

# The conserved elongation factor Spn1 is required for normal transcription, histone modifications, and splicing in *Saccharomyces cerevisiae*

Natalia I. Reim<sup>1,†</sup>, James Chuang<sup>1,†</sup>, Dhawal Jain<sup>2,†</sup>, Burak H. Alver<sup>2</sup>, Peter J. Park<sup>2</sup> and Fred Winston<sup>1,\*</sup>

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA and <sup>2</sup>Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA

Received July 14, 2020; Revised August 20, 2020; Editorial Decision August 24, 2020; Accepted August 27, 2020

## ABSTRACT

**Spn1/Iws1 is a conserved protein involved in transcription and chromatin dynamics, yet its general *in vivo* requirement for these functions is unknown. Using a Spn1 depletion system in *Saccharomyces cerevisiae*, we demonstrate that Spn1 broadly influences several aspects of gene expression on a genome-wide scale. We show that Spn1 is globally required for normal mRNA levels and for normal splicing of ribosomal protein transcripts. Furthermore, Spn1 maintains the localization of H3K36 and H3K4 methylation across the genome and is required for normal histone levels at highly expressed genes. Finally, we show that the association of Spn1 with the transcription machinery is strongly dependent on its binding partner, Spt6, while the association of Spt6 and Set2 with transcribed regions is partially dependent on Spn1. Taken together, our results show that Spn1 affects multiple aspects of gene expression and provide additional evidence that it functions as a histone chaperone *in vivo*.**

## INTRODUCTION

Spn1/Iws1 (hereafter called Spn1) is an evolutionarily conserved protein that is essential for viability in *Saccharomyces cerevisiae* and human cells and plays a role in mammalian and plant development (1–6). Spn1 was first identified in *S. cerevisiae* as a protein that interacts physically and genetically with factors involved in transcription (1,7–9), and it was later found to be part of the general RNA polymerase II (RNAPII) elongation complex on actively transcribed genes (10). Several studies have suggested that Spn1 is involved in transcription and co-transcriptional processes, including transcriptional activation (1,11,12), elongation (1,8–10), pre-mRNA processing (3,5,13–15), and mRNA

nuclear export (3,13,14). The recent demonstration that Spn1 binds histones and assembles nucleosomes *in vitro* suggested that Spn1 functions as a histone chaperone (16).

Spn1 directly interacts with Spt6 (17,18), another essential histone chaperone that associates with elongating RNAPII (10,19,20), and some evidence suggests that the two factors contribute to a common function. For example, some *spn1* and *spt6* mutations exhibit similar genetic interactions with mutations in genes encoding other factors involved in transcription and chromatin dynamics (21,22). In addition, combining *spn1* and *spt6* mutations or disrupting the Spn1-Spt6 interaction by mutation results in synthetic lethality, poor growth, or phenotypes associated with defective transcription and chromatin organization ((17,18,23) and our unpublished results). However, there is also evidence for separate functions of Spn1 and Spt6 based on distinct recruitment patterns to transcribed chromatin (10) and different functional requirements at the yeast *CYCI* gene (1).

Studies have suggested a role for Spn1 in histone H3K36 methylation. In mammalian cells, it was proposed that Spn1 is required for H3K36 trimethylation by mediating the interaction between Spt6 and the histone methyltransferase SETD2 during elongation by RNAPII (5,14). In yeast, while Spt6 is clearly required for Set2-dependent H3K36 methylation (24–30), the requirement for Spn1 is less certain, as mutations disrupting the Spn1-Spt6 interaction did not cause a complete loss of H3K36me3 and H3K36me2 ((30) and our unpublished results).

A functional connection between Spn1 and H3K36 methylation in yeast came from recent genetic evidence that showed that null mutations in genes encoding Set2 or components of the Rpd3S histone deacetylase complex (*set2Δ*, *rpd3Δ*, *eam3Δ*, *rcol1Δ*) suppress the temperature sensitive phenotype of a *spn1* mutation, *spn1-K192N* (22). Rpd3S is recruited and activated by H3K36me2/3 and the Ser5-phosphorylated RNAPII CTD (24,31–35). This suppres-

\*To whom correspondence should be addressed. Tel: +1 617 432 7768; Email: winston@genetics.med.harvard.edu

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

sion of *spn1-K192N* (22) suggested that one role of Spn1 is to oppose the functions of the Rpd3S histone deacetylase complex and, furthermore, that increased histone acetylation can compensate for impairment of Spn1 function.

In contrast to the activation of Rpd3S by H3K36 methylation, Set1-dependent H3K4 methylation has been proposed to oppose the function of Rpd3S, similar to the role proposed for Spn1 (22,36). By this model, H3K4me3 would inhibit deacetylation by Rpd3S in the region immediately downstream of promoters, where H3K4me3 is primarily found (37–41). In line with the idea that H3K36 methylation and H3K4 methylation have opposing functions, studies showed that loss of H3K4 methylation exacerbated the Ts<sup>-</sup> phenotype of *spn1-K192N*, while increased levels of H3K4 methylation suppressed *spn1-K192N* (21,22). In summary, genetic analysis suggests that the Set2-Rpd3S pathway opposes the function of Spn1, while Set1-dependent H3K4 methylation promotes it.

In this work, we have performed experiments to understand the requirement for Spn1 in transcription and co-transcriptional processes *in vivo*. Using an inducible Spn1 depletion system and RNA-seq, we found that loss of Spn1 results in widespread changes in mRNA levels. Furthermore, using ChIP-seq, we found that Spn1 depletion causes redistribution of H3K36me3, H3K36me2, and H3K4me3 genome-wide, as well as histone loss at highly transcribed genes. Additional experiments showed that Spt6 is crucial for the association of Spn1 with RNAPII, while Spn1 is not strongly required for the interaction of Spt6 with RNAPII, but rather plays a role in optimal Spt6 recruitment to chromatin. Finally, our results showed that Spn1 is required for efficient splicing of ribosomal protein transcripts. Together, our results reveal that Spn1 has widespread roles in the control of transcription and chromatin.

## MATERIALS AND METHODS

### Yeast strains

All yeast strains are listed in Supplementary Table S1, and all primers used for strain constructions are listed in Supplementary Table S2. *S. cerevisiae* liquid cultures were grown in YPD (1% yeast extract, 2% peptone and 2% glucose) at 30°C. *S. pombe* liquid cultures were grown in YES (0.5% yeast extract, 3% glucose, 225 mg/L each of adenine, histidine, leucine, uracil, and lysine) at 32°C. For Spn1 and Spt6 depletion experiments using an auxin-inducible degron (AID) (42), strains expressing the *Oryza sativa* *TIR1* gene and *SPN1-AID* or *SPT6-AID* were constructed as previously described (43). For Spt6 ChIP-seq, a 5' HA-tagged allele of *SPT6* was crossed into the *SPN1-AID* background. For Set2 ChIP-seq, Set2 was epitope-tagged by transformation of a DNA fragment amplified from pFA6a-3xHA-NatMX6 (FB2517) (44), which fuses the HA epitope to the C-terminal end of Set2. The *SET2-HA* fusion did not cause a detectable growth phenotype. For co-immunoprecipitation experiments, a triple Flag-tagged allele of *RPB3* was crossed into the *SPN1-AID* and *SPT6-AID* backgrounds. In all cases, possible growth effects of epitope tags were assessed by colony size. For 5' regulatory region replacement experiments, the *YLR454W* promoter

and 5' UTR were replaced by those of *UBI4* or *GCV3*. To do this, the *UBI4* and *GCV3* 5' regulatory regions were amplified from genomic DNA of a wild-type strain (FY92) using a forward primer with a 5' 20-bp overhang that would allow for later amplification of the NatMX cassette from a pFA6a plasmid (this 20-bp sequence is the reverse complement of the reverse primer sequence used for making deletions in (44)). The reverse primer for 5' regulatory region amplification included a 3' overhang with homology to the 40-bp sequence beginning at the *YLR454W* start codon. The resulting amplicon containing homology to the pFA6a-NatMX plasmid, the regulatory region of interest, and homology to the region beginning at the *YLR454W* ATG was recovered using the E.Z.N.A. Cycle Pure Kit and used as a reverse primer in a second PCR reaction in which the NatMX cassette was amplified from the pFA6a plasmid. The forward primer for this reaction included 40 bp of homology to the region from -323 to -284 relative to the *YLR454W* start codon and the Longtine forward primer sequence (44) for making deletion cassettes from pFA6a. The reverse primer for the first PCR reaction was used additionally in this second PCR to increase product yield. The final PCR product for each construct was ethanol-precipitated, gel-purified using the E.Z.N.A. Gel Purification Kit, and used to transform *S. cerevisiae* strain FY3227 (Spn1-AID). Transformants were selected on YPD plates that contained 100 µg/ml nourseothricin, and insertion of each cassette in the correct orientation and location was checked by two PCR reactions (PCR-1 using NR179 and FO5520, PCR-2 using FO7627 and NR180). The PCR-2 reaction amplified a region from the NatMX cassette to the first 379 bp of *YLR454W* ORF; this product was used to sequence across the 5' regulatory region of each construct using primer FO6606. A control strain in which the NatMX cassette alone was inserted upstream of the endogenous *YLR454W* promoter without deleting any *YLR454W* sequences was also constructed. In this case, only one PCR reaction using the pFA6a-NatMX plasmid as a template was required. The *YLR454W* 5' regulatory region was defined as the region from -283 to -1 relative to the ATG and included the TFIIB (-208 to -103 relative to the ATG) and TSS (-148 to -90 relative to the ATG) peaks previously identified (19). The *UBI4* 5' regulatory region was defined as the region between -567 and -1 relative to the ATG and included the TFIIB (-220 to -80 relative to the ATG) and TSS (-111 to -87 relative to the ATG) peaks, as well as other regulatory elements previously described (45). Two different 5' regulatory region amplicons were made for *GCV3*: one in which it was defined as the region between -232 and -1 relative to the annotated ATG and one in which it was defined as the region between -232 and +90 relative to the ATG (this ATG remained in frame with the *YLR454W* ATG in the final strain). The latter construct was made because the TFIIB peak mapped in Doris *et al.* spanned ~40 bp into coding region of *GCV3* (TFIIB peak: -114 to +43 relative to the ATG; TSS peak: -15 to -1 relative to the ATG). In summary, the *UBI4* and *GCV3* 5' regulatory region amplicons, marked with NatMX upstream, were inserted immediately upstream of the *YLR454W* start codon, while the region from -283 to -1 was deleted in the process.

### Auxin-induced depletion of Spn1 and Spt6

Spn1 depletion was optimized as previously described (43) by testing final concentrations of the auxin 3-indoleacetic acid (IAA), ranging from 25–500  $\mu$ M and treatment times ranging from 0 to 120 min. Viability was measured by counting colony-forming units per OD<sub>600</sub> unit plated in the 25- $\mu$ M IAA time-course. Optimization of the time for Spt6 depletion was done for the 25- $\mu$ M IAA treatment. For all Spn1 and Spt6 depletion experiments in this study, IAA was added to a final concentration of 25  $\mu$ M followed by growth for 90 min.

### Western blotting and antibodies

For checking Spn1 and Spt6 protein levels in all depletion experiments, OD-normalized aliquots of the cultures grown for the main experiment were saved and processed later on. For assessing levels of proteins and histone modifications assayed in ChIP-seq experiments, triplicate 20-ml cultures for the wild type, Spn1-noAID (contains TIR1 but no AID tag on Spn1), and Spn1-AID strains were grown to OD<sub>600</sub>  $\approx$  0.7 and treated with DMSO or IAA as described in the ChIP-seq methods. Culture volumes were normalized by their OD<sub>600</sub> to an OD equivalent (OD<sub>600</sub>  $\times$  volume of the culture) of 12 for harvesting. Protein extracts were prepared as described previously (27), and 12–15  $\mu$ l were loaded on 10% or 15% SDS-PAGE gels for western blotting. The following dilutions of primary antibodies were used: 1:1000 anti-Rpb1 (Millipore, 8WG16); 1:6000 anti-Spn1 (generously provided by Laurie Stargell); 1:10 000 anti-Spt6 (generously provided by Tim Formosa); 1:6000 anti-Set2 (generously provided by Brian Strahl); 1:1000 anti-H3K36me3 (Abcam, ab9050); 1:1000 anti-H3K36me2 (Abcam, ab9049); 1:2000 anti-H3 (Abcam, ab1791); 1:8000 anti-Pgk1 (Life Technologies, #459250). The following dilutions of secondary antibodies were used: 1:10 000 goat anti-rabbit IgG (Licor IRDye 680RD) and 1:20 000 goat anti-mouse IgG (Licor IRDye 800CW). Quantification of western blots was done using ImageJ 1.48v.

