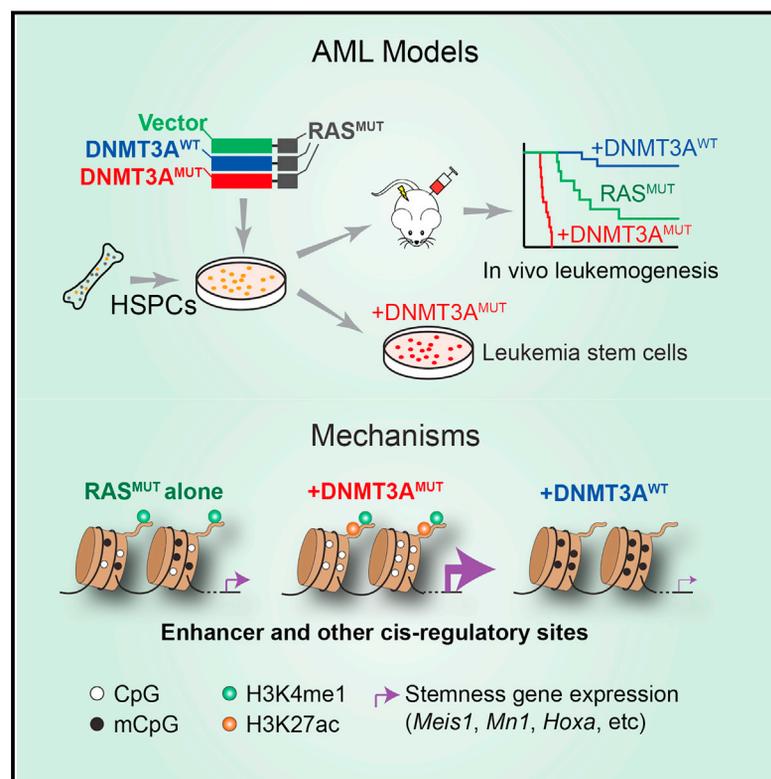


Cancer Cell

Epigenetic Perturbations by Arg882-Mutated DNMT3A Potentiate Aberrant Stem Cell Gene-Expression Program and Acute Leukemia Development

Graphical Abstract



Authors

Rui Lu, Ping Wang, Trevor Parton, ..., Paul A. Wade, Deyou Zheng, Gang Greg Wang

Correspondence

deyou.zheng@einstein.yu.edu (D.Z.), greg_wang@med.unc.edu (G.G.W.)

In Brief

Lu et al. establish that Arg882-mutated DNMT3A contributes to acute myeloid leukemia (AML) pathogenesis through epigenetic activation of leukemia-related genes. Inhibition of Dot1l reverses mutant DNMT3A-induced gene expression, indicating a potential therapeutic strategy for AMLs harboring this mutation.

Highlights

- DNMT3A^{R882H} promotes acute leukemogenicity in the presence of mutant NRAS
- DNMT3A^{R882H} induces focal DNA hypomethylation at cis elements of key stemness genes
- DNMT3A^{R882H} potentiates stemness gene expression via enhancer/promoter activation
- DNMT3A^{R882H}-induced gene activation programs are sensitive to Dot1l blockade

Accession Numbers

GSE71475

Epigenetic Perturbations by Arg882-Mutated DNMT3A Potentiate Aberrant Stem Cell Gene-Expression Program and Acute Leukemia Development

Rui Lu,^{1,2} Ping Wang,³ Trevor Parton,¹ Yang Zhou,⁴ Kaliopi Chrysovergis,⁵ Shira Rockowitz,⁶ Wei-Yi Chen,⁷ Omar Abdel-Wahab,⁸ Paul A. Wade,⁵ Deyou Zheng,^{3,6,*} and Gang Greg Wang^{1,2,9,*}

¹Lineberger Comprehensive Cancer Center

²Department of Biochemistry and Biophysics

University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA

³Department of Neurology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

⁴Department of Pathology and Laboratory Medicine, McAllister Heart Institute, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA

⁵Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, National Institute of Health, Research Triangle Park, NC 27709, USA

⁶Departments of Genetics and Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

⁷Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei 11221, Taiwan

⁸Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

⁹Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Correspondence: deyou.zheng@einstein.yu.edu (D.Z.), greg_wang@med.unc.edu (G.G.W.)

<http://dx.doi.org/10.1016/j.ccell.2016.05.008>

SUMMARY

DNA methyltransferase 3A (*DNMT3A*) is frequently mutated in hematological cancers; however, the underlying oncogenic mechanism remains elusive. Here, we report that the *DNMT3A* mutational hotspot at Arg882 (*DNMT3A*^{R882H}) cooperates with *NRAS* mutation to transform hematopoietic stem/progenitor cells and induce acute leukemia development. Mechanistically, *DNMT3A*^{R882H} directly binds to and potentiates transactivation of stemness genes critical for leukemogenicity including *Meis1*, *Mn1*, and *Hoxa* gene cluster. *DNMT3A*^{R882H} induces focal epigenetic alterations, including CpG hypomethylation and concurrent gain of active histone modifications, at *cis*-regulatory elements such as enhancers to facilitate gene transcription. CRISPR/Cas9-mediated ablation of a putative *Meis1* enhancer carrying *DNMT3A*^{R882H}-induced DNA hypomethylation impairs *Meis1* expression. Importantly, *DNMT3A*^{R882H}-induced gene-expression programs can be repressed through Dot1l inhibition, providing an attractive therapeutic strategy for *DNMT3A*-mutated leukemias.

INTRODUCTION

DNA methylation provides a critical epigenetic means for defining cellular identity and regulating functional output of gene-regulatory elements such as promoters and enhancers (Jones, 2012; Schubeler, 2015). Recently, DNA methyltransfer-

ase 3A (*DNMT3A*), a de novo DNA methyltransferase gene, was found mutated in ~20%–30% of human acute myeloid leukemias (AMLs) and ~10%–20% of various other hematological cancers (Cancer Genome Atlas Research Network, 2013; Ley et al., 2010; Patel et al., 2012; Yan et al., 2011; Yang et al., 2015). *DNMT3A* mutations also associate well with clonally

Significance

Recurrent *DNMT3A* mutations at Arg882 are found in hematological malignancies and disorders; however, due to a lack of relevant disease models, molecular mechanisms by which *DNMT3A* mutations influence leukemogenesis remain largely undefined. Through establishment and characterization of murine leukemia and leukemia stem cell models, we show that *DNMT3A*^{R882H} mutation potentiates transactivation of stemness genes required for acute leukemogenicity. Integrated epigenomic profiling of murine models further reveals the underlying epigenetic alterations induced by *DNMT3A*^{R882H}, which are enriched at gene-regulatory sites and resemble those seen in human patients. Pharmacological inhibition of Dot1l suppresses *DNMT3A*^{R882H}-associated gene activation and acute leukemogenesis. Our findings not only promote mechanistic understandings of *DNMT3A* mutation-associated clonal and malignant hematopoiesis but also provide a therapeutic avenue for *DNMT3A*-mutated leukemias.

derived hematopoiesis at premalignant stages (Genovese et al., 2014; Jaiswal et al., 2014; Shlush et al., 2014; Xie et al., 2014) and often coexist with a secondary lesion that “hits” either the FLT3-RAS kinase pathway, an epigenetic regulator (*IDH1/2*, *TET2*), or *NPM1* in AML patients (Cancer Genome Atlas Research Network, 2013; Ley et al., 2010; Patel et al., 2012; Yang et al., 2015). These clinical findings suggest that *DNMT3A* mutation acts as a founder lesion and requires an additional genetic event to induce malignant development. Consistently, mice with *Dnmt3a* knockout in the bone marrow produced phenotypically normal hematopoietic stem cells (HSCs); only after rounds of transplantation did *Dnmt3a*-null HSCs display self-renewal advantages (Challen et al., 2012). Mice with *Dnmt3a* mutation alone did not develop frank AML but showed increased susceptibility to malignant development upon acquisition of additional mutations (Celik et al., 2015; Chang et al., 2015; Mayle et al., 2015; Xu et al., 2014).

Mutational hotspot at Arg882 (R882), a residue located within the homodimerization interface of DNMT3A, accounts for the majority (~60%) of *DNMT3A* mutations found in AMLs (Ley et al., 2010; Yang et al., 2015). Due to a primarily heterozygous nature of *DNMT3A* R882 mutation, it was thought to act in a dominant-negative and/or haploinsufficient manner (Holz-Schietinger et al., 2012; Kim et al., 2013; Russler-Germain et al., 2014). Clinical evidence supports this notion, as AML patients with *DNMT3A* R882 mutation exhibited focal DNA hypomethylation (Russler-Germain et al., 2014). Despite these advances, there is a lack of relevant AML animal models for studying *DNMT3A* R882 mutation. Molecular pathways and mechanisms by which *DNMT3A* mutation contributes to AML pathogenesis remain undefined. Targeted approaches for the treatment of *DNMT3A*-mutated AMLs remain to be developed.

RESULTS

***DNMT3A* Hotspot Mutation Enhances Sensitivity of Hematopoietic Stem/Progenitor Cells to Transformation In Vitro**

Previous reports indicate that hotspot mutations of *DNMT3A* such as *DNMT3A*^{R882H} act in a dominant-negative manner by disrupting formation of a DNMT3A-associated tetramer complex required for efficient DNA methylation (Holz-Schietinger et al., 2012; Kim et al., 2013; Russler-Germain et al., 2014). These studies prompted us to ask whether ectopic expression of human *DNMT3A*^{R882H} in murine hematopoietic stem/progenitor cells (HSPCs) could establish a transformation phenotype in a colony-forming unit (CFU) and replating assay (Figure S1A). Initially, we found a lack of CFU-promoting effect by *DNMT3A*^{R882H} alone (Figures 1A–1C). We then asked whether *DNMT3A*^{R882H} could enhance sensitivity of HSPCs to transformation in the presence of a second oncogenic lesion. Toward this end, we used a bicistronic retroviral system to coexpress either wild-type (WT) or R882H-mutant (RH) *DNMT3A*, together with other mutations known to coexist with *DNMT3A* mutation in human AMLs: *NRAS* (*NRAS*^{G12D}), *NPM1* (*NPM1c*), or *IDH1* (*IDH1*^{R132H}) (Figure S1A) (Ley et al., 2010; Patel et al., 2012; Shih et al., 2012). Following viral transduction and drug selection, we obtained highly pure HSPCs with comparable levels of oncogene expression for CFU assays (Figures 1C, S1B, and S1C). We

did not observe a CFU-promoting effect of *DNMT3A*^{R882H} in the presence of *NPM1c* or *IDH1*^{R132H} (Figure S1D). However, a significant increase in CFUs was seen after replating of HSPCs coexpressing *DNMT3A*^{R882H} and *NRAS*^{G12D} (hereafter referred to as “RH-RAS”), relative to those with either oncogene alone (Figures 1A and 1B). In contrast to *DNMT3A*^{R882H}, *DNMT3A*^{WT} did not promote colony formation (Figures 1A and 1B). Post replating, HSPCs expressing *NRAS*^{G12D} alone produced tiny and diffuse colonies of differentiated cells whereas those with RH-RAS gave rise to large, compact colonies that mainly comprised undifferentiated progenitors (Figures 1A [inset], 1B, S1E, and S1F). Importantly, cells expressing RH-RAS as derived from serially replated colonies were able to propagate and maintain their immature progenitor status in long-term liquid culture (Figures 1D and S1G), suggesting acquisition of indefinite self-renewal capability by these cells. These data have shown that, in contrast to *DNMT3A*^{WT}, R882-mutated *DNMT3A* promotes aberrant self-renewal of HSPCs and enhances their sensitivity to transformation in vitro. In addition, *NRAS*^{G12D} genetic background provides a useful platform for dissecting the role of *DNMT3A* mutation in AML development.

***DNMT3A*^{R882H} Acts in Concert with Activated RAS to Induce Murine AMLs In Vivo**

The observed in vitro effect of *DNMT3A*^{R882H} on aberrant HSPC self-renewal and immortalization indicates that it could cooperate with *NRAS*^{G12D} to cause malignant transformation in vivo. Thus, we transplanted murine HSPCs freshly transduced with *DNMT3A* (either WT or RH) and/or *NRAS*^{G12D} to syngeneic mice. *NRAS*^{G12D} alone induced a myeloproliferative disease with incomplete penetrance (Figures 1E and S1H). *DNMT3A*^{R882H} alone did not cause detectable diseases over a 12-month monitoring period; however, in the presence of *NRAS*^{G12D}, it significantly accelerated development of leukemia with a shorter latency phenotype and full penetrance (Figure 1E). RH-RAS-induced leukemia was also characterized by hepatosplenomegaly (Figures 1F, 1G, and S1I), leukemic infiltration to bone marrow, spleen, and liver (Figures 1H and S1J), and elevated counts of peripheral white blood cells and blasts (Figures 1I and S1K–S1O; Table S1). Leukemia induced by RH-RAS expressed virally transduced genes at a level comparable with progenitors immortalized by RH-RAS in vitro (Figure 1J) and displayed an immature myeloid (AML) immunophenotype (*Mac-1*⁺/*c-Kit*^{low}/*Cd34*^{low}/*Gr1*⁻/*Cd3e*⁻/*Cd19*⁻/*Ter119*⁻; Figures 1K, S1P, and S1Q; Table S1). Whole-exome sequencing of three independent murine AMLs identified no recurrent mutation of additional genes (Figure S1R), suggesting that *DNMT3A*^{R882H} and *NRAS*^{G12D} are sufficient to drive AML development. Interestingly, unlike *DNMT3A*^{R882H}, *DNMT3A*^{WT} suppressed leukemogenesis in vivo (Figure 1E), suggesting that normal *DNMT3A* activities oppose AML pathogenesis.

***DNMT3A* Hotspot Mutation Produces Leukemia-Initiating Stem Cells Ex Vivo in the Presence of *NRAS*^{G12D}**

To further verify the cell transformation effect of RH-RAS, we used a previously described liquid cultivation system (Wang et al., 2007, 2009) and were able to recapitulate HSPC immortalization with RH-RAS only, and not either oncogene alone or

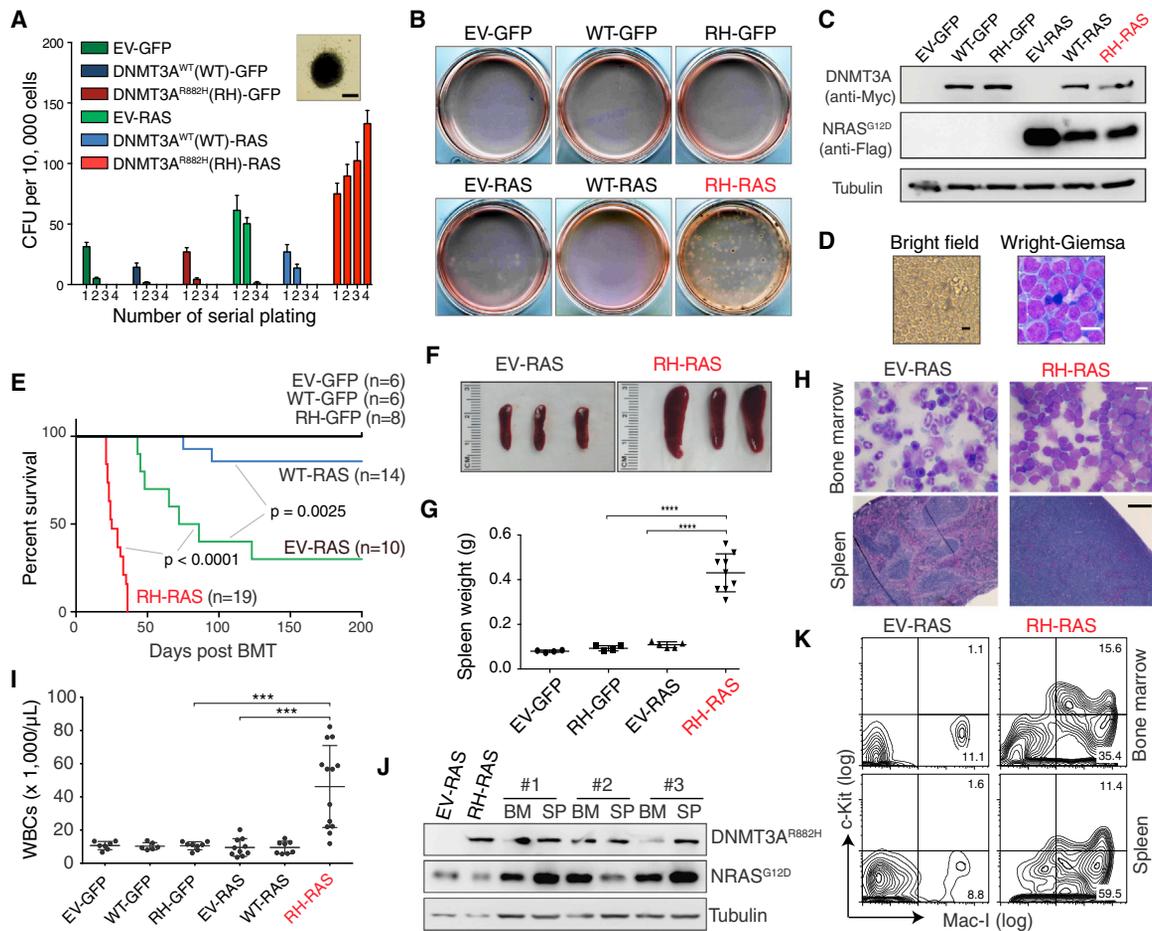


Figure 1. DNMT3A^{R882H} Acts in Concert with Mutant RAS to Transform Murine HSPCs Ex Vivo and Induce AMLs In Vivo

(A) Colony-forming unit (CFU) assay using murine HSPCs expressing empty control (EV), wild-type (WT), or R882H mutant (RH) DNMT3A in combination with GFP or NRAS^{G12D} (RAS). Inset shows a typical colony expressing RH-RAS at the fourth replating. Scale bar, 1 mm.

(B) Images of CFU assay at the fourth replating.

(C) Immunoblot of DNMT3A (Myc-tagged) and NRAS (Flag-tagged) in HSPCs after infection.

(D) Microscopic image and Wright-Giemsa staining of RH-RAS-coexpressing cells derived from the fourth replating after long-term culture with the SCF cytokine in vitro. Scale bars, 10 μm.

(E) Kaplan-Meier survival curve of mice after bone marrow transplantation (BMT) of HSPCs freshly transduced with indicated genes. The p values were calculated by log-rank test.

(F and G) Spleen size (F, n = 3) and weight (G, n = 4–9) of indicated cohorts 3–4 weeks post BMT. The p values were calculated by Student’s t test.

(H) Wright-Giemsa staining of bone marrow (upper) and H&E staining of spleen (bottom) of indicated cohorts 4 weeks post BMT. Scale bar, 10 μm (upper) and 200 μm (bottom).

(I) White blood cell (WBC) counts in peripheral blood of indicated cohorts (n = 6–13) 4 weeks post BMT. The p values were calculated by Student’s t test.

(J) Immunoblot of DNMT3A (Myc) and NRAS (Flag) proteins in bone marrow (BM) and spleen (SP) cells from mice with leukemia induced by RH-RAS coexpression. The first two lanes were loaded with samples of in vitro infected HSPCs.

(K) Fluorescence-activated cell sorting (FACS) analysis of Mac-1 and c-Kit with bone marrow and spleen cells of indicated cohorts 4 weeks post BMT. Error bar denotes ±SD; ***p < 0.001, ****p < 0.0001. See also Figure S1 and Table S1.

coexpression of DNMT3A^{WT} with NRAS^{G12D} (Figure S2A). Similarly to those derived from CFU assays, RH-RAS-immortalized progenitors stably maintained their progenitor identity in vitro in the presence of SCF or Flt3 ligand, and presented with expression of immature myeloid (c-Kit⁺/Mac-1^{low}/Gr1⁻) and stem cell antigens (Cd34^{low}/Flt3⁺/Sca1^{low/-}) as well as a lack of other lineage markers (Figures 2A, 2B, S2B, and S2C). Exposure of these progenitors to myeloid-promoting cytokines decreased cell proliferation (Figure S1G) and induced terminal myeloid differentiation (c-Kit⁻/Mac-1^{high}/F4-80^{high}; Figures 2B and S2D),

demonstrating their myeloid differentiation capability. Engraftment with each of three independent RH-RAS-immortalized progenitor lines induced murine AMLs (Figures S2E–S2H) that can be propagated in vivo with sequential transplantation (Figure 2C). Importantly, as few as 50–500 of these cells were sufficient to cause AML (Figure 2D), illustrating their leukemia-initiating stem cell (LSC) characteristic (hereafter called “LSCs^{RH-RAS}”). To further characterize LSCs^{RH-RAS}, we profiled their transcriptome and genome-wide occupancy of H3K4me1, a histone mark demarcating lineage-specific enhancers (Lara-Astiaso

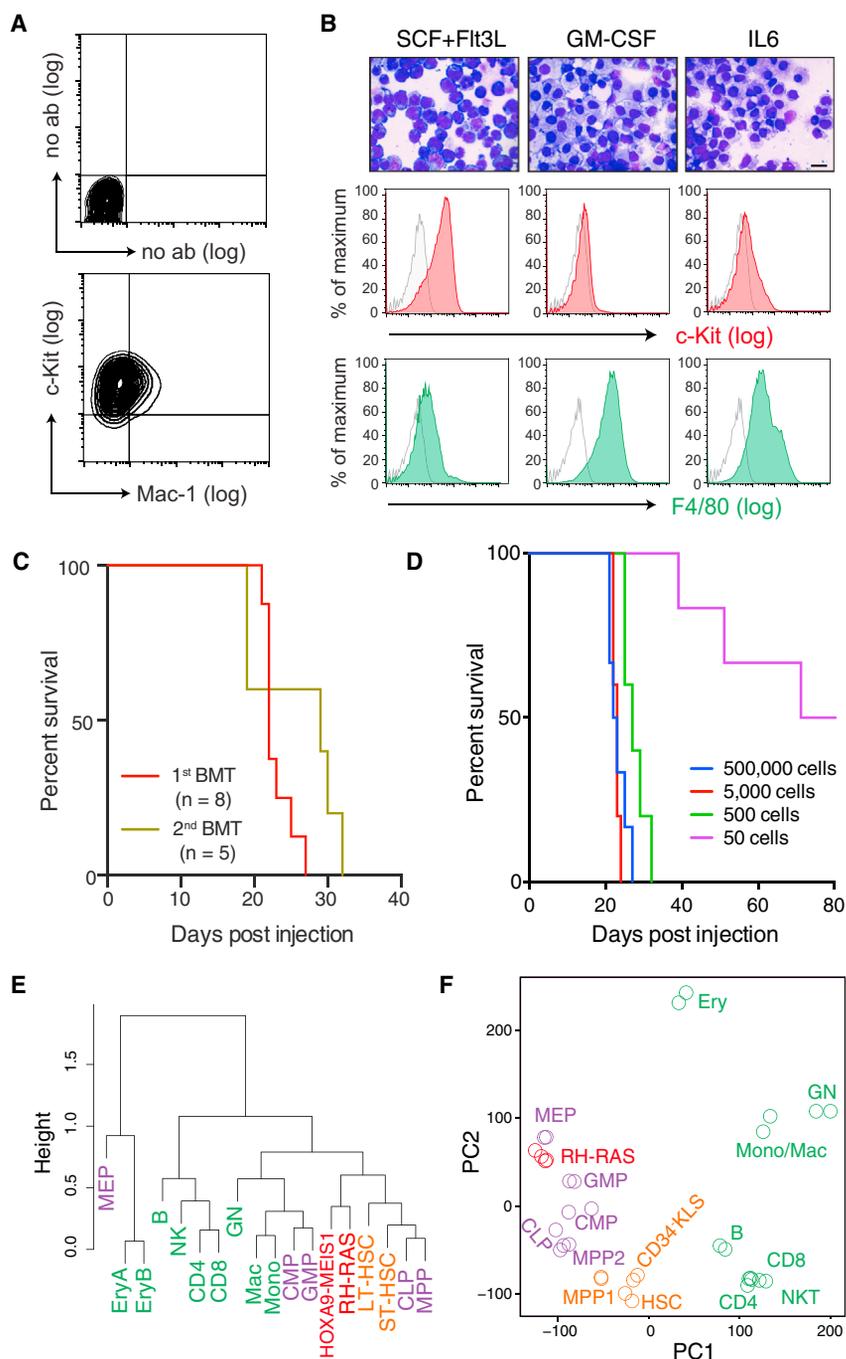


Figure 2. R882-Mutated DNMT3A Establishes Leukemia-Initiating Stem Cells Ex Vivo in the Presence of Activated RAS

(A) FACS analysis of in vitro immortalized progenitors by RH-RAS using a liquid culture system.

(B) Wright-Giemsa staining (upper) and FACS analysis of RH-RAS-immortalized progenitors 14 days after cultivation with indicated cytokines. FACS control, non-specific immunoglobulin G (gray trace). Scale bar, 10 μ m.

(C) Kaplan-Meier curve of mice receiving primary or secondary BMT with RH-RAS-induced leukemia.

(D) Kaplan-Meier curve of mice (n = 5–6) receiving BMT of the indicated numbers of RH-RAS immortalized cells.

(E) Hierarchical clustering of genome-wide H3K4me1 profiles of LSCs^{RH-RAS}, AML-causing leukemia-initiating stem cell (LSC) lines produced by overexpressed HOXA9 plus MEIS1 (HOXA9-MEIS1), and various normal blood cell types. LT-HSC, long-term HSC; ST-HSC, short-term HSC; MPP, multipotent progenitor; CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-monocyte progenitor; MEP, megakaryocyte-erythroid progenitor; Mac, macrophage; Mono, monocyte; GN, granulocyte; B, B220⁺/CD19⁺ B cell; CD4/8, CD4/8⁺ T cell; NK, natural killer cell; EryA and EryB, Ter119⁺/CD71⁺ erythroid cell with high and low forward scatter, respectively.

(F) Principal component (PC) analysis of transcriptome profiles of LSCs^{RH-RAS} and various normal blood cell types. CD34⁺/KLS, Cd34⁺/c-Kit⁺/Lin⁻/Scal⁺ HSC; MPP1, Flk2⁻ multipotent progenitor; MPP2, Flk2⁺ multipotent progenitor; NKT, natural killer T cell; Ery, erythroid cell. Other abbreviations as in (E).

