Expanded View Figures

Figure EV1. Generation of Spt5-depleted (Spt5^{dep}) mice and characterization of Spt5^{dep} mouse embryonic fibroblasts (MEFs).

A Diagram of the targeting vector obtained from the Sanger Consortium. Exons 14–16 of the *Supt5h* gene are flanked by loxP sites. Additionally, the intron between exons 13 and 14 has a selection cassette comprising of a FLP recombination target (FRT) site, a splice acceptor (SA) and a β-galactosidase cassette (shown for simplicity as LacZ), followed by a phosphoglycerate kinase (*Pgk*) promoter-driven neomycin resistance (NeoR) cassette, a SV40 poly-adenylation (pA) site (not shown) and a second FRT site. Following germline transmission, the entire region between the two FRT sites was deleted by crossing with mice expressing Flpe recombinase to generate *Supt5h*$^{Fl/+}$ mice which were then used to generate homozygous *Supt5h*$^{Fl/Fl}$ mice. Heterozygous *Supt5h*$^{-/+}$ mice were generated by crossing *Supt5h*$^{Fl/+}$ mice with EIIA-cre mice wherein Cre recombinase is active in the germline. *Supt5h*$^{-/+}$ mice were bred with *Rosa26*$^{Cre\text{-}ERT2}$ mice, which express the Cre recombinase fused to the tamoxifen-inducible estrogen receptor domain constitutively from the *Rosa26* locus. Finally, *Supt5h*$^{Fl/Fl}$ mice were crossed with *Supt5h*$^{-/+}$ *Rosa26*$^{Cre\text{-}ERT2/Cre\text{-}ERT2}$ mice to give the final genotype, *Supt5h*$^{Fl/-}$*Rosa26*$^{Cre\text{-}ERT2/+}$, used in this study to generate Spt5-depleted (Spt5^{dep}) primary MEFs.

B Overview of the experimental scheme to generate WT (*Rosa26*$^{Cre\text{-}ERT2/+}$) and Spt5^{dep} (*Supt5h*$^{Fl/-}$*Rosa26*$^{Cre\text{-}ERT2/+}$) primary MEFs. MEFs were prepared from d13.5 embryos. After three passages, MEFs were treated with 2 μM 4-hydroxytamoxifen (4-HT) for 72 h and then harvested for all downstream analyses.

C Viability curve for WT and Spt5^{dep} primary MEFs treated with 4-HT for 72 h as described in (B). Cells were counted using the Trypan Blue dye exclusion approach at the indicated time-points. The data represent MEFs prepared from four mice for each genotype. For each time point, the mean with standard deviation is plotted.

D Multiplex PCR-based genotyping from genomic DNA of WT and Spt5^{dep} primary MEFs. For each genotype, MEFs from two different mice were analyzed.

E UCSC genome browser snapshot of the *Supt5h* locus showing Pol II ChIPseq, GROseq, and mRNAseq profiles. The floxed segment comprising exons 14–16 is boxed. WT and Spt5^{dep} tracks are overlaid in blue and red representing WT and Spt5^{dep} tracks, respectively, and black representing the overlap between them.

F The left panel shows the mRNAseq read counts in the floxed exons 14–16 of *Supt5h* mRNA and in the remaining *Supt5h* exons in WT and Spt5^{dep} samples. In the right panel box plot, the ratio of the *Supt5h* exons 14–16 to remaining *Supt5h* exons is depicted for WT and Spt5^{dep} samples.

G Western blot analysis from nuclear extracts of two independent WT and Spt5^{dep} primary MEFs preparations treated with 4-HT for 72 h as described in (B). Blots were probed for Spt5 and histone H3 as a loading control. A twofold titration of nuclear extract was performed for each sample.

H ChIP-qPCR analysis for Spt5 occupancy at promoters and gene bodies of several long genes in WT and Spt5^{dep} MEFs treated with 4-HT for 72 h. Within gene bodies, a proximal (GB proximal; < 15 kb from the TSS) and distal (GB distal; > 50 kb from the TSS) region was chosen for this analysis. The negative control region represents a gene desert on chromosome 1 that is transcriptionally inert. The data were normalized to the input DNA in all cases and represented as percent of input. The plot represents three independent experiments with mean and standard deviations shown.

Source data are available online for this figure.

