

# Sox4 Is a Key Oncogenic Target in C/EBP $\alpha$ Mutant Acute Myeloid Leukemia

Hong Zhang,<sup>1,2</sup> Meritxell Alberich-Jorda,<sup>1,2,3</sup> Giovanni Amabile,<sup>1,2</sup> Henry Yang,<sup>4</sup> Philipp B. Staber,<sup>1,2,5</sup> Annalisa DiRuscio,<sup>1,2</sup> Robert S. Welner,<sup>1,2</sup> Alexander Ebralidze,<sup>1,2</sup> Junyan Zhang,<sup>1,2</sup> Elena Levantini,<sup>1,2,6</sup> Véronique Lefebvre,<sup>7</sup> Peter J.M. Valk,<sup>8</sup> Ruud Delwel,<sup>8</sup> Maarten Hoogenkamp,<sup>9</sup> Claus Nerlov,<sup>10</sup> Jörg Cammenga,<sup>11</sup> Borja Saez,<sup>1,12,13</sup> David T. Scadden,<sup>1,12,13</sup> Constanze Bonifer,<sup>9</sup> Min Ye,<sup>1,2,\*</sup> and Daniel G. Tenen<sup>1,4,\*</sup>

<sup>1</sup>Harvard Stem Cell Institute, Harvard Medical School, Boston, MA 02215, USA

<sup>2</sup>Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

<sup>3</sup>Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague 14220, Czech Republic

<sup>4</sup>Cancer Science Institute, National University of Singapore, Singapore 117599, Singapore

<sup>5</sup>Division of Hematology and Hemostaseology, Comprehensive Cancer Centre Vienna, Medical University of Vienna, Vienna 1090, Austria

<sup>6</sup>Institute of Biomedical Technologies, National Research Council, Pisa 56124, Italy

<sup>7</sup>Department of Cellular and Molecular Medicine, Orthopaedic and Rheumatologic Research Center, Cleveland Clinic Lerner Research Institute, Cleveland, OH 44195, USA

<sup>8</sup>Erasmus University Medical Center, Rotterdam 3015 GE, the Netherlands

<sup>9</sup>School of Cancer Sciences, Institute of Biomedical Research, University of Birmingham, Birmingham B15 2TT, UK

<sup>10</sup>MRC Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK

<sup>11</sup>Department of Molecular Medicine and Gene Therapy and Department of Hematology, Lund University, Lund 22184, Sweden

<sup>12</sup>Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

<sup>13</sup>Stem Cell and Regenerative Biology Department, Harvard University, Cambridge, MA 02138, USA

\*Correspondence: [mye@bidmc.harvard.edu](mailto:mye@bidmc.harvard.edu) (M.Y.), [csidgt@nus.edu.sg](mailto:csidgt@nus.edu.sg) (D.G.T.)

<http://dx.doi.org/10.1016/j.ccr.2013.09.018>

## SUMMARY

Mutation or epigenetic silencing of the transcription factor C/EBP $\alpha$  is observed in ~10% of patients with acute myeloid leukemia (AML). In both cases, a common global gene expression profile is observed, but downstream targets relevant for leukemogenesis are not known. Here, we identify Sox4 as a direct target of C/EBP $\alpha$  whereby its expression is inversely correlated with C/EBP $\alpha$  activity. Downregulation of Sox4 abrogated increased self-renewal of leukemic cells and restored their differentiation. Gene expression profiles of leukemia-initiating cells (LICs) from both Sox4 overexpression and murine C/EBP $\alpha$  mutant AML models clustered together but differed from other types of AML. Our data demonstrate that Sox4 overexpression resulting from C/EBP $\alpha$  inactivation contributes to the development of leukemia with a distinct LIC phenotype.

## INTRODUCTION

Acute myeloid leukemia (AML) is characterized by a differentiation block and aberrant clonal growth of hematopoietic blasts. It has been classified into individual subtypes with respect to morphology, immunophenotype, and genetic abnormalities. In recent years, genome-wide gene expression profiling has further identified distinct subsets (Valk et al., 2004), which may reflect the underlying biology of these subtypes and potentially reveal critical downstream targets for therapeutic intervention.

Transcription factor CEBPA is differentially translated into two isoforms of 42 kDa and 30 kDa (Lin et al., 1993). Two-thirds of AML cases with acquired point mutations of CEBPA have one allele harboring N-terminal frame-shift mutations leading to increased 30 kDa isoform and the other allele harboring C-terminal in-frame insertions or deletions resulting in deficient DNA-binding and/or homodimerization activities (Gombart et al., 2002; Pabst et al., 2001). CEBPA double-mutant cases and cases where CEBPA has been epigenetically silenced demonstrate similar gene expression signatures, suggesting a common mechanism of disease (Valk et al., 2004).

### Significance

The alteration of the activity of transcription factors crucial for blood cell development with a concomitant deregulation of their target genes is a recurrent pattern in leukemia. However, the precise nature and the role of their downstream targets are still poorly understood. This also holds true for AML with loss of C/EBP $\alpha$  activity. Here, we identified Sox4 as a direct target of transcriptional repression by C/EBP $\alpha$  and functionally validated it as a central mediator of C/EBP $\alpha$  deficiency. Our data provide an explanation for the high levels of SOX4 expression in a human AML subtype carrying mutated or silent CEBPA and highlight SOX4 as a potential therapeutic target for a subtype of AML with currently limited therapeutic options.

