nature cell biology

NUP98–NSD1 links H3K36 methylation to *Hox-A* gene activation and leukaemogenesis

Gang G. Wang[1,2], Ling Cai[3], Martina P. Pasillas[1] and Mark P. Kamps[1,2,4]

Nuclear receptor-binding SET domain protein 1 (NSD1) prototype is a family of mammalian histone methyltransferases (NSD1, NSD2/MMSET/WHSC1, NSD3/WHSC1L1) that are essential in development and are mutated in human acute myeloid leukemia (AML)[1,2], overgrowth syndromes[3], multiple myeloma[4] and lung cancers[5]. In AML, the recurring t(5;11)(q35;p15.5) translocation fuses NSD1 to nucleoporin-98 (NUP98)[6]. Here, we present the first characterization of the transforming properties and molecular mechanisms of NUP98–NSD1. We demonstrate that NUP98–NSD1 induces AML *in vivo,* sustains self-renewal of myeloid stem cells *in vitro,* and enforces expression of the *HoxA7, HoxA9, HoxA10* and *Meis1* proto-oncogenes. Mechanistically, NUP98–NSD1 binds genomic elements adjacent to *HoxA7* and *HoxA9*, maintains histone H3 Lys 36 (H3K36) methylation and histone acetylation, and prevents EZH2-mediated transcriptional repression of the *Hox-A* locus during differentiation. Deletion of the NUP98 FG-repeat domain, or mutations in NSD1 that inactivate the H3K36 methyltransferase activity or that prevent binding of NUP98–NSD1 to the *Hox-A* locus precluded both *Hox-A* gene activation and myeloid progenitor immortalization. We propose that NUP98–NSD1 prevents EZH2-mediated repression of *Hox-A* locus genes by colocalizing H3K36 methylation and histone acetylation at regulatory DNA elements. This report is the first to link deregulated H3K36 methylation to tumorigenesis and to link NSD1 to transcriptional regulation of the *Hox-A* locus.

The epigenetic activation of mammalian *Hox* genes by trithorax proteins and their silencing by polycomb proteins control important aspects of differentiation and proliferation during embryogenesis. In haematopoiesis, expression of *HoxA7, HoxA9* and *HoxA10* promote stem-cell self-renewal, and the downregulation of these genes coincides with terminal differentiation[7,8]. In 70–80% of human AML, a disease characterized by uncontrolled expansion of myeloid-lineage progenitors, *HoxA7*, *HoxA9* and *HoxA10* are overexpressed with *Meis1*, a Hox cofactor gene[9,10]. In a mouse leukaemia model, coexpression of *HoxA9* (or *HoxA7*) plus *Meis1* enforces myeloid progenitor expansion, causes rapid AML[11], and activates transcription of leukaemia-associated genes such as *Cd34, Flt3* and *Erg1* (ref. 12). The mechanisms that activate and/or maintain transcription of *Hox-A* genes in most cases of human AML is unknown, with exception of 5–10% of cases that contain chromosomal translocations involving *MLL* (mixed lineage leukaemia)[13], H3K4 histone methyltransferase (HMT) of the trithorax class that positively regulates *Hox-A* locus transcription in a complex with Menin, RbBP5, Ash2L, WDR5 and the MOF H4K16 acetyltransferase[14–16]. Chromosomal translocations fuse the portion of MLL that contains AT-hooks and binds Menin to heterologous transactivation domains that bind the H3K79 HMT hDOT1L, linking H3K79 methylation to leukaemogenesis and persistent activation of *Hox-A* protooncogenes[17–19].

About 5% of human AMLs harbour the t(5;11)(q35;p15.5) translocation[6], which generates *NUP98–NSD1*, a chimeric gene encoding the FG-repeat domain of NUP98, which binds the histone acetyltransferase CBP/p300 (ref. 20), fused to the carboxy-terminal 60% of NSD1, which contains five PHD fingers, a PHD finger-like Cys–His rich domain (CH), a PWWP domain, and a Set2-family HMT domain reported to target H3K36 (ref. 21; Fig. 1a). Loss of the reciprocal translocation product, NSD1–NUP98, implicates NUP98–NSD1 as the sole initiating oncoprotein[1,6]. Methylated H3K36 (H3mK36) is enriched in regions of active transcription[22] it has also been linked to transcriptional elongation[23], shown to antagonize transcriptional silencing[24], suppress erroneous transcription initiation[25] and is required for transcriptional activation of the *Arabiodopsis* flowering regulator, *FLOWERING LOCUS C*[26]. A number of clues suggest that NSD1 is linked to developmental regulation through its positive impact on *Hox* gene expression: first, PHD fingers 2 and 3 of NSD1 are most strongly related to consecutive PHD fingers in MLL[21], a well-known positive regulator of the *Hox* locus; second, the *Caenorhabditis elegans* homologue of NSD1, MES-4, also methylates H3K36 (ref. 27) and antagonizes the function of the nematode EZH2-complex[27,28], which is responsible for *Hox*-gene silencing during

[1]Department of Pathology, University of California at San Diego, School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA. [2]Biomedical Sciences Graduate Program, University of California at San Diego, School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA. [3]Division of Biological Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.
[4]Correspondence should be addressed to M.P.K (e-mail: mkamps@ucsd.edu)

Received 28 December 2006; accepted 30 May 2007; published online 24 June 2007; DOI: 10.1038/ncb1608

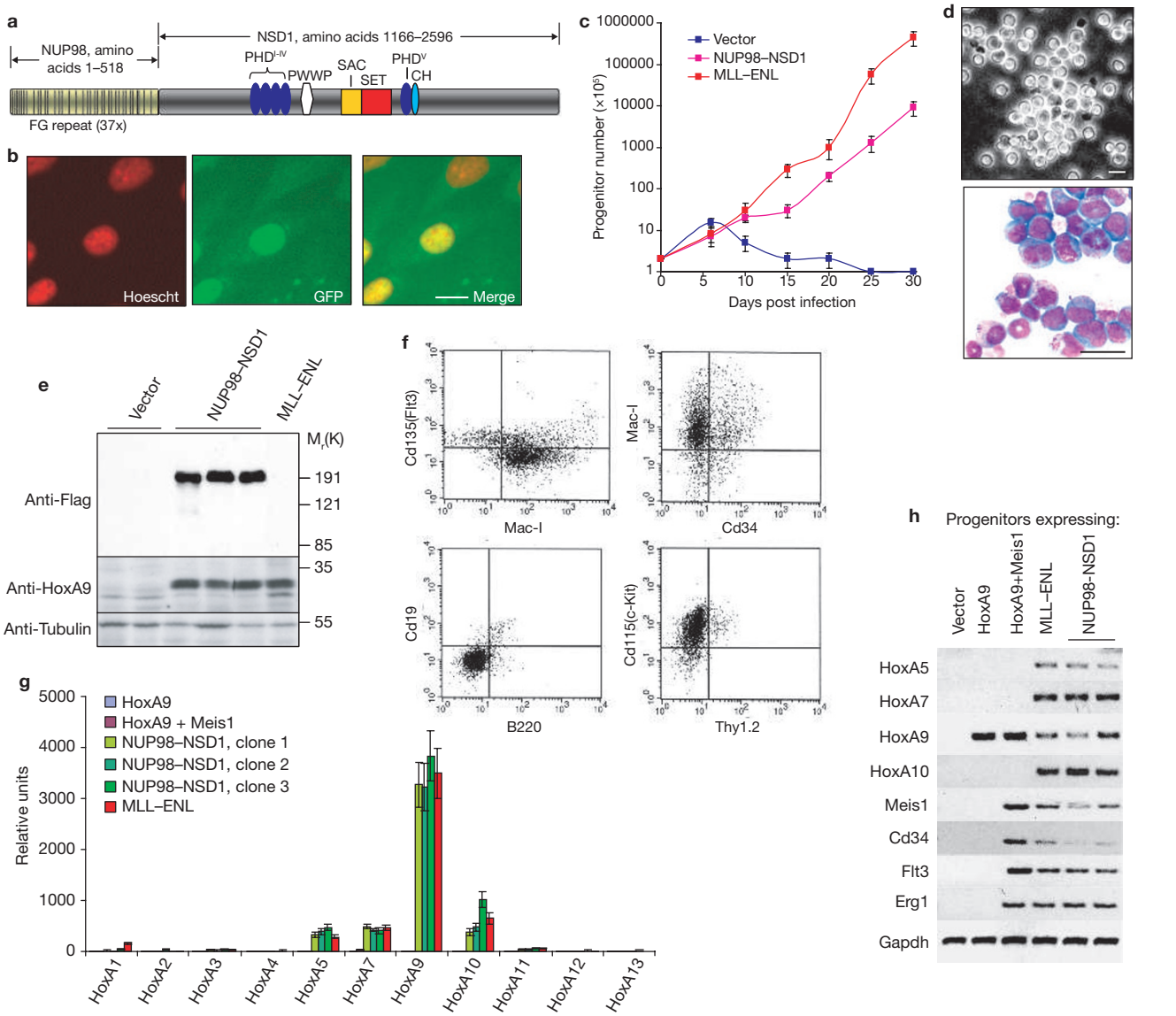


Figure 1 NUP98–NSD1 enforces self-renewal of myeloid progenitors and maintains expression of a subset of *Hox-A* genes, *Meis1* and their downstream target genes. (**a**) Schematic representation of the structure of NUP98–NSD1. Residues 1–518 of NUP98 are fused to residues 1164–2588 of NSD1. Each FG repeat in NUP98 is designated by a bar. (**b**) Fluorescence localization of GFP-tagged NUP98–NSD1 (GFP), DNA (Hoescht), and the merged image in NIH3T3 fibroblasts. (**c**) Altered proliferation kinetics of haematopoietic progenitors after infection by MSCV retrovirus encoding NUP98–NSD1 or MLL–ENL. At day 1, 200,000 drug-resistant Lin-enriched progenitors were plated in medium containing stem cell factor (SCF) plus Flt3 ligand (FL). The error bars indicate the s.d. of progenitor numbers in five independent cultures. (**d**) Light micrographs and Wright-Giemsa staining of myeloid progenitors immortalized by NUP98–NSD1. (**e**) Western-blot analysis of NUP98–NSD1 (Flag-tagged, anti-Flag) and HoxA9 in total cell lysates from cultures infected by empty retroviral vector, or immortalized by Flag-tagged NUP98–NSD1 or by MLL–ENL. Tubulin was used as a loading control. An uncropped image of the scan is shown in the Supplementary Information, Fig. S4a. (**f**) FACS analysis of progenitors immortalized by NUP98–NSD1 four weeks after retroviral infection. (**g**) Expression levels of *Hox-A* locus genes quantified by Affymetrix microarrays, using total RNA from progenitors immortalized by retroviral HoxA9 (three clones), coexpressed HoxA9 plus Meis1 (three clones), NUP98–NSD1 (three clones) or MLL–ENL (one clone). The error bars indicate the s.d. from 11 independent oligonucleotide-based assessments of transcript abundance. The *HoxA9*-specific probes on Affymetrix arrays detect a 3′-UTR region not contained in the retroviral *HoxA9* cDNA. Array data is deposited in the GEO database as specified in the Methods. (**h**) RT–PCR measurement of *HoxA5, HoxA7, HoxA9, HoxA10* and *Meis1*, and downstream targets *Cd34*, *Flt3* and *Erg1* in progenitors immortalized by NUP98–NSD1. The cDNA templates used are from control progenitors expressing empty vector and progenitors immortalized by retroviral HoxA9, HoxA9 plus Meis1, MLL–ENL or NUP98–NSD1 (two cultures). *HoxA9*- and *Meis1*-specific RT–PCR primers target coding sequences, amplifying both retroviral and endogenous transcripts. *Gapdh* served as internal template control. The scale bars represent 10 μm in **b** and 20 μm in **d**.

