

GENETICS

Tumor suppressor CEBPA interacts with and inhibits DNMT3A activity

Xiufei Chen^{1,2†}, Wenjie Zhou^{1,3†}, Ren-Hua Song^{4†}, Shuang Liu^{5†}, Shu Wang⁶, Yujia Chen^{1,3}, Chao Gao^{1,3}, Chenxi He^{1,3}, Jianxiong Xiao^{1,3}, Lei Zhang¹, Tianxiang Wang^{1,3}, Peng Liu^{1,3}, Kunlong Duan^{1,3}, Zhouli Cheng^{1,3}, Chen Zhang^{1,3}, Jinye Zhang^{1,3}, Yiping Sun^{1,3}, Felix Jackson⁷, Fei Lan^{1,3}, Yun Liu⁸, Yanhui Xu¹, Justin Jong-Leong Wong⁴, Pu Wang^{1,3}, Hui Yang⁹, Yue Xiong^{10‡}, Tong Chen^{6*}, Yan Li^{5*}, Dan Ye^{1,3,11*}

DNA methyltransferases (DNMTs) catalyze DNA methylation, and their functions in mammalian embryonic development and diseases including cancer have been extensively studied. However, regulation of DNMTs remains under study. Here, we show that CCAAT/enhancer binding protein α (CEBPA) interacts with the long splice isoform DNMT3A, but not the short isoform DNMT3A2. CEBPA, by interacting with DNMT3A N-terminus, blocks DNMT3A from accessing DNA substrate and thereby inhibits its activity. Recurrent tumor-associated CEBPA mutations, such as preleukemic CEBPA^{N321D} mutation, which is particularly potent in causing AML with high mortality, disrupt DNMT3A association and cause aberrant DNA methylation, notably hypermethylation of PRC2 target genes. Consequently, leukemia cells with the CEBPA^{N321D} mutation are hypersensitive to hypomethylation agents. Our results provide insights into the functional difference between DNMT3A isoforms and the regulation of de novo DNA methylation at specific loci in the genome. Our study also suggests a therapeutic strategy for the treatment of CEBPA-mutated leukemia with DNA-hypomethylating agents.

INTRODUCTION

DNA methylation of cytosine-guanine (CpG) dinucleotides is a covalent modification involved in the regulation of gene expression and plays crucial functions in mammalian development and diseases (1, 2). In mammalian genomes, DNA methylation is usually excluded from CpG-rich promoters of actively transcribed genes, and the correlation between DNA methylation and gene silencing increases with the density of CpG islands (CGIs) at gene promoters (3). There are two de novo DNA methyltransferases (DNMTs) in mammals, DNMT3A and DNMT3B, which both contain a highly conserved catalytic domain in the C terminus and recognize unmethylated CpG. Genetic studies in mouse models have linked the function of

Dnmt3a and *Dnmt3b* genes to diverse biological pathways. For instance, mice with whole-body knockout (KO) of *Dnmt3a* gene are normal at birth but die at about 4 weeks of age (4). *Dnmt3b*-null mouse embryos encounter rostral neural tube defects and growth impairment and die in utero (4), suggesting that *Dnmt3b* may play a more substantial role during embryonic development than *Dnmt3a*. Catalytically inactive *Dnmt3b* can restore a majority of methylation and expression changes and rescue mouse embryonic defect by accessory functions to support other *Dnmts* (5). In addition, *Dnmt3a* is also essential for paternal and maternal genome imprinting and germ cell development (6, 7). Tissue-specific KO of *Dnmt3a* has revealed its important regulatory roles in adipose (8), bone (9), muscle (10, 11), T cells (12), and the innate immune system (13). Although the function of DNMT enzymes in mammalian development and diseases has been extensively studied, it is still unclear how these epigenetic enzymes are regulated at specific genomic regions such as CpG-rich promoters and thereby precisely regulate gene expression by DNA methylation.

Alternate promoter usage at the *DNMT3A* locus gives rise to two isoforms: DNMT3A1 (also known as DNMT3A) and DNMT3A2. The long splice isoform DNMT3A contains a unique N-terminal domain containing 219 and 223 amino acid residues in mouse and human, respectively (14), and is expressed broadly. The short isoform DNMT3A2 is expressed restrictively in embryonic stem cells (ESCs) and tissues with active proliferation and differentiation. Both isoforms of *Dnmt3a* are equally active in in vitro assays (14, 15), but they have different chromatin binding preferences in mouse ESCs (16). *Dnmt3a* is substantially more active than *Dnmt3a2* in increasing DNA methylation at the major and minor satellite repeats in *Dnmt3a/3b/3l* triple-KO mouse ESCs (17), and isoform-specific function of *Dnmt3a* is also required during germ cell development in mice (18). The current knowledge on the function of the N-terminal region uniquely present in DNMT3A is very limited. This domain contains a nuclear localization signal sequence, which may account

¹Huashan Hospital, Fudan University, and Shanghai Key Laboratory of Medical Epigenetics, International Co-laboratory of Medical Epigenetics and Metabolism (Ministry of Science and Technology), and Molecular and Cell Biology Lab, Institutes of Biomedical Sciences, Shanghai Medical College of Fudan University, Shanghai, China. ²Target Discovery Institute, NDM Research Building, Oxford Ludwig Institute of Cancer Research, Oxford University, Old Road Campus, Roosevelt Drive, Oxford OX3 7FZ, UK. ³Key Laboratory of Metabolism and Molecular Medicine, Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai, China. ⁴Epigenetics and RNA Biology Program, Centenary Institute, The University of Sydney, Camperdown 2050, Australia. ⁵MOE Key Laboratory of Model Animals for Disease Study, Department of Rheumatology and Immunology, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Chemistry and Biomedicine Innovation Center (ChemBIC), Model Animal Research Center, Nanjing University Medical School, Nanjing University, Nanjing, China. ⁶Department of Hematology, Huashan Hospital, Fudan University, Shanghai, China. ⁷Department of Computer Science, University of Oxford, 15 Parks Rd, Oxford OX1 3QD, UK. ⁸MOE Key Laboratory of Metabolism and Molecular Medicine, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences and Zhongshan Hospital, Fudan University, Shanghai, China. ⁹Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai, China. ¹⁰Lineberger Comprehensive Cancer Center, Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599, USA. ¹¹Department of General Surgery, Huashan Hospital, Fudan University, Shanghai, China.

*Corresponding author. Email: chentong@fudan.edu.cn (T.C.); yanli@nju.edu.cn (Yan Li); yedan@fudan.edu.cn (D.Y.)

†These authors contributed equally to this work.

‡Present address: Cullgen Inc., 12671 High Bluff Drive, San Diego, CA 92130, USA.

for the exclusive nuclear localization of Dnmt3a and its heterochromatin enrichment (17), and can also facilitate the interaction with other chromatin-remodeling proteins, such as lysine methyltransferase G9a (19, 20).

The transcription factor CCAAT/enhancer binding protein α (CEBPA) controls lineage-specific gene expression and acts as a master regulator of terminal differentiation of several cell types including granulocytes (21). The *CEBPA* gene comprises a single exon with an internal translational start site, resulting in the full-length CEBPA (42 kDa; p42) and a shorter N-terminal-truncated isoform (30 kDa; p30). As compared to full-length p42, p30 lacks the first 117 amino acids including the transactivating domain 1 (TAD1) and exhibits lower transactivation potential (22, 23). The *CEBPA* gene is mutated in a subset of acute myeloid leukemia (AML), and most *CEBPA*-mutated AML patients harbor a unique biallelic mutation pattern: an N-terminal frameshift mutation in one allele that generates a dominant negative p30 isoform (23) and a C-terminal missense mutation in the other allele that targets the basic region-leucine zipper domain (bZIP) and affects DNA binding and/or homo- or heterodimerization with other CEBP family members (24, 25). Among tumor-associated missense mutations targeting the C-terminal region of CEBPA, N321D mutation is particularly potent to cause AML with high mortality in murine models (26). In an effort to determine how mutations of CEBPA contribute to AML development, we found that CEBPA mutations are associated with CpG hypermethylation, especially Polycomb repressive complex 2 (PRC2) target genes. We found that CEBPA interacts with and acts as a novel inhibitor of DNMT3A and further demonstrated that AML-derived C-terminal mutations in CEBPA disrupt DNMT3A interaction and relieve DNMT3A inhibition, leading to DNA hypermethylation and sensitizing leukemia cells to hypomethylation agents.

RESULTS

CEBPA physically interacts with DNMT3A

Aberrant DNA methylation is a well-established cancer hallmark, including regional hypermethylation at CGIs of tumor suppressor genes (27, 28). A previous study by Sinha and colleagues (29) reported genetic drivers of DNA hypermethylation in AML patients, including isocitrate dehydrogenase 2 (IDH2), CEBPA, and Wilms tumor protein 1 (*WT1*). In accord, we found that mutations in *IDH2*, *CEBPA*, and *WT1* led to increased methylation of 43,417, 42,537, and 33,077 CpG sites, respectively, in The Cancer Genome Atlas (TCGA) AML dataset consisting of 194 patients (Fig. 1A). As compared to DNMT3A mutation, which causes 29,056 hypomethylated CpG sites, much fewer hypomethylated CpG sites were found in AML patients with mutations in *IDH2* (803), *CEBPA* (1901), or *WT1* (1808), confirming that mutations in these three genes are associated with DNA hypermethylation in AML. We and others have previously demonstrated that ten-eleven translocation 2 (TET2) DNA dioxygenase is catalytically inactivated by D-2-hydroxyglutarate, an oncometabolite produced by the mutated IDH (30, 31), which occurs in a mutually exclusive manner with TET2 mutations (32); the transcription factor WT1 recruits TET2 to enable target promoter demethylation and that AML-associated recurrent mutations in WT1 or TET2 disrupt their interaction and cause focal CGI hypermethylation (32, 33). In AML cases examined, mutations in *CEBPA* are linked to the second largest number of hypermethylated CpG sites, but the underlying mechanism is yet to be elucidated.

To determine how CEBPA affects DNA methylation, we applied a dual-luciferase reporter assay system and identified DNMT3A as a potential CEBPA-interacting protein, which exhibited the strongest luciferase reporter activation followed by DNMT3B (Fig. 1B). The protein association of Flag-tagged CEBPA with Myc-tagged DNMT3A and DNMT3B, but not DNMT1 and DNMT3L, was confirmed by coimmunoprecipitation (co-IP) and Western blotting (Fig. 1C). We then focused on DNMT3A, given that its interaction with CEBPA was the strongest among all DNA methylation regulators examined. We observed that CEBPA selectively bound to DNMT3A, but not DNMT3A2, when co-overexpressed in human embryonic kidney (HEK) 293T cells (Fig. 1, D and E), implying that the N terminus of DNMT3A might be crucial for the interaction with CEBPA. By constructing a series of truncated proteins of DNMT3A, we found that CEBPA specifically interacted with the N-terminal region of DNMT3A (amino acid residues 1 to 291), but not the other regions, such as the Pro-Trp-Trp-Pro domain (PWWP; amino acid residues 292 to 481), the ATRX-DNMT3-DNMT3L domain (ADD; amino acid residues 482 to 633), and the catalytic domain (CD; amino acid residues 634 to 912) (fig. S1A). In situ proximity ligation assay revealed the physical closeness of endogenous CEBPA and DNMT3A in K562 leukemia cells (fig. S1B). The interaction between endogenous DNMT3A and CEBPA proteins was confirmed in K562 and HL-60 cells via IP coupled with Western blotting (fig. S1C). To further confirm the specificity of endogenous co-IP, we found that CEBPA was immunoprecipitated by the DNMT3A antibody in wild-type HL-60 cells, but not in *DNMT3A*-KO cells (Fig. 1F). In addition, in vitro pull-down assay demonstrated the direct protein interaction between CEBPA and DNMT3A (fig. S1, D to F). Collectively, these data suggest that CEBPA physically interacts with DNMT3A and that the N-terminal domain of DNMT3A is primarily responsible for this interaction.

