SUPPLEMENTARY INFORMATION

DNMT3A R882 mutations promote anthracycline resistance in acute myeloid leukemia through impaired nucleosome remodeling

Olga A. Guryanova, Kaitlyn Shank, Barbara Spitzer, Luisa Luciani, Richard P. Koche, Francine E. Garrett-Bakelman, Chezi Ganzel, Benjamin H. Durham, Abhinita Mohanty, Gregor Hoermann, Sharon A. Rivera, Alan G. Chramiec, Elodie Pronier, Lennart Bastian, Matthew D. Keller, Daniel Tovbin, Evangelia Loizou, Abby R. Weinstein, Adriana Rodriguez Gonzalez, Yen Lieu, Jacob M. Rowe, Friederike Pastore, Anna Sophia McKenney, Andrei V. Krivtsov, Wolfgang R. Sperr, Justin R. Cross, Christopher E. Mason, Martin S. Tallman, Maria E. Arcila, Omar Abdel-Wahab, Scott A. Armstrong, Stefan Kubicek, Philipp B. Staber, Mithat Gönen, Elisabeth M. Paietta, Ari M. Melnick, Stephen D. Nimer, Siddhartha Mukherjee, and Ross L. Levine

This .pdf file includes supplementary methods, supplementary figures 1-8, supplementary references, and supplementary tables 1-4 and 6 (supplementary tables 5 and 7-8 are available as .xlsx files online):

Extended Data Figure 8. Uncropped images of Western blotting experiments presented in main and supplementary figures.

Extended Data Table 1. Minimal residual disease analysis by mutational status in 144 evaluable patients from the E1900 cohort.

Extended Data Table 2. Molecular persistence of *DNMT3A*^{R882} mutations in 9 paired diagnosis and post-induction samples from the E1900 cohort.

Extended Data Table 3. Clinical characterization of the primary *FLT3*^{ITD} and *NPM1*^{c}-positive AML samples with and without *DNMT3A* mutations used to evaluate response to daunorubicin and doxorubicin *in vitro*.

Extended Data Table 4. Clinical characterization of the primary AML samples used in Comet assay.

Extended Data Table 5. Candidate DNMT3A interacting proteins identified by mass-spectroscopy following DNMT3A peptide pull-down.

Extended Data Table 6. Summary statistics for ERRBS assay on LSK cells derived from *Dnmt3a*^{wt} and *Dnmt3a*^{mut} mice.

Extended Data Table 7. Differentially methylated cytosines (DMCs) in LSK cells from *Dnmt3a*^{mut} mice compared to wild-type controls, identified by MethyKit.

Extended Data Table 8. Differentially expressed genes in LSK cells from *Dnmt3a*^{mut} mice compared to wild-type controls, identified by RNA-seq.

Supplementary methods

Quantitative PCR

Cell were washed in cold PBS twice and harvested in Trizol (Invitrogen), total RNA was isolated according to manufacturer's protocol. Verso cDNA Synthesis Kit (Thermo Scientific) was used for reverse transcription on 1 µg total RNA, primed with random hexamers. Quantitative PCR was performed using VeriQuest Fast SYBR qPCR Master Mix with ROX (USB) on a Quant Studio Flex 6 machine (Applied Biosystems) with the following primers:

Gene	Forward primer	Reverse primer
MDR1	5'-CAG CAA AGG AGG CCA ACA TAC	5'-TGA GGC TGT CTA ACA AGG GCA
CDKN1A	5'-TGT CCG TCA GAA CCC ATG C	5'-AAA GTC GAA GTT CCA TCG CTC
BAX	5'-CCC GAG AGG TCT TTT TCC GAG	5'-CCA GCC CAT GAT GGT TCT GAT
PUMA	5'-GAC CTC AAC GCA CAG TAC GAG	5'-AGG AGT CCC ATG ATG AGA TTG T
L32	5'-TGT CTT GAA TGT GGT CAC CTG A	5'-CTG CAG TCT CTT GCA CAC CT

DNA methylation and gene expression analyses

Total DNA and RNA were isolated from FACS-sorted LSK cells derived from *Dnmt3a*$^{+/m}$ (*Mx1-Cre*$^{+}$, n=4), *Dnmt3a*$^{+/-}$ (*Mx1-Cre*$^{-}$, n=2) and control *Dnmt3a*$^{+/+}$ (*Mx1-Cre*$^{+}$, n=3) 6 months old mice using AllPrep Kit (Qiagen). RNA was subjected to Ribo-Zero rRNA Removal Kit RNA-seq preparation (Illumina) and sequenced on an Illumina 2500 using a paired end 50bp approach. The sequencing libraries were mapped to the mouse genome (mm10) using Star(v.2.3.0e)[1] with default parameters. We then computed read counts with HTSeq v0.5.3p3[2] using parameters "intersection-nonempty" and the Mus_musculus.GRCm38.75 gene model from ENSEMBL. Differential expression analysis was performed using the Bioconductor DESeq2[3] package with default parameters. A rank-list was generated based on a score for each gene, calculated as

−log10(FDR)×sign(log2FoldChange). Gene-set enrichment analysis[4,5] was performed using this rank-list and the Hallmark gene set collection from MSigDB v5.0[4].

Enhanced Reduced Representation Bisulfite Sequencing (ERRBS) was used to determine cytosine methylation patterns at base pair resolution. Briefly, ERRBS library preparation was performed by *Msp*I restriction enzyme digestion of high molecular weight genomic DNA, followed by end repair, size selection, bisulfite conversion and library amplification as previously described[6,7]. Libraries were sequenced on a HiSeq 2000 Illumina machine, 50-bp single-end reads, using a dark sequencing approach (sequencing without imaging (dark) for first 5 bases (insert base positions 1-4), followed by sequencing with imaging (light) for the rest of the read bases (insert base positions 5-50), followed by sequencing with imaging (light) for the first 5 bases (insert base position 1-5)). Sequenced reads were trimmed for adaptor sequences. Reads that pass filter were considered for mapping. If both matched reads were passing filter, the last base of R2 was compared to the first base of R1 (one base overlap; insert base position 5); the base with the best quality score was retained and R1 was concatenated to R2 for use in alignment. Sequences were then mapped to mm10 using a whole genome bismark alignment software approach with a maximum of 2 mismatches in a directional manner allowed. Only uniquely aligned reads were retained. Methylation status was determined from the ratio of C's to T's at CpG sites covered at a minimum of 10X as previously described[7] (for summary please see Extended Data Table 6). Differentially methylated cytosines (DMCs; sites with 2 or more CpGs covered in either a wild-type control or *Dnmt3a*$^{+/m}$ animal were considered (n = 1767524 CpGs), minimum of 25% methylation difference and adjusted p-value for Fisher's exact test (q)<0.01) were determined using the methylKit R package[8] (Extended Data Table 7). No differentially methylated regions were detected using the eDMR R package[9] (criteria used: areas containing a minimum of 3 DMCs with

uniform direction of methylation change that had a minimum of 20% methylation difference and an adjusted *p*-value for Fisher's exact test ($q<0.01$) for significance). Differentially upregulated genes ($q<0.05$, average normalized read count >50, Extended Data Table 8) were compared to genes whose promoters contained at least 3 hypomethylated DMCs.

Raw and processed data were uploaded into the GEO database under the following accession number: GSE72737.

LC-MS analysis of intracellular daunorubicin and daunorubicinol

All experiments were performed in quadruplicate. MEFs derived from *Dnmt3a +/+*, *+/–*, or *–/m* animals were plated on 10 cm^2 dishes (about 1×10^6 cells/dish) and allowed to attach overnight. Then cells were exposed to 300 ng/ml daunorubicin for 4, 8, or 16h, or left untreated. Cells were washed twice in ice-cold PBS, dislodged by scraping in a small amount of liquid, harvested by centrifugation, and snap-frozen on dry ice. Concentration of daunorubicin was determined in a previous pilot experiment as the lowest concentration producing reliable measurements. High-performance liquid chromatography with tandem mass-spectroscopy was utilized to determine relative intracellular levels of daunorubicin and its major metabolite daunorubicinol. Briefly, cell pellets were resuspended in 80% aqueous methanol and protein was precipitated by incubation at –80°C. Extracts were clarified by centrifugation, dried in a Genevac benchtop evaporator (SP Industries), and reconstituted in 60% acetonitrile with 5 mM ammonium formate. Liquid chromatography was performed using a 1290 HPLC system (Agilent Technologies). Daunorubicin and daunorubicinol were detected by an Agilent 6550 Q-TOF with a JetStream source. Identification was achieved by the combination of accurate mass and retention time relative to authentic standards, and peak areas were calculated using Agilent MassHunter software.