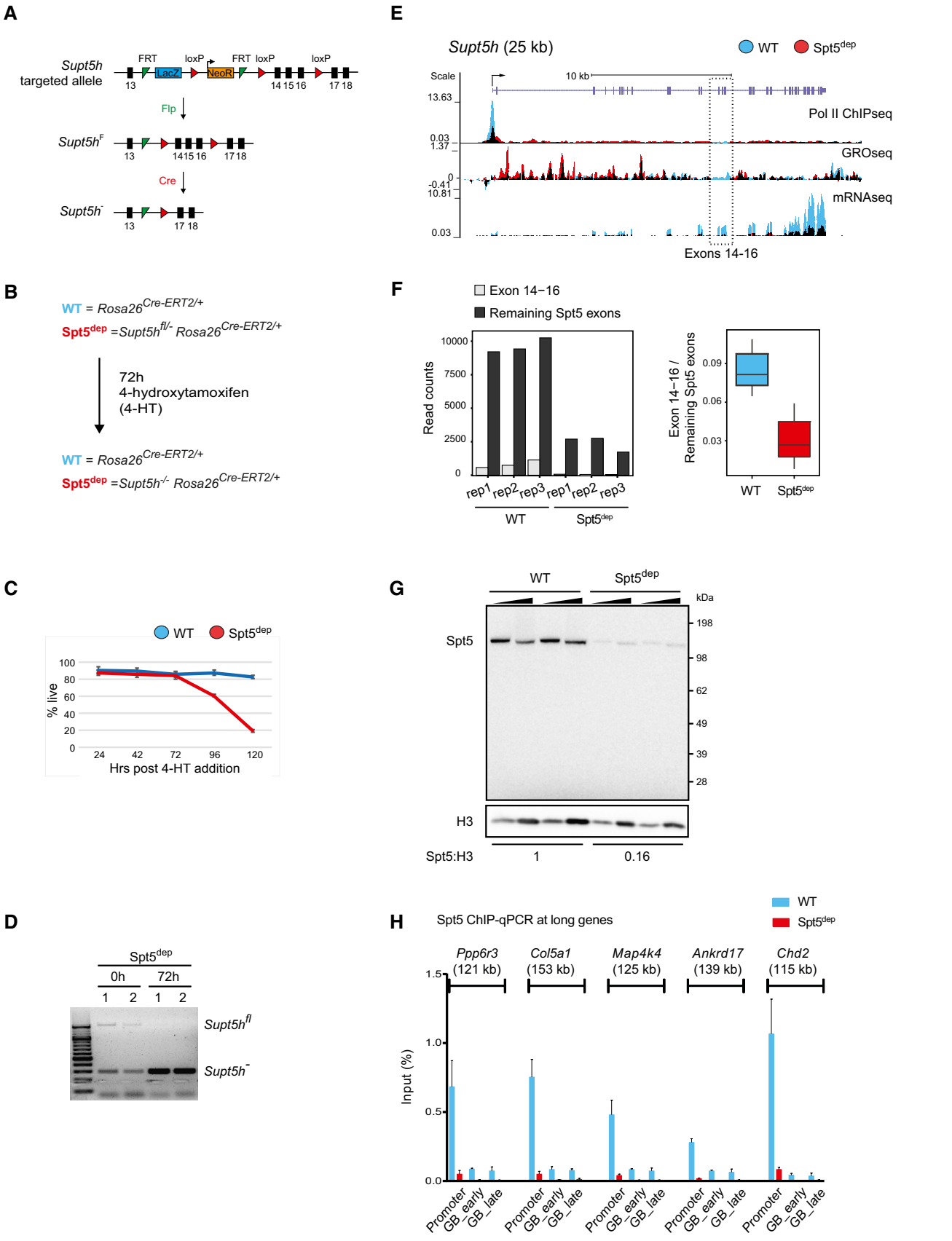

Figure EV1.

Figure EV2. Correlation between replicate datasets and classification of the expressed gene list.

A Replicate correlation plots for GROseq and Pol II, Ser5P-Pol II and Ser2P-Pol II ChIPseq datasets from WT and Spt5^{dep} MEFs used in this study. Spearman's correlation coefficient (*r*) is indicated within the plots.