AML-derived mutations in CEBPA C terminus disrupt DNMT3A binding

To map the binding regions between CEBPA and DNMT3A, we constructed a series of CEBPA truncations and observed the retention of hemagglutinin (HA)-tagged DNMT3A in the immunoprecipitates of Flag-tagged p30 and bZIP domain (amino acid residues 278 to 358), but not the region containing TAD1 (amino acid residues 1 to 120) or TAD2 (amino acid residues 121 to 277) (Fig. 2A), indicating that the C-terminal bZIP domain of CEBPA mediates the interaction with DNMT3A. Besides CEBPA (α), the CEBP family transcription factors contain CEBPB (β), CEBPG (γ), CEBPD (δ), and CEBPE (ϵ), all of which have the conserved bZIP domain in their C terminus (fig. S2, A and B). Moreover, we found that DNMT3A N terminus is conserved in mouse and human (fig. S2C) and that CEBP proteins interacted with the N-terminal domain (amino acid residues 1 to 291) of both human and mouse DNMT3A when co-overexpressed in HEK293T cells (fig. S2, D and E). Collectively, these findings suggest that the bZIP domain of CEBP proteins mediates the interaction with DNMT3A through its N-terminal domain.

According to the CEBPA-DNA cocystal structure (24), the bZIP domain of CEBPA can be divided into a DNA contacting basic region (BR; amino acid residues 278 to 309) and a leucine zipper coiled-coil interface for dimerization (LZ; amino acid residues 310 to 358) (fig. S3, A and B). Among the TCGA AML patients, the

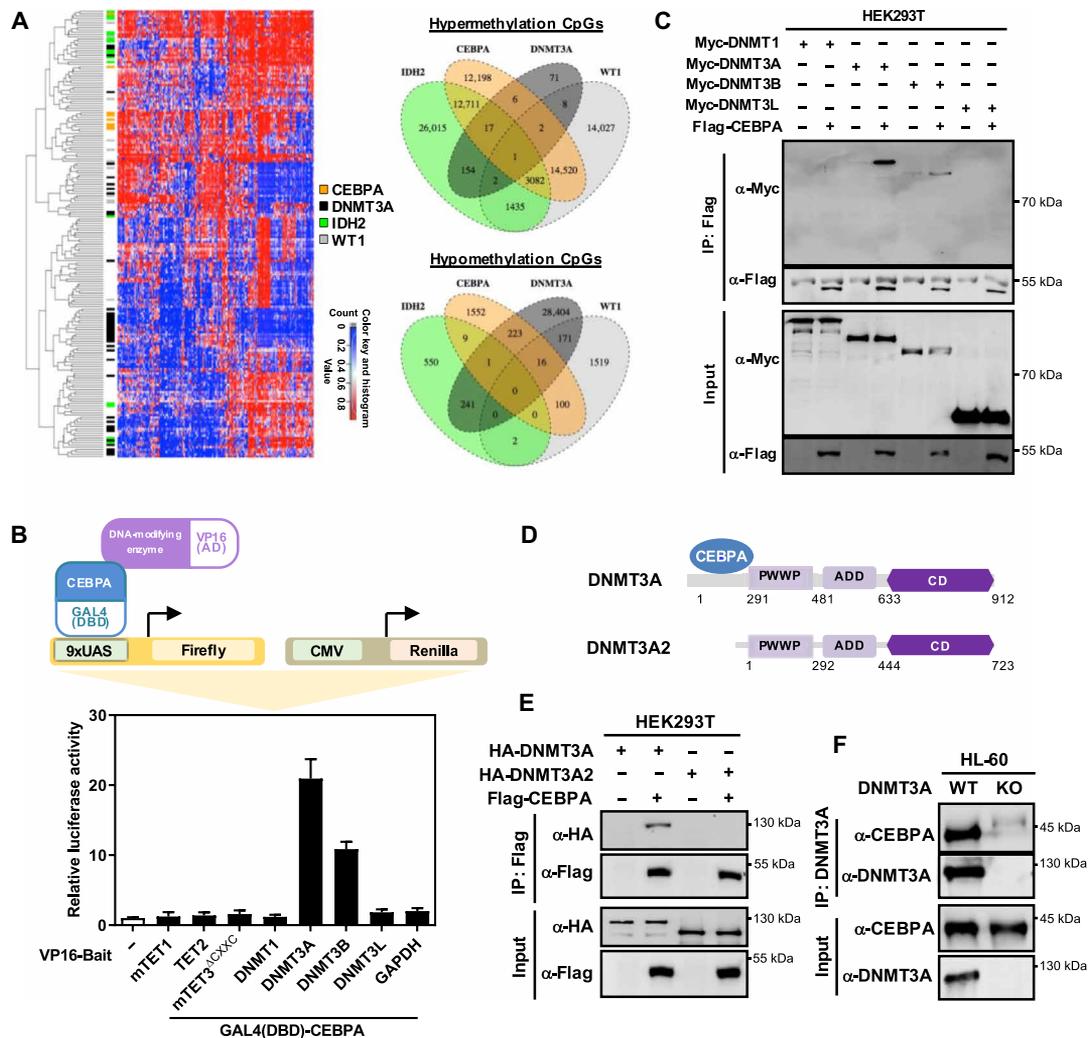


Fig. 1. CEBPA physically interacts with DNMT3A. (A) Heatmap and Venn diagram showing different DNA methylation in 194 sequenced AML patients from the TCGA cohort, including mutations that are mostly linked to DNA hypermethylation (WT1, IDH2, and CEBPA) and hypomethylation (DNMT3A). (B) Schematic representation of the mammalian two-hybrid system and dual-luciferase reporter system used to search for CEBPA-interacting DNA-modifying enzymes in HEK293K cells. Average values of triplicated results with SD are shown. (C) Myc-tagged DNMT proteins, including DNMT1, DNMT3A, DNMT3B, and DNMT3L, were coexpressed with Flag-CEBPA in HEK293T. Protein-protein interaction was examined by IP-Western using the indicated antibodies. (D) Schematic representation of two isoforms of DNMT3A. (E) HA-tagged DNMT3A or DNMT3A2 was coexpressed with Flag-CEBPA in HEK293T cells. Protein-protein interaction was examined by IP-Western using the indicated antibodies. (F) Endogenous DNMT3A protein was purified by IP with an anti-DNMT3A antibody from HL-60 cells with or without knockout (KO) of DNMT3A, followed by Western blot to detect CEBPA. WT, wild type.

C-terminal bZIP domain contains hotspot mutations in CEBPA (Fig. 2B), but only a few mutations in this region have been functionally characterized (26). Notably, of a total of 19 AML-derived CEBPA mutants targeting the bZIP domain tested, almost all reduced or abolished the CEBPA-DNMT3A interaction (Fig. 2, C and D, and fig. S3, C to F). These findings suggest that disrupting CEBPA-DNMT3A interaction may be a shared mechanism underlying tumor-derived C-terminal mutations in CEBPA.

CEBPA inhibits DNMT3A by decreasing its accessibility to DNA substrate

It was previously reported that the N-terminal domain of DNMT3A was the major determinant of its accessibility to DNA substrate (34), which led us to speculate that CEBPA might interfere with the DNA binding activity of DNMT3A. In vitro electrophoretic mobility

shift assay (EMSA) demonstrated that DNMT3A bound to DNA with no sequence selectivity, while CEBPA selectively recognized sequences containing the specific motif (GCAAT) (fig. S4, A and B). The dimerization-defective CEBPA^{N321D} mutant failed to bind this DNA motif (fig. S5A), and this mutation in CEBPA also failed to influence the binding of DNMT3A N terminus (amino acid residues 1 to 291) to random DNA (fig. S5B). CEBPA remarkably reduced the binding of either DNMT3A N terminus or full-length protein to random DNA (fig. S5, C and D), and this inhibitory effect on the DNA binding activity of DNMT3A N terminus or full-length protein was in a CEBPA dose-dependent manner (Fig. 3, A and B). Considering that CEBPA mutations are heterozygous in most clinical cases of AML, we found that the DNA binding activity of DNMT3A N terminus was diminished by the coexistence of wild-type CEBPA and N321D mutant proteins in vitro (Fig. 3C).

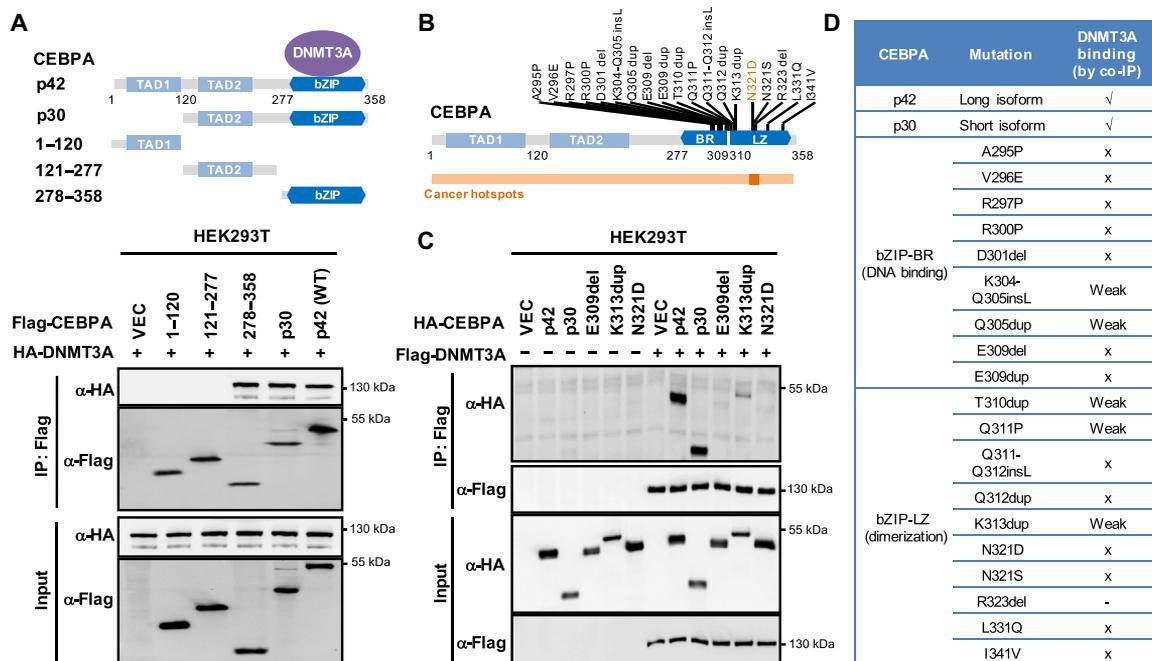


Fig. 2. AML-derived mutations in CEBPA C terminus disrupt DNMT3A binding. (A) Wild-type CEBPA and its deletion mutants, as shown in the schematic illustration (top), were coexpressed with HA-DNMT3A in HEK293T cells. Protein-protein interaction was examined by IP-Western using the indicated antibodies. (B and C) HA-tagged wild-type CEBPA or AML-derived mutants, as shown in the schematic illustration (B), were transiently coexpressed with Flag-DNMT3A in HEK293T cells. Protein-protein interaction was examined by IP-Western using the indicated antibodies. (D) Summary of tested AML-derived CEBPA mutations that affect DNMT3A binding, including 19 mutations in the BR and LZ regions of CEBPA bZIP domain.