mammalian development; third, knockout of *Nsd1* in mice produces developmental defects and precludes expression of *HoxA1* and *HoxB1* at E7.5 (ref. 21). If NSD1-mediated H3K36 methylation regulates normal *Hox* gene expression, then deregulated NUP98–NSD1 activity could link persistent methylation of H3K36 to enforced expression of *Hox-A* genes in AML. To investigate this hypothesis, the leukaemic potential of retrovirally expressed NUP98–NSD1 was evaluated in an *in vitro* myeloid-progenitor transformation assay. NUP98–NSD1 protein was localized to

the nucleus (Fig. 1b). Unlike marrow-derived progenitors transduced by control empty retrovirus, which proliferated transiently and differentiate into neutrophils, macrophages, and mast cells, those infected with NUP98–NSD1 retrovirus proliferated indefinitely as undifferentiated progenitors (Fig. 1c, d). NUP98–NSD1 did not induce proliferation in NIH3T3 fibroblasts or produce factor-independence in haematopoietic progenitors (data not shown); therefore, the principal transforming property of NUP98–NSD1 is disallowing cellular differentiation, therein enforcing progenitor self-renewal. NUP98–NSD1-immortalized progenitors exhibited a myeloblast morphology (Fig. 1d), expressed full-length NUP98–NSD1 (Fig. 1e), and displayed myeloid surface antigens ($Cd34^{low}$/$Flt3^{low}$/c-Kit^{+}/Mac-I^{+}/$B220^{-}$/$Cd19^{-}$/$Thy1.2^{-}$; Fig. 1f). To gain insight into the genetic mechanism of myeloid progenitor immortalization, microarrays were used to compare gene-expression profiles between progenitors immortalized by NUP98–NSD1 or by the well-studied leukaemia proteins, HoxA9 plus Meis1, which do not activate the *Hox-A* locus or the *Meis1* gene[29]. Differentiation genes expressed in the small fraction of maturing granulocytes and macrophages in NUP98–NSD1 progenitor cultures were selected against by eliminating genes expressed during differentiation of myeloid progenitors conditionally immortalized by ER–HoxA9 (ref. 30). Only six genes — *HoxA5*, *HoxA7, HoxA9*, *HoxA10*, *Meis1* and *Rab38* — were upregulated fivefold or more in NUP98–NSD1 progenitors, demonstrating a high targeting specificity by NUP98–NSD1 on the *Hox-A* locus and on *Meis1* (Fig. 1e, g and 1h). No other *Hox-A* genes were expressed. Sustained transcription of *HoxA7, HoxA9, HoxA10* and *Meis1* also occured in progenitors immortalized by MLL–ENL (Fig. 1g, h) or MLL–AF9 (ref. 31). Downstream targets of HoxA9 and Meis1, such as *Cd34*, *Flt3* and *Erg1,* were also transcribed (Figs 1f, h and Fig. 2f). This suggested that NUP98–NSD1 enforces self-renewal and arrests differentiation by activating transcription of *Hox-A* genes and *Meis1*.

The observation that the potent leukaemia proto-oncogenes, *HoxA9* and *Meis1,* were coexpressed in NUP98–NSD1-immortalized progenitors suggested that NUP98–NSD1 would cause AML in a mouse adoptive-transfer experiment. Each of 18 mice reconstituted with lineage-negative marrow progenitors infected with NUP98–NSD1 retrovirus died of AML (average lifespan of 126 days post-transplant), whereas those reconstituted with vector-transduced marrow progenitors remained healthy (Fig. 2a). NUP98–NSD1 AML presented with >80% progenitors in marrow, 3–10-fold enlarged spleens, 2–10-fold enlarged lymph nodes and 3–40-fold increases in circulating white blood cells (Fig. 2b and see Supplementary Information, Fig. S1a–e and Table S1). Leukaemic blasts were myeloid ($Cd34^{variable}$/$Flt3^{variable}$/c-$Kit^{variable}$/Mac-I^{+}/$B220^{-}$/$Cd19^{-}$/$Thy1.2^{-}$; Fig. 2c), contained the same retroviral integrations as briefly expanded parental progenitor populations (Fig. 2e), expressed NUP98–NSD1 (Fig. 2d and see Supplementary Information Fig. S1f), and expressed endogenous *Meis1*, *HoxA5*, *HoxA7*, *HoxA9* and *HoxA10,* and their downstream targets *Flt3* and *Erg1* (Fig. 2d, f). Thus, NUP98–NSD1 targets the *Hox-A* locus and *Meis1* genes both in the progenitor immortalization model and in the mouse AML model.

The mechanistic similarity of NUP98–NSD1 to MLL oncoproteins prompted us to investigate whether NUP98–NSD1, like MLL oncoproteins, binds the *Hox-A* locus directly. NUP98–NSD1 activated transcription of a *HoxA7* promoter–*luciferase* reporter in both HEK293 and NIH3T3 cells (Figs 3a and 4e). Using chromatin immunoprecipitation (ChIP) analysis[17], anti-NUP98 or anti-Flag antibodies, but not control antibodies (nonspecific IgG or anti-HA antibody), precipitated genomic fragments exhibiting two enrichment peaks preceding the *HoxA7* and *HoxA9* transcriptional start sites from progenitors immortalized by Flag-tagged NUP98–NSD1, but not from control progenitors immortalized by MLL–ENL (Fig. 3b, c). Positive control anti-dimethylated-H3K4 and negative control anti-trimethylated-H3K9 antibodies demonstrated uniformity of ChIP conditions. NUP98–NSD1 did not bind proximal *Hox-A* genes (Fig. 3b), distal *Hox-A* genes or *HoxC8*, each of which is silent in NUP98–NSD1-immortalized progenitors, and did not bind the *Cd34* or *Flt3* promoters (data not shown). Thus, specific binding of NUP98–NSD1 to the *Hox-A* locus paralleled specific activation of *Hox-A*-locus genes in NUP98–NSD1 progenitors.

NUP98–NSD1 harbours at least two histone-modifying activities — the CBP/p300 acetyltransferase is recruited by NUP98 domain FG-repeats[20] (confirmed by coimmunoprecipitation; see Supplementary Information, Figs S1f and S2a, b), and H3K36 methyltransferase activity is catalysed by the NSD1 SET domain. ChIP was used to examine epigenetic changes induced by NUP98–NSD1 on the *HoxA5–A10* locus in haematopoietic progenitors. To discriminate between changes induced by NUP98–NSD1 and those intrinsic to normal undifferentiated myeloid progenitors, early haematopoietic progenitors expressing an empty retroviral vector and progenitors immortalized by coexpressed *HoxA9* plus *Meis1* were used as negative controls. Based on our array data, progenitors immortalized by HoxA9 plus Meis1 are blocked at a differentiation stage almost identical to those immortalized by NUP98–NSD1 or MLL–ENL[31], yet because their endogenous *Hox-A* locus is silent[29], they serve as negative controls for NUP98–NSD1-specific mechanisms of *Hox-A* locus activation. ChIP revealed strong histone acetylation and H3mK36 throughout the *HoxA7–A9* locus in NUP98–NSD1 progenitors, but not in progenitors expressing empty vector or in those immortalized by HoxA9 plus Meis1 (Fig. 3d). Consistent with the model that acetylation results from recruitment of CBP/p300 to NUP98, ChIP identified p300 on the *HoxA7–A9* locus in NUP98–NSD1 progenitors but not in control progenitors (Fig. 3e). ChIP also revealed strong H3K27 trimethylation on the *HoxA7–A9* locus in control progenitors, but not in progenitors immortalized by NUP98–NSD1 (Fig. 3d). In HeLa or 293T cells, recruitment of EZH2 complexes to the *HoxA9* locus coincides with appearance of H3mK27 and downregulation of *HoxA9* (refs 32, 33). Consistent with this observation, EZH2 was not detected at the *HoxA7–A9* locus in NUP98–NSD1-transformed progenitors but was abundant in control progenitors (Fig. 3e). Methylation of H3K9 or H4K20 in the *HoxA7–A9* locus was not detected in these samples (data not shown). Collectively, these data indicate that NUP98–NSD1 maintains active expression of *Hox-A* locus target genes by colocalizing H3mK36 and H3/H4 acetylation at regulatory elements, which prevents binding of EZH2-complexes, antagonizing initiation of H3mK27-mediated transcriptional silencing.

We constructed mutations in NUP98–NSD1 to locate sequences that are essential for binding the *Hox-A* locus, activating transcription of a *HoxA7* promoter–reporter, enforcing expression of endogenous *HoxA9*, and enforcing self-renewal of myeloid progenitors (Fig. 4a, b). All wild-type functions of NUP98–NSD1 were preserved in NUP98–$NSD1^{\Delta C}$, which lacks residues 2142-2588 (truncating the protein after the KEQSSE sequence) at the C-terminus (see Supplementary Information, Figs S1f and S3a). Sequences homologous to this C-terminal region are absent in NSD3 or in NUP98–NSD3, another AML translocation protein[2]. This