C/EBP $\alpha$  regulates the expression of myeloid lineage-specific genes and cell-cycle regulators, affect the self-renewal and myeloid lineage commitment of hematopoietic stem cells (HSCs), and induce growth arrest (Nerlov, 2004). However, the 30 kDa isoform fails to induce differentiation of granulocytes and to block cell proliferation (Nerlov, 2004). *Cebpa* knockout mice die at birth with a complete lack of mature granulocytes, while adult mice with induced loss of C/EBP $\alpha$  demonstrate a block from common myeloid progenitors (CMP) to granulocyte monocyte progenitors (GMP) and accumulation of myeloid blasts (Ye et al., 2013; Zhang et al., 2004). Knockin mice carrying engineered biallelic *CEBPA* mutations as found in human AML developed leukemia (Bereshchenko et al., 2009), but the key molecular downstream events required to trigger leukemogenesis remain unclear.

Sox4 belongs to the Sox (SRY-related HMG-box) transcription factor family (Jafarnejad et al., 2013). T cell development in Sox4-deficient embryos is severely impaired (Kuwahara et al., 2012), and mice receiving Sox4<sup>-/-</sup> fetal liver cells exhibit a block at the pro-B cell stage (Schilham et al., 1996). SOX4 is upregulated in various types of human solid tumors and is a frequent target of retroviral insertional mutagenesis in many murine B cell lymphoma and myeloid leukemia models (Jafarnejad et al., 2013). Its overexpression is associated with clonal dominance of HSC (Kustikova et al., 2007), stem/progenitor cell repopulation advantage (Deneault et al., 2009), a block in differentiation of myeloid progenitor 32DC13 cells (Boyd et al., 2006), and induction of myeloid leukemia (Du et al., 2005; Kvinlaug et al., 2011). However, the precise role of Sox4 gene in AML and how it is involved in specific AML subtypes is poorly understood.

## RESULTS

### A shRNA Screen Identifies Sox4 as a Mediator of Enhanced Replating Ability and Decreased Differentiation of *Cebpa*-Deficient Cells in Culture

Previous studies have revealed that disruption of C/EBP $\alpha$  in the hematopoietic system resulted in abnormal expansion and an altered transcription program of hematopoietic stem cells (Ye et al., 2013; Zhang et al., 2004). To identify the downstream effectors, we performed genome-wide gene expression profiling and verified expression changes of the top 30 candidates of up-regulated genes upon loss of C/EBP $\alpha$  in LSK cells (Lin<sup>-</sup>Sca1<sup>+</sup>c-kit<sup>+</sup>) (Figure 1A; Table S1 available online). We then functionally evaluated the effect of knocking down these genes on *Cebpa* KO cells (Mx1-Cre<sup>+</sup> *Cebpa*<sup>loxP/loxP</sup>; from here on referred to as *Cebpa* KO following Cre-mediated deletion) after serially replating in methylcellulose cultures, a cell culture assay that has been correlated with the ability to induce leukemia in mice (Huntly et al., 2004; Lavau et al., 1997; Moran-Crusio et al., 2011). We transduced *Cebpa* KO LSK cells with lentiviruses carrying either a mix of scrambled small hairpin RNA (shRNA) (control) or a pool of five shRNAs all targeting one specific candidate and assessed their capability to undergo serial replating (Figure S1A). Among the 30 candidates, shRNA-mediated knockdown of Sox4 exhibited the strongest reduction of serial replating capability of *Cebpa* KO LSK cells, with only a few colonies formed after two rounds of replating and none at the fourth round, while scrambled controls maintained colony formation even after four rounds

of replating (Figures 1B and S1B). Moreover, Sox4 shRNA-transduced *Cebpa* KO LSK cells differentiated into macrophages (Figure 1C), as compared to morphologically immature control colonies (aberrantly differentiated leukemic blast cells). Flow cytometry analysis also revealed significantly increased expression of the mature myeloid marker Mac1 (Figure 1D). Genotyping PCR results confirmed complete excision of *Cebpa* alleles in these cells (Figures S1C and S1D).

Because knocking down of Sox4 impaired the outgrowth capability of *Cebpa* KO cells, we asked if Sox4 overexpression was sufficient to confer replating ability to wild-type cells. We therefore overexpressed Sox4 in wild-type LSK and progenitor cells (common myeloid and granulocyte-macrophage progenitors) and again performed serial replating assays. Aberrant Sox4 expression resulted in enhanced replating ability for more than four rounds in all tested cell types (Figures 1E and S1E). Moreover in cytokine-supplemented myeloid liquid culture of Sox4 transduced cells, myeloid differentiation was impaired (Figures 1F and S1F).

