SETDB1 mediated histone H3 lysine 9 methylation suppresses MLL-fusion target expression and leukemic transformation

James Ropa,[1,2] Nirmalya Saha,[1*] Hsiangyu Hu,[1*] Luke F. Peterson,[3] Moshe Talpaz[3] and Andrew G. Muntean[1]

[1]Department of Pathology, University of Michigan Medical School; [2]Department of Computational Medicine and Bioinformatics, University of Michigan Medical School and [3]Department of Internal Medicine/Division of Hematology/Oncology, University of Michigan School of Medicine and Comprehensive Cancer Center, Ann Abor, MI, USA

**NS and HH contributed equally as co-second authors.*

 doi:10.3324/haematol.2019.223883

Received: April 4, 2019.

Accepted: September 25, 2019.

Pre-published: September 26, 2019.

Correspondence: *ANDREW G. MUNTEAN* - andrewmu@umich.edu

Supplemental Methods and Figures

Leukemia colony formation assay

Cells were retrovirally transduced as described above. They were then seeded at a density of $1x10^3$ cells in 2mL semi-solid methylcellulose medium for mouse cells (STEMCELL M3234) containing selection antibiotics (1mg/mL G418 and 1ug/mL puromycin or 100ug/mL hygromycin), 10 ng/mL IL-3, GM-CSF, and IL-6 and 100ng/uL SCF (R&D). Colonies were counted, pooled, and replated in the same way for two additional rounds of colony formation and counting. 1x images of the 5-phenyl tetrazolium chloride (INT) stained dishes were taken after round 2.

HSPC colony formation

Mouse bone marrow was harvested and lineage depletion was performed using the Hematopoeitic Progenitor Isolation Kit (STEMCELL) or CD34+ cells were isolated from mobilized blood. One group of mouse cells were transduced with shRNA against SETDB1 as previously described. These cells were pretreated for 4 days with puromycin before being plated in colony formation assays. The other group of mouse cells and the human cells were treated with the indicated dose of UNC0638 or DMSO vehicle control in liquid culture containing cytokines: 100ng/mL SCF and 10ng/mL IL-3. Cells were incubated for two days, retreated and given fresh media and cytokines, and incubated for another two days. For normal mouse HSPC colony formation assays, 10,000 cells were plated in semi-solid methylcellulose in the presence of cytokines: 100ng/mL SCF and 10ng/mL IL-3, IL-6, and GM-CSF. For normal mouse HSPC colony formation assays, 10,000 cells were plated in semi-solid methylcellulose in the presence of cytokines: 50ng/mL SCF, 10ng/mL GM-CSF, IL-6 and IL-3 and 1U/mL EPO. Colonies were allowed to form for 10 days (mouse) or 14 days (human) and were then counted. Cells were pooled from colonies and 20,000 cells were replated in methylcellulose. Colonies were counted again after 10 days (mouse) or 14 days (human). Each mouse harvest was considered a biological replicate: n=3 for shRNA, n=4 for UNC0638. Each patient sample was considered a biological replicate for CD34+ cell assays: n=2.

Apoptosis/ Chemosensitivity assays

Cells were seeded at a density of 25,000 cells in a total volume of 100 µl. Cells were treated with Daunorubicin (Tocris) at a final concentration of 0.0001 µM, 0.001 µM, 0.004 µM, 0.007 µM, 0.01 µM, 0.1 µM and 1 µM for 18 hours followed by treatment with MTT. IC50 values were determined for MA9+EV and MA9+SETDB1 cell lines. Apoptosis analyses were then performed after treating with the IC50 concentrations. Apoptosis assay was performed by flow cytometry analysis of Annexin V/DAPI staining.

Flow cytometry/ AML transformation assay

For assays comparing Ckit+ normal hematopoietic cells to primary AML samples, 5 mice age 10-12 weeks were harvested per group (n=2 groups of 5 mice). Bone marrow was flushed, lineage depletion performed using Hematopoeitic Progenitor Isolation Kit (STEMCELL). Lin- cells from all 5 mice were pooled and stained using anti-Ckit-APC (BioLegend). Cells were washed and subjected to flow cytometry in the presence of DAPI using the University of Michigan Flow Cytometry core. Live Ckit+ cells were sorted and RNA was harvested from these cells. An equivalent number of live primary AML cells were harvested for RNA.

For assays involving cell surface markers of HSPCs or AML transformation after treatment with UNC0638, 1 mouse per replicate (n=4) was harvested. Lin- mouse bone marrow was isolated as described above and treated in culture for 4 days with the indicated dose of UNC0638 or DMSO vehicle control. Cells were stained with anti-Ckit conjugated to APC fluorophore or anti-Cd11b conjugated to PE fluorophore (BioLegend). Separately, cells treated with UNC0638 were spinfected with MigR1-MLL-AF9 as described above. These cells were then monitored for GFP positivity for 10-13 days starting the day after transduction. Cells were monitored for GFP expression by flow cytometry until 100% GFP was achieved. Flow cytometry was performed to analyze the cell populations. Shown are representative flow plots for UNC0638 or vehicle treatments. Biological replicates (n=4) were also graphed. All additional biological replicates for the AML transformation assay are shown in Supplemental figures.

Statistical Analysis

Statistical tests used to analyze data were determined prior to the experiments being performed. These tests are indicated in figure legends. N values indicate biological replicates. Data assumed to be normally distributed (i.e. qPCR) were analyzed using two sample t-tests comparing raw delta Ct (test gene compared to housekeeping gene) of control groups to test groups. For visualization, average delta delta Ct values were graphed (test gene compared to housekeeping gene relative to the control group) with standard deviations. For colony formation assays, we were testing for reproducible changes in test groups relative to control. Therefore, each test group had to be compared to its respective control group (experiments performed at the same time with cells from the same mice and the same lots of media/ serum/ cytokines). An appropriate statistical test to test for a reproducible change in test groups relative to control groups is generalized linear modeling followed by ANOVA. We report the p-values for the main effect (whether the change is the same for all biological replicates), unless there is a significant interaction between the groups (indicating that the changes are not in the same direction or are significantly different in magnitude). If there is an interaction, no comment can be made on statistical significance. Graphs were generated in R or Prism (Graphpad).

