Cell Reports, Volume *32*

Supplemental Information

Binding of the Treslin-MTBP Complex to Specific Regions of the Human Genome Promotes the Initiation of DNA Replication

Akiko Kumagai and William G. Dunphy

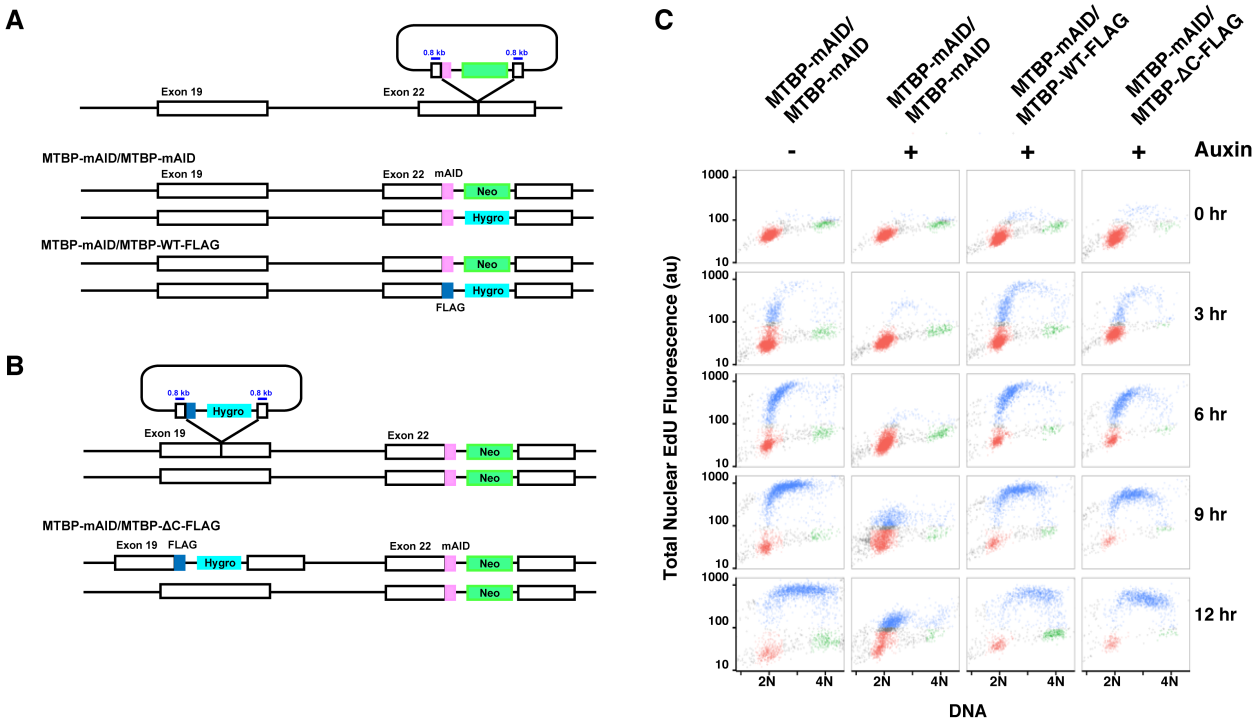

Figure S1. Construction and Characterization of Auxin-Regulated Cell Lines, Related to Figure 1

(A) For preparation of MTBP-mAID/MTBP-mAID and MTBP-mAID/MTBP-WT-FLAG cell lines, DLD-1 cells were transfected with pX330-Cas9-MTBP-Exon22 and combinations of either pBluescript-MTBP-mAID-Neo and pBluescript-MTBP-mAID-Hygro or pBluescript-MTBP-mAID-Neo and pBluescript-MTBP-WT-FLAG-Hygro, respectively. Clones resistant to G418 and Hygromycin were selected for genotype testing.

(B) For preparation of MTBP-mAID/MTBP-ΔC-FLAG cells, DLD-1 cells were transfected with pX330-Cas9-MTBP-Exon22 and pBluescript-MTBP-mAID-Neo. G418-resistant clones were selected for insertion of mAID at both alleles of MTBP. Cells with mAID at both alleles were transfected with pX330-Cas9-MTBP-Exon19 and pBluescript-MTBP-ΔC-FLAG-Hygro. Clones resistant to both G418 and Hygromycin were screened for the presence of one copy each of MTBP-mAID and MTBP-ΔC-FLAG.

(C) MTBP-mAID/MTBP-mAID, MTBP-mAID/MTBP-WT-FLAG, and MTBP-mAID/MTBP-ΔC-FLAG cells were arrested with Cdk4/6i for 20 hr. At this point, cells were maintained in Cdk4/6i (t=0) or washed to remove Cdk4/6i and thereafter incubated in the absence (column 1) or the presence of auxin (columns 2-4). Cells were incubated for 20 min with EdU at the indicated times. Incorporated EdU was labeled with Alexa 488 azide and DNA was stained with DAPI. Total nuclear EdU fluorescence (au) and nuclear DNA content were determined with CellProfiler. Red, G1; Blue, S-phase; and Green, G2/M. At least 1,500 nuclei were examined for each time point. Representative of two experiments.

