An MLL-dependent network sustains hematopoiesis

Erika L. Artinger,a Bibhu P. Mishraa, Kristin M. Zaffutoa, Bin E. Li,b Elaine K. Y. Chungb, Adrian W. Mooreb, Yufei Chenb, Chao Chengd,e,f, and Patricia Ernst*a,b,c,d,e,f

Departments of aGenetics and bMicrobiology and Immunology, Institute for Quantitative Biomedical Sciences, and cNorris Cotton Cancer Center, Geisel School of Medicine at Dartmouth, Hanover, NH 03755; and dDisease Mechanism Research Core, RIKEN Brain Science Institute, Wako City, Satima 351-0198, Japan

Edited by Janet D. Rowley, The University of Chicago, Chicago, IL, and approved May 14, 2013 (received for review January 21, 2013)

The histone methyltransferase Mixed Lineage Leukemia (MLL) is essential to maintain hematopoietic stem cells and is a leukemia protooncogene. Although clustered homeobox genes are well-characterized targets of MLL and MLL fusion oncoproteins, the range of MLL-regulated genes in normal hematopoietic cells remains unknown. Here, we identify and characterize part of the MLL-dependent transcriptional network in hematopoietic stem cells with an integrated approach by using conditional loss-of-function models, genomewide expression analyses, chromatin immunoprecipitation, and functional rescue assays. The MLL-dependent transcriptional network extends well beyond the previously appreciated Hox targets, is comprised of many characterized regulators of self-renewal, and contains target genes that are both dependent and independent of the MLL cofactor, Menin. Interestingly, PR-domain containing 16 emerged as a target gene that is uniquely effective at partially rescuing MLL-deficient hematopoietic stem and progenitor cells. This work highlights the tissue-specific nature of regulatory networks under the control of MLL/Trithorax family members and provides insight into the distinctions between the participation of MLL in normal hematopoiesis and in leukemia.

proliferation | HSC | epigenetics

Epigenetic regulation is an important mechanism by which gene expression fidelity is maintained during development. The trithorax-group (trx-G) and Polycomb-group (Pc-G) genes encode epigenetic factors that act as opposing regulators of clustered homeobox (Hox) gene expression and of axial patterning in most metazoans (1, 2). In addition, numerous studies implicate Pc-G and trx-G homologs in mammals in the maintenance of broader gene expression programs in embryonic and tissue stem cells and in cancer (1, 2). Because of the reversible nature of epigenetic lesions in cancer, targeting oncogenes and tumor suppressors that use epigenetic mechanisms is a promising approach for targeted therapy (3).

The human protooncogene Mixed Lineage Leukemia (MLL) was the first mammalian trx homolog identified because of its characteristic rearrangement in ~70% of infant leukemia. Rearrangement of the human MLL gene by chromosomal translocation also occurs at a lower frequency in childhood acute lymphoblastic leukemia (ALL), acute myelogenous leukemia (AML), and treatment-related and de novo AML in adults (4, 5). Most translocations produce MLL fusion oncoproteins that retain the chromatin-targeting N terminus and acquire a transcriptional effector domain from the C-terminal partner. Partner proteins frequently recruit protein complexes that differentially recruit protein complexes that control and function in MLL-dependent genes involved in maintaining HSCs, we analyzed differentially expressed transcripts after MLL deletion. Lineage-negative, stem cell antigen-1 (Sca-1)+, c-Kit+, Cdx4+/− (LSK/CD48−/+) HSC-enriched cells from the bone marrow (BM) of polyinosinic:polycytidylic acid (pI:pC)-injected control Mll+/− or Mll−/− animals were purified 6 d after the first pI:pC injection, the optimal timing for MLL deletion, cell yield, and down-regulation of homeobox protein a9 (Hoxa9), a bona fide Mll target gene (13). Assessment of normalized gene expression differences between control and Mll-deficient LSK/CD48−/+ cells revealed 1,935 differentially expressed genes using Significance Analysis of Microarrays (which does not impose a fold cutoff; Fig. 1) (20). Functional classification of genes differentially expressed in Mll-deficient HSCs compared with controls resulted in three global observations: (i) more genes were up-regulated than down-regulated, (ii) a subset...
of erythroid-specific genes were up-regulated, and (iii) the largest category of annotated down-regulated genes was comprised of transcriptional regulators.

Among the up-regulated genes, the largest group corresponds to HSC proliferation and ribosome or mitochondrial biogenesis (Fig. 1A and Dataset S1). Up-regulation of genes involved in ribosome biogenesis reflected the greater proportion of cycling Mll-deficient LSK/CD48<sup>neg</sup> cells (45% G<sub>0</sub> in Mll-deleted cells versus 75% G<sub>0</sub> in controls; ref. 13). Ten percent in this category and 17% in the mitochondrial group were also identified in proliferating HSCs (21), (Dataset S1). Thus, many of the up-regulated genes reflect the expected changes based on the proliferation state of Mll-deficient LSK/CD48<sup>neg</sup> cells. Unexpectedly, 5% of the genes that were up-regulated in Mll-deficient LSK/CD48<sup>neg</sup> cells encode erythroid-specific proteins including transcriptional regulators such as GATA1 binding protein 1 (Gal1) and Kruppel-like factor 1 (Klf1), as well as spectrin, Kell protein (Kel), Erythropoietin receptor (EpoR), and hemoglobin biosynthesis genes (Dataset S1). Gene set enrichment analysis (GSEA) also identified a GATA1-induced gene signature and a tendency toward erythroid identity (Fig. S1A and B). The up-regulation of erythroid genes was validated by using an independent in vitro Mll deletion system, illustrating that the scale of gene up-regulation was consistent with derepression rather than full induction of erythroid genes (Fig. S1 C and D). Furthermore, this derepression was not sufficient to impart erythroid fate as demonstrated by colony assay (Fig. S1E). Derepression of erythroid genes likely occurs through an indirect mechanism, thus we focused on the down-regulated genes as potential Mll effectors in the maintenance of HSCs.

**Identifying an Mll-Dependent Transcriptional Network.** Transcriptional regulators comprised the largest single annotated category of down-regulated genes in Mll-deleted LSK/CD48<sup>neg</sup> cells (Fig. 1B and Dataset S2). Because many of these regulators are highly expressed in HSCs relative to more differentiated cell types (22), we asked whether Mll-deficient HSCs exhibit a global shift in cell fate by assessing the relatedness of our gene expression data to other hematopoietic populations (23, 24). This analysis showed an enrichment of erythroid identity as described earlier, but did not suggest that HSCs were generally differentiated, because HSC and multipotent progenitor signatures were equivalently enriched by GSEA (Fig. S1F). Mll itself (Fig. S1G) and well-characterized Mll targets such as Hoxa9 were down-regulated although the majority of the genes in this category were not previously known to be Mll targets (Fig. 1C). We confirmed the Mll dependence for all annotated transcription factors >2.5-fold down-regulated by quantitative RT-PCR (RT-qPCR) using independently sorted samples from Mxl-cre;Mll<sup><sup>-/-</sup></sup> animals (Fig. 1D), as well as cells in which Mll was deleted in vitro by using 4-hydroxytamoxifen (4-OHT; Fig. 1E). Each inducible knockout model has its characteristic limitations, so to discover genes that were truly Mll-dependent, we only pursued genes down-regulated in both Mxl-cre and Er<sup>-cre</sup> systems. Of the annotated transcription factors down-regulated >2.5-fold (Fig. 1C), MDS and Evi1 complex locus (Mecom), Prdm16, Pax5, B cell leukemia homeobox protein 1 (Pbx1), Eyes absent homolog 1 (Eya1) and Hoxa9 were consistently Mll-dependent (Fig. 1E). Tripartite motif-containing 30b (Trim30b) is not characterized, so we focused on the other five genes for the following studies.