RNAseq analysis

RNA was harvested from MLL-AF9+EV control cells or MLL-AF9+SETDB1 cells using the Qiagen RNeasy Plus Kit. mRNA libraries were generated by the University of Michigan sequencing core and sequenced on an Illumina HiSeq2500. Sequencing files were obtained from the University of Michigan Sequencing Core. Following is a brief description of the RNA-sequencing analysis pipeline. Tool version, authors, and any parameters used that are not defaults of the software/ tools used are listed in Table S1. Quality of sequencing data was analyzed using FastQC. Reads were trimmed for quality using Cutadapt. Indices for alignment were generated using the GENCODE M20 release of the GRCm38 mouse genome assembly with STAR. Reads were aligned to the genome using STAR. Aligned reads were counted by gene (features = exons) using HTSeq, which only keeps uniquely aligned reads by default. Separately, raw aligned count files for THP-1 cells subjected to Crispr-Cas9 KD of SETDB1 were downloaded from GEO (GSE103409)[1,2]. These counts were analyzed for differential gene expression using two separate programs implemented in R: DESeq2 and the edgeR-voom-eBayes pipeline (edgeR). DESeq2 and edgeR both perform computational normalizations based on effective library size, with the base assumption that most genes are not differentially expressed. Both analyses yielded similar results with DESeq2 yielding a more stringent same to same comparison for genes in treated versus control samples, so the significance of differential expression of individual genes was based on DESeq2 analysis. edgeR is more suited to piping into gene set analysis (GSA) using CAMERA, so it was used for downstream GSA and

generation of barcode plots. For overlap analysis between MLL-AF9 + SETDB1 cells and sgRNA-SETDB1 THP1 cells published dataset, only genes that were significant in both sgRNA6 and sgRNA9 datasets were included[1]. Counts were taken from the experimental set up analyzing expression 4 days after induction of SETDB1 KD with sgRNA[1]. Stringency for significant genes in MLL-AF9 + SETDB1 cells was relaxed due to the low number of differentially expressed genes in the THP1 cells with our stringent parameters. In this analysis, alpha=0.05 was the only cutoff used, without a fold-change threshold. The following analyses were all implemented in R using the listed packages: The MA plot was generated using ggplot2. MA plot is a visualization application to represent differential expression of genes in RNAseq data computed based on log-intensity ratios of treated relative to control (M) versus log-intensity averages (A) of all samples in the comparison. Venn diagrams were produced using VennDiagram. CAMERA and ROAST were used for gene set analyses using curated MSigDB[3] groups of gene sets or individual gene sets, respectively. Barcode plots were produced by modifying code from the function barcodeplot (limma). All analyses were performed in R version 3.5.1.

Chromatin Immunoprecipitation sequencing (ChIP-seq) prep and analysis

ChIP experiments were performed as previously described[4] with slight modifications. Briefly, $3x10^7$ MLL-AF9+EV control cells or MLL-AF9+SETDB1 overexpression cells (n=2 each) were crosslinked with 1% formaldehyde, quenched with 125mM Glycine, lysed with 1% SDS and sonicated on a Bioruptor Pico sonicator (Diagenode). Cleared lysates were immunoprecipitated with 4ug antibodies using Protein G dynabeads (Invitrogen). All antibodies were validated for specificity using Histone peptide arrays from EpiCypher (Antibodies used: anti-H3K9me3: Active Motif; anti-H3K9ac: EpiCypher; anti-H3K79me2: Abcam). The IPs were washed with a low salt buffer, a high salt buffer, and a lithium chloride wash buffer. Protein-DNA complexes were eluted in 1% SDS, were decrosslinked in high salt and treated with RNaseA and ProteinaseK (Invitrogen). DNA was purified with a Qiagen PCR purification kit. Libraries were prepared using the iDeal Library Preparation kit (Diagenode) according manufactuerer's recommendations. Library amplification was optimized by monitoring amplification cycles using SYBR I fluorophore and qPCR. Paired-end 38bp sequencing was performed on the NextSeq500 (Illumina) by the University of Michigan Sequencing Core.

Sequencing files were obtained from the University of Michigan Sequencing Core. Following is a brief description of the ChIP-seq analysis pipeline. Tool version, authors, and any parameters used that are not defaults of the software/ tools used are listed in Table S1. Quality control was performed using FastQC. Adapters were trimmed and reads were trimmed for quality using AfterQC, which specifically recognizes poly(G) sequences that are artifacts of NextSeq sequencers. Reads were aligned to the GRCh38 genome assembly using Bowtie2. SAMtools was used to sort the aligned reads, filter out mitochondrial reads, and keep only reads that were confidently uniquely mapping (0.995 probability of mapping uniquely). Picard was used to remove PCR duplicates. After reads were aligned, peaks were called using MACS2. Peak analysis was done in two separate ways. First, consensus peaks were determined for MA9+EV or MA9+SETDB1. Only peaks that were called in both replicates for a given sample were kept for this analysis. Next, overlap analysis was performed using ChIPpeakAnno to determine peaks that are found in both samples, or in only one sample. Finally, these peaks were annotated to promoter regions (-5000bp, +2000bp from transcription start site) for H3K9ac ChIP or for any peak overlapping a gene locus for H3K9me3 or H3K79me2 ChIP. The second analysis used DiffBind to analyze differential signals for the different experiments. DiffBind keeps any peaks that are found in at least two samples in the entire experiment as consensus peaks, and then re-counts the reads for all samples at those peak regions. After obtaining differential binding results, the peaks were annotated as described above. Overlap analyses were performed with this differential binding analysis to determine peaks that were significantly

changed in the different ChIP conditions, as well as RNA-seq and ATAC-seq. Signal tracks were generated using DeepTools and normalizing by reads per genomic content. Integrated Genome Browser (IGB) was used to visualize sequencing tracks.

Assay for Transposase-Accessible Chromatin (ATAC-seq)

Cells were pretreated with DNase I, harvested, and shipped to Active Motif for ATAC-seq library prep and sequencing. Sequencing files were obtained from Active Motif. Analysis was performed the same as ChIP-seq with two modifications. First, MACS2 parameters were amended to center reads on the site of transposition (Table S1). Second, signal tracks were generated from MACS2 (using deepTools to convert pileup files to bigwig) in order to preserve the signal pattern after this read shifting so that visualized peaks are at the actual point of transposition.

