Cancer Cell
Article

H3K79 Methylation Profiles Define Murine and Human MLL-AF4 Leukemias

Andrei V. Krivtsov,[1,5] Zhaohui Feng,[1,5] Madeleine E. Lemieux,[2] Joerg Faber,[1] Sridhar Vempati,[1,2] Amit U. Sinha,[1,2] Xiaobo Xia,[2] Jonathan Jesneck,[2] Adrian P. Bracken,[4] Lewis B. Silverman,[1,2] Jeffery L. Kutok,[3] Andrew L. Kung,[1,2] and Scott A. Armstrong[1,2,*]
[1]Division of Hematology/Oncology, Children's Hospital Boston
[2]Department of Pediatric Oncology, Dana-Farber Cancer Institute
[3]Department of Pathology, Brigham and Women's Hospital
Harvard Medical School, Boston, MA 02115, USA
[4]Department of Genetics, The Smurfit Institute, Trinity College, Dublin 2, Ireland
[5]These authors contributed equally to this work
*Correspondence: scott.armstrong@childrens.harvard.edu
DOI 10.1016/j.ccr.2008.10.001

SUMMARY

We created a mouse model wherein conditional expression of an Mll-AF4 fusion oncogene induces B precursor acute lymphoblastic (ALL) or acute myeloid leukemias (AML). Gene expression profile analysis of the ALL cells demonstrated significant overlap with human *MLL*-rearranged ALL. ChIP-chip analysis demonstrated histone H3 lysine 79 (H3K79) methylation profiles that correlated with Mll-AF4-associated gene expression profiles in murine ALLs and in human *MLL*-rearranged leukemias. Human *MLL*-rearranged ALLs could be distinguished from other ALLs by their H3K79 profiles, and suppression of the H3K79 methyltransferase DOT1L inhibited expression of critical MLL-AF4 target genes. We thus demonstrate that ectopic H3K79 methylation is a distinguishing feature of murine and human MLL-AF4 ALLs and is important for maintenance of MLL-AF4-driven gene expression.

INTRODUCTION

Leukemias mediated by *MLL* rearrangements possess unique clinical and biological features. *MLL* rearrangements can be found in acute lymphoblastic leukemia (ALL), acute myeloid leukemias (AML), and acute biphenotypic or mixed-lineage leukemia (MLL). *MLL* translocations are present in over 70% of cases of infant leukemias (Biondi et al., 2000) and in general account for approximately 5% of ALLs, 5%–10% of AMLs, and a significant portion of acute biphenotypic or mixed-lineage leukemias (Huret et al., 2001). Patients with *MLL*-rearranged ALL have a particularly unfavorable prognosis as compared to patients with other forms of ALL (Chen et al., 1993). To date, more than 50 MLL fusion partner genes have been reported (Ayton and Cleary, 2001; Krivtsov and Armstrong, 2007). The t(4;11)(q21;q23) encodes MLL-AF4 and is the most frequent *MLL* translocation found in ALL.

Multiple mouse models have been developed that recapitulate *MLL*-rearranged AML (reviewed in Ayton and Cleary, 2001). However, development of models that faithfully recapitulate MLL fusion-mediated ALL has proven more difficult. Constitutive *Mll-AF4* knockin results in mixed lymphoid/myeloid hyperplasia and mature B cell neoplasms in mice (Chen et al., 2006). Conditional expression of *Mll-AF4* based on interchromosomal recombination in lymphoid lineages produces mature B cell lymphomas (Metzler et al., 2006). Thus, further development of murine models of *MLL*-rearranged ALL is needed.

MLL is a mammalian homolog of *Drosophila* trithorax and possesses multiple functional domains, including amino-terminal AT hooks that bind DNA and a carboxy-terminal Su(var)3-9, Enhancer-of-zeste, Trithorax (SET) domain that methylates lysine 4 of histone H3 (H3K4). H3K4 methylation is associated with transcriptional activation (reviewed in Shilatifard, 2006),

SIGNIFICANCE

The t(4;11) encodes an MLL-AF4 fusion protein and predicts a particularly poor prognosis when found in acute lymphoblastic leukemia (ALL). Recent studies suggest certain MLL fusion proteins enhance gene expression by recruitment of the histone H3 lysine79 (H3K79) methyltransferase DOT1L. We demonstrate that H3K79 methylation is enhanced at many loci in leukemia cells from a murine model of Mll-AF4 and in human MLL-AF4 leukemia cells, and that this elevation is correlated with enhanced gene expression. Furthermore, suppression of H3K79 methylation leads to inhibition of gene expression in MLL-AF4 cells. These data demonstrate that inhibition of DOT1L may be a therapeutic approach in this disease and that this mouse model should be useful for assessment of therapeutic approaches for *MLL*-rearranged ALL.

and MLL positively regulates expression of clustered homeobox (*HOX*) and other genes during development, at least in part via H3K4 methylation (Milne et al., 2002; Nakamura et al., 2002). Mll has important roles in development, including hematopoietic development (Hess et al., 1997; Jude et al., 2007).

MLL translocations invariably encode fusion proteins that have lost the H3K4 methyltransferase (SET) domain. However, MLL fusion proteins retain the ability to bind *HOX* genes and other promoter regions and are associated with enhanced gene expression (Armstrong et al., 2002; Guenther et al., 2005; Rozovskaia et al., 2001; Yeoh et al., 2002; Zeisig et al., 2004). Several mechanisms have been proposed as to how MLL fusions may deregulate gene expression, including recruitment of abnormal histone modification activities (Cheung et al., 2007; Krivtsov and Armstrong, 2007; Okada et al., 2005). For example, MLL-AF10 and MLL-ENL have been shown to recruit the non-SET domain methyltransferase DOT1L, which promotes methylation of histone H3 lysine 79 (H3K79) on the *HOXA9* promoter (Mueller et al., 2007; Okada et al., 2005; Zeisig et al., 2005). Since H3K79 methylation is linked to positive transcriptional regulation (Schubeler et al., 2004; Shilatifard, 2006), DOT1L-mediated methylation of H3K79 may contribute to increased expression of *HOXA9* in MLL-AF10- and MLL-ENL-induced leukemias. Furthermore, AF10, ENL, and other MLL fusion partners such as AF4 and AF9 are normally found in nuclear complexes associated with DOT1L (Bitoun et al., 2007; Mueller et al., 2007; Okada et al., 2005; Zeisig et al., 2005; Zhang et al., 2006). Thus, aberrant recruitment of DOT1L to the promoters of MLL target genes may be a common feature of many oncogenic MLL fusion proteins. However, the extent of H3K79 methylation changes and the specificity of these epigenetic changes for *MLL*-rearranged leukemias have not been defined.

Here, we report the development of a murine model in which conditional expression of MLL-AF4 induces both ALL and AML. Genome-wide assessment of gene expression and H3K79 methylation demonstrates that this model faithfully recapitulates human ALL resulting from *MLL-AF4* translocation and identifies ectopic H3K79 methylation as an important part of MLL-AF4-driven gene expression and transformation.

RESULTS

Generation of a Conditional *Mll-AF4* Knockin Mouse

We used a conditional expression approach that proved successful for development of an Mll-Cbp myelodysplasia/AML model to create a model in which the Mll-AF4 fusion product was conditionally expressed from the endogenous *Mll* locus. We engineered a conditional Mll-AF4^{stop} targeting construct by replacing the *Cbp* cDNA in the previously reported Mll-Cbp^{stop} targeting vector (Wang et al., 2005) with a cDNA encoding the C-terminal portion of human AF4 (Figure 1A). This generated a targeting construct that placed the human AF4 sequence in the murine *Mll* exon 8, downstream of a transcriptional stop site flanked by *LoxP* sites, which allows conditional expression of an *Mll-AF4* fusion RNA upon expression of Cre recombinase (Figure 1B). The construct was electroporated into CJ7 mouse embryonic stem cells (Swiatek and Gridley, 1993), and clones possessing the targeted allele were selected by Southern blot (Figure 1C) and used to achieve germline transmission of the knockin allele. Mice heterozygous for the *Mll-AF4*^{stop} conditional allele were born at slightly less than Mendelian frequency, presumably due to heterozygosity for *Mll* (Yu et al., 1995). Founder mice were backcrossed to C57BL6/129 F1 mice. We used Mll-AF4-specific primers to confirm the absence of the fusion RNA in bone marrow from mice heterozygous for the *Mll-AF4*^{stop} allele.

Mll-AF4 Expression Enhances Serial Replating of Lymphoid Progenitors

First, we determined whether expression of Mll-AF4 could transform lymphoid cells in vitro. We collected bone marrow from *Mll-AF4*^{stop} (MA4) heterozygous mice (three experiments, five mice each) 5 days after 5-fluorouracil (5-FU) treatment and transduced cells with retroviruses encoding either GFP control (MIG) or Cre-GFP (Cre) to initiate expression of Mll-AF4. Cells were then cultured in semisolid media supplemented with IL-7, SCF, and FLT3. After 14 days of culture, over 90% of cells in both Cre and MIG groups expressed CD19 and B220. Weekly replating of 1×10^4 MIG-transduced cells exhausted their colony-forming potential by the third week, while similar replating of Cre-transduced cells did not exhaust their replating potential for at least 6 weeks (see Figure S1A available online). To assess for gene expression changes associated with Mll-AF4 expression, we extracted RNA from 1×10^5 cells at the end of the second week of plating and amplified, labeled, and hybridized labeled RNA to Affymetrix mouse 430 A2.0 microarrays. Supervised analysis identified *HoxA5*, *HoxA9*, *Runx2*, *Meis1*, and *Mylk* among the 20 most upregulated genes in the cells expressing Mll-AF4 (Figure S1B). These genes are also found as central members of early gene expression changes associated with MLL-AF9 expression in myeloid cells and in human *MLL*-rearranged lymphoblastic leukemias (Armstrong et al., 2002; Krivtsov et al., 2006). Thus, conditional expression of Mll-AF4 in lymphoid cells leads to in vitro transformation and gene expression changes associated with human MLL fusion leukemias.

Mll-AF4 Expression Induces Acute Leukemias

The initial strategy we used to activate Mll-AF4 expression utilized a self-excising retrovirus that transiently expresses Cre recombinase in transduced cells (Silver and Livingston, 2001). We either collected bone marrow (BM) from heterozygous MA4 mice 5 days after 5-FU treatment or used 5-FU-untreated bone marrow depleted for cells expressing CD3, CD4, CD8α, F4/80, B220, Gr1, and TER119 (Lin$^-$). BM cells were transduced with retroviruses encoding either "hit-and-run Cre" (HR-Cre) or GFP (MIG) and transplanted into lethally or sublethally irradiated syngeneic recipients. The expression of Mll-AF4 in HR-Cre cells but not MIG-transduced cells was confirmed by RT-PCR (Figure 1D). Mice (n = 43) expressing the *Mll-AF4* allele as a result of retroviral transduction of Cre developed a fatal disease (Table 1; Figures S2A and S2B) consistent with acute leukemia including bone marrow replacement, splenomegaly, and variable lymphadenopathy (Table 1; Figure 2A). Mice transplanted with BM from MA4 mice (n = 20) transduced with MIG did not develop leukemia (Figure S2A).

