoform 5' ends used, suggesting good agreement of our data with previous work (Supplemental Fig. S1D; Pelechano et al. 2013). Nonetheless, we attained an even greater reduction in signal dispersion by centering PRO-cap signal around the most preferred base (Supplemental Fig. S1E). Plotting PRO-seq signal relative to TSSs confirms that average 5' boundaries' global RNA polymerase profiles more closely reflect the observed TSS compared with previous annotations in both fission and budding yeast (Supplemental Fig. S3). Moreover, PRO-cap signal in *S. cerevisiae* was found to be strictly enriched downstream from previously characterized TATA-like motifs and TFIIB positions (Supplemental Fig. S4; Rhee and Pugh 2012).

To further evaluate our observed TSSs, we assessed the effect of centering available MNase-sequencing data around the preferred base (DeGennaro et al. 2013; Hu et al. 2014). Composite MNase-seq profiles around the observed TSSs in *S. pombe* show a more highly phased pattern when compared with annotated TSSs (Fig. 1F). Centering MNase-seq data from *S. cerevisiae* around the observed TSSs produced a modest reduction in phasing pattern of the MNase-seq reads, relative to the longest TIF-seq isoform 5' ends (Supplemental Fig. S1F). Together, our results indicate that PRO-cap-defined TSSs are truly an accurate representation of transcriptional start sites in *S. pombe* and *S. cerevisiae* that can improve our analysis of early transcription elongation in yeast, especially in *S. pombe*, where less mapping has been performed to date.

Budding and fission yeast exhibit vastly different polymerase distributions across genes

Using PRO-seq-based maps of RNA Pol II in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, we sought to identify general features of transcription that are species specific. To minimize

the effects of transcription emanating from neighboring genes, we identified a set of filtered genes at least 1 kb from any other genes on the same strand (*S. pombe*: $n=874$; *S. cerevisiae*: $n=1101$). We then generated scaled composite profiles in which the region between 300 bp downstream from the observed TSS and 300 bp upstream of the cleavage and polyadenylation signal (CPS) is scaled to the same number of bins, while everything outside this gene-center region is not scaled (10 bp bins). Importantly, PRO-seq signal increases immediately downstream from the observed TSS (Fig. 2A,B), further indicating PRO-cap-based observed TSSs are genuine start sites for these genes.

Similar to previous studies of transcription in *S. cerevisiae* (Steinmetz et al. 2006; Churchman and Weissman 2011), we observed a relatively uniform distribution of RNA Pol II across transcription units, with transcription extending ~200–300 bp beyond the CPS on average (Fig. 2A). Notably, NET-seq profiles exhibit a modest decline in read density downstream from the TSS, which was not present in our PRO-seq data (Supplemental Fig. S5A). This distinction may reflect differences in the forms of nascent RNA captured by each technique. For instance, NET-seq captures all Pol II-associated RNAs with a 3' OH, such as those in backtracked complexes. Pol II backtracking may occur more frequently during early elongation (Churchman and Weissman 2011) and not be captured by PRO-seq, which requires complexes capable of running on. Nonetheless, in contrast to the Pol II distributions of *S. cerevisiae* measured by either technique, our composite RNA Pol II profiles across genes in *S. pombe* show two more prominent peaks of Pol II density. The first peak is immediately downstream from the TSS, whereas a second, broader peak occurs downstream from the 3' CPS. Furthermore, Pol II in *S. pombe* appears to transcribe much farther beyond the CPS when compared with *S. cerevisiae* (Fig. 2B). Visualization of PRO-seq density for representative genes, YBR078W in *S. cerevisiae* and SPBC32F12.10 in *S. pombe*, illustrates these same characteristics (Fig. 2C,D). Regarding intron–exon junctions, Pol II density around fission yeast splice sites exhibited features similar to those described in *Drosophila* (Kwak et al. 2013), whereas the rarity of introns in budding yeast precluded a more thorough comparison (Supplemental Fig. S6A,B).

To compare the enrichment of RNA Pol II density downstream from the TSS between *S. pombe* and *S. cerevisiae*, we calculated a 5' pausing index for each gene considered in the composite profiles. The 5' pausing index was defined as the ratio of read density within the first 100 bp downstream from the TSS over the gene body (TSS + 200 to CPS). On average, *S. pombe* genes exhibited significantly higher pausing indexes when compared with *S. cerevisiae* (Fig. 2E). Similarly, to assess the amount of pausing or slowing beyond the 3' CPS, we computed a 3' termination index for these filtered genes. The post-CPS termination index was defined as the ratio of the highest density 100 bp window within 500 bp downstream from the CPS to the gene body (defined above). We found that *S. pombe* genes on average also have a significantly higher termination index than *S. cerevisiae*.

To investigate the significance of promoter-proximal pause-like distributions observed in *S. pombe*, we systematically classified genes as being paused or not paused. Genes classified as paused were required to show a significant enrichment (Bonferroni corrected $P < 0.01$) of mappable reads within the promoter-proximal region (TSS to +100 bp) relative to the gene body, as assessed by Fisher's exact test. Not paused genes were required to have a P -value > 0.99 , suggestive of no enrichment of promoter-proximal reads. Based on the criteria above, of the 3214 filtered *S. pombe* genes (Methods), we identified 901 genes as being paused and

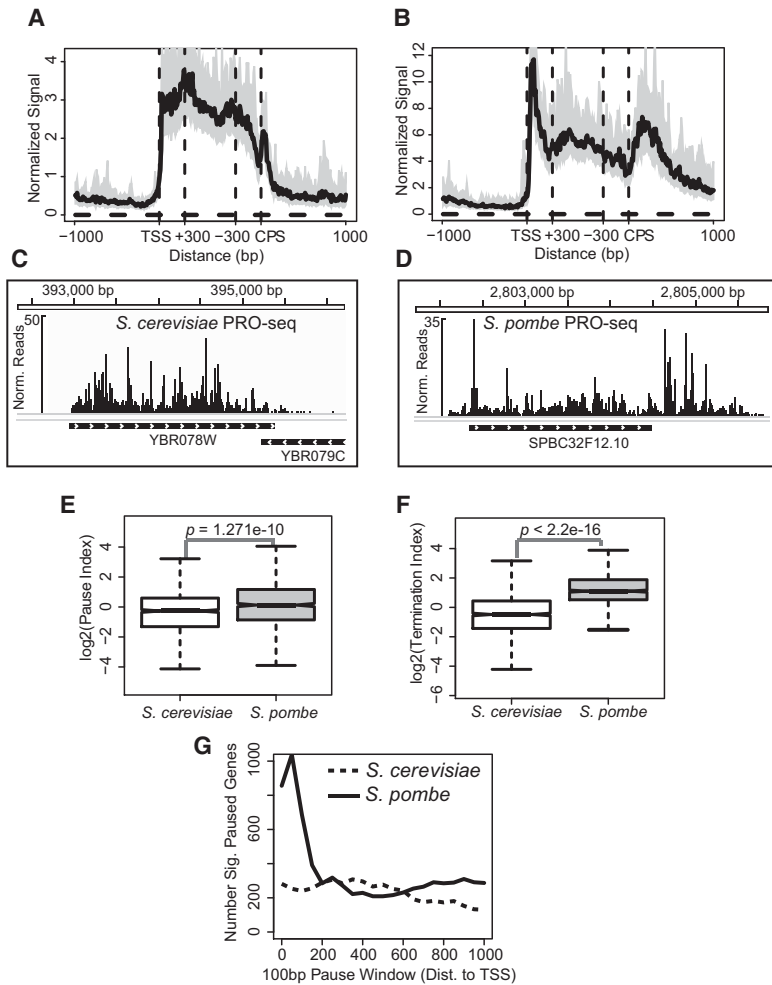


Figure 2. PRO-seq reveals distinct transcription elongation profiles in *S. cerevisiae* and *S. pombe*. (A,B) Median PRO-seq read count across all active and filtered genes in *S. cerevisiae* (A) or *S. pombe* (B) that are separated from neighboring genes on the strand by at least 1 kb (*S. cerevisiae*: $n = 1101$; *S. pombe*: $n = 874$). Bins between the +300 and -300 bp marks are scaled based on gene length, whereas upstream of and downstream from this center region, 10-bp bins were used. The shaded regions around the curves represent the 12.5% and 87.5% quantiles. (C,D) Representative genes from *S. cerevisiae* (C) and *S. pombe* (D) with PRO-seq read counts plotted above. (E,F) Box plots of pausing index (E) or termination index (F) values calculated for all genes that were included in A and B. *P*-values represent results of Student's *t*-test. (G) A test for enrichment of pausing near the promoter versus other gene regions. Reads were counted within a sliding 100-bp window from 0 to 1000 bp from the TSS of all filtered, active genes and divided by the counts within the remaining mappable gene length. Fisher's exact test was used to determine the number of significantly paused genes (adjusted $P < 0.01$).

