Cell Stem Cell
Article

Hierarchical Maintenance of *MLL* Myeloid Leukemia Stem Cells Employs a Transcriptional Program Shared with Embryonic Rather Than Adult Stem Cells

Tim C.P. Somervaille,[1,4] Christina J. Matheny,[1] Gary J. Spencer,[4] Masayuki Iwasaki,[1] John L. Rinn,[2] Daniela M. Witten,[3] Howard Y. Chang,[2] Sheila A. Shurtleff,[5] James R. Downing,[5] and Michael L. Cleary[1,*]
[1]Department of Pathology
[2]Program in Epithelial Biology
[3]Department of Statistics
Stanford University, Stanford, CA 94305, USA
[4]Cancer Research UK Leukaemia Biology Group, Paterson Institute for Cancer Research, University of Manchester, Manchester M20 4BX, UK
[5]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA
*Correspondence: mcleary@stanford.edu
DOI 10.1016/j.stem.2008.11.015

SUMMARY

The genetic programs that promote retention of self-renewing leukemia stem cells (LSCs) at the apex of cellular hierarchies in acute myeloid leukemia (AML) are not known. In a mouse model of human AML, LSCs exhibit variable frequencies that correlate with the initiating *MLL* oncogene and are maintained in a self-renewing state by a transcriptional subprogram more akin to that of embryonic stem cells (ESCs) than to that of adult stem cells. The transcription/chromatin regulatory factors Myb, Hmgb3, and Cbx5 are critical components of the program and suffice for *Hoxa*/*Meis*-independent immortalization of myeloid progenitors when coexpressed, establishing the cooperative and essential role of an ESC-like LSC maintenance program ancillary to the leukemia-initiating *MLL*/*Hox*/*Meis* program. Enriched expression of LSC maintenance and ESC-like program genes in normal myeloid progenitors and poor-prognosis human malignancies links the frequency of aberrantly self-renewing progenitor-like cancer stem cells (CSCs) to prognosis in human cancer.

INTRODUCTION

The cancer stem cell (CSC) model posits that many human malignancies consist of two functionally distinct cell types: (1) CSCs, which are self-renewing cells with the capacity to initiate, sustain, and expand the disease; and (2) non-self-renewing progeny cells, derived from CSCs through differentiation, which likely make up the bulk of the tumor and account for disease symptomatology (Jordan et al., 2006; Clarke et al., 2006). For malignancies to be cured, it may be necessary and sufficient to exclusively eliminate CSCs. Consequently, there is considerable interest in further understanding the biological and molecular properties of these cells, by comparison with both their non-self-renewing downstream progeny and their normal adult stem cell counterparts.

CSCs (also called tumor-propagating or tumor-initiating cells) were first formally described in human acute myeloid leukemia (AML) as rare cells that share an immunophenotype with normal hematopoietic stem cells (HSCs) (Bonnet and Dick, 1997). This paradigm has recently been substantially revised based on two significant observations. First, in murine models of human leukemia induced by the *MLL-AF9* oncogene, self-renewing leukemia stem cells (LSCs) may account for up to a quarter of cells within the leukemia clone and exhibit mature myeloid immunophenotypes (Somervaille and Cleary, 2006). A high frequency of LSCs was also observed in *Sfpi*$^{-/-}$ murine AMLs (Kelly et al., 2007). Second, protocols that enhance engraftment of human leukemia cells in xenogeneic transplant assays demonstrate the presence of LSCs in leukemia cell subpopulations previously considered to be devoid of them (Taussig et al., 2008).

Since LSCs may be more numerous and mature than originally proposed, the nature and generality of the hierarchical organization of malignancies has recently been questioned (Kelly et al., 2007). However, consistent with the CSC model, only a subset of AML cells have clonogenic potential in in vitro assays (Buick et al., 1977; Somervaille and Cleary, 2006), and human AML blast cells undergo differentiation in vivo to mature granulocytes (Fearon et al., 1986), as may murine LSCs initiated by *MLL-AF9* (Somervaille and Cleary, 2006). To further elucidate the hierarchical disposition of AML, a major goal is to identify transcriptional programs, genes, and pathways that specifically correlate with and promote the retention of LSCs within the self-renewing compartment of leukemias (Clarke et al., 2006). It is not known whether such LSC maintenance programs are synonymous with programs responsible for leukemia initiation—for example, *Hoxa*/*Meis* in *MLL* leukemogenesis (Ayton and Cleary, 2003; Krivtsov et al., 2006; Wong et al., 2007). It is also not clear whether they share features with transcriptional programs expressed in adult or embryonic stem cells (ESCs) (Ben-Porath et al., 2008; Wong et al., 2008) or whether there is a relationship with genes and pathways implicated in the function of AML stem cells such as NFκB, phosphatidylinositide-3-kinase, *CTNNB1*, *Pten*

(reviewed in Jordan et al., 2006), *Bmi1* (Lessard and Sauvageau, 2003), and *Junb* (Steidl et al., 2006).

To address these issues, we investigated the genetic determinants that maintain LSC frequencies and leukemia cell hierarchies using a mouse model that faithfully recapitulates many of the pathologic features of AML induced by chromosomal translocations of the *MLL* gene (Lavau et al., 1997; Somervaille and Cleary, 2006), which occur in about 5%–10% of human AMLs (Look, 1997). Confirming recent speculation that CSC frequency may differ between distinct tumor types (Kelly et al., 2007; Adams et al., 2007; Kennedy et al., 2007), LSC frequency in AML was found to vary substantially according to the initiating *MLL* oncogene. This feature, and the observation that LSC frequency varies within the leukemia cell hierarchy, was used to derive a transcriptional program for LSC hierarchical maintenance. The program indicates that *MLL* LSCs are maintained in a self-renewing state by co-option of a transcriptional program that shares features with ESCs and is transiently expressed in normal myeloid precursors rather than HSCs or mature neutrophils. Furthermore, the shared transcriptional features of LSCs, ESCs, normal midmyeloid lineage cells, and a diverse set of poor-prognosis human malignancies support the broader conclusion that CSCs may be aberrantly self-renewing downstream progenitor cells whose frequency in human malignant disease correlates with and dictates prognosis.

RESULTS

Distinct Molecular Subtypes of *MLL* Leukemia Are Associated with Different LSC Frequencies and Leukemia Cell Hierarchies

LSC frequencies and leukemia cell hierarchies were characterized in cohorts of mice where AML had been initiated using *MLL* oncogenes (*MLL-AF9*, *MLL-ENL*, *MLL-AF10*, *MLL-AF1p*, and *MLL-GAS7*) that function through distinct molecular mechanisms (Somervaille and Cleary, 2006; Lavau et al., 1997; DiMartino et al., 2002; So et al., 2003a, 2003b) (see Table S1 available online). Relative LSC frequencies were assessed by determining AML colony-forming cell (CFC) frequencies in the spleen and bone marrow (BM) of leukemic mice. AML CFCs are a surrogate measure for LSCs in both *MLL-AF9* leukemia (Somervaille and Cleary, 2006) and other *MLL* leukemia molecular subtypes (Figure S1), as shown by secondary transplantation of cells derived from singly isolated colonies. In leukemic mice in which AML had been initiated by *MLL-ENL* or *MLL-AF9*, the frequencies of AML CFCs/LSCs in both spleen and BM were significantly higher than in mice in which leukemia had been initiated by *MLL-AF10*, *MLL-AF1p*, or *MLL-GAS7* (Figure 1A and data not shown). Concordant with these data, approximately seven times as many *MLL-AF1p* leukemia cells were required to initiate secondary AML by comparison with *MLL-ENL* leukemia cells in limit dilution analyses (Figure 1B). FACS analyses showed that greater than 99.5% of BM cells in all leukemias expressed one or more myeloid-specific antigens, whereas less than 0.5% of cells were Lin^- (Table S2), demonstrating that LSCs in the different molecular subtypes of AML are predominantly myeloid Lin^+ cells.

There was also a marked disparity in the hierarchical organization of the respective leukemias. Those initiated by *MLL-AF10* or *MLL-AF1p* displayed extensive mature myeloid lineage differentiation (Figures 1C and 1D) that was part of the leukemia clone (Figure S2), and were classified as either acute myeloproliferative disease-like myeloid leukemia or acute myelomonocytic leukemia according to the Bethesda criteria (Kogan et al., 2002). Conversely, leukemias initiated by *MLL-AF9* or *MLL-ENL* showed less differentiation, and most (16 of 18) were classified as acute monocytic leukemia (Table S3). Consistent with this, their explanted colonies in CFC assays were significantly more "blast-like" (type 1, Lavau et al. [1997]), containing a higher proportion of cells displaying immature cytological and functional features (Figure 1E and data not shown). *MLL-GAS7* leukemias exhibited the lowest AML CFC frequencies but morphologic appearances more akin to *MLL-ENL* and *MLL-AF9* leukemias (Figures 1A and 1D), indicating that functional and morphologic features of AML cellular hierarchies are not necessarily correlated. The spectrum of disease manifestations was not due to variations in the frequency of distinct clones or the number of retroviral integrations per clone (Figure S3). Therefore, leukemia cell hierarchy and the frequency of self-renewing Lin^+ LSCs vary significantly according to the initiating *MLL* oncogene.

Derivation of a Genetic Program that Maintains *MLL* LSCs within the Self-Renewing Fraction of the Leukemia Cell Hierarchy

A genetic program that distinguishes LSCs from their differentiated progeny within leukemia cell hierarchies and promotes maintenance of self-renewal was derived based on the hypothesis that it would be (1) differentially expressed between leukemias containing LSCs at high versus low frequencies and (2) concordantly regulated in LSC-enriched versus LSC-depleted cell populations within a given leukemia.