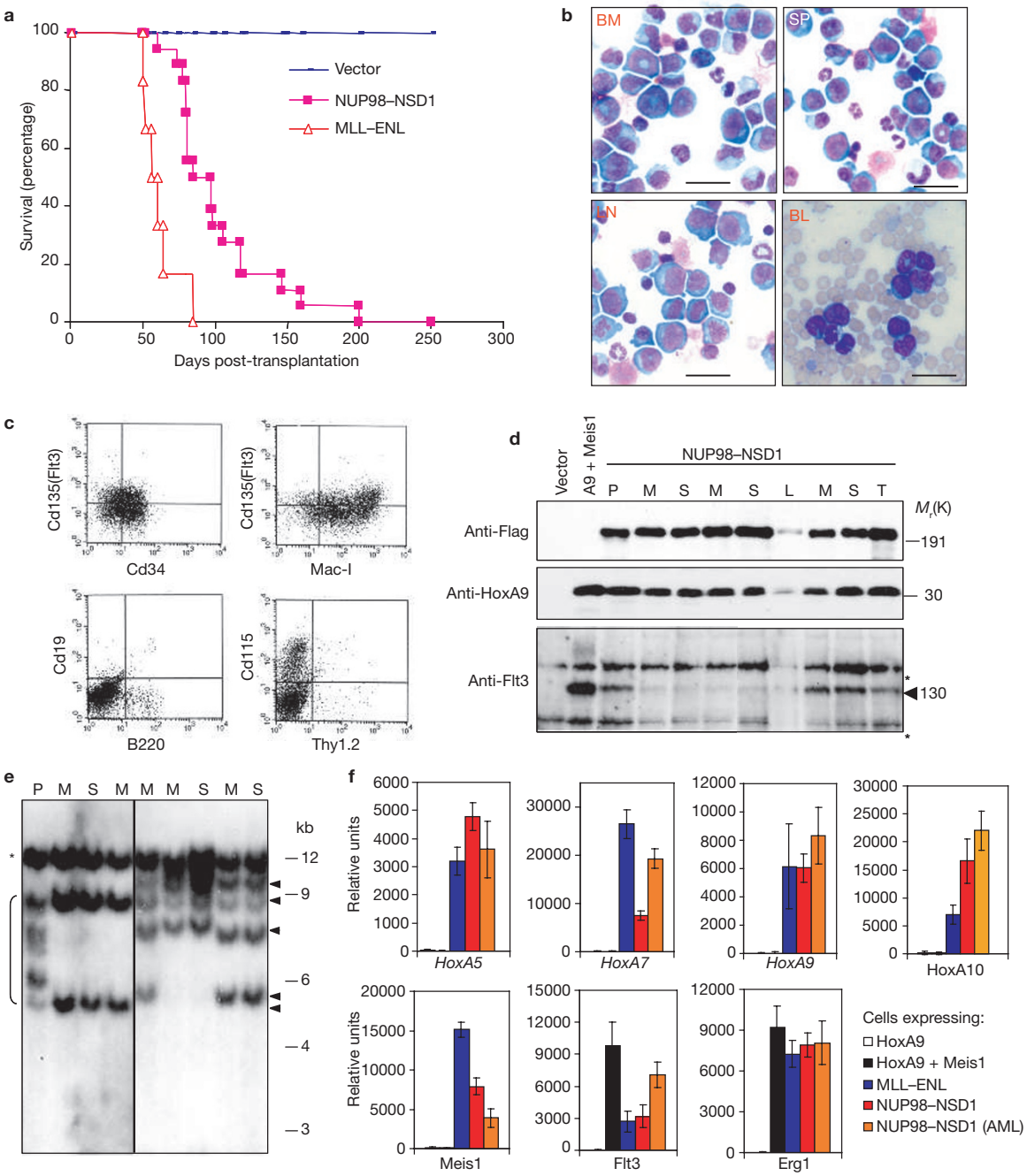


Figure 2 NUP98–NSD1 causes AML characterized by expression of a subset of *Hox-A* genes and *Meis1*. (**a**) Kinetics of leukaemia induction in the recipient mice after transplantation of marrow progenitors infected with retrovirus encoding empty vector ($n = 8$), NUP98–NSD1 ($n = 18$) or MLL–ENL ($n = 6$). (**b**) Wright-Giemsa staining of leukaemia cells extracted from bone marrow (BM), spleen (SP), lymph nodes (LN) and circulating blood (BL) of mice with NUP98–NSD1-induced AML. The scale bars represent 20 μm. (**c**) FACS analysis of NUP98–NSD1 AML cells exhibit a myeloid phenotype (Mac1$^+$ Cd19$^-$ B220$^-$ Thy1.2$^-$) characterized by coexpression of early stem cell antigens Flt3 (Cd135), c-Kit (Cd115) and Cd34. (**d**) Western-blot analysis of Flag-tagged NUP98–NSD1, HoxA9 and Flt3 in drug-resistant Lin$^-$ progenitors infected with retrovirus encoding Flag-tagged NUP98–NSD1 (P) or in leukaemic cells extracted from bone marrow (M), spleen (S), lymph nodes (L) or thymus (T) of mice bearing NUP98–NSD1-induced AML. Asterisks indicate nonspecific bands. Progenitors infected with empty vector or with a vector encoding *HoxA9* plus *Meis1* (A9 + Meis1) serve as negative and positive controls, respectively. An uncropped image of the anti-HoxA9 blot is shown in the Supplementary Information Fig. S4b. (**e**) Southern-blot analysis of genomic integration sites of NUP98–NSD1-retrovirus, using an *NSD1* cDNA probe. Samples include parental retrovirus-infected marrow progenitors (P) and their derivative leukaemic cells extracted from bone marrow (M) or spleen (S). The bracket designates multiple genomic bands harbouring integrated retrovirus in the parental population of infected marrow progenitors, arrowheads indicate those in derivative AML cells, and an asterisk indicates the endogenous *NSD1* locus. The two sets of samples in Figure 2E were run on the same gel. The second set was derived from mice injected with a second population of NUP98-NSD1-transduced progenitors. (**f**) Real-time RT–PCR analysis of transcript abundance of the endogenous *Hox-A* genes (*HoxA5, HoxA7, HoxA9* and *HoxA10*), *Meis1, Flt3* and *Erg1* in progenitors immortalized by HoxA9, HoxA9 plus Meis1, MLL–ENL or NUP98–NSD1 and in total marrow from leukaemia induced by NUP98–NSD1. 3′-UTR probes were designed for assessment of *HoxA9* and *Meis1* expression to measure endogenous gene expression and exclude detection of retroviral transcripts.

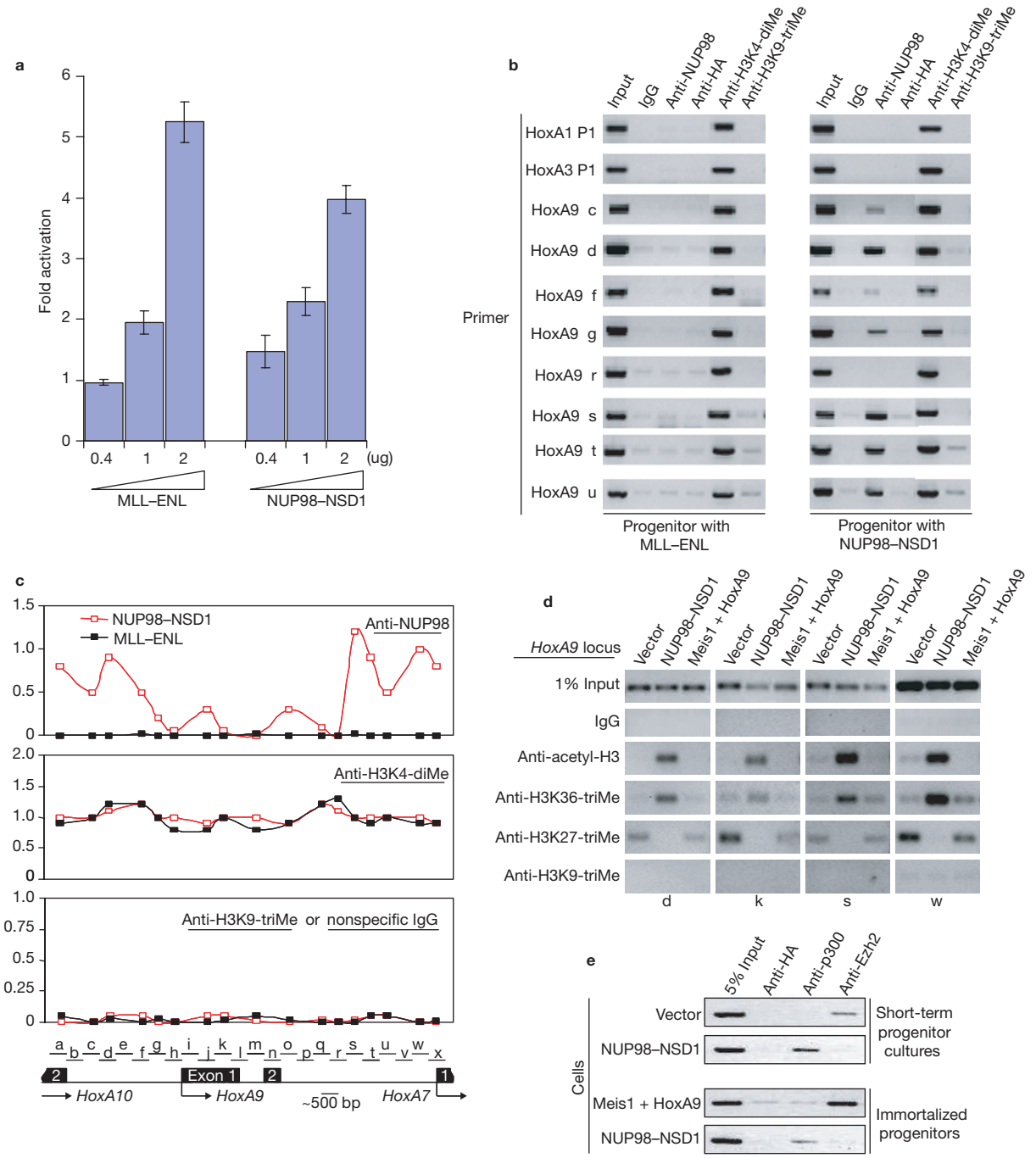


Figure 3 NUP98–NSD1 directly activates *HoxA* gene transcription and prevents H3K27-mediated represssion. (**a**) Transactivation of a *HoxA7* promoter luciferase reporter 72 h after transfection of 0.4 μg, 1 μg or 2 μg of MLL–ENL or NUP98–NSD1 plasmids into HEK293 cells. The error bars indicate the s.d. from triplicate analysis. (**b**) ChIP analysis of genomic fragments from the proximal promoter of *HoxA1* or *HoxA3* and sequences spanning 3′ of *HoxA10* to 5′ of *HoxA7*. Immunoprecipitation was performed using negative control antibodies (Abs; nonspecific IgG or anti-HA), antibodies against NUP98 (anti-NUP98 or anti-Flag), and antibodies against dimethylated-H3K4 (H3K4-diMe) or trimethylated-H3K9 (H3K9-triMe) from progenitors immortalized by NUP98–NSD1 or MLL–ENL. Five percent of sonicated chromatin was used as an input control. The position of *HoxA10–A7* primer sets is indicated in **c**. Uncropped images of the scans for primer sets d and t are shown in the Supplementary Information, Fig. S4c. (**c**) Quantification of NUP98–NSD1 binding and histone modifications H3K4-diMe and H3K9-triMe across the *HoxA10–A7* locus in progenitors immortalized by NUP98–NSD1 or by MLL–ENL, deduced by ChIP analysis (signals from **b**). Each point represents one ChIP primer set whose genomic location is shown in the lower panel. (**d**) Markers of active transcription, histone acetylation and H3K36-triMe, characterize the *HoxA9* locus in NUP98–NSD1-transformed progenitors, whereas the repression marker, H3K27-triMe, characterizes the same locus in early Lin⁻ marrow progenitors expressing empty vector or those transformed by retroviral coexpression of HoxA9 plus Meis1. Non-specific IgG was used as an immunoprecipitation control, and 1% of chromatin DNA used as an input control. The letter under each panel represents the primer set used for ChIP, as shown in **c**. (**e**) ChIP reveals recruitment of p300 to the *HoxA10–A7* locus in progenitors expressing NUP98–NSD1, and of polycomb protein EZH2 on the *HoxA10–A7* locus in control Lin⁻ progenitors infected with empty retrovirus or immortalized by HoxA9 plus Meis1 (ChIP result using *HoxA9* promoter set d). Uncropped images of the scans for primer sets d and t are shown in the Supplementary Information, Fig. S4d.

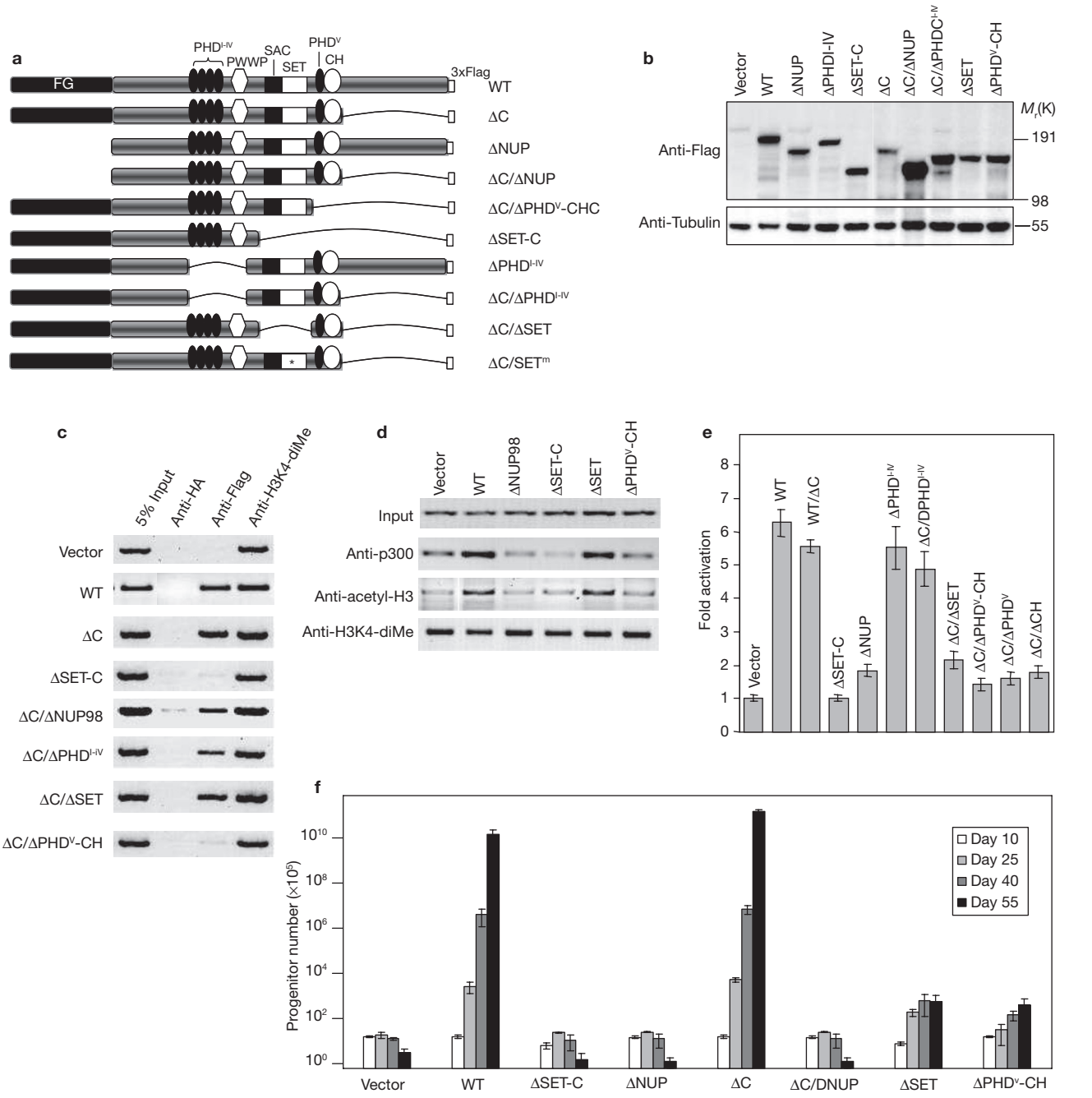


Figure 4 Analysis of NUP98-NSD1 domains required to bind chromatin activate *Hox-A* gene transcription and enforce myeloid progenitor self-renewal. (**a**) Schematic representation of deletions and point mutations in NUP98–NSD1. All constructs were tagged with a 3×Flag tag at the C-terminus. (**b**) Stable expression of mutants in drug-selected progenitors at levels similar to or higher than NUP98–NSD1$^{\Delta C}$, as detected by western blotting with anti-Flag antibodies. Anti-tubulin was used for the loading control. (**c**) ChIP analysis of binding of NUP98–NSD1 mutants to the *HoxA9* promoter in short-term haematopoietic progenitor cultures, using anti-Flag antibodies (results using primer set d). Five percent of total chromatin or ChIP with anti-H3K4-diMe are the input and positive controls, and ChIP with anti-HA is the negative control. Representative full scans are shown in the Supplementary Information, Fig. S4e. (d) ChIP analysis of bound CBP/p300 and of H3 acetylation on the *HoxA9* promoter in haematopoietic progenitors 5 days after transduction with retrovirus expressing wild-type or mutant NUP98–NSD1, using anti-p300 and anti-acetyl-H3 antibodies and primer set d. Five percent of total chromatin or ChIP with anti-H3K4-diMe are the input and positive controls. (**e**) Transactivation of a *HoxA7*-promoter reporter in NIH3H3 fibroblasts stably expressing wild-type and deletion mutants of NUP98–NSD1. Luciferase activity normalized to internal renilla luciferase control and compared to the basal activity of the *HoxA7*-reporter. Student's *t*-test, $P < 0.00001$ for wild-type ($n = 5$) versus deletion forms exhibiting loss of function (all mutants with exception of ΔC or ΔPHD^{1-IV}). Error bars indicate s.d. for triplicate analysis for all time points for each mutant. (**f**) Capability of NUP98–NSD1 mutants to enforce self-renewal of primary myeloid progenitors in medium containing SCF and FL. After retroviral infection and drug selection, progenitors were quantified at day 10, day 25, day 40 and day 55. The error bars indicate s.d. for four different cultures.