Several of the transcriptional regulators identified above individually play critical roles in HSC homeostasis. For example, the proteins encoded by the Pbx1, Prdm16, and Mecom genes act to restrain HSC proliferation and/or promote self-renewal (25–29), as has been demonstrated for Mll (13, 18). Interestingly, Mecom and Prdm16 were not Mll-dependent in fibroblasts or in Mll knockout embryos overall, despite coexpression of Mll and these genes (Fig. S2).

**Mll Binds Directly to the Promoter Regions of a Subset of Mll-Dependent Genes.** Mll and its homolog Trihotax typically act to maintain expression of their direct target genes (30), thus we evaluated the down-regulated transcription factors as potential direct Mll targets. To assess whether Mll acts directly to promote expression of the identified transcriptional regulators, we used a mini-ChIP procedure optimized for ∼10<sup>6</sup> BM cells (31). Based on previous results demonstrating Mll binding near transcription start

---

**Fig. 1.** Identification of Mll-regulated genes in HSCs. General overview of genes up-regulated (A) or down-regulated (B) in Mll-deficient LSK/CD48<sup>neg</sup> cells compared with controls. Cells were sorted from pLPC-infected control Mll<sup>F/F</sup> or Mxl-cre;Mll<sup>F/F</sup> animals at day six. Gene Ontology assignments were based on the criteria in Datasets S1 and S2. (C) The top down-regulated transcription factors in Mll-deficient LSK/CD48<sup>neg</sup> cells listed by fold reduction (see also Dataset S2). (D) RT-qPCR validating down-regulated genes in independent control Mll<sup>F/F</sup> (blue) or Mll-deficient (red) LSK/CD48<sup>neg</sup> cells, n = 8 animals per genotype; ND, not detected. (E) RT-qPCR validation of transcripts in LSK cells sorted from control Er<sup>-cre</sup>;Mll<sup>F/F</sup> (blue) or Er<sup>-cre</sup>;Mll<sup>-/-</sup> (red) cultured for 72 h after initiating Mll deletion. Relative expression levels were determined by normalizing to Gapdh, n = 4 animals per genotype. Error bars represent 95% confidence interval (CI). *P < 0.05, **P < 0.01. Er<sup>-cre</sup>, estrogen receptor<sup>2β</sup> mutant fused to cre recombinase.
sites (TSS) in cell lines (32, 33), we assessed MLL binding within 2 kb of the TSS by using 3–5 amplicons per gene. MLL-dependence was similarly observed in the BM lineage-negative (lin−) population and LSK cells (Fig. S3A). Control ChIP experiments demonstrated MLL binding to the Hoxa9 but not Gapdh TSS regions (Fig. S3B). Using lin− BM cells, we observed specific MLL binding around each TSS of the Mecom locus [both Myelodysplastic syndrome 1 (Mds1) and Ectropic virus integration site 1 (Evi1) promoter regions], as well as the Prdm16, Pbx1, and Eya1 genes (Fig. 2 and Fig. S3 C–G). Interestingly, we did not observe MLL binding to the Early growth response 1 (Egr1) promoter (Fig. 2B and Fig. S3H), consistent with the observation that this gene was not MLL-dependent in both model systems (Fig. 1E).

Therefore, we conclude that like Hoxa9, the expression of Mecom, Prdm16, Pbx1, and Eya1 is maintained directly by MLL in normal lin− BM cells.

Only a Subset of MLL-Dependent Genes Are Affected by Men1 Deletion. MLL itself does not harbor sequence-specific DNA binding motifs. One important chromatin-targeting mechanism occurs through an N-terminal interaction with Menin and p75/lens epithelium-derived growth factor (LEDGF), thought to be essential for targeting wild-type MLL to promoter regions based on studies using MLL fusion oncoproteins (34). To understand how the MLL complex localizes to its targets in HSCs, we assessed the Menin dependence of Egr1, Hoxa9, Prdm16, Mecom, Pbx1, and Eya1. Consistent with a previous study (35), we found that Hoxa9 expression was reduced in Menin (Men1)− deficient LSK cells. Interestingly, Mecom and Eya1 were slightly reduced, but the latter was not statistically significant (Fig. S4A). Despite efficient excision of Men1 (Fig. 3B), Prdm16 and Pbx1 levels were not affected (Fig. S4A), suggesting that a subset of HSC-specific MLL-dependent genes do not require Menin. These data demonstrate that the MLL complex differentially requires the Menin chromatin-targeting cofactor to regulate distinct classes of target genes.

Structure of the MLL-Dependent Transcriptional Network. We considered that some of the MLL-dependent transcriptional regulators act in interconnected pathways to modulate HSC function. For example, it has been reported that overexpression of Evil up-regulates Pbx1 in c-Kit−enriched BM cells (36). To identify potential expression interrelationships and determine whether the identified genes represent a linear or branched pathway downstream of MLL, we overexpressed Hoxa9, Prdm16, Eya1, Pbx1, or Mecom isoforms (Mds1−Evil and Evil) in wild-type or MLL-deficient LSK cells and assessed the effect on other genes in this network 48 h later. Focusing first on the effects of overexpression in wild-type cells, we found that Hoxa9 could increase levels of Prdm16, Evil could increase both Prdm16 and Hoxa9, and Prdm16 could increase Hoxa9 levels. For MLL-deficient LSK cells infected with the empty retrovirus, we observed reduced expression of Hoxa9, Prdm16, Mecom, Pbx1, and Eya1 (Fig. 4, empty) as observed in unmanipulated MLL-deficient LSK cells (Fig. 1). However, reexpression of Hoxa9, Prdm16, Eya1, or Pbx1 did not restore expression of the other tested genes to wild-type levels in MLL-deficient LSK cells (Fig. 4). In contrast, expression of either of the Mecom isoforms altered the expression of other genes in this network in MLL-deficient LSK cells. Evil expression increased Prdm16 and Hoxa9 transcripts in MLL-deficient LSK cells back to the wild-type levels (Fig. S4 A and B). Mds1-Evil suppressed Prdm16. Hoxa9, Pbx1, and Eya1 expression in wild-type cells to the low levels observed in MLL-deficient LSK cells (Fig. 4 A, B, D, and E), consistent with previous observations that Mds1-Evil and Evil have opposing activities on hematopoietic differentiation and cytokine-stimulated growth (37, 38). These data illustrate that overexpression of individual transcription factors can influence the expression levels of other regulators in this network primarily in wild-type LSK cells, yet in most cases cannot restore normal levels of any of the network genes in MLL-deficient cells. The exception is Evil, which is capable of restoring the expression of two of the five genes in this network in MLL-deficient LSK cells. Taken together, these data exclude that these transcriptional regulators are organized in a linear pathway downstream of MLL and, instead, suggest that they each perform independent functions as downstream effectors of MLL.

