Kinetic measurement of gene-specific RNAPII transcription elongation rates

Haiyue Liu[1] and Lea H. Gregersen[1,*,#]

[1]Center for Gene Expression, Department of Cellular and Molecular Medicine, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark.

* Correspondence and # lead contact: leag@sund.ku.dk

Supplemental materials

Supplemental Fig S1. Quality control, background reads and wave front model.

Supplemental Fig S2. DRB/TT$_{chem}$-seq2 sequencing statistics.

Supplemental Fig S3. Examples of read coverage profiles and linear regression analyses.

Supplemental Fig S4. Downsampling analyses of DRB/TT$_{chem}$-seq2 dataset.

Supplemental Fig S5. Comparing gene features and histone modifications to RNAPII elongation rates in lncRNAs.

Supplemental Fig S6. Genes of different lengths exhibiting similar RNAPII elongation rates.

Supplemental Fig S7. Characteristics of RNAPII elongation rates.

Supplemental Fig S8. Pairwise comparisons of RNAPII elongation rates.

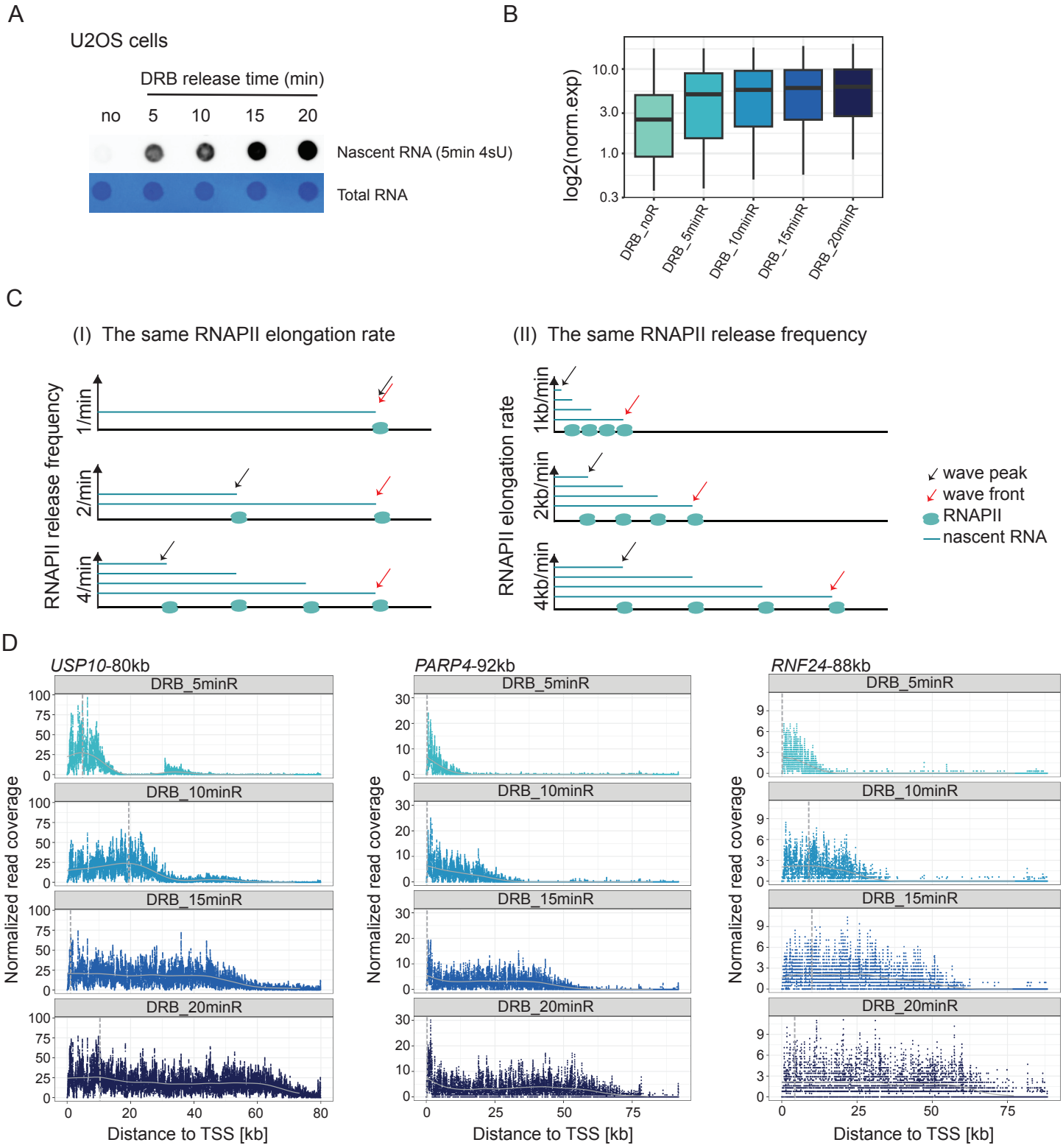
Supplemental Fig S9. Enrichment of histone modifications and elongation factors in genes with different elongation rates.

Supplemental Fig S10. Enrichment profiles of H3K36me3 for genes that have been estimated with elongation rates across studies.

