An MLL-dependent network sustains hematopoiesis

Erika L. Artinger[a], Bibhu P. Mishra[a], Kristin M. Zaffuto[a], Bin E. Li[a], Elaine K. Y. Chung[b], Adrian W. Moore[b], Yufei Chen[a], Chao Cheng[a,c,d], and Patricia Ernst[a,d,e,1]

Departments of [a]Genetics and [e]Microbiology and Immunology, [c]Institute for Quantitative Biomedical Sciences, and [d]Norris Cotton Cancer Center, Geisel School of Medicine at Dartmouth, Hanover, NH 03755; and [b]Disease 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 metazoa (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 supressors that use epigenetic mechanisms is a promising an 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 result in increased histone H3 lysine 79 dimethylation at MLL-fusion targets, overexpression of these target genes, and leukemic transformation (6). Because many of the chromatin-targeting motifs are shared between oncogenic MLL fusions and wild-type MLL, targeting of MLL-fusion oncoproteins will also require a thorough understanding of normal MLL-dependent regulatory pathways.

Wild-type MLL exists in cells as part of a large multiprotein, chromatin-associated complex that contains chromatin remodeling and histone acetylation/methylation activities (7, 8). MLL itself is thought to regulate genes in part through a highly conserved histone methyltransferase motif, the Su(var)3-9, Enhancer of Zeste, and Trithorax (SET) domain. MLL, like Trithorax, maintains precise domains of *Hox* gene expression during embryo development (9, 10). In addition, MLL has been shown to regulate other tissue-specific processes in immune, hematopoietic, vascular, and neural cell types (11–14). Germ-line disruption of *Mll* is generally embryonic lethal with multiple developmental defects (9, 15–17); however, conditional deletion of *Mll* in specific cell types revealed unique functions. For example, hematopoietic-specific deletion of *Mll* demonstrated that it is essential for maintaining hematopoietic stem and progenitor cells (HSPCs), but dispensable for lineage-committed precursors (13, 18, 19). The breadth of target genes regulated by MLL in specific tissues is largely unknown, although *Hox* genes are consistently down-regulated in many *Mll*-deficient cell types (9, 13, 14).

In this study, we investigate the molecular circuitry underlying the critical role of *Mll* in maintaining hematopoiesis as a means to understand trx-G function in normal and pathologic gene regulation. We used inducible loss-of-function models to identify hematopoietic stem cells (HSC)-specific MLL-regulated genes and delineated a network of transcriptional regulators that are direct transcriptional targets of MLL. We then tested reexpression of a subset of these genes in *Mll*-deficient hematopoietic cells to determine the epistatic relationships among transcriptional targets, to identify cross-regulatory relationships, and assess their individual ability to restore function in *Mll*-deficient cells. These studies reveal a coherent MLL pathway that coordinates self-renewal, proliferation, and lineage-specific gene expression fidelity in HSCs. Furthermore, this work distinguishes the MLL-dependent transcriptional network from that controlled by MLL fusion oncoproteins in leukemia.

Results

Short-Term Consequences of *Mll* Deletion in HSCs. To identify *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$^+$, CD48$^-$ (LSK/CD48^{neg}) HSC-enriched cells from the bone marrow (BM) of polyinosinic:polycytidylic acid (pI:pC)-injected control *Mll*$^{F/F}$ or *Mx1-cre;Mll*$^{F/F}$ 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^{neg} 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

Author contributions: E.L.A. and P.E. designed research; E.L.A., B.P.M., K.M.Z., B.E.L., Y.C., and P.E. performed research; E.K.Y.C., A.W.M., and C.C. contributed new reagents/analytic tools; E.L.A., B.P.M., K.M.Z., B.E.L., E.K.Y.C., A.W.M., and P.E. analyzed data; and E.L.A. and P.E. wrote the paper.

Conflict of interest statement: P.E. is a shareholder of Amgen stock.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE47205).

[1]To whom correspondence should be addressed. E-mail: patricia.ernst@dartmouth.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301278110/-/DCSupplemental.

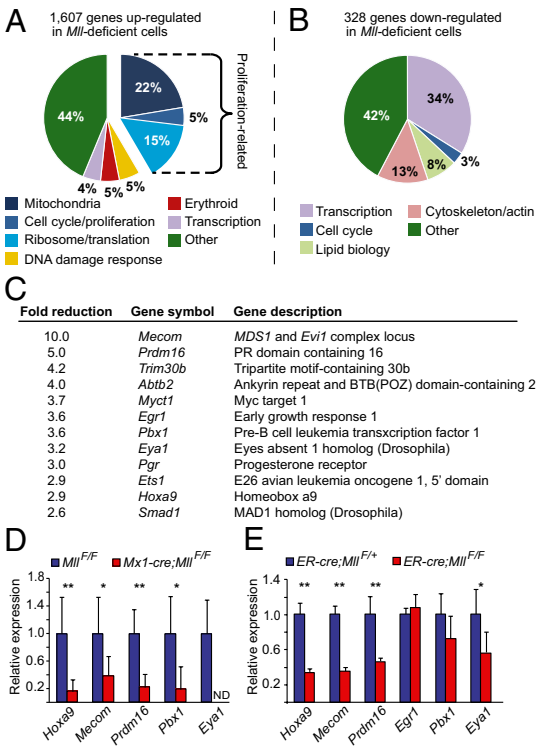


Fig. 1. Identification of *Mll*-regulated genes in HSCs. General overview of genes up-regulated (*A*) or down-regulated (*B*) in *Mll*-deficient LSK/CD48^{neg} cells compared with controls. Cells were sorted from pI:pC-injected control *Mll$^{F/F}$* or *Mx1-cre;Mll$^{F/F}$* animals at day six. Gene Ontology assignments were based on the criteria in Datasets S1 and S2. (*C*) The top down-regulated transcription factors in *Mll*-deficient LSK/CD48^{neg} cells listed by fold reduction (see also Dataset S2). (*D*) RT-qPCR validating down-regulated genes in independent control *Mll$^{F/F}$* (blue) or *Mll*-deficient (red) LSK/CD48^{neg} cells, $n = 8$ animals per genotype; ND, not detected. (*E*) RT-qPCR validation of transcripts in LSK cells sorted from control *ER-cre;Mll$^{F/+}$* (blue) or *ER-cre;Mll$^{F/F}$* animals (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). $*P \leq 0.07$, $**P \leq 0.05$. ER-cre, estrogen receptor^{T2} mutant fused to cre recombinase.

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. 1*A* and Dataset S1). Up-regulation of genes involved in ribosome biogenesis reflected the greater proportion of cycling *Mll*-deficient LSK/CD48^{neg} cells (45% G_0 in *Mll*-deleted cells versus 75% G_0 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^{neg} cells. Unexpectedly, 5% of the genes that were up-regulated in *Mll*-deficient LSK/CD48^{neg} cells encode erythroid-specific proteins including transcriptional regulators such as GATA binding protein 1 (*Gata1*) 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. S1 *A* 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. S1*E*). 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^{neg} cells (Fig. 1*B* 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. S1*F*). *Mll* itself (Fig. S1*G*) 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. 1*C*). 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 *Mx1-cre;Mll$^{F/F}$* animals (Fig. 1*D*), as well as cells in which *Mll* was deleted in vitro by using 4-hydroxytamoxifen (4-OHT; Fig. 1*E*). 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 *Mx1-cre* and *ER-cre* systems. Of the annotated transcription factors down-regulated >2.5-fold (Fig. 1*C*), MDS and Evi1 complex locus (*Mecom*), *Prdm16*, Pre-B cell leukemia homeobox protein 1 (*Pbx1*), Eyes absent homolog 1 (*Eya1*) and *Hoxa9* were consistently *Mll*-dependent (Fig. 1*E*). Tripartate 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 *Trithorax* 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 5×10^4 BM cells (31). Based on previous results demonstrating MLL binding near transcription start

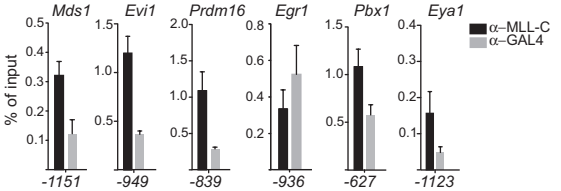


Fig. 2. MLL binds directly to the promoter regions of a subset of genes identified by expression array. ChIP results demonstrating specific enrichment at the *Mecom* locus (*Mds1* and *Evi1* start sites) and the *Prdm16, Pbx1*, and *Eya1* promoter regions. Anti-MLL C-terminal (black) or control (anti-GAL4, gray) antibodies were used for ChIP, and enrichment was determined by using quantitative PCR assays. Amplicon position is indicated relative to the TSS for each gene. Results using additional primers surrounding the TSS are shown in Fig. S4. Data represents averages ± SEM for two to four PCR replicates and are representative of at least four independent experiments.