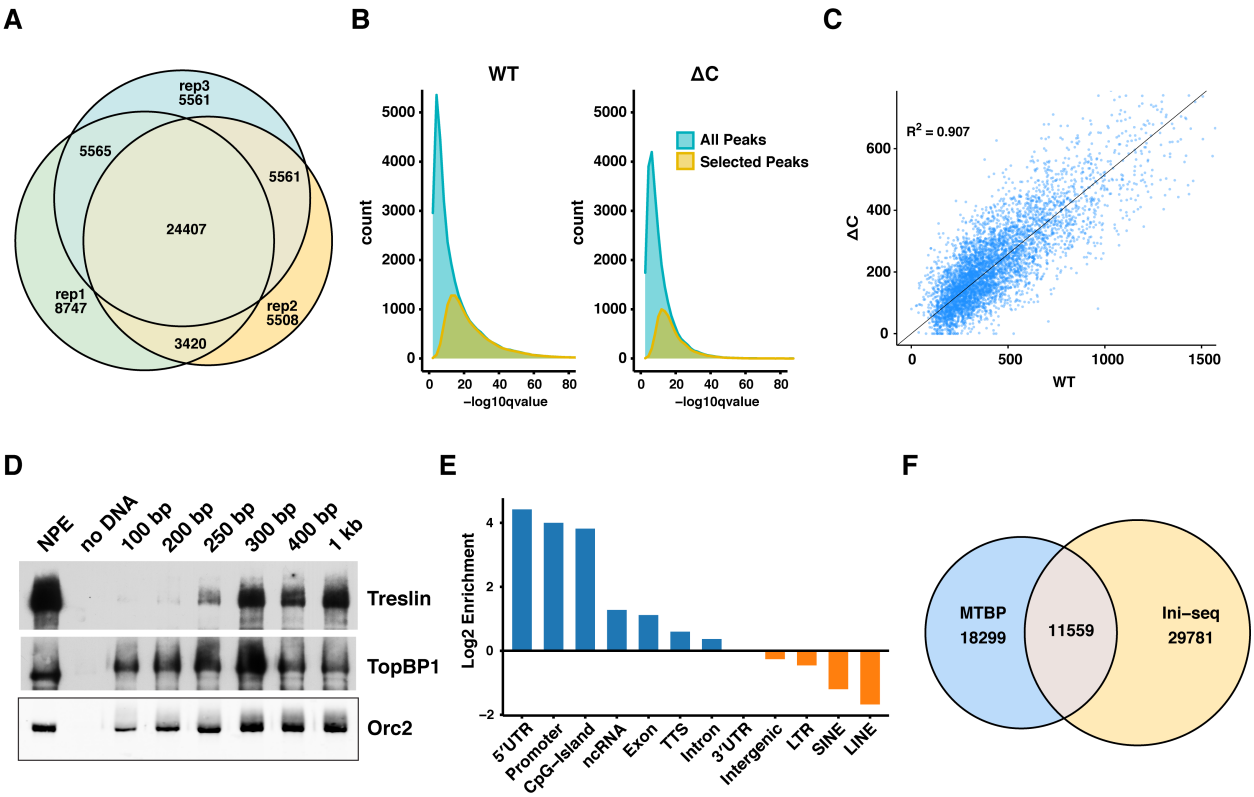

Figure S2. Characterization of MTBP Binding Sites, Related to Figure 2

(A) Venn diagram of three replicates of MTBP-WT CUT&RUN experiments. Peaks were called on each of three replicates using MACS2. Peaks that were called at least two out of three times were selected.

(B) All of the peaks identified in pooled files by MACS2 for MTBP-WT (left) and MTBP-ΔC (right) are shown in blue and selected reproducible peaks are shown in yellow. For MTBP-WT and MTBP-ΔC, more than 99.7% and 99.0% of the selected peaks, respectively, have a q-value (FDR) of $< 10^{-6}$.

(C) Read scores at each peak for MTBP-WT and MTBP-ΔC were generated with Deeptools multiBigWigSummary and plotted. Each dot corresponds to a peak.

(D) Binding of *Xenopus* Treslin, TopBP1, and Orc2 to biotinylated fragments of double-stranded DNA of the indicated lengths bound to streptavidin beads in *Xenopus* NPE fractions. DNA fragments of overlapping sequence were generated from pBluescript by PCR.

(E) Genome ontology analysis using HOMER AnnotatePeaks is plotted. Genome segments were determined according to the GENCODE basic gene annotation file (v32).

(F) Venn diagram of MTBP peaks and Ini-seq peaks showing that 38.7% of the MTBP binding sites overlap with initiation sites that were detected by Ini-seq.