Immunofluorescence analysis

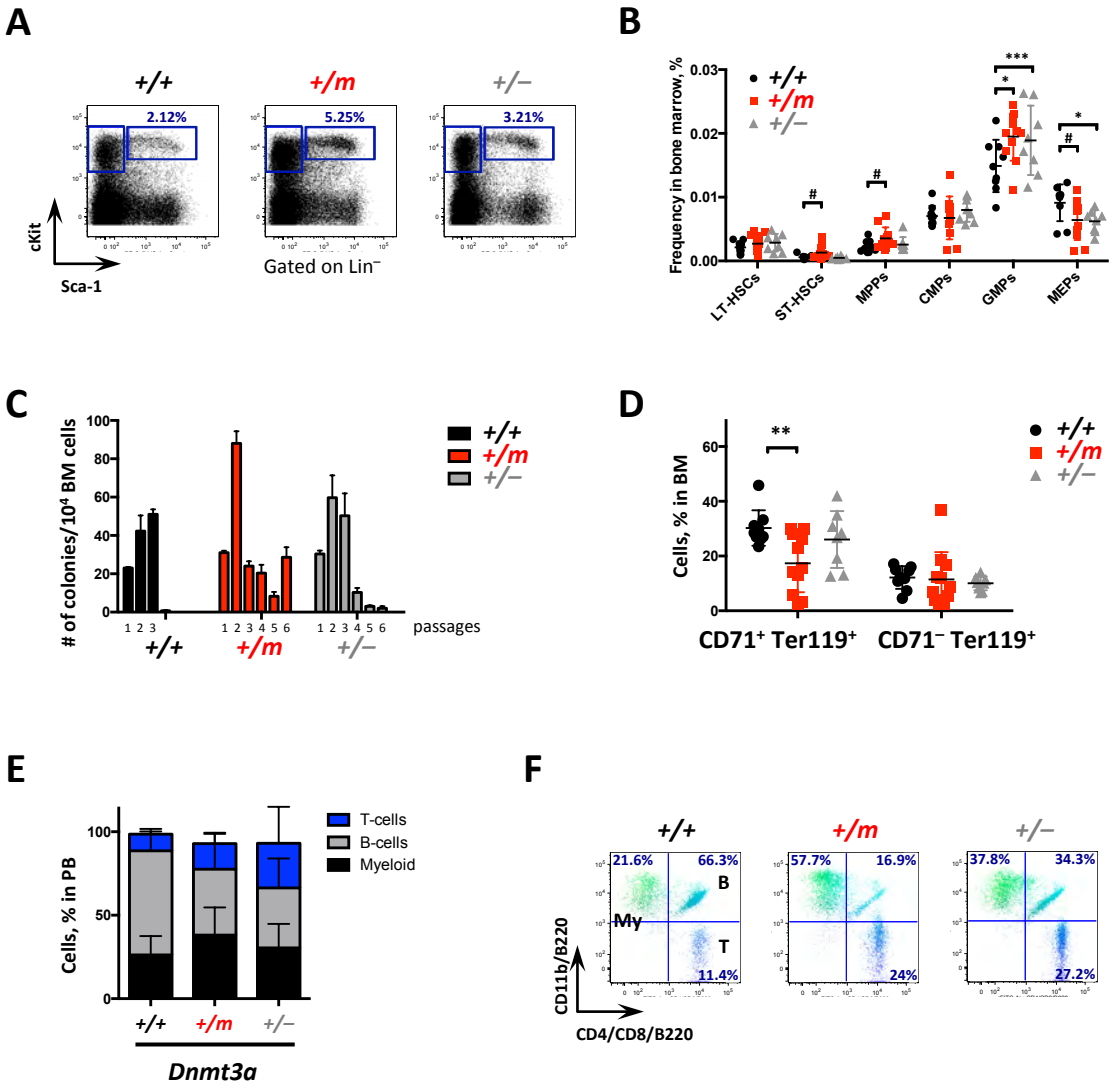
For immunofluorescent staining MEF cells were plated on glass coverslips and allowed to attach for 24 hours and then exposed to 250 ng/ml daunorubicin for 1 or 16 hours, or left untreated. Cells were fixed in 4% formaldehyde in PBS for 15 min at room temperature (RT), permeabilized by 0.5% Triton X-100 in PBS supplemented with 0.225M sucrose for 15 minutes RT, and blocked in 2% bovine serum albumin (BSA) for 45 minutes RT. Coverslips were stained overnight at +4°C in blocking buffer with pH2A.X-specific antibody (1:400, Millipore #05-636), washed 3 times with PBS before adding secondary AlexaFluor488-conjugated antibodies in blocking buffer for 1 hour RT protected from light. DAPI was used to counterstain for nuclei. Finally, the cells were washed 3 times in PBS and mounted using Fluoromount G (Southern Biotech).

Chromatin immunoprecipitation followed by sequencing (ChIP-seq)

ChIP-seq was performed as described[10]. 1-2×10^6 MEF cells were crosslinked using 1% formaldehyde, washed in 1 ml of wash buffer (20 mM Tris pH8.0, 167mM NaCl, 5mM EDTA, 1% SDS) and lysed in lysis buffer (20 mM Tris pH8.0, 167mM NaCl, 5mM EDTA, 0.66% SDS, 1% Triton X-100). Chromatin was sheared to 400-700 bp using Covaris E220 sonicator and used for chromatin immunoprecipitation (ChIP) with Dnmt3a-specific antibodies ab13888 mouse mAbs and ab2850 rabbit pAbs (Abcam) and Protein-A (Life Technologies) magnetic beads. After washing, the immune complexes were eluted and DNA fragments recovered after de-crosslinking at 65°C for 4 hours and used for TruSeq libraries generated using NEBNext kit (New England BioLabs). Libraries were sequenced using HiSeq2500 (Illumina) to generate approximately 4×10^7 single end (SE)50 reads per sample. Reads were aligned using Bowtie2 followed by peak calling

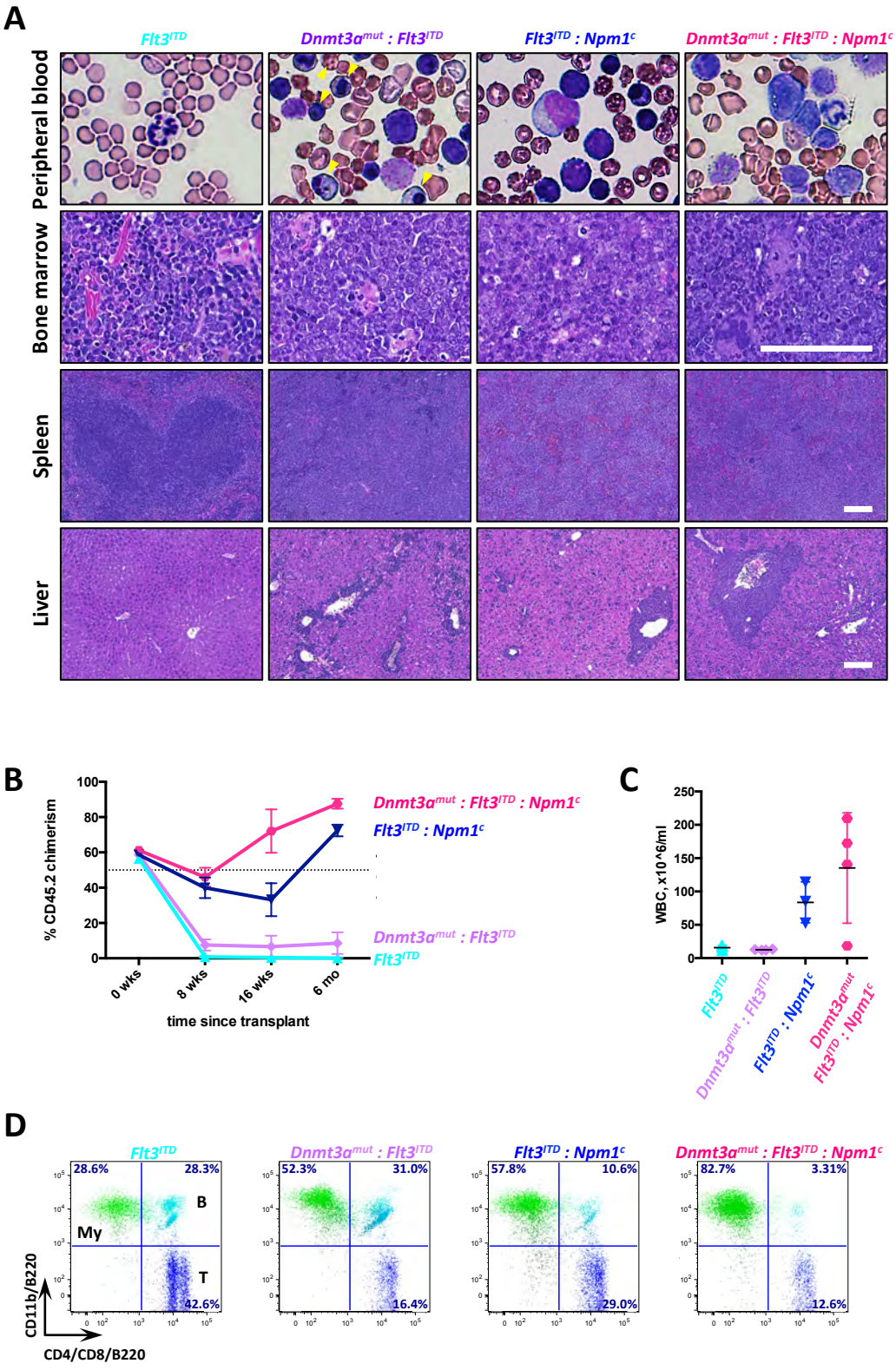
using MACS2. Raw and processed data were uploaded into the GEO database under the following accession number: GSE72883.