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^{neg}) population and LSK cells (Fig. S3*A*). Control ChIP experiments demonstrated MLL binding to the *Hoxa9* but not *Gapdh* TSS regions (Fig. S3*B*). Using lin^{neg} BM cells, we observed specific MLL binding around each TSS of the *Mecom* locus [both Myelodysplastic syndrome 1 (*Mds1*) and Ecotropic 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. 2*B* and Fig. S3*H*), consistent with the observation that this gene was not *Mll*-dependent in both model systems (Fig. 1*E*). Therefore, we conclude that like *Hoxa9*, the expression of *Mecom*, *Prdm16*, *Pbx1*, and *Eya1* is maintained directly by MLL in normal lin^{neg} 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* (*Men*) *1*-deficient LSK cells. Interestingly, *Mecom* and *Eya1* were slightly reduced, but the latter was not statistically significant (Fig. 3*A*). Despite efficient excision of *Men1* (Fig. 3*B*), *Prdm16* and *Pbx1* levels were not affected (Fig. 3*A*), 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 *Evi1* 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-Evi1* and *Evi1*) 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*, *Evi1* 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. *Evi1* expression increased *Prdm16* and *Hoxa9* transcripts in *Mll*-deficient LSK cells back to the wild-type levels (Fig. 4 *A* and *B*). *Mds1-Evi1* 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-Evi1* and *Evi1* 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 *Evi1*, 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.

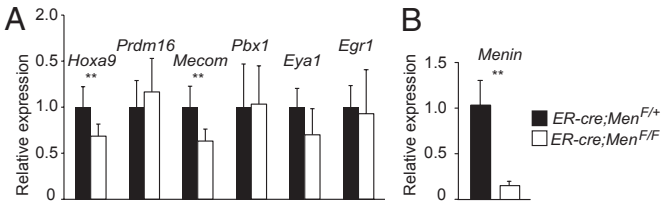


Fig. 3. Menin loss affects some but not all MLL targets in LSK cells. (*A*) RT-qPCR of *Mll*-regulated genes in LSK cells sorted from control *ER-cre;Men1*$^{F/+}$ (black) or *ER-cre;Men1*$^{F/F}$ cells (white) cultured for 72 h after initiating *Menin* deletion. Expression levels were normalized to rRNA. (*B*) Menin transcript levels in LSK cells treated as in *A*. Error bars represent 95% CI; $n = 4$–8 animals per genotype. $^{**}P \leq 0.05$.

***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*$^{F/F}$ 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 reexpression 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*$^{F/F}$ or *Mx1-cre;Mll*$^{F/F}$ mice, then engrafted into lethally irradiated recipients together with uninfected wild-type BM cells. After stable engraftment, *Mll* excision was induced by pI:pC injection and the persistence of *Mll*-deficient BM cells expressing the reintroduced gene was determined 2 wk later (Fig. 5*A*). Thus, in this assay, "rescue" is defined as the selective persistence of retrovirus-infected cells within the population of *Mll*-deleted cells (Fig. S4*A*). 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 from chimeric animals as expected (Fig. S4 *B* and *C*). *Hoxa9* overexpression resulted in the expansion of donor-derived cells in chimeras (*Hoxa9* versus empty) but also *Hoxa9* expressing *Mll*-deficient cells were protected from attrition, as evidenced by their overrepresentation in the *Mll*-deficient population (Fig. 5*B*, red versus blue). Surprisingly, *Prdm16* reexpression resulted in the most significant rescue of *Mll*-deficient cells. Despite its greater ability to influence other network genes, reexpression of *Evi1* only marginally protected *Mll*-deficient cells from attrition, and *Mds1-Evi1*, *Pbx1*, and *Eya1* had no specific activity in this assay (Fig. 5*B*). Because of the low contribution of *Evi1*-expressing cells in chimeras, we considered in this case that overexpression may suppress hematopoiesis overall, but we found that a retrovirus producing ~10-fold less Evi1 produced similar results (Fig. S4 *E–H*). Complete *Mll* deletion in the persisting cells of chimeras was confirmed by a quantitative genomic PCR assay (Fig. S4*D*). We found that retroviral overexpression of the individual genes resulted in a similar contribution to lymphoid and myeloid lineages, with the exception being the suppression of B-lymphopoiesis by *Prdm16* (Fig. S4*I*) as has been noted (26). Taken together, these data suggest that in addition to *Hoxa9*, *Prdm16* is an important direct target of MLL in HSCs and is capable of partially rescuing *Mll*-deficient hematopoietic cells from attrition in BM chimeras without restoring the entire transcriptional network.

***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

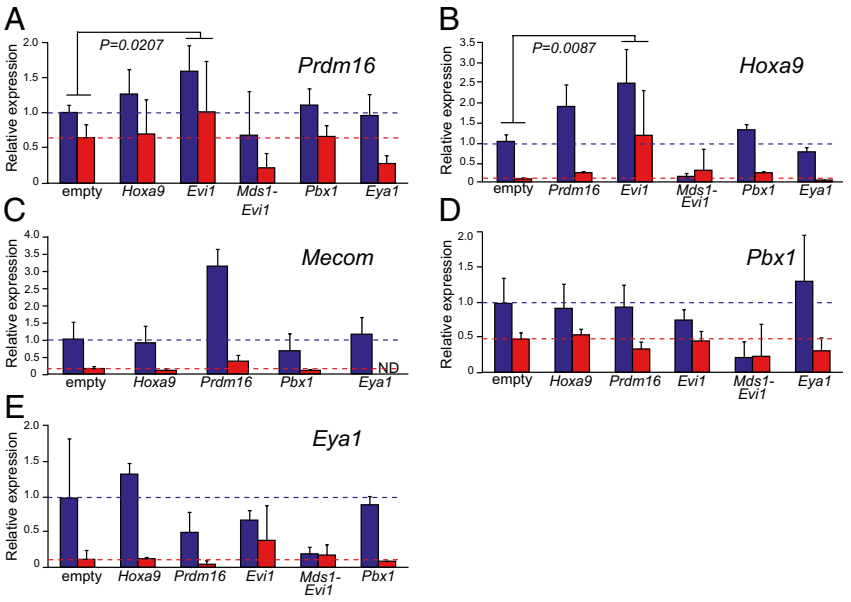


Fig. 4. Effect of reexpression of individual *Mll* targets on others in the network. RT-qPCR of genes in LSK cells reexpressing the cDNA indicated below each set of bars. Cells were produced in vivo by pI:pC injection, sorted 6 d later, then infected with a retrovirus without an added cDNA (empty) or cDNA as indicated. Two days later, retrovirally infected cells were sorted and RT-qPCR assays were performed. (*A*) *Prdm16* expression levels in control $Mll^{F/F}$ (blue) or *Mll*-deficient (red) LSK cells infected with the retrovirus indicated below each set of bars. Expression levels were normalized to the average expression level empty retrovirus-infected $Mll^{F/F}$ cells and to *Gapdh* in each sample. Expression of *Hoxa9* (*B*), *Mecom* transcripts (*C*), *Pbx1* (*D*), and *Eya1* (*E*) were analyzed and normalized as in *A*. Dashed lines indicate the average expression level in wild-type or *Mll*-deficient, empty retrovirus infected cells; four to five animals per genotype were used for each experiment, and error bars represent 95% CI. *P* values are shown for the comparison between pairs of empty vector and *Evi1*-expressing cells, calculated with the paired Student *t* test. ND, not detected.