To quantitatively assess the effect of CEBPA on impairing the DNA binding activity of DNMT3A, we measured the superimposed fluorescence polarization of DNMT3A with the presence of CEBPA (i.e., p42) by using a FAM-labeled random DNA sequence containing one CG dinucleotide. As shown, the N terminus of DNMT3A exhibited lower dissociation constant (K_d) values than the full-length protein for binding with random DNA substrate in vitro (Fig. 3, D and E), which is in line with previous studies reporting that the N terminus allowed Dnmt3a to anchor in the nucleus and recognize nucleic acids or nucleoproteins (34, 35). The DNA binding affinity of DNMT3A N terminus was reduced by CEBPA recombinant protein, with the increment of K_d value from $1.0 \pm 0.1 \mu\text{M}$ to $7.2 \pm 0.5 \mu\text{M}$ (Fig. 3D). Likewise, the DNA binding affinity of full-length DNMT3A was reduced by wild-type CEBPA, as the K_d value was increased from $7.6 \pm 0.4 \mu\text{M}$ to $117.5 \pm 35.3 \mu\text{M}$ (Fig. 3E). The inhibitory effect on the DNA binding activity of DNMT3A N terminus or full-length protein was, however, not observed when using the recombinant protein of CEBPA^{N321D} mutant (Fig. 3D and fig. S5E).

Next, we evaluated the effect of CEBPA on regulating DNMT3A enzyme activity in vitro. DNMT3A is known to be structurally stimulated by histone H3K4me0, which disrupts ADD-CD interaction and induces a large movement of the ADD domain and releases the autoinhibition form of DNMT3A (36). In the presence of H3K4me0 peptide and random DNA substrate, the enzymatic activity of DNMT3A was significantly suppressed ($P < 0.001$) by recombinant wild-type CEBPA, but not the N321D mutant (Fig. 3F). The DNMT3A D529A mutant within the ADD link, which can hamper ADD-CD interaction and thus release autoinhibition of DNMT3A (36), was still significantly inhibited ($P < 0.001$) by

recombinant wild-type CEBPA, but not N321D mutant (fig. S6A), implying that CEBPA inhibits DNMT3A activity without causing H3K4me0-mediated DNMT3A protein structural change. Furthermore, we found that the inhibitory effect of CEBPA on DNMT3A activity was more profound when the specific motif (GCAAT) of CEBPA was used as DNA substrate (Fig. 3G). In contrast, CEBPA failed to inhibit the activity of DNMT3A2 or DNMT3A^{PWWP-ADD-CD} (fig. S6, B and C), both of which lack the N-terminal CEBPA-interacting domain. Together, these in vitro results support a model that CEBPA inhibits DNMT3A activity by decreasing the accessibility of DNMT3A to DNA substrate.

CEBPA decreases CGI methylation at target gene promoters

Approximately 10% of AML patients show mutations in the CEBPA coding region, including N-terminal frameshift mutations and/or C-terminal in-frame insertions/deletion mutations in the bZIP domain (Fig. 4A). Notably, total CpG methylation was significantly increased ($P < 0.05$) in C-terminal CEBPA-mutated patients, but not in patients harboring N-terminal mutations, as compared to patients with wild-type CEBPA (Fig. 4B). In agreement with the findings in AML patient samples, overexpression of wild-type CEBPA, but not the N321D mutant, led to decreased global 5mC in cells (Fig. 4C and fig. S7, A and B). CEBPA did not affect either DNMT3A homodimer formation or DNMT3A-DNMT3L interaction (fig. S7, C and D), both of which are important for stimulating DNMT3A activity (37, 38). Moreover, we found that both wild-type CEBPA and the N321D mutant localized in the nucleus when expressed in cells (fig. S8A). However, the chromatin/DNA binding of the CEBPA^{N321D} mutant was remarkably reduced as compared to

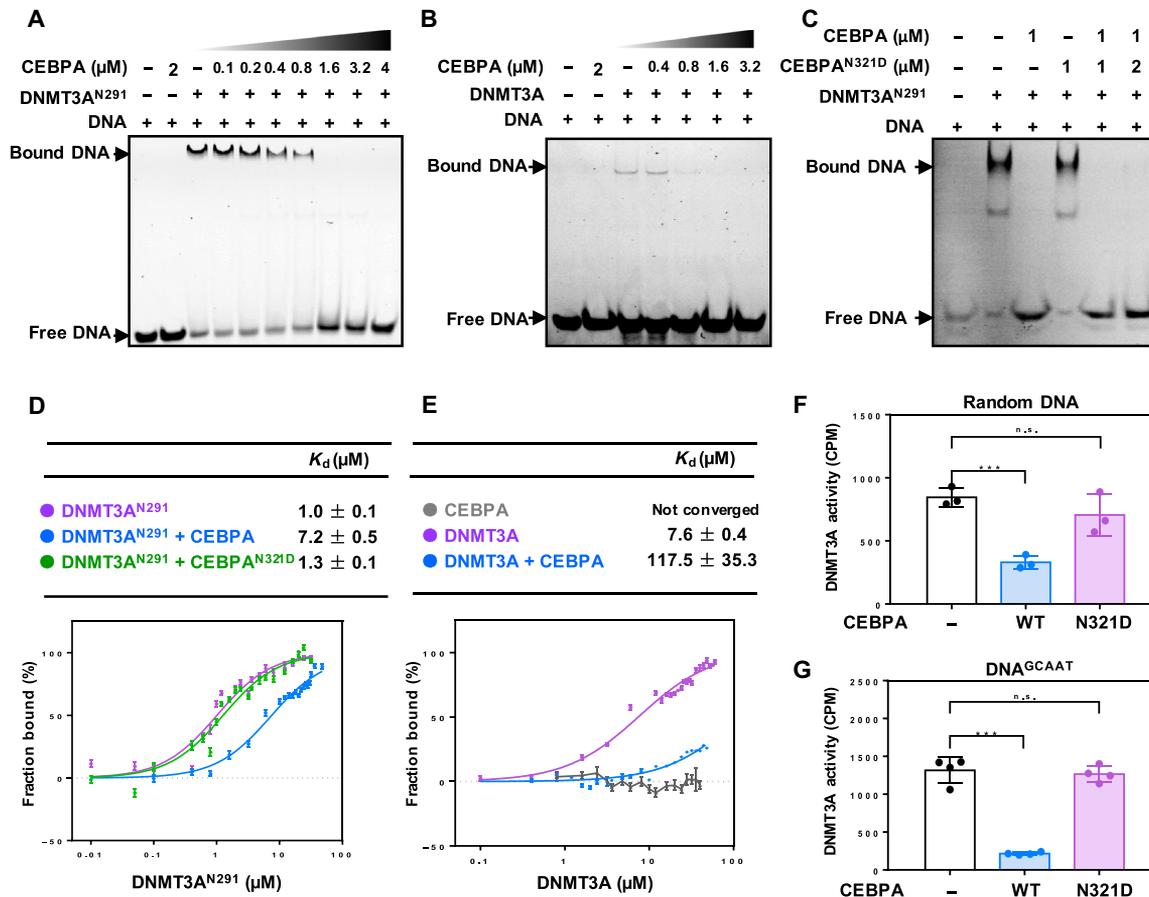


Fig. 3. CEBPA inhibits the activity of DNMT3A by decreasing its accessibility to substrate DNA. (A and B) EMSA for DNA binding activity of N terminus (residues 1 to 291) (A) or full length (B) of DNMT3A in the absence or presence of CEBPA. See Materials and Methods for more details. Data are representative of three independent experiments. (C) EMSA showing the DNA binding ability of DNMT3A N terminus in the absence or presence of both wild-type CEBPA and N321D mutant proteins at the mix ratios of 1:1 or 1:2. (D and E) Superimposed fluorescence polarization plots for DNA binding affinities of the N terminus (D) or full length (E) of DNMT3A in the absence or presence of wild-type CEBPA or the N321D mutant. Random DNA substrate (forward, 5'-6-FAM/TGGATATCTAGGGGCGCTATGATATCT-3'; reverse, 5'-AGATATCATAGC-GCCCTAGATATCCA-3') was used. See Materials and Methods for more details. (F and G) In vitro methyltransferase activity assays for DNMT3A activity measured in the presence of wild-type CEBPA or N321D mutant protein, using random DNA substrate (F) or CEBPA binding motif (GCAAT) containing DNA (G). See Materials and Methods for more details. n.s., not significant; CPM, counts per minute.

wild-type CEBPA (fig. S8B). The capability of CEBPA^{N321D} mutant to form dimers with the full-length CEBPA (p42) or its truncated isoform (p30) was also impaired in cells (fig. S8C), which is in line with a previous study showing that dimerization of bZIP proteins occurs in the absence of DNA and is a prerequisite for DNA binding (39).

How is DNMT3A regulated by CEBPA at specific genomic regions in the cell? An appealing model is that CEBPA interacts with DNMT3A N terminus and subsequently hampers the accessibility of DNMT3A to DNA, including sites containing the CEBPA binding motif, and thus prevents CGI methylation at target gene promoters. To test this hypothesis, we generated stable HL-60 cells with depletion of endogenous CEBPA and reinduction of lentiviral vectors expressing empty vector (VEC) or Flag-tagged wild-type CEBPA or DNMT3A-binding defective N321D mutant to a comparable level as endogenous CEBPA protein (fig. S8D). In these stable cells, coupled chromatin IP and quantitative polymerase chain reaction (ChIP-qPCR) analysis confirmed the binding of wild-type

CEBPA, but not the N321D mutant, to transcription start sites of target genes, including *GATA3*, *CTNNA1*, and *HOXA5*, but not *GAPDH* (a negative control) (Fig. 4D). Bisulfite (BS)-PCR and sequencing analysis showed that 5mC levels at the CpG-rich promoter regions of *GATA3*, *CTNNA1*, and *HOXA5* were remarkably lower in stable cells expressing wild-type CEBPA than those expressing the CEBPA^{N321D} mutant or empty vector (Fig. 4E). As a result, put-back of CEBPA, but not the N321D mutant, led to up-regulation of target gene expression in stable HL-60 cells with endogenous CEBPA depletion (Fig. 4F). In addition, we also re-analyzed the Illumina 450K methylation array and RNA sequencing (RNA-seq) results from the TCGA-AML cohort and confirmed that down-regulation of selected CEBPA target genes (e.g., *GATA3*, *CTNNA1*, and *HOXA5/9*) was associated with CpG hypermethylation in CEBPA-mutated AML patient samples (fig. S9, A to C). Collectively, these results indicate that CEBPA inhibits DNMT3A and prevents CGI methylation to induce target gene expression in both cultured leukemia cells and AML patients.