**Prdm16 Exhibits a Unique Capacity to Partially Rescue MLL-Deficient Cells.** One to two weeks after inducing cre, the attrition of BM cells in Mx1-cre;Mll−/− animals results in animal death accompanied by multiple defects in HSPCs (13). To evaluate the relative functional importance of the identified Mll targets, we assessed whether re-expression of individual genes could rescue MLL-deficient cell attrition from BM chimeras. To this end, the Mll target genes identified above were overexpressed individually in sorted LSK cells from uninduced control Mll+/− or Mx1-cre;Mll−/− mice, then engrafted into lethally irradiated recipients together with uninfected wild-type BM cells. After stable engraftment, Mll excision was induced by plpC injection and the persistence of MLL-deficient BM cells expressing the reintroduced gene was determined 2 wk later (Fig. 5A). Thus, in this assay, “rescue” is defined as the selective persistence of retrovirus-infected cells within the population of Mll-deleted cells (Fig. S4C). The use of Mll itself as a positive control was precluded by the large size of the Mll transcript (>11 kb), because it could not be packaged into a retrovirus.

Upon Mll deletion, uninfected or empty retrovirus-expressing donor cells were lost rapidly, and Mll−/− hematopoietic cells from untreated lethally irradiated recipients were lost with a time course consistent with previous observations that Mds1-Evil and Evil have opposing activities on hematopoietic differentiation and cytokine-stimulated growth (37, 38). These data illustrate that overexpression of individual transcription factors can influence the expression levels of other regulators in this network primarily in wild-type LSK cells, yet in most cases cannot restore normal levels of any of the network genes in MLL-deficient cells. The exception is Evil, which is capable of restoring the expression of two of the five genes in this network in MLL-deficient LSK cells. Taken together, these data exclude that these transcriptional regulators are organized in a linear pathway downstream of MLL and, instead, suggest that they each perform independent functions as downstream effectors of MLL.

**Prdm16 Can Correct the Intrinsic Proliferation Defect of MLL-Deficient HSCs.** To determine the mechanism by which Prdm16 partially rescued Mll-deleted BM cells, we examined the consequences of
**Prdm16** reexpression on LSK cell proliferation. We demonstrated that more **Mll**-deleted LSK cells are in S phase compared with wild-type, and that the CD48<sup>Ho</sup> subset of these cells were largely in G<sub>1</sub>/S rather than G<sub>0</sub> (13). Thus, we first assessed whether we could recapitulate any aspects of the hyperproliferative phenotype in vitro, then assessed the impact of Prdm16 in this setting.

To directly assess proliferation kinetics in vitro, wild-type (**Mlf<sup>F/F</sup>**) or **Mll**-deleted (**Mx1-cre;Mll<sup>F/F</sup>**) LSK/CD48<sup>Ho</sup> cells were sorted from Flrt-gC-treated animals, deposited into wells as single cells and cultured in serum-free medium containing cytokines to maintain HSC identity and function (39) (Fig. 6A). Importantly, the percentage of surviving clones was similar between wild-type and **Mll**-deleted cells (**Fig. S5A**), confirming previous observations that apoptosis is not induced in **Mll**-deleted HSPCs (13). Integrating individual observations for 158 wild-type and 240 **Mll**-deleted LSK/CD48<sup>Ho</sup> cells, we found that the proliferation kinetics of the latter were consistently more advanced than wild type (Fig. 6E). After 48 h, the mode (greatest number of cells) of **Mll**-deleted LSK/CD48<sup>Ho</sup> clones had progressed approximately one-half a division further than the wild-type clones (Fig. 6E), and by 72 h, the mode was one full cell division ahead (Fig. 6D). To address the possibility that **Mll**-deficient LSK/CD48<sup>Ho</sup> cells exhibit earlier cell division because more are initially in G<sub>1</sub>/S compared with wild type, we performed higher resolution studies examining the initial three cell divisions (Fig. 6E). We found that **Mll**-deficient LSK/CD48<sup>Ho</sup> cells enter the cell cycle earlier at all cell divisions; in fact, **Mll**-deficient cells had a shorter cell cycle (~1 h) than wild-type cells (**Fig. S5B**). Therefore, **Mll**-deficiency results in a cell-intrinsic increase in proliferation that is recapitulated in vitro in conditions that maintain HSC identity. This system likely models the increased proportion of LSK cells in S phase we observed in vivo but does not represent the defect in maintaining G<sub>0</sub> (13).

To investigate whether **Prdm16** reexpression influenced the proliferation phenotype observed in **Mll**-deficient cells, we sorted LSK cells from control **ER-cre;Mlf<sup>F/F</sup>** and **ER-cre;Mlf<sup>F/F</sup>** mice, retrovirally introduced **Prdm16**, and concurrently incubated with 4-OHT to induce **Mll** deletion (**Fig. 6F**). **ER-cre;Mlf<sup>F/F</sup>** cells infected with an empty control retrovirus displayed greater cell accumulation than the **ER-cre;Mlf<sup>F/F</sup>** control cells, consistent with the single cell observations. However, **Prdm16** reexpression restored the growth of **Mll**-deficient LSK cells to within the normal range of the control LSK cells (**Fig. 6G**). Together, these data suggest that the mechanism by which **Prdm16** can correct **Mll** deficiency is, in part, by restraining proliferation within HSPCs.

**Discussion**

Using two complementary conditional knockout models (**Mx1-cre and ER-cre**), we have identified genes that are consistently **Mll** dependent in HSC-enriched cell populations. The acute nature of **Mll** deletion and the use of highly purified cells resulted in the identification of a succinct list of transcriptional regulators with a high level of reproducibility and enrichment for genes that control self-renewal and proliferation specifically in HSCs. Thus, we refer to this set of genes as core components of the MLL HSC-specific transcriptional network. Among the down-regulated genes, **Prdm16**, **Mecom**, **Pbx1**, **Eya1**, and **Hoxa9** emerged as a series of
interconnected Mll-regulated transcriptional nodes, with Prdm16 exhibiting the greatest activity to replace Mll function in HSCs. We tested these genes individually by overexpression to uncover dominant nodes downstream of Mll, but our data are consistent with the concept that this network functions coordinately to sustain HSC homeostasis through diverse functions, hence the inability of any individual gene to completely replace Mll in the gene expression or functional assays. In fact, each of these genes has distinct targets and loss-of-function phenotypes (25, 27–29, 40). Ultimately, identification of the minimal network of genes sufficient to replace Mll function will require simultaneous expression of physiologic levels of multiple genes.