### RNA-seq

Duplicate 100-ml cultures for the wild type (FY3225), Spn1-noAID (FY3226), and Spn1-AID (FY3227) strains were grown in YPD at 30°C to  $\sim 3.5 \times 10^7$  cells/ml, diluted back to OD<sub>600</sub>  $\approx$  0.3 in pre-warmed YPD, divided into two equal volumes, treated with DMSO or IAA at a final concentration of 25  $\mu$ M, and grown for an additional 90 min. An OD-normalized aliquot of each culture was processed to verify depletion by western blot. Cells were counted and harvested, and each sample was spiked in to obtain a final concentration of 10% *S. pombe* cells by cell number and flash-frozen. Total RNA was prepared by hot acid phenol extraction (46), and a portion of each sample was treated with RQ1 DNase and used for subsequent steps (Promega, M6101). rRNA was removed using a RiboZero kit for yeast (Illumina, MRZY1324) as per manufacturer instructions, except all volumes were halved. Library preparation was carried out according to a previously published method, except fragmentation was performed prior

to linker ligation for RNA-seq (47). Briefly, alkaline fragmentation (2 $\times$  fragmentation buffer: 100 mM NaCO<sub>3</sub> pH 9.2, 2 mM EDTA pH 8.0) was performed to obtain 40–70-nucleotide fragments that were subsequently size-selected and dephosphorylated with T4 PNK (NEB, M0201S) for ligation of a 3' DNA linker containing a random hexamer molecular barcode to allow detection of PCR duplicates (48). First-strand cDNA was synthesized with SuperScriptIII (Invitrogen, #56575) using a primer with sequence complementary to the 3' linker and with an additional adapter sequence. cDNA was circularized with CircLigase (Epicentre, CL4111K). DNA libraries were amplified by PCR for 6–10 cycles with Phusion HF DNA Polymerase (NEB, M0530) using Illumina TruSeq indexing primers for multiplexing. Libraries were gel-purified and run on Agilent Bioanalyzer and/or TapeStation to assess size distribution, quality, and concentration. RNA-seq libraries were sequenced on the Illumina NextSeq 500 platform at the Biopolymers Facility at Harvard Medical School.

### RT-qPCR

Spn1-AID cells were cultured and treated with DMSO or IAA as described for RNA-seq. Total RNA was prepared by hot acid phenol extraction, DNase-treated, and ethanol-precipitated after each step. cDNA was synthesized from 1  $\mu$ g of RNA using qScript XLT cDNA SuperMix (Quanta Biosciences, #95161), which contains a blend of oligo d(T) and random primers. A no-RT (no reverse transcription) control, where no cDNA SuperMix was added to the RNA, was included for each sample and used for qPCR. The 20- $\mu$ l RT reactions were brought up to 100  $\mu$ l with water, and three 1:3 serial dilutions were prepared for each sample. Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent Technologies, #600882) was used for quantitative PCR. Primers are listed in Supplementary Table S2. Triplicate intercept values for regions of interest were divided by triplicate intercept values of *S. pombe pma1*<sup>+</sup> (primers FO7083 and FO7084) if *S. pombe* cells were used as spike-in. For intron retention analysis, unspliced transcripts were measured with a forward primer spanning the 3' intron-exon junction, while spliced transcripts were measured with a forward primer spanning the exon–exon junction. The reverse primer was the same for both transcript species. The ratio of unspliced to spliced transcripts was calculated for six genes in duplicate. For the 5' regulatory region replacement experiments, changes in *YLR454W* transcript levels following Spn1 depletion were measured using primers NR183 and NR184, which amplified the region from +994 to +1091 relative to the start codon.

### ChIP-seq

The Spn1-AID strain (FY3227) was used for ChIP of Rpb1, Rpb1 CTD modifications, histone H3, histone modifications, and Spn1 (Spn1-AID-3xV5). Spn1-AID strains containing HA-tagged Spt6 (FY3228) or Set2 (FY3229) were used for those respective ChIPs. For *S. cerevisiae* strains, duplicate cultures were grown in YPD at 30°C to OD<sub>600</sub>  $\approx$  0.65, diluted back to OD<sub>600</sub>  $\approx$  0.3–0.4 in pre-warmed YPD, divided into two equal volumes, treated with DMSO

or IAA at a final concentration of 25  $\mu\text{M}$ , and allowed to grow for 90 min. An OD-normalized aliquot of each culture was processed to verify depletion by western blot. Cultures were processed as previously described (27), and the final pellet after cell lysis was resuspended in 500  $\mu\text{l}$  LB140 and sonicated in a QSonica Q800R machine for 15 min (30 s on, 30 s off, 70% amplitude) to achieve 100–500-bp DNA fragments. For *S. pombe* strains, cultures were grown at 32°C in YES to  $\text{OD}_{600} \approx 0.7\text{--}0.8$  and processed similarly to the *S. cerevisiae* cultures, except for the following steps. Bead-beating was done for 10–14 min, chromatin was concentrated at the last resuspension in LB140, and samples were sonicated for 20 min. Protein concentrations of all *S. cerevisiae* and *S. pombe* chromatin samples were measured by Bradford (49). For Spt6 and Set2 ChIPs, 1 mg of chromatin was used as input, whereas 500  $\mu\text{g}$  of chromatin was used as input for all other ChIPs. As a spike-in control, *S. pombe* chromatin was added to a final concentration of 10% of total mass, and the volume was brought up to 800  $\mu\text{l}$  with WB140 buffer. Samples for Rpb1, Spt6, Set2, H3, and H3K36me3 ChIP were spiked in with an *S. pombe* strain encoding an HA-tagged protein (FWP568). Samples for Rpb1, Rpb1 Ser5-P, Rpb1 Ser2-P, H3, H3K36me2, H3K4me3, and Spn1 were spiked in with an *S. pombe* strain encoding a V5-tagged protein (FWP485). Before addition of antibody, 5% (40  $\mu\text{l}$ ) was set aside as input and stored for later processing, starting at the decrosslinking step later on. The following antibodies were used for IPs: 10  $\mu\text{g}$  anti-HA (Abcam, ab91110), 10  $\mu\text{l}$  anti-Rpb1 (Millipore, 8WG16), 8  $\mu\text{g}$  anti-H3 (Abcam, ab1791), 6  $\mu\text{g}$  anti-H3K36me3 (Abcam, ab9050), 6  $\mu\text{g}$  anti-H3K36me2 (Abcam, ab9049), 6  $\mu\text{g}$  anti-H3K4me3 (Abcam, ab8580), 10  $\mu\text{g}$  anti-Ser5-P (Active Motif, 3E8), 10  $\mu\text{g}$  anti-Ser2-P (Active Motif, 3E10), 8  $\mu\text{l}$  anti-V5 (Invitrogen, R960-25). Immunoprecipitation reactions were incubated at 4°C with end-over-end rotation overnight (~16–18 h). Protein G Sepharose beads (GE Healthcare, 4 Fast Flow) were washed with WB140, and a 50% slurry was subsequently made in WB140. Fifty-microliter aliquots of slurry were transferred into new Eppendorf tubes, and the WB140 was removed. The immunoprecipitation reactions were added directly to the beads and incubated at 4°C with end-over-end rotation for 4 h. Bead-washing and elution were done as previously described (27). The saved inputs were thawed from storage, and 60  $\mu\text{l}$  LB140 and 100  $\mu\text{l}$  TES were added. The 200- $\mu\text{l}$  eluates and inputs were incubated at 65°C overnight to reverse crosslinking. RNase A/T1 and Proteinase K digestions, DNA purification, and library preparation were done as previously described (27). The number of PCR cycles for library amplification was determined empirically for each sample and ranged from 11–19 cycles. Samples were again purified twice using 0.7 $\times$  volume SPRI beads and run on Agilent Bioanalyzer to assess size distribution, quality, and concentration. ChIP-seq libraries were sequenced on the Illumina NextSeq 500 platform at the Harvard Bauer Core Facility.

### Co-immunoprecipitation

Rpb3-Flag (FY3224), Spn1-AID (FY3227), Rpb3-Flag Spn1-AID (FY3230), Spt6-AID (FY3122), and Rpb3-Flag

Spt6-AID (FY3231) cells were cultured in triplicate and treated with DMSO or IAA as described for ChIP-seq. Depletions of Spn1 and Spt6, as well as levels of other proteins assayed in this experiment, were verified by western blot. After collecting the cell pellets and washing them with water, the cells were resuspended in an amount of IP buffer (20 mM HEPES pH 7.6, 20% glycerol, 1 mM DTT, 1 mM EDTA, 125 mM KoAc, 1% NP-40) proportional to the OD of the culture harvested (with a culture at  $\text{OD}_{600}$  0.65 receiving 500  $\mu\text{l}$  of buffer). The cells were lysed by bead-beating at 4°C for 6 min with 2-min incubations on ice after every minute of bead-beating. The lysates were collected, centrifuged at 14 000 rpm for 10 min at 4°C, and the cleared extracts were transferred to new tubes. Protein concentrations of all extracts were measured using Pierce BCA Protein Assay Kit (Thermo Fisher, #23225). Samples for immunoprecipitation were prepared at a final concentration of 15  $\mu\text{g}/\mu\text{l}$  protein in IP buffer up to 500  $\mu\text{l}$ . As input, 40  $\mu\text{l}$  of each sample was set aside, 20  $\mu\text{l}$  3 $\times$  SDS buffer (240 mM Tris-HCl pH 6.8, 6% SDS, 16%  $\beta$ -mercaptoethanol, 4% SDS, 0.6 mg/mL bromophenol blue, 30% glycerol) was added, and this was frozen until ready to use. Anti-Flag M2 affinity gel beads (Sigma, A2220) were washed once in IP buffer and prepared as a 10% bead slurry in IP buffer. One hundred-microliter aliquots of slurry (containing a total of 10  $\mu\text{l}$  packed beads) were prepared in separate tubes, and 400  $\mu\text{l}$  of the normalized extracts were added to each tube and incubated at 4°C with end-over-end rotation for 2 h. Following the IP, the beads were washed three times with 500  $\mu\text{l}$  of IP buffer for 2 min per wash at 4°C. For elution, the beads were resuspended in 45  $\mu\text{l}$  of 3 $\times$  SDS buffer and boiled for 5 min. The beads were then spun down, and the supernatant was transferred to a new tube and either stored or used for western blots. The following antibodies were used for western blots: 1:8000 anti-Flag (Sigma, F3165); 1:6000 anti-Spn1 (provided by Laurie Stargell); 1:10 000 anti-Spt6 (provided by Tim Formosa); 1:6000 anti-Set2 (provided by Brian Strahl); 1:2000 anti-V5 (Invitrogen, R960-25) to check depletions.

### Genome builds, annotations, and gene sets analyzed

The genome builds used were *S. cerevisiae* R64–2–1 (50) and *S. pombe* ASM294v2 (51). *S. cerevisiae* transcript coordinates were generated from TIF-seq (52) and TSS-seq data, as previously described (19). For quantitative comparisons, we have analyzed the full set of 5091 verified coding genes. For heatmap and metagene average plots, where the intent is to analyze how data varies as a function of position along a gene, we have limited the genes being analyzed to the set of 3087 non-overlapping, verified coding genes in order avoid any influence of overlapping genes.

### RNA-seq library processing

The random hexamer molecular barcode on the 5' end of each read was removed and processed using a custom Python script. Cutadapt (53) was then used to remove adapter sequences and low-quality base calls from the 3' end of the read, as well as poly-A sequences from the 5' end of the read. Reads were aligned to the combined *S. cerevisiae*

and *S. pombe* genome using TopHat2 (54) without a reference transcriptome, and uniquely mapping alignments were selected using SAMtools (55). Alignments mapping to the same location as another alignment with the same molecular barcode were identified as PCR duplicates and removed using a custom Python script. Coverage corresponding to the 3'-most base of the RNA fragment sequenced, as well as the entire interval for each alignment, was extracted using BEDTools (56) genomcov and normalized to the total number of alignments uniquely mapping to the *S. pombe* genome. The quality of raw, cleaned, non-aligning, and uniquely aligning non-duplicate reads was assessed using FastQC (57).

#### ***Ab initio* transcript annotation from RNA-seq data**

StringTie (58) was used to perform *ab initio* transcript annotation for each RNA-seq sample, with minimum transcript length of 50 bp, minimum transcript coverage of 2.5, and minimum separation between transcripts of 12 nt. Transcript annotations for all samples (two replicates each of DMSO- and IAA-treated wild type, Spn1-noAID, and Spn1-AID samples) were merged, and those annotations not overlapping known transcripts were combined with the annotations of known transcripts to generate a unified transcript annotation. Transcripts from this list were classified as antisense if their 5' end overlapped a known transcript on the opposite strand.

#### **RNA-seq differential expression analysis**

For differential expression analyses of verified coding genes, RNA-seq alignments overlapping the transcript annotation of these genes were counted using BEDTools (56). To compare Spn1-AID treated with IAA versus DMSO, the transcript counts from these two conditions were used to perform differential expression analysis using DESeq2 (59) with a null hypothesis of no change between conditions and a false discovery rate of 0.1. All differential expression analyses were normalized to the spike-in control by using size factors obtained from *S. pombe* transcript counts for each sample. To obtain changes in verified coding transcript abundance due to Spn1 depletion while accounting for possible effects from expressing the TIR1 ubiquitin ligase, AID-tagging Spn1, or treating yeast cells with IAA, a separate multi-factor differential expression analysis was performed (Supplementary Figure S2A). Transcript counts from all RNA-seq samples (two replicates each of DMSO- or IAA-treated wild type, Spn1-noAID, and Spn1-AID strains) were used to perform a multi-factor differential expression analysis using DESeq2 (59) at a false discovery rate of 0.1. The design formula for the multi-factor analysis was a generalized linear model with no interaction terms in which the variables represented the presence of TIR1, the presence of the AID tag on Spn1, the presence of IAA in the media, and the depletion of Spn1. To obtain changes for all annotated transcripts, including antisense transcripts (Supplementary Figure S2E), a multi-factor differential expression analysis of the same design was performed using counts measured over the unified transcript annotation described above.