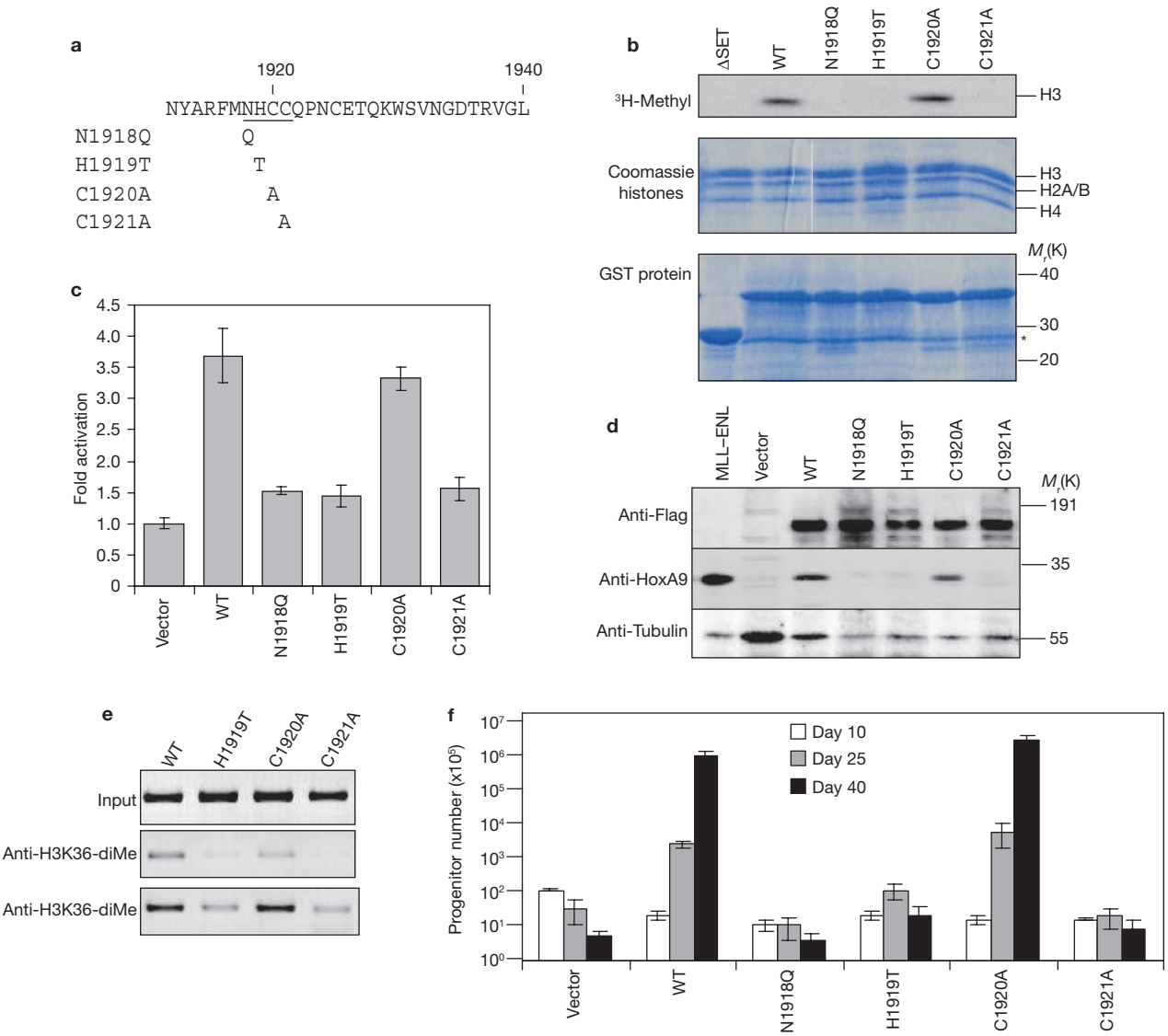


Figure 5 Point mutations in the NSD1 SET domain abolish upregulation of *HoxA9* transcription and the ability to enforce self-renewal of myeloid progenitors. (**a**) Schematic representation of the location of point mutations in the NSD1 SET domain. (**b**) Impact of NSD1 SET-domain mutants on histone H3 methylation *in vitro*. An autoradiogram of ^{3}H-methyl incorporated into histone H3 after incubation with the wild-type GST–NSD1 SET/PHD^{V}-CH fusion protein or its SET domain mutations is shown (top). Quantification of histone substrates and GST fusion proteins visualized by Coomassie blue staining is also shown. The star indicates degradation products of the GST fusion substrate (relative molecular mass of 35,000). An uncropped image of the methylation data is shown in Supplementary Information, Fig. S4f. (**c**) Impact of SET domain mutants on transactivation of the *HoxA7* promoter–luciferase reporter by NUP98–NSD1 in HEK293 cells. Student's *t*-test, $P = 0.0002$ for wild-type versus mutants N1918Q, H1919T or C1921A. Error bars indicate s.d. (n=4 for wild-type and each mutant). (**d**) Impact of SET domain point mutants on activation of endogenous *HoxA9* transcription by NUP98–NSD1 (3×Flag-tagged), introduced by retroviral transduction into Lin^{-} marrow progenitors and measured after a 5-day drug selection (see Methods for details). Cells were probed with anti-Flag, anti-HoxA9 and anti-tubulin antibodies. Progenitors infected with MLL–ENL retrovirus and empty vector served as positive and negative controls of *HoxA9* upregulation, respectively. The faint HoxA9 band in negative controls represents the low levels of endogenous HoxA9 produced in normal differentiating progenitors. Uncropped images of the scans of the westerns blots for NUP98–NSD1 and HoxA9 are shown in the Supplementary Information Fig. S4e. (**e**) ChIP analysis of H3mK36 on the *HoxA9* promoter in haematopoietic progenitors transduced with wild-type or SET domain point mutants of NUP98–NSD1, using anti-dimethyl-H3K36 and anti-trimethyl-H3K36 antibodies. The drug-selected progenitors were obtained as in Fig. 4c using primer set d. Five percent of total chromatin was used as input control. (**f**) Impact of SET domain point mutants on myeloid progenitor self-renewal and differentiation arrest. After retroviral infection and selection for drug-resistance, progenitor expansion was measured at day 10, day 25 and day 40. Progenitor expansion for each construct was evaluated for four cultures and the error bars represent s.d.

suggests that NUP98–NSD3 uses the same transforming mechanism as NUP98–NSD1. The fifth PHD finger (PHD^{V}) and adjacent C5HCH finger (CH) were the sole domains required for tight binding of NUP98–NSD1 to DNA probes 5′ of the *HoxA7* and *HoxA9* genes (Fig. 4c). The ΔPHD^{V}–CH mutant abolished both the increased recruitment of CBP/p300 to the *HoxA9* promoter and the increased histone acetylation of the *HoxA9* promoter (Fig. 4d), prevented transcriptional activation of the *HoxA7* promoter (Fig. 4e), and precluded immortalization of myeloid

progenitors (Fig. 4f), indicating that the PHD^{V}–CH motifs harbour a chromatin-targeting ability underlying gene activation and leukaemogenesis. PHD fingers I–IV were unnecessary for NUP98–NSD1 binding to the *Hox-A* locus or for transactivation of the *HoxA7* promoter. The NUP98 domain was not required for binding to the *HoxA9* promoter (Fig. 4c); however, its elimination prevented NUP98–NSD1 from binding CBP/p300 (see Supplementary Information, Figs S1f and S2C), and prevented additional recruitment of CBP/p300 to the *HoxA9* promoter (Fig. 4d) it also prevented additional histone acetylation on the *HoxA9* promoter (Fig. 4d), *HoxA7* promoter activation (Fig. 4e) and myeloid progenitor immortalization (Fig. 4f). Thus, the transforming abilities of NUP98–NSD1 are strictly dependent on a NUP98 function that correlates with binding CBP/p300 and with *Hox-A* locus acetylation although we cannot exclude the possibility that NUP98 performs another essential role. The impact of NUP98–NSD1 mutants on preventing H3K27 methylation could not be measured in this transient assay because many progenitors unaffected by NUP98–NSD1 produce robust H3K27 methylation in cultures infected by wild-type or mutant forms of NUP98–NSD1 alike (see Supplementary Information, Figs S1f and S2d).

Deletion of the SET domain (ΔC/ΔSET) did not alter binding of NUP98–NSD1 to the *HoxA9* promoter (Fig. 4c), but precluded both *HoxA7* promoter activation (Fig. 4e) and immortalization of myeloid progenitors (Fig. 4f) indicating that H3K36 methylation controls both target gene activation and leukaemogenesis. To more rigorously test this methylation-dependent transformation hypothesis, point mutations reported to disrupt (N1918Q, H1919T and C1921A) or augment (C1920A) HMT activity were introduced within the NUP98–NSD1 SET domain (Fig. 5a, b), and their transforming properties were examined. NUP98–NSD1 mutants N1918Q, H1919T or C1921A failed to methylate H3K36 *in vitro* (Fig. 5b) and failed to activate transcription of the *HoxA7*-promoter reporter (Fig. 5c) it also failed to increase H3mK36 within the *HoxA9* promoter (Fig. 5e), to stimulate transcription of the endogenous *HoxA9* gene (Fig. 5d) and to immortalize myeloid progenitors (Fig. 5f). Mutant C1920A retained all these activities at levels comparable to wild-type NUP98–NSD1 (Fig. 5b–f). Therefore, H3K36 methylation by NSD1 is an essential epigenetic modification through which it activates *Hox-A* locus transcription and enforces myeloid progenitor self-renewal.

In summary, we established mouse models of myeloid leukaemogenesis by the human translocation protein NUP98–NSD1, and investigated the genetic pathways and epigenetic mechanisms by which NUP98–NSD1 deregulates proliferation and differentiation (see Supplementary Information, Fig. S3b). We demonstrated that *HoxA5–A10* are downstream targets of NUP98–NSD1 and that NUP98–NSD1 binds chromatin adjacent to *HoxA7* and *HoxA9* through its PHD^{V}–CH domain. Although PHD$^{I–IV}$ of NSD1 is not required for binding the *Hox-A* locus, its deletion prevents NUP98–NSD1 from establishing long-term immortalization of haematopoietic progenitors (data not shown), indicating these PHD domains also perform critical activities in target gene regulation. Interestingly, mutation of the three PHD fingers in MES-4 also dissociates it from autosomes[27]. The PHD fingers in NSD1 could bind modified forms of histones, as recently observed for NURF–BPTF[34] and ING2 (ref. 35). Binding of both NSD1 and MLL proteins to the *Hox-A* locus raises the possibility that their cognate modifications — H3mK36 and H3mK4 — may cooperatively regulate *Hox-A* locus transcription during embryogenesis. As mutations in the PHD fingers and SET domain of NSD1 compromised progenitor self-renewal by NUP98–NSD1, we predict that NSD1 point mutations in Soto's syndrome, which cluster in the PHD fingers and SET domain[3] prevent NSD1 from regulating self-renewal and differentiation genes by interfering with its ability to bind chromatin and to modify histones or associate with cofactors, resulting in mental retardation and cerebral gigantism. Finally, we demonstrated that H3K36 methylation by NUP98–NSD1 is required for activation of *Hox-A* locus transcription and enforced self-renewal of myeloid progenitors. This is the first example of oncogenic involvement of an H3K36-specific histone methyltransferase, and joins the H3K79-specific methyltransferase hDOT1L as the second example in which HMT activity mediates human tumorigenesis.

