Expanded View Figures

Figure EV1. The transcriptome and RNA metabolism in human cells hardly change upon CTD shortening.

A Spearman correlation of read counts in RefSeq genes in biological replicates of TT-seq experiments.

B Spearman correlation of read counts in RefSeq genes in biological replicates of total RNA- seq experiments.

C MA plot showing changes in RNA synthesis in TT-seq datasets upon expression of the α-amanitin-resistant full-length CTD variant in U2OS cells in steady state conditions. Wild-type U2OS cells are used as a control, and the data were normalized using spike-in counts. 16,214 expressed genes annotated in RefSeq were analyzed, 7 genes were significantly upregulated (adjusted P value < 0.05, $\log_2$ fold change ≥ 1, depicted in red) and 5 genes were significantly downregulated (adjusted P value < 0.05, $\log_2$ fold change ≤ 1, depicted in red).

D Gene ontology analysis of significantly downregulated genes (adjusted P value < 0.05, $\log_2$ fold change ≤ 1) upon CTD shortening. GO categories with FDR≤ 0.05 are shown.

E Proliferation curve generated for cells expressing the α-amanitin-resistant RPB1-52R and RPB1-25R CTD variants, as well as wild-type U2OS cells. Error bars show standard deviation of 3 biological replicates. Statistical significance was estimated using two-way ANOVA, followed by the Tukey multiple comparisons test. The difference in growth rate between RPB1-52R and RPB1-25R cells is significant (adjusted P value $= 0.0014322$).

F RNA synthesis of transposable elements in TT-seq datasets upon expression RPB1-25R CTD variant in U2OS cells in steady state conditions. Wild-type U2OS cells are used as a control and the data were normalized using spike-in counts. 1,026 expressed transposable elements were analyzed. Differentially expressed transposons are in red. Five were significantly upregulated (adjusted P value < 0.05, $\log_2$ fold change ≥ 1) and 9 were significantly downregulated (adjusted P value < 0.05, $\log_2$ fold change ≤ 1).

