BMB Reports - Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-16-156

Title: Epigenetic regulation of long noncoding RNA UCA1 by SATB1 in breast cancer

Article Type: Article

Keywords: SATB1; breast cancer; UCA1; histone methylation; epigenetic regulation

Corresponding Author: Hyoung-Pyo Kim

Authors: Jong-Joo Lee[1], MiKyoung Kim[1], Hyoung-Pyo Kim[1,*]

Institution: [1]Department of Environmental Medical Biology, Institute of Tropical Medicine and [2]Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine,

Manuscript Type: Article

Title: Epigenetic regulation of long noncoding RNA UCA1 by SATB1 in breast cancer

Author's name: Jong-Joo Lee[1,2], Mikyoung Kim[1], and Hyoung-Pyo Kim[1,2,*]

Affiliation: [1]Department of Environmental Medical Biology, Institute of Tropical Medicine, Yonsei University College of Medicine, Seoul, Korea

[2]Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, Seoul, Korea

Running Title: Epigenetic regulation of UCA1 by SATB1

Keywords: *SATB1, breast cancer, UCA1, histone methylation, epigenetic regulation*

[*]**Corresponding author:**

Hyoung-Pyo Kim, Department of Environmental Medical Biology, Institute of Tropical Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea

Tel: +82-2-2228-1842, Fax: +82-2-363-8676, E-mail: kimhp@yuhs.ac

ABSTRACT

Special AT-rich sequence binding protein 1 (SATB1) is a nuclear matrix-associated DNA-binding protein that functions as a chromatin organizer. SATB1 is highly expressed in aggressive breast cancer cells and promotes growth and metastasis by reprograming gene expression. Through genome-wide cross-examination of gene expression and histone methylation, we identified SATB1 target genes for which expression is associated with altered epigenetic marks. Among the identified genes, long noncoding RNA urothelial carcinoma-associated 1 (*UCA1*) was upregulated by SATB1 depletion. Upregulation of *UCA1* coincided with increased H3K4 trimethylation (H3K4me3) levels and decreased H3K27 trimethylation (H3K27me3) levels. Our study showed that SATB1 binds to the upstream region of *UCA1 in vivo*, and that its promoter activity increases with SATB1 depletion. Furthermore, simultaneous depletion of *SATB1* and *UCA1* potentiated suppression of tumor growth and cell survival. Thus, SATB1 repressed the expression of oncogenic *UCA1*, suppressing growth and survival of breast cancer cells.

INTRODUCTION

Breast cancer is the most frequently diagnosed life-threatening cancer in women and the leading cause of cancer death in American women after lung cancer (1). Formation and progression of breast cancer is driven by epigenetic alterations as well as progressive genetic abnormalities (2). Epigenetic alterations involve aberrations such as DNA methylation, histone modifications, and nucleosome remodeling. These epigenetic events modulate chromatin structure, which in turn causes aberrant transcriptional regulation that results in changes in the expression patterns of genes implicated in cellular proliferation, survival, and differentiation (3). Among epigenetic alterations, histone modifications play a fundamental role in carcinogenesis; recent advances in epigenomic analyses show that patterns of histone marks are profoundly altered in cancer cells (3).

Special AT-rich sequence binding protein 1 (SATB1) is a nuclear matrix-associated protein that binds to the ATC-rich DNA sequences of base unpairing regions (BURs) (4). Structurally, SATB1 consists of an N-terminal PDZ-like domain, a C-terminal homeodomain, and tandem CUT domains in the center (5). The PDZ-like domain contributes to the DNA-binding ability of SATB1 through oligomerization and modulates the association of SATB1 with other proteins via post-translational modifications. The CUT domains and the homeodomain mediate the sequence-specific binding of SATB1 to its DNA targets. Through BUR-binding, SATB1 tethers multiple genomic loci to the nuclear matrix to build the appropriate higher-order chromatin structure (6, 7). In addition, SATB1 recruits various transcription factors and chromatin modifying enzymes to regulate gene expression through histone modifications and nucleosomal remodeling at SATB1-bound matrix-associated regions (8). Therefore, SATB1 functions as a chromatin organizer that integrates global

epigenetic and transcriptional programs, which are essential for cellular phenotypes and differentiation.

By controlling the expression of transcription factors such as GATA binding protein 3 (GATA3), Sfpi1, or Nanog, SATB1 functions in thymocyte development, T helper type 2 lineage commitment of naïve T cells, and embryonic stem cell differentiation (9, 10). Interestingly, high expression levels of SATB1 in metastatic breast cancer cells and in aggressive and poorly differentiated breast cancer tissues, as well as non-expression (or non-detection) in normal adjacent tissues, suggest that SATB1 has a role in the metastatic phenotype of breast cancer cells (11). Depletion of SATB1 in aggressive breast cancer cells reverses the metastatic potential, whereas ectopic expression of SATB1 in non-aggressive cells induces metastatic and tumorigenic activities (11). These findings strongly suggest that SATB1 can function as a determinant of breast cancer metastasis. Moreover, a Kaplan-Meier survival analysis revealed a correlation between high expression levels of SATB1 with shorter survival rates in breast cancer patients, and a multivariate analysis confirmed that SATB1 is an independent prognostic marker in patients (12). Additional studies have associated SATB1 with the development of various cancers, including bladder, colorectal, gastric, liver, ovarian, pancreatic, and prostate cancer (13).

Although alterations in SATB1 expression are shown to disrupt the pattern and levels of histone acetylation, which promote tumor growth and metastasis through deregulation of gene expression (11), the effect of SATB1 on chromatin structure in a genome-wide scale, and its correlation with gene expression profiles have not been explored in breast cancer cells. In this study, we used chromatin immunoprecipitation-coupled to deep sequencing (ChIP-seq) to profile genome-wide locations of H3K4 trimethylation (H3K4me3) and H3K27 trimethylation (H3K27me3) epigenetic marks in MDA-MB-231 aggressive breast cells. Furthermore, we identified a list of SATB1 target genes that showed significant alterations on

these histone modifications after SATB1 depletion and examined their correlation with gene expression. Among the identified genes, we revealed that long noncoding RNA urothelial carcinoma-associated 1 (*UCA1*) was upregulated by SATB1 depletion and regulated via an epigenetic mechanism to control growth and apoptosis of aggressive breast cancer cells.

RESULTS

Genome-wide changes in H3K4me3 and H3K27me3 levels by SATB1 depletion

To investigate the role of SATB1 in tumor growth and metastasis in MDA-MB-231 breast cancer cells, we knocked down SATB1 expression by using lentiviral vector-mediated RNA interference of SATB1. Depletion of SATB1 in *SATB1* short hairpin RNA (shRNA) (KD) cells was validated at messenger RNA (mRNA) and protein levels (Figure S1A and S1B). We performed mRNA sequencing (mRNA-seq) using the SATB1-knockdown system and identified genes with significant change in expression levels (fold change > 2, p-value < 0.05) (Figure 1A). Differentially expressed genes (DEG) included 222 downregulated genes and 470 upregulated genes. To validate mRNA-seq data, we measured several altered genes by quantitative real-time polymerase chain reaction (qRT-PCR) (Figure S1C). To assess whether SATB1 can affect genome-wide distribution of histone modification, we performed ChIP-seq using antibodies against H3K4me3, which is present in promoters of actively transcribed genes, and against H3K27me3, which is associated with the repression of transcription. ChIP-seq profiles in both control shRNA (CTL) and *SATB1* shRNA (KD) cells revealed thousands of discrete genomic regions that are enriched with either H3K4me3 or H3K27me3 marks. Consistent with a previous report (14), genomic regions with H3K4me3 or H3K27me3 modifications were identified preferentially at promoter or intergenic regions, respectively (Figure S2A).

The proximal sequences around the transcription start site (TSS) are essential elements of gene regulation (15). Therefore, we examined uniquely mapped tags of H3K4me3 and H3K27me3 at promoter regions, which were defined in this study as 3.0 kb upstream and downstream of the TSS (±3.0 kb around the TSS). According to the estimation provided by mRNA-seq analysis, highly transcribed genes exhibited high H3K4me3 levels but very low

H3K27me3 levels around the TSS. In contrast, silent genes were depleted of H3K4me3 marks and exhibited high levels of H3K27me3 (Figure S2B and S2C). These results confirmed the positive correlation between transcription activity and H3K4me3 levels at promoter regions, and the negative correlation between transcription activity and H3K27me3 levels at promoter regions. To investigate potential changes in H3K4me3 and H3K27me3 levels caused by SATB1 depletion, we compared genome-wide enrichment of H3K4me3 and H3K27me3 at gene promoter regions for control shRNA (CTL) and *SATB1* shRNA (KD) cells. By using the edgeR (empirical analysis of digital gene expression data in R) approach, we identified 2,975 and 1,408 gene promoter regions with differentially higher and lower levels (fold change >1.5), respectively, of H3K4me3 in SATB1-knockdown cells (Figure 1B). We also identified 5,327 and 1,384 gene promoter regions with differentially higher and lower levels of H3K27me3 in SATB1-knockdown cells (fold change >1.5), respectively. (Figure 1C). Next, we focused on DEGs to identify SATB1 target genes for which expressions are regulated by altered histone methylation. Among the 222 genes upregulated by SATB1 depletion, we identified 33 genes with higher H3K4me3 levels, 39 genes with lower H3K27me3 levels, and 6 genes with both higher H3K4me3 and lower H3K27me3 levels (Figure 1D, Table S2). Among the 470 genes downregulated by SATB1 depletion, we found 99 genes with lower H3K4me3 levels, 179 genes with higher H3K27me3 levels, and 45 genes with both lower H3K4me3 and higher H3K27me3 levels (Figure 1E, Table S3).

Repression of long noncoding RNA *UCA1* by SATB1

We identified long noncoding RNA (lncRNA) urothelial carcinoma-associated 1 (*UCA1*) as one of the epigenetically regulated SATB1 target genes (Figure 2A). *UCA1* was originally identified in bladder transitional cell carcinoma (16) and is known to play an important role in the occurrence and development of many tumor and non-tumor diseases (17). Integration

of mRNA-seq and ChIP-seq demonstrated that upregulation of *UCA1* lncRNA after SATB1 depletion coincided with an increased level of H3K4me3 and a decreased level of H3K27me3 at the promoter (Figure 2A). These results were validated independently by performing qRT-PCR (Figure 2B) and ChIP-qPCR (Figure 2C and 2D). An increased level of H3K27 acetylation (H3K27Ac), another active gene regulatory histone modification marker, at UCA1 promoter region was also increased by SATB1 depletion (Figure 2E). However, the enrichment of Ezh2, which catalyzes trimethylation of H3K27, was not changed at the UCA1 promoter region (Figure 2F).

To further understand the molecular basis of SATB1-mediated *UCA1* repression, we examined whether *UCA1* is directly regulated by SATB1. Analysis of ChIP-seq data against SATB1 (unpublished data from our laboratory) revealed prominent SATB1 binding to the promoter and the 3.0-kb upstream region of *UCA1* (Figure 3A). Binding of SATB1 to these regions was independently validated by performing ChIP-qPCR (Figure 3B). Furthermore, SATB1 depletion abrogated the enrichment of SATB1 on the *UCA1* locus, further confirming the validity of the ChIP-seq data (Figure 3B). To test whether the promoter and the 3.0-kb upstream region of *UCA1* contain SATB1 response elements, the regions were separately subcloned into a luciferase reporter construct. Transfection of these luciferase constructs into MDA-MB-231 cells revealed a significant increase of promoter activities by SATB1 depletion (Figure 3C). The data indicated that SATB1 represses *UCA1* expression by directly binding to the *UCA1* promoter and 3.0-kb upstream regions.