Extended Data Figure 1



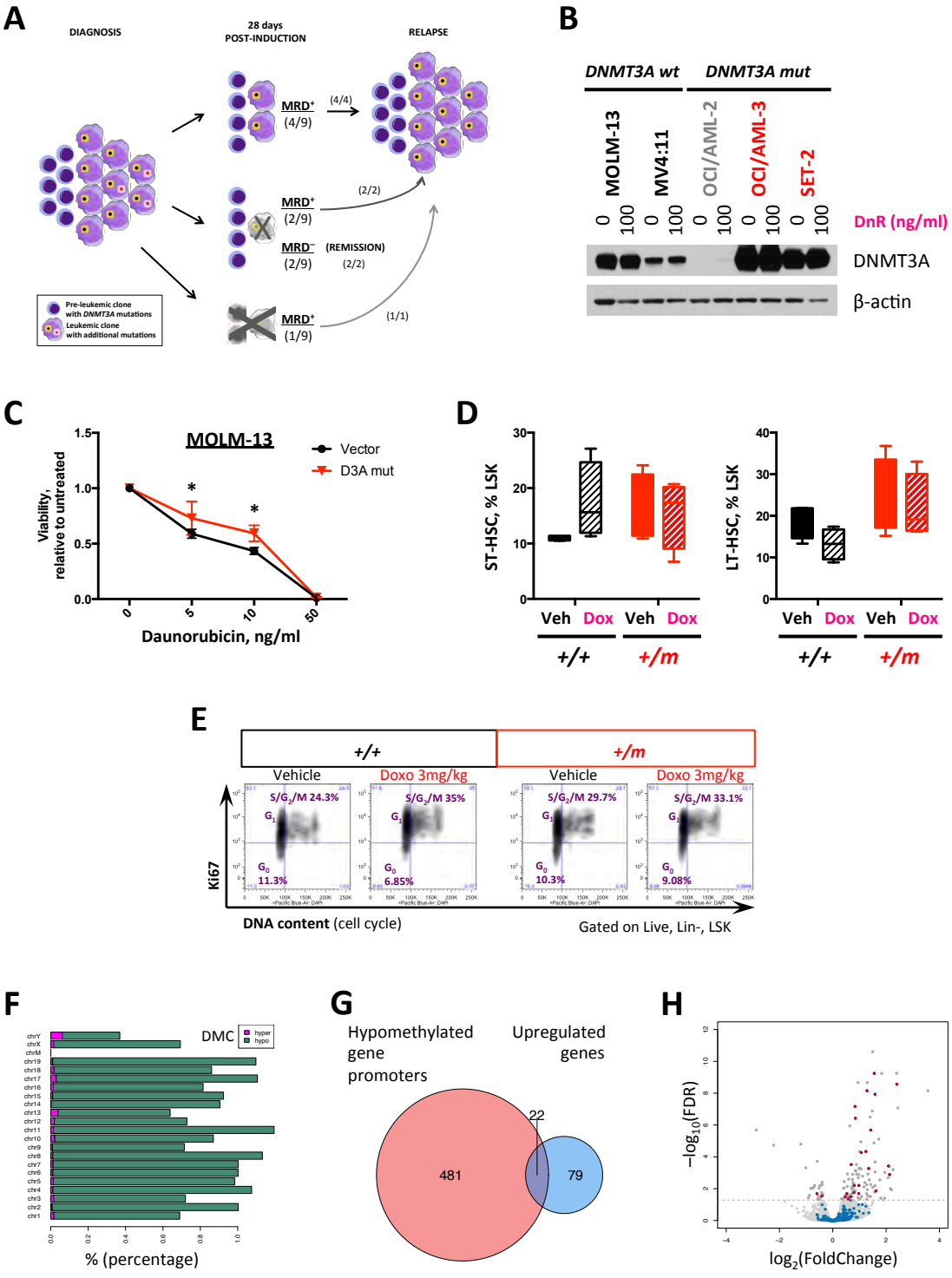
Extended data Figure 1. Characterization of the steady-state hematopoietic phenotype in aged *Dnmt3a*^{mut} mice. (A) Enumeration of the immature Lineage$^-$Sca1$^+$cKit$^+$ (LSK) cell population in the bone marrow of aged *Dnmt3a*^{mut} (*+/m*), wild-type (*+/+*), and haploinsufficient (*+/-*) mice. Representative flow plots for data in Figure 1E. (B) Bone marrow frequencies of the indicated stem/progenitor cell populations in aged animals (n=8-11, $^{\#}$ p<0.065, * p<0.05, *** p<0.001). (C) Serial colony-forming ability of BM cells in methylcellulose assay. (D) Erythroid differentiation analysis in *Dnmt3a*^{mut} bone marrow cells (*+/m*) compared to wild-type (*+/+*) and haploinsufficient (*+/–*) mice (n=8-11, * p<0.01). (E) Analysis of the major differentiated cell lineages (myeloid, B-cells, T-cells) in the peripheral blood of aged *Dnmt3a*^{mut} mice (*+/m*) compared to wild-type (*+/+*) and haploinsufficient (*+/–*) animals (n=8-11, $^{\#}$ p=0.083). (F) Representative flow plots for data in (E).

Extended Data Figure 2



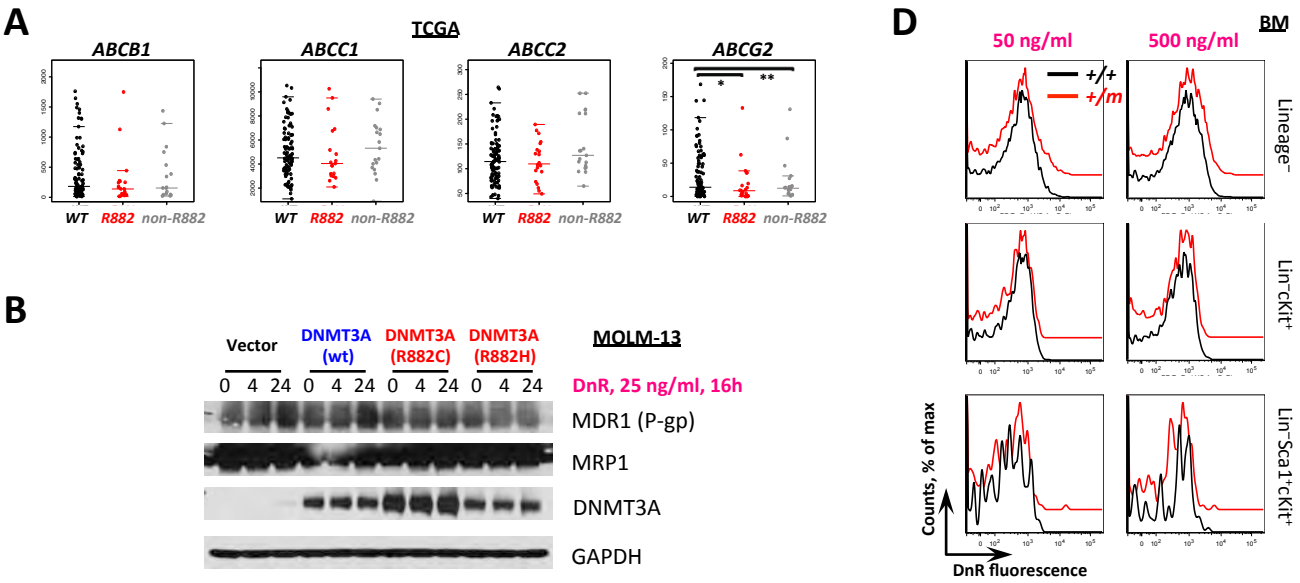
Extended Data Figure 2. *Dnmt3a*^{mut} cooperates with *Flt3*^{ITD} and *Npm1*^{c} to cause AML in a genetic mouse model. (A) Histopathology of H/E-stained bone marrow, spleen and liver sections from mice with indicated genotypes. Bar – 100 μm. (B, C) Peripheral blood CD45.2 chimerism (B) and white blood cell counts (C) in CD45.1 recipients transplated with 1:1 ratio of test CD45.2 BM cells of indicated genotypes and CD45.1 wt competitor after 6 months. (D) Flow cytometric analysis of donor-derived mature lineages in the peripheral blood of CD45.1 recipients non-competitively transplanted with CD45.2 cells of indicated genotypes.