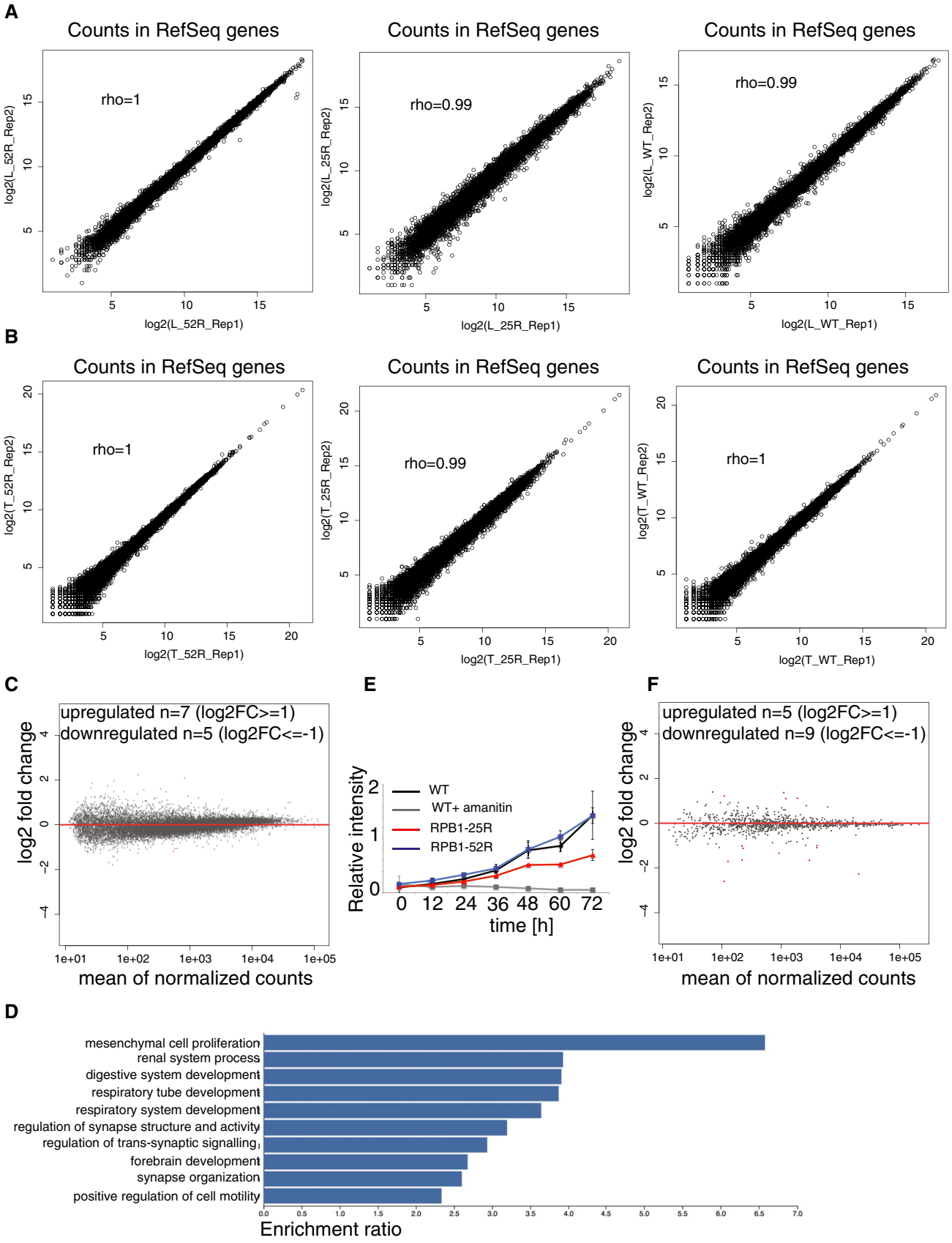

A
Counts in RefSeq genes
rho=1
log2(L_52R_Rep2)
log2(L_52R_Rep1)
Counts in RefSeq genes
rho=0.99
log2(L_25R_Rep2)
log2(L_25R_Rep1)
Counts in RefSeq genes
rho=0.99
log2(L_WT_Rep2)
log2(L_WT_Rep1)
B
Counts in RefSeq genes
rho=1
log2(T_52R_Rep2)
log2(T_52R_Rep1)
Counts in RefSeq genes
rho=0.99
log2(T_25R_Rep2)
log2(T_25R_Rep1)
Counts in RefSeq genes
rho=1
log2(T_WT_Rep2)
log2(T_WT_Rep1)
C
upregulated n=7 (log2FC>=1)
downregulated n=5 (log2FC<=-1)
log2 fold change
mean of normalized counts
E
Relative intensity
WT
WT+ amanitin
RPB1-25R
RPB1-52R
time [h]
F
upregulated n=5 (log2FC>=1)
downregulated n=9 (log2FC<=-1)
log2 fold change
mean of normalized counts
D
mesenchymal cell proliferation
renal system process
digestive system development
respiratory tube development
respiratory system development
regulation of synapse structure and activity
regulation of trans-synaptic signalling
forebrain development
synapse organization
positive regulation of cell motility
Enrichment ratio

Figure EV1.

Figure EV2. CTD shortening leads to minor changes in Pol II pausing at steady state.

A, B Spearman correlation of read counts in the first constitutive exon of genes encoding major isoforms in biological replicates of PRO-seq data.

C, D Scatter plot of pausing duration (d) and productive initiation frequencies (I) in genes encoding major isoforms in RPB1-52R and RPB1-25R cells. The gray-shaded area depicts impossible combinations of I and d according to published kinetic theory (Ehrensberger *et al*, 2013).

E, F Histogram of distances from the PAS to the transcription termination site (TTS) of genes encoding major isoforms.