B Classification of all expressed genes used in this study. The heatmaps show gene body (GB) densities of Pol II and Ser2P-Pol II as well as mRNAseq from WT MEFs ordered by GROseq GB density (TSS +0.5 kb to TTS). The genes were grouped into three equal groups termed groups 1, 2, and 3, which comprise genes with highest, intermediate, and lowest GROseq GB density, respectively.

C Replicate correlation plots for the triplicate mRNAseq datasets from WT and Spt5^{dep} MEFs used in this study. Spearman's correlation coefficient (*r*) is indicated.

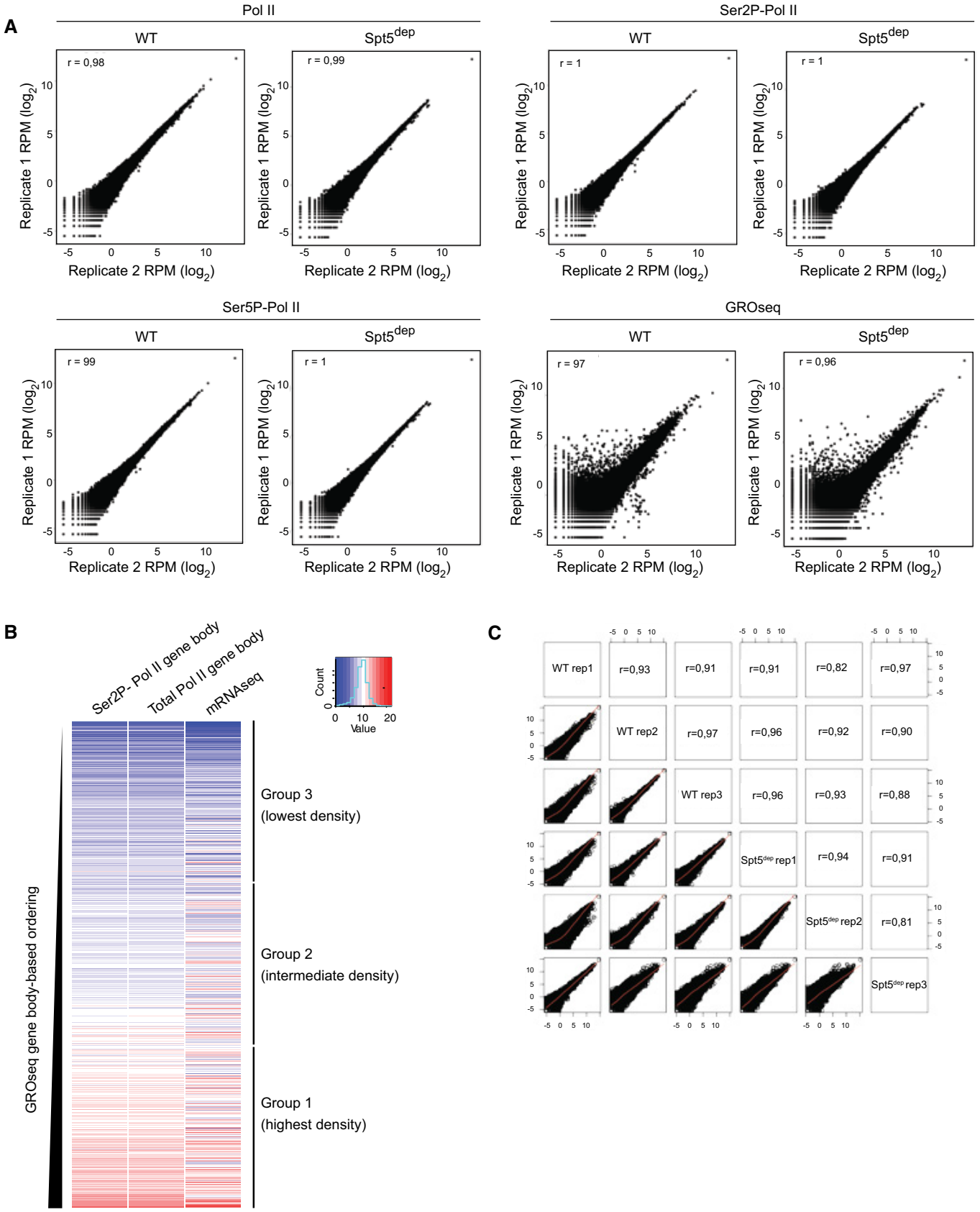

Figure EV2.