In contrast to the findings for SATB1-mediated *UCA1* repression, SATB1 expression remained unaltered in MDA-MB-231 cells expressing shRNA against *UCA1*, which suggests that SATB1 is the upstream regulator of *UCA1* expression (Figure 4A). MDA-MB-231 cells simultaneously infected with lentiviruses expressing shRNA against *SATB1* and *UCA1*, showed a significant decrease of expression for both genes (Figure 4A). In agreement with

previous reports (11, 18), depletion of either *SATB1* or *UCA1* suppressed the growth of MDA-MB-231 cells (Figure 4B), which supports the oncogenic roles of both genes. Interestingly, cell viability was nearly abrogated by simultaneous depletion of *SATB1* and *UCA1* (Figure 4B). In addition, we analyzed apoptosis of MDA-MB-231 cells after knockdown of *SATB1* or *UCA1* expression. Depletion of either *SATB1* or *UCA1* triggered significant apoptosis of MDA-MB-231 cells, and simultaneous depletion of both genes further increased apoptosis (Figure 4C). The data indicated that SATB1 depletion results in upregulation of oncogenic lncRNA *UCA1*, which promotes growth and survival of MDA-MB-231 breast cancer cells.

DISCUSSION

SATB1 is a global chromatin organizer that integrates higher-order chromatin architecture with regulation of gene expression and modulates the accessibility of many gene loci to chromatin remodeling enzymes and transcription factors. Several studies have demonstrated the critical role of SATB1 in cancer in which enhanced expression of SATB1 promotes aberrant growth and metastasis of various epithelial tumors by reprogramming global transcriptional profiles. In this study, we investigated the effect of SATB1 on histone modification status in a genome-wide scale to better understand its epigenetic mechanism in breast cancer. Two histone methylation marks, activating H3K4me3 and silencing H3K27me3, were analyzed by ChIP-seq. We focused on the promoter regions for the enrichment of these epigenetic marks, and thousands of genes were found to have differential levels of H3K4me3 and H3K27me3 after SATB1 depletion. These changes in H3K4me3 and H3K27me3 levels support the suggested role of SATB1 as a global chromatin organizer and epigenetic factor. By integrating chromatin modification (ChIP-seq) with gene expression (mRNA-seq), we identified SATB1 target genes for which expression is associated with histone modifications. Depletion of SATB1 led to upregulation of genes that coincided with higher H3K4me3 levels and lower H3K27me3 levels at promoter regions. In addition, SATB1 depletion resulted in downregulation of genes for which expression was associated with lower H3K4me3 levels and higher H3K27me3 levels. However, many genes differentially expressed by SATB1 depletion did not show significantly altered histone modifications at their promoters. SATB1 may regulate the expression of these genes by modulating chromatin structure at distal regulatory elements other than the promoter region, or by mediating long-range chromatin interaction between such regulatory elements and their cognate promoters.

Here, we revealed that *UCA1* is a novel SATB1 target gene, which is regulated via an epigenetic mechanism in aggressive breast cancer cells. *UCA1* is an lncRNA that is aberrantly expressed in a broad range of cancers and plays oncogenic roles in tumor growth and metastasis. Recent studies show that *UCA1* expression can be upregulated by several transcription factors, including Ets-2, C/EBPα, and HIF-1α (19-21). In addition, transforming growth factor beta (TGF-β) treatment of breast cancer cells is reported to induce *UCA1* expression by recruiting a transcriptional complex composed of TAZ, YAP, TEAD and SMAD2/3 (22). In contrast, repression of *UCA1* expression is mediated in human foreskin fibroblasts by CAPERα/ TBX3 transcriptional complex (23).

Previous studies have focused on the role of SATB1 in mRNA expression but its role in lncRNA transcription has not been studied. This study seems to be the first report describing the function of SATB1 in lncRNA transcription. In addition to UCA1, RNA-seq analysis revealed that expression of other lncRNAs, such as HCP5 and LOC100506844, are positively regulated by SATB1. Further studies will be necessary to dissect how these lncRNAs are regulated by SATB1.

Upregulation of *UCA1* after SATB1 depletion coincided with increased levels of activating histone marks, H3K4me3 and H3K27Ac, and a decreased level of repressive H3K27me3 at the promoter. This result suggests that SATB1 represses *UCA1* expression by closing the chromatin structure at the promoter region of *UCA1*. ChIP assay showed *in vivo* occupancy of SATB1 at the promoter and 3.0-kb upstream regions of *UCA1*. Moreover, transient transfection experiments using luciferase reporter constructs confirmed the repressive activity of these SATB1 binding sites. Because enrichment of the histone methyltransferase EZH2 at the *UCA1* promoter region was not affected by SATB1 depletion, the change in H3K27me3 level at the promoter may not be dependent on EZH2. H3K27 demethylases UTX and JMJD3 has been discovered to actively demethylate H3K27me3 (24), and UTX has been shown to be

part of the H3K4me3 methyltransferase complex (25). Whether SATB1 can affect the recruitment of these H3K27 demethylases and H3K4me3 methyltransferase complex to the UCA1 locus remains to be elucidated.

The role of *UCA1* in cell proliferation and apoptosis has been studied in various cancers. In breast cancer cells, *UCA1* interacts with heterogeneous nuclear ribonucleoprotein I (hnRNP I) and suppresses p27 protein expression, leading to increased cell proliferation (18). In addition, *UCA1* interacts directly with miR-143 and modulates breast cancer cell growth and apoptosis (26). Consistent with those reports, our study indicated that depletion of *UCA1* results in attenuated growth and induced apoptosis in MDA-MB-231 cells. Depletion of SATB1 resulted in induced *UCA1* expression, and simultaneous depletion of SATB1 and *UCA1* potentiated the suppression of growth and survival of breast cancer cells. These effects indicate that upregulated *UCA1* may partially rescue the growth and survival of SATB1-knockdown MDA-MB-231 cells. Our results strongly suggest that targeting SATB1 by itself will not be sufficient to treat breast cancer, but may require combination strategies of anti-SATB1 with gene targets such as *UCA1*.

MATERIALS AND METHODS

Detailed experimental procedures are described in Supplementary Information.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health, Welfare and Family Affairs, Korea (1120370 to H.-P. Kim), and a National Research Foundation (NRF) of Korea grant (MEST; NRF-2011-0030086, 2012M3A9B4028272, and 2016R1A2B4014183 to H.-P. Kim)

DISCLOSURES

The authors declare that they have no conflicts of interest.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D (2011) Global cancer statistics. CA Cancer J Clin 61, 69-90
2. Hanahan D and Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144, 646-674
3. Allis CD and Jenuwein T (2016) The molecular hallmarks of epigenetic control. Nat Rev Genet 17, 487-500
4. Dickinson LA, Joh T, Kohwi Y and Kohwi-Shigematsu T (1992) A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. Cell 70, 631-645
5. Wang Z, Yang X, Chu X et al. (2012) The structural basis for the oligomerization of the N-terminal domain of SATB1. Nucleic Acids Res 40, 4193-4202
6. Kumar PP, Bischof O, Purbey PK et al. (2007) Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol 9, 45-56
7. Cai S, Han HJ and Kohwi-Shigematsu T (2003) Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nat Genet 34, 42-51
8. Yasui D, Miyano M, Cai S, Varga-Weisz P and Kohwi-Shigematsu T (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. Nature 419, 641-645
9. Burute M, Gottimukkala K and Galande S (2012) Chromatin organizer SATB1 is an important determinant of T-cell differentiation. Immunol Cell Biol 90, 852-859
10. Savarese F, Davila A, Nechanitzky R et al. (2009) Satb1 and Satb2 regulate embryonic stem cell differentiation and Nanog expression. Genes Dev 23, 2625-2638

11. Han HJ, Russo J, Kohwi Y and Kohwi-Shigematsu T (2008) SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. Nature 452, 187-193

12. Kohwi-Shigematsu T, Poterlowicz K, Ordinario E, Han HJ, Botchkarev VA and Kohwi Y (2013) Genome organizing function of SATB1 in tumor progression. Semin Cancer Biol 23, 72-79

13. Mir R, Pradhan SJ and Galande S (2012) Chromatin organizer SATB1 as a novel molecular target for cancer therapy. Curr Drug Targets 13, 1603-1615

14. Barski A, Cuddapah S, Cui K et al. (2007) High-resolution profiling of histone methylations in the human genome. Cell 129, 823-837

15. Rosenbloom KR, Dreszer TR, Long JC et al. (2012) ENCODE whole-genome data in the UCSC Genome Browser: update 2012. Nucleic Acids Res 40, D912-917

16. Wang F, Li X, Xie X, Zhao L and Chen W (2008) UCA1, a non-protein-coding RNA up-regulated in bladder carcinoma and embryo, influencing cell growth and promoting invasion. FEBS Lett 582, 1919-1927

17. Xue M, Chen W and Li X (2015) Urothelial cancer associated 1: a long noncoding RNA with a crucial role in cancer. J Cancer Res Clin Oncol

18. Huang J, Zhou N, Watabe K et al. (2014) Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1). Cell Death Dis 5, e1008

19. Wu W, Zhang S, Li X, Xue M, Cao S and Chen W (2013) Ets-2 regulates cell apoptosis via the Akt pathway, through the regulation of urothelial cancer associated 1, a long non-coding RNA, in bladder cancer cells. PLoS One 8, e73920

20. Xue M, Li X, Wu W et al. (2014) Upregulation of long non-coding RNA urothelial carcinoma associated 1 by CCAAT/enhancer binding protein alpha contributes to bladder cancer cell growth and reduced apoptosis. Oncol Rep 31, 1993-2000

21. Xue M, Li X, Li Z and Chen W (2014) Urothelial carcinoma associated 1 is a hypoxia-inducible factor-1alpha-targeted long noncoding RNA that enhances hypoxic bladder cancer cell proliferation, migration, and invasion. Tumour Biol 35, 6901-6912

22. Hiemer SE, Szymaniak AD and Varelas X (2014) The transcriptional regulators TAZ and YAP direct transforming growth factor beta-induced tumorigenic phenotypes in breast cancer cells. J Biol Chem 289, 13461-13474

23. Kumar PP, Emechebe U, Smith R et al. (2014) Coordinated control of senescence by lncRNA and a novel T-box3 co-repressor complex. Elife 3

24. Agger K, Cloos PA, Christensen J et al. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature 449, 731-734

25. Issaeva I, Zonis Y, Rozovskaia T et al. (2007) Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Mol Cell Biol 27, 1889-1903

26. Tuo YL, Li XM and Luo J (2015) Long noncoding RNA UCA1 modulates breast cancer cell growth and apoptosis through decreasing tumor suppressive miR-143. Eur Rev Med Pharmacol Sci 19, 3403-3411

FIGURE LEGENDS

Figure 1. Genome-wide changes in gene expression and histone methylation induced by SATB1 depletion. MDA-MB-231 cells were infected with lentiviruses expressing shRNA against *SATB1* or containing empty pLKO.1 vector; drug-resistant cells were selected. (A) Scatter plot shows differentially expressed genes from control shRNA (CTL) and *SATB1* shRNA (KD) cells. Significantly changed genes (fold change > 2, p-value < 0.05) that are upregulated in *SATB1* shRNA cells (red) or downregulated in *SATB1* shRNA cells (green) are indicated. (B and C) Scatter plots show genes with differential enrichment of H3K4me3 (B) and H3K27me3 (C) within the promoter region (3 kb either side of the TSS) for control shRNA (CTL) and *SATB1* shRNA (KD) cells. Significantly changed genes (fold change > 1.5, p-value < 0.05) that are upregulated in *SATB1* shRNA cells (red) or downregulated in *SATB1* shRNA cells (green) are indicated. (D) Venn diagram shows overlap of SATB1 knockdown-induced genes that were upregulated in RNA expression, had greater H3K4me3 enrichment at the promoter region, and had lesser H3K27me3 enrichment at the promoter region. (E) Venn diagram shows overlap of SATB1 knockdown-induced genes that were downregulated in RNA expression, had lesser enrichment with H3K4me3 at the promoter region, and had greater enrichment with H3K27me3 at the promoter region.