Extended Data Figure 3



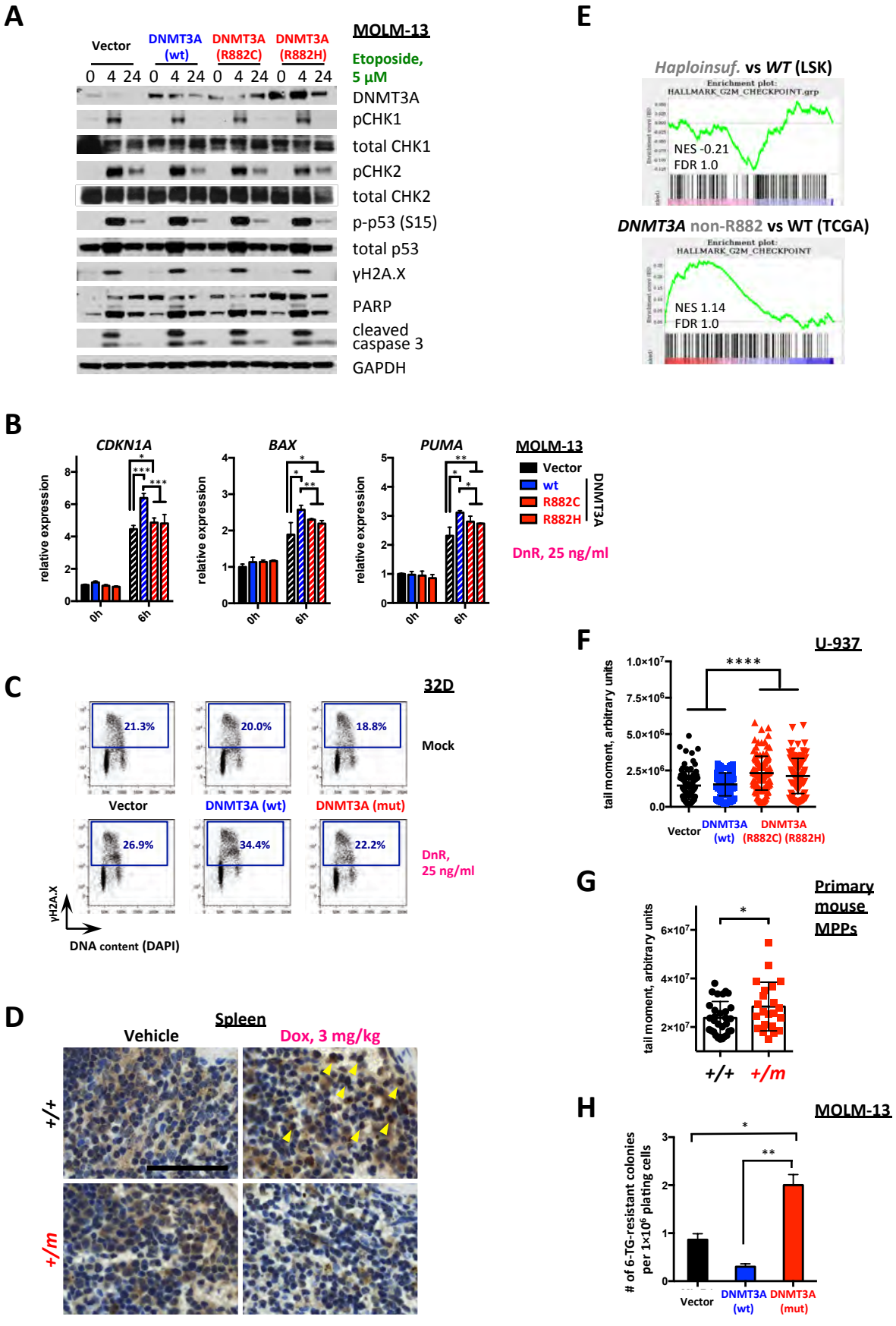
Extended data Figure 3. Expression of *DNMT3A*^{mut} leads to anthracycline resistance. (A) Summary of hybridization capture-based next-generation DNA sequencing using MSKCC-IMPACT gene panel 28 days post-inductions in 9 randomly selected patients with *DNMT3A* R882 mutations at diagnosis (ECOG E1900 AML cohort; average sequencing depth 2246×). In 4 patients all mutations were found persisting after first cycle of chemotherapy; 4 patients cleared non-*DNMT3A* mutations while *DNMT3A* R882 mutations persisted; all mutations were cleared in 1 patient. See also Extended Data Table 2. (B) DNMT3A protein levels in a panel of cell lines with different *DNMT3A* status (red – R882 mutation, gray – non-R882 mutation, black – no mutation), with and without daunorubicin exposure, analyzed by Western blotting. (C) Daunorubicin sensitivity of MOLM-13 cells retrovirally expressing *DNMT3A*^{mut} compared to empty vector control (* p<0.05). (D, E) Relative LT-HSC and ST-HSC frequencies (D) and cell cycle profiles of LSK cells (E) in the bone marrow of *Dnmt3a*^{mut} and wild-type mice 24 hours after one-time *in vivo* treatment with 3 mg/kg doxorubicin. (F-H) Changes in DNA methylation as measured by ERRBS (F) and their effect on gene expression by RNA-seq (G, H) in LSK cells FACS-sorted from *Dnmt3a*^{mut} animals (n=4) compared to wild-type (n=3). 16484 differentially methylated cytosines (DMCs) were identified: 16238 sites were hypomethylated and 246 hypermethylated in animals expressing mutant *Dnmt3a* compared to wild-type controls (F, Extended Data Table 7). 3856 DMCs annotated to 1273 coding gene promoters (TSS ± 2000 bp), of which 503 contained ≥3 hypomethylated DMCs. The Venn diagram shows overlap of genes with ≥3 hypomethylated DMCs in their promoters with upregulated genes (q<0.05, average normalized read count >50, Extended Data Table 8) in *Dnmt3a*^{mut} mice compared to *Dnmt3a* wild-type (G, n=22 genes, p=3.75×10^{-11}, Fisher's exact t-test). Gene ontology enrichment analysis of the overlapping 22 genes did not demonstrate any significantly enriched biological processes. The volcano plot showing all genes detected by RNA-seq (H). Colored dots represent genes with ≥3 DMCs (hypo- or hypermethylated) in their promoters; dotted line indicates significance cutoff set at q=0.05.

Extended Data Figure 4



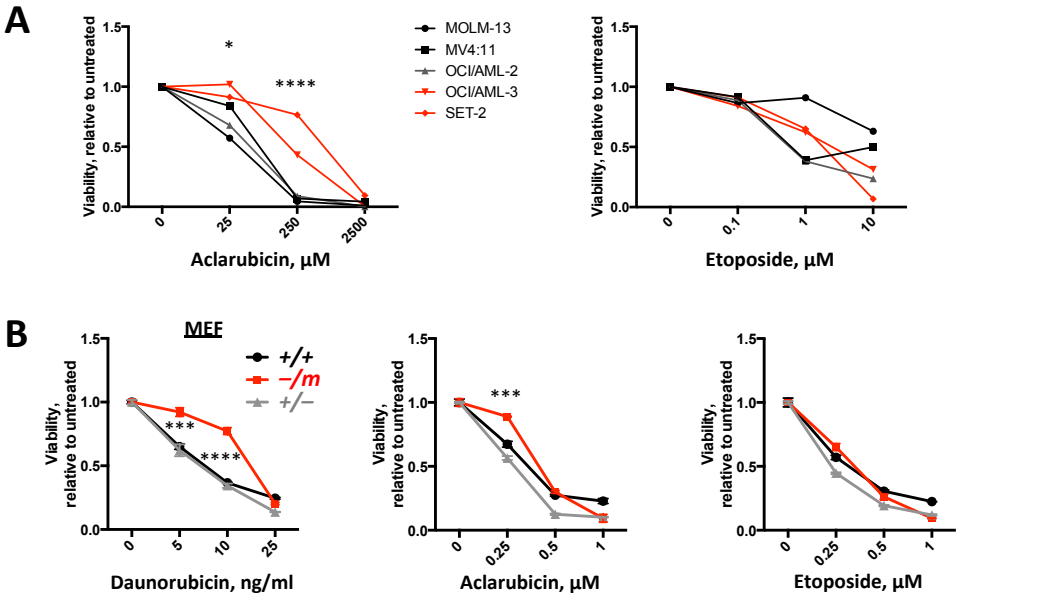
Extended Data Figure 4. *DNMT3A* mutational status does not affect cellular efflux, metabolism, or compartmentalization of daunorubicin. (A) Expression (normalized read counts) of *ABCB1* (MDR1), *ABCC1* (MRP1), *ABCC2* (MRP2), and *ABCG2* (BCRP) – the major ABC-type transporters implicated in chemotherapy resistance in multiple cancer types[11,12] – in AML samples with different *DNMT3A* mutational status. RNA-seq data were retrieved from the TCGA database (* $p<0.05$, ** $p<0.01$). (B) Immunoblotting analysis of the protein levels of MDR1 (P-gp) and MRP1 in MOLM-13 cells ectopically expressing wildtype or mutant forms of *DNMT3A*, or empty vector, overtime after daunorubicin exposure. This is a part of the same experiment also shown in Figure 3B. (C) Intracellular levels of daunorubicin measured over time by its fluorescence in MOLM-13 cells retrovirally expressing two R882 *DNMT3A* mutant isoforms or wild-type, or empty vector control. Cells were continuously exposed to indicated concentrations of daunorubicin. Before addition of daunorubicin cells were pre-treated with verapamil, a wide-spectrum inhibitor of ion channels, to determine the function of drug efflux transporters[13]. (D) Intracellular levels of daunorubicin as evidenced by its autofluorescence in bone marrow cells harvested from *Dnmt3a*^{R878H} and wild-type mice treated *in vitro* as indicated. (E) Intracellular levels of daunorubicin and its major metabolite daunorubicinol (DnR-ol) were measured by mass-spectroscopy in MEF cells with different *Dnmt3a* genotypes at indicated time points after addition of the drug to the culture media. (F) Intracellular compartmentalization[13] of daunorubicin (red) was examined in MEF cells with different *Dnmt3a* status at indicated time points by fluorescence microscopy. Cells were immunostained for γH2A.X (green) to control for DNA damage signaling, and counterstained with DAPI (blue) to visualize nuclei. Bar – 100 μm.