2133 not paused. The relative amount of active genes identified as paused in *S. pombe* (28%) is thus lower than that observed for humans (41%), using the same criteria (Core et al. 2008). Although there were few functional class (Thomas et al. 2003) enrichments for paused or not paused genes, highly paused genes (defined below) were significantly enriched for structural constituents of the ribosome ($P = 3.84 \times 10^{-4}$), which are often highly paused in other species (Min et al. 2011).

Broad peaks of RNA Pol II have been observed beyond the 3' CPS of many genes in *S. pombe* (Coudreuse et al. 2010; Sansó et al. 2012; Castel et al. 2014); yet to our knowledge, no group has identified a prominent peak downstream from the TSS. We thus compared PRO-seq profiles for highly paused, not paused, and all genes with publicly available Pol II ChIP-seq data in fission

yeast (DeGennaro et al. 2013). Both data types showed broad Pol II enrichment downstream from the CPS for all gene types (Supplemental Fig. S5B, right). Furthermore, consistent with our gene classification based on PRO-seq, highly paused genes also showed a promoter-proximal Pol II ChIP-seq peak that was greater than that of the other gene groups, further suggesting an enrichment of elongating Pol II downstream from the TSS for these genes (Supplemental Fig. S5B, left).

To assess whether the observed pause peaks are specific to the promoter-proximal region, we repeated the calculation of pausing index and significance using a sliding 100-bp window starting every 50 bases between the TSS and 1000 bp downstream (20 measurements per gene). We then tallied the number of genes that were identified as significantly paused within each window relative to the remaining gene length. In *S. pombe*, we found that the most promoter-proximal regions have significantly larger number of genes identified as paused, but in *S. cerevisiae*, there is not a significant difference in any given window (Fig. 2G). These results strongly indicate the existence of a previously undescribed form of promoter-proximal pausing in *S. pombe* that is absent in *S. cerevisiae*.

Average Pol II density across transcription units of fission and budding yeast reveals large, previously uncharacterized, differences in transcription mechanisms between these divergent yeast. Although our data in *S. cerevisiae* corroborate the wealth of available information describing Pol II distributions in this system (Churchman and Weissman 2011; Jordán-Pla et al. 2014), applying the same technique in *S. pombe* has identified Pol II dynamics that resemble mechanisms pervasive in metazoans.

Paused and not paused genes in *S. pombe* exhibit distinct Pol II patterns around their +1 nucleosome

Having observed the distribution of RNA Pol II in fission yeast, we further investigated differences between paused and not paused genes. Paused and not paused genes, identified using the combined replicate data, were further refined as high-confidence paused and not paused genes by requiring that they are also significantly paused ($P < 0.05$) or not paused ($P > 0.95$) in each biological replicate (paused: $n = 656$; not paused: $n = 2088$). An example of PRO-seq profiles for a highly paused gene and a not paused gene are displayed in Figure 3, A and B, respectively. Average PRO-seq profiles centered on the observed TSS in *S. pombe* show a striking signal enrichment, specific to paused genes,

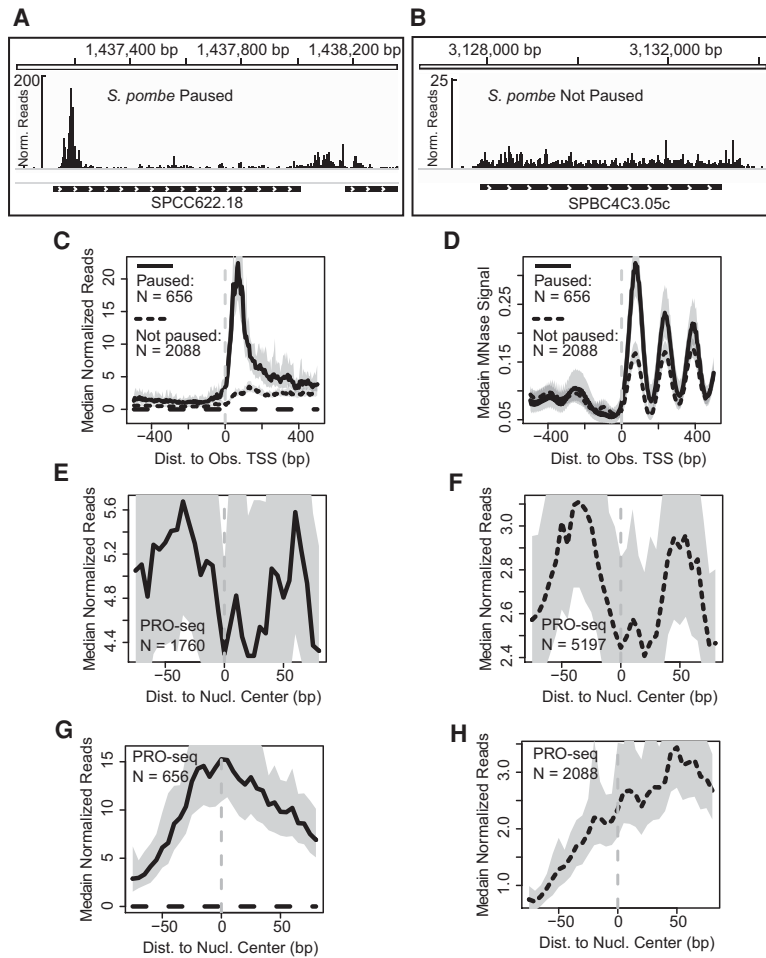


Figure 3. Pol II distributions at paused genes in *S. pombe* are coupled with increased nucleosome occupancy or positioning. (A,B) Browser images of PRO-seq read counts across *S. pombe* genes classified as either high-confidence paused (A) or not paused (B). (C) Median PRO-seq read counts within 1 kb upstream of and downstream from the observed TSS of paused and not paused genes. (D) Median MNase-seq read coverage within 1 kb upstream of and downstream from the observed TSS of paused and not paused genes. (E,F) PRO-seq signal around gene-body nucleosome centers for paused (E) and not paused genes (F). (G,H) PRO-seq signal around +1 nucleosome centers for paused (G) and not paused genes (H). For metagene plots in C–H, medians reflect 5-bp bins, and the 12.5% and 87.5% quantiles are shown as shaded regions. All panels represent profiles of combined wild-type biological replicates.

largely confined within the first 100 bp downstream from the TSS (Fig. 3C).

The transcription process is tightly coupled to the chromatin environment (Venkatesh and Workman 2015), and nucleosomes are known to form a barrier to transcriptional elongation, slowing or stalling RNA polymerases as they unwind the coiled DNA around the histone octamer (Hall et al. 2009). We analyzed an existing MNase-seq data set to explore the organization of nucleosomes relative to Pol II initiation and pause sites. Interestingly, average MNase profiles relative to the observed TSSs revealed a more highly phased nucleosome pattern for paused genes than for not paused genes (Fig. 3D). The largest difference in MNase-seq signal was at the +1 nucleosome position, possibly suggesting that this nucleosome is more precisely positioned or maintains higher occupancy within paused genes.

To assess the distribution of elongating polymerases around nucleosomes in *S. pombe*, we displayed average PRO-seq density

relative to previously identified nucleosome centers (Givens et al. 2012). For both paused and not paused genes, we parsed nucleosomes within gene boundaries based on their relative positions to the TSS. Nucleosomes were split into two groups: +1 nucleosomes (first nucleosome downstream from TSS) and gene-body nucleosomes (nucleosomes downstream from +1 center). Within the gene body, Pol II shows a similar profile around nucleosomes whether the gene is paused or not paused (Fig. 3E,F). The distributions around gene-body nucleosomes reflect known properties of transcription through chromatin, where Pol II slows down when it encounters the strong force of the DNA histone interaction, and speeds up when those contacts are broken as it approaches the DNA near the dyad (Hall et al. 2009). Plotting the average PRO-seq density within 160 bp around +1 nucleosome centers for paused and not paused genes revealed a remarkable difference in Pol II distribution relative to the dyad axis. Whereas Pol II shows no enrichment upstream of the +1 nucleosome center within not paused genes (Fig. 3H), Pol II is highly enriched around the dyad axis of the +1 nucleosome of paused genes (Fig. 3G). This data may suggest a role for histone-DNA interactions in specifying pause sites in *S. pombe*. Conversely, Pol II pausing may be independent of the nucleosome barrier, but restrict movement of the +1 nucleosome (Struhl and Segal 2013), inducing a more rigid positioning of downstream nucleosomes also (Jiang and Pugh 2009). The Pol II distributions around the +1 and gene-body nucleosomes (+2 to +4) indicate a general difference between the way Pol II negotiates its transit through the first and subsequent nucleosomes, as well as an interplay between paused Pol II and the +1 nucleosome that distinguishes paused and not paused genes.

sequent nucleosomes, as well as an interplay between paused Pol II and the +1 nucleosome that distinguishes paused and not paused genes.

