Supplemental Data

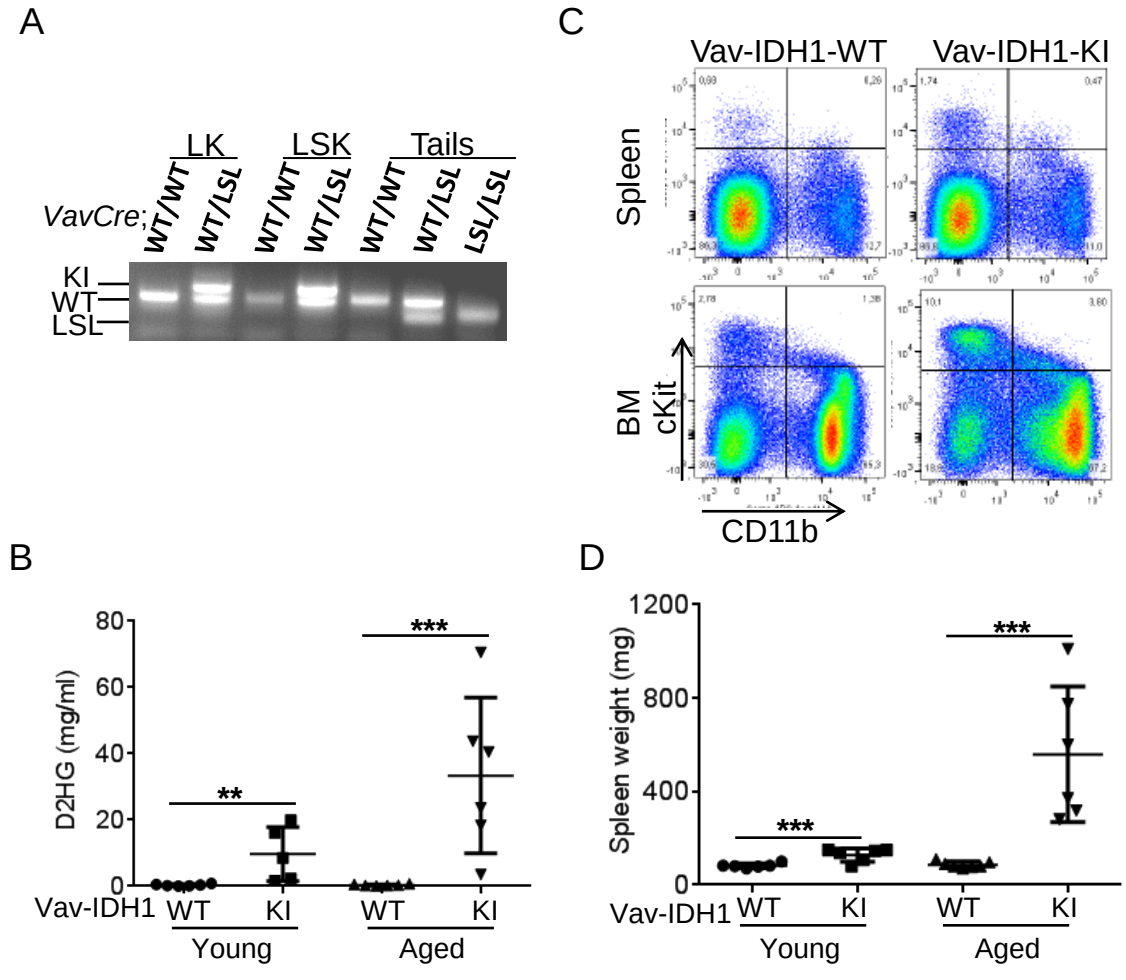


Figure S1, related to Figure 2. Vav-IDH1-KI mice show high urinary D2HG levels plus splenomegaly and anemia. (A) Genotyping PCR of genomic DNA from bone marrow (BM) LK cells and LSK cells and tails of *Vav-Cre*;*Idh1* WT/WT, WT/LSL, and LSL/LSL mice. Bands associated with the IDH1 R132Q knock-in (KI), WT and LSL alleles are indicated. The IDH1 R132Q mutation has been observed in human chondrosarcoma and causes the same enzymatic gain-of-function as all other IDH1 R132 mutations identified to date (Hirata et al., 2015). (B) LC-MS determination of urinary D2HG levels in young (3-4 month old) and aged (7-10 month old) Vav-IDH1-KI (KI) and control (WT) mice. Data for individual mice are shown as well as the group mean ± SD. (C) Representative flow cytometric analyses of cKit vs. CD11b expression in spleen and BM cells from aged Vav-IDH-WT and Vav-IDH-KI mice. (D) Weights of spleens from the young and aged Vav-IDH1-KI and control littermate mice. Data for individual mice are shown as well as the group mean ± SD. **$p<0.01$, ***$p<0.005$, unpaired Student's t-test.