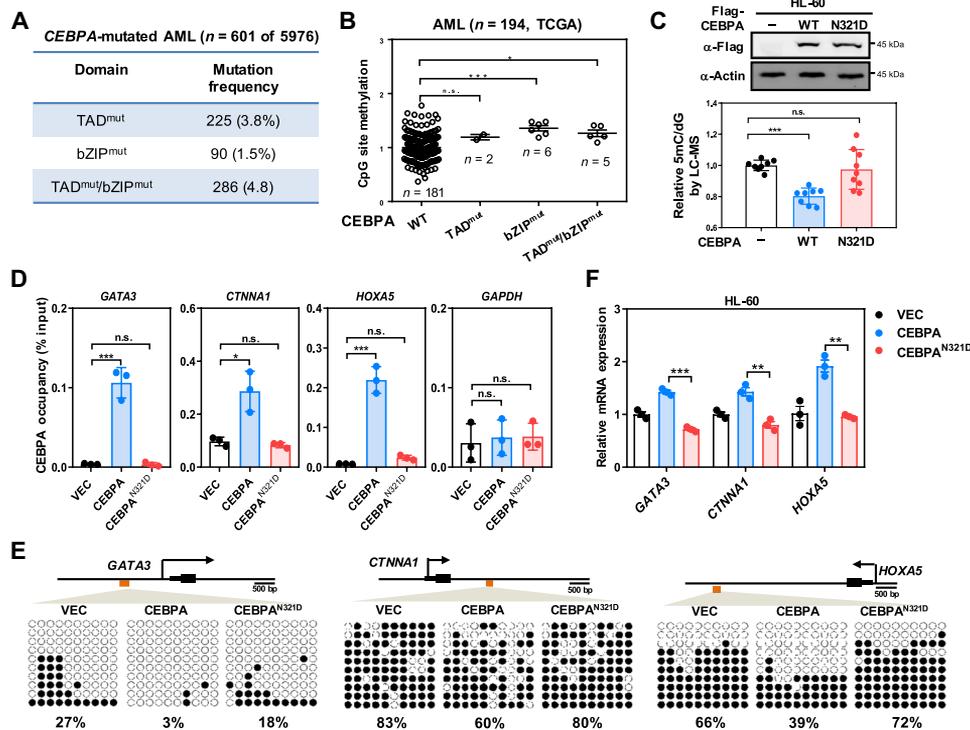


Fig. 4. CEBPA decreases CGI methylation at target gene promoters. (A) Summary of CEBPA mutations in total 5976 AML patients according to the TCGA database and other literatures. (B) Reanalysis of DNA methylation in 194 sequenced AML patients by dividing the TCGA cohort into four groups, including CEBPA nonmutation (i.e., WT), single N-terminal mutations (i.e., TAD^{mut}), single C-terminal mutations (i.e., bZIP^{mut}), and double mutations (i.e., both TAD^{mut}/bZIP^{mut}). The patient numbers for each group are shown. Asterisks denote statistical significance with two-tailed unpaired Student's *t* test. (C) In HL-60 cells, Flag-tagged wild-type CEBPA or N321D mutant protein was ectopically expressed, followed by Western blot to detect overexpressed proteins (top). In these cells, genomic DNA was isolated and then subjected to LC-MS/MS analysis to quantify global 5mC levels. See Materials and Methods for more details. (D) Stable HL-60 cells with knockdown of endogenous CEBPA and reinduction of lentiviral vectors expressing empty vector (VEC) or Flag-tagged wild-type CEBPA or the N321D mutant to a comparable level as endogenous CEBPA protein. In these stable cells, the occupancy of Flag-CEBPA at the promoter regions of indicated target genes was determined by ChIP-qPCR. See Materials and Methods for more details. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was included as a negative control. (E) In stable HL-60 cells with CEBPA knockdown and put-back as described in (E), 5mC levels at CEBPA binding sites at the promoter regions of indicated CEBPA target genes were determined by BS-PCR. See Materials and Methods for more details. 5mC level was calculated as the percentage of methylated cytosines. (F) In stable HL-60 cells in (E), the mRNA expression of indicated CEBPA target genes was determined by quantitative reverse transcription (qRT-PCR). Average values of triplicated results with SD are shown. Asterisks denote statistical significance with two-tailed unpaired Student's *t* test. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001 for the indicated comparison.

AML-derived mutations in CEBPA C terminus cause DNA hypermethylation of PRC2 targets

To systematically investigate the impact of CEBPA mutation on promoter CGI methylation and gene regulation, we next carried out gene set enrichment analysis (GSEA) in the TCGA AML dataset and found that hypermethylated genes associated with CEBPA mutations, hereafter referred to as CEBPA^{mut} genes, were overrepresented in multiple gene sets of PRC2 component targets and genes associated with histone H3 lysine 27 trimethylated markers (H3K27me3) ($P < 3.51 \times 10^{-85}$ for multiple sets) (fig. S10A). There was a significant overlap between CEBPA^{mut} genes and high CpG-containing promoters as defined by the hypergeometric test ($P = 1.62 \times 10^{-21}$) (fig. S10B). Supporting the negative correlation between DNA methylation and gene expression, three-fifths of CEBPA^{mut} genes overlap with down-regulated genes in the TCGA AML cohort ($P = 7.64 \times 10^{-15}$, hypergeometric test) (fig. S10C).

We analyzed the hypermethylated genes associated with CEBPA C-terminal mutations, hereafter referred to as CEBPA^{C-mut} genes, and again found that PRC2 component target genes and H3K27me3-associated

genes were significantly hypermethylated (Fig. 5, A and B). There was a significant overlap between CEBPA^{C-mut} genes and high CpG-containing promoters ($P = 3.65 \times 10^{-30}$, hypergeometric test) (Fig. 5C). Again, supporting the negative correlation between DNA methylation and gene expression, two-thirds of CEBPA^{C-mut} genes overlap with down-regulated genes in the TCGA AML cohort ($P = 3.05 \times 10^{-16}$, hypergeometric test) (Fig. 5D). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that CEBPA^{C-mut} genes were enriched in multiple signaling pathways critical for the pathogenesis of AML, including the Wnt signaling, the stem cell pluripotency-regulating pathway, and the Rap1 signaling pathway (Fig. 5E and table S1).

Of all hypermethylated promoters associated with CEBPA mutations or C-terminal mutations in the TCGA AML cohort, ~97% contained at least one CEBPA binding motif (G/CCAAT) (fig. S10, D and E). Of the total PRC2 target genes, 44.8% (291 of 649) and 41.8% (271 of 649) were hypermethylated at their CpG-containing promoters in AML patients with CEBPA mutations and C-terminal mutations, respectively (table S2). These results indicate that

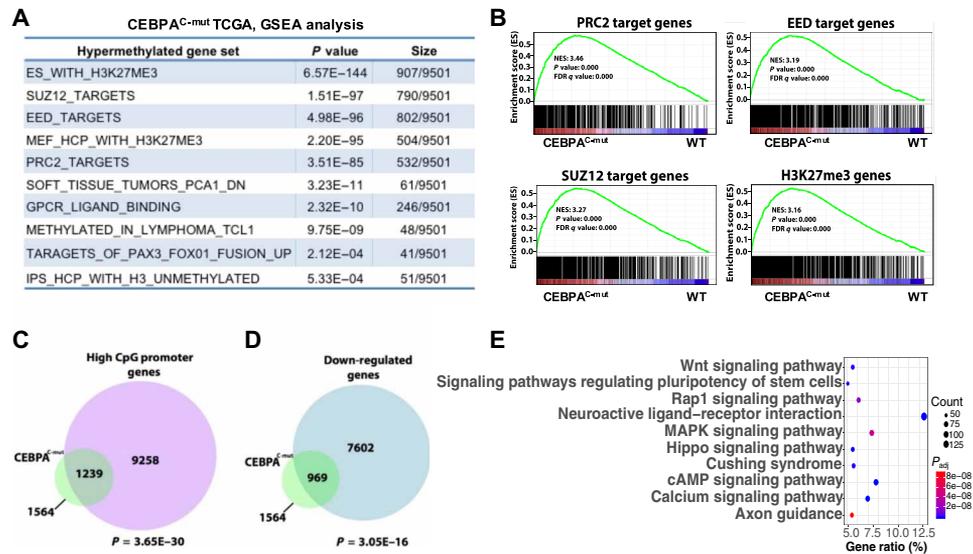


Fig. 5. AML-derived C-terminal mutations in CEBPA are associated with DNA hypermethylation of PRC2 targets. (A) Summary of gene sets enriched in the list of CEBPA C-terminal mutation–associated hypermethylated genes (CEBPA^{C-mut}) from the TCGA dataset. Gene sets were obtained from the Molecular Signatures Database (MSigDB). Hypergeometric test *P* values are indicated, and gene numbers are also illustrated. (B) Enrichment plots from GSEA showing significant enrichment of gene sets associated with CEBPA C-terminal mutations in TCGA AML patients. *P* values were determined using a hypergeometric test. (C) Venn diagram showing overlap of high CpG promoter genes and CEBPA^{C-mut} genes that are hypermethylated PRC2 target genes in TCGA AML patients. Overlap of these gene sets was measured by a hypergeometric test. (D) Venn diagram showing overlap of the CEBPA^{C-mut} genes from the TCGA dataset and genes that are significantly down-regulated in AML. Overlap of these gene sets was measured by a hypergeometric test. (E) Pathway enrichment of CEBPA^{C-mut} genes that are hypermethylated PRC2 target genes derived from the TCGA dataset. cAMP, adenosine 3',5'-monophosphate; MAPK, mitogen-activated protein kinase; NES, normalized enrichment score.

CEBPA-DNMT3A interaction regulates a considerable proportion of PRC2 target genes by modulating their promoter methylation.

Leukemia cells with CEBPA^{N321D} mutation are susceptible to DNA-hypomethylating agents

5-Azacytidine (5-Aza) and 5-aza-2'-deoxycytidine [decitabine (DAC)] are U.S. Food and Drug Administration (FDA)–approved drugs and can reverse DNA hypermethylation patterns and restore the expression of tumor suppressor genes silenced by DNA methylation, thereby exerting antitumoral effects in the clinic for treating myelodysplastic syndrome (MDS) and AML (40, 41). The finding that AML-derived CEBPA mutations were associated with DNA hypermethylation encouraged us to examine the susceptibility of CEBPA-mutated leukemia cells to DNMT inhibitors (DNMTis). As expected, genomic 5mC was reduced by 5-Aza or DAC treatment in K562 cells expressing wild-type CEBPA or CEBPA^{N321D} mutant (fig. S11A). In addition, we found that benzidine-positive cells corresponding to hemoglobin synthesis were increased by 5-Aza or DAC treatment in K562 cells and that erythroid differentiation seemed to be impaired in cells expressing CEBPA^{N321D} mutant as compared to those expressing wild-type CEBPA when treated with either 5-Aza or DAC for 4 days (fig. S11B). K562 cells expressing CEBPA^{N321D} mutant underwent significantly increased apoptosis as compared to those expressing wild-type CEBPA when treated with either 5-Aza or DAC (fig. S11C). The 50% inhibitory concentrations (IC₅₀) of 5-Aza was threefold lower in CEBPA^{N321D}-expressing K562 cells than those expressing wild-type CEBPA (IC₅₀, 0.3 and 1.5 μM, respectively) (fig. S11D). Likewise, the IC₅₀ value of DAC was also significantly lower in CEBPA^{N321D}-expressing K562 cells than those expressing wild-type CEBPA (IC₅₀, 15.4 and >100 μM,

respectively) (fig. S11D). By monitoring cell proliferation over 4 days in both K562 and HL-60 stable cells treated with or without 5-Aza or DAC, we found a more profound reduction in cell proliferation at day 4 in CEBPA^{N321D}-expressing cells compared to those expressing wild-type CEBPA (Fig. 6A). The effect of CEBPA N321D mutation on increasing cellular susceptibility to DNMTis was eliminated in DNMT3A-KO K562 and HL-60 cells (Fig. 6A), suggesting that AML-derived hotspot mutation in CEBPA sensitizes leukemia cells to DNA-hypomethylating agents in a DNMT3A-dependent manner.