#### **Gene set enrichment analysis**

The set of genes with altered expression during the environmental stress response (ESR) (60) was separated into ESR-upregulated and ESR-downregulated genes based on the median fold change of expression across stress conditions. Gene set enrichment analysis (61) was performed for these gene sets as well as for gene ontology sets curated by the Saccharomyces Genome Database (62), with genes ranked by their RNA-seq fold-change in the Spn1-AID strain treated with IAA versus DMSO.

#### **Intron retention analysis**

Intron retention was defined as the proportion of unspliced RNA for an intron and was estimated from counts of RNA-seq alignments. Spliced and unspliced alignments were counted for 252 non-mitochondrial introns not overlapping 5' UTRs (50). Alignments overlapping an intron on the same strand were counted as spliced if the alignment indicated matches of at least 4 nt with both the 5' and 3' exon sequences, separated by a skipped region over the coordinates of the intron. Alignments overlapping an intron on the same strand were counted as unspliced if the alignment indicated a continuous match spanning at least 4 nt on both sides of the 5' or 3' splice junction, or if the alignment was contained entirely within the intron boundaries. Alignments overlapping annotated features contained within intron boundaries were excluded to avoid conflating differential expression of these features with changes in intron retention. Spliced and unspliced alignment counts were pooled between replicates. To address high variance in the raw intron retention proportion  $\frac{\text{unspliced}}{\text{unspliced} + \text{spliced}}$  associated with low coverage introns, intron retention was estimated using empirical Bayes shrinkage towards a beta-binomial prior. The prior distribution was fit by maximum likelihood estimation using the intron alignment counts for the 217/252 introns with at least one spliced alignment in any of the conditions tested. To determine the significance of changes in intron retention upon Spn1 depletion, Monte Carlo sampling from the posterior intron retention distributions was used to determine the probability that intron retention in the Spn1-AID strain treated with IAA was more extreme than intron retention in all control conditions (wild type DMSO, wild type IAA, Spn1-noAID DMSO, Spn1-noAID IAA, and Spn1-AID DMSO).

#### **ChIP-seq library processing**

Reads were demultiplexed using fastq-multx (63), allowing one mismatch to the index sequence and A-tail. Cutadapt (53) was then used to remove index sequences and low-quality base calls from the 3' end of the read. Reads were aligned to the combined *S. cerevisiae* and *S. pombe* genome using Bowtie 2 (64), and alignments with a mapping quality of at least 5 were selected using SAMtools (55). The median fragment size estimated by MACS2 (65) cross-correlation over all samples of a factor was used to generate coverage of fragments and fragment midpoints by extending alignments to the median fragment size or by shifting the 5' end of alignments by half the median fragment size,

respectively. The quality of raw, cleaned, non-aligning, and uniquely aligning reads was assessed using FastQC (57).

### ChIP-seq normalization

For ChIP-seq coverage from IP samples, spike-in normalization was accomplished by scaling coverage proportionally to the normalization factor  $\frac{N_{\text{input, spike-in}}}{N_{\text{IP, spike-in}}(N_{\text{input, experimental}})}$ , where  $N_{\text{IP, spike-in}}$  is the number of *S. pombe* alignments in the IP sample,  $N_{\text{input, spike-in}}$  is the number of *S. pombe* alignments in the corresponding input sample, and  $N_{\text{input, experimental}}$  is the number of *S. cerevisiae* alignments in the input sample. Coverage from input samples was normalized to  $N_{\text{input, experimental}}$ . Relative estimates of the total abundance of each ChIP target on chromatin were also obtained by multiplying the normalization factor with the number of *S. cerevisiae* alignments in an IP sample. Coverage of the relative ratio of IP over input was obtained by first smoothing normalized IP and input fragment midpoint coverage using a Gaussian kernel with 20-bp bandwidth and then taking the ratio. Coverage of the relative ratio of one factor to another (e.g. H3K36me3 over H3) was obtained as follows: For each factor, coverage of IP over input in each sample was standardized using the genome-wide mean and standard deviation over all samples, weighted such that the non-depleted and Spn1-depleted conditions had equal contribution. In the cases where ChIP for both factors was performed using the same chromatin preparation, the standardized coverage of the normalizing factor was subtracted from the matched coverage of the factor to be normalized. In the cases where ChIP for each factor was performed using different chromatin preparations (Spt6 and Set2 normalized to Rpb1), standardized coverages were averaged across replicates before subtracting coverage of the normalizing factor from the coverage of the factor to be normalized.

### ChIP-seq differential occupancy analysis

For differential occupancy analyses of single factors over verified coding genes, IP and input fragment midpoints overlapping the transcript annotation of these genes were counted using BEDTools (56). These counts were used to perform a differential occupancy analysis using DESeq2 (59) at a false discovery rate of 0.1. The design formula used was a generalized linear model with variables for sample type (IP or input), condition (non-depleted or Spn1-depleted), and the interaction of sample type with condition. Fold changes were extracted from the coefficients of the interaction of sample type with condition and represent the change in IP signal observed upon Spn1 depletion, corrected for changes in input signal. When normalizing to the spike-in control, size factors obtained from *S. pombe* counts over peaks called with MACS2 (65) were used for each sample. All single-factor differential occupancy results shown are normalized to spike-in, except for H3 differential occupancy results, where the additional variability introduced by normalization to spike-in would have complicated identification of the set of genes with significantly reduced H3. For differential occupancy analysis of H3 over the first 500 bp of verified coding genes (Supplementary Figure S4B),

the same approach was taken except using counts over the relevant regions. For differential occupancy analyses considering one factor normalized to another (e.g. changes in Rpb1-normalized Spt6), IP and input counts for both factors were used to perform a differential binding analysis using DESeq2 (59) at a false discovery rate of 0.1. The design formula used was a generalized linear model with all crossings of variables for sample type (IP or input), ChIP target, and condition (non-depleted or Spn1-depleted). Fold changes were extracted from the coefficients of the interaction of sample type, ChIP target, and condition. These values represent changes in input-normalized IP signal of one factor upon Spn1 depletion, corrected for changes in input-normalized IP signal of the other factor. All factor-normalized differential occupancy analysis results shown are normalized to spike-in by using size factors obtained from counts over *S. pombe* transcripts.

## RESULTS

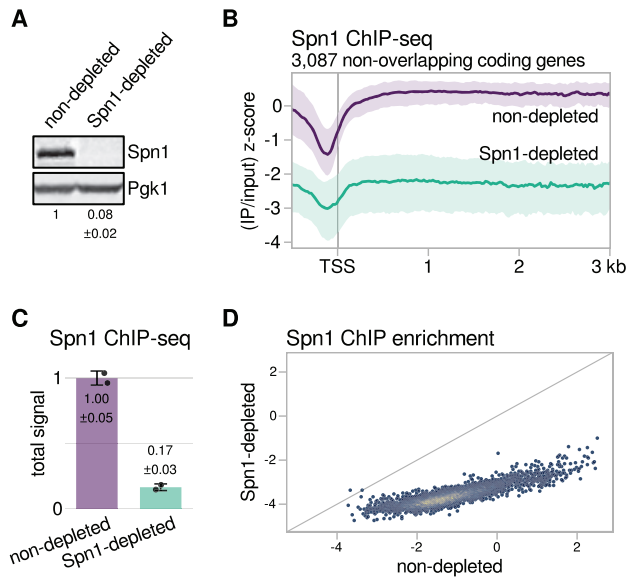
### An auxin-inducible degron system efficiently depletes Spn1 in *S. cerevisiae*

To study the requirement for the essential protein Spn1, we constructed an *S. cerevisiae* strain for the depletion of Spn1 using an auxin-inducible degron (see Materials and Methods). Treatment of this strain with the auxin 3-indoleacetic acid (IAA) for 90 min (Spn1-depleted condition) resulted in the reduction of Spn1 protein levels to about 8% of those observed when the strain was treated with DMSO (non-depleted condition) (Figure 1A). Importantly, these conditions did not impair viability (Supplementary Figure S1A).

To verify that Spn1 depletion reduced the amount of Spn1 on chromatin, we performed chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) for Spn1 in Spn1-depleted and non-depleted cells, using *S. pombe* chromatin as a spike-in control to allow the detection of global changes (see Materials and Methods). In agreement with a previous study (10), Spn1 in non-depleted cells was primarily enriched over gene bodies (Figure 1B) at levels strongly correlated with levels of the RNAPII subunit Rpb1 (Supplementary Figure S1B). Additionally, we noted that the ratio of Spn1 to Rpb1 over coding genes decreased slightly with increasing Rpb1 levels (Supplementary Figure S1C). In Spn1-depleted cells, the total Spn1 ChIP signal was ~17% that of non-depleted cells (Figure 1C). The reduction in Spn1 levels was observed for virtually all coding genes (Figure 1D) and occurred uniformly over gene bodies (Figure 1B). Taken together, our results show that the auxin-inducible degron system effectively depletes Spn1 from chromatin genome-wide and allows for interrogation of its essential functions during transcription.

### Depletion of Spn1 causes a widespread reduction of mRNA levels

Although there is evidence that Spn1 is involved in transcription, the genome-wide requirement for Spn1 in regulating mRNA levels was unknown. To investigate this, we performed RNA sequencing (RNA-seq) of Spn1-depleted and non-depleted cells, using *S. pombe* cells as a spike-in control. The major effect that we observed was a widespread



**Figure 1.** An auxin-inducible degron efficiently depletes Spn1 from chromatin. **(A)** Western blot showing levels of Spn1 protein in the Spn1 degron strain after 90 min of treatment with DMSO (non-depleted) or 25  $\mu$ M IAA (Spn1-depleted). Pgk1 was used as a loading control. The values below the blots indicate the mean  $\pm$  standard deviation of normalized Spn1 abundance quantified from two biological replicates. **(B)** Average Spn1 ChIP enrichment over 3087 non-overlapping, verified coding genes aligned by TSS in non-depleted and Spn1-depleted conditions. For each gene, the spike-in normalized ratio of IP over input signal is standardized to the mean and standard deviation of the non-depleted signal over the region, resulting in standard scores that allow genes with varied expression levels to be compared on the same scale. The solid line and shading are the median and interquartile range of the mean standard score over two replicates. **(C)** Bar plot showing total levels of Spn1 on chromatin in non-depleted and Spn1-depleted conditions, estimated from ChIP-seq spike-in normalization factors (see Materials and Methods). The values and error bars for each condition indicate the mean  $\pm$  standard deviation of two replicates. **(D)** Scatterplot showing Spn1 ChIP enrichment in Spn1-depleted versus non-depleted conditions for 5091 verified coding genes. Enrichment values are the relative log<sub>2</sub> enrichment of IP over input.

decrease in mRNA abundance in Spn1-depleted versus non-depleted cells, with a median change to around 73% of non-depleted transcript levels and over 1500 mRNAs detected as differentially reduced (Figure 2A; Supplementary Table S3). We also detected 33 mRNAs with increased levels (Figure 2A; Supplementary Table S3). The changes were independent of transcript levels in non-depleted cells (Figure 2A) and gene length. By analyzing RNA-seq data from additional control conditions, we verified that the changes in transcript abundance were caused primarily by Spn1 depletion rather than by the presence of the degron tag on Spn1, the expression of degron system components, or the effects of auxin on yeast cells (see Materials and Methods; Supplementary Figure S2A). We verified the differential expression of seven genes by RT-qPCR (Supplementary Figure S2B), including the strong induction of *SER3* previously observed in *spn1* mutants (66).