Table S1: List of software and tools for analysis of Next Generation Sequencing

Software/Tool	Version	Citation	Use	Parameters
SRAtoolkit	2.9.4		Download sequencing data from GEO	
FastQC	0.11.7	Babraham Bioinformatics[5]	Quality statistics	
Cutadapt	1.18	Martin, 2011[6]	Quality filtering of fastq files for RNA-seq/ ATAC-seq; adapter trimming	-q 20,20 --minimum-length 30
AfterQC	0.9.6	Chen, et al. 2017[7]	Quality filtering of fastq files for ChIP-seq; adapter trimming	-q 20 -u 19 -p 19 -s 25
STAR	2.6.0c	Dobin, et al. 2013[8]	Alignment of RNA-seq data	genomeGenerate: --sjdbOverhang 49 Alignment: --outSAMtype BAM SortedByCoordinate
Bowtie2	2.3.4.3	Langmead & Salzberg 2012[9]	Alignment of ChIP-seq data	--very-sensitive -X 2000
SAMtools	1.9	Li, et al 2009[10]	Sorting, indexing, quality filtering of aligned BAM files	view -b -q 25 -f 0x2
Picard	2.18.19	Broad Institute[11]	Removing PCR duplicates from ATAC-seq/ChIP-seq	
MACS2	2.1.2	Zhang, et al. 2008[12]	Calling peaks for ATAC-seq/ ChIP-seq	Narrow ChIP-seq peaks (H3K9ac): -t <input> -f BAMPE -g mm --keep-dup all Broad ChIP-seq peaks (H3H9me3, H3K79me2): -t <input> -f BAMPE -g mm --

				scale-to-large --broad --keep-dup all ATAC-seq peaks: -f BAM -g mm --nomodel --shift -100 --extsize 200 --keep-dup all
deepTools (bamCoverage)	3.1.3	Ramirez, et al. 2014[13]	Generation of normalized bigwig files for sequencing tracks	RNA-seq: --binSize 10 --scaleFactor <sizeFactor from DESeq2> --normalizeUsing None ChIP-seq/ATAC-seq: --binSize 10 --normalizeUsing RPGC --effectiveGenomeSize 2150570000 --extendReads
HTSeq	0.11.0	Anders, et al. 2014[14]	Counting reads assigned to features (exons) for RNA-seq	-s reverse -r pos
Integrated Genome Browser	9.0.2	Nicol, et al 2009[15]	Visualization of sequencing tracks	
DESeq2	3.8	Love, et al. 2014[16]	Differential gene expression analysis	alpha = 0.05; lfcThreshold = log2(1.5)
edgeR	3.8	Robinson, et al. 2010[17] McCarthy, et al. 2012[18]	Differential gene expression analysis	
limma	3.8	Ritchie, et al 2015[19]	Barcode plots; library normalization quality control	
ChIPpeakAnno	3.8	Zhu, et al. 2010[20,21]	Establishing consensus peak sets and gained/lost peaks in ChIP-seq/ ATAC-seq analyses	Peaks were annotated to promoters (-5000, +2000 from TSS) for ATAC and H3K9ac; Peaks were annotated to any overlapping gene feature for H3K79me2 and H3K9me3
DiffBind	3.8	Stark, et al 2011[22,23]	Analyzing differential peak signals for ChIP-seq/ATAC-seq	method = DBA_DESEQ2; bFullLibrarySize=TRUE
CAMERA		Wu and Smyth 2012[24]	Differential gene expression analysis to generate statistics for barcode plots	inter.gene.cor=0.01

ROAST		Wu, et al. 2010[25]	(see CAMERA, for stand alone gene sets)	
ggplot2	3.1.0	Wickham, et al	Generation of graphs	
VennDiagram	1.6.2	Boutros, Paul	Generation of venn diagrams	

Table S2: Primers used for qPCR		
Target	F-primer (5'-3')	R-primer (5'-3')
HA-hSETDB1 (exogenous)	TGTTCCTGATTATGCCGGGT	TTCCTCAACCACTGCCTGTT
mSETDB1 (endogenous)	CTTCTGGCTCTGACGGTG	GGAAGCCATGTTGGTTGATT
hMLL (exogenous)	ACCCCATCAGCAAGAGAGG	TTCGTGGAGGAGGCTCAC
Id2	GTCCTTGCAGGCATCTGAAT	CTCCTGGTGAAATGGCTGA
Cd80	TATTGCTGCCTTGCCGTTAC	GACCAGGCCCAGGATGATA
Nab2	AGGACAGCGCCAGTCTGT	GTGTCTGCTGCAAGATGTGG
Itgam	CCCATGACCTTCCAAGAGAA	ACACCGGCTTGTGCTGTAGT
β-Actin	GCCCTGAGGCTCTTTTCCAG	TGCCACAGGATTCCATACCC
GAPDH	CTGCACCACCAACTGCTTAG	GTCTTCTGGGTGGCAGTGAT

Table S3: Target sequences of shRNA (obtained from Genetic Perturbation Platform https://portals.broadinstitute.org/gpp/public/gene/search_clones)	
shSetdb1_1	CCCGAGGCTTTGCTCTTAAAT
shSetdb1_2	TGGCACAAAGGCACCCTTATT
shSetdb1_3	TGCTATCTGGGAACCATATTG
shSetdb1_4	CCCATGAGAAACGAACAGTAT

Description of other supplemental tables:

Table S4: RNA-sequencing data. Excel file containing the following sheets: 1) DESeq2 differential expression analysis for all genes that were not filtered out by low counts. 2) DESeq2 analysis for only significantly differentially expressed genes. 3) Overlap of significantly downregulated genes and promoter regions that have reduced H3K9me3 in AML patient samples relative to CD34+ cells pictured in Figure 5F (Muller-Tidow, et al. 2010)[26]. 4) Overlap of significantly downregulated genes and genes that are significantly upregulated by sgRNA mediated knockdown of SETDB1 in human THP1 cells pictured in Figure 5G (Cuellar, et al 2017)[1].

Table S5: ChIP-sequencing/ ATAC-sequencing data. Excel file containing the following sheets: 1) Consensus peaks for ATAC-seq data. 2) DiffBind analysis for ATAC-seq data. 3) Consensus peaks for H3K9me3 ChIP data. 4) DiffBind analysis for H3K9me3 ChIP data. 5) Consensus peaks for H3K9ac ChIP data. 6) DiffBind analysis for H3K9ac ChIP data. 7) Consensus peaks for H3K79me2 ChIP data. 8) DiffBind analysis for H3K79me2 ChIP data. 9) Overlap of genes that have reduced expression, reduced promoter HK9ac signal and reduced promoter ATAC-seq signal with overexpression of SETDB1 compared to EV control, pictured in Figure 6C.

