Supplementary data file

SUPPLEMENTARY MATERIALS AND METHODS

Antibodies

Antibodies used in ChIP-Seq and ChIP analyses were as follows: ZNF143, rabbit polyclonal raised against a C-terminal epitope of ZNF143 (51); ZNF76, mouse monoclonal (Santa Cruz Biotechnology, cat. #sc-81147); ICN1, rabbit polyclonal raised against amino acids 2278-2470 of human Notch1 (45); HA, 12CA5 antibody, (Roche Applied Science), RNA polymerase II CTD (Abcam, ab5408). Western blot analysis was performed with the antibodies above or antibodies against ICN1, rabbit monoclonal, Val1744 (Cell Signaling Technology, cat. #sc-4147); CSL/RBPJk, goat polyclonal (Santa Cruz Biotechnology, cat. #sc-55021); ronin/THAP11 goat polyclonal (Santa Cruz Biotechnology, cat. #sc-165683); ZNF76, rabbit polyclonal raised against amino acids 501-515 of human ZNF76 (16); α-tubulin, mouse monoclonal (Sigma, T6557).

Cell proliferation assay

CyQUANTCell® NF Cell Proliferation Assay Kit (Invitrogen) was used according to manufacturer's instructions to measure the proliferation of cells after siRNA transfection. HeLa S3 cells were transfected with 2.5 pmol of specific or non-specific siRNA (described in the main Material and Methods section) in 96 well plates culture vessels (2000 cells per well) using Lipofectamine 2000 reagent (Invitrogen). The fluorescence was measured 24, 48 and 72h post transfection on the GloMax®-Multi Detection System (Promega).

Constructs

Construct ZNF143-3xHA-pcDNA5/FRT/TO (pNG26). The 3X HA TAG was amplified by PCR from pGHA-BcK using forward (fw) and reverse (re) primers (fw,

CTTAAGTACCCATACGATGTTCC bearing an AflII site; re, GCTAGCGGATCCTTAAGCGTAATCTGGAACGT carrying NheI and BamHI sites) and ligated to pGMTeasy (Promega). Site-directed mutagenesis obliterated the BamHI (GGATCC to GCATCC) and AflII sites (CTTAAG to CTTGAG) located in the Tag ORF generated pNG08. The AflII-NheI fragment from pNG08 containing the 3xHA TAG was inserted into pNG07 to produce pNG09. pNG07 was constructed by ligating to pGMTEasy the PCR product (containing the full length ZNF143, 638 aa, without stop codon) obtained from pSK(-)–ZNF143 (16) using forward (fw) and reverse (re) primers (fw, GGATCCGCCGCCACCATGTTGTTAGCCCAAATAAATCGAGATTCTCAGGGAATG GCAGAGTTTCCTGGAGGAGG bearing BamHI site and Kozak ATG; re, GCTAGCCTTAAGATCATCCAACCCTGGCGTTTC carries NheI and AflII sites). The BamHI fragment of pNG09 was ligated to pcDNA5/FRT/TO (Invitrogen) to generate ZNF143-3xHA-pcDNA5/FRT/TO.

ZNF143-pcDNA5/FTR/TO (pNG27). pNG19 containing the full length ZNF143 (638 aa) was obtained by ligating to pGMTEasy the PCR amplification product from pNG07 using forward (fw) and reverse (re) primers (fw, GGATCCGCCGCCACCATGTTGTTAGCCCAAATAAATCGAGATTCTCAGGGAATG GCAGAGTTTCCTGGAGGAGG bearing BamHI site and Kozak ATG; re, CGGGATCCTTAATCATCCAACCCTG bearing stop codon and BamHI site). The ZNF143-pcDNA5/FTR/TO (pNG27) was then obtained by ligating the BamHI fragment from pNG19 to pcDNA5/FRT/TO.

To generate construct ZNF76-pcDNA5/FRT/TO (pNG24), the region coding the full length ZNF76 (570 aa) was amplified from pOTB7-ZNF76-BC002549 (Imagenes GmbH) by PCR using forward (fw) and reverse (re) primers (fw,

GGATCCGCCGCCACCATGGAGAGCTTGGGGC bearing BamHI site and Kozak ATG; re, GGATCCTCAGCAGCCACTCTCCG carrying BamHI site). The resulting BamHI fragment was ligated to BamHI-cut pcDNA5/FRT/TO to generate ZNF76-pcDNA5/FRT/TO. To generate construct THAP11-3xHA-pcDNA5/FRT/TO, the region coding the full length human THAP11 (without stop codon) was amplified by PCR from clone IRAUp969F0765D (Imagenes GmbH) using forward (fw) and reverse (re) primers (fw, CTTAAGGCCACCATGCCTGGCTTTACGTGCTGC bearing AflII site and Kozak ATG; re, CTTAAGCATTCCGTGCTTCTTGCGG carry AflII site) and inserted into pGMTeasy to generate E482. The AflII fragment from E482 was ligated to pcDNA5/FRT/TO/3xHA to generate ronin-3xHA-pcDNA5/FRT/TO. To generate construct GST-THAP11 DBD, the coding region of the THAP11 DBD (aa 1-90) was amplified by PCR from clone IRAUp969F0765D (Imagenes GmbH) using forward (fw) and reverse (re) primers (fw, GCTCCGCCATGGCTatgCCTGGCTTTACGTGCTG bearing, NcoI site and Kozak ATG; re, ATCCGGCTCGAGTACTTTGCGCTCATTGACGC carries XhoI site). The resulting NcoI-XhoI fragment was ligated in frame into VK91 downstream of GST (30).

ChIP-Seq data analysis

Peak calling. Peak detection was performed using the Model-based Analysis of ChIP-Seq (MACS) software (52) under default MACS settings with a bandwidth of 200bp and the use of a dynamic λ local (10 kb window).

Peak analysis. The utility Peak Splitter from PeakAnalyzer (53) was used for accurately subdividing experimentally-derived peak regions containing more than one site of signal enrichment. The PeakAnnotator component from PeakAnalyzer was used to identify and report the functional elements proximal to peak loci. The BEDTools (54) utilities allowed us to address common genomic tasks such as finding

feature overlaps.

Integrative analysis of ChIP-Seq data. Refseq genes, ncRNA genes and repeated element (RepeatMasker) annotations of peaks were performed using corresponding GTF (hg19) tables from the UCSC TableBrowser (55) and Biomart Ensembl (56) servers. Table data from other works were collected from publication supplemental materials (25, 34). Raw ChIP-Seq data were downloaded from the Gene Expression Omnibus (GEO) or Sequence Read Archive (SRA) databases as follows: Pol II ChIP-Seq data on HeLa-S3 cells, Encode project, NCBI GEO GSE12781 (57); H3K4me3 and H3K36me3 data on HeLa cells (58) NCBI GEO GSE20303; H3K4me3 and Pol II data on K562 cells, Encode project (6), NCBI GEO GSM733680, GSM608165 and GSE30227; H3K27ac and H3K4me1 NCBI GEO GSM733684 and GSM798322 (6) Pol II on mESC, NCBI GEO GSE12241 (59). When required, sequence alignments were performed on hg19 or mm9 by the Bowtie software (29). Multiple enrichment analysis of ChIP-Seq data was performed with the Seq-Miner software (60). Correlation, association and peak distribution studies were performed using Cistrome (http://cistrome.org/) (49).

Motif searches. For de novo motif search, we used the MEME suite 4.6.1 (50) via the MEME-ChIP web service analysis (http://meme.sdsc.edu/) and also by running MEME locally on 100 bp regions centered at the pic maxima. To determine the central motif enrichment, we used the CentriMo tool from the MEME suite**.**

Functional annotation analysis. Gene ontology analysis was conducted using the Functional Annotation Chart and Functional Annotation Clustering services of the DAVID version 6.7 (http://david.abcc.ncifcrf.gov/) (38). To look for enriched biological processes, we used the list of genes associated with the set of peaks located at +/-2 kbp of a TSS.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Western blot detection of ZNF143, ZNF76, THAP11, RBPJ and ICN1 in various human and murine whole-cell lysates. ChIP-qPCR showing ZNF143 occupancy of known ZNF143 binding sites. **(A)** Detection of ZNF143, ZNF76, THAP11, RBPJ in the indicated cell lines; FLP143-HA (denoted by a star) and FLP76 correspond to inducible stable cell lines expressing ZNF143-HA and ZNF76, respectively. **(B)** Detection of Zfp143 (mouse ZNF143), Zfp523 (mouse ZNF76), Thap11 and RBPJ in the indicated cell lines. **(C)** Detection of ICN1 in the indicated cell lines. α-tubulin levels are shown as loading control. **(D)** Western blot detection of ZNF76 after immunoprecipitation (IP) of whole cell extract from $4x10^7$ T-REx HEK293 cells; no Ab, IP control without antibody; S, supernatant of IP; Ctrl ZNF76, extract from stable cell line expressing ZNF76. **(E)** Anti-ZNF143-ChIP-qPCR and fold enrichment observed at previously identified ZNF143 binding sites versus close regions devoid of binding sites.

Figure S2. Agreement between the data sets from ChIP-Seq experiments performed using anti-ZNF143 antibody in mouse cell lines and anti-HA in FLP143-HA cell line expressing ZNF143-HA. **(A)** Biological replicates (HeLa-rep1 and HeLa-rep2) of ZNF143 ChIP-Seq analysis in HeLa-S3 cells, **(B)** Data from ChiP-Seq experiments performed using anti-ZNF143 in 293T-Rex and anti-HA in FLP143-HA cell line expressing ZNF143-HA. Correlation plot represents the enrichment correlation on the union of peaks (p-value 10^{-5} and minimum height of 20 reads). Scale represents normalized number of reads. **(C)** Correlation plots of ZNF143 binding events in mouse cell lines. Grid of pairwise scatter plots with associated Pearson correlation coefficient. Enrichment at all associated loci (63482 loci for a p-value threshold of

10^5) is included. The regression line is indicated in red and the Pearson correlation coefficients are indicated in the squares on the top right.

Figure S3. Distribution and correlation of ZNF143 binding events in different human and mouse cells. (A) Pie charts showing distribution of all ZNF143 binding events in different genomic regions. Legend was as described in Figure 1B. (B) Correlation plots of ZNF143 binding events located at +/- 2kbp from a TSS. Grid of pairwise scatter plots with associated Pearson correlation coefficient obtained using Cistrome software (49). Enrichments at the 3122 human and 3385 mouse loci are listed in Supplemental Tables S4 and S5, respectively. The regression line is indicated in red and the Pearson correlation coefficients are indicated in the squares on the top right.