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 $CD48^{neg}$ subset of these cells were largely in G_1/S rather than G_0 (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 ($Mll^{F/F}$) or *Mll*-deleted (*Mx1-cre;*$Mll^{F/F}$) LSK/$CD48^{neg}$ cells were sorted from pI:pC-injected animals, deposited into wells as single cells and cultured in serum-free medium containing cytokines to maintain HSC identity and function (39) (Fig. 6*A*). Importantly, the percentage of surviving clones was similar between wild-type and *Mll*-deleted cells (Fig. S5*A*), 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/$CD48^{neg}$ cells, we found that the proliferation kinetics of the latter were consistently more advanced than wild type (Fig. 6*E*). After 48 h, the mode (greatest number of cells) of *Mll*-deleted LSK/$CD48^{neg}$ clones had progressed approximately one-half a division further than the wild-type clones (Fig. 6*C*), and by 72 h, the mode was one full cell division ahead (Fig. 6*D*). To address the possibility that *Mll*-deficient LSK/$CD48^{neg}$ cells exhibit earlier cell division because more are initially in G_1/S compared with wild type, we performed higher resolution studies examining the initial three cell divisions (Fig. 6*E*). We found that *Mll*-deficient LSK/$CD48^{neg}$ 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. S5*B*). 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_0 (13).

To investigate whether *Prdm16* reexpression influenced the proliferation phenotype observed in *Mll*-deficient cells, we sorted LSK cells from control *ER-cre;*$Mll^{F/+}$and *ER-cre;*$Mll^{F/F}$ mice, retrovirally introduced *Prdm16*, and concurrently incubated with 4-OHT to induce *Mll* deletion (Fig. 6*F*). *ER-cre;*$Mll^{F/F}$ cells infected with an empty control retrovirus displayed greater cell accumulation than the *ER-cre;*$Mll^{F/+}$ 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. 6*G*). Together, these data suggest that the mechanism by which *Prdm16* can correct *Mll* deficiency is, in part, by restraining proliferation within HSPCs.

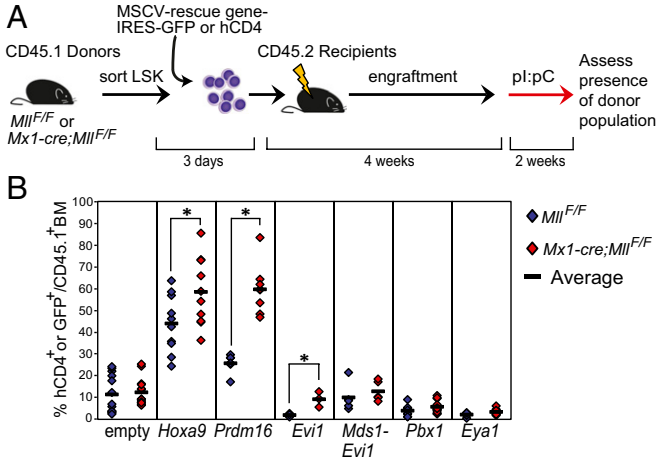


Fig. 5. Reexpression of *Prdm16* partially rescues *Mll* deficiency. (*A*) Experimental scheme to determine effects of reexpression of *Mll*-dependent genes. LSK cells were sorted from control $Mll^{F/F}$ or *Mx1-cre;*$Mll^{F/F}$ donor animals then infected with the indicated retrovirus. The entire pool of infected and uninfected cells was transplanted into irradiated recipients, which were analyzed 6 wk later. (*B*) Results of reexpression of each individual gene in control $Mll^{F/F}$ (blue) or *Mll*-deficient LSK cells (red); each point represents an individual recipient animal, $n = 3$–10 recipients per condition. The percentage of donor-type ($CD45.1^+$) BM cells that are GFP^+ or $hCD4^+$ 2 wk after *Mll* deletion is shown. Data are representative of three independent experiments. $*P \leq 0.05$ was calculated by using the Wilcoxon rank-sum test.

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

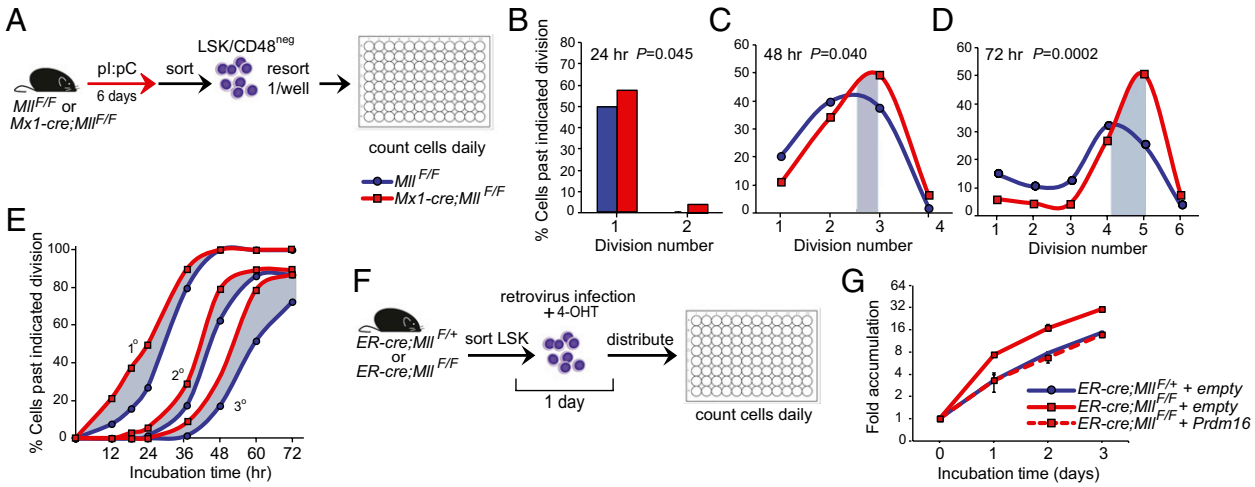


Fig. 6. The intrinsic proliferation defect of *Mll*-deficient HSCs is corrected by reexpression of *Prdm16*. (*A*) Scheme to determine proliferation kinetics of individual LSK/CD48^{neg} cells. *Mll* deletion was performed in vivo, and double-sorted LSK/CD48^{neg} cells were deposited at 1 cell per well. Cell divisions were scored every 24 h. (*B–D*) Cumulative proliferation data from individual control *Mll*$^{F/F}$ (blue) or *Mll*-deficient LSK/CD48^{neg} cells (red). Data represent 158 control *Mll*$^{F/F}$ and 240 *Mll*-deficient cells; $n = 3–5$ animals per genotype. The difference between modes of each line is indicated by gray fill. The Pearson's $\chi 2$ test was performed to determine statistical significance, shown on *B–D*. (*E*) Higher-resolution proliferation kinetics of control *Mll*$^{F/F}$ (blue) or *Mll*-deficient LSK/CD48^{neg} cells (red). Cells were prepared as in *A*, $n = 2–3$ animals per genotype, 93 control *Mll*$^{F/F}$ and 38 *Mll*-deficient cells. The percentage of cells past the first, second, and third divisions are graphed separately (1°, 2°, 3°). (*F*) Scheme to determine the impact of *Prdm16* reexpression in *Mll*-deficient LSK cells. (*G*) Accumulation of LSK cells expressing an empty (solid) or hCD4-*Prdm16* retrovirus (dashed). LSK cells were sorted from control *ER-cre;Mll*$^{F/+}$ (blue) or *ER-cre;Mll*$^{F/F}$ animals (red), cultured in 4-OHT during the retroviral infection to induce *Mll* deletion then enumerated every 24 h for 3 d. Data represent averages ± 95% CI, $n = 4$ animals per genotype, 3 replicates per time point.

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 Recombinase. *Mx1-cre;Mll*$^{F/F}$ animals and *cre* induction have been described (13). *Men1*$^{F/F}$ 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 *ER-cre* strain.