MLL leukemia molecular subtypes were therefore grouped into two classes according to their relative LSC frequencies: high LSC frequency leukemias (*MLL-AF9* and *MLL-ENL*) and low LSC frequency leukemias (*MLL-AF10*, *MLL-AF1p*, and *MLL-GAS7*). Comparison of global transcriptional profiles in 34 distinct BM AML samples representative of these two groups (12 and 22, respectively) revealed 2943 differentially expressed probe sets (at significance level $p \leq 0.01$ [unpaired t test]; median false discovery rate [FDR] 12.3%; Figure 2A; Table S4).

Leukemia cells from mice with *MLL-AF10* AML were then fractionated into separate subpopulations on the basis of c-kit expression, which correlates with *MLL* LSC frequency (Somervaille and Cleary, 2006). The sorted AML subpopulations exhibited substantial differences in their frequencies of AML CFCs/LSCs (mean 14-fold; Figure 2B) and morphologic features (Figure 2C), consistent with a leukemia cell hierarchy with maturation through to terminally differentiated neutrophils. Comparison of global transcriptional profiles of the LSC-enriched (c-kit^{hi}) versus LSC-depleted (c-kit^{lo}) populations from within individual leukemias revealed 5173 differentially expressed probe sets (at significance level $p \leq 0.01$ [unpaired t test]; Figure 2D and Table S5).

The respective data sets from the above analyses were intersected to identify core gene sets that positively or negatively correlate with LSC potential (Figures 2E and 2F). Expression of 204 probe sets was found to be positively correlated with high LSC frequency across the different molecular subtypes of *MLL* leukemia and concordantly downregulated between LSC-enriched versus LSC-depleted populations (Figure 2E and Table

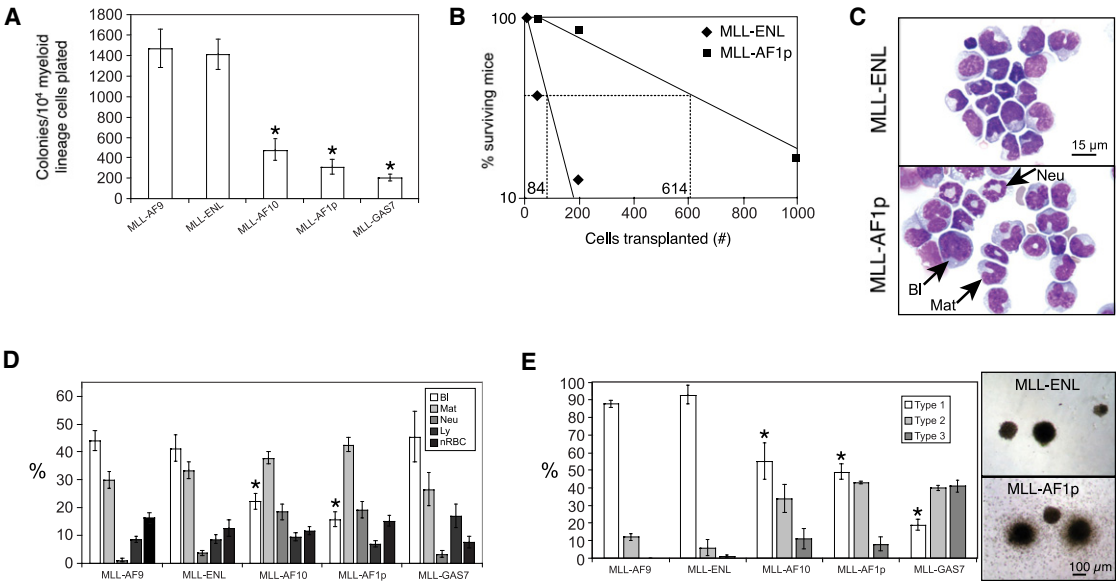

Figure 1. Variable LSC Frequencies and Cellular Hierarchies in Murine AMLs Induced by Different *MLL* Oncogenes

(A) Bar graph shows frequency of CFCs in the spleens of leukemic mice (mean ± SEM; n = 7 leukemias for each fusion, except *MLL-GAS7*, where n = 11; *p ≤ 0.001 by comparison with *MLL-AF9*).

(B) Limit dilution analyses show the estimated number of secondarily transplanted leukemic *MLL-ENL* or *MLL-AF1p* splenocytes required to initiate AML in sublethally irradiated recipient mice (n = 8 recipients for each cell dose). Mice were followed for 180 days, with the longest disease latencies being 66 days for *MLL-ENL* and 67 days for *MLL-AF1p*.

(C) In representative BM cytospin images, *MLL-AF1p*-initiated AML cells are more differentiated by comparison with *MLL-ENL*-initiated AML cells. Scale bar applies to both images.

(D) Bar graph shows percentage cell type in spleen as determined by morphologic analysis of cytospin preparations (mean ± SEM; n = 8–13 for each cohort; *p ≤ 0.05 for blast cell percentage versus *MLL-AF9*). Bl, blast cell; Mat, maturing myelomonocytic cell; Neu, neutrophil; Ly, lymphocyte; nRBC, nucleated erythroid precursor cell.

(E) Bar graph shows proportion of myeloid leukemia colonies of each morphologic type (defined in Lavau et al., 1997) (mean ± SEM; n = 3–5 separate leukemias for each fusion; *p < 0.05 by comparison with *MLL-AF9* for type 1 colony frequency). Representative images are shown on the right. Scale bar applies to both images.

S6). Conversely, the expression of 466 probe sets was negatively correlated with LSC frequency across the different molecular subtypes of *MLL* leukemia and concordantly upregulated from LSC-enriched to LSC-depleted populations (Figure 2F and Table S7; $p < 0.001$ for both intersections using Pearson's chi-square test). The 185 genes (204 probe sets) whose expression positively correlated with LSC frequency, along with the 375 genes (466 probe sets) whose downregulation correlated with LSC frequency, are hereafter collectively termed the LSC maintenance program.

LSCs in Murine *MLL* Myeloid Leukemia Are Metabolically Active, Proliferating Cells

In contrast to the assumption that AML LSCs are predominantly quiescent (Jordan et al., 2006), gene ontology (GO) annotations of upregulated LSC maintenance program genes were consistent with LSCs being metabolically active, cycling cells (Figure 3A). Furthermore, gene set enrichment analysis (GSEA; Subramanian et al., 2005) revealed significant enrichments of expression of curated gene sets for cell cycle, metabolism, and Myc targets in high LSC frequency cell populations by comparison with their respective low LSC frequency populations (Table S8). Cell-cycle analyses performed on AML subpopulations sorted for different levels of c-kit expression (Figure 3B), which correlates with LSC frequency (Figure 2B and Somervaille and Cleary, 2006), showed that fractions containing the highest frequency of LSCs also contained the highest proportion of cycling cells (Figure 3B). Thus, LSCs in this model of AML are predominantly metabolically active, cycling cells.

The LSC Maintenance Program Represents a Committed Myeloid, as Opposed to an HSC, Transcriptional Program

Upregulated LSC maintenance program genes displayed highest relative expression during normal hematopoiesis in midmyeloid lineage cells (granulocyte-macrophage progenitors and myeloblasts) versus lower levels in KSL (c-kit$^+$ Sca1$^+$ Lin$^-$) cells and terminally differentiated neutrophils (Figure 3C and Figure S4). Conversely, an inverse pattern of expression was observed for genes negatively correlated with LSC potential (Figure 3D). Thus, a critical effect of *MLL* oncogenes is to convert a normal midmyeloid lineage program into an AML LSC maintenance program. Consistent with these observations, GSEA revealed that leukemia cell populations containing high frequencies of LSCs exhibited enriched expression of genes expressed in normal lineage-committed hematopoietic progenitors (Ivanova et al., 2002), but not genes expressed in long-term HSCs (Table S8).

Notably, *Hoxa* and *Meis1*, which are prominently expressed in KSL cells, did not feature in the LSC maintenance signature, nor did KSL transcripts previously shown to be shared with *MLL-AF9*

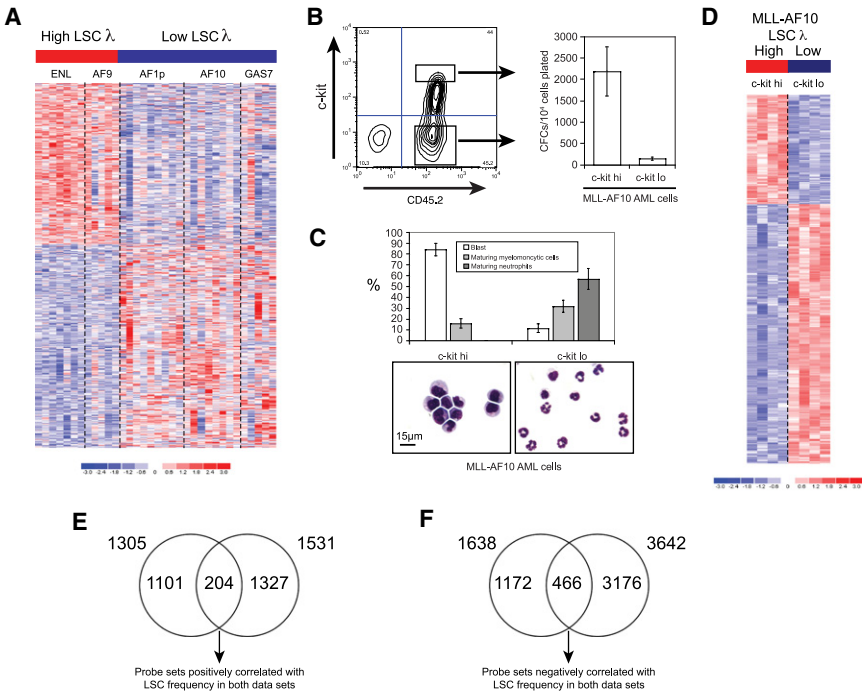

Figure 2. Derivation of an *MLL* LSC Maintenance Signature

(A) Image represents the 2943 probe sets differentially expressed between 12 high LSC frequency AMLs (*MLL-ENL* and *MLL-AF9*) and 22 low LSC frequency AMLs (*MLL-AF1p*, *MLL-AF10* and *MLL-GAS7*). Color scale indicates normalized expression values.