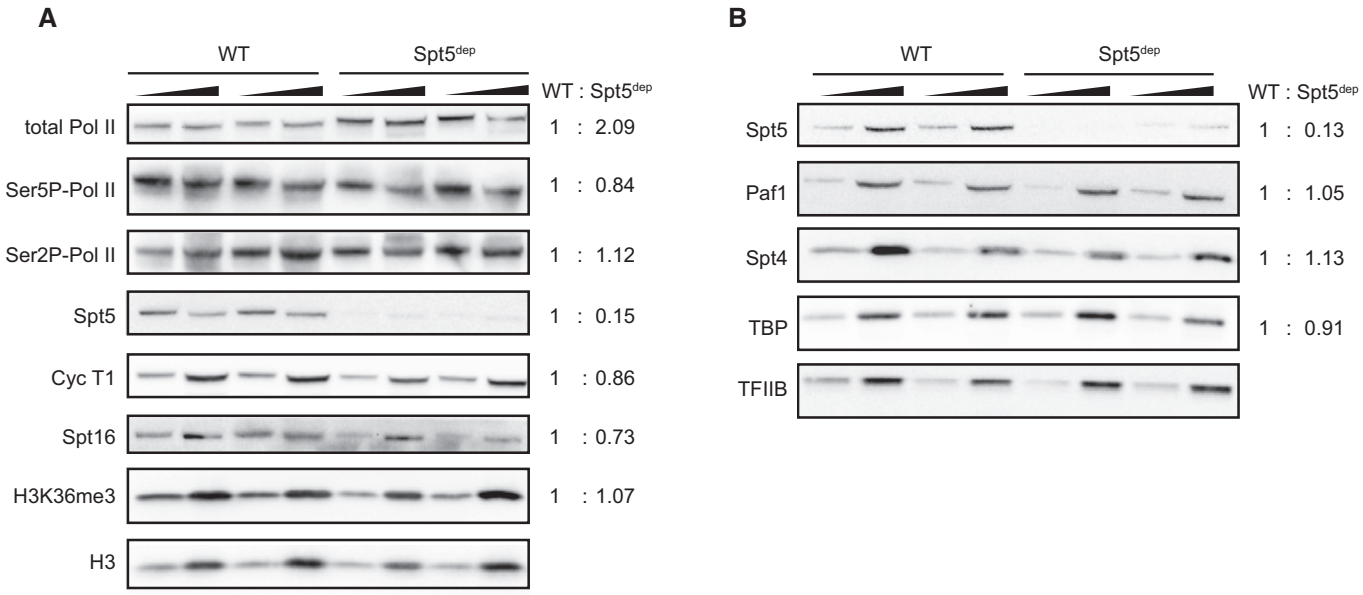

Figure EV3. Western blot analysis for selected transcription initiation and elongation factors.

A Western blot analysis from nuclear extracts of two independent WT and Spt5^{dep} primary MEFs preparations treated with 4-HT for 72 h. Blots were probed for total Pol II, Ser5P phosphorylated Pol II, Ser2P phosphorylated Pol II, Spt5, cyclin T1, Spt16, H3K36me3 and histone H3 as a loading control. A twofold titration was performed for each sample. The mean of the titration was normalized to the mean of the loading control. The mean of the two H3-normalized WT replicates was set as 1, and the ratio of the mean of both H3-normalized Spt5^{dep} replicates was calculated accordingly.

B Western blot analysis from extracts of two independent WT and Spt5^{dep} primary MEFs preparations treated with 4-HT for 72 h. Blots were probed for Spt5, Paf1, Spt4, TBP, and TFIIB as a loading control. A twofold titration of extracts was performed for each sample. Quantification was done as described in (A).

Source data are available online for this figure.

Figure EV4. Examples of long and short genes showing the impact of Spt5 depletion on nascent transcription and Pol II occupancy.

A UCSC genome browser snapshots of five long genes in WT and Spt5^{dep} MEFs. WT and Spt5^{dep} tracks are overlaid in blue and red representing WT and Spt5^{dep} signal respectively, and black representing the overlap between them. The insets provide magnified views of the gene body region.

B UCSC genome browser snapshots of five short genes in WT and Spt5^{dep} MEFs represented as in (A). The insets provide magnified views of the gene body region.