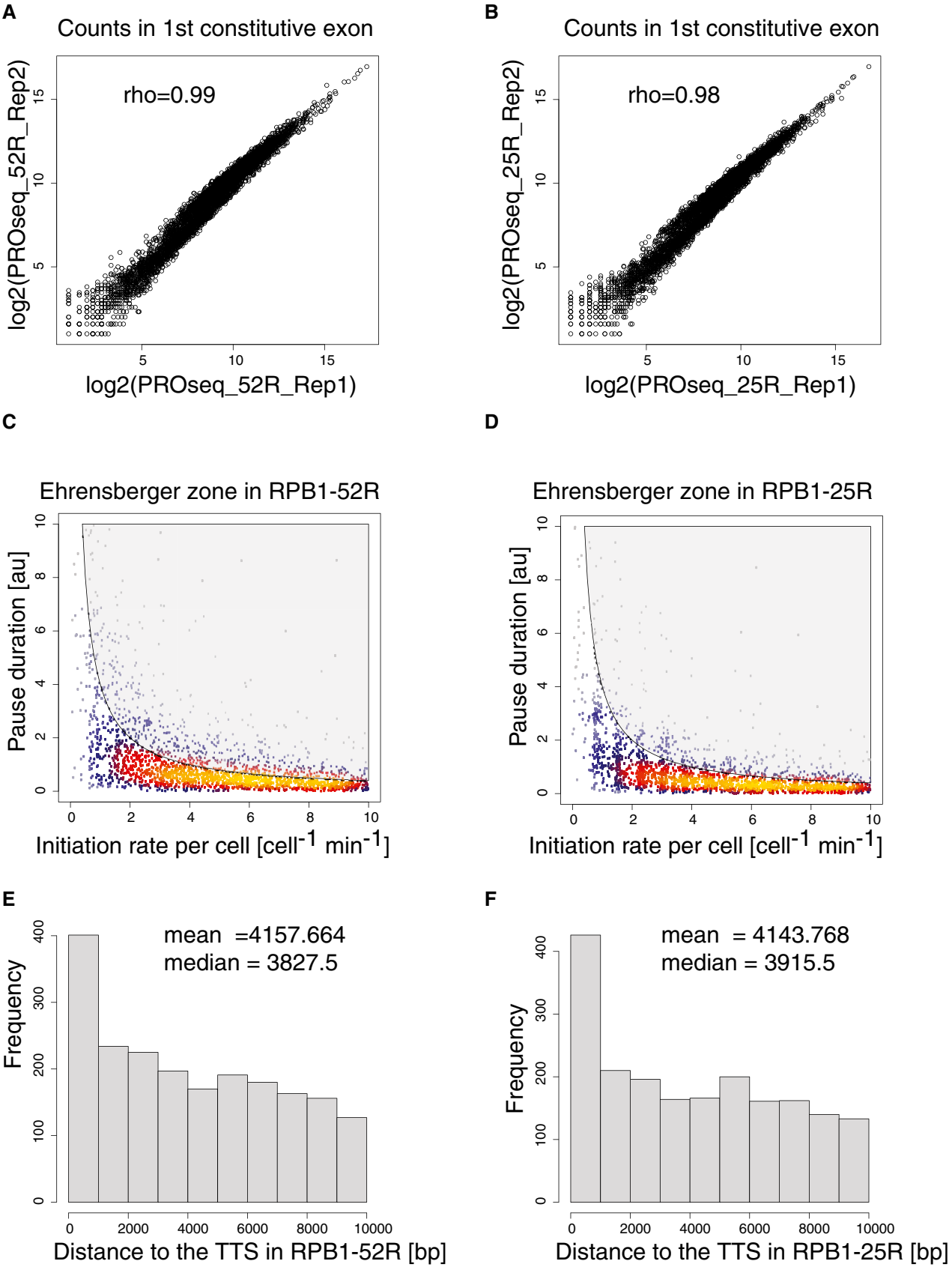
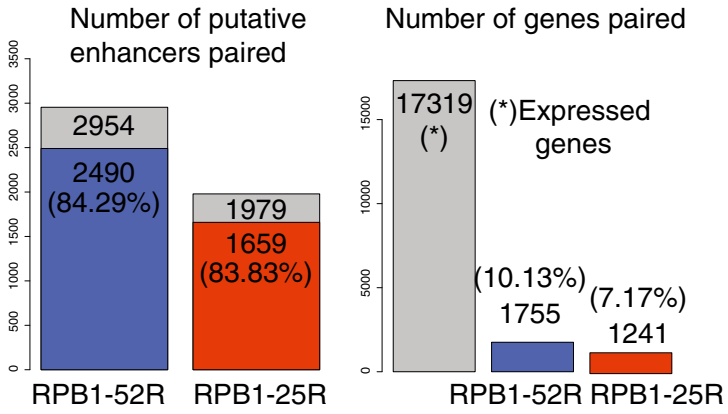