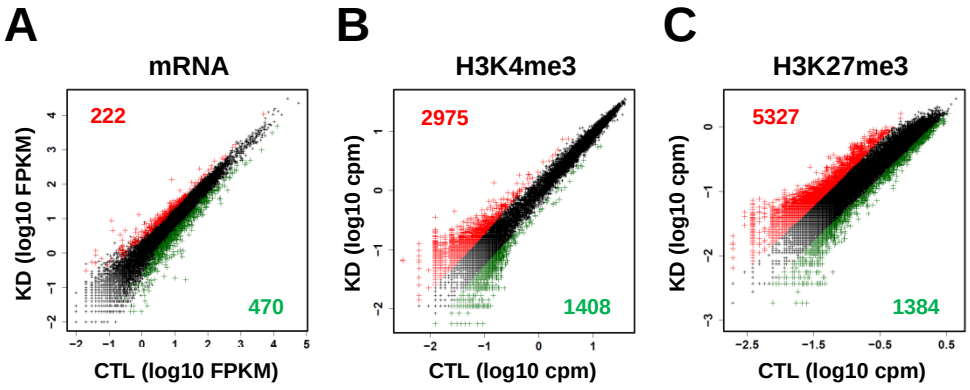
Figure 2. Upregulation of *UCA1* induced by SATB1 depletion coincides with altered histone methylation. (A) Genomic snapshot of the *UCA1* locus. Density of mRNA-seq reads and ChIP-seq reads of H3K4me3 and H3K27me3 in control shRNA (CTL) and *SATB1* shRNA (KD) cells are shown. (B) qRT-PCR validation of *UCA1* expression in control shRNA (CTL) and *SATB1* shRNA (KD) cells (mean ± standard error of the mean [SEM] of three

independent biological replicates) (C-F) ChIP-qPCR validation of H3K4me3 (C), H3K27me3 (D), H3K27Ac (E), and EZH2 (F) levels at the *UCA1* locus.

Figure 3. Direct regulation of *UCA1* by SATB1. (A) Genomic snapshot of the *UCA1* locus showing ChIP-seq reads for SATB1 in MDA-MB-231 cells. (B) ChIP-qPCR validation for SATB1 binding at the *UCA1* locus in control shRNA (CTL) and *SATB1* shRNA (KD) cells. (C) Luciferase constructs containing *UCA1* promoter with or without 3.0-kb upstream region were transfected into control shRNA (CTL) and *SATB1* shRNA (KD) cells. Values are mean ± SEM of results from three experiments.

Figure 4. Effect of SATB1 and *UCA1* on breast cancer cell growth and apoptosis. MDA-MB-231 cells were infected with lentiviruses expressing shRNA against *SATB1* or *UCA1*; drug-resistant cells were selected. (A) qRT-PCR was performed to analyze mRNA levels of *SATB1* and *UCA1* and normalized against *18S* ribosomal RNA (rRNA) levels. (B) Viable cells were counted by Trypan blue staining after depletion of *SATB1* and/or *UCA1*. (C) Flow cytometric analysis of apoptotic cells induced by depletion of *SATB1* and/or *UCA1*. For each sample, the bar indicates standard deviation (SD) from three experiments.

Figure 1


A
mRNA
222
470
KD (log10 FPKM)
CTL (log10 FPKM)
B
H3K4me3
2975
1408
KD (log10 cpm)
CTL (log10 cpm)
C
H3K27me3
5327
1384
KD (log10 cpm)
CTL (log10 cpm)

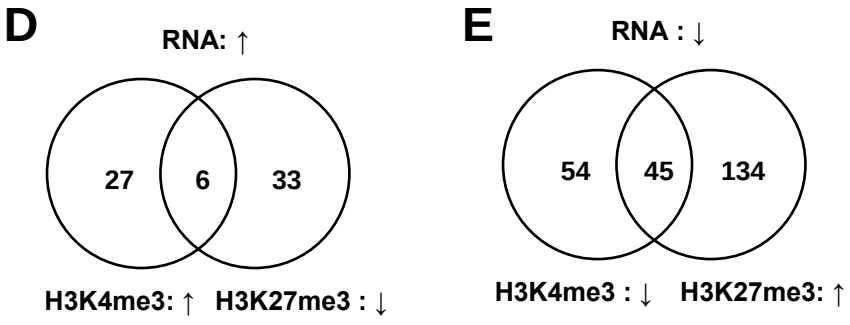

D
RNA: ↑
27
6
33
H3K4me3: ↑
H3K27me3 : ↓
E
RNA : ↓
54
45
134
H3K4me3 : ↓
H3K27me3: ↑

Figure 2

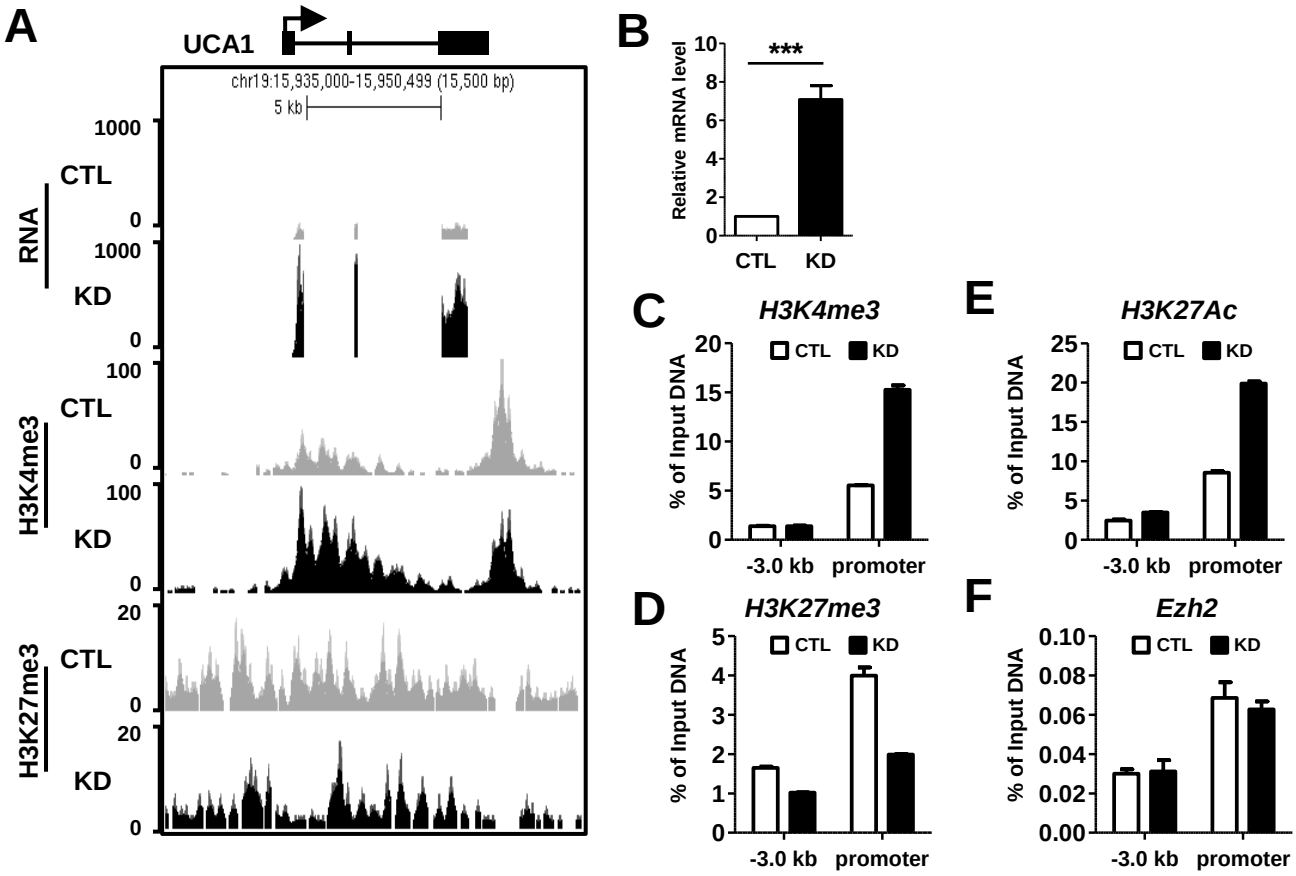

A
UCA1
chr19:15,935,000-15,950,499 (15,500 bp)
5 kb
RNA
CTL
KD
H3K4me3
H3K27me3
1000
100
20
0
B
Relative mRNA level
***
CTL
KD
C
H3K4me3
% of Input DNA
-3.0 kb
promoter
D
H3K27me3
E
H3K27Ac
F
Ezh2

Figure 3

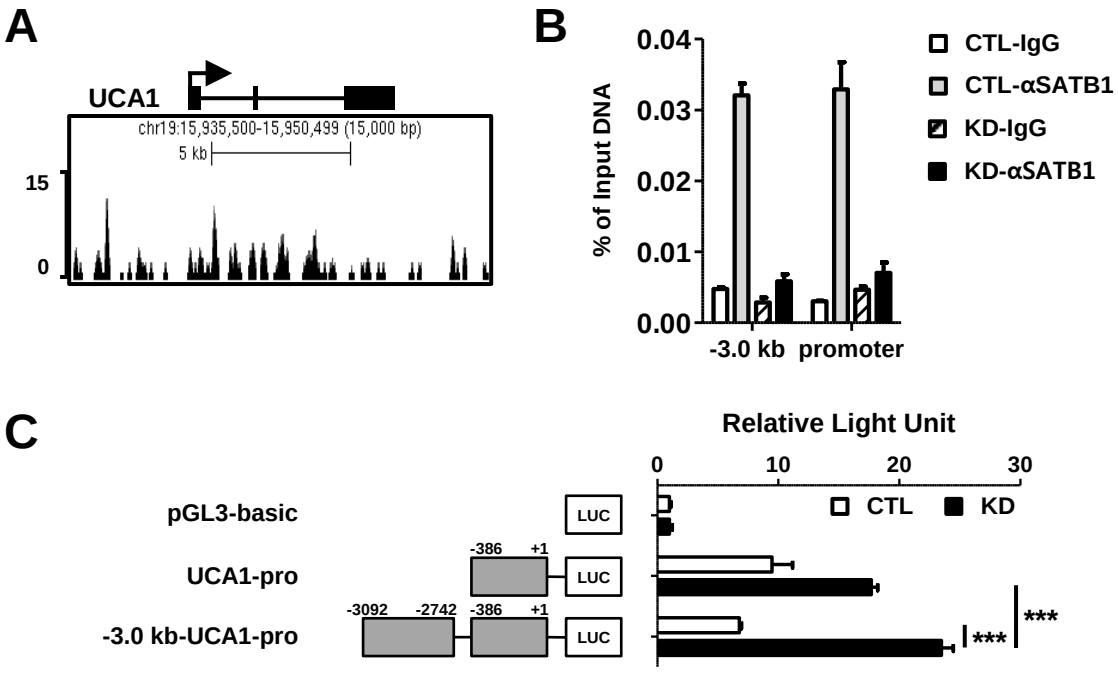
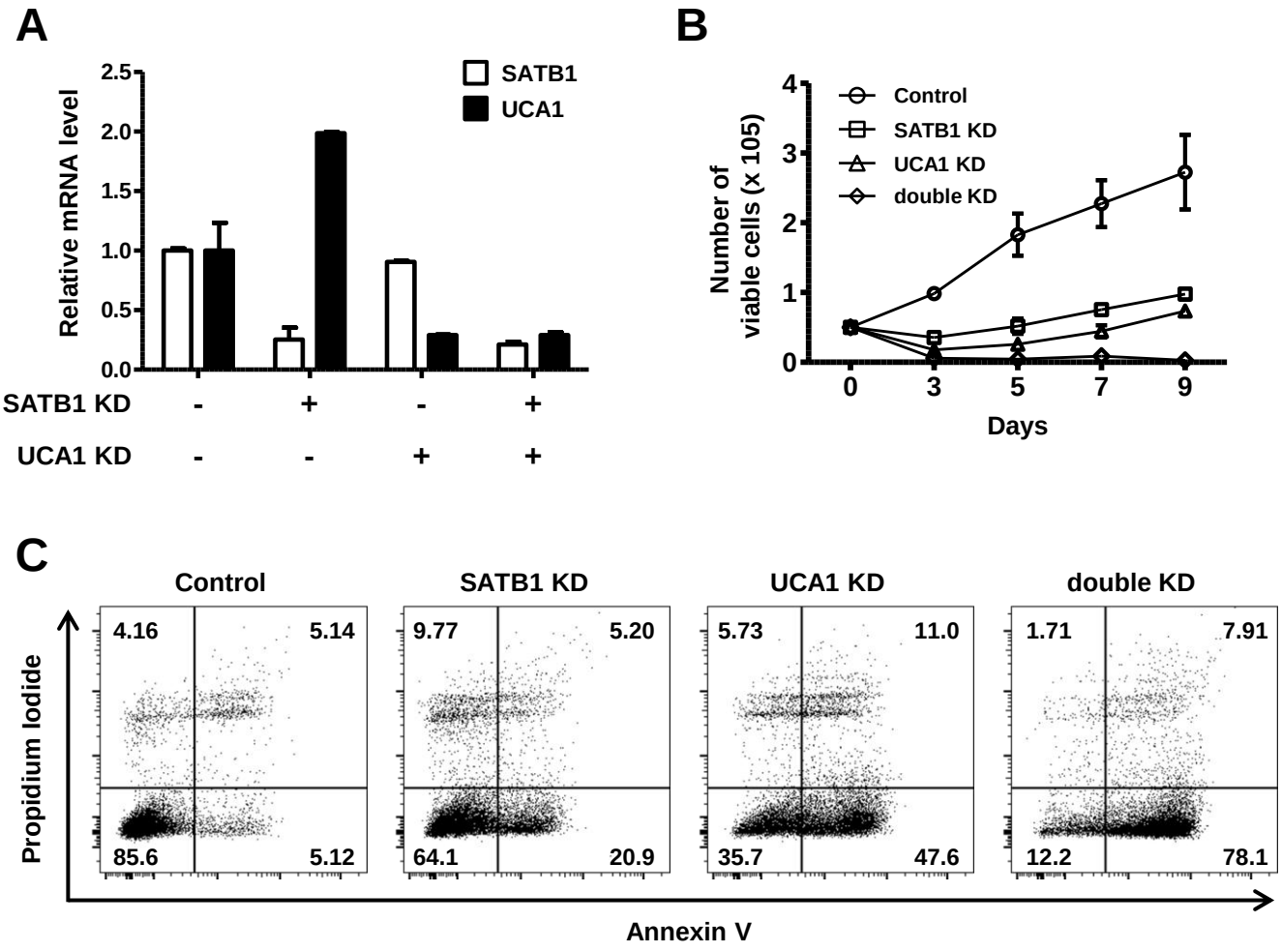