Flow Cytometry, Cell Sorting, and Culture. Flow cytometry and cell sorting were performed on a FACSCalibur and FACSAria, 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 cotransfection, and sorted LSK cells were infected by using retronectin (Takara). Retrovirally infected cells were cotransplanted into lethally irradiated (950 Rads, split dose) C57BL/6J female mice. For proliferation assays, LSK and LSK/CD48^{neg} cells were cultured in HSC expansion medium [StemSpan Serum Free Expansion Medium (SFEM); 300 ng/mL recombinant murine (rm) SCF, 20 ng/mL rmIL-11, and 4 ng/mL rmFlt3L; StemCell Techologies and R&D Systems]. To induce deletion using the *ER-cre* strain, HSC expansion medium was supplemented with 300–400 nM 4-OHT (Sigma).

ChIP. Rabbit polyclonal anti-MLL C terminus (50) or anti-Gal4 (Santa Cruz; SC-577) antibodies were used for ChIP by using lin^{neg} or LSK cells (31) with refinements as indicated in *SI Materials and Methods*. Primer sequences and genomic positions are described in Dataset S3.

Microarray Sample Preparation and Data Analyses. Affymetrix microarray analyses were performed by using sorted LSK/CD48^{neg} cells from five *Mll*$^{F/F}$ or *Mx1-cre*; *Mll*$^{F/F}$ mice 6 d after cre induction. Detailed methods and bioinformatic analyses are found in *SI Materials and Methods* and Dataset S4.

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.

ACKNOWLEDGMENTS. We thank R. Mako Saito, Chris Vakoc, Steve Smale, Hanna Mikkola, Emmanuelle Passegué, and Adolfo Ferrando for critical comments; Thomas Milne and Joanna Attema for advice on ChIP; and Drs. Perkins, Morishita, 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 Scientist (B) and a RIKEN Brain Sciences 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 Lauri Strauss Leukemia Foundation.

1. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* 128(4):735–745.
2. Sparmann A, van Lohuizen M (2006) Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6(11):846–856.
3. Deshpande AJ, Bradner J, Armstrong SA (2012) Chromatin modifications as therapeutic targets in MLL-rearranged leukemia. *Trends Immunol* 33(11):563–570.
4. Daser A, Rabbitts TH (2004) Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. *Genes Dev* 18(9):965–974.
5. Krivtsov AV, Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* 7(11):823–833.
6. Marschalek R (2011) Mechanisms of leukemogenesis by MLL fusion proteins. *Br J Haematol* 152(2):141–154.
7. Nakamura T, et al. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell* 10(5):1119–1128.
8. Yokoyama A, et al. (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* 24(13):5639–5649.
9. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 378(6556):505–508.
10. Ingham PW (1985) A clonal analysis of the requirement for the trithorax gene in the diversification of segments in Drosophila. *J Embryol Exp Morphol* 89:349–365.
11. Lim DA, et al. (2009) Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature* 458(7237):529–533.
12. Yamashita M, et al. (2006) Crucial role of MLL for the maintenance of memory T helper type 2 cell responses. *Immunity* 24(5):611–622.
13. Jude CD, et al. (2007) Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* 1(3):324–337.
14. Diehl F, Rössig L, Zeiher AM, Dimmeler S, Urbich C (2007) The histone methyltransferase MLL is an upstream regulator of endothelial-cell sprout formation. *Blood* 109(4):1472–1478.
15. Yagi H, et al. (1998) Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* 92(1):108–117.
16. Ayton P, et al. (2001) Truncation of the Mll gene in exon 5 by gene targeting leads to early preimplantation lethality of homozygous embryos. *Genesis* 30(4):201–212.
17. Yokoyama A, et al. (2011) Proteolytically cleaved MLL subunits are susceptible to distinct degradation pathways. *J Cell Sci* 124(Pt 13):2208–2219.
18. McMahon KA, et al. (2007) Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* 1(3):338–345.
19. Gan T, Jude CD, Zaffuto K, Ernst P (2010) Developmentally induced Mll1 loss reveals defects in postnatal haematopoiesis. *Leukemia* 24(10):1732–1741.
20. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98(9):5116–5121.
21. Venezia TA, et al. (2004) Molecular signatures of proliferation and quiescence in hematopoietic stem cells. *PLoS Biol* 2(10):e301.
22. Seita J, et al. (2012) Gene Expression Commons: An open platform for absolute gene expression profiling. *PLoS ONE* 7(7):e40321.
23. Novershtern N, et al. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144(2):296–309.
24. He S, Kim I, Lim MS, Morrison SJ (2011) Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. *Genes Dev* 25(15):1613–1627.
25. Ficara F, Murphy MJ, Lin M, Cleary ML (2008) Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. *Cell Stem Cell* 2(5):484–496.
26. Deneault E, et al. (2009) A functional screen to identify novel effectors of hematopoietic stem cell activity. *Cell* 137(2):369–379.
27. Aguilo F, et al. (2011) Prdm16 is a physiologic regulator of hematopoietic stem cells. *Blood* 117(19):5057–5066.
28. Goyama S, et al. (2008) Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell* 3(2):207–220.
29. Zhang Y, et al. (2011) PR-domain-containing Mds1-Evi1 is critical for long-term hematopoietic stem cell function. *Blood* 118(14):3853–3861.
30. Schuettengruber B, Martinez AM, Iovino N, Cavalli G (2011) Trithorax group proteins: Switching genes on and keeping them active. *Nat Rev Mol Cell Biol* 12(12):799–814.
31. Weishaupt H, Attema JL (2010) A method to study the epigenetic chromatin states of rare hematopoietic stem and progenitor cells; MiniChIP-Chip. *Biol Proced Online* 12(1):1–17.
32. Guenther MG, et al. (2005) Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci USA* 102(24):8603–8608.
33. Wang QF, et al. (2011) MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood* 117(25):6895–6905.
34. Yokoyama A, et al. (2005) The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* 123(2):207–218.
35. Maillard I, et al. (2009) Menin regulates the function of hematopoietic stem cells and lymphoid progenitors. *Blood* 113(8):1661–1669.
36. Shimabe M, et al. (2009) Pbx1 is a downstream target of Evi-1 in hematopoietic stem/progenitors and leukemic cells. *Oncogene* 28(49):4364–4374.
37. Sitailo S, Sood R, Barton K, Nucifora G (1999) Forced expression of the leukemia-associated gene EVI1 in ES cells: A model for myeloid leukemia with 3q26 rearrangements. *Leukemia* 13(11):1639–1645.
38. Sood R, Talwar-Trikha A, Chakrabarti SR, Nucifora G (1999) MDS1/EVI1 enhances TGF-beta1 signaling and strengthens its growth-inhibitory effect but the leukemia-associated fusion protein AML1/MDS1/EVI1, product of the t(3;21), abrogates growth-inhibition in response to TGF-beta1. *Leukemia* 13(3):348–357.
39. Uchida N, Dykstra B, Lyons KJ, Leung FY, Eaves CJ (2003) Different in vivo repopulating activities of purified hematopoietic stem cells before and after being stimulated to divide in vitro with the same kinetics. *Exp Hematol* 31(12):1338–1347.
40. Chuikov S, Levi BP, Smith ML, Morrison SJ (2010) Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. *Nat Cell Biol* 12(10):999–1006.
41. Avagyan S, Aguilo F, Kamezaki K, Snoeck HW (2011) Quantitative trait mapping reveals a regulatory axis involving peroxisome proliferator-activated receptors, PRDM16, transforming growth factor-β2 and FLT3 in hematopoiesis. *Blood* 118(23):6078–6086.
42. Liu H, et al. (2010) Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. *Nature* 467(7313):343–346.
43. Pinheiro I, et al. (2012) Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* 150(5):948–960.
44. Armstrong SA, et al. (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 30(1):41–47.
45. Ferrando AA, et al. (2003) Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: Dominance of HOX dysregulation. *Blood* 102(1):262–268.
46. Ross ME, et al. (2004) Gene expression profiling of pediatric acute myelogenous leukemia. *Blood* 104(12):3679–3687.
47. Arai S, et al. (2011) Evi-1 is a transcriptional target of mixed-lineage leukemia oncoproteins in hematopoietic stem cells. *Blood* 117(23):6304–6314.
48. Morishita K (2007) Leukemogenesis of the EVI1/MEL1 gene family. *Int J Hematol* 85(4):279–286.
49. Grembecka J, et al. (2012) Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. *Nat Chem Biol* 8(3):277–284.
50. Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ (2003) Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol Cell Biol* 23(1):186–194.