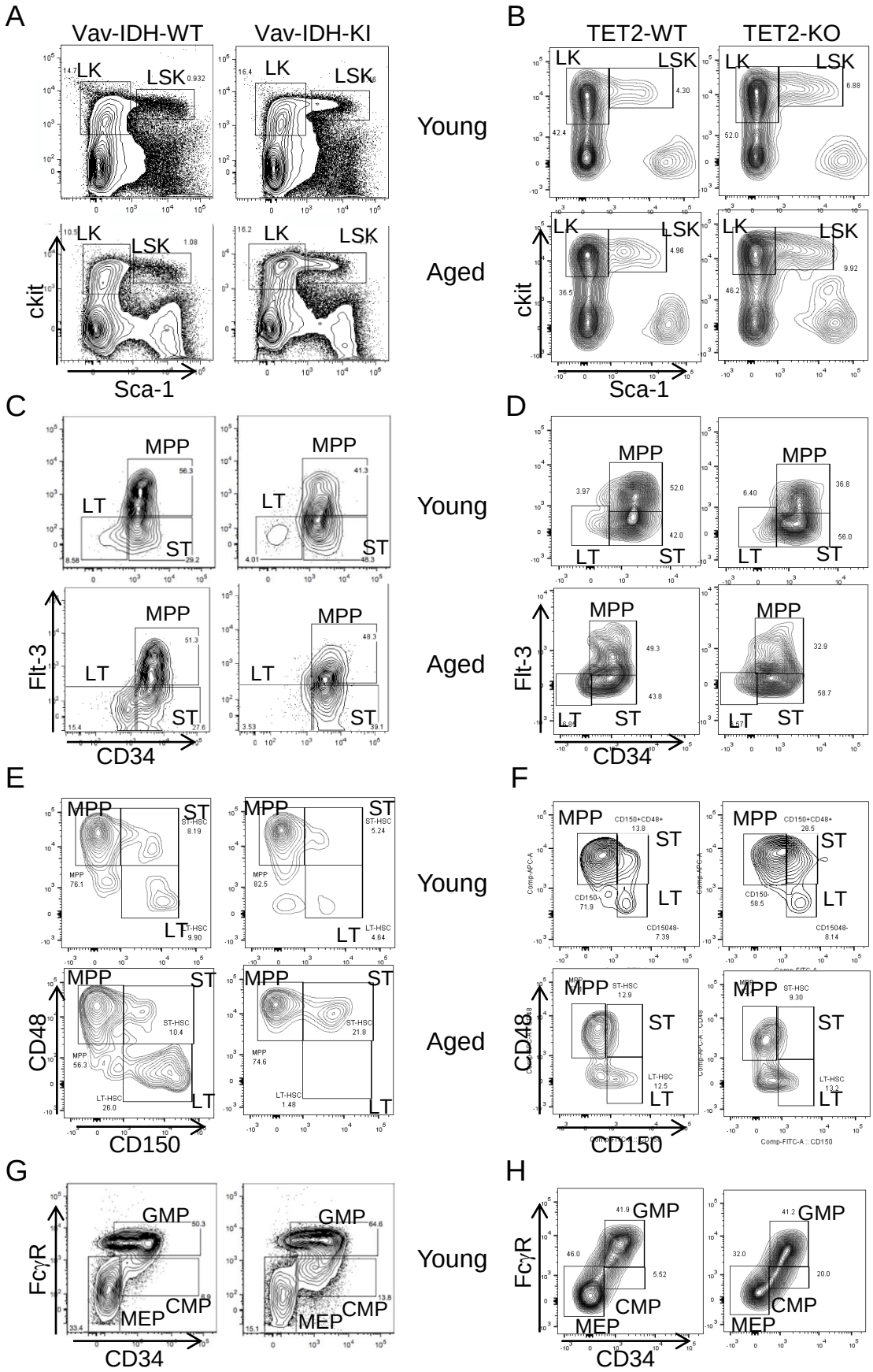
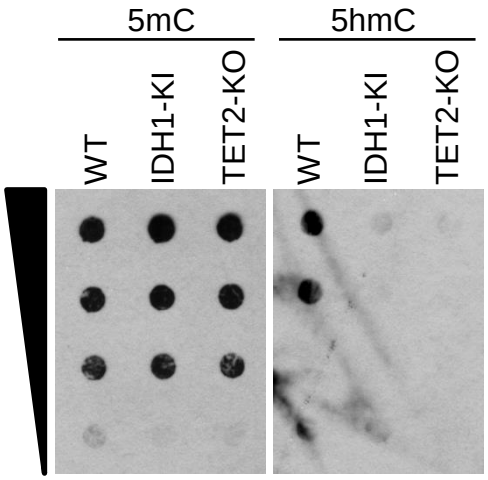

A
Vav-IDH-WT
Vav-IDH-KI
LK
LSK
Young
Aged
ckit
Sca-1
B
TET2-WT
TET2-KO
C
MPP
LT
ST
Flt-3
CD34
D
E
ST-HSC
LT-HSC
CD48
CD150
F
CD150+CD48+
CD150-
CD15048-
G
GMP
MEP
CMP
FcγR
H

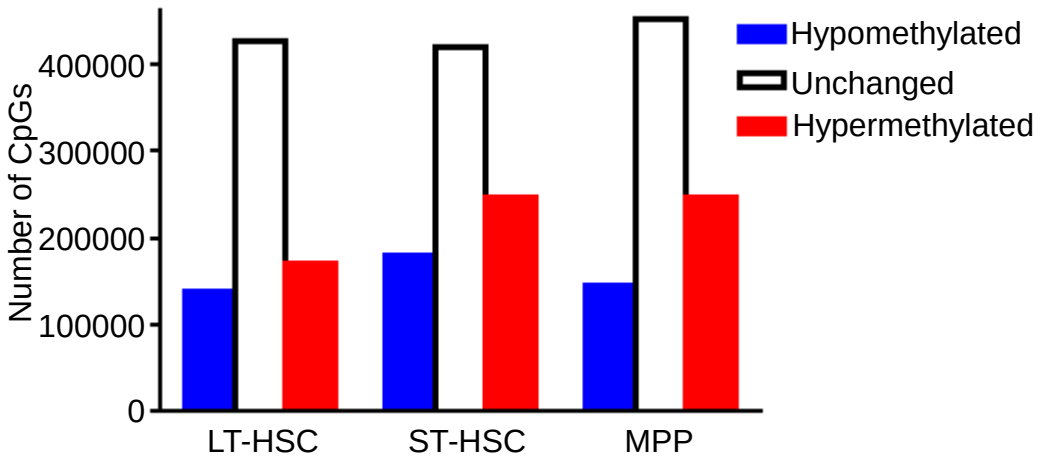
Figure S2, related to Figure 3. Comparative flow cytometric analyses of Vav-IDH1-KI and TET2-KO mice.

(A, B) Representative flow cytometric analyses of cKit vs. Sca1 expression by LK ($Lin^-Sca1^-cKit^+$) and LSK ($Lin^-Sca1^+cKit^+$) cells among viable Lin^- cells from young and aged Vav-IDH-WT and Vav-IDH-KI mice (A), and young and aged TET2-WT and TET2-KO mice (B). (C, D) Representative flow cytometric analyses of Flt-3 vs. CD34 expression by LT-HSC ($CD34^-Flt\text{-}3^-$), ST-HSC ($CD34^+Flt\text{-}3^-$) and MPP ($CD34^+Flt\text{-}3^+$) among viable LSK cells from young and aged Vav-IDH-WT and Vav-IDH-KI mice (C), and young and aged TET2-WT and TET2-KO mice (D). (E, F) Representative flow cytometric analyses of CD48 vs. CD150 expression among viable LSK cells from young and aged Vav-IDH1-KI and control Vav-IDH1-WT mice (E), and young and aged TET2-WT and TET2-KO mice (F). LT-HSC, $CD150^+CD48^-$; ST-HSC, $CD150^+CD48^+$; MPP, $CD150^-CD48^+$. (G, H) Representative flow cytometric analyses of FcγR vs. CD34 expression by GMP ($Fc\gamma R^+CD34^+$), CMP ($Fc\gamma R^-CD34^+$) and MEP ($Fc\gamma R^-CD34^-$) among viable LK cells from young Vav-IDH1-KI (KI) and Vav-IDH1-WT mice (G), and young TET2-WT and TET2-KO mice (H). In all cases, at least 3 mice/genotype were examined.