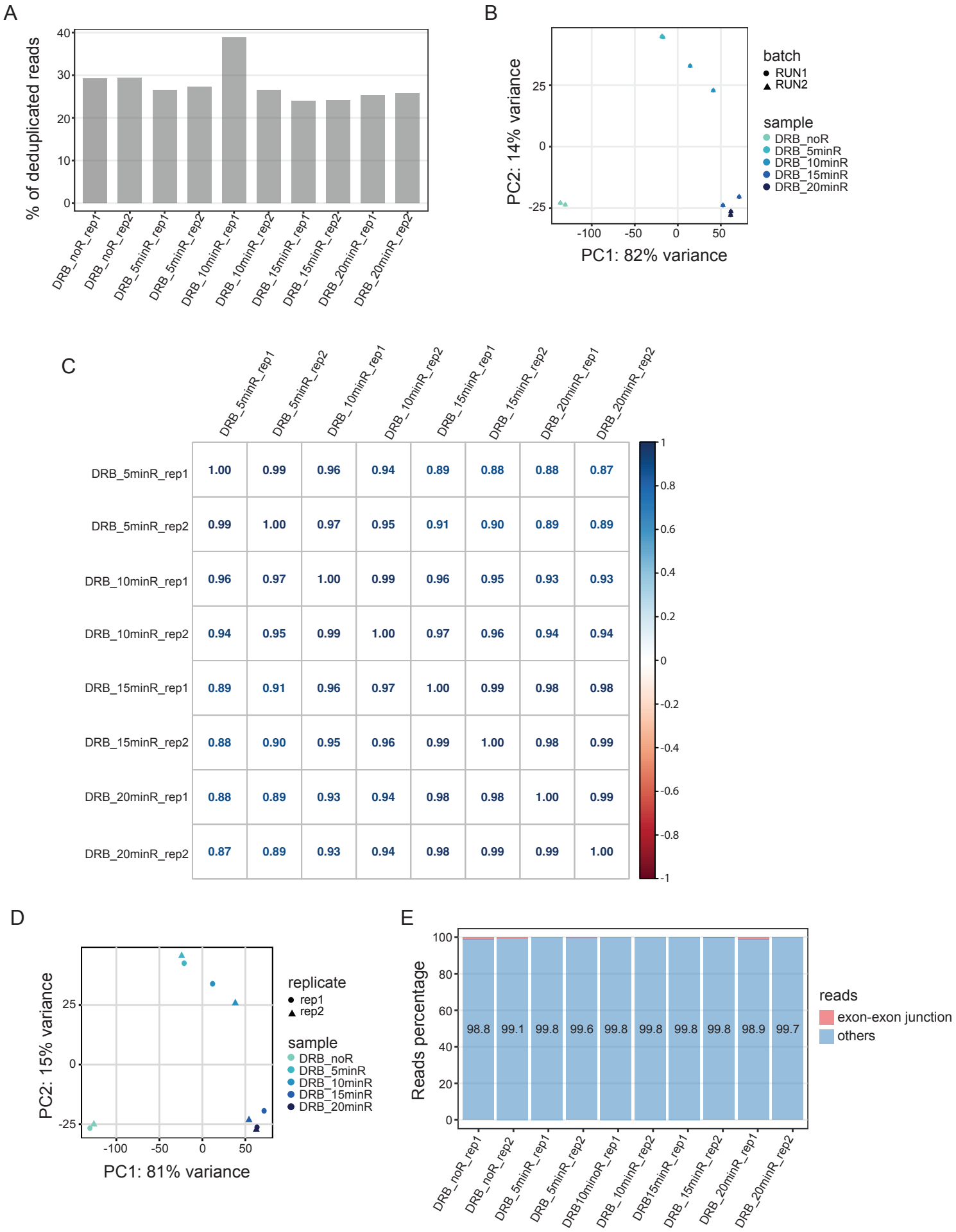
Supplemental Table S1. RNAPII speed estimates from metabolic labeling experiments.

Supplemental Table S2. ENCODE ChIP-seq datasets.