Given the mechanisms by which MLL family members regulate gene expression, one surprising finding was the large number of up-regulated genes in Mll-deficient HSCs. However, the majority of these genes reflect the enhanced proliferation that we observe in Mll-deficient HSC-enriched populations in vivo, a finding that we also observe at single-cell resolution in the current study. The direct connection between Mll and enhanced proliferation in HSCs could be explained by three mechanistically distinct hypotheses. First, Pbx1, Mecom, and Prdm16 have all been suggested to suppress HSC proliferation, based on the analysis of hematopoietic populations in the corresponding knockout animals (25, 27, 29). Thus, the reduction in these three factors would be predicted to result in unrestrained proliferation, specifically in HSCs. Interestingly, responsiveness to TGFβ signaling is attenuated in hematopoietic cells from each of these knockouts (25, 29, 41), suggesting that the overall effect may have a significant impact on TGFβ signaling (Fig. S5 C and D). Alternatively, a distinct mechanism has been proposed to link Mll to proliferation in the setting of DNA damage. In this case, DNA damage-induced delay in origin of replication activation is enforced by wild-type MLL (42). In our conditional knockout system, it is possible that the loss of MLL (even in the absence of overt DNA damage) also results in unrestrained origin activation, a more rapid S phase, and shorter overall cell cycle duration. Finally, a recent demonstration that Mds1-Evi1 and Prdm16 are H3K9 monomethylases (43) suggests that global derepression of heterochromatinized genes could potentially have a broad impact on the suppression of proliferation or erythropoiesis in Mll-deficient HSCs.

By identifying this transcriptional network, we discovered three important features of this HSC-specific Mll pathway. First, some (e.g., Hoxa9, Mecom), but not all (e.g., Pbx1, Prdm16), of the direct Mll target genes also require the cofactor Menin. This finding illustrates that MLL uses distinct chromatin-targeting motifs for distinct categories of its direct target genes. Second, the genes identified here as Mll dependent in HSCs are not universally regulated by Mll in other tissues, with the exception of Hoxa9. This observation suggests that tissue-specific targeting and restriction mechanisms are behind the tissue-specific activity of MLL family members. Third, we note that not all of the HSC-specific Mll target genes are up-regulated in leukemia, possibly reflecting the distinction between the chromatin targeting/activation mechanisms used by fusion oncoproteins in contrast to those used by wild-type MLL. For example, it is clear that Hoxa9 is consistently overexpressed in MLL translocation leukemia, whether T-cell ALL (T-ALL), B-cell ALL (B-ALL), or AML (44–46). Evi1 and Eya1 have recently been implicated as targets of MLL fusion oncoproteins in some leukemia subsets (33, 47), but they are not consistently up-regulated in either ALL or AML harboring an MLL rearrangement. Prdm16 is not up-regulated in MLL-translocation leukemia yet can be activated by retroviral insertion in leukemia by translocation in other contexts, therefore has leukemogenic potential (48). Thus, our data begin to delineate a normal and reversible HSC-specific maintenance pathway, of which a selective portion is subverted to result in leukemia. Interestingly, Hoxa9, Mecom, and possibly Eya1 are the Mll-dependent genes we found to be affected by Menin loss, providing an intriguing connection between chromatin-targeting mechanism and leukemogenic versus normal HSC regulatory networks. The selective dependence on particular protein–protein interactions may render leukemia-specific gene programs driven by MLL-fusion oncogenes more sensitive to inhibitors than normal HSCs, as suggested by the study of compounds that disrupt the Menin–MLL interaction (49). Our work illustrates that MLL family members control exquisitely tissue-specific gene programs despite their ubiquitous expression.
patterns, underscoring the complexity of mechanisms that must be used to regulate diverse gene expression programs in vivo.

Materials and Methods

Mice and in Vivo Induction of CRE Reombinase. Mx1-cre;Mll\# animals and cre induction have been described (13). Mfn1\# mice (kind gift of Matthew L. Meyerson, Harvard Medical School, Boston, MA) were back-crossed by using the DartMouse speed congenic facility then crossed to the Ert-cre strain.

Flow Cytometry, Cell Sorting, and Culture. Flow cytometry and cell sorting were performed on a FACS Calibur and FACS aria, respectively (BD Biosciences). Data were analyzed by using Flowjo software (Tree Star). Fluorochrome-labeled antibodies and procedures are detailed in SI Materials and Methods.

Plasmids, Retroviral Infection, Cell Culture, and Transplantation. Murine stem cell virus (MSCV)-based retroviral expression plasmids were constructed by using cDNAs obtained or cloned as described in SI Materials and Methods. Viral supernatants were prepared by cotenatation, and sorted LSK cells were infected by using retroencon (Takara). Retrovirally infected cells were cotransplanted into lethally irradiated (950 Rad, split dose) C57Bl/6J female mice. For proliferation assays, LSK and LSKDCD48 cells were cultured in HSC expansion medium [StemSpan Serum Free Expansion Medium (SFEM); 300 ng/mL recombinant murine (r) SCF, 20 ng/mL rm-IL-7, and 4 ng/mL rmFt3L; StemCell Technologies and R&D Systems]. To induce deletion using the Ert-cre strain, HSC expansion medium was supplemented with 300–400 nM 4-OOHT (Sigma).

ChIP. Rabbit polyclonal anti-MLL C terminus (50) or anti-Gal4 (Santa Cruz; SC-218) or anti-MLL C terminus (50) or anti-Gal4 (Santa Cruz; SC-218). Unless indicated otherwise, the unpaired Student t test was used to determine significance, and error bars represent 95% CI. Statistical analyses were performed by using Excel (Microsoft) or Prism (GraphPad) Software.

ACKNOWLEDGMENTS. We thank R. Mako Saito, Chris Vakoc, Steve Smale, Hanna Mikkoika, Emmanuelle Passegué, and Adolfo Ferrando for critical comments; Thomas Milne and Joanna Attema for advice on ChIP; and Drs. Perkins, Morisihita, and Spiegelman for providing plasmids. E.L.A. and B.P.M. were partially supported by the Lady Tata Memorial Trust. E.K.Y.C. is a Japan Society for the Promotion of Science Foreign Postdoctoral Fellow, and A.W.M. was supported by a Japan Society for the Promotion of Science Grants-in-Aid Young Scientists (A) and a RIKEN Brain Science Institute core grant. This work was supported in part by National Institutes of Health Grants HL090036 and RR16437, American Cancer Society Grant RSG-10-242-LIB, and funds from the Gabrielle’s Angel Foundation for Cancer Research and Laura Strauss Leukemia Foundation.