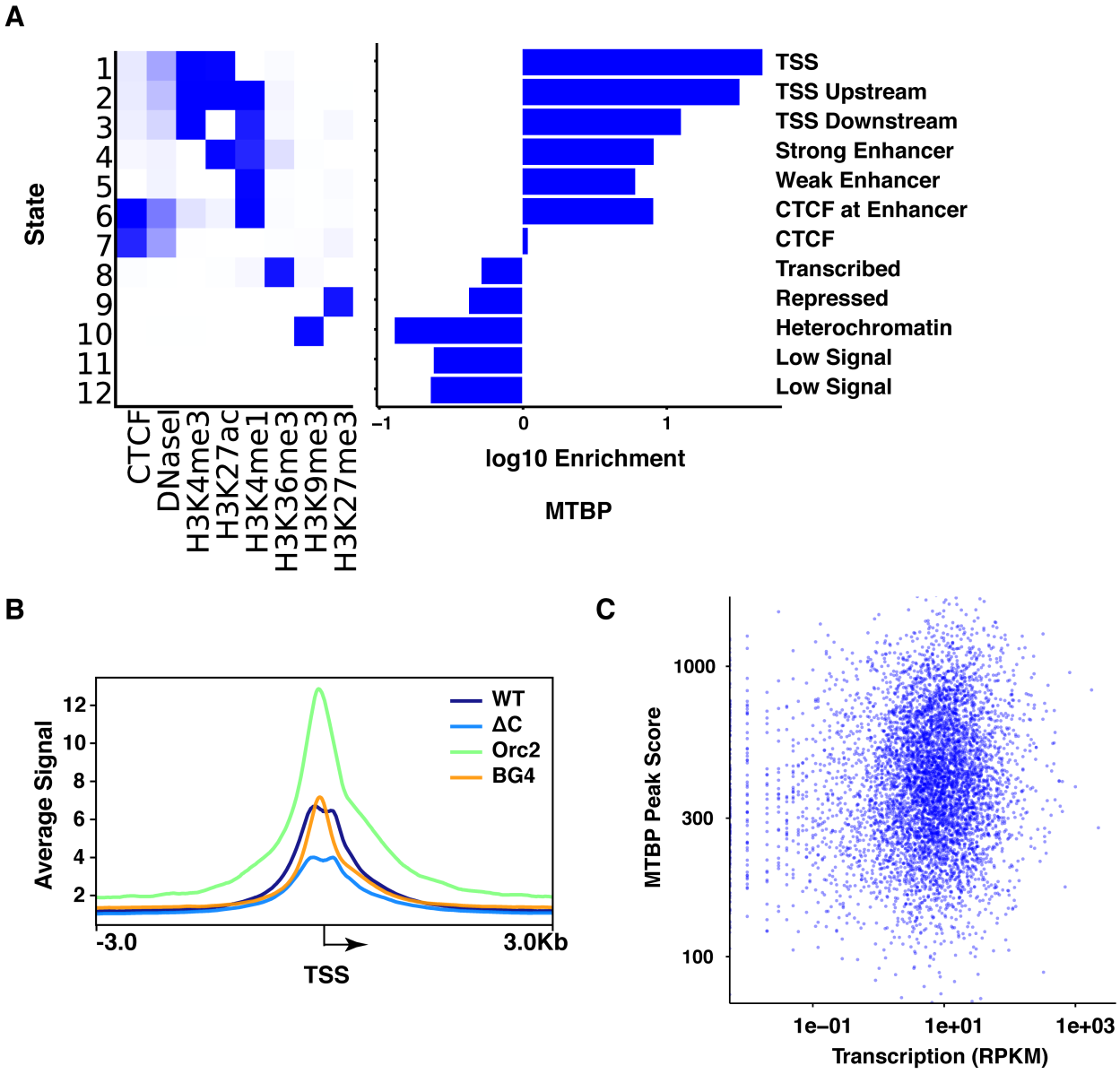

Figure S3. MTBP Is Enriched at Transcriptional-Regulatory Elements, Related to Figure 3

(A) ChromHMM analysis of MTBP peaks. With the chromatin-state discovery tool ChromHMM and publicly available DNase-seq and ChIP-seq datasets (H3K4me3, H3K27ac, H3K4me1, H3K36me3, H3K9me3, H3K27me3, and CTCF) from ENCODE, 200 bp genomic segments were classified into twelve states and annotated as shown on the right.

(B) MTBP, MTBP-ΔC, Orc2 ChIP-seq, and BG4 ChIP-seq average reads around TSSs were plotted.

(C) Peak scores of MTBP peaks at TSSs and corresponding reads from RNA-seq (RPKM) from DLD-1 cells were plotted.