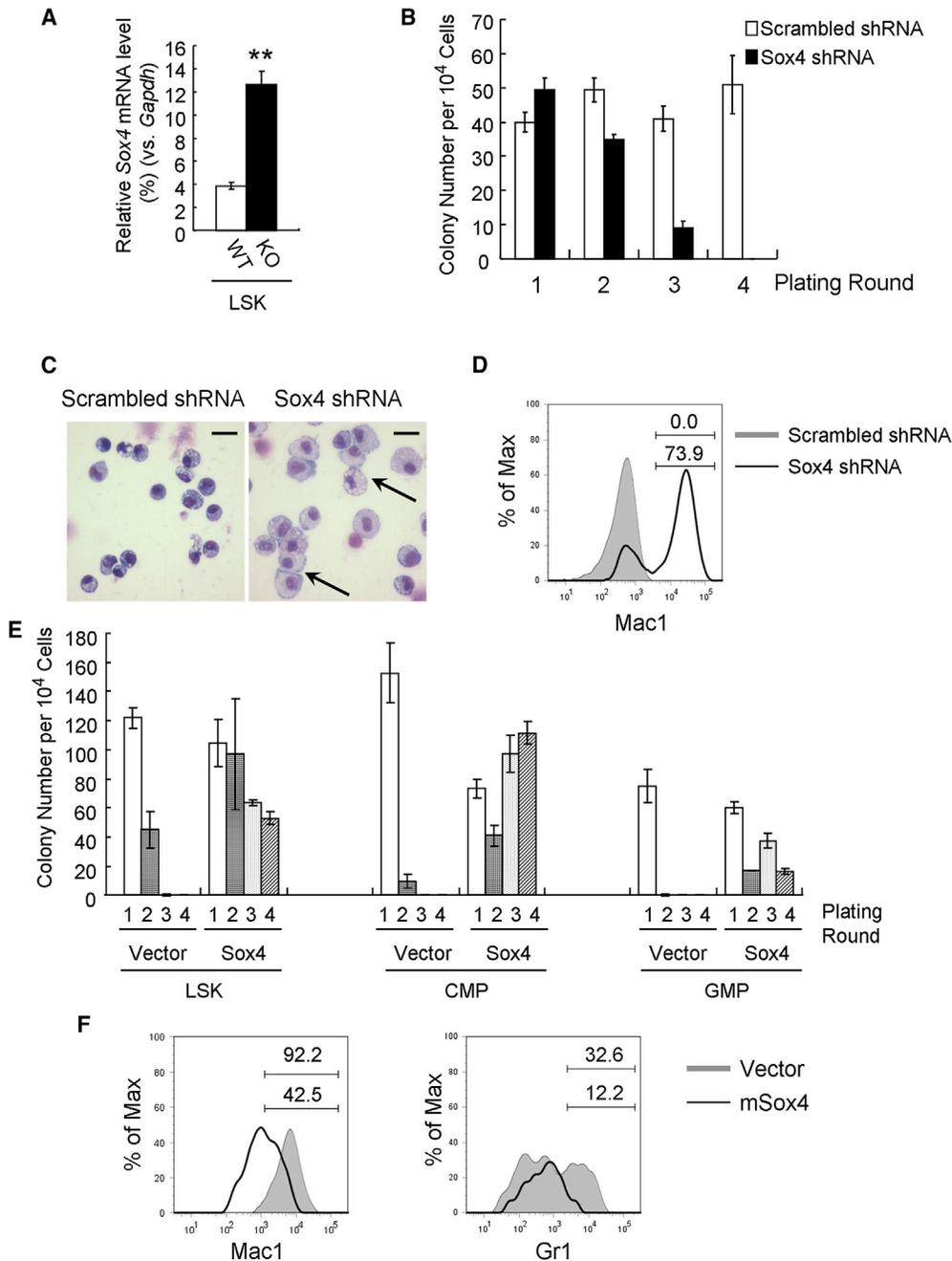
Collectively, our experiments identified the oncogene Sox4 as a factor mediating increased serial-replating ability and blocked differentiation of *Cebpa*-deficient progenitors.

### Sox4 Expression Is Directly Regulated by C/EBP $\alpha$

To understand the precise regulatory interaction between *Cebpa* and Sox4 during hematopoiesis, we examined their expression patterns during different stages of blood cell differentiation. Quantitative PCR (qPCR) analysis of adult wild-type mice revealed that expression of *Cebpa* increased from SLAMF<sup>+</sup>-HSC (enriched for HSCs), LSK cells (enriched for both HSCs and multipotent progenitors), common myeloid progenitors (CMP) to granulocyte-macrophage progenitors (GMP), and finally dropped in mature myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>) (Figure 2A). In contrast, Sox4 demonstrated an inverse expression pattern during myeloid lineage commitment, with highest expression in LSK, and gradually decreasing toward myeloid maturation with barely detectable levels in terminally differentiated myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>) (Figure 2A).

We next examined whether the reciprocal expression of *Cebpa* and Sox4 during myeloid commitment was due to a direct transcriptional repression of Sox4 by C/EBP $\alpha$ . Chromatin immunoprecipitation sequencing (ChIP-seq) analysis of murine bone marrow derived primary macrophages indeed revealed two adjacent C/EBP $\alpha$ -binding regions upstream of the murine Sox4 gene (Figure S2A). A major binding region was approximately 500 bp upstream of the Sox4 transcriptional start site (TSS) (-500 bp~+1 bp relative to TSS) and a minor one was observed within the 5' UTR region of Sox4 open reading frame (+1 bp~+200 bp relative to TSS) (Figures 2B and S2A). Luciferase assays using a reporter carrying a DNA fragment with the Sox4 promoter (-500 bp to 889 bp relative to the TSS) revealed that C/EBP $\alpha$  repressed the transcription activity from the Sox4 promoter in a dose-dependent manner (Figure 2B). Such repression was not due to the cellular saturation or toxicity of C/EBP $\alpha$  protein, since a parallel experiment indicated that within a similar dose range, C/EBP $\alpha$  transactivated a luciferase reporter carrying six consensus C/EBP $\alpha$  binding sites (Figure S2B).

Previous studies from our laboratory demonstrated that C/EBP $\alpha$  inhibits *Myc* and *Cebpg* transcription through interfering



**Figure 1. Sox4 Is Required for Abnormal Serial-Replating Ability and Myeloid Differentiation Block of *Cebpa* KO Stem/Progenitor Cells**

(A) *Sox4* expression was analyzed by qPCR in LSK cells ( $\text{Lin}^- \text{c-kit}^+ \text{Sca1}^+$ ) sorted from  $\text{Mx1-Cre}^- \text{Cebpa}^{\text{loxP/loxP}}$  (WT) and  $\text{Mx1-Cre}^- \text{Cebpa}^{\text{loxP/loxP}}$  (KO) mice 14 days after poly:I:C injection. Relative gene expression levels were determined as % *Gapdh*.

(B) *Cebpa* KO LSK cells were transduced with lentiviruses harboring either scrambled shRNAs or Sox4 shRNAs and replated in methylcellulose plus puromycin. The bar chart shows the colony number for four rounds.