It is not known how H3K36 methylation by NSD1 contributes to *Hox-A* locus activation. MES-4, the nematode homologue of NSD1, associates specifically with autosomes and/or euchromatin, and antagonizes the function of the EZH2-complex homologues MES-2/MES-3/MES-6 during development[28]. MES-4 is not involved in transcriptional elongation, as are other SET2-related HMTs that catalyse H3mK36 in transcribed regions[27]. Thus, by analogy, NSD1 could have a similar anti-repression function, using H3mK36 to block EZH2-mediated silencing of the *Hox-A* locus. □

METHODS

Plasmid construction and retroviral expression system. The *NUP98–NSD1* cDNA was generated by ligating the *NUP98* cDNA sequences encoding amino acids 1–518 to *NSD1* cDNA sequences encoding amino acid 1164–2588 (kindly provided by P. Chambon, Strasbourg, France). The *NUP98–NSD1* cDNA was then cloned into the Mscv–Neo retroviral expression vector. Flag-tagged or GFP-tagged versions were generated by inserting the 3×Flag tag or GFP sequences into a unique *Nsi*I site generated by site-directed mutagenesis at the C-terminal codon. Internal deletions were generated after excising cDNA regions flanked by two *Mfe*I sites created by mutagenesis.

***In vitro* haematopoietic progenitor proliferation and immortalization assays.** Protocols for purification and culture of primary haematopoietic progenitors were described previously[12]. Briefly, 3×10^5 of enriched Lin$^-$ bone marrow progenitors from Balb/c or B/6 mice were subjected to two rounds of spinoculation infection using 1 ml of retroviral supernatant with a viral titre of $2–5 \times 10^5$ ml^{-1}, followed by 4 days of drug resistance selection (0.8–1 mg ml^{-1} G418), and then cultured in medium containing SCF (generated by SCF-producing cell lines) and FLt3 ligard (Sigma, St. Louis, MO). Proliferation kinetics were evaluated by plating $1–3 \times 10^5$ drug-resistant Lin$^-$ progenitors in a 12-well tissue culture plate and splitting every 3–4 days in fresh medium. Wright-Giemsa staining and FACS analysis were performed as previously described[12].

Affymetrix microarray analysis. Total RNA was extracted from stably transformed progenitors cultured *in vitro* and the expression levels of mRNA transcripts quantified using the Affymetrix GeneChip Mouse Genome 430 2.0 array, as previously described[12]. The GEO database accession numbers: for progenitors immortalized by HoxA9 (GSM190542, GSM190546, GSM190547); for progenitors immortalized by coexpressed HoxA9 plus Meis1 (GSM190548, GSM190549, GSM190550); for progenitors immortalized by NUP98-NSD1 (GSM190551, GSM190552, GSM190553); and for progenitors immortalized by MLL-ENL (GSM190554).

Semi-quantitative RT–PCR and real-time qPCR. Total RNA was extracted from immortalized progenitors, reverse-transcribed, and used as a template for RT–PCR and quantitative PCR, as previously described[12].

***In vivo* leukaemogenesis assay**. The leukaemic potential of oncogenes was evaluated in sub-lethally irradiated syngeneic Balb/c mice (500 rads) or B/6 mice (700 rads) followed by tail-vein injection with 5×10^5 Lin$^-$ progenitors that were prestimulated in a cytokine-rich medium, infected with retrovirus and subjected to drug resistance selection, as previously described[12]. Mice exhibiting a leukae-

mia phenotype (lethargy and splenomegaly) were sacrificed, and cells isolated from leukaemic tissues were subjected to further analysis.

Luciferase reporter assay. Transactivation of the *HoxA7* promoter by NUP98–NSD1 was evaluated using a pGL3 luciferase reporter containing a 2.8 kb fragment preceding the *HoxA7* transcription initiation site (a gift from R. Slany, Erlangen, Germany), which was previously used to evaluate transactivation by *MLL* fusion oncoproteins. The *HoxA7*-promoter reporter (100–500 ng) was cotransfected with 1 ng pRL-TK (Promega, Madison, WI) into NIH3T3 cell lines stably expressing empty retroviral vector or either wild-type or mutant NUP98–NSD1. For luciferase assay performed in HEK293 cells, 2 μg Mscv plasmids encoding NUP98–NSD1 or MLL–ENL were cotransfected with 500 ng *HoxA7* reporter plus 1 ng pRL-TK. At 48–72 h post transfection, cell lysate was prepared and quantified for firefly and renilla luciferase activity using a dual luciferase reporter system (Promega) and LMaxII luminometer (Molecular Devices, Sunnyvale, CA). After normalization against renilla luciferase units, transcriptional activation by NUP98–NSD1 was expressed as fold-increase in comparison to basal luciferase levels produced in cells transfected with the *HoxA7*-reporter and empty Mscv vector.

ChIP analysis. ChIP analysis was performed according to a previously described protocol[12,17]. Anti-acetyl-histone H3, anti-dimethylated H3K4, anti-trimethylated H3K9, anti-trimethylated H3K27, anti-dimethylated H3K36 and anti-trimethylated H4K20 were purchased from Upstate, Charlottesville, VA. Other antibodies included anti-Flag (M2; Sigma, St. Louis, MO), anti-HA (MMS101; Covance, Berkeley, CA), anti-Myc (9E10), anti-trimethylated H3K36 (Abcam, Cambridge, UK), anti-p300 (sc-584, sc-585; Santa Cruz Biotech, CA) and anti-EZH2 (cat#4905; Cell Signaling, Charlottesville, VA). Direct binding of NUP98–NSD1 to the *Hox-A* locus was examined using twenty-four primers that sampled every 500 nucleotides spanning the *HoxA7* to *HoxA10* gene locus[17], six primer sets that sampled outlying proximal (*HoxA1–A3*) or distal (*HoxA11–A13*) *Hox-A* regions or *HoxC8*, and six primer sets that sampled proximal promoter regions of *Cd34* and *Flt3* (see Supplementary Information, Methods, for primer sequences). For short-term cultured progenitors, such as those used for analysis of NUP98–NSD1 deletion mutants, progenitors were collected following a 5-day drug selection of retrovirally infected cultures of normal marrow Lin^- progenitors. Stable immortalized progenitors were collected from similar cultures at least 30 days after retroviral infection and drug-resistance selection.

***In vitro* HMT assay.** *In vitro* HMT reactions were carried out in a volume of 50 μl containing 1× HMT buffer (Upstate, Charlottesville, VA), 20 μg chicken core histones (Upstate) as substrates, 300 nCi of S-methyl-^{3}H-adenosyl-L-methionine (Perkin-Elmer, Waltham, MA) as methyl donor, and 10–20 μg of freshly made GST-fusion proteins immobilized on glutathione–Sepharose. After incubation for 4 h to overnight at 30–37 °C, reactions were stopped by boiling in SDS sample buffer, and proteins were separated by electrophoresis on 15% SDS–PAGE gels and visualized by staining with Coomassie blue. Gels were treated with Fluro-Enhance solution (Research Products International, Mt. Prospect, IL), dried, and exposed for one week at –80 °C.

Note: Supplementary Information is available on the Nature Cell Biology website.

ACKNOWLEDGEMENTS

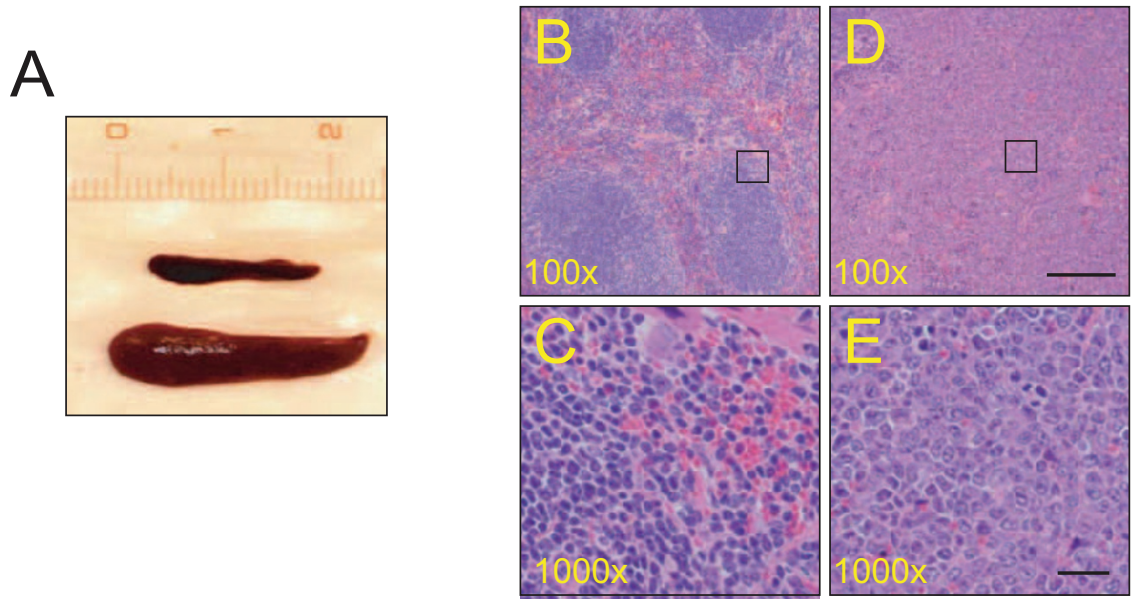
This work is supported by Public Health Service grant NIH CA56876 awarded to M.P.K. We thank D. Young for assistance with FACS analysis, J. Aguilera for assistance with mouse radiation and C. Yost for assistance with haematological analysis (UCSD Cancer Center Shared Resources). Special thanks to P. Chambon, J. Van Deursen and R. Slany for providing plasmids and antibodies. We also thank J.-L. Luo and S. Zhang for helping with bone marrow transplantation and fluorescence microscopy.