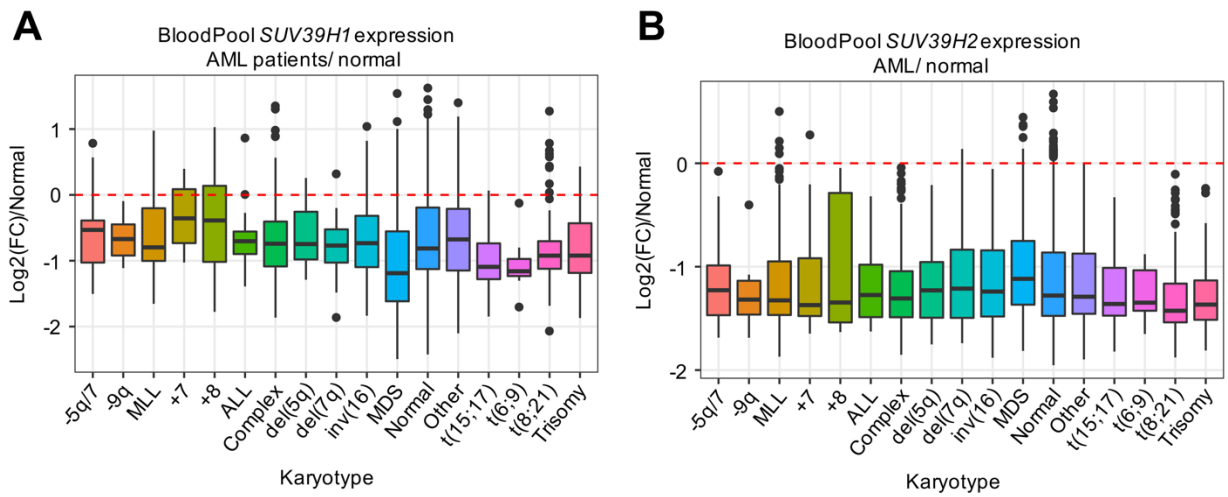

Figure S1: AML patients have low relative expression of SUV39H1/2 regardless of karyotype. A/B) BloodPool data for AML patients' SUV39H1/2 expression relative to nearest normal counterpart divided by karyotype (n=1991).

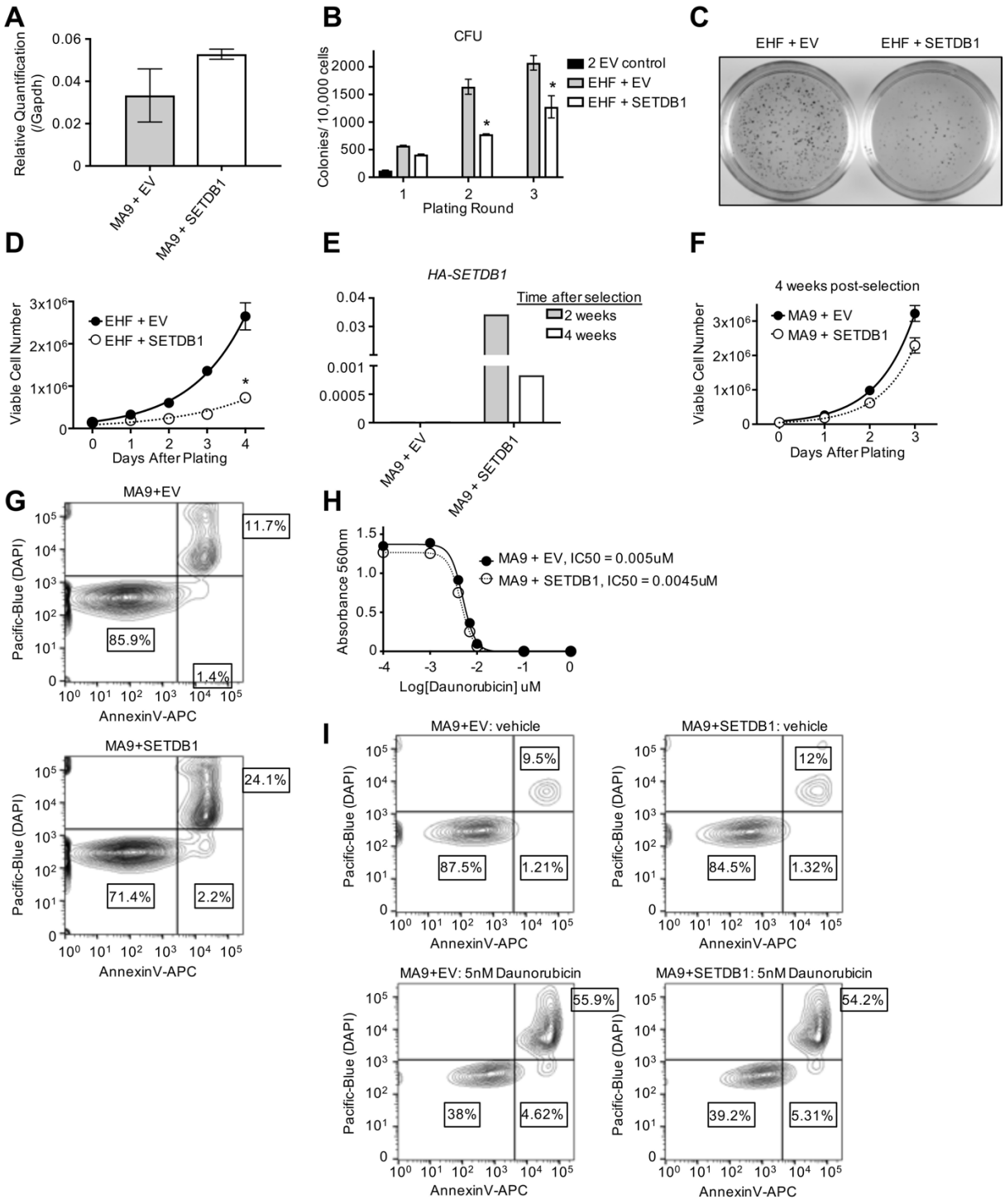

Figure S2: Overexpression of SETDB1 delays AML growth. A) qPCR using primers for hMLL to determine expression levels of MLL-AF9 in MA9+EV or MA9+SETDB1 cells (n=2) from ex vivo experiments in Figure 2. B) Mouse lin- bone marrow was retrovirally transduced with the indicated plasmid vectors (EHF= E2A-HLF) and plated in methylcellulose. Colonies were counted after 7 days and re-plated, for a total of three rounds. Shown is one representative experiment of n=2. C) Representative INT staining of colony assay plates for from B. D) Lin- bone marrow cells retrovirally transduced with EHF in the presence or absence of SETDB1

overexpression, selected for 2 weeks, then proliferation was monitored by viable cell count daily. Shown is one representative experiment of n=2. E) qPCR using primers for HA-SETDB1 to determine expression of exogenous SETDB1 in MA9 cells with or without SETDB1 overexpression 2 and 4 weeks after selection media is withdrawn. F) Proliferation assay of MA9 cells 4 weeks after selection media is withdrawn. G) Representative Apoptosis analysis flow plot. H) Determination of IC50 for Daunorubicin on MA9 cells using MTT assay. Cells were seeded at a density of 25000 cells in a total volume of 100 µl. Cells were treated with Daunorubicin (Tocris) at a final concentration of 0.0001 µM, 0.001 µM, 0.004 µM, 0.007 µM, 0.01 µM, 0.1 µM and 1 µM for 18 hours followed by treatment with MTT. IC50 determined by the MTT assay is listed on the figure. n=1. (I) Representative flow cytometry plot for apoptotic assay showing the effect of Daunorubicin treatment at IC50 on MA9 cells upon overexpression of SETDB1.Statistics: Generalized linear modeling followed by ANOVA where each EHF+SETDB1 replicate was paired to the EHF+EV control from the same biological replicate. Main effect is reported if there are no significant interactions. (B/D); MA9 = MLL-AF9; EHF=E2A-HLF; EV= Empty vector control; n=biological replicates; * $p < 0.05$.