A **Homer *de novo* Motif Results**

Total target sequences = 26627
Total background sequences = 25375

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1		1e-1789	-4.120e+03	30.95%	10.54%	491.6bp (302.2bp)	AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer(0.950)
2		1e-437	-1.008e+03	21.77%	11.99%	500.9bp (310.5bp)	RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq(GSE29180)/Homer(0.966)
3		1e-411	-9.472e+02	39.25%	26.97%	523.7bp (322.8bp)	TEAD3/MA0808.1/Jaspar(0.970)
4		1e-296	-6.833e+02	28.65%	19.28%	527.5bp (319.9bp)	FOXP3/MA0850.1/Jaspar(0.912)
5		1e-249	-5.735e+02	35.62%	26.22%	513.5bp (318.0bp)	CDX1/MA0878.1/Jaspar(0.938)
6		1e-173	-3.999e+02	43.24%	34.90%	563.9bp (329.6bp)	TFCP2/MA0145.3/Jaspar(0.865)
7		1e-163	-3.758e+02	29.39%	22.21%	544.8bp (323.1bp)	Tcf3(HMG)/mES-Tcf3-ChIP-Seq(GSE11724)/Homer(0.928)

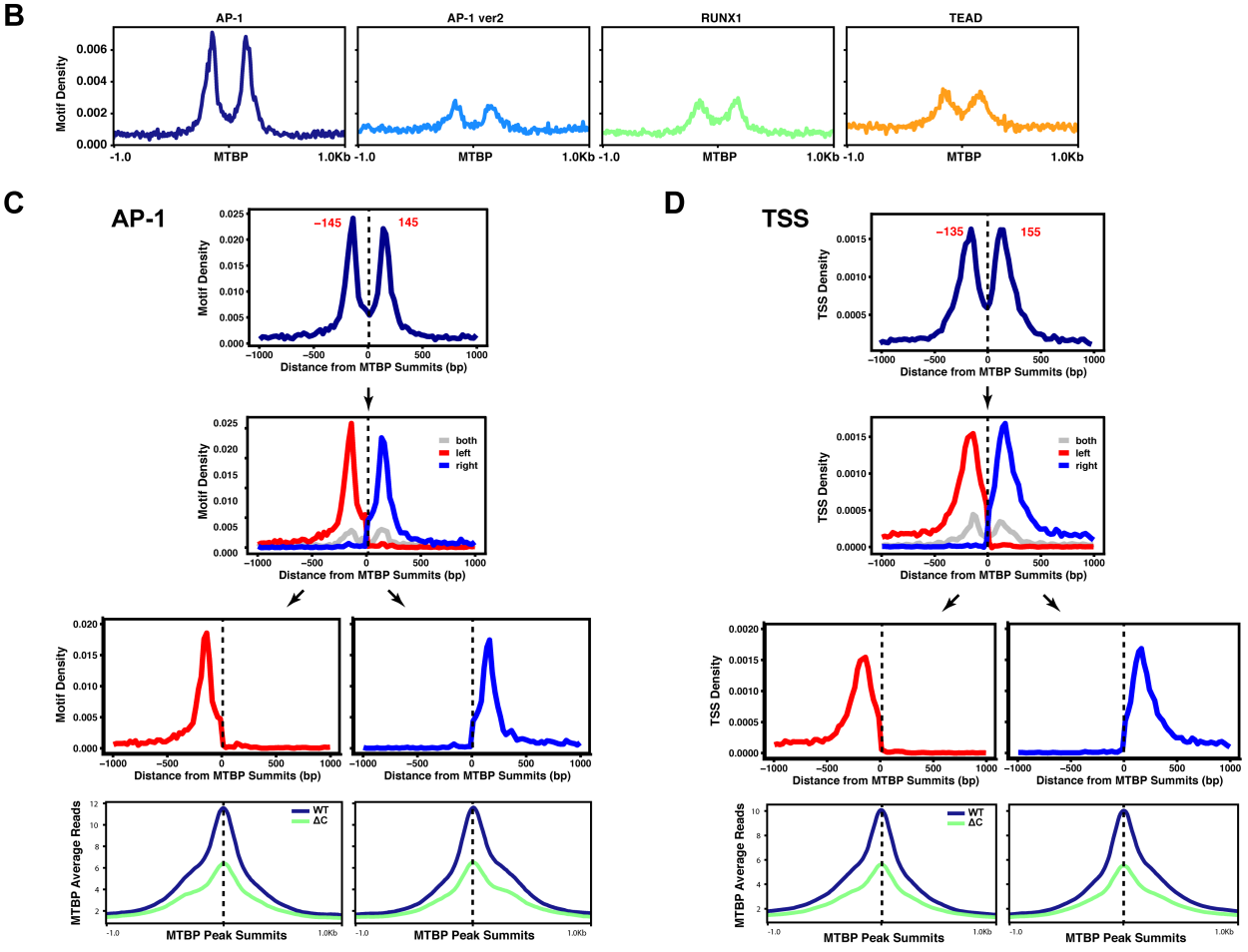

Figure S4. Identification and Characterization of Motifs in MTBP Binding Sites, Related to Figure 4

(A) Results of motif-searching analysis. In HOMER, findMotifsGenome.pl was used to search for motifs enriched in the MTBP peaks.