Figure EV2.

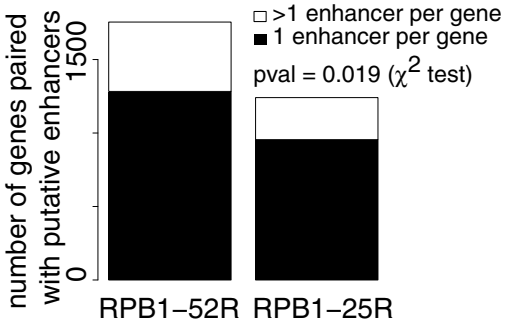
A

Rank	Motif	TF name	p value
1		FRA1	1e-258
2		FOS	1e-248
3		FRA2	1e-242
4		JUNB	1e-240
5		ATF3	1e-238
6		BATF	1e-224
7		JUN-AP1	1e-222
8		AP-1	1e-218
9		FOSL2	1e-218
10		BACH2	1e-125

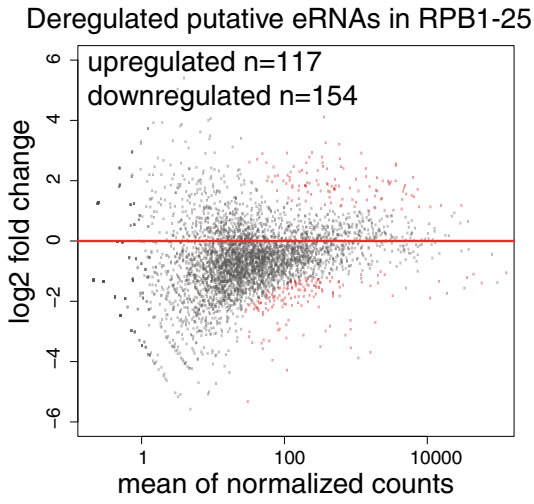
B



C



D



E

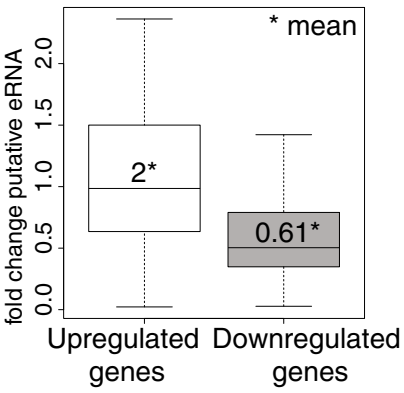


Figure EV3.

Figure EV3. CTD shortening in human cells alters transcribed putative enhancers.

A Transcription factor-binding sites enriches in the area of ± 50 bp from the TSS of putative eRNAs annotated using TT-seq data in RPB1-52R and RPB1-25R cells. *P* values were determined by the hypergeometric test.