Last, we performed in vivo transplantation experiments by establishing K562 cell line–derived xenograft (CDX) model in NOD-Prkdc^{em26Cd52}-Il2rg^{em26Cd22}-Kit^{em1Cin(V831M)}/Gpt (NCG-X) mice (Fig. 6B). Upon doxycycline (DOX) treatment, the ectopic expression of wild-type CEBPA or the N321D mutant was induced without affecting cell proliferation under cultured condition (fig. S12A). At day 14 after CDX, DOX was supplied in drinking water to induce the expression of wild-type CEBPA or the N321D mutant in NCG-X mice (fig. S12B). To verify a correct K562 chimeric ratio, we collected peripheral blood samples and confirmed that the count and percentages of mouse white blood cells (mWBCs) and K562 cells were not affected by DOX treatment (fig. S12, C to E). Moreover, DOX treatment did not change the body weight of mice xenografted with K562 expressing wild-type CEBPA or the N321D mutant (fig. S12F), and animal survival was also not affected (fig. S12, G and H). At day 19 after CDX, xenografted mice received repeated injections of DAC or phosphate-buffered saline (PBS) every 3 days, and blood samples were collected on the date of sacrifice when an individual animal had lost ≥25% of body weight. Phenotypically, we found that WBC counts in xenografted NCG-X mice were around 1 × 10⁵/50 μl

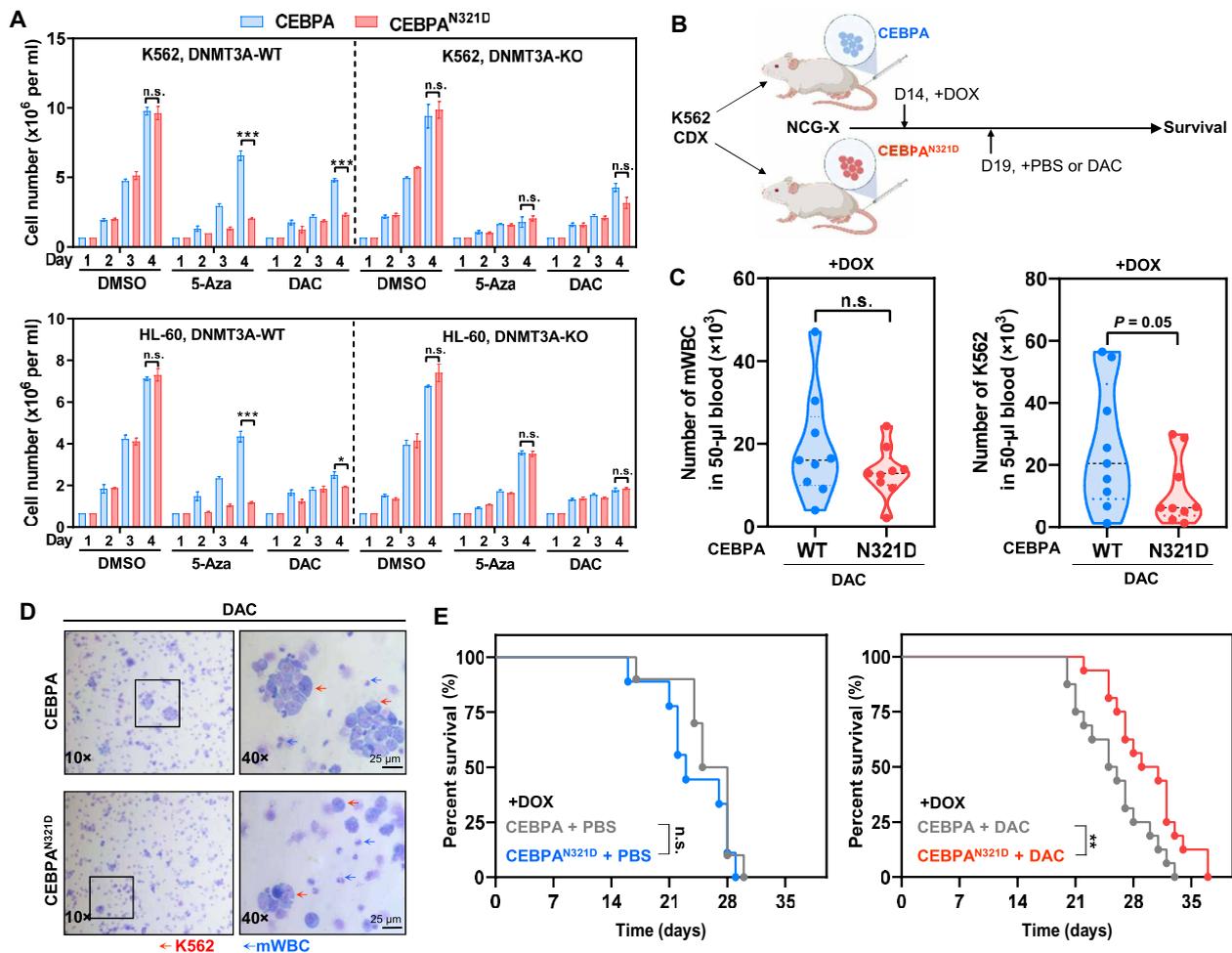


Fig. 6. CEBPA^{N321D} mutation sensitizes leukemia cells to DNMTi. (A) K562 and HL-60 stable cells with knockdown of endogenous CEBPA and reinduction of lentiviral vectors expressing empty vector (VEC) or Flag-tagged wild-type CEBPA or N321D mutant to a comparable level as endogenous CEBPA protein. Moreover, DNMT3A was deleted by CRISPR-Cas9-mediated genome editing technique in these stable cells. For each group, 2×10^6 cells were initially grown in 3 ml of culture medium without or with 5-Aza (1 μ M) or DAC (1 μ M), and the cell numbers were monitored during the period of 4 days by cell counting. (B) Strategy to generate K562 CDX mouse model. See Materials and Methods for more details. (C) Numbers of mWBC and K562 cells in peripheral blood samples of xenografted mice after DAC therapy ($n = 9$ for each group). The data were collected on the date of sacrifice when an individual animal had lost $\geq 25\%$ of body weight, and the results are representative of two independent experiments. (D) Representative data for Giemsa staining in peripheral blood from xenografted mice after DAC therapy (blue arrow, mWBC; red arrow, K562). (E) Kaplan-Meier survival curves and the log-rank test in xenografted mice with K562 overexpressing wild-type CEBPA or CEBPA^{N321D} with DAC therapy ($n = 9$ to 10 and $n = 16$ per group for PBS and DAC treatment, respectively). Average values of triplicated results with SD are shown. Asterisks denote statistical significance with two-tailed unpaired Student's *t* test. * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$ for the indicated comparison.

(fig. S12D) and that the counts were reduced by fivefold to less than $2 \times 10^4/50 \mu\text{l}$ in DAC-treated moribund mice (Fig. 6C). NCG-X mice xenografted with K562 expressing CEBPA^{N321D} exhibited fewer human leukemia cells in peripheral blood than those expressing wild-type CEBPA after DAC treatment (Fig. 6, C and D), and the survival of these mice was significantly ($P < 0.01$) prolonged compared to controls expressing wild-type CEBPA (Fig. 6E). These results provide in vivo evidence to support that the CEBPA^{N321D} mutation sensitizes leukemia cells to DNMTi chemotherapy.

DISCUSSION

DNA methylation is established and maintained by DNMT enzymes, including DNMT3A and DNMT3B, which are responsible for

de novo DNA methylation (42), and DNMT1, which is critical for the maintenance of methylation patterns after DNA replication (43, 44). Like most epigenetic modifying enzymes, DNMT enzymes have little intrinsic DNA sequence specificity, and they bind to many sites across the genome. DNA methylation within the genome is not random, and regional DNA methylation forms a basis for regulation of tissue-specific gene transcription (45). Previous studies have shown that DNMT enzymes require chromatin-remodeling proteins or other factors to gain access to preferred genomic loci. For instance, DNMTs can form complexes with several sequence-specific binding proteins such as PML-RAR (46), Myc/Miz-1 (47), and PU.1 (48) and are delivered to gene promoters for targeted DNA methylation. Here, we show a previously unrecognized mechanism of DNMT regulation—binding with a sequence-specific transcription

factor—by which the de novo DNMT activity of DNMT3A is inhibited at specific sequences in the genome.

Global epigenetic reprogramming during early development sets key developmental genes in a bivalent state marked by both gene-activating H3K4me3 and gene-repressive H3K27me3 modifications. An important observation of this epigenetic regulation is that bivalent histone methylation and DNA methylation are compatible throughout most of the genome, except for the PRC2 target genes enriched with CGIs, where these two marks are mutually exclusive (49). This is believed to poise bivalently marked developmental genes in a temporarily silenced, but not permanently repressed, state, which is subsequently resolved during differentiation into either active or repressive state. It is largely unclear how these epigenetically primed states are established specifically at the promoters of PRC2 target genes. In the current study, we speculate that CEBPA may be a part of this epigenetic reprogramming (fig. S13): CEBPA interacts with DNMT3A on specific DNA and prevents DNMT3A from accessing to and catalyzing CpG methylation, thereby keeping DNMT3A activity in a “poised” state near CEBPA binding sites. Once tumor-associated mutations in the bZIP domain of CEBPA occur to disrupt CEBPA-DNA binding, it relieves DNMT3A inhibition and causes hypermethylation of PRC2 target genes. Given that about 40% of PRC2 target genes are hypermethylated in AML patients with CEBPA mutations, it is possible that additional factors may regulate PRC2 target genes by modulating their promoter CGI methylation. Besides CEBPA, other members of the CEBP family can also form protein complexes with DNMT3A, and these transcription factors are crucial regulators of progenitor differentiation in a range of cell types, including granulocytes, adipocytes, mammary epithelial cells, and granulosa cells, and play a central role in liver metabolism, adipogenesis, hematopoiesis, inflammatory processes, and tumorigenesis (50). Whether CEBP-DNMT3A interaction represents a general mechanism for the regulation of PRC2 target genes remains unknown and needs further investigation. This line of research may provide new insights into the regulation of DNMT and de novo site-specific DNA methylation in a wide range of biological processes.

Aberrant DNA methylation due to DNMT dysfunction is an important molecular event during leukemogenesis. FDA-approved DNMTi have been widely used as the standard of care for treating hematological malignancies. However, only 40 to 50% of patients with MDS or AML demonstrated a positive clinical response to DNMTi chemotherapy (51, 52). In addition, the responses are universally transient with loss of response within 2 years and being associated with very poor prognosis (53, 54). Therefore, there is an unmet medical need for the identification of more biomarkers to precisely predict patients' response to DNA-hypomethylating agents and to provide better guidelines for AML treatment, and this relies on further mechanistic insight into the preferential regulation of DNMT particularly at specific genomic loci. CEBPA mutation in C-terminal bZIP domain is abundant in biallelic mutation and is a strong favorable prognostic factor in de novo AML (55). In this study, we show that AML-associated mutations in the CEBPA C terminus, rather than the N terminus, cause CGI hypermethylation and dysregulation of PRC2 target genes, including those critical for self-renewal and cancer progression. Leukemia cells with the hotspot N321D mutation in CEBPA are more susceptible to DNMTi treatment than those with wild-type CEBPA both under culture conditions and in a mouse xenograft model. In terms of clinical manifestation,

AML with *biallelic* CEBPA mutation is more frequent in children or young adults with favorable prognosis, where DNMTi is not included into the standard treatment at this age (56, 57). Our findings not only shed light on the pathological significance of CEBPA-DNMT3A interaction and reinforce the importance of the bZIP domain of CEBPA in DNMT3A regulation but also suggest that frequently occurring C-terminal mutations of CEBPA may be a useful clinical “signature” and predictor of sensitivity to DNA hypomethylation agents with potential therapeutic implications.

MATERIALS AND METHODS

Antibodies

Antibody specific to CEBPA (Santa Cruz Biotechnology, sc-166258), DNMT3A (Abcam, 2850 for IP), DNMT3A (Hua'an ET1609-31 for Western blot), DNMT3A (Proteintech, 20954-1-AP for immunofluorescent staining), β -actin (Cell Signaling Technology), Flag (Shanghai Genomic), HA (Santa Cruz Biotechnology), or Myc (Hua'an) was purchased commercially. Secondary antibodies for goat anti-mouse immunoglobulin G (IgG) light chain specific (Jackson ImmunoResearch) and mouse anti-rabbit IgG light chain specific (Jackson ImmunoResearch) were also purchased commercially.

Cell culture and transfection

HEK293T, HeLa, HL-60, and K562 cell lines were obtained from the American Type Culture Collection. HEK293T and HeLa cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen), and HL-60 and K562 cells were cultured in 1640 (Gibco), both supplemented with 10% fetal bovine serum (FBS; Gibco) at 37°C with 5% CO₂. Plasmid transfection was carried out by either PEI (polyethylenimine) (1 mg/ml; pH 7.0; Polysciences Inc.) or Lipofectamine 2000 (Invitrogen) according to the manufacturer's instruction.

DNA methylation analysis in AML patients

DNA methylation was analyzed for 194 AML patient samples in the TCGA Research Network (TCGA) dataset. A sample was included in the CEBPA C-terminal mutant group if it has one or more C-terminal mutations on different alleles of *CEBPA*. There were 11 patients with CEBPA C-terminal mutation and 183 patients without CEBPA C-terminal mutation. In a separate analysis, a sample was included in the CEBPA mutant group if it has one or more mutations in CEBPA. There were 13 patients with CEBPA mutation and 181 patients without CEBPA mutation.