In response to various environmental stresses, yeast cells are known to alter the transcript levels of approximately 870 environmental stress response (ESR) genes (60). To look for a possible relationship between Spn1 depletion and the

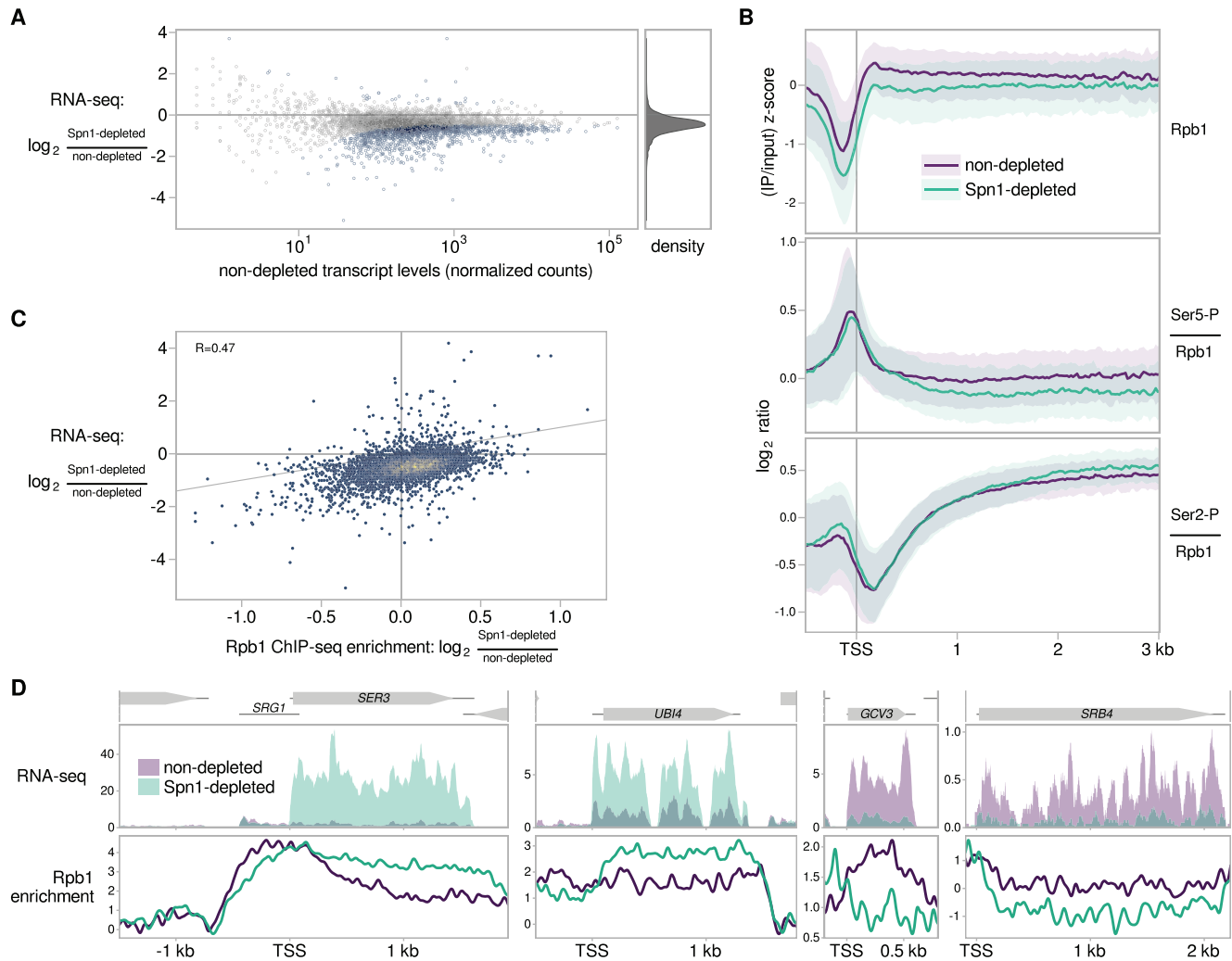
ESR, we performed gene set enrichment analysis. Our results revealed an association between Spn1-dependent transcriptional changes and the ESR (Supplementary Figure S2C), consistent with the fact that Spn1 is essential for viability. Other gene sets with significant associations were largely subsets of the ESR genes, including gene sets related to ribosomes, translation, and rRNA processing. In spite of these associations, the transcriptional changes in response to Spn1 depletion are more global in scope than the ESR, do not correlate with the gene expression changes of any individual ESR stress (data not shown), and do not correlate with a related transcriptional signature of slow growth (67) (Supplementary Figure S2D). This suggests that the Spn1-dependent changes we observe are not only explained by stress responses.

Impaired function of the Spn1 binding partner Spt6 causes widespread derepression of antisense transcripts (19,26,68–71). Therefore, we were interested to see whether a similar effect was observed after Spn1 depletion. In contrast to *spt6* mutants, we detected only four antisense transcripts with significantly increased levels in Spn1-depleted cells (Supplementary Figure S2E). Though this may be a slight underestimate due to the difficulty of annotating lowly expressed transcripts, it is clear that antisense transcripts are much more limited for Spn1 depletion as compared to *spt6* mutants.

To test whether the widespread changes in sense transcript levels observed after Spn1 depletion might be driven by changes in the level or state of RNAPII, we performed ChIP-seq for the largest RNAPII subunit, Rpb1, as well as for two major Rpb1 C-terminal domain (CTD) post-translational modifications: serine 5 phosphorylation (Ser5-P) and serine 2 phosphorylation (Ser2-P). Our results show that the relative distribution of Rpb1 over transcribed regions was not greatly affected by Spn1 depletion (Figure 2B). However, we observed that the modest changes in Rpb1 levels in Spn1-depleted cells correlate positively with changes in transcript levels as measured by RNA-seq (Figure 2C,D; see Materials and Methods). This suggests that the Spn1-dependent effects on transcript levels can be explained, at least in part, by changes in RNAPII occupancy. The modest effect on Rpb1 levels leaves open the possibility that Spn1 may also affect mRNA levels post-transcriptionally. There was little effect of Spn1 depletion on either Ser5-P or Ser2-P levels, with Ser5-P retaining its expected preferential enrichment near transcription start sites and Ser2-P retaining its expected progressive enrichment over gene bodies (72) (Figure 2B, Supplementary Figure S2F).

### 5' regulatory regions influence Spn1-dependent regulation of transcript levels

Although Spn1 is thought of primarily as a transcription elongation factor, it was also shown to function at the promoter of the *S. cerevisiae* *CYC1* gene (11,16). Therefore, we wanted to determine whether 5' regulatory regions, including promoters, might play a role in Spn1-dependent regulation of transcript levels. To test this possibility, we made gene fusions of the 5' regulatory regions of *UBI4* or *GCV3*—two genes respectively greatly upregulated or



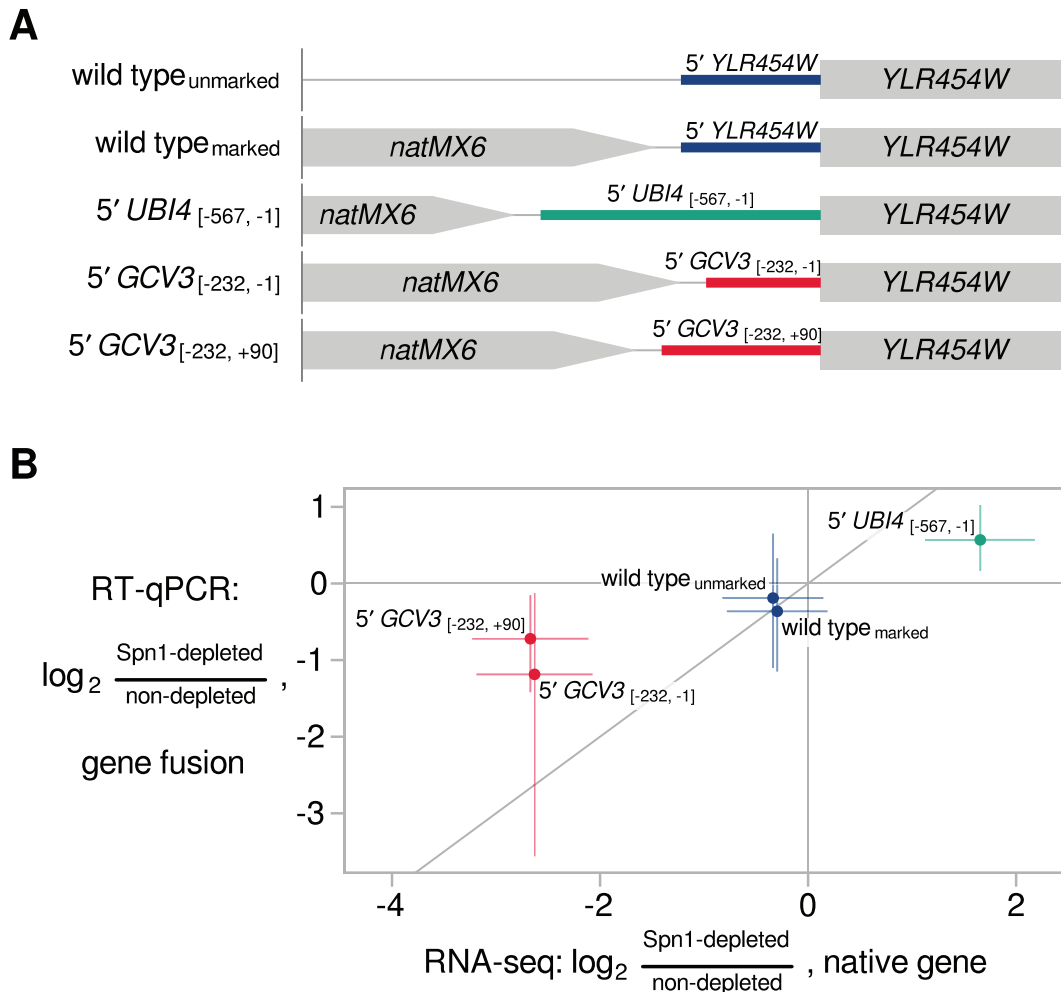
**Figure 2.** Effects of Spn1 depletion on mRNA levels and RNAPII. (A) Scatterplot showing changes in transcript abundance as measured by RNA-seq upon Spn1 depletion versus non-depleted transcript abundance for 5091 verified coding transcripts. Transcripts with significant changes are colored blue, and the relative density of fold-change values is shown in the right panel. (B) Top panel: Average Rpb1 ChIP enrichment over 3087 non-overlapping, verified coding genes in non-depleted and Spn1-depleted conditions. Spike-in normalized standard scores are calculated as in Figure 1B. The solid line and shading are the median and interquartile range of the mean standard score over the two replicates generated from the same chromatin preparations used to perform ChIP-seq for Rpb1 Ser5-P and Rpb1 Ser2-P. Bottom two panels: Average Rpb1-normalized Rpb1 Ser5-P and Rpb1 Ser2-P ChIP enrichment in non-depleted and Spn1-depleted conditions for the same genes. The solid line and shading are the median and interquartile range of the mean spike-in normalized ratio over two replicates. (C) Scatterplot showing change in transcript abundance upon Spn1 depletion versus change in Rpb1 enrichment for 5091 verified coding genes. The line  $y = x$  is drawn for comparison, and the Pearson correlation coefficient is shown. (D) Spike-in normalized sense strand RNA-seq coverage and spike-in normalized Rpb1 ChIP enrichment over four genes, two with elevated mRNA levels and two with decreased mRNA levels, upon Spn1 depletion.

downregulated in Spn1-depleted cells (Figure 2D, Supplementary Figure S2B)—to the coding region of *YLR454W*, a gene whose transcript levels were minimally affected by Spn1 depletion (Figure 3A). The sequence used as the 5' regulatory region of a gene included the 5' untranslated region (UTR) of the gene, in addition to a promoter region defined using TSS-seq and TFIIB ChIP-nexus data (see Materials and Methods). These fusions and controls were constructed in a Spn1 degtron strain.

We then used RT-qPCR to measure the change in *YLR454W* transcript levels upon Spn1 depletion in the fusion and control strains (Figure 3B). For the 5' *UBI4*-

*YLR454W* fusion, transcript levels increased in response to Spn1 depletion ~48% relative to the non-depleted condition. For the 5' *GCV3*-*YLR454W* fusions, transcript levels decreased about 56% or 39%, depending on the particular sequence used for the *GCV3* 5' regulatory region (see Materials and Methods). The direction of these changes is consistent with the RNA-seq and RT-qPCR data for the native *UBI4* and *GCV3* genes, although the magnitude of the *YLR454W* transcript changes in the fusions was smaller. Therefore, our results suggest that the 5' regulatory regions of these genes contribute, at least partially, to the Spn1 dependence of their transcript levels.





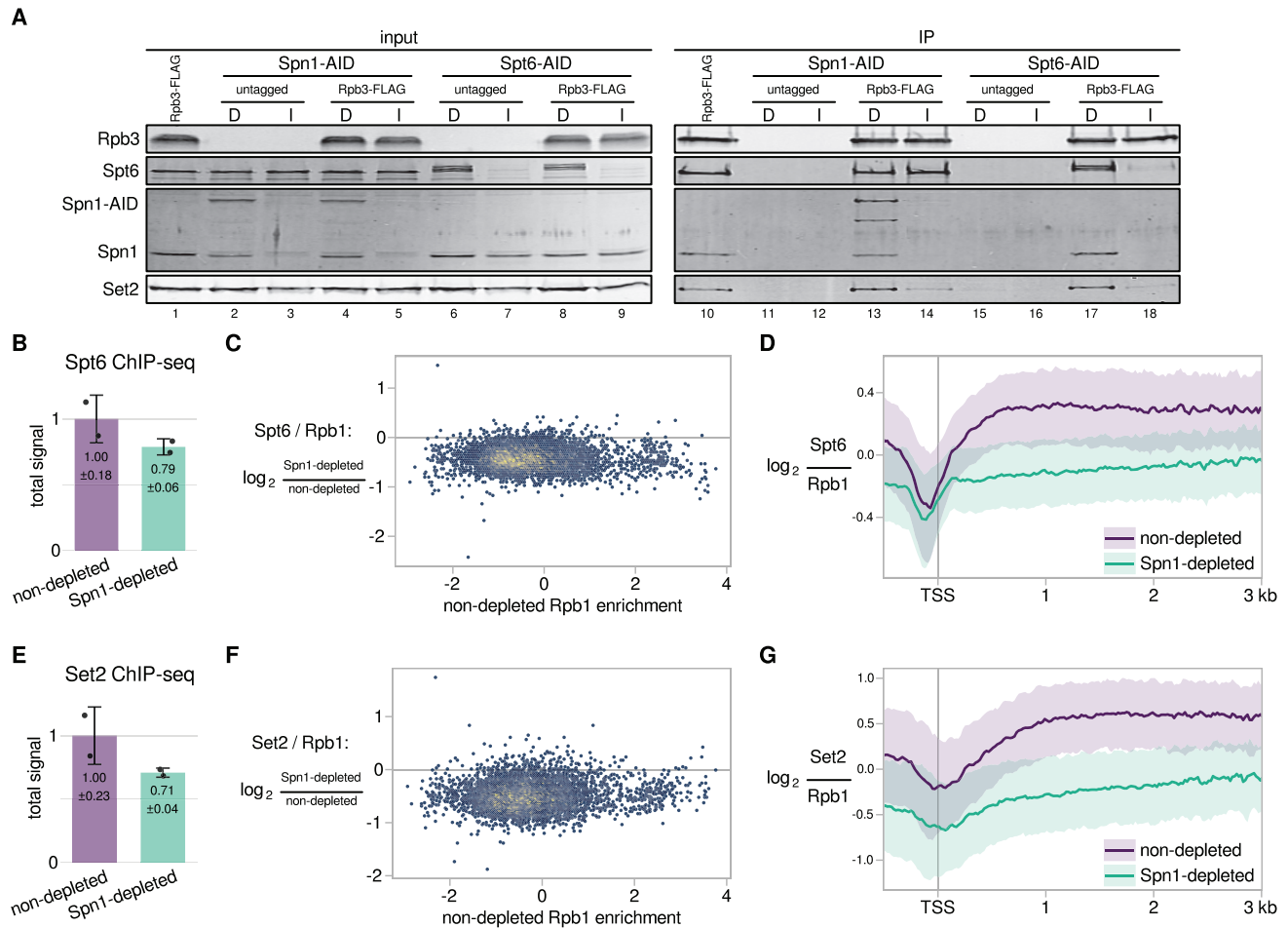
**Figure 3.** 5' regulatory regions influence Spn1-dependent regulation of transcript levels. (A) Schematic of gene fusions in which the 5' regulatory regions of two Spn1-dependent genes replaced the 5' regulatory region of the *YLR454W* gene. The nucleotide coordinates of the native *UBI4* and *GCV3* regulatory regions are shown, relative to the start of the coding sequence (19,45). All fusions were made in the Spn1 degon background. (B) Scatterplot summarizing results of the 5' regulatory region replacement experiment. The y-axis shows RT-qPCR measurements of the change in *YLR454W* transcript abundance upon Spn1 depletion in the gene fusion strains. The x-axis shows RNA-seq measurements of the change in transcript abundance upon Spn1 depletion for the native gene whose 5' regulatory region was used to drive *YLR454W* expression in the fusion strains. The line  $y = x$  is drawn for comparison. Error bars indicate 95% confidence intervals, and points corresponding to the same native gene are slightly jittered by adding a small amount of random noise along the x-axis for clarity.