Deletion of Spt4 increases Pol II density genome-wide in both budding and fission yeast

The striking resemblance of Pol II distributions in *S. pombe* to promoter-proximal pausing in metazoans led us to postulate that there is a pausing factor conserved in yeast playing a role in this process. Because both *S. cerevisiae* and *S. pombe* lack NELF subunits, the next logical candidate was the Spt4-Spt5 complex. To investigate functional divergence of this complex between *S. cerevisiae* and *S. pombe*, we prepared PRO-seq libraries from strains lacking Spt4. Wild-type and *spt4Δ* PRO-seq libraries from each species were prepared from two biological replicates. The absence of all

transcription and messenger RNA from the Spt4 locus confirms the deletion of the gene in each yeast (Supplemental Fig. S7A,B).

Anticipating the possibility of global transcriptional changes upon deletion of Spt4, we used an approach to normalize sequencing depth between samples based on a spike-in reference control. In this approach, we took advantage of the fact that >99% of mappable 36-nt sequences derived from *S. pombe* and *S. cerevisiae* can be distinguished. Similar to an approach developed previously for microarray analysis of labeled mRNA (Sun et al. 2012), we included an exogenous, whole yeast spike-in control in our experiments, which was carried out as follows: All samples and biological replicates to be compared were required to have precisely the same starting OD₆₀₀ before harvesting. Immediately prior to harvesting, a fixed volume of the divergent “spike-in” yeast culture was added to each sample. After preparing and sequencing PRO-seq libraries from the mixed samples, we aligned sequencing reads to a combined genome of *S. cerevisiae* and *S. pombe* with the requirement that sequences align to only one location to remove any ambiguous reads. The proportion of spike-in to experimental reads was highly similar between biological replicates (Supplemental Fig. S7G,H), and normalization factors were calculated from the relative number of uniquely mapped reads from the spiked-in genome between samples (Supplemental Table S1). To verify that this approach accurately scales the sequencing depth of each sample, we compared the normalized gene-body read densities between biological replicates, resulting in Spearman’s $\rho > 0.97$ (Supplemental Fig. S7C–F). Importantly, replicate comparisons are nearly symmetrical about the diagonal line, $x = y$, indicating this normalization accurately accounts for any technical differences in library preparation and sequencing.

To assess the effects of deletion of Spt4 on transcription in *S. cerevisiae* and *S. pombe*, we performed a differential expression analysis using read counts within the gene-body region. After spike-in normalization, we observed a global increase in the gene-body read density in *spt4Δ* strains in both yeasts. Using the DESeq2 R package (Love et al. 2014), we identified 1572 genes in *S. cerevisiae* (46% of filtered, active genes) and 2291 genes in *S. pombe* (67%) as having significantly higher gene-body Pol II levels ($P < 0.01$) (Fig. 4A,B). In contrast, only 47 (1.4%) and 156 (4.8%) genes were found to have significantly lower read counts in *S. cerevisiae* and *S. pombe*, respectively. On average, Pol II density within gene bodies of *S. pombe* and *S. cerevisiae* transcription units increased by a factor of 1.81 and 1.47, respectively. Looking more specifically at the YGR204W locus in *S. cerevisiae* (Fig. 4C) and the SPCC622.18-1 locus in *S. pombe* (Fig. 4D), we clearly observed a reproducible increase in the signal intensity within the gene body in *spt4Δ* biological replicates compared with wild type. Thus, the use of spike-in controls in the preparation of these data is invaluable. This approach circumvents the assumptions

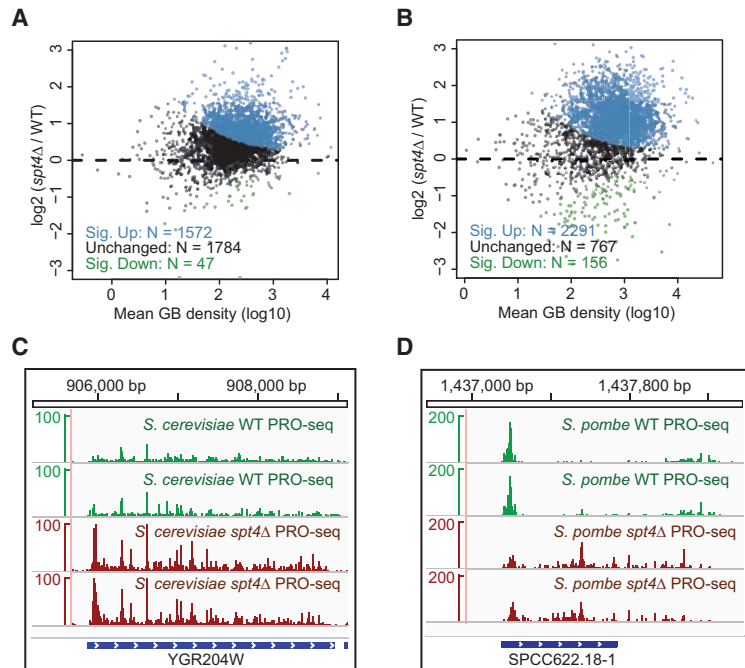


Figure 4. Deletion of Spt4 results in genome-wide increase in Pol II density within gene bodies of *S. cerevisiae* and *S. pombe*. (A,B) MA plots showing the DESeq2-based differential expression analysis of spike-in normalized PRO-seq read-counts within the gene bodies of *S. cerevisiae* (A) and *S. pombe* genes (B). (C,D) Example of *spt4Δ* affected gene in *S. cerevisiae* (C) and *S. pombe* (D). Browser tracks correspond to PRO-seq data derived separately from two biological replicates of WT (green) and *spt4Δ* (red) in budding and fission yeast.

of an equal number of RNA sequence reads per cell or that genes on average are unchanged, both of which are violated in mutants globally affecting transcription (Sun et al. 2012).

Our spike-in-normalized PRO-seq data also lends itself to a precise analysis of noncoding RNA. Antisense transcription within the gene body, as well as upstream, divergent transcription at promoters has become increasingly appreciated in yeast with the advent of high-throughput assays of RNA (Wilhelm et al. 2008; Neil et al. 2009; Xu et al. 2009). Evidence for alterations in such noncoding transcription have been reported for mutations in the elongation factor, Spt6, in *S. pombe* (DeGennaro et al. 2013). Subsets of genes used in the analysis of upstream, antisense or intragenic, antisense transcription were filtered to remove possible influence of nearby functional transcription. Unlike what has been observed in mammals (Core et al. 2008), composite profiles of PRO-cap and PRO-seq density around transcription start sites showed little indication of widespread upstream, divergent transcription at promoters in either yeast (Supplemental Fig. S8A,B,E,F), although *S. cerevisiae* tended to have higher ratios of upstream, antisense to downstream, sense transcription than *S. pombe* (cf. Supplemental Fig. S8C,G). Interestingly, intragenic, antisense to sense transcription ratios were more affected in *S. pombe* than in *S. cerevisiae* by the deletion of Spt4 (cf. Supplemental Fig. S8D,H). Indeed, we found that most levels of antisense transcription in fission yeast were unchanged (71%) or decreased (19%) in *spt4Δ*, suggesting that in *S. pombe*, antisense Pol II transcription may not depend on Spt4 in the same manner.

The general increase in sense-transcribing Pol II levels within a large proportion of genes likely reflects the ubiquitous role of Spt4 in positively affecting transcription processivity and elongation rate in these divergent yeasts (Wada et al. 1998; Rondón

et al. 2003; Hartzog and Fu 2013). Elevated levels of RNA abundance are often interpreted as a higher level of gene activity or increased initiation rates. However, increases in Pol II density could also reflect a reduction of Pol II elongation rate in the case of a fixed initiation rate (Danko et al. 2013; Jonkers et al. 2014).