Raw IDAT files from Infinium HumanMethylation450 BeadChip array were downloaded and used to obtain raw signal intensities by the RnBeads Bioconductor package (58, 59). The rule-based filtering criteria were applied to the preprocess step. For each probe, beta values were calculated using $(\text{Meth}/[\text{Meth}+\text{Unmeth}])$ and further normalized using the beta-mixture quantile normalization (60). DNA methylation differences were analyzed using the limma method (61), and *P* values were corrected for multiple testing using the false discovery rate (FDR) method. Differential methylated probes were defined as having a change in methylation (absolute delta beta) of >0.1 and an FDR-adjusted *P* value of <0.05. KEGG pathway analysis was conducted using the clusterProfiler Bioconductor package.

Gene set enrichment analysis

The hypermethylated CpG sites in the CEBPA C-terminal mutant or CEBPA mutant group were annotated to genes using the annotation

file provided by Illumina (Gene Expression Omnibus: GPL13534). The genes were ranked according to the strength of the association between the mutation and a CpG site corresponding to the gene. A lower *P* value of differential methylation translated into a higher gene rank.

By using the GSEA software from Broad Institute, the ranked gene list was compared to the gene sets in the gene set collection C2 from Molecular Signatures Database (MSigDB) for enrichment analysis. Analysis was run with 2000 permutations and a hypergeometric (classic) test. Normalized enrichment score and FDR-adjusted *P* values were measured to obtain enrichments with statistical significance. All gene sets that have FDR < 0.05 were considered as significantly enriched.

Gene promoter CpG density analysis

We selected the genes from the top 10 enriched gene sets as CEBPA C-terminal mutant or CEBPA mutant genes and used the set of 10,497 genes that were studied and classified as having high CpG promoter genes (62). The overlap between CEBPA^{C-mut} or CEBPA^{mut} genes and high CpG promoter genes was determined using a hypergeometric test.

RNA-seq analysis in AML patients

For TCGA AML patients, normalized gene quantification data were downloaded to perform differential expression analysis. Comparison between CEBPA C-terminal mutant or CEBPA mutant samples and the remaining samples was performed using *t* test. The overlap between CEBPA^{C-mut} or CEBPA^{mut} genes and down-regulated genes was determined using a hypergeometric test.

Mammalian two-hybrid assay

UAS-luciferase reporter plasmid, cytomegalovirus (CMV)–Renilla control plasmid, human DNA-modifying enzymes fusing with VP16 transactivation domain, and CEBPA fusing with Gal4-DBD were cotransfected into HEK293T cells. After transfection for 30 hours, the luciferase reporter activity was detected using a commercial kit (Promega, E1910) with a Turner BioSystems luminometer reader (Promega). The individual luciferase raw data are shown in table S3.

Generation of stable cells

To generate CEBPA knockdown and rescue stable cells, the lentivirus PLKO.1-shCEBPA and PCDH-Flag-CEBPA/CEBPA^{N321D} plasmids were cotransfected into HEK293T cells with packing plasmids PSPAX2 and PMD2G at a 4:3:1 ratio. After 48 hours, virus was collected to infect HL-60 and K562 cells for another 24 hours. Polybrene (8 µg/ml) was mixed to increase infection efficiency. Infected cells were selected in puromycin (2 µg/ml) for 1 week. The short hairpin RNA (shRNA) sequences targeting CEBPA are 5'-GTAGAAGTCGGCCGACTCCA-3' and 5'-CAGTTCAGATC-GCGCACTG-3'.

To generate DNMT3A KO K562 or HL-60 cells, the lentiCRISPR (Addgene, plasmid #52961) DNMT3A KO plasmids were cotransfected into HEK293T cells with packing plasmids PSPAX2 and PMD2G at a 4:3:1 ratio. After 48 hours, virus was collected to infect HL-60 and K562 cells for another 24 hours. Polybrene (8 µg/ml) was mixed to increase infection efficiency. Infected cells were selected in puromycin (2 µg/ml) for 1 week. The single-guide RNA sequences used for targeting DNMT3A were 5'-GCATGATGCGCGCCCAAGG-3'

and 5'-GCAGAGGACGAGCCGGGACG-3'. After that, single cells were sorted into 96-well plates by using fluorescence-activated cell sorting, and KO clones were validated by Western blot analysis.

Co-IP and Western blot

Cells were lysed in 1 ml of NP-40 buffer containing 50 mM tris-HCl (pH 7.5), 500 mM NaCl, 0.3% NP-40, and protease inhibitor cocktail (Roche) with rotation at 4°C for 60 min and then centrifuged at 12,000 rpm for 15 min at 4°C. IP was carried out either by incubating Flag beads (Sigma-Aldrich) at 4°C with lysate for 3 hours or by incubating cell lysate with the DNMT3A primary antibody for 1 hour, followed by incubating with protein A/G beads (GE) for another 3 hours at 4°C. Then, samples were harvested by washing beads three times and boiled in water bath for 10 min.

Western blotting was performed after loading the samples into 10% SDS–polyacrylamide gel electrophoresis (PAGE) gel and running for about 1 hour at 120 V. The membrane was then blocked with 5% fat-free milk (BD Biosciences). After incubation with primary antibodies overnight at 4°C and secondary antibodies for 1 hour at room temperature, the membrane was washed three times and analyzed by Typhoon FLA 9500 image scanning (GE Healthcare).

Immunofluorescence assay

Cells were fixed with 4% phenylmethylsulfonyl fluoride (Sangon) overnight at 4°C and then permeabilized with 0.4% Triton X-100 for 15 to 30 min at room temperature. Next, the cells were incubated with blocking buffer [3% bovine serum albumin (BSA) in PBS; Sigma-Aldrich] for 1 hour at 4°C, followed by incubation with primary antibodies against HA [dilution at 1:500; Santa Cruz Biotechnology] and Flag [dilution at 1:500; MCB (The Molecular and Cell Biology Research Lab)] overnight at 4°C, and then the Alexa Fluor 594 (red) and 488 (green)–conjugated secondary antibodies (Invitrogen) for 1 hour at room temperature. Cell nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI; Invitrogen). Images were captured using a Leica fluorescence microscope.

Proximity ligation assay

Briefly, 2×10^5 cells were harvested and washed with prechilled PBS. Cells were then fixed and permeabilized for 10 min using 80% prechilled methanol and centrifuged to remove the methanol at 500g for 5 min at 4°C. Afterward, the cells were blocked in blocking solution (PBS and 1% BSA) and incubated with the anti-DNMT3A and anti-CEBPA primary antibodies (dilution at 1:100 in blocking solution). The next steps of the assay were performed according to the manufacturer's instruction (DUO92101; Sigma-Aldrich). Cell nuclei were stained with DAPI (Invitrogen). Images were captured using a Leica fluorescence microscope.

Protein purification

DNMT3A protein was purified from HEK293T cells at 24 hours after transfection, and cells were harvested and lysed in buffer containing 50 mM tris-Cl (pH 8.0), 500 mM NaCl, and 0.01% 2-mercaptoethanol. The fusion protein was immunoprecipitated by Flag beads and then cleaved by PreScission protease. Moreover, DNMT3A N terminus (residues 1 to 291) and D529A mutant and wild-type and N321D mutant CEBPA proteins were purified from *Escherichia coli* strain BL21 (DE3). The transformed *E. coli* cells were grown at 37°C to reach an OD₆₀₀ (optical density at 600 nm) of

0.6 to 0.8 and were then induced with 1 mM isopropyl- β -D-thiogalactopyranoside at 16°C for 20 hours. The cell lysate after mechanical disruption and centrifugation was applied for glutathione S-transferase purification. Then, the elution was digested with PreScission protease. The recombinant proteins were stored at -80°C until further analysis.

Electrophoretic mobility shift assay

Briefly, FAM-labeled DNA (20 nM) and indicated amounts of purified proteins were incubated in a 20- μl reaction system [20 mM Hepes, 100 mM KCl, 5% glycerol, and BSA (0.5 mg/ml) (pH 7.5)] for 45 min at room temperature. The samples were subjected to a 6 to 18% gradient PAGE and analyzed by Typhoon FLA 9500 image scanning (GE Healthcare).

The FAM-labeled substrate DNA was commercially synthesized. Random DNA substrates are as follows: 5'-6-FAM/TGGA-TATCTAGGGCGCTATGATATCT-3' (forward) and 5'-AGA-TATCATAGCGCCCTAGATATCCA-3' (reverse). The CEBPA motif (GCAAT)-containing DNA substrates are as follows: 5'-6-FAM/CTAGGGCTTGCGCAATCTATATTCG-3' (forward) and 5'-CGAATATAGATTGCGCAAGCCCTA-3' (reverse).

DNA affinity and fluorescence polarization assay

FAM-labeled double-stranded DNA (10 nM) substrate was incubated with increased amounts of DNMT3A with wild-type or N321D mutant CEBPA proteins for 45 min at room temperature in the same reaction buffer as EMSA assay. Fluorescence polarization was measured with a Synergy 4 microplate reader (BioTek) at room temperature. Each reaction was performed in triplicate. The curves were fitted using GraphPad Prism 7.

In vitro DNMT3A activity assay

Briefly, biotin-labeled DNA was commercially synthesized as a substrate. Proteins of 0.5 mM DNMT3A or DNMT3A D529A/DNMT3A2/DNMT3A-CD and 0.5 mM DNMT3L were preincubated with the same amount of CEBPA or CEBPA^{N321D} protein on ice for 30 min. Then, the mixed proteins were added to the reaction buffer, which contains substrate DNA (100 ng), 2.5 mM [methyl-³H] SAM (NET155V), 5% glycerol, 0.01% 2-mercaptoethanol, BSA (0.5 mg/ml), 1 μg of H3K4me0 peptide, and 25 mM Tris-HCl (pH 7.5). The reactions were incubated in 37°C water bath for 30 min. Next, the reaction was stopped by adding 500 μl of cold wash buffer (500 mM NaCl and 1 mM EDTA in PBST). The DNA was collected by streptavidin beads, washed three times with cold washing buffer, and then subjected to liquid scintillation counting (PerkinElmer). Each reaction was performed in triplicate.

Dot-blot assay

DNA dot-blot assay was performed as described previously with some modifications (31). Briefly, genomic DNA was spotted on a nitrocellulose membrane (Whatman). The membrane was placed under an ultraviolet lamp for 20 min to cross-link DNA, and then blocked with 5% milk in Tris-buffered saline (TBS)-Tween 20 for 1 hour, followed by incubation with the anti-5mC antibody (Active Motif) at 4°C overnight. After incubation with a horseradish peroxidase-conjugated secondary antibody (GenScript) for 1 hour at room temperature, the membrane was washed with TBS-Tween 20 three times and then scanned with a Typhoon scanner (GE

Healthcare), and the quantification of 5mC was done by ImageQuant software (GE Healthcare).

LC-MS/MS analysis

Genomic DNA from cells was digested with DNA Degradase Plus (Zymo Research) at 37°C for 5 hours. The digested samples were then subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis using a Shimadzu LC (LC-20AB pump) system coupled with TSQ-vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific). A C18 column (250 mm by 2.1 mm inside diameter, 3 μm particle size, ULTIMATE) was used. The mass spectrometer was optimized and set up in selected reaction monitoring scan mode for monitoring the [M + H]⁺ of 5mC [mass/charge ratio (*m/z*), 242.1 \rightarrow 126.1] and deoxyguanosine (dG; *m/z*, 268.1 \rightarrow 152.3). The standard nucleosides of dG and 5mC were commercially purchased (Thermo Fisher Scientific). The Analyst Software Xcalibur 2.2 sp1 was used for analysis, and global 5mC level was quantified as 5mC/dG.