### The association of Spn1 with RNAPII is dependent on Spt6

Spn1 directly interacts with Spt6, and both proteins are part of the RNAPII elongation complex (8–10,14,17,18). To assess the possible dependency of one factor on the other for association with RNAPII, we performed immunoprecipitation of the epitope-tagged RNAPII subunit Rpb3 under conditions where we depleted either Spn1 or Spt6. In the Spn1-depleted condition, Spt6 co-immunoprecipitated with Rpb3 at levels similar to the non-depleted condition (Figure 4A, compare lanes 13 and 14, Spt6 panel). Therefore, the association of Spt6 with RNAPII is independent of Spn1, consistent with the fact that Spt6 directly interacts with RNAPII (73,74). In contrast, in Spt6-depleted conditions, the interaction of Spn1 with RNAPII was lost (Figure 4A, compare lanes 17 and 18, Spn1 panel). Therefore, the association of Spn1 with RNAPII is Spt6-dependent,

supporting previous findings from studies of Spn1 (Iws1) in human cell lines (13,14).

Although our co-immunoprecipitation experiments indicated that Spn1 is not required for the association of Spt6 with RNAPII, it was possible that Spn1 was required for Spt6 recruitment at certain locations in the genome. For example, previous studies showed that Spn1 is required for the recruitment of Spt6 upon induction of the *CYC1* gene (11). Therefore, to investigate the Spn1-dependence of Spt6 recruitment to chromatin genome-wide, we performed Spt6 ChIP-seq using a Spn1 degon strain expressing epitope-tagged Spt6. In the non-depleted condition, Spt6 and Rpb1 occupancy correlated strongly, as expected (10,20,75) (Supplementary Figure S3A), with the ratio of Spt6 to Rpb1 decreasing slightly with increasing Rpb1 occupancy (Supplementary Figure S3B). Like in the case of Spn1 (Supplementary Figure S1C), this suggests that Spt6 does not in-



**Figure 4.** Requirement of Spn1 for Spt6 and Set2 association with RNAPII and chromatin. (A) Western blots showing the results of co-immunoprecipitation experiments to analyze the effects of Spn1 or Spt6 depletion on the binding of Spt6, Spn1, and Set2 to Rpb3-FLAG. Spn1 or Spt6 degon strains (Spn1-AID or Spt6-AID) were treated for 90 min with DMSO (lanes marked 'D') or 25  $\mu$ M IAA (lanes marked 'I'), after which Rpb3-FLAG was immunoprecipitated and co-immunoprecipitation of Spn1, Spt6, and Set2 was assayed by western blots using native antibodies. Rpb3-FLAG levels were measured using anti-FLAG antibody, and an Rpb3-FLAG strain lacking the degon tag on Spn1 or Spt6 was included as an additional non-depleted control. Shown is one representative western blot of three biological replicates with similar results. (B) Bar plot showing total levels of Spt6 on chromatin in non-depleted and Spn1-depleted conditions, estimated from ChIP-seq spike-in normalization factors (see Materials and Methods). The values and error bars for each condition indicate the mean  $\pm$  standard deviation of two replicates. (C) Scatterplot showing changes in Rpb1-normalized Spt6 ChIP enrichment upon Spn1 depletion versus non-depleted Rpb1 enrichment for 5091 verified coding transcripts. Rpb1 enrichment values are the relative  $\log_2$  enrichment of IP over input. (D) Average Rpb1-normalized Spt6 ChIP enrichment over 3087 non-overlapping, verified coding genes aligned by TSS in non-depleted and Spn1-depleted conditions. The solid line and shading are the median and interquartile range of the ratio calculated from the mean of two Spt6 and four Rpb1 replicates. (E) Same as in (B) but for Set2. (F) Same as in (C) but for Set2. (G) Same as in (D) but for Set2.

crease in equal proportion to RNAPII on genes as a function of transcription frequency. Following Spn1 depletion, Spt6 protein levels were unaffected (Supplementary Figure S3C), while the total amount of Spt6 on chromatin was reduced to  $\sim$ 79% of non-depleted levels (Figure 4B). The decrease in Spt6 on chromatin occurred uniformly over genes, with a net effect of reducing the ratio of Spt6 to Rpb1 over gene bodies (Figure 4C,D). These results show that Spn1 is required for the optimal recruitment of Spt6 to chromatin.

#### The association of Set2 with RNAPII and chromatin is dependent on Spn1

As both Spt6 and Spn1 are involved in the recruitment and/or activity of the histone methyltransferase Set2

(5,14,24,25,27,28), we also measured the association of Set2 with Rpb3 by co-immunoprecipitation in both Spn1-depleted and Spt6-depleted conditions. Our results showed that both proteins are required for Set2 association with RNAPII (Figure 4A, compare lane 13 to 14 and lane 17 to 18, Set2 panel). In the case of Spt6 depletion, the diminished interaction of Set2 with Rpb3 can likely be explained by both the effect on Spn1 and the decreased Set2 protein levels that occur when Spt6 is mutant or depleted, as previously observed by others (27,28), although this effect was modest in our results here (Figure 4A, compare lane 6 to 7 and lane 8 to 9, Set2 panel).

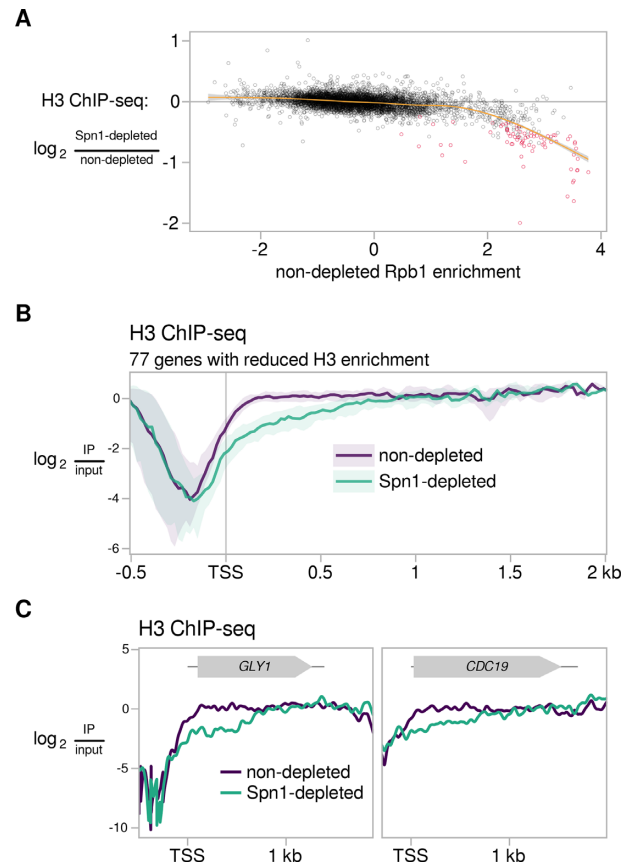
To examine the Spn1-dependence of Set2 recruitment to chromatin across the genome, we performed Set2 ChIP-seq using a Spn1 degon strain expressing epitope-tagged

Set2. In the non-depleted condition, we observed that the level of Set2 over coding genes positively correlated with Rpb1 levels (Supplementary Figure S3A), while the ratio of Set2 to Rpb1 strongly negatively correlated with Rpb1 levels (Supplementary Figure S3B). This suggests that the level of Set2 over genes varies much less with the level of transcription compared to elongation factors like Spt6 and Spn1 (Supplementary Figs. S3C and S1C). Following Spn1 depletion, the total amount of Set2 on chromatin was reduced to approximately 71% of non-depleted levels (Figure 4E), and the ratio of Set2 to Rpb1 was reduced over gene bodies (Figure 4F,G). The reduction in Set2 recruitment to chromatin measured by ChIP-seq appeared to be less severe than the reduction in Set2-RNAPII association observed by co-immunoprecipitation. This may be accounted for by technical differences between the two assays, such as the use of crosslinking in ChIP-seq, and may reflect the stabilization of Set2 on chromatin by factors besides RNAPII (27,28), including other elongation factors (76,77), nucleosomes/DNA (78–80), and mRNA (81). Thus, we conclude that Spn1 is required for the optimal recruitment of Set2 to chromatin, primarily by maintaining the Set2-RNAPII interaction.

### H3 levels are reduced near the 5' end of a small set of genes after Spn1 depletion

As Spn1 binds to histones H3/H4 *in vitro* (16) and has a modest role in the recruitment of Spt6 to chromatin (Figure 4B–D), we investigated a possible role for Spn1 in the control of histone occupancy and distribution *in vivo*. To do this, we performed ChIP-seq for histone H3 in Spn1-depleted and non-depleted conditions. We observed that the effect of Spn1 depletion on H3 occupancy over coding genes varied with expression level: while most genes had minimally affected H3 occupancy and distribution after Spn1 depletion, a small set of 77 mostly highly expressed genes had significantly decreased H3 occupancy (Supplementary Figure S4A, Figure 5A). Furthermore, the decrease in H3 occupancy at these genes occurred specifically over the first ~500 bp of the gene (Figure 5B,C). Since a 5'-specific effect might be masked for long genes when quantifying signal over the entire gene body, we then quantified the change in H3 occupancy over the first 500 bp of genes. By this approach, we observed the same relationship between expression level and H3 occupancy change and identified a total of 134 genes with significantly decreased H3 occupancy near the 5' end (Supplementary Figure S4B).

The 5' reduction in H3 at highly expressed genes upon Spn1 depletion was reminiscent of the 5' reduction in H3 over a broader set of genes previously observed upon Spt6 depletion (82) and in an *spt6* mutant with decreased Spt6 protein stability (23). Because of this and the known histone chaperone function of Spt6, we examined whether the reduction in H3 after Spn1 depletion could be explained by reduced levels of Spt6 on chromatin (Figure 4B–D). We found that for those genes with reduced H3 levels, the H3 changes did not correlate with changes in Rpb1-normalized Spt6 signal (Supplementary Figure S4C). Additionally, the 5' reduction in H3 at those genes was coincident with a 5' reduction in Rpb1 (Supplementary Figure S4C,D). These



**Figure 5.** Histone H3 levels are reduced at the 5' end of a small set of genes after Spn1 depletion. (A) Scatterplot showing change in H3 enrichment upon Spn1 depletion versus non-depleted Rpb1 enrichment for 5091 verified coding genes. Rpb1 enrichment values are the relative log<sub>2</sub> enrichment of IP over input. Genes with significant decreases in H3 enrichment are colored red, and a cubic regression spline is overlaid. (B) Average H3 ChIP enrichment in non-depleted and Spn1-depleted conditions for the 77 genes with significantly decreased H3 ChIP enrichment over the entire gene upon Spn1 depletion. The solid line and shading are the median and interquartile range of the mean enrichment over four replicates. (C) H3 ChIP enrichment in non-depleted and Spn1-depleted conditions for two genes with significantly decreased H3 ChIP enrichment upon Spn1 depletion. The mean enrichment over four replicates is shown.

results suggest that the changes in H3 occupancy at this set of genes upon Spn1 depletion are not the indirect effect of changes in Spt6 recruitment and may be related to changes in transcription, though we cannot say if the relationship is causal.

