

Cell Reports, Volume 41

Supplemental information

**Distinct patterns of RNA polymerase II and
transcriptional elongation characterize
mammalian genome activation**

Kenichiro Abe, Tamas Schauer, and Maria-Elena Torres-Padilla

Figure S1

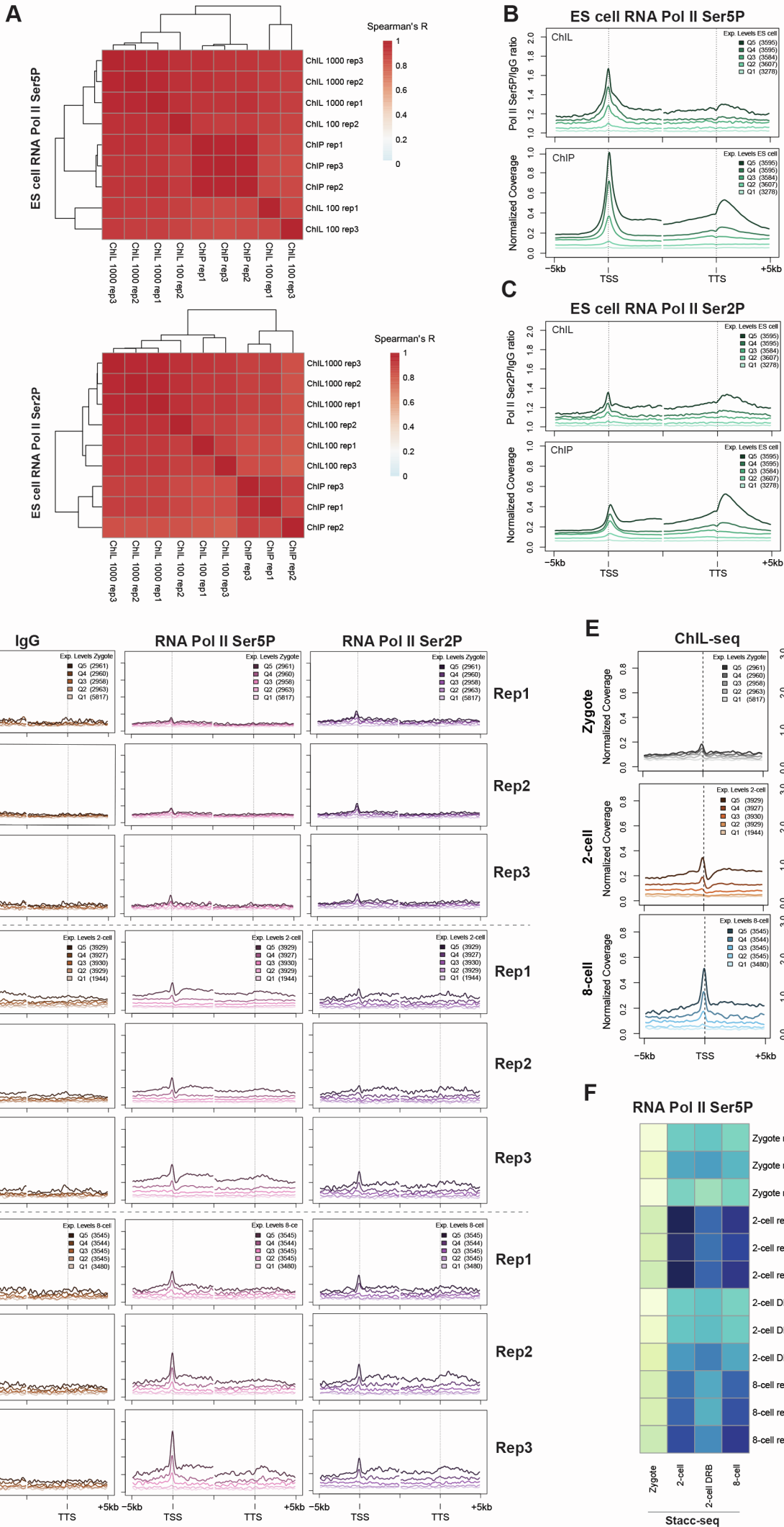


Figure S1. ChIL-seq maps phosphorylated Pol II, related to Figure 1

(A) Correlation heatmaps between Pol II Ser5P (top) or Pol II Ser2P (bottom) ChIL-seq data from 100 or 1000 ES cells and ChIP-seq data of bulk ES cells. Correlations are calculated based on read counts at the gene body from 3 biological replicates.

(B and C) Metaplots comparing Pol II Ser5P (B) and Pol II Ser2P (C) ChIL-seq data of 100 ES cells (top) and ChIP-seq data of bulk ES cells (bottom). Metaplots depict ChIL-seq mean coverage normalized by IgG and ChIP-seq mean coverage centered at TSS and TTS. Maternal genes⁴⁵ were removed and genes were grouped in quintiles based on their expression level⁴⁶. The first quintile includes genes with zero expression. The average of 3 biological replicates is shown.