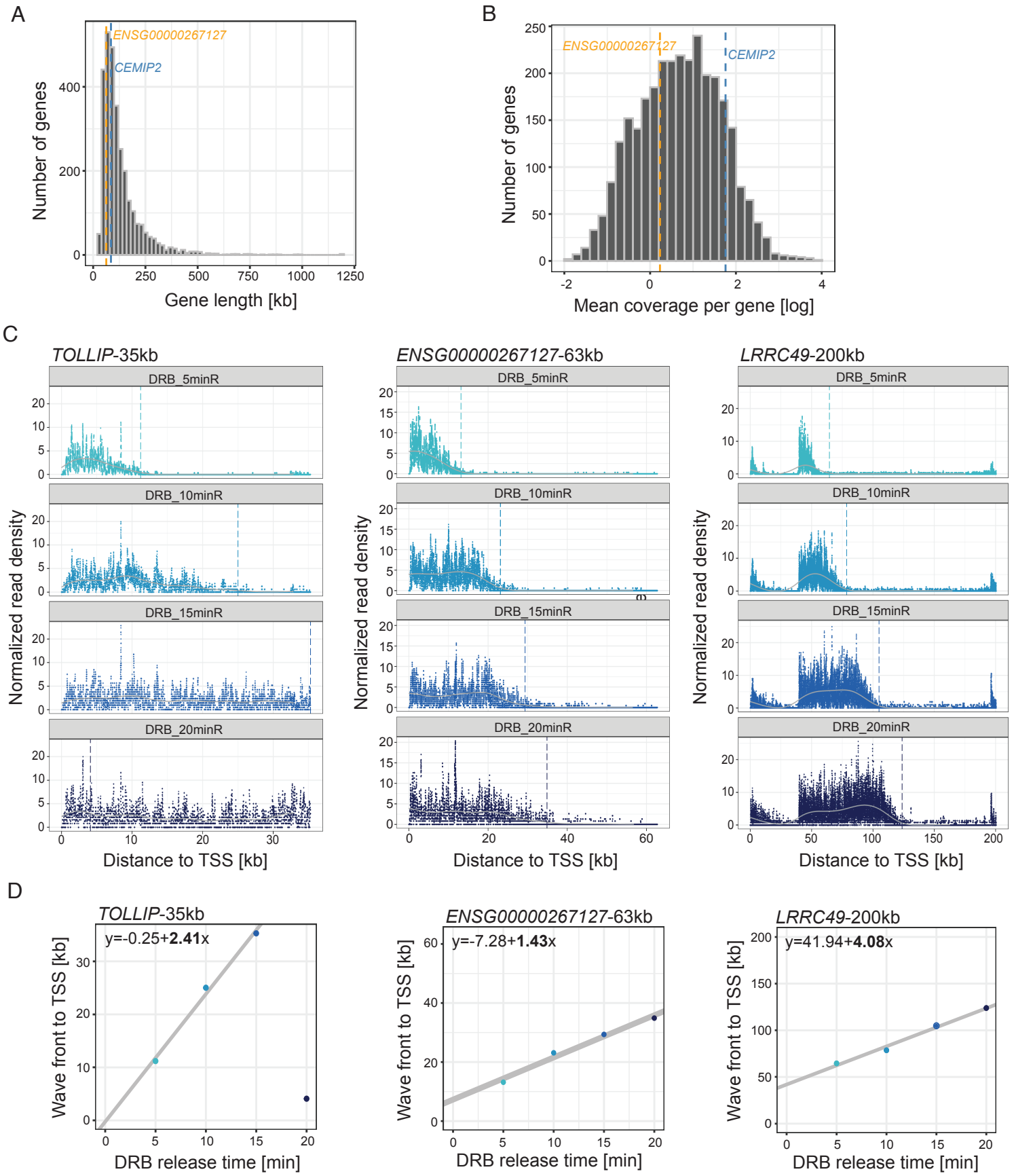
Supplemental references



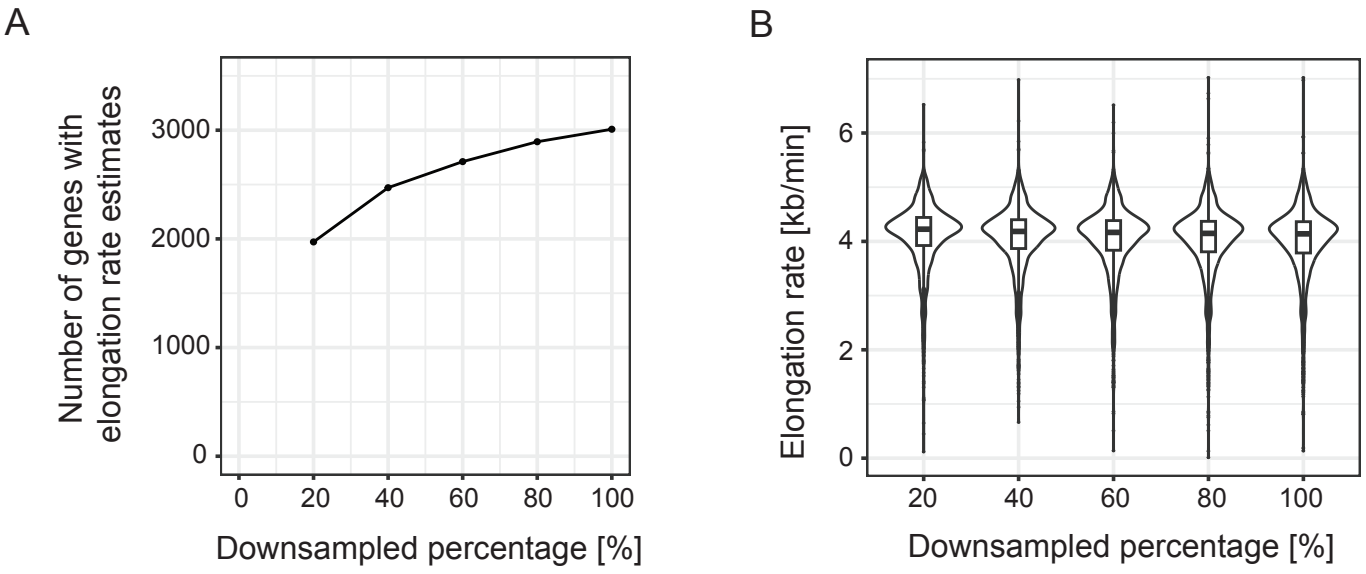
Supplemental Fig S1. Quality control, background reads and wave front model. A. Dot blot analysis of 4sU incorporation into RNA in U2OS cells. Cells were treated with 100 μM DRB for 3.5 hrs. DRB release samples were collected after 5, 10, 15 and 20 minutes post release. For all samples, RNA was labeled with 1 mM 4sU for 5 minutes immediately prior to harvest. Methylene blue staining was used as a loading control. B. Transcript levels detected in the DRB/TT$_{chem}$-seq2 experiment. C. Schematic illustration highlighting the differences between wave fronts and wave peaks under the assumption of constant RNAPII release frequency and elongation rate. (I) With the same elongation rates but varying release frequencies, RNAPII reaches the same position within gene bodies, differing only in the number of released polymerases. This model results in different wave peaks but consistent wave fronts. (II) With the same RNAPII release frequency but varying elongation rates, RNAPII travels different distances within gene bodies, producing distinct wave fronts by the different elongation rates. The number of released polymerases is the same. D. Example coverage profiles of genes where wave peak positions do not consistently progress over time in the DRB/TT$_{chem}$-seq2 experiment. Grey dashed lines indicate the wave peak positions.



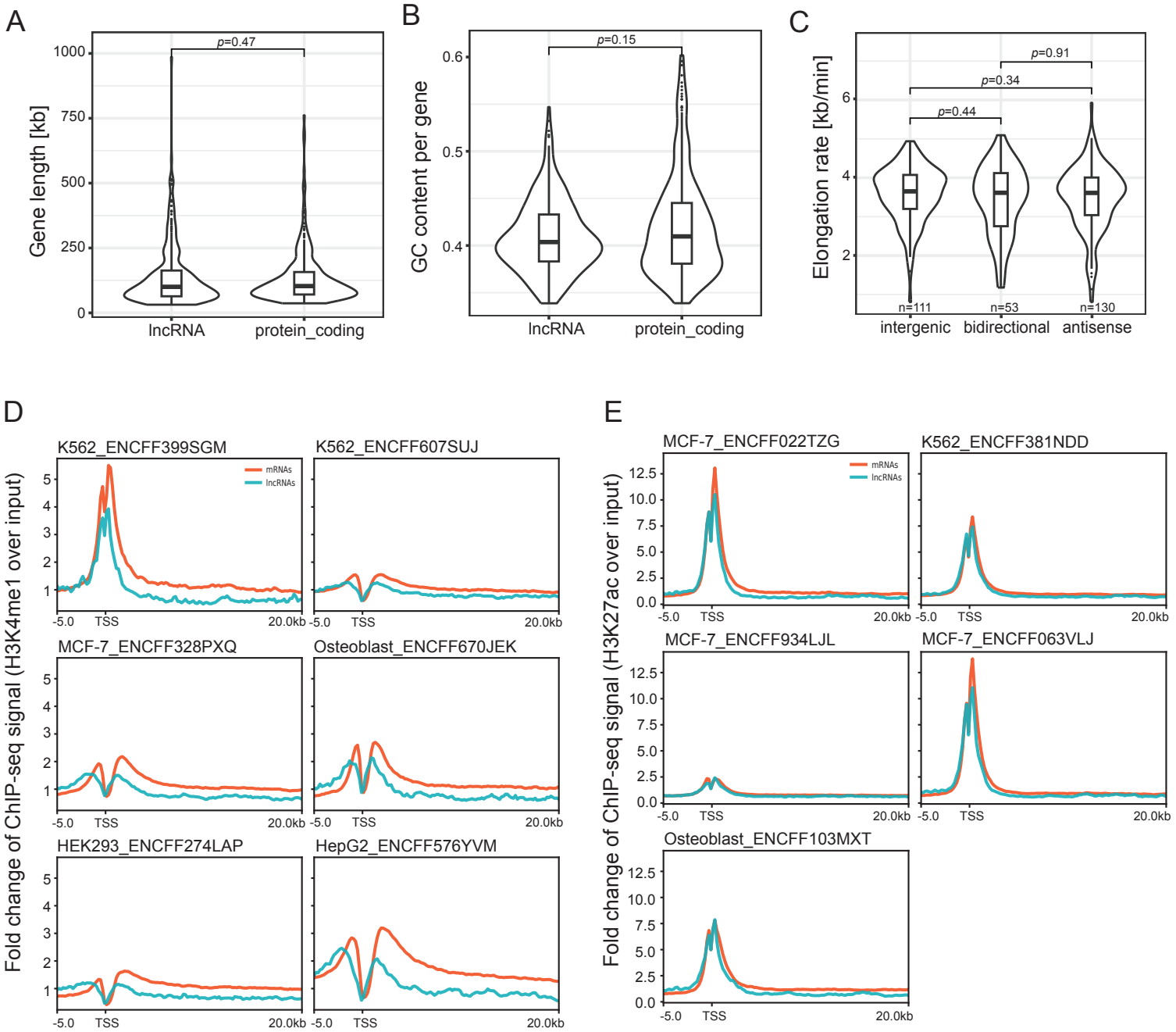
Supplemental Fig S2. DRB/TT$_{chem}$-seq2 sequencing statistics. A. Bar plots showing the percentage of reads identified as PCR duplicates. B. PCA plot illustrating clustering of technical replicates from two independent DRB/TT$_{chem}$-seq2 sequencing runs. C. Pairwise Pearson correlations between samples and replicates. The values displayed in the plot are the Pearson correlation coefficients comparing the common set of expressed gene (n=27,063) in log(TPM+1) scale. D. PCA plot showing clustering of biological replicates in DRB/TT$_{chem}$-seq2 data.