Figure S4. Examples of high-confidence ChIP-Seq peaks common to HeLa-S3, K562, T-Rex-293 and HPB-ALL cells and identified at known or at new ZNF143 binding sites. **(A)** Divergent TP53-WRAP53 gene pair. **(B)** SCARNA2 gene. **(C)** BUB1B gene. Chromosome coordinates are shown on top. Gene annotations derived from the UCSC Genome Browser Database are shown at the bottom with arrows indicating direction of transcription. Data are shown as UCSC browser views.

Figure S5. Chromosomal distribution and enrichment of the human (A) and mouse (B) ZNF143 binding events located at +/- 2kbp from a TSS. **Left part of A and B.** The distributions of peaks are given in percentage of events. The distribution of mappable regions in the chromosome (genome background) is given as a reference. P-values are shown for the significance of the relative enrichment of ChIP regions

with respect to the genome background. **Right part of A and B**. The ratio of ZNF143 peaks compared to the number of genes is represented for each chromosome. The gray line represents the average ratio of binding events per number of genes at the scale of the genome.

Figure S6. Distribution in different families of 512 ZNF143 peaks associated to ncRNA genes.

Figure S7. Alignment of promoter elements for the Pol II snRNA genes (upper part of part A, and B) and Pol III type 3 genes (lower part of part A) associated to ZNF143 binding events. The coordinates on the left or right of the motifs refer to the first or last nucleotide of the sequence shown relative to the TSS (+1). Motifs in italics are in inverted orientation. Underlined nucleotides are common to octamer and SBS1 motifs. In part B, the motifs are depicted as sequence logo.

Figure S8. ZNF143 occupancy on SBS1 and SBS2 motifs. Box blots showing ChIP-Seq score (number of reads) distribution of ZNF143 binding events containing: only increasing number of SBS1 motifs (A), and only increasing number of SBS2 motifs (B).

Figure S9. Functional annotation analysis of ZNF143 related genes. Functional categories enriched among the genes located at +/- 2 kbp of the 3122 identified human ZNF143 binding events (Table S4), as reported by the DAVID web-based functional annotation software (38). Binding events containing either an SBS1 or an SBS2 motif, at least one motif and without identified motif are shown. Values refer to

the fold-enrichment score compared to a background set of all human genes.

Figure S10. The ZNF143 and ZNF76 binding events are identical. Examples of ZNF143 and ZNF76 binding events. UCSC Genome Browser views of the indicated genomic locations. In each case, the upper track shows the peak obtained with ZNF76 and the lower one that obtained with ZNF143. The Y axis indicates the tag number at the peak summit. Left and right part, genomic ZNF143 and ZNF76 locations with only SBS2 or only SBS1 motif, respectively.

Figure S11. Measurements of the DNA-binding affinities of ZNF143 for SBS1 and SBS2 probes. Saturation curve data were generated by incubating a constant amount of protein with variable concentrations of probes until the equilibrium occurred. The bound fraction was quantitated following separation by gel electrophoresis. The results of one representative experiment for each probe are shown. Two other independent determinations gave similar results. The given Kd value is the mean of the three experiments.

Figure S12. The MLT1J loci are occupied by ZNF143, ICN1 and THAP11 in vivo. **(A)** Association of THAP11 with MLT1J element. Anti-HA-ChIP-qPCR and enrichment value of genomic regions containing MLT1J repeat element or SBS2 motif in AHSA1 promoter. **(B)** Examples of ZNF143 and ICN1 binding events on MLT1J elements. UCSC Genome Browser views of the indicated genomic locations. The upper track shows the peak obtained in ChIP-Seq with ZNF143, the middle one with ICN1 and the lower one with input DNA. The Y axis indicates the tag number at the peak summit.

Figure S13. ICN1 binding regions colocalize on the ZNF143 binding events. Examples of common ZNF143 and ICN1 binding events in HPB-ALL cells. UCSC Genome Browser views of the indicated genomic locations. The upper track shows the peak obtained in ChIP-Seq with ICN1, the middle one with ZNF143 and the lower one with input DNA in HPB-ALL cells. The Y axis indicates the tag number at the peak summit.

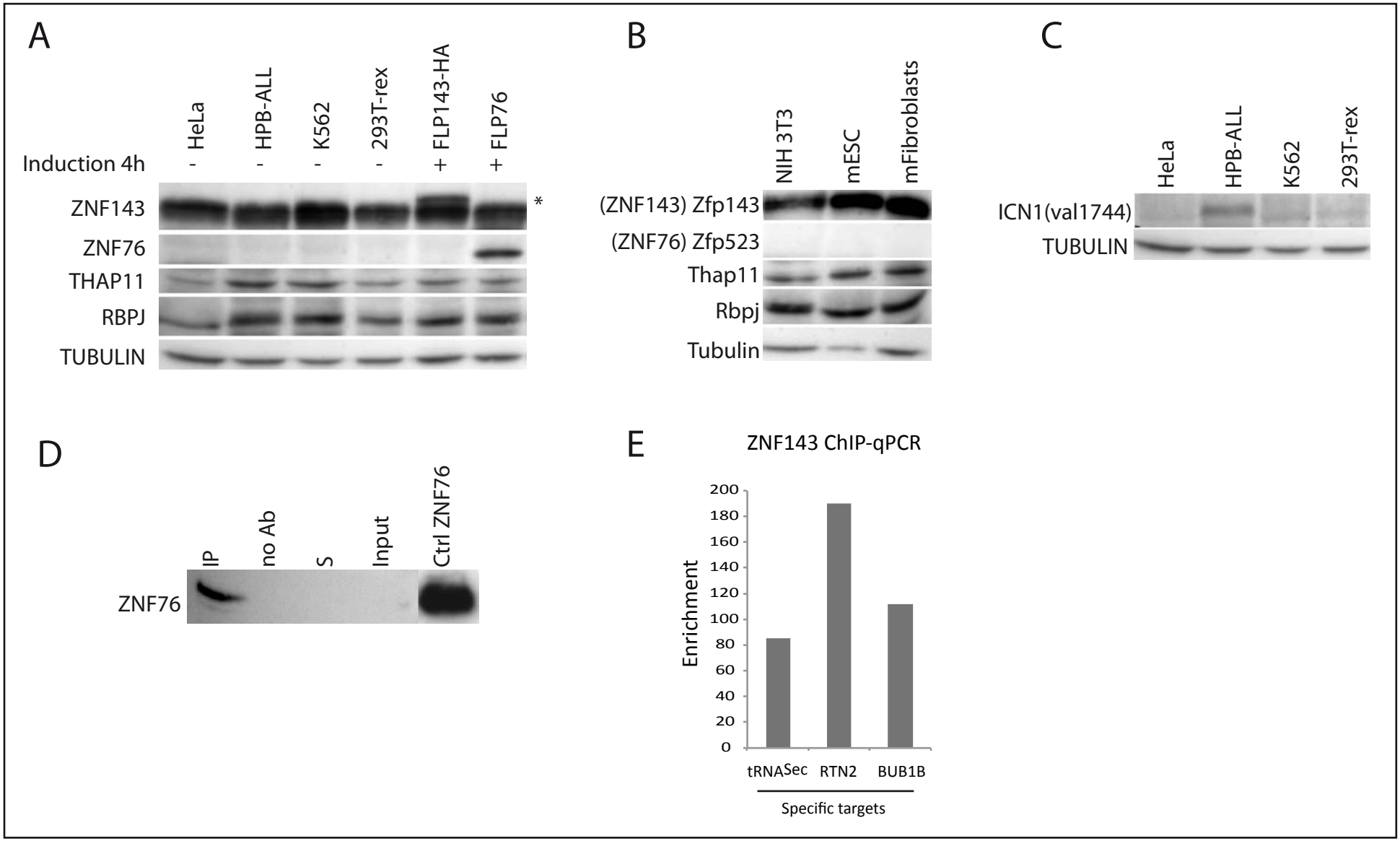
Figure S14. Effect of ZNF143 knockdown on the expression of a set of SBS1 and SBS2 genes. **(A)** Bar Chart representing the relative expression of a set of 51 genes obtained from RNA-Seq and RT-qPCR analysis after knockdown of ZNF143. In red is the relative expression ratio detected by RT-qPCR using RNA from cells transfected with ZNF143 specific siRNA compared to transfection with siRNA control. GAPDH was used for normalization. Values represent the mean +/- SD of two or three independent experiments. In blue is the relative expression ratio obtained from the deep RNA sequencing data after ZNF143 knockdown. **(B)** Chart showing the functional annotation clustering (38) of genes having more than +/- 0.5 (log2) fold change in expression ratio after knockdown of ZNF143. Overrepresented biological processes of down-regulated genes (positively regulated genes) compared to the total set of genes located at +/- 2kb from a ZNF143 ChIP-Seq peak are represented in red, the up-regulated (negatively regulated) are in blue.

Figure S15 Heatmap enrichment profile of H3K27ac (data from GSM733684), H3K4me1 (data from GSM798322) and ZNF143 in HeLa-S3 cells on **(A)** all ZNF143 binding events; **(B)** on ZNF143 peaks in promoters and on **(C)** ZNF143 peaks

outside promoters**.**

Figure S16. Association of ZNF143 occupancy in promoters with RNA polymerase II and various histones marks. (A) Heatmap enrichment profile of ZNF143, Pol II, H3K4me3, H3K27me3 and H3K9me3 (data from GSE12781 and GSE20303 for Pol II and histone modification) in HeLa-S3 cells on the ZNF143 binding events identified at +/-2kbp from a TSS (Table S4). Shown are ZNF143, Pol II and histone modification localization between + / -1kbp around the summit of peaks. The Heatmap represents chip-seq enrichment scores near ZNF143 peaks intervals. Intervals are ordered using k-means clustering (using 5 classes) on all the 5 Chip-seq signals. **(B)** Average H3K36me3 gene profiles associated to human ZNF143 binding events located at +/-2kbp from a TSS. All ZNF143 binding events, binding events containing only SBS1 or only SBS2 are indicated in red, green and blue, respectively. **(C)** Average profile of ZNF143 peak around the center of ZNF143 binding events identified in HeLa cells at +/- 2kpb from a TSS (Table S4). All ZNF143 binding events, binding events containing only SBS1 or only SBS2 are indicated in red, green and blue**,** respectively. **(D)** Average profile of H3K4me3 (data from GSM733680 and GSM608165) and Pol II (data from GSE30227 and GSE12241) around the center of ZNF143 binding events identified in K562 cells and mESC at +/-2kbp from a TSS (Table S4). All ZNF143 binding events, binding events containing only SBS1 or only SBS2 are indicated in red, green and blue, respectively.