COMPETING FINANCIAL INTERESTS

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturecellbiology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. La Starza, R. *et al.* Cryptic insertion producing two *NUP98/NSD1* chimeric transcripts in adult refractory anemia with an excess of blasts. *Genes Chromosomes Cancer* **41,** 395–399 (2004).
2. Rosati, R. *et al.* NUP98 is fused to the *NSD3* gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). *Blood* **99,** 3857–3860 (2002).
3. Tatton-Brown, K. *et al.* Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with *NSD1* aberrations. *Am. J. Hum. Genet.* **77,** 193–204 (2005).
4. Keats, J. J. *et al.* Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. *Blood* **105,** 4060–4069 (2005).
5. Tonon, G. *et al.* High-resolution genomic profiles of human lung cancer. *Proc. Natl Acad. Sci. USA* **102,** 9625–9630 (2005).
6. Cerveira, N. *et al.* Frequency of NUP98–NSD1 fusion transcript in childhood acute myeloid leukaemia. *Leukemia* **17,** 2244–2247 (2003).
7. Sauvageau, G. *et al.* Differential expression of homeobox genes in functionally distinct $CD34^+$ subpopulations of human bone marrow cells. *Proc. Natl Acad. Sci. USA* **91,** 12223–12227 (1994).
8. Thorsteinsdottir, U. *et al.* Overexpression of the myeloid leukemia-associated *Hoxa9* gene in bone marrow cells induces stem cell expansion. *Blood* **99,** 121–129 (2002).
9. Lawrence, H. J. *et al.* Frequent co-expression of the *HOXA9* and *MEIS1* homeobox genes in human myeloid leukemias. *Leukemia* **13,** 1993–1999 (1999).
10. Afonja, O. *et al. MEIS1* and *HOXA7* genes in human acute myeloid leukemia. *Leuk. Res.* **24,** 849–855 (2000).
11. Kroon, E. *et al.* Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17,** 3714–3725 (1998).
12. Wang, G. G., Pasillas, M. P. & Kamps, M. P. 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 (2006).
13. Hess, J. L. MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol. Med.* **10,** 500–507 (2004).
14. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121,** 859–872 (2005).
15. Dou, Y. *et al.* Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121,** 873–885 (2005).
16. Yokoyama, A. *et al.* The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123,** 207–218 (2005).
17. Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121,** 167–178 (2005).
18. Okada, Y. *et al.* Leukaemic transformation by CALM–AF10 involves upregulation of *Hoxa5* by hDOT1L. *Nature Cell Biol.* **8,** 1017–1024 (2006).
19. Milne, T. A., Martin, M. E., Brock, H. W., Slany, R. K. & Hess, J. L. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res.* **65,** 11367–11374 (2005).
20. Kasper, L. H. *et al.* CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98–HOXA9 oncogenicity. *Mol. Cell Biol.* **19,** 764–776 (1999).
21. Rayasam, G. V. *et al.* NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* **22,** 3153–3163 (2003).
22. Pokholok, D. K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122,** 517–527 (2005).
23. Xiao, T. *et al.* Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev.* **17,** 654–663 (2003).
24. Tompa, R. & Madhani, H. D. Histone H3 lysine 36 methylation antagonizes silencing in *Saccharomyces cerevisiae* independently of the Rpd3S histone deacetylase complex. *Genetics* **175,** 585–593 (2007).
25. Carrozza, M. J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123,** 581–592 (2005).
26. Zhao, Z., Yu, Y., Meyer, D., Wu, C. & Shen, W. H. Prevention of early flowering by expression of *FLOWERING LOCUS C* requires methylation of histone H3 K36. *Nature Cell Biol.* **7,** 1256–1260 (2005).
27. Bender, L. B. *et al.* MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the *C. elegans* germ line. *Development* **133,** 3907–3917 (2006).
28. Fong, Y., Bender, L., Wang, W. & Strome, S. Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of *C. elegans*. *Science* **296,** 2235–2238 (2002).
29. Wang, G. G., Pasillas, M. P. & Kamps, M. P. 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 (2005).
30. Wang, G. G. *et al.* Quantitative production of macrophages or neutrophils *ex vivo* using conditional Hoxb8. *Nature Methods* **3,** 287–293 (2006).
31. Zeisig, B. B. *et al.* Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol. Cell Biol.* **24,** 617–628 (2004).
32. Cao, R. & Zhang, Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED–EZH2 complex. *Mol. Cell* **15,** 57–67 (2004).
33. Cha, T. L. *et al.* Akt-mediated phosphorylation of EZH2 suppresses methylation of lysine 27 in histone H3. *Science* **310,** 306–310 (2005).
34. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442,** 86–90 (2006).
35. Shi, X. *et al.* ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442,** 96–99 (2006).



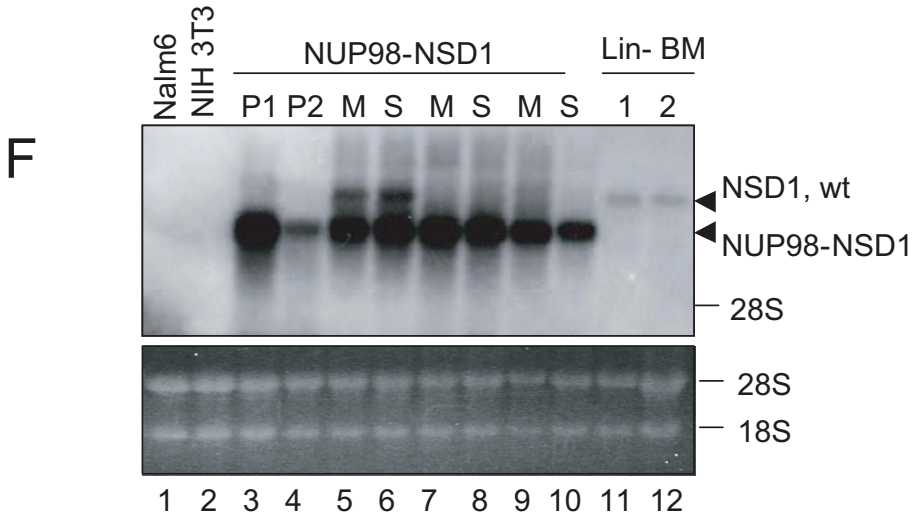


Figure S1. Phenotypic analysis of AML induced by NUP98-NSD1 in bone marrow transplantation models. (a) Typical spleen size in a mouse presenting with NUP98-NSD1-induced AML (lower) compared with that from a cohort control injected with Lin- progenitors expressing empty vector (upper). (b-e) Hematoxylin and Eosin (H&E) staining of spleen sections demonstrates that the follicles characteristic of normal splenic architecture (b and c) are disrupted in the spleens of mice bearing NUP98-NSD1 AML (d and e). Scale bars represent 250 micrometers in panels B and D, and 30 micrometers in panels c and e. (f) Northern blot using an NSD1 probe verifying expression of retroviral NUP98-NSD1 in infected parental progenitors (P1 and P2; lane 3-4) and in leukemic progenitors exacted from bone marrow (M) or spleen (S) of AML-bearing mice injected with P1 or P2 (lanes 5-10). Endogenous NSD1 was detected in each of two uninfected cultures of Lin$^-$ bone marrow progenitors (lane 11-12), but not in Nalm-6 pre-B cell lines (lane 1) or in NIH-3T3 fibroblasts (lane 2). Bottom panel represents the agarose gel image of RNA samples demonstrating equal loading of RNA. The position of 28S and 18S RNA are indicated.

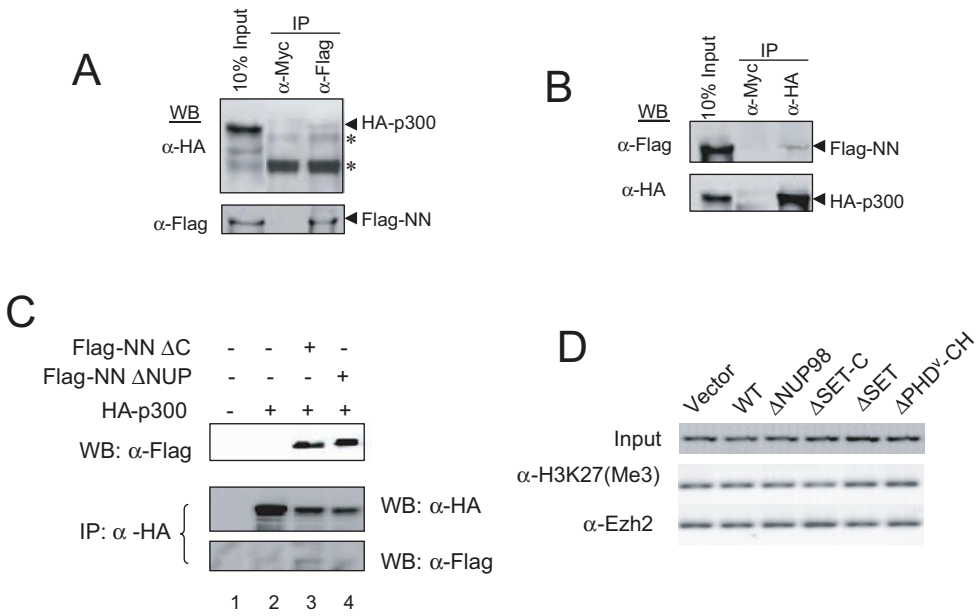


Figure S2. (a-c) Interaction of CBP/p300 with NUP98-NSD1 is mediated by NUP98 sequences. (a) α-Flag immunoprecipitation (IP) of NUP98-NSD1 (NN, Flag-tagged) co-precipitates p300 (HA-tagged), identified by α-HA western blotting, in cotransfected 293T cells. Co-precipitation is not observed in the α-Myc sera control. Stars represent nonspecific bands. The bottom panel represents an α-Flag western blot demonstrating that NUP98-NSD1 is efficiently immunoprecipitated. (b) α-HA IP of HA-p300 coprecipitates NUP98-NSD1 (NN, Flag-tagged), which is not precipitated by α-Myc sera. The bottom panel represents an α-HA western blot demonstrating that HA-p300 is efficiently immunoprecipitated. The positions of HA-p300 and NUP98-NSD1 are indicated at right. (c) Deletion of NUP98 (NN-ΔNUP98) abolishes the interaction of NUP98-NSD1 with CBP/p300 in HEK293 cells, while deletion of the C-terminal 50 KD (NN-ΔC) does not. The top panel demonstrates equal expression of NUP98-NSD1 by α-Flag western blot in total cell lysate, the middle panel demonstrates equal IP efficiency of HA-p300, and the bottom panel examines the presence of NUP98-NSD1 in precipitated samples by α-Flag western blotting. (d) ChIP analysis of Ezh2 binding and H3K27 methylation on the *HoxA9* promoter in the early hematopoietic progenitors transduced with wild-type or mutant NUP98-NSD1, using α-Ezh2 and anti-H3K27-Me3 antibodies. The drug-selected progenitors were obtained as in Fig. 4d and the result using ChIP primer set "d" shown. 5% of total chromatin or ChIP with α-H3K4-Me^2 are the input and positive controls.

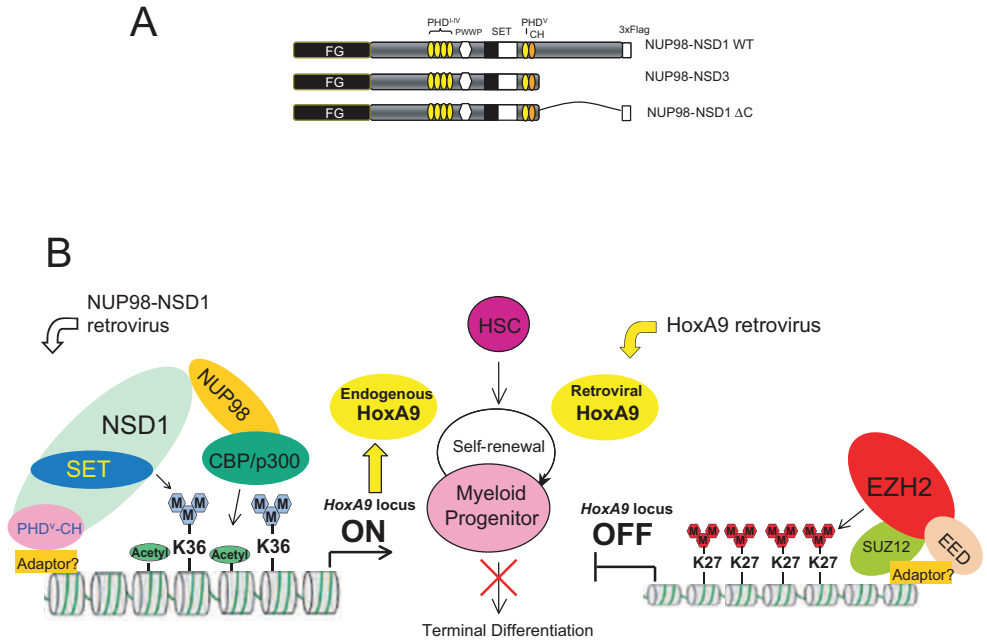
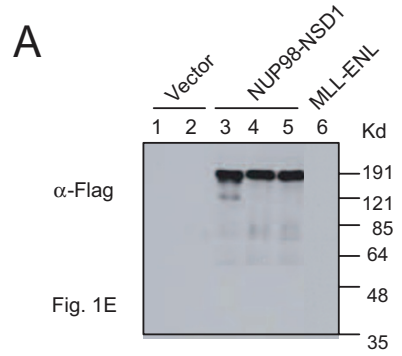
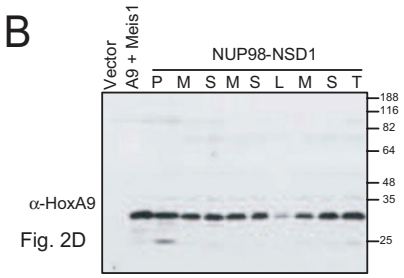
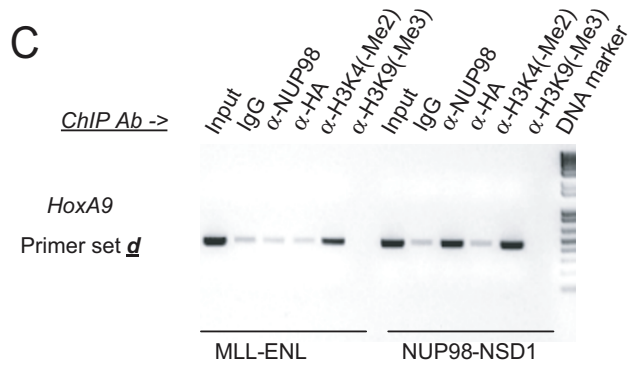
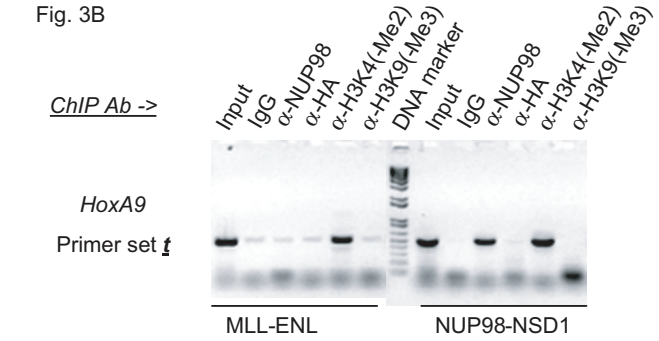