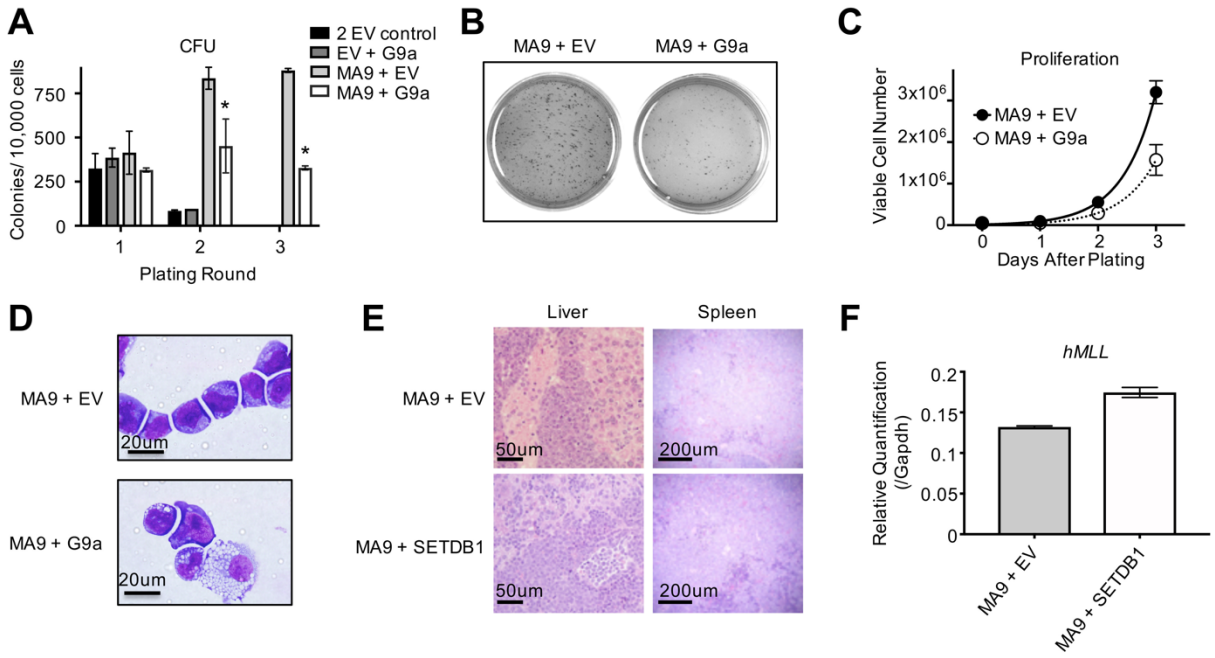

Figure S3: Overexpression of G9a delays AML growth. A) Mouse lin- bone marrow was retrovirally transduced with the indicated plasmid vectors and plated in methylcellulose. Colonies were counted after 7 days and re-plated, for a total of three rounds. Shown is one representative experiment of n=3. B) Representative INT staining of colony assays from A. C) Proliferation assay of MA9 cells with or without G9a overexpression. Representative of n=2 D) Cytospin and Hema3 staining of MA9 cells with or without G9a overexpression (n=2). E) H&E staining of liver (40x) and spleen (10x) sections from MA9+EV or MA+SETDB1 leukemic mice, taken after mice are euthanized at the moribund stage. F) qPCR using primers for hMLL to determine expression levels of MLL-AF9 in MA9+EV or MA9+SETDB1 cells (n=2) from in vivo experiment in Figure 2. Statistics: Generalized linear modeling followed by ANOVA where each MA9+G9a replicate was paired to the MA9+EV control from the same biological replicate. Main effect is reported if there are no significant interactions. (A); MA9 = MLL-AF9; EV= Empty vector control; n=biological replicates; * $p < 0.05$