(D) Metaplots comparing IgG (left), Pol II Ser5P (middle) and Pol II Ser2P (right) ChIL-seq mean coverage centered at TSS and TTS in zygote, 2-cell and 8-cell stage embryos. Three biological replicates are plotted independently. Maternal RNA genes⁴⁵ were removed and genes were grouped in quintiles based on their expression levels⁴⁶. The first quintile includes genes with zero expression.

(E) Metaplots comparing Pol II Ser5P ChIL-seq (left) and published Pol II Ser5P Stacc-seq³⁵ (right) mean coverage centered at TSS in embryos. Maternal RNA genes⁴⁵ were removed and genes were grouped in quintiles based on their expression levels in the given stage⁴⁶. The first quintile includes genes with zero expression. The average of 3 biological replicates of ChIL-seq and 2 biological replicates of Stacc-seq are shown. We note that some differences between the Stacc-seq and ChIL-seq could arise due to slightly different time points at which embryos were collected and the different antibodies for Pol II Ser5P used.

(F) Correlation heatmap in embryo datasets between Pol II Ser5P ChIL-seq and published Stacc-seq generated using an antibody against phosphorylated Ser5P³⁵. Spearman's correlation coefficients were calculated based on gene body counts from on 3 biological replicates of ChIL-seq and 2 biological replicates of Stacc-seq.