A



C



B

Sample	Total # of reads	Alignment rate	# of uniquely aligned reads	Bisulfite conversion rate	Total # of CpGs*	Mean coverage per CpG*
WT LT-HSC	25775716	59.97%	15458344	98.60%	958850	51
WT ST-HSC	24695720	61.93%	15294790	99.00%	936607	42
WT MPP	25349828	57.50%	14575955	98.68%	931370	48
KI LT-HSC	24700909	57.12%	14108078	99.21%	854387	63
KI ST-HSC	23043855	61.35%	14137102	99.24%	1040407	37
KI MPP	28008363	60.89%	17055021	98.77%	1069270	49

*After 10x coverage filter

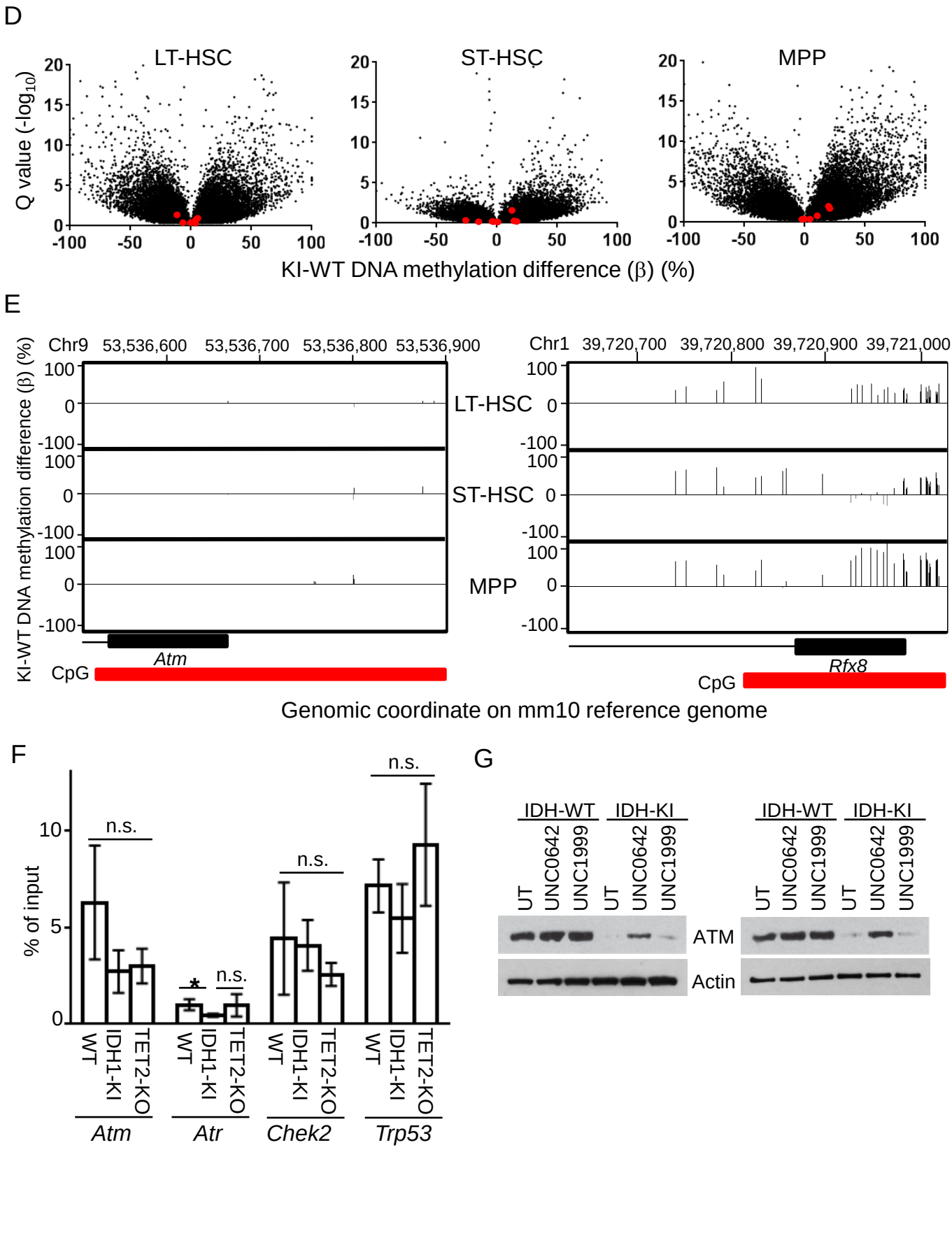

D
LT-HSC
ST-HSC
MPP
Q value (-log10)
KI-WT DNA methylation difference (β) (%)
E
Chr9 53,536,600 53,536,700 53,536,800 53,536,900
Chr1 39,720,700 39,720,800 39,720,900 39,721,000
KI-WT DNA methylation difference (β) (%)
LT-HSC
ST-HSC
MPP
Atm
Rfx8
CpG
Genomic coordinate on mm10 reference genome
F
% of input
n.s.
*
WT
IDH1-KI
TET2-KO
Atm
Atr
Chek2
Trp53
G
IDH-WT
IDH-KI
UT
UNC0642
UNC1999
ATM
Actin

Figure S3, related to Figure 5. The *Atm* promoter CpG island is not hypermethylated in Vav-IDH1-KI cells.

(A) Dot blot detection of 5hmC and 5mC in Lin^- cells from aged Vav-IDH-WT (WT), Vav-IDH-KI (IDH1-KI), and TET2-KO mice. (B) Total number of reads and aligned reads, alignment and bisulfite conversion rate, total number of CpGs covered, and mean CpG coverage from enhanced reduced representation bisulfite sequencing (eRRBS) of genomic DNA from LT-HSC, ST-HSC and MPP isolated from aged Vav-IDH1-KI (KI) and Vav-IDH1-WT (WT) (n=6 mice pooled/group). (C) Summary of numbers of hypermethylated, unchanged and hypomethylated CpG sites in genomic DNA isolated from LT-HSC, ST-HSC and MPP of Vav-IDH1-KI (KI) mice compared to littermate controls (WT) (n=6 mice pooled/group). (D) CpG sites (100,000) were randomly chosen from the set of CpG sites sequenced in each pair of WT and KI samples and plotted by difference in β values (KI – WT percentage DNA methylation) vs. the significance of the change (-log10(Q) value). Red dots represent CpG sites in the CpG island of the *Atm* promoter. (E) eRRBS profile of DNA methylation at the *Atm* promoter in the Vav-IDH1-KI and control LT-HSC, ST-HSC, and MPP populations examined in (D). The x-axis presents the genomic coordinates on the mm10 reference genome at positions 53536500-53536900 on chromosome 9 near the *Atm* gene (*Atm*) (left panel) and positions 39720600-39721020 on chromosome 1 near the *Rfx8* gene (*Rfx8*) (right panel). The *Rfx8* gene and its CpG island profile are shown an example of a differentially DNA methylated region: note the DNA hypermethylation of the *Rfx8* promoter in Vav-IDH1-KI cells compared to control cells. The y-axes show the change in DNA methylation levels (β values) between Vav-IDH1-KI and control LT-HSC, ST-HSC and MPP populations. Positive values indicate hypermethylation in Vav-IDH1-KI cells while negative values indicate hypomethylation. (F) hMeDIP-qPCR analysis of the indicated DDR genes in LSK cells from Vav-IDH-WT (WT), Vav-IDH-KI (IDH1-KI), and TET2-KO mice (n=3/group). Data are the mean ± SD. *$p<0.05$ by unpaired Student's t-test. n.s., not significant. (G) Immunoblot to detect the indicated proteins in LSK cells that were isolated from pooled Vav-IDH1-WT or Vav-IDH1-KI mice (n=5~10 pooled/genotype) and cultured for 3 days with/without 1 μM H3K9 methylation inhibitor UNC0642 or 1 μM H3K27 methylation inhibitor UNC1999. Two independent experiments were presented. Actin, loading control.