In order to confirm leukemia development in a purely genetic model, we crossed MA4 mice with mice transgenic for *Mx1-Cre* (Kuhn et al., 1995). *Mx1-Cre* × MA4 mice and control littermates received three intraperitoneal injections of polyinosinic/polycytidylic

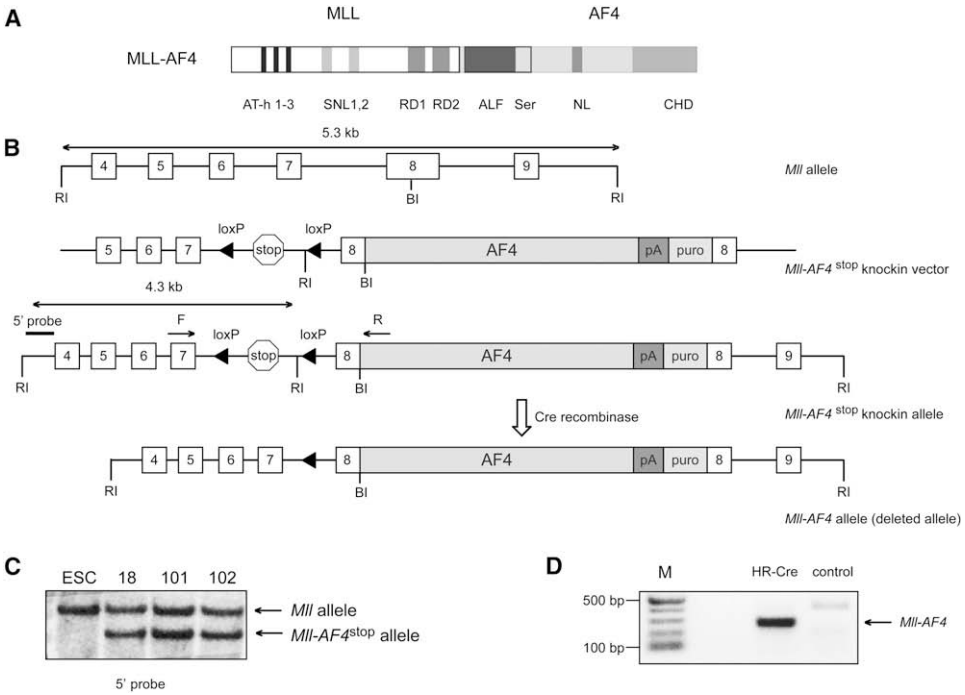


Figure 1. Generation of Mice with a Conditional *Mll-AF4*^{stop} Allele

(A) Schematic representation of the MLL-AF4 fusion protein including a number of protein motifs. AT-h 1–3, AT hooks; SNL1 and SNL2, speckled nuclear localization sites; RD1 and RD2, transcriptional repression domain consisting of two functional subunits; ALF, homologous regions among the AF4/FMR2 family members; Ser, serine-rich regions that contain a transactivation domain; NL, nuclear localization signal; CHD, C-terminal homology domain conserved between the FMR2, AF4, and LAF4 family of transcription factors.

(B) A diagrammatic description of the *Mll-AF4*^{stop} knockin allele. *Mll* exons are numbered 4 to 9. The polyadenylation (pA) site and puromycin (puro) resistance cassette are shown. EcoRI (RI) and BamHI (BI) sites and *Mll* forward (F) and *AF4* reverse (R) primers used for *Mll-AF4* fusion RNA detection are shown.

(C) Southern blot analysis of EcoRI-digested genomic DNA from wild-type and targeted embryonic stem cells (ESC) probed with the 5′ probe shown in (B).

(D) RT-PCR analysis of *Mll-AF4* transcript expression in total bone marrow from *Mll-AF4*^{stop} mice retrovirally transduced with either "hit-and-run" Cre (HR-Cre) or MSCV-MIG (control). M, molecular weight marker in base pairs (bp).

acid (pIpC). *Mll-AF4* transcript expression was confirmed by RT-PCR. Fourteen of 22 *Mx1-Cre* × MA4 mice developed a disease consistent with acute leukemia, and three died from unknown causes within 250 days from the initial pIpC treatment. Mice that possessed only *Mx1-Cre* or *Mll-AF4*^{stop} did not develop leukemia (Table 1; Figure S2C). The median latency of leukemias

Table 1. Characteristics of Mice with Mll-AF4-Induced Leukemias

	Latency in Days (Mean Range)	SP Weight in mg (Mean Range)	WBC × 1000/μl (Mean Range)	HCT % (Mean Range)	Platelets × 1000/μl (Mean Range)	% of Leukemia Cells in BM
Normal		70–110	5–10	45–47	800–1400	
tBM TR (n = 37)						
ALL (n = 11)	122.8 (71–220)	425.5 (200–670)	138.0 (17–476)	32.1 (20–40.9)	447.9 (144–920)	85.3 (52–99)
AML (n = 23)	75.6 (25–221)	404.8 (150–800)	124.8 (8–389)	24.7 (8.9–41)	310.3 (70–824)	86.0 (54–98)
MLL (n = 3)	92 (81–105)	666.7 (390–810)	224.3 (155–281)	23.7 (17.7–26)	201 (99–312)	59 (39–75)
Lin$^-$ TR (n = 7)						
ALL (n = 4)	67.5 (62–75)	287.5 (200–450)	459.8 (236–748)	25.5 (16.1–39)	373.5 (251–496)	88.3 (67–98)
AML (n = 2)	103 (98–108)	515 (470–560)	72.5 (23–122)	33.6 (29.9–37)	394.5 (367–422)	83.5 (80–87)
Alive (n = 1)						
Mx1-Cre (n = 22)						
ALL (n = 8)	152.3 (120–242)	691 (260–2100)	43.0 (5–92.6)	30.3 (5.2–44)	640 (102–1683)	46.6 (24–86)
AML (n = 6)	143.5 (67–240)	485.0 (300–910)	84.3 (6–279)	23.5 (10.4–40)	732 (400–1463)	92.2 (84–98)
No leukemia (n = 3)						
Alive (n = 5)						

SP, spleen; WBC, white blood cells; HCT, hematocrit; BM, bone marrow; TR, transgenic.

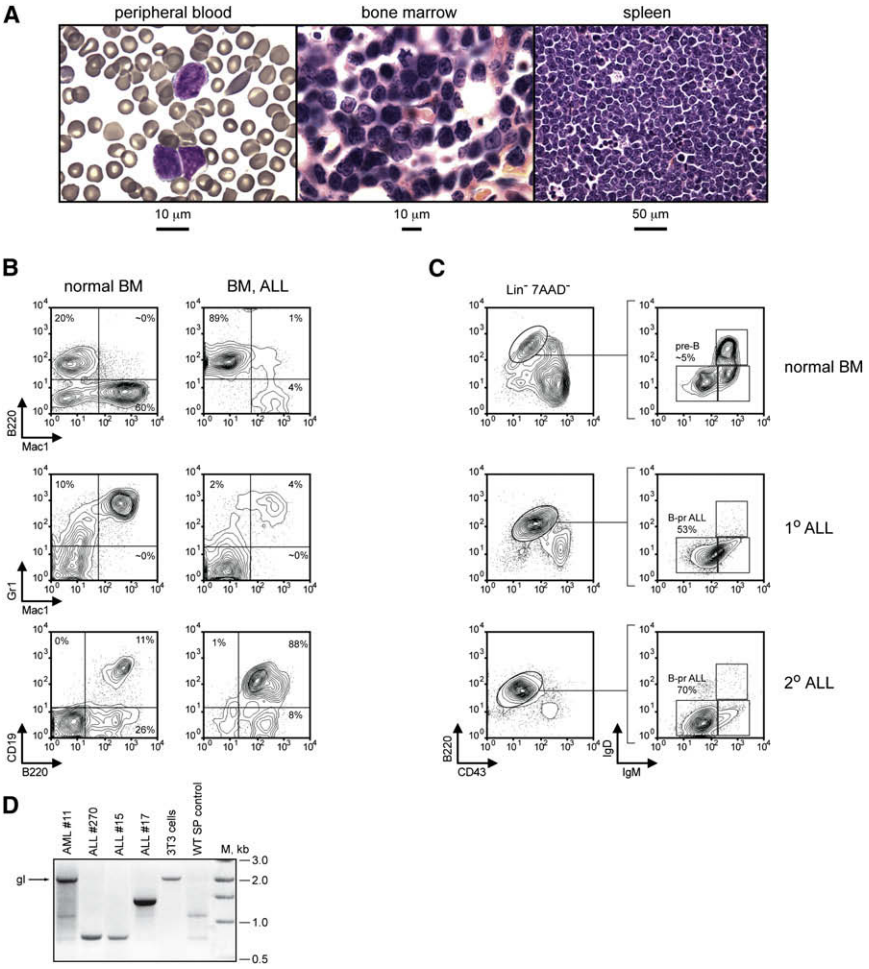


Figure 2. Characterization of Mll-AF4-Mediated B Precursor ALLs

(A) Histopathologic analysis of B precursor (B-pr) ALL in *Mll-AF4* mice. Cells consistent with lymphoblasts are found in the peripheral blood (left). The bone marrow (center) and spleen (right) are infiltrated with leukemia cells.

(B) FACS analysis of normal bone marrow (BM) (left) and leukemic BM (right) using antibodies against B220, Gr1, Mac1, and CD19.

(C) Detailed immunophenotypic analysis of normal BM (top), leukemic BM from a mouse with primary ALL (1°, middle), and ALL BM from a secondary recipient mouse (2°, bottom).

(D) PCR-based analysis showing germline or rearranged immunoglobulin DJ_H heavy-chain loci in AML cells (11), ALL cells (270, 15, and 17), NIH 3T3 cells, or normal splenocytes (WT SP). The arrow indicates germline (gl) DJ_H configuration. M, molecular weight marker in kilobases (kb).

originated using this approach was longer (131 days) than either of the transplant models (82 and 71 days), and these mice tended to have less BM involvement than mice that received the retrovirally delivered Cre (Table 1; Figures S2A–S2C).