A
UCA1
chr19:15,935,500-15,950,499 (15,000 bp)
5 kb
15
0
B
% of Input DNA
0.04
0.03
0.02
0.01
0.00
-3.0 kb
promoter
CTL-IgG
CTL-αSATB1
KD-IgG
KD-αSATB1
C
Relative Light Unit
0
10
20
30
CTL
KD
pGL3-basic
UCA1-pro
-3.0 kb-UCA1-pro
LUC
-386
+1
-3092
-2742
***
***

Figure 4


A
Relative mRNA level
2.5
2.0
1.5
1.0
0.5
0.0
SATB1
UCA1
SATB1 KD - + - +
UCA1 KD - - + +
B
Number of viable cells (x 105)
4
3
2
1
0
Control
SATB1 KD
UCA1 KD
double KD
0 3 5 7 9
Days
C
Control
SATB1 KD
UCA1 KD
double KD
Propidium Iodide
4.16 5.14 85.6 5.12
9.77 5.20 64.1 20.9
5.73 11.0 35.7 47.6
1.71 7.91 12.2 78.1
Annexin V

SUPPLEMENTARY MATERIALS AND METHODS

Cell culture

MDA-MB-231 breast cancer cells were maintained in Dulbecco's modified Eagle's medium (Welgene, Deagu, Korea) supplemented with 10% heat-inactivated fetal bovine serum (Welgene) and 100 U/ml penicillin/streptomycin (Welgene).

Plasmid constructs

Oligonucleotides (shown in Table S1) for special AT-rich sequence binding protein 1 (*SATB1*) short hairpin RNA (shRNA) or urothelial carcinoma-associated 1 (*UCA1*) shRNA were annealed and subcloned to pLKO.1-neo vector (Addgene, Cambridge, MA, USA) or pLKO.1-puro vector (Sigma-Aldrich, Saint Louis, MO, USA), respectively, after digestion with AgeI and EcoRI. To generate luciferase reporter constructs, *UCA*1 promoter regions were amplified by primers listed in Table S1 and subcloned into pGL3-Basic luciferase reporter vector (Promega, Madison, WI, USA).

Production and transduction of lentivirus

Lentiviruses were produced by transfecting 293FT cells with shRNA lentiviral vectors using polyethylenimines (Polysciences, Warrington, PA, USA). Virus-containing supernatants were collected 48 h after transfection and added to MDA-MB-231 cells with 8 μg/ml polybrene (Sigma-Aldrich). Twenty-four hours after infection, cells were treated with G418 (1 mg/ml) or puromycin (2 μg/ml) for 3 days to eliminate uninfected cells.

Transient transfections and luciferase assays

Transient transfections of MDA-MB-231 cells were performed using polyethylenimines. After 24 h, luciferase activities were analyzed using a luminometer (TD-20/20 luminometer, Turner Designs, Sunnyvale, CA, USA) and a dual luciferase assay system kit (Promega).

Quantitative real-time polymerase chain reaction

Total RNA was isolated using RiboEx column (GeneAll Biotechnology, Seoul, Korea), and first-strand complementary DNAs (cDNAs) were made using the PrimeScript RT Kit (Takara Bio Inc., Shiga, Japan). Quantitative real-time polymerase chain reaction (qRT-PCR) was performed on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) using SYBR® Premix Ex Taq™ II (Takara Bio Inc.). For each sample, duplicate test reactions were analyzed for the expression of the gene of interest, and results were normalized to *18S* ribosomal RNA (rRNA). Primer sequences are provided in Table S1.

Western blot analysis

Cells were lysed in 100 μL cell lysis buffer containing 50 mM Tris-HCl (pH 7.4), 1 mM ethylenediaminetetraacetic acid (EDTA), 0.1% sodium dodecyl sulfate (SDS), 140 mM NaCl, 0.2% deoxycholic acid, 2% triton X-100, and protease inhibitor cocktail (Sigma-Aldrich). Whole cell lysates were resolved by SDS-polyacrylamide gel electrophoresis (PAGE), transferred to a polyvinylidene fluoride membrane, and blocked with 5% skim milk. The membrane was incubated with antibodies against SATB1 (BD Biosciences, San Jose, CA, USA) and β-actin (Sigma-Aldrich), and then with horseradish peroxidase-conjugated secondary antibody. Target proteins were visualized with SuperSignal West Pico Chemiluminescent Substrate (Pierce, Rockford, IL, USA) using an LAS-4000 imaging system (GE Healthcare, Piscataway, NJ, USA).

Cell proliferation assay

To assess cell proliferation, cells were seeded into 24-well plates at a density of 5×10^4 per dish, and grown for 3, 5, 7, and 9 days in standard culture medium. Proliferation of each cell line was measured by Trypan blue (Invitrogen, Carlsbad, CA, USA) exclusion analysis after incubation with 0.05% trypsin-EDTA (Invitrogen) for 5 min.

Quantitation of apoptosis by Annexin V flow cytometry

Cells were washed twice with phosphate buffered saline (PBS) and resuspended in 1× binding buffer (10 mM HEPES, pH 7.4, 140 mM NaCl, and 2.5 mM $CaCl_2$). Afterwards, cells were incubated with 4 μl Annexin V-APC (e-Bioscience, San Diego, CA, USA) and 4 μl of 100 μg/ml propidium iodide (Sigma-Aldrich) for 15 min at room temperature. Cells were analyzed by flow cytometry using a FACSVerse system (BD Biosciences). Data analysis was performed using FlowJo software (Treestar, Ashland, OR, USA).

mRNA-seq and data analysis

The mRNA sequencing (mRNA-seq) library was prepared using a TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA). The library was sequenced using a HiSeq 2000 system (Illumina) to generate 101 bp paired-end reads. Reads were quality-trimmed and filtered using an NGS QC Toolkit v2.3 (1) to remove reads with low-quality bases (quality score > 20). High-quality reads were mapped to human reference (build hg19) genome using RSEM with Bowtie2 v2.0.0-beta7 (2). The expression level of each transcript was quantified as FPKM (fragments per kilobase of exon per million fragments mapped), and the EBSeq package (3) was used to select differentially expressed genes. mRNA-seq data were visualized using the University of California, Santa Cruz Genome Browser (4).

ChIP-seq and data analysis

Chromatin immunoprecipitation (ChIP) assays were performed as previously described (5) using antibodies specific for H3K4me3 (Abcam, Cambridge, UK) and H3K27me3 (Abcam). For ChIP sequencing, 140 bp genomic libraries were generated from the input and chromatin-immunoprecipitated DNA. The libraries were sequenced using a HiSeq 2000 system to generate 101 bp paired-end reads. Reads were quality-trimmed and filtered using an NGS QC Toolkit v2.3 (1) to remove reads with low-quality bases (quality score > 20). High-quality reads were mapped to human reference (build hg19) genome using Bowtie2 v2.0.0-beta7 (6). Read counts were calculated for each gene in the 3.0-kb region surrounding the transcription start site (TSS) using HTSeq v0.6.1 (7) and converted to log2 CPM (counts-per-million) values using a bioconductor package, edgeR v3.6.8 (8). To eliminate biases among libraries, normalization was performed using TMM (trimmed mean of M-values) (9). Peaks were called using HOMER with default options (10). To investigate the genomic distribution of H3K4me3 and H3K27me3 peaks, we categorized the following genomic features: promoter (±3.0 kb around the TSS), exon, intron, and intergenic region, and assigned each peak to one of the categories.

Statistical analysis

Results are expressed as means ± standard deviation (SD). Student's t-test was used for statistical analysis. Relationships were considered statistically significant when the p-value was less than 0.05.

SUPPLEMENTARY REFERENCES

1. Patel, R. K. and Jain, M. (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* **7,** e30619.
2. Li, B. and Dewey, C. N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* **12,** 323.
3. Leng, N., Dawson, J. A., Thomson, J. A., Ruotti, V., Rissman, A. I., Smits, B. M., Haag, J. D., Gould, M. N., Stewart, R. M. and Kendziorski, C. (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29,** 1035-1043.
4. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome research* **12,** 996-1006.
5. Park, J. H., Choi, Y., Song, M. J., Park, K., Lee, J. J. and Kim, H. P. (2016) Dynamic Long-Range Chromatin Interaction Controls Expression of IL-21 in CD4+ T Cells. *Journal of immunology (Baltimore, Md. : 1950)* **196,** 4378-4389.
6. Langmead, B. and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* **9,** 357-359.
7. Anders, S., Pyl, P. T. and Huber, W. (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166-169.
8. Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139-140.
9. Robinson, M. D. and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* **11,** R25.

10. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38,** 576-589.

Table S1. Primers used for plasmid construction, qRT-PCR,and ChIP-qPCR assay.