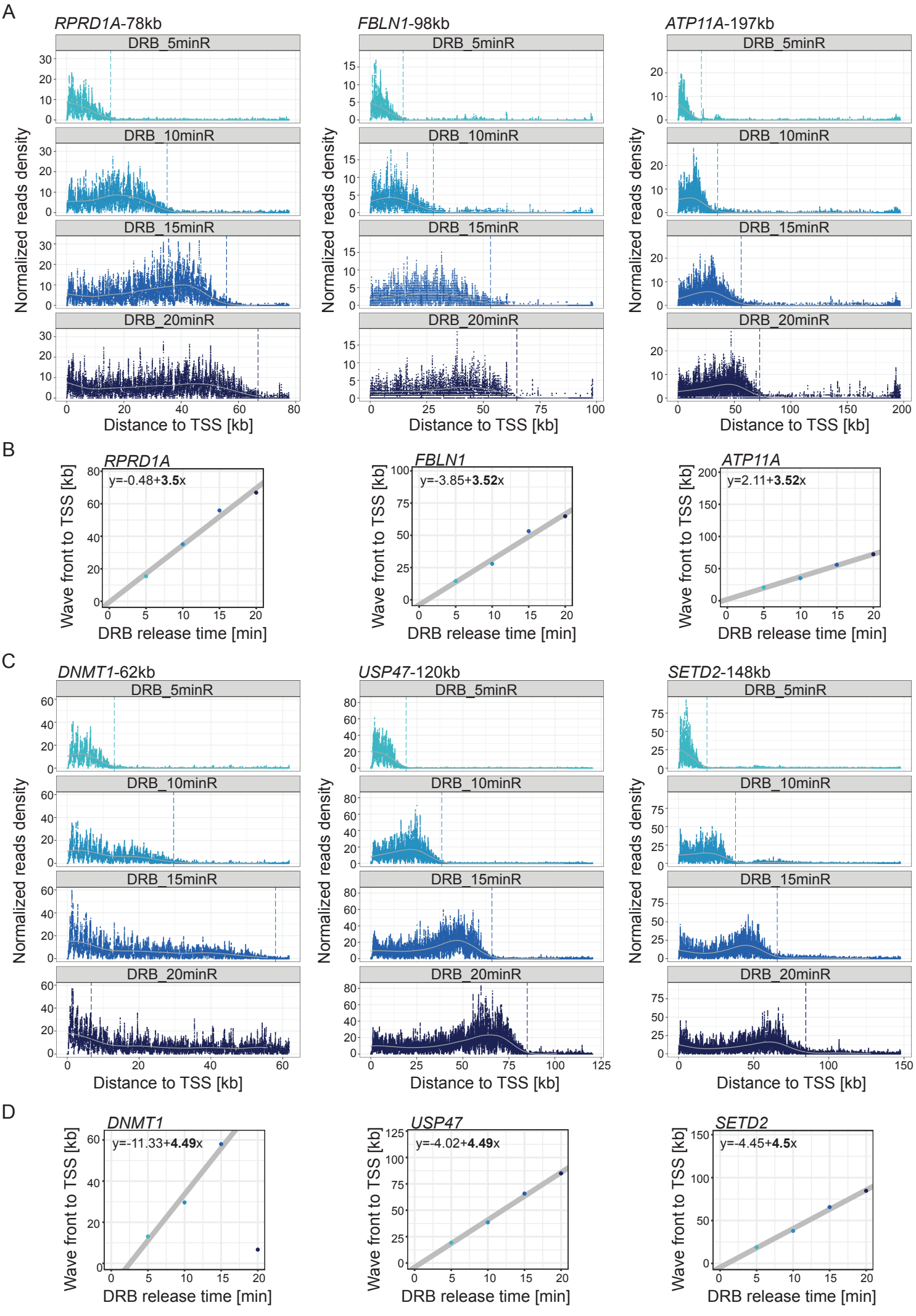
Supplemental Fig S3. Examples of read coverage profiles and linear regression analyses. A. Gene length distributions of the protein-coding and lncRNA genes included in the analysis. The lengths of the two example genes shown in Figure 2B and S3C are highlighted. B. Mean coverage distributions of the analyzed genes. Coverage levels of the two example genes from Figure 2B and S3C are highlighted. C. Example read coverage profiles for three genes of varying lengths: short (35 kb), medium (63 kb), long (200 kb). Dashed vertical lines indicate the identified wave front positions. D. Linear regression of RNAPII traveled distance as a function of time for the example genes shown in panel C. Note that only the first three DRB release times were used for fitting when the fourth point does not advance with time.