Figure S2

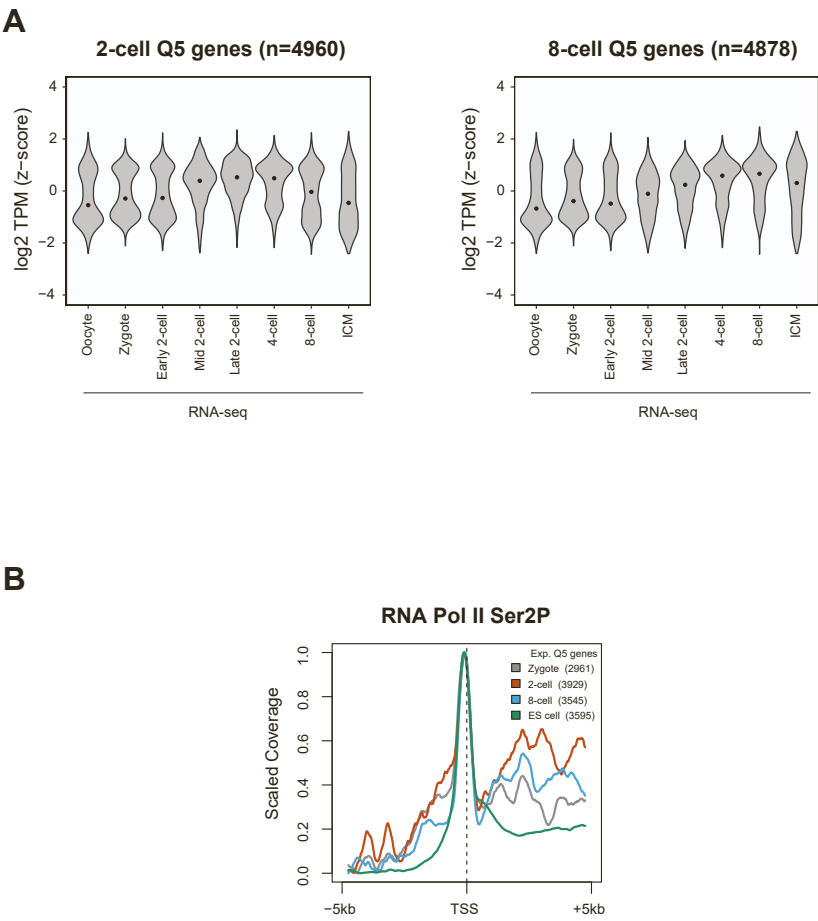


Figure S2. Dynamics of phosphorylated Pol II accumulation during development, related to Figure 2

(A) Violin plots indicating mean expression levels (z-score of transcripts per million [TPM]) of genes in the highest expression quintile of 2-cell (left) or 8-cell (right) stage embryos throughout development.^{46,73} ‘n’ indicates the number of genes.

(B) Scaled metaplots of Pol II Ser2P ChIL-seq mean coverage centered at TSS comparing embryos and ES cells. Maternal RNA genes⁴⁵ were removed and genes in the highest expression quintile are included. The average of 3 biological replicates is shown.