Figure S1


A
Induction 4h
- HeLa
- HPB-ALL
- K562
- 293T-rex
+ FLP143-HA
+ FLP76
ZNF143
ZNF76
THAP11
RBPJ
TUBULIN
*
B
NIH 3T3
mESC
mFibroblasts
(ZNF143) Zfp143
(ZNF76) Zfp523
Thap11
Rbpj
Tubulin
C
HeLa
HPB-ALL
K562
293T-rex
ICN1(val1744)
TUBULIN
D
IP
no Ab
S
Input
Ctrl ZNF76
ZNF76
E
ZNF143 ChIP-qPCR
Enrichment
200
180
160
140
120
100
80
60
40
20
0
tRNASec
RTN2
BUB1B
Specific targets

FigureS2

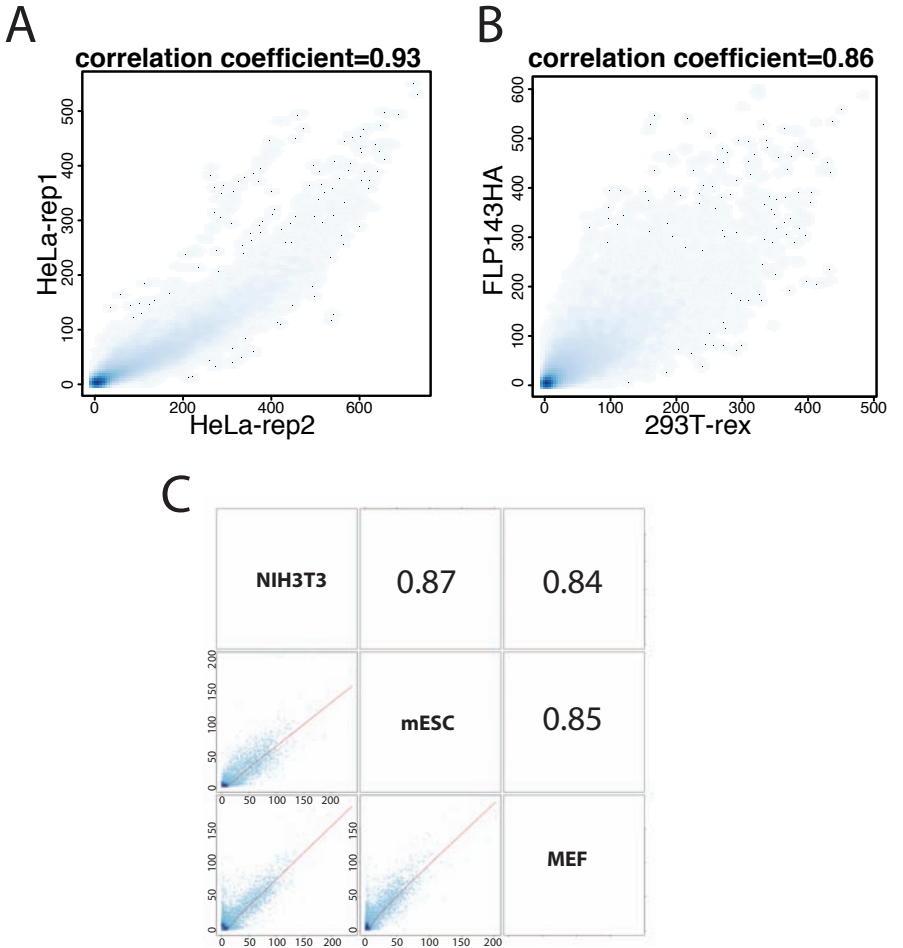
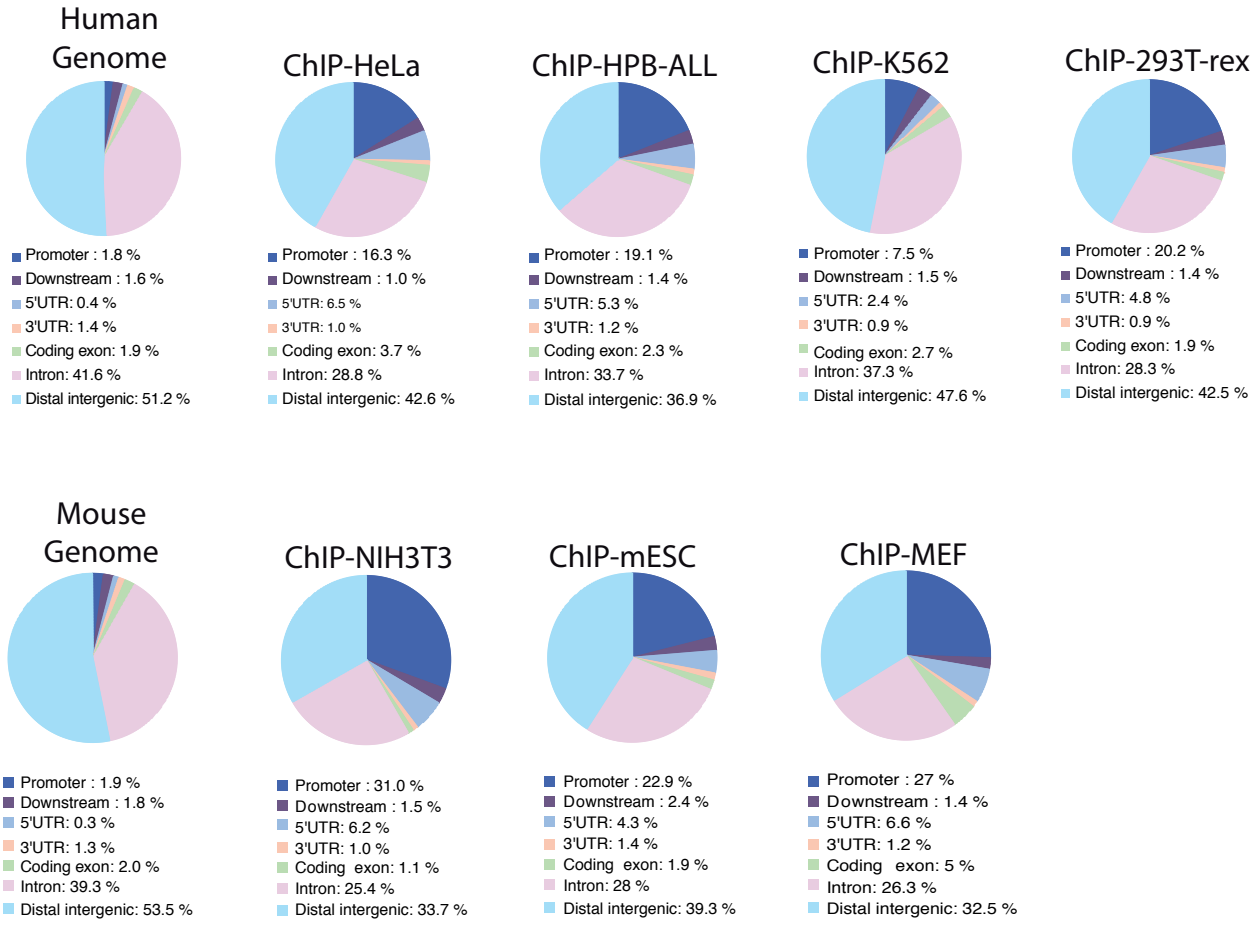

A
correlation coefficient=0.93
HeLa-rep1
HeLa-rep2
0
100
200
300
400
500
600
B
correlation coefficient=0.86
FLP143HA
293T-rex
0
100
200
300
400
500
600
C
NIH3T3
0.87
0.84
mESC
0.85
MEF
0
50
100
150
200

FIGURE S3

A


Human Genome
Promoter : 1.8 %
Downstream : 1.6 %
5'UTR: 0.4 %
3'UTR: 1.4 %
Coding exon: 1.9 %
Intron: 41.6 %
Distal intergenic: 51.2 %
ChIP-HeLa
Promoter : 16.3 %
Downstream : 1.0 %
5'UTR: 6.5 %
3'UTR: 1.0 %
Coding exon: 3.7 %
Intron: 28.8 %
Distal intergenic: 42.6 %
ChIP-HPB-ALL
Promoter : 19.1 %
Downstream : 1.4 %
5'UTR: 5.3 %
3'UTR: 1.2 %
Coding exon: 2.3 %
Intron: 33.7 %
Distal intergenic: 36.9 %
ChIP-K562
Promoter : 7.5 %
Downstream : 1.5 %
5'UTR: 2.4 %
3'UTR: 0.9 %
Coding exon: 2.7 %
Intron: 37.3 %
Distal intergenic: 47.6 %
ChIP-293T-rex
Promoter : 20.2 %
Downstream : 1.4 %
5'UTR: 4.8 %
3'UTR: 0.9 %
Coding exon: 1.9 %
Intron: 28.3 %
Distal intergenic: 42.5 %
Mouse Genome
Promoter : 1.9 %
Downstream : 1.8 %
5'UTR: 0.3 %
3'UTR: 1.3 %
Coding exon: 2.0 %
Intron: 39.3 %
Distal intergenic: 53.5 %
ChIP-NIH3T3
Promoter : 31.0 %
Downstream : 1.5 %
5'UTR: 6.2 %
3'UTR: 1.0 %
Coding exon: 1.1 %
Intron: 25.4 %
Distal intergenic: 33.7 %
ChIP-mESC
Promoter : 22.9 %
Downstream : 2.4 %
5'UTR: 4.3 %
3'UTR: 1.4 %
Coding exon: 1.9 %
Intron: 28 %
Distal intergenic: 39.3 %
ChIP-MEF
Promoter : 27 %
Downstream : 1.4 %
5'UTR: 6.6 %
3'UTR: 1.2 %
Coding exon: 5 %
Intron: 26.3 %
Distal intergenic: 32.5 %