Primers used for lentiviral vector construction		
shRNA-SATB1		
Forward (5' to 3')	CCGGGGATTTGGAAGAGAGTGTCTTCAAGAGAGACACTCTCTTCCAAATCCTTTTT	
Reverse (5' to 3')	AATTAAAAAGGATTTGGAAGAGAGTGTCTCTCTTGAAGACACTCTCTTCCAAATCC	
shRNA-UCA1		
Forward (5' to 3')	CCGGGAGAGCCGATCAGACAAACAACTCGAGTTGTTTGTCTGATCGGCTCTCTTTTTG	
Reverse (5' to 3')	AATTCAAAAAGAGAGCCGATCAGACAAACAACTCGAGTTGTTTGTCTGATCGGCTCTC	
Primers used for luciferase reporter plasmid construction		
location	Forward (5' to 3')	Reverse (5' to 3')
UCA1 promoter	CAGCTAGCTCTCAGGCTGT	GACTCGAGCAGGGTCGATTG
UCA1 -3.0 kb	TCGGTACCGAAAGTGTTAAACAAAACTT	CTACGCGTTAAGAACTCCTGCTTT
Primers used for qRT-PCR		
Gene	Forward (5' to 3')	Reverse (5' to 3')
GAPDH	GCACCGTCAAGGCTGAGAAC	ATGGTGGTGAAGACGCCAGT
SATB1	GTGGAAGCCTTGGGAATCC	CTGACAGCTCTTCTTCTAGTT
CLDN1	CCCCAGTGGAGGATTTACTCCTA	GCAATGTGCTGCTCAGATTCA
INHBA	GGAGGGCAGAAATGAATGAA	CCTTGGAAATCTCGAAGTGC
ADAMTS6	CCTGGGTTCAGCTTTTACCA	ATGACCTTTCGTCCCACAAG
CLSPN	GAGTCAGAAGCCAGGTGGAG	TGACTGTCCTCCCCAATTTC
S100A4	GATGAGCAACTTGGACAGCAA	CTGGGCTGCTTATCTGGGAAG
TGFB1	CAACAATTCCTGGCGATACCT	GCTAAGGCGAAAGCCCTCAAT
VEGFB	GAGATGTCCCTGGAAGAACACA	GAGTGGGATGGGTGATGTCAG
OASL	TTTCTGAGGCAGGAGCATTT	GCCCACCTTGACTACCTTCA
UCA1	CGGGTAACTCTTACGGTGGA	TGGTCCATTGAGGCTGTAGA
Primers used for ChIP-qPCR assay		
location	Forward (5' to 3')	Reverse (5' to 3')
UCA1 -3.0 kb	ATGAGAGGCCACTGTTTTGG	GAGGATTTGCCTCCTTCCTT
UCA1 promoter	TTCACTTTGGGCAATCTTCC	TTATATTGCGGCCAAGCAGT

Table S2. The list of genes which are potentially activated by SATB1.

NCBI Accession No.	Gene Symbol	mRNA (log2 FC)	H3K4me3 (log2 FC)	H3K27me3 (log2 FC)	Gene description
NM_001202470	RPS10-NUDT3	-11.88	-0.05	0.49	RPS10-NUDT3 readthrough (RPS10-NUDT3)
NR_033351	GOLGA8F	-10.39	0.00	0.53	golgin A8 family, member G (GOLGA8G), non-coding RNA.,
NR_037860	EGFL8	-10.36	0.99	1.01	EGF Like Domain Multiple 8
NM_001163990	RAB37	-10.31	-0.51	-0.76	RAB37, member RAS oncogene family (RAB37), transcript variant 3
NM_012385	NUPR1	-6.89	-0.59	0.21	nuclear protein, transcriptional regulator, 1 (NUPR1), transcript variant 2
NM_152565	ATP6V0D2	-5.75	0.32	0.66	ATPase, H+ transporting, lysosomal 38kDa, V0 subunit d2 (ATP6V0D2)
NM_001216	CA9	-4.42	-1.45	0.04	carbonic anhydrase IX (CA9)
NM_001200047	NMNAT3	-4.23	-0.53	-0.62	nicotinamide nucleotide adenylyltransferase 3 (NMNAT3), nuclear gene encoding mitochondrial protein, transcript variant 2
NM_139027	ADAMTS13	-4.23	-0.40	-0.73	ADAM metallopeptidase with thrombospondin type 1 motif, 13 (ADAMTS13), transcript variant 1
NM_152772	TCP11L2	-3.98	-0.24	0.91	T-Complex 11 Like 2
NM_198213	OASL	-3.88	-0.39	-0.90	2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 2
NR_051988	FBXW10	-3.74	0.37	0.64	F-box and WD repeat domain containing 10 (FBXW10), transcript variant 2
NM_198593	C1QTNF1	-3.61	-1.04	0.27	C1q and tumor necrosis factor related protein 1 (C1QTNF1), transcript variant 1
NM_001352	DBP	-3.54	-0.69	1.75	D site of albumin promoter (albumin D-box) binding protein (DBP)
NM_005252	FOS	-3.41	-0.27	0.63	FBJ murine osteosarcoma viral oncogene homolog (FOS)
NM_001198793	ARPC4-TTLL3	-3.28	-0.15	1.65	ARPC4-TTLL3 readthrough (ARPC4-TTLL3)tubulin tyrosine ligase-like family, member 3 (TTLL3), transcript variant 2, non-coding RNA.,tubulin tyrosine ligase-like family, member 3 (TTLL3), transcript variant 1
NM_203403	LURAP1L	-3.21	-0.33	0.53	leucine rich adaptor protein 1-like (LURAP1L)
NM_002596	CDK18	-3.19	-0.51	-0.96	cyclin-dependent kinase 18 (CDK18), transcript variant 2
NM_006472	TXNIP	-3.18	-0.39	0.88	thioredoxin interacting protein (TXNIP)
NM_031847	MAP2	-3.13	1.44	0.45	microtubule-associated protein 2 (MAP2), transcript variant 5
NR_040662	HCP5	-2.93	-1.75	0.28	HLA complex P5 (non-protein coding) (HCP5), non-coding RNA.,
NM_020682	AS3MT	-2.91	-0.71	1.33	C10orf32-AS3MT readthrough (C10orf32-AS3MT), non-coding RNA.,arsenic (+3 oxidation state) methyltransferase (AS3MT)
NM_003641	IFITM1	-2.74	-0.97	0.47	interferon induced transmembrane protein 1 (IFITM1)
NM_016233	PADI3	-2.74	-0.99	-0.08	peptidyl arginine deiminase, type III (PADI3)
NM_001295	CCR1	-2.71	-0.87	0.55	chemokine (C-C motif) receptor 1 (CCR1)
NM_001198595	STON1	-2.68	-0.39	0.96	stonin 1 (STON1), transcript variant 1
NM_002214	ITGB8	-2.67	-0.60	0.39	integrin, beta 8 (ITGB8)
NM_001217	CA11	-2.67	-0.17	0.79	carbonic anhydrase XI (CA11)
NM_001136019	FCGRT	-2.66	-0.79	0.01	Fc fragment of IgG, receptor, transporter, alpha (FCGRT), transcript variant 2
NM_001081955	RGS9	-2.61	-0.47	-0.05	regulator of G-protein signaling 9 (RGS9), transcript variant 1
NM_001135242	NDRG1	-2.60	-0.51	0.66	N-myc downstream regulated 1 (NDRG1), transcript variant 2
NM_006981	NR4A3	-2.59	-0.19	0.82	nuclear receptor subfamily 4, group A, member 3 (NR4A3), transcript variant 3
NM_002121	HLA-DPB1	-2.55	-1.73	-0.09	major histocompatibility complex, class II, DP beta 1 (HLA-DPB1)
NM_001964	EGR1	-2.54	0.00	0.98	early growth response 1 (EGR1)
NM_019058	DDIT4	-2.54	-0.48	2.53	DNA-damage-inducible transcript 4 (DDIT4)
NM_001199989	RASD1	-2.43	-0.47	0.03	RAS, dexamethasone-induced 1 (RASD1), transcript variant 1

NM_005345	HSPA1A	-2.39	0.40	0.91	heat shock 70kDa protein 1A (HSPA1A)
NM_024677	NSUN7	-2.38	-0.17	0.82	NOP2/Sun domain family, member 7 (NSUN7)
NM_015526	CLIP3	-2.32	0.58	0.58	CAP-GLY domain containing linker protein 3 (CLIP3), transcript variant 2
NM_000213	ITGB4	-2.32	-0.37	1.25	integrin, beta 4 (ITGB4), transcript variant 1
NM_001243799	TSC22D1	-2.28	-0.19	1.09	TSC22 domain family, member 1 (TSC22D1), transcript variant 5
NM_001015053	HDAC5	-2.23	-0.19	0.49	histone deacetylase 5 (HDAC5), transcript variant 1
NM_001164489	ADAM8	-2.22	-1.17	1.03	ADAM metallopeptidase domain 8 (ADAM8), transcript variant 2
NM_177938	P4HTM	-2.20	-0.20	1.72	prolyl 4-hydroxylase, transmembrane (endoplasmic reticulum) (P4HTM), transcript variant 1
NM_016618	KRCC1	-2.20	-0.12	0.66	lysine-rich coiled-coil 1 (KRCC1)
NM_006120	HLA-DMA	-2.18	-0.99	-0.46	major histocompatibility complex, class II, DM alpha (HLA-DMA)
NM_014033	METTL7A	-2.18	-1.10	0.66	methyltransferase like 7A (METTL7A)
NM_020801	ARRDC3	-2.16	-0.14	0.89	arrestin domain containing 3 (ARRDC3)
NM_197974	BTN3A3	-2.15	0.18	0.59	butyrophilin, subfamily 3, member A3 (BTN3A3), transcript variant 2
NM_001178102	LOX	-2.13	-0.49	0.01	lysyl oxidase (LOX), transcript variant 2
NM_007021	C10orf10	-2.13	-1.05	0.74	chromosome 10 open reading frame 10 (C10orf10)
NM_145047	OSCP1	-2.12	-0.18	0.57	organic solute carrier partner 1 (OSCP1), transcript variant 1
NM_001015881	TSC22D3	-2.12	-0.34	0.60	TSC22 domain family, member 3 (TSC22D3), transcript variant 3
NM_002032	FTH1	-2.11	-0.14	1.14	ferritin, heavy polypeptide 1 (FTH1)
NM_001098634	RBM47	-2.10	-0.24	0.41	RNA binding motif protein 47 (RBM47), transcript variant 2
NM_002612	PDK4	-2.10	-0.21	1.10	pyruvate dehydrogenase kinase, isozyme 4 (PDK4), nuclear gene encoding mitochondrial protein
NM_024496	IRF2BPL	-2.09	-0.65	-0.06	interferon regulatory factor 2 binding protein-like (IRF2BPL)
NM_025149	ACSF2	-2.09	-0.38	0.92	Acyl-CoA Synthetase Family Member 2
NM_002118	HLA-DMB	-2.09	-0.05	0.41	major histocompatibility complex, class II, DM beta (HLA-DMB)
NM_022168	IFIH1	-2.06	-0.48	0.94	interferon induced with helicase C domain 1 (IFIH1)
NM_130851	BMP4	-2.06	-0.85	0.54	bone morphogenetic protein 4 (BMP4), transcript variant 1
NM_001253908	AKR1C3	-2.05	-0.14	0.68	aldo-keto reductase family 1, member C3 (AKR1C3), transcript variant 2
NM_198277	SLC37A2	-2.05	-0.42	0.16	solute carrier family 37 (glycerol-3-phosphate transporter), member 2 (SLC37A2), transcript variant 2
NM_021158	TRIB3	-2.03	0.14	1.05	Tribbles Pseudokinase 3
NM_019111	HLA-DRA	-2.03	-0.99	0.21	major histocompatibility complex, class II, DR alpha (HLA-DRA)
NM_000598	IGFBP3	-2.03	-0.52	0.92	insulin-like growth factor binding protein 3 (IGFBP3), transcript variant 2
NM_006065	SIRPB1	-2.02	-1.22	0.17	signal-regulatory protein beta 1 (SIRPB1), transcript variant 1
NM_001897	CSPG4	-2.01	-1.04	-0.90	chondroitin sulfate proteoglycan 4 (CSPG4)
NM_004655	AXIN2	-2.00	-0.32	1.18	axin 2 (AXIN2)
NM_002125	HLA-DRB5	-1.99	0.29	0.65	major histocompatibility complex, class II, DR beta 5 (HLA-DRB5)
NM_001166266	LTBP1	-1.99	-0.59	0.21	latent transforming growth factor beta binding protein 1 (LTBP1), transcript variant 5
NM_005841	SPRY1	-1.98	-0.65	-0.26	sprouty homolog 1, antagonist of FGF signaling (Drosophila) (SPRY1), transcript variant 3
NM_014568	GALNT5	-1.98	-1.07	0.76	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5 (GalNAc-T5) (GALNT5)
NM_005347	HSPA5	-1.97	0.05	0.48	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa) (HSPA5)
NM_001170820	IFITM10	-1.96	-1.19	0.90	interferon induced transmembrane protein 10 (IFITM10)
NM_173642	RIMKLA	-1.95	-0.38	-0.10	ribosomal modification protein rimK-like family member A (RIMKLA)