Supporting Information

Artinger et al. 10.1073/pnas.1301278110

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*$^{F/F}$ animals were extensively back-crossed to the B6.SJL strain (B6.SJL-*Ptprc*^{a} *Pep3*^{b}*/BoyJ,* stock no. 002014; Jackson Laboratory). *ER-cre* mice have an estrogen receptor^{T2} mutant fused to cre recombinase knock-in at the Rosa locus (Jackson Laboratory; stock no. 004847; ref. 1). The Dartmouse facility (Geisel School of Medicine at Dartmouth) was used to back-cross to the B6.SJL strain until >93% strain-specific SNPs were represented in the breeding animals. Murine embryo fibroblasts (MEFs) used in Fig. S2*A* were prepared from embryonic day (E)14.5 embryos by using standard methods and *Mll*$^{F/F}$ intercrosses. MEFs were then infected with a MIG-cre retrovirus, GFP^{+} cells were sorted, *Mll* deletion was confirmed, and MEFs were passaged at least 20 times before quantitative RT-PCR (RT-qPCR). For Fig. S2*B*, *ER-cre;Mll*$^{F/+}$ and *Mll*$^{F/F}$ animals were intercrossed to produce the control *ER-cre;Mll*$^{F/+}$and *ER-cre;Mll*$^{F/F}$ MEFs by using similar methods.

Cell Culture and in Vitro Induction of cre Recombinase. Lineage-negative, Sca-1^{+}, c-Kit^{+} (LSK) and LSK/CD48^{neg} cells were cultured in hematopoietic stem cell (HSC) expansion medium, which is defined as StemSpan Serum Free Expansion Medium (Stem Cell Technologies) plus 300 ng/mL rmSCF, 20 ng/mL rm IL-11, 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 Chris Lowrey, Geisel School of Medicine at Dartmouth, Lebanon, NH) and 400 nM 4-OHT for 24 h, the medium was replaced with HSC expansion medium plus Epo for an additional 24 h. Cells were seeded at 1×10^4 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^{+} *Mx1-cre;Mll*$^{F/F}$ or *Mll*$^{F/F}$ donors. To express genes in LSK cells, 96-well suspension plates were coated with 25 mg/mL Retronectin (Takara Bio) for at least 4 h and then loaded with 200 μL of retroviral supernatant. After 3–4 hours incubation, at 32 degrees, excess supernatant was removed, and cells were added and centrifuged for 90 min at 380 × *g*, room temperature. After 48 h of culture in HSC expansion medium, 5×10^3 to 5×10^4 infected LSK cells were mixed with 4.5×10^5 C57BL/6J (CD45.2^{+}) Sca-1–depleted bone marrow (BM) cells ("carrier"). Carrier cells were prepared by staining BM with anti–Sca-1 PE-labeled antibody followed by depletion of Sca-1–labeled cells using anti-PE magnetic beads (Miltenyi). Infected LSK cells and Sca-1–depleted BM cells were injected periorbitally into lethally irradiated female C57BL/6J recipients. Four weeks after engraftment, chimeric mice were injected with four doses of pI:pC every other day and euthanized for analysis 2 wk after the first injection. Donor contribution was determined by flow cytometry using anti-CD45.1 and anti-CD45.2 antibodies. *Mll* deletion efficiency was determined by quantitative PCR assays using genomic DNA as described (2) or a custom Taqman assay using cDNA.

Liquid Culture Proliferation Assays. For single-cell liquid culture assays, LSK/CD48^{neg} cells from pI:pC-injected control *Mll*$^{F/F}$ or *Mx1-cre;Mll*$^{F/F}$ 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_2 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. 6*G*, 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 incubated HSC expansion medium in 96-well plates. To obtain an accurate count of cells at every time point, 1,000 15.7-μm polystyrene polybeads (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 FACSCalibur to enumerate hCD4^{+} 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 FACSAria at the DartLab Flow Cytometry Shared Resource at the Geisel School of Medicine at Dartmouth. Fluorochrome-labeled antibodies used were as follows:

Antibody	Company	Clone	Fluorochrome
B220	Invitrogen	RM2600	Unlabeled
CD19	Invitrogen	RM7700	Unlabeled
CD3	eBiosciences	17A2	Unlabeled
CD4	Invitrogen	MCD0400	Unlabeled
CD8	Invitrogen	MCD0800	Unlabeled
Ter119	Invitrogen	MTER00	Unlabeled
Gr1	Invitrogen	RM3000	Unlabeled
Mac1	Invitrogen	RM2800	Unlabeled
IL7Ra	eBiosciences	A7R34	Unlabeled
CD71	BD Pharmingen	C2	FITC
Ter119	BD Pharmingen	TER-119	PE
Goat anti-Rat F(ab)$_2$	Invitrogen	N/A	Cy5-R-PE
Sca-1	BD Pharmingen	E13-161.7	FITC
c-Kit	Biolegend	2B8	APC
CD48	Biolegend	HM48-1	FITC
B220	Biolegend	RA6-8B2	APC
CD3	BD Pharmingen	145–2C11	PE
Mac1	BD Pharmingen	M1/70	APC
Gr1	BD Pharmingen	RB6-8C5	FITC
hCD4	Biolegend	OKT4	APC
FcgRIII	eBiosciences	93	FITC
CD34	Biolegend	RAM34	PE
CD45.1	Biolegend	A20	PE
CD45.2	Biolegend	104	FITC

APC, allophycocyanin; BD, Beckton, Dickinson; Cy5-R-PE, R-Phycoerythrin-Cyanine 5; F(ab)$_2$, fragment-antigen binding; FcgRIII, Fc gamma receptor, type III; FITC, fluorescein isothiocyanate; hCD4, human anti-CD4; PE, phycoerythrin.

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 linage antibodies. Lineage$^+$ 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-Evi1* expression was accomplished by using the MIGR1.ME retrovirus obtained from Archibald Perkins (University of Rochester Medical School, Rochester, NY) (3). The murine *Evi1* 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 HpaI site of MSCV resulting in MSCV.*Evi1* (Fig. S4*E*). This retrovirus expresses the Evi1a or p135 isoform as described (3, 5). Protein expression from this retroviral vector was confirmed by using anti-Evi1 from Santa Cruz Biotechnology [(C20)-R; Fig. S4*J*]. 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 BglII-XhoI sites. Protein expression was confirmed by immunoblot using Millipore anti-Hoxa9 (07-178, Fig. S4*K*). The murine *Prdm16* 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′-TCGAGCCATGGACTACAAGGACGACGATGACAAGG-3′ and 5′-TCGACCTTGTCATCGTCGTCCTTGTAGTCGATGGC-3′. The hCD4-*Prdm16* retrovirus was constructed by amplification of the Flag-tagged ORF with the following oligos: 5′-GAATTCATGGACTACAAAGACGATGAC-3′ and 5′-TTAATTAATCATTGCATATGCCTCCGG-3′, isolating the resulting fragment by using pCR-Blunt (Invitrogen), then excising the EcoRI-PacI fragment and inserting it into the EcoRI-PacI sites of MSCV.hCD4. Protein expression from the resulting plasmids was confirmed by using anti-Flag rabbit antibody (Bethyl A190-102A; Fig. S4*L*). The *Eya1* ORF was amplified by PCR from mouse BM cells by using oligos 5′-GCAGGTCTATGGAAATGCAGGATCTAACC-3′ and 5′-TTAATTAATTACAGGTACTCTAATTCCAAGGCGC-3′. Expression was confirmed by immunoblot using an antibody from Aviva (ARP39974-P050; Fig. S4*M*). The murine pre-B cell leukemia homeobox protein 1 (*Pbx1*) *a* cDNA was obtained from Invitrogen (IMAGE Consortium clone 5701148, NM_183355). Confirmation of the 50 KDa Pbx1a isoform was performed by immunoblot using anti-Pbx1 rabbit antibody (Cell Signaling, 4342S; Fig. S4*N*). The *Evi1* ORF was PCR-amplified from the Moroshita cDNA by using the following oligos: 5′-GCACTTTAATTAAGCGCCTGGGGAA-3′ and 5′-TCACGACGCGTAACCTTGACAATGTC-3′, and the resulting 3-kb fragment inserted into the MSCV.hCD4 plasmid at the PacI and MluI sites. The bicistronic MSCV.hCD4 retroviral plasmid ("low-dose MSCV", used in Fig. S4*F*) was constructed by inserting the hCD4 cassette just downstream of the MSCV LTR then replacing the polylinker and IRES element downstream of hCD4. This plasmid was used for cloning low-dose retroviral plasmids expressing the corresponding cDNAs. All transferred or amplified cDNAs were confirmed by sequencing and Western blotting. To make virus, 293T cells were co-transfected with the MSCV-based plasmid and a Ψ-ecotropic packaging plasmid by using FuGENE6 (Roche). Supernatants were collected 48 h later, filtered using 0.45-μm syringe filters (Acrodisc) and stored at −80 °C until use.