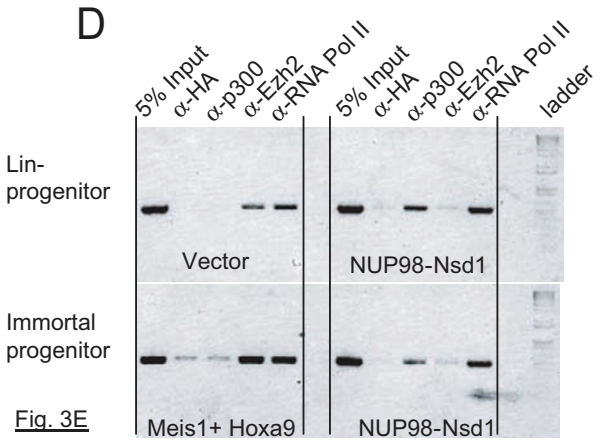
Figure S3. (a) Comparison of domain conservation between NUP98-NSD1 and NUP98-NSD3. The ΔC deletion of NUP98-NSD1 removes the same region of NSD1 that is missing in NUP98-NSD3. (b) Model for enforced transcription of the endogenous *HoxA9* gene by NUP98-NSD1, resulting in sustained progenitor self-renewal. In NUP98-NSD1-expressing progenitors, transactivation of the *HoxA9* gene is maintained by co-localization of H3K36 methylation catalyzed by the NSD1 SET domain and histone acetylation catalyzed by NUP98-associated CBP/p300. These combined epigenetic marks prevent EZH2-mediated H3K27 methylation, which silences the *Hox-A* locus during normal myeloid differentiation and which down-regulates *Hox-A* locus transcription in myeloid progenitors immortalized by retroviral expression of *HoxA9* plus *Meis1*.

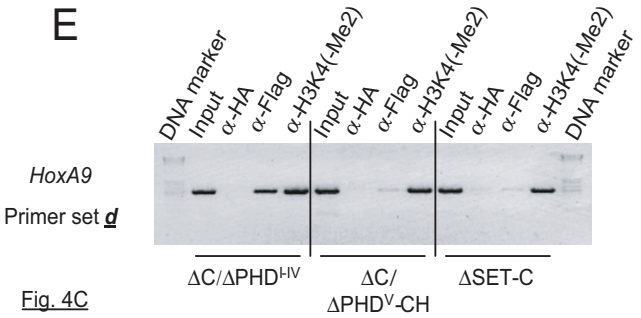












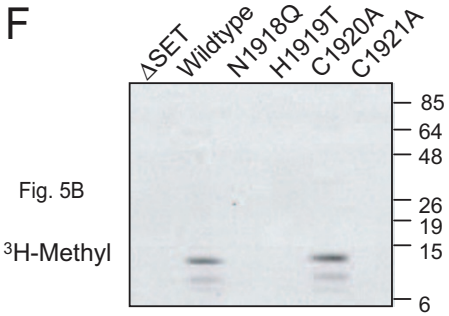
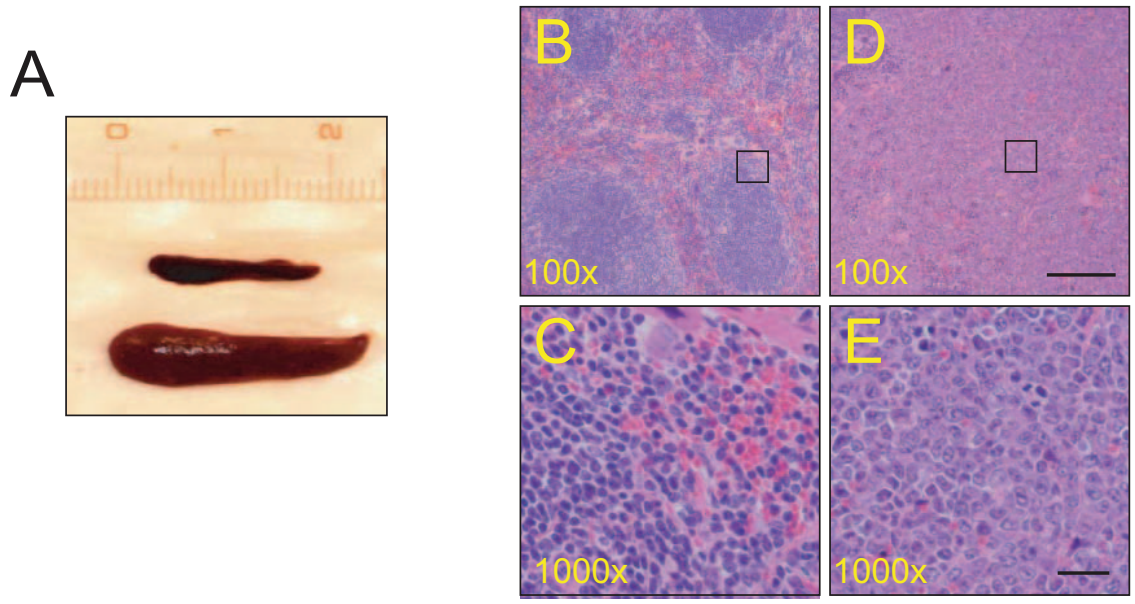


Figure S4. Full scans of (a) immunoblot detection of NUP98-NSD1 shown in Fig. 1e, (b) immunoblot detection of HoxA9 in progenitors immortalized by NUP98-NSD1 shown in Fig. 2d, (c) representative ChIP analysis using primer sets "d" and "t" to interrogate binding of NUP98-NSD1 and histone modifications shown in Fig. 3b, (d) ChIP analysis of p300 and Ezh2 binding to *HoxA9* shown in Fig. 3e, (e) ChIP analysis of binding of NUP98-NSD1 mutants to *HoxA9*, (f) autoradiogram of histone methylation by the wild-type and mutant versions of NSD1 shown in Fig. 5b.

Supplementary Table 1. Phenotypes of AML induced by NUP98-NSD1.

Mouse ID	WBC[1] 103/uL	Myeloid[2] (%)	Lymphoid (%)	Hematocrit %	Spleen (mg)	Node (mg)	Thymus (mg)	Latency[3] (days)	FACS[4]: bone marrow (positive%)					FACS: spleen (positive%)				
									MacI+	B220+	CD19+	CD34+	Flt3+	MacI+	B220+	CD19+	CD34+	Flt3+
1	24.6	84	16	40.8	605	31	50	63	82	21	1	37	73	N/D[5]	N/D	N/D	N/D	N/D
2	137	92	8	38.6	1300	20	66	65	70	28	2	64	51	69	19	1	65	75
3	160	16	84	18.1	616	18.3	111	65	62	28	2	82	87	68	24	10	64	80
4	137	77	23	43.1	495	25	27	147	82	12	1	58	41	73	18	10	44	22
5	113	70	30	42.3	685	51	41	147	83	6	1	32	60	84	10	10	16	21
6	17.9	76	24	42.9	280	20	32	132	60	65	0	70	80	50	44	23	43	40
7	112	98	2	30.2	661	40	40	160	81	4	1	25	30	70	8	8	18	16
8	14.9	51	48	29.3	450	15	30	200	76	8	1	64	20	64	14	10	55	12

[1]WBC (1,000/ul), white blood cell count in one microliter of peripheral circulating blood;
[2]Percentage of myeloid lineage cells out of total white blood cells that were found in circulating blood;
[3]Latency of AML, duration from day of bone marrow transplantation to day of identification of leukemia;
[4]FACS analysis of leukemia cells extracted from AML mice, showing the percentage of cells scored as positive signals using antigen-specific antibodies (for example, MacI-positive cells comprise 82% of total cells in bone marrow of mouse ID#1);
[5]N/D, not determined.



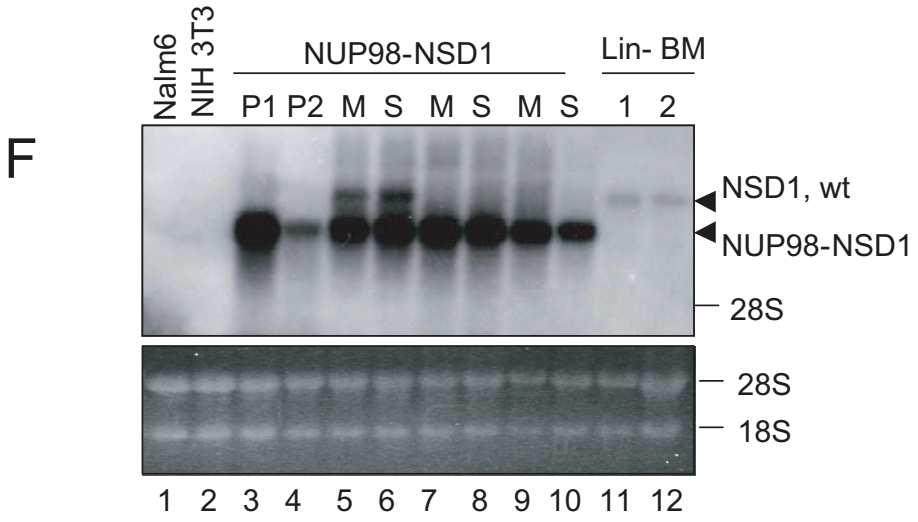


Figure S1. Phenotypic analysis of AML induced by NUP98-NSD1 in bone marrow transplantation models. (a) Typical spleen size in a mouse presenting with NUP98-NSD1-induced AML (lower) compared with that from a cohort control injected with Lin- progenitors expressing empty vector (upper). (b-e) Hematoxylin and Eosin (H&E) staining of spleen sections demonstrates that the follicles characteristic of normal splenic architecture (b and c) are disrupted in the spleens of mice bearing NUP98-NSD1 AML (d and e). Scale bars represent 250 micrometers in panels B and D, and 30 micrometers in panels c and e. (f) Northern blot using an NSD1 probe verifying expression of retroviral NUP98-NSD1 in infected parental progenitors (P1 and P2; lane 3-4) and in leukemic progenitors exacted from bone marrow (M) or spleen (S) of AML-bearing mice injected with P1 or P2 (lanes 5-10). Endogenous NSD1 was detected in each of two uninfected cultures of Lin^- bone marrow progenitors (lane 11-12), but not in Nalm-6 pre-B cell lines (lane 1) or in NIH-3T3 fibroblasts (lane 2). Bottom panel represents the agarose gel image of RNA samples demonstrating equal loading of RNA. The position of 28S and 18S RNA are indicated.