Genome-wide increases in nascent transcription in *spt4Δ* are not reflected in mRNA abundance

To test the possibility of a global elongation rate defect in *S. cerevisiae* and *S. pombe* lacking Spt4, we investigated the effects of the mutation on mRNA abundance. We postulated that under conditions in which Pol II experiences a compromised elongation rate, with little change in recruitment and initiation, messenger RNA would continue to be made at roughly the same rate. To assay the relative abundance of all expressed transcripts, we prepared strand-specific poly-A selected RNA-seq libraries from the same wild-type and *spt4Δ* strains of budding and fission yeast. Again, to control for the possibility of genome-wide differences in transcript abundance per cell, we used the same culture-level divergent yeast spike-in approach as used for the PRO-seq experiments (Supplemental Table S1). The deletion of Spt4 was confirmed by absence of any reads from the locus (Supplemental Fig. S7A,B), and mRNA levels were found to be highly reproducible between biological replicates (Supplemental Fig. S9A–D).

Using the gene-body region to quantify gene activity from PRO-seq or mRNA-seq data, we observed a moderate correlation between nascent transcription and mRNA abundance in wild-type strains (*S. cerevisiae*: Spearman's $\rho=0.572$; *S. pombe*: Spearman's $\rho=0.574$) (Supplemental Fig. S9E,F). We next asked if changes in transcription, resulting from the deletion of Spt4, consequently affect mRNA abundance. Here, we compared the distribution of fold changes in gene-body read-density from *spt4Δ* to wild-type libraries between PRO-seq and mRNA-seq experiments. Although PRO-seq signal increases within gene bodies upon deletion of Spt4 in both species, changes in mRNA levels were either largely unchanged (*S. cerevisiae*) or slightly reduced (*S. pombe*) (Fig. 5A,B). Although the deletion of Spt4 altered mRNA abundance of some genes in both species (Supplemental Fig. S10A,B), these changes in gene-body read density were poorly correlated with changes in PRO-seq density (*S. cerevisiae*: Spearman's $\rho=0.248$; *S. pombe*: Spearman's $\rho=0.359$) (Supplemental Fig. S9G,H). Overall, the effects of deleting Spt4 on Pol II density within a gene is in stark contrast with the messenger RNA abundance, sup-

porting a model in which Spt4 increases the elongation rate of RNA Pol II transcription in vivo without necessarily altering Pol II recruitment or initiation.

To determine whether the effects of Spt4 deletion in *S. pombe* depended on the extent of observed promoter-proximal pausing, we binned genes in deciles based on the magnitude of their pausing index. Gene characteristics, such as expression level and termination index, appeared to be unrelated to pausing index (Supplemental Fig. S11A,B,E). Furthermore, there were no considerable relationships between how paused a gene is and the effect of Spt4 deletion on gene-body Pol II density (Supplemental Fig. S11C) or mRNA abundance (Supplemental Fig. S11D). We surmise that Spt4, as a complex with Spt5, influences various transcriptional features independent of the degree of promoter-proximal pausing.

Deletion of Spt4 reduces promoter-proximal pausing and elongation beyond the cleavage and polyadenylation signal in *S. pombe*

Although in metazoans, the Spt4-Spt5 complex acts as a pausing factor as well as a positive elongation factor (Wada et al. 1998; Wu et al. 2003), most research in yeast has focused on its role in positively influencing transcription (Hartzog and Fu 2013). To assess the existence of additional roles for the Spt4-Spt5 complex in yeast, we investigated the effect of Spt4 deletion on the distribution of Pol II at gene 5' and 3' ends in both species. To limit the effects of nearby genes, we restricted all analyses in this section to genes >1 kb in length and at least 1 kb away from neighboring genes on the same strand (*S. pombe*: $n=714$; *S. cerevisiae*: $n=816$). Plotting the median PRO-seq signal around the observed TSS of *S. pombe* genes for wild-type and *spt4Δ* strains revealed a dramatic shift of the Pol II density downstream in mutants (Fig. 6A). Whereas wild-type PRO-seq profiles display the pause-like distributions immediately adjacent to the TSS, *spt4Δ* strains lack this focused, promoter-proximal peak. Instead, they display a more gradual accumulation of density that is much greater than that of the wild type within the gene body. To evaluate the generality of this phenotype for all 714 genes, we generated heatmaps displaying the fold change in PRO-seq read density ($\log_2[spt4Δ/WT]$) within 10 bp bins from -1 kb to $+1$ kb around the observed TSS (Fig. 6B). We found that the majority of genes in the mutant show a decreased level of Pol II relative to wild type solely within the first ~ 150 bases. Beyond this promoter-proximal region,

spt4Δ strains showed heightened PRO-seq density compared with wild type. Furthermore, this mutant enrichment is most pronounced early within the gene body, as if early elongation complexes were leaking into the downstream regions rather than undergoing a pause. To quantify the significance of this shift in Pol II density, we calculated a pausing index for each gene in wild-type and *spt4Δ* strains. Consistent with the changes in average PRO-seq profiles in Figure 6A, we observed a highly significant decrease in average pausing index for all genes used in this analysis (Fig. 6C).

In light of the fact that Spt4-Spt5 tracks with elongating Pol II across the entire length of genes in eukaryotes

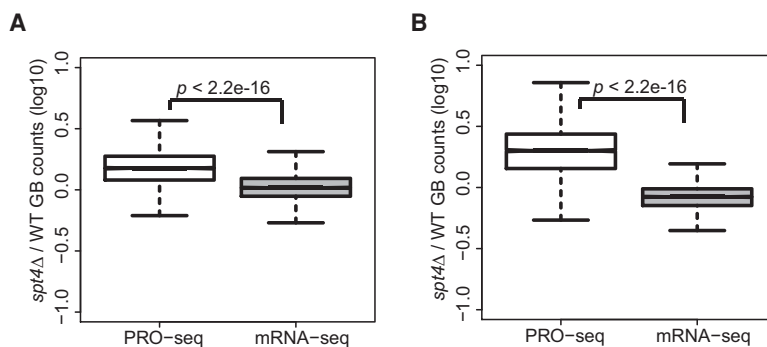


Figure 5. Global increase in Pol II density in *spt4Δ* does not result in increased transcript abundance in *S. cerevisiae* or *S. pombe*. (A,B) Box plots comparing the fold change in gene-body PRO-seq density resulting from the deletion of Spt4 with the corresponding change in poly-A selected RNA-seq in *S. cerevisiae* (A) and *S. pombe* (B). *P*-values represent the results of Student's *t*-test.

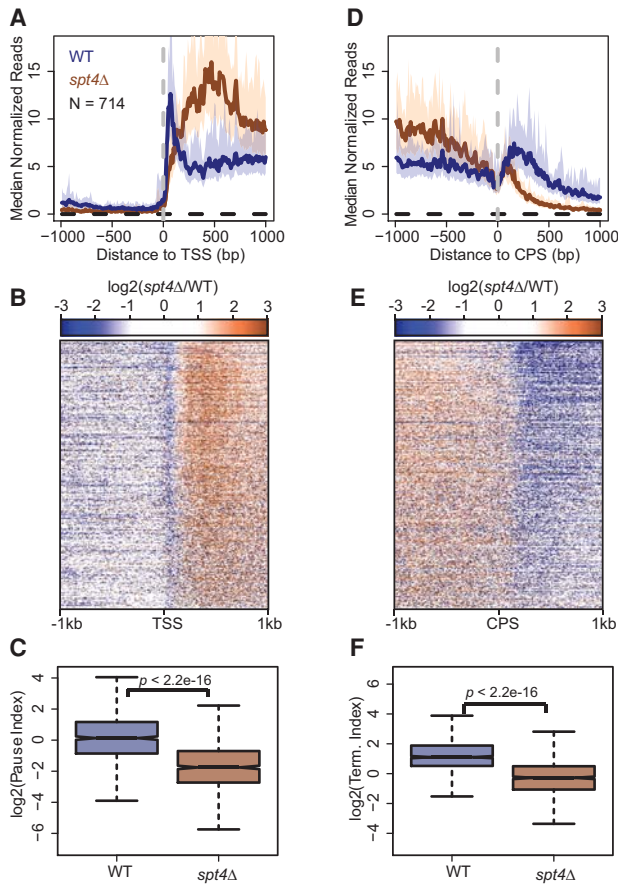


Figure 6. 5' and 3' ends of genes exhibit a loss of Pol II density as a result of Spt4 deletion in *S. pombe*. (A,D) Median PRO-seq signal around the observed TSS (A) or CPS (D) of active, filtered genes longer than 1 kb and separated from the boundaries of neighboring genes on the same strand by at least 1 kb. Medians reflect 10-bp bins, and the 12.5% and 87.5% quantiles are shown in light shaded regions. (B,E) Heatmaps of \log_2 fold change of mutant versus wild type per 10-bp bin around the TSS (B) or CPS (E) for all genes used in A and D. Genes within heatmaps are sorted by decreasing amount of wild-type PRO-seq reads within the first 500 bp downstream from the TSS. (C) Box plots showing the distribution of pausing index values for WT and *spt4Δ* in *S. pombe*. (F) Box plots showing the distribution of termination index values for WT and *spt4Δ* in *S. pombe*. P-values represent the results of Student's *t*-test.