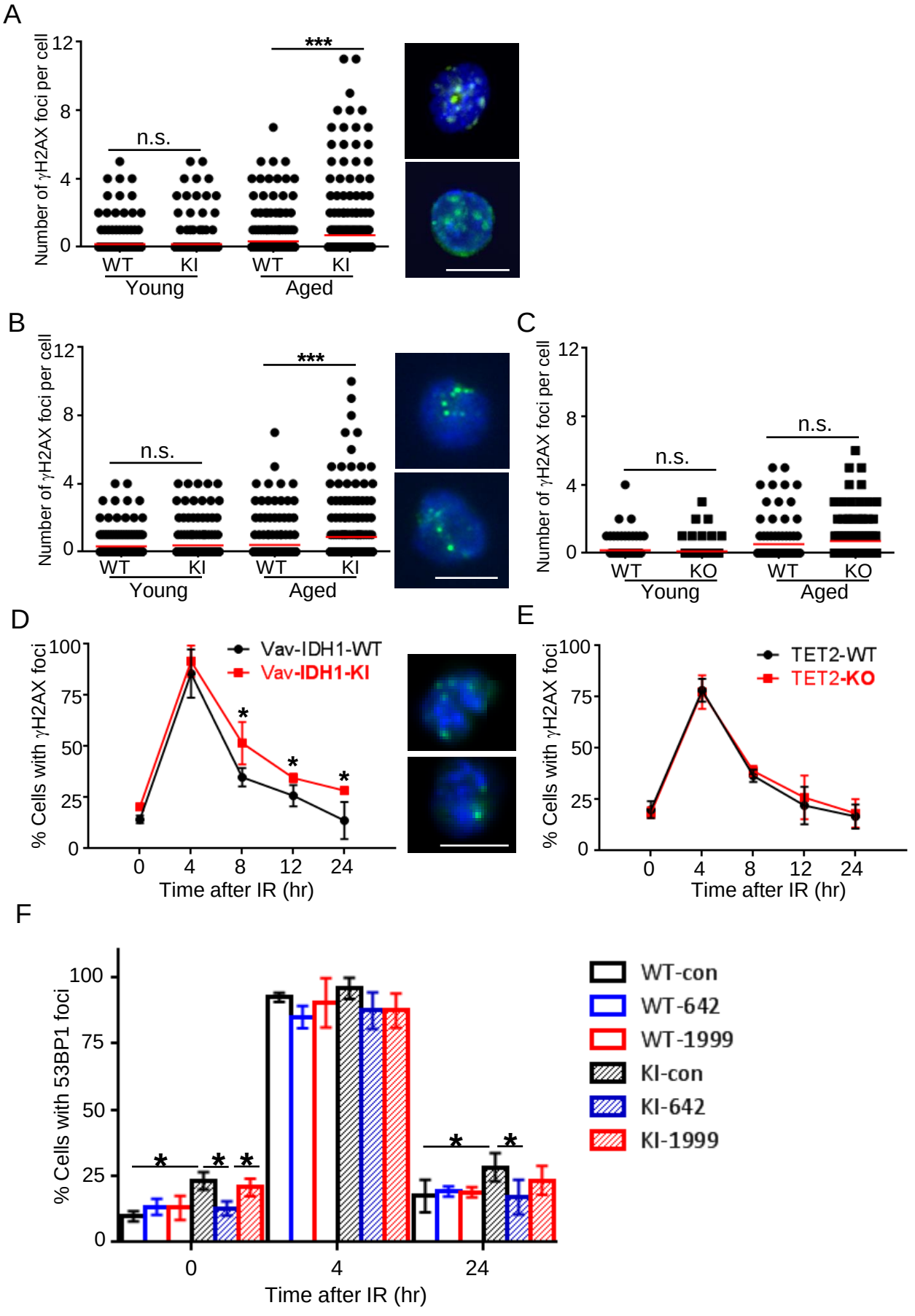

A
Number of γH2AX foci per cell
12
8
4
0
***
n.s.
WT
KI
WT
KI
Young
Aged
B
Number of γH2AX foci per cell
12
8
4
0
***
n.s.
WT
KI
WT
KI
Young
Aged
C
Number of γH2AX foci per cell
12
8
4
0
n.s.
n.s.
WT
KO
WT
KO
Young
Aged
D
% Cells with γH2AX foci
100
75
50
25
0
Vav-IDH1-WT
Vav-IDH1-KI
*
*
*
0
4
8
12
24
Time after IR (hr)
E
% Cells with γH2AX foci
100
75
50
25
0
TET2-WT
TET2-KO
0
4
8
12
24
Time after IR (hr)
F
% Cells with 53BP1 foci
100
75
50
25
0
*
*
*
*
*
0
4
24
Time after IR (hr)
WT-con
WT-642
WT-1999
KI-con
KI-642
KI-1999

Figure S4, related to Figure 6. Accumulation of γH2AX foci and impaired DNA damage repair capacity in aged Vav-IDH1-KI, but not TET2-KO, LT-HSC, and inhibitor of H3K9 methylation rescues ATM protein levels and DDR phenotype in IDH1-KI cells.