See also Figure S2.

et al., 2014). Unsupervised clustering of H3K4me1 profiles of LSCs^{RH-RAS} and various hematopoietic cell lineages revealed a similarity of LSCs^{RH-RAS} to HSPCs such as HSC and myeloid progenitors, when compared with differentiated cell types (Figures 2E and S2I); similar results were found in their transcriptome comparison (Figures 2F and S2J). Notably, a closer similarity was seen when comparing LSCs^{RH-RAS} with leukemic progenitors we and others previously produced using either HOXA9 plus MEIS1 (Wang et al., 2005), MLL translocation (Bernt et al., 2011), NUP98-NSD1 (Wang et al., 2007), or NUP98-JARID1A (Wang et al., 2009) (Figures 2E, S2I, and S2J), implying a com-

monality of pathways underlying leukemogenicity by these oncogenes. **R882-Mutated DNMT3A Potentiates Abnormal Transcription of Stem Cell Genes Including a Meis1-Mn1-Hoxa Regulatory Node** Next, we sought to understand the molecular basis underlying indefinite self-renewal shown by LSCs^{RH-RAS}. First, we asked whether they carry self-renewal or stemness gene-expression programs, a known feature of LSCs (Abramovich et al., 2005; Eppert et al., 2011; Krivtsov et al., 2006). By transcriptome analysis, we identified 54 genes uniquely expressed in LSCs^{RH-RAS} and primitive HSPCs with self-renewal capabilities, relative to differentiating and mature hematopoietic cell types (Figure 3A and Table S2). The stem cell signature genes expressed in LSCs^{RH-RAS} are only part of HSC stemness gene programs (~10%, Figure S3A); we further verified enrichment of the LSC^{RH-RAS} stemness signature in self-renewing HSCs with independent datasets (Figures S3B and S3C). The top LSC^{RH-RAS} stemness genes included *Hoxa9*, *Mn1*, *Hoxa5*, and *Meis1* (Figure 3A), which encode a set of transcription factors (TFs) and

cofactors crucial for sustaining self-renewal of normal HSCs and leukemic LSCs (Heuser et al., 2011; Huang et al., 2012; Wang et al., 2006; Wong et al., 2007). Gene targets of Meis1-Mn1-Hoxa, *Ft3* and *Sox4* (Heuser et al., 2011; Huang et al., 2012; Wang et al., 2005), were also among top stemness genes identified (Figure 3A), indicating activity of this TF regulatory circuitry in LSCs^{RH-RAS}. Moreover, activation of Meis1 and Hoxa in LSCs^{RH-RAS} was found to be comparable with that in LSCs defined by other deregulated chromatin factors such as MLL-AF9, NUP98-JARID1A, or NUP98-NSD1, while *Mn1* and *Mycn* showed unique expression in LSCs^{RH-RAS} (Figures S3D and S3E).

As LSCs^{RH-RAS} carry both *DNMT3A* and *NRAS* mutations, we next asked which stemness gene signatures are dependent on *DNMT3A*^{R882H}. We performed microarray studies using HSPCs after transduction of *NRAS*^{G12D} alone or with coexpressed *DNMT3A*, either WT or R882H mutant (hereafter referred to as EV-RAS, WT-RAS, or RH-RAS). These HSPCs were collected 12 and 16 days after viral transduction when their proliferation rates were comparable (Figure S2A). Among the 54 LSC^{RH-RAS} stemness genes, nine were found to be upregulated by *DNMT3A*^{R882H} at both time points, including *Meis1*, *Mn1*, and *Hoxa* (Figures 3B and S3F). Consistently, gene set enrichment analysis (GSEA) found that gene sets associated with AML development, undifferentiated myeloid cells, and NUP98-HOXA9 targets were significantly enriched in HSPCs with RH-RAS (Figure 3C). Conversely, gene sets associated with myeloid differentiation showed reduced expression in HSPCs expressing RH-RAS, relative to EV-RAS (Figure 3D), whereas the same gene sets showed enhanced expression in HSPCs expressing WT-RAS (Figure 3D), thus suggesting opposite effects of WT and R882-mutated *DNMT3A* on regulating genes crucial for HSPC self-renewal versus differentiation. We verified unique upregulation of *Meis1*, *Hoxa*, and *Mn1* in RH-RAS HSPCs (Figures 3E and S3G) and their induced AMLs (Figures 3F, 3G, and S3H). To functionally assess whether the activated Meis1-Mn1-Hoxa circuitry is essential for RH-RAS-induced AML development, we introduced independent small hairpin RNAs (shRNAs) of *Meis1* or *Mn1* into LSCs^{RH-RAS} (Figure 3H) and found that knocking down either gene significantly impaired *in vitro* growth of LSCs^{RH-RAS} (Figures 3I and S3I) as well as their *in vivo* leukemogenic function (Figures 3J, S3J, and S3K).

Together, these data reveal a role of R882-mutated *DNMT3A* in potentiating abnormal activation of stemness genes such as *Meis1*, *Mn1*, and *Hoxa*, which are required for mutant *DNMT3A*-mediated AML progression.

ChIP-Seq Reveals Context-Dependent Targeting of R882-Mutated DNMT3A into the LSC Genome

The LSC^{RH-RAS} cellular model described above provides an ideal system for dissecting the molecular mechanism underlying *DNMT3A*^{R882H}-mediated oncogenesis. Mutant *DNMT3A* proteins are exclusively nuclear (Figure S4A); thus, we first mapped their genome-wide occupancy in LSCs^{RH-RAS} by chromatin immunoprecipitation sequencing (ChIP-seq) using antibodies against the Myc tag fused to *DNMT3A*^{R882H} (Figure S1A). Myc-*DNMT3A*^{R882H} ChIP-seq gave robust and specific signals (Figure 4A); as a negative control, Myc ChIP-seq using cells without

Myc-*DNMT3A*^{R882H} expression did not detect any peaks (Figure S4B). We identified 13,705 genomic regions with significant *DNMT3A*^{R882H} binding (i.e., *DNMT3A*^{R882H} peaks, Table S3) in LSCs^{RH-RAS}, which spread over promoter and inter- or intragenic regions (Figure S4C). *DNMT3A*^{R882H} exhibited a broad binding pattern with an average peak size of ~17 kb (Figures S4D and 4B, with an example peak at *Lig1*). Interestingly, *DNMT3A*^{R882H} binding was stronger at intermediately transcribed genes, relative to lowly or highly expressed genes (Figure 4A), and positively correlated to CpG dinucleotide density except at CpG islands (CGIs) where *DNMT3A*^{R882H} has a sharp drop in overall binding (Figure 4C). *DNMT3A*^{R882H} binding regions also showed depletion of H3K4me3 (Figure 4D), a histone modification known to suppress *DNMT3A* binding due to an intrinsic histone H3 “reader” activity of *DNMT3A*’s ADD domain (Guo et al., 2015; Noh et al., 2015). Intriguingly, 76.1% of *DNMT3A*^{R882H} peaks were found in close proximity to and significantly overlapped with peaks of H3K4me1, a histone mark demarcating enhancer elements (observed/expected = 10.2, $p < 10^{-300}$; Figure 4E), as exemplified by those identified in an intron region of *Lig1* and an intergenic region of *Vegfa* (Figures 4B [inset] and S4E [boxed areas]). Ontology analysis of *DNMT3A*^{R882H} peaks revealed their significant enrichment at genes related to normal and malignant hematopoiesis, PML-RAR α targets, and *MLL* rearrangement-associated genes (Figures 4F and S4F). Notably, key AML-promoting or stemness genes upregulated by *DNMT3A*^{R882H} such as *Meis1*, *Mn1*, *Hoxa*, and *Mycn* were all found directly bound by *DNMT3A*^{R882H} (Figures 4G, 4H, S4G, and S4H). Collectively, our genome-wide profiling of *DNMT3A*^{R882H} has revealed a CpG content and “histone code”-dependent targeting of R882-mutated *DNMT3A* into cancer cell genomes; we have also identified a previously unappreciated overlap of *DNMT3A*^{R882H} with putative enhancer and *cis*-regulatory sites (marked by H3K4me1) at numerous developmental genes including a Meis1-Mn1-Hoxa node.

R882-Mutated DNMT3A Induces Focal Hypomethylation of CpG Sites Enriched with Gene-Regulatory Elements

We next aimed to delineate *DNMT3A*^{R882H}-induced epigenetic perturbations during AML progression. By enhanced reduced representation bisulfite sequencing (eRRBS), we mapped global DNA methylation patterns of murine HSPCs 16 days after transduction of EV-RAS, RH-RAS, or WT-RAS. Analysis of eRRBS data, which had 11–12 \times coverage for 6.5 million CpGs in all samples, revealed no significant changes in global CpG methylation (Figures S5A–S5C) except a moderate change at CpG shores (Figure S5C). By pairwise comparison of CpG methylation, we identified 12,889 differentially methylated CpG sites (DMCs) in HSPCs expressing RH-RAS relative to EV-RAS, with most DMCs (80.8%) hypomethylated (Figure 5A, left; hereafter termed “*DNMT3A*^{R882H}-associated hypo-DMCs”); in contrast, DMCs associated with *DNMT3A*^{WT} are largely hypermethylated (hyper-DMCs, 80.6%; Figure 5A, right). *DNMT3A*^{R882H}-associated hypo-DMCs were found mainly in intron, intragenic, and promoter regions, while *DNMT3A*^{WT}-induced hyper-DMCs were enriched in promoters and CGIs (Figures S5D and S5E). Importantly, *DNMT3A*^{R882H}-associated hypo-DMCs were significantly enriched at genomic regions with H3K4me1 (Figure 5B) or with *DNMT3A*^{R882H} binding (Figures

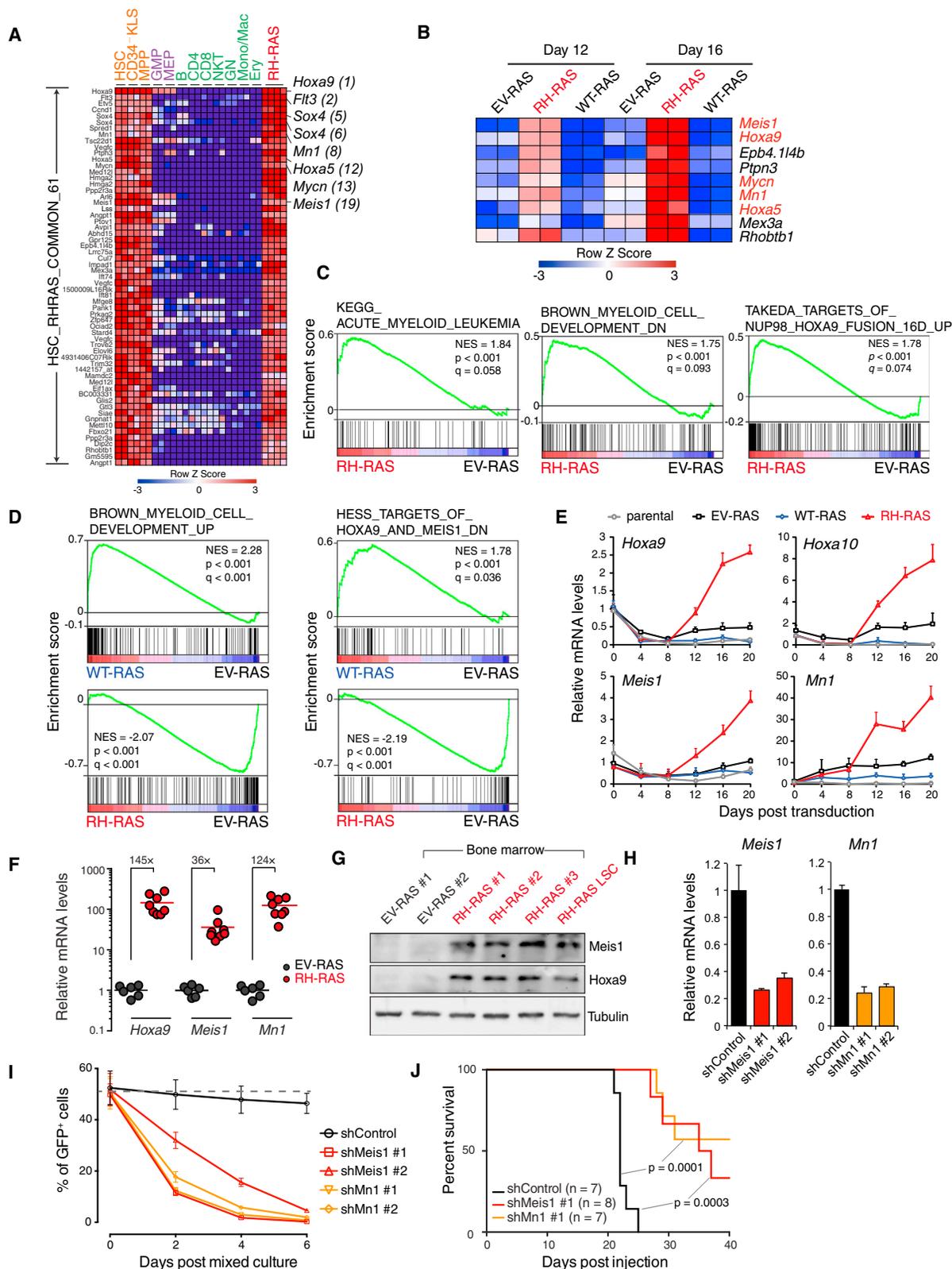


Figure 3. DNMT3A^{R882H} Potentiates Aberrant Activation of Stemness Genes Including a Critical *Meis1-Mn1-Hoxa* Regulatory Node

(A) Heatmap of 61 probes (54 genes) showing unique expression in both self-renewing HSPCs (HSC, Cd34⁺KLS, and MPP) and LSCs^{RH-RAS} but not in differentiating (purple) or mature (green) blood cell types. Probes are ranked by higher expression in LSCs^{RH-RAS} relative to differentiating and mature cells. Example genes are highlighted along with their respective rankings.

(legend continued on next page)

5B and 5C). DNMT3A^{R882H}-associated hypo-DMCs were also found to be enriched with the binding site of the ETS family of TFs (Erg and Spi1/PU.1) and other hematopoietic TFs (Runx1 and Mycn; Figure 5D). In contrast, DNMT3A^{R882H}-associated hyper-DMCs exhibited none of these features and, instead, correlated negatively to DNMT3A^{R882H} binding (Figures 5B–5D), suggesting that creation of hyper-DMCs is due to an indirect effect of DNMT3A^{R882H}.

Consistent with DMCs, differentially methylated regions (DMRs) identified in HSPCs co-transduced with DNMT3A^{R882H} relative to control were mainly hypomethylated (hereafter called “DNMT3A^{R882H}-associated hypo-DMRs”; Table S4, n = 1,199) while DNMT3A^{WT}-associated DMRs were mainly hypermethylated (hyper-DMRs) (Figures 5E and S5F). These two sets of DMRs showed a significant overlap, including those found at DNMT3A^{R882H}-deregulated stemness genes (*Meis1*, *Mn1*, *Hoxa7*, and *Mycn*), further highlighting that WT and R882-mutated DNMT3A have opposing effects on DNA methylation of crucial AML-promoting genes (Figure 5F). In addition, DNMT3A^{R882H}-associated hypo-DMRs were enriched at genes related to transcriptional regulation, hematopoietic development, and cancer (Figures 5G and S5G). Consistent with results in DMCs, H3K4me1 and DNMT3A^{R882H} binding was significantly enriched at DNMT3A^{R882H}-associated hypo-DMRs (Figures 5H and S5H). Taken together, our results show that R882-mutated DNMT3A is sufficient to induce CpG hypomethylation at putative *cis*-regulatory sites of key stemness genes that we have functionally validated as essential for AML progression in murine models.

DNMT3A^{R882H}-Induced DNA Hypomethylation Identified in Murine Models Mirrors What Was Seen in Human AMLs with DNMT3A R882 Mutation

A focal CpG hypomethylation phenotype seen in the above murine model is reminiscent of what was observed in human AMLs with DNMT3A mutation (Russler-Germain et al., 2014). To assess whether our murine model mimics human disease, we first identified regions in the human genome that are homologous (i.e., conserved) to DNMT3A^{R882H}-associated hypo-DMRs defined in the murine model. We then found that, relative to randomized control, CpGs located in such conserved human genomic sites showed a significant reduction in their methylation levels among human AML samples with DNMT3A R882 mutation, relative to those with normal DNMT3A (Figure 5I; $p < 2.2 \times 10^{-16}$). Despite a relatively limited coverage of CpGs by the 450K-array platform used in the human AML study (Russ-

ler-Germain et al., 2014), genes with hypo-DMRs identified in AML patients carrying DNMT3A R882 mutation also had a significant overlap with those that gain DNMT3A^{R882H}-associated hypo-DMRs in our murine model (Table S5; $p < 0.05$). We identified 119 genes showing CpG hypomethylation in both human AMLs and murine LSC models, which again include stemness and AML-promoting genes *MEIS1*, *HOXA7*, and *MN1* (Figures 5J and S5I). We subsequently verified differential CpG methylation of DMRs at these genes in murine cells by direct bisulfite sequencing (Figures 5K and S5J), and further showed that a consistent hypomethylation pattern exists at conserved DMRs in human AMLs with DNMT3A R882 mutation, relative to those with non-R882 mutated or normal DNMT3A (Figures 5L and S5K).

Hypo-DMRs Induced by DNMT3A^{R882H} Facilitate Gain of Histone Acetylation at Gene-Regulatory Sites

Because DNMT3A^{R882H} binding and induced hypo-DMRs showed significant overlap with H3K4me1, a histone mark demarcating gene-regulatory regions such as enhancers and proximal elements close to promoters (Rada-Iglesias et al., 2011), we performed ChIP-seq profiling of H3K27ac, a histone modification correlating to enhancer/promoter activity, with the samples we used for eRRBS. Intriguingly, we found that introducing DNMT3A^{R882H} to HSPCs caused an overall gain of H3K27ac at DNMT3A^{R882H}-associated hypo-DMRs (Figure 6A, left) whereas no overall change in H3K4me1 was seen for these hypo-DMRs (Figure 6B, left); in contrast, expression of DNMT3A^{WT} decreased overall H3K27ac and H3K4me1 at these hypo-DMRs (Figures 6A and 6B, left). As a control, DNMT3A^{R882H}-associated hyper-DMRs did not show such changes (Figures 6A and 6B, right). Consistently, similar histone modification changes were seen at regions in close proximity to DNMT3A-associated DMCs (Figures S6A and S6B). Importantly, DMRs at key stemness or AML genes such as *Meis1*, *Mn1*, *Hoxa*, and *Mycn* all exhibited significant gain of H3K27ac in DNMT3A^{R882H}-expressing HSPCs as well as loss of H3K27ac in DNMT3A^{WT}-expressing HSPCs at their putative *cis*-regulatory sites (Figures 6C and S6C–S6H). By ChIP-qPCR, we verified the observed changes of H3K27ac and H3K4me1 at a panel of DMRs after transduction of DNMT3A^{R882H} versus DNMT3A^{WT} into HSPCs (Figures 6D and S6I). Furthermore, expression of DNMT3A^{R882H} enhanced binding of the H3K27 acetyltransferase p300 to hypo-DMRs at stemness genes (Figure 6E), suggesting that CpG hypomethylation facilitates recruitment of H3K27ac

- (B) Of the 54 self-renewal genes, genes showing consistently higher expression in HSPCs 12 and 16 days after transduction of RH-RAS relative to EV-RAS.
 (C) GSEA shows enrichment of AML-associated genes (left), genes downregulated upon myeloid differentiation (middle), and NUP98-HOXA9 targets (right) in HSPCs with RH-RAS versus EV-RAS.
 (D) GSEA shows enrichment of differentiation gene sets in WT-RAS or RH-RAS HSPCs relative to EV-RAS. Left, myeloid differentiation genes; right, genes downregulated upon activation of HOXA9 and MEIS1.
 (E) qRT-PCR of indicated genes in murine HSPCs after transduction of EV-RAS, WT-RAS, or RH-RAS.
 (F) qRT-PCR of indicated genes in mouse bone marrow 21 days post BMT of HSPCs with EV-RAS (n = 6) or RH-RAS (n = 8).
 (G) Immunoblot of *Meis1* and *Hoxa9* in bone marrow of mice 21 days post BMT of HSPCs with EV-RAS or RH-RAS. The last lane was loaded with LSC^{RH-RAS} samples.
 (H) qRT-PCR showing shRNA-mediated *Meis1* or *Mn1* knockdown in LSCs^{RH-RAS}.
 (I) Relative proliferation of indicated shRNA-expressing LSCs^{RH-RAS} (GFP⁺) versus parental cells (GFP⁻). These GFP⁻ and GFP⁺ cells were mixed in a 1:1 ratio at day 0, followed by measurement of percentage of GFP⁺ cells.
 (J) Kaplan-Meier curve of mice engrafted with indicated shRNA-expressing LSCs^{RH-RAS}. The p values were calculated by log-rank test. Error bars denote \pm SD. See also Figure S3 and Table S2.

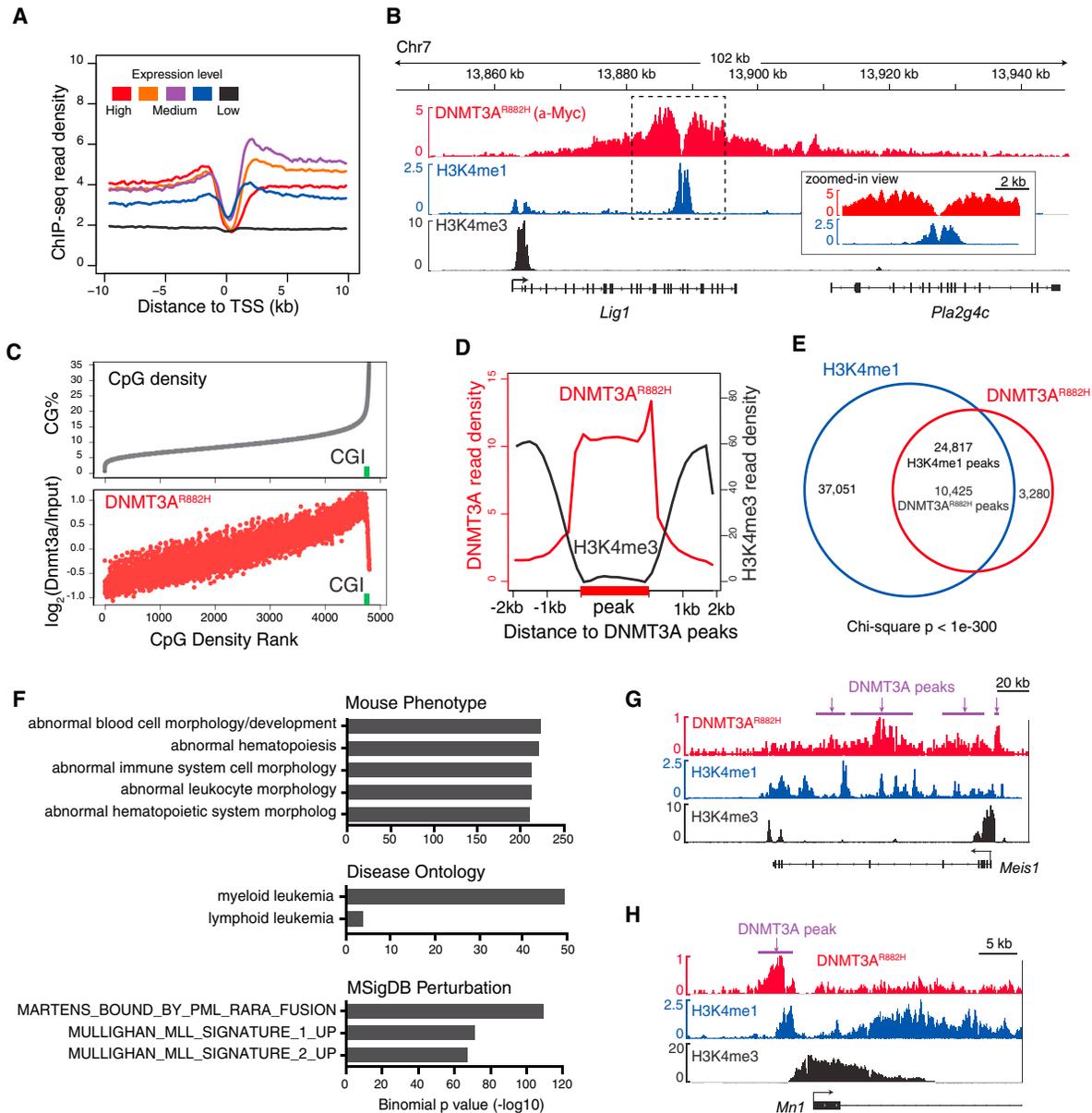


Figure 4. ChIP-Seq Reveals Chromatin Context-Dependent Binding of R882-Mutated DNMT3A to Genomic Regions, Including Stemness Genes such as a *Meis1-Mn1-Hoxa* Node

(A) DNMT3A^{R882H} ChIP-seq profiles across transcription start site (TSS) of genes with different expression levels in LSCs^{RH-RAS}.
 (B) Example ChIP-seq profiles for DNMT3A^{R882H}, H3K4me1, and H3K4me3 at the *Lig1* gene. Box shows a zoomed-in view of dashed-box region showing overlap of DNMT3A^{R882H} and H3K4me1 peaks.
 (C) Correlation of DNMT3A^{R882H} binding and CpG density. Shown is percentage of CpG density (gray) and DNMT3A^{R882H} ChIP-seq reads (red) at 1-kb windows of the entire genome ranked by CpG density. Green square, CpG island (CGI).
 (D) Plot of averaged DNMT3A^{R882H} (red) and H3K4me3 (black) ChIP-seq signals at DNMT3A^{R882H} peaks (labeled in bold on x axis) and surrounding regions (± 2 kb).
 (E) Venn diagram shows significant overlap of DNMT3A^{R882H} and H3K4me1 peaks in LSCs^{RH-RAS}.
 (F) Genomic Regions Enrichment of Annotations Tool (GREAT) analysis shows enrichment of indicated gene signatures among DNMT3A^{R882H} peaks.
 (G and H) ChIP-seq profiles of DNMT3A^{R882H}, H3K4me1 and H3K4me3 at *Meis1* (G) and *Mn1* (H). Purple bars, DNMT3A^{R882H} peak calls.
 See also [Figure S4](#) and [Table S3](#).

“writers.” In addition, overall gain of H3K27ac at hypo-DMRs was found to be significant regardless of expression changes of their associated genes ([Figure S6J](#)), indicating that H3K27ac

gain at hypo-DMRs is not merely a consequence of gene activation, as exemplified by that found at hypo-DMRs of *Kdm2b*, *Sirt4*, and *Pax5* ([Figures S6K–S6M](#)).