NM_014067	MACROD1	-1.95	0.16	1.81	MACRO domain containing 1 (MACROD1)
NM_181642	SPINT1	-1.95	-0.25	0.72	serine peptidase inhibitor, Kunitz type 1 (SPINT1), transcript variant 1
NM_021643	TRIB2	-1.94	-0.50	-1.15	tribbles homolog 2 (Drosophila) (TRIB2), transcript variant 1tribbles homolog 2 (Drosophila) (TRIB2), transcript variant 2, non-coding RNA.,
NM_002775	HTRA1	-1.93	-0.46	-0.01	HtrA serine peptidase 1 (HTRA1)
NM_001334	CTSO	-1.93	0.12	1.48	cathepsin O (CTSO)
NM_033215	PPP1R3F	-1.92	-0.17	1.04	protein phosphatase 1, regulatory subunit 3F (PPP1R3F), transcript variant 1
NM_001025158	CD74	-1.91	-1.30	0.87	CD74 molecule, major histocompatibility complex, class II invariant chain (CD74), transcript variant 2
NM_182487	OLFML2A	-1.88	-0.27	0.82	olfactomedin-like 2A (OLFML2A)
NM_005165	ALDOC	-1.87	-0.81	0.67	aldolase C, fructose-bisphosphate (ALDOC)
NM_033260	FOXQ1	-1.86	-0.38	0.53	forkhead box Q1 (FOXQ1)
NM_005346	HSPA1B	-1.86	0.05	1.57	heat shock 70kDa protein 1B (HSPA1B)
NM_001243186	PIM1	-1.85	-0.50	0.18	pim-1 oncogene (PIM1), transcript variant 1
NM_006813	PNRC1	-1.83	-0.24	1.03	proline-rich nuclear receptor coactivator 1 (PNRC1)
NR_037193	SLC3A2	-1.82	0.01	0.64	Solute Carrier Family 3 Member 2
NM_001202522	DDR1	-1.82	-0.43	1.70	discoidin domain receptor tyrosine kinase 1 (DDR1), transcript variant 3
NM_012309	SHANK2	-1.82	0.06	0.63	SH3 and multiple ankyrin repeat domains 2 (SHANK2), transcript variant 2
NM_006382	CDRT1	-1.82	-1.73	0.33	CMT1A duplicated region transcript 1 (CDRT1)
NR_033663	EPOR	-1.81	-0.26	0.59	erythropoietin receptor (EPOR), transcript variant 1
NM_203434	IER5L	-1.81	-0.42	0.01	immediate early response 5-like (IER5L)
NM_001547	IFIT2	-1.80	0.19	1.18	interferon-induced protein with tetratricopeptide repeats 2 (IFIT2)
NM_021724	NR1D1	-1.80	-0.35	1.21	nuclear receptor subfamily 1, group D, member 1 (NR1D1)
NM_152495	CNIH3	-1.79	-0.48	1.47	cornichon homolog 3 (Drosophila) (CNIH3)
NM_002029	FPR1	-1.79	-0.10	1.20	formyl peptide receptor 1 (FPR1), transcript variant 1
NM_002124	HLA-DRB1	-1.78	-1.73	0.54	major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), transcript variant 1
NM_016946	F11R	-1.78	-0.48	0.37	F11 receptor (F11R)
NM_016154	RAB4B	-1.77	-0.03	1.51	RAB4B, member RAS oncogene family (RAB4B)
NM_017709	FAM46C	-1.76	-0.35	0.76	family with sequence similarity 46, member C (FAM46C)
NM_018110	DOK4	-1.76	-0.45	-0.75	docking protein 4 (DOK4)
NR_002578	GAS5	-1.75	-0.19	0.55	growth arrest-specific 5 (non-protein coding) (GAS5), non-coding RNA.
NM_001172660	ZFYVE28	-1.75	-1.27	0.90	zinc finger, FYVE domain containing 28 (ZFYVE28), transcript variant 6
NM_006931	SLC2A3	-1.75	-0.53	-0.38	solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3)
NM_020453	ATP10D	-1.74	-0.21	1.15	ATPase, class V, type 10D (ATP10D)
NM_000104	CYP1B1	-1.74	-0.31	1.01	cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1)
NM_004878	PTGES	-1.74	-0.57	-0.42	prostaglandin E synthase (PTGES)
NR_030775	SHISA4	-1.74	-0.41	-0.10	shisa homolog 4 (Xenopus laevis) (SHISA4), transcript variant 1
NM_183376	ARRDC4	-1.74	-0.22	0.51	arrestin domain containing 4 (ARRDC4)
NM_000203	IDUA	-1.74	-0.12	1.46	iduronidase, alpha-L- (IDUA)
NM_001143778	ASAP3	-1.73	-0.38	0.03	ArfGAP with SH3 domain, ankyrin repeat and PH domain 3 (ASAP3), transcript variant 2
NM_001122769	LCA5	-1.73	-0.11	0.81	Leber congenital amaurosis 5 (LCA5), transcript variant 2

NM_006343	MERTK	-1.72	-0.43	-0.84	c-mer proto-oncogene tyrosine kinase (MERTK)
NM_019554	S100A4	-1.71	-1.38	1.00	S100 calcium binding protein A4 (S100A4), transcript variant 2
NM_001005914	SEMA3B	-1.71	-1.15	1.69	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B (SEMA3B), transcript variant 1
NM_001205254	OCLN	-1.71	-0.28	0.65	occludin (OCLN), transcript variant 1
NM_001252607	THBS3	-1.71	-0.17	0.94	thrombospondin 3 (THBS3), transcript variant 2
NM_001207012	PGF	-1.71	-1.01	-0.48	placental growth factor (PGF), transcript variant 2
NM_001242702	MKX	-1.71	-0.32	0.55	mohawk homeobox (MKX), transcript variant 2
NM_005737	ARL4C	-1.70	-0.71	0.44	ADP Ribosylation Factor Like GTPase 4C
NM_003995	NPR2	-1.67	0.31	1.11	natriuretic peptide receptor B/guanylate cyclase B (atrionatriuretic peptide receptor B) (NPR2)
NM_152288	ORAI3	-1.66	-0.39	0.62	ORAI calcium release-activated calcium modulator 3 (ORAI3)
NM_197958	LARP6	-1.65	0.02	0.65	La ribonucleoprotein domain family, member 6 (LARP6), transcript variant 2
NM_001145524	YPEL3	-1.65	-0.59	0.62	yippee-like 3 (Drosophila) (YPEL3), transcript variant 1
NM_000757	CSF1	-1.65	-0.60	0.30	colony stimulating factor 1 (macrophage) (CSF1), transcript variant 1
NM_181726	ANKRD37	-1.63	-0.22	0.55	ankyrin repeat domain 37 (ANKRD37)
NM_173211	TGIF1	-1.63	-0.35	0.69	TGFB-induced factor homeobox 1 (TGIF1), transcript variant 6
NM_001122	PLIN2	-1.63	-0.22	0.99	Perilipin 2
NM_198282	TMEM173	-1.63	-0.49	0.44	Transmembrane Protein 173
NM_198580	SLC27A1	-1.62	-0.12	0.83	solute carrier family 27 (fatty acid transporter), member 1 (SLC27A1)
NM_152556	C7orf60	-1.62	-0.22	0.97	chromosome 7 open reading frame 60 (C7orf60)
NM_005261	GEM	-1.61	-0.03	0.49	GTP binding protein overexpressed in skeletal muscle (GEM), transcript variant 2
NM_001113755	TYMP	-1.61	-0.03	1.75	thymidine phosphorylase (TYMP), nuclear gene encoding mitochondrial protein, transcript variant 5
NM_001145398	TEF	-1.61	-1.15	0.04	thyrotrophic embryonic factor (TEF), transcript variant 2thyrotrophic embryonic factor (TEF), transcript variant 1
NM_018083	ZNF358	-1.60	-0.77	0.23	zinc finger protein 358 (ZNF358)
NM_000115	EDNRB	-1.60	1.71	0.78	endothelin receptor type B (EDNRB), transcript variant 1
NM_001078650	TMEM134	-1.59	-0.36	0.94	transmembrane protein 134 (TMEM134), transcript variant 5, non-coding RNA.
NM_138393	REEP6	-1.58	-0.36	0.67	receptor accessory protein 6 (REEP6)
NM_001130101	NR1H3	-1.58	-0.28	2.32	nuclear receptor subfamily 1, group H, member 3 (NR1H3), transcript variant 1
NM_003020	SCG5	-1.57	-0.36	1.43	secretogranin V (7B2 protein) (SCG5), transcript variant 2
NM_005978	S100A2	-1.57	0.41	1.04	S100 calcium binding protein A2 (S100A2)
NM_001100176	HOOK2	-1.56	-0.34	1.22	hook homolog 2 (Drosophila) (HOOK2), transcript variant 2
NM_001166103	SPINT2	-1.56	-0.34	2.94	serine peptidase inhibitor, Kunitz type, 2 (SPINT2), transcript variant b
NM_021732	AVPI1	-1.55	-0.64	-2.12	arginine vasopressin-induced 1 (AVPI1)
NM_002133	HMOX1	-1.55	-0.26	0.57	heme oxygenase (decycling) 1 (HMOX1)
NR_038269	LOC100506844	-1.53	-0.16	0.38	uncharacterized LOC100506844 (LOC100506844), non-coding RNA.
NM_005036	PPARA	-1.53	-0.42	0.50	peroxisome proliferator-activated receptor alpha (PPARA), transcript variant 5
NM_006470	TRIM16	-1.53	-0.21	1.18	tripartite motif containing 16 (TRIM16)
NR_003662	RPSAP58	-1.53	-0.32	1.91	Ribosomal Protein SA Pseudogene 58
NR_037948	BSCL2	-1.53	0.10	0.63	Berardinelli-Seip Congenital Lipodystrophy 2 (Seipin)
NM_180991	SLCO4C1	-1.52	-0.38	-0.55	solute carrier organic anion transporter family, member 4C1 (SLCO4C1)
NM_017633	FAM46A	-1.51	-0.19	0.88	family with sequence similarity 46, member A (FAM46A)