(Mayer et al. 2010; Rahl et al. 2010), we were curious about the effect of Spt4 deletion on transcription around the cleavage and polyadenylation signal (CPS). Similar to our analysis of PRO-seq signal around the observed TSS, we calculated the median read density around the annotated CPS for all 714 genes. Strikingly, although the *spt4Δ* *S. pombe* strain shows elevated levels of elongating Pol II leading up to the CPS, immediately after this element, we begin to observe a much more rapid decline in polymerase density in the mutants relative to wild type (Fig. 6D). Looking at fold change in polymerase density across this surrounding region for each individual gene shows that this decreased transcription beyond the CPS in *spt4Δ* is broadly observable (Fig. 6E). Finally, we assessed the significance of this change in post-CPS transcription by calculating a termination index for each gene in wild-type and *spt4Δ* *S. pombe* strains. The combined effect of higher read density within the gene body and lower signal beyond the CPS in *spt4Δ*, produced a significantly lower average termination index relative to wild

type (Fig. 6F). Thus, in fission yeast, transcription beyond the CPS is dramatically reduced in the absence of Spt4.

Profiles of transcription elongation differ greatly between budding and fission yeast, yet upon deletion of Spt4, both show a similar increase in Pol II levels within gene bodies. To address whether *S. cerevisiae* is similarly affected by deletion of Spt4 at the 5' and 3' gene boundaries, we performed the same set of analyses on our PRO-seq samples from wild-type and *spt4Δ* strains (Supplemental Fig. S12). Surprisingly, genes in *S. cerevisiae* exhibit an enormous increase in PRO-seq density at the 5' ends of genes upon deletion of Spt4 (Supplemental Fig. S12A,B). This is precisely the opposite effect of that observed in *S. pombe* and results in a significant increase in average pausing index (Supplemental Fig. S12C). Beyond the CPS, somewhat similar to *S. pombe*, there appears to be a slight reduction in PRO-seq signal as a result of Spt4 deletion (Supplemental Fig. S12D,E). However, this modest reduction does not register as a significant reduction in average termination index (Supplemental Fig. S12F).

The highly contrasting effect of Spt4 deletion in budding and fission yeast at the 5' ends of genes indicates divergent roles for this elongation factor in early transcription elongation. Whereas Spt4 appears to be solely required to enhance productive elongation in *S. cerevisiae*, our results in *S. pombe* reveal an additional function of Spt4 in promoting early Pol II pausing. Perhaps as evidence of shared defects in elongation, Pol II requires Spt4 for extended transcription beyond the CPS in both species.

Discussion

We have adapted the Precision Run-On sequencing (PRO-seq) and 5' cap sequencing (PRO-cap) protocols for use in both budding and fission yeast. Precisely mapping 5' ends of capped nascent RNAs reveals highly preferred transcription start sites for a large fraction of coding genes and noncoding transcription units, enhancing current annotations for many applications. Our PRO-seq profiles provide comprehensive base-pair resolution maps of elongating RNA polymerases genome-wide in two evolutionarily distant yeast strains. These data provide the highest resolution, strand-specific look at transcription across the fission yeast genome to date.

The detailed transcriptional landscapes of *S. cerevisiae* and *S. pombe* prove to be very dissimilar, with metazoan-like features of transcription unique to fission yeast. Consistent with previous investigations of transcription in *S. cerevisiae*, Pol II across annotated genes is largely confined within the well-annotated 5' and 3' gene boundaries, with a relatively uniform distribution across transcription units (Steinmetz et al. 2006). In contrast, we found *S. pombe* to exhibit elongation continuing well beyond the 3' cleavage and polyadenylation signal of most genes. Immediately downstream from the *S. pombe* 3' CPS, Pol II appears to experience an elongation rate-reducing step, resulting in heightened polymerase density over a broad region. Although delayed termination is less well characterized in yeast, studies in a wide variety of metazoans have revealed similar patterns (Core et al. 2008, 2012; Krusi et al. 2013; Jonkers et al. 2014). Our findings support the conclusion that the kinetics of transcription termination and its regulation in *S. pombe* is more similar to that of metazoans than *S. cerevisiae*.

Another intriguing characteristic of transcription profiles in fission yeast is the prominent enrichment in Pol II immediately downstream from the TSS, resembling promoter-proximal pausing in metazoans (Fuda et al. 2009; Adelman and Lis 2012). We found that 28% of our filtered, actively transcribed *S. pombe* genes could

be classified as significantly paused when considering the first 100 transcribed bases. These paused genes tend to show a much more highly positioned or high-occupancy +1 nucleosome, with Pol II density peaking near the dyad axis. Interestingly, this near-overlapping association of Pol II with the +1 nucleosome dyad is unlike that observed in *Drosophila*, in which most pausing occurs upstream of the +1 nucleosome (Kwak et al. 2013). This difference could be the result of NELF, which might have evolved in metazoans to capture Pol II rapidly in a more promoter-proximal paused state (Li et al. 2013) and thereby strengthen the potential for upstream factor regulation. Considering the proximity of the +1 nucleosome to the TSS and the tight relationship between the Pol II and nucleosome occupancy within this region in *S. pombe*, it is tempting to speculate that they are connected (Fig. 3G). On one hand, the first nucleosome may act as a barrier to early transcription elongation, causing a slower rate of transcription shortly after initiation (Weber et al. 2014). Alternatively, a frequently paused polymerase at this position could restrict the localization of the +1 nucleosome (Struhl and Segal 2013). A rigidly positioned +1 nucleosome could then restrict the mobility of downstream, gene-body nucleosomes, resulting in the observed increase in phasing (Fig. 3D; Jiang and Pugh 2009).

In metazoans, two complexes, NELF and DSIF (Spt4-Spt5), are both critical for pausing (Wu et al. 2003). The observation of pause-like distributions in fission yeast is thus curious because *S. pombe* lack any recognizable NELF subunits that might help to establish a pause (Narita et al. 2003). Both *S. pombe* and *S. cerevisiae*, however, possess Spt4-Spt5 and CDK9 (Bur1) homologs (Hartzog and Fu 2013). Intriguingly, deleting Spt4 in *S. pombe* causes a dramatic reduction in Pol II density within the promoter-proximal region. Moreover, heightened Pol II density shifts into the gene body, suggesting Spt4, most likely in complex with Spt5, is necessary for preventing the premature release of the Pol II from these sites during early transcription elongation. This possibility is supported by findings in metazoans indicating pausing is only modestly reduced upon RNAi knock down of NELF (Core et al. 2012; Gilchrist et al. 2012).

The apparent similarities between *S. pombe* and metazoan pausing raise many additional mechanistic and functional questions. As in metazoans, Spt4-Spt5-mediated promoter-proximal pausing in *S. pombe* might provide a regulatory step or checkpoint to ensure proper capping of the nascent RNA or modification of the elongation complex (Pei and Shuman 2002; Schwer et al. 2009; Viladevall et al. 2009; Schneider et al. 2010). Indeed, CDK9 and its Cyclin partner Pch1 in fission yeast interact with nascent RNA 5' capping machinery (Pei et al. 2003, 2006). Alternatively, paused Pol II could be capable of holding certain *S. pombe* genes poised for expression, synchronizing the expression of sets of genes, or allowing the integration of regulatory signals at distinct steps of recruitment and release of paused Pol II (Adelman and Lis 2012). In the future, determining whether pause release in *S. pombe* relies on the kinase activity of CDK9 will be imperative for comparing this phenomenon with that of metazoans.