B

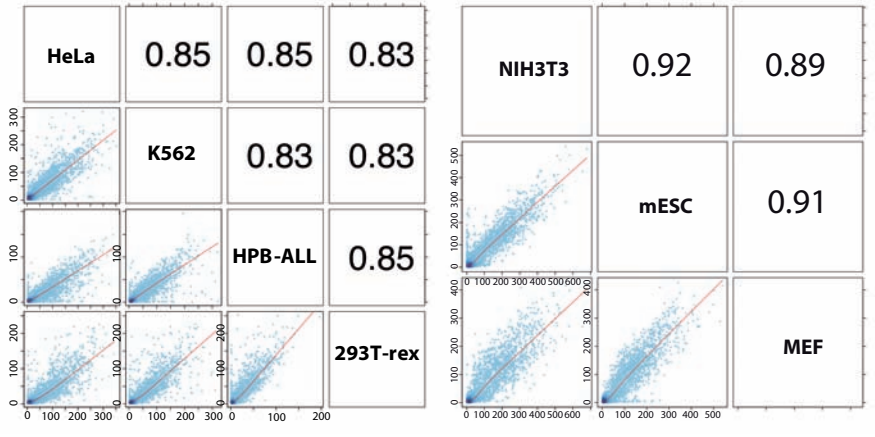

HeLa
0.85
0.85
0.83
K562
0.83
0.83
HPB-ALL
0.85
293T-rex
NIH3T3
0.92
0.89
mESC
0.91
MEF

Figure S4

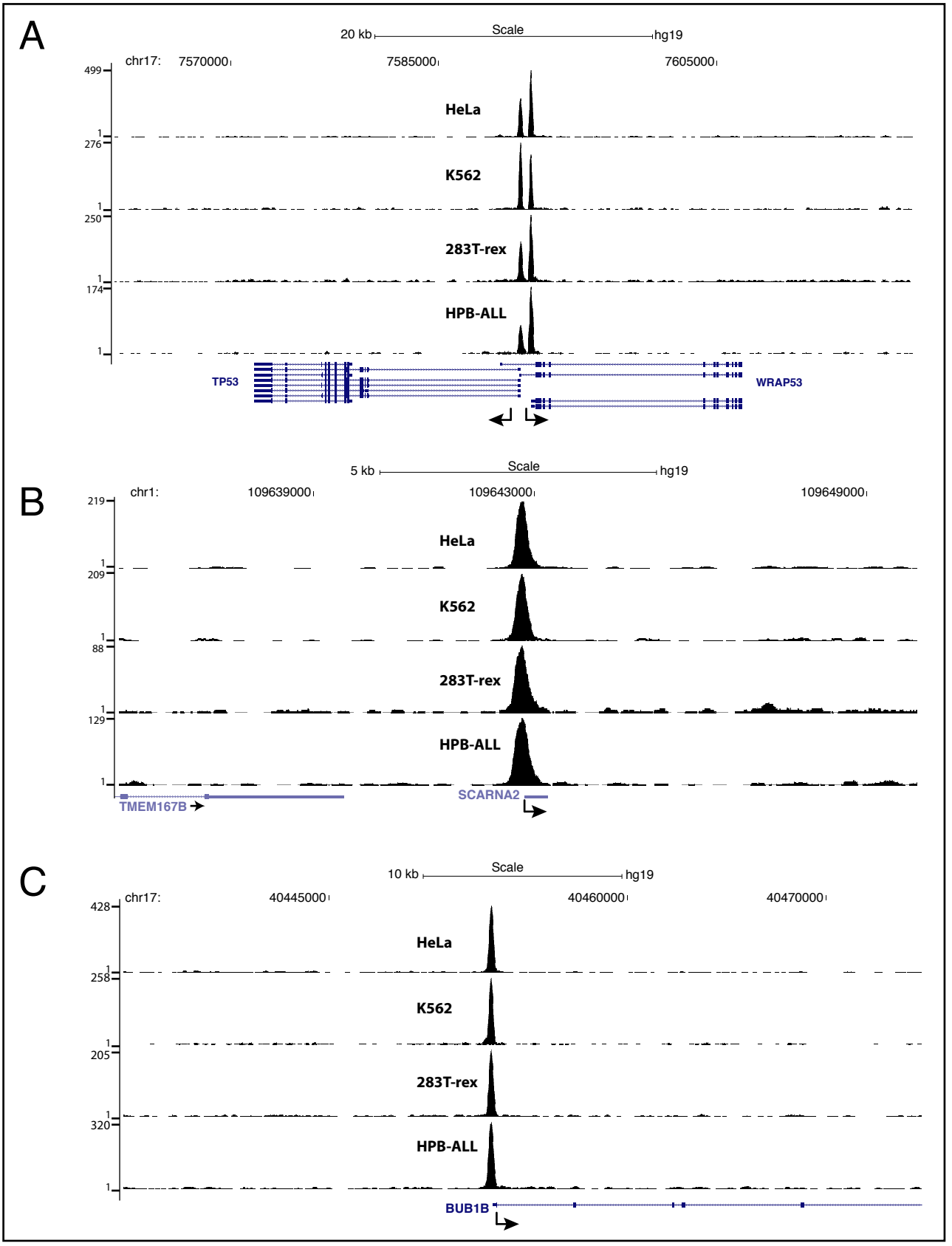

A
20 kb
Scale
hg19
chr17:
7570000
7585000
7605000
499
1
276
HeLa
1
250
K562
1
174
283T-rex
1
HPB-ALL
1
TP53
WRAP53
B
5 kb
Scale
hg19
chr1:
109639000
109643000
109649000
219
1
209
HeLa
1
88
K562
1
129
283T-rex
1
HPB-ALL
1
TMEM167B
SCARNA2
C
10 kb
Scale
hg19
chr17:
40445000
40460000
40470000
428
1
258
HeLa
1
205
K562
1
320
283T-rex
1
HPB-ALL
1
BUB1B

Figure S5

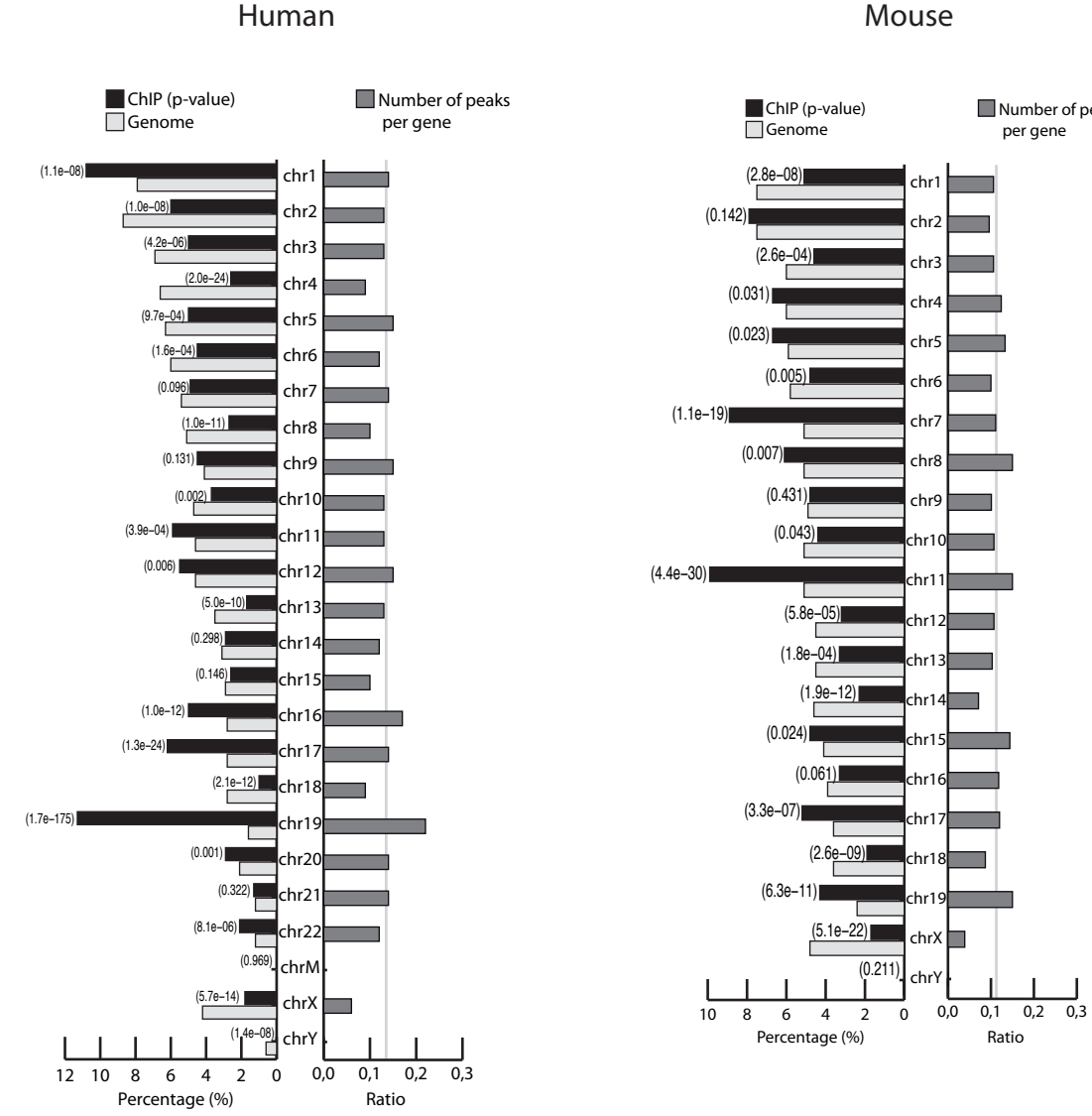

Human
ChIP (p-value)
Genome
Number of peaks per gene
(1.1e−08) chr1
(1.0e−08) chr2
(4.2e−06) chr3
(2.0e−24) chr4
(9.7e−04) chr5
(1.6e−04) chr6
(0.096) chr7
(1.0e−11) chr8
(0.131) chr9
(0.002) chr10
(3.9e−04) chr11
(0.006) chr12
(5.0e−10) chr13
(0.298) chr14
(0.146) chr15
(1.0e−12) chr16
(1.3e−24) chr17
(2.1e−12) chr18
(1.7e−175) chr19
(0.001) chr20
(0.322) chr21
(8.1e−06) chr22
(0.969) chrM
(5.7e−14) chrX
(1.4e−08) chrY
12 10 8 6 4 2 0
Percentage (%)
0,0 0,1 0,2 0,3
Ratio
Mouse
ChIP (p-value)
Genome
Number of peaks per gene
(2.8e−08) chr1
(0.142) chr2
(2.6e−04) chr3
(0.031) chr4
(0.023) chr5
(0.005) chr6
(1.1e−19) chr7
(0.007) chr8
(0.431) chr9
(0.043) chr10
(4.4e−30) chr11
(5.8e−05) chr12
(1.8e−04) chr13
(1.9e−12) chr14
(0.024) chr15
(0.061) chr16
(3.3e−07) chr17
(2.6e−09) chr18
(6.3e−11) chr19
(5.1e−22) chrX
(0.211) chrY
10 8 6 4 2 0
Percentage (%)
0,0 0,1 0,2 0,3
Ratio