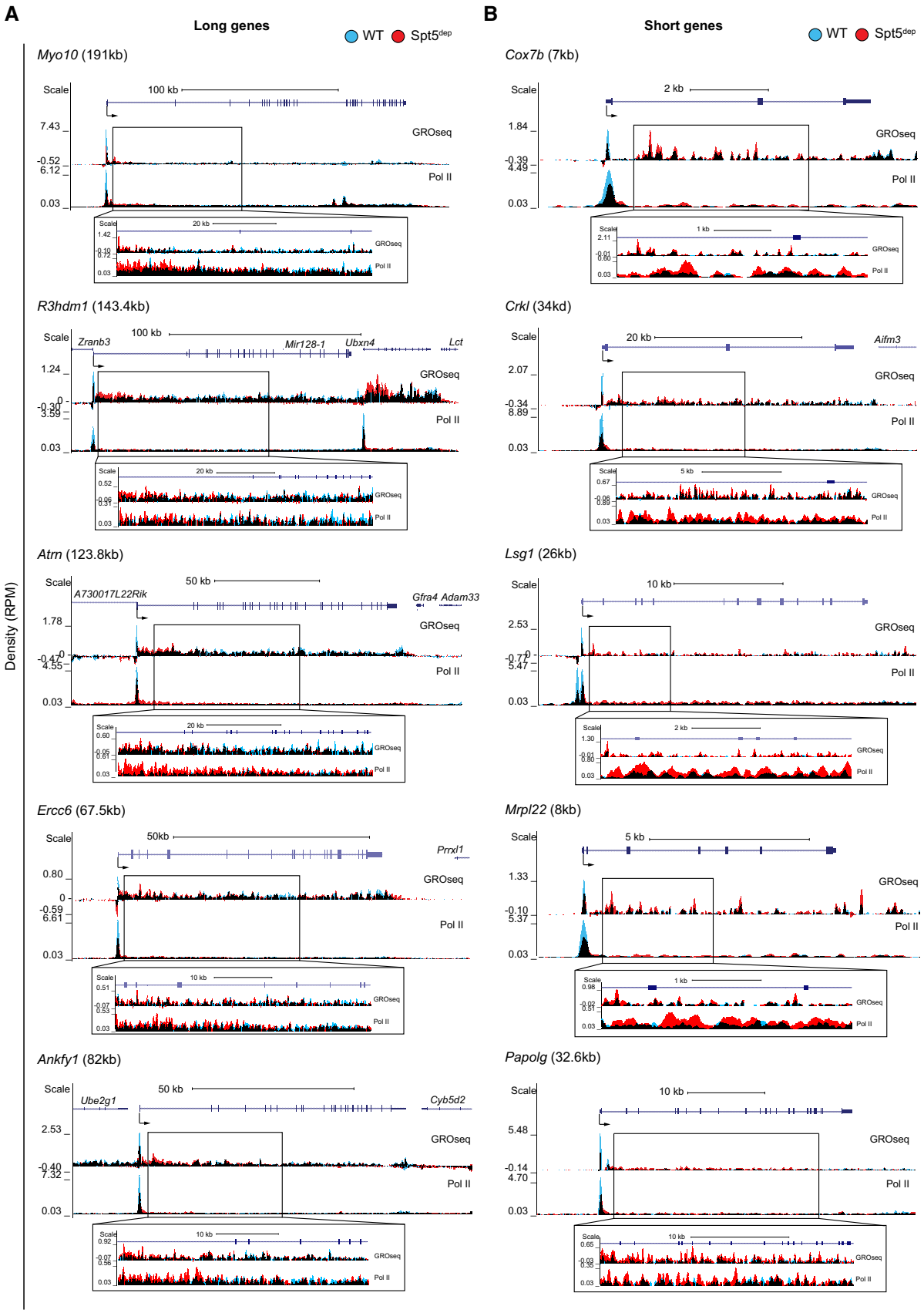

Figure EV4.