Histopathologic evaluation of tissues from moribund mice demonstrated bone marrow, splenic, and liver infiltration with immature hematopoietic cells (Figure 2A). The 31 cases with a morphologic appearance most consistent with myeloid leukemia possessed immunophenotypes consistent with AML: $Mac1^+Gr1^+B220^-CD19^-$ (n = 26) (Figure S2D), $Mac1^-Gr1^+B220^-CD19^-$ (n = 3), or $Mac1^+Gr1^-B220^-CD19^-$ (n = 2). The 26 cases whose morphologic appearances were most consistent with a lymphoid leukemia possessed immunophenotypes that included expression of the lymphoid markers B220 and/or CD19. Sixteen of 26 lymphoid-appearing leukemias were $B220^+CD19^+Mac1^-$ (Figure 2B). Southern and PCR analyses of genomic DNA from spleen cells isolated from leukemic mice demonstrated that all $CD19^+B220^+$ leukemias analyzed possessed a clonal DJ_H rearrangement (Figure 2D; Figures S2E and S2F). This type of leukemia was therefore consistent with typical B cell ALL. None of the AMLs analyzed possessed DJ_H rearrangements (Figure 2D; Figures S2E and S2F). Three lymphoid-appearing leukemias were $B220^+CD19^-Mac1^+$ (Figure S2D), and based on the coexpression of B220 and Mac1, these leukemias were termed mixed-lineage leukemias (MLL). One of three possessed a single DJ_H configuration as assessed by a PCR-based approach (Figure S2E). Three leukemias were $B220^+CD19^-Mac1^-$, one of which possessed a clonal DJ_H rearrangement in support of a very early lymphoid arrest, and thus were designated pro-B ALL (Figures S2D and S2E). Four lymphoid-appearing leukemias were $CD19^+B220^-Mac1^-$. All lymphoid leukemias expressed the *Mll-AF4* fusion RNA (Figure S2G).

We assessed the ability of the leukemias to initiate disease in secondary recipient mice to functionally assess the presence of leukemia-initiating cells (LICs). BM cells from mice with AML (n = 1), MLL (n = 2), or B precursor (B-pr) ALL (n = 4) could initiate leukemia in secondary recipient mice (n = 21) when 3×10^4 cells were injected (data not shown). We performed limiting dilution transplant experiments to determine the frequency of LICs in the $B220^+CD19^+$ B-pr ALLs (n = 2) and found that all mice (n = 22) that received as few as 1×10^3 sorted $B220^+CD19^+$ leukemia cells developed an identical ALL (Figure 2C), whereas 3 of 9 mice that received 100 cells developed leukemia. We also further characterized the MLLs (n = 2) for LICs and found that all secondary recipient mice (n = 24) that received 5×10^4 sorted $B220^+Mac1^+$ cells developed an identical MLL (data not shown), whereas none of the mice (n = 20) that received 1×10^3 cells developed leukemia. These data confirm the development of an acute leukemia in this model system and demonstrate the presence of LICs in leukemias with different immunophenotypes.

We further characterized the $CD19^+B220^+$ leukemias and defined the precise stage of B-lymphoid development that most closely matched the leukemias. We performed immunophenotypic analysis that assessed B220, CD43, IgM, and IgD expression to compare the ALL cells to various stages of B cell

development. Bone marrow cells taken from wild-type mice demonstrated appropriate percentages of B cells at various stages of differentiation (Figure 2C). However, the majority of lymphoid cells present in BM from mice with primary or secondary B-pr ALL were $B220^+CD43^-IgM^-IgD^-$, most consistent with mouse pre-B cells (Hardy, 1990). Thus, the leukemias represent an expansion of cells that by immunophenotypic analysis are arrested at the pre-B cell stage of development, thus confirming their identity as B-pr ALL.

Murine Mll-AF4 ALL Recapitulates Human *MLL*-Rearranged ALLs

To more definitively determine the developmental stage of leukemia cells, we compared the global gene expression pattern of the expanded leukemia cells to normal B cells. We isolated RNA from 16 primary B-pr ALLs and normal lymphocyte populations (Hardy, 1990) including pro-B cells ($Lin^-B220^+CD43^+$), pre-B cells ($Lin^-B220^+CD43^-IgM^-IgD^-$), immature B cells ($Lin^-B220^+CD43^-IgM^+IgD^-$), and mature B cells ($Lin^-B220^+CD43^-IgM^+IgD^+$) and amplified and hybridized RNA to Affymetrix 430 A2.0 microarrays. Unsupervised hierarchical clustering analysis demonstrated that 14 ALLs were more similar to pre-B cells and that two ALLs were more similar to pro-B cells than to immature B or mature B cells (Figure 3A), providing further evidence for leukemia cell expansion at an early stage of B cell development. The relationship between the ALL samples and normal B cells was further tested by consensus clustering, which demonstrated that ALLs 117 and 122 clustered separately and that the other ALLs were most similar to pre-B cells (Figure S3A).

Supervised gene expression analysis demonstrated that murine B-pr ALL cells express high levels of certain *HoxA* cluster genes but do not differentially express *HoxB*, *HoxC*, or *HoxD* cluster genes (Figure 3B). The elevated expression of *HoxA9* was confirmed by quantitative PCR (qPCR) (Figure S3B). The murine ALL cells also expressed high levels of *Meis1* as measured by microarrays and qPCR (Figure S3C), similar to human *MLL*-rearranged ALL (Armstrong et al., 2002).

Next, to assess globally whether mouse Mll-AF4 ALLs recapitulate gene expression profiles observed in human *MLL*-rearranged ALLs, we compared gene expression signatures of mouse Mll-AF4 ALL to human *MLL*-rearranged ALL. We identified the top 500 probe sets (386 genes) with increased expression in murine Mll-AF4 B-pr ALL as compared to pre-B cells using a signal-to-noise statistic and converted them to human homologs found on Affymetrix U133 arrays (259 genes). Gene set enrichment analysis (GSEA) demonstrated strong enrichment ($p < 0.014$) of the murine Mll-AF4 signature in human *MLL*-rearranged ALLs as compared to *MLL*-germline ALLs (Ross et al., 2003; Subramanian et al., 2005) (Figure 3C). These data demonstrate that the ALLs in our mouse model recapitulate gene expression programs found in human MLL-AF4-induced disease.

H3K79 Methylation in Murine Mll-AF4 ALL

Since AF4 associates with the DOT1L methyltransferase (Bitoun et al., 2007), we hypothesized that Mll-AF4 may be associated with ectopic H3K79 methylation. First, we determined that *Dot1L* is ubiquitously expressed in B cell progenitors and B-pr ALL (Figure S4A). We used chromatin immunoprecipitation (ChIP) to assess histone methylation near the promoters of *HoxA* genes. We sorted pre-B cells and leukemia cells with a similar immunophenotype and performed ChIP with antibodies directed against dimethylated H3K79 (H3K79me2), trimethylated lysine 27 (H3K27me3), trimethylated lysine 36 (H3K36me3), or trimethylated lysine 4 (H3K4me3). We assessed the ChIP DNA by qPCR using primers that spanned the promoter region of *HoxA9* and compared the amount of precipitated DNA as a percentage of input in normal versus leukemia cells (Figure S4B). We found slightly elevated H3K4me3, no obvious change in H3K27me3 or H3K36me3, and dramatically enhanced H3K79me2 associated with *HoxA* gene promoters in the leukemia cells (Figures 4A and 4C). Next, we assessed expression levels of *HoxA9* in pre-B ALLs and normal B cell progenitors and found a strong correlation between the amount of H3K79me2 and *HoxA9* expression (Figure 4B). We expanded our assessment and found enhanced H3K79me2 across much of the *HoxA* cluster (Figure 4C). Given the data demonstrating a potential role for DOT1L in some *MLL*-rearranged leukemias and these data demonstrating enhanced H3K79me2 in Mll-AF4 ALL cells, we focused further assessment on H3K79 methylation.

Genome-wide assessment of DNA binding of MLL through the use of ChIP followed by microarray analysis (ChIP-chip) demonstrated association of MLL with the promoter regions in thousands of genes (Guenther et al., 2005). This prompted the question of whether the MLL-AF4 fusion influences a select set of genes such as *HoxA* genes or whether there are more widespread histone methylation abnormalities in Mll-AF4 leukemias. We performed ChIP-chip analysis on three mouse Mll-AF4 B-pr ALLs and three normal pre-B samples using an antibody against H3K79me2 and Affymetrix 1R mouse promoter arrays. These studies verified an increase in H3K79me2 across the *HoxA* cluster in leukemia cells (Figure 4D). A second group of six murine Mll-AF4 ALLs confirmed H3K79me2 across the *HoxA* loci but not the *HoxB*, *HoxC*, or *HoxD* clusters (Figures S5A–S5C). An expanded view across the genome revealed thousands of loci with significant H3K79me2 in both normal pre-B cells and ALL cells, as would be expected given a normal role for H3K79me2 in epigenetic regulation. However, on a comparative basis, there were 1186 promoter regions in which H3K79me2 was increased in the leukemia cells in comparison to normal pre-B cells (B-pr ALL H3K79 signature), whereas only 285 promoter regions had H3K79me2 that was higher in normal pre-B cells (pre-B H3K79 signature) (Figure 4E).

As H3K79me2 has been associated with positive regulation of transcription (Schubeler et al., 2004; Shilatifard, 2006), we used GSEA to determine whether H3K79me2 was correlated with gene expression. Genes associated with increased H3K79me2 in Mll-AF4-induced leukemia cells (B-pr ALL H3K79 signature) were highly enriched for genes with elevated mRNA expression in leukemia cells ($p < 0.005$; Figure 4F). Likewise, genes with elevated H3K79me2 within normal pre-B cells (pre-B H3K79 signature) were associated with elevated expression in normal pre-B cells ($p < 0.005$; Figure 4G). These analyses demonstrate that ectopic H3K79me2 in Mll-AF4 leukemias is associated with enhanced gene expression.