(B) *MLL-AF10* leukemic splenocytes were FACS sorted for high- or low-level c-kit expression, as indicated in the representative plot. Bar graph shows clonogenic potentials of the sorted subpopulations in semisolid culture (mean ± SEM; n = 4). AML was initiated by transplantation of immortalized CD45.2$^+$ c-kit$^+$ BM stem and progenitor cells.

(C) Bar graph and representative images show morphology of sorted leukemia cell populations (mean ± SEM; n = 4). Scale bar applies to both images.

(D) Image represents the 5173 probe sets differentially expressed between high LSC frequency (c-kit^{hi}) and low LSC frequency (c-kit^{lo}) *MLL-AF10* leukemia cell subpopulations. Color scale indicates normalized expression values.

(E and F) Venn diagrams show the intersection of probe sets positively (E) and negatively (F) correlated with LSC frequency from the analyses shown in (A) and (D).

LSCs (Krivtsov et al., 2006; Figure S5). Although required for leukemic transformation of BM stem and progenitor cells by *MLL* oncogenes (Ayton and Cleary, 2003; So et al., 2004; Wong et al., 2007), *Hoxa*/*Meis1* and *Mef2c* levels did not vary concordantly with LSC frequency (Figure 3E). Indeed, spontaneous terminal leukemia cell differentiation and loss of self-renewal potential in *MLL-AF10* leukemias (Figures 2B and 2C) occurred in the presence of substantially elevated *Hoxa* and *Meis1* expression, which did not vary significantly between LSC-enriched and LSC-depleted subpopulations (Figure 3E). Thus, *MLL* LSCs are metabolically active, aberrantly self-renewing, committed myeloid cells, rather than HSCs aberrantly expressing myeloid lineage antigens, and undergo loss of self-renewal potential in vivo in the presence of relatively high-level expression of *Hoxa* and *Meis1* transcripts.

MLL LSCs Share Transcriptional Similarities with ESCs and Poor-Prognosis Human Malignancies, Including Pediatric Leukemia

Poor prognosis in a diverse set of human malignancies is associated with the expression of an ESC-like program (Wong et al., 2008; Ben-Porath et al., 2008). GSEA demonstrated that expression of ESC-like program genes was substantially enriched in the high versus low LSC frequency *MLL* leukemia cell populations, in contrast to the reverse enrichment of an adult stem cell core program (Wong et al., 2008) in LSC-depleted populations (Figure 4A and Table 1). Enrichment for an ESC-like program was not driven by proliferation genes, since the significance of the correlation was maintained when genes with cell-cycle-functional annotations were excluded from the analysis (Table 1). The ESC-like state in *MLL* LSCs was not induced by genes such as *Pou5f1*, *Nanog*, or *Sox2*, which were not significantly expressed (data not shown), nor coordinate expression of their downstream transcriptional effecters (Boyer et al., 2005), which were not enriched in *MLL* LSCs (Table 1).

Expression of ESC-like program genes was also transiently induced during normal myeloid differentiation, whereas expression of adult stem cell core genes was suppressed upon exit from the KLS compartment (Figure 4B). This showed that the normal midmyeloid program co-opted by MLL oncoproteins to maintain LSCs in a self-renewing state exhibits shared features with ESCs. Furthermore, expression of gene sets associated with poor prognosis in a diverse set of human cancers, including pediatric myeloid leukemia (Yagi et al., 2003), was enriched in high versus low LSC frequency cell populations (Figure 4C and Table 1). Taken together, these observations suggest a link between prognosis in human cancer and the frequency of aberrantly self-renewing progenitor cells downstream of the tissue stem cell (Figure 4D).

Genes that regulate transcription and chromatin are candidate critical drivers of the shared transcriptional and biological properties of both self-renewing LSCs and ESCs. In a leading-edge analysis of the expression of ESC-like program genes with transcription or chromatin annotations in high versus low LSC frequency populations (Figure 4E), there was prominent placement of genes found within the LSC maintenance signature (e.g., *Hmgb3*, *Cbx5*, *Mtf2*, *Orc2l*) or of genes with close homologs within the signature (e.g., *Mybl2*, *Ilf3*). Thus, regulators of transcription or chromatin, or their close homologs, that are expressed in ESCs are also expressed in *MLL* LSCs.

Validation of the *MLL* Leukemia Maintenance Program Indicates Critical Collaborative Roles for *Myb*, *Cbx5*, and *Hmgb3*

To validate the role of the maintenance program in LSC self-renewal, three transcription or chromatin regulators were

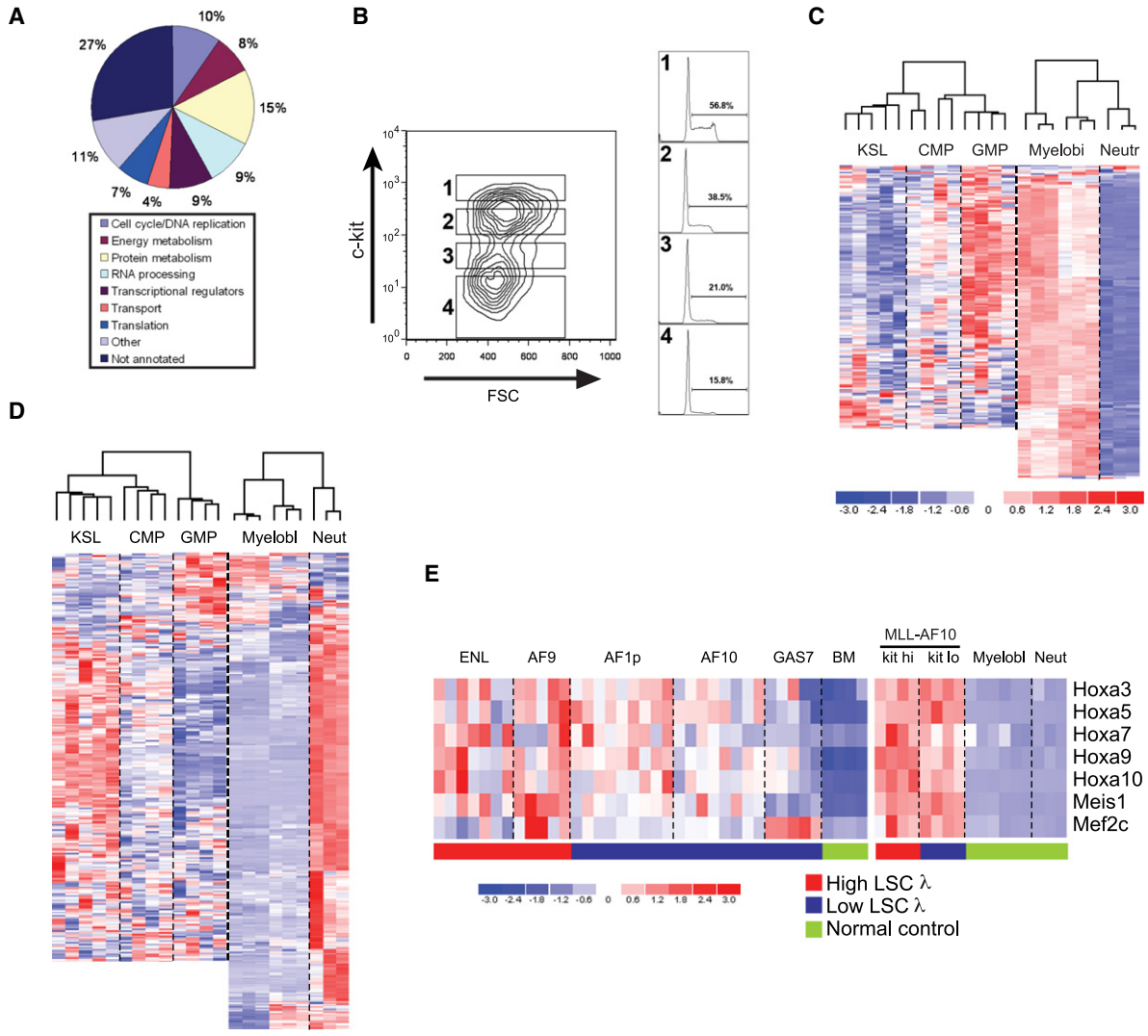

Figure 3. Biological Features of *MLL* LSCs

(A) GO analysis of upregulated LSC maintenance signature genes based on their biological process annotations.

(B) Mac1$^+$ splenocytes from a mouse with *MLL-AF9* AML were FACS sorted using the gates indicated in the representative plot. Sorted cells within each fraction were then analyzed for their cell-cycle status (right hand plots).

(C and D) Unsupervised cluster analysis of the expression patterns of LSC maintenance signature genes positively (C) or negatively (D) correlated with LSC function, in defined populations of normal hematopoietic cells. Values from KSL, CMP, and GMP cells (Krivtsov et al., 2006) and normal myeloblasts and neutrophils (this study) were utilized. Not all of the genes in the signature are represented on the Affymetrix 430A 2.0 chip used by Krivtsov et al. Color scale applies to both images and represents normalized expression values.