(A) Left: Quantitation of γH2AX foci counted by confocal microscopy in LSK cells (~300 cells/group from 3 mice) isolated from young and aged Vav-IDH1-KI (KI) and control (WT) mice **(**~300 cells/group from 3 mice**)**. Right: Two representative images of γH2AX foci in LSK cells from aged Vav-IDH1-KI mice. (B) Left: Quantitation of γH2AX foci in LT-HSC (~200 cells/group from 3 mice) isolated from young and aged Vav-IDH1-KI (KI) and control (WT) mice analyzed as for (A). Right: Two representative images of γH2AX foci in LT-HSC from aged Vav-IDH1-KI mice. (C) Quantitation of γH2AX foci in LT-HSC (~100 cells/group from 3 mice) isolated from young and aged TET2-KO and control (WT) mice analyzed as for (B). (A-C) Data are values for individual cells. Red lines, group means. ***$p<0.005$ by unpaired Student's t-test. n.s., not significant. (D, E) Kinetics of resolution of γH2AX foci over the indicated period after exposure to 0.5 Gy ionizing radiation (IR) in LT-HSC (~100 cells/group from 3 mice) from aged Vav-IDH-KI and control Vav-IDH-WT mice (D, left), and aged TET2-KO mice and control TET2-WT mice (E). (D, right) Two representative images of γH2AX foci in LT-HSC from aged Vav-IDH1-KI mice at 24 hr post-IR. (A,B,D) Scale bars, 5 μm. (F) Kinetics of 53BP1 foci formation during the indicated period in LT-HSC that were isolated from aged Vav-IDH1-KI (KI) and Vav-IDH1-WT (WT) mice (n=3/group), cultured with/without 1 μM UNC0642 (642) or 1 μM UNC1999 (1999) for 3 days, and exposed to IR (0.5 Gy). (D-F) Data are the mean ± SD. *$p<0.05$ by unpaired Student's t-test.

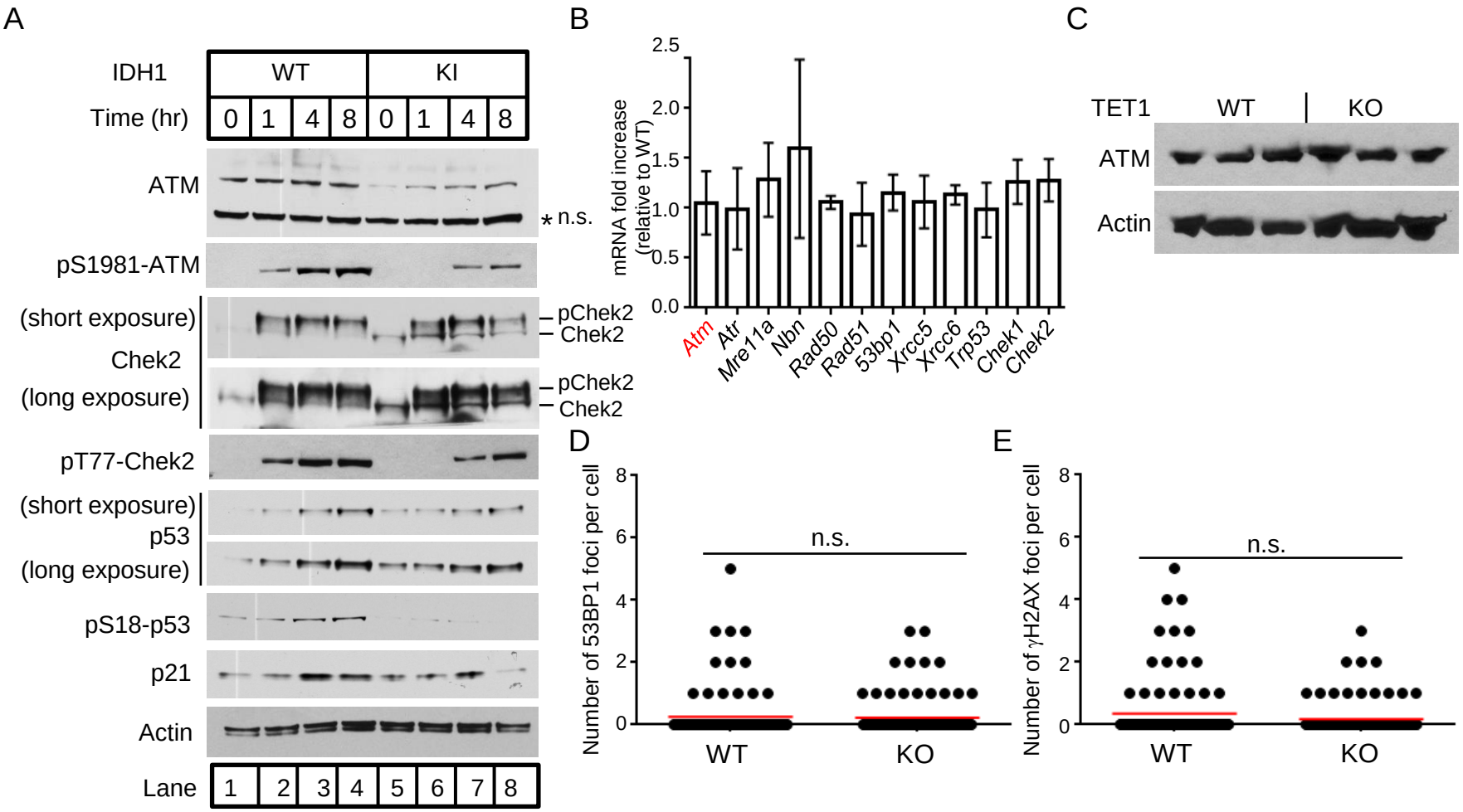

A
IDH1
WT
KI
Time (hr)
0 1 4 8 0 1 4 8
ATM
* n.s.
pS1981-ATM
(short exposure)
Chek2
(long exposure)
pChek2
Chek2
pChek2
Chek2
pT77-Chek2
(short exposure)
p53
(long exposure)
pS18-p53
p21
Actin
Lane
1 2 3 4 5 6 7 8
B
mRNA fold increase (relative to WT)
2.5
2.0
1.5
1.0
0.5
0.0
Atm
Atr
Mre11a
Nbn
Rad50
Rad51
53bp1
Xrcc5
Xrcc6
Trp53
Chek1
Chek2
C
TET1
WT
KO
ATM
Actin
D
Number of 53BP1 foci per cell
8
6
4
2
0
n.s.
WT
KO
E
Number of γH2AX foci per cell
8
6
4
2
0
n.s.
WT
KO

Figure S5, related to Figure 7. Mutant IDH1 impairs activation of ATM, Chek2 and p53 and the ATM downregulation and DNA damage accumulation caused by mutant IDH1 are TET1-independent. (A) Immunoblot to detect the indicated proteins in lysates of primary MEF that were prepared from *Creert*;*Idh1* LSL/WT (KI) and control littermate (WT) mice, subjected to IR (5 Gy) in culture, and incubated at 37°C for the indicated times. Short and long exposures of radiographs are indicated. Actin, loading control. *n.s. indicates a non-specific band. (B) qRT-PCR determination of the indicated mRNAs in LSK cells isolated from young *Tet1* KO mice relative to WT controls (n=3/group). Data are the mean ± SD. (C) Immunoblot to detect the indicated proteins in lysates of Lin^- cells isolated from aged *Tet1* KO and WT mice. (D, E) Quantitation of 53BP1 (D) and γH2AX (E) foci in LT-HSC (~100 cells/group from 3 mice) isolated from aged *Tet1* KO (KO) and *Tet1* WT (WT) mice. Data are values for individual LT-HSC. Red lines, group means. n.s., not significant.