H3K79 Methylation in Human *MLL*-Rearranged ALLs

To assess whether similar abnormalities in H3K79me2 exist in human *MLL*-rearranged ALLs, we assessed histone methylation

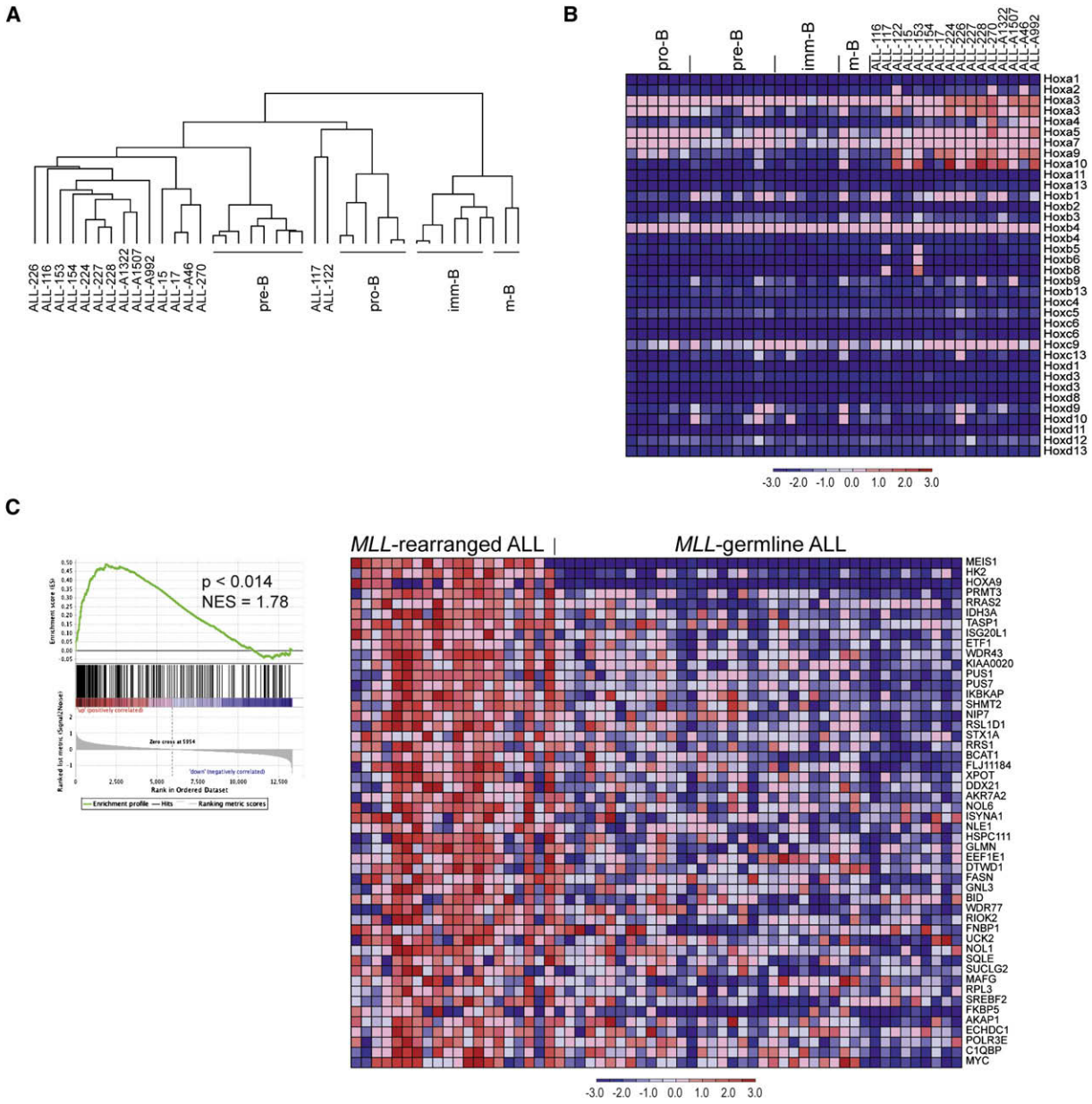


Figure 3. Gene Expression in Normal and Malignant Mouse B Cells and Comparison to Human Leukemias

(A) Hierarchical clustering using 10,218 filtered (minimum fold 3; minimum delta 100; background 20; ceiling 20,000) probe sets.

(B) Expression of *HoxA*, *B*, *C*, and *D* cluster genes in normal developing B lymphocytes and B-pr ALL cells.

(C) Gene set enrichment analysis (GSEA) of gene expression in human *MLL*-rearranged ALL (n = 20) as compared to *MLL*-germline ALL (n = 40) (Ross et al., 2003) using the top 386 genes identified as highly expressed in murine Mll-AF4 ALL as a gene set. Left: GSEA enrichment plot. Right: the top 50 genes showing increased expression in human *MLL*-rearranged leukemias.

in *MLL*-rearranged ALLs as compared to $Lin^{-}CD34^{+}CD19^{+}$ BM cells. First, we performed ChIP-qPCR to assess H3K79me2 associated with *HOXA* cluster loci. We found elevated H3K79me2 associated with multiple *HOXA* genes (Figure 5A). Next, we performed ChIP-chip analyses on five *MLL*-rearranged ALLs and five BM $Lin^{-}CD34^{+}CD19^{+}$ cell samples using Affymetrix 1R human promoter arrays. The data demonstrate elevation of H3K79me2 associated with multiple *HOXA* cluster genes similar to that seen in the mouse leukemias (Figure 5B; Figure S6A). We assessed genome-wide differences in H3K79me2 in *MLL*-rearranged ALL as compared to normal $CD34/CD19^{+}$ cells just as we did for murine cells. There were many genomic regions associated with H3K79me2 in both cell types. We found elevated H3K79me2 associated with 1378 promoter regions in the leukemias and elevated H3K79me2 associated with 562 promoter regions in the normal samples, thus demonstrating that more genes are associated with increased H3K79me2 in *MLL*-rearranged leukemias as compared to normal cells (Figure 5C). Since

the general trend in H3K79me2 appeared similar in mouse and human cells, we were interested to see whether there was overlap between the genes associated with H3K79me2 in mouse and human cells. We compared the B-pr ALL H3K79 signature derived from mouse and human cells and identified 369 genes associated with increased H3K79me2 in both mouse and human MLL-AF4 leukemias (Figure 5D). We also identified 48 genes associated with increased H3K79me2 in normal B-pr cells in mouse and human (Figure 5D). These data show that the epigenetic differences between *MLL*-rearranged ALL and normal B-pr cells are similar between our mouse model and the human disease, with more genes associated with enhanced H3K79me2 in leukemia cells than in normal cells.

Next, we reasoned that genes found to be associated with H3K79me2 in both mouse and human MLL-AF4 leukemias may represent a set of MLL-AF4 genes and should be functionally associated with increased gene expression in human *MLL*-rearranged ALLs as compared to ALLs that are not driven by MLL fusion oncoproteins. We performed GSEA using the 369 genes found in the mouse/human H3K79me2 overlap as a gene set and gene expression data from *MLL*-rearranged and germline ALL samples (Ross et al., 2003). The analysis demonstrated significant correlation between elevated H3K79me2 and elevated mRNA expression in *MLL*-rearranged ALL in comparison to other ALLs ($p < 0.02$; Figure 5E). The 48 genes that were common to mouse and human normal pre-B cells were not associated with gene expression changes in either *MLL*-rearranged or other ALLs ($p < 0.49$; Figure 5F). These data demonstrate that ectopic H3K79me2 is associated with gene expression in human leukemias just as in mouse leukemias.

H3K79me2 Profiles Distinguish Human *MLL*-Rearranged ALL from Other ALLs

Having demonstrated that we can identify histone methylation abnormalities associated with the presence of MLL-AF4 in mouse as well as human *MLL*-rearranged leukemias, we wondered whether the histone methylation abnormalities were sufficiently different between *MLL*-rearranged and *MLL*-germline leukemias such that global patterns of H3K79me2 could be used to distinguish *MLL*-rearranged ALLs from *MLL*-germline ALLs and normal B cell progenitors. We assessed genome-wide patterns of H3K79me2 in five *MLL*-rearranged ALL samples, seven *MLL*-germline ALL samples, and five normal controls. Similar to normal pre-B cells and *MLL*-rearranged leukemia, ALL samples with germline *MLL* had thousands of loci with H3K79me2. We developed a method to use ChIP-chip histone methylation data to cluster samples. After individually processing each promoter array with an enrichment p value cutoff of 10^{-5}, we associated the resulting ChIP-chip regions with RefSeq genes if the site fell within a window extending from −500 to +2500 bp of the transcriptional start site. We retained only those genes that had H3K79me2 in at least 3 of the 17 samples. This filtering identified 5438 distinct genes with H3K79me2. Remarkably, the H3K79 epigenetic profile of *MLL*-rearranged leukemia samples was clearly distinct from normal CD34/CD19$^+$ cells and from *MLL*-germline ALLs when clustered with hierarchical clustering (Figure 6A) or when projected using principal component analysis (PCA) (Figure 6B). Thus, *MLL*-rearranged ALLs have a global H3K79me2 profile that is sufficiently distinct to distinguish *MLL*-rearranged ALLs from other lymphoblastic leukemias, just as profiles of mRNA expression can distinguish these classes (Armstrong et al., 2002). Next, we identified the genes that had the highest mean histone methylation score in *MLL*-rearranged samples compared to normal CD34/CD19$^+$ cells and *MLL*-germline ALL cells. This analysis identified a number of genes previously associated with *MLL*-rearranged leukemias, including *HOXA* genes, *MEIS1*, *RUNX2*, *PROM1*, and *FLT3* (Figures 6C and 6D; Figure S6B) (Armstrong et al., 2002, 2003; Krivtsov and Armstrong, 2007; Krivtsov et al., 2006; Rozovskaia et al., 2001). A number of other genes of interest, including *BCL2* and *CEBPα*, have previously been implicated in leukemias, but not specifically those with *MLL* rearrangements. *CEBPα* is of particular interest given its role in specification of myeloid cell fate and the propensity for myeloid gene expression in MLL-AF4 ALLs. These data demonstrate a unique H3K79me2 profile for *MLL*-rearranged leukemias as compared to other ALLs, which includes a number of loci previously identified as playing important roles in *MLL*-rearranged leukemias.

DOT1L Suppression Inhibits H3K79me2 and *HOXA* Expression in MLL-AF4 Cells

We have demonstrated a strong correlation between H3K79me2 and aberrant *HOXA* gene expression in MLL-AF4 cells. Thus, we were interested to know whether continued expression of the H3K79 methyltransferase DOT1L was required to maintain aberrant *HOX* gene expression, as a first step toward assessing DOT1L as a potential therapeutic target in this disease. We developed two lentiviral-based shRNA constructs directed toward *DOT1L* and transduced two different t(4;11) (MLL-AF4) cell lines. Forty-eight hours after transduction with a viral titer sufficient to transduce >90% of the cells, *DOT1L* expression was assessed by quantitative RT-PCR, which demonstrated suppression of *DOT1L* RNA to approximately 20%–30% of control levels (Figures 7A and 7B). Suppression of DOT1L protein expression was confirmed in one of the two cell lines (Figure 7B). Seventy-two hours after transduction, we assessed H3K79me2 across the *HOXA* cluster and found diminished H3K79me2 in cells transduced with the *DOT1L*-directed shRNA (Figures 7C and 7D). This reduction of H3K79 across the *HOXA* locus resulted in decreased expression of both *HOXA5* and *HOXA9* (Figures 7E and 7F). These data demonstrate that continued expression of DOT1L is required for H3K79me2 and *HOXA* gene expression in MLL-AF4 cells.