(E) Heat map shows expression of *Hoxa*, *Meis1*, and *Mef2c* genes in microarray analyses of *MLL* leukemias with high or low frequencies of LSCs (left panel), and in sorted subpopulations of *MLL-AF10* leukemias with high (c-kit^{hi}) or low (c-kit^{lo}) frequencies of LSCs (right panel). Expression values for whole BM (left panel) and prospectively isolated normal myeloblasts or neutrophils (right panel) are shown for comparison (normal control).

selected for further study: *Myb*, a transcription factor required for in vitro transformation of murine BM progenitor cells by *MLL-ENL* (Hess et al., 2006); *Cbx5* (HP1α), an epigenetic repressor that binds the histone H3-K9 methylation mark to promote heterochromatin formation (Bannister et al., 2001); and *Hmgb3*, a high-mobility group box protein that regulates self-renewal in normal HSCs (Nemeth et al., 2006). Consistent with their potential roles in maintaining LSC self-renewal, the expression of all three genes was substantially reduced concomitant with loss of clonogenic activity and induction of macrophage differentiation in transformed hematopoietic progenitors following inactivation of an *MLL-ENL-ER*-conditional oncoprotein (Ayton and Cleary, 2003) by withdrawal of 4-hydroxytamoxifen (Figure 5A).

When assessed across 23 distinct AMLs, *Myb* transcript levels correlated significantly with LSC frequencies (Figure 5B). Furthermore, *Myb* knockdown reduced AML CFC/LSC frequencies (Figure 5C) in direct proportion to the extent of knockdown (Figure 5D) without inducing apoptosis (Figure S6), and surviving colonies predominantly featured type 2 or type 3 morphology, consistent with *Myb* reduction promoting terminal leukemia

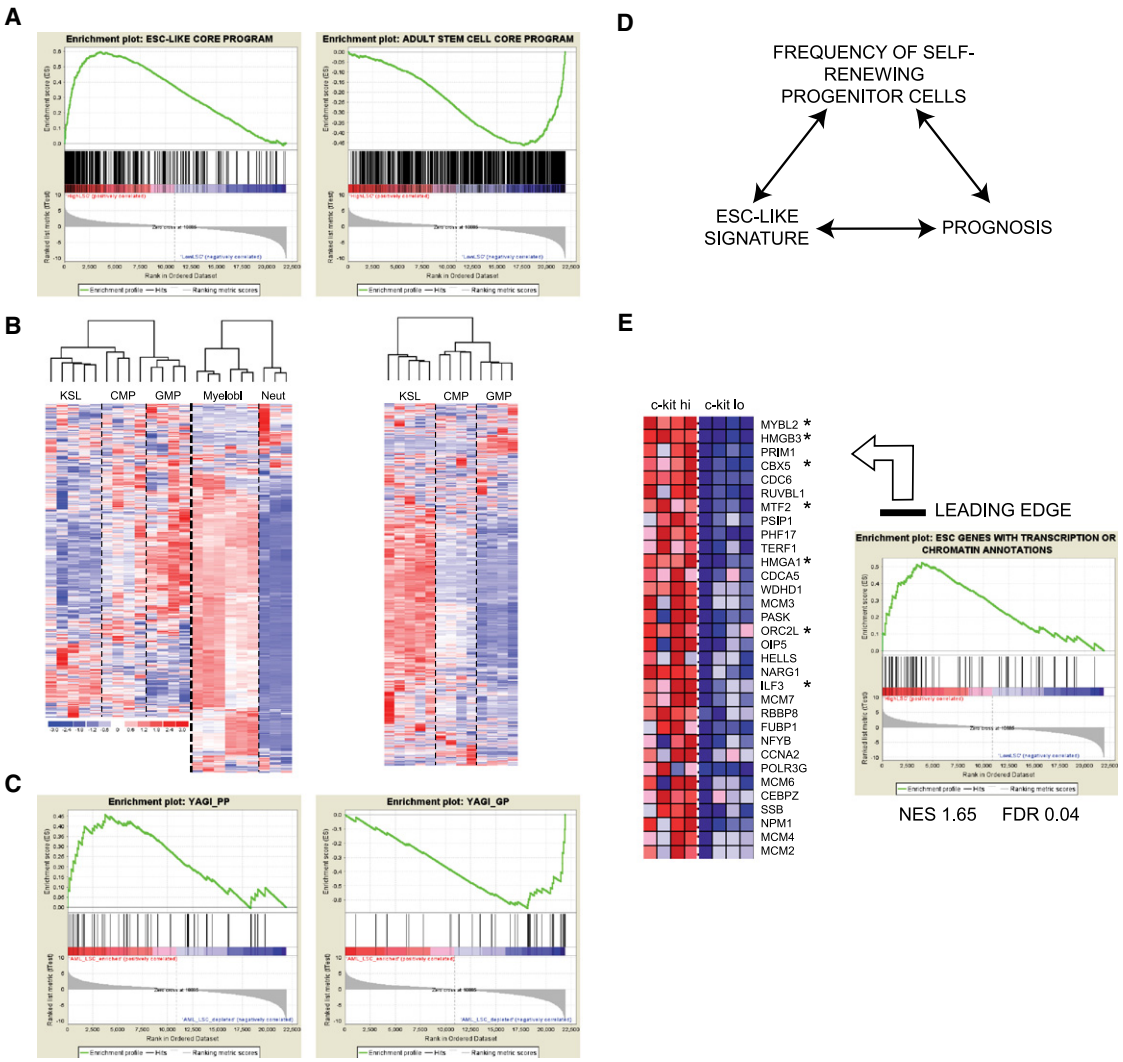

Figure 4. Shared Transcriptional Features of *MLL* LSCs, ESCs, Normal Myeloid Progenitors, and Poor-Prognosis Human Malignancies, Including Pediatric Leukemia

(A) GSEA plots show that expression of an ESC-like core gene module (Wong et al., 2008; Table S9) is enriched in high (c-kit^{hi}) versus low (c-kit^{lo}) LSC frequency *MLL-AF10* leukemia cell populations (left panel), whereas the reverse correlation is observed for an adult tissue stem cell module (right panel).

(B) Unsupervised cluster analysis of the expression pattern of ESC-like core module genes (left panel) and adult tissue stem cell module genes (right panel) (Wong et al., 2008) in the indicated defined populations of normal hematopoietic cells (Krivtsov et al., 2006, and this study). Color scale indicates normalized expression values.

(C) GSEA plots show expression of poor-prognosis genes in pediatric myeloid leukemia (Yagi et al., 2003; Table S9) is enriched in high (c-kit^{hi}) versus low (c-kit^{lo}) LSC frequency *MLL-AF10* leukemia cell populations (left panel), whereas the reverse correlation is observed for good prognosis genes (right panel).

(D) Suggested link between an ESC-like transcriptional signature, prognosis, and CSC frequency in human malignancy.

(E) Leading-edge analysis of GSEA plot shows enriched expression of ESC-like genes with transcription or chromatin annotations (Ben-Porath et al., 2008; Table S9) in high (c-kit^{hi}) versus low (c-kit^{lo}) LSC frequency *MLL-AF10* leukemia cell populations. Those marked with an asterisk, or their close homologs, are found within the LSC maintenance signature.

cell differentiation (Figures 5C and 5E). Mice secondarily transplanted with *Myb* knockdown AML cells displayed reduced leukemia incidence (50%) with considerably longer latencies compared to controls (Figure 5F). Knockdown of *Cbx5* or *Hmgb3*, but not three other upregulated LSC maintenance program genes (*E2f6*, *Serbp1*, and *Pebp1*; data not shown), also significantly reduced the frequency of AML CFCs/LSCs (Figures 6A and 6B) in direct proportion to the degree of knockdown (Figure 6C), also without inducing apoptosis (Figure S6).

By contrast, enforced expression of genes negatively correlated with LSC potential (*Klf3*, *Klf13*, but not *Smad4*; Figure S7) abrogated AML CFC/LSC potential.

Neither *Myb*, *Cbx5*, or *Hmgb3*, however, induced serial replating in semisolid culture when individually transduced into c-kit^{+} BM stem and progenitor cells (Figure 6D and data not shown). The clonogenic potentials of transduced cells were either weakly enhanced (*Myb*) or not enhanced (*Cbx5* or *Hmgb3*) in the third round of culture by comparison with cells transduced with

Table 1. Gene Set Enrichment Analyses

Gene Set Name	High versus Low LSC Frequency AML Comparison (Figure 2A) NES	FDR	c-kit^{hi} versus c-kit^{lo} AML Comparison (Figure 2D) NES	FDR
Stem Cell Gene Sets				
CORE ESC-LIKE GENE MODULE (Wong et al., 2008)	1.69	0.02	1.75	0.03
CORE ESC-LIKE GENE MODULE WITHOUT PROLIFERATION GENES	1.76	0.03	1.83	0.02
ADULT TISSUE STEM MODULE	−1.13	0.36	−1.74	0.03
ES expressed 1 (Assou et al., 2007; Ben-Porath et al., 2008)	1.26	0.19	1.66	0.03
ES expressed 2 (Ben-Porath et al., 2008)	1.52	0.09	1.58	0.03
NOS TFs (Boyer et al., 2005; Ben-Porath et al., 2008)	1.0	0.45	−1.38	0.17
Human Cancer Gene Sets				
YAGI_AML_PROGNOSIS (POOR PROG SET)	1.39	0.11	1.42	0.07
YAGI_AML_PROGNOSIS (GOOD PROG SET)	−1.53	0.07	−1.76	0.02
BRCA_PROGNOSIS_NEG	1.65	0.19	1.51	0.20
VANTVEER_BREAST_OUTCOME_GOOD_VS_POOR_DN	1.60	0.20	1.42	0.23
CANCER_UNDIFFERENTIATED_META_UP	1.43	0.26	1.54	0.19
ZHAN_MULTIPLE_MYELOMA_SUBCLASSES_DIFF	1.66	0.20	1.45	0.23