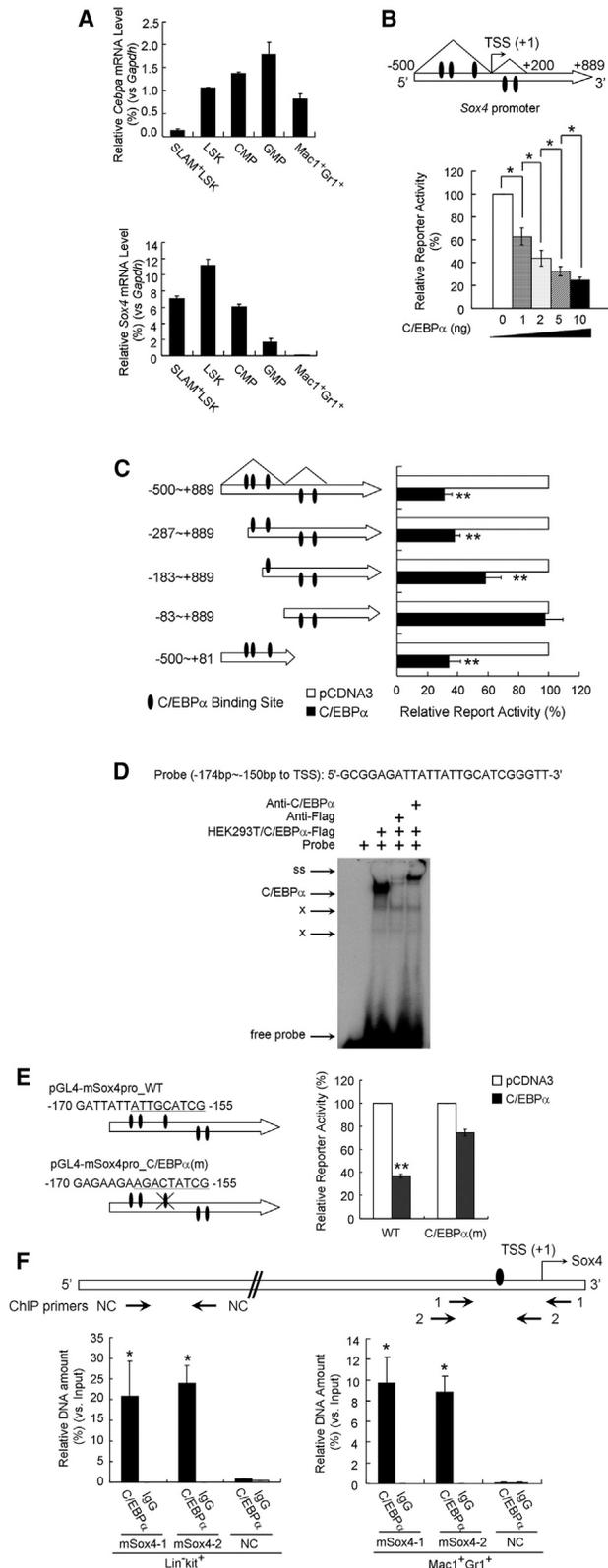
(C) Wright-Giemsa staining of cells from the first plating round in Figure 1B. Data are representative of three independent experiments. Macrophages are indicated with arrows. Scale bar, 20  $\mu$ m.

(D) Flow cytometry analysis of cells from the first plating round in Figure 1B. The shadow histogram indicates scrambled shRNA-infected cells, and the black line indicates Sox4 shRNA-infected cells. Percent cells are shown for the indicated gates. Data are representative of three independent experiments.

(E) Wild-type LSK, CMP ( $\text{Lin}^- \text{c-kit}^+ \text{Sca1}^- \text{CD34}^+ \text{FcRgII/III}^{\text{lo}}$ ), and GMP ( $\text{Lin}^- \text{c-kit}^+ \text{Sca1}^- \text{CD34}^+ \text{FcRgII/III}^{\text{high}}$ ) were transduced with an empty retrovirus (Vector) or a retrovirus expressing Sox4 (*Sox4*) and replated as Figure 1B.

(F) Wild-type LSK cells were transduced as in Figure 1E and grown in liquid culture plus puromycin and cytokines. Seven days later, Mac1 (left) and Gr1 (right) expression was analyzed by flow cytometry. Shadow histograms indicate cells infected with empty virus, and black lines indicate cells overexpressing Sox4. Percent cells are shown for the indicated gates. Data are representative of three independent experiments.

Error bars indicate the mean  $\pm$  SEM of three independent experiments. \*\*p < 0.001. See also Figure S1 and Table S1.



**Figure 2. Sox4 Is a Direct Target of Repression by C/EBP $\alpha$**

(A) qPCR analysis of *Cebpa* (top) and *Sox4* (bottom) levels in SLAMF1<sup>+</sup>-LSK, LSK, CMP, GMP, and myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>). Relative gene expression levels were determined as % *Gapdh*.

with E2F1 transactivation on these promoters (Alberich-Jordà et al., 2012; Johansen et al., 2001). We therefore first examined whether C/EBP $\alpha$  mediated *Sox4* repression through a similar mechanism. However, luciferase assays revealed that neither E2F1 nor other E2F family members were able to activate the *Sox4* promoter (Figures S2C and S2D), although they were all fully capable of transactivating a reporter carrying consensus E2F binding sites. Increasing the amount of E2F1 did not antagonize the repression of *Sox4* promoter by C/EBP $\alpha$  (Figure S2E). Furthermore, to determine whether C/EBP $\alpha$ -mediated repression of E2F1 activity suppressed *Sox4* mRNA expression, we transduced *Cebpa* KO LSK cells with either wild-type *Cebpa* or a series of *Cebpa* mutants (Porse et al., 2001). These experiments showed that C/EBP $\alpha$  mutants BRM2 and BRM5, which were defective in inhibiting E2F1-dependent transcription but still bound efficiently to C/EBP $\alpha$  sites, still repressed aberrant *Sox4* expression in *Cebpa* KO LSK cells at levels comparable to that of wild-type C/EBP $\alpha$  (Figure S2F). In these experiments, C/EBP $\alpha$  wild-type and mutant constructs were expressed at similar levels (Figure S2F). Taken together, these observations ruled out a role of E2F1 in the repression of *Sox4* by C/EBP $\alpha$ .