### Spn1 is required for the normal distribution of co-transcriptional histone modifications

Given previous studies that suggested a role for Spn1 in H3K36 methylation (3,5,14) and the change we observed in Set2 recruitment after Spn1 depletion (Figure 4E–G), we used ChIP-seq to measure the genome-wide occupancy of H3K36me2 and H3K36me3 in Spn1-depleted and non-depleted conditions. In these experiments, we also included analysis of H3K4 trimethylation (H3K4me3), given the genetic relationship between Spn1 and the enzymes required for the methylation and demethylation of H3K4 (21,22).

For H3K36 methylation, our results showed that, while the total level of each histone modification on chromatin was not greatly affected by Spn1 depletion (Supplementary Figure S5A,B), the distribution of each modification was markedly changed (Figure 6A,B, Supplementary Figure S5C). In the non-depleted condition, the H3K36 methylation marks were distributed as expected: H3K36me2 and H3K36me3 were primarily enriched within gene bodies, with H3K36me3 reaching its maximum levels further downstream of the TSS than H3K36me2 (40,83) (Figure 6A,B). In contrast, depletion of Spn1 resulted in prominent 3' shifts in the distributions of both H3K36me2 and H3K36me3 over the vast majority of coding genes, including those with and without altered RNA levels (Figure 6A,B). More specifically, in the non-depleted condition, the median H3-normalized H3K36me2 signal reached 90% of its maximum value within approximately 450 bp of the TSS, while in the Spn1-depleted condition, 90% of the maximum value was not reached until about 650 bp (Figure 6A, top). For H3-normalized H3K36me3 in the non-depleted condition, the median signal reached 90% of its maximum value within ~690 bp of the TSS, while in the Spn1-depleted condition, 90% of the maximum value was not reached until about 1390 bp (Figure 6B, top). Though these changes in distribution occurred at most genes, the magnitude of the changes increased with increasing expression level (Supplementary Figure S5D).

The genome-wide distribution of H3K4me3 was also strongly affected by Spn1 depletion. While in the non-depleted condition, H3K4me3 was primarily localized over the 5' regions of genes, as expected (38–40,84), depletion of Spn1 caused H3K4me3 signal to spread toward the 3' end for the vast majority of genes, into regions normally enriched for H3K4me2 and H3K4me1 (40) (Figure 6C, Supplementary Figure S5C). In the non-depleted condition, the median H3-normalized H3K4me3 signal over coding genes decreased to 10% of its maximum value within about 1530 bp of the TSS, whereas in the Spn1-depleted condition, the median signal did not decrease to 10% of its maximum value until about 2430 bp (Figure 6C, top). As with H3K36 methylation, the magnitude of the change in H3K4me3 distribution over genes increased with increasing expression level (Supplementary Figure S5D). It is possible that the change in H3K4me3 pattern upon Spn1 depletion is caused by altered recruitment or activity of either Set1, the H3K4 methyltransferase, or Jhd2, the demethylase, though this remains to be tested. Altogether, we conclude that Spn1 is required to maintain the normal pattern of H3K36 and H3K4 methylation across the genome.

### Spn1 depletion results in increased intron retention in ribosomal protein transcripts

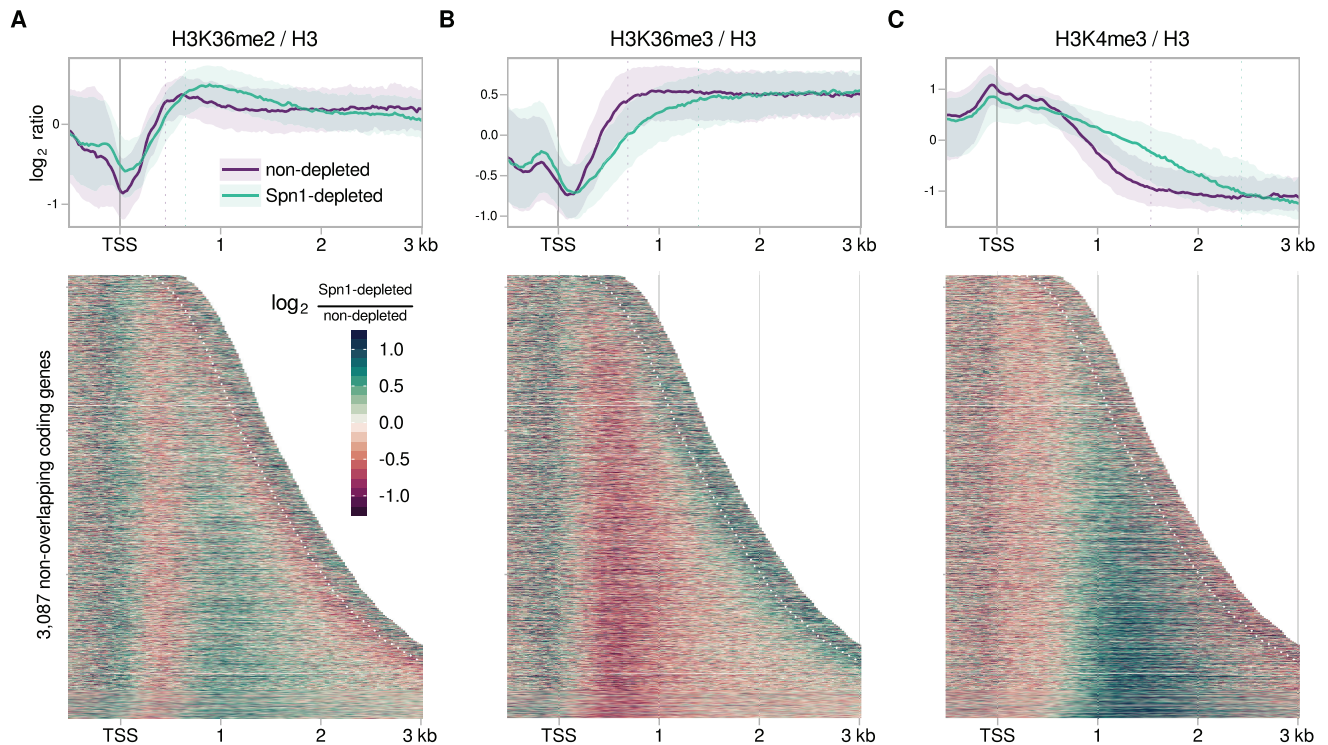
Since both Spn1 and H3K36 methylation have been previously implicated in splicing (3,5,85–89), we looked for changes in splicing efficiency after Spn1 depletion by using our RNA-seq data to estimate intron retention (see Materials and Methods). We observed a general tendency for increased intron retention after Spn1 depletion (Figure 7A), particularly in ribosomal protein transcripts, whose high expression levels gave us greater power to detect changes in

intron retention. Out of 252 introns analyzed, we found 59 introns with significantly increased retention upon Spn1 depletion and two introns with significantly decreased retention (Figure 7A). Of the introns with significantly increased retention after Spn1 depletion, 27/59 showed increases of at least 10%, and 54/59 were from ribosomal protein transcripts. To verify our intron retention estimates, we used RT-qPCR to measure the ratio of unspliced to spliced transcripts for four genes that showed increased intron retention and two genes without significant changes (Figure 7B). In all cases, the changes in unspliced transcript abundance observed by this method were in agreement with the results from the transcriptome-wide analysis.

H3K36 methylation has been linked to splicing in mammalian cells by studies showing that chromodomain and PWWP domain-containing proteins such as MRG15 can bind SETD2-dependent H3K36me3 to bridge the association of splicing factors and chromatin, thus co-transcriptionally regulating alternative splicing (85,88,89). Eaf3, the yeast ortholog of MRG15, also binds to both H3K36me3 and H3K36me2 (24,33,34,90) and mediates the recruitment of splicing factors (87), supporting previous results that implicated H3K36 di- and/or trimethylation in splicing regulation in yeast (86). Therefore, we wanted to determine if there were Spn1-dependent changes in either H3K36me2 or H3K36me3 that were specific to intron-containing genes. Sixty-five percent of the 137 ribosomal protein (RP) genes contain introns, giving us an opportunity to compare two sets of genes similar in expression and length but differing by the presence or absence of introns. In the non-depleted condition, intron-containing RP genes showed a slightly higher level of H3K36me2 over the intron compared to the downstream exon (Figure 7C). Following Spn1 depletion, redistribution of this modification toward the 3' end resulted in a marked decrease in signal over the intron, a pattern not observed at a similar distance from the TSS in RP genes without introns. Additionally, RP genes with introns had a slightly greater decrease in H3K36me3 over most of the gene body compared to RP genes without introns. We did not detect major differences in the distribution of other factors or modifications assayed between the two sets of genes, apart from a very mild increase in Rpb1 Ser5-P over the introns, which has been observed when pre-spliceosome formation is perturbed (91,92) (Supplementary Figure S6A). We obtained very similar results when we performed this analysis comparing intron-containing RP genes to non-RP genes without introns matched by length and expression (data not shown). Finally, the overlap of decreases in H3K36me2 and regions encoding introns observed for RP genes was not seen for non-RP intron-containing genes, as the decreases in H3K36me2 over non-RP intron-containing genes occurred further 3' of the TSS and 5' splice site (Supplementary Figure S6B). Thus, for intron-containing RP genes, distinct changes in H3K36me2 and H3K36me3 correlate with increased intron retention.

## DISCUSSION

In this study, we have uncovered previously unknown *in vivo* roles for Spn1, an essential and conserved factor. We have



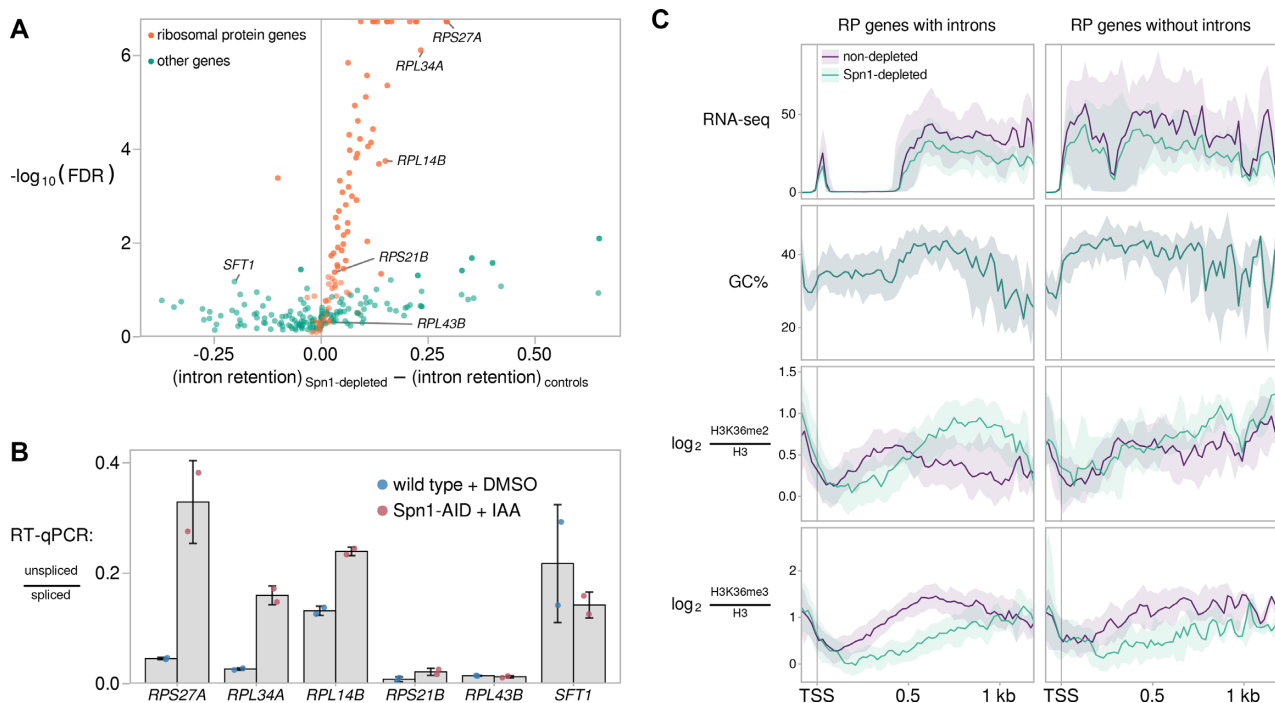
**Figure 6.** Spn1 is required for normal localization of histone H3 modifications along coding genes. (A) Top: Average H3-normalized H3K36me2 for 3087 non-overlapping, verified coding genes aligned by TSS in non-depleted and Spn1-depleted conditions. The solid line and shading are the median and interquartile range of the mean ratio over two replicates. Vertical dotted lines indicate the distances at which the median signal in each condition reached 90% of its maximum value. Bottom: Heatmap of the change in H3-normalized H3K36me2 enrichment upon Spn1 depletion for the same genes, aligned by TSS and arranged by transcript length. Data is shown for each gene up to 300 bp 3' of the cleavage and polyadenylation site, which is indicated by the white dotted line. (B) Same as in (A) but for H3K36me3. (C) Same as in (A) but for H3K4me3, with vertical dotted lines instead indicating the distances at which the median signal in each condition dropped to 10% of its maximum value.

shown that Spn1 is widely required for normal transcript levels, as depletion of Spn1 significantly reduces mRNA levels of over 1500 genes in a manner that correlates with changes in Rpb1 occupancy. In addition, depletion of Spn1 alters the distribution of H3K36 and H3K4 methylation along most genes and impairs efficient splicing, particularly in ribosomal protein transcripts. Furthermore, Spn1 is required for normal levels of histone H3 over the 5' ends of a small set of mostly highly transcribed genes. This provides evidence that Spn1 functions *in vivo* as a histone chaperone at a set of genes, supporting recent *in vitro* studies (16). Finally, we have shown that the association of Spn1 with RNAPII is dependent on its binding partner, Spt6. Together, these results establish that Spn1 plays a genome-wide role in transcription and chromatin dynamics.