Detailed Microarray Sample Preparation and Data Analysis. Total RNA was prepared from 1,500 to 10^4 LSK/CD48^{neg} cells sorted from five *Mll*$^{F/F}$ mice and five *Mx1-cre*; *Mll*$^{F/F}$ 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 aRNA 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 SuiteVersion 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%, a target percentile of 90%, and 100 permutations. 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 Analyses (GSEA) (17) was performed with software available from the Broad Institute (http://www.broadinstitute.org/gsea/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/CD48^{neg} and wild-type CD48^{neg} 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 genes 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 CD48$^+$ LSK cells, in fetal liver HSC compared with BM HSCs, and in BM CD48$^+$ LSK cells compared with BM HSCs.

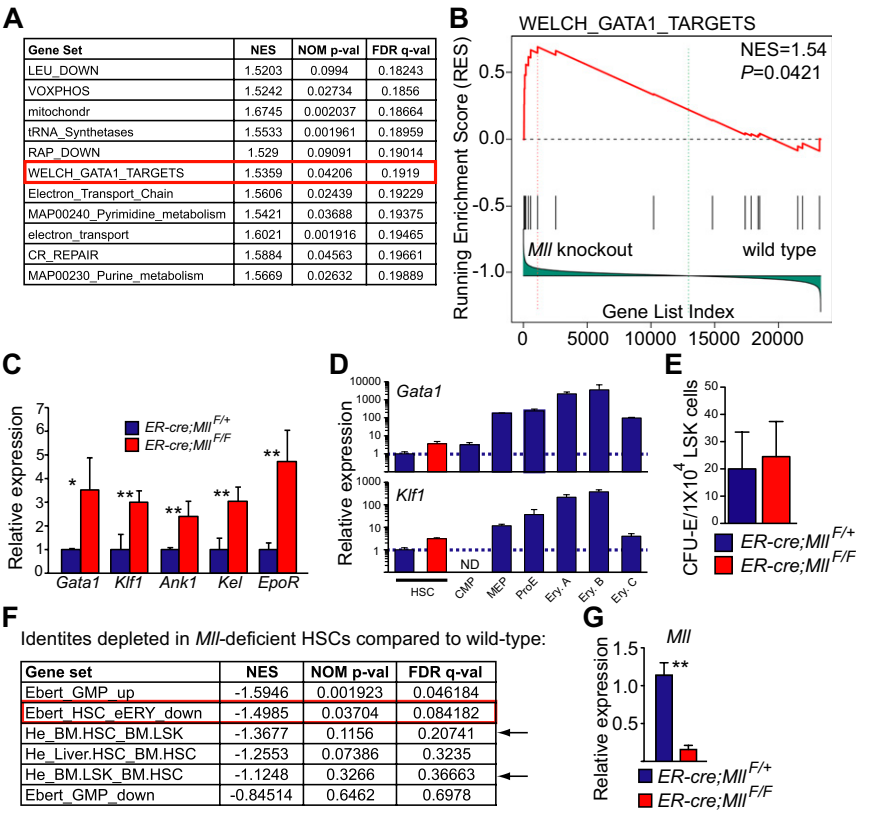
Quantitative PCR and Detection of *Mll* and *Menin* Transcripts. Total RNA from the sorted populations indicated was isolated as described above. If necessary, mRNA was amplified with the RiboAmp RNA Amplification Kit (Arcturus) or the MessageAMP II aRNA Amplification Kit for one or two rounds of amplification. cDNA was reverse transcribed by using SuperScript III (Invitrogen). *Mll* transcripts were quantified by using a custom Taqman assay the following primers 5′-TTCTCGTCAAATAGCCCTGC-3′, 5′-CTACTCTTGTCCTTCTCCACG-3′ and probe: 5′-FAM-TCTCTTCCCATGGTTCACCCCAG-TAMRA-3′. *Men1* transcripts were quantified by using 2× SYBR Master Mix (Bio-Rad) and the following primers: forward, 5′-TCC CTC TTC AGC TTC ATC ACA -3′ and reverse, 5′-ACCCAAGCATGATCTTCAGCA-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 *rRNA* 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, 1× protease inhibitor complex; Roche) at 5×10^4 cells per 20 μL. Low-retention surface 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* at 4 °C for 5 min, and the supernatant was diluted 10-fold with 2× 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 1× 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 1× 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 2× SYBR green master mix (Bio-Rad). Enrichment of MLL at genomic loci was expressed as the percent input by using the following formula: % of total input = $100 \times 2\hat{}$ [Ct (ChIP) – (Ct input – $\log_2$ (input dilution factor))] (21).

In Situ Hybridization. Embryos were generated by crossing $Mll^{\Delta/+}$ 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.

1. Badea TC, Wang Y, Nathans J (2003) A noninvasive genetic/pharmacologic strategy for visualizing cell morphology and clonal relationships in the mouse. *J Neurosci* 23(6): 2314–2322.
2. Jude CD, et al. (2007) Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* 1(3):324–337.
3. Zhang Y, et al. (2011) PR-domain-containing Mds1-Evi1 is critical for long-term hematopoietic stem cell function. *Blood* 118(14):3853–3861.
4. Morishita K, et al. (1988) Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. *Cell* 54(6):831–840.
5. Goyama S, et al. (2008) Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell* 3(2):207–220.
6. Ernst P, Mabon M, Davidson AJ, Zon LI, Korsmeyer SJ (2004) An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Curr Biol* 14(22):2063–2069.
7. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98(9):5116–5121.
8. Akashi K, et al. (2003) Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. *Blood* 101(2):383–389.
9. Terskikh AV, Miyamoto T, Chang C, Diatchenko L, Weissman IL (2003) Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood* 102(1): 94–101.
10. Forsberg EC, et al. (2005) Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet* 1(3):e28.
11. Venezia TA, et al. (2004) Molecular signatures of proliferation and quiescence in hematopoietic stem cells. *PLoS Biol* 2(10):e301.
12. Chambers SM, et al. (2007) Hematopoietic fingerprints: An expression database of stem cells and their progeny. *Cell Stem Cell* 1(5):578–591.
13. Forsberg EC, et al. (2010) Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. *PLoS ONE* 5(1):e8785.
14. Ficara F, Murphy MJ, Lin M, Cleary ML (2008) Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. *Cell Stem Cell* 2(5): 484–496.
15. Kiel MJ, et al. (2005) SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121(7):1109–1121.
16. Ivanova NB, et al. (2002) A stem cell molecular signature. *Science* 298(5593):601–604.
17. Subramanian A, et al. (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102(43):15545–15550.
18. Novershtern N, et al. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144(2):296–309.
19. He S, Kim I, Lim MS, Morrison SJ (2011) Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. *Genes Dev* 25(15):1613–1627.
20. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25(4):402–408.
21. Weishaupt H, Attema JL (2010) A method to study the epigenetic chromatin states of rare hematopoietic stem and progenitor cells; MiniChIP-Chip. *Biol Proced Online* 12(1):1–17.
22. Kinameri E, et al. (2008) Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. *PLoS ONE* 3(12): e3859.
23. Grove EA, Tole S, Limon J, Yip L, Ragsdale CW (1998) The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. *Development* 125(12):2315–2325.