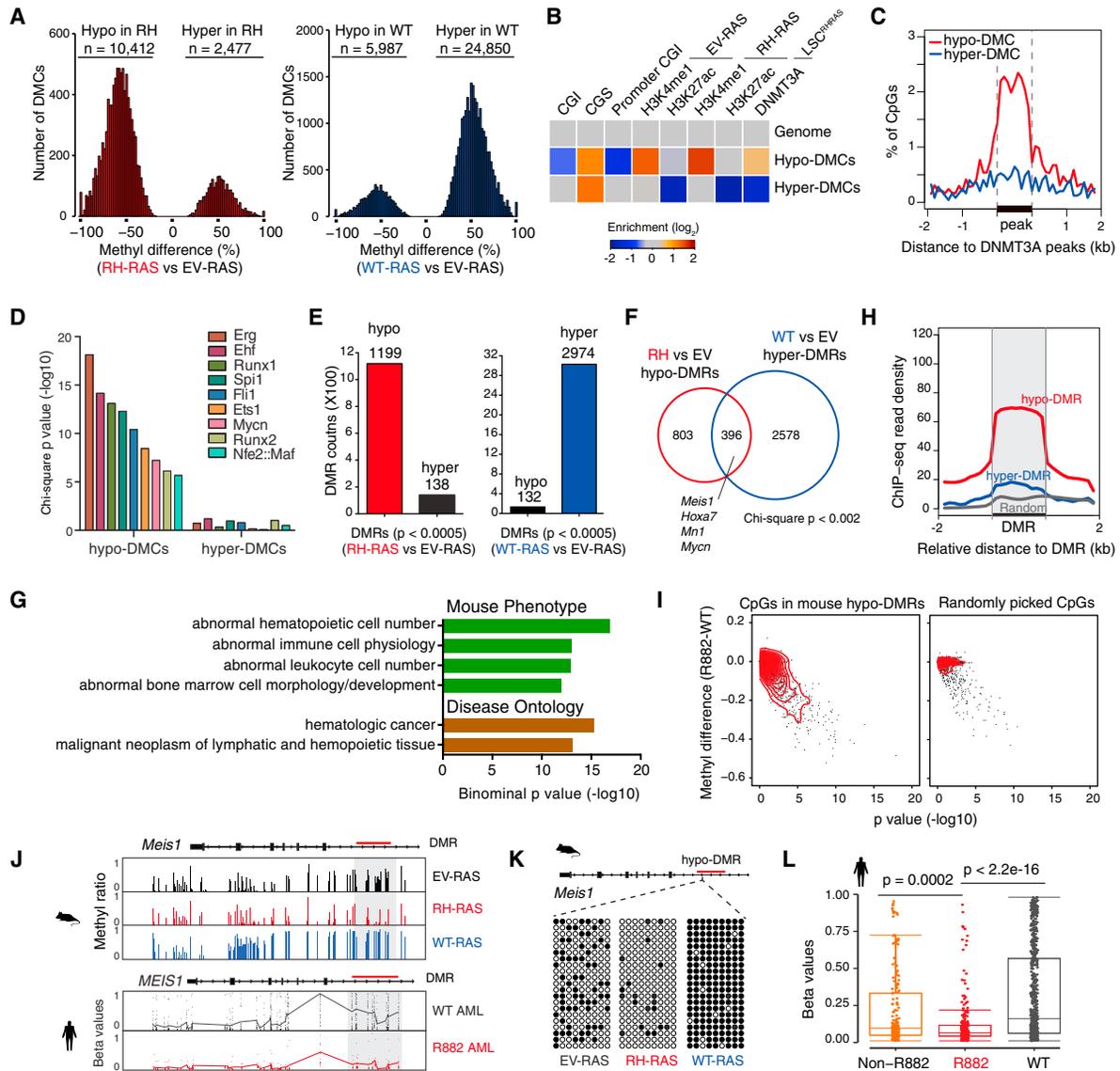


Figure 5. DNMT3A^{R882H} Induces Focal CpG Hypomethylations Enriched at H3K4me1-Demarcated, Gene-Regulatory Sites in HSPCs

(A) Distribution of DMCs (defined by $q < 0.05$) in the genome of murine HSPCs transduced with RH-RAS or WT-RAS, relative to EV-RAS.
 (B) Heatmap showing enrichment of DNMT3A^{R882H}-associated DMCs at indicated genomic regions or ChIP-seq peaks in comparison with genome average. The enrichment value was calculated as $\log_2(\text{observed}/\text{expected})$ of the DMC numbers. CGI, CpG island; CGS, CpG shore.
 (C) Distribution of DNMT3A^{R882H}-associated DMCs across DNMT3A^{R882H} ChIP-seq peaks (shown by a bold bar on x axis). y Axis shows percentage of DMCs located at 100-bp window of genomic regions centered on DNMT3A^{R882H} peaks.
 (D) Enrichment of indicated TF binding motifs in DNMT3A^{R882H}-associated hypo-DMCs and hyper-DMCs.
 (E) Summary of DMRs identified in the HSPCs with RH-RAS or WT-RAS, relative to EV-RAS.
 (F) Venn diagram showing overlap of DNMT3A^{R882H} and DNMT3A^{WT}-associated DMRs.
 (G) GREAT annotation of DNMT3A^{R882H}-associated hypo-DMRs.
 (H) H3K4me1 profiles at DNMT3A^{R882H}-associated hypo-DMRs, hyper-DMRs, and random control regions. Plotted across DMRs (labeled by a bold line on x axis) were averaged H3K4me1 ChIP-seq read densities in EV-RAS cells.
 (I) Scatterplots showing methylation changes of selected CpGs in human AMLs with DNMT3A R882 mutation relative to DNMT3A WT AMLs. Mean methylation differences (y axis) and p value (x axis) for each CpG between two AML patient groups were plotted. Left, CpGs in the human genome homologous to DNMT3A^{R882H}-associated hypo-DMRs identified in murine HSPCs; right, randomly picked CpG controls.
 (J) DNA methylation profiles of Meis1 in indicated murine HSPCs and MEIS1 in human AMLs with WT (n = 50) or R882-mutated (n = 20) DNMT3A. Faded points show individual CpG methylation β values and connected lines indicate the mean methylation levels at each CpG site. Gray box and red bar represent a hypo-DMR in intron 6.
 (K) Bisulfite sequencing of the Meis1 intron 6 DMR in indicated murine HSPC samples.

(legend continued on next page)

Because DNMT3A^{R882H}-induced hypo-DMRs can be found outside of gene-regulatory regions, we focused on those overlapping with a peak of H3K4me1 (a total of 777 DMRs) or H3K27ac (333 DMRs) in at least one cell condition and found that, in either case, 9- to 11-fold more DMRs showed enhanced H3K27ac levels than those with decreased H3K27ac (Figure 6F). These results indicate that DNA hypomethylation facilitates H3K27ac gain at gene-regulatory sites but also acts in a context-dependent manner. Consistently, more hypo-DMRs with gained H3K27ac were observed at regions showing a greater loss of CpG methylation (Figure 6G), supporting the degree of DNA hypomethylation as a contributing factor that fine-tunes functional output of gene-regulatory sites. Moreover, genes with increased H3K27ac at their hypo-DMRs were found to be enriched in HSPCs expressing RH-RAS relative to EV-RAS (Figure 6H), from which we identified 57 genes as both epigenetically altered and transcriptionally activated by DNMT3A^{R882H} (thus hereafter termed “DNMT3A^{R882H} signature genes,” Figure 6I and Table S6). Notably, these DNMT3A^{R882H} signature genes included DNMT3A^{R882}-associated stemness genes studied above (a *Meis1-Mn1-Hoxa* node and *Mycn*) as well as other putative AML-promoting genes such as *Id2*, *Bcl2*, and *Runx3* (Figure 6I).

The *Meis1* Intron 6 Enhancer Carrying DNMT3A^{R882H}-Induced CpG Hypomethylation Is Crucial for *Meis1* Gene Activation in LSCs

To demonstrate a causal role of DNMT3A^{R882H}-induced focal DNA hypomethylation in gene-expression regulation, we cloned sequences from a panel of hypo-DMRs into a CpG-free reporter system designed to assess putative gene-regulatory activity and its relationship to CpG methylation (Schmidl et al., 2009). We found that all tested hypo-DMRs possess strong expression-enhancing activity in the absence of their CpG methylation (Figure 6J). CpG methylation of these hypo-DMRs completely abolished their expression-enhancing activities (Figure 6J), demonstrating a hypomethylation-dependent activation of *cis*-regulatory elements harbored within hypo-DMRs. To further verify DMR-associated enhancer activity in LSCs^{RH-RAS}, we closely examined a hypo-DMR located in the intron 6 of *Meis1* (Figure 6C, green bar) because *Meis1* is a critical effector gene for DNMT3A^{R882H}-associated AML progression (Figures 3H–3J) and this hypo-DMR is also found to be conserved in human AMLs with DNMT3A R882 mutation (Figures 5J–5L). Notably, this hypo-DMR is positive for H3K4me1 (Figure 6C) and has a significant overlap with a previously reported *MEIS1* enhancer in human cells (Xiang et al., 2014). First, we carried out chromosome conformation capture (3C), a surrogate assay for scoring enhancer usage and promoter association, and indeed detected a long-range looping interaction of the intron 6 hypo-DMR with the *Meis1* promoter in LSCs^{RH-RAS} (Figure 6K). To further determine the role of this putative intron 6 enhancer in DNMT3A^{R882H}-induced *Meis1* gene activation, we employed

the CRISPR/Cas9-based genomic editing technology. Cas9 and a pair of single guide RNAs (sgRNA) targeting boundaries of the *Meis1* hypo-DMR were transduced into LSCs^{RH-RAS} (Figure 6L). PCR and direct sequencing confirmed sgRNA-mediated specific deletion of the hypo-DMR in five independent LSC^{RH-RAS} lines (Figures 6M, 6N, and S6N). In all cases, ablation of this putative enhancer significantly reduced *Meis1* expression (Figure 6O). Consistently, among human AMLs with DNMT3A R882 mutation, lower DNA methylation at the *MEIS1* intron 6 correlated with higher expression of *MEIS1* (Figure 6P). It is also worth noting that 54.5% (6 of 11) of DNMT3A WT AMLs display significant DNA methylation of *MEIS1* intron 6 and yet express *MEIS1* at high levels (Figure 6P), indicating that different gene activation mechanisms exist in these AML cases. Together, using *Meis1* as a paradigm example, we show that focal CpG hypomethylation induced by DNMT3A R882 mutation promotes enhancer activation and expression of key AML genes.

Dot11 Inactivation Suppresses DNMT3A^{R882H}-Associated LSC Properties and Aberrant Activation of Stemness Gene Programs

To explore the potential strategy for reversing DNMT3A^{R882H}-induced gene deregulation and thus treating DNMT3A-mutated leukemia, we conducted compound treatment studies with a collection of epigenetic regulator inhibitors and identified that LSCs^{RH-RAS} showed a significantly higher sensitivity to a Dot11 inhibitor, SGC0946, relative to control cells without DNMT3A mutation, i.e., LSCs expressing *NRAS*^{G12D} plus oncogenic TFs (Figure S7A). Dot11, a histone H3 lysine 79 (H3K79) methyltransferase, belongs to a transcription elongation regulatory complex that engages acetylated histones at *cis*-regulatory sites (Li et al., 2014). Genomic profiling of H3K79 dimethylation (H3K79me2) detected its overall elevation at DNMT3A^{R882H}-associated hypo-DMRs in HSPCs (Figure 7A), as exemplified by those at *Meis1*, *Hoxa*, *Mn1*, and *Mycn* (Figures 7B and S7B). We confirmed H3K79me2 gain at these genes by ChIP-qPCR (Figure S7C). Next, we asked whether pharmacological inhibition of Dot11 could reverse DNMT3A^{R882H}-induced gene activation. We first confirmed SGC0946-mediated suppression of H3K79me2 in LSCs^{RH-RAS} (Figure S7D), followed by microarray profiling. Notably, after SGC0946 treatment, we detected significant downregulation of DNMT3A^{R882H} signature genes (Figures 7C and 7D) and concurrent upregulation of myeloid differentiation genes in LSCs^{RH-RAS} (Figures 7D and S7E). Although *Hoxa* and *Meis1* were shown as part of MLL-AF9 target genes that are dependent on Dot11 (Chen et al., 2015), the DNMT3A^{R882H} signature genes displayed a greater sensitivity to Dot11 inhibitors than MLL-AF9 targets in LSCs^{RH-RAS} (Figure 7E); conversely, the DNMT3A^{R882H} signature genes do not show overall response to Dot11 inhibitors in MLL-AF9-transformed AML cells (Figure S7F). These analyses indicate that DNMT3A R882 mutation confers a unique dependency on the Dot11 enzymatic activity in AML. We further verified downregulation of DNMT3A^{R882H}-associated

(L) Box plots of methylation β values of all CpGs (shown as dots in box plot) at *MEIS1* intron 6 in human AMLs with R882-mutated DNMT3A ($n = 20$) relative to AMLs with either non-R882 mutated ($n = 15$) or WT ($n = 50$) DNMT3A. Horizontal line, median; box, interquartile range; whiskers extend to $1.5 \times$ the interquartile range. The p values were calculated by Mann-Whitney U test. See also Figure S5; Tables S4 and S5.

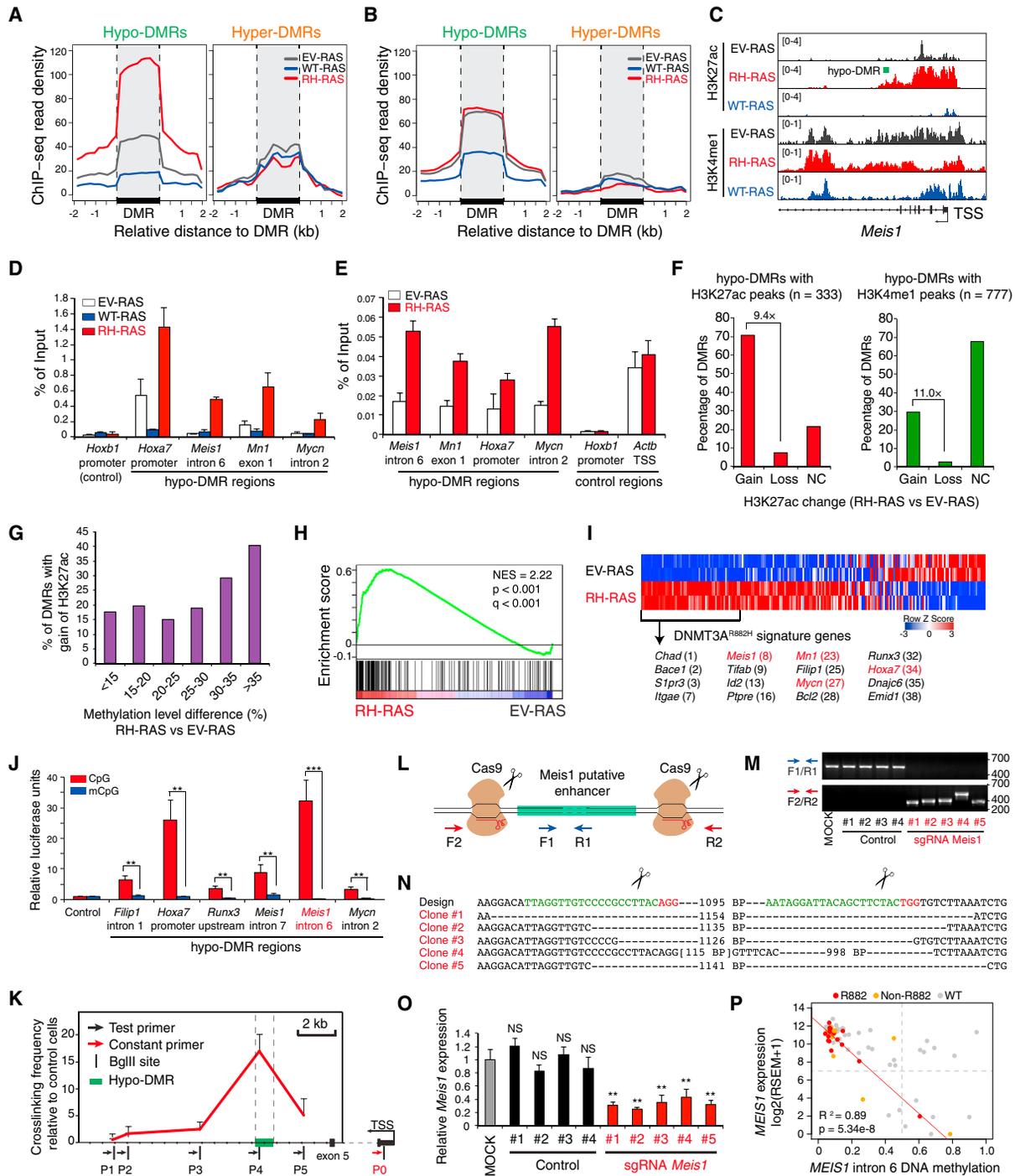


Figure 6. DNMT3A^{R882H}-Associated Hypo-DMRs Gain Epigenetic Alterations Associated with Gene Activation

(A and B) H3K27ac (A) and H3K4me1 (B) profiles at DNMT3A^{R882H}-associated DMRs (bold bar on x axis) and the surrounding regions. Averaged ChIP-seq read densities in HSPCs with EV-RAS, RH-RAS, or WT-RAS were plotted.

(C) H3K27ac and H3K4me1 profiles at *Meis1* intron 6 in indicated HSPCs. Green bar, hypo-DMR.

(D and E) ChIP-qPCR of H3K27ac (D) and p300 binding (E) at hypo-DMRs in indicated HSPCs.

(F) Percentage of DNMT3A^{R882H}-associated hypo-DMRs showing indicated H3K27ac changes in HSPCs with RH-RAS versus EV-RAS. Gain, increased H3K27ac; Loss, reduced H3K27ac; NC, no significant H3K27ac change. The total DMRs used for calculation were hypo-DMRs carrying H3K27ac (left) or H3K4me1 (right) in at least one cell condition.

(G) Percentage of DNMT3A^{R882H}-associated hypo-DMRs (n = 1,199) showing H3K27ac gain in HSPCs with RH-RAS versus EV-RAS, when these hypo-DMRs are divided based on degree of DNA methylation reduction (x axis) shown in the same samples.

(H) GSEA shows that genes with gain of H3K27ac at hypo-DMRs are enriched in HSPCs 16 days after transduction of RH-RAS, relative to EV-RAS.

(legend continued on next page)

stemness genes *Hoxa*, *Meis1*, *Mn1*, and *Mycn* after treatment with SGC0946 (Figures 7F, 7G, and S7G) or knockdown of Dot1l (Figures 7H, S7H, and S7I). In response to Dot1l inactivation, multiple murine and human AML lines bearing *DNMT3A* mutation showed suppressed in vitro growth (Figures 7I, S7J, and S7K) and concurrent cell differentiation (Figures 7J, 7K, and S7L–S7N). *DNMT3A*-mutated human AML lines also had decreased *HOXA* or *MEIS1* expression upon DOT1L blockade (Figures S7O and S7P). In contrast, various murine and human leukemia lines established by oncogenic TFs were all insensitive to Dot1l inhibition (Figures 7I and S7K). Also, enforced expression of *HOXA9* plus *MEIS1* reversed sensitivity of LSCs^{RH-RAS} to Dot1l inhibition (Figure 7L), demonstrating a crucial role of these TFs in *DNMT3A*^{R882H}-mediated oncogenic effects. Importantly, knockdown of Dot1l in LSCs^{RH-RAS} or their pretreatment with Dot1l inhibitors significantly delayed in vivo AML progression and prolonged the survival of engrafted mice (Figure 7M). Collectively, we show that expression of *DNMT3A*^{R882H} confers Dot1l dependency in AML and that reversing *DNMT3A*^{R882H}-induced gene activation by Dot1l inhibition may provide a potential therapeutic means for the treatment of AMLs with *DNMT3A* mutation.

DISCUSSION

In this study, we report a set of ex vivo LSC and in vivo murine AML model systems for studying the functionality of *DNMT3A* R882 mutation in AML pathogenesis. Using these human disease-mimicking models, we have (1) defined a causal role of *DNMT3A*^{R882H} in promoting AML transformation in vitro and in vivo; (2) identified *DNMT3A*^{R882H}-deregulated gene pathways, including a *Meis1*-*Mn1*-*Hoxa* TF node that we functionally validated as essential for *DNMT3A*^{R882H}-mediated AML progression; (3) shown that *DNMT3A*^{R882H} directly binds to gene-regulatory sites, notably enhancers, inducing focal DNA hypomethylation and concurrent gain of histone acetylation; (4) determined a critical role of the epigenetically altered enhancer and *cis*-regulatory elements for *DNMT3A*^{R882H}-associated gene activation; and (5) demonstrated, importantly, that pharmacological inhibition of Dot1l reverses the mutant *DNMT3A*-associated gene activation, thus providing a potential therapeutic avenue for the affected AMLs.

The molecular pathways identified in this study help explain several important biological phenomena related to *DNMT3A* mutation and hematological disease. First, as the *Meis1*-*Mn1*-*Hoxa*

circuitry is crucial for both normal expansion of HSCs and malignant transformation of LSCs (Argiropoulos and Humphries, 2007; Heuser et al., 2011), deregulation of this TF node by R882-mutated *DNMT3A* provides a molecular explanation not only for malignant hematopoiesis but also for clonal hematopoiesis, a phenotype strongly associated with *DNMT3A* mutation (Genovese et al., 2014; Jaiswal et al., 2014; Xie et al., 2014). In addition, these findings help explain a mutually exclusive pattern for *DNMT3A* mutation and *MLL* rearrangement in AMLs (Cancer Genome Atlas Research Network, 2013; Patel et al., 2012) because the latter itself is a strong inducer of *Meis1* and *Hoxa* activation (Chi et al., 2010).

Our results also demonstrate the requirement of cooperation between *DNMT3A* mutation and the activated kinase such as RAS for AML induction. RAS mutation alone induces a hyperproliferative phenotype but does not support self-renewal, which is in agreement with previous studies (Zhang et al., 2009); RAS activation was also known to induce cell senescence, a barrier of cancer development (Campisi and d'Adda di Fagagna, 2007). On the other hand, *DNMT3A* mutation confers aberrant HSPC self-renewal, blocks differentiation programs, and yet lacks a pro-proliferation effect; besides a *Meis1*-*Mn1*-*Hoxa* node we have functionally confirmed as essential for *DNMT3A*^{R882H}-associated AML, other downstream targets of *DNMT3A*^{R882H}, such as pro-survival (*Bcl2*), anti-differentiation (*Id2*), and stemness (*Mycn*) genes, might be equally crucial for AML progression. These findings suggest that synergy between *DNMT3A* and kinase mutations is likely due to their differential effects on pathways relating to AML development. However, it is also possible that the two mutations may affect distinctive sets as well as the same sets of downstream effectors via genetic or epigenetic mechanisms. A similar synergy is most likely to exist between *DNMT3A* mutation and the activated FLT3, which acts upstream of RAS and coexists with the former in human AMLs as well.

Our studies clearly show that *DNMT3A* mutation-induced CpG hypomethylations are not random: they are significantly enriched at gene-regulatory sites, notably, putative enhancers marked by H3K4me1 as well as the binding sites of master hematopoietic TFs. Precise mechanisms by which CpG methylation of these *cis*-regulatory sites regulates gene expression remain to be fully studied. For example, despite a large number of DMCs found to be associated with either *DNMT3A* or *TET2* mutation in AML, a relatively small number of genes show changes in their expression (Russler-Germain et al., 2014; Shih et al., 2015). A possible

(I) Heatmap shows expression of genes in (H) ranked by higher expression in HSPCs with RH-RAS, relative to EV-RAS. The significantly upregulated genes in RH-RAS HSPCs are defined as “*DNMT3A*^{R882H} signature genes” (n = 57), with selected ones listed along with their respective rankings (bottom).

(J) Quantification of expression-enhancing activity of *DNMT3A*^{R882H}-associated hypo-DMRs with the embedded CpGs either non-methylated (CpG) or methylated (mCpG) using a CpG-free luciferase reporter system. The reporter without any DMR insertion was used as control. The p values were calculated by Student's t test.

(K) 3C assay shows looping interaction of the *Meis1* intron 6 hypo-DMR (P4) to gene promoter (P0), relative to other tested sites.

(L and M) Scheme (L) and PCR validation (M) of CRISPR/Cas9-mediated deletion of the *Meis1* intron 6 DMR. MOCK, parental LSC^{RH-RAS}; Control, no sgRNA; sgMeis1, a pair of sgRNAs that target the DMR boundaries.

(N) Sequencing of the genomic PCR products from F2/R2 primers shows CRISPR/Cas9-induced deletion of the *Meis1* intron 6 DMR.

(O) Expression levels of *Meis1* in LSC^{RH-RAS} lines shown in (M). The p values were calculated by Student's t test by comparing with MOCK.

(P) Impact of DNA methylation levels in *MEIS1* intron 6 in cytogenetically normal human AMLs grouped by *DNMT3A* WT (n = 45), non-R882 (n = 13), and R882 mutations (n = 16). Plotted were mean methylation β values of CpGs at *MEIS1* intron 6 and log₂-transformed expression values of RNA sequencing by expectation maximization (RSEM). R² and p values shown were determined with data of R882-mutant AMLs.

Error bar denotes \pm SD. **p < 0.01, ***p < 0.001; NS, not significant. See also Figure S6; Tables S4 and S6.