We have shown that a major function of Spn1 is to maintain proper localization of H3K36me3, H3K36me2, and H3K4me3, as Spn1 depletion causes redistribution of these three transcription-coupled histone modifications toward the 3' ends of genes. Mislocalization of these histone marks has been previously observed in other studies, although in some cases the shifts were in the 5' direction, opposite to the shifts observed after Spn1 depletion. For example, loss of the chromatin remodeler Chd1 resulted in 5' shifts of both H3K36me3 and H3K4me3 (93,94). Interestingly, the loss of Chd1 also resulted in more efficient splicing of

ribosomal protein transcripts, in contrast to the less efficient splicing that we observed after Spn1 depletion. Thus, Spn1 and Chd1 appear to play opposite roles. In the case of Spn1 depletion, it is plausible that reduced Set2 activity caused by reduced recruitment contributes to the changes in H3K36 methylation pattern. By this model, reduced Set2 leads to 3' shifts of H3K36 methylation by reducing the co-transcriptional methylation rate and thus increasing the distance along a gene needed to reach a given methylation state.

Changes in the rate of transcription have also been shown to cause shifts of H3K36 and H3K4 methylation. In yeast, an RNAPII slow mutant with a reduced polymerization rate showed a 5' shift of H3K4me2 and H3K4me3, while an RNAPII fast mutant with an increased polymerization rate showed a 3' shift of these modifications (84,95). In mammalian cells, a 5' shift of H3K36me3 was also observed in an RNAPII mutant with slowed transcription rate (96), opposite to what we observed after Spn1 depletion. In both the yeast and mammalian studies, the histone modification changes correlated with shifts of Rpb1 Ser2-P: a 3' shift in slow mutants and a 5' shift in the fast mutant (84,96). In contrast, we did not detect a significant change in Rpb1 Ser2-P after Spn1 depletion; therefore, the histone modification 3' shifts that we observe might not be caused by an increased rate of transcription, although this remains to be tested.



**Figure 7.** Depletion of Spn1 results in increased intron retention in ribosomal protein transcripts. (A) Volcano plot showing changes in intron retention upon Spn1 depletion for 252 introns encoded from within coding regions of nuclear genes. Intron retention is defined as the proportion of unspliced transcripts and was estimated from RNA-seq data (see Materials and Methods). Labeled introns were further assayed by RT-qPCR, as shown in (B). (B) RT-qPCR measurements of the ratio of unspliced to spliced transcripts in a wild-type strain treated with DMSO for 90 min and the Spn1 degran strain treated with 25  $\mu\text{M}$  IAA for 90 min. Error bars indicate the mean  $\pm$  standard deviation of two replicates. (C) Visualization of average changes in H3K36 methylation state at ribosomal protein genes with and without introns. For each group of genes, the following data are plotted: sense strand RNA-seq coverage, GC% in a 21-bp window, and H3-normalized H3K36me2 and H3K36me3 ChIP enrichment. The solid line and shading are the median and interquartile range over the genes in each group. The introns of RP genes tend to be close to the 5' end of the gene, the region where RNA-seq signal is seen as depleted (top left panel).

Recent studies showed that the histone chaperones Spt6 and FACT (composed of Spt16 and Pob3) are also required for the normal distribution of several histone modifications, as shifts toward the 3' end occurred in yeast *spt6* and *spt16* mutants (95), as well as after depletion of FACT in *Drosophila* cells (97). The yeast study suggested that these shifts were a consequence of transcription-dependent loss of histones on chromatin due to impaired nucleosome reassembly in the wake of transcription. The decreases in histone H3 occupancy that we observed after Spn1 depletion were much less extensive than those observed in the *spt6* and *spt16* mutants. Therefore, although Spn1 is likely required for nucleosome reassembly at highly expressed genes, a defect in nucleosome assembly is unlikely to completely explain the widespread shifts in histone modifications observed after Spn1 depletion.

We found that Spn1 depletion causes both increased retention of many ribosomal protein introns and reduced H3K36me2 and H3K36me3 over the regions encoding these introns. The possibility that the reduced H3K36 methylation causes the defects in splicing is supported by previous studies showing that loss of H3K36 methylation impairs splicing from yeast to mammals (5,85–89,93,98). We observed that Spn1 depletion caused ribosomal protein genes to have reduced H3 and H3K36me3 regardless of whether the gene contained an intron; however, ribosomal protein genes with introns had a distinct redistribution

of H3K36me2 after Spn1 depletion, resulting in decreased levels of this mark before the 3' splice site and increased levels thereafter. Our results, then, suggest that H3K36me2 is a Spn1-dependent mark that distinguishes ribosomal protein genes with introns from those without. Given that defects in splicing have also been observed to alter the recruitment of SETD2 and the distribution of H3K36me3 at intron-containing genes in mammalian cells (99,100), further experiments will be needed to determine whether the direct effect of Spn1 depletion is on H3K36 methylation patterns or splicing.

Our studies have also revealed new insights into the molecular relationship between Spn1 and Spt6. The importance of this interaction is highlighted by the fact that *spn1* or *spt6* mutations that impair this interaction cause severe mutant phenotypes (17,18). Other studies previously suggested that the two proteins do not contribute to a common essential function since they differ based on both genetic and molecular properties (10,11,18). Our results extend these distinctions, as depletion of Spn1 causes different effects on transcription compared to mutation or depletion of Spt6, particularly with respect to the level of antisense transcription (19,26,71). The critical nature of the Spn1–Spt6 interaction can be explained by our findings that the association of Spn1 with RNAPII is strongly dependent on Spt6, and that Spn1 is required for optimal recruitment of Spt6 to chromatin.

Additionally, our results suggest that Spn1 regulation occurs, at least in part, via the 5' regulatory regions of genes. While this is not expected for an elongation factor, there is precedent. Previous studies of Spt6 similarly showed effects at 5' regulatory regions (20,82,101). While a specific role for Spn1 has been demonstrated at the 5' regulatory region of the yeast *CYCI* gene (11), the effect that we detected after Spn1 depletion may be more general. Depletion or mutation of Spt6 causes general defects in chromatin structure, including at nucleosome-depleted regions (19,26,71,82). A similar effect after Spn1 depletion could conceivably increase or decrease the level of transcription initiation.

In summary, our studies have demonstrated that Spn1 is widely required *in vivo* for normal transcript levels, splicing, and chromatin-related functions. Given its physical and genetic interactions with two other essential histone chaperones, Spt6 and FACT (21,22,102), and its genetic interactions with additional histone chaperones, including Asf1 and the Hir complex (16), it appears that Spn1 is part of a network of histone chaperones required for normal gene expression. Deciphering the individual roles of these factors and how they function together remains a challenge for future studies.

## DATA AVAILABILITY

The RNA-seq and ChIP-seq data sets are available in the GEO repository, accession number GSE153567. They can be accessed at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153567>. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Data. All data analyses were managed using the Snakemake workflow management system (103). An archive containing code and raw data suitable for reproducing the analyses presented in this manuscript is available at Zenodo (doi.org/10.5281/zenodo.3901642). Additionally, updated versions of the Snakemake pipelines used are available at [github.com/winston-lab](https://github.com/winston-lab). A UCSC browser session displaying uploaded sequence tracks can be viewed at the following site: [https://genome.ucsc.edu/s/jameschuang/Spn1\\_depletion](https://genome.ucsc.edu/s/jameschuang/Spn1_depletion).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Olga Viktorovskaya for suggesting and helping with the co-immunoprecipitation experiments. We thank Raja Gopalakrishnan, Olga Viktorovskaya, and James Warner for helpful comments on the manuscript.

*Author contributions:* N.I.R. and F.W. conceived the study and designed the experiments. N.I.R. performed all the experiments. J.C., D.J., and B.A. performed the computational analyses with support from P.J.P. for D.J. and B.A. N.I.R., J.C., and F.W. wrote the manuscript.

## FUNDING

National Institutes of Health [F31GM112370 to N.I.R., R01GM120038 to F.W.]. Funding for open access charge: NIH.

*Conflict of interest statement.* None declared.

## REFERENCES

- Fischbeck, J.A., Kraemer, S.M. and Stargell, L.A. (2002) SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant. *Genetics*, **162**, 1605–1616.
- Liu, Z., Zhou, Z., Chen, G. and Bao, S. (2007) A putative transcriptional elongation factor hIws1 is essential for mammalian cell proliferation. *Biochem. Biophys. Res. Commun.*, **353**, 47–53.
- Oqani, R.K., Lin, T., Lee, J.E., Kang, J.W., Shin, H.Y. and Il Jin, D. (2019) Iws1 and Spt6 regulate trimethylation of histone H3 on lysine 36 through Akt signaling and are essential for mouse embryonic genome activation. *Sci. Rep.*, **9**, 3831.
- Orlacchio, A., Stark, A.E., Foray, C., Amari, F., Sheetz, T., Reese, E., Tessari, A., La Perle, K., Palmieri, D., Tschlis, P.N. *et al.* (2018) Genetic ablation of interacting with Spt6 (Iws1) causes early embryonic lethality. *PLoS One*, **13**, e0201030.
- Sanidas, I., Polytarchou, C., HatziaPOSTOLOU, M., Ezell, S.A., Kottakis, F., Hu, L., Guo, A., Xie, J., Comb, M.J., Iliopoulos, D. *et al.* (2014) Phosphoproteomics screen reveals akt isoform-specific signals linking RNA processing to lung cancer. *Mol. Cell*, **53**, 577–590.
- Widiez, T., El Kafafi, S., Girin, T., Berr, A., Ruffel, S., Krouk, G., Vayssieres, A., Shen, W.H., Coruzzi, G.M., Gojon, A. *et al.* (2011) High nitrogen insensitive 9 (HNI9)-mediated systemic repression of root NO<sub>3</sub>- uptake is associated with changes in histone methylation. *PNAS*, **108**, 13329–13334.
- 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.
- Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S. and Greenblatt, J.F. (2002) RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol. Cell. Biol.*, **22**, 6979–6992.
- Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R. 3rd and Hartzog, G.A. (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol. Cell. Biol.*, **23**, 1368–1378.
- Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J. and Cramer, P. (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.*, **17**, 1272–1278.
- Zhang, L., Fletcher, A.G., Cheung, V., Winston, F. and Stargell, L.A. (2008) Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. *Mol. Cell. Biol.*, **28**, 1393–1403.
- Pujari, V., Radebaugh, C.A., Chodaparambil, J.V., Muthurajan, U.M., Almeida, A.R., Fischbeck, J.A., Luger, K. and Stargell, L.A. (2010) The transcription factor Spn1 regulates gene expression via a highly conserved novel structural motif. *J. Mol. Biol.*, **404**, 1–15.
- Yoh, S.M., Cho, H., Pickle, L., Evans, R.M. and Jones, K.A. (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev.*, **21**, 160–174.
- Yoh, S.M., Lucas, J.S. and Jones, K.A. (2008) The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev.*, **22**, 3422–3434.
- Kanno, T., Venhuizen, P., Wu, M.T., Chiou, P., Chang, C.L., Kalyana, M., Matzke, A.J.M. and Matzke, M. (2020) A collection of Pre-mRNA splicing mutants in *Arabidopsis thaliana*. *G3*, **10**, 1983–1996.
- Li, S., Almeida, A.R., Radebaugh, C.A., Zhang, L., Chen, X., Huang, L., Thurston, A.K., Kalashnikova, A.A., Hansen, J.C.,