The *Sox4* promoter contains multiple C/EBP $\alpha$  binding sites. We therefore asked whether C/EBP $\alpha$  repressed the *Sox4* promoter through direct DNA binding. We generated a series of truncation mutants of the *Sox4* promoter corresponding to the location of the putative C/EBP $\alpha$  binding sites (Figures 2C and S2G). Luciferase assays with these truncation mutants revealed that a fragment between -183 bp and -83 bp upstream of the

(B) Top: schematic representation of the *Sox4* promoter extending from -500 bp to +889 bp relative to TSS of murine *Sox4* gene. Triangles denote the ChIP-Seq C/EBP $\alpha$  binding peaks as shown in Figure S2A. Bottom: HEK293 cells were transfected with the *Sox4* promoter reporter and increasing amounts of pCDNA3-C/EBP $\alpha$ . Luciferase activities were normalized to Renilla activities and presented as percentage of empty vector.

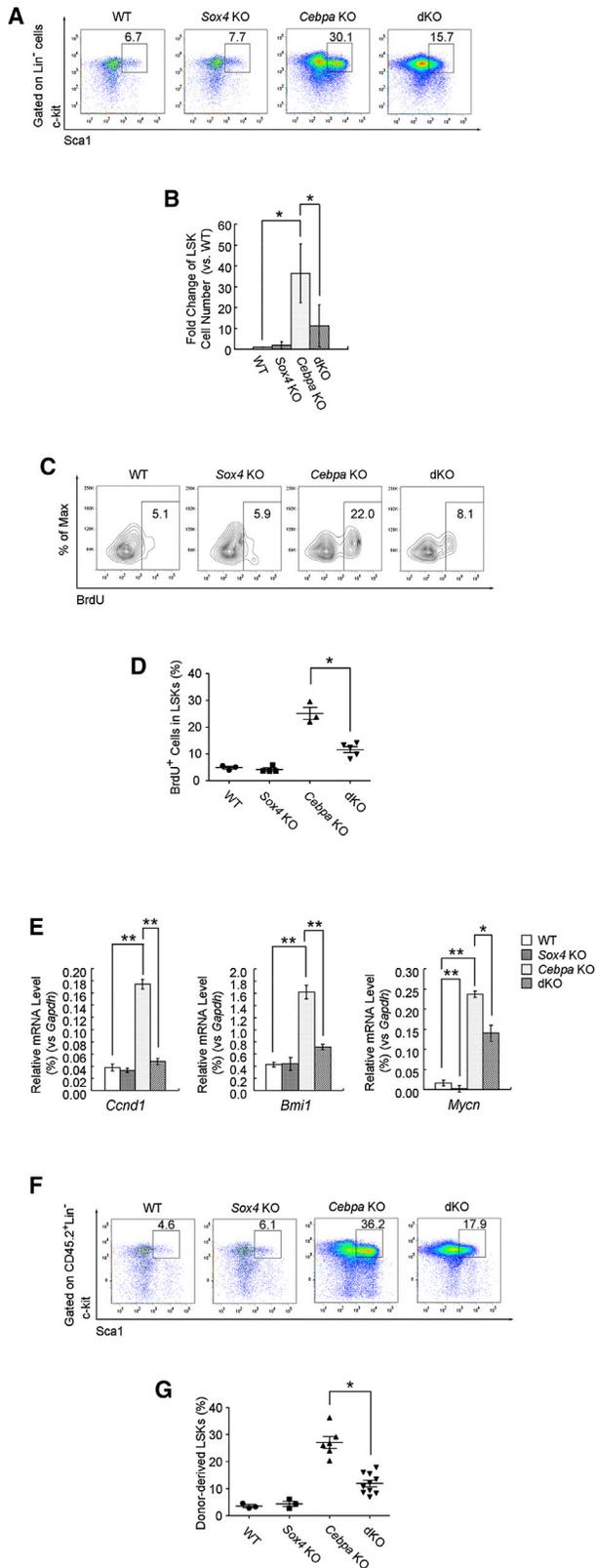
(C) Left: schematic representation of truncated constructs of the *Sox4* promoter reporter. Right: luciferase values of reporter constructs with an empty vector (white bar) or C/EBP $\alpha$ -expressing constructs (black bars).

(D) Electrophoretic mobility shift assay using a probe containing a potential C/EBP $\alpha$  binding site from the *Sox4* promoter was performed with nuclear extracts from HEK293T cells overexpressing Flag-tagged C/EBP $\alpha$  (HEK293T/C/EBP $\alpha$ -Flag) and Flag antibody (anti-Flag) or C/EBP $\alpha$  antibody (anti-C/EBP $\alpha$ ). C/EBP $\alpha$ , C/EBP $\alpha$  complex; ss, supershifted band; x, migration of nonspecific protein complexes binding to the probes.

(E) Left: schematic representation of luciferase reporters carrying wild-type *Sox4* promoter (WT) or *Sox4* promoter with mutated C/EBP $\alpha$  binding sites [C/EBP $\alpha$ (m)] marked with a cross. The putative C/EBP $\alpha$  binding sites and the mutated version are underlined. Right: luciferase values of reporter constructs carrying wild-type or site-mutated *Sox4* promoter with an empty vector (white bar) or C/EBP $\alpha$ -expressing constructs (black bars).