Figure S6

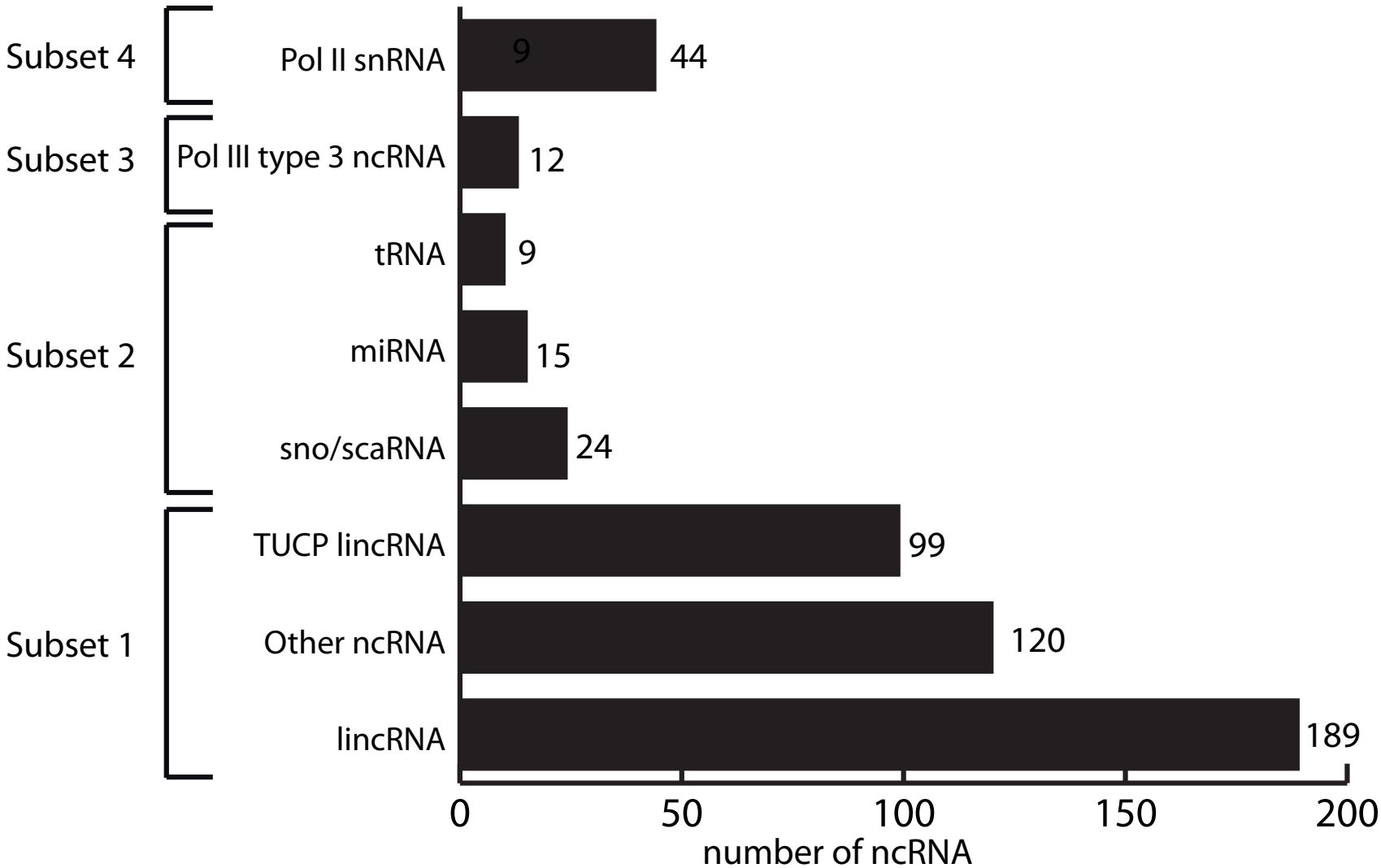

Subset 4
Subset 3
Subset 2
Subset 1
Pol II snRNA
Pol III type 3 ncRNA
tRNA
miRNA
sno/scaRNA
TUCP lincRNA
Other ncRNA
lincRNA
9
44
12
9
15
24
99
120
189
0
50
100
150
200
number of ncRNA

Figure S7

A

Promoters of Pol II snRNAs occupied by ZNF143

	Octamer		SBS1		PSE	
U2 ENST478444	-233	atgcaaat	-209	*gctcaggagctatgggaa*	ttaccgtgacctcaggatga	-44
U2 ENST411315	-222	atgcaaac	-250	*ggccacgctctgtgggaa*	tcaccgcgactttaatatgg	-42
U4B	-227	*attagcat*	-215	ctcccatagtgctttgct	tcacctttgcgaaataggaa	-42
U4C	-229	*atttgcat*	-216	tccccagcgtcccaagcg	tcaccctcaatgtaatggta	-43
U4ATAC	-228	atgcaaat	-206	*cgcgcggtcttctgggta*	ttaccacaaccctaccaggt	-41
U5 ENST362477	-219	*atttgcat*	-212	tacccacaatgcattgcg	gtaccattacccgtttagga	-41
U5 ENST362698	-203	*atttgcat*	-196	tacccataatacactgca	acaccatcagtgtactagga	-41
U5 ENST363286	-212	*atttgcat*	-205	tacccataatgcactgct	tcaccatcactatactagga	-41
U5 ENST364102	-208	*atttgcat*	-201	tacccatagtaccctgca	gcaccatcggcgtactagga	-41
U5 ENST364931	-222	*atttgcat*	-215	tacccacaatgcattgtg	ttaccattagcctgttggga	-42
U5 ENST365574	-216	*atttgcat*	-209	tacccatggtgcaccacg	caaccataagtgtgttaagt	-41
U5 ENST363299	-214	*atttacat*	-207	tacccacaatgcactgca	taaccattgctaatctagta	-41
U5 ENST362507	-226	*atttgcat*	-220	atcccagaatgcattgta	gcaccgtaagtagagggaga	-41
U8	-249	atgtaaat	-203	*atcggagcattctgggaa*	tcaccctagcttgtaacgga	-41
U11	-219	*atttgcat*	-241	ttcccagcaagccctgag	tcacctttaccaaaaaatgg	-43
U12	-205	*atctgcat*	-236	tccccgtgagccttcgcg	tcatcatacgcctaatgaga	-41

Promoters of Pol III type 3 ncRNAs occupied by ZNF143

	Octamer		SBS1		PSE		TATA
7SK	-236	atttagcat	-215	*tgcaaggcattctggata*	tgacc-taagtgtaaagttg	-47	tttatatag-24
tRNASec	-211	atgtaaat	-231	ttcccagaatgcgtggcg	cgaccataactctaaaaggt	-47	cttatatag-24
RPPH1	-97	*atttgcat*	-87	*cgctatgtgttctgggaa*	tcaccataaacgtgaaatgt	-49	tcttataag-25
U6(1)	-221	*atttgcat*	-239	ttcccatgattccttcat	ttaccgtaacttgaaagtat	-46	tttatatat-23
U6(2)	-241	*ttttgcat*	-228	cacccacaatccaccgcg	ttatcctaaccaaaagatga	-46	cttaaaata-24
U6(7)	-221	*atttgcat*	-244	tacccagggtgccccggg	ttaccgtaaggaaaacaaat	-46	cttataaga-22
U6(8)	-222	*atttgcat*	-245	tacccagggtgccccggg	tcaccgtaagtagaataggt	-46	cttataagg-24
U6(9)	-224	*atttgcat*	-199	ttcccagcagccaccgta	tcggcctatgtgtacagaca	-46	ctttaaata-24
Y1	-229	atgctaat	-231	ttcccaggatgcttagca	tcactgtaaggggaaaatga	-47	ctttaaata-25
Y3	-242	atgcaaat	-234	cacccacaatgctttgag	tcaccgtaactatggtagag	-48	cttatataa-24
Y4	-183	*gtttgcat*	-218	ttcccatcatgcaactac	tcatcctaacttatttagag	-44	tttataaaa-23
Y5(hsa-mir-1975)	-201	atgcagaa	-228	tacccacaatgcatttgt	tctccttacctagaaaagac	-45	tataaataa-24