- Luger, K. *et al.* (2018) The elongation factor Spn1 is a multi-functional chromatin binding protein. *Nucleic Acids Res.*, **46**, 2321–2334.
17. Diebold, M.L., Koch, M., Loeliger, E., Cura, V., Winston, F., Cavarelli, J. and Romier, C. (2010) The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. *EMBO J.*, **29**, 3979–3991.
  18. McDonald, S.M., Close, D., Xin, H., Formosa, T. and Hill, C.P. (2010) Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. *Mol. Cell*, **40**, 725–735.
  19. Doris, S.M., Chuang, J., Viktorovskaya, O., Murawska, M., Spatt, D., Churchman, L.S. and Winston, F. (2018) Spt6 is required for the fidelity of promoter selection. *Mol. Cell*, **72**, 687–699.
  20. Ivanovska, I., Jacques, P.E., Rando, O.J., Robert, F. and Winston, F. (2011) Control of chromatin structure by spt6: different consequences in coding and regulatory regions. *Mol. Cell Biol.*, **31**, 531–541.
  21. Lee, K.Y., Chen, Z., Jiang, R. and Meneghini, M.D. (2018) H3K4 methylation dependent and independent chromatin regulation by JHD2 and SET1 in budding Yeast. *G3*, **8**, 1829–1839.
  22. Lee, K.Y., Ranger, M. and Meneghini, M.D. (2018) Combinatorial genetic control of Rpd3S through histone H3K4 and H3K36 methylation in budding Yeast. *G3*, **8**, 3411–3420.
  23. Dronamraju, R., Kerschner, J.L., Peck, S.A., Hepperla, A.J., Adams, A.T., Hughes, K.D., Aslam, S., Yoblinski, A.R., Davis, I.J., Mosley, A.L. *et al.* (2018) Casein kinase II phosphorylation of Spt6 enforces transcriptional fidelity by maintaining Spn1-Spt6 interaction. *Cell Rep.*, **25**, 3476–3489.
  24. Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P. *et al.* (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, **123**, 581–592.
  25. Chu, Y., Sutton, A., Sternglanz, R. and Prelich, G. (2006) The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2. *Mol. Cell Biol.*, **26**, 3029–3038.
  26. Yegennaro, C.M., Alver, K.B., Marguerat, S., Stepanova, E., Davis, C.P., Bahler, J., Park, P.J. and 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.
  27. Gopalakrishnan, R., Marr, S.K., Kingston, R.E. and Winston, F. (2019) A conserved genetic interaction between Spt6 and Set2 regulates H3K36 methylation. *Nucleic Acids Res.*, **47**, 3888–3902.
  28. Youdell, M.L., Kizer, K.O., Kisseleva-Romanova, E., Fuchs, S.M., Duro, E., Strahl, B.D. and Mellor, J. (2008) Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol. Cell Biol.*, **28**, 4915–4926.
  29. Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R. and Winston, F. (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol.*, **6**, e277.
  30. Dronamraju, R. and Strahl, B.D. (2014) A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res.*, **42**, 870–881.
  31. Drouin, S., Laramée, L., Jacques, P.E., Forest, A., Bergeron, M. and Robert, F. (2010) DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. *PLoS Genet.*, **6**, e1001173.
  32. Govind, C.K., Qiu, H., Ginsburg, D.S., Ruan, C., Hofmeyer, K., Hu, C., Swaminathan, V., Workman, J.L., Li, B. and Hinnebusch, A.G. (2010) Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. *Mol. Cell*, **39**, 234–246.
  33. Joshi, A.A. and Struhl, K. (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol. Cell*, **20**, 971–978.
  34. Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J. *et al.* (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, **123**, 593–605.
  35. Li, B., Jackson, J., Simon, M.D., Fleharty, B., Gogol, M., Seidel, C., Workman, J.L. and Shilatifard, A. (2009) Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *J. Biol. Chem.*, **284**, 7970–7976.
  36. McDaniel, S.L., Fligor, J.E., Ruan, C., Cui, H., Bridgers, J.B., DiFiore, J.V., Guo, A.H., Li, B. and Strahl, B.D. (2016) Combinatorial histone readout by the dual plant homeodomain (PHD) fingers of roo1 mediates Rpd3S chromatin recruitment and the maintenance of transcriptional fidelity. *J. Biol. Chem.*, **291**, 14796–14802.
  37. Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T. and Schreiber, S.L. (2002) Methylation of histone H3 Lys 4 in coding regions of active genes. *PNAS*, **99**, 8695–8700.
  38. Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N. and Rando, O.J. (2005) Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol.*, **3**, e328.
  39. Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell*, **11**, 709–719.
  40. Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E. *et al.* (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, **122**, 517–527.
  41. Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J. and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature*, **419**, 407–411.
  42. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. and Kanemaki, M. (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods*, **6**, 917–922.
  43. Shetty, A., Reim, N.I. and Winston, F. (2019) Auxin-inducible degron system for depletion of proteins in *Saccharomyces cerevisiae*. *Curr. Protoc. Mol. Biol.*, **128**, e104.
  44. Longtine, M.S., McKenzie, A. 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*, **14**, 953–961.
  45. Simon, J.R., Treger, J.M. and McEntee, K. (1999) Multiple independent regulatory pathways control UBI4 expression after heat shock in *Saccharomyces cerevisiae*. *Mol. Microbiol.*, **31**, 823–832.
  46. Collart, M.A. and Oliviero, S. (2001) Preparation of yeast RNA. *Curr. Protoc. Mol. Biol.*, doi:10.1002/0471142727.mb1312s23.
  47. Churchman, L.S. and Weissman, J.S. (2012) Native elongating transcript sequencing (NET-seq). *Curr. Protoc. Mol. Biol.*, **14**, 11–17.
  48. Mayer, A., di Iulio, J., Maleri, S., Eser, U., Vierstra, J., Reynolds, A., Sandstrom, R., Stamatoyannopoulos, J.A. and Churchman, L.S. (2015) Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell*, **161**, 541–554.
  49. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.
  50. Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S. *et al.* (2014) The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3*, **4**, 389–398.
  51. Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature*, **415**, 871–880.
  52. Pelechano, V., Wei, W. and Steinmetz, L.M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*, **497**, 127–131.
  53. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, **17**, 10–12.
  54. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
  55. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data



- Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
56. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
57. Andrews, S. (2014) FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics*, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
58. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T. and Salzberg, S.L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.*, **33**, 290–295.
59. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
60. 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. Biol. Cell*, **11**, 4241–4257.
61. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*, **102**, 15545–15550.
62. Hong, E.L., Balakrishnan, R., Dong, Q., Christie, K.R., Park, J., Binkley, G., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G. et al. (2008) Gene Ontology annotations at SGD: new data sources and annotation methods. *Nucleic Acids Res.*, **36**, D577–D581.
63. Aronesty, E. (2013) Comparison of sequencing utility programs. *Open Bioinformatics J.*, **7**, 1–8.
64. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
65. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
66. Hainer, S.J., Pruneski, J.A., Mitchell, R.D., Monteverde, R.M. and Martens, J.A. (2011) Intergenic transcription causes repression by directing nucleosome assembly. *Genes Dev.*, **25**, 29–40.
67. O'Duibhir, E., Lijnzaad, P., Benschop, J.J., Lenstra, T.L., van Leenen, D., Groot Koerkamp, M.J., Margaritis, T., Brok, M.O., Kemmeren, P. and Holstege, F.C. (2014) Cell cycle population effects in perturbation studies. *Mol. Syst. Biol.*, **10**, 732.
68. Dronamraju, R., Heppner, A.J., Shibata, Y., Adams, A.T., Magnuson, T., Davis, I.J. and Strahl, B.D. (2018) Spt6 association with RNA polymerase II directs mRNA turnover during transcription. *Mol. Cell*, **70**, 1054–1066.
69. Gouot, E., Bhat, W., Rufange, A., Fournier, E., Paquet, E. and Nourani, A. (2018) Casein kinase 2 mediated phosphorylation of Spt6 modulates histone dynamics and regulates spurious transcription. *Nucleic Acids Res.*, **46**, 7612–7630.
70. Uwimana, N., Collin, P., Jeronimo, C., Haibe-Kains, B. and Robert, F. (2017) Bidirectional terminators in *Saccharomyces cerevisiae* prevent cryptic transcription from invading neighboring genes. *Nucleic Acids Res.*, **45**, 6417–6426.
71. van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R. and Nislow, C. (2013) A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. *PLoS Genet.*, **9**, e1003479.
72. Buratowski, S. (2009) Progression through the RNA polymerase II CTD cycle. *Mol. Cell*, **36**, 541–546.
73. Sdano, M.A., Fulcher, J.M., Palani, S., Chandrasekharan, M.B., Parnell, T.J., Whitby, F.G., Formosa, T. and Hill, C.P. (2017) A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *eLife*, **6**, e28723.
74. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. and Cramer, P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature*, **560**, 607–612.
75. Burugula, B.B., Jeronimo, C., Pathak, R., Jones, J.W., Robert, F. and Govind, C.K. (2014) Histone deacetylases and phosphorylated polymerase II C-terminal domain recruit Spt6 for cotranscriptional histone reassembly. *Mol. Cell Biol.*, **34**, 4115–4129.
76. Chu, Y., Simic, R., Warner, M.H., Arndt, K.M. and Prelich, G. (2007) Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.*, **26**, 4646–4656.
77. Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M. et al. (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol. Cell*, **11**, 721–729.
78. Du, H.N. and Briggs, S.D. (2010) A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *J. Biol. Chem.*, **285**, 11704–11713.
79. Du, H.N., Fingerman, I.M. and Briggs, S.D. (2008) Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. *Genes Dev.*, **22**, 2786–2798.
80. Wang, Y., Niu, Y. and Li, B. (2015) Balancing acts of SRI and an auto-inhibitory domain specify Set2 function at transcribed chromatin. *Nucleic Acids Res.*, **43**, 4881–4892.
81. Sayou, C., Millan-Zambrano, G., Santos-Rosa, H., Petfalski, E., Robson, S., Houseley, J., Kouzarides, T. and Tollervy, D. (2017) RNA binding by histone methyltransferases Set1 and Set2. *Mol. Cell Biol.*, **37**, e00165-17.
82. Perales, R., Erickson, B., Zhang, L., Kim, H., Valiquett, E. and Bentley, D. (2013) Gene promoters dictate histone occupancy within genes. *EMBO J.*, **32**, 2645–2656.
83. Rao, B., Shibata, Y., Strahl, B.D. and Lieb, J.D. (2005) Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. *Mol. Cell Biol.*, **25**, 9447–9459.
84. Soares, L.M., He, P.C., Chun, Y., Suh, H., Kim, T. and Buratowski, S. (2017) Determinants of histone H3K4 methylation patterns. *Mol. Cell*, **68**, 773–785.
85. Luce, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
86. Sorenson, M.R., Jha, D.K., Ucles, S.A., Flood, D.M., Strahl, B.D., Stevens, S.W. and Kress, T.L. (2016) Histone H3K36 methylation regulates pre-mRNA splicing in *Saccharomyces cerevisiae*. *RNA Biol.*, **13**, 412–426.
87. Leung, C.S., Douglass, S.M., Morselli, M., Obusan, M.B., Pavlyukov, M.S., Pellegrini, M. and Johnson, T.L. (2019) H3K36 methylation and the chromodomain protein Eaf3 are required for proper cotranscriptional spliceosome assembly. *Cell Rep.*, **27**, 3760–3769.
88. Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F., Xu, W., Mu, S., Wen, H., Qiu, J. et al. (2014) BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol. Cell*, **56**, 298–310.
89. Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R. and Bickmore, W.A. (2012) Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.*, **8**, e1002717.
90. Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q. and Ding, J. (2006) Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res.*, **34**, 6621–6628.
91. Alexander, R.D., Innocente, S.A., Barrass, J.D. and Beggs, J.D. (2010) Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell*, **40**, 582–593.
92. Chathoth, K.T., Barrass, J.D., Webb, S. and Beggs, J.D. (2014) A splicing-dependent transcriptional checkpoint associated with pre-spliceosome formation. *Mol. Cell*, **53**, 779–790.
93. Lee, Y., Park, D. and Iyer, V.R. (2017) The ATP-dependent chromatin remodeler Chd1 is recruited by transcription elongation factors and maintains H3K4me3/H3K36me3 domains at actively transcribed and spliced genes. *Nucleic Acids Res.*, **45**, 8646.
94. Radman-Livaja, M., Quan, T.K., Valenzuela, L., Armstrong, J.A., van Welsom, T., Kim, T., Lee, L.J., Buratowski, S., van Leeuwen, F., Rando, O.J. et al. (2012) A key role for Chd1 in histone H3 dynamics at the 3' ends of long genes in yeast. *PLoS Genet.*, **8**, e1002811.
95. Jeronimo, C., Poitras, C. and Robert, F. (2019) Histone recycling by FACT and Spt6 during transcription prevents the scrambling of histone modifications. *Cell Rep.*, **28**, 1206–1218.

96. Fong,N., Saldi,T., Sheridan,R.M., Cortazar,M.A. and Bentley,D.L. (2017) RNA Pol II dynamics modulate Co-transcriptional chromatin modification, CTD phosphorylation, and transcriptional direction. *Mol. Cell*, **66**, 546–557.
97. Tettey,T.T., Gao,X., Shao,W., Li,H., Story,B.A., Chitsazan,A.D., Glaser,R.L., Goode,Z.H., Seidel,C.W., Conaway,R.C. *et al.* (2019) A role for FACT in RNA polymerase II Promoter-Proximal pausing. *Cell Rep.*, **27**, 3770–3779.
98. Simon,J.M., Hacker,K.E., Singh,D., Brannon,A.R., Parker,J.S., Weiser,M., Ho,T.H., Kuan,P.F., Jonasch,E., Furey,T.S. *et al.* (2014) Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome Res.*, **24**, 241–250.
99. de Almeida,S.F., Grosso,A.R., Koch,F., Fenouil,R., Carvalho,S., Andrade,J., Levezinho,H., Gut,M., Eick,D., Gut,I. *et al.* (2011) Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.*, **18**, 977–983.
100. Kim,S., Kim,H., Fong,N., Erickson,B. and Bentley,D.L. (2011) Pre-mRNA splicing is a determinant of histone H3K36 methylation. *PNAS*, **108**, 13564–13569.
101. Adkins,M.W. and Tyler,J.K. (2006) Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Mol. Cell*, **21**, 405–416.
102. McCullough,L., Connell,Z., Petersen,C. and Formosa,T. (2015) The abundant histone chaperones Spt6 and FACT collaborate to assemble, inspect, and maintain chromatin structure in *Saccharomyces cerevisiae*. *Genetics*, **201**, 1031–1045.
103. Koster,J. and Rahmann,S. (2018) Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*, **28**, 2520–2522.