Because Spt4 is a subunit of a ubiquitous elongation factor and can affect transcription broadly, we found that proper normalization of our PRO-seq and mRNA-seq data was critical for accurately assessing genome-wide changes in transcription between wild-type and mutant yeast strains. Using a cell culture spike-in approach, we observed global increases in sense-transcribing RNA polymerase density across genes in *spt4Δ* strains of *S. cerevisiae* and *S. pombe* relative to wild type. However, effects on antisense transcription appeared to be species specific. In *S. cerevisiae*, anti-

sense-to-sense ratios were minimally affected by the deletion of Spt4. In contrast, the relative amount of antisense transcription in *S. pombe* appeared to globally decrease, perhaps indicating little influence of Spt4-Spt5 over cryptic, intragenic, antisense transcription events in fission yeast. We found that mRNA levels in budding and fission were largely either unchanged or slightly decreased in the absence of Spt4, despite the increase in gene-body Pol II (Supplemental Fig S9). One explanation could be that RNAs may in fact be increasingly produced in *spt4Δ* strains, yet buffered through surveillance mechanisms such as preRNA degradation. However, in light of several lines of evidence, we favor a hypothesis in which, in the absence of Spt4, Pol II exhibits a reduced elongation rate. First, early experiments demonstrated that recombinant human Spt4 and Spt5, but not Spt5 alone, could increase rates of in vitro transcription reactions (Wada et al. 1998). Second, reduced elongation rates are known to decrease the distance transcribed by Pol II beyond the 3' cleavage and polyadenylation signal (CPS) (Fong et al. 2015). We found that in *S. pombe*, transcription termination occurs within a much shorter distance from the CPS in the absence of Spt4 (this was also observed to a lesser extent in *S. cerevisiae*, but may be more difficult to detect because termination occurs normally only a short distance from the CPS). Compatible with the torpedo model of transcription termination (Connelly and Manley 1988), the 5'-3' exonucleases, Dhp1 (*S. cerevisiae*: Rat1), may be better able to catch up to and displace the slower Pol II from the DNA. Importantly, since Spt4 acts in complex with Spt5 to affect elongation, we surmise that many of the phenotypes we have observed in *spt4Δ* strains are a consequence of perturbation of this elongation factor complex and its association with Pol II, rather than the independent action of Spt4. Further experiments are required to distinguish independent roles of each subunit of the Spt4-spt5 complex in transcription elongation.

Apart from having a role in maintaining transcription elongation rates, the Spt4-Spt5 complex is also thought to affect the processivity of Pol II (Mason and Struhl 2005; Hartzog and Fu 2013). The increase in gene-body Pol II in *spt4Δ* strains does not preclude the possibility of defects in Pol II processivity, but rather indicates that the more dominant effect of Spt4 deletion is on Pol II elongation rate. The modest decreases in levels of many mRNAs in *S. pombe* could reflect a combination of decreased rates of transcription and reduced processivity by Pol II in the absence of Spt4.

Using PRO-seq, adapted for determining the transcription profiles in budding and fission yeast, provides novel insight into the evolution of mechanisms governing their transcriptional landscapes. In this study, we identified divergence in the role of Spt4-Spt5 between budding and fission yeast. *S. cerevisiae* appears to rely on Spt4 primarily to maintain a steady and rapid elongation rate. In contrast, *S. pombe* additionally possesses a previously undescribed role for this complex in slowing or stalling Pol II between the TSS and the +1 nucleosome. The ability with which PRO-seq can be applied in yeast ultimately enhances our ability to combine genome-wide snapshots of transcription elongation with the powerful genetic tools inherent to these model systems.

Methods

Yeast strains

S. cerevisiae strains, WT and *spt4Δ*, are in the background W303-1a and W303-1a, respectively. WT and *spt4Δ* strains in *S. pombe* are in the background 972 h-.

Precision Run-On sequencing in yeast

The protocol described in Kwak et al. (2013) was modified slightly to accommodate the physiology of yeast (García-Martínez et al. 2004). Instead of nuclei isolation, both *S. cerevisiae* and *S. pombe* were permeabilized to allow for the efflux of endogenous NTPs and uptake of biotin-11-NTPs (Epicentre). Yeast cultures, grown overnight, were diluted to an $OD_{600} = 0.1$ and grown to mid-log phase ($OD_{600} = 0.5-0.6$) in either YPD (*S. cerevisiae*) or YES media (*S. pombe*). Ten-milliliter cultures of equal cell concentration were spun down and media was removed. Cells were then resuspended in 10 mL cold H_2O . Cultures were pelleted again, resuspended in 10 mL 0.5% sarkosyl at 4°C, and incubated on ice for 20 min. Cells were then spun at a reduced RCF (400g) for 5 min at 4°C. After removal of permeabilization buffer, yeast pellets were resuspended in 120 μ L of 2.5 \times transcription buffer (50 mM Tris-HCl, pH 7.7, 500 mM KCl, 12.5 mM $MgCl_2$) with 6 μ L 0.1 M DTT and 3.75 μ L of each 1 mM biotin-NTP added immediately before use. After suspending yeast in the aforementioned transcription mix, the volume was brought to 285 μ L with DEPC-treated H_2O . Finally, 15 μ L 10% sarkosyl was added, and the reaction was placed at 30°C and allowed to run on for 5 min. RNA was extracted using a hot phenol approach (Collart and Oliviero 2001); after the run-on reaction, cells were pelleted at 400g for 5 min at 4°C and quickly resuspended in 500 μ L acid phenol. An equal volume of AES buffer (50 mM NaAc pH 5.3, 10 mM EDTA, 1% SDS) was added and placed for 5 min at 65°C with periodic vortexing, followed by 5 min on ice. Two hundred microliters chloroform was added and mixed followed by centrifugation at 14,000g for 5 min (4°C). To the aqueous layer (200 mM), 3 M NaOAc was added followed by ethanol precipitation with a 3 \times volume of 100% ethanol. The RNA pellet was air dried before being resuspended in 20 μ L DEPC-treated water. The PRO-seq or PRO-cap protocol (Kwak et al. 2013) was then followed, beginning with base-hydrolysis, through to sequencing.

Spike-in approach for Run-On (PRO-seq) or RNA isolation (RNA-seq)

Preparation of libraries with a spiked-in reference organism was carried out by first growing a culture of the reference organism separate from the experimental cultures to $OD_{600} = 0.5$. After experimental samples had reached the desired OD_{600} and all samples were adjusted to have precisely equal numbers of cells, a fixed amount of the reference culture was added to each sample for a given experiment. All samples and biological replicates for a given experiment were prepared side by side such that all replicates received exactly the same amount of reference culture. In the case of either RNA-seq or PRO-seq, the spike-in approach described above was applied in order to control for every step of the sequencing library preparation during which technical variation may be introduced, beginning with the pelleting of cell cultures.

PRO-seq/cap sequencing alignment

Raw sequencing files were processed by first removing reads that do not pass Illumina's quality filter. Full-length or partial adapter sequences (5'-TGGAATTCTCGGGTGCCAAGG-3') were then removed using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were then trimmed to a maximum length of 36 nt (minimum: 15 nt) and, in the case of PRO-seq, reverse-complemented using the FASTX-Toolkit. Finally, all nascent sequencing alignments were performed using Bowtie (version 1.0.0) (Langmead et al. 2009). Processed reads were aligned allowing for two mismatches, while requiring a unique alignment to the genome. For PRO-seq experiments, due to the existence of reads

originating from separate species resulting from our spike-in procedure, reads were initially aligned to ribosomal DNA sequences from each species. Because of the high level of conservation of these sequences, ribosomal mapping reads were not considered in a species-specific manner. The remaining reads were then mapped to a combined genome consisting of all chromosomes from both *S. cerevisiae* (sacCer3=S288C_reference_genome_R64-1-1_20110203) and *S. pombe* (version: ASM294v2) facilitating the removal of reads with an ambiguous origin. Although we are aligning reads of various lengths (15–36 nt), the majority of reads in each library were 36 nt long. The method of aligning to a combined *S. pombe* and *S. cerevisiae* genome prevents genome-indistinguishable, shorter reads from being included in downstream analysis. Unique reads were parsed based on their species of origin. Bedgraph files were created by recording only the most 3' base of each read, which represents the position of the Pol II active site, for PRO-seq data, or the most 5' base of reads from PRO-cap data. The counts at each position in the bedgraph file were normalized based on the relative amount of reads aligning to the spike-in genome (PRO-seq), or based on the number of reads per million mapped reads (PRO-cap) (for displaying in genome browsers). Bedgraphs were ultimately converted to bigWig-formatted files for downstream analysis.