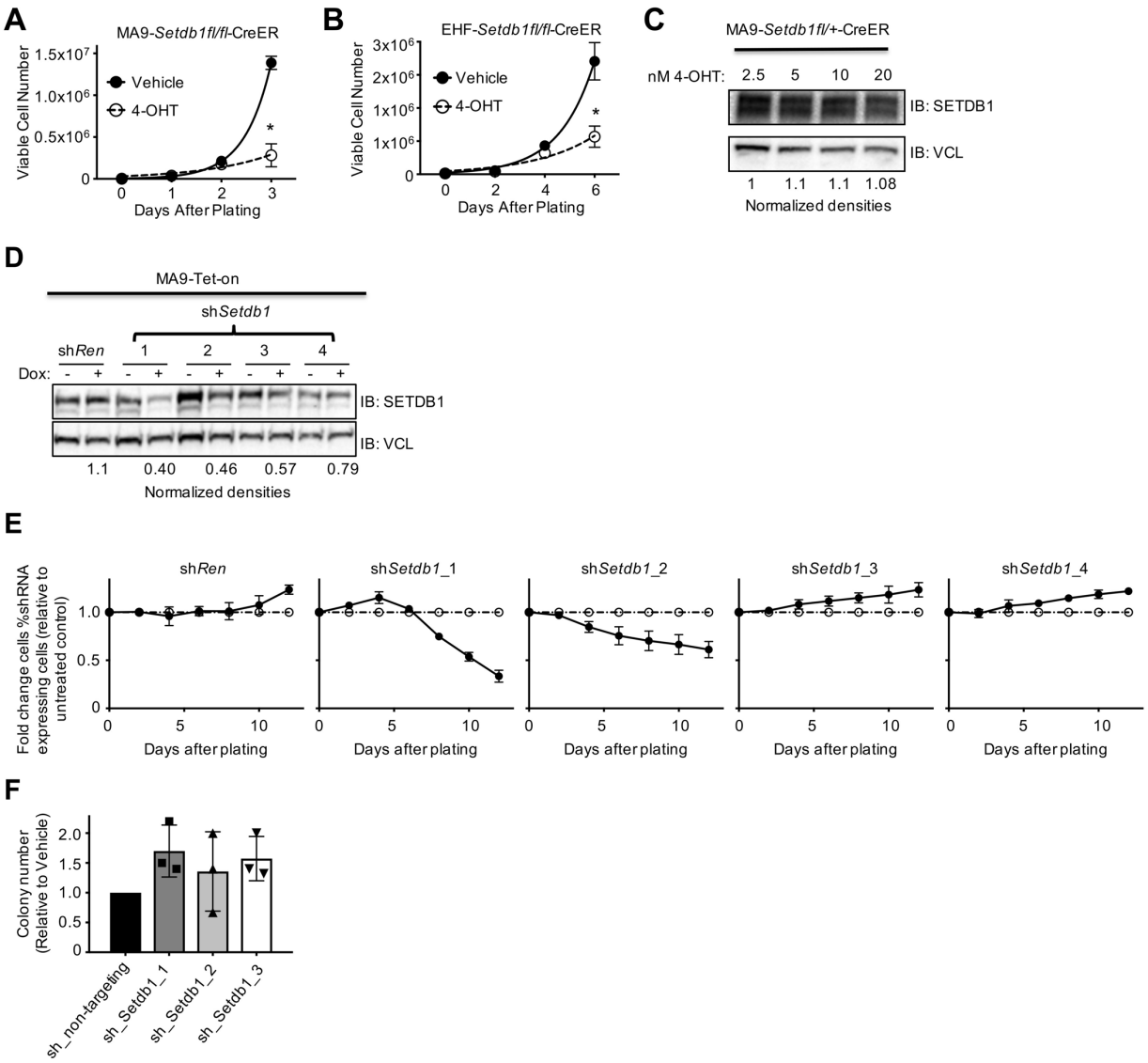

Figure S4: Deletion and KD of Setdb1 have varying effects on established AML cell lines. A/B) Proliferation assays of MA9 or EHF cells following treatment with 5nM 4-OHT to induce genetic excision of Setdb1. 1 representative experiment of n=2. C) Western blot after inducing heterozygous deletion of Setdb1 using the Setdb1fl/+ CreER tamoxifen inducible system. Shown below are the normalized densities of the SETDB1 bands compared to vehicle. G) MLL-AF9 cells were transduced with docycycline inducible shRNA targeting SETDB1. Western blot shows protein levels of SETDB1 in MLL-AF9 Tet-on cells that have been induced to express shRNA targeting Setdb1 by treating with 7nM doxycycline. Shown below are the normalized Setdb1 band densities relative to the non-targeting control. E) After selection, transduced cells were plated at a 1:1 ratio with the parent cell line and treated with doxycycline to induce expression of the shRNA or vehicle control, separately. shRNA expressing cells express GFP and were monitored by FACS. Shown are the changes in shRNA expressing cell percentages normalized to the untreated group. n=2. F) Lin- mouse bone marrow was isolated and spinfected with constitutive shRNA targeting Setdb1. Cells were pre-selected with puromycin for 7 days, then plated in methylcellulose. Colonies were counted after 7 days. Statistics: Generalized linear modeling followed by ANOVA where each *Setdb1fl/fl*CreER + OHT replicate

was paired to the + Vehicle control from the same biological replicate. Main effect is reported if there are no significant interactions (A/B). MA9 = MLL-AF9, EHF=E2A-HLF; EV= Empty vector control; n=biological replicates; * $p < 0.05$.