B Bar plots showing statistics of our enhancer-promoter pairing strategy. The left bar plot shows a fraction of annotated putative enhancers that could be paired to promoters using our strategy. 2,490 putative enhancers in RPB1-52R cells (out of all 2,954 enhancers annotated in RPB1-52R cells) could be paired with promoters using our strategy (84.29% of all the putative enhancers we annotated in RPB1-52R could be paired to promoters). In RPB1-25R cells, 83.83% of annotated putative enhancers could be paired to promoters. The right panel shows a fraction of active promoters in U2OS cells that could be paired to putative enhancers using our strategy. There are 17,319 active promoters in U2OS cells (gray bar). In RPB1-52R cells, 1,755 of active promoters could be paired with putative enhancers (which constitute 10.13% of all active promoters in RPB1-52R cells). In the case of RPB1-25R cells, 1,241 of all active promoters could be paired with putative enhancers (which constitutes 7.17% of all active promoters).

C Bar plot showing a number of putative enhancers paired per gene in RPB1-52R and RPB1-25R cells. *P* value = 0.019 (χ^2 test).

D MA plot showing changes in putative eRNAs synthesis upon CTD shortening in U2OS cells in steady state conditions. RPB1-52R cells are used as a control and the data were normalized using spike-in counts. 4,335 putative eRNAs were analyzed, 117 putative eRNAs were significantly upregulated (adjusted *P* value < 0.05, $\log_2$ fold change ≥ 1, depicted in red) and 154 putative eRNAs were significantly downregulated (adjusted *P* value < 0.05, $\log_2$ fold change ≤ 1, depicted in red).

E Boxplot showing changes in putative eRNA synthesis if paired with genes that were significantly upregulated (adjusted *P* value < 0.05, $\log_2$ fold change ≥ 1) (left box) or significantly downregulated (adjusted *P* value < 0.05, $\log_2$ fold change ≤ 1) (right box) upon CTD shortening. Box limits are the first and third quartiles, the band inside the box is the median. The ends of the whiskers extend the box by 1.5 times the interquartile range. Two independent biological replicates were analyzed.

Figure EV4. CTD shortening impairs transcription induction.

A Schematic representation of the experimental procedure used to study the effects of CTD shortening on TPA-induced transcription.

B, C Spearman correlation of read counts in RefSeq genes of biological replicates in TT-seq (B) and RNA-seq (C) in RPB1-52R and RPB1-25R cells upon 15 and 30 min of TPA treatment (200 nM) and the respective DMSO controls.

D Bar plot showing number of RefSeq genes that change RNA synthesis in TT-seq datasets in RPB1-52R and RPB1-25R upon 15- and 30-min treatment with TPA (200 nM). Genes showing adjusted *P* value < 0.05 and $\log_2$ fold change ≥ 1 were considered as significantly upregulated and genes showing adjusted *P* value < 0.05 and $\log_2$ fold change ≤ 1 were considered as significantly downregulated.

E Exemplary genome browser view showing TT-seq read coverage over *FOS* gene in RPB1-52R and RPB1-25R cells upon 15 min of TPA treatment (200 nM).