B

Promoters of 27 U1 Pol II snRNA occupied by ZNF143

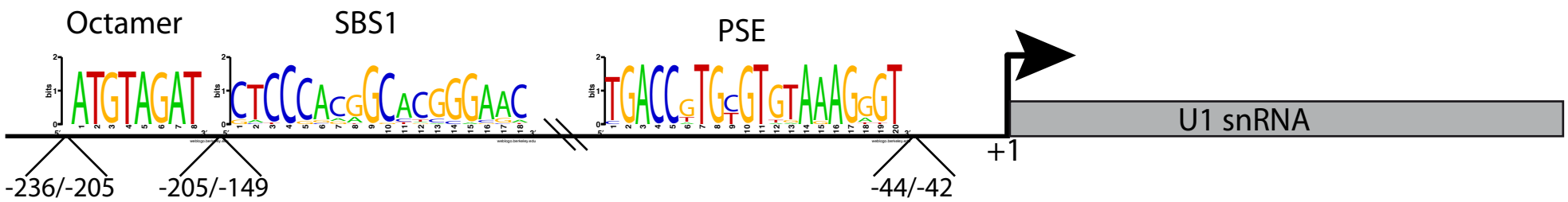


Figure S8

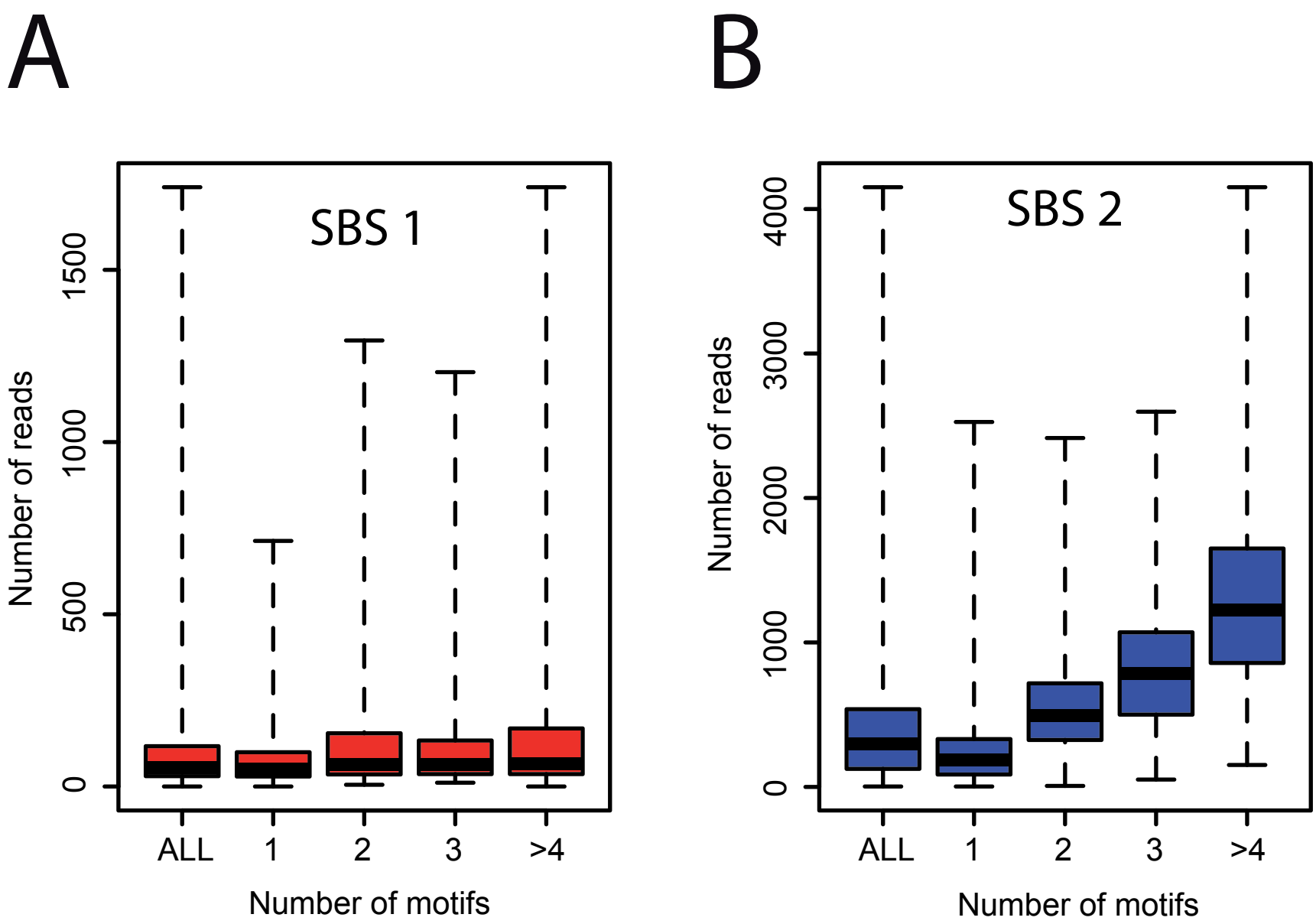
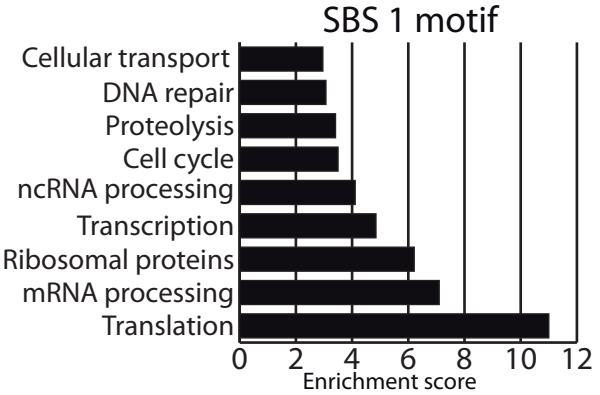
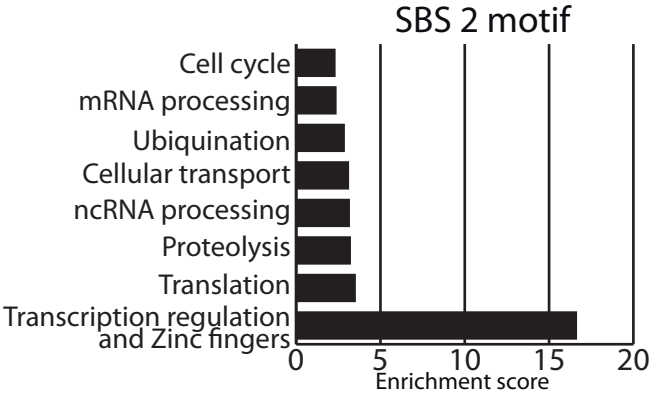
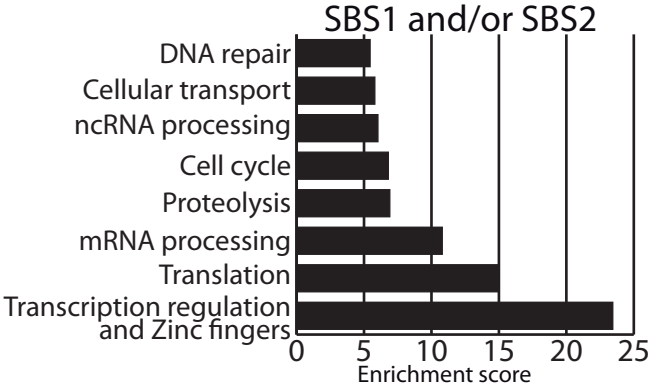

A
SBS 1
Number of reads
0
500
1000
1500
ALL
1
2
3
>4
Number of motifs
B
SBS 2
Number of reads
0
1000
2000
3000
4000
ALL
1
2
3
>4
Number of motifs

Figure S9


SBS 1 motif
Cellular transport
DNA repair
Proteolysis
Cell cycle
ncRNA processing
Transcription
Ribosomal proteins
mRNA processing
Translation
0 2 4 6 8 10 12
Enrichment score


SBS 2 motif
Cell cycle
mRNA processing
Ubiquination
Cellular transport
ncRNA processing
Proteolysis
Translation
Transcription regulation and Zinc fingers
0 5 10 15 20
Enrichment score


SBS1 and/or SBS2
DNA repair
Cellular transport
ncRNA processing
Cell cycle
Proteolysis
mRNA processing
Translation
Transcription regulation and Zinc fingers
0 5 10 15 20 25
Enrichment score

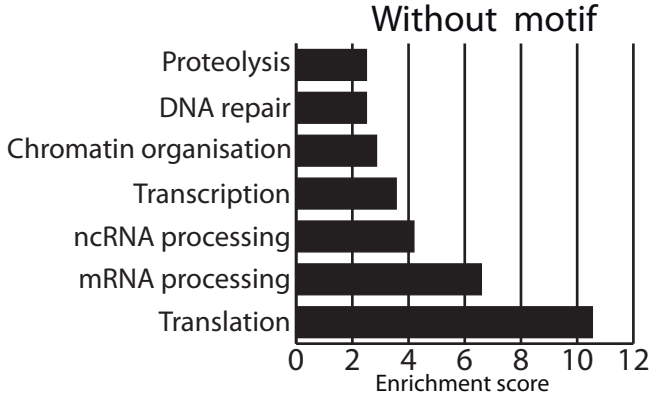
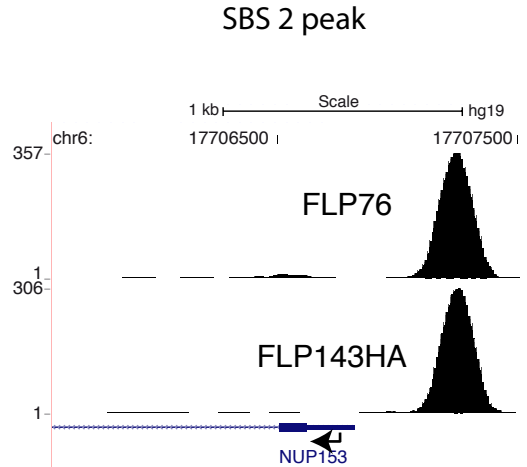

Without motif
Proteolysis
DNA repair
Chromatin organisation
Transcription
ncRNA processing
mRNA processing
Translation
0 2 4 6 8 10 12
Enrichment score

Figure S10


SBS 2 peak
1 kb
Scale
hg19
chr6:
17706500
17707500
357
FLP76
1
306
FLP143HA
1
NUP153

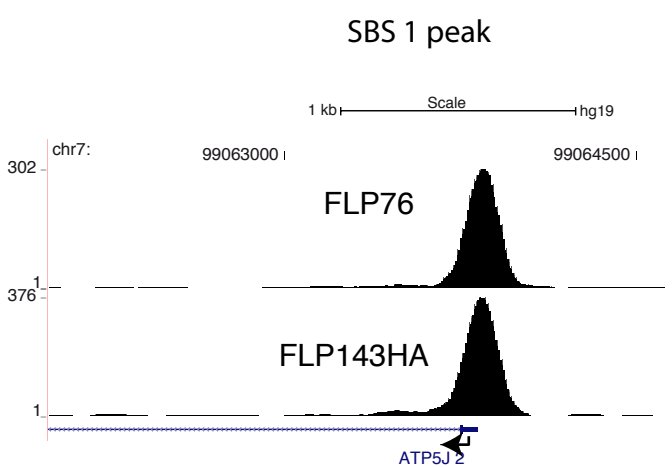
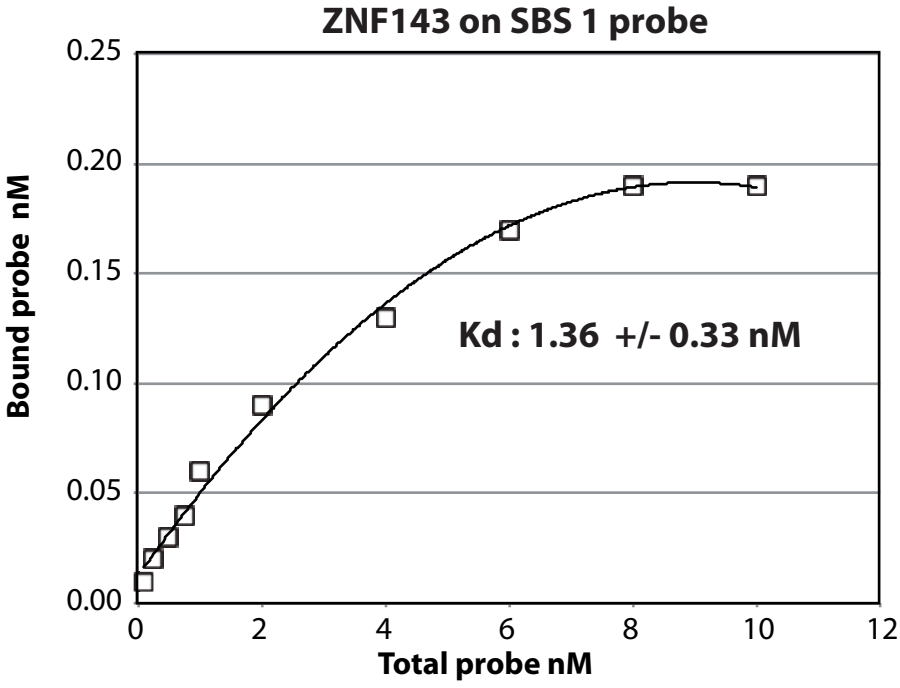