DISCUSSION

Translocations involving the *MLL* gene result in both AML and ALL in humans. *MLL-AF9* and *MLL-AF4* are the most common translocations resulting in AML and ALL, respectively. Although a number of retroviral and genetically engineered mouse models of MLL fusion-mediated AML have been developed, models recapitulating MLL-AF4-mediated ALL have been more elusive. Ectopic expression of MLL-AF4 through retroviral transduction has been difficult, and other genetically engineered mouse models have resulted in myelodysplasia or mature B cell lymphomas. We report here the generation of a conditional knockin model in which the MLL-AF4 fusion product is expressed within the context of the endogenous *MLL* locus. Upon conditional activation, mice develop AML or ALL, the latter with an

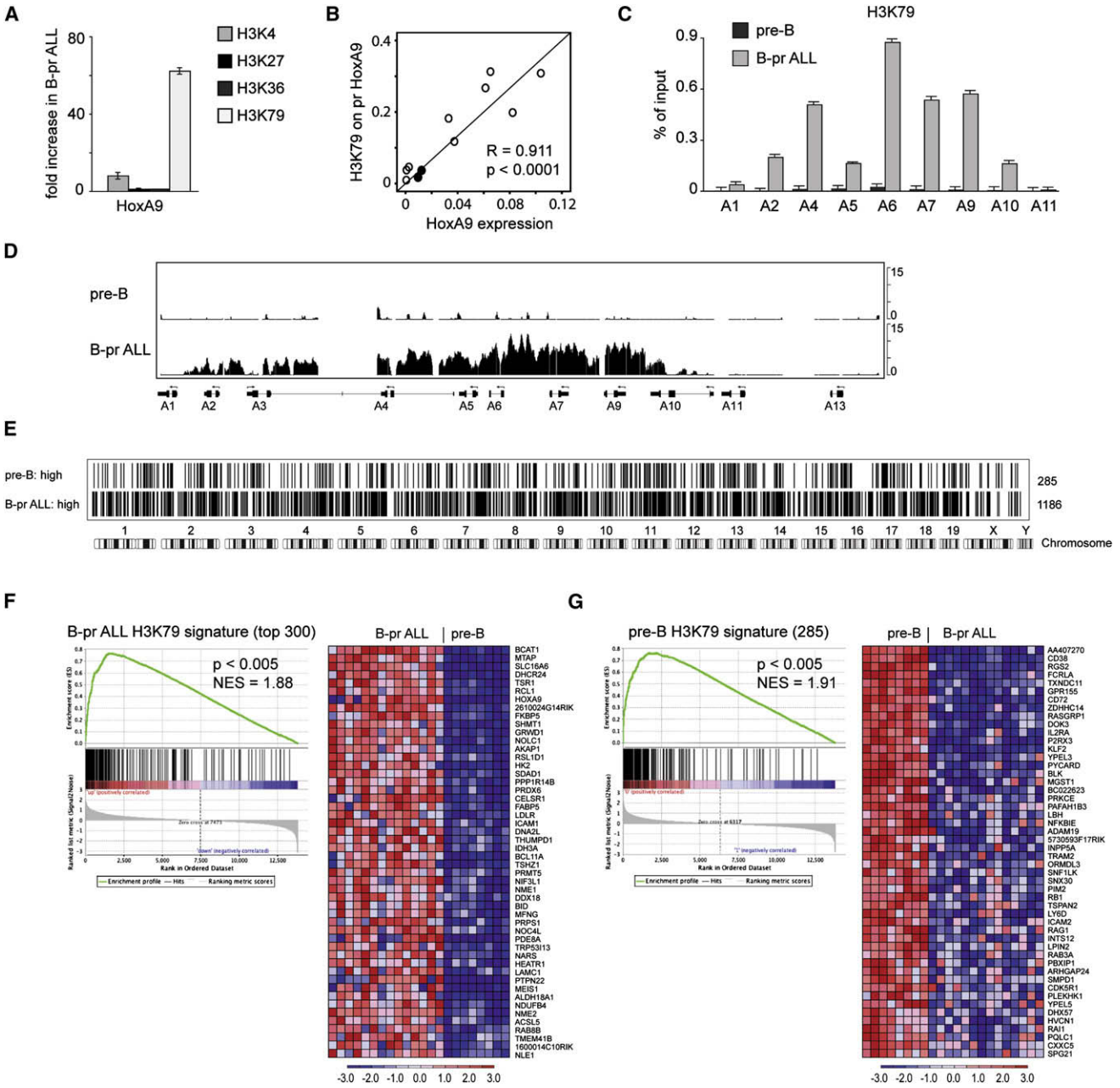


Figure 4. ChIP Analysis of Histone Methylation in Mouse MII-AF4 B Precursor ALL and Normal Pre-B Cells

(A) Fold increase in trimethyl H3K4, trimethyl H3K27, trimethyl H3K36, and dimethyl H3K79 marks associated with the *Hoxa9* promoter in B-pr ALL cells over pre-B cells as assayed by ChIP-qPCR. H3K4 p = 0.05, H3K79 p = 0.005; two-tailed t test. Error bars represent ±SD between four independent experiments.

(B) Correlation of *HoxA9* expression with H3K79me2 content on the *HoxA9* promoter. *HoxA9*/*Gapdh* ratio (HoxA9 expression) was assessed by RT-qPCR, and H3K79me2 content (percent of input) on the *HoxA9* promoter was assessed by ChIP-qPCR in pre-B (●, n = 2) and B-pr ALL (○, n = 9) samples. Pearson correlation R = 0.911; p < 0.0001 by two-tailed t test.

(C) Enrichment of H3K79me2 associated with *HoxA* cluster promoter regions in B-pr ALL and normal pre-B cells expressed as a percentage of input. Error bars represent ±SD of triplicates in one of three independent experiments.

(D) Identically scaled average tracks from ChIP-chip analysis representing *HoxA* cluster loci in normal pre-B (n = 3) and B-pr ALL (n = 3) cells.

(E) Graphical (bar) representation of regions associated with H3K79me2 in normal pre-B and B-pr ALL cells. Averaged model-based analysis of tiling (MAT) scores from genome-wide promoter analysis of three pre-B cell and three B-pr ALL samples were used to define genomic regions possessing significant H3K79 signal ($p < 10^{-5}$) over background. The bars represent 285 genes that were identified to be associated with H3K79me2 in pre-B cells but not in B-pr ALL and 1186 genes associated with H3K79me2 in B-pr ALL but not in pre-B cells.

(F) Correlation of H3K79me2 marks with gene expression in B-pr ALL. GSEA was performed using the top 300 genes by MAT score (out of 1186 genes) in the B-pr ALL H3K79 signature as a gene set. This gene set was assessed for enrichment in the genes more highly expressed in B-pr ALL as compared to normal mouse

immunophenotype and gene expression profile consistent with acute B precursor cell leukemia. The gene expression signature of murine Mll-AF4 ALL cells was found to be highly enriched in the gene expression profile of human ALL samples with *MLL* rearrangements as compared to ALL samples with germline *MLL*. Therefore, this murine model recapitulates human MLL-AF4-mediated ALL in terms of both disease phenotype and overall patterns of gene expression. We note that our mice have a propensity toward AML development, which differs from the strong association between t(4;11) and ALL in humans. Future studies will determine whether this is due to expression of Mll-AF4 in different cells of origin or to microenvironmental influences as recently shown for MLL-AF9 (Wei et al., 2008). Also, the genetic background may influence the leukemia incidence and phenotype. The development of a faithful model of MLL-AF4 ALL will allow detailed characterization of the mechanisms of MLL-AF4-mediated leukemogenesis, including the mechanisms of transformation, the cell (or cells) of origin, the phenotypes of cancer stem cells in leukemias expressing various lineage markers, and the mechanisms associated with the development of mixed myeloid/lymphoid (mixed-lineage) disease. Also, a faithful model of MLL-AF4 ALL will allow for assessment of putative therapeutics in a well-defined model system.

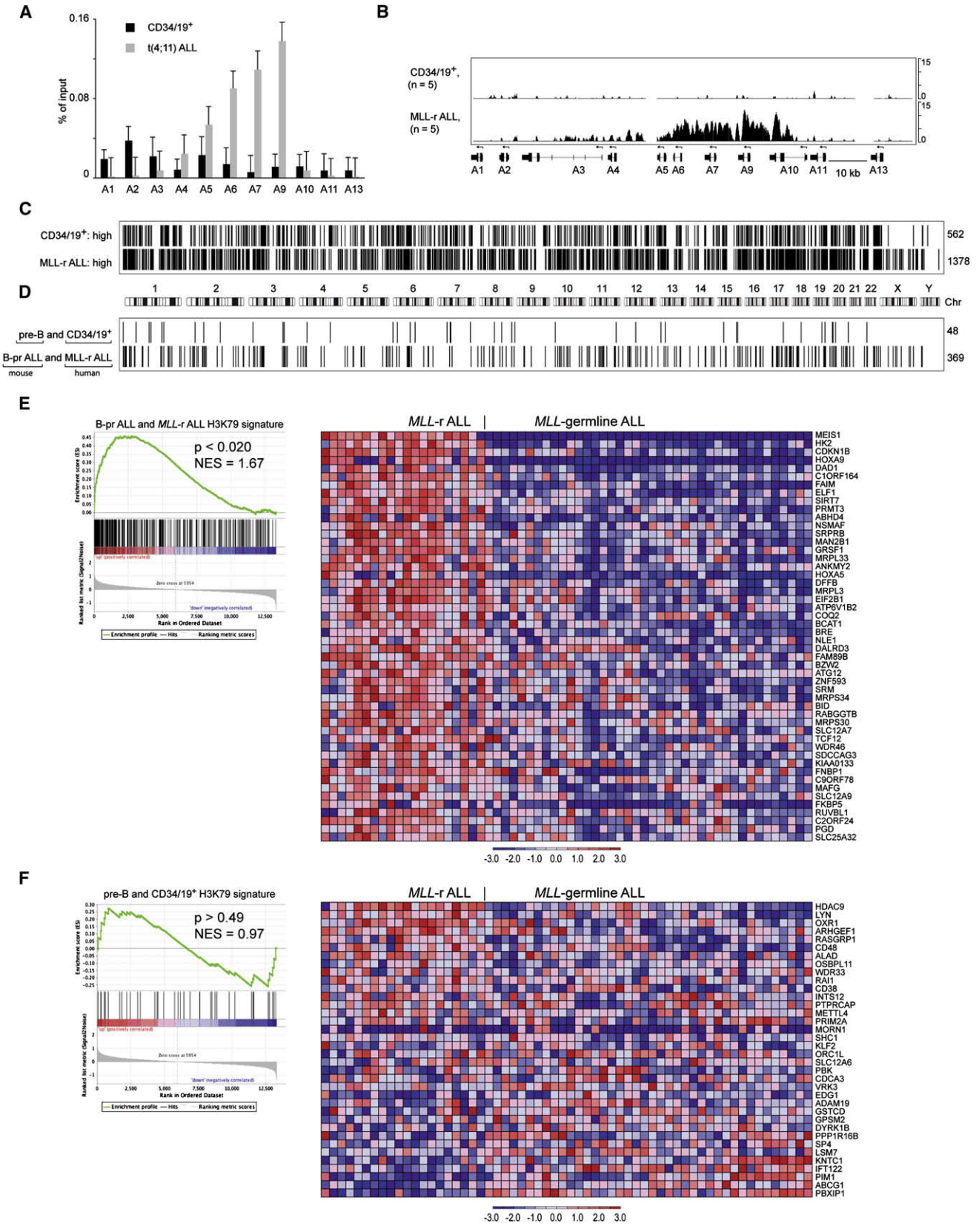
Recent studies have identified association of multiple MLL fusion partners including AF4, AF9, and AF10 with DOT1L, a histone H3K79 methyltransferase (Bitoun et al., 2007; Mueller et al., 2007; Okada et al., 2005; Zeisig et al., 2005; Zhang et al., 2006). We used our murine model to ask whether abnormalities in H3K79me2 were present in Mll-AF4-mediated ALL. Genome-wide assessment revealed enhanced H3K79me2 in comparison to matched normal cells at many loci across the genome, including across the *HOXA* cluster. This prompted a similar analysis in human ALL samples, where we confirmed widespread abnormalities in H3K79me2 as a characteristic distinguishing *MLL*-rearranged from *MLL*-germline ALLs. These results demonstrate that in both murine and human disease mediated by *MLL* rearrangement, there are widespread changes in H3K79me2, including on genes that are critical for leukemogenesis such as *HOXA9* and *MEIS1*. Future studies will determine which subset of these genes are direct MLL-AF4 target genes. The results described here further underscore the utility of this murine model system for providing mechanistic insight into human disease.