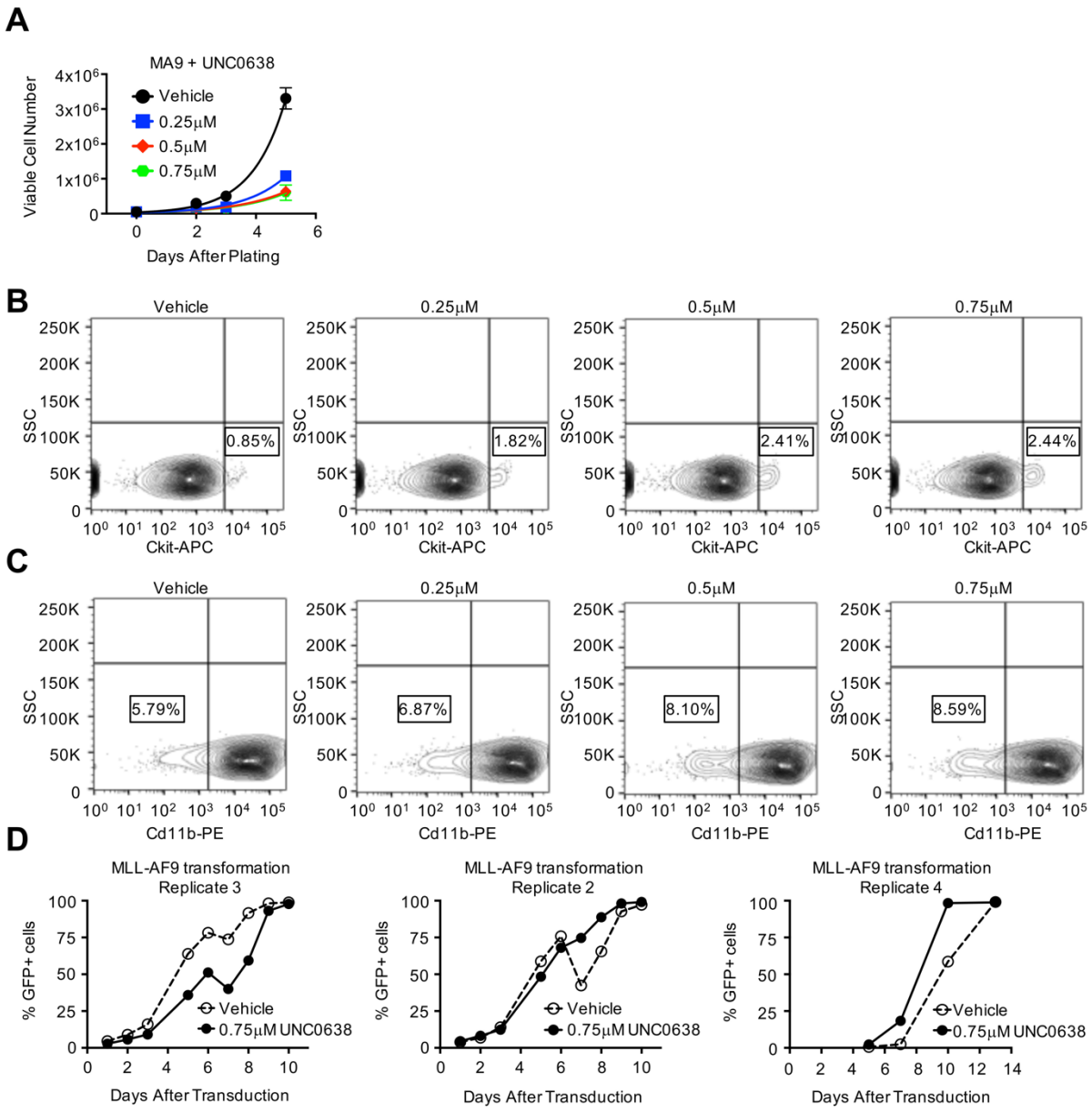

Figure S5: Treatment with UNC0638 preserves primitive cells amenable to transformation. A) Proliferation assay monitoring MA9 cell growth after treatment with UNC0638. B-D) Lin- mouse bone marrow was treated with the indicated doses of UNC0638 for 4 days. B/C) Cells were stained with anti-Ckit conjugated to APC fluorophore (B) or anti-Cd11b conjugated to PE fluorophore (C). Flow cytometry was performed to analyze Ckit+ or Cd11b- populations. B/C) Shown are representative plots for all UNC0638 concentrations tested, n=3 for each concentration. D) After treatment with UNC0638, cells were spinfected with MigR1-MLL-AF9, which also expresses a GFP reporter. Cells were monitored for GFP expression by flow cytometry until 100% GFP was achieved. Shown are the three additional biological replicates that were not shown in Figure 4, for a total of n=4.

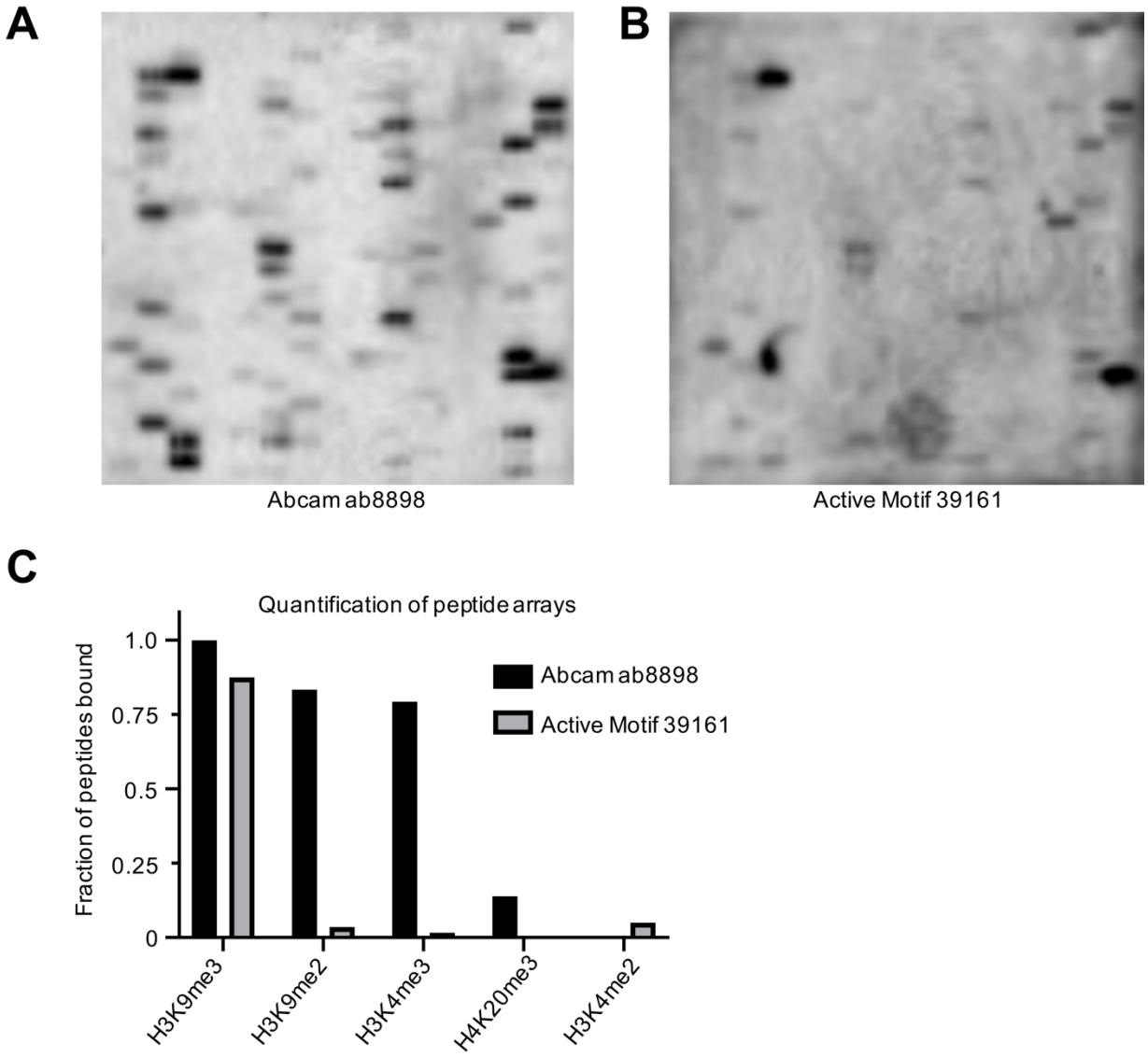

Figure S6: Antibody validation for H3K9me3 ChIP-seq. A/B) Histone peptide arrays using the indicated antibodies. Blots were developed for 5 minutes and imaged at the same time. C) Quantification of the peptide arrays. Shown are the fraction of peptides containing the target (H3K9me3) or peptides that did not contain the target (indicated on the x axis) and the fraction of those peptides that were bound relative to how many were present in the array.