(F) ChIP-qPCR analysis of specific binding of C/EBP $\alpha$  to the upstream proximal region of murine *Sox4* gene in wild-type primary stem/progenitor cells (Lin<sup>-</sup>c-kit<sup>+</sup>) or myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>). Top: schematic of *Sox4* gene with three qPCR primer sets (arrows below). Oval denotes the C/EBP $\alpha$  binding site identified in Figure 2E. Primer sets "mSox4-1" and "mSox4-2" were used to amplify regions -171 to +77 bp and -215 to -16 bp (relative to the TSS) of *Sox4* gene and primer set "NC" were used as a negative control. Bottom: relative DNA enrichment was measured by qPCR and is presented as percentage of input chromatin. Black bars represent ChIP-qPCR signals from the C/EBP $\alpha$  pull-down (C/EBP $\alpha$ ) while white bars are from immunoglobulin G control (IgG).

Error bars indicate the mean  $\pm$  SEM of three independent experiments. \*p < 0.01; \*\*p < 0.001. See also Figure S2.



**Figure 3. Sox4 Mediates the Aberrant Progenitor Expansion following Loss of C/EBP $\alpha$**

(A) Flow cytometry analysis of LSK frequency (Lin<sup>-</sup>c-kit<sup>+</sup>Sca1<sup>+</sup>) in bone marrow of Mx1-Cre<sup>-</sup>Sox4<sup>loxP/loxP</sup>Cebpa<sup>loxP/loxP</sup> (WT), Mx1-Cre<sup>+</sup>Sox4<sup>loxP/loxP</sup>

Sox4 TSS was essential for C/EBP $\alpha$ -mediated repression (Figure 2C). EMSA combined with supershift assay and competitor assay further revealed specific association of C/EBP $\alpha$  protein with potential *cis* elements at -174 bp to -150 bp upstream of the Sox4 TSS (Figures 2D and S2H). Mutation of a highly evolutionally conserved C/EBP $\alpha$  binding site in this region resulted in loss of repression of the Sox4 promoter by C/EBP $\alpha$  (Figures 2E and S2I). To further verify direct binding of C/EBP $\alpha$  to the identified region within the Sox4 promoter, we performed ChIP assays on primary cells isolated from adult wild-type mice. Using two sets of primers amplifying the -183~-83 bp region (relative to the TSS), we confirmed binding of endogenous C/EBP $\alpha$  to this region in both a stem/progenitor cell enriched population (Lin<sup>-</sup>c-kit<sup>+</sup>) and mature myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>) (Figure 2F). Moreover, the binding of C/EBP $\alpha$  was specific to this region, since no enrichment of a distal region of the Sox4 gene was observed. In summary, our data demonstrate that C/EBP $\alpha$  negatively regulates Sox4 transcription via direct DNA binding.

### Sox4 Mediates the Abnormal Expansion of LSK Cells in the Absence of C/EBP $\alpha$ In Vivo

To functionally dissect the role in which dysregulated Sox4 expression is contributing to the Cebpa knockout phenotype in vivo, we designed an inducible Sox4/Cebpa double-knockout model. We generated conditional Sox4-deficient mice alone (Mx1-Cre<sup>+</sup>Sox4<sup>loxP/loxP</sup>; referred to as Sox4 KO following Cre-mediated deletion) (Penzo-Méndez et al., 2007) and inducible Sox4/Cebpa double knockouts (Mx1-Cre<sup>+</sup>Sox4<sup>loxP/loxP</sup>Cebpa<sup>loxP/loxP</sup>; referred to as dKO after deletion) (Figure S3A). Flow cytometry analysis of bone marrow from mice 21 days after induction of Cre recombinase-mediated deletion with poly:C demonstrated that Cebpa KO mice exhibited a remarkable expansion of LSK cells, both in terms of percentage of total bone marrow (4-fold; Figure 3A) and in absolute number (18-fold; Figure 3B). LSK cells in Sox4 KO mice did not display any gross phenotypical differences compared to wild-type LSKs (Figures 3A and 3B). However, in concordance with the opposing

Cebpa<sup>+/+</sup> (Sox4 KO), Mx1-Cre<sup>+</sup>Sox4<sup>+/+</sup>Cebpa<sup>loxP/loxP</sup> (Cebpa KO), and Mx1-Cre<sup>+</sup>Sox4<sup>loxP/loxP</sup>Cebpa<sup>loxP/loxP</sup> (dKO) mice 21 days after poly:C treatment. Representative fluorescence-activated cell sorting (FACS) plots of six independent experiments are shown.

(B) Comparison of LSK cell number in the bone marrow of WT, Sox4 KO, Cebpa KO, and dKO mice. Numbers are presented as fold change relative to the WT.

(C) Flow cytometry of BrdU incorporation in LSK cells of single and double knockouts. Percent BrdU<sup>+</sup> cells are shown for the indicated gates. One representative experiment of three is shown.

(D) Quantitative analysis of BrdU incorporation in LSK cells of single and double knockouts.

(E) qPCR analysis of Ccnd1, Bmi1, and Mycn expression in LSK cells from single and double knockouts. Relative mRNA levels were determined as % Gapdh.

(F) Flow cytometry analysis of donor-derived LSK frequency (CD45.2<sup>+</sup>Lin<sup>-</sup>c-kit<sup>+</sup>Sca1<sup>+</sup>) in the bone marrow of recipient mice transplanted with LSK cells from mice with indicated phenotypes. Analysis was performed 4 months posttransplantation. One representative recipient of ten recipients is shown.

(G) Quantitative analysis of donor-derived LSK frequency in the bone marrow of recipient mice.

Error bars indicate the mean  $\pm$  SEM of three independent experiments. \*p < 0.01; \*\*p < 0.001. See also Figure S3.

roles of Sox4 and C/EBP $\alpha$ , LSK expansion of *Cebpa* KO mice was partially rescued in dKO mice with a 2-fold proportional (Figure 3A) and a 3-fold absolute reduction in LSK numbers (Figure 3B).