NES, normalized enrichment score. FDR, false discovery rate. NOS TFs, transcription factors regulated by Nanog, Pou5f1, and Sox2. Gene sets are from the indicated publications (and listed in Table S9) or http://www.broad.mit.edu/gsea.

control retroviruses. In pairwise combinations, *Hmgb3* slightly enhanced the weak third round clonogenic potentials induced by expression of *Myb* alone, which was dramatically enhanced by further addition of *Cbx5*. Thus, cells cotransduced by all three genes displayed robust serial replating (Figure 6D and data not shown). Colonies were exclusively type 1 and contained myeloblasts (c-kit$^+$ Mac1$^+$ Gr1$^+$ immunophenotype) that readily adapted to growth in liquid culture (data not shown). Effective bypass of the upstream *MLL* transformation pathway was confirmed by lack of *Hoxa* and *Meis1* expression in cells immortalized by the combined actions of *Myb*, *Cbx5*, and *Hmgb3* (Figure 6E). Thus, individual genes of the program are necessary but not sufficient for LSC maintenance. However, they suffice for *Hoxa*/*Meis1*-independent immortalization when simultaneously coexpressed, thereby establishing their cooperative roles in the LSC maintenance program.

DISCUSSION

The CSC model predicts that malignancies are organized into cellular hierarchies characterized by differing self-renewal potentials, such that a significant fraction of cells in the tumor clone is incapable of self-renewal despite shared genetic features with the CSC fraction (Clarke et al., 2006). Recent studies, however, have demonstrated that CSCs may be considerably more frequent than predicted by xenogeneic transplantation assays and have raised questions as to the validity and generality of the CSC model as well as the existence and extent of CSC hierarchical relationships (Kelly et al., 2007). Identification of molecular pathways that regulate the self-renewing fraction of CSCs within a malignancy would illustrate the nature of potential CSC hierarchies and additionally hold out promise for the development of therapies that specifically target the critical cellular components of malignant disease. Our current studies demonstrate that *MLL* LSCs are actively cycling downstream myeloid precursors occupying the apex of a leukemia cell hierarchy, and maintained in a self-renewing state by a genetic program more akin to embryonic than adult stem cells. This contrasts with previous suggestions that LSCs are synonymous with HSCs or exclusively maintained by corrupted HSC-specific programs (Bonnet and Dick, 1997; Krivtsov et al., 2006). Despite the downstream character of LSCs in this model of AML, our studies provide unequivocal support for a hierarchical leukemia cell organization and, in fact, further suggest that the relative abundance of self-renewing cells in human cancers may be an important prognostic factor.

MLL Oncogenes Differentially Specify LSC Hierarchies and Frequencies

MLL oncogenes are structurally heterogeneous due to differences in their respective fusion partners, which have previously been suggested to "instruct" certain biological properties of leukemia such as disease latency and lineage of the leukemic blasts (Lavau et al., 1997, 2000; Cano et al., 2007; Wong et al., 2007). Consistent with this, our comparative analysis of several molecularly distinct *MLL* oncogenes under identical experimental conditions showed that MLL fusion partners determine the frequencies of LSCs, which vary by up to 7-fold in a murine model of AML. At a cellular level, this reflects the abilities of *MLL* oncogenes to differentially specify the probability of self-renewal versus differentiation, which in turn may directly affect the cellular hierarchy and hematologic features of disease. At a molecular level, this likely reflects quantitative and/or qualitative differences in the transcriptional effecter functions of MLL fusion partners, which sustain divergent global transcriptional profiles that distinguish low versus high LSC frequency AMLs.

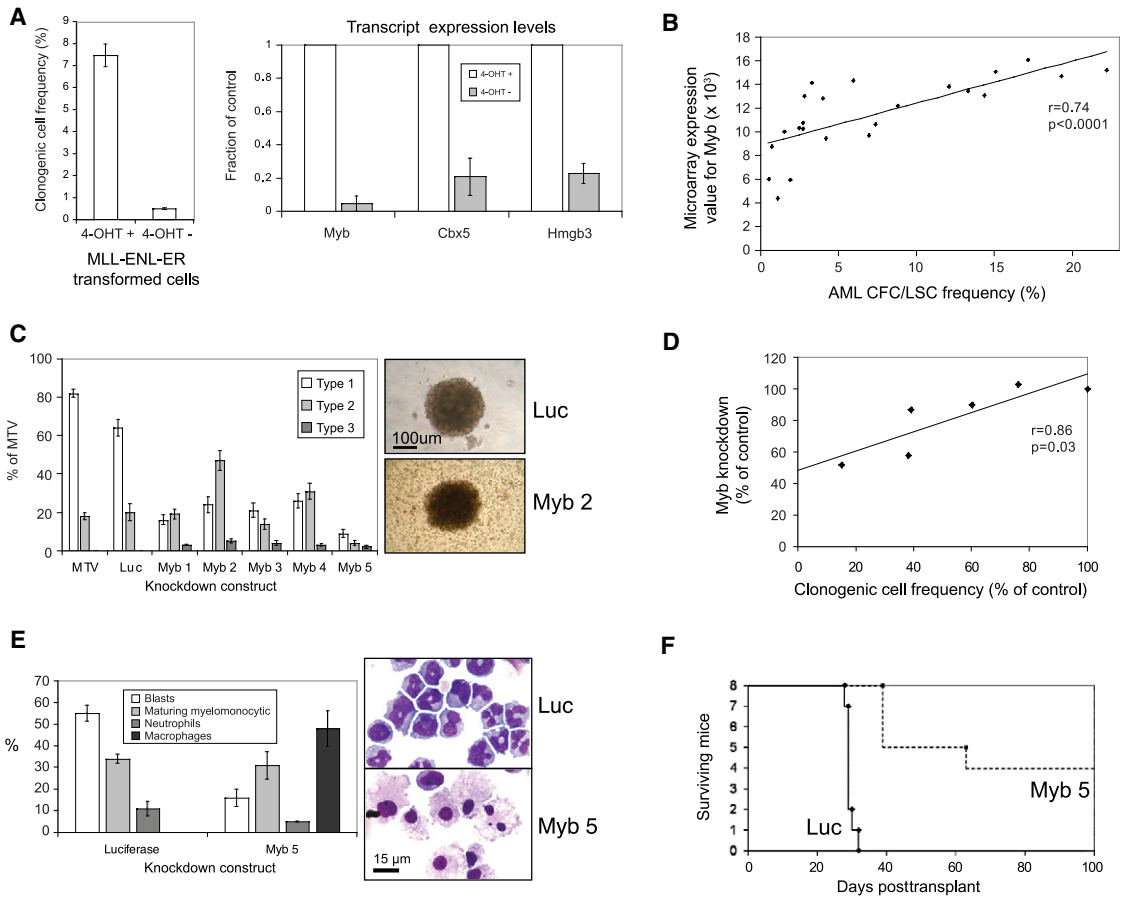

Figure 5. *Myb* Is a Critical Component of the *MLL* LSC Maintenance Signature

(A) Bar graph (left panel) shows clonogenic potentials of c-kit$^+$ BM stem and progenitor cells transformed by *MLL-ENL-ER* and then cultured in the presence or absence of 4-OHT for seven days (mean ± SEM; n = 4). Bar graph (right panel) shows expression of *Myb*, *Cbx5*, and *Hmgb3* transcripts in cells deprived of 4-OHT by comparison with control cells exposed to 4-OHT (mean ± SEM; n = 4).

(B) Graph shows that *Myb* transcript levels (microarray expression value) correlate significantly with LSC frequencies (see Figure 1A) within individual leukemias (n = 23 separate leukemias; n = 3–6 for each of the five *MLL* molecular subtypes in this study).

(C) Bar graph shows CFC potentials of *MLL-ENL* AML cells transduced with lentiviruses containing shRNA constructs targeting *Myb* or luciferase transcripts, by comparison with control cells transduced with an empty lentivirus (MTV), stratified by colony types (indicated in the representative images; mean ± SEM; n = 4). Transduced AML cells were FACS sorted for mCherry expression 48 hr following lentiviral transduction.

(D) Graph demonstrates that the extent of *Myb* knockdown measured by quantitative PCR in transduced cells 48 hr following lentiviral transduction correlated significantly with the inhibition of AML CFC/LSC formation for each of the five *Myb* knockdown constructs.

(E) Bar graph (mean ± SEM; n = 4) and representative images show morphologic analysis of *Myb* knockdown and control *MLL-ENL* AML cells 48 hr following transduction.

(F) Survival curves of sublethally irradiated syngeneic mice transplanted with 2.5–25 × 10^4 *MLL-ENL* AML cells transduced with lentiviruses targeting *Myb* or luciferase expression.