ChIP-qPCR assays

Briefly, 5×10^6 cells were collected and cross-linked with 1% paraformaldehyde for 10 min and immediately stopped by adding 0.125 M glycine. After sonication at 4°C for 25 min (Bioruptor, 90% power, 15 s on, 45 s off), chromatin was immunoprecipitated at 4°C for 3 hours with the antibody against Flag (Sigma-Aldrich) and DNMT3A (Abcam) and for another 2 hours with the protein A + G beads (GE). The binding DNA was washed with 3 \times high-salt buffer, 2 \times low-salt buffers, and 1 \times TE (Tris-EDTA) buffer. The product DNA was de-cross-linked and collected with the PCR Purification Kit (QIAGEN) for further analysis by real-time qPCR. Primer sequences are listed in table S4.

BS-PCR assay

BS-PCR experiments were performed with a commercial EZ DNA Methylation-Gold kit following the manufacturer's instruction. Briefly, 500 ng of genomic DNA extracted from cells was used for BS conversion and PCR amplification. The target DNA products were inserted to a T vector, and 10 clones from each sample were randomly picked for sequencing. Primer sequences are listed in table S4.

RNA isolation and qRT-PCR analysis

Total RNA was extracted from cultured cells by TRIzol reagent (Invitrogen) following the manufacturer's instruction. RNA was reversely transcribed with oligo(dT) primers. Diluted complementary DNA was then used for real-time PCR with gene-specific primers in the presence of TB Green Advantage qPCR Premix (Takara) by using a QuantStudio 6 Flex Real-Time PCR system (Applied Biosystems). β -Actin was used as a housekeeping control. Primer sequences are listed in table S4.

Benzidine staining

K562 cells were treated with hypomethylating agents or hemin (25 μM) for 96 hours. Cells (1×10^6) were harvested and washed twice with prechilled PBS, and suspended cells (in 50 μl) were incubated with 10 μl of freshly prepared benzidine solution (2 mg/ml in glacial acetic acid containing 0.59% H₂O₂) for 10 min at room temperature. The stained cells were imaged with a microscope, and the percentages of benzidine-positive cells were calculated.

Cell proliferation, apoptosis, and viability assays

Cells were seeded in six-well plates in RPMI 1640 medium at the density of 5×10^4 cells per well and then treated with hypomethylating agents with indicated concentrations for indicated time periods. Cell numbers were quantified with an automated counter (Costar). In addition, cells were also collected and stained with the FITC Apoptosis Detection Kit (BD Biosciences) following the manufacturer's instruction. The stained cells were detected by BD Accuri C6 to calculate the percentage of apoptotic cells (annexin V positive) and viable cells (both propidium iodide and annexin V negative). Each experiment was performed in triplicate.

CDX model

Severe immunodeficient NCG-X mice were purchased from GemPharmatech. Male animals were housed in individual ventilated cages under specific pathogen-free condition and used for experiments at 8 to 10 weeks of ages. To generate the CDX model, 6×10^6 K562 cells expressing Ptripz-CEBPA^{N321D} or Ptripz-CEBPA were suspended in IMDM (Iscove Modified Dulbecco Media) (Sangon) containing 1% penicillin and streptomycin and were then intravenously injected into NCG-X mice via tail vein. At day 14 after injection, mice were continuously treated with DOX [2 mg/ml; MCE (MedChemExpress)] in drinking water to induce the expression of wild-type CEBPA or the N321D mutant until the end of experiment. Afterward, DOX-treated mice were intraperitoneally given PBS or DAC (0.4 mg/kg per 36 hours; Selleck). The body weight of mice was measured twice a week from days 0 to 19 and daily since day 20. Upon sacrifice, mice were euthanized, and peripheral blood was collected for further analysis. The Institutional Animal Care and Use Committee at Model Animal Research Center of Nanjing University approved all animal procedures used in this study (AP#: LY-01), which complied with all relevant ethical regulations.

Flow cytometry analysis

Peripheral blood of mice was collected in an anticoagulant tube and stored at 4°C. To verify proper K562 chimeric ratio, 5 μ l of 123count eBeads counting beads (Invitrogen) was added into 50 μ l of peripheral blood, which was then incubated with anti-mouse CD45-VioGreen (Miltenyi, clone REA737) and anti-human CD45-phycoerythrin-Cy7 (BioLegend, clone 2D1) for 30 min at 4°C in darkness. After that, FACS Lysing Solution (BD Biosciences) was used according to the manufacturer's instruction. Cells were washed and suspended with PBS containing 2% FBS, 0.5 mM EDTA, and 1% penicillin and streptomycin. Fortessa (BD Biosciences) and Penton (Agilent) were used for acquisition, and FlowJo v10.6.2 (TreeStar) was applied for data analysis.

Giemsa staining

Mouse peripheral blood samples were lysed by 1 \times red blood cell lysis buffer (eBioscience), cytospinned onto microscope adhesion slides, and then stained with a rapid Giemsa staining kit (Sangon) according to the manufacturer's instruction. Images of cells were captured with a stereoscopic SOPTOP IC X 41 microscope.

Statistics

Two-tailed unpaired Student's *t* test and hypergeometric test were applied for statistical analyses as indicated in the figure legends. With the exception of publicly available large population datasets, all other data were obtained from triplicated independent experiments,

with SD of the mean (mean \pm SD) displayed. The value of *P* < 0.05 was considered to be statistically significant.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abl5220>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- M. V. C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **20**, 590–607 (2019).
- K. D. Robertson, DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610 (2005).
- D. M. Jeziorska, R. J. S. Murray, M. de Gobbi, R. Gaentzsch, D. Garrick, H. Ayyub, T. Chen, E. Li, J. Telenius, M. Lynch, B. Graham, A. J. H. Smith, J. N. Lund, J. R. Hughes, D. R. Higgs, C. Tufarelli, DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7526–E7535 (2017).
- M. Okano, D. W. Bell, D. A. Haber, E. Li, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
- P. Nowialis, K. Lopusna, J. Opavska, S. L. Haney, A. Abraham, P. Sheng, A. Riva, A. Natarajan, O. Guryanova, M. Simpson, R. Hlady, M. Xie, R. Opavsky, Catalytically inactive Dnmt3b rescues mouse embryonic development by accessory and repressive functions. *Nat. Commun.* **10**, 4374 (2019).
- M. Kaneda, M. Okano, K. Hata, T. Sado, N. Tsujimoto, E. Li, H. Sasaki, Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900–903 (2004).
- M. Kaneda, R. Hirasawa, H. Chiba, M. Okano, E. Li, H. Sasaki, Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. *Genes Cells* **15**, 169–179 (2010).
- D. You, E. Nilsson, D. E. Tenen, A. Lyubetskaya, J. C. Lo, R. Jiang, J. Deng, B. A. Dawes, A. Vaag, C. Ling, E. D. Rosen, S. Kang, Dnmt3a is an epigenetic mediator of adipose insulin resistance. *eLife* **6**, e30766 (2017).
- K. Nishikawa, Y. Iwamoto, Y. Kobayashi, F. Katsuoka, S.-I. Kawaguchi, T. Tsujita, T. Nakamura, S. Kato, M. Yamamoto, H. Takayanagi, M. Ishii, DNA methyltransferase 3a regulates osteoclast differentiation by coupling to an S-adenosylmethionine-producing metabolic pathway. *Nat. Med.* **21**, 281–287 (2015).
- M. Naito, M. Mori, M. Inagawa, K. Miyata, N. Hashimoto, S. Tanaka, H. Asahara, Dnmt3a regulates proliferation of muscle satellite cells via p57Kip2. *PLoS Genet.* **12**, e1006167 (2016).
- Y. Hatazawa, Y. Ono, Y. Hirose, S. Kanai, N.-L. Fujii, S. Machida, I. Nishino, T. Shimizu, M. Okano, Y. Kamei, Y. Ogawa, Reduced Dnmt3a increases Gdf5 expression with suppressed satellite cell differentiation and impaired skeletal muscle regeneration. *FASEB J.* **32**, 1452–1467 (2018).
- B. H. Ladle, K.-P. Li, M. J. Phillips, A. B. Pucsek, A. Haile, J. D. Powell, E. M. Jaffee, D. A. Hildeman, C. J. Gamper, De novo DNA methylation by DNA methyltransferase 3a controls early effector CD8+ T-cell fate decisions following activation. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10631–10636 (2016).
- X. Li, Q. Zhang, Y. Ding, Y. Liu, D. Zhao, K. Zhao, Q. Shen, X. Liu, X. Zhu, N. Li, Z. Cheng, G. Fan, Q. Wang, X. Cao, Methyltransferase Dnmt3a upregulates HDAC9 to deacetylate the kinase TBK1 for activation of antiviral innate immunity. *Nat. Immunol.* **17**, 806–815 (2016).
- T. Chen, Y. Ueda, S. Xie, E. Li, A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.* **277**, 38746–38754 (2002).
- T. Chen, Y. Ueda, J. E. Dodge, Z. Wang, E. Li, Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**, 5594–5605 (2003).
- M. Manzo, J. Wirz, C. Ambrosi, R. Villaseñor, B. Roschitzki, T. Baubec, Isoform-specific localization of DNMT3A regulates DNA methylation fidelity at bivalent CpG islands. *EMBO J.* **36**, 3421–3434 (2017).
- Y. Zeng, R. Ren, G. Kaur, S. Hardikar, Z. Ying, L. Babcock, E. Gupta, X. Zhang, T. Chen, X. Cheng, The inactive Dnmt3b3 isoform preferentially enhances Dnmt3b-mediated DNA methylation. *Genes Dev.* **34**, 1546–1558 (2020).
- S. La Salle, J. M. Trasler, Dynamic expression of DNMT3a and DNMT3b isoforms during male germ cell development in the mouse. *Dev. Biol.* **296**, 71–82 (2006).
- S. Epsztejn-Litman, N. Feldman, M. Abu-Remalch, Y. Shufaro, A. Gerson, J. Ueda, R. Deplus, F. Fuks, Y. Shinkai, H. Cedar, Y. Bergman, De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).