Preparation and alignment of RNA-seq libraries

Total RNA was isolated from each sample culture (with spike-in reference) using the hot phenol approach followed by ethanol precipitation. The integrity of each sample was assessed on a denaturing poly-acrylamide gel (7 M urea PAGE), and the quantity of total RNA was measured. Three micrograms of total RNA from each sample was used for library preparation. Sequencing libraries were prepared using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina). Raw sequencing of 75 nt reads was conducted, and raw sequencing files were aligned using the STAR alignment suite (version 2.3.0) (Dobin et al. 2013). Uniquely mapping reads with no more than two mismatches were reported and used in this analysis. Sequences were aligned to a combined genome of *S. pombe* (version: ASM294v2) and *S. cerevisiae* (sacCer3=S288C_reference_genome_R64-1-1_20110203.tgz), and reads were separated by species-specific chromosome names. Bedgraph files were created by recording only the most 3' base of each read. For calculating the number of reads within desired regions, custom scripts were used to count the amount of read 3' ends within provided coordinates. The counts at each position in a bedgraph file were normalized based on the relative amount of reads aligning to the spike-in genome.

Mappability tracks

When determining the read density for genes or other transcription units, the length of the region was defined as the number of "mappable" bases within the unit. Uniquely mappable positions for 36-mer sequence alignments were determined by first computing all possible 36-mer sequences across the genome of interest and then aligning them back to the genome. By restricting the reported hits to perfect and unique alignments, we were able to generate mappability files that describe the mappable regions of the genome. In the case of samples receiving spiked-in libraries (i.e., *S. pombe* cultures spiked into *S. cerevisiae*), mappability tracks were created by taking all possible 36-mer sequences derived from the *S. cerevisiae* genome and aligning them to a species-combined genome. By aligning *S. cerevisiae* 36-mer sequences to a composite genome and reporting only unique hits, we are able to identify all 36-mer sequences that (1) uniquely align to the *S.*

cerevisiae genome and are thus not repeat elements, and (2) are distinguishable from the spiked-in genome.

Gene sets

For *S. cerevisiae*, the original (annotated) gene list consisted of all transcription units identified by Pelechano et al. (2013), considering only the longest major transcript isoform for each model. For *S. pombe*, a general transfer format (version: ASM294v1.16) file for gene annotations was downloaded from ftp://ftp.ensembl.genomes.org/pub/fungi/release-16/gtf/schizosaccharomyces_pombe/. Gene models used in this study were filtered and refined based on the following criteria. Genes had to be nonoverlapping with any other annotated transcription units on the same strand. Gene transcription was defined as the normalized read depth across mappable bases within the gene body. A background read depth was estimated based on the read density within a set of intergenic regions. Genes were only considered active if their gene-body read density was greater than expected ($P < 0.01$) based on a Poisson distribution with a background sequencing rate. Inactive genes were omitted from downstream analysis. Moreover, only genes for which we were able to identify an observed TSS were used in our analysis—*S. cerevisiae*: $n = 4398$ (84% of not-overlapping transcription units) and *S. pombe*: $n = 4674$ (80% of not-overlapping transcription units). In order to avoid instances in which upstream transcription interferes with a gene's perceived transcription, we calculated a run-through index for each gene. The run-through index was defined as the ratio of reads within the upstream window from -300 bases to TSS, relative to a downstream region $+250$ to $+550$. Genes were only used if their upstream run-through index was less than 1. Although the figures in this work are based on PRO-cap-based reannotation of early annotations (ASM294v1.16), we found no effect of using an updated set of annotations for *S. pombe* (ASM294v2.30) on our results and figures. Only 23 of the shared genes appear to have 5' or 3' coordinate changes. All genes for which we called an observed TSS are listed in Supplemental Table S2 (*S. cerevisiae*) and Supplemental Table S3 (*S. pombe*). In Supplemental Tables S2 and S3, normalized, background-subtracted PRO-cap signal for each gene is provided, and filtered genes used in this work are distinguished from unused genes.

Observed TSS identification

Our approach for calling transcription start sites with PRO-cap data is modeled after a similar analysis used for Start-seq data in mouse (Scruggs et al. 2015). Using the starting lists of TSS annotations, we developed an algorithm to search for the position within ± 250 bases of each annotation possessing the highest PRO-cap read count on the corresponding strand. Importantly, we assess the background signal of our PRO-cap libraries with a TAP-minus replicate experiment. Thus, to identify the true most preferred base near the annotated TSS, we specifically looked for the position with the highest background-subtracted signal. In cases in which two or more bases shared the same signal intensity, the base closest to the annotation was chosen. Although situations arose in which multiple bases appeared to be used as TSSs, we restricted our analysis to the most preferred site to simplify downstream analysis. An observed TSS was required to have at least five reads more than the relative background signal.

Pausing and termination index analysis

Pausing index was quantified as the read density per mappable base within the first 100 bp downstream from the observed TSS divided by the read density per mappable base within the gene body

(TSS + 200 to CPS). The termination index was calculated by identifying the 100-bp window 0 to +500 bp downstream from the CPS with the maximum read density per mappable base and dividing by the read density per mappable base within the gene body. For determining the number of paused genes within each WT yeast strain, we used the combined data from biological replicates to calculate pausing index. We further tested the significance of pausing within each gene by using Fisher's exact test to assess the likelihood of obtaining each ratio by chance. Using Bonferroni's correction, Fisher's exact test P -values were adjusted based on the number of genes analyzed. We called a gene paused if it had an adjusted P -value < 0.01 and not paused if $P > 0.99$. For calling high-confidence paused and not paused genes, we further restricted the respective gene sets based on analysis of individual biological replicates. For both biological replicates, a gene was required to have an adjusted P -value < 0.05 (high-confidence paused) or $P > 0.095$ (high-confidence not paused).

MNase data

Aligned and processed MNase-seq files were downloaded from GEO (*S. pombe* accession: GSE49575; *S. cerevisiae* accession: GSM1143089) for use in our analysis (DeGennaro et al. 2013; Hu et al. 2014). Nucleosome centers were previously identified by independent groups, i.e., *S. pombe* (Givens et al. 2012) and *S. cerevisiae* (Weiner et al. 2010). To ensure that the nucleosome centers used accurately reflected the distinct MNase-seq data, we assessed average MNase profiles relative to the centers. We found that the MNase-seq profiles of data used in this work corresponded nearly perfectly (within ~ 5 bp from MNase peaks on average) to the lists of nucleosome centers.

Data access

The raw and processed sequencing files from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE76142.

Acknowledgments

Research reported in this publication was supported by the National Institutes of Health (R01GM25232). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Cornell Genomics Facility for overseeing the sequencing of our libraries.

Author contributions: G.T.B. conducted the experiments and data analysis; G.T.B. and J.T.L. conceived the experiments and analysis and wrote the manuscript; and V.G.C. and I.X.W. devised early PRO-seq adaptations for yeast and generated preliminary data in the strains used.

References

- Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**: 720–731.
- Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, Fong N, Kiemlele L, Hansen K, Davis R, Lykke-Andersen J, et al. 2012. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol Cell* **46**: 311–324.
- Castel SE, Ren J, Bhattacharjee S, Chang AY, Sánchez M, Valbuena A, Antequera F, Martienssen RA. 2014. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. *Cell* **159**: 572–583.
- Churchman LS, Weissman JS. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**: 368–373.