(B) Locations of AP-1 (TGAG/CTCA), a variation of AP-1 (version 2; TGAG/CTAA), RUNX (AACCACA), and TEAD (A/GCATTCC) motifs were plotted around MTBP peak summits.

(C) AP-1 motif peaks are 145 bp (± 10 bp) away from the MTBP peak summits. Each peak file was examined and separated into three classes according to whether the AP-1 motif was on the left, right, or both sides. The bottom figure shows the two profiles of MTBP peaks that have an AP-1 site on the left or right, respectively.

(D) TSSs are located around 145 bp (± 10 bp) away from MTBP peak summits. Analysis was performed in the same manner as in panel C.

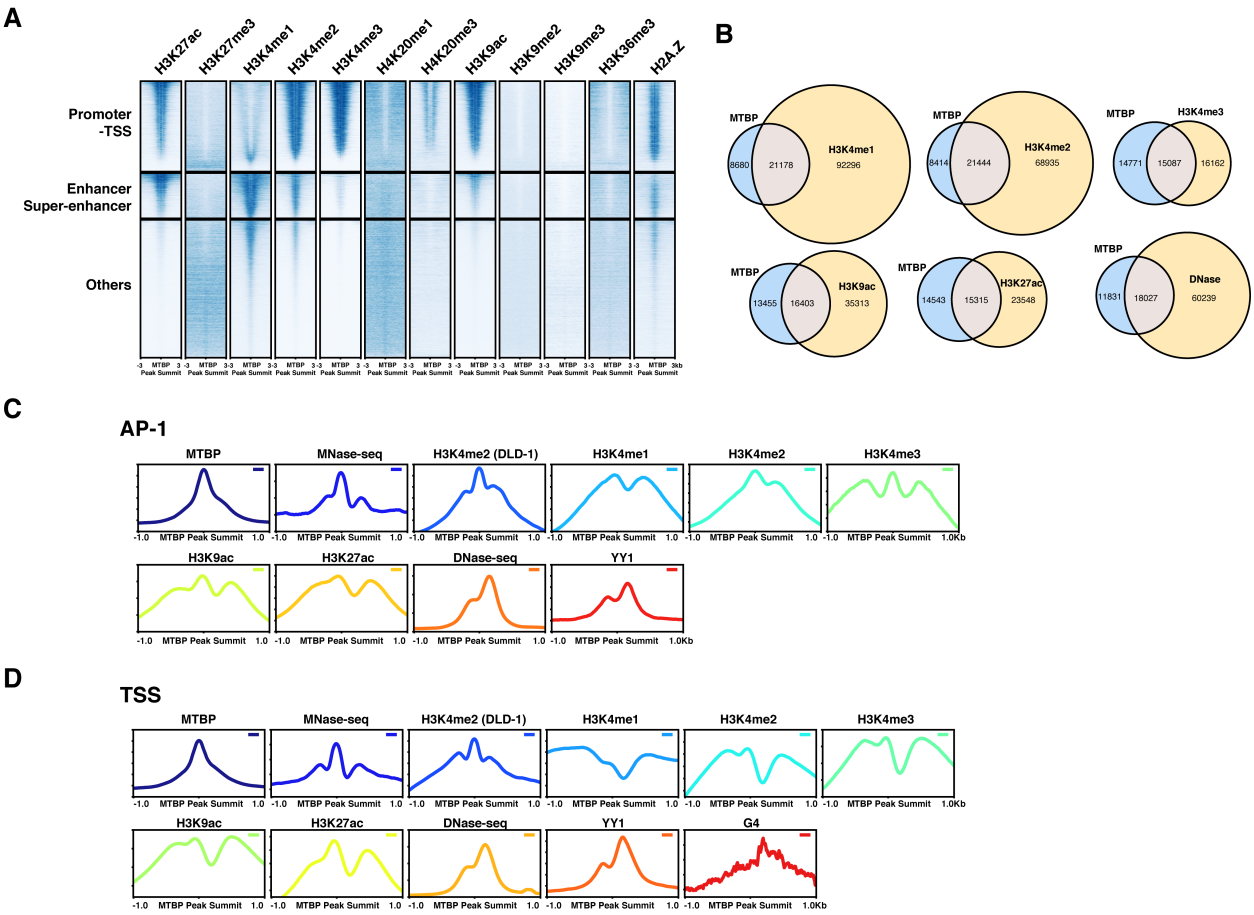

Figure S5. Analysis of Histone Markers at the MTBP Binding Sites, Related to Figure 5

(A) Heatmap analysis of read distributions for various histone markers from HCT-116 cells around the summits of MTBP peaks.

(B) Venn diagram of DNase-seq and various ChIP-seq peaks that overlap with MTBP peaks.