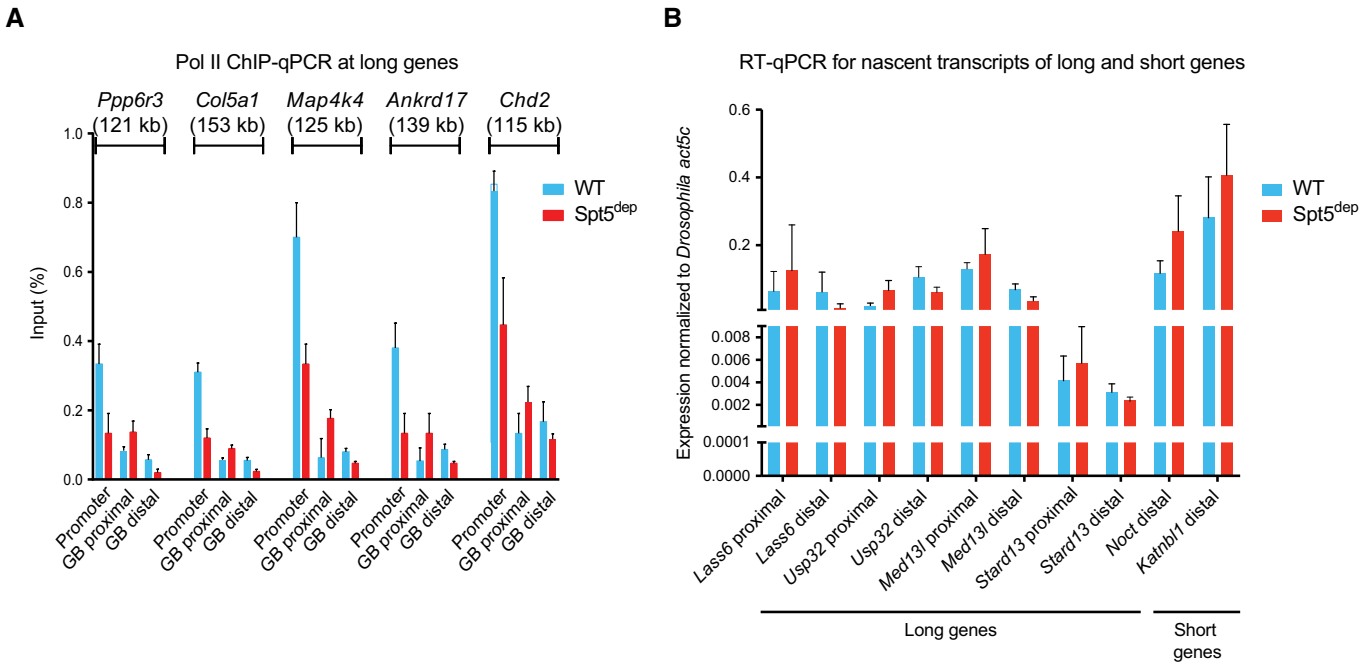

Figure EV5. ChIP-qPCR and RT–qPCR validation of Pol II occupancy, nascent transcription, and mRNA abundance between WT and Spt5^{dep} cells.

A ChIP-qPCR at promoters and gene bodies (GB) of five long genes in WT and Spt5^{dep} cells. Two primer pairs probing a proximal (< 15 kb from the TSS; GB proximal) and distal (< 50 kb from the TSS, GB distal) region of the GB were used for each gene. The data were normalized to the input DNA in all cases and represented as percent of input. The plot represents three independent experiments with mean and standard deviations shown.

B RT–qPCR analysis in WT and Spt5^{dep} cells for nascent transcript abundance using primers in intronic regions within four long and two short genes. For the long genes, the two primer pairs shown probe a proximal and distal region of the gene. Equal numbers of *Drosophila* S2 cells were spiked in to each sample prior to RNA extraction, and the *Drosophila*-specific *act5a* transcript was used for normalization between WT and Spt5^{dep} qPCR values. The plot shows the mean of three independent WT and Spt5^{dep} samples with standard deviations.

Figure EV6. Analysis of GROseq antisense transcripts in Spt5^{dep} cells.

A Spt5^{dep}/WT ratio heatmaps with static binning spanning the entire length of genes up to 100 kb long ordered by increasing gene length (n = 3,473). The sense and antisense ratio heatmaps are shown on the left and right, respectively. Read densities are in reads per million (RPM). The TTS is marked with a black line. The arrow above the plot marks the ~17 kb region downstream of the TSS where, on average, the sense strand Spt5^{dep} and WT signals intersect (see also Fig 3A–C).

B Ratio heatmaps in (A) focused on the TSS −2 kb to +20 kb region for genes > 25 kb (n = 1,263).