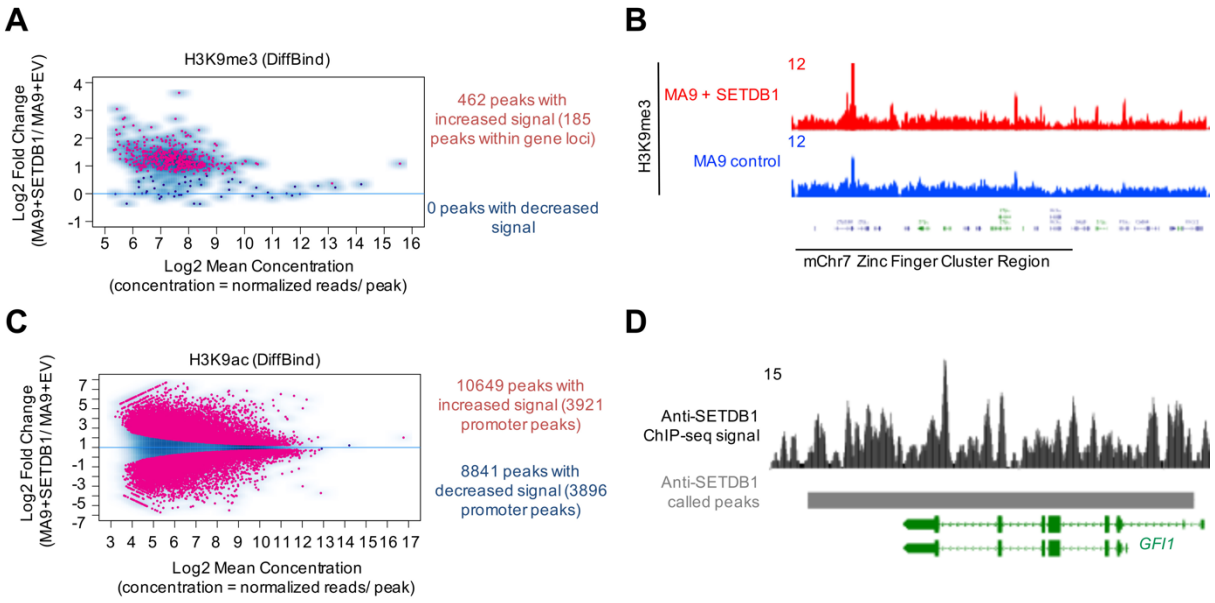

Figure S7: H3K9ac is a useful alternate to looking at changes in the epigenome to H3K9me3. A) DiffBind analysis showing H3K9me3 ChIP-seq peaks that demonstrate differential binding. 552 consensus peaks were found in the analysis. B) H3K9me3 ChIP-seq track of the mouse Chromosome 7 Zinc Finger Cluster, which is known to be enriched for H3K9me3. C) DiffBind analysis showing H3K9ac ChIP-seq peaks that demonstrate differential binding. D) ENCODE data showing the anti-SETDB1 ChIP-seq signal and called peaks at the *GFI1* locus.

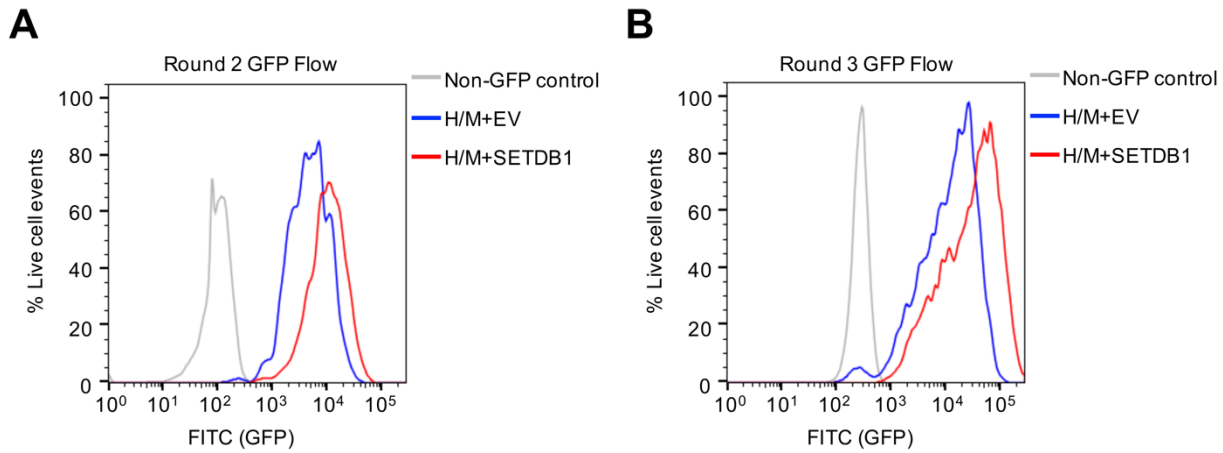

Figure S8: Exogenous MEIS1 is upregulated upon SETDB1 overexpression in HOXA9/MEIS1 transformed cells. A/B) Representative histograms showing GFP positivity for cells from Hoxa9/Meis1 colony formation assays harvest after replating round 2. GFP is expressed off the exogenous MigR1 vector, off of which *Meis1* is also expressed. (A) and round 3 (B). Cells were harvested after the indicated round and subjected to flow cytometry.

References

1. Cuellar TL, Herzner A-M, Zhang X, et al. Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. J Cell Biol. 2017;216(11):3535–3549.
2. Barrett T, Suzek TO, Troup DB, et al. NCBI GEO: mining millions of expression profiles--database and tools. Nucleic Acids Res. 2005;33(Database issue):D562-6.
3. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739–1740.
4. Lin X, Tirichine L, Bowler C. Protocol: Chromatin immunoprecipitation (ChIP) methodology to investigate histone modifications in two model diatom species. Plant Methods. 2012;8(1):48.
5. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. [cited 2019 Feb 12];Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
6. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j. 2011;17(1):10.
7. Chen S, Huang T, Zhou Y, Han Y, Xu M, Gu J. AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. BMC Bioinformatics. 2017;18(Suppl 3):80.
8. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
9. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–359.
10. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–2079.
11. Picard Tools - By Broad Institute [Internet]. [cited 2019 Feb 12];Available from: http://broadinstitute.github.io/picard/
12. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
13. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 2014;42(Web Server issue):W187-91.
14. Anders S, Pyl PT, Huber W. HTSeq — a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–169.
15. Nicol JW, Helt GA, Blanchard SG, Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics. 2009;25(20):2730–2731.
16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
17. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–140.
18. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40(10):4288–4297.
19. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
20. Zhu LJ, Gazin C, Lawson ND, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010;11:237.
21. Zhu LJ. Integrative analysis of ChIP-chip and ChIP-seq dataset. Methods Mol Biol.

2013;1067:105–124.
22. Stark R, Brown G. DiffBind: Differential binding analysis of ChIP- Seq peak data. Bioconductor Vignettes. 2018;
23. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012;481(7381):389–393.
24. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 2012;40(17):e133.
25. Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK. ROAST: rotation gene set tests for complex microarray experiments. Bioinformatics. 2010;26(17):2176–2182.
26. Müller-Tidow C, Klein H-U, Hascher A, et al. Profiling of histone H3 lysine 9 trimethylation levels predicts transcription factor activity and survival in acute myeloid leukemia. Blood. 2010;116(18):3564–3571.