SBS 1 peak
1 kb
Scale
hg19
chr7:
99063000
99064500
302
FLP76
1
376
FLP143HA
1
ATP5J 2

Figure S11


ZNF143 on SBS 1 probe
Bound probe nM
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
Total probe nM
Kd : 1.36 +/- 0.33 nM

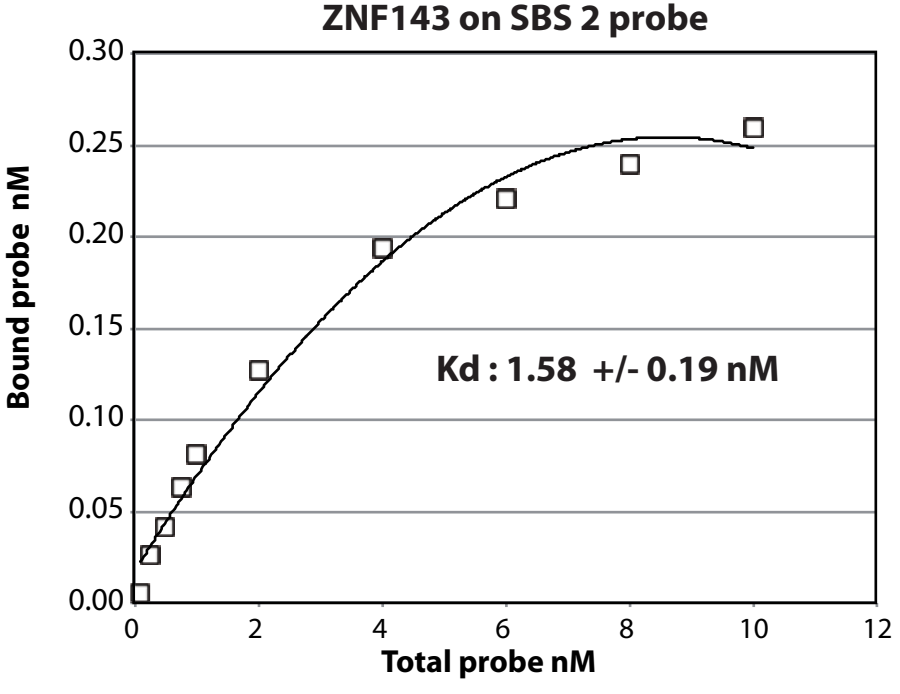
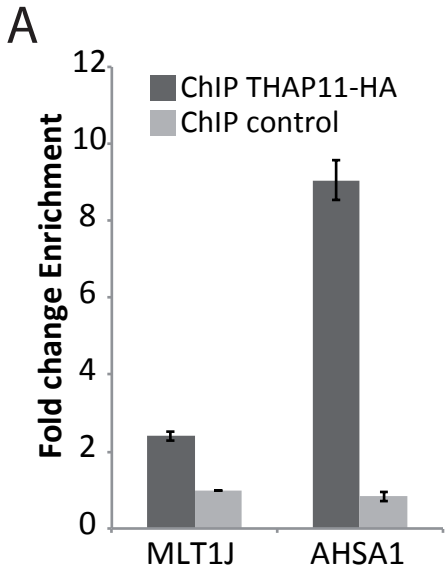

ZNF143 on SBS 2 probe
Bound probe nM
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
Total probe nM
Kd : 1.58 +/- 0.19 nM

Figure S12


A
12
10
8
6
4
2
0
Fold change Enrichment
ChIP THAP11-HA
ChIP control
MLT1J
AHSA1

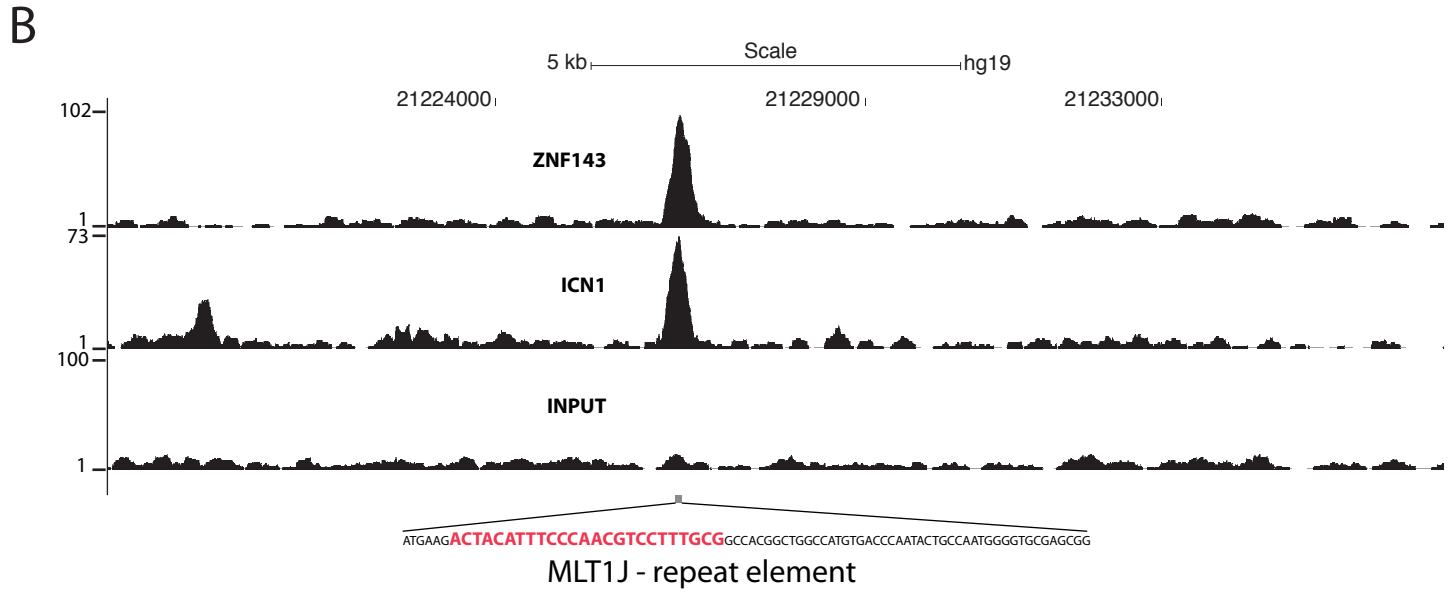

B
5 kb
Scale
hg19
21224000
21229000
21233000
102
1
ZNF143
73
1
ICN1
100
1
INPUT
ATGAAGACTACATTTCCCAACGTCCTTTGCGGCCACGGCTGGCCATGTGACCCAATACTGCCAATGGGGTGCGAGCGG
MLT1J - repeat element

Figure S13

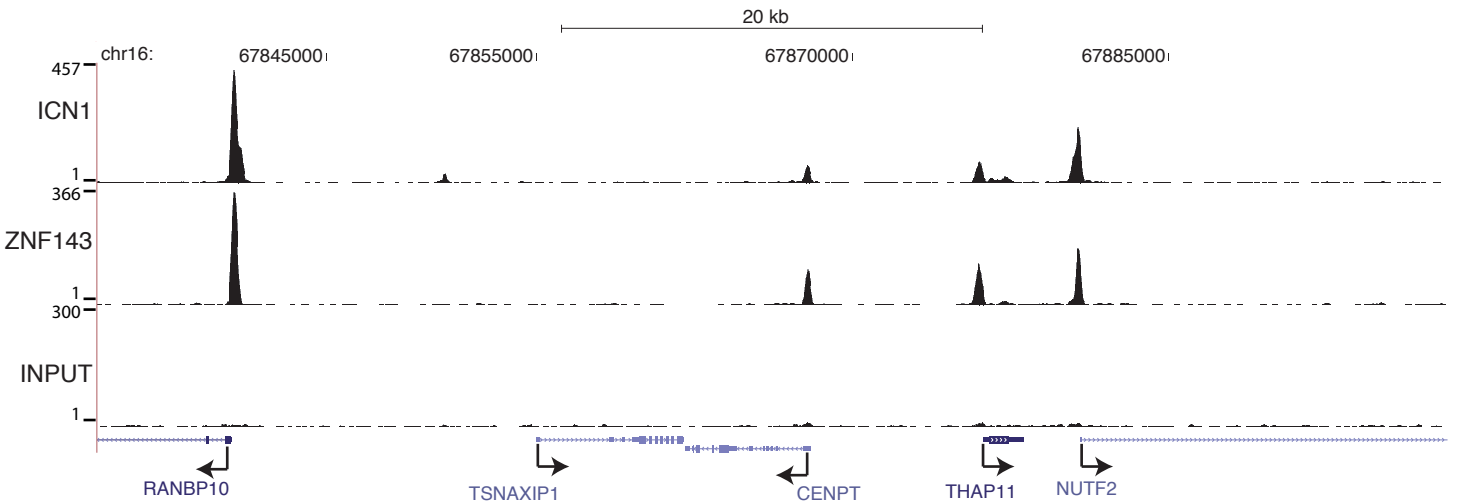
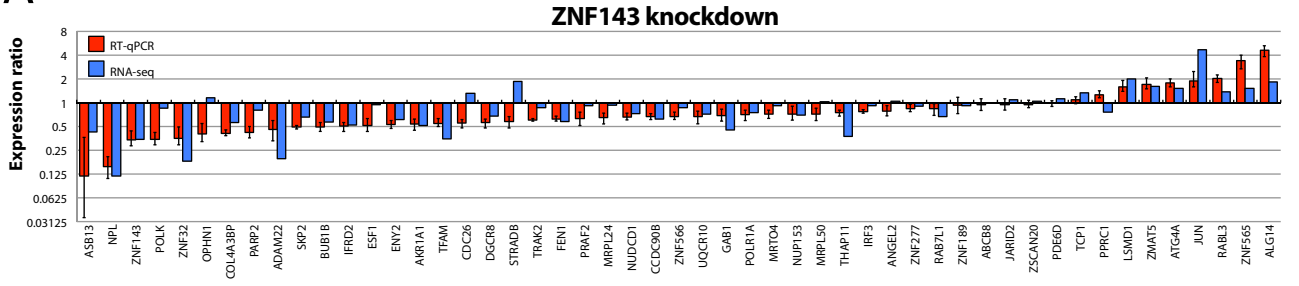