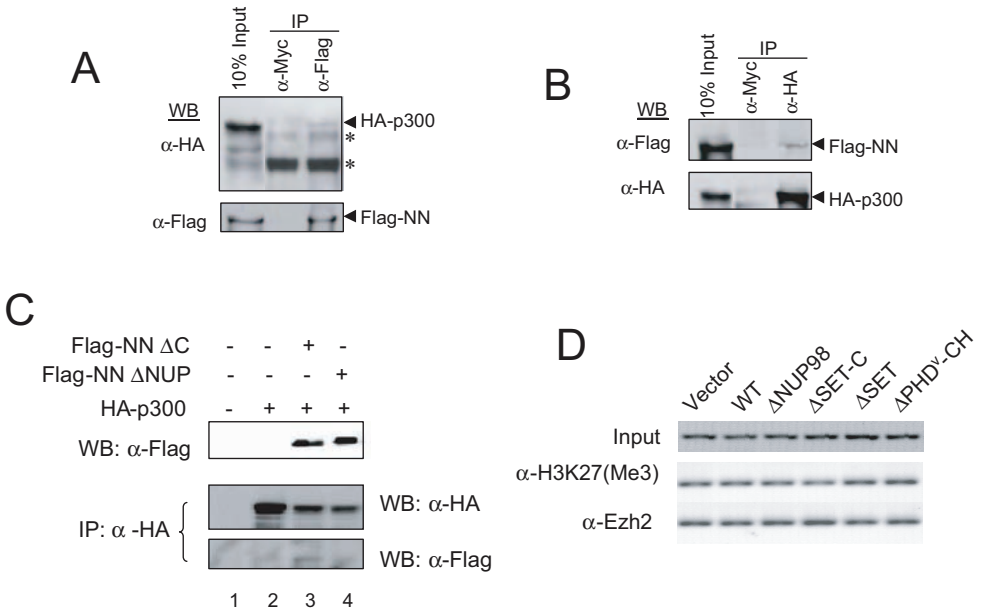


Figure S2. (a-c) Interaction of CBP/p300 with NUP98-NSD1 is mediated by NUP98 sequences. (a) α-Flag immunoprecipitation (IP) of NUP98-NSD1 (NN, Flag-tagged) co-precipitates p300 (HA-tagged), identified by α-HA western blotting, in cotransfected 293T cells. Co-precipitation is not observed in the α-Myc sera control. Stars represent nonspecific bands. The bottom panel represents an α-Flag western blot demonstrating that NUP98-NSD1 is efficiently immunoprecipitated. (b) α-HA IP of HA-p300 coprecipitates NUP98-NSD1 (NN, Flag-tagged), which is not precipitated by α-Myc sera. The bottom panel represents an α-HA western blot demonstrating that HA-p300 is efficiently immunoprecipitated. The positions of HA-p300 and NUP98-NSD1 are indicated at right. (c) Deletion of NUP98 (NN-ΔNUP98) abolishes the interaction of NUP98-NSD1 with CBP/p300 in HEK293 cells, while deletion of the C-terminal 50 KD (NN-ΔC) does not. The top panel demonstrates equal expression of NUP98-NSD1 by α-Flag western blot in total cell lysate, the middle panel demonstrates equal IP efficiency of HA-p300, and the bottom panel examines the presence of NUP98-NSD1 in precipitated samples by α-Flag western blotting. (d) ChIP analysis of Ezh2 binding and H3K27 methylation on the *HoxA9* promoter in the early hematopoietic progenitors transduced with wild-type or mutant NUP98-NSD1, using α-Ezh2 and anti-H3K27-Me3 antibodies. The drug-selected progenitors were obtained as in Fig. 4d and the result using ChIP primer set "d" shown. 5% of total chromatin or ChIP with α-H3K4-Me^2 are the input and positive controls.

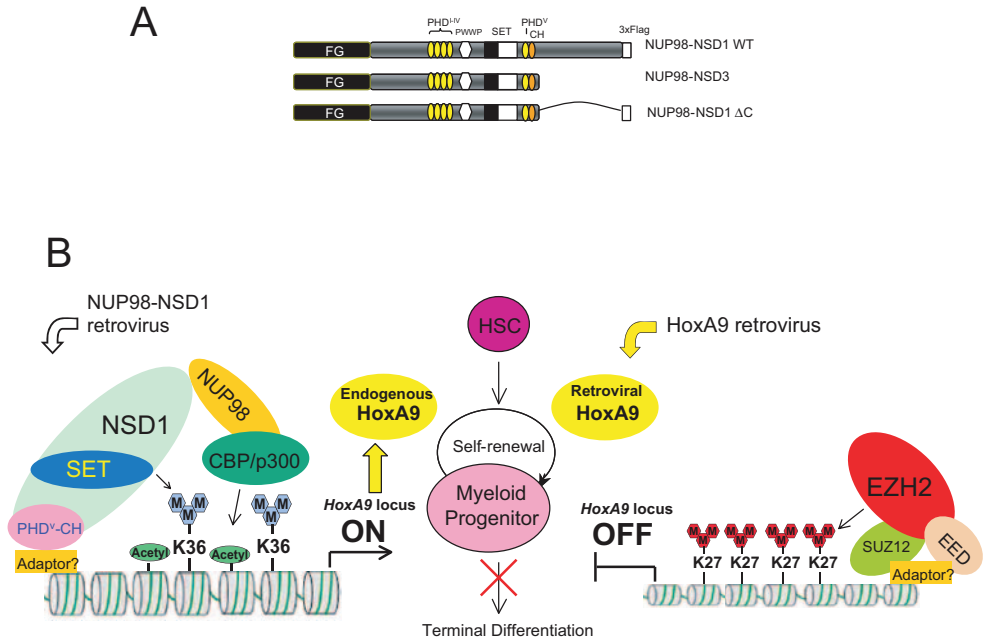
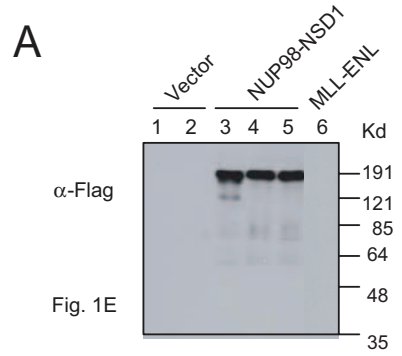
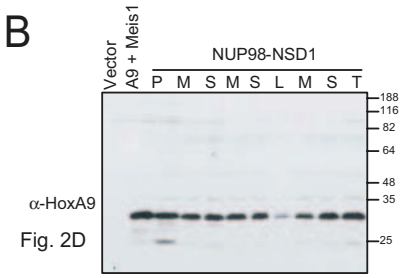
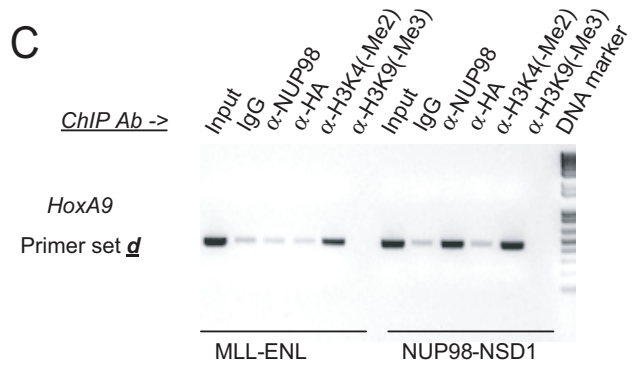
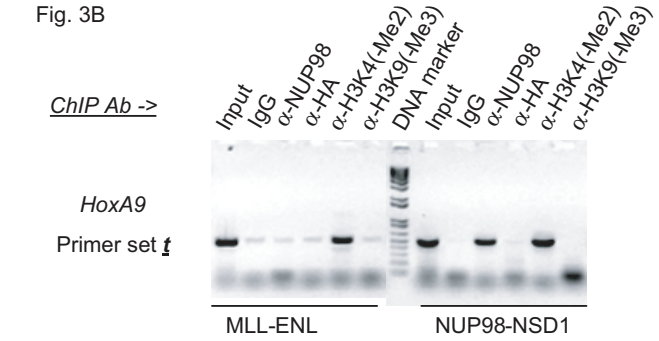


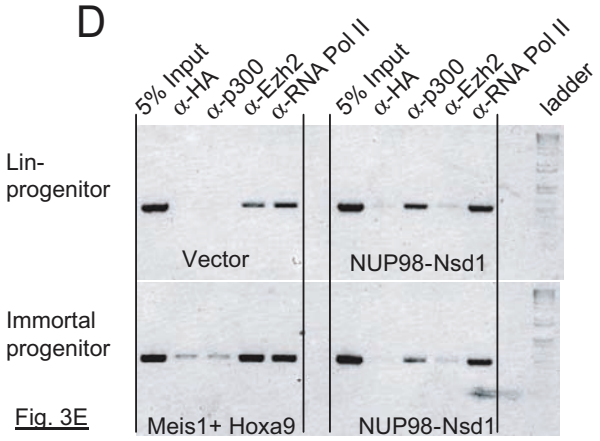
Figure S3. (a) Comparison of domain conservation between NUP98-NSD1 and NUP98-NSD3. The ΔC deletion of NUP98-NSD1 removes the same region of NSD1 that is missing in NUP98-NSD3. (b) Model for enforced transcription of the endogenous *HoxA9* gene by NUP98-NSD1, resulting in sustained progenitor self-renewal. In NUP98-NSD1-expressing progenitors, transactivation of the *HoxA9* gene is maintained by co-localization of H3K36 methylation catalyzed by the NSD1 SET domain and histone acetylation catalyzed by NUP98-associated CBP/p300. These combined epigenetic marks prevent EZH2-mediated H3K27 methylation, which silences the *Hox-A* locus during normal myeloid differentiation and which down-regulates *Hox-A* locus transcription in myeloid progenitors immortalized by retroviral expression of *HoxA9* plus *Meis1*.

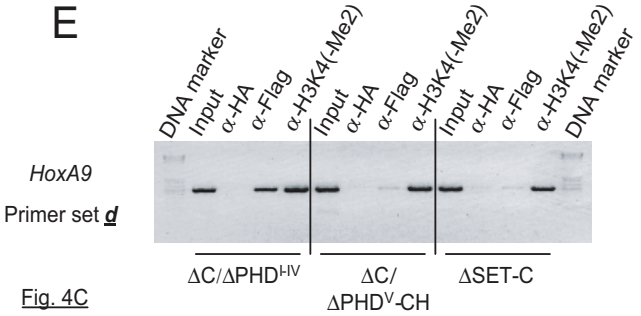












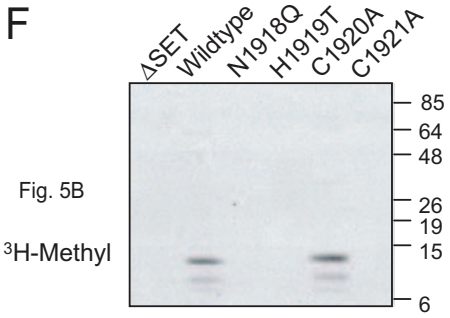


Figure S4. Full scans of (a) immunoblot detection of NUP98-NSD1 shown in Fig. 1e, (b) immunoblot detection of HoxA9 in progenitors immortalized by NUP98-NSD1 shown in Fig. 2d, (c) representative ChIP analysis using primer sets "d" and "t" to interrogate binding of NUP98-NSD1 and histone modifications shown in Fig. 3b, (d) ChIP analysis of p300 and Ezh2 binding to *HoxA9* shown in Fig. 3e, (e) ChIP analysis of binding of NUP98-NSD1 mutants to *HoxA9*, (f) autoradiogram of histone methylation by the wild-type and mutant versions of NSD1 shown in Fig. 5b.

Supplementary Table 1. Phenotypes of AML induced by NUP98-NSD1.

Mouse ID	WBC[1] 103/uL	Myeloid[2] (%)	Lymphoid (%)	Hematocrit %	Spleen (mg)	Node (mg)	Thymus (mg)	Latency[3] (days)	FACS[4]: bone marrow (positive%)					FACS: spleen (positive%)				
									MacI+	B220+	CD19+	CD34+	Flt3+	MacI+	B220+	CD19+	CD34+	Flt3+
1	24.6	84	16	40.8	605	31	50	63	82	21	1	37	73	N/D[5]	N/D	N/D	N/D	N/D
2	137	92	8	38.6	1300	20	66	65	70	28	2	64	51	69	19	1	65	75
3	160	16	84	18.1	616	18.3	111	65	62	28	2	82	87	68	24	10	64	80
4	137	77	23	43.1	495	25	27	147	82	12	1	58	41	73	18	10	44	22
5	113	70	30	42.3	685	51	41	147	83	6	1	32	60	84	10	10	16	21
6	17.9	76	24	42.9	280	20	32	132	60	65	0	70	80	50	44	23	43	40
7	112	98	2	30.2	661	40	40	160	81	4	1	25	30	70	8	8	18	16
8	14.9	51	48	29.3	450	15	30	200	76	8	1	64	20	64	14	10	55	12

[1]WBC (1,000/ul), white blood cell count in one microliter of peripheral circulating blood;
[2]Percentage of myeloid lineage cells out of total white blood cells that were found in circulating blood;
[3]Latency of AML, duration from day of bone marrow transplantation to day of identification of leukemia;
[4]FACS analysis of leukemia cells extracted from AML mice, showing the percentage of cells scored as positive signals using antigen-specific antibodies (for example, MacI-positive cells comprise 82% of total cells in bone marrow of mouse ID#1);
[5]N/D, not determined.