(C) Various signal profiles around MTBP summits that have AP-1 sites. The peaks are oriented so that AP-1 sites are on the right of MTBP summits.

(D) Various signal profiles and G4 motifs around MTBP summits near TSSs. The peaks are oriented so that TSSs are on the right of MTBP summits.

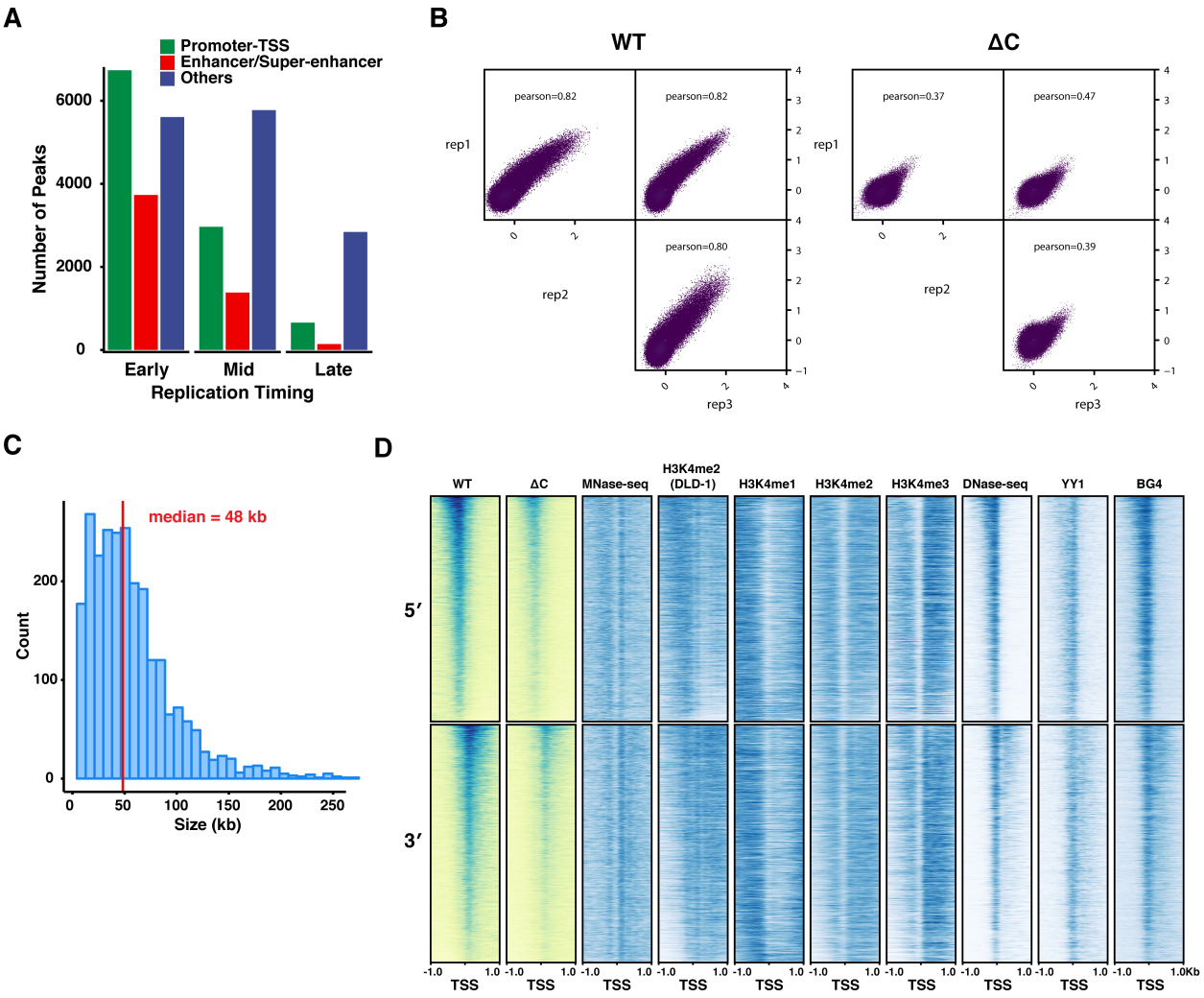

Figure S6. Analysis of EdU-seq Patterns in DLD-1 Cells, Related to Figure 6

(A) Distributions of different classes of MTBP binding sites (Promoter-TSS, Enhancer/Super-enhancer, and Others) in replication timing domains (Early, Mid, and Late).

(B) Comparison of the read counts for the three replicates of EdU-seq for MTBP-WT (left three panels) and MTBP-ΔC (right three panels). Analysis was performed using Deeptools multiBigWigSummary.

(C) Size distribution of the initiation zones for MTBP-WT determined by EdU-seq.