We next assessed the impact of C/EBP $\alpha$  and Sox4 on LSK proliferation utilizing 5-bromodeoxyuridine (BrdU) incorporation in vivo. The proliferative fraction of LSKs was remarkably enhanced (~4-fold) in *Cebpa* KO mice but was lower in the dKO mice, comparable to wild-type (Figures 3C and 3D). Gene expression analysis further identified several candidate genes potentially mediating the differences in proliferation between *Cebpa* KO and dKO LSK cells (Figures 3E, S3B, and S3C). In *Cebpa* KO LSK cells, among the major cell-cycle regulators, *Cnnd1* was upregulated whereas *p21* and *p57* (Pietras et al., 2011) were downregulated compared to the wild-type and Sox4 KO counterparts. Increased expression was also observed for *Bmi1* and *Mycn*, both of which are implicated in HSC proliferation and self-renewal in the absence of C/EBP $\alpha$  (Ye et al., 2013; Zhang et al., 2004). However, the additional disruption of Sox4 significantly impeded such deregulation in the dKO LSK cells, providing a putative molecular mechanism of how Sox4 induce LSK expansion in *Cebpa* KO mice.

To investigate if Sox4 modulates the *Cebpa* KO HSC phenotype in a cell-intrinsic manner, LSK cells from wild-type mice or conditional deficient mice (CD45.2<sup>+</sup>) 21 days post-polyI:C injection were transplanted into lethally irradiated recipients (CD45.1<sup>+</sup>). Four months after transplantation, flow cytometry confirmed aberrant LSK expansion in recipients of *Cebpa* KO LSK cells and a partial rescue in recipients of dKO LSK cells (Figures 3F and 3G). Genotyping of LSK cells isolated from the conditional deficient mice or donor-derived LSK cells isolated from the recipients verified complete excision of the *Cebpa* and Sox4 alleles (Figures S3D–S3F). These data demonstrate an HSC cell-intrinsic role for Sox4 on proliferation induced by loss of C/EBP $\alpha$ .

### Targeting Sox4 Restores Myeloid Differentiation of *Cebpa* KO Cells In Vivo

Knockdown of Sox4 enables *Cebpa* KO LSK cells to undergo myeloid differentiation in vitro (Figures 1C and 1D). To test the effect of Sox4 on the impaired myeloid differentiation of *Cebpa* KO LSKs in vivo, we examined whether the disruption of Sox4 could restore myeloid differentiation in *Cebpa* KO mice. Flow cytometry analysis of bone marrow cells from wild-type or Sox4 KO mice 21-day post-polyI:C injection revealed a similar proportion of Mac1<sup>+</sup>Gr1<sup>+</sup> myeloid cells, whereas in *Cebpa* KO mice the myeloid population was barely detectable (Figure S4A; deletion efficiency shown in Figure S4B). Although a Mac1<sup>lo/+</sup>Gr1<sup>-</sup> population appeared in the dKO mice, they were morphologically immature blasts, as with the Mac1<sup>-</sup>Gr1<sup>-</sup> population (data not shown).

In contrast, dKO LSK cells were able to give rise to a significant population of mature myeloid (Mac1<sup>+</sup>Gr1<sup>+</sup>) 4 months following transplantation (Figures 4A and 4B). Cytology of sorted donor-derived Mac1<sup>+</sup>Gr1<sup>+</sup> cells confirmed the presence of morphologically differentiated mature myeloid cells, similar to wild-type cells (Figure 4C). Genotyping analysis confirmed complete excision of *Cebpa* and Sox4 in the donor-derived Mac1<sup>+</sup>Gr1<sup>+</sup> population (Figure S4C).

We also transplanted *Cebpa* KO LSK cells that had been transduced with GFP containing retroviruses expressing either Sox4 shRNAs (shRNA 672 and shRNA 928) or scrambled shRNA (Figures S4D and S4E). Mature myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>) were only present in the GFP<sup>+</sup> donor population of mice receiving Sox4 shRNA-transduced *Cebpa* KO LSK cells (Figures 4D and 4E). Genotyping of donor-derived GFP<sup>+</sup>Mac1<sup>+</sup>Gr1<sup>+</sup> cells confirmed a complete excision of *Cebpa* allele (Figure S4F). Taken together, these data demonstrate that the disruption of Sox4 restores myeloid programming of *Cebpa* KO LSKs in vivo, further substantiating the notion that Sox4 is epistatic to *Cebpa*.