Statistical Analyses. Unless indicated otherwise, the unpaired Student t test was used to determine significance, and error bars represent 95% CI. Statistical analyses were performed by using Excel (Microsoft) or Prism (GraphPad) Software.

Statistical Analyses. Unless indicated otherwise, the unpaired Student t test was used to determine significance, and error bars represent 95% CI. Statistical analyses were performed by using Excel (Microsoft) or Prism (GraphPad) Software.
Artinger et al. 10.1073/pnas.1301278110

Supporting Information

SI Materials and Methods

**Animal Strain Details.** All animal procedures were approved by the Institutional Animal Care and Use Committee of Dartmouth College. Mx1-cre;Mll mice were extensively back-crossed to the B6.SJL strain (B6.SJL-Ptpε<sup>[a]</sup> Pepκ<sup>[b]</sup>/BoyJ, stock no. 002014; Jackson Laboratory). ER-cre mice have an estrogen receptor<sup>12</sup> mutant fused to cre recombinase knock-in at the Rosa locus (Jackson Laboratory; stock no. 004847; ref. 1). The Dartmouth facility (Geisel School of Medicine at Dartmouth) was used to back-cross to the B6.SJL strain until >95% strain-specific SNPs were represented in the breeding animals. Murine embryo fibroblasts (MEFs) used in Fig. S24 were prepared from embryonic day (E)14.5 embryos by using standard methods and MBP<sup>0/0</sup> intercrosses. MEFs were then infected with a MIG-cre retrovirus, GFP<sup>+</sup> cells were sorted, Mll deletion was confirmed, and MEFs were passaged at least 20 times before quantitative RT-PCR (RT-qPCR).

**Cell Culture and in Vitro Induction of cre Recombinase.** Lineage-negative, Sca-1<sup>-</sup>, c-kit<sup>+</sup> (LSK) and LSK/CD48<sup>neg</sup> cells were cultured in hematopoietic stem cell (HSC) expansion medium, which is defined as StemSpan Serum Free Expansion Medium (StemCell Technologies) plus 300 ng/mL rmSCF, 20 ng/mL rm IL-7, and 4 ng/mL rmFlt3L (R&D Systems). To induce deletion using the ER-cre strain, HSC expansion medium was supplemented with 300–400 nM 4-hydroxytamoxifen (4-OHT) for 24 h (Sigma) to induce maximal deletion without harming cell viability. For CFU-E assays, sorted LSK cells were cultured in HSC expansion medium containing 4 U/mL Erythropoietin (Epo, Procrit; courtesy of Genentech, South San Francisco, CA) and 400 nM 4-OHT for 24 h, the medium was re-supplemented with HSC expansion medium plus Epo for an additional 24 h. Cells were seeded at 1 × 10<sup>5</sup> cells per 35-mm dish in M3434 (StemCell Technologies) supplemented with 2 U/mL Epo. After 3–4 d, CFU-E was scored.

**Viral Infection, Transplantation, and Rescue Assay.** LSK cells were sorted from CD45.1<sup>+</sup> Mx1-cre;Mll<sup>F/F</sup> or Mll<sup>0/0</sup> donors. To express genes in LSK cells, 96-well suspension plates were coated with 200 μg/mL Retronectin (Takara Bio) for at least 4 h and then incubated at 32°C, excess supernatant was removed, and cells were added and centrifuged for 90 min at 380 × g. Cell viability was determined by an accurate count of cells at every time point, 1,000 15.7-μm poly styrene polystyrene beads (Polysciences) were added to each well immediately after harvest. For isolation of HSCs, engrafted female C57BL/6J recipients. Four weeks after engraftment, depleted BM cells were injected peri-orbitally into lethally irradiated female C57BL/6J recipients. For isolation of HSCs-enriched populations, single-cell liquid culture assays, LSK/CD48<sup>neg</sup> cells from pI:pC-injected control Mll<sup>F/F</sup> or Mx1-cre;Mll<sup>F/F</sup> mice were sorted into collection tubes containing Hanks’ Balanced Salt Solution (HBSS; Mediatech) with 20% (vol/vol) FBS and then resorted at one per well into the individual wells of U-bottom 96-well plates (Nunc) containing 100 μL of HSC expansion medium. After sorting, plates were centrifuged briefly at 380 × g then incubated at 37°C in 5% (vol/vol) CO<sub>2</sub> for 2 h then scored for the presence of a single cell. The percentage of responding clones was calculated as the percentage of visually confirmed cells that ultimately divide at least once during 72 h of culture. For Fig. 6G, LSK cells were infected with retroviruses as described above in the presence of 300 nM 4-OHT for 36 h to induce Mll deletion, then ~500 retrovirally infected LSK cells were cultured HSC expansion medium in 96-well plates. To obtain an accurate count of cells at every time point, 1,000 15.7-μm poly styrene polystyrene beads (Polysciences) were added to each well immediately before harvest then the mixture was stained with anti-human CD4 antibody (anti-hCD4). Samples were collected for 30 s at the low setting of a FACS Calibur to enumerate hCD4<sup>+</sup> cells. An exact determination of cell number in each well was calculated by using a ratio of the number of beads collected in 30 s to the total number of beads seeded in each well.

**Flow Cytometry and Cell Sorting.** Sorting experiments were performed on a FACS Aria at the DartLab Flow Cytometry Shared Resource at the Geisel School of Medicine at Dartmouth. Fluorochrome-labeled antibodies used were as follows:

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Company</th>
<th>Clone</th>
<th>Fluorochrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>B220</td>
<td>Biolegend</td>
<td>RM2600</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>CD19</td>
<td>Biolegend</td>
<td>RM7700</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>CD3</td>
<td>eBiosciences</td>
<td>17A2</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>CD4</td>
<td>Biolegend</td>
<td>MCD0400</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>CD8</td>
<td>Biolegend</td>
<td>MCD0800</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>Ter119</td>
<td>Biolegend</td>
<td>MTER000</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>Gr1</td>
<td>Biolegend</td>
<td>RM3000</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>Mac1</td>
<td>Biolegend</td>
<td>RM2800</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>IL7Rα</td>
<td>eBiosciences</td>
<td>A7R34</td>
<td>Unlabeled</td>
</tr>
<tr>
<td>CD71</td>
<td>BD Pharmingen</td>
<td>C2</td>
<td>FITC</td>
</tr>
<tr>
<td>Ter119</td>
<td>BD Pharmingen</td>
<td>TER-119</td>
<td>PE</td>
</tr>
<tr>
<td>Goat anti-Rat F(ab)₂</td>
<td>Biolegend</td>
<td>N/A</td>
<td>Cy5-R-PE</td>
</tr>
<tr>
<td>Sca-1</td>
<td>BD Pharmingen</td>
<td>E13-161.7</td>
<td>FITC</td>
</tr>
<tr>
<td>c-Kit</td>
<td>Biolegend</td>
<td>28B8</td>
<td>APC</td>
</tr>
<tr>
<td>CD48</td>
<td>Biolegend</td>
<td>HM48-1</td>
<td>FITC</td>
</tr>
<tr>
<td>B220</td>
<td>Biolegend</td>
<td>RA6-8B2</td>
<td>APC</td>
</tr>
<tr>
<td>CD3</td>
<td>BD Pharmingen</td>
<td>145–2C11</td>
<td>PE</td>
</tr>
<tr>
<td>Mac1</td>
<td>BD Pharmingen</td>
<td>M170</td>
<td>APC</td>
</tr>
<tr>
<td>Gr1</td>
<td>BD Pharmingen</td>
<td>RB6-8C5</td>
<td>FITC</td>
</tr>
<tr>
<td>hCD4</td>
<td>Biolegend</td>
<td>OKT4</td>
<td>APC</td>
</tr>
<tr>
<td>FgRIII</td>
<td>eBiosciences</td>
<td>93</td>
<td>FITC</td>
</tr>
<tr>
<td>CD34</td>
<td>Biolegend</td>
<td>RAM34</td>
<td>PE</td>
</tr>
<tr>
<td>CD45.1</td>
<td>Biolegend</td>
<td>A20</td>
<td>PE</td>
</tr>
<tr>
<td>CD45.2</td>
<td>Biolegend</td>
<td>104</td>
<td>FITC</td>
</tr>
</tbody>
</table>