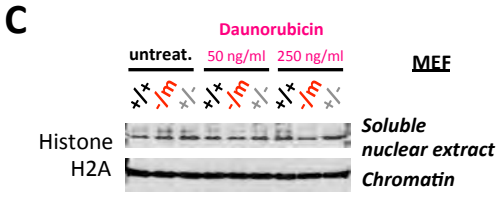
Extended Data Figure 5

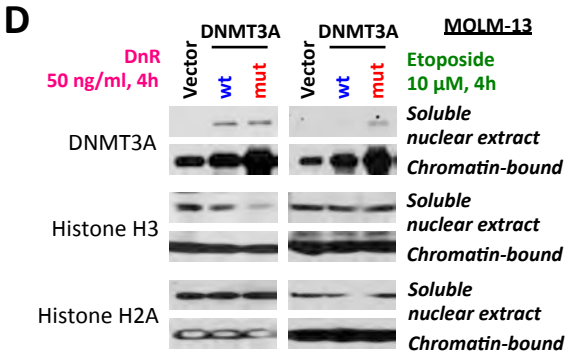


Extended Data Figure 5. Cells with mutant DNMT3A have attenuated response to anthracycline-induced DNA damage. (A) Analysis of DNA damage response pathways in MOLM-13 cells retrovirally expressing *DNMT3A*^{mut}, wild-type or empty vector control at different time points after DNA double-strand breaks induced by etoposide treatment. (B) Quantitative RT-PCR analysis of expression of p53-target genes *CDKN1A*, *BAX*, and *PUMA* after daunorubicin treatment in cells ectopically expressing wild-type or mutant forms of DNMT3A, compared to empty vector control (* p<0.05, ** p<0.01, *** p<0.001). (C) Analysis of γH2A.X signal in 32D cells retrovirally expressing wild-type *DNMT3A* or *DNMT3A*^{mut} after 16 hour exposure to daunorubicin (25 ng/ml) *in vitro*, measured by intracellular flow cytometry. (D) γH2A.X levels in spleens of *Dnmt3a*^{mut} and wild-type mice after one-time doxorubicin treatment (3 mg/kg) *in vivo* as detected by immunohistochemistry. Arrowheads mark γH2A.X-positive nuclei; bar – 100 μm. (E) GSEA analysis of the G2/M checkpoint signature from the Hallmark collection in the bone marrow LSK cells from *Dnmt3a* haploinsufficient mice compared to *Dnmt3a*^{WT} animals and in peripheral blood mononuclear cells of AML patients carrying non-R882 *DNMT3A* mutations compared to *DNMT3A*^{WT} cases from TCGA; gene expression by RNA-sequencing. (F, G) Alkaline comet assay on U-937 cells ectopically expressing *DNMT3A* wild-type or two *DNMT3A* R882 mutants (F) and on sorted MPPs derived from *Dnmt3a*^{mut} or wild-type animals (G) after 16 hour exposure to 25 ng/ml daunorubicin *in vitro* (* p<0.05, **** p<0.0001, Mann-Whitney rank sum test). U-937 cells are characterized by a *TP53* mutation. (H) Random mutagenesis measured by number of 6-TG-resistant colonies in MOLM-13 cells retrovirally expressing DNMT3A^{mut} or wild-type, or empty vector control, grown in semisolid media (* p<0.05, ** p<0.01). Colony counts were normalized by plating efficiency.

Extended Data Figure 6

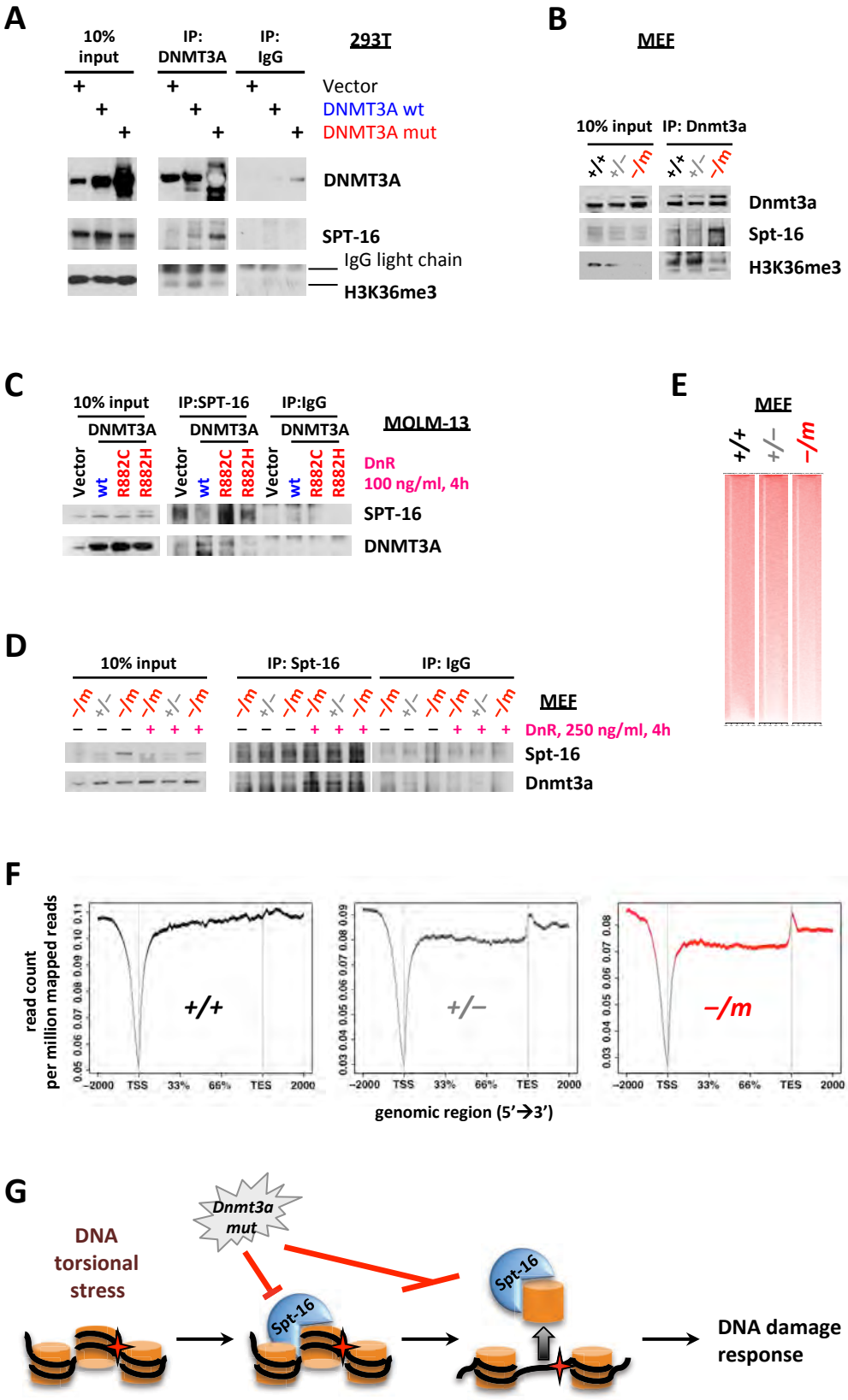






Extended Data Figure 6. Cells with mutant DNMT3A are differentially resistant to DNA torsional stress induced by anthracyclines rather than DNA double-strand breaks induced by topoisomerase inhibitor etoposide. (A) Response to DNA torsional stress inducer aclarubicin (left panel) and to DNA dauble-strand break inducing topoisomerase II inhibitor etoposide in *DNMT3A* R882-mutant cell lines (red, OCI/AML-3 and SET-2) compared to a non-R882 cell line (grey, OCI/AML-2) and *DNMT3A* wild-type cells (black, MOLM-13 and MV4:11) measured as viability relative to untreated control (* $p<0.05$, **** $p<0.0001$, pair-wise comparisons between *DNMT3A* R882-mutant cells and cells of other genotypes). (B) Response to anthracyclines daunorubicin (left panel) and aclarubicin (middle panel) and to topoisomerase II inhibitor etoposide (right panel) in MEFs expressing a single copy of mutant *Dnmt3a* (red, *–/m*) compared to MEFs with one or two wild-type copies of *Dnmt3a* (grey, *+/–*, or black, *+/+*, respectively; *** $p<0.001$, pair-wise comparisons between *–/m* cells and cells of other genotypes). (C) Distribution between chromatin-bound and free nuclear states of histone H2A after 4 hours of daunorubicin treatment in indicated doses in mouse embryonic fibroblasts derived from animals of indicated genotypes. (D) Distribution between chromatin-bound and free nuclear states of DNMT3A and histones H3 and H2A after 4 hour exposure to 50 ng/ml daunorubicin (left) or 10 μM etoposide in MOLM-13 cells retrovirally transduced with wild-type or mutant versions of DNMT3A, or with empty vector control. Immunoblotting analysis.