In both human and murine ALL, we found that ectopic H3K79me2 was highly associated with increased mRNA expression. Wild-type MLL possesses H3K4 methyltransferase activity, a modification associated with transcriptional "priming" of the promoters of genes important for appropriate developmental cell fate decisions (Guenther et al., 2007). Since the methyltransferase domain of MLL is invariably lost in MLL fusion proteins, including MLL-AF4, the question arises as to how MLL fusions promote enhanced expression of its target genes such as *HOXA* genes. We found that H3K4me3 is still associated with *HOX* genes in our murine MLL-AF4 model, and preliminary experiments suggest this to also be the case in human ALL with MLL fusions (A.V.K., unpublished data). In both human disease and our murine model, the second *MLL* allele remains germline, and thus wild-type MLL protein, or perhaps other H3K4 methyltransferases, may continue to regulate H3K4me3 at these loci. However, H3K4me3 is insufficient to fully activate transcription, which appears to require the presence of other histone modifications such as H3K79me2 for transcriptional elongation to proceed (Guenther et al., 2007). This leads to a model wherein MLL-AF4 recruits DOT1L to MLL target genes and promotes methylation of H3K79 at loci with existing H3K4 methylation (i.e., by wild-type MLL or other H3K4 methyltransferases), thus stimulating transcriptional elongation of genes that are normally primed but not fully transcribed.

Prior studies have established that various cancer subtypes can be distinguished on the basis of specific oncoproteins or global mRNA expression patterns. In this study, we have demonstrated that widespread differences in an epigenetic histone modification can likewise distinguish different cancer subtypes, and malignant versus normal cells. The ability to distinguish cancer subtypes on the basis of epigenetic profiles does not necessarily offer any advantages over mRNA profiling in terms of disease diagnosis. However, the tight linkage between oncoprotein (MLL-AF4), specific epigenetic changes (H3K79me2), and altered gene expression has important therapeutic implications. Specifically, many oncoproteins are DNA-binding proteins and are generally not readily amenable to targeting by small molecules or biologics. Pharmacologic inhibition of transcription factor function, including disrupting protein-protein or protein-DNA interactions, remains a largely experimental undertaking with few examples of success. A common strategy to bypass this impasse has been to look for downstream transcriptional targets that may be more druggable (e.g., enzymes and/or receptors). Since oncoproteins may alter the expression of hundreds or thousands of genes, in most cases inhibiting the activity of a single druggable downstream target may not fully reverse the malignant phenotype. Inhibition of epigenetic-modifying enzymes offers an alternative approach that is facilitated by the fact that enzymes are generally more amenable to drug development. Targeting the epigenetic link between an oncoprotein and downstream gene expression changes may more broadly impact the multitude of gene expression changes that in aggregate contribute to the malignant phenotype. In the case of diseases driven by MLL fusion proteins, development of a DOT1L inhibitor may impact expression of a diversity of critical genes including *HOXA9*, *MEIS1*, *FLT3*, and *BCL2*. Indeed, our data demonstrate that suppression of DOT1L expression leads to decreased expression of *HOXA* cluster genes. Furthermore, a DOT1L inhibitor might be applicable to other diseases characterized by ectopic H3K79me2. In this regard, the ability to identify disease subtypes that have widespread ectopic H3K79me2 would have important implications for patient selection and stratification in testing a DOT1L inhibitor. Inasmuch as epigenetic modifications are

pre-B cells. Left: GSEA enrichment plot. Right: expression of genes corresponding to the top 50 differentially expressed genes from the B-pr ALL H3K79 signature. Normalized enrichment score (NES) = 1.88; $p < 0.005$.

(G) Correlation of H3K79me2 marks with gene expression in pre-B cells. GSEA was performed using 285 genes in the pre-B H3K79 signature as a gene set. Left: GSEA enrichment plot. Right: expression of genes corresponding to the top 50 differentially expressed genes from the B-pr ALL H3K79 signature. NES = 1.91; $p < 0.005$.


A
0.16
0.08
0
% of input
CD34/19+
t(4;11) ALL
A1 A2 A3 A4 A5 A6 A7 A9 A10 A11 A13
B
CD34/19+, (n = 5)
MLL-r ALL, (n = 5)
15
0
15
0
A1 A2 A3 A4 A5 A6 A7 A9 A10 A11 10 kb A13
C
CD34/19+: high
MLL-r ALL: high
562
1378
D
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y Chr
pre-B and CD34/19+
B-pr ALL and MLL-r ALL
mouse
human
48
369
E
B-pr ALL and MLL-r ALL H3K79 signature
p < 0.020
NES = 1.67
Enrichment score (ES)
Ranked list metric (Signal2Noise)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
MLL-r ALL | MLL-germline ALL
MEIS1
HK2
CDKN1B
HOXA9
DAD1
C1ORF164
FAIM
ELF1
SIRT7
PRMT3
ABHD4
NSMAF
SRPRB
MAN2B1
GRSF1
MRPL33
ANKMY2
HOXA5
DFFB
MRPL3
EIF2B1
ATP6V1B2
COQ2
BCAT1
BRE
NLE1
DALRD3
FAM89B
BZW2
ATG12
ZNF593
SRM
MRPS34
BID
RABGGTB
MRPS30
SLC12A7
TCF12
WDR46
SDCCAG3
KIAA0133
FNBP1
C9ORF78
MAFG
SLC12A9
FKBP5
RUVBL1
C2ORF24
PGD
SLC25A32
-3.0 -2.0 -1.0 0.0 1.0 2.0 3.0
F
pre-B and CD34/19+ H3K79 signature
p > 0.49
NES = 0.97
MLL-r ALL | MLL-germline ALL
HDAC9
LYN
OXR1
ARHGEF1
RASGRP1
CD48
ALAD
OSBPL11
WDR33
RAI1
CD38
INTS12
PTPRCAP
METTL4
PRIM2A
MORN1
SHC1
KLF2
ORC1L
SLC12A6
PBK
CDCA3
VRK3
EDG1
ADAM19
GSTCD
GPSM2
DYRK1B
PPP1R16B
SP4
LSM7
KNTC1
IFT122
PIM1
ABCG1
PBXIP1
-3.0 -2.0 -1.0 0.0 1.0 2.0 3.0

likely to be linked to the maintenance of gene expression in many human malignancies, these strategies may be applicable to a wide variety of cancer subtypes.

EXPERIMENTAL PROCEDURES

Human Samples

Normal bone marrow samples as well as diagnostic leukemia samples (peripheral blood or bone marrow) were obtained with informed consent from individuals treated according to protocols approved by the Institutional Review Board at the Dana-Farber Cancer Institute between 2000 and 2007. For details, see Supplemental Experimental Procedures.

Generation of Conditional *Mll-AF4* Knockin Mice

Animals were maintained at the Animal Research Facility at Children's Hospital Boston. Animal experiments were approved by the Institutional Animal Care and Use Committee. The *Mll-AF4*^{stop} targeting vector was assembled in the plasmid vector Litmus 28 (New England Biolabs). We removed *Cbp* cDNA from a previously described (Wang et al., 2005) *Mll-Cbp*^{stop} targeting construct using BamHI digestion and ligated in a portion of AF4 (aa 347–1210; GenBank accession number L13773) using 5′ BamHI and 3′ blunt ends. For further details, see Supplemental Experimental Procedures. *Mx1-Cre* × MA4 mice were generated by crossing *Mll-AF4*^{stop} (MA4) males with females transgenic for *Mx1-Cre*. *Mx1-Cre* × MA4 mice (6–8 weeks old) and control littermates were injected intraperitoneally with 400 μg of polyinosinic/polycytidylic acid (pIpC; Sigma-Aldrich) three times over a 6 day interval.

Bone Marrow Reconstitution Assay and Retroviral Infections

Bone marrow was collected from tibias and femurs of MA4 mice treated with or without 5-FU (Sigma-Aldrich). Red blood cells were lysed using RBC lysis buffer (Puregene), and the remaining cells were washed twice in PBS (Invitrogen) and transduced with "hit-and-run" Cre (Silver and Livingston, 2001) retrovirus (produced as described previously in Krivtsov et al., 2006) overnight in IMDM supplemented with 10% FBS (Invitrogen); 1/100 penicillin/streptomycin (Invitrogen); 10 ng/ml each of IL-3, IL-6, and IL-7; 50 ng/ml each SCF and FLT3 (PeproTech); and 7 μg/ml polybrene (Sigma-Aldrich). Syngeneic recipient mice were lethally irradiated (2 × 700 rads) or sublethally irradiated (600 rads). Retrovirally transduced cells were transplanted retro-orbitally. Detection of Mll-AF4 is described in Supplemental Experimental Procedures.