20 kb
chr16:
67845000
67855000
67870000
67885000
457
ICN1
1
366
ZNF143
1
300
INPUT
1
RANBP10
TSNAXIP1
CENPT
THAP11
NUTF2

Figure S14

A



B

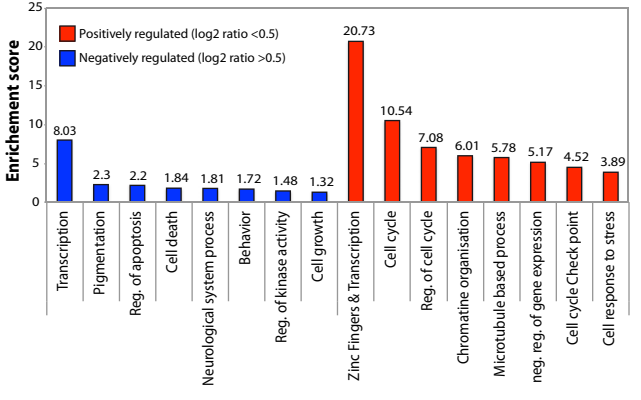


Figure S15

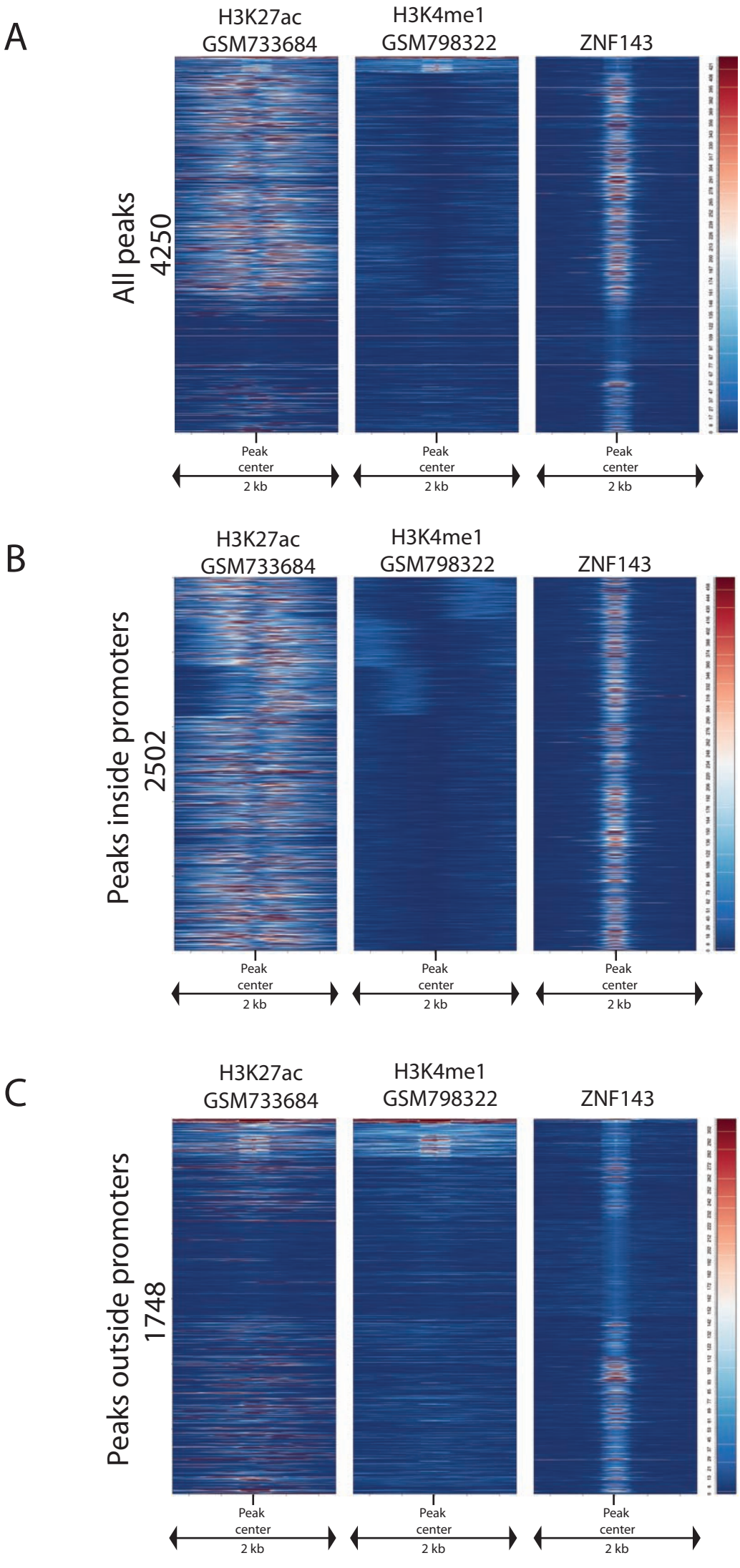
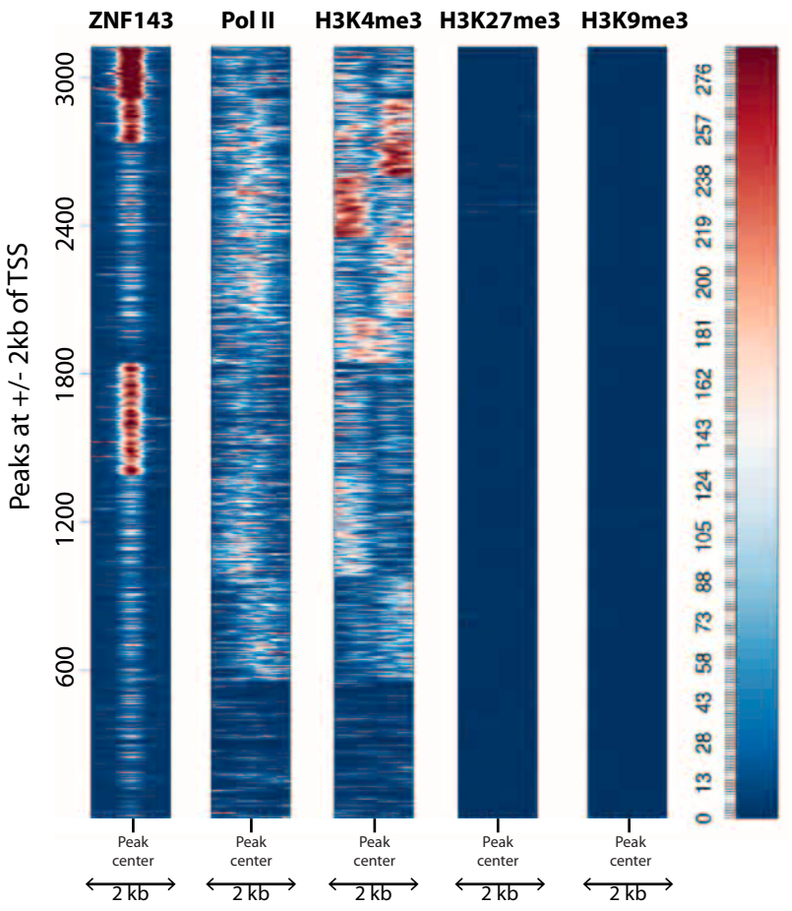
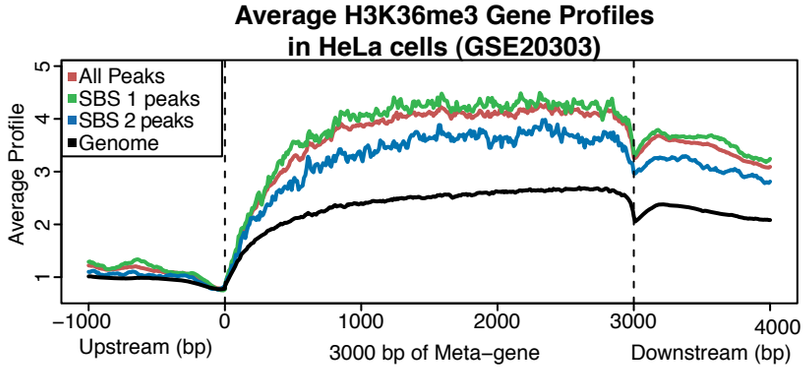

A
H3K27ac
GSM733684
H3K4me1
GSM798322
ZNF143
All peaks
4250
Peak
center
2 kb
B
H3K27ac
GSM733684
H3K4me1
GSM798322
ZNF143
Peaks inside promoters
2502
Peak
center
2 kb
C
H3K27ac
GSM733684
H3K4me1
GSM798322
ZNF143
Peaks outside promoters
1748
Peak
center
2 kb

Figure S16

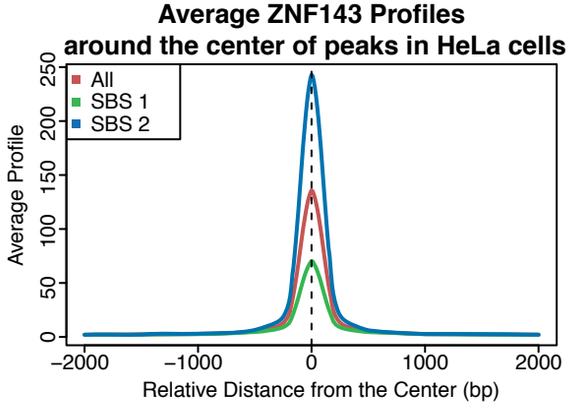
A



B



C



D