Extended Data Figure 7

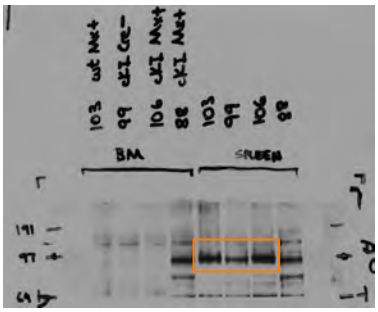


Extended Data Figure 7. Direct protein-protein interaction between Dnmt3a and Spt-16. (A, B) Analysis of DNA-independent (nuclease-treated extracts) protein-protein interaction between DNMT3A and SPT-16 by co-immunoprecipitation (co-IP) followed by Western blotting. Co-IP with DNMT3A-specific antibodies in 293T cells retrovirally expressing DNMT3A^{R882} or wild-type, or empty vector control (A). Co-IP with Dnmt3a-specific antibodies using nuclease-treated extracts from *Dnmt3a* wild-type MEFs (*+/+*), or MEF cells endogenously expressing either one wild-type copy of *Dnmt3a* (*+/–*) or one mutant copy of *Dnmt3a* (*–/m*) (B). Binding to tri-methylated histone H3K36me3, of which DNMT3A is a known reader, was used as a positive control. (C, D) Analysis of DNA-independent protein-protein interaction between endogenous SPT-16 and DNMT3A with and without daunorubicin-induced DNA torsional stress. Co-IP with SPT-16-specific antibodies using extracts from MOLM-13 cells retrovirally expressing either wild-type or two different mutants of DNMT3A, or empty vector control, treated with daunorubicin (C), or from MEF cells endogenously expressing either one wild-type copy of *Dnmt3a* (*+/–*) or one mutant copy of *Dnmt3a* (*–/m*), with and without daunorubicin treatment (D). Whole protein lysates were pre-treated with DNA- and RNA-non-discriminating nuclease to remove nucleic-acid-dependent interactions. (E, F) ChIP-sequencing analysis of DNA binding of Dnmt3a in mouse embryonic fibroblasts that express 2 wild-type copies of *Dnmt3a* (*+/+*), one wild-type copy of *Dnmt3a* (*+/–*), or one mutant copy of *Dnmt3a* (*+/m*). Heat maps (tornado plots) across gene bodies (E), and ChIP signal histograms (read count per million mapped reads) across gene bodies defined from annotated transcription start sites to transcription termination sites, ±2kb (F). Please note differences in scale. (G) Proposed model of the molecular mechanism whereby DNMT3A^{mut} contributes to anthracycline resistance through impaired nucleosome remodeling. Following anthracycline binding and resulting distortion of the DNA topology, adjacent regions

undergo chromatin remodeling starting with nucleosome eviction to provide a landing pad for the DNA damage signaling and repair machinery. Release of the histones from the DNA is being facilitated by nucleosome chaperones such as the FACT complex subunit SPT-16. Presence of mutant DNMT3A attenuates recruitment of SPT-16 to DNA distortion sites through direct protein-protein interaction, possibly through sequestration of SPT-16 away from chromatin, leading to impaired histone clearance. This results in the failure to initiate DNA damage signaling events driving apoptosis and/or DNA repair.

Extended Data Figure 8. Uncropped images of Western blotting experiments presented in main and supplementary figures.

Figure 1C. *



* #88 – animal from contaminated line, developed T-ALL

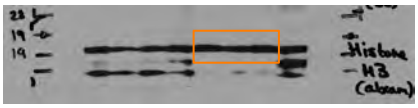
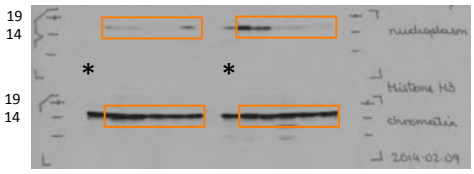


Figure 4B.

* Cell line did not authenticate



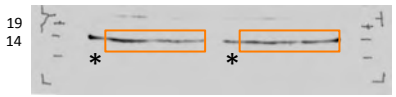


Figure 4C.

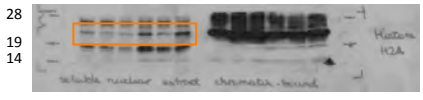
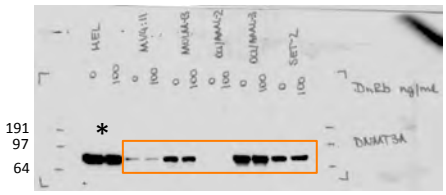
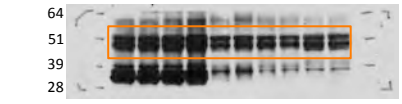
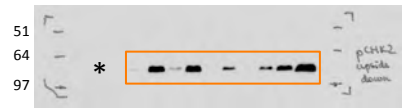


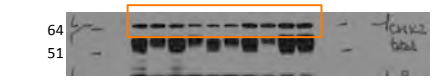
Figure 3A.

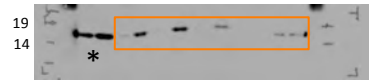


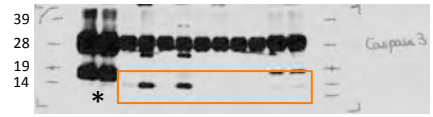
* Cell line did not authenticate











Extended Data Figure 8 (cont'd)

Figure 3B.

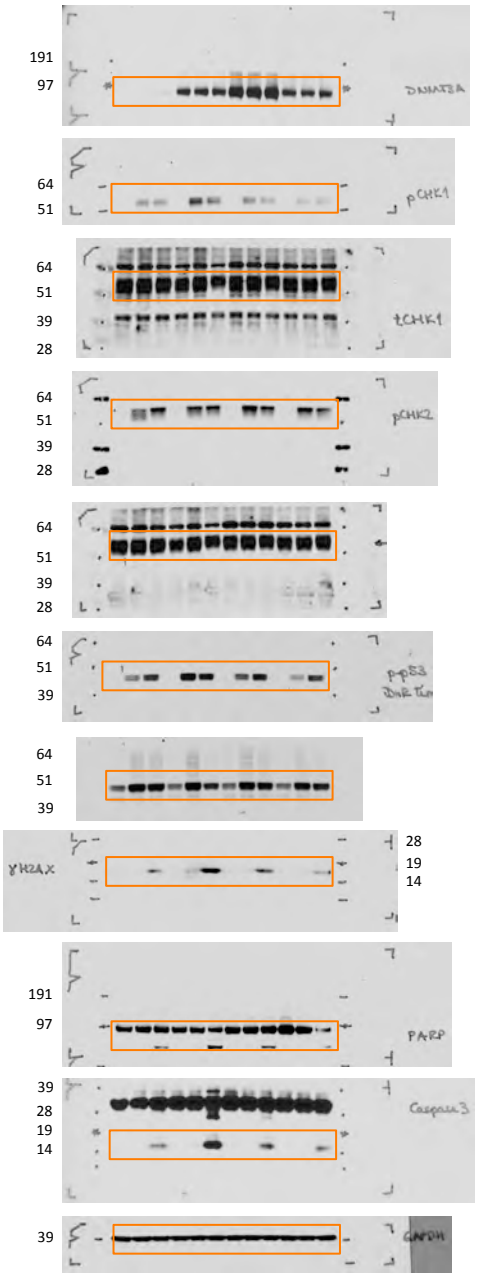


Figure 4D.

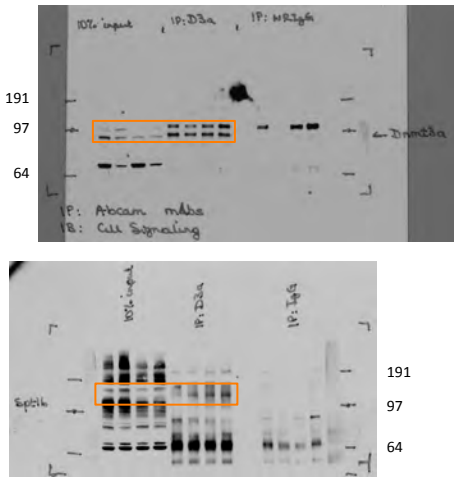


Figure 4E.