Our studies also provide support for the CSC model itself, which has recently been questioned (Kelly et al., 2007). For example, *MLL-AF10* leukemic clones consisted of approximately 5% self-renewing LSCs as well as up to 35% terminally differentiated mature neutrophils, reflecting previous reports of neutrophilic differentiation within the leukemia clone in human AML (Fearon et al., 1986). Our observation that LSC frequencies in experimentally induced AML can vary substantially, even when initiated by broadly similar oncogenes, provides direct experimental confirmation of recent proposals (Adams et al., 2007; Kennedy et al., 2007) and further suggests that variation in LSC frequency may also be a feature of human leukemia. Consistent with this, clonogenic cell frequencies vary over a 1000-fold range (Buick et al., 1977) in AMLs, which also exhibit markedly different engraftment abilities in immune-deficient mice (Pearce et al., 2006), a functional readout of LSC potential. Thus, AMLs are hierarchically organized, but the LSC compartment is not necessarily a numerically fixed, prospectively definable cellular subfraction.

The Biology of *MLL* LSCs Is Informed by Their Maintenance Signature

LSCs can be partially enriched based on immunophenotypic and other markers, but current approaches do not enable their prospective isolation to levels of purity comparable with those achieved for isolation of normal murine HSCs. Thus, as an

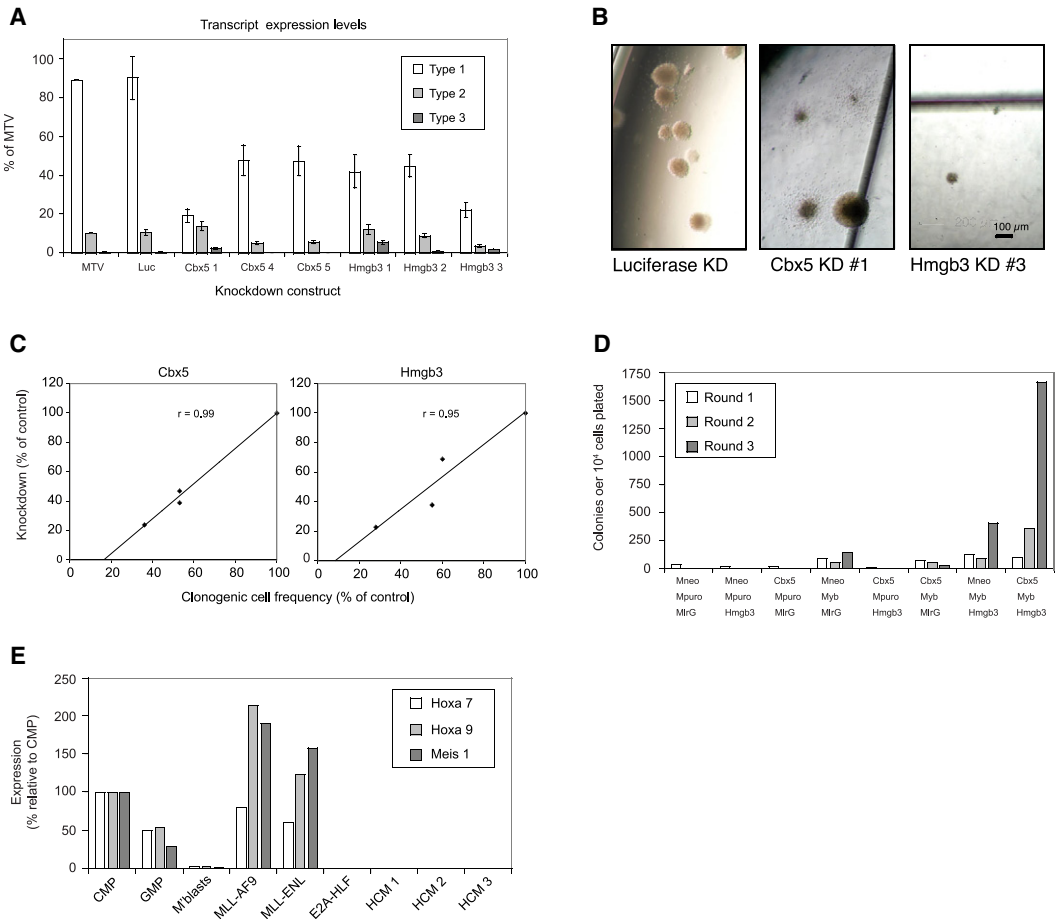

Figure 6. Requirement of *Hmgb3* and *Cbx5* for the Cooperative Maintenance of *MLL* LSC Potentials

(A and B) Bar graph shows clonogenic/LSC potentials (mean ± SEM) of *MLL-ENL* AML cells transduced with lentiviruses containing shRNA constructs targeting *Cbx5*, *Hmgb3*, or luciferase expression, by comparison with control cells transduced with an empty lentivirus (MTV). Transduced AML cells were FACS sorted for mCherry expression 48 hr following lentiviral infection. Colony types are indicated (A), and representative images are shown (B). Scale bar applies to all images.

(C) Graphs demonstrate that the extent of *Cbx5* and *Hmgb3* knockdown measured by quantitative PCR in transduced cells 48 hr following lentiviral transduction correlated with the inhibition of AML-CFC/LSC potential.

(D) Bar graph shows serial replating potentials of c-kit$^+$ BM stem and progenitor cells triple cotransduced with all possible combinations of *Myb*, *Cbx5*, *Hmgb3*, and control viruses. Mneo, MSCV Neo; Mpuro, MSCV Puro; MIrG, MSCV IRES-GFP.

(E) Bar graph shows relative expression of *Hoxa7*, *Hoxa9*, and *Meis1*, as determined by quantitative PCR, in prospectively sorted normal murine BM progenitor subsets (CMP, GMP, and c-kit$^+$ Mac1$^+$ myeloblasts) and in c-kit$^+$ BM stem and progenitor cells immortalized by the indicated oncogenes, or coexpression of *Myb*, *Cbx5* and *Hmgb3* (HCM, three separate lines are shown).

alternative approach to identify a transcriptional program that defines LSCs, and because cell-type-specific transcripts from minority cell populations can readily be detected in bulk tumor samples (Perou et al., 2000), we intersected gene sets correlating with LSC frequency derived from two different experimental methodologies (Figures 2A and 2D), each associated with low FDRs.

The derived LSC maintenance program revealed several novel attributes of *MLL* LSCs. In contrast to the prevailing notion that AML LSCs are predominantly quiescent (Jordan et al., 2006), *MLL* LSCs are predominantly metabolically active, cycling cells as demonstrated by GO and GSEA analyses of the signature, and confirmed experimentally by cell-cycle analysis of sorted leukemia cell subpopulations. Furthermore, upregulated LSC maintenance program genes are maximally expressed at a mid-myeloid stage during normal differentiation (GMPs and myeloblasts) and mostly downregulated in mature neutrophils and the KSL compartment, which contains HSCs. Together with the observation that *MLL* LSC transcriptional profiles are enriched for expression of gene sets expressed in normal lineage-committed hematopoietic progenitor cells rather than long-term HSCs, these data emphasize that the LSC in *MLL* leukemias is an aberrantly self-renewing downstream myeloid cell, rather than a tissue stem cell aberrantly expressing myeloid lineage antigens (Passegue et al., 2003; Somervaille and Cleary, 2006).

Although *Hoxa* and *Meis1* expression is induced by MLL oncoproteins (Krivtsov et al., 2006) and required for initiation of *MLL* leukemia (Ayton and Cleary, 2003; So et al., 2004; Wong et al., 2007), these genes were not part of the LSC maintenance program because their expression levels did not vary concordantly between leukemia cell populations containing different frequencies of LSCs. Indeed, *MLL-AF10* LSCs underwent terminal granulocytic differentiation with concomitant loss of

self-renewal potential in the presence of continued high-level *Hoxa* and *Meis1* transcript levels, although the possibility of posttranslational regulation is not excluded. However, there is also no overlap of the maintenance program with a KSL/LSC shared signature (Krivtsov et al., 2006), defined by transcripts with high-level coexpression in normal KSL cells and *MLL-AF9* LSCs but low-level expression in committed myeloid progenitors (CMP, GMP, and MEP), which likely represents an *MLL* leukemia initiation program. Thus, the genetic programs required for leukemia initiation may be quite distinct from those that function to maintain LSCs within the self-renewing fraction and regulate their differentiation. It appears that a critical function of the MLL oncoprotein, through its subordinate effecters *Hoxa* and *Meis1*, is to transmute a normal midmyeloid lineage transcriptional program into an LSC maintenance program that is not synonymous with an adult stem cell program.

The Nexus of *MLL* LSCs, ESCs, Myeloid Progenitors, and Poor-Prognosis Human Malignancies

Module map analysis indicates that stem cell expression programs are clustered into two distinct categories, ESC-like programs versus adult tissue stem cell-like programs (Wong et al., 2008), and that there is likely not a shared program for self-renewal in all types of stem cells, in contrast to previous suggestions (Ivanova et al., 2002). Expression of ESC-like core program genes is substantially enriched in self-renewing ESCs and *MLL* LSCs, as well as in diverse poor-prognosis, poorly differentiated human malignancies, which may contain frequent CSCs. However, our analyses also showed transient upregulation of ESC-like program genes during normal myeloid differentiation, with inverse regulation of adult stem cell core program genes, emphasizing that these dissimilar programs are not necessarily specific to self-renewing cells.

The ESC-like state in *MLL* LSCs was not mediated by *Nanog*, *Pou5f1*, or *Sox2*, critical regulators of ESC fate that are not significantly expressed in *MLL* LSCs. Nor was it directly controlled by *Myc*, which induces the ESC-like signature and regulates CSC frequency in experimentally induced human epithelial tumors (Wong et al., 2008), because LSC frequencies across distinct *MLL* leukemia subtypes did not correlate significantly with its expression (data not shown).