NM_001018109	PIR	-1.51	-0.15	0.53	pirin (iron-binding nuclear protein) (PIR), transcript variant 1
NM_005962	MXI1	-1.51	-0.08	0.40	MAX interactor 1, dimerization protein (MXI1), transcript variant 2
NM_004163	RAB27B	-1.49	-0.41	0.96	RAB27B, member RAS oncogene family (RAB27B)
NM_006645	STARD10	-1.49	-0.01	1.55	StAR-related lipid transfer (START) domain containing 10 (STARD10)
NM_001190452	MTRNR2L1	-1.48	0.25	1.80	MT-RNR2-like 1 (MTRNR2L1)
NM_003236	TGFA	-1.48	-0.46	0.14	transforming growth factor, alpha (TGFA), transcript variant 1
NM_198881	TBC1D8B	-1.48	-0.59	-0.30	TBC1 domain family, member 8B (with GRAM domain) (TBC1D8B), transcript variant 1
NM_018004	TMEM45A	-1.47	-0.10	0.63	transmembrane protein 45A (TMEM45A)
NM_006484	DYRK1B	-1.47	-0.54	0.04	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1B (DYRK1B), transcript variant a
NM_002353	TACSTD2	-1.46	-0.22	1.45	tumor-associated calcium signal transducer 2 (TACSTD2)
NM_001127692	PCCA	-1.45	-0.22	1.21	propionyl CoA carboxylase, alpha polypeptide (PCCA), transcript variant 3
NM_018976	SLC38A2	-1.44	-0.08	2.23	solute carrier family 38, member 2 (SLC38A2)
NM_004567	PFKFB4	-1.43	-0.65	1.41	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4 (PFKFB4)
NM_000147	FUCA1	-1.43	-0.45	-0.39	fucosidase, alpha-L- 1, tissue (FUCA1)
NM_020639	RIPK4	-1.43	-0.45	1.30	receptor-interacting serine-threonine kinase 4 (RIPK4)
NM_002970	SAT1	-1.42	0.17	0.56	spermidine/spermine N1-acetyltransferase 1 (SAT1), transcript variant 1spermidine/spermine N1-acetyltransferase 1 (SAT1), transcript variant 2, non-coding RNA.,
NM_000920	PC	-1.40	-0.07	0.39	pyruvate carboxylase (PC), nuclear gene encoding mitochondrial protein, transcript variant 1
NM_001731	BTG1	-1.39	-0.20	1.12	B-cell translocation gene 1, anti-proliferative (BTG1)
NM_003650	CST7	-1.38	-0.59	-0.11	cystatin F (leukocystatin) (CST7)
NM_001025092	MBP	-1.38	1.67	1.75	myelin basic protein (MBP), transcript variant 8
NM_018295	TMEM140	-1.38	0.32	0.67	transmembrane protein 140 (TMEM140)
NM_005384	NFIL3	-1.37	-0.31	0.57	nuclear factor, interleukin 3 regulated (NFIL3)
NM_057159	LPAR1	-1.36	-0.31	0.45	lysophosphatidic acid receptor 1 (LPAR1), transcript variant 1
NM_021109	TMSB4X	-1.35	-0.24	0.56	thymosin beta 4, X-linked (TMSB4X)
NM_032995	ARHGEF4	-1.35	-0.43	0.73	Rho guanine nucleotide exchange factor (GEF) 4 (ARHGEF4), transcript variant 1
NM_001130726	CCDC149	-1.34	-0.10	0.45	coiled-coil domain containing 149 (CCDC149), transcript variant 2
NM_022060	ABHD4	-1.34	-0.24	0.77	abhydrolase domain containing 4 (ABHD4)
NM_001080424	KDM6B	-1.34	-0.17	1.44	lysine (K)-specific demethylase 6B (KDM6B)
NM_199076	CNNM2	-1.34	-0.16	1.77	cyclin M2 (CNNM2), transcript variant 2
NM_020747	ZNF608	-1.34	-0.44	1.61	zinc finger protein 608 (ZNF608)
NM_001017917	CYB561	-1.34	0.20	0.87	cytochrome b561 (CYB561), transcript variant 1
NM_178314	RILPL1	-1.33	-0.05	0.45	Rab interacting lysosomal protein-like 1 (RILPL1)
NM_004614	TK2	-1.32	-0.71	1.82	Thymidine Kinase 2, Mitochondrial
NM_000963	PTGS2	-1.32	-0.42	-0.68	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2)
NM_001040437	C6orf48	-1.32	-0.07	0.87	Chromosome 6 Open Reading Frame 48
NM_001621	AHR	-1.32	-0.25	0.69	aryl hydrocarbon receptor (AHR)
NM_001256718	SNAP91	-1.31	-0.11	0.50	synaptosomal-associated protein, 91kDa (SNAP91), transcript variant 5
NM_001101802	PHF21A	-1.31	-0.24	1.06	PHD finger protein 21A (PHF21A), transcript variant 2
NM_004995	MMP14	-1.30	-0.42	3.75	matrix metallopeptidase 14 (membrane-inserted) (MMP14)

NM_182491	ZFAND2A	-1.30	-0.33	1.00	zinc finger, AN1-type domain 2A (ZFAND2A)
NM_177433	MAGED2	-1.28	-0.49	-0.32	melanoma antigen family D, 2 (MAGED2), transcript variant 2
NM_002588	PCDHGC3	-1.27	0.05	0.81	protocadherin gamma subfamily C, 3 (PCDHGC3), transcript variant 1
NM_005952	MT1X	-1.26	-0.50	-0.58	metallothionein 1X (MT1X)
NM_017885	HCFC1R1	-1.26	-0.20	1.42	Host Cell Factor C1 Regulator 1
NM_001124	ADM	-1.25	-0.45	-1.82	adrenomedullin (ADM)
NM_001145418	TTC28	-1.25	-0.38	0.19	tetratricopeptide repeat domain 28 (TTC28)
NM_001498	GCLC	-1.24	-0.10	0.62	glutamate-cysteine ligase, catalytic subunit (GCLC), transcript variant 1
NM_003485	GPR68	-1.23	0.10	1.42	G protein-coupled receptor 68 (GPR68), transcript variant 2
NM_032883	TOX2	-1.23	-0.80	0.48	TOX high mobility group box family member 2 (TOX2), transcript variant 1
NM_002276	KRT19	-1.23	-0.53	1.82	keratin 19 (KRT19)
NM_003359	UGDH	-1.23	-0.24	1.62	UDP-glucose 6-dehydrogenase (UGDH), transcript variant 1
NM_003299	HSP90B1	-1.22	-0.03	0.57	heat shock protein 90kDa beta (Grp94), member 1 (HSP90B1)
NM_002204	ITGA3	-1.21	-0.31	0.70	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor) (ITGA3), transcript variant a
NM_001145646	APH1B	-1.21	-0.26	1.12	APH1B gamma secretase subunit (APH1B), transcript variant 2
NM_001005336	DNM1	-1.20	-0.29	0.38	Dynamin 1
NM_001243733	VEGFB	-1.20	-0.33	0.69	vascular endothelial growth factor B (VEGFB), transcript variant VEGFB-186
NM_198859	PRICKLE2	-1.20	0.96	2.15	prickle homolog 2 (Drosophila) (PRICKLE2)
NM_007257	PNMA2	-1.19	-0.39	0.64	paraneoplastic Ma antigen 2 (PNMA2)
NM_002332	LRP1	-1.19	-0.41	0.14	low density lipoprotein receptor-related protein 1 (LRP1)
NM_001100829	TMEM170B	-1.18	-0.29	0.80	transmembrane protein 170B (TMEM170B)
NM_022842	CDCP1	-1.18	-0.21	0.79	CUB domain containing protein 1 (CDCP1), transcript variant 2
NM_002056	GFPT1	-1.18	0.12	0.78	glutamine--fructose-6-phosphate transaminase 1 (GFPT1), transcript variant 1
NM_005529	HSPG2	-1.17	-0.24	0.94	heparan sulfate proteoglycan 2 (HSPG2)
NM_012257	HBP1	-1.17	-0.14	1.26	HMG-box transcription factor 1 (HBP1), transcript variant 1
NM_005065	SEL1L	-1.15	-0.34	0.91	sel-1 suppressor of lin-12-like (C. elegans) (SEL1L), transcript variant 1
NM_025139	ARMC9	-1.15	-0.39	0.45	armadillo repeat containing 9 (ARMC9), transcript variant 1
NM_005711	EDIL3	-1.15	-0.36	0.83	EGF-like repeats and discoidin I-like domains 3 (EDIL3)
NM_001142596	P4HA1	-1.14	-0.20	0.60	prolyl 4-hydroxylase, alpha polypeptide I (P4HA1), transcript variant 1
NM_001430	EPAS1	-1.13	-0.04	1.79	endothelial PAS domain protein 1 (EPAS1)
NM_001221	CAMK2D	-1.12	-0.41	1.48	calcium/calmodulin-dependent protein kinase II delta (CAMK2D), transcript variant 2
NM_080491	GAB2	-1.12	-0.22	1.38	GRB2-associated binding protein 2 (GAB2), transcript variant 1
NM_182646	CPEB2	-1.12	-0.25	1.23	cytoplasmic polyadenylation element binding protein 2 (CPEB2), transcript variant F
NM_005723	TSPAN5	-1.12	0.11	1.65	tetraspanin 5 (TSPAN5)
NM_005013	NUCB2	-1.11	-0.14	0.45	nucleobindin 2 (NUCB2)
NM_001423	EMP1	-1.10	-0.46	1.29	epithelial membrane protein 1 (EMP1)
NM_001161779	PDP1	-1.10	-0.12	0.39	pyruvate dehyrogenase phosphatase catalytic subunit 1 (PDP1), nuclear gene encoding mitochondrial protein, transcript variant 5pyruvate dehyrogenase phosphatase catalytic subunit 1 (PDP1), nuclear gene encoding mitochondrial protein, transcript variant 3
NM_014631	SH3PXD2A	-1.09	-0.62	-0.79	SH3 and PX domains 2A (SH3PXD2A)
NM_001127401	YPEL5	-1.08	-0.28	0.51	yippee-like 5 (Drosophila) (YPEL5), transcript variant 2

Table S3. The list of genes which are potentially repressed by SATB1.

NCBI Accession No.	Gene Symbol	mRNA (log2 FC)	H3K4me3 (log2 FC)	H3K27me3 (log2 FC)	Gene description
NM_001031749	LYPD5	10.66	0.01	-0.59	LY6/PLAUR domain containing 5 (LYPD5), transcript variant A
NM_002192	INHBA	6.25	0.31	-1.63	inhibin, beta A (INHBA)
NM_007036	ESM1	5.21	1.01	-1.10	endothelial cell-specific molecule 1 (ESM1), transcript variant 2
NR_002312	RPPH1	4.75	0.41	0.26	ribonuclease P RNA component H1 (RPPH1), RNase P RNA.
NM_000211	ITGB2	3.26	1.47	-1.43	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit) (ITGB2), transcript variant 1
NM_031246	PSG2	2.93	3.19	-0.48	pregnancy specific beta-1-glycoprotein 2 (PSG2)
NM_001005752	GJB3	2.90	0.47	0.20	gap junction protein, beta 3, 31kDa (GJB3), transcript variant 1
NM_024786	ZDHHC11	2.47	-1.00	-0.45	zinc finger, DHHC-type containing 11 (ZDHHC11)
NM_001267048	KIAA1324	2.43	0.84	0.12	KIAA1324 (KIAA1324), transcript variant 1
NM_004591	CCL20	2.38	1.53	0.22	chemokine (C-C motif) ligand 20 (CCL20), transcript variant 2
NM_001024858	SPTB	2.34	0.07	-0.91	spectrin, beta, erythrocytic (SPTB), transcript variant 2
NR_015379	UCA1	2.33	1.42	-0.54	urothelial cancer associated 1 (non-protein coding) (UCA1), non-coding RNA.
NM_014365	HSPB8	2.29	1.36	-0.75	heat shock 22kDa protein 8 (HSPB8)
NM_001025266	C3orf70	2.25	-0.22	-0.64	chromosome 3 open reading frame 70 (C3orf70)
NM_001511	CXCL1	2.22	-0.18	-0.89	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) (CXCL1), transcript variant 1
NM_001178129	SEMA3E	2.07	1.44	0.22	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E (SEMA3E), transcript variant 2
NM_033066	MPP4	2.07	1.18	0.58	membrane protein, palmitoylated 4 (MAGUK p55 subfamily member 4) (MPP4)
NM_022341	PDF	2.00	0.13	-1.38	peptide deformylase (mitochondrial) (PDF), nuclear gene encoding mitochondrial protein
NM_001124758	SPNS2	1.98	0.31	-0.96	spinster homolog 2 (Drosophila) (SPNS2)
NR_002147	MYH16	1.98	0.75	1.44	myosin, heavy chain 16 pseudogene (MYH16), non-coding RNA.
NM_002564	P2RY2	1.95	0.42	0.52	purinergic receptor P2Y, G-protein coupled, 2 (P2RY2), transcript variant 1
NM_000494	COL17A1	1.91	3.01	-0.26	collagen, type XVII, alpha 1 (COL17A1)
NM_020814	04-Mar	1.90	0.91	1.07	membrane-associated ring finger (C3HC4) 4, E3 ubiquitin protein ligase (MARCH4)
NM_017671	FERMT1	1.89	0.24	-0.62	fermitin family member 1 (FERMT1)
NM_000393	COL5A2	1.82	2.18	0.41	collagen, type V, alpha 2 (COL5A2)
NM_014677	RIMS2	1.81	-0.73	-0.51	regulating synaptic membrane exocytosis 2 (RIMS2), transcript variant 2
NM_001165979	PLCE1	1.81	1.92	0.26	phospholipase C, epsilon 1 (PLCE1), transcript variant 2
NM_014317	PDSS1	1.79	-0.20	-0.51	prenyl (decaprenyl) diphosphate synthase, subunit 1 (PDSS1)
NM_024758	AGMAT	1.78	0.36	-0.87	agmatine ureohydrolase (agmatinase) (AGMAT)
NM_001789	CDC25A	1.73	0.04	-1.04	cell division cycle 25A (CDC25A), transcript variant 1
NM_001168319	EDN1	1.73	0.14	-1.20	endothelin 1 (EDN1), transcript variant 1
NR_026880	MGC12916	1.71	-0.06	-1.02	uncharacterized protein MGC12916 (MGC12916), non-coding RNA.
NM_178232	HAPLN3	1.66	-0.18	-0.68	hyaluronan and proteoglycan link protein 3 (HAPLN3)
NM_152765	C8orf46	1.65	0.54	0.72	chromosome 8 open reading frame 46 (C8orf46)
NM_031476	CRISPLD2	1.64	0.08	-0.52	cysteine-rich secretory protein LCCL domain containing 2 (CRISPLD2)
NM_031464	RPS6KL1	1.60	0.19	-0.50	ribosomal protein S6 kinase-like 1 (RPS6KL1)
NR_024374	LOC642846	1.58	0.44	-0.27	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11-like (LOC642846), non-coding RNA.