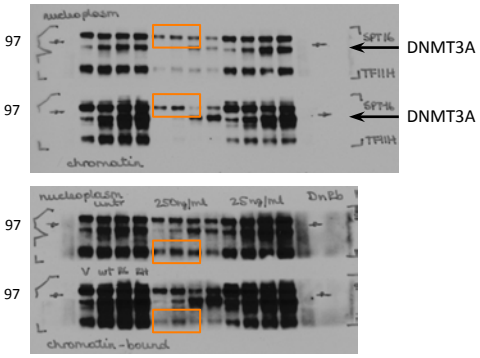


Figure 4F.

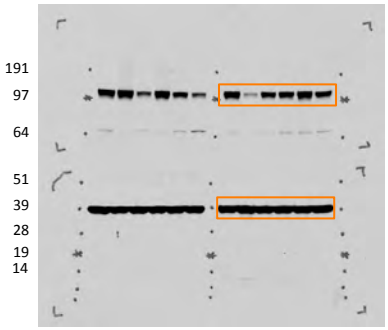
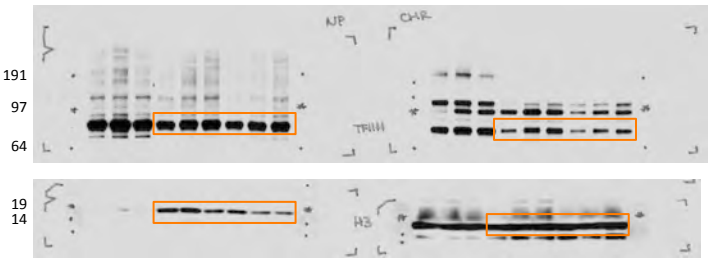
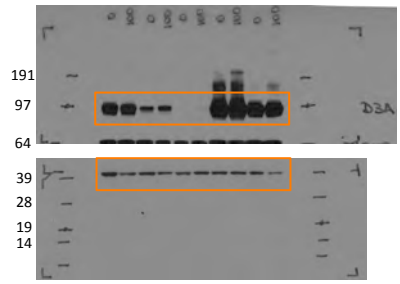


Figure 4G.

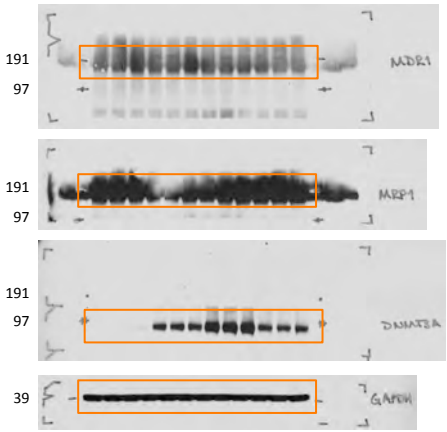


Extended Data Figure 8 (cont'd)

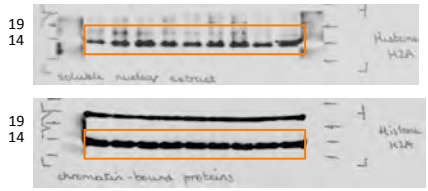
Extended Data Fig. 3B.



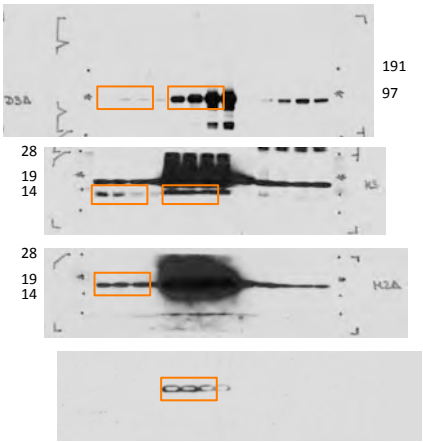
Extended Data Fig. 4B.

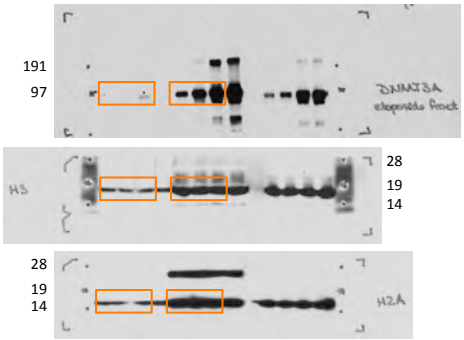


Extended Data Fig. 6C.

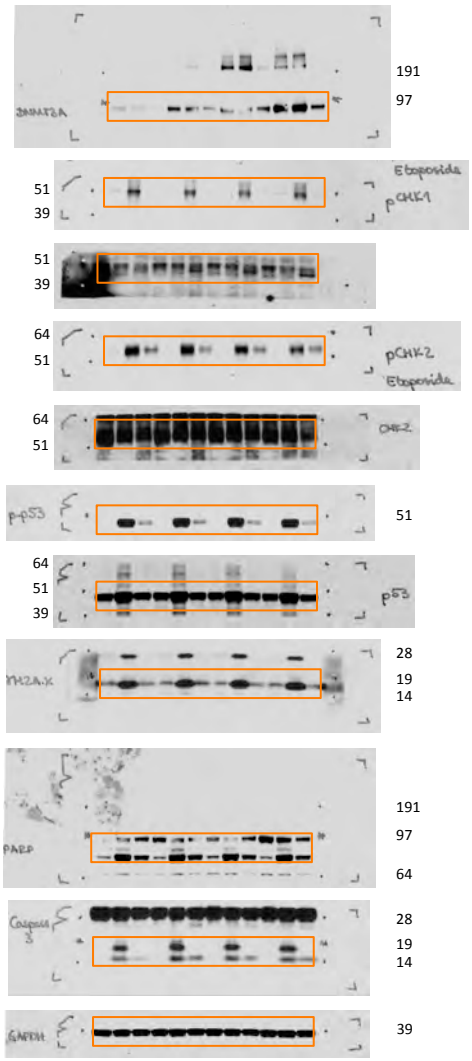


Extended Data Fig. 6D.



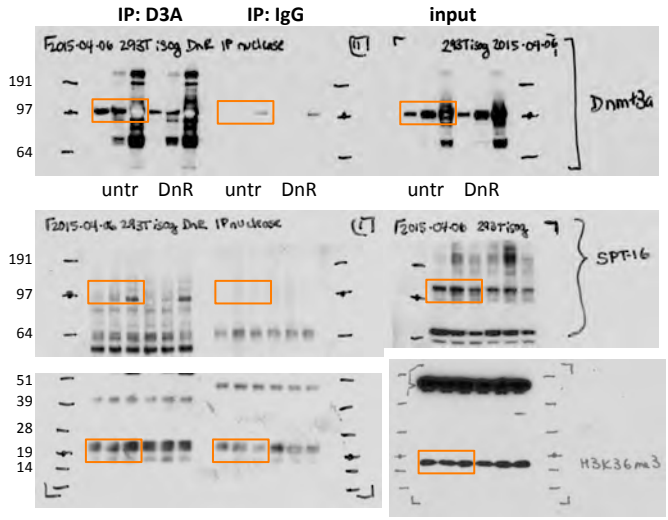


Extended Data Fig. 5A.

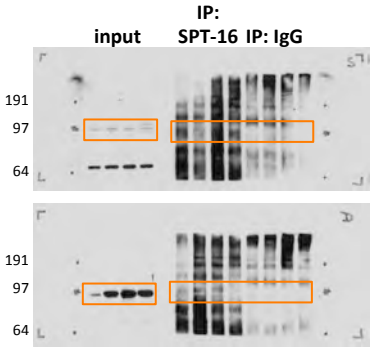


Extended Data Figure 8 (cont'd)

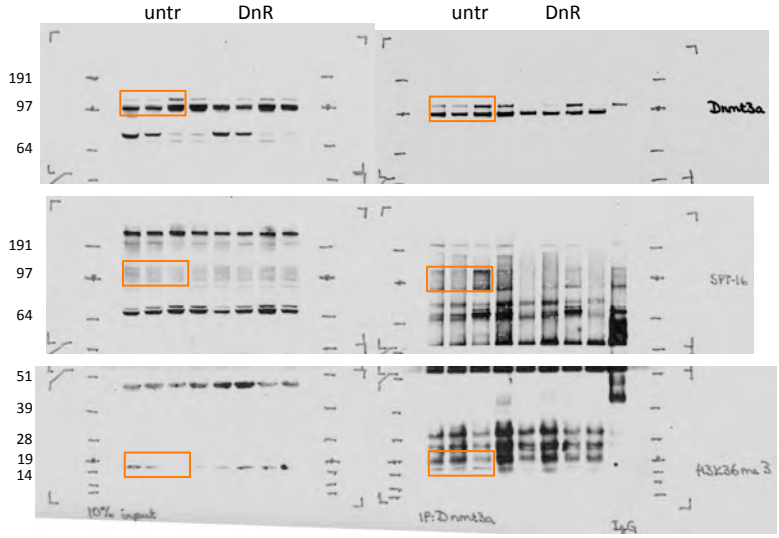
Extended Data Fig. 7A.