20. A. G. Kotini, A. Mpakali, T. Agaloti, Dnmt3a1 upregulates transcription of distinct genes and targets chromosomal gene clusters for epigenetic silencing in mouse embryonic stem cells. *Mol. Cell. Biol.* **31**, 1577–1592 (2011).
21. H. S. Radomska, C. S. Huettner, P. Zhang, T. Cheng, D. T. Scadden, D. G. Tenen, CCAAT/enhancer binding protein alpha is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors. *Mol. Cell. Biol.* **18**, 4301–4314 (1998).
22. F. T. Lin, O. A. MacDougald, A. M. Diehl, M. D. Lane, A 30-kDa alternative translation product of the CCAAT/enhancer binding protein alpha message: Transcriptional activator lacking antimetabolic activity. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9606–9610 (1993).
23. T. Pabst, B. U. Mueller, P. Zhang, H. S. Radomska, S. Narravula, S. Schnittger, G. Behre, W. Hiddemann, D. G. Tenen, Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat. Genet.* **27**, 263–270 (2001).
24. M. Miller, J. D. Shuman, T. Sebastian, Z. Dauter, P. F. Johnson, Structural basis for DNA recognition by the basic region leucine zipper transcription factor CCAAT/enhancer-binding protein alpha. *J. Biol. Chem.* **278**, 15178–15184 (2003).
25. X.-M. Wen, J. Lin, J. Yang, D.-M. Yao, Z.-Q. Deng, C.-Y. Tang, G.-F. Xiao, L. Yang, J.-C. Ma, J.-B. Hu, W. Qian, J. Qian, Double CEBPA mutations are prognostically favorable in non-M3 acute myeloid leukemia patients with wild-type NPM1 and FLT3-ITD. *Int. J. Clin. Exp. Pathol.* **7**, 6832–6840 (2014).
26. K. Togami, J. Kitaura, T. Uchida, D. Inoue, K. Nishimura, K. C. Kawabata, R. Nagase, S. Horikawa, K. Izawa, T. Fukuyama, F. Nakahara, T. Oki, Y. Harada, H. Harada, H. Aburatani, T. Kitamura, A C-terminal mutant of CCAAT-environment-binding protein alpha (C/EBPalpha-Cm) downregulates Csf1r, a potent accelerator in the progression of acute myeloid leukemia with C/EBPalpha-Cm. *Exp. Hematol.* **43**, 300–308.e1 (2015).
27. P. A. Jones, S. B. Baylin, The epigenomics of cancer. *Cell* **128**, 683–692 (2007).
28. A. P. Feinberg, B. Tycko, The history of cancer epigenetics. *Nat. Rev. Cancer* **4**, 143–153 (2004).
29. S. Sinha, D. Thomas, L. Yu, A. J. Gentles, N. Jung, M. R. Corces-Zimmerman, S. M. Chan, A. Reinisch, A. P. Feinberg, D. L. Dill, R. Majeti, Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition. *Blood* **125**, 316–326 (2015).
30. R. Chowdhury, K. K. Yeoh, Y. M. Tian, L. Hillringhaus, E. A. Bagg, N. R. Rose, I. K. H. Leung, X. S. Li, E. C. Y. Woon, M. Yang, M. A. McDonough, O. N. King, I. J. Clifton, R. J. Klose, T. D. W. Claridge, P. J. Ratcliffe, C. J. Schofield, A. Kawamura, The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep.* **12**, 463–469 (2011).
31. W. Xu, H. Yang, Y. Liu, Y. Yang, P. Wang, S.-H. Kim, S. Ito, C. Yang, P. Wang, M.-T. Xiao, L.-X. Liu, W.-Q. Jiang, J. Liu, J.-Y. Zhang, B. Wang, S. Frye, Y. Zhang, Y.-H. Xu, Q.-Y. Lei, K.-L. Guan, S.-M. Zhao, Y. Xiong, Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).
32. Y. Wang, M. Xiao, X. Chen, L. Chen, Y. Xu, L. Lv, P. Wang, H. Yang, S. Ma, H. Lin, B. Jiao, R. Ren, D. Ye, K.-L. Guan, Y. Xiong, WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol. Cell* **57**, 662–673 (2015).
33. R. Rampal, A. Alkalin, J. Madzo, A. Vasanthakumar, E. Pronier, J. Patel, Y. Li, J. Ahn, O. Abdel-Wahab, A. Shih, C. Lu, P. S. Ward, J. T. Tsai, T. Hricik, V. Tosello, J. E. Tallman, X. Zhao, D. Daniels, Q. Dai, L. Ciminio, I. Aifantis, C. He, F. Fuks, M. S. Tallman, A. Ferrando, S. Nimer, E. Paietta, C. B. Thompson, J. D. Licht, C. E. Mason, L. A. Godley, A. Melnick, M. E. Figueroa, R. L. Levine, DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep.* **9**, 1841–1855 (2014).
34. I. Suetake, Y. Mishima, H. Kimura, Y.-H. Lee, Y. Goto, H. Takeshima, T. Ikegami, S. Tajima, Characterization of DNA-binding activity in the N-terminal domain of the DNA methyltransferase Dnmt3a. *Biochem. J.* **437**, 141–148 (2011).
35. S. Jeong, G. Liang, S. Sharma, J. C. Lin, S. H. Choi, H. Han, C. B. Yoo, G. Egger, A. S. Yang, P. A. Jones, Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol. Cell. Biol.* **29**, 5366–5376 (2009).
36. X. Guo, L. Wang, J. Li, Z. Ding, J. Xiao, X. Yin, S. He, P. Shi, L. Dong, G. Li, C. Tian, J. Wang, Y. Cong, Y. Xu, Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**, 640–644 (2015).
37. D. Jia, R. Z. Jurkowska, X. Zhang, A. Jeltsch, X. Cheng, Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449**, 248–251 (2007).
38. Z. M. Zhang, R. Lu, P. Wang, Y. Yu, D. Chen, L. Gao, S. Liu, D. Ji, S. B. Rothbart, Y. Wang, G. G. Wang, J. Song, Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* **554**, 387–391 (2018).
39. C. M. Virbasius, S. Wagner, M. R. Green, A human nuclear-localized chaperone that regulates dimerization, DNA binding, and transcriptional activity of bZIP proteins. *Mol. Cell* **4**, 219–228 (1999).
40. A. Koch, S. C. Joosten, Z. Feng, T. C. de Ruijter, M. X. Draht, V. Melotte, K. M. Smits, J. Veeck, J. G. Herman, L. van Neste, W. van Criekinge, T. de Meyer, M. van Engeland, Analysis of DNA methylation in cancer: Location revisited. *Nat. Rev. Clin. Oncol.* **15**, 459–466 (2018).
41. C. B. Yoo, P. A. Jones, Epigenetic therapy of cancer: Past, present and future. *Nat. Rev. Drug Discov.* **5**, 37–50 (2006).
42. M. Okano, S. Xie, E. Li, Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19**, 219–220 (1998).
43. T. H. Bestor, Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.* **11**, 2611–2617 (1992).
44. E. Li, T. H. Bestor, R. Jaenisch, Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
45. R. Illingworth, A. Kerr, D. DeSousa, H. Jorgensen, P. Ellis, J. Stalker, D. Jackson, C. Clee, R. Plumb, J. Rogers, S. Humphray, T. Cox, C. Langford, A. Bird, A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22 (2008).
46. L. D. Croce, V. A. Raker, M. Corsaro, F. Fazi, M. Fanelli, M. Faretta, F. Fuks, F. L. Coco, T. Kouzarides, C. Nervi, S. Minucci, P. G. Pelicci, Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**, 1079–1082 (2002).
47. C. Brenner, R. Deplus, C. Didelot, A. Loriot, E. Viré, C. de Smet, A. Gutierrez, D. Danovi, D. Bernard, T. Boon, P. Giuseppe Pelicci, B. Amati, T. Kouzarides, Y. de Launoit, L. di Croce, F. Fuks, Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.* **24**, 336–346 (2005).
48. M. Suzuki, T. Yamada, F. Kihara-Negishi, T. Sakurai, E. Hara, D. G. Tenen, N. Hozumi, T. Oikawa, Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. *Oncogene* **25**, 2477–2488 (2006).
49. A. B. Brinkman, H. Gu, S. J. J. Bartels, Y. Zhang, F. Matarese, F. Simmer, H. Marks, C. Bock, A. Gnirke, A. Meissner, H. G. Stunnenberg, Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138 (2012).
50. C. Nerlov, The C/EBP family of transcription factors: A paradigm for interaction between gene expression and proliferation control. *Trends Cell Biol.* **17**, 318–324 (2007).
51. Y. G. Lee, I. Kim, S.-S. Yoon, S. Park, J. W. Cheong, Y. H. Min, J.-O. Lee, S.-M. Bang, H. G. Yi, C. S. Kim, Y. Park, B.-S. Kim, Y.-C. Mun, C.-M. Seong, J. Park, J. H. Lee, S.-Y. Kim, H. G. Lee, Y.-K. Kim, H.-J. Kim; for the Korean Society of Haematology AML/MDS working party, Comparative analysis between azacitidine and decitabine for the treatment of myelodysplastic syndromes. *Br. J. Haematol.* **161**, 339–347 (2013).
52. J. Schecter, N. Galili, A. Raza, MDS: Refining existing therapy through improved biologic insights. *Blood Rev.* **26**, 73–80 (2012).
53. E. Jabbour, G. Garcia-Manero, N. Batty, J. Shan, S. O'Brien, J. Cortes, F. Ravandi, J. P. Issa, H. Kantarjian, Outcome of patients with myelodysplastic syndrome after failure of decitabine therapy. *Cancer* **116**, 3830–3834 (2010).
54. A. Gil-Perez, G. Montalban-Bravo, Management of myelodysplastic syndromes after failure of response to hypomethylating agents. *Ther. Adv. Hematol.* **10**, 2040620719847059 (2019).
55. S. Wakita, M. Sakaguchi, I. Oh, S. Kako, T. Toya, Y. Najima, N. Doki, J. Kanda, J. Kuroda, S. Mori, A. Satake, K. Usuki, T. Ueki, N. Uoshima, Y. Kobayashi, E. Kawata, K. Tajika, Y. Nagao, K. Shono, M. Shibusawa, J. Tadokoro, K. Kayamori, M. Hagihara, H. Uchiyama, N. Uchida, Y. Kubota, S. Kimura, H. Nagoshi, T. Ichinohe, S. Kurosawa, S. Motomura, A. Hashimoto, H. Muto, E. Sato, M. Ogata, K. Mitsushashi, J. Ando, A. Marumo, I. Omori, Y. Fujiwara, K. Terada, S. Yui, K. Arai, T. Kitano, M. Miyata, A. Kurosawa, A. Mizoguchi, N. Komatsu, T. Fukuda, K. Ohashi, Y. Kanda, K. Inokuchi, H. Yamaguchi, Prognostic impact of CEBPA bZIP domain mutation in acute myeloid leukemia. *Blood Adv.* **6**, 238–247 (2021).
56. I. De Kouchkovsky, M. Abdul-Hay, Acute myeloid leukemia: A comprehensive review and 2016 update. *Blood Cancer J.* **6**, e441 (2016).
57. D. A. Arber, The 2016 WHO classification of acute myeloid leukemia: What the practicing clinician needs to know. *Semin. Hematol.* **56**, 90–95 (2019).
58. F. Müller, M. Scherer, Y. Assenov, P. Lutsik, J. Walter, T. Lengauer, C. Bock, RnBeads 2.0: Comprehensive analysis of DNA methylation data. *Genome Biol.* **20**, 55 (2019).
59. Y. Assenov, F. Müller, P. Lutsik, J. Walter, T. Lengauer, C. Bock, Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* **11**, 1138–1140 (2014).
60. A. E. Teschendorff, F. Marabita, M. Lechner, T. Bartlett, J. Tegner, D. Gomez-Cabrero, S. Beck, A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189–196 (2013).
61. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, *limma* powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
62. S. Saxonov, P. Berg, D. L. Brutlag, A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1412–1417 (2006).

Acknowledgments: We thank members of the Molecular and Cell Biology laboratory of Fudan University for discussions and support throughout this study. We also thank the staff members working at IBS, Fudan University for help on MS analyses. We acknowledge the TCGA for the publicly available data, and we acknowledge G.-L. Xu and A. Lu for providing [methyl-3H] SAM (NET155V), M. A. Goodell for providing Bio-Dnmt3a1 ES cells, and L.-r. Wang

for generating CEBPAN321D knock-in cells in K562. **Funding:** This study was supported by the National Key R&D Program of China (2020YFA0803202 and 2016YFA0501800 to D.Y.; no. 2019YFA0802900 to Y. Li), the NSFC (nos. 31871431 and 31821002 to D.Y.; no. 81870081 to T.C.; no. 32070042 to Y. Li), Shanghai Collaborative Innovation Program on Regenerative Medicine and Stem Cell Research of Shanghai Municipal Health Commission (2019CXJQ01 to T.C.), the Innovative research team of high-level local universities in Shanghai (SSMU-ZLXC20180501 to D.Y.), and Program for Innovative Talents and Entrepreneur in Jiangsu (to Y.L.). J.J.-L.W. is a fellow of the Cancer Institute of NSW Australia. **Author contributions:** Y. Xiong and D.Y. conceived and supervised the project. X.C. and W.Z. performed most of the experiments and wrote the draft with D.Y. R.-H.S., J.J.-L.W., and Y.L. (Yun Liu) conducted bioinformatic analysis of methylation and RNA-seq data from the TCGA database. L.Z. and C.Z. performed LC-MS/MS analysis to quantify genomic 5mC. J.X. and M.-H. Li assisted in in vitro DNMT3A activity assay.

S.L. generated K562 CDX model and DNMTi treatment under the supervision of Y.L. (Yan Li). S.W., Y.C., C.H., C.G., T.W., P.L., K.D., and Z.C. assisted in cell culture and Western-IP experiments. J.Z. and Y.S. provided and managed the experimental materials and equipment. T.C., H.Y., P.W., Y. Xu, and F.L. guided some critical assays. F.J. helped with English grammar checking of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 21 July 2021
Accepted 4 December 2021
Published 26 January 2022
10.1126/sciadv.abl5220