APC, allophycocyanin; BD, Beckton, Dickinson; Cy5-R-PE, R-Phycyoerythrin-Cyanine 5; F(ab)₂, fragment-antigen binding; FgRIII, Fc gamma receptor, type III; FITC, fluorescein isothiocyanate; hCD4, human anti-CD4; PE, phycerythrin.

For isolation of HSCs-enriched populations, single-cell suspensions were made from the hind limb bones (and pelvis for some experiments) by crushing with a sterile mortar and pestle. Lineage
staining was performed with a mixture of unlabeled lineage antibodies. Lineagé cells were depleted before sorting by using sheep anti-rat magnetic beads (Invitrogen) then stained with goat anti-rat Cy5-R-PE, c-kit APC, CD48 PE, and Sca-1 FITC. Retrovirally infected cells were detected by using anti-hCD4. Peripheral blood was collected from the peri-orbital sinus or cardiac puncture was collected in EDTA-coated tubes and analyzed immediately.

Cloning and Validation of Retroviral Plasmids. Mds1-Evil expression was accomplished by using the MIGR1ME retrovirus obtained from Archibald Perkins (University of Rochester Medical School, Rochester, NY) (3). The murine Evil cDNA was obtained from Kazuhiro Moroshita (Miyazaki University School of Medicine, Kihara, Japan) (4) as a pBluescript clone. The 4.5-kb EcoRI fragment was excised and inserted as a blunt fragment into the Hpal site of MSCV resulting in MSCV.Evil (Fig. S4E). This retrovirus expresses the Evil1 or p135 isoform as described (3, 5). Protein expression from this retroviral vector was confirmed by using anti-Evil from Santa Cruz Biotechnology [(20)-R; Fig. S4J]. The human Hoxa9 cDNA was obtained from Origene (IMAGE consortium clone 2987818, accession no. NM_152739, corresponds to the canonical 2 exon Hoxa9 ORF); the MIG-based Hoxa9 retrovirus described in Ernst et al. (6) was used to excise the cDNA as a BamHI-XhoI fragment, which was inserted into MSCV.hCD4 at the BglIII-Xhol sites. Protein expression was confirmed by immunoblot using Millipore anti-Hoxa9 (07-178, Fig. S4K).

The murine Pdm16 cDNA was obtained from Bruce Spiegelman (Harvard Medical School, Boston, MA) via Addgene. The cDNA was excised by using Xho-EcoRI and inserted into MSCV.hCD4 at the Xho and Hpa sites. This strategy resulted in excising the original Flag tag, but this was reintroduced by using the following annealed oligos: 5′-CTGCGCTTAGACTACAAGGACGATGACAAAAGG-3′ and 5′-TCGACCTTGTACGTGCTCTTTGATGTCGATGCG-3′. The hCD4-Pdm16 retrovirus was constructed by amplification of the Flag-tagged ORF with the following oligos: 5′-GAATTCATGGACTACAAAGACGATGAC-3′ and 5′-CTCGTCAAATAGC-3′, which were used in Fig. S4K. The murine Prdm16 cDNA was constructed by inserting the hCD4 cassette just downstream of the MSCV PucI and MluI sites. The bicistronic MSCV.hCD4 retroviral 3-kb fragment inserted into the MSCV.hCD4 plasmid at the SacII-XhoI fragment, which was inserted into MSCV.hCD4 at the BglIII-Xhol sites. Protein expression was confirmed by immunoblot using Millipore anti-Prdm16 (07-178, Fig. S4K).

Detailed Microarray Sample Preparation and Data Analysis. Total RNA was prepared from 1,500 to 10⁴ LSK/CD48́ cells sorted from five MLC-6 mice and five M13-cre; MLC-6 d after cre induction. Sorted cells were centrifuged and resuspended in TRIzol (Invitrogen), and total RNA was further purified with RNeasy columns (Qiagen) by following the manufacturer’s recommendations. RNA quantity and quality was determined by using an Agilent Technologies 2100 Bioanalyzer. RNA was amplified by using the MessageAMP II RNA Amplification Kit (Ambion), labeled with using the BioArray HighYield RNA Transcript Labeling Kit (T7, Enzo Life Sciences), fragmented and hybridized to GeneChip Mouse Genome 430 2.0 Arrays at the Dartmouth Medical School Genomics Shared Resource (http://geiselmed.dartmouth.edu/dgml). Raw intensity data for each probe set was collected with Microarray Suite Version 5.0 software (Affymetrix). GC-Normalized Robust Multi-Array Averaging (GCRMA) normalization and expression value calculation were performed by using BRB Array Tools Version 4.1. Significant changes in gene expression were identified by subjecting unfiltered expression values to Significance Analysis of Microarrays (7), accessed through BRB Array Tools, with a false discovery rate of 10% and 95% confidence intervals. For Gene Ontology assignment, probe sets were manually annotated into functional categories by using a combination of hematopoietic lineage and proliferation fingerprints (8–16). Gene Ontology analysis and functional annotation clustering (DAVID; http://david.abcc.ncifcrf.gov). Gene Set Enrichment Analysis (GSEA) (17) was performed with software available from the Broad Institute (http://www.broadinstitute.org/igsa/index.jsp). In the analysis, differential expression of genes in the C2 collection of gene sets from the MSigDB database was investigated by comparing Mll-deficient LSK/CD48́ and wild-type CD48́ cells. For specific comparisons with purified hematopoietic populations, we performed GSEA by using relevant gene sets compiled by Novershtern et al. (18) and by He et al (19). The former contains gene sets that are up- or down-regulated in purified human hematopoietic populations. The latter contains gene sets that are preferentially expressed in BM HSC compared with BM CD48́ LSK cells, in fetal liver HSC compared with BM HSCs, and in BM CD48́ LSK cells compared with BM HSCs.