Supplemental Experimental Procedures

D2HG measurement

D2HG was measured by liquid chromatography/mass spectrometry as previously described (Sasaki et al., 2012).

Histology

Histological examinations were performed as previously described (Sasaki et al., 2012). Briefly, spleens were fixed in formalin at 4°C and paraffin-embedded. Sternums were fixed in formalin, washed in PBS, and decalcified overnight in 4% formic acid/4% HCl. Sternums were then rinsed and neutralized in an ammonium hydroxide solution, washed in PBS, and fixed in 10% formalin overnight. Sternums were then treated as other tissues for processing and embedding in paraffin. Sections were stained with hematoxylin and eosin (H&E) using a standard protocol.

Flow cytometry

Flow cytometry was performed as previously described (Beerman et al., 2014; Hirata et al., 2015). Antibodies (Abs) used to detect cell surface markers were: PE-Cy7-anti-cKit (eBioscience, Cat.No: 25-1171-82) 1:100 or 200; FITC-anti-Sca1 (eBioscience, Cat.No: 11-5981-82) 1:50; PE-anti-Sca1 (eBioscience, Cat.No: 12-5981-83) 1:100; PE-anti-CD150 (Biolegend, Cat.No: 115904) 1:100; FITC-anti-CD150 (eBioscience, Cat.No: 11-1501-80) 1:100; APC-anti-CD48 (eBioscience, Cat.No: 17-0481-82) 1:100; APC-Cy7-streptavidin (Biolegend, Cat.No: 405208), 1:50; anti-APC-Cy7-anti-cKit (Biolegend, Cat.No: 105826) 1:200; APC-anti-Sca1 (Biolegend, Cat.No: 108112) 1:200; FITC-anti-CD34 (eBioscience, Cat.No: 11-0341-81) 1:33; Alexa700-anti-CD34 (eBioscience, Cat.No: 56-0341-82) 1:33; PE-anti-Flt3 (eBioscience, Cat.No: 12-1351-83) 1:200; APC-anti-Flt3 (Biolegend, Cat.No: 135310) 1:50; Pacific Blue-anti-CD3 (Biolegend, Cat.No: 100214) 1:200; Pacific Blue-anti-CD4 (Biolegend, Cat.No: 100428) 1:200; Pacific Blue-anti-CD8a (Biolegend, Cat.No: 100725) 1:200; Pacific Blue-anti-B220 (Biolegend, Cat.No: 103227) 1:200; Pacific Blue-anti-Mac1 (Biolegend, Cat.No: 101224) 1:200; Pacific Blue-anti-Gr1 (Biolegend, Cat.No: 108430) 1:200; Pacific Blue-anti-IL7ra (eBioscience, Cat.No:48-1271-82) 1:200; Pacific Blue-anti-Ter119 (Biolegend, Cat.No: 116232) 1:200; and PerCP-Cy5.5-anti-CD16/32 (BD Bioscience, Cat.No: 560540) 1:100.

CyTOF mass cytometry

Lineage negative (Lin^-) BM cells (2-3 x 10^6 per mouse) from one LysM-IDH1-KI mouse and a control littermate at age 3-4 months (young; Figure 1A) or 7-10 months (aged; Figure 1B) were fixed, permeabilized, and immunostained as previously described (Bendall et al., 2011). All CyTOF Abs were carrier-free and conjugated to the appropriate metal isotope using the MaxPar-X8 conjugation kit (Fluidigm Sciences). Events (~2-4 x 10^5 per sample) were acquired on a CyTOF I instrument (Fluidigm Sciences) and data were normalized to internal metal isotope bead standards as previously described (Bendall et al., 2011). Mass cytometry data plots, heatmaps, histograms and fold-change analyses were performed using Cytobank (www.cytobank.org). Antibodies used in this study recognized: pATM (Millipore; clone 10H11.E12), pChek2 (Cell Signaling Technology, clone C13C1), p53 (Cell Signaling Technology, clone 1C12), γH2AX (Millipore, clone JBW301), cleaved caspase-3 (BD Bioscience, clone C92-605), cleaved PARP (BD Bioscience, clone F21-852), pHistone H3 (Biolegend, clone HTA28), Ki67 (eBioscience, clone SolA15), pRb (BD Bioscience, clone J112-906), pMPM2 (Millipore, clone MPM-2), pSrc (BD Bioscience, clone K98-37), pSTAT5 (BD Bioscience, clone 47), pSTAT3 (BD Bioscience, clone 4), pAkt (Cell Signaling Technology, clone D9E), pERK (Cell Signaling Technology, clone D13.14.4E), β-catenin (Cell Signaling Technology, clone 6B3), pS6 (BD Bioscience, clone N7-548), pCREB (Cell Signaling Technology, clone 87G3), acetylated H4 (Millipore, polyclonal), Sca1 (Biolegend, clone E13-161.7), cKit (Fluidigm Science, clone 2B8), CD34 (Biolegend, HM34), and CD150 (Biolegend, clone TC15-12F12.2).

Cleaved caspase-3

LT-HSC (~500 cells/sample) were irradiated with 1 Gy and incubated at 37°C for 4 hr. Irradiated LT-HSC were plated onto poly-D-lysine-coated coverslips for 30 min, fixed with 4% paraformaldehyde/PBS for 30 min, permeabilized with 0.1% Triton X-100/PBS for 20 min, and incubated overnight at 4°C with rabbit anti-cleaved caspase-3 Ab (Cell Signaling Technology, Cat.No: #9661, 1:1,000). After 1 hr at RT, cells were washed with PBS, incubated with goat anti-rabbit Cy3 Ab (Jackson Immuno Research, 1:750) for 35 min, washed and counterstained with DAPI (0.2 μg/ml) for 10 min. Coverslips were treated and examined as above.

Drug treatments of cultured LSK cells and LT-HSC

Sorted LSK cells (5,000~10,000 cells/sample) or LT-HSC (300~1,000 cells/sample) from pooled Vav-IDH1-WT or Vav-IDH1-KI mice were cultured in Iscove's medium containing 10% fetal calf serum (FCS), 55 μM 2-

mercaptoethanol, 10 ng/ml SCF (Peprotech, Cat.No: #250-03), 10 ng/ml TPO (Peprotech, Cat.No: 315-14), 10 ng/ml Flt3 ligand (Peprotech, Cat.No: 250-31L), 10 ng IL-3 (Peprotech, Cat.No: #213-13), and 10 ng/ml GM-CSF (Peprotech, Cat.No: 315-03). Cells were treated for 48 or 72 hr with/without 1 μM AGI-5198 (the kind gift of Agios Pharmaceuticals), 0.25 or 1 μM UNC0642 (Sigma; Cat.No: SML1037) or 1 μM UNC1999 (Sigma; Cat.No: SML0778).