- Collart MA, Oliviero S. 2001. Preparation of yeast RNA. *Curr Protoc Mol Biol*. doi: 10.1002/0471142727.mb1312s23.
- Connelly S, Manley JL. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev* **2**: 440–452.
- Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**: 1845–1848.
- Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. 2012. Defining the status of RNA polymerase at promoters. *Cell Rep* **2**: 1025–1035.
- Coudreuse D, van Bakel H, Dewez M, Soutourina J, Parnell T, Vandenhaute J, Cairns B, Werner M, Hermand D. 2010. A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in *S. pombe*. *Curr Biol* **20**: 1053–1064.
- Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14**: 1188–1190.
- Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. 2013. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* **50**: 212–222.
- DeGennaro CM, Alver BH, Marguerat S, Stepanova E, Davis CP, Bähler J, Park PJ, Winston F. 2013. Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol Cell Biol* **33**: 4779–4792.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
- Dujon B. 2010. Yeast evolutionary genomics. *Nat Rev Genet* **11**: 512–524.
- Eser P, Wachutka L, Maier KC, Demel C, Boroni M, Iyer S, Cramer P, Gagneur J. 2016. Determinants of RNA metabolism in the *Schizosaccharomyces pombe* genome. *Mol Syst Biol* **12**: 857.
- Fong N, Brannan K, Erickson B, Kim H, Cortazar MA. 2015. Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol Cell* **60**: 256–267.
- Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**: 186–192.
- García-Martínez J, Aranda A, Pérez-Ortín JE. 2004. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell* **15**: 303–313.
- Gilchrist DA, Fromm G, dos Santos G, Pham LN, McDaniel IE, Burkholder A, Fargo DC, Adelman K. 2012. Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. *Genes Dev* **26**: 933–944.
- Givens RM, Lai WK, Rizzo JM, Bard JE, Mieczkowski PA, Leatherwood J, Huberman JA, Buck MJ. 2012. Chromatin architectures at fission yeast transcriptional promoters and replication origins. *Nucleic Acids Res* **40**: 7176–7189.
- Gromak N, West S, Proudfoot NJ. 2006. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol Cell Biol* **26**: 3986–3996.
- Hahn S, Young ET. 2011. Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* **189**: 705–736.
- Hall MA, Shundrovsky A, Bai L, Fulbright RM, Lis JT, Wang MD. 2009. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat Struct Mol Biol* **16**: 124–129.
- Hartzog GA, Fu J. 2013. The Spt4–Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim Biophys Acta* **1829**: 105–115.
- Hu Z, Chen K, Xia Z, Chavez M, Pal S, Seol JH, Chen CC, Li W, Tyler JK. 2014. Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. *Genes Dev* **28**: 396–408.
- Jiang C, Pugh BF. 2009. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**: 161–172.
- Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**: e02407.
- Jordán-Pla A, Gupta I, de Miguel-Jiménez L, Steinmetz LM, Chávez S, Pelechano V, Pérez-Ortín JE. 2014. Chromatin-dependent regulation of RNA polymerases II and III activity throughout the transcription cycle. *Nucleic Acids Res* **43**: 787–802.
- Keogh MC, Podolny V, Buratowski S. 2003. Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. *Mol Cell Biol* **23**: 7005–7018.
- Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ. 2013. Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. *eLife* **2**: e00808.
- Kwak H, Fuda NJ, Core LJ, Lis JT. 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**: 950–953.
- Laitem C, Zaborowska J, Isa NF, Kufs J, Dienstbier M, Murphy S. 2015. CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II-transcribed genes. *Nat Struct Mol Biol* **22**: 396–403.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Li J, Liu Y, Rhee HS, Ghosh SKB, Bai L, Pugh BF, Gilmour DS. 2013. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. *Mol Cell* **50**: 711–722.
- Li H, Hou J, Bai L, Hu C, Tong P, Kang Y, Zhao X, Shao Z. 2015. Genome-wide analysis of core promoter structures in *Schizosaccharomyces pombe* with DeepCAGE. *RNA Biol* **12**: 525–537.
- Love JD, Vivino AA, Minton KW. 1985. Detection of low-level gene induction using *in vitro* transcription heat shock genes. *Gene Anal Tech* **2**: 100–107.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.
- Mason PB, Struhl K. 2005. Distinction and relationship between elongation rate and processivity of RNA polymerase II *in vivo*. *Mol Cell* **17**: 831–840.
- Maxwell CS, Kruesi WS, Core LJ, Kurhanewicz N, Waters CT, Lewarch CL, Antoshechkin I, Lis JT, Meyer BJ, Baugh LR. 2014. Pol II docking and pausing at growth and stress genes in *C. elegans*. *Cell Rep* **6**: 455–466.
- Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* **17**: 1272–1278.
- McKinlay A, Araya CL, Fields S. 2011. Genome-wide analysis of nascent transcription in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **1**: 549–558.
- Min IM, Waterfall JJ, Core LJ, Munroe RJ, Schimenti J, Lis JT. 2011. Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev* **25**: 742–754.
- Narita T, Yamaguchi Y, Yano K, Sugimoto S, Chanarat S, Wada T, Kim DK, Hasegawa J, Omori M, Inukai N, et al. 2003. Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. *Mol Cell Biol* **23**: 1863–1873.
- Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**: 1038–1042.
- Pei Y, Shuman S. 2002. Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *J Biol Chem* **277**: 19639–19648.
- Pei Y, Schwer B, Shuman S. 2003. Interactions between fission yeast Cdk9, its cyclin partner Pch1, and mRNA capping enzyme Pct1 suggest an elongation checkpoint for mRNA quality control. *J Biol Chem* **278**: 7180–7188.
- Pei Y, Du H, Singer J, Stamour C, Granitto S, Shuman S, Fisher RP. 2006. Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIF-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 *in vivo*. *Mol Cell Biol* **26**: 777–788.
- Pelechano V, Steinmetz LM. 2013. Gene regulation by antisense transcription. *Nat Rev Genet* **14**: 880–893.
- Pelechano V, Chávez S, Pérez-Ortín JE. 2010. A complete set of nascent transcription rates for yeast genes. *PLoS One* **5**: e15442.
- Pelechano V, Wei W, Steinmetz LM. 2013. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**: 127–131.
- Peterlin BM, Price DH. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23**: 297–305.
- Ping YH. 2000. DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J Biol Chem* **276**: 12951–12958.
- Porrua O, Libri D. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**: 190–202.
- Proudfoot NJ. 1989. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci* **14**: 105–110.
- Ptashne M, Gann A. 1997. Transcriptional activation by recruitment. *Nature* **386**: 569–577.
- Rahl PB, Lin CY, Seila AC, Flynn RA, McQuinn S. 2010. c-Myc regulates transcriptional pause release. *Cell* **141**: 432–445.
- Rhee HS, Pugh BF. 2012. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**: 295–301.
- Rodríguez-Gil A, García-Martínez J, Pelechano V, Muñoz-Centeno Mde LC, Geli V, Pérez-Ortín JE, Chávez S. 2010. The distribution of active RNA polymerase II along the transcribed region is gene-specific and controlled by elongation factors. *Nucleic Acids Res* **38**: 4651–4664.
- Rondón AG, García-Rubio M, González-Barrera S, Aguilera A. 2003. Molecular evidence for a positive role of Spt4 in transcription elongation. *EMBO J* **22**: 612–620.
- Sansó M, Lee KM, Viladevall L, Jacques PÉ, Pagé V, Nagy S, Racine A, St Amour CV, Zhang C, Shokat KM, et al. 2012. A positive feedback loop links opposing functions of P-TEFb/Cdk9 and histone H2B

- ubiquitylation to regulate transcript elongation in fission yeast. *PLoS Genet* **8**: e1002822.
- Schneider S, Pei Y, Shuman S, Schwer B. 2010. Separable functions of the fission yeast Spt5 carboxyl-terminal domain (CTD) in capping enzyme binding and transcription elongation overlap with those of the RNA polymerase II CTD. *Mol Cell Biol* **30**: 2353–2364.
- Schwer B, Schneider S, Pei Y, Aronova A, Shuman S. 2009. Characterization of the *Schizosaccharomyces pombe* Spt5-Spt4 complex. *RNA* **15**: 1241–1250.
- Scruggs BS, Gilchrist DA, Nechaev S, Muse GW. 2015. Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin. *Mol Cell* **58**: 1101–1112.
- Stargell LA, Struhl K. 1996. Mechanisms of transcriptional activation *in vivo*: two steps forward. *Trends Genet* **12**: 311–315.
- Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* **24**: 735–746.
- Struhl K, Segal E. 2013. Determinants of nucleosome positioning. *Nat Struct Mol Biol* **20**: 267–273.
- Sun M, Schwalb B, Schulz D, Pirkl N, Etzold S, Larivière L, Maier KC, Seizl M, Tresch A, Cramer P. 2012. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res* **22**: 1350–1359.
- Swanson MS, Malone EA, Winston F. 1991. *SPT5*, an essential gene important for normal transcription in *Saccharomyces cerevisiae*, encodes an acidic nuclear protein with a carboxy-terminal repeat. *Mol Cell Biol* **11**: 3009–3019.
- Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **13**: 2129–2141.
- Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**: 178–189.
- Viladevall L, St Amour CV, Rosebrock A, Schneider S, Zhang C, Allen JJ, Shokat KM, Schwer B, Leatherwood JK, Fisher RP. 2009. TFIIF and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast. *Mol Cell* **33**: 738–751.
- Wada T. 1998. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription *in vitro*. *EMBO J* **17**: 7395–7403.
- Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, et al. 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**: 343–356.
- Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* **53**: 819–830.
- Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. 2010. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. *Genome Res* **20**: 90–100.
- Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bähler J. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**: 1239–1243.
- Wu CH, Yamaguchi Y, Benjamin LR, Horvat-Gordon M, Washinsky J, Enerly E, Larsson J, Lambertsson A, Handa H, Gilmour D. 2003. NELF and DSIF cause promoter proximal pausing on the *hsp70* promoter in *Drosophila*. *Genes Dev* **17**: 1402–1414.
- Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**: 1033–1037.

Received January 18, 2016; accepted in revised form April 19, 2016.