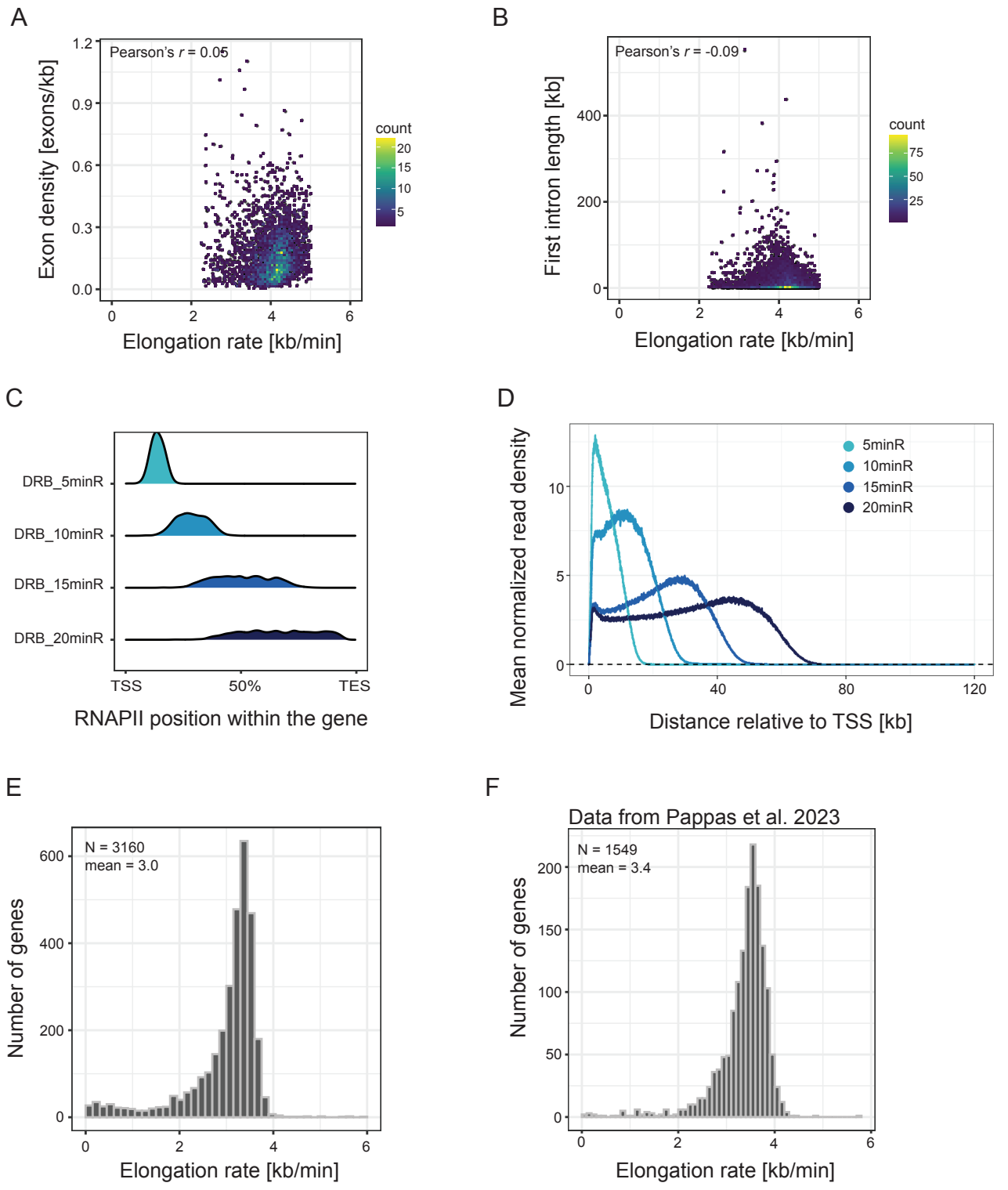
Supplemental Fig S4. Downsampling analyses of DRB/TT$_{chem}$-seq2 dataset. A. Point plots with line connections showing the number of genes with RNAPII elongation rates estimates when downsampling the sequencing reads to certain fractions. B. Violin plots displaying the values and distributions of the RNAPII elongation rates estimated for those genes shown in A at each downsampled fraction.