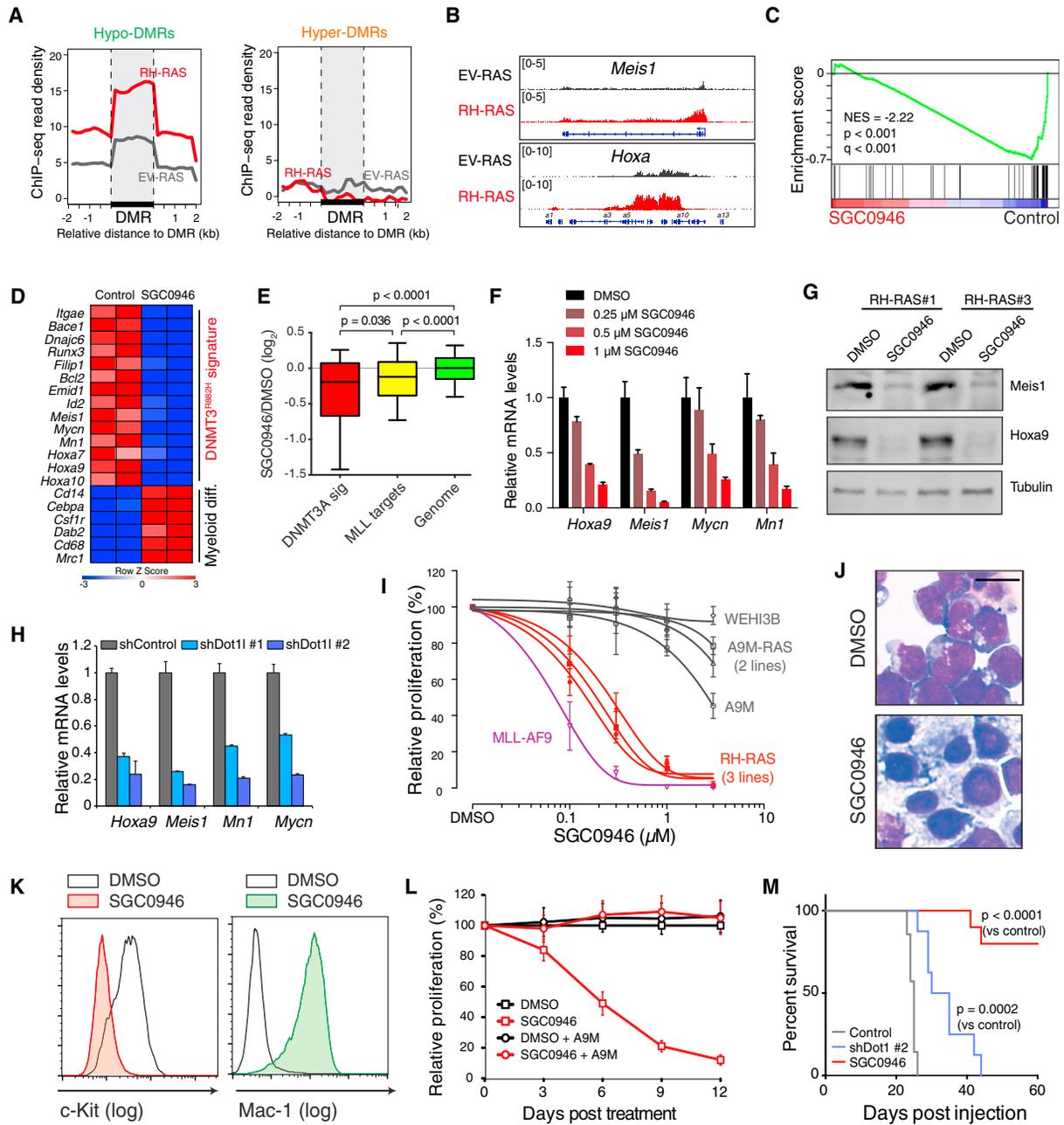


Figure 7. Dot1l Inhibition Reverses DNMT3A^{R882H}-Mediated Aberrant Transactivation of Stem Cell Genes, thereby Suppressing Acute Leukemogenicity

(A) Averaged H3K79me2 ChIP-seq signals at DNMT3A^{R882H}-associated hypo-DMRs and hyper-DMRs in HSPCs with RH-RAS or EV-RAS.
 (B) H3K79me2 profiles at *Meis1* and *Hoxa* in indicated HSPCs.
 (C) GSEA shows downregulation of DNMT3A^{R882H} signature genes in LSCs^{RH-RAS} after treatment with 1 μ M SGC0946 for 4 days.
 (D) Heatmap shows downregulation of DNMT3A^{R882H} signature genes and upregulation of myeloid differentiation genes in SGC0946-treated LSCs^{RH-RAS} versus mock treatment.
 (E) Box plots show relative expression of DNMT3A signature genes (n = 54), MLL-AF9 gene targets (n = 129), and all genes in the genome in SGC0946-treated LSCs^{RH-RAS}, relative to mock treatment. Horizontal line, median; box, interquartile range; whiskers, 10–90 percentiles. The p values were calculated by Mann-Whitney U test.
 (F and G) qRT-PCR (F) and immunoblot (G) of indicated genes and proteins in LSCs^{RH-RAS} 6 days after treatment with SGC0946.
 (H) Expression of indicated genes in LSCs^{RH-RAS} transduced with *Dot1l* shRNAs or vector control.
 (I) Relative growth of LSCs^{RH-RAS} and other AML lines established by MLL-AF9, *Hoxa9* plus *Meis1* (A9M), A9M plus NRAS^{G12D} (A9M-RAS), and *Hoxb8* plus *Meis2* (WEHI3B) after a 12-day treatment with SGC0946 versus DMSO.
 (J and K) Wright-Giemsa staining (J) and FACS analysis (K) of LSCs^{RH-RAS} 6 days after treatment with DMSO or 1 μ M SGC0946. Scale bar, 10 μ m.

(legend continued on next page)

explanation is that the effect of CpG methylation on gene expression is context dependent (Baubec and Schubeler, 2014): degree of CpG methylation change, density, or genomic location of CpG, methyl-CpG “readers,” and TF binding are all possible factors affecting the ultimate effect of DNA methylation on gene expression. Unlike histone (de)acetylation, CpG (de) methylation at distal *cis*-regulatory sites such as enhancers may act as a permissive mechanism influencing gene expression, rather than a strong and instructive one controlling levels of gene activation and transcription. Nevertheless, using reporter assays and CRISPR/Cas9-mediated enhancer editing, we have determined the role of select hypo-DMRs in the activation of associated target genes such as *Meis1*.

This study also provides useful information on how to treat *DNMT3A*-mutated AMLs. Pharmacological blockade of Dot1l reversed *DNMT3A* mutation-induced gene activation, resulting in an impaired AML pathogenesis. In the future, examination of other “druggable” factors would likely identify additional therapeutic strategies for the treatment of *DNMT3A*-mutated AMLs. Therefore, in addition to elucidating the underlying oncogenic mechanisms, the *ex vivo* and *in vivo* model systems presented herein should be useful for exploring AML therapeutics.

EXPERIMENTAL PROCEDURES

The detailed procedures of plasmid construction, cell culture, antibody and immunoblot, flow cytometry, microarray analysis, ChIP-seq, eRRBS, Exome-seq, qRT-PCR, ChIP-qPCR, 3C-qPCR, shRNA-mediated knock-down, luciferase reporter assay, CRISPR/Cas9-mediated genomic editing, as well as the detailed information for computational and statistical analysis of deep sequencing data are described in [Supplemental Experimental Procedures](#).

In Vitro CFU Assay with Serial Replating

Following lineage-negative (Lin^-) enrichment and retroviral transduction, 30,000 infected HSPCs were plated in the semi-solid methylcellulose cultivation system (Methocult; Stem Cell Technologies), followed by CFU counting and replating for every 10–14 days according to the manufacturer’s protocol.

Animal Studies and In Vivo Leukemogenic Assay

All animal experiments were approved by and performed in accord with the guidelines of the Institutional Animal Care and Use Committee at the University of North Carolina. Leukemogenic potentials of transduced HSPCs were evaluated by bone marrow transplantation into sublethally irradiated syngeneic mice. In brief, 200,000 of bone marrow-derived Lin^- HSPCs following procedures of cytokine stimulation, retroviral transduction, and drug selection were injected via tail vein to recipient mice as described previously (Wang et al., 2009).

Statistical Analysis

Data are presented as the mean \pm SD of three independent experiments unless otherwise noted. Statistical analysis was performed with Student’s *t* test for comparing two sets of data with assumed normal distribution. We used a Mann-Whitney U test for data not showing a normal distribution, chi-square test for categorical variables, and log-rank test for Kaplan-Meier survival curves to determine statistical significance. $p < 0.05$ was considered significant.

ACCESSION NUMBERS

The microarray, eRRBS-seq, and ChIP-seq data reported in this paper have been deposited in the Gene Expression Omnibus with accession number GEO: GSE71475.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and six tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ccell.2016.05.008>.

AUTHOR CONTRIBUTIONS

R.L. designed the research, performed experiments and computational analysis, interpreted data, and wrote the manuscript. P.W., S.R., and D.Z. performed computational analysis. T.P., Y.Z., and W.C. helped with animal modeling, viral transduction, and immunostaining studies, respectively. K.C. and P.A.W. participated in DNA methylome studies. O.A. helped with human AML cell studies. D.Z. supervised the computational studies. G.G.W. supervised the work, designed the research, interpreted data, and wrote the manuscript.

ACKNOWLEDGMENTS

We graciously thank Drs. Y. Xiong and K. Humphries for providing constructs, M. Kamps and M. Minden for leukemia lines, M. Torres for *Meis1* antibody, D. Bauer and F. Zhang for CRISPR/Cas9 systems, M. Rehli for a CpG-free reporter, and J. Bear for an shRNA vector used in this study. Thanks to Drs. D. Allison and L. Cai and other members of the Wang laboratory for helpful discussion and technical support. We thank UNC’s Genomics Core, Animal Studies Core, Flow Core, and HTSF core for their support of this work. This work is supported by NCI K99/R00 grant CA151683 to G.G.W., a DoD grant CA130247 to G.G.W., and grants of Gabrielle’s Angel Foundation to O.A. and G.G.W. G.G.W. is an American Society of Hematology Scholar and a Kimmel Scholar. R.L. is a Lymphoma Research Foundation Postdoctorate Fellow. UNC Cores including the flow cytometry facility are supported in part by the North Carolina Biotech Center Institutional Support Grant 2012-IDG-1006 and UNC Cancer Center Core Support Grant P30 CA016086.

Received: July 21, 2015

Revised: March 3, 2016

Accepted: May 19, 2016

Published: June 23, 2016

REFERENCES

- Abramovich, C., Pineault, N., Ohta, H., and Humphries, R.K. (2005). Hox genes: from leukemia to hematopoietic stem cell expansion. *Ann. N. Y. Acad. Sci.* **1044**, 109–116.
- Argiropoulos, B., and Humphries, R.K. (2007). Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766–6776.
- Baubec, T., and Schubeler, D. (2014). Genomic patterns and context specific interpretation of DNA methylation. *Curr. Opin. Genet. Dev.* **25**, 85–92.
- Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., et al. (2011). MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78.
- Campisi, J., and d’Adda di Fagnagna, F. (2007). Cellular senescence: when bad things happen to good cells. *Nat. Rev. Mol. Cell Biol.* **8**, 729–740.

(L) Effect of SGC0946 on growth of LSCs^{RH-RAS} transduced with vector or Hoxa9 plus *Meis1* (A9M). Relative proliferation was normalized to DMSO-treated cells.

(M) Survival of mice engrafted with LSCs^{RH-RAS}, either mock-treated, stably transduced with a *Dot1l* shRNA, or pretreated with 1 μM SGC0946 *ex vivo* for 6 days.

The *p* values were calculated by log-rank test.

Error bars denote \pm SD. See also [Figure S7](#) and [Table S6](#).

- Cancer Genome Atlas Research Network. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 368, 2059–2074.
- Celik, H., Mallaney, C., Kothari, A., Ostrander, E.L., Eultgen, E., Martens, A., Miller, C.A., Hundal, J., Klco, J.M., and Challen, G.A. (2015). Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. *Blood* 125, 619–628.
- Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* 44, 23–31.
- Chang, Y.I., You, X., Kong, G., Ranheim, E.A., Wang, J., Du, J., Liu, Y., Zhou, Y., Ryu, M.J., and Zhang, J. (2015). Loss of Dnmt3a and endogenous Kras(G12D/+) cooperate to regulate hematopoietic stem and progenitor cell functions in leukemogenesis. *Leukemia* 29, 1847–1856.
- Chen, C.W., Koche, R.P., Sinha, A.U., Deshpande, A.J., Zhu, N., Eng, R., Doench, J.G., Xu, H., Chu, S.H., Qi, J., et al. (2015). DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia. *Nat. Med.* 21, 335–343.
- Chi, P., Allis, C.D., and Wang, G.G. (2010). Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nat. Rev. Cancer* 10, 457–469.
- Eppert, K., Takenaka, K., Lechman, E.R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K.H., Poepl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* 17, 1086–1093.
- Genovese, G., Kahler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* 371, 2477–2487.
- Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* 517, 640–644.
- Heuser, M., Yun, H., Berg, T., Yung, E., Argiropoulos, B., Kuchenbauer, F., Park, G., Hamwi, I., Palmqvist, L., Lai, C.K., et al. (2011). Cell of origin in AML: susceptibility to MN1-induced transformation is regulated by the MEIS1/AbdB-like HOX protein complex. *Cancer Cell* 20, 39–52.
- Holz-Schietinger, C., Matje, D.M., and Reich, N.O. (2012). Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J. Biol. Chem.* 287, 30941–30951.
- Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G., MacDonald, J., Cezard, T., Bilenky, M., et al. (2012). Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood* 119, 388–398.
- Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* 371, 2488–2498.
- Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492.
- Kim, S.J., Zhao, H., Hardikar, S., Singh, A.K., Goodell, M.A., and Chen, T. (2013). A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood* 122, 4086–4089.
- Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* 442, 818–822.
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science* 345, 943–949.
- Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J., et al. (2010). DNMT3A mutations in acute myeloid leukemia. *N. Engl. J. Med.* 363, 2424–2433.
- Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., Ren, Y., Jin, Q., Dent, S.Y., Li, W., et al. (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* 159, 558–571.
- Mayle, A., Yang, L., Rodriguez, B., Zhou, T., Chang, E., Curry, C.V., Challen, G.A., Li, W., Wheeler, D., Rebel, V.I., and Goodell, M.A. (2015). Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood* 125, 629–638.
- Noh, K.M., Wang, H., Kim, H.R., Wenderski, W., Fang, F., Li, C.H., Dewell, S., Hughes, S.H., Melnick, A.M., Patel, D.J., et al. (2015). Engineering of a histone-recognition domain in Dnmt3a alters the epigenetic landscape and phenotypic features of mouse ESCs. *Mol. Cell* 59, 89–103.
- Patel, J.P., Gonen, M., Figueroa, M.E., Fernandez, H., Sun, Z., Racevskis, J., Van Vlierberghe, P., Dolgalev, I., Thomas, S., Aminova, O., et al. (2012). Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* 366, 1079–1089.
- Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283.
- Russler-Germain, D.A., Spencer, D.H., Young, M.A., Lamprecht, T.L., Miller, C.A., Fulton, R., Meyer, M.R., Erdmann-Gilmore, P., Townsend, R.R., Wilson, R.K., and Ley, T.J. (2014). The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell* 25, 442–454.
- Schmidl, C., Klug, M., Boeld, T.J., Andreesen, R., Hoffmann, P., Edinger, M., and Rehli, M. (2009). Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Res.* 19, 1165–1174.
- Schubeler, D. (2015). Function and information content of DNA methylation. *Nature* 517, 321–326.
- Shih, A.H., Abdel-Wahab, O., Patel, J.P., and Levine, R.L. (2012). The role of mutations in epigenetic regulators in myeloid malignancies. *Nat. Rev. Cancer* 12, 599–612.
- Shih, A.H., Jiang, Y., Meydan, C., Shank, K., Pandey, S., Barreyro, L., Antony-Debre, I., Viale, A., Socci, N., Sun, Y., et al. (2015). Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell* 27, 502–515.
- Shlush, L.I., Zandi, S., Mitchell, A., Chen, W.C., Brandwein, J.M., Gupta, V., Kennedy, J.A., Schimmer, A.D., Schuh, A.C., Yee, K.W., et al. (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* 506, 328–333.
- Wang, G.G., Pasillas, M.P., and Kamps, M.P. (2005). Meis1 programs transcription of FLT3 and cancer stem cell character, using a mechanism that requires interaction with Pbx and a novel function of the Meis1 C-terminus. *Blood* 106, 254–264.
- Wang, G.G., Pasillas, M.P., and Kamps, M.P. (2006). Persistent transactivation by meis1 replaces hox function in myeloid leukemogenesis models: evidence for co-occupancy of meis1-pbx and hox-pbx complexes on promoters of leukemia-associated genes. *Mol. Cell Biol.* 26, 3902–3916.
- Wang, G.G., Cai, L., Pasillas, M.P., and Kamps, M.P. (2007). NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat. Cell Biol.* 9, 804–812.
- Wang, G.G., Song, J., Wang, Z., Dormann, H.L., Casadio, F., Li, H., Luo, J.L., Patel, D.J., and Allis, C.D. (2009). Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* 459, 847–851.
- Wong, P., Iwasaki, M., Somerville, T.C., So, C.W., and Cleary, M.L. (2007). Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes Dev.* 21, 2762–2774.
- Xiang, P., Wei, W., Lo, C., Rosten, P., Hou, J., Hoodless, P.A., Bilenky, M., Bonifer, C., Cockerill, P.N., Kirkpatrick, A., et al. (2014). Delineating MEIS1 cis-regulatory elements active in hematopoietic cells. *Leukemia* 28, 433–436.
- Xie, M., Lu, C., Wang, J., McLellan, M.D., Johnson, K.J., Wendl, M.C., McMichael, J.F., Schmidt, H.K., Yellapantula, V., Miller, C.A., et al. (2014).

Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* *20*, 1472–1478.

Xu, J., Wang, Y.Y., Dai, Y.J., Zhang, W., Zhang, W.N., Xiong, S.M., Gu, Z.H., Wang, K.K., Zeng, R., Chen, Z., and Chen, S.J. (2014). DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *Proc. Natl. Acad. Sci. USA* *111*, 2620–2625.

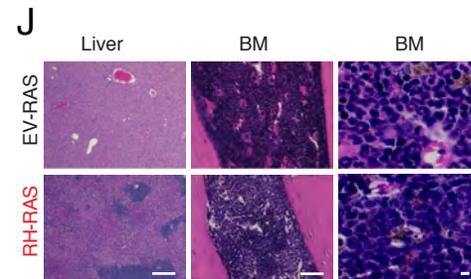
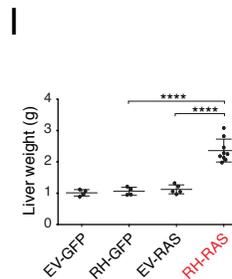
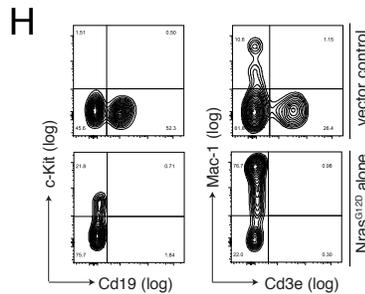
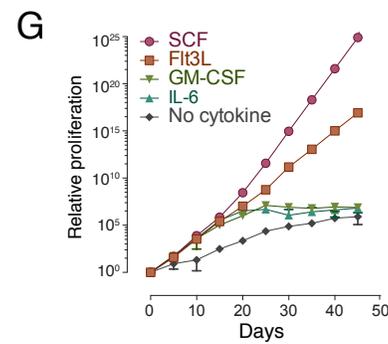
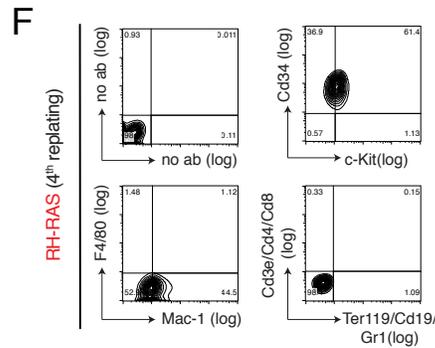
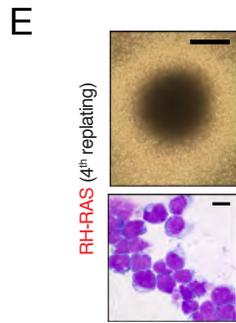
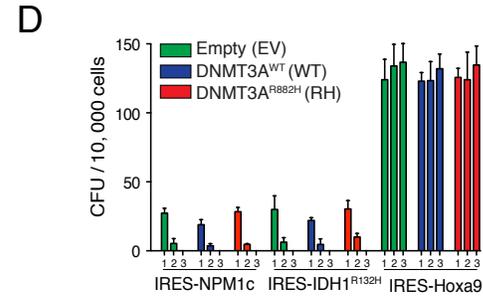
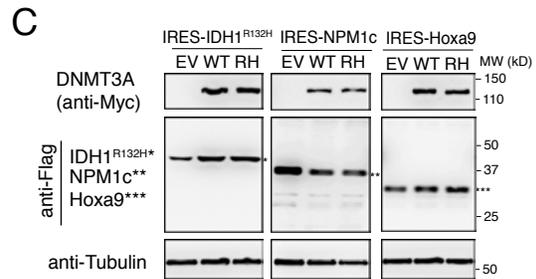
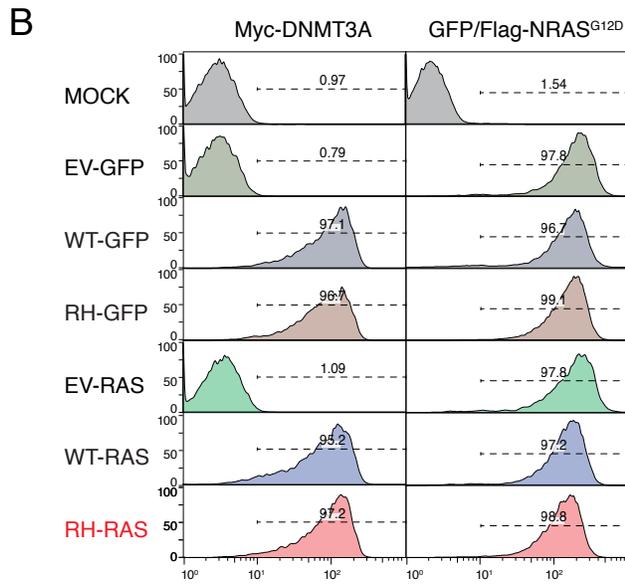
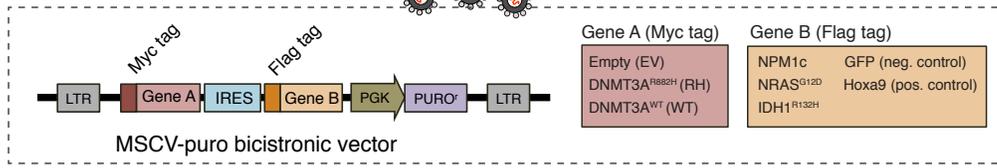
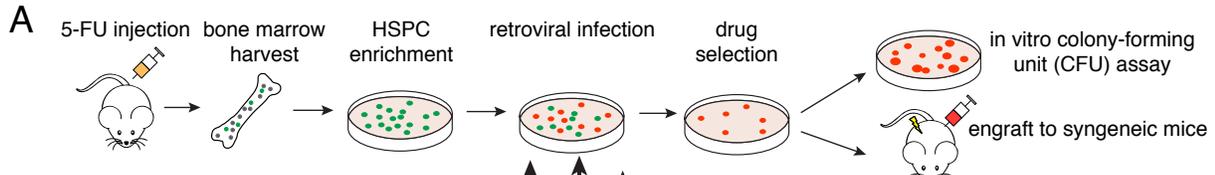
Yan, X.J., Xu, J., Gu, Z.H., Pan, C.M., Lu, G., Shen, Y., Shi, J.Y., Zhu, Y.M., Tang, L., Zhang, X.W., et al. (2011). Exome sequencing identifies somatic mu-

tations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat. Genet.* *43*, 309–315.

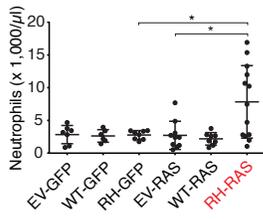
Yang, L., Rau, R., and Goodell, M.A. (2015). DNMT3A in haematological malignancies. *Nat. Rev. Cancer* *15*, 152–165.

Zhang, J., Wang, J., Liu, Y., Sidik, H., Young, K.H., Lodish, H.F., and Fleming, M.D. (2009). Oncogenic Kras-induced leukemogenesis: hematopoietic stem cells as the initial target and lineage-specific progenitors as the potential targets for final leukemic transformation. *Blood* *113*, 1304–1314.