NM_002185	IL7R	1.57	1.70	0.88	interleukin 7 receptor (IL7R)
NM_173529	C18orf54	1.56	-0.44	-0.58	Chromosome 18 Open Reading Frame 54
NM_001126122	SLC25A19	1.56	-0.14	-0.52	solute carrier family 25 (mitochondrial thiamine pyrophosphate carrier), member 19 (SLC25A19), nuclear gene encoding mitochondrial protein, transcript variant 2solute carrier family 25 (mitochondrial thiamine pyrophosphate carrier), member 19 (SLC25A19), nuclear gene encoding mitochondrial protein, transcript variant 3
NM_002016	FLG	1.54	2.18	0.41	filaggrin (FLG)
NM_001144935	FGF1	1.53	0.48	0.15	fibroblast growth factor 1 (acidic) (FGF1), transcript variant 9
NM_014900	COBLL1	1.52	-0.20	-0.85	cordon-bleu WH2 repeat protein-like 1 (COBLL1), transcript variant 2cordon-bleu WH2 repeat protein-like 1 (COBLL1), transcript variant 1cordon-bleu WH2 repeat protein-like 1 (COBLL1), transcript variant 3cordon-bleu WH2 repeat protein-like 1 (COBLL1), transcript variant 4
NM_001018113	FANCB	1.51	-0.05	-1.06	Fanconi anemia, complementation group B (FANCB), transcript variant 2
NM_152680	TMEM154	1.49	0.82	-0.03	transmembrane protein 154 (TMEM154)
NM_000999	RPL38	1.49	-0.11	-0.80	ribosomal protein L38 (RPL38), transcript variant 1
NM_001079821	NLRP3	1.48	0.39	0.41	NLR family, pyrin domain containing 3 (NLRP3), transcript variant 1
NM_001080471	PEAR1	1.45	0.66	-0.54	platelet endothelial aggregation receptor 1 (PEAR1)
NM_012099	CD3EAP	1.44	-0.15	-0.40	CD3e Molecule Associated Protein
NM_178550	C1orf110	1.39	0.50	0.56	chromosome 1 open reading frame 110 (C1orf110)
NM_014503	UTP20	1.38	-0.32	-0.81	UTP20, small subunit (SSU) processome component, homolog (yeast) (UTP20)
NM_002053	GBP1	1.33	0.42	1.07	guanylate binding protein 1, interferon-inducible (GBP1)
NM_001846	COL4A2	1.31	0.01	-0.49	collagen, type IV, alpha 2 (COL4A2)
NM_001173465	KIF21A	1.31	-0.02	-1.43	kinesin family member 21A (KIF21A), transcript variant 4
NM_001173517	MAP7D3	1.29	-0.10	-0.88	MAP7 domain containing 3 (MAP7D3), transcript variant 1
NM_001172309	NEXN	1.26	0.18	-0.55	nexilin (F actin binding protein) (NEXN), transcript variant 2
NR_027350	MIR17HG	1.25	-0.04	-1.03	miR-17-92 cluster host gene (non-protein coding) (MIR17HG), transcript variant 1, non-coding RNA.
NM_020297	ABCC9	1.24	1.18	-0.33	ATP-binding cassette, sub-family C (CFTR/MRP), member 9 (ABCC9), transcript variant SUR2B
NM_004666	VNN1	1.24	1.06	0.14	vanin 1 (VNN1)
NM_024577	SH3TC2	1.22	0.54	-0.12	SH3 domain and tetratricopeptide repeats 2 (SH3TC2)
NM_022164	TINAGL1	1.22	0.27	-0.91	tubulointerstitial nephritis antigen-like 1 (TINAGL1), transcript variant 1
NM_005562	LAMC2	1.20	0.56	0.59	laminin, gamma 2 (LAMC2), transcript variant 1
NM_021128	POLR2L	1.19	-0.17	-1.47	polymerase (RNA) II (DNA directed) polypeptide L, 7.6kDa (POLR2L)
NM_003487	TAF15	1.18	0.01	-0.54	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa (TAF15), transcript variant 1
NM_001821	CHML	1.18	0.55	0.34	choroideremia-like (Rab escort protein 2) (CHML)
NM_001082959	SCARB1	1.14	0.56	-0.15	scavenger receptor class B, member 1 (SCARB1), transcript variant 1scavenger receptor class B, member 1 (SCARB1), transcript variant 2scavenger receptor class B, member 1 (SCARB1), transcript variant 1

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Knockdown of SATB1 expression in MDA-MB-231 cells and validation of mRNA-seq. (A) qRT-PCR was performed to analyze mRNA levels of *SATB1* in control shRNA (CTL) and *SATB1* shRNA (KD) cells and normalized against *18S* rRNA levels. (B) Cell lysates were separated by SDS-PAGE and subjected to western blotting using antibodies against SATB1 or β-actin. (C) qRT-PCR confirmed expression levels of several differentially expressed genes that were identified by mRNA-seq. Error bars indicate SD from three experiments.

Figure S2. Genome-wide distribution of H3K4me3 and H3K27me3 marks and correlation with gene expression. (A) Pie chart shows the distribution of H3K4me3 and H3K27me3 marks in control shRNA (CTL) and *SATB1* shRNA (KD) cells for different gene features. (B and C) Genes were categorized based on expression levels observed in mRNA-seq analysis; enrichments of H3K4me3 (B) and H3K27me3 (C) within the promoter region (3 kb either side of the TSS) were collected at corresponding genes, and box plots were plotted for each category.

Figure S1

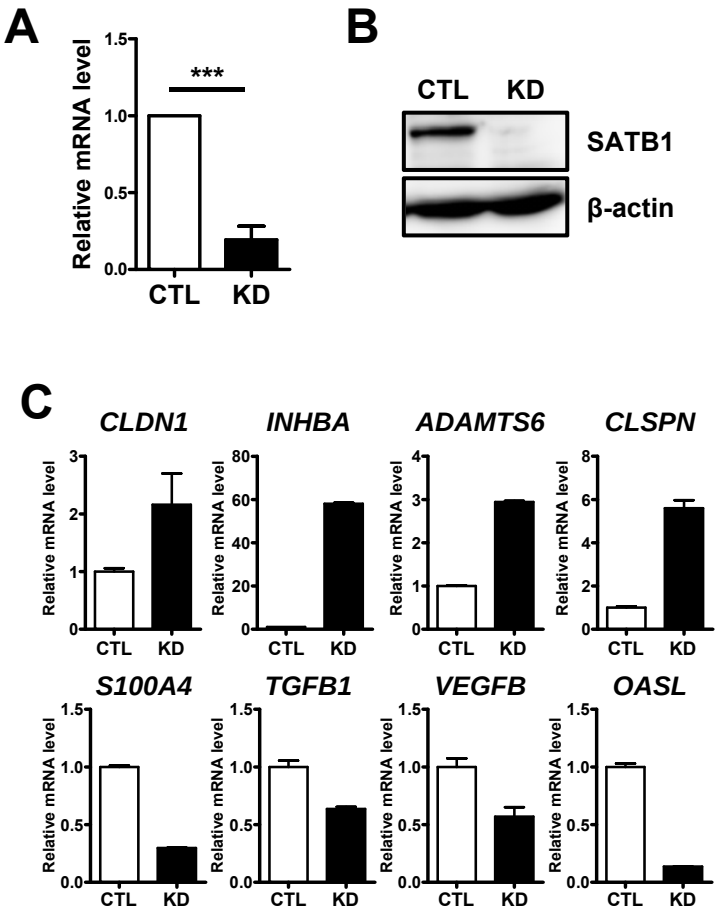
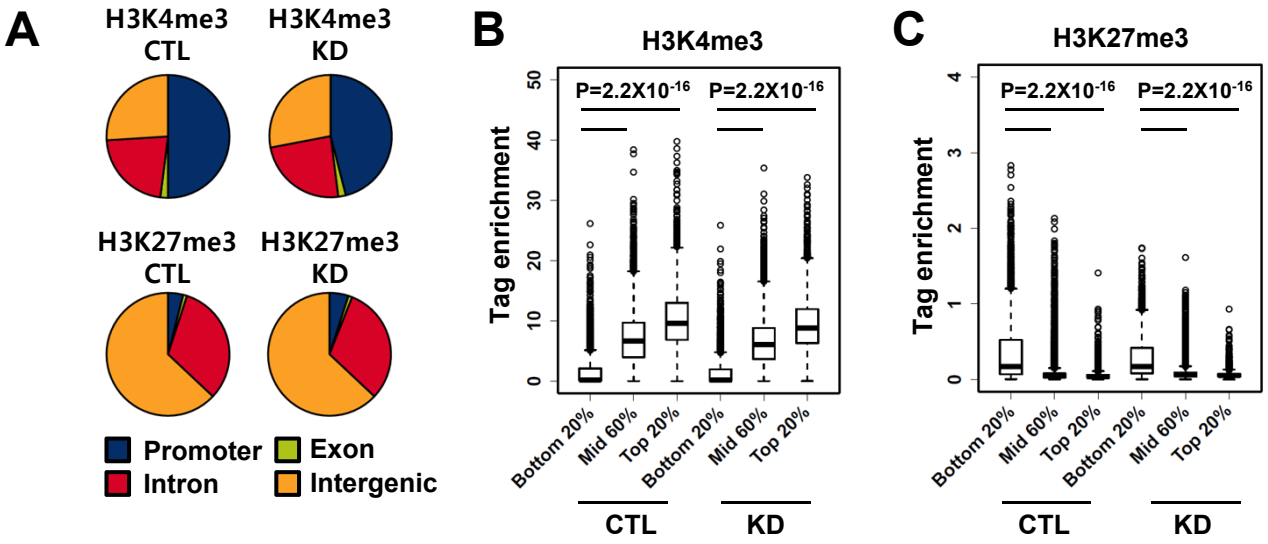

A
Relative mRNA level
1.5
1.0
0.5
0.0
***
CTL
KD
B
CTL
KD
SATB1
β-actin
C
CLDN1
INHBA
ADAMTS6
CLSPN
S100A4
TGFB1
VEGFB
OASL
Relative mRNA level
CTL
KD

Figure S2


A
H3K4me3
CTL
H3K4me3
KD
H3K27me3
CTL
H3K27me3
KD
Promoter
Intron
Exon
Intergenic
B
H3K4me3
Tag enrichment
P=2.2X10-16
P=2.2X10-16
Bottom 20%
Mid 60%
Top 20%
Bottom 20%
Mid 60%
Top 20%
CTL
KD
C
H3K27me3
Tag enrichment
P=2.2X10-16
P=2.2X10-16
Bottom 20%
Mid 60%
Top 20%
Bottom 20%
Mid 60%
Top 20%
CTL
KD