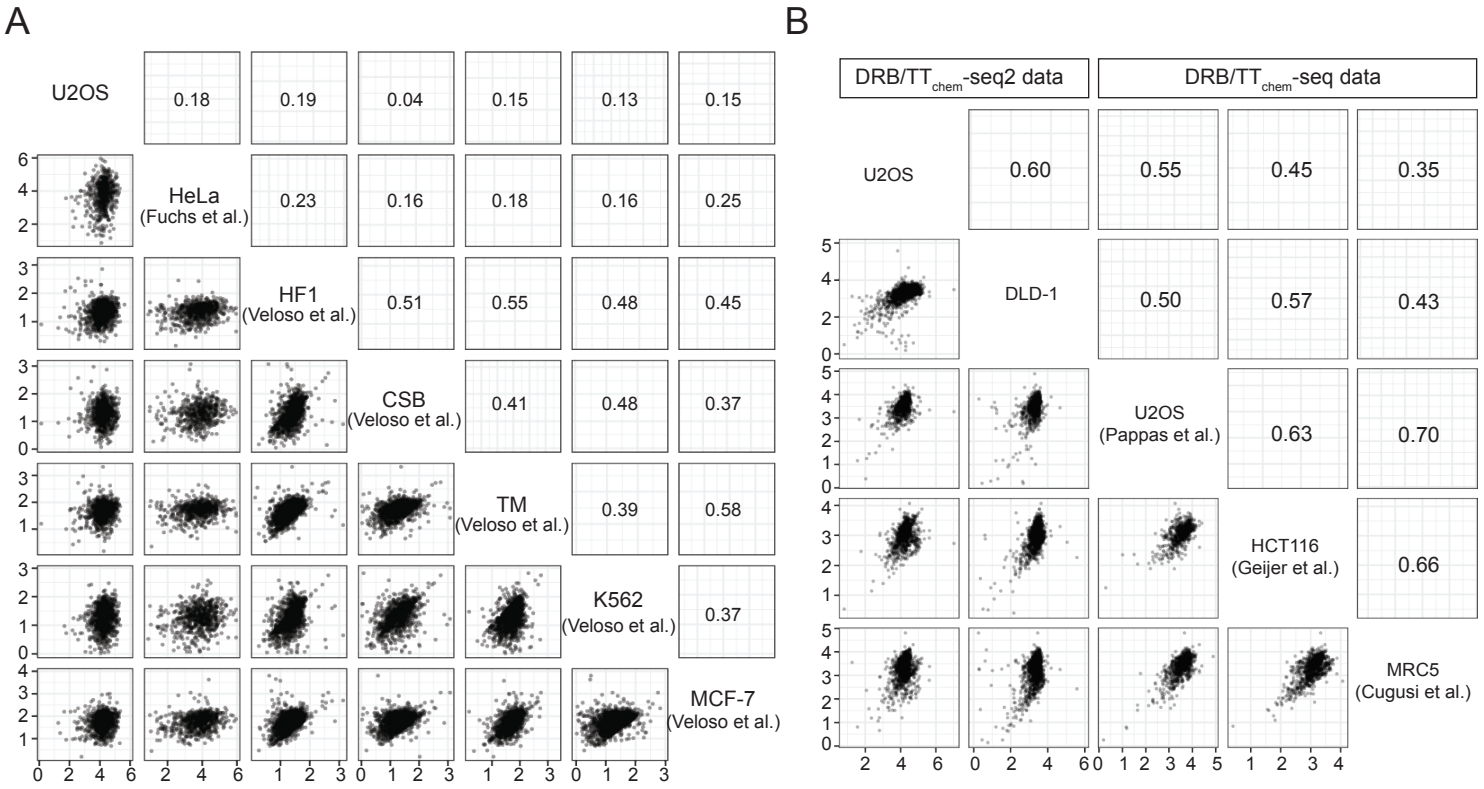
Supplemental Fig S5. Comparing gene features and histone modifications to RNAPII elongation rates in lncRNAs. A. Violin plots comparing the distributions of gene lengths of lncRNA genes and mRNA genes that have RNAPII elongation rate estimates. The number of protein coding genes were sampled to match the number of lncRNA genes (i.e., n=300 for both groups). B. Comparison of GC content distributions in lncRNAs and mRNAs with RNAPII elongation rate estimates (n=300 for both groups). C. Comparing the elongation rate distributions among the three subgroups of lncRNA genes classified based on their locations relative to mRNAs. D. Comparison of H3K4me1 enrichment profiles over gene bodies between lncRNAs and mRNAs with RNAPII elongation rate estimates. The histone ChIP-seq data was derived from ENCODE. E. Same as D but comparing H3K27ac profiles.


A
RPRD1A-78kb
FBLN1-98kb
ATP11A-197kb
DRB_5minR
DRB_10minR
DRB_15minR
DRB_20minR
Normalized reads density
Distance to TSS [kb]
B
RPRD1A
y=-0.48+3.5x
FBLN1
y=-3.85+3.52x
ATP11A
y=2.11+3.52x
Wave front to TSS [kb]
DRB release time [min]
C
DNMT1-62kb
USP47-120kb
SETD2-148kb
D
DNMT1
y=-11.33+4.49x
USP47
y=-4.02+4.49x
SETD2
y=-4.45+4.5x

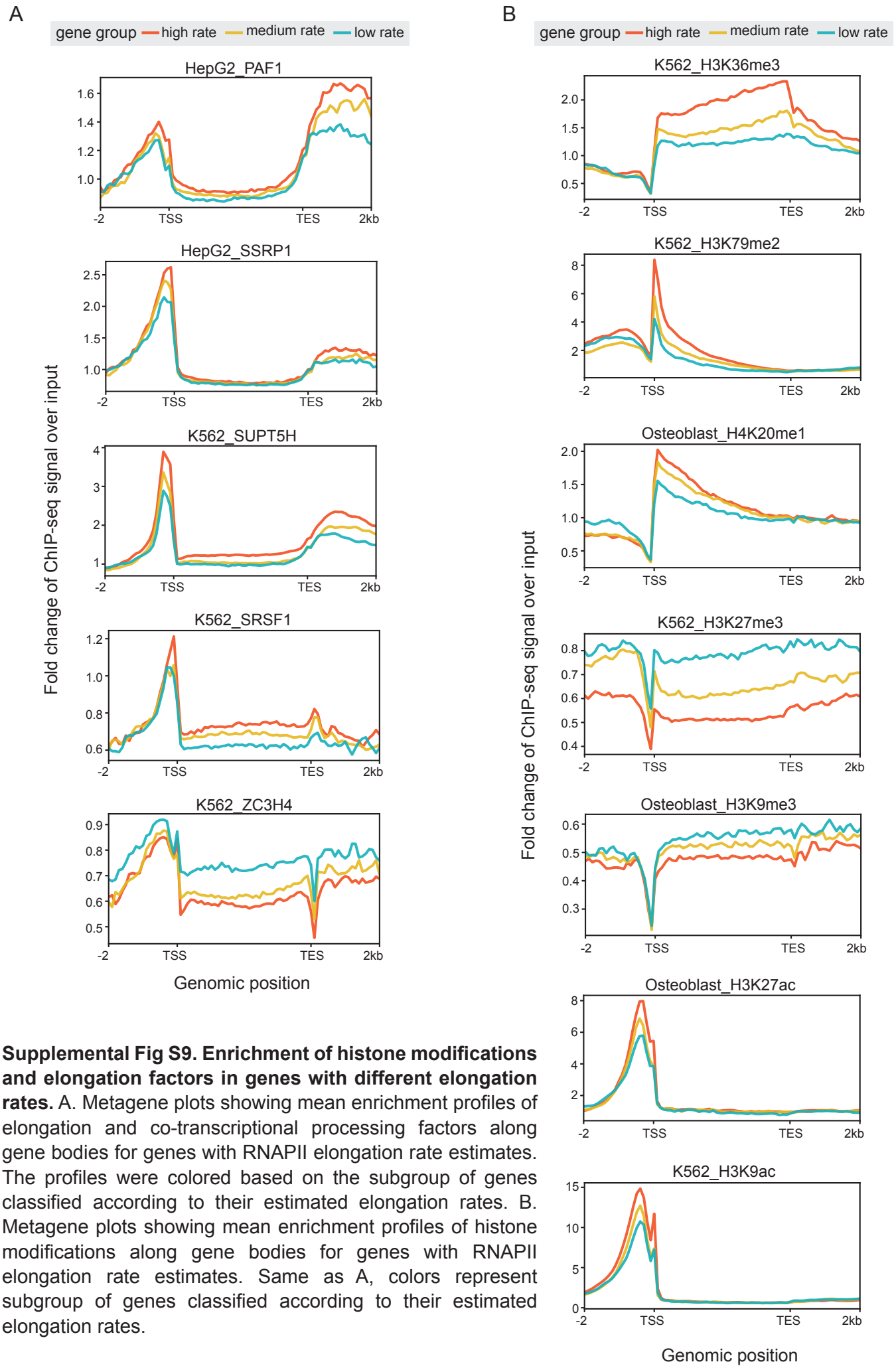
Supplemental Fig S6. Genes of different lengths that exhibit similar RNAPII elongation rates. A. Read coverage profiles of *RPRD1A*, *FBLN1* and *ATP11*, which show similar elongation rates despite differing gene lengths. Dashed vertical lines show the identified wave front. B. Linear regression profiles corresponding to the genes shown in panel A. The slope of each regression line represents the RNAPII elongation rate. C. Read coverage profiles of *DNMT1*, *USP47*, and *SETD2*, which also have similar elongation rates but differ in lengths. D. Linear regression profiles for the genes shown in panel C.