(D) Heatmap analysis of MTBP-WT, MTBP-ΔC, MNase-seq, H3K4me2 CUT&RUN from DLD-1 cells, DNase-seq, and various ChIP-seq peaks around TSSs containing MTBP at either the 5′ or 3′ side, respectively. MTBP peaks were segregated according to whether their summits were on the 5′ or 3′ side. Heatmaps were oriented so that transcription was from left to right.

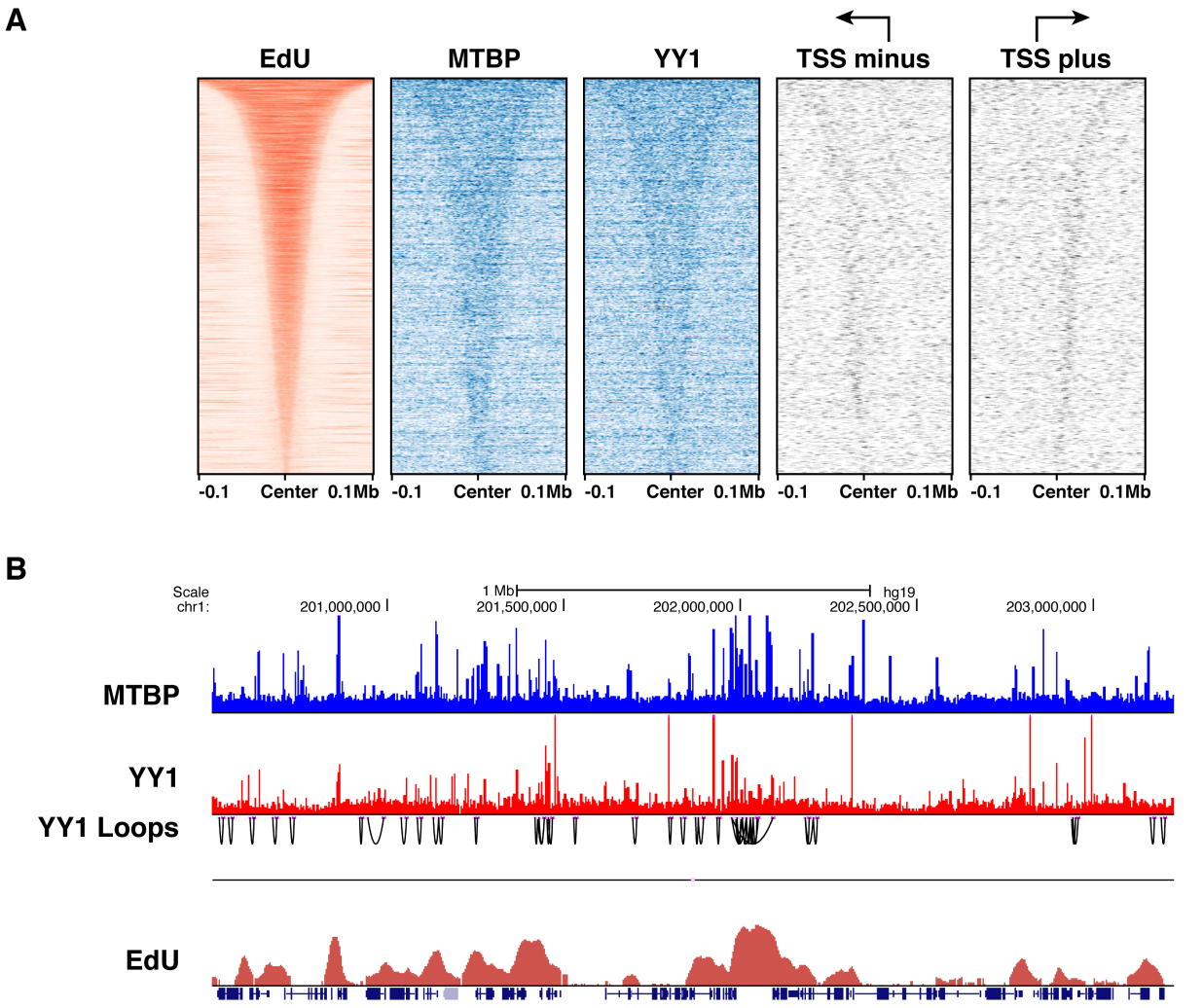

Figure S7. Characterization of Early-Replicating Zones in DLD-1 Cells, Related to Figure 7
(A) Heatmaps of reads of EdU-seq, MTBP, and YY1 as well as locations of TSSs were centered around the midpoint of initiation zones. TSSs were plotted according to the direction of transcription as indicated with arrows. Locations of TSSs were extended 2 kb upstream and downstream for visibility.
(B) Genome browser profiles of MTBP, YY1, YY1 loops, and EdU-seq. Profiles were mapped to hg19. Loops connecting adjacent bins were removed. YY1 loops with PET scores of 10 and above are shown.