Gene set enrichment analysis (GSEA) of mouse microarray data

Raw microarray data from LSK cells of LysM-IDH1-KI (KI) and control littermate (WT) mice (n=5/group) were obtained as previously described (Sasaki et al., 2012). Data were quintile-normalized, and one WT and one KI sample were discarded as outliers based on hierarchical clustering. The remaining WT (n=4) and KI (n=4) samples were tested for enrichment using GSEA version 2.0.12 from the Broad Institute. Curated gene sets were downloaded from MSigDB by searching for the keywords "DNA AND (damage OR repair)" or "p53". Default settings were used except for 10,000 gene set permutations. The top hit was the KEGG p53 signaling pathway.

Assay for Transposase-Accessible Chromatin (ATAC)-qPCR

Chromatin structure was examined using the "assay for transposase-accessible chromatin" coupled to PCR (ATAC-PCR) approach as previously described (Buenrostro et al., 2013) with minor modifications. Briefly, nuclei were prepared from $5x10^4$ LSK cells from pooled Vav-IDH1-WT or Vav-IDH1-KI mice (n=3) and subjected to transposition using the Nextera DNA Sample Preparation kit (Illumina, Cat.No: FC-121-1030). Transposed DNA was purified using a DNA mini-kit (Qiagen) according to the manufacturer's protocol. PCR amplification of transposed DNA was performed using SYBR Green I (Life Technologies) according to the manufacturer's protocol. The reaction was monitored in real-time using a Bio-Rad CFX96 instrument and samples were removed prior to the plateau phase of PCR. Amplified samples were purified using Agencourt AMPure XP beads (Beckman-Coulter) according to the manufacturer's protocol and excess primers and dsDNA <100bp were removed. Quantitative PCR was performed using the following primers for sequences containing the promoter and first exon of genes: *Atm* (Forward, TAGCAGCTGGTTCTGTACGC; Reverse, CGACTTCTTCCTTTGGGGCA), *Atr* (Forward, CTCTCGGTTGCGGTTGGAA; Reverse, CCGTGGAGTAGGTCACGC), *Chek2* (Forward,

ACTGCCGCGATTAGTCTCAA; Reverse, CAAGACGTCGAGTCCCGTTC) and *Trp53* (Forward, AGCCAGGATGGTCCCAATGAA; Reverse, GCGACTATCCAGCTTTGTGCC).

Chromatin immunoprecipitation (ChIP)-qPCR

ChIP was performed using the ChIP-IT Express Enzymatic kit (Active Motif, Cat.No:140603). Briefly, LSK cells (~$2x10^6$ cells) from pooled Vav-IDH1-WT or Vav-IDH1–KI mice (n=3/genotype) were cross-linked by 0.5% paraformaldehyde for 5 min. Enzymatically sheared chromatin was immunoprecipitated by incubation overnight with 2 µg anti-H3K9me3 Ab (Abcam, Cat.No: ab8898) or control IgG. Immunoprecipitated samples were exposed to RNase A and protease K. DNA was recovered using the NucleoSpin Gel and PCR clean-up kit (Macherey Nagel, Cat.No: 740609.250). Purified immunoprecipitated DNA samples and 10% of input DNA samples were subjected to qPCR using primers for the *Atm* promoter (Forward, GGAAAGAGGGACGTGCAGAA; Reverse, AGAACGGCCGTCATCTTCTC).

Primary mouse embryonic fibroblasts (MEF)

Primary MEF were prepared from E13.5 mouse embryos and cultured in DMEM containing 10% FCS plus 55µM 2-mercaptoethanol as previously described (Inoue et al., 2013). For inducible Cre-mediated recombination of the *Idh1*-R132H-*LSL* allele, primary MEF (~$5x10^4$ cells/well in 6-well plates) from *Cre-ERT:Idh1* WT/WT or *Cre-ERT*:*Idh1* LSL/WT mice were exposed to 1µM 4-hydroxytamoxifen (4-OHT; Sigma) for 24-30 hr.

Quantitative real-time RT-PCR

Real-time RT-PCR was performed as previously described (Inoue et al., 2013). Each sample was assayed in triplicate and data were normalized to housekeeping genes (18S rRNA or *Gapdh*). Results were calculated using the comparative threshold cycle method ($2\text{-}\Delta\Delta^{Ct}$). The following primers were used: *Atm* (forward 5'-TTAAGCATTCCTCCCAAGCA-3', reverse 5'- GCAAGCATGGTTCTTTGTGA-3'), *Atr* (forward 5'-AGTTGGCCAGTGCTACTCCAGA-3', reverse 5'-GGTCGGCTGAGCGTCAGTTTTCTT-3'), *53bp1* (forward 5'-TCAGCGTGCTATCTCGACAC-3', reverse 5'-TCCTTGTTCTACAGCAGATTGTTC-3'), *Mrn11a* (forward 5'-CCTCTTATCCGACTACGGGTG-3', reverse 5'-ACTGCTTTACGAGGTCTTCTACT-3'), *Nbn* (forward 5'-GATCAGTCAATCAGTCGAAACCA-3', reverse 5'-CCAAGGGCTCGTATTCTACTCTG-3'), *Rad50* (forward

5’-TGATAAGTTGTCTTGGGGTTTCC-3’, reverse 5’-CTGTGTCTGACGCACCTGT-3’), *Rad51* (forward 5’-CAACGAAGCGCGTTCGAGCC-3’, reverse 5’-GCATAGCCATGACTGTCCCGCG-3’), *Chek1* (forward 5’-GGGGTGGTTTATCTTCATGG-3’, reverse 5’-GCCAAGCCAAAGTCAGAGAT-3’), *Chek2* (forward 5’-CTATGGGCTCTTCAGGATGG-3’, reverse 5’-CCGTCCTTCTCAACAGTGGT-3’), *Trp53* (forward 5’-ACTGCATGGACGATCTGTTG-3’, reverse 5’- GCCATAGTTGCCCTGGTAAG-3’), *Xrcc5* (forward 5’-TGTCCAACGACAGGTATTTTCG-3’, reverse 5’-AAGGGCATTATCAGTGCCATC-3’), *Xrcc6* (forward 5’-ATGTCAGAGTGGGAGTCCTAC-3’, reverse 5’-TCGCTGCTTATGATCTTACTGGT-3’), *Tet2* (forward 5’-AGAGAAGACAATCGAGAAGTCGG-3’, reverse 5’-CCTTCCGTACTCCCAAACTCAT-3’), *18S rRNA* (forward 5’- CGGCGACGACCCATTCGAAC-3’, reverse 5’-GAATCGAACCCTGATTCCCCGTC-3’), and *Gapdh* (forward 5’- GGAAGCTCACTGGCATGGCC-3’, reverse 5’- CCTGCTTCACCACCTTCTTG-3’).