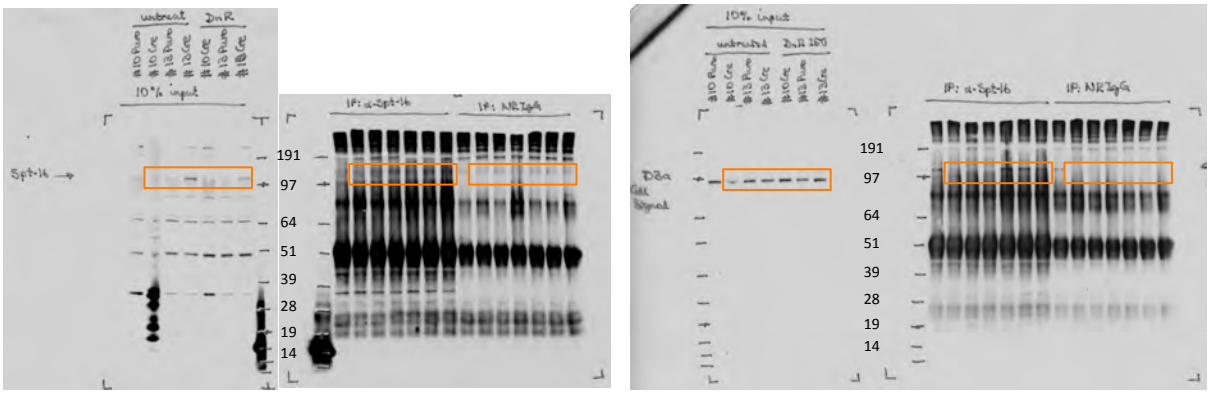
Extended Data Fig. 7C.



Extended Data Fig. 7B.



Extended Data Fig. 7D.



Supplementary references

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
2. Anders S, Pyl PT, Huber W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.
3. 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.
4. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
5. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-273.
6. Akalin A, Garrett-Bakelman FE, Kormaksson M, et al. Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet*. 2012;8(6):e1002781.
7. Garrett-Bakelman FE, Sheridan CK, Kacmarczyk TJ, et al. Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *J Vis Exp*. 2015(96).
8. Akalin A, Kormaksson M, Li S, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 2012;13(10):R87.
9. Li S, Garrett-Bakelman FE, Akalin A, et al. An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinformatics*. 2013;14 Suppl 5:S10.
10. Deshpande AJ, Deshpande A, Sinha AU, et al. AF10 regulates progressive H3K79 methylation and HOX gene expression in diverse AML subtypes. *Cancer Cell*. 2014;26(6):896-908.
11. Baguley BC. Classical and Targeted Anticancer Drugs: An Appraisal of Mechanisms of Multidrug Resistance. *Methods Mol Biol*. 2016;1395:19-37.
12. Gottesman MM, Lavi O, Hall MD, Gillet JP. Toward a Better Understanding of the Complexity of Cancer Drug Resistance. *Annu Rev Pharmacol Toxicol*. 2016;56:85-102.
13. Ihlefeld K, Vienken H, Claas RF, et al. Upregulation of ABC transporters contributes to chemoresistance of sphingosine 1-phosphate lyase-deficient fibroblasts. *J Lipid Res*. 2015;56(1):60-69.

Extended Data Table 1. Minimal residual disease analysis by mutational status in 144 evaluable patients from the E1900 cohort.

Mutation	Number of Patients WT and No MRD	Number of Patients Mutant and No MRD	Number of Patients WT and MRD	Number of Patients Mutant and MRD	Proportion Mutant Among No MRD	Proprotion Mutant Among MRD	Proportion MRD Among WT	Proportion MRD Among Mutant	**Adjusted *p*-value**
***DNMT3A* all**	40	4	68	30	9.09%	30.61%	62.96%	88.24%	**0.005**
***DNMT3A* R882**	42	2	76	22	4.55%	22.45%	64.41%	91.67%	**0.007**
***DNMT3A* other**	42	2	89	9	4.55%	9.18%	67.94%	81.82%	**0.503**
IDH	37	8	83	16	17.78%	16.16%	69.17%	66.67%	**0.813**
TET2	41	4	89	10	8.89%	10.10%	68.46%	71.43%	**1.000**
ASXL1	43	1	97	2	2.27%	2.02%	69.29%	66.67%	**1.000**
FLT3	32	13	63	36	28.89%	36.36%	66.32%	73.47%	**0.450**
NPM1	33	12	68	31	26.67%	31.31%	67.33%	72.09%	**0.695**

Extended Data Table 2. Molecular persistence of *DNMT3A*^{R882} mutations in 9 paired diagnosis and post-induction samples from the E1900 cohort.

Case ID	Presentation		Follow-up			Outcome	
	***DNMT3A* mut**	**Other mutations**	**VAF (*DNMT3A*)**	**MRD**	**Other mutations cleared?**	**EFS**	
15429	R882C	*NPM1*	0.219	+	Y	457	
20889	R882C	TET2 (2 mut), NPM1	0.375	+	N	134	
14787	R882H	*IDH1*, *NPM1*, *FLT3*	0	+	Y	106	
15492	R882H	*NPM1*, *FLT3*	0.155	+	N	244	
15705	R882H	*NRAS*, *NPM1*	0.082	-	Y	1926	complete remission
16308	R882H	*TET2* (2 mut), *NPM1*, *FLT3*	0.227	+	N	531	
16446	R882H	*IDH1*, *NPM1*, *KRAS*	0.221	+	Y	283	
16758	R882H	*NRAS*, *NPM1*	0.092	-	Y	1638	complete remission
20164	R882H	*NPM1*, *ASXL1*	0.071	+	N	238	

EFS – event-free survival

Extended Data Table 3. Clinical characterization of the primary *FLT3^{ITD}* and *NPM1^{c}*-positive AML samples with and without *DNMT3A* mutations used to evaluate response to daunorubicin and doxorubicin *in vitro*.

Sample ID	Karyotype	DNMT3A	FLT3-ITD	NPM1	Induction	Outcome of induction	Consolidation	OS	Comment
1	Normal	neg	pos	MUT	3+5+7	CR	1 x FLAG, 3 x IDAC	229	
2	N/A	neg	pos	MUT	3+5+7	CR	1 x FLAG, 3 x IDAC	270	
3	Normal	neg	pos	MUT	3+5+7, 2+5+5	CR	4 x IDAC	342	
4	Normal	neg	pos	MUT	3+5+7	CR	3 x IDAC	394	
5	Normal	neg	pos	MUT	3+5+7	CR	1 x IDAC	259	allo HSCT
6	Normal	MUT	pos	MUT	3+5+7	CR	3 x HiDAC	423	
7	45,X,-Y	MUT	pos	MUT	3+5+7, MiDAC, FLAG	NR		211	
8	Normal	MUT	pos	MUT	3+5+7, MiDAC	CR	1 x HiDAC	197	allo HSCT
9	Normal	MUT	pos	MUT	3+5+7	CR	2 x HiDAC	1628	allo HSCT

Extended Data Table 4. Clinical characterization of the primary AML samples used in Comet assay.

Sample ID	AML status	Treatment status	Karyotype	Blast %	*DNMT3A* mutation	Other mutations
114	Newly diagnosed	Untreated	Partial del(9p)	71.0%	Negative	*IDH2* exon 4; *NPM1* exon 11; *FLT3* ITD
289	Newly diagnosed	Untreated	Normal	52.8%	Negative	*IDH2* exon 4; *NPM1* exon 11; *FLT3* ITD
17	Newly diagnosed	Untreated, post-MPN AML	Partial del(21q)	84.3%	Exon 23 p.R882H (c.2645G>A)	*TET2* exon 10; *JAK2* exon 14; *RUNX1* exon 4
814	Newly diagnosed	Untreated s/p 7+3 (Ida+AraC);	Tri(13)	83.9%	Exon 23 p.R882C (c.2644C>T)	None detected
208	Refractory	3 cycles consolidation HiDAC; FLAG	N/A	90.0%	Exon 20 p.V778E (c.2333T>A)	*NPM1* exon 11; *IDH2* exon 4
278	Newly diagnosed	Untreated	Normal	32.5%	Exon 10 p.Q420X (c.1204C>T)	*NRAS* exon 2, *IDH1* exon 4, *NPM1* exon 11

Extended Data Table 6. Summary statistics for ERRBS assay on LSK cells derived from *Dnmt3a*^{wt} and *Dnmt3a*^{mut} mice.

Animal ID	Genotype	CpG number	Mean CpG coverage	Mapping efficiency	Conversion rate	Total number of reads	Aligned reads with unique matches
820	*wt Mx+*	1960026	163.12	69.2	99.88	169896422	117517219
830	*cKI Mx+*	1837033	190.5	68.5	99.92	170026268	116416635
829	*cKI Mx+*	1836807	193.45	69.1	99.92	171457184	118456647
823	*wt Mx+*	1861561	186.8	68.8	99.9	170295184	117207919
825	*wt Mx+*	1886051	180.44	69.1	99.9	175019619	121022038
824	*cKI Mx+*	1973664	171.41	68.6	99.93	170118037	116720044
827	*cKI Mx+*	1883402	184.02	68.7	99.92	169500983	116504751