Flow Cytometry

Labeling of cells for fluorescence-activated cell sorting (FACS) was performed on ice in PBS supplemented with 0.5% horse serum (Invitrogen) using conjugated antibodies purchased from BD PharMingen, eBioscience, or Invitrogen unless otherwise indicated. Details of the antibodies used can be found in Supplemental Experimental Procedures.

Transduction with shRNA Constructs

shRNAs cloned into hairpin pLKO.1 lentiviral vector were synthesized by the Harvard/MIT Broad Institute RNA consortium utilizing a complex selection algorithm to minimize off-target effects (http://www.broad.mit.edu/rnai/trc/). The target sequences for suppression of human *DOT1L* (NM_032482) were DOT1L shRNA1, 5′-CGCCAACACGAGTGTTATATT-3′; DOT1L shRNA2, 5′-CCGCAAGAAGAAGCTAAACAA-3′; and nontargeting GFP control shRNA, 5′-GCAAGCTGACCCTGAAGTTCA-3′. Details of lentiviral transduction can be found in Supplemental Experimental Procedures.

RNA Amplification and Microarray Data Analysis

1–5 × 10^5 cells of different populations were used to isolate RNA. Twenty nanograms of total RNA was amplified and labeled using Ovation RNA Amplification System V2 (NuGEN). Expression data were analyzed with GenePattern release 3.1 and the GSEA 2.0 software package (http://www.broad.mit.edu/tools/software.html). Details of the gene expression analysis can be found in Supplemental Experimental Procedures.

Chromatin Immunoprecipitation, DNA Amplification, Data Analysis, and PCR Analysis

3–10 × 10^5 FACS-sorted cells were fixed in 1% formaldehyde with gentle rotation for 10 min, and the reaction was stopped by addition of glycine (125 mM final concentration). The fixed cells were washed twice with PBS, resuspended in SDS buffer, and further processed as described previously (Bracken et al., 2006) using anti-dimethyl H3K79 (Abcam 3594), anti-trimethyl H3K4 (Upstate 05-745), anti-trimethyl H3K36 (Abcam 9050), or anti-trimethyl H3K27 (Upstate 07-449). Eluted DNA fragments were used directly for qPCR with *HoxA* promoter-specific primers (Bracken et al., 2006) and Sybr GreenER qPCR mix using an ABI 7700 Sequence Detection System. Alternatively, ChIP DNA fragments were amplified using ligation-mediated PCR as described elsewhere (http://chiponchip.org/protocol_itm3.html) (experiment 1 in Figure 4) or using a WGA1 GenomePlex Whole Genome Amplification kit (Sigma) (experiment 2 in Figure S5) and subjected to hybridization with Affymetrix 1R mouse or 1R human promoter arrays.

Identification of ChIP-Chip Regions

M.E.L. directed computational analyses. The model-based analysis of tiling (MAT) array algorithm (Johnson et al., 2006) was used to identify genomic regions ("hits") and highest mean histone methylation scores (MAT scores) associated with those regions on Affymetrix promoter tiling arrays enriched by ChIP. MAT was run with the following parameters to capture regions of increased signal intensity: bandwidth = 300, maximum gap = 300, minimum probes = 10, window p value cutoff = 1 × 10^{-5}. The resulting hits were all false discovery rate (FDR) < 5, so no additional filtering was used. The MAT library and mapping files were based on the March 2006 Human Genome Assembly (HG18) or the February 2006 Mouse Genome Assembly (MM8) as appropriate. Hits flagged by MAT as mapping to repeat regions were excluded from consideration in all cases.

Linking ChIP Hits to RefSeq Genes

The University of California, Santa Cruz RefGene tables (ftp://hgdownload.cse.ucsc.edu/goldenPath) for the human and mouse genomes (HG18 and MM8, respectively), which contain transcription start site (TSS) and end

Figure 5. ChIP-Chip Analysis of Histone Methylation in Human *MLL*-Rearranged ALL and Normal Pre-B Cells

(A) Enrichment of H3K79 dimethylation marks associated with *HOXA* cluster promoters in t(4;11) ALL and human Lin$^-$CD34$^+$CD19$^+$ (CD34/19$^+$) BM cells expressed as percentage of input as assessed by ChIP-qPCR. Error bars represent ±SD of triplicates in one of three independent experiments.

(B) Identically scaled average tracks from ChIP-chip analysis of H3K79me2 modifications associated with *HOXA* cluster loci in human CD34/19$^+$ cells (n = 5) or *MLL*-rearranged ALL (MLL-r ALL) (n = 5).

(C) Graphical (bar) representation of regions associated with H3K79me2 in normal CD34/CD19$^+$ and *MLL*-rearranged ALL cells. 562 genes were identified in CD34/19$^+$ cells but not in *MLL*-rearranged ALL (CD34/19$^+$: high), whereas 1378 genes with significant H3K79 signal were present in *MLL*-rearranged ALL but not CD34/19$^+$ cells (MLL-r ALL: high).

(D) Forty-eight genes had higher H3K79 ChIP-chip signals in both mouse pre-B and human CD34/19$^+$ cells as compared to mouse and human leukemias. 369 genes had higher H3K79 ChIP-chip signals in both mouse Mll-AF4 and human *MLL*-rearranged ALL as compared to normal B cells. $p < 0.0001$ by Fisher's exact test.

(E) GSEA analysis of the 369-gene signature with elevated H3K79me2 found in both human *MLL*-rearranged ALL and mouse B-pr ALL demonstrated enrichment of the gene expression signature in *MLL*-rearranged ALL as compared to *MLL*-germline ALL (NES = 1.63, $p < 0.019$). Left: GSEA enrichment plot. Right: heat map of expression for the top 50 probe sets in *MLL*-rearranged ALL as compared to *MLL*-germline ALL (Ross et al., 2003).

(F) GSEA analysis of the 48 genes with elevated H3K79me2 found in both human CD34/CD19$^+$ BM and mouse pre-B cells found no enrichment of the signature in *MLL*-rearranged ALL as compared to *MLL*-germline ALL (NES = 0.97, $p > 0.49$). Left: GSEA enrichment plot. Right: heat map of expression for 36 probe sets in *MLL*-rearranged ALL as compared to *MLL*-germline ALL (Ross et al., 2003).

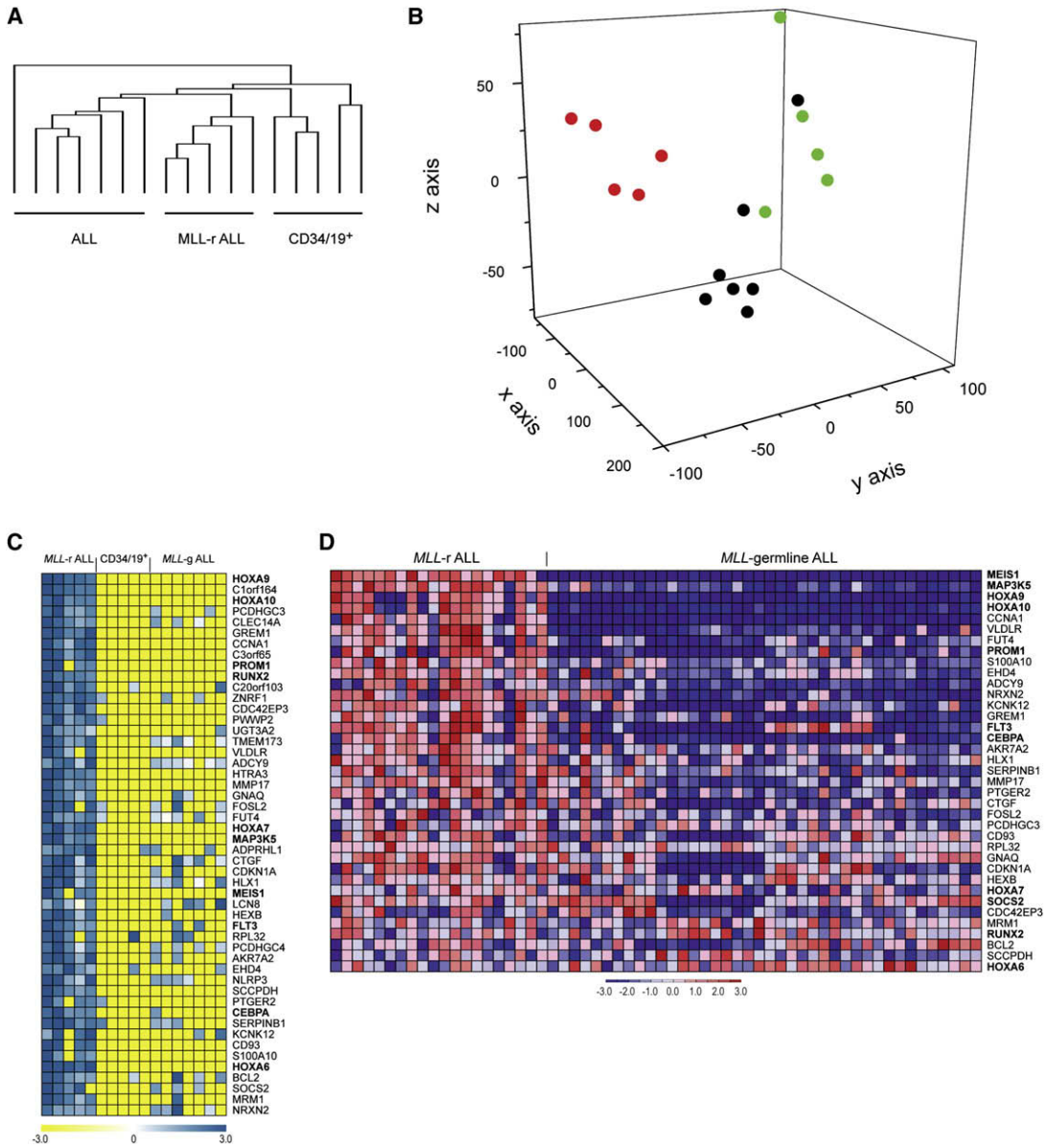


Figure 6. Unsupervised and Supervised Analysis of ChIP-Chip Data for H3K79 Methylation in *MLL*-Rearranged ALL, *MLL*-Germline ALL, and CD34/19$^+$ Cells

(A) Hierarchical clustering of 5438 genes associated with H3K79me2 demonstrates that *MLL*-rearranged ALL samples cluster together and are separate from other *MLL*-germline ALLs and normal B-pr cells.

(B) Principal component analysis (PCA) of the same 5438 genes associated with H3K79me2 separates *MLL*-rearranged ALL (red) from *MLL*-germline ALL (black) and CD34/19$^+$ cells (green).

(C) The top 50 genes associated with H3K79me2 in *MLL*-rearranged ALL compared to the rest of the samples. Genes were ranked based on difference of mean MAT scores between *MLL*-rearranged (*MLL*-r) ALL samples and *MLL*-germline (*MLL*-g) ALL plus CD34/19$^+$ cells.