We propose a unifying model whereby the ESC-like program, which coordinately maintains proliferation and inhibits differentiation, is subordinate to diverse upstream regulators that critically sustain its expression. Upstream positive regulators of the program might include *Nanog*, *Pou5f1*, and *Sox2* in ESCs; *Myc* in experimentally induced human skin cancer (Wong et al., 2008); and *Hoxa*/*Meis* in *MLL* LSCs. This model is supported by our finding that coexpression of *Myb*, *Hmgb3*, and *Cbx5*, genes that are expressed in both ESCs and *MLL* LSCs, is sufficient to induce *Hoxa*/*Meis*-independent immortalization of myeloid progenitors. Although expression of the ESC-like program is induced during normal myeloid differentiation, the absence of sustained *Hoxa*/*Meis* expression in committed myeloid progenitors/precursors may permit its downregulation, allowing terminal differentiation. Conversely, an inability to normally extinguish the ESC-like program in hematopoietic progenitors is leukemogenic. This tentative model pathogenically links inappropriate expression of the initiating *MLL*/*Hox*/*Meis* subprogram, which is typically a feature of HSCs, with a downstream ESC-like program in maintaining the aberrant self-renewal of myeloid progenitors that serve as LSCs in *MLL*-associated AML, although how the two programs interlink is not clear.

The co-ordinate expression of ESC-like genes, rather than adult stem cell program genes, in poor-prognosis human malignancies and *MLL* LSCs further suggests that the frequency of CSCs may be more generally linked to prognosis than previously appreciated, and that these cells are more akin to aberrantly self-renewing downstream tissue progenitor cells than their normal tissue stem cell counterparts. Consistent with this possibility, adverse clinical outcome in human AML correlates with the ability of AML cells to engraft immune-deficient mice (Pearce et al., 2006), a functional assay for LSC potential that may also measure their relative frequencies. Furthermore, human AML cells with downstream progenitor cell immunophenotypes are able to initiate leukemia in immune-deficient mice (Taussig et al., 2008).

Cooperative Roles of LSC Maintenance Genes as Downstream Effecters in *MLL* Leukemogenesis

Knockdown of upregulated LSC maintenance program genes (*Myb*, *Hmgb3*, and *Cbx5*) abrogated AML CFC/LSC potential in proportion to the extent of knockdown, validating their individual contributions. Furthermore, their simultaneous coexpression robustly induced BM stem and progenitor cell immortalization independent of *Hoxa*/*Meis1*, whereas expression of the genes individually did not. These results illustrate the cooperative and essential nature of an ESC-like LSC maintenance program, which lies downstream of the initiating *MLL* oncogene and its *Hoxa*/*Meis1* effecters.

Although *Hmgb3* regulates self-renewal in HSCs (Nemeth et al., 2006), and is frequently highly expressed in poor-prognosis breast cancer (Ben-Porath et al., 2008), the mechanisms by which it collaborates with *Myb*, a key regulator of normal and leukemic hematopoiesis (Ramsay and Gonda, 2008), and *Cbx5*, which normally functions to promote formation of heterochromatin, to mediate LSC maintenance are currently unclear. Furthermore, it is also unclear whether downregulation of LSC maintenance program genes in vivo, which permits LSC differentiation and loss of self-renewal, is stochastic or deterministic. Nevertheless, our findings emphasize the potential for therapeutic approaches that target genes and pathways that are of greater importance to the function of LSCs than HSCs, since LSCs in this model are distinguished by biological differences from normal HSCs, which are necessary for regeneration of hematopoiesis following chemotherapy.

EXPERIMENTAL PROCEDURES

Mice

C57BL/6 mice were purchased from either Jackson Laboratories (Bar Harbor, Maine) or the Stanford University Research Animal Facility and maintained on an inbred background. For experiments requiring use of congenic marker analysis, B6.SJL-*Ptprc*^{a} *Pepc*^{b}/BoyJ mice, purchased from Jackson Laboratories, were used as transplant recipients. All experiments on mice in this study were performed with the approval of and in accordance with Stanford University's Administrative Panel on Laboratory Animal Care.

Plasmids

Retroviral vectors containing *MLL-ENL* (Lavau et al., 1997), *MLL-AF10* (DiMartino et al., 2002), *MLL-GAS7* (So et al., 2003a), *MLL-AF1p* (So et al., 2003b),

MLL-ENL-ER (Ayton and Cleary, 2003) (all in pMSCV Neo), or *MLL-AF9* (pMSCV Puro; Somervaille and Cleary, 2006), and preparation of retroviral supernatants, have been previously described. Other retroviral vectors used included the following: HA-Cbx5 (NM_007626) in pMSCVNeo; Myc-tagged Myb (NM_010848) in pMSCVPuro; and HA-Hmgb3 (NM_008253), HA-Klf3 (NM_008453), HA-Klf13 (NM_021366) and HA-Smad4 (NM_008540) (all in pMSCV-ires-GFP). The pSicoR lentiviral vector (Ventura et al., 2004) was adapted to express the monomeric red fluorescent protein mCherry (Shaner et al., 2004). shRNAs were designed using pSicoOligomaker v1.5 (http://web.mit.edu/ccr/labs/jacks/protocols/pSico.html), and sequences are listed in Table S10. Lentiviral supernatants were manufactured as described (Wong et al., 2007).

BM Stem/Progenitor Cell Transduction, Transplantation, and Colony-Forming Assays

C-kit$^+$ BM stem and progenitor cells were isolated from donor mice and transduced with retroviruses as previously described (Somervaille and Cleary, 2006). For triple transductions, BM cells were spinoculated sequentially with each retroviral supernatant for 30 min at 32°C and 1350 g. Transduced cells (3–10 × 10^5) were transplanted by retro-orbital injection immediately following spinoculation into lethally irradiated (900 cGy) syngeneic recipient mice, together with an additional radioprotective dose of 2 × 10^5 nucleated whole BM cells. An aliquot of the postspinoculation cells was cultured in methylcellulose medium with 20 ng/ml SCF, 10 ng/ml IL-6, 10 ng/ml GM-CSF, and 10 ng/ml IL-3 (Peprotech, Rocky Hill, New Jersey), in the presence or absence of G418 or puromycin drug selection, to determine retrospectively the transplanted dose of transduced CFCs.

For congenic transplantation analyses, postspinoculation CD45.2$^+$ cells were serially replated in semisolid culture for three rounds, as previously described (Somervaille and Cleary, 2006). Pooled, washed cells (8 × 10^5) were then transplanted into sublethally irradiated CD45.1$^+$ recipients (450 cGy).

For mouse leukemia CFC assays and limit dilution secondary transplantation assays, splenocytes frozen in 10% DMSO were thawed and then rested in RPMI with 20% fetal calf serum and 20% WEHI-conditioned medium for 4 hr. FACS-sorted Mac1$^+$ PI$^-$ cells were then plated in semisolid culture, as above, for 6 days prior to colony enumeration or injected immediately in varying doses into sublethally irradiated syngeneic recipients. For secondary transplantation of cells derived from a single AML CFC, individual AML colonies were plucked from semisolid culture after 5 days and replated once, and then, after a further 5 days, 2.5 × 10^5 cells were washed and injected into sublethally irradiated syngeneic recipients. For secondary transplantation of AML cells transduced with *Myb* knockdown or control lentiviruses, cells were sorted by FACS 48 hr following transduction for red fluorescence and then transplanted into sublethally irradiated syngeneic mice.

Flow Cytometry

Analyses were performed, as previously described (Somervaille and Cleary, 2006), using an LSR Model 1a flow cytometer (BD Biosciences, San Jose, California). Some analyses and all cell-sorting experiments were performed using a dual-laser FACS Vantage (BD Biosciences). The following antibodies were used (from eBioscience, San Diego, California, or BD Biosciences where indicated): Mac1-FITC (eBioscience) or Cy7APC (BD Biosciences) (clone M1/70), Gr1-PE or Cy7PE (clone RB6-8C5, eBioscience), c-kit-APC (clone 2B8, eBioscience), F4/80-PE (Caltag Laboratories, Burlingame, California), CD45.2-FITC (clone 104; BD Biosciences), and CD45.1-PE (clone A20; BD Biosciences).

Quantitative Polymerase Chain Reaction Assays

Quantitative PCR was performed as described (Somervaille and Cleary, 2006) using the following primer sets from Applied Biosystems (Foster City, California): *Myb* (Mm00501741_m1), *Cbx5* (Mm00483092_m1), *Hmgb3* (Mm01281712_g1), *Hoxa7* (Mm00657963_m1), *Hoxa9* (Mm00439364_m1), *Meis1* (Mm00487664_m1) and *Actb* (Mm00607939_s1).

Microarray and Bioinformatics Analyses

For whole *MLL* leukemias, RNA was extracted from at least 5 × 10^6 BM leukemia cells by using Trizol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, California). For sorted *MLL-AF10* leukemia cell and normal myeloid cell subpopulations, RNA was prepared using an RNeasy Mini kit (QIAGEN, Valencia, California). RNA quality was confirmed using an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, California). For whole *MLL* leukemias, cDNA and cRNA were synthesized, fragmented, and hybridized to Affymetrix Mouse Genome 430 2.0 microarrays (Affymetrix, Santa Clara, CA) according to manufacturer's instructions. For sorted *MLL-AF10* leukemia cell and normal myeloid cell subpopulations, RNA was amplified using a GeneChip Two-Cycle Target Labeling and Control Reagents kit (Affymetrix), and cRNA was hybridized to Affymetrix Mouse Genome 430 2.0 microarrays (Affymetrix), as per the manufacturer's instructions. Hybridized chips were scanned by using the GeneChip Scanner 3000, with GeneChip operating software (v1.1.1) (Affymetrix). Raw data are available for download from Gene Expression Omnibus (GSE13690; GSE13692; GSE13693).