Dot blotting

Dot blotting to detect 5hmC and 5mC were performed as previously described (Chen et al., 2013). Briefly, genomic DNA was prepared from Lin^- cells isolated from Vav-IDH1-WT, Vav-IDH1-KI, and Tet2-KO mice using the QIAamp DNA Mini-kit (Qiagen). Genomic DNA was denatured in 0.1N NaOH at 70°C for 4 min, followed by neutralization with chilled 0.2 M Tris-HCl (pH8). Two-fold serial dilutions of denatured genomic DNA were blotted onto Hybond N^+ membranes (Amersham) in an assembled BIO-DOT apparatus (Bio-Rad). After blotting, membranes were air-dried, cross-linked (1200 $\mu J/cm^2$) using a UV Stratalinker 2400 (Stratagene), blocked with 5% skim milk, and incubated overnight at 4°C with mouse anti-5mC Ab (Millipore, clone 33D3, Cat.No:MABE146) or rabbit anti-5hmC Ab (Active Motif, Cat.No: 39769).

Hydroxymethylated DNA immunoprecipitation (hMeDIP)

hMeDIP was performed as previously described (Taiwo et al., 2012). Briefly, genomic DNA was prepared from LSK cells isolated from Vav-IDH1-WT, Vav-IDH1-KI, and Tet2-KO mice (n=3/genotype) using the QIAamp DNA Mini-kit (Qiagen). Genomic DNA was sheared to 300bp using a Covaris M220 instrument and DNA concentrations were normalized to 100 ng. hMeDIP was conducted using the Diagenode hMeDIP kit using the manufacturer’s protocol as follows. Sheared DNA was combined in a 200 µl PCR tube with buffer containing spiked-in control methylated, unmethylated and hydroxymethylated DNA. This mixture was incubated at 95˚C for 10 min, and then

transferred to an ice water bath for 10 min. For each sample, an aliquot of 10 µl was reserved as the input control and stored at 4˚C. Another 100 µl of the sample was subjected to immunoprecipitation by mixing with anti-5-hMC Ab and incubating in a 4˚C rotator for 2 hr. Washed beads were added and incubation continued at 4˚C for 17 hr. Beads were washed and DNA was extracted using the Diagenode iPure kit following the manufacturer's protocol. Input controls were also subjected to DNA extraction using the Diagenode iPure kit. Samples and input controls were eluted in a final volume of 50 µl and subjected to quantitative PCR using two sets of primers to test enrichment, as previously described (Taiwo et al., 2012). The first set of primers, for the mouse *Sf1* gene, was provided by the hMeDIP kit and served as an internal positive control for enrichment. The second set of primers for each sample contained sequences of the promoter and first exon of one of the following genes: *Atm* (Forward, TAGCAGCTGGTTCTGTACGC; Reverse, CGACTTCTTCCTTTGGGGCA), *Atr* (Forward, CTCTCGGTTGCGGTTGGAA; Reverse, CCGTGGAGTAGGTCACGC), *Chek2* (Forward, ACTGCCGCGATTAGTCTCAA; Reverse, CAAGACGTCGAGTCCCGTTC) and *Trp53* (Forward, AGCCAGGATGGTCCCAATGAA; Reverse, GCGACTATCCAGCTTTGTGCC).

DNA methylation

Genome methylation was determined using reduced representation bisulfite sequencing (RRBS) carried out according to previously published protocols (Gu et al., 2011) enhanced with slight modifications. This enhanced RRBS (eRRBS) protocol was performed on 10 ng DNA purified from LT-HSC, ST-HSC, or MPP isolated from BM of aged male Vav-IDH1-KI mice and control littermates (Vav-IDH1-WT) (6 mice pooled per group) using a FACS Aria instrument (Beckton Dickinson). For ligation, methylated adapters from the TruSeq Kit (Illumina) were diluted 1:4, and 1.5 µl was used in the reaction. After bisulfite conversion and before fragment selection, gels were cut between 150-400 bp for fragment size selection. Final libraries were purified using Agencourt AMpure XP magnetic beads (Beckman Coulter) according to the manufacturer's protocol. Libraries were sequenced at the Princess Margaret Genomics Centre using a HiSeq2000 instrument. Five libraries per lane were multiplexed and 30% PhiX was added to increase library complexity. Reads were aligned to the mm10 reference genome and CpG methylation calls were made using Bismark. 10x coverage filter was applied and the differential DNA methylation and significance levels were analyzed using the methylKit package in R software. DNA methylation was visualized using the UCSC Genome Browser.

Human AML data analysis

Human AML samples in The Cancer Genome Atlas (TCGA) clinical database (Cancer Genome Atlas Research, 2013) were queried for the expression of 12 DNA damage and repair genes using cBio portal (http://www.cbioportal.org/public-portal/). For Figure 4F, *IDH2*- and *TET2*-mutated samples were excluded due to similarities with *IDH1*-mutated AML. For Figure 4G, *IDH1*- and *IDH2*-mutated samples were excluded due to similarities with *TET2*-mutated AML. For the remaining non-mutated *IDH1* (n=127) and *IDH1*-mutated (n=13) AML samples, or non-mutated *TET2* (n=128) and *TET2*-mutated (n=13) AML samples, mRNA expression values determined by RNASeq v2 RSEM (RNA-Seq version 2 by Expectation-Maximization) for each gene were transformed to z-scores for box-plotting. P values were determined using the two-tailed Wilcoxon rank-sum test and adjusted with the Benjamini and Hochberg method.

Comparison of gene expression signatures between TCGA database AML samples and homologous recombination-defective cells

RSEM-normalized mRNA expression levels from RNASeq v2 data of human TCGA AML samples was downloaded from FireBrowse (http://www.firebrowse.org/) and $\log_2$ transformed. Samples with non-normal cytogenetics, or with mutations in *IDH2* or *DNMT3A*, or in *IDH1* and *TET2*, were excluded. Differential gene expression between the *IDH1* (n=5) and *TET2* (n=7) AML samples was calculated using the limma package. The directions of fold change of the genes in *IDH1* vs. *TET2* AML were compared against a gene expression signature of homologous recombination-defective cells (Peng et al., 2014). The association between the directions of fold changes was determined using Fisher's exact test.

Supplemental References

Bendall, S. C., Simonds, E. F., Qiu, P., Amir el, A. D., Krutzik, P. O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A., Ornatsky, O. I., *et al.* (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687-696.

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nature protocols *6*, 468-481.