(D) Expression of genes identified in Figure 6C. Of the 50 genes, 37 had corresponding probe sets on U133A arrays.

position information for RefSeq genes, were linked with the gene_info table (downloaded May 24, 2007) from the National Center for Biotechnology Information (NCBI) (ftp://ftp.ncbi.nih.gov/gene/DATA/) via reference RNA accession number. Chromosomal positions were then used to associate ChIP hits with RefSeq gene IDs. Specifically, hits falling within a window of −0.5 kb to +2.5 kb of a given RefSeq TSS were annotated as being associated with that gene. This window is more restrictive on the 5′ side than the regions tiled on the promoter arrays (7.5 kb–10 kb 5′ to 2.5 kb 3′ of TSS).

Linking Human and Mouse Genes

Homologous human and mouse genes were linked via the NCBI HomoloGene database. Mouse and human gene symbols and IDs were extracted from each homology group that included both human and mouse homologs.

Clustering and Principal Component Analysis of ChIP hits

Individual samples were processed using MAT as above, and the resulting ChIP hits were linked to RefSeq genes as before. A list of represented genes, G, was generated from the nonredundant union of all genes associated with

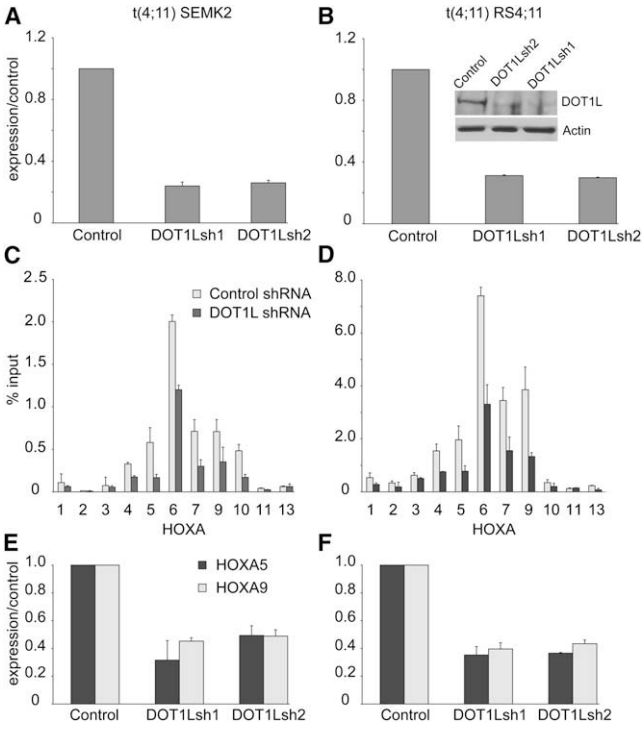


Figure 7. shRNA-Mediated *DOT1L* Suppression in Human MLL-AF4 Cell Lines

(A and B) SEMK2 (A) or RS4;11 (B) ALL cell lines possessing a t(4;11) were transduced with either a control shRNA or one of two different shRNAs that target *DOT1L*, and *DOT1L* RNA expression was assessed by quantitative RT-PCR and western blot for RS4;11 cells 72 hr after transduction (inset). Error bars represent ±SD of duplicates in one of two independent experiments.

(C and D) Assessment of H3K79me2 associated with *HOXA* cluster genes using ChIP 96 hr after treatment of SEMK2 (C) or RS4;11 (D) cells with either a control or *DOT1L*-directed shRNA (DOT1Lsh1). Error bars represent ±SD of duplicates in one of two independent experiments.

(E and F) Assessment of *HOXA5* and *HOXA9* expression in SEMK2 (E) or RS4;11 (F) cells 96 hr after lentiviral transduction with either control or one of two different shRNA constructs. Error bars represent ±SD of duplicates in one of two independent experiments.

a ChIP hit in any sample. A matrix, M, was constructed such that for each gene i in G and for each sample j, $M_{i,j} = 0$ if the j^{th} sample had no ChIP hit associated with gene i or $M_{i,j}$ = MAT score of the ChIP hit in sample j associated with gene i. Hierarchical clustering by sample of the resulting matrix was performed using the clustering tool of GenePattern 2.0. Principal component analysis (PCA) was performed using an R script (http://www.r-project.org/) based on the PCA Methods library from the Bioconductor project (http://bioconductor.org). PCA images were generated using Origin 7.5.

ACCESSION NUMBERS

Microarray data reported herein have been deposited at the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE12363.

SUPPLEMENTAL DATA

The Supplemental Data include Supplemental Experimental Procedures, Supplemental References, and six figures and can be found with this article online at http://www.cancercell.org/supplemental/S1535-6108(08)00328-0.

ACKNOWLEDGMENTS

We would like to thank Y. Fujiwara for help with embryonic stem cell modification and blastocyst injections. This work was supported in part by the National Cancer Institute (grants K08CA92551, 5U01CA105423, and 5P01CA0684841), the Leukemia & Lymphoma Society, and the Damon Runyon Cancer Research Foundation.

Received: February 1, 2008
Revised: August 11, 2008
Accepted: October 2, 2008
Published: November 3, 2008

REFERENCES

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R., and Korsmeyer, S.J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat. Genet. *30*, 41–47.

Armstrong, S.A., Kung, A.L., Mabon, M.E., Silverman, L.B., Stam, R.W., Den Boer, M.L., Pieters, R., Kersey, J.H., Sallan, S.E., Fletcher, J.A., et al. (2003). Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. Cancer Cell *3*, 173–183.

Ayton, P.M., and Cleary, M.L. (2001). Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene *20*, 5695–5707.

Biondi, A., Cimino, G., Pieters, R., and Pui, C.H. (2000). Biological and therapeutic aspects of infant leukemia. Blood *96*, 24–33.

Bitoun, E., Oliver, P.L., and Davies, K.E. (2007). The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum. Mol. Genet. *16*, 92–106.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. *20*, 1123–1136.

Chen, C.S., Sorensen, P.H., Domer, P.H., Reaman, G.H., Korsmeyer, S.J., Heerema, N.A., Hammond, G.D., and Kersey, J.H. (1993). Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome. Blood *81*, 2386–2393.

Chen, W., Li, Q., Hudson, W.A., Kumar, A., Kirchhof, N., and Kersey, J.H. (2006). A murine Mll-AF4 knock-in model results in lymphoid and myeloid deregulation and hematologic malignancy. Blood *108*, 669–677.

Cheung, N., Chan, L.C., Thompson, A., Cleary, M.L., and So, C.W. (2007). Protein arginine-methyltransferase-dependent oncogenesis. Nat. Cell Biol. *9*, 1208–1215.

Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E., and Young, R.A. (2005). Global and Hox-specific roles for the MLL1 methyltransferase. Proc. Natl. Acad. Sci. USA *102*, 8603–8608.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Hardy, R.R. (1990). Development of murine B cell subpopulations. Semin. Immunol. *2*, 197–206.

Hess, J.L., Yu, B.D., Li, B., Hanson, R., and Korsmeyer, S.J. (1997). Defects in yolk sac hematopoiesis in Mll-null embryos. Blood *90*, 1799–1806.

Huret, J.L., Dessen, P., and Bernheim, A. (2001). An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners. Leukemia *15*, 987–989.

Johnson, W.E., Li, W., Meyer, C.A., Gottardo, R., Carroll, J.S., Brown, M., and Liu, X.S. (2006). Model-based analysis of tiling-arrays for ChIP-chip. Proc. Natl. Acad. Sci. USA *103*, 12457–12462.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell *1*, 324–337.

Krivtsov, A.V., and Armstrong, S.A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. Nat. Rev. Cancer *7*, 823–833.

Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature *442*, 818–822.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427–1429.

Metzler, M., Forster, A., Pannell, R., Arends, M.J., Daser, A., Lobato, M.N., and Rabbitts, T.H. (2006). A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology. Oncogene *25*, 3093–3103.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol. Cell *10*, 1107–1117.

Mueller, D., Bach, C., Zeisig, D., Garcia-Cuellar, M.P., Monroe, S., Sreekumar, A., Zhou, R., Nesvizhskii, A., Chinnaiyan, A., Hess, J.L., et al. (2007). A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood *110*, 4445–4454.

Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C.M., and Canaani, E. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol. Cell *10*, 1119–1128.

Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V.M., Su, L., Xu, G., and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. Cell *121*, 167–178.

Ross, M.E., Zhou, X., Song, G., Shurtleff, S.A., Girtman, K., Williams, W.K., Liu, H.C., Mahfouz, R., Raimondi, S.C., Lenny, N., et al. (2003). Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood *102*, 2951–2959.

Rozovskaia, T., Feinstein, E., Mor, O., Foa, R., Blechman, J., Nakamura, T., Croce, C.M., Cimino, G., and Canaani, E. (2001). Upregulation of Meis1 and HoxA9 in acute lymphocytic leukemias with the t(4: 11) abnormality. Oncogene *20*, 874–878.

Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev. *18*, 1263–1271.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu. Rev. Biochem. *75*, 243–269.

Silver, D.P., and Livingston, D.M. (2001). Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. Mol. Cell *8*, 233–243.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Swiatek, P.J., and Gridley, T. (1993). Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene Krox20. Genes Dev. *7*, 2071–2084.

Wang, J., Iwasaki, H., Krivtsov, A., Febbo, P.G., Thorner, A.R., Ernst, P., Anastasiadou, E., Kutok, J.L., Kogan, S.C., Zinkel, S.S., et al. (2005). Conditional MLL-CBP targets GMP and models therapy-related myeloproliferative disease. EMBO J. *24*, 368–381.

Wei, J., Wunderlich, M., Fox, C., Alvarez, S., Cigudosa, J.C., Wilhelm, J.S., Zheng, Y., Cancelas, J.A., Gu, Y., Jansen, M., et al. (2008). Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. Cancer Cell *13*, 483–495.

Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., et al. (2002). Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell *1*, 133–143.

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. Nature *378*, 505–508.

Zeisig, B.B., Milne, T., Garcia-Cuellar, M.P., Schreiner, S., Martin, M.E., Fuchs, U., Borkhardt, A., Chanda, S.K., Walker, J., Soden, R., et al. (2004). Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. Mol. Cell. Biol. *24*, 617–628.

Zeisig, D.T., Bittner, C.B., Zeisig, B.B., Garcia-Cuellar, M.P., Hess, J.L., and Slany, R.K. (2005). The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. Oncogene *24*, 5525–5532.

Zhang, W., Xia, X., Reisenauer, M.R., Hemenway, C.S., and Kone, B.C. (2006). Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCalpha in an aldosterone-sensitive manner. J. Biol. Chem. *281*, 18059–18068.