Supplemental Fig S7. Characteristics of RNAPII elongation rates. A. Scatter plot depicting elongation rates against exon densities for genes with RNAPII elongation rate estimates. Only genes containing multiple exons were retained in the plot. B. Scatter plot showing the first intron length of each gene against RNAPII elongation rates for genes filtered with linear regression adjusted *p*-value ≥0.95 intron number ≥1. C. Density plots showing the distribution of RNAPII relative positions within gene bodies at the end of each DRB release times. D. Metagene plot showing the averaged read density for each DRB release time points in DLD-1 data. The read density was averaged across the genes with RNAPII elongation rate estimates (N=3,168). E. Histogram of the RNAPII elongation rate estimates in the DLD-1 dataset. Elongation rate values larger than 6 kb/min (8 genes) were excluded from the plot. F. Distribution of RNAPII elongation rates estimated using computational method implemented in DRB/TT$_{chem}$-seq2 on DRB/TT$_{chem}$-seq data from Pappas et al., 2023.

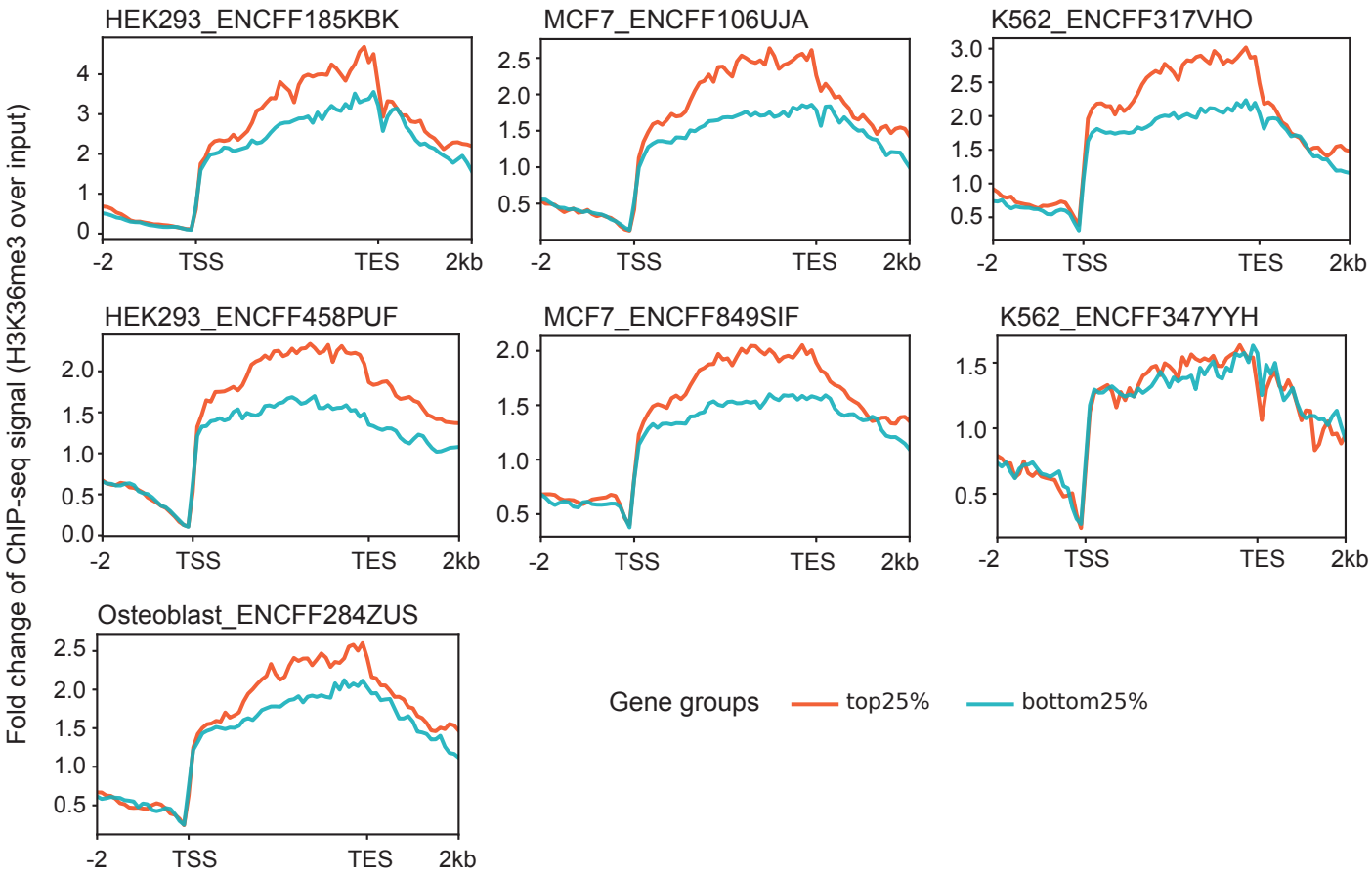


Supplemental Fig S8. Pairwise comparisons of RNAPII elongation rates. A. Pairwise comparisons of RNAPII elongation rates from our study and those reported previously using similar methods. Lower left panel shows scatter plots of the pairwise comparisons. Upper right panel shows Pearson correlation coefficients among the comparisons. B. Pairwise comparisons of RNAPII elongation rates among two cell lines from our study and those estimates from published DRB/TT_{chem}-seq datasets using the updated computational pipeline. Lower left panel shows scatter plots of the pairwise comparisons. Upper right panel shows Pearson correlation coefficients among the comparisons.