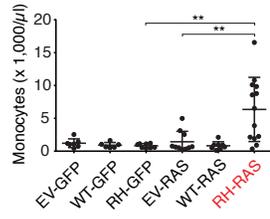
Supplemental Data



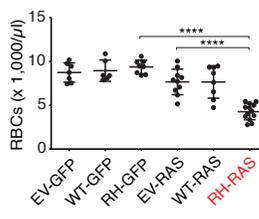
K



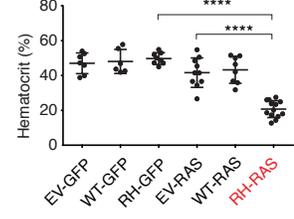
L



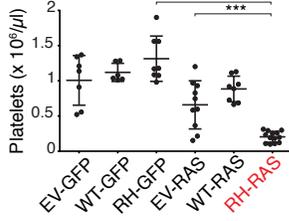
M



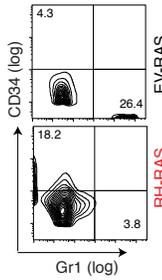
N



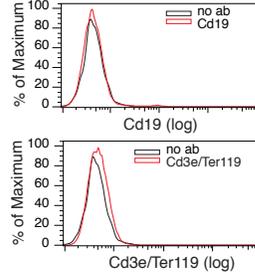
O



P



Q



R

Chromosome	Position	Ref	Alt	Gene	Snps_Eff_Effect	Snps_Eff_HGVs	AA_Change	Normal_Alt	Normal_Total	Tumor_Alt	Tumor_Total
Tumor 1											
chr2	180937146	C	T	<i>Srms</i>	downstream_gene_variant	c.1937G>A	No	0	43	19	50
chr3	127023244	G	T	<i>Ank2</i>	intron_variant	c.12+18806C>A	No	0	21	8	32
chr7	103517976	G	T	<i>Odz4</i>	intron_variant	c.90+14055G>T	No	0	48	9	40
chr8	121973844	A	G	<i>Necab2</i>	intron_variant	c.211-138A>G	No	0	45	9	59
chr8	121973870	G	A	<i>Necab2</i>	intron_variant	c.211-112G>A	No	0	32	8	50
chr10	5035165	T	C	<i>Syne1</i>	intron_variant	c.6819+312T>C	No	0	31	11	38
chr12	73371371	G	A	<i>Rtn1</i>	intron_variant	c.242-4684C>T	No	0	47	20	54
chr14	63495968	T	C	<i>Wdly2</i>	intron_variant	c.138-8679T>C	No	0	35	15	36
chr15	91191312	A	G	<i>Slc2a13</i>	intron_variant	c.1002-10751T>C	No	0	74	35	85
chr17	12940301	C	A	<i>Airn</i>	intron_variant	n.27+6011C>A	No	0	59	23	59
chr17	25025673	T	G	<i>Spsb3</i>	intron_variant	c.55-358T>G	No	0	66	32	73
chr18	12698079	T	C	<i>Lama3</i>	intron_variant	c.2458-99T>C	No	0	59	14	53
Tumor 2											
chr3	64503579	T	G	<i>Vmn2r7</i>	intron_variant	c.1369-6550A>C	No	0	46	18	48
chr4	18844775	A	G	<i>Cnbd1</i>	intron_variant	c.432-10487T>C	No	0	53	12	67
chr5	3344315	A	C	<i>Cdk6</i>	5_prime_UTR_variant	c.-52A>C	No	0	27	9	34
chr5	3344338	C	T	<i>Cdk6</i>	5_prime_UTR_variant	c.-29C>T	No	0	38	9	39
chr5	31067469	A	C	<i>Dpysl5</i>	intron_variant	c.262-12224A>C	No	0	62	46	80
chr5	129513237	C	T	<i>Sbx2</i>	intron_variant	c.30+1032G>A	No	0	85	43	76
chr6	60083270	G	A	<i>Gm5570</i>	intron_variant	n.901-15143G>A	No	0	35	19	40
chr7	8031898	T	G	<i>AC190408.2</i>	intron_variant	c.1102-49867A>C	No	0	49	18	42
chr7	13472971	G	T	<i>Zfp128</i>	intron_variant	c.290-2374G>T	No	0	183	24	209
chr7	106330699	G	T	<i>Dgal2</i>	intron_variant	c.121+203C>A	No	0	33	12	43
chr8	81902933	G	C	<i>Smad1</i>	intron_variant	c.-181-6670C>G	No	0	25	8	27
chr8	84127582	C	G	<i>Inpp4b</i>	intron_variant	c.-103+130864C>G	No	0	53	25	55
chr8	85517955	G	T	<i>Rnf150</i>	intron_variant	c.732+3459G>T	No	0	43	16	68
chr8	122865623	G	A	<i>Galnt2</i>	intron_variant	c.21+430784G>A	No	0	57	13	55
chr9	119815293	A	G	<i>Wdr48</i>	intron_variant	c.481+678A>G	No	0	25	13	43
chr10	33305020	T	del	<i>Clvs2</i>	intron_variant	c.236+10530delA	No	0	24	11	24
chr10	61961300	T	C	<i>Srgn</i>	intron_variant	c.74-629A>G	No	0	169	92	185
chr11	67887869	T	C	<i>Sbx8</i>	intron_variant	c.386-34910T>C	No	0	50	21	50
chr14	43563174	C	A	<i>Gm3676</i>	intron_variant	c.491-50G>T	No	0	69	33	59
chr14	52075131	A	G	<i>Vmn2r89</i>	intron_variant	c.207-228A>G	No	0	41	27	47
chr15	81208662	T	C	<i>Sl13</i>	intron_variant	c.269-232A>G	No	0	22	10	32
chr15	98639564	G	T	<i>B130046B21Rik</i>	intron_variant	n.193+1909G>T	No	0	40	14	45
chr16	20735595	C	G	<i>Chrd</i>	intron_variant	c.1092+23C>G	No	0	61	9	46
chr16	37288253	A	del	<i>Stxbp5l</i>	intron_variant	c.666+120delT	No	0	28	11	28
chr16	90248725	G	A	<i>Srsf15</i>	intron_variant	c.1681+33C>T	No	0	46	26	61
chr17	25385402	C	T	<i>Balap3</i>	intron_variant	c.1155+15G>A	No	0	21	8	30
chr18	63544895	C	T	<i>Fam38b</i>	intron_variant	c.-2934+965G>A	No	0	35	13	23
chr19	44155623	C	T	<i>Chuk</i>	intron_variant	c.194+813G>A	No	0	72	17	61
chrX	6082010	G	A	<i>Shroom4</i>	intron_variant	c.-36-15901G>A	No	0	53	18	63
Tumor 3											
chr1	8679206	A	T	<i>Sntg1</i>	intron_variant	c.143-7096T>A	No	0	26	9	36
chr1	74519008	G	A	<i>Usp37</i>	intron_variant	c.1026-748C>T	No	0	22	9	15
chr2	76758370	G	T	<i>Tin</i>	intron_variant	c.355+1238C>A	No	0	32	12	26
chr2	77075014	A	del	<i>Sestd1</i>	intron_variant	c.215-2218delT	No	0	30	11	38
chr3	5243764	G	C	<i>Zfx4</i>	missense_variant	c.2049G>C.p.Arg683Ser	Yes	0	33	17	36
chr4	115565831	T	G	<i>6430628N08Rik</i>	intron_variant	n.296+2265T>G	No	0	21	10	28
chr4	47843058	C	T	<i>Irf2</i>	intron_variant	c.-7+17768C>T	No	0	54	25	52
chr8	116993479	G	A	<i>Wwox</i>	intron_variant	c.130-19319G>A	No	0	49	8	61
chr9	3020484	C	A	<i>Gm16869</i>	intron_variant	n.936-2402G>T	No	0	158	8	161
chr9	3022139	A	T	<i>Gm16869</i>	intron_variant	n.936-4057T>A	No	0	102	15	140
chr10	33335118	A	G	<i>Clvs2</i>	intron_variant	c.389+7232T>C	No	0	47	25	55
chr11	77295772	G	A	<i>Ankrd13b</i>	intron_variant	c.115-4492C>T	No	0	127	61	123
chr14	34170775	C	T	<i>Arhgap22</i>	intron_variant	c.122-1529C>T	No	0	77	35	76
chr14	73955585	C	T	<i>Sucla2</i>	intron_variant	c.90+2881C>T	No	0	47	15	47

Figure S1, related to Figure 1. DNMT3A^{R882H} acts in concert with mutant RAS to transform murine HSPCs ex vivo and induce AMLs in vivo.

(A) Scheme of in vitro and in vivo leukemic assays (top). Bottom, diagram of a home-made, MSCV retrovirus-based bicistronic co-expression system that carries either empty control (EV), DNMT3A^{WT} (WT) or DNMT3A^{R882H} (RH) upstream of an internal ribosome entry site (IRES) sequence, followed by cDNA of either GFP, NRAS^{G12D}, IDH1^{R132H}, NPM1c, or Hoxa9. Myc and Flag tags were fused in-frame with the first (gene A) and second (gene B) cDNA, respectively.

(B) Flow cytometry analysis of Myc-tagged DNMT3A (left) and Flag-tagged NRAS^{G12D} or GFP (right) expression to assess cell purity and gene transduction rates of murine HSPCs after viral infection and drug selection. HSPCs were infected with retrovirus encoding a bicistronic vector of GFP alone (EV-GFP), GFP with DNMT3A^{WT} (WT-GFP) or DNMT3A^{R882H} (RH-GFP), NRAS^{G12D} alone (EV-RAS), or NRAS^{G12D} with DNMT3A^{WT} (WT-RAS) or DNMT3A^{R882H} (RH-RAS). MOCK, non-infected HSPCs.

(C) Immunoblot of indicated proteins in murine HSPCs post-transduction of retrovirus as shown on top of the blot. EV, empty; WT, DNMT3A^{WT}; RH, DNMT3A^{R882H}. Tubulin was used as a protein loading control.

(D) In vitro serial replating assays that score colony-forming units (CFU) of murine HSPCs post-transduction of retrovirus carrying indicated coexpressed oncogenes.

(E-F) Representative microscopic image (panel E, top), Wright–Giemsa staining (panel E, bottom) and flow cytometry analysis (panel F) of cells prepared from the 4th replating of HSPC colonies transduced with DNMT3A^{R882H}-IRES-NRAS^{G12D} (RH-RAS). Scale bars in panel E represent 1 mm and 10 μ m in top and bottom, respectively.

(G) Proliferation of RH-RAS-immortalized progenitors derived from CFU assays in liquid culture medium with indicated cytokines. SCF, stem cell factor; Flt3L, Flt3 ligand; GM-SCF, granulocyte-macrophage colony-stimulating factor; IL-6, interleukin 6.

(H) Flow cytometry analysis of indicated markers using spleen cell populations isolated from recipient mice 3 months post-engraftment of HSPCs transduced with empty vector control (upper; EV-GFP) or NRAS^{G12D} alone (EV-RAS, bottom).

(I) Size of liver in indicated recipient cohorts (n = 4-9) 4 weeks post-engraftment of retrovirally transduced HSPCs (refer to also panel A for the retrovirus-encoded genes).

(J) Hematoxylin and eosin (H&E) staining of liver (left) and bone marrow (middle and right) sections of recipient cohorts 3-4 weeks post-engraftment of HSPCs transduced with EV-RAS (upper) or RH-RAS (bottom). Scale bars in the left, middle and right represent 100, 100 and 10 μ m, respectively.

(K-O) Summary of various examined peripheral blood counting parameters among indicated recipient cohorts (n = 6-13) 4 weeks post-engraftment of retrovirally transduced HSPCs (refer to also panel A for identity of the retrovirus-encoded genes).

(P) Flow cytometry analysis of bone marrow cells from mice 4 weeks post-engraftment of HSPCs transduced with RH-RAS, in comparison to those transduced with EV-RAS.

(Q) Flow cytometry analysis of bone marrow cells from murine leukemias induced by RH-RAS. Non-specific IgG (no ab; black) was used as antibody control.

(R) Summary of somatic nucleotide polymorphisms, insertion and deletions (SNP/Indels) from 3 independent RH-RAS-induced murine AMLs as determined by whole exome capture sequencing. Ref, reference nucleotide; Alt, alternate nucleotide; Snp_Eff_Effect, effect of this variant; Snp_Eff_HGVS, HGVS notation of this variant; AA_Change, effect of this variant on amino acid change; Normal_Alt, alternate counts in normal; Normal_Total, total counts in normal; Tumor_Alt, alternate counts in tumor; Tumor_Total, total counts in tumor; del, deletion.

Error bar, +/- SD; *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.0001.

Table S1, related to Figure 1. Summary of phenotypic analysis of murine AMLs induced by coexpression of DNMT3A^{R882H} and NRAS^{G12D} (RH-RAS).

Mouse ID	WBC* (k/ μ l)	Hematocrit (%)	Spleen (mg)	Lymph node (mg)	Liver (mg)	FACS: bone marrow (positive %)**					FACS: Spleen (positive %)				
						Mac1 ⁺	c-Kit ⁺	Cd19 ⁺	Cd3e ⁺	CD34 ⁺	Mac1 ⁺	c-Kit ⁺	Cd19 ⁺	Cd3e ⁺	CD34 ⁺
1	35.44	20.2	358	44	1436	58.5	16.1	3.6	0.9	18.4	70.9	10.4	N/D***	N/D	N/D
2	58.8	24.9	521	54	1890	52.1	10.8	4.7	0.9	13.0	58.9	10.9	2.5	1.1	41.5
3	75.82	26	478	75	2250	69.7	20.9	3.4	1.0	18.9	78.7	19.5	2.2	1.2	40.7
4	59.36	23.4	561	97	1567	78.5	21.5	8.0	1.6	20.9	61.7	15.7	2.5	2.5	31.5
5	56.72	16.5	382	77	1546	51.9	19.5	5.0	1.1	19.2	70.9	14.4	4.1	1.8	39.2
6	76.2	23.8	432	47	1742	55.4	12.4	5.4	1.2	16.8	74.6	13.2	3.3	1.6	38.4

* , WBC (1,000/ μ l), white blood cell count in one microliter of peripheral circulating blood

** , FACS analysis of leukemic mice showing the percentage of cells scored as positive signals using antigen-specific antibodies

*** , N/D, not determined

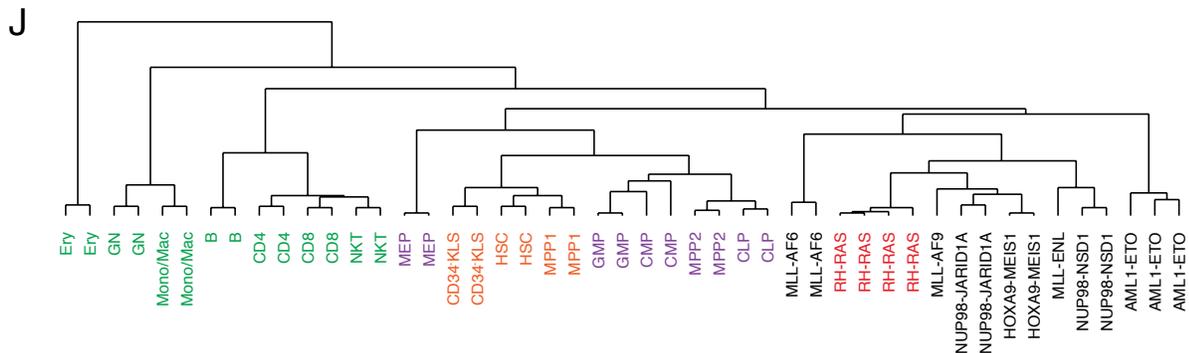
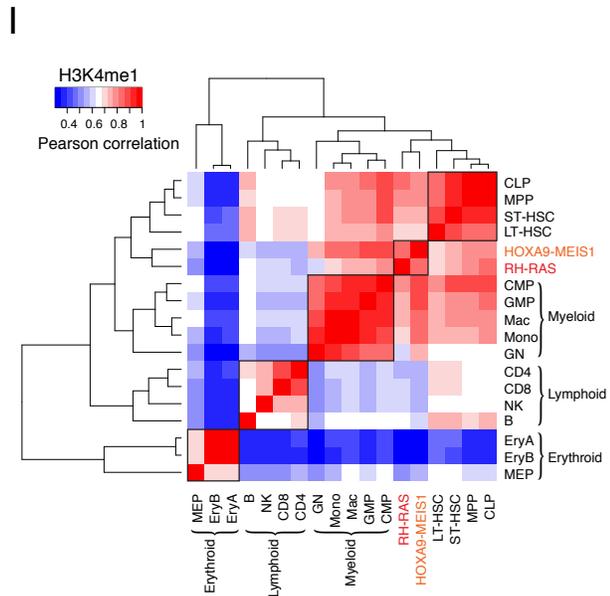
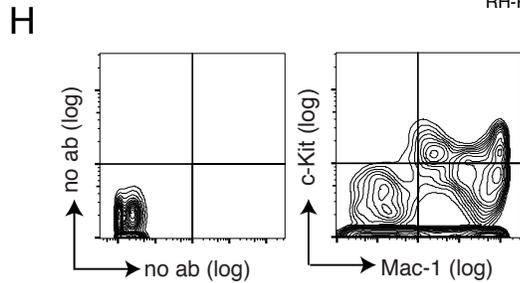
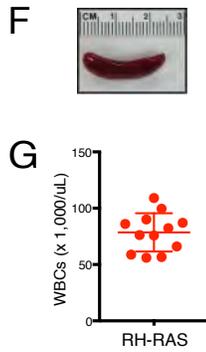
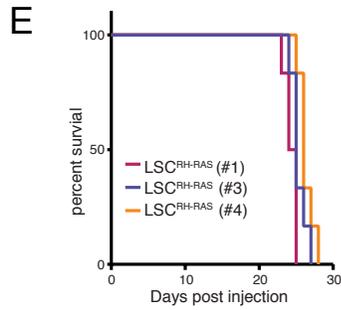
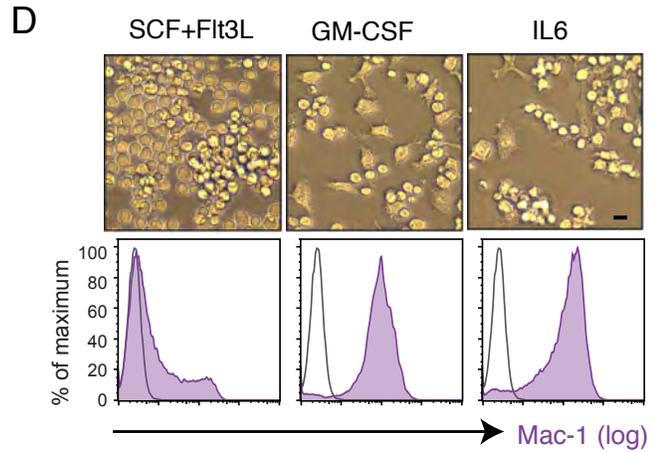
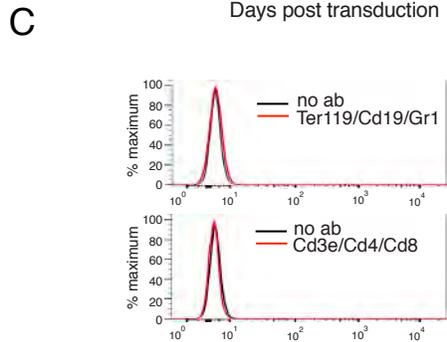
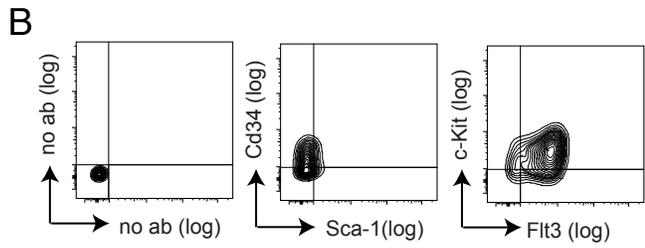
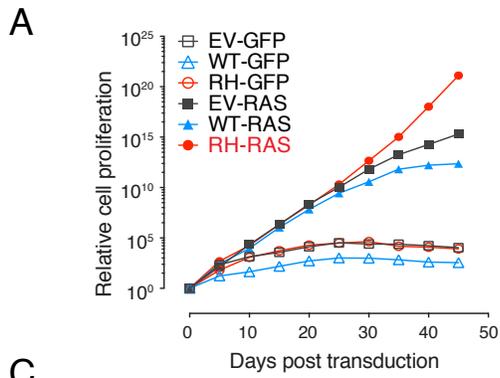


Figure S2, related to Figure 2. R882-mutated DNMT3A establishes leukemia-initiating stem cells (LSCs) ex vivo in the presence of activated RAS.

(A) Proliferation kinetics of murine HSPCs post-transduction of indicated genes.

(B-C) FACS analysis of indicated antigens of RH-RAS immortalized progenitors. Non-specific IgG (no ab) was used as antibody control.

(D) Typical microscopic images (upper) and FACS analysis of Mac-1 (bottom; purple) two weeks post-cultivation of RH-RAS immortalized progenitors in liquid medium with indicated cytokines. Non-specific IgG (open) was used as antibody control in FACS. Scale bar, 10 μ m.

(E) Kaplan-Meier survival curve of mice (n = 6 for each progenitor line) engrafted with each of 3 independent RH-RAS immortalized progenitor lines demonstrates a LSC property of these cells.

(F-G) Enlarged spleen (panel F) and increased counts in peripheral white blood cells (WBC, panel G, n = 12; relative to normal WBC counts of \sim 10,000 per μ L) among mice that developed AML post-engraftment of the ex vivo derived LSC^{RH-RAS} lines.

(H) Flow cytometry analysis of murine AMLs produced by LSCs^{RH-RAS}.

(I) Heatmap showing hierarchical clustering of samples based on their similarities of genome-wide H3K4me1 ChIP-Seq profiles. A non-overlapping 5-kb window was used to count histone ChIP-Seq reads across the mouse genome. The colors represent the scaled Pearson correlation coefficients of the resultant vectors for individual cell lines. The used H3K4me1 ChIP-Seq datasets are those we produced in this current study for LSCs^{RH-RAS} (RH-RAS; red font) and for a murine AML line established by leukemic TFs HOXA9 plus MEIS1 (HOXA9-MEIS1; orange) (Wang et al., 2005), as well as the publicly available datasets of various normal blood cell types (black) as defined in a previous study (Lara-Astiaso et al., 2014). LT-HSC, long-term HSC; ST-HSC, short-term HSC; MPP, multipotent progenitor; CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-monocyte progenitor; MEP, megakaryocyte-erythroid progenitor; Mac, macrophage; Mono, monocyte; GN, granulocyte; B, B220⁺/CD19⁺ B-cell; CD4/8, CD4/8⁺ T-cell; NK, natural killer cell; EryA and EryB, Ter119⁺/CD71⁺ erythroid cell with high and low forward scatter (FSC), respectively.

(J) Hierarchical clustering of samples based on similarity of genome-wide gene expression profiling. The used microarray data are those we generated in this current study for LSCs^{RH-RAS} and murine leukemogenic (AML) lines carrying MLL-fusion proteins (such as MLL-AF9 or MLL-ENL), as well as publicly available datasets of either normal blood cell types (Bock et al., 2012; Ji et al., 2010; Konuma et al., 2011) or murine leukemogenic (AML) lines established by TF mutations, such as HOXA9 plus MEIS1 (Wang et al., 2005) and AML-ETO (Lo et al., 2012), or by a deregulated chromatin regulator, such as MLL-AF6 (Deshpande et al., 2013), NUP98-JARID1A/KDM5A (Wang et al., 2009) or NUP98-NSD1 (Wang et al., 2007). Besides what is described in panel I, CD34⁻KLS represents Cd34⁻/cKit⁺/Lineage⁻/Scal⁺ HSC; MPP1, Flk2⁻ multipotent progenitor; MPP2, Flk2⁺ multipotent progenitor; NKT, natural killer T-cell; Ery, erythroid cell.

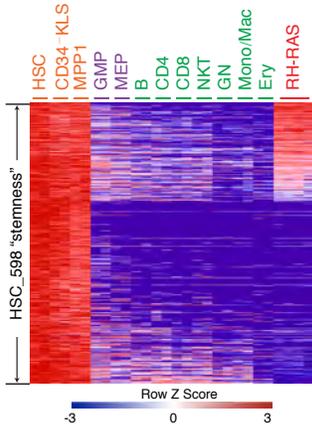
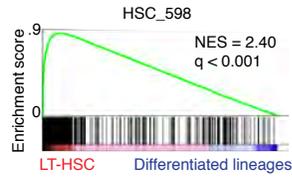
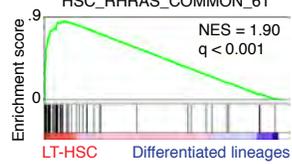
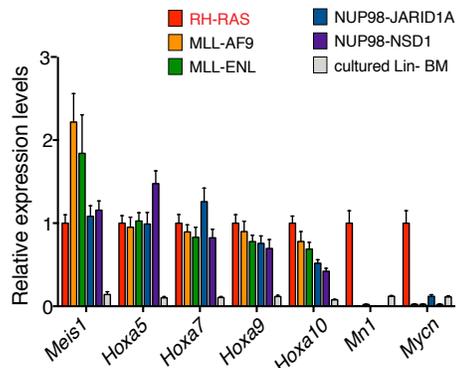
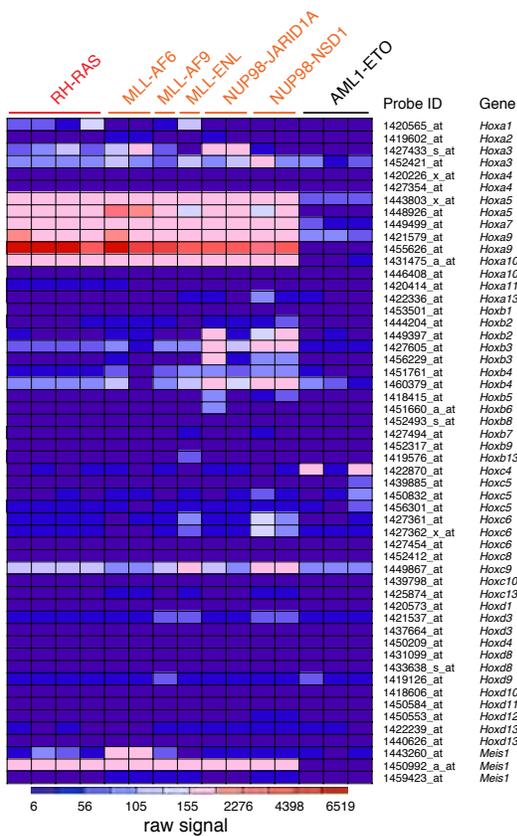
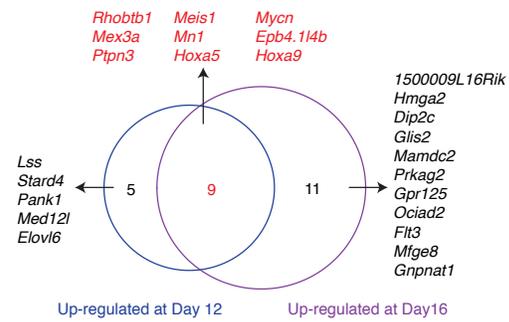
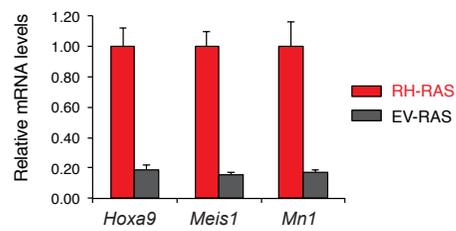
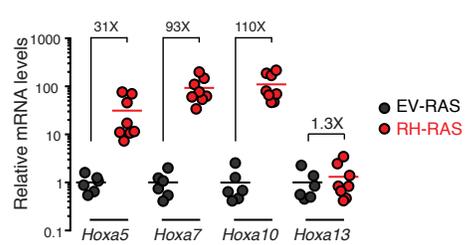
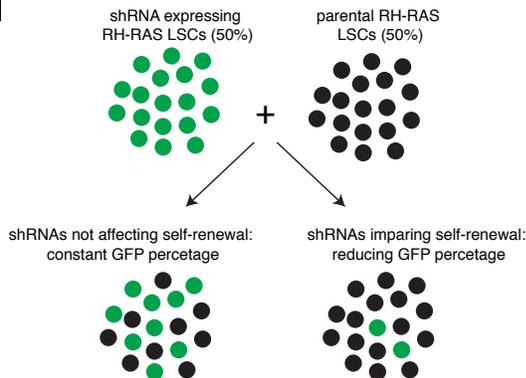
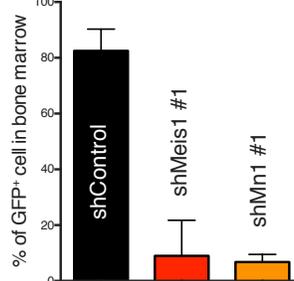
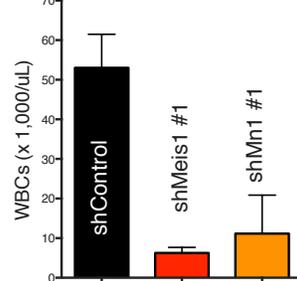
A**B****C****E****D****F****G****H****I****J****K**

Figure S3, related to Figure 3. DNMT3A^{R882H} potentiates activation of ‘stemness’ genes including a critical *Meis1-Mn1-Hoxa* regulatory node

(A) Heatmap showing relative expression of 598 gene probes (464 genes) uniquely expressed among primitive self-renewing HSPCs (orange font; i.e., HSC, Cd34-/KLS and MPP; also refer to Figures S2I-J for details of cell identity and data source) relative to differentiating (purple) and mature (green) blood cell types (from GMP to Ery), followed by ranking based on higher gene expression in LSCs^{RH-RAS} (red) relative to differentiating and mature blood cell types.

(B-C) GSEA analysis with different datasets from a previous study (Chambers et al., 2007) verifies differential expression of the 598 ‘stemness’ probes identified as specific to primitive self-renewing HSPCs (panel B), as well as the 61 probes identified as unique to both self-renewing HSPC and LSCs^{RH-RAS} (panel C), in LT-HSCs relative to differentiated blood cells.

(D) Heatmap showing expression levels of indicated gene probes using genome-wide gene expression data generated by this current or previously published study for LSCs^{RH-RAS} and various murine leukemogenic (AML) progenitor lines established by MLL-AF6 (Deshpande et al., 2013), MLL-AF9, MLL-ENL, NUP98-JARID1A/KDM5A (Wang et al., 2009), NUP98-NSD1 (Wang et al., 2007), or AML-ETO (Lo et al., 2012). The scale bar shows color-coded values for microarray hybridization signals.