Figure S3

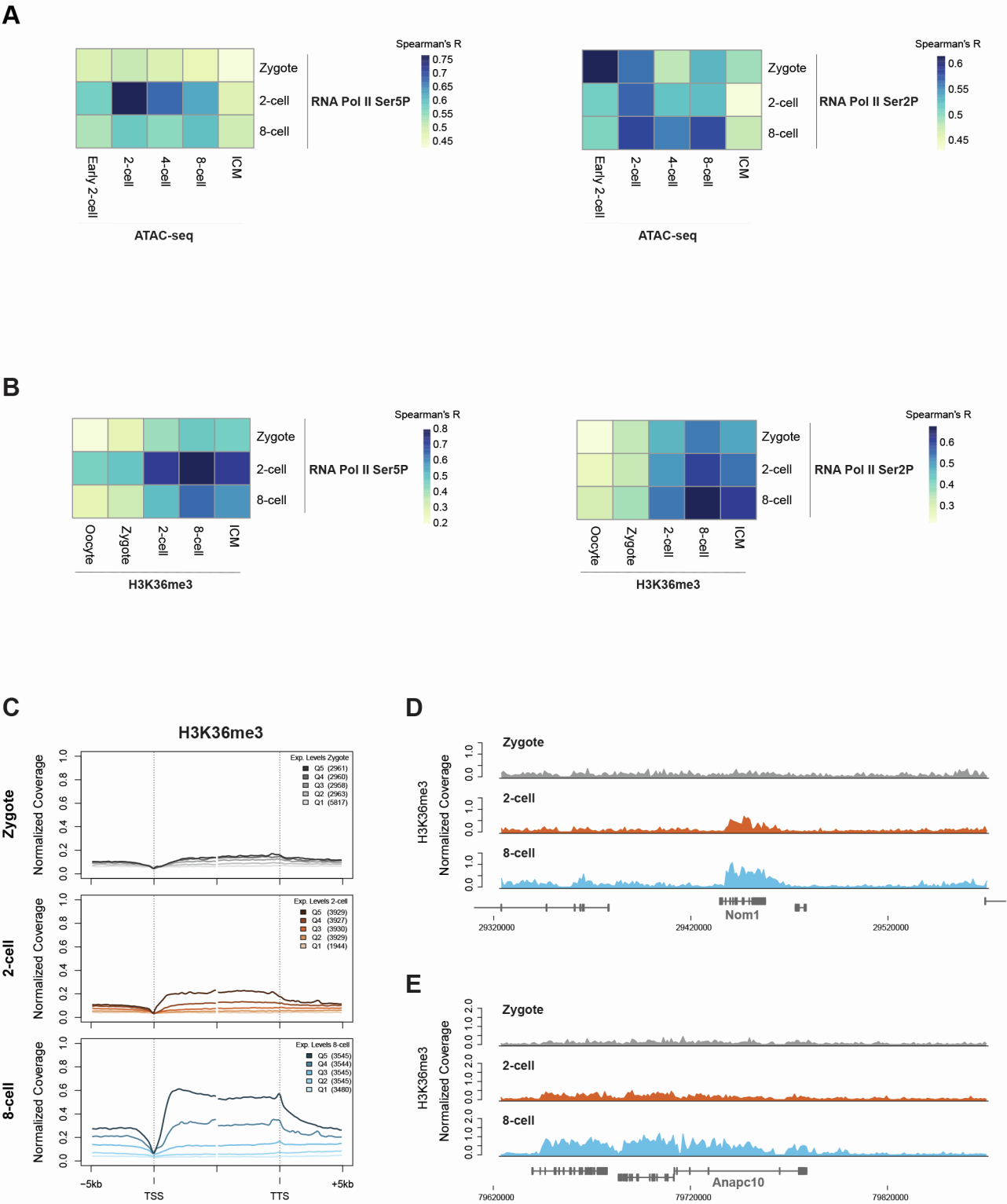


Figure S3. Pol II Ser5 and Ser2 phosphorylation patterns at stage-specific genes, related to Figure 3

(A and B) Correlation heatmaps between Pol II Ser5P (left) or Pol II Ser2P (right) ChIL-seq and ATAC-seq (A) or H3K36me3 ChIP-seq (B) embryonic datasets^{47,76} read counts at the gene body at the indicated developmental stages. Spearman's correlation coefficients were calculated on the mean of 3 biological replicates of ChIL-seq, 2 biological replicates of ATAC-seq or 2 biological replicates of H3K36me3 ChIP-seq.

(C) Metaplots of H3K36me3 ChIP-seq datasets⁴⁷ mean coverage centered at TSS and TTS. Genes were grouped in quintiles based on their expression levels in the given stage⁴⁶. The first quintile includes genes with zero expression. Maternal RNA genes⁴⁵ were removed from the analysis. The average of 2 biological replicates is shown.

(D) Example genomic regions showing H3K36me3 ChIP-seq⁴⁷ coverage in zygote, 2-cell and 8cell stage embryos at an example gene (Nom1) displaying highest Pol II Ser2P across the gene body in 2-cell stage embryos, corresponding to its transcriptional activation. The average of 2 biological replicates is shown.

(E) Example genomic regions showing H3K36me3 ChIP-seq⁴⁷ coverage at an example of '4-cell transient' gene (Anapc10) in zygotes, 2- and 8-cell stage embryos. The average of 2 biological replicates is shown.