Quantitative PCR and Detection of Mll and Menin Transcripts. Total RNA from the sorted populations indicated was described as above. If necessary, mRNA was amplified with the RiboAmp RNA Amplification Kit (Arcturus) or the MessageAMP II aRNA Amplification Kit for one or two rounds of amplification. cDNA was reverse transcribed by using SuperScript III (Invitrogen). Mll transcripts were quantified by using a custom Taqman assay the following primers 5′-TTCTCTGCAAATAGCCCTGC-3′, 5′-CTCCTTGTCTTCTTCAACGG-3′, and 5′-TTTGAATTTACATTGATCCCATGCG-3′, and probe: 5′-FAM-TCTCTTCCATGTTCCACCCGAG-TAMRA-3′. Men1 transcripts were quantified by using 2x SYBR Master Mix (Bio-Rad) and the following primers: forward, 5′-TCCCTTCTTCACGTTACAACA-3′ and reverse, 5′-ACCCAAGCATGATCCTCAG-3′. Relative expression levels of transcripts were determined by use of the ΔΔCT method (20) with data from duplicate or triplicate reactions normalized to Gapdh or RNA4 transcripts, as specified in the figure legends. Primers or assays for other genes are shown in Dataset S4.

Anti-MLL Chip. Cells were fixed in PBS containing 0.5 mM ethylene glycol-bis(succinic acid N-hydroxysuccinimide ester) (EGS; Thermo Scientific) for 20 min on a nutator. Cells were centrifuged, fixative removed, pellet resuspended in 1% formaldehyde and incubated 10 min, followed by centrifugation and resuspension in 50 mM glycine/PBS, a 10-min incubation, then a PBS incubation for 10 min. All steps were performed at room temperature at a cell density of 1 million per mL. Fixed cell pellets were either processed immediately or stored at −80 °C. To shear the chromatin,
the cell pellet was resuspended in lysis buffer (Tris at pH 7.5, 1 mM EDTA 1% SDS, 1x protease inhibitor complex; Roche) at 5 × 10⁷ cells per 20 μL. Low-surface-retention barrier tips were used for all steps (CLP Neptune). Sonication was performed for 10 cycles (30 s with 30-s rest) by using a Bioruptor UCD-200 (Diagenode). Sonicated chromatin was centrifuged at 13,000 g × 4 °C for 5 min, and the supernatant was diluted 10-fold with 2x RIPA buffer [20 mM Tris at pH 7.5, 2 mM EDTA, 2% (vol/vol) Triton X-100, 0.1% SDS, 0.2% sodium deoxycholate, and 200 mM NaCl]. For each ChIP reaction, 200 μL of diluted chromatin was incubated with 1 μg of antibody overnight at 4 °C, then 7.5 μL each of protein A and protein G Dynabeads (Invitrogen), previously washed in 1x RIPA buffer, were added to each immunoprecipitation and incubated for additional 2 h at 4 °C. The bead-protein complexes were washed three times with 200 μL of 1x RIPA buffer and once with 200 μL of TE (10 mM Tris at pH 7.5 and 1 mM EDTA). Genomic DNA was eluted from the ChIP and input samples for 3 h at 65 °C in 300 μL of elution buffer (20 mM Tris at pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 μg/mL proteinase K) by using an Eppendorf Thermomixer at 1,000 rpm. Samples were phenol/ chloroform extracted, ethanol precipitated with 10 μg each linear acrylamide, and glycogen then was centrifuged at 13,000 × g for 20 min at 4 °C. Pellets were air dried and resuspended in 15 μL of TE containing 0.1 mM EDTA. ChIP enrichment was determined by quantitative PCR using 2x SYBR green master mix (Bio-Rad). Enrichment of ML1 at genomic loci was expressed as the percent input by using the following formula: % of total input = 100 × 2^\[\text{Ct (ChIP)} – \text{Ct input} – \log_2 \text{(input dilution factor)}\]] (21).

**In Situ Hybridization.** Embryos were generated by crossing *Mll*Δ/Δ animals. The presence of a vaginal plug at 8 a.m. the next morning was defined as E0.5. E10.5 embryos were dissected from the yolk sac, which was used for genotyping, and the embryo was fixed in 4% paraformaldehyde/PBS at pH 7.4. Embryos were subjected to whole mount in situ hybridization as described (22, 23). Embryos were photographed with a Nikon DS-L1 camera mounted on an Olympus SZX16 stereomicroscope.

Fig. S1. Up-regulation of erythroid genes, not fate, in Mll-deficient HSC-enriched populations. (A) GSEA was used to identify gene sets with significant concordant gene expression differences compared with Mll-deficient LSK/CD48<sup>−/−</sup> cells versus wild-type CD48<sup>−/−</sup> cells. Shown are all gene sets with a false discovery rate (FDR) <20% in the C2 collection of curated gene sets; NES, normalized enrichment score. (B) GSEA plot showing the enrichment of Gata1-induced genes (1) in the Mll-deficient LSK/CD48<sup>−/−</sup> dataset. (C) RT-qPCR validation of select up-regulated, erythroid-specific genes in LSK cells sorted from control ER-cre;Mll<sup>−/−</sup> (blue) or ER-cre;Mll<sup>F/F</sup> animals (red) cultured for 24 h in HSC expansion medium with 400 nM 4-OHT, then an additional 24 h in HSC expansion medium. Data represents relative expression levels normalized to Gapdh, error bars represent 95% CI, *P ≤ 0.07, **P ≤ 0.02, ***P ≤ 0.01. (D) GSEA was used to identify gene sets with significant gene expression differences compared with Mll-deficient HSC to illustrate the scale of derepression compared with the level of induction observed during erythropoiesis. HSC transcript levels are compared with levels in common myeloid progenitors (CMP), myeloid-erythroid progenitors (MEP), proerythroblasts (proE), and erythroblast fractions A and C. (E) Average CFU-E from LSK cells sorted from control ER-cre;Mll<sup>F/F</sup> or ER-cre;Mll<sup>−/−</sup> animals. LSK cells were cultured in HSC expansion medium plus erythropoietin and 4-OHT for 24 h, HSC expansion medium for an additional 24 h then plated in semisolid medium (M3434; StemCell Technologies) for colony enumeration 3 d later. Data represent average CFU-E ± SD, n = 4 mice per genotype. (F) GSEA analyses comparing purified hematopoietic populations enriched in the Mll-dependent gene set using human (3) and murine (4) purified populations. Gene sets are ordered by P value (NOM p-val), a significantly related erythroid expression is outlined in red. Arrows highlight that Mll-deficient HSC are not significantly closer to BM HSC (LSK/CD48<sup>−/−</sup>CD150<sup>−</sup>) than they are to BM LSK/CD48<sup>−/−</sup> progenitor cells. (G) Efficiency of Mi1 deletion in control ER-cre;Mll<sup>F/F</sup> (blue) or ER-cre;Mll<sup>−/−</sup> animals (red) cultured for 24 h in HSC expansion medium with 400 nM 4-OHT, then an additional 24 h in HSC expansion medium. Data represents relative Mi1 expression levels normalized to Gapdh. Error bars represent 95% CI, n = 4 animals per genotype; *P ≤ 0.07, **P ≤ 0.02.