(E) RT-qPCR of *Meis1*, *Hoxa*, *Mn1* and *Mycn* in murine leukemogenic progenitor lines transformed by various mutated epigenetic regulators such as RH-RAS, MLL fusion proteins (MLL-AF9 and MLL-ENL), NUP98-JARID1A/KDM5A (Wang et al., 2009) or NUP98-NSD1 (Wang et al., 2007). Y-axis represents mean \pm SD of relative gene expression from at least 2-4 independent lines after normalization. In vitro cultured lineage-negative (Lin⁻; cultured for > 2 weeks) bone marrow (BM) cells serve as control.

(F) Venn diagram shows part of the identified 54 LSC^{RH-RAS} ‘stemness’ genes consistently upregulated among murine HSPCs at both day 12 and 16 post-transduction of RH-RAS, relative to EV-RAS (also refer to Figure 3B).

(G) RT-qPCR analysis of *Hoxa9*, *Meis1* and *Mn1* expression with samples from the 2nd serial replating of murine HSPCs transduced with EV-RAS or RH-RAS.

(H) RT-qPCR of *Hoxa* expression in murine leukemias induced by RH-RAS (n = 8) or EV-RAS (n = 6).

(I) Scheme of cell competition assays used to assess effects of shRNA-mediated gene knockdown on cell growth. GFP-positive (shRNA-expressing) cells were mixed at a 1:1 ratio with un-transduced (GFP-negative) control cells, followed by in vitro cultivation and measurement of percentage of GFP-positive cells every 2 days.

(J-K) Percentage of GFP-positive cells in bone marrow (panel J) and peripheral WBC counts (panel K) in recipient mice 3 weeks post-transplantation of LSCs^{RH-RAS} carrying the stably expressed vector control or shRNA specific to either *Meis1* or *Mn1*.

Error bar, +/- SD.

Table S2, related to Figure 3 (provided as a separate Excel file). Microarray analysis for genes showing a unique high expression among primitive self-renewing HSPCs in comparison to non-self-renewing differentiating and mature blood cell subtypes, as well as genes showing a unique high expression among both normal self-renewing HSPCs and LSCs^{RH-RAS}.

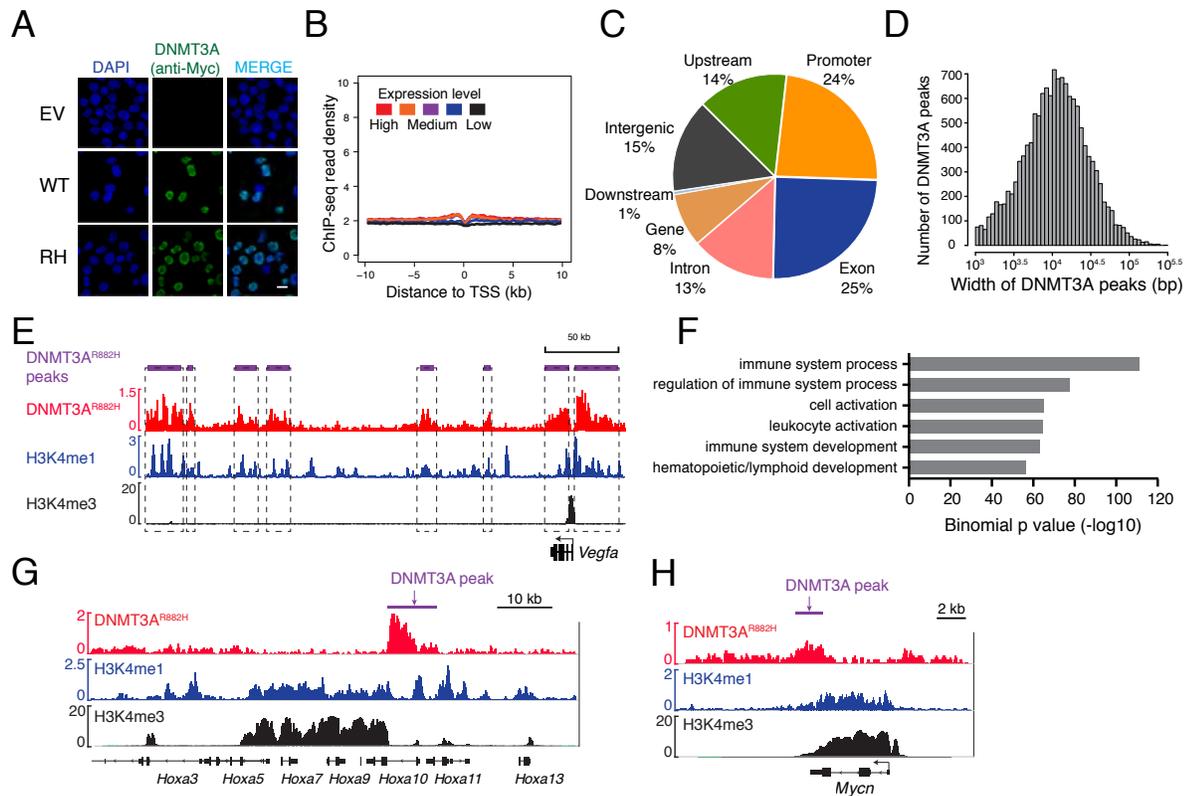


Figure S4, related to Figure 4. ChIP-Seq reveals chromatin context-dependent binding of R882-mutated DNMT3A to genomic regions, including ‘stemness’ genes such as a *Meis1-Mn1-Hoxa* node.

(A) Anti-Myc immunofluorescence of Myc-tagged DNMT3A^{R882H} and DAPI staining after transduction of either empty control (EV), DNMT3A^{WT} (WT) or DNMT3A^{R882H} (RH). Scale bar, 10 μ m.

(B) Average signal densities from anti-Myc ChIP-Seq experiments using EV-RAS-transduced HSPCs, which do not express Myc-DNMT3A^{R882H}, show no enrichment across genes at all expression levels.

(C) Pie-chart showing distribution of the identified 13,705 DNMT3A^{R882H} ChIP-Seq peaks among indicated genomic regions in LSCs^{RH-RAS}.

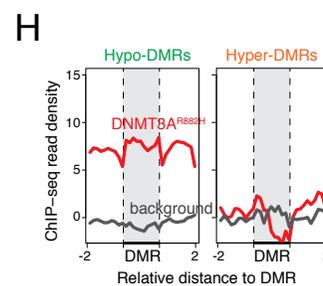
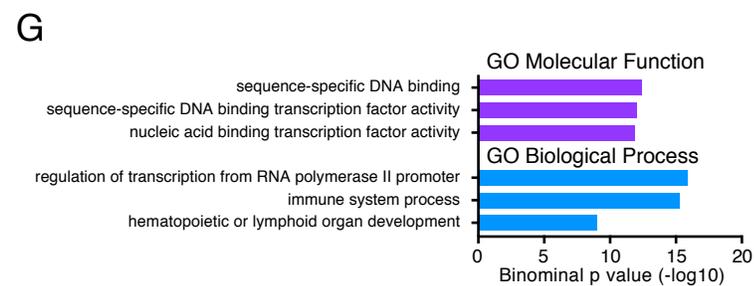
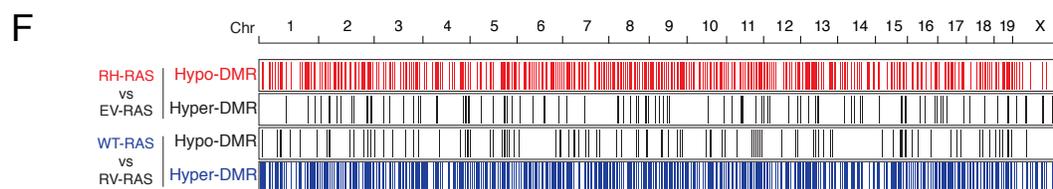
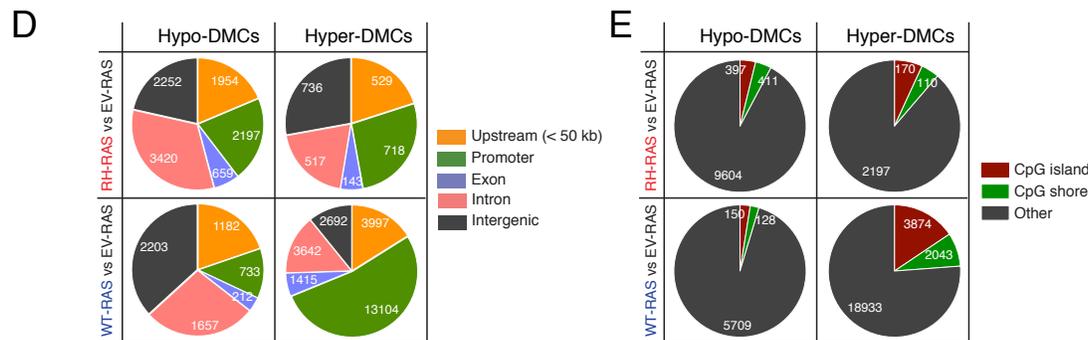
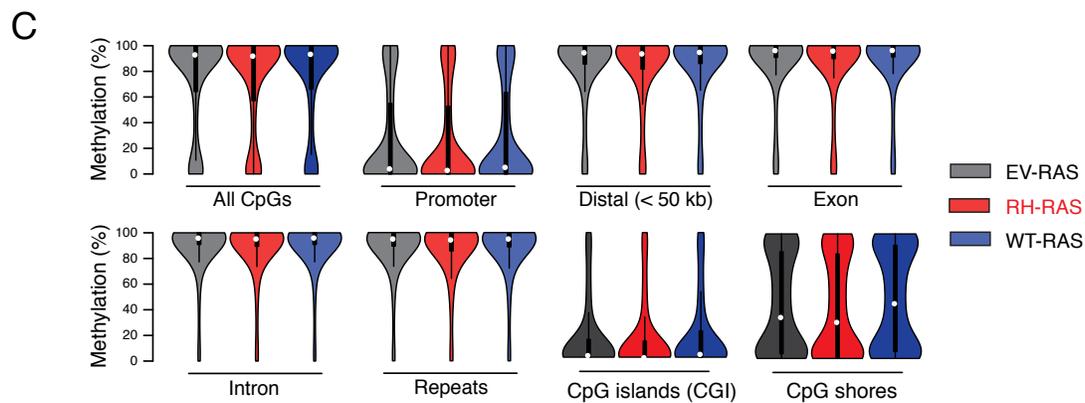
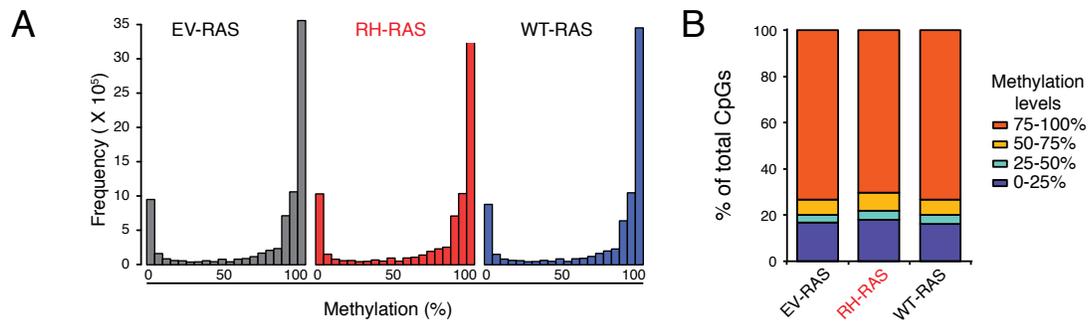
(D) Distribution plot of DNMT3A^{R882H} ChIP-Seq peaks identified in LSCs^{RH-RAS} based on the peak width (x-axis, log scale) shows a generally broad binding pattern on chromatin, with a mean size of ~17 kb.

(E) ChIP-Seq peaks of DNMT3A^{R882H} (top), H3K4me1 (middle) and H3K4me3 (bottom) at an intragenic region close to *Vegfa* in LSCs^{RH-RAS}. Boxed areas highlight close proximity and significant overlap of DNMT3A^{R882H} binding and H3K4me1 peaks.

(F) Functional annotation of DNMT3A^{R882H} ChIP-Seq peaks by GREAT tools (with a setting of “single nearest gene, 200.0 kb max extension”). Shown are the top over-represented categories belonging to Gene Ontology (GO) biological processes.

(G-H) ChIP-Seq profiles of DNMT3A^{R882H}, H3K4me1 and H3K4me3 at the LSC^{RH-RAS} ‘stemness’ genes *Hoxa* cluster (panel G) and *Mycn* (panel H) in LSCs^{RH-RAS}. Shown in y-axis is ChIP-Seq read coverage normalized to a read depth of 1 million reads. Shown on top (purple) is position of the called DNMT3A^{R882H} ChIP-Seq peaks.

Table S3, related to Figure 4 (provided as a separate Excel file). Summary of DNMT3A^{R882H} ChIP-Seq peaks identified in murine LSCs^{RH-RAS}.



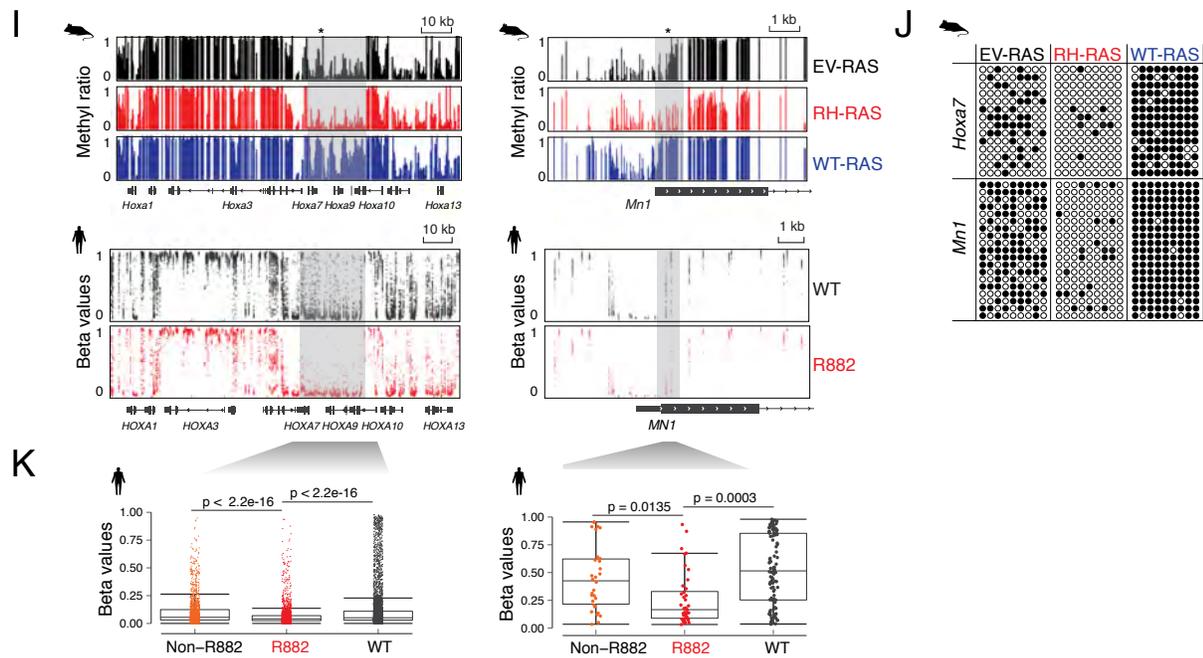


Figure S5, related to Figure 5. DNMT3A^{R882H} induces focal CpG hypomethylations enriched at H3K4me1-demarcated gene-regulatory sites in murine HSPCs.

(A-B) Distribution (panel A) and bar plots (panel B) of absolute methylation levels for a total of ~6.5 million CpG sites detected by eRRBS with >10× coverage among indicated HSPC samples.

(C) Violin plots showing distribution of absolute methylation levels at CpG sites within indicated genomic features among EV-RAS, RH-RAS or WT-RAS HSPCs. White dots are the median and box lines are the first and third quartile of the data. Repeats, repetitive sequence.

(D-E) Pie charts showing distribution of indicated DMCs among various genomic regions.

(F) Distribution of indicated DMRs among chromosomes. Each vertical line represents a DMR.

(G) GREAT annotation of DNMT3A^{R882H}-associated hypo-DMRs. Shown are the top over-represented categories belonging to GO Molecular Function and GO Biological Process.

(H) DNMT3A^{R882H} ChIP-Seq profiles at DNMT3A^{R882H}-associated hypo-DMRs (left), hyper-DMRs (right) and their background controls (grey).

(I) DNA methylation profiles of *Hoxa* (left) and *Mn1* (right) in murine HSPCs (upper panel), as well as those of the human gene homologue (bottom panel) among AML patient samples carrying either WT or R882-mutated *DNMT3A* based on the TCGA dataset (Cancer Genome Atlas Research, 2013; Russler-Germain et al., 2014). Faded points represent individual CpG methylation beta values (WT, n = 50; R882, n = 20). Grey box, DMR detected by both mouse and human studies.

(J) Individual bisulfite sequencing verifies two hypo-DMRs at LSC^{RH-RAS} ‘stemness’ genes *Hoxa7* and *Mn1* (as indicated with asterisks in panel I) in murine HSPCs after indicated gene transduction.

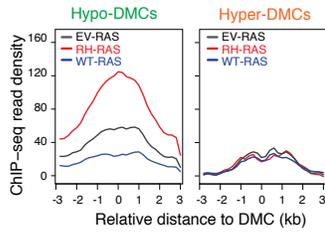
(K) Box plots of methylation beta values of pooled CpGs from grey marked regions at *HOXA7-HOXA10* and *MN1* in human AMLs carrying WT (n = 50), R882-mutated (n = 20) or non-R882-mutated (n = 15) *DNMT3A* based on the TCGA studies. Horizontal line, median; box, interquartile range; whiskers extend to 1.5× the interquartile range. The p values were calculated by Mann-Whitney U test.

Table S4, related to Figures 5 and 6 (provided as a separate Excel file). Summary of DNMT3A^{R882H}- and DNMT3A^{WT}-associated DMRs identified by eRRBS studies in murine HSPCs.

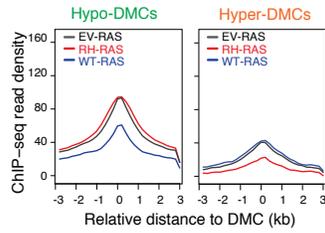
Table S5, related to Figure 5. List of genes having DNMT3A^{R882H}-associated DMRs identified in murine HSPCs, as well as those associated with hypo-DMRs identified among human normal-karyotype (NK) AML patients carrying R882-mutated DNMT3A in comparison to WT DNMT3A.

<i>ABCC4</i>	<i>DLG4</i>	<i>LHX9</i>	<i>RSPO1</i>
<i>ABLIM1</i>	<i>EBF1</i>	<i>LRRC8D</i>	<i>RTKN</i>
<i>ACOXL</i>	<i>EDNRB</i>	<i>MAN1C1</i>	<i>RUSC1</i>
<i>ADAMTSL5</i>	<i>EFHD1</i>	<i>MARVELD2</i>	<i>SBNO2</i>
<i>ADARB1</i>	<i>ELK3</i>	<i>MBNL1</i>	<i>SCARF2</i>
<i>ADCY4</i>	<i>F3</i>	<i>MEGF11</i>	<i>SCRT2</i>
<i>AHDC1</i>	<i>FAM171A2</i>	<i>MEIS1</i>	<i>SCUBE1</i>
<i>ALDH7A1</i>	<i>FOXK1</i>	<i>MN1</i>	<i>SEMA4B</i>
<i>AMIGO3</i>	<i>FSCN1</i>	<i>MSRB3</i>	<i>Sept9</i>
<i>APC2</i>	<i>FZD5</i>	<i>NAV2</i>	<i>SIK3</i>
<i>APLP1</i>	<i>GABBR1</i>	<i>NCKAP5L</i>	<i>SIPA1L3</i>
<i>ATF6B</i>	<i>GATA4</i>	<i>NCOR2</i>	<i>SIX3</i>
<i>ATP9A</i>	<i>GBX2</i>	<i>NFATC1</i>	<i>SLC11A2</i>
<i>AXL</i>	<i>GLTP</i>	<i>NFIA</i>	<i>SLC23A2</i>
<i>B4GALNT1</i>	<i>HLCS</i>	<i>NGB</i>	<i>SLC38A1</i>
<i>BAHCC1</i>	<i>HLF</i>	<i>NKX2-5</i>	<i>SMAD3</i>
<i>CALCA</i>	<i>HOXA7</i>	<i>NXPH4</i>	<i>SYNE2</i>
<i>CHN2</i>	<i>HOXB4</i>	<i>ONECUT1</i>	<i>TBX5</i>
<i>CIT</i>	<i>HOXD13</i>	<i>PANX2</i>	<i>TFAP2A</i>
<i>CLCF1</i>	<i>HS3ST2</i>	<i>PAX3</i>	<i>TIAM1</i>
<i>CLDN5</i>	<i>HSF4</i>	<i>PBX1</i>	<i>TNKS1BP1</i>
<i>CTBP2</i>	<i>IRX4</i>	<i>PCDHGA11</i>	<i>TRAF1</i>
<i>CXCL12</i>	<i>ITGB2</i>	<i>PCDHGA2</i>	<i>TRIM14</i>
<i>CYP1B1</i>	<i>ITPK1</i>	<i>PMEPA1</i>	<i>TRPS1</i>
<i>CYP26C1</i>	<i>KCNG3</i>	<i>POMC</i>	<i>TSPAN4</i>
<i>DAB2IP</i>	<i>KCNIP2</i>	<i>PRRT1</i>	<i>TTC7B</i>
<i>DHX35</i>	<i>KDM2B</i>	<i>PTPRS</i>	<i>USP2</i>
<i>DLC1</i>	<i>KIFC2</i>	<i>RAI1</i>	<i>WWTR1</i>
<i>DLEU2</i>	<i>LASP1</i>	<i>RILP</i>	<i>ZMIZ1</i>
<i>DLEU7</i>	<i>LEF1</i>	<i>RIN3</i>	

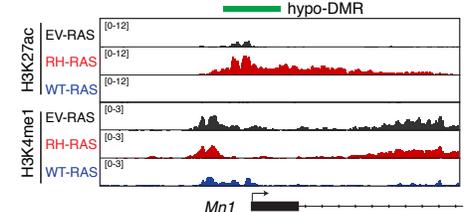
A



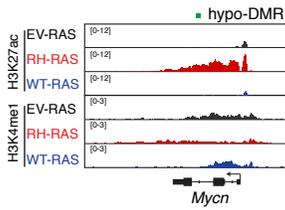
B



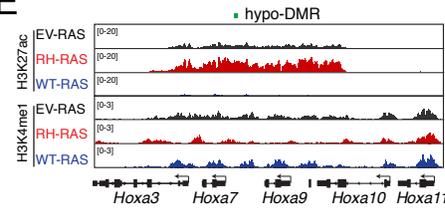
C



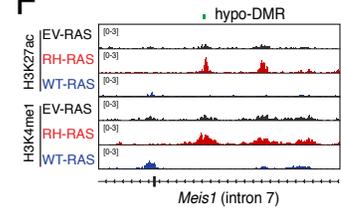
D



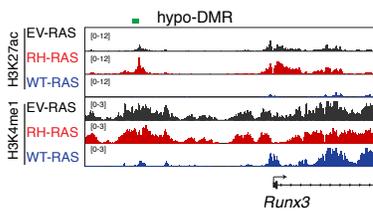
E



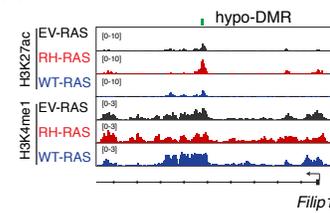
F



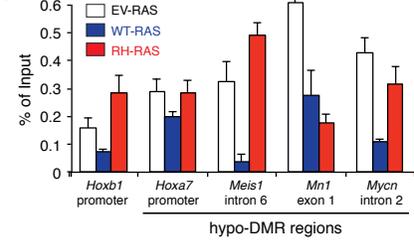
G



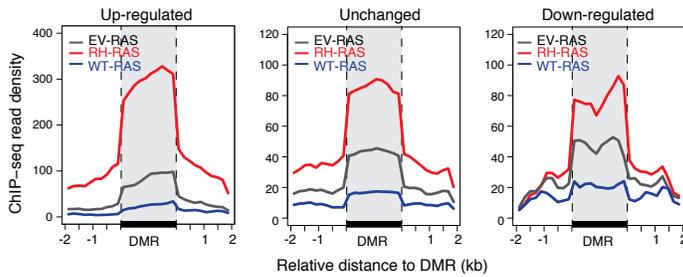
H



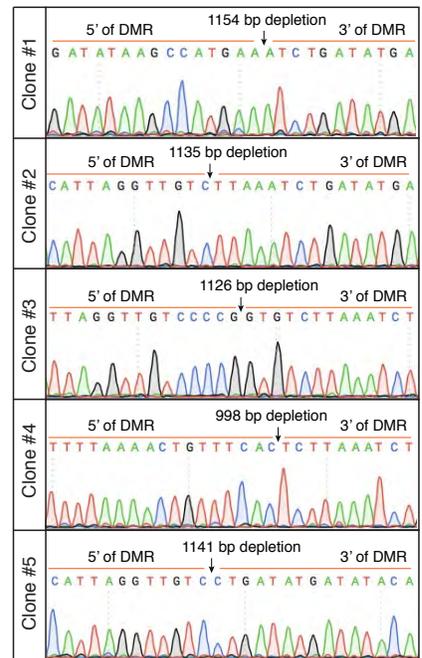
I



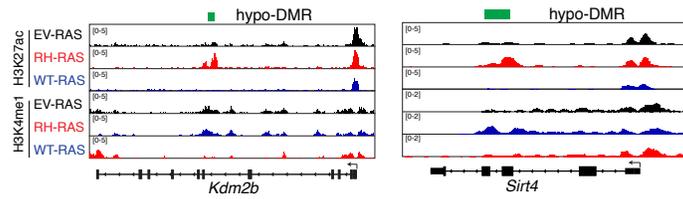
J



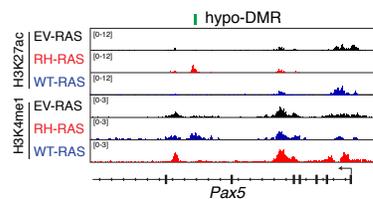
N



K



L



M

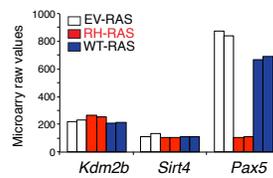


Figure S6, related to Figure 6. DNMT3A^{R882H}-associated hypo-DMRs gain epigenetic alterations associated with gene activation.

(A-B) Enrichment of H3K27ac (panel A) and H3K4me1 (panel B) ChIP-Seq signals at DNMT3A^{R882H}-associated DMCs (with DMCs positioned at the center of x-axis). Plotted were average histone ChIP-Seq read densities at 100-bp bins within +/-3 kb of DMCs in murine HSPCs co-transduced with DNMT3A^{R882H} (RH-RAS, red) in comparison to mock transduction (EV-RAS, grey) or co-transduced with DNMT3A^{WT} (WT-RAS, blue).