Figure S4

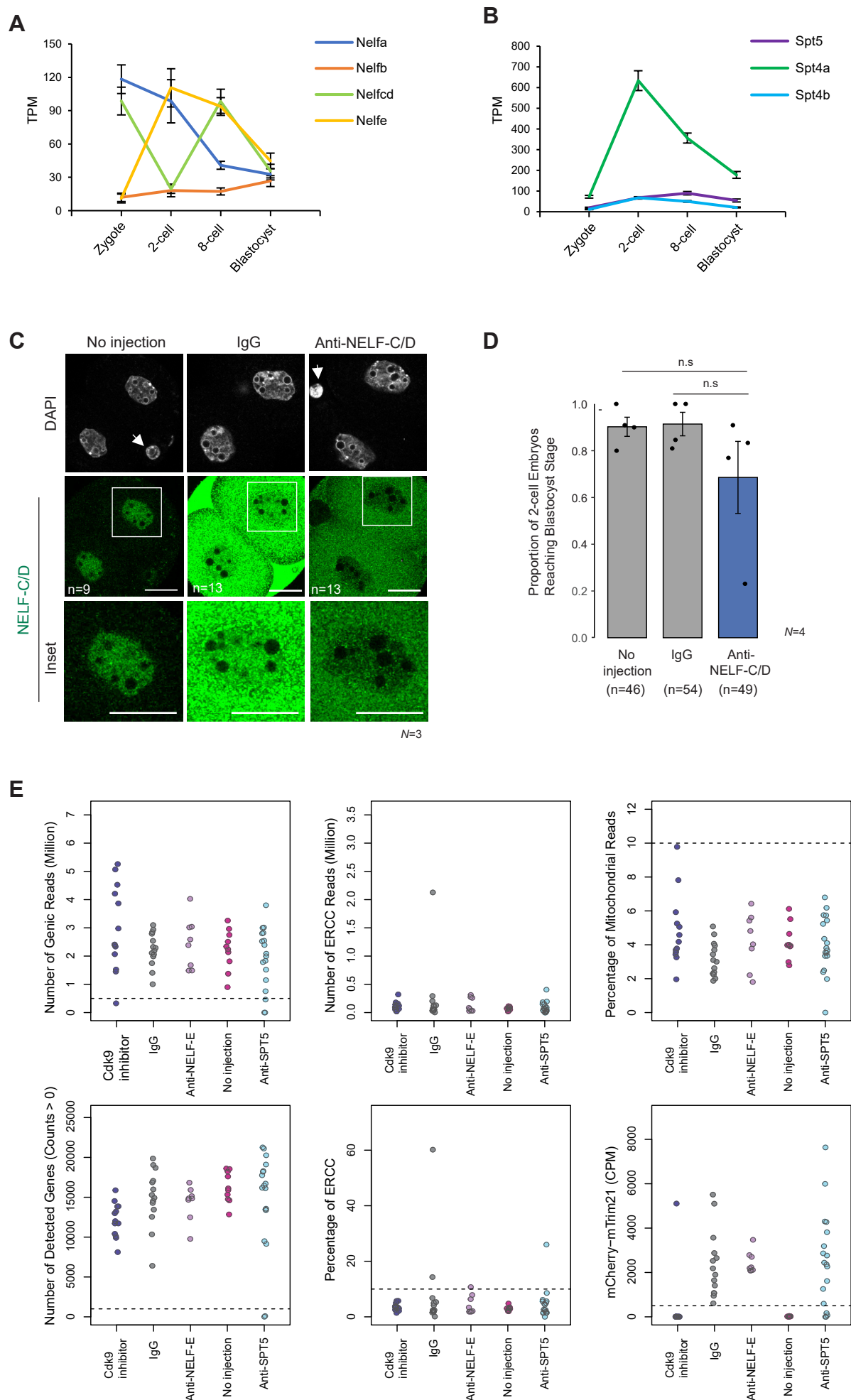


Figure S4. Expression of NELF and DSIF components, Trim-Away for NELF-C/D and quality control analysis of RNA-seq datasets, related to Figure 4

(A and B) Expression levels (transcripts per million [TPM]) of NELF (A) and DSIF subunits (B) from published embryonic RNA-seq datasets at the indicated stages⁴⁶. Error bars indicate mean \pm standard error of the mean (SEM) from single cells.

(C) Immunostaining using anti-NELF-C/D antibody in 2-cell stage embryos after Trim-Away for NELF-C/D or negative controls (Trim-Away for IgG and non-injected embryos). 'n' indicates the total number of embryos analyzed and 'N' the number of independent experiments. Scale bars indicate 50 μ m.