Whole *MLL* leukemia BM data sets were normalized using Microarray Analysis Software v5.0 (Affymetrix). All other normalizations (including downloaded CEL file data from Krivtsov et al. [2006]), together with all data set comparisons and cluster analyses, were performed using dChip 2007 (DNA-Chip Analyzer) software (Li and Wong, 2001; http://biosun1.harvard.edu/complab/dchip/manual.htm). A perfect match/mismatch (PM/MM) model was used for the calculation of expression values. Differentially regulated probe sets in whole *MLL* leukemias, and in sorted leukemia cell subpopulations, were determined using $p \leq 0.01$ (unpaired t test) as the sole comparison criterion. Median FDR was assessed using 1000 data permutations. Probe set lists were intersected using a web-based tool (http://jura.wi.mit.edu/bioc/tools/compare.html). For certain analyses, expression values for defined gene sets were extracted from normalized data sets using the Excel function VLOOKUP. Unsupervised cluster analysis, with masking of redundant probe sets to exclude gene duplicates, was performed using the resultant tab-delimited text files. HUGO gene names of core ESC-like gene module genes and adult stem cell tissue module genes were converted into 430 2.0 Affymetrix probe set identifiers using a webtool (http://www.affymetrix.com).

GO analysis was performed using MGI Gene Ontology Term Finder (http://proto.informatics.jax.org/prototypes/GOTools/web-docs/MGI_Term_Finder.html).

Gene set enrichment analyses (Subramanian et al., 2005) were performed using GSEA v2.0 software (http://www.broad.mit.edu/gsea) with a tTest metric for gene ranking and 1000 data permutations.

SUPPLEMENTAL DATA

The Supplemental Data include seven figures and ten tables and can be found with this article online at http://www.cell.com/cell-stem-cell/supplemental/S1934-5909(08)00617-6.

ACKNOWLEDGMENTS

We thank Maria Ambus, Cita Nicholas, Jude Aldaco, and Letha Phillips for technical support; Scott Kogan for consultation on leukemia classifications; Bill Wong for MLL-ENL-ER cells; Robert Tibshirani, Francesca Ficara, and Crispin Miller for helpful discussions; and Julien Sage and Patrick Viatour for the mCherry construct. T.C.P.S. was supported by a Leukaemia Research Fund (UK) Senior Clinical Fellowship and by Cancer Research UK grant number C147/A6058. C.J.M. was supported by Public Health Service grant T32 CA09151 awarded by the National Cancer Institute, Department of Health and Human Services. J.L.R. is a Fellow and H.Y.C. is the Kenneth G. and Elaine A. Langone Scholar of the Damon Runyon Cancer Research Foundation. We acknowledge support from the American Cancer Society (H.Y.C.), Children's Health Initiative of the Packard Foundation, PHS grants CA55029 and CA116606, and the Leukemia and Lymphoma Society.

Received: January 8, 2008
Revised: September 17, 2008
Accepted: November 25, 2008
Published: February 5, 2009

REFERENCES

Adams, J.M., Kelly, P.N., Dakic, A., Nutt, S.L., and Strasser, A. (2007). Response to Comment on "Tumor growth need not be driven by rare cancer stem cells". Science *318*, 1722.

Assou, S., Le Carrour, T., Tondeur, S., Strom, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Reme, T., Hugnot, J.P., et al. (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells *25*, 961–973.

Ayton, P.M., and Cleary, M.L. (2003). Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. Genes Dev. *17*, 2298–2307.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. (2008). An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat. Genet. *40*, 499–507.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat. Med. *3*, 730–737.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Buick, R.N., Till, J.E., and McCulloch, E.A. (1977). Colony assay for proliferative blast cells circulating in myeloblastic leukaemia. Lancet *1*, 862–863.

Cano, F., Drynan, L.F., Pannell, R., and Rabbitts, T.H. (2007). Leukaemia lineage specification caused by cell-specific Mll-Enl translocations. Oncogene *27*, 1945–1950.

Clarke, M.F., Dick, J.E., Dirks, P.B., Eaves, C.J., Jamieson, C.H., Jones, D.L., Visvader, J., Weissman, I.L., and Wahl, G.M. (2006). Cancer stem cells—perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer Res. *66*, 9339–9344.

DiMartino, J.F., Ayton, P.M., Chen, E.H., Naftzger, C.C., Young, B.D., and Cleary, M.L. (2002). The AF10 leucine zipper is required for leukemic transformation of myeloid progenitors by MLL-AF10. Blood *99*, 3780–3785.

Fearon, E.R., Burke, P.J., Schiffer, C.A., Zehnbauer, B.A., and Vogelstein, B. (1986). Differentiation of leukemia cells to polymorphonuclear leukocytes in patients with acute nonlymphocytic leukemia. N. Engl. J. Med. *315*, 15–24.

Hess, J.L., Bittner, C.B., Zeisig, D.T., Bach, C., Fuchs, U., Borkhardt, A., Frampton, J., and Slany, R.K. (2006). c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. Blood *108*, 297–304.

Ivanova, N.B., Dimos, J.T., Schaniel, C., Hackney, J.A., Moore, K.A., and Lemischka, I.R. (2002). A stem cell molecular signature. Science *298*, 601–604.

Jordan, C.T., Guzman, M.L., and Noble, M. (2006). Cancer stem cells. N. Engl. J. Med. *355*, 1253–1261.

Kelly, P.N., Dakic, A., Adams, J.M., Nutt, S.L., and Strasser, A. (2007). Tumor growth need not be driven by rare cancer stem cells. Science *317*, 337.

Kennedy, J.A., Barabe, F., Poeppl, A.G., Wang, J.C., and Dick, J.E. (2007). Comment on ''Tumor growth need not be driven by rare cancer stem cells''. Science *318*, 1722.

Kogan, S.C., Ward, J.M., Anver, M.R., Berman, J.J., Brayton, C., Cardiff, R.D., Carter, J.S., de Coronado, S., Downing, J.R., Fredrickson, T.N., et al. (2002). Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice. Blood *100*, 238–245.

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.

Lavau, C., Szilvassy, S.J., Slany, R., and Cleary, M.L. (1997). Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. EMBO J. *16*, 4226–4237.

Lavau, C., Luo, R.T., Du, C., and Thirman, M.J. (2000). Retrovirus-mediated gene transfer of MLL-ELL transforms primary myeloid progenitors and causes acute myeloid leukemias in mice. Proc. Natl. Acad. Sci. USA *97*, 10984–10989.

Lessard, J., and Sauvageau, G. (2003). Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature *423*, 255–260.

Li, C., and Wong, W.H. (2001). Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc. Natl. Acad. Sci. USA *98*, 31–36.

Look, A.T. (1997). Oncogenic transcription factors in the human acute leukemias. Science *278*, 1059–1064.

Nemeth, M.J., Kirby, M.R., and Bodine, D.M. (2006). Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. Proc. Natl. Acad. Sci. USA *103*, 13783–13788.

Passegue, E., Jamieson, C.H., Ailles, L.E., and Weissman, I.L. (2003). Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? Proc. Natl. Acad. Sci. USA *100* (*Suppl 1*), 11842–11849.

Pearce, D.J., Taussig, D., Zibara, K., Smith, L.L., Ridler, C.M., Preudhomme, C., Young, B.D., Rohatiner, A.Z., Lister, T.A., and Bonnet, D. (2006). AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. Blood *107*, 1166–1173.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. Nature *406*, 747–752.

Ramsay, R.G., and Gonda, T.J. (2008). MYB function in normal and cancer cells. Nat. Rev. Cancer *8*, 523–534.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. *22*, 1567–1572.

So, C.W., Karsunky, H., Passegue, E., Cozzio, A., Weissman, I.L., and Cleary, M.L. (2003a). MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. Cancer Cell *3*, 161–171.

So, C.W., Lin, M., Ayton, P.M., Chen, E.H., and Cleary, M.L. (2003b). Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. Cancer Cell *4*, 99–110.

So, C.W., Karsunky, H., Wong, P., Weissman, I.L., and Cleary, M.L. (2004). Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. Blood *103*, 3192–3199.

Somervaille, T.C., and Cleary, M.L. (2006). Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. Cancer Cell *10*, 257–268.

Steidl, U., Rosenbauer, F., Verhaak, R.G., Gu, X., Ebralidze, A., Otu, H.H., Klippel, S., Steidl, C., Bruns, I., Costa, D.B., et al. (2006). Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. Nat. Genet. *38*, 1269–1277.

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., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Taussig, D.C., Miraki-Moud, F., Anjos-Afonso, F., Pearce, D.J., Allen, K., Ridler, C., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S.G., et al. (2008). Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. Blood *112*, 568–575.

Ventura, A., Meissner, A., Dillon, C.P., McManus, M., Sharp, P.A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc. Natl. Acad. Sci. USA *101*, 10380–10385.

Wong, P., Iwasaki, M., Somervaille, T.C., So, C.W., and Cleary, M.L. (2007). Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. Genes Dev. *21*, 2762–2774.

Wong, D.J., Liu, H., Ridky, T.W., Cassarino, D., Segal, E., and Chang, H.Y. (2008). Module map of stem cell genes guides creation of epithelial cancer stem cells. Cell Stem Cell *2*, 333–344.

Yagi, T., Morimoto, A., Eguchi, M., Hibi, S., Sako, M., Ishii, E., Mizutani, S., Imashuku, S., Ohki, M., and Ichikawa, H. (2003). Identification of a gene expression signature associated with pediatric AML prognosis. Blood *102*, 1849–1856.