(C-H) Histone ChIP-Seq profiles of indicated genes among HSPCs post-transduction of EV-RAS, RH-RAS or WT-RAS. Green bars indicate gained hypo-DMRs upon DNMT3A^{R882H} transduction (RH-RAS) relative to control (EV-RAS) as identified by eRRBS.

(I) ChIP-qPCR of H3K4me1 binding at indicated genomic regions among HSPCs post-transduction of EV-RAS, RH-RAS or WT-RAS.

(J) Enrichment of H3K27ac ChIP-Seq signals at DNMT3A^{R882H}-associated hypo-DMRs (shown in bold on x-axis) associated with three gene subgroups showing either up-regulation (n = 144), down-regulation (n = 89) or no significant changes (n = 821) in their expression levels in murine HSPCs 16 days post-transduction of RH-RAS, relative to EV-RAS (with significant gene expression change defined by fold-of-change > 1.3 and p value < 0.05). Plotted were averaged H3K27ac ChIP-Seq read densities across hypo-DMRs in murine HSPCs post-transduction of EV-RAS, RH-RAS or WT-RAS.

(K-L) H3K4me1 and H3K27ac ChIP-Seq profiles at *Kdm2b* and *Sirt4* (panel K), as well as *Pax5* (panel L) in murine HSPCs post-transduction of EV-RAS, RH-RAS or WT-RAS. Green bars indicate gained hypo-DMRs upon DNMT3A^{R882H} transduction (RH-RAS), relative to control (EV-RAS), as identified by eRRBS.

(M) Microarray analysis shows no expression up-regulation of *Kdm2b*, *Sirt4* and *Pax5*, the genes with significant gain of H3K27ac at their hypo-DMR as shown in the panels K and L.

(N) Sequencing results verified CRISPR/Cas9-induced genomic deletion of a hypo-DMR found at the *Meis1* intron 6 in LSC^{RH-RAS} lines. The genomic PCR products were generated with the two DMR-flanking primers (i.e. F2 and R2 as shown in Figures 6L-M) followed by direct sequencing.

Error bar, +/- SD.

Table S6, related to Figures 6 and 7. List of DNMT3A^{R882H} signature genes (defined as genes carrying DNMT3A^{R882H}-induced epigenetic changes [CpG hypomethylation and H3K27ac gain] and up-regulated by DNMT3A^{R882H} in HSPCs with RH-RAS, relative to EV-RAS).

<i>2610307P16Rik</i>	<i>Hoxa7</i>	<i>Prrt2</i>
<i>Alox5ap</i>	<i>Id2</i>	<i>Ptpre</i>
<i>Anxa2</i>	<i>Itgae</i>	<i>Pxn</i>
<i>Bace1</i>	<i>Itpk1</i>	<i>Rab27a</i>
<i>Baiap2</i>	<i>Itsn1</i>	<i>Rab3ip</i>
<i>Bcl2</i>	<i>Klhl2</i>	<i>Rin3</i>
<i>Bsn</i>	<i>Limk1</i>	<i>Rora</i>
<i>Chad</i>	<i>Lpar2</i>	<i>Runx3</i>
<i>Chst12</i>	<i>Meis1</i>	<i>S1pr3</i>
<i>Chst14</i>	<i>Mn1</i>	<i>Selplg</i>
<i>Cpne2</i>	<i>Mrvi1</i>	<i>Sh2d5</i>
<i>Dnajc6</i>	<i>Mta3</i>	<i>Slc9a9</i>
<i>Dpp4</i>	<i>Mycn</i>	<i>Snx18</i>
<i>Emid1</i>	<i>Nfix</i>	<i>Tarm1</i>
<i>Ephb2</i>	<i>Nrip1</i>	<i>Tifab</i>
<i>Erap1</i>	<i>Pan3</i>	<i>Trerf1</i>
<i>Filip1</i>	<i>Pde4a</i>	<i>Tubb6</i>
<i>Gne</i>	<i>Plec</i>	<i>Uck2</i>
<i>Golm1</i>	<i>Plxnc1</i>	<i>Zfp217</i>

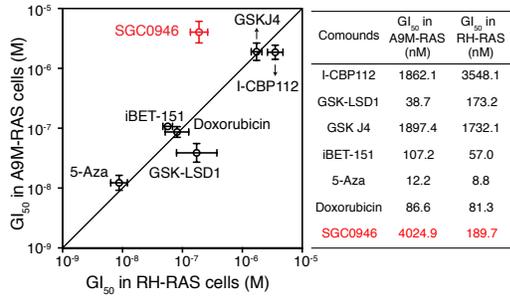
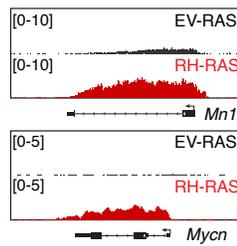
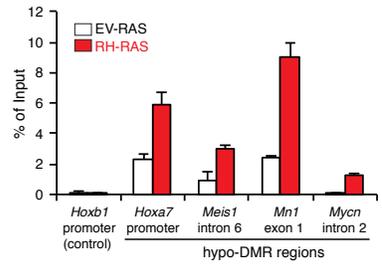
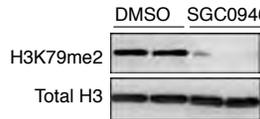
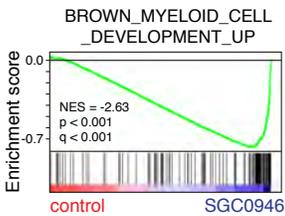
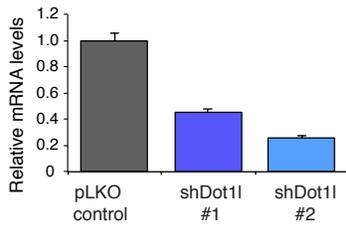
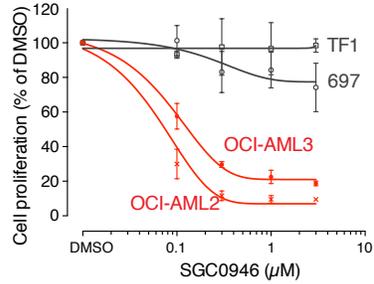
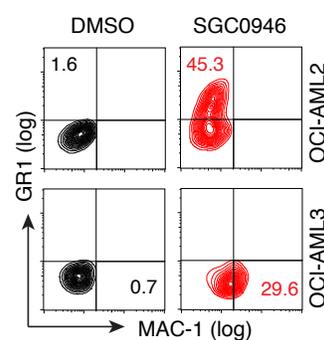
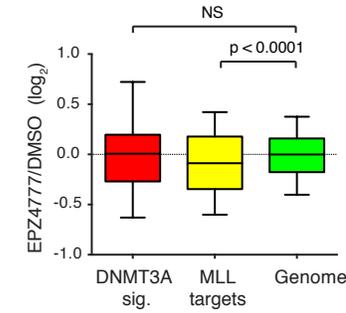
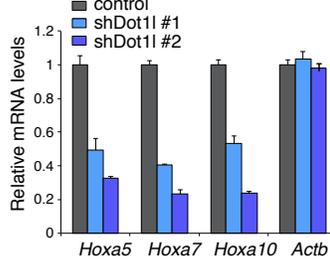
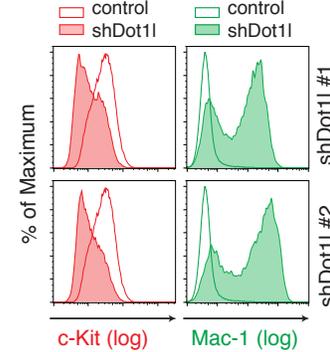
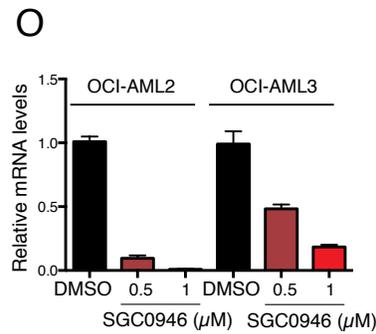
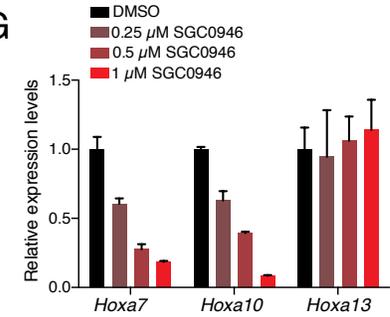
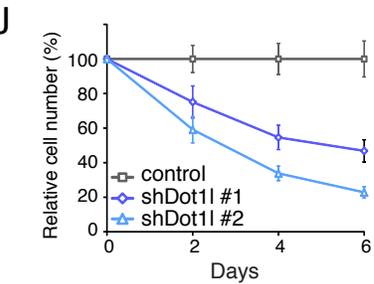
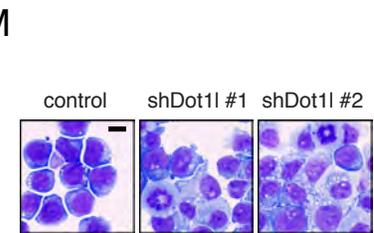
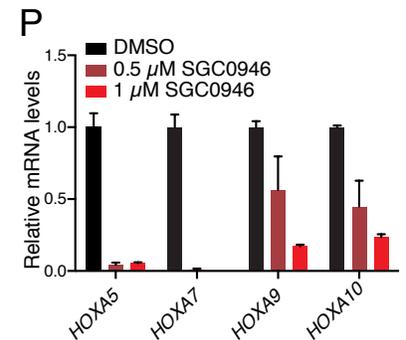
A**B****C****D****E****H****K****N****F****I****L****O****G****J****M****P**

Figure S7, related to Figure 7. Dot1l inhibition reverses DNMT3A^{R882H}-mediated aberrant trans-activation of stem cell genes, thereby suppressing acute leukemogenicity.

(A) Plots (left) and summary table (right) of the half maximal inhibition of cell proliferation (GI_{50}) of indicated compounds in suppressing in vitro growth of LSCs^{RH-RAS} and a control line without DNMT3A^{R882H}, i.e., LSCs established by coexpression of NRAS^{G12D} and Hoxa9 plus Meis1. Data were collected 12 days post-treatment. Compounds include bromodomain inhibitor iBET-151, a DNA demethylating agent 5-Aza, topoisomerase inhibitor Doxorubicin, LSD1 inhibitor GSK-LSD1, CBP/p300 inhibitor I-CBP112, Dot1l inhibitor SGC0946, and the jumonji H3K27 demethylase inhibitor GSK-J4.

(B) H3K79me2 ChIP-Seq profiles at indicated LSC^{RH-RAS} 'stemness' genes in murine HSPCs transduced with EV-RAS or RH-RAS.

(C) ChIP-qPCR of H3K79me2 levels at indicated genomic sites in murine HSPCs post-transduction of EV-RAS or RH-RAS.

(D) Immunoblot of H3K79me2 and total histone H3 in LSCs^{RH-RAS} 4 days post-treatment with DMSO or 1 μ M SGC0946.

(E) GSEA analysis reveals a significant enrichment of myeloid differentiation gene programs in LSCs^{RH-RAS} 4 days post-treatment with 1 μ M SGC0946, relative to DMSO.

(F) Boxplots showing relative expression of DNMT3A^{R882H} signature genes (n = 54), MLL-AF9 target genes (n = 129), and all genes in the genome among MLL-AF9-transformed leukemic lines post-treatment with the Dot1l inhibitor EPZ4777 (3 μ M) for 6 days, relative to DMSO. Expression data were obtained from GEO datasets GSE61013 (Chen et al., 2015). Horizontal line, median; box, interquartile range; whiskers, 10 to 90 percentiles. The p values were calculated by Mann-Whitney U test.

(G) RT-qPCR shows a dose-dependent suppression of *Hoxa7* and *Hoxa10*, but not *Hoxa13* (an already repressed *Hoxa* gene), in LSCs^{RH-RAS} 6 days post-treatment with indicated concentrations of SGC0946, relative to DMSO.

(H-I) RT-qPCR detects relative expression of *Dot1l* (panel H) and *Hoxa* genes (panel I) in LSC^{RH-RAS} lines stably transduced with either of the two tested *Dot1l*-specific shRNAs, in comparison to pLKO vector control. *Actb* serves as a control for gene expression.

(J) Relative in vitro growth of LSC^{RH-RAS} lines after stable transduction with either of the two tested *Dot1l*-specific shRNAs in comparison to pLKO vector control.

(K) Relative growth of two DNMT3A-mutated human AML cell lines, OCI-AML2 and OCI-AML3, and two human leukemia lines with no DNMT3A mutation, TF-1 (carrying an activating NRAS mutation) and 697 (a pre-B ALL line with translocation of a HOX/MEIS1 cofactor gene, *E2A-PBX1*) 12 days post-treatment with various concentrations of SGC0946 in comparison to DMSO.

(L-M) FACS analysis for c-Kit and Mac-1 (panel L) and Wright-Giemsa staining (panel M; scale bar, 5 μ m) of LSC^{RH-RAS} lines after stable transduction with *Dot1l*-specific shRNAs, relative to pLKO vector control.

(N) FACS analysis for GR1 and MAC-1 expression in the OCI-AML2 and OCI-AML3 human AML cell lines 6 days post-treatment with 1 μ M of SGC0946 in comparison to DMSO.

(O) RT-qPCR of *MEIS1* in OCI-AML2 or OCI-AML3 human AML cells 6 days post-treatment with indicated concentrations of SGC0946, relative to DMSO.

(P) RT-qPCR of *HOXA* expression in OCI-AML2 human AML cells 6 days post-treatment with indicated concentrations of SGC0946, relative to DMSO.

Error bar, +/- SD. NS, not significant.

Supplemental Experimental Procedures

Plasmid Construction

MYC-tagged human DNMT3A isoform 1 (also known as DNMT3A1, NCBI accession NP_783328) was generated by PCR-mediated ligation of a MYC tag to N-terminus of DNMT3A1 and then cloned to a MSCV retroviral expression vector (Clontech). For gene co-expression, DNMT3A1 cDNA was inserted to upstream of an internal ribosome entry site (IRES) sequence, followed by either GFP, NRAS^{G12D}, IDH1^{R132H} or NPM1c, in a home-made MSCV-based bicistronic system. DNMT3A1, NRAS^{G12D} and NPM1 plasmids were obtained from Addgene and IDH1^{R132H} was kindly provided by Dr. Yue Xiong (UNC at Chapel Hill). Point mutation was generated by site-directed mutagenesis and all used plasmids were confirmed by sequencing. MSCV expression plasmids for Hoxa9 alone or in combination with co-expressed Meis1 were used as previously described (Wang et al., 2005). All plasmids used were confirmed by sequencing.

Cell lines and Tissue Culture

Cell lines of HEK293, NIH3T3, and TF-1 were obtained from ATCC and maintained using recommended culture conditions. Maintenance and cultivation in vitro of murine leukemogenic progenitor lines established by various leukemic factors such as HOXA9 plus MEIS1 (Wang et al., 2006; Wang et al., 2005), MLL-AF9 (Xu et al., 2015), NUP98-NSD1 (Wang et al., 2007) or NUP98-JARID1A/KDM5A (Wang et al., 2009) were described before. A WEHI3B murine AML line that shows high Hoxb8 and Meis2 coexpression (Fujino et al., 2001) and a human pre-B ALL cell line, 697 (Kamps et al., 1991), are obtained from Dr. Mark Kamps (UCSD). Two *DNMT3A*-mutated human AML lines OCI-AML2 and OCI-AML3 (Tiacci et al., 2012) are kind gifts of Dr. Mark Minden (Princess Margaret Hospital).

Retroviral Production

MSCV-based retrovirus was packaged and produced in HEK293 followed by retroviral titration with NIH3T3 cell infection as previously described (Wang et al., 2006).

Purification, Retroviral Transduction, and Cultivation of Primary Murine Hematopoietic Stem/Progenitor Cells (HSPCs)

Bone marrow cells are harvested from femur and tibia of wild-type balb/C mice pre-treated with 5-Fluorouracil (5-FU), followed by a lineage-negative (Lin⁻) enrichment protocol to remove cells expressing differentiation antigens as we described before (Wang et al., 2006; Wang et al., 2009). Briefly, 500,000 of Lin⁻ enriched hematopoietic stem/progenitor cells (HSPCs) were obtained per mouse with ~10% of them c-Kit⁺Lin⁻Sca1⁺ (KLS) HSPCs. Before retroviral infection, Lin⁻-enriched HSPCs were pre-stimulated in the base medium OptiMEM (Invitrogen, 31985), 15% of FBS (Invitrogen, 16000-044), 1% of antibiotics and 50 μ M of β -mercaptoethanol) complemented with a cytokine cocktail that contains 10 ng/mL each of SCF (Miltenyi), Flt3 ligand (Flt3L; Sigma), IL3 (Peprotech) and IL6 (Peprotech) for three days as described (Wang et al., 2007; Wang et al., 2006; Wang et al., 2009). Post-infection with the concentrated retrovirus (Retro-X concentrator, Clontech) and drug selection (2 μ g/mL puromycin for the first 2-4 days followed by an extended seven-day selection using 1 μ g/mL puromycin for the MSCV-puro vector used in the study), rates of purity and exogenous oncogene expression in those retrovirus-transduced, drug-selected murine HSPCs were assessed by flow cytometry of MYC and/or FLAG-tagged proteins, followed by subsequent functional readout assays such as CFU assays in the semi-solid culture system or the growth assay in the liquid

culture medium. For the latter, the above base medium is added with either recombinant murine SCF alone or together with Flt3L; alternatively, we also routinely use home-made cell culture supernatants of mSCF-producer (mSCF-CHO cells, gift of Dr. Mark Kamps, UCSD) and mFlt3-producer (SP2.0-mFlt3L cells, gift of Dr. Robert Rottapel, University of Toronto) cell lines as source of murine SCF and Flt3L, respectively. To test differentiation potency in the presence of myeloid-promoting cytokines such as GM-SCF or IL6 (PeproTech), progenitor cells were rinsed with PBS and seeded in the same base medium that contains 10 ng/mL of each tested cytokine. Cell splitting and replating with fresh medium were carried out every 3-4 days to keep cell number <1.5 million per mL (in a 6- or 12-well plate). In vitro cultured HSPCs were routinely monitored under microscopy and cellular morphology examined by Wright-Giemsa staining and FACS as described (Wang et al., 2006).

In vivo Leukemogenic Assay

To determine potential leukemia-initiating stem cell (LSC) properties possessed by ex vivo immortalized progenitor lines, 0.1 million of these cells were engrafted to syngeneic mice via tail vein injection (The UNC Cancer Center Animal Studies Core). Mice were monitored with complete blood counting (CBC) of the obtained peripheral blood samples and abdomen palpation for early signs of early leukemia such as lethargy, increased white blood cell (WBC) counts and enlarged spleen (Xu et al., 2015). Mice exhibiting full-blown leukemic phenotypes were euthanized followed by pathological and histological analyses as described (Wang et al., 2007; Wang et al., 2006; Wang et al., 2005; Xu et al., 2015). LSCs established ex vivo were used for primary (1st) transplantation assays, and the produced primary leukemias in independent mice were used for secondary (2nd) transplantation.

Flow Cytometry (FACS) Analysis

Cells were blocked with BD FcBlock (2.4G2) and stained on ice with fluoro-conjugated antibodies (1:100 dilution of c-Kit^{FITC/APC}, Sca-1^{PE-Cy7}, Cd34^{APC/FITC}, Mac-1^{APC}, Gr1^{PE/FITC}, Cd19^{PE/FITC}, Cd3e^{FITC}, Ter119^{FITC}, Cd4^{FITC}, Cd8^{FITC}; purchased from BD Biosciences and eBioscience Inc.) and analyzed on the Beckman Coulter CyAn ADP or Becton Dickinson LSR II machine (UNC Flow Cytometry Core Facility) as described before (Wang et al., 2009; Xu et al., 2015). For intracellular staining of FLAG- or MYC-tagged proteins, cells were first prefixed by the fixation/permeability buffer (BD Biosciences) or 100% methanol, respectively, and stained with Fluor-conjugated, DYKDDDDK or c-Myc Alexa antibodies (1:100 dilution, R&D Systems), respectively. Data were analyzed using FlowJo software.

Microarray Analysis

Total RNA was extracted using RNeasy Plus Mini Kit (Qiagen) followed by quality check with Bioanalyzer; subsequent sample preparation and hybridization to either Mouse Genome 430 2.0 Arrays (Affymetrix) or Mouse Gene 2.0 ST Arrays (Affymetrix) were carried out according to manufacturer's instruction and by UNC Functional Genomics Core as described before (Wang et al., 2007; Xu et al., 2015). Raw CEL data from Mouse Gene 2.0 ST Arrays were processed with GeneSpring software X12.6 (Agilent Technologies, Inc.) as described before (Xu et al., 2015).

Raw Data of Mouse Genome 430 2.0 Array studies – Gene expression data generated by this current study included those of four independently derived RH-RAS LSC lines (termed as LSC^{RH-RAS} #1, #2, #3 and #4), two RH-RAS LSC lines (LSC^{RH-RAS} #1 and #3) treated with 1 μ M SGC0946 for 4 days and one MLL-AF9-transformed progenitor line. The other datasets in the same Mouse Genome 430 2.0 platform for various normal and leukemogenic hematopoietic cell lines are listed in the following list.

Published Expression Data Used In This Study	GSE datasets	Reference
HSC and MEP	GSE38557	(Bock et al., 2012)
MPP1, MPP2, CLP, CMP and GMP	GSE20244	(Ji et al., 2010)
Cd34 KLS, Ery, GN, Mono/Mac, B, CD4, CD8 and NKT	GSE27787	(Konuma et al., 2011)
LT-HSC and various differentiated lineages	GSE6506	(Chambers et al., 2007)
MLL-AF6 AML	GSE43067	(Deshpande et al., 2013)
AML1-ETO AML	GSE15195	(Lo et al., 2012)
NUP98-JARID1A AML	This study	(Wang et al., 2009)
NUP98-NSD1, HOXA9/MEIS1 and MLL-ENL AML	GSE10071	(Wang et al., 2007)

Data Preprocessing, Hierarchical clustering and Principal component analysis (PCA) – Raw CEL files were subjected to a series of normalization and quality control procedures, similar to a previously described method (Bock et al., 2012). First, CEL files from all experiments were jointly normalized through the Robust Multi-array Average (RMA) method (Irizarry et al., 2003) to remove systematic variations. RMA normalization was performed using the ‘affy’ package of the Bioconductor suite (<http://www.bioconductor.org/>). Second, for most cell types, the two replicates with the highest pairwise correlation were selected and used. Third, we performed RMA normalization with these used CEL files to generate final datasets for further analysis. Normalized data were then processed and filtered by floor = 20, ceiling = 20000, min fold change = 3, min delta = 100 for hierarchical clustering and PCA analysis. Hierarchical clustering was performed by average linkage using uncentered Pearson correlation on GenePattern platform (Reich et al., 2006). PCA analysis was performed using the ‘prcomp’ package in R (www.r-project.org).

Identification of ‘stemness gene’ signatures – Base on hierarchical clustering and PCA analysis, we used gene expression data of HSC, MPP1 and CD34 KLS from different sources (see the list above; defined as “self-renewing group”) to minimize potential unwanted batch effects; gene expression datasets of differentiating or mature blood cell types (i.e., GMP, MEP, B, CD4, CD8, NKT, GN, Mono/Mac, and Ery) represent the “non-self-renewing group”. HSC “stemness” genes were defined by two standards: (1) by comparing “self-renewing group” with each category in the “non-self-renewing group”, a gene probe should have at least two-fold overexpression and $p < 0.05$ with standard two sample t-test, and (2) by comparing “self-renewing” vs. “non-self-renewing” groups, a gene probe should meet statistical significance $p < 0.001$ with standard two sample t-test. From these, we identified 598 gene probes as HSC “stemness” genes. Next, those 598 HSC “stemness” gene probes that meet the same criteria by comparing RH-RAS LSCs to the “non-self-renewing group” will then be defined as “LSC^{RH-RAS} ‘stemness’ genes/programs”.

Gene Set Enrichment Analysis (GSEA)

GSEA analysis was performed using GSEA2-2.2.0 software (Subramanian et al., 2005) for testing enrichment of curated gene sets (C2) or customized gene sets as performed before (Xu et al., 2015).

RT-qPCR Analysis

Reverse transcription of RNA was performed using the random hexamer and High Capacity cDNA Reverse Transcription Kits (Applied Biosystems). Usually the PCR amplicon (size ~75-150 bp) is designed to span over large intron regions using NCBI Primer-BLAST. Quantitative PCR was performed in triplicate using SYBR green master mix reagent (Applied Biosystem) on an ABI 7900HT

fast real-time PCR system. The detailed primer sequences for RT-qPCR are provided in the following lists.

Mouse Genes	Forward	Reverse	Reference
<i>Actb</i>	ACCAACTGGGACGACATGGA	GGTCTCAAACATGATCTGGGTCAT	
<i>Hoxa5</i>	GCAAGCTGCACATTAGTCAC	GCATGAGCTATTTGATCCT	(Wang et al., 2009)
<i>Hoxa7</i>	CGGGCTTATACAATGTCAACAG	AAATGGAATTCCTTCTCCAGTTC	(Wang et al., 2009)
<i>Hoxa9</i>	ACAATGCCGAGAATGAGAGC	CAGCGTCTGGTGTGTTTGTGT	
<i>Hoxa10</i>	CTCCAGCCCCTTCAGAAAAC	TGTAAGGGCAGCGTTTCTTC	
<i>Hoxa13</i>	CCAAATGTAAGTCCCAAG	CCTATAGGAGCTGGCGTCTG	
<i>Meis1</i>	AAGGTGATGGCTTGGACAAC	TGTGCCAACTGCTTTTTCTG	
<i>Mn1</i>	TGATGGCAGAACACAGCACT	CTTGTGGTGGGGTGGTCAT	
<i>Mycn</i>	CGGAGAGGATACCTTGAGCG	AGTGGTTACCGCCTTGTGT	
<i>Dot1l</i>	GCGGAACCGTTGGAGGTAAT	TTCACAGTGGCTCCATGTCC	

Human Genes	Forward	Reverse	Reference
<i>GAPDH</i>	AACATCATCCCTGCCTCTACTGG	GTTTTTCTAGACGGCAGGTCAGG	
<i>HOXA5</i>	TCTACCCCTGGATGCGCAAG	AATCCTCCTTCTGCGGGTCA	
<i>HOXA7</i>	TCTGCAGTGACCTCGCCAAA	AGCGTCTGGTAGCGCGTGTA	
<i>HOXA9</i>	AAAAACAACCCAGCGAAGGC	ACCGCTTTTTCCGAGTGGAG	
<i>HOXA10</i>	CCTCCGAGAGCAGCAAAGC	CAGCGCTTCTCCGACCACT	
<i>MEIS1</i>	TGACCGTCCATTACGAAACCT	CCAGTCCAACCGAGCAGTAAG	(Abdel-Wahab et al., 2012)

Antibodies and Immunoblotting

Antibodies used for immunoblotting were α -Flag (Sigma; M2), α -MYC (Sigma, 9E10), α -Hoxa9 (Wang et al., 2007; Wang et al., 2005), α -Meis1 (a gift of Dr. Miguel Torres, Centro Nacional de Investigaciones Cardiovasculares, Spain), α -H3K79me2 (Abcam, ab3594) and α -Tubulin (Sigma). Total protein samples were prepared by cell lysis with either SDS-containing Laemmli sample buffer or RIPA lysis buffer as described (Wang et al., 2007; Wang et al., 2005) followed by brief sonication; extracted samples equivalent to 100,000 cells were loaded to SDS-PAGE gels for immunoblotting analysis.