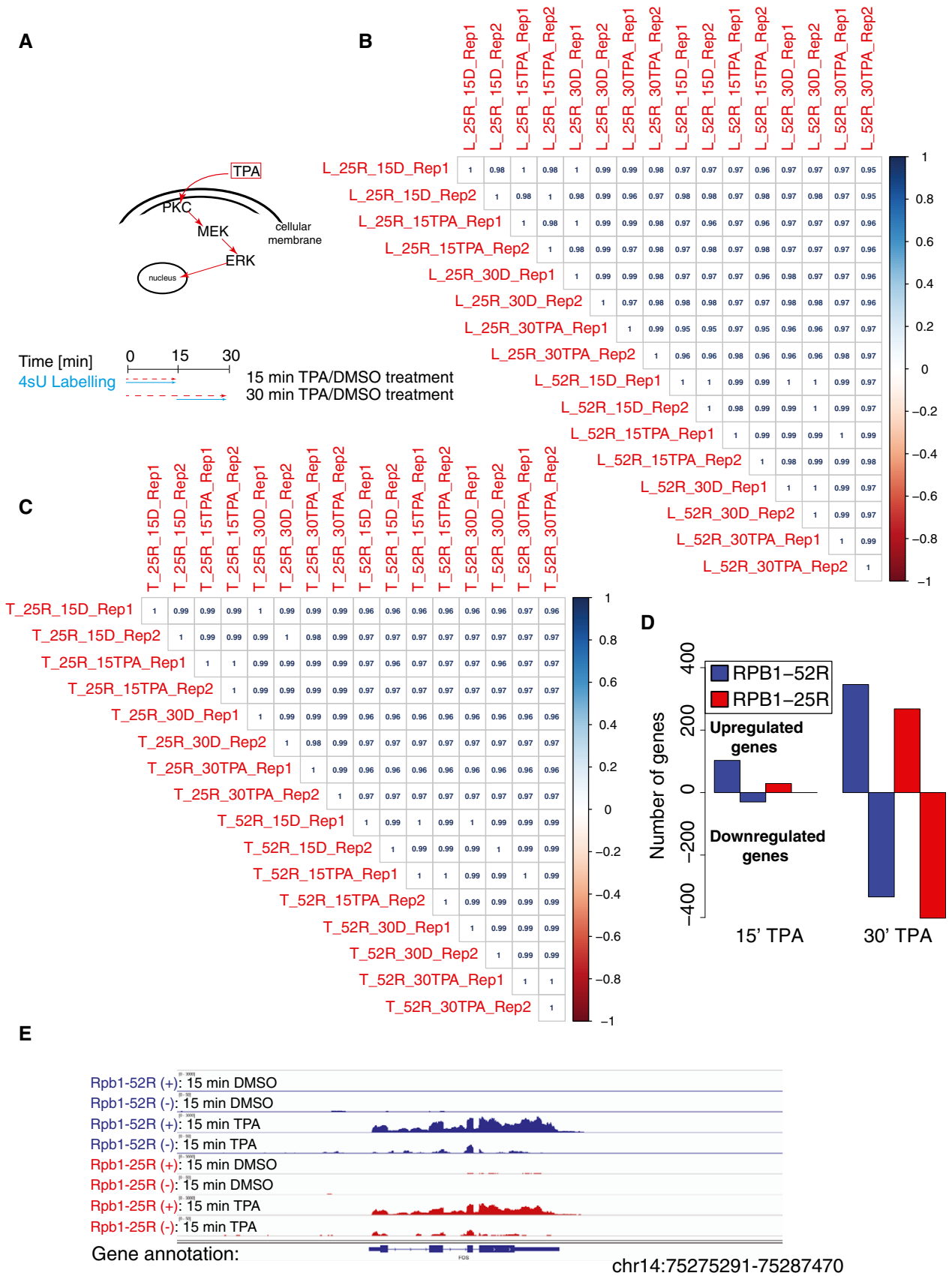


Figure EV4.

Figure EV5. CTD shortening delays transcription induction and changes transcription of putative enhancers.

A Barplot showing number of annotated putative eRNAs upon 200 nM TPA or DMSO (solvent control) treatment in RPB1-52R and RPB1-25R cells.

B Bar plot showing number of differentially expressed putative eRNAs in TT-seq datasets in RPB1-52R and RPB1-25R upon 15- and 30-min treatment with TPA (200 nM). Putative eRNAs showing adjusted *P* value < 0.05 and $\log_2$ fold change ≥ 1 were considered as significantly upregulated and putative eRNAs showing adjusted *P* value < 0.05 and $\log_2$ fold change ≤ 1 were considered as significantly downregulated.

C MA plot showing changes in putative eRNAs synthesis in RPB1-52R and RPB1-25R cells upon 15 min or 30 min of TPA treatment. DMSO samples for respective cell lines were used as controls, and the data were normalized using library size normalization. Putative eRNAs showing adjusted *P* value < 0.05, $\log_2$ fold change ≥ 1, depicted in red) were considered significantly upregulated and putative eRNAs showing adjusted *P* value < 0.05, $\log_2$ fold change ≤ 1, depicted in red) were considered significantly downregulated.

D Exemplary genome browser view showing TT-seq read coverage over *GADD45* gene as well a region giving rise to a putative eRNA mapped to it in RPB1-52R and RPB1-25R cells upon 15 min of TPA treatment (200 nM).