(D) Developmental progression plotted as the proportion of control (non-injected or embryos under Trim-Away against IgG) or Trim-Away NELF-C/D embryos reaching the blastocyst stage. 'n' indicates the total number of embryos analyzed and 'N' the number of independent. Graphs show the mean \pm standard error of the mean (SEM) obtained in individual experiments. P-values were obtained by analysis of variance.

(E) Quality filter of single embryo RNA-seq. Top-left, total sum of read counts per gene (minimum threshold 5×10^5 reads); top-middle, total sum of read counts per ERCC spike-in; top-right, percentage of read counts at mitochondrial genes (maximum threshold 10%); bottom-left, number of detected genes with a read count greater than zero (minimum threshold 1.000 genes); bottom-middle, percentage of read counts at ERCC spike-in (maximum threshold 10%); read counts per million (CPM) aligned to the mCherry-mTrim21 construct (minimum threshold: 500 counts in injected samples). Thresholds are indicated as dashed line. The experiments were conducted 3 times for the collection of non-injected embryos and NELF-E depleted embryos and 5 times for the collection of IgG controls, SPT5 depleted embryos and CDK9 inhibition.

Figure S5

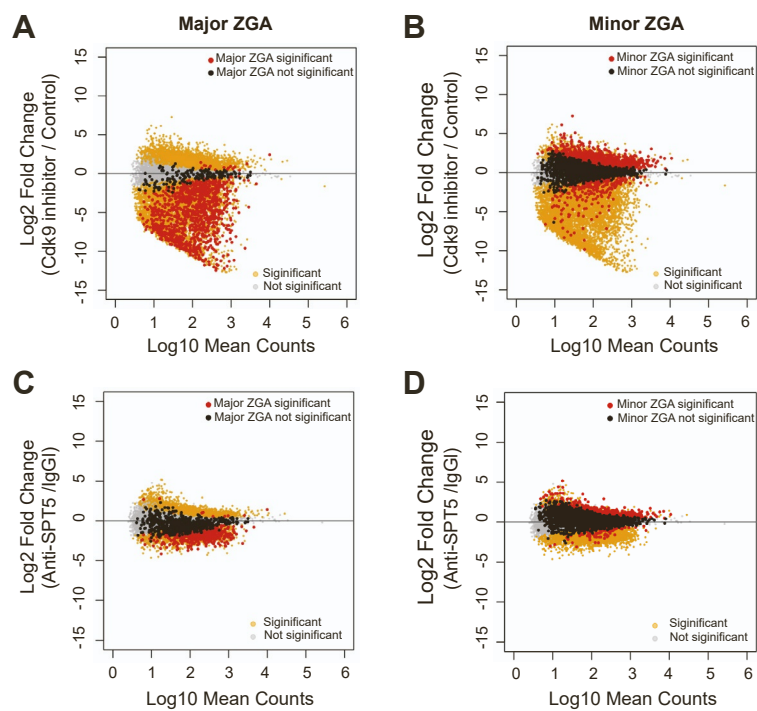


Figure S5. Differentially expressed major or minor ZGA genes upon CDK9 inhibition and SPT5 depletion, related to Figure 5

(A and B) MA plots comparing the log₂ fold change between 2-cell stage embryos treated with BAY1251125 and control, untreated embryos in relation to mean RNA-seq. Differentially expressed major ZGA (A) or minor ZGA (B) gene groups are marked in red ($\text{padj} < 0.05$), non-differential genes within the gene group in black, differentially expressed but not within the corresponding gene group in orange and non-differentially expressed and not in the gene group in grey.

(C and D) MA plots comparing the log₂ fold change between 2-cell stage embryos upon SPT5 depletion by Trim-Away and control embryos injected with IgG in relation to mean RNA-seq. Differentially expressed major ZGA (C) or minor ZGA (D) gene groups are marked in red ($\text{padj} < 0.05$), non-differential genes within the gene group in black, differentially expressed but not within the corresponding gene group in orange and non-differentially expressed and not in the gene group in grey.