Gene Set	NES	NOM p-val	FDR q-val
LEU_DOWN	1.5203	0.0994	0.18243
VOXPHOS	1.5242	0.02734	0.1856
mitochondr	1.6745	0.002037	0.18664
tRNA_Synthetases	1.5533	0.001961	0.18959
RAP_DOWN	1.529	0.09091	0.19014
WELCH_GATA1_TARGETS	1.5359	0.04206	0.1919
Electron_Transport_Chain	1.5606	0.02439	0.19229
MAP00240_Pyrimidine_metabolism	1.5421	0.03688	0.19375
electron_transport	1.6021	0.001916	0.19465
CR_REPAIR	1.5884	0.04563	0.19661
MAP00230_Purine_metabolism	1.5669	0.02632	0.19889

Gene set	NES	NOM p-val	FDR q-val
Ebert_GMP_up	-1.5946	0.001923	0.046184
Ebert_HSC_eERY_down	-1.4985	0.03704	0.084182
He_BM.HSC_BM.LSK	-1.3677	0.1156	0.20741
He_Liver.HSC_BM.HSC	-1.2553	0.07386	0.3235
He_BM.LSK_BM.HSC	-1.1248	0.3266	0.36663
Ebert_GMP_down	-0.84514	0.6462	0.6978

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^{neg} cells versus wild-type CD48^{neg} 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^{neg} dataset. (*C*) RT-qPCR validation of select up-regulated, erythroid-specific genes in LSK cells sorted from control *ER-cre;Mll$^{F/+}$* (blue) or *ER-cre;Mll$^{F/F}$* animals (red) cultured for 24 h in HSC expansion medium with 400 nM 4-OHT, then an additional 48 h in HSC expansion medium. Data represents relative expression levels normalized to *Gapdh*. Error bars represent 95% CI, $n = 4$ animals per genotype; $*P \leq 0.07$, $**P \leq 0.05$. (*D*) RT-qPCR results measuring *Gata1* and *Klf1* expression in LSK/CD48^{neg} (HSC) of wild-type (blue) and *Mll*-deficient (red) HSCs 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–C (2); ND, not detected. Expression levels from purified populations were normalized to *Gapdh* levels and reflect averages ± 95% CI, $n = 2$ animals. (*E*) Average CFU-E from LSK cells sorted from control *ER-cre;Mll$^{F/+}$* or *ER-cre;Mll$^{F/F}$* 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 ± 95% CI, $n = 4$ 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 data are outlined in red. Arrows highlight that *Mll*-deficient HSC are not significantly closer to BM HSC (LSK/CD48^{neg}CD150^{+}) than they are to BM LSK/CD48^{+} progenitor cells. (*G*) Efficiency of *Mll* deletion in control *ER-cre;Mll$^{F/+}$* (blue) or *ER-cre;Mll$^{F/F}$* 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 *Mll* expression levels normalized to *Gapdh*. Error bars represent 95% CI, $n = 4$ animals per genotype; $*P \leq 0.07$, $**P \leq 0.05$.

1. Welch JJ, et al. (2004) Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* 104(10):3136–3147.
2. Koulnis M, et al. (2011) Identification and analysis of mouse erythroid progenitors using the CD71/TER119 flow-cytometric assay. *J Vis Exp* (54):e2809, 10.3791/2809.
3. Novershtern N, et al. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144(2):296–309.
4. He S, Kim I, Lim MS, Morrison SJ (2011) Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. *Genes Dev* 25(15): 1613–1627.

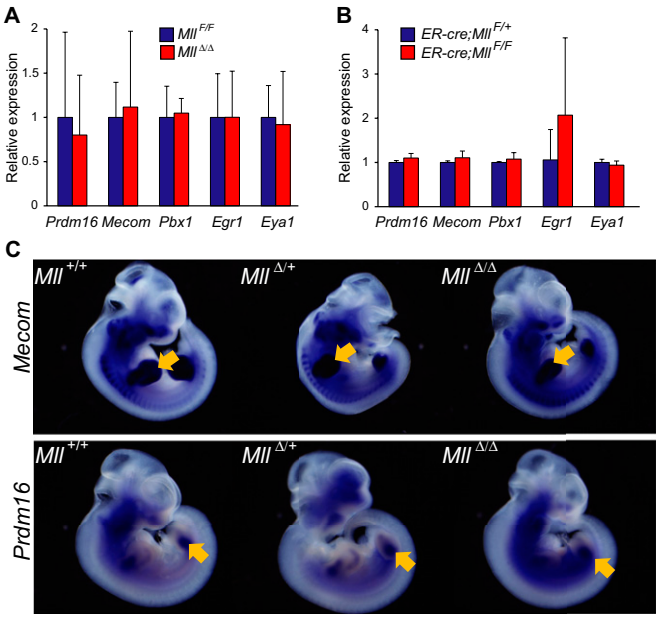


Fig. S2. HSC-specific *Mll* target genes are not *Mll*-regulated in nonhematopoietic tissues. (*A*) RT-qPCR results comparing gene expression levels in wild-type (blue) and *Mll*-deficient murine embryo fibroblasts (MEFs) ($Mll^{\Delta/\Delta}$, red) generated by infecting $Mll^{F/F}$ MEFs with a cre-expressing retrovirus. (*B*) RT-qPCR results using MEFS prepared from *ER-cre;*$Mll^{F/+}$ or *ER-cre;*$Mll^{F/F}$ 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, $Mll^{\Delta/+}$, and $Mll^{\Delta/\Delta}$ E9.5 embryos. Yellow arrows indicate limb buds, an example of normal expression of *Prdm16* and *Mecom* in $Mll^{\Delta/\Delta}$ embryos.

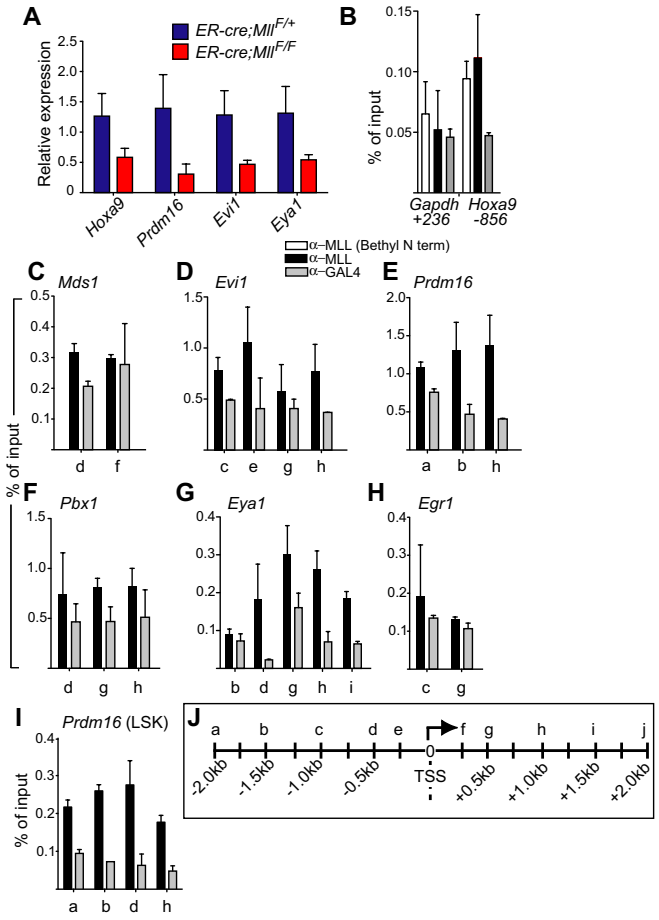


Fig. S3. Additional ChIP experiments support specific enrichment around certain TSS regions. (*A*) Putative MLL target genes are also *Mll* dependent in the total lineage-negative (Lin^{neg}) BM population. Lin^{neg} BM cells were enriched from *ER-cre*;$Mll^{F/+}$ (control, blue) or *ER-cre*;$Mll^{F/F}$ 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 5×10^4 lin^{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 5×10^4 lin^{neg} BM cells using primers surrounding the TSS of the indicated genes. (*I*) ChIP-qPCR results using 5×10^4 sorted LSK cells. (*J*) General diagram illustrating the position of amplicons shown in *C–I*; for specific positions, see Dataset S3.