C Ratio heatmaps generated as in (A) for the TSS −/+500 bp region. The sense strand heatmap was subjected to k-means clustering, depicted on the left, with the corresponding antisense heatmap on the right.

D Spt5^{dep}/WT ratio box plots of GROseq antisense reads in the TSS upstream (TSS −500 bp) and TSS downstream (TSS +500 bp) regions.

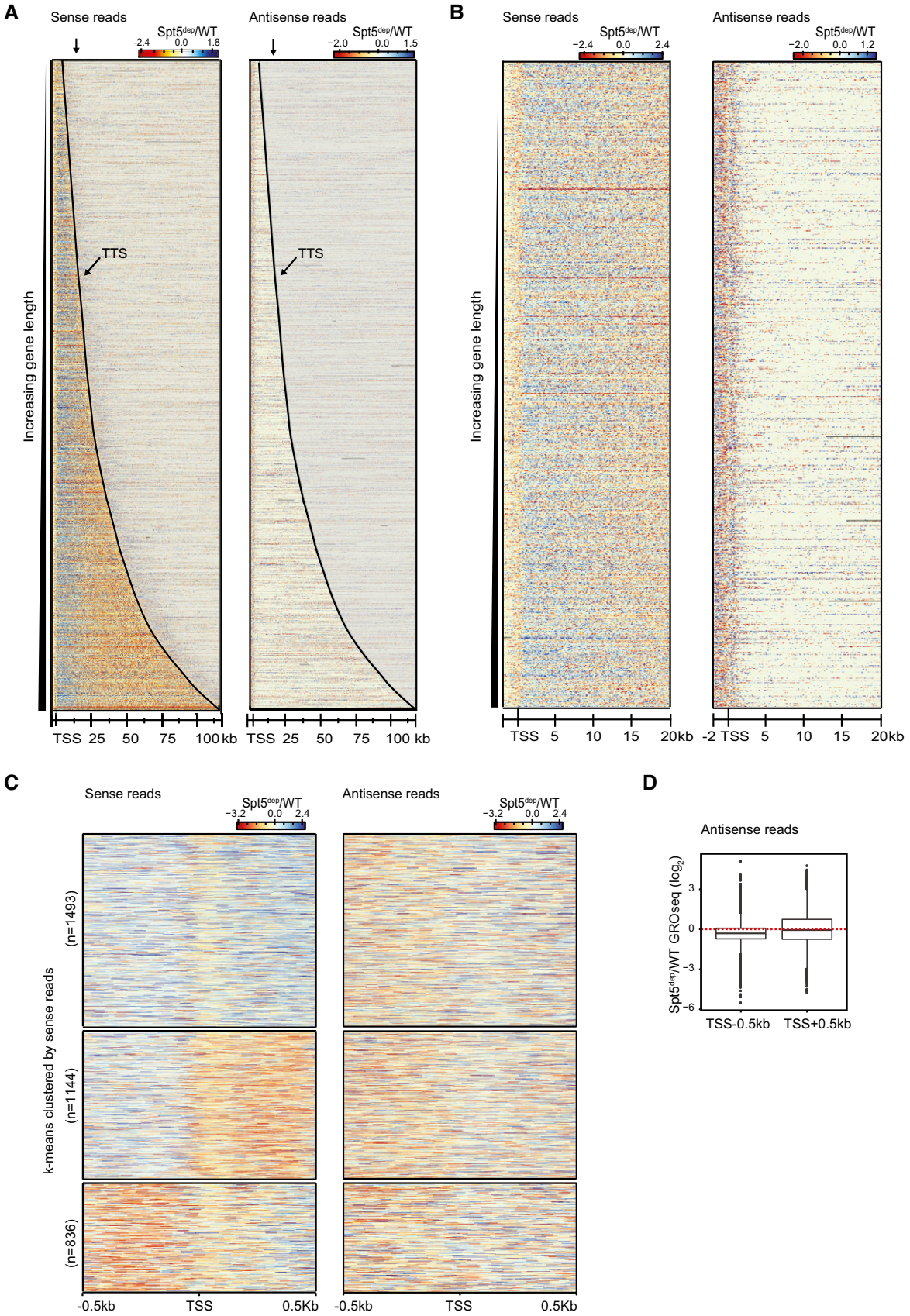

Figure EV6.