Chromatin immunoprecipitation (ChIP) Followed by Deep Sequencing (ChIP-Seq)

Chromatin samples used for ChIP-Seq were prepared using a previously described protocol (Goldberg et al., 2010; Lee et al., 2006), followed by antibody enrichment, library generation, and parallel sequencing using an Illumina HiSeq-2500 Sequencer (UNC High-throughput Sequencing Facility). Chromatin samples extracted from ~100 million of cells were used for ChIP-Seq of Myc-tagged DNMT3A with the 9E10 anti-Myc antibody (Sigma); other antibodies used for ChIP-Seq were α -H3K4me1 (Abcam, 8895), α -H3K4me3 (Abcam, 8580), α -H3K27me3 (Upstate, 07-449), α -H3K27ac (Abcam, 4729), and α -H3K79me2 (Abcam, 3594). ChIP-Seq data for various normal blood cell subtypes (Lara-Astiaso et al., 2014) were obtained from GEO database (GSE60103).

ChIP-Seq Data Analysis

Briefly, all sequencing reads were mapped to the mouse genome (mm9) using the BWA alignment software (Li and Durbin, 2009), and unique reads mapped to a single best-matching location with no more than two mismatches were kept for data analysis. Duplicated reads, likely generated by PCRs,

were removed using SAMtools (Xu et al., 2015). Bedgraph signal track files for each ChIP-Seq experiment were generated using MACS2 software (Zhang et al., 2008), followed by displaying in the Integrative Genomics Viewer (IGV, Broad Institute) (Robinson et al., 2011). The MACS2 software was also used for peak identification with data from input as controls. DNMT3A^{R882H} ChIP-Seq peaks were called by a window approach as described previously (Goldberg et al., 2010; Xu et al., 2015). In-house scripts were used to assign peaks to annotated genomic features, defined as “promoter proximal” (± 2 kb of transcription start site, TSS), “promoter distal” or “upstream” (-50kb to -2kb of TSS), introns and exons, using the mouse RefSeq annotation as reference (both coding and non-coding were considered). In the case of broad peaks, we assigned a peak as “gene” if it extends beyond the start and end of an annotated gene. The CpG island annotation from the UCSC browser was used to associate peaks to CpG islands (CGIs), and the 1kb regions flanking CGIs were defined as CpG shores. In all analyses, 1-bp intersection was considered as peak overlapping. In the generation of profiles of ChIP-Seq read densities, a 100-bp window was used to determine ChIP-Seq read counts unless specified otherwise. To account for different peak (or DMR) sizes, peaks (or DMRs) were broken into 10-equal size bins and ChIP-Seq reads in individual bins were summed and normalized to yield reads per 100-bp. All ChIP-Seq profiles were normalized to a read depth of 10 millions for comparison, if necessary. The mapping information of the ChIP-Seq studies is provided in the following list.

Sample / Epitope	Total reads	Uniquely mapped reads
LSC ^{RH-RAS} / Input	18,324,401	15,535,177
LSC ^{RH-RAS} / Myc-DNMT3A ^{R882H} (9E10)	29,195,863	21,976,593
LSC ^{RH-RAS} / H3K4me1	32,441,556	29,241,182
LSC ^{RH-RAS} / H3K4me3	37,160,636	24,298,214
HOXA9-MEIS1 / Input	13,813,627	8,567,831
HOXA9-MEIS1 / H3K4me1	26,034,975	24,547,306
EV-RAS / Input	24,486,819	21,306,596
EV-RAS / H3K4me1	12,754,638	11,736,817
EV-RAS / H3K27ac	23,362,273	21,913,284
EV-RAS / H3K79me2	65,883,113	58,669,486
EV-RAS / Input (matched for H3K79me2)	63,142,797	56,204,919
WT-RAS / Input	19,919,007	17,486,043
WT-RAS / H3K4me1	25,580,239	23,773,579
WT-RAS / H3K27ac	12,547,468	11,497,378
RH-RAS / Input	36,252,235	31,338,073
RH-RAS / H3K4me1	67,194,048	59,841,459
RH-RAS / H3K27ac	13,021,014	11,684,066
RH-RAS / H3K79me2	64,436,265	56,961,688
RH-RAS / Input (matched for H3K79me2)	62,786,927	56,304,371

ChIP Followed by qPCR (ChIP-qPCR)

The detailed procedure for ChIP-qPCR was described before (Wang et al., 2007; Wang et al., 2009), and the primer sequences used for ChIP-qPCR are provided in the following list. ChIP signals were produced from 3 independent experiments followed by normalization to input signals and presented as mean \pm SD.

Mouse Genes	Primer location	Forward	Reverse	Reference
<i>Hoxa7</i>	upstream/DMR	TGGTGGGCTTCAGCTATTGG	TTCGGGTAGGAATTGGTGGC	
<i>Hoxb1</i>	promoter	GGGACTGCCAAACTCTGGC	CATGTGATCTCTCCAGGCC	(Bernt et al., 2011)
<i>Meis1</i>	intron 6/DMR	ATCTCTGCTCTCTCCCGAG	CTTCCCGGGCATATCTGGTC	
<i>Mn1</i>	exon 1/DMR	GAGGGTGACGAACCAAGGAG	GCTGCCCTTCAGAGTCAGAG	
<i>Mycn</i>	intron 2/DMR	ACTGCCAGGCTAGGAGAGAA	CAATTGTTCCGCTTTCGGT	

Genomic Regions Enrichment of Annotations Tool (GREAT) Analysis

GREAT analysis (McLean et al., 2010) for DNMT3A^{R882H} ChIP-Seq peaks and DNMT3A^{R882H} associated hypo-DMRs was performed at its website (<http://great.stanford.edu>) using a setting of single nearest gene (200.0 kb max extension) with curated regulatory domains included.

Enhanced Reduced Representation Bisulfite Sequencing (eRRBS)

eRRBS was carried out with a previously described protocol with slight modification (Garrett-Bakelman et al., 2015). Briefly, ~300 ng of genomic DNA were digested with three different enzymes (80 units of MspI, 40 units of Bfal and 40 units of MseI) to enhance genomic fragmentation and coverage. The produced fragments were ligated to pre-annealed adapters containing 5'-methyl-cytosine instead of cytosine, followed by overhang fill-in, 3'-terminal-A extension and purification. Bisulfite treatment of the fragments was done using the EZ DNA Methylation–Lightning kit (Zymo Research). Amplified eRRBS libraries was quality checked with Agilent 2200 TapeStation, followed by deep sequencing on the Illumina HiSeq-2000 genome analyzer with 50 bp SE parameters. Differential methylation of CpGs was determined by the methylKit software (Akalin et al., 2012). A cut-off of $q < 0.05$ was used to define differentially methylation CpGs (DMCs). A slight window approach was used to call differentially methylated regions (DMRs), similar to the previously described method (Russler-Germain et al., 2014). Adjacent Cs within 300 bp apart were merged to form genomic regions and genomic regions < 400 bp were discarded. For the rest, we determined the % of hypo- or hyper-Cs within the regions. If a region contained at least 10 Cs and at least 20% of its Cs exhibited hypo- or hyper-methylation, we performed paired t-test to compute the statistics of methylation difference between samples. At the end, a cutoff of p value < 0.0005 was used to define DMRs. The overlapping of DMRs with genomic features was computed by treating DMRs as ChIP-Seq peaks. The sequencing and mapping information of eRRBS studies is provided in the following list.

eRRBS Sample	Total Read #	Mapped Read #	Unique CpGs
EV-RAS	511,071,807	328,254,404	17,169,526
RH-RAS	493,228,644	324,267,522	16,635,379
WT-RAS	523,165,129	320,339,494	16,190,740

Motif Analysis of DMCs

For enrichment of transcription factor binding motifs, 50-bp sequences centered on individual CpGs were used for predicting transcription factor binding sites presented in the Jaspar database (vertebrate TFs only) (Mathelier et al., 2014) by the software FIMO with default parameter (Grant et al., 2011). The number of Cs matching to a TF in all Cs tested by eRRBS and the corresponding numbers in DMCs were compared by Chi-square test for significant difference.

Analysis of Human Methylation Data and Cross-Comparisons with Mouse eRRBS Data

Human AML patient DNA methylation data are downloaded from The Cancer Genome Atlas (TCGA)

(<http://cancergenome.nih.gov/cancersselected/acutemyeloidleukemia>). To compare DNA methylation data between human cancer samples and our mouse data, we used the tool liftover from the UCSC genome browser to find the conserved mouse regions of each individual probe (extended to 100 bp) on the Infinium HumanMethylation450 BeadChip platform used in the human/TCGA study. Conserved regions smaller than 150 bp in the mouse genome were kept and the average methylation level of CpGs in the regions were used to compare the human and mouse methylation. In addition, the conserved regions were intersected with mouse DMRs. For a gene having hypo-DMRs identified among murine HSPCs post transduction of RH-RAS relative to EV-RAS and among human normal-karyotype (NK) AML patients carrying R882-mutated *DNMT3A* relative to WT *DNMT3A* ones, we consider the gene with a common DMR in both human and mouse data sets.

Bisulfite Sequencing

Genomic DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen) and bisulfite-conversion carried out using EZ DNA methylation gold kit (Zymo Research) according to manufacturer's instructions. Purified DNA (50 ng) was used as template for PCRs with region-specific primers designed by Methprimer software (Li and Dahiya, 2002). PCR product was gel purified and cloned into pCR4-TOPO vector (Invitrogen), followed by transformation and mini-preps. Plasmids isolated from at least 10 colonies for each sample were sequenced. The detailed primer sequences for bisulfite sequencing are provided in the following list.

Gene	Location	Forward	Reverse
<i>Meis1</i>	intron 6	TTTTGGAGTTTAAATTGTTTAGATT	AAATCCCCTAACCATACCTAATAAC
<i>Hoxa7</i>	promoter	GGGGATTTTGAATTTTTTTAGTTTT	TAACATAACCAACAACCTCCCTTAC
<i>Mn1</i>	exon 1	GGTTAGGGTTTTGGTTTAGAGGTAGACCCCACTTTAAAAACAACCTTC	

Chromosome Conformation Capture (3C) Followed by qPCR

3C DNA samples were prepared as described previously (Hagege et al., 2007). Briefly, 10 millions cells were cross-linked with 2% formaldehyde for 5 min at room temperature followed by quench in 0.125 M glycine for 5 min. Cells were then washed with ice-cold PBS and lysed in 3C lysis buffer (10 mM Tris-HCl, pH 8.0; 10 mM NaCl; 0.2% NP-40; and 1X Roche Complete Protease Inhibitor Cocktail) for 30 min at 4°C. Nuclei were pelleted by centrifugation at 1000 × g and digested with 600U of high concentration BglII (New England Biolabs) in 1× NEB buffer 3 plus 0.3% SDS and 1.8% Triton-X at 37°C for overnight. After enzyme inactivation with 1.6% SDS at 65°C, samples were brought up to a final volume of 7 ml in 1× NEB T4 ligase buffer plus 1% Triton-X, and added with 6000U of T4 DNA ligase (New England Biolabs). Samples were incubated at 16°C for 4 hr followed by 30-min incubation at room temperature. After ligation, samples were treated with RNase A and proteinase K, extracted with a phenol-chloroform protocol, and precipitated by ethanol. About 100 ng of 3C DNA sample was used as template with a primer pair (one test primer and the constant primer) for qPCR analysis. For testing *Meis1* enhancer-promoter 'looping' in the RH-RAS LSC cells, we employed a negative control cell line, *HOXA9/MEIS1*-coexpressing AML lines (Wang et al., 2005) that are known to have a repressed endogenous *Meis1* gene locus (Wang et al., 2007; Wang et al., 2009). The qPCR signal was first normalized with that of *Gapdh* for input normalization, followed by a second normalization to that of negative cell control (i.e., cells with a silenced endogenous *Meis1*) for calculating relative crosslinking frequency to demonstrate 'looping' specificity. The detailed primer sequences for 3C-qPCR at *Meis1* locus are provided in the following list.

Primer	Location	Sequence (5' to 3')
P0 (constant)	chr11:18,925,586-18,925,605	CCTTGGTGCAAGGACTCTTC
P1 (test)	chr11:18,900,080-18,900,101	TGTTGTCTCCAAGAATTTCCAA
P2 (test)	chr11:18,900,816-18,900,835	GGCAGAGGAAGCTTGAAAAA
P3 (test)	chr11:18,904,495-18,904,514	TGGCTATCCTGGCACTTTCT
P4 (test)	chr11:18,907,527-18,907,545	AAAGTTCCAGGCCCATTT
P5 (test)	chr11:18,909,825-18,909,844	GACTCTCCCGCTGACACTTC

Whole Exome Sequencing and Data Analysis

Genomic DNA was extracted from bone marrow of murine leukemic and normal mice. Whole-exome captures and 50 × sequencing experiments were carried out by Otogenetics (Norcross, GA). The sequencing reads were aligned to mm10 reference genome by using BWA (Li and Durbin, 2009), PCR-produced tag duplicates removed by Picard Tools, and base quality score recalibration and Indel (insert/deletion) realignment using the Genome Analysis Toolkit GATK (McKenna et al., 2010). SNP and Indel discovery and genotyping were performed on all samples simultaneously with GATK Lite 2.3.9 using mm10 reference genome as reference. The called SNP/Indels were further filtered for high-quality variants (total counts ≥ 20, alternate counts ≥ 8). Known SNPs from dbSNP137 and variants present in the normal control samples were removed for identifying tumor-specific variants. Lastly, identified variants matching sequences of transduced human *DNMT3A* or *NRAS* cDNA were discarded. The SNP/Indels were annotated with snpEff v3.6 (Cingolani et al., 2012).

Luciferase Reporter Analysis

pCpGL-CMV/EF1a, a CpG-free, luciferase-containing reporter vector designed for assessing regulatory effects of CpG methylations on the introduced cis-element, was kindly provided by Dr. Rehli (University Hospital Regensburg University Hospital Regensburg, Regensburg, Germany). To specifically assess their putative enhancer activity, the hypo-DMR genomic regions were obtained by genomic PCR followed by cloning into the PstI and SpeI sites of pCpGL-CMV/EF1a vector to replace the CMV enhancer with the hypo-DMR sequences. These generated reporter plasmids were either mock-treated or methylated in vitro with recombinant CpG methyltransferase M.SssI (New England Biolabs) for 4 h at 37°C, followed by plasmid purification with QIAquick PCR Purification Kit (Qiagen). 3 × 10⁵ Hela cells were transfected with 1.0 µg of each reporter plasmid and 0.2 µg of pRL-TK Renilla luciferase control reporter using the standard Lipofectamine 2000 transfection protocol (Invitrogen). Triplicate transfections were harvested 36 hr post-transfection. Cell lysates were assayed for firefly and Renilla luciferase activities using the Dual-Luciferase Reporter Assay System (Promega). The firefly luciferase activity for each transfection was normalized to that of Renilla luciferase. Relative luciferase units are generated by normalization to those of Renilla and then to empty vector controls. The detailed primer sequences for cloning of each tested DMR are provided in the following list.

Gene	DMR location	Forward Primer (with PstI site)	Reverse Primer (with SpeI site)
<i>Hoxa7</i>	Upstream	ATCGctgcagGCTCAGAGGCCTGGTGAAT	ATCGactagtCCAACGCTGTCCCAGAACTA
<i>Meis1</i>	Intron 7	ATCGctgcagTCTATGGCTGGGTTGTGCTG	ATCGactagtTGAAGAAAACGCCTCCTGCT
<i>Meis1</i>	Intron 6	ATCGctgcagTCCTAGGGGTTACGGTCTT	ATCGactagtCATTAGGTTGTCCCCGCCTT
<i>Runx3</i>	Upstream	ATCGctgcagAGGAGAGCCAGGTGTAGAGG	ATCGactagtGGTTAGACCACAGGCTGGAC
<i>Mycn</i>	Intron 2	ATCGctgcagTGTCTGTGCCTTGACAGCTT	ATCGactagtCCAGTCTGCCCATGGATTT
<i>Filip1</i>	Intron 1	ATCGctgcagTACACTTGGTCCAGCAAGGC	ATCGactagtGAGGCAAGAGCTGATGCAGA

CRISPR/Cas9-mediated genomic editing of putative enhancers

LentiCas9-Venus and LentiGuide-Crimson plasmids were kindly provided by Drs. Daniel Bauer (Harvard Medical School) and Feng Zhang (MIT and Broad Institute). CRISPR/Cas9-mediated deletion of a specific genomic region was performed as previously described (Canver et al., 2015). Briefly, two sgRNA sequences that specifically target the boundary sequences of a putative *Meis1* enhancer located in intron 6 were designed using an online CRISPR design tool (crispr.mit.edu). Top ranked sgRNAs (with minimal predicted off-targets) at either side of target region (sgRNA-5' or sgRNA-3'; see also a list below) were selected and cloned into a BsmBI site of LentiGuide-Crimson vector. To generate a single vector with tandem sgRNA expression cassette (i.e., sgRNA-5' plus sgRNA-3'), LentiGuide-Crimson carrying sgRNA-5' was first digested with PspXI and XmaI, and then inserted with PspXI/XmaI-digested PCR products (which contain sgRNA-3') amplified by using LentiGuide-Crimson with sgRNA-3' as the template and specific primers (5'-GGCCGGCC-gctcgaggGAGGGCCTATTTCC-3' and 5'-CCGGCCGGcccgggTTGTGGATGAATACTGCCATTT-3'). The detailed sgRNA sequences are provided below.

sgRNA	Position	Target sequence
sgRNA-5'	5' of <i>Meis1</i> DMR at intron 6	TTAGGTTGTCCCCGCCTTAC
sgRNA-3'	3' of <i>Meis1</i> DMR at intron 6	AATAGGATTACAGCTTCTAC

Stable LentiCas9-Venus expressing cell lines were produced by infection with LentiCas9-Venus lentivirus, followed by Venus sorting on FACS Aria II (BD Biosciences). Lentivirus carrying either the LentiGuide-Crimson empty vector or that with a tandem sgRNA expression cassette was prepared and used for infecting cell lines with stable LentiCas9-Venus expression, followed by sorting of cells expressing both Venus and E2-Crimson 48 hr post-infection. The sorted single cells (200-300 in total) were plated into 96-well plates. After growth for 1-2 weeks, single-cell colonies carrying the desired CRISPR/cas9-induced genomic targeting/deletion were first screened out by PCR using primer pairs either flanking or located within target genomic region (which serves positive or negative selection primers), and further confirmed by direct sequencing of genomic PCR products flanking target genomic region. The detailed genotyping PCR primer sequences are provided below.

Primer Location	Forward Primer	Reverse Primer	Product Size
Within DMR	GTTCCCAGGCCCATTTGAGA	CACTACCGGATGTCGCCTTT	579 bp
Outside of DMR	ACCCACTGCTGGTTGTATCC	AAGACACCGAGGTTGCCATT	~ 382 bp

Compound Treatment

SGC0946 and other used compounds in the study was kindly provided by Drs. Matthieu Schapira and Cheryl Arrowsmith at Structural Genomics Consortium (SGC) in Toronto and used as described before (Yu et al., 2012).

shRNA Knockdown

The pLKO.1 lentiviral shRNA plasmids for *Dot1l* knockdown were obtained from Sigma. Lentivirus production and infection was performed using provider's protocols. An LSC^{RH-RAS} cell line (#1) generated using a puromycin-free MSCV vector similar to the one shown in Figure S1A was used for *Dot1l* knockdown. Stable shRNA-expressing cell lines were established by puromycin (1-2 µg/mL) selection. The detailed target sequences for *Dot1l*-specific shRNAs are provided below.

shRNA	Clone ID	Target sequence	Vector
shDot1l #1	TRCN0000125102	CGGCAGAATCGTATCCTCAAA	pLKO.1
shDot1l #2	TRCN0000125100	GCTGACCTACAATGACCTGAT	pLKO.1

The lentivirus-based shRNA plasmids for knocking down *Meis1* or *Mn1* were generated by cloning a 19-mer shRNA sequence into HpaI and XhoI sites in LentiLox 5.0 vector (kind gift from Dr. James Bear, UNC at Chapel Hill). Stable shRNA-expressing cell lines were obtained by GFP sorting on FACS Aria II (BD Biosciences). The detailed target sequences for *Meis1* or *Mn1*-specific shRNAs are provided below.

shRNA	Target sequence	Vector	Reference
shMeis1 #1	GGATAACAGCAGTGAGCAA	LentiLox 5.0	(Kumar et al., 2009)
shMeis1 #2	GCGTGGCATCTTTCCCAAA	LentiLox 5.0	
shMn1 #1	GGTACATGCCACCTGACAA	LentiLox 5.0	
shMn1 #2	GCTTGAACATGGAGCCCTA	LentiLox 5.0	

Supplemental References

- Abdel-Wahab, O., Adli, M., LaFave, L. M., Gao, J., Hricik, T., Shih, A. H., Pandey, S., Patel, J. P., Chung, Y. R., Koche, R., et al. (2012). ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer cell* 22, 180-193.
- Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., and Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* 13, R87.
- Bock, C., Beerman, I., Lien, W. H., Smith, Z. D., Gu, H., Boyle, P., Gnirke, A., Fuchs, E., Rossi, D. J., and Meissner, A. (2012). DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Molecular cell* 47, 633-647.
- Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., Chen, D. D., Schupp, P. G., Vinjamur, D. S., Garcia, S. P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192-197.
- Chambers, S. M., Boles, N. C., Lin, K. Y., Tierney, M. P., Bowman, T. V., Bradfute, S. B., Chen, A. J., Merchant, A. A., Sirin, O., Weksberg, D. C., et al. (2007). Hematopoietic fingerprints: an expression database of stem cells and their progeny. *Cell stem cell* 1, 578-591.
- Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80-92.
- Deshpande, A. J., Chen, L., Fazio, M., Sinha, A. U., Bernt, K. M., Banka, D., Dias, S., Chang, J., Olhava, E. J., Daigle, S. R., et al. (2013). Leukemic transformation by the MLL-AF6 fusion oncogene requires the H3K79 methyltransferase Dot1l. *Blood* 121, 2533-2541.
- Fujino, T., Yamazaki, Y., Largaespada, D. A., Jenkins, N. A., Copeland, N. G., Hirokawa, K., and Nakamura, T. (2001). Inhibition of myeloid differentiation by Hoxa9, Hoxb8, and Meis homeobox genes. *Exp Hematol* 29, 856-863.
- Garrett-Bakelman, F. E., Sheridan, C. K., Kacmarczyk, T. J., Ishii, J., Betel, D., Alonso, A., Mason, C. E., Figueroa, M. E., and Melnick, A. M. (2015). Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *J Vis Exp*, e52246.
- Goldberg, A. D., Banaszynski, L. A., Noh, K. M., Lewis, P. W., Elsaesser, S. J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678-691.
- Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017-1018.
- Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forne, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature protocols* 2, 1722-1733.

- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264.
- Ji, H., Ehrlich, L. I., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M. J., Irizarry, R. A., Kim, K., et al. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* 467, 338-342.
- Kamps, M. P., Look, A. T., and Baltimore, D. (1991). The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials. *Genes & development* 5, 358-368.
- Konuma, T., Nakamura, S., Miyagi, S., Negishi, M., Chiba, T., Oguro, H., Yuan, J., Mochizuki-Kashio, M., Ichikawa, H., Miyoshi, H., et al. (2011). Forced expression of the histone demethylase Fbxl10 maintains self-renewing hematopoietic stem cells. *Exp Hematol* 39, 697-709 e695.
- Kumar, A. R., Li, Q., Hudson, W. A., Chen, W., Sam, T., Yao, Q., Lund, E. A., Wu, B., Kowal, B. J., and Kersey, J. H. (2009). A role for MEIS1 in MLL-fusion gene leukemia. *Blood* 113, 1756-1758.
- Lee, T. I., Johnstone, S. E., and Young, R. A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nature protocols* 1, 729-748.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760.
- Li, L. C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18, 1427-1431.
- Lo, M. C., Peterson, L. F., Yan, M., Cong, X., Jin, F., Shia, W. J., Matsuura, S., Ahn, E. Y., Komeno, Y., Ly, M., et al. (2012). Combined gene expression and DNA occupancy profiling identifies potential therapeutic targets of t(8;21) AML. *Blood* 120, 1473-1484.
- Mathelier, A., Zhao, X., Zhang, A. W., Parcy, F., Worsley-Hunt, R., Arenillas, D. J., Buchman, S., Chen, C. Y., Chou, A., Ienasescu, H., et al. (2014). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic acids research* 42, D142-147.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20, 1297-1303.
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495-501.
- Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J. P. (2006). GenePattern 2.0. *Nature genetics* 38, 500-501.

- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 15545-15550.
- Tiacci, E., Spanhol-Rosseto, A., Martelli, M. P., Pasqualucci, L., Quentmeier, H., Grossmann, V., Drexler, H. G., and Falini, B. (2012). The NPM1 wild-type OCI-AML2 and the NPM1-mutated OCI-AML3 cell lines carry DNMT3A mutations. *Leukemia* *26*, 554-557.
- Wang, G. G., Calvo, K. R., Pasillas, M. P., Sykes, D. B., Hacker, H., and Kamps, M. P. (2006). Quantitative production of macrophages or neutrophils ex vivo using conditional Hoxb8. *Nature methods* *3*, 287-293.
- Wang, J., Liu, Y., Li, Z., Wang, Z., Tan, L. X., Ryu, M. J., Meline, B., Du, J., Young, K. H., Ranheim, E., et al. (2011). Endogenous oncogenic Nras mutation initiates hematopoietic malignancies in a dose- and cell type-dependent manner. *Blood* *118*, 368-379.
- Xu, B., On, D. M., Ma, A., Parton, T., Konze, K. D., Pattenden, S. G., Allison, D. F., Cai, L., Rockowitz, S., Liu, S., et al. (2015). Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia. *Blood* *125*, 346-357.
- Yu, W., Chory, E. J., Wernimont, A. K., Tempel, W., Scopton, A., Federation, A., Marineau, J. J., Qi, J., Barsyte-Lovejoy, D., Yi, J., et al. (2012). Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nat Commun* *3*, 1288.
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* *9*, R137.