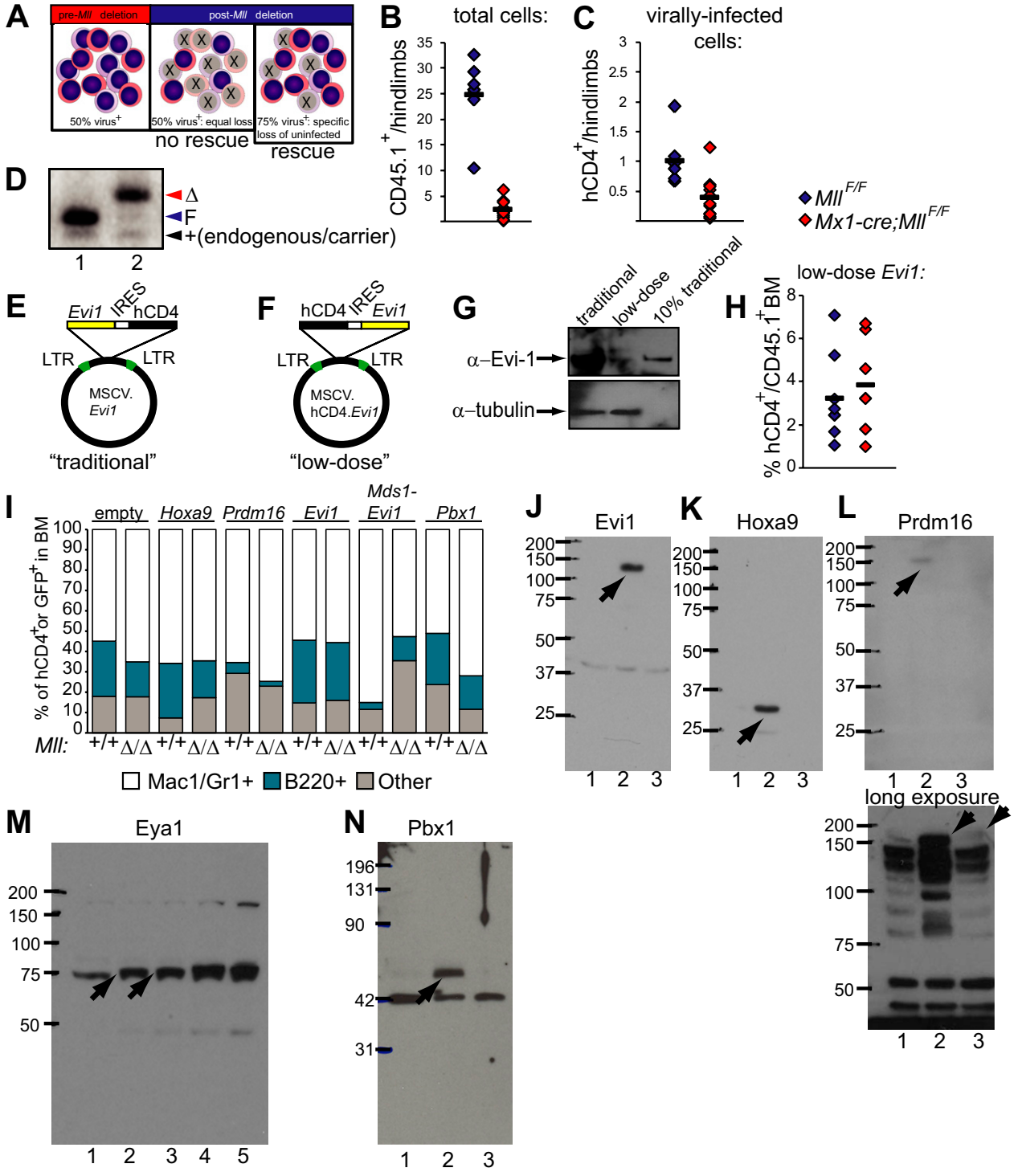
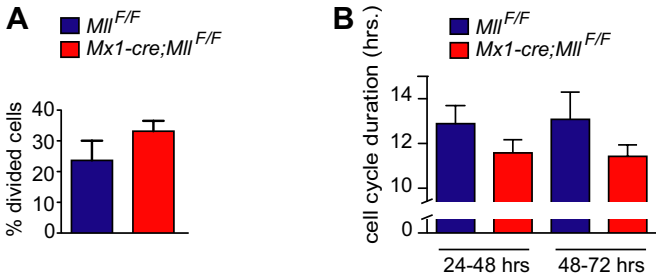


Fig. S4. Features of assays used to assess whether reexpression of *Mll* target genes can partially rescue *Mll*-deficient BM cells. (*A*) Scheme for assessing the ability of individual genes to rescue *Mll*-deficient BM cells based on the selective enrichment of retrovirus-infected cells ("Virus$^+$") 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% in the example) Virus$^+$ cells within the CD45.1$^+$ pool. (*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 *Mll*$^{F/F}$ (lane 1) or *Mx1-cre;Mll*$^{F/F}$ 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*$^+$, *Mll*^{F}, and *Mll*$^\Delta$ 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

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 pI: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 retrovirus 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.



C Gene sets downregulated in *Mll*-deficient HSCs:

Gene Set	NES	NOM p-val	FDR q-val
ST_G_alpha_i_Pathway_	-1.6047	0.01378	0.3324
nfkbPathway	-1.5162	0.03409	0.35688
tcrPathway	-1.6086	0.03536	0.37029
nthiPathway	-1.5176	0.06066	0.37583
il7Pathway	-1.5204	0.06836	0.39747
ST_ADRENERGIC	-1.5541	0.0251	0.39805
EMT_DOWN	-1.4085	0.02303	0.40423
tollPathway	-1.4908	0.01136	0.40935
TGF_Beta_Signaling_Pathway	-1.5323	0.02524	0.40996

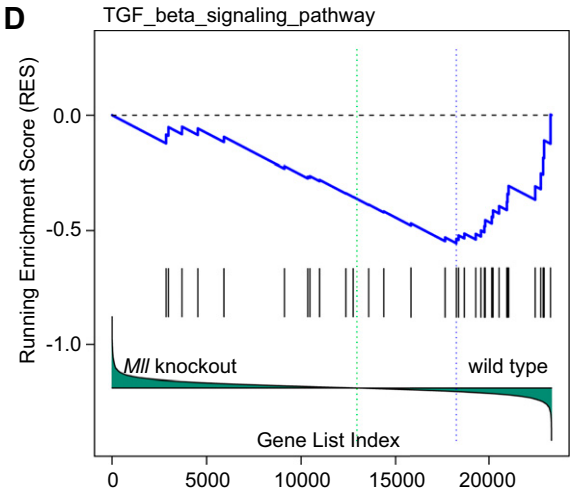


Fig. S5. Overall features of the single cell proliferation assays and TGFβ GSEA analysis. (*A*) The percentage of clones that ultimately completed at least one cell division comparing control (*Mll*$^{F/F}$) or *Mll*-deleted (*Mx1-cre;Mll*$^{F/F}$) LSK/CD48^{neg} single cells ("divided cells") is shown. 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 S1

Dataset S2. Genes down-regulated in *Mll*-deficient LSK/CD48neg cells

Dataset S2

Dataset S3. Primers used for ChIP-qPCR

Dataset S3

Dataset S4. Primers used for RT-qPCR

Dataset S4