Fig. S2. HSC-specific Mll target genes are not Mll-regulated in nonhematopoietic tissues. (A) RT-qPCR results comparing gene expression levels in wild-type (blue) and Mll-deficient murine embryo fibroblasts (MEFs) (MllΔ/Δ, red) generated by infecting MllΔ/Δ MEFs with a cre-expressing retrovirus. (B) RT-qPCR results using MEFs prepared from ER-cre;MllΔ/+ or ER-cre;MllΔ/Δ embryos (n = 3). Deletion of Mll was performed in vitro by culturing in 4-OHT for 48 h; RT-qPCR assays were performed by using the primers described in Dataset S4. Error bars represent 95% CI. (C) In situ hybridization to detect Mecom or Prdm16 transcripts and (indicated on the side of the images) in wild-type, MllΔ/Δ, and MllΔ/Δ E9.5 embryos. Yellow arrows indicate limb buds, an example of normal expression of Prdm16 and Mecom in MllΔ/Δ embryos.
Additional ChIP experiments support specific enrichment around certain TSS regions. (A) Putative MLL target genes are also Mll dependent in the total lineage-negative (Lin\textsuperscript{neg}) BM population. Lin\textsuperscript{neg} BM cells were enriched from ER-cre;Mll\textsuperscript{F/F} (control, blue) or ER-cre;Mll\textsuperscript{F/+} animals (red). Cells were then cultured in 300 nM 4-OHT for 48 h, RNA was prepared, and RT-qPCR assays were performed as described in Fig. 1. (B) Anti-MLL–N-terminal (white), C-terminal (black), or control (anti-GAL4, gray) antibodies were used to immunoprecipitate fixed, sheared protein–DNA complexes, then qPCR assays were performed to determine the relative enrichment for each IP. The amplicon location relative to the TSS is indicated below each set of bars. Control ChIP-qPCR using 5 × 10\textsuperscript{4} lin\textsuperscript{neg} BM cells. Gapdh (negative control) and Hoxa9 (positive control) enrichment was determined by using qPCR as described in SI Materials and Methods. (C–H) ChIP-qPCR results from 5 × 10\textsuperscript{4} lin\textsuperscript{neg} BM cells using primers surrounding the TSS of the indicated genes. (I) ChIP-qPCR results using 5 × 10\textsuperscript{5} sorted LSK cells. (J) General diagram illustrating the position of amplicons shown in C–I; for specific positions, see Dataset S3.
Fig. S4. Features of assays used to assess whether reexpression of Mll target genes can partially rescue Mll-deficient BM cells. (A) Scheme for assessing the ability of individual genes to rescue Mll-deficient BM cells based on the selective enrichment of retrovirus-infected cells ("Virus+") after Mll deletion within the donor population (CD45.1+). In the hypothetical case illustrated, no rescue is shown as equivalent loss of Virus+ and uninfected CD45.1 cells after Mll deletion. Rescue would be represented by increased (50–75%) Virus+ cells within the CD45.1+ population. (B and C) Rescue assay was performed as in Fig. 5 by using empty retrovirus-infected/engrafted cells. Two weeks after pI:pC injection, the CD45.1+ cell number was determined within total BM (B) or hCD4+ (retrovirus-infected) within the CD45.1+ population (C). Cell numbers are expressed as millions per 2 hindlimbs (femurs and tibia); averages are represented by black bars, and individual symbols represent individual recipients. (D) BM cells from chimeras engrafted with Hoxa9 expressing MllF/F (lane 1) or Mx1-cre;MllF/F cells (lane 2) were harvested 2 wk after pI:pC injection, and quantitative genomic PCR was performed to determine the extent of Mll deletion. The Mll+, MllF, and MllΔ PCR bands are indicated by arrowheads. Data are representative of chimeras shown in Fig. 5. (E and F) Diagram of bicistronic MSCV retroviral vectors to deliver the standard (E; traditional) dose of expressed gene and an attenuated dose of the gene of interest by placing it 3′ of the IRES element (F; low-dose). (G) Western blot from 293T cell lysates transfected with the traditional or low-dose versions of Evi1-expressing retroviral vectors. We estimate that ~10 times less Evi1 protein is expressed from the low-dose virus. (H) Rescue data performed as in Fig. 5 illustrating similar overall engraftment levels and marginal rescue efficiency of low-dose Evi1. Legend continued on following page.
from the low-dose Evi1 virus. (I) Lineage distribution of retrovirally infected BM cells in chimeras at the time of analysis in Fig. 5. Average percentage of retrovirus-infected donor-type (CD45.1+) BM cells that are Mac-1/Gr-1+, B220+, or neither (other) 2 wk after pl:pC injection. The Mll genotype is shown below each bar. (J–N) Immunoblot analyses of proteins expressed by the MSCV plasmids used to make retroviral supernatant. 293T lysates transfected with the indicated retroviruses were resolved by SDS/PAGE, transferred to PVDF membranes, and probed with the antibodies indicated in SI Materials and Methods. Arrows indicate specific bands. Lane 1, nontransfected lysate; lane 2, lysate from cells transfected with traditional orientation MSCV vectors; for J–L, lane 3 represents the low dose version of the virus; for M, lanes 2–3 are duplicates and 4–5 represent Flag-tagged Eya1. Below L is a long exposure of a duplicate gel to show the low-dose expression level. In N, lanes 1 and 3 are negative controls.

**Fig. S5.** Overall features of the single cell proliferation assays and TGFβ GSEA analysis. (A) The percentage of clones that ultimately completed at least one cell division comparing control (Mll/F/F) or Mll-deleted (Mx1-cre;Mll/F/F) LSK/CD48neg single cells (“divided cells”) is shown. The data are presented as a percentage of the total number of wells that were confirmed to have received a single cell after sorting; n = 592 for control and n = 697 for Mll-deficient cells. Data represent plate averages and error bars 95% CI. (B) The average cell cycle length was calculated by comparing the slopes of the average growth curves of the actively dividing clones analyzed in Fig. 6 between the time points indicated. Error bars represent 95% CI. (C) GSEA analysis of the C2 gene sets in Mll-deficient HSCs versus wild-type cells. All of the identified gene sets <40% FDR are shown and are listed by FDR q-value. (D) GSEA plot of the TGFβ pathway (annotated by Biocarta) in the Mll-dependent gene set.

**Dataset S1.** Genes up-regulated in Mll-deficient LSK/CD48neg cells

**Dataset S2.** Genes down-regulated in Mll-deficient LSK/CD48neg cells
Dataset S3. Primers used for ChIP-qPCR

Dataset S4. Primers used for RT-qPCR