Table S1. Statistical Tests Related to Figure 4

Summary of Kolmogorov-Smirnov tests performed using ks.test() in R

MTBP peaks within 10 kb upstream and 10 kb downstream of TSSs and 20 kb upstream and 20 kb downstream from the centers of enhancers were selected. Distributions of the peaks near TSSs and near enhancers with or without G4 DNA and with or without the TGAG/CTCA (AP-1) motif were compared.

	Features	No. of Peaks	D	p-value
TSS	G4 No G4	6,425 10,651	0.18672	$< 2.2e^{-16}$
Enhancer	G4 No G4	3,694 7,570	0.041602	0.0003708
TSS	AP-1 No AP-1	4,079 12,997	0.12454	$< 2.2e^{-16}$
Enhancer	AP-1 No AP-1	3,667 7,597	0.07467	$2.106e^{-12}$

Table S2. Statistical Tests Related to Figures 4 and 5

Summary of permutation analyses of overlaps between MTBP peaks and various genomic features using regioneR
p-value = 0.001 (1,000 permutations)

Feature	No. of Features	Overlaps (observed)	Overlaps (random) (Mean ± sd)	Z-score
G4/DNase I	11,919	5,244	128 ± 14	359.3
AP-1/DNase I	15,065	5,528	158 ± 15	350.7
YY1	17,804	9,302	231 ± 17	538.0
H3K4me2 (DLD-1)	62,503	29,598	1716 ± 39	706.2
H3K4me1	116,231	17,224	1,438 ± 38	412.2
H3K4me2	91,319	21,046	1,178 ± 33	597.3
H3K4me3	31,914	14,664	518 ± 24	598.2
H3K9ac	54,532	15,639	664 ± 25	605.5
H3K27ac	40,985	14,568	611 ± 25	568.6

Z-score is the difference between observed overlaps and the mean of random overlaps divided by the standard deviation of random overlaps. For G4/DNase I and AP-1/DNase I, DNase I-hypersensitive sites were extended by 50 bp upstream and downstream and either G4 motifs or AP-1 motifs that intersected with extended DNase I-hypersensitive sites were selected for testing overlaps with MTBP peaks. MTBP summits (± 100 bp from the peak summits) were used for testing overlaps with various histone modifications.

Table S3. Guide Oligonucleotides, Related to STAR Methods

MTBP Exon 22 Top	CACCGATACCCAGTCAATCACCTAG
MTBP Exon 22 Bottom	AAACCTAGGTGATTGACTGGGTATC
MTBP Exon 19 Top	CACCGCTTGCTACAAAGACCAGTTC
MTBP Exon 19 Bottom	AAACGAACTGGTCTTTGTAGCAAGC

Table S4. Oligonucleotides for Cloning of Genomic MTBP Fragments for Donor Templates, Related to STAR Methods

For Exon 22 donor 5′	TGTTGTAGTCTGTCCCTAATCAC
For Exon 22 donor 3′	CACATACACACATATGCATAACAAC
For Exon 19 donor 5′	ATACAGGTTGTGGTGTTTCTGTTC
For Exon 19 donor 3′	TTTCAAAGAGACGCTGGCTGC

Table S5. Oligonucleotides for Construction of Donor Templates, Related to STAR Methods

MTBP-mAID-Hygro and MTBP-FLAG-Hygro_A	GGTGATTGATTGGGTCCTTGAAAAGACAA GCAAGAAAGGATCCGGTGCAGGCGCC
MTBP-mAID-Hygro and MTBP-FLAG-Hygro_B	TACCCAGTCAATCACCCCTAGTGAACCTC TTCGAGGGACC
MTBP-mAID-Hygro and MTBP-FLAG-Hygro_C	GGTGATTGACTGGGTATTAGAAAAG
MTBP-mAID-Hygro and MTBP-FLAG-Hygro_D	GACCCAATCAATCACCTAGGGGGAGAAA AAATATGATTACAGAAG
MTBP-mAID-Neo_A	GGTGATTGATTGGGTCCTTGAAAAGACAA GCAAGAAAGGATCCGGTGCAGGCGCC
MTBP-mAID-Neo_B	TACCCAGTCAATCACCGATCCAACGACCC AACACCGTGCGTT
MTBP-mAID-Neo_C	GGTGATTGACTGGGTATTAGAAAAG
MTBP-mAID-Neo_D	GACCCAATCAATCACCTAGGGGGAGAAA AAATATGATTACAGAAG
MTBP-ΔC-FLAG-Hygro_A	CAGTGGTCAGAAGTCAATGCATGAAAGTA AGGGATCCGGTGCAGGCGCC
MTBP-ΔC-FLAG-Hygro_B	ATACTTTTTTGACCTGAACCCTAGTGAAC CTCTTCGAGGGACC
MTBP-ΔC-FLAG-Hygro_C	GAAGAGGTTCACTAGGGTTCAGGTCAAAA AAGTATGCATGAATC
MTBP-ΔC-FLAG-Hygro_D	GACTTCTGACCACTGCTGGTCTTTGTAGC AAGTTTTG