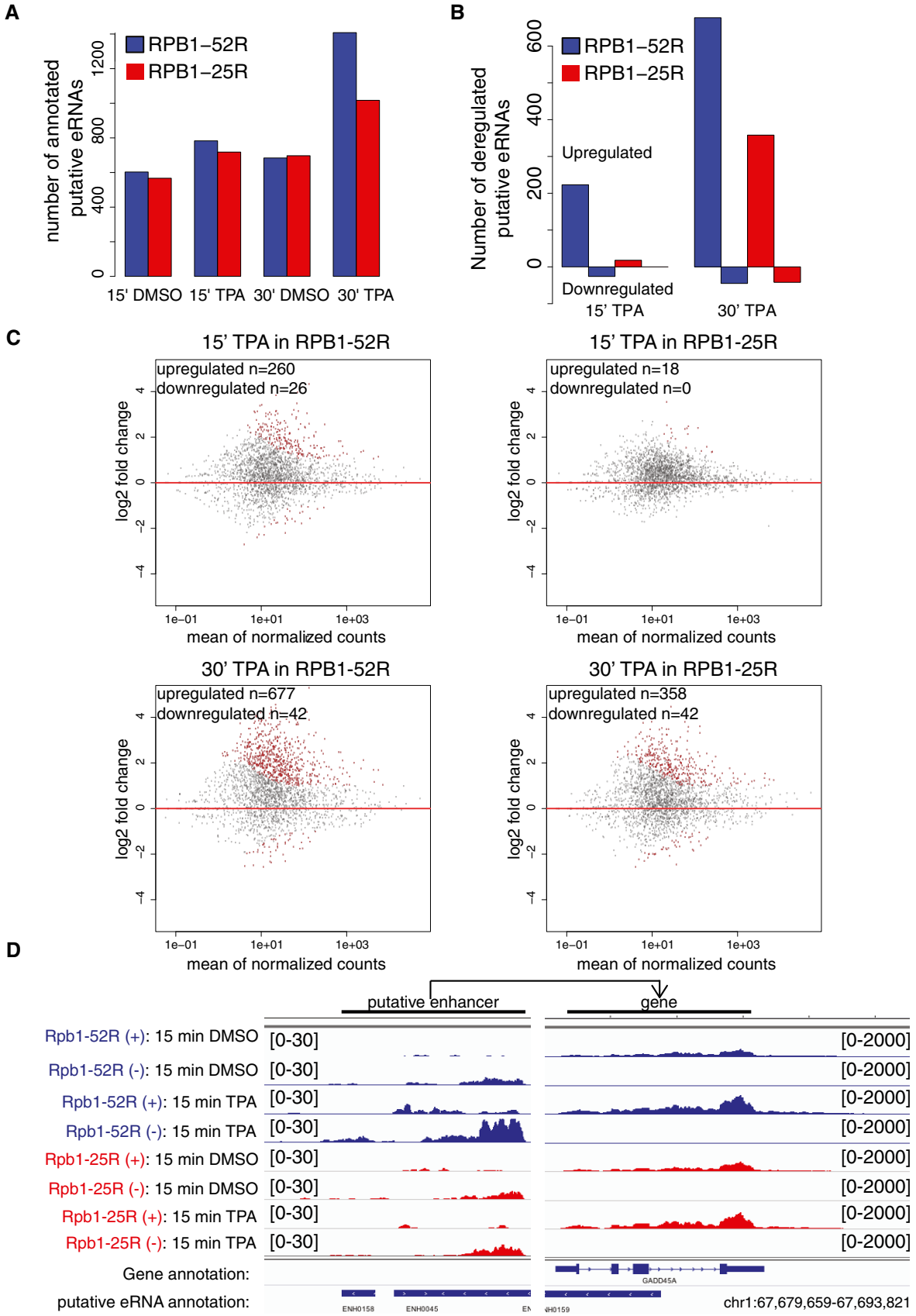


Figure EV5.