Supplemental Fig S9. Enrichment of histone modifications and elongation factors in genes with different elongation rates. A. Metagene plots showing mean enrichment profiles of elongation and co-transcriptional processing factors along gene bodies for genes with RNAPII elongation rate estimates. The profiles were colored based on the subgroup of genes classified according to their estimated elongation rates. B. Metagene plots showing mean enrichment profiles of histone modifications along gene bodies for genes with RNAPII elongation rate estimates. Same as A, colors represent subgroup of genes classified according to their estimated elongation rates.

A



Supplemental Fig S10. Enrichment profiles of H3K36me3 for genes that have been estimated with elongation rates across studies. A. Enrichment profiles of H3K36me3 for genes with RNAPII elongation rate estimates reported in Fuchs et al., and our U2OS cells. Only genes from the top 25% highest and bottom 25% lowest elongation rates were included in the analysis. Colors represent the two subgroups.

Supplemental Table S1. RNAPII speed estimates from metabolic labeling experiments.

RNAPII speed [kb/min]	Experimental setup	Detection of RNAPII progression	Cell line	# of genes characterized with elongation rates	Reference
2.1 2.8	GRO-seq	Wave end (i.e., wave front)	MCF-7 AC16	140 26	(Danko *et al*, 2013)
2.0	GRO-seq	Wave progression	mESC	1333	(Jonkers *et al*, 2014)
1.25 1.75 1.25 1.25 1.75	BruDRB-seq	Wave end (i.e., wave front)	HF1 TM CS-B K562 MCF-7	2,702 2,469 1,932 2,270 2,399	(Veloso *et al*, 2014)
2.3	DRB/TTchem-seq	Wave peak	HEK293	378	(Gregersen *et al*, 2020)
3.5	4sUDRB-seq	Wave front	HeLa	1,577	(Fuchs *et al*, 2014)
3.2	4sU-FP-seq	Wave front	HEK293	982	(Liang *et al*, 2018)
3.13	DRB/GRO-seq	Wave front	HEK293	273	(Saponaro *et al*, 2014)
3.27	4sUDRB-seq	Wave front	U2OS	2,163	(Baluapuri *et al*, 2019)

Adapted from (Muniz *et al*, 2021)

Supplemental Table S2. ENCODE ChIP-seq datasets

ChIP target	bigwig accession	Cell line
H3K36me3	ENCFF347YYH	K562
	ENCFF317VHO	K562
	ENCFF185KBK	HEK293
	ENCFF458PUF	HEK293
	ENCFF849SIF	MCF-7
	ENCFF106UJA	MCF-7
	ENCFF284ZUS	Osteoblast
H3K79me2	ENCFF544AVW	K562
	ENCFF105VBM	MCF-7
	ENCFF656MWZ	Osteoblast
H4K20me1	ENCFF605FAF	K562
	ENCFF382HTX	MCF-7
	ENCFF967MBK	Osteoblast
	ENCFF330AIV	HepG2
H3K27me3	ENCFF139KZL	K562
	ENCFF242ENK	K562
	ENCFF725VID	MCF-7
	ENCFF210IBW	MCF-7
	ENCFF659QTD	Osteoblast
H3K9me3	ENCFF631FTI	MCF-7
	ENCFF085EUT	MCF-7
	ENCFF430KTS	Osteoblast
	ENCFF526FQB	HEK293

	ENCFF754ROM	HepG2
H3K27ac	ENCFF022TZG	HepG2
	ENCFF381NDD	K562
	ENCFF934LJL	MCF-7
	ENCFF063VLJ	MCF-7
	ENCFF103MXT	Osteoblast
H3K9ac	ENCFF286WRJ	K562
	ENCFF583BKU	K562
	ENCFF862AJX	MCF-7
	ENCFF040RHK	HepG2
PAF1	ENCFF825ZDL	HepG2
SSRP1	ENCFF517CGF	K562
SUPT5H	ENCFF498YTJ	K562
SRSF1	ENCFF704FKY	K562
	ENCFF498YTJ	K562
ZC3H4	ENCFF024ZVX	K562
	ENCFF597RGZ	HepG2
POLR2A	ENCFF124WLE	K562
	ENCFF806LCJ	K562
	ENCFF937ZPS	K562
	ENCFF496FVA	K562
	ENCFF348QLX	K562
	ENCFF444MEV	K562
	ENCFF018UUW	MCF-7
	ENCFF473DMJ	HepG2

	ENCFF323HAZ ENCFF761IJZ	HepG2 HepG2
POLR2AphosphoS5	ENCFF907FLD ENCFF801BXZ	K562 HepG2
POLR2AphosphoS2	ENCFF711PAO ENCFF434PYZ	K562 K562
POLR2B	ENCFF331WPU	K562

Supplemental references:

Baluapuri A, Hofstetter J, Dudvarski Stankovic N, Endres T, Bhandare P, Vos SM, Adhikari B, Schwarz JD, Narain A, Vogt M, *et al* (2019) MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation. *Mol Cell* 74: 674-687.e11

Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A & Kraus WL (2013) Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. *Mol Cell* 50: 212–222

Fuchs G, Voichek Y, Benjamin S, Gilad S, Amit I & Oren M (2014) 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol* 15: R69

Gregersen LH, Mitter R & Svejstrup JQ (2020) Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. *Nat Protoc* 15: 604–627

Jonkers I, Kwak H & Lis JT (2014) Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* 3: e02407

Liang K, Smith ER, Aoi Y, Stoltz KL, Katagi H, Woodfin AR, Rendleman EJ, Marshall SA, Murray DC, Wang L, *et al* (2018) Targeting Processive Transcription Elongation via SEC Disruption for MYC-Induced Cancer Therapy. *Cell* 175: 766-779.e17

Muniz L, Nicolas E & Trouche D (2021) RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *EMBO J* 40

Saponaro M, Kantidakis T, Mitter R, Kelly GP, Heron M, Williams H, Söding J, Stewart A & Svejstrup JQ (2014) RECQL5 Controls Transcript Elongation and Suppresses Genome Instability Associated with Transcription Stress. *Cell* 157: 1037–1049

Veloso A, Kirkconnell KS, Magnuson B, Biewen B, Paulsen MT, Wilson TE & Ljungman M (2014) Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res* 24: 896–905