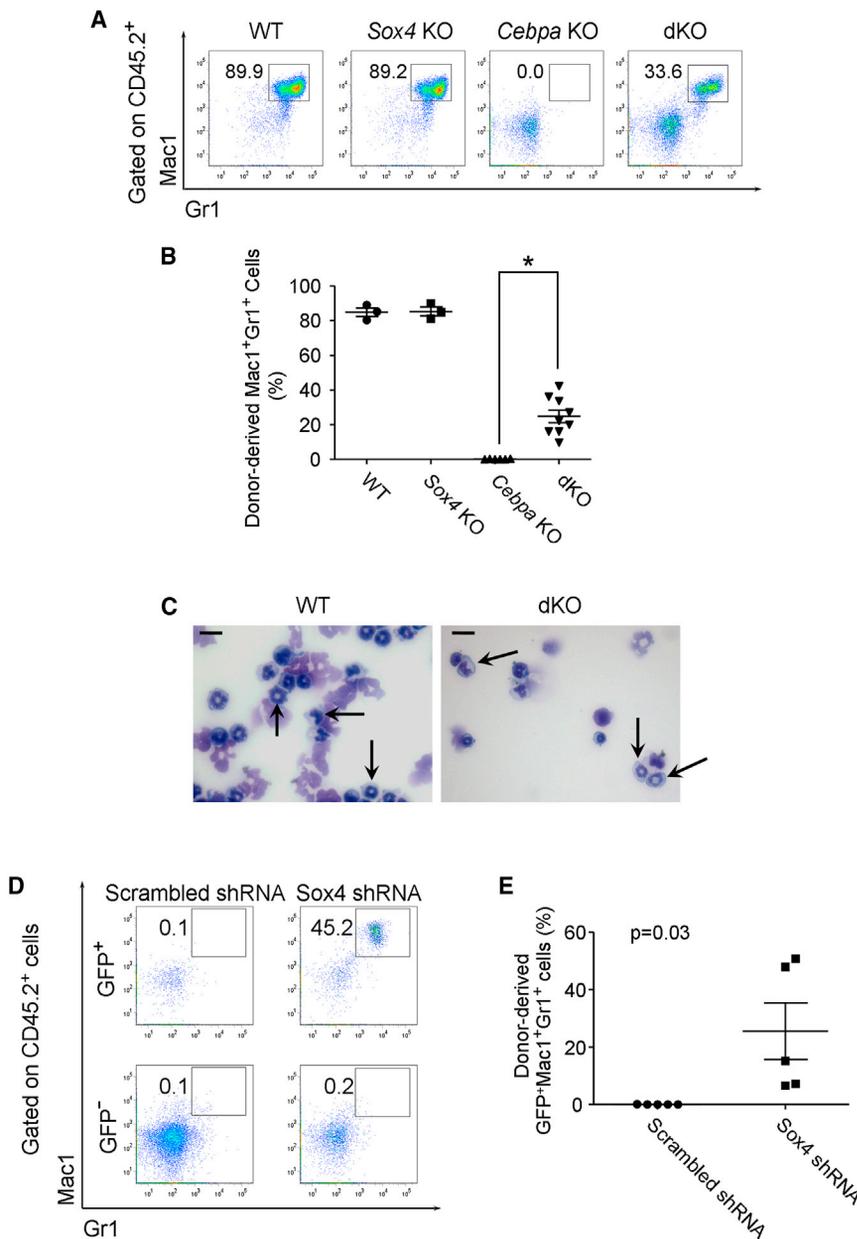
### Sox4 Is Crucial for the Oncogenic Activity of Mutated C/EBP $\alpha$

In a murine model of human AML expressing biallelic *CEBPA* mutations (K/L mice), only a distinct phenotypically defined group of cells (leukemia-initiating cells, or LICs) gives rise to leukemia (Bereshchenko et al., 2009). To examine the role of Sox4 in those cells, we examined its expression in cellular subpopulations of leukemic K/L mice and observed that Sox4 mRNA was exclusively increased in the LIC fraction (fraction II: Lin<sup>lo</sup>Sca1<sup>-</sup>Mac1<sup>lo</sup>c-kit<sup>+</sup>), whereas it appeared to be decreased or absent in the other fractions (fraction I: Lin<sup>lo</sup>Sca1<sup>+</sup>; fraction III: Lin<sup>lo</sup>Sca1<sup>-</sup>Mac1<sup>+</sup>c-kit<sup>lo/-</sup>) (Figures 5A and S5A). It had previously been shown that mice transplanted with Sox4-transduced 5-FU treated bone marrow cells develop myeloid leukemia (Du et al., 2005). We confirmed this result by transplanting Sox4-transduced LSK cells and showed that mice died with a median latency of 4 months (Figures S5B and S5C). Flow cytometry revealed that Sox4-induced leukemia cells shared key phenotypical features with the model for biallelic *Cebpa* mutants (K/L leukemia; Figure 5B). Most importantly, as described for K/L leukemia, in Sox4-induced leukemia, only cells from fraction II gave rise to leukemia in recipient mice (Figures 5C and 5D).

To determine if Sox4 contributes to leukemogenesis not only in *Cebpa*-deficient AML but also AML with mutated *Cebpa*, we examined the effect of Sox4 ablation on the self-renewal ability of K/L LICs. As with *Cebpa* KO cells, LICs from leukemic K/L mice exhibited an enhanced replating ability (Figure S5D). In addition, sorted K/L LIC cells were lentivirally transduced with either scrambled shRNAs ( $\phi$ -1 and  $\phi$ -2) or Sox4 shRNAs (#80, #81, and #82) (Figure S5E). Cells transduced with Sox4 shRNA exhibited decreased ability to undergo serial replating and barely gave rise to colonies in the second round of plating, while the cells transduced with scrambled shRNA still gave rise to colonies even after the third round of plating (Figure 5E), demonstrating that increased Sox4 levels are crucial for the self-renewal properties of *Cebpa* mutant LICs.

### A Common Gene Expression Signature Defines *Cebpa* KO, Sox4 AML, and K/L AML

Hierarchical clustering analysis of gene expression profiles further revealed that the LIC populations from K/L and Sox4 murine leukemia models clustered together and were distinct from their immunophenotypic counterparts (fraction II from wild-type mice [WT FrII]) as well as from the leukemic stem cell-enriched L-GMP from the MLL-AF9 and MOZ-TIF2 murine AML models (MF9 L-GMP and MT2 L-GMP) (Krivtsov et al.,



**Figure 4. Loss of Sox4 Restores Myeloid Differentiation of Cebpa-deficient LSK Cells In Vivo**

(A) Flow cytometry analysis of donor-derived myeloid cell (CD45.2<sup>+</sup>Mac1<sup>+</sup>Gr1<sup>+</sup>) in the bone marrow of recipient mice transplanted with LSK cells from mice with indicated phenotypes. Analysis was performed 4 months posttransplantation. Representative FACS plots of six independent experiments are shown.

(B) Quantitative analysis of donor-derived myeloid cell frequency in the bone marrow of recipient mice. Error bars indicate the mean  $\pm$  SEM of six independent experiments. \* $p < 0.01$ .

(C) Wright-Giemsa staining of donor-derived Mac1<sup>+</sup>Gr1<sup>+</sup> cells in the bone marrow of recipient mice. Granulocytic cells are indicated with arrows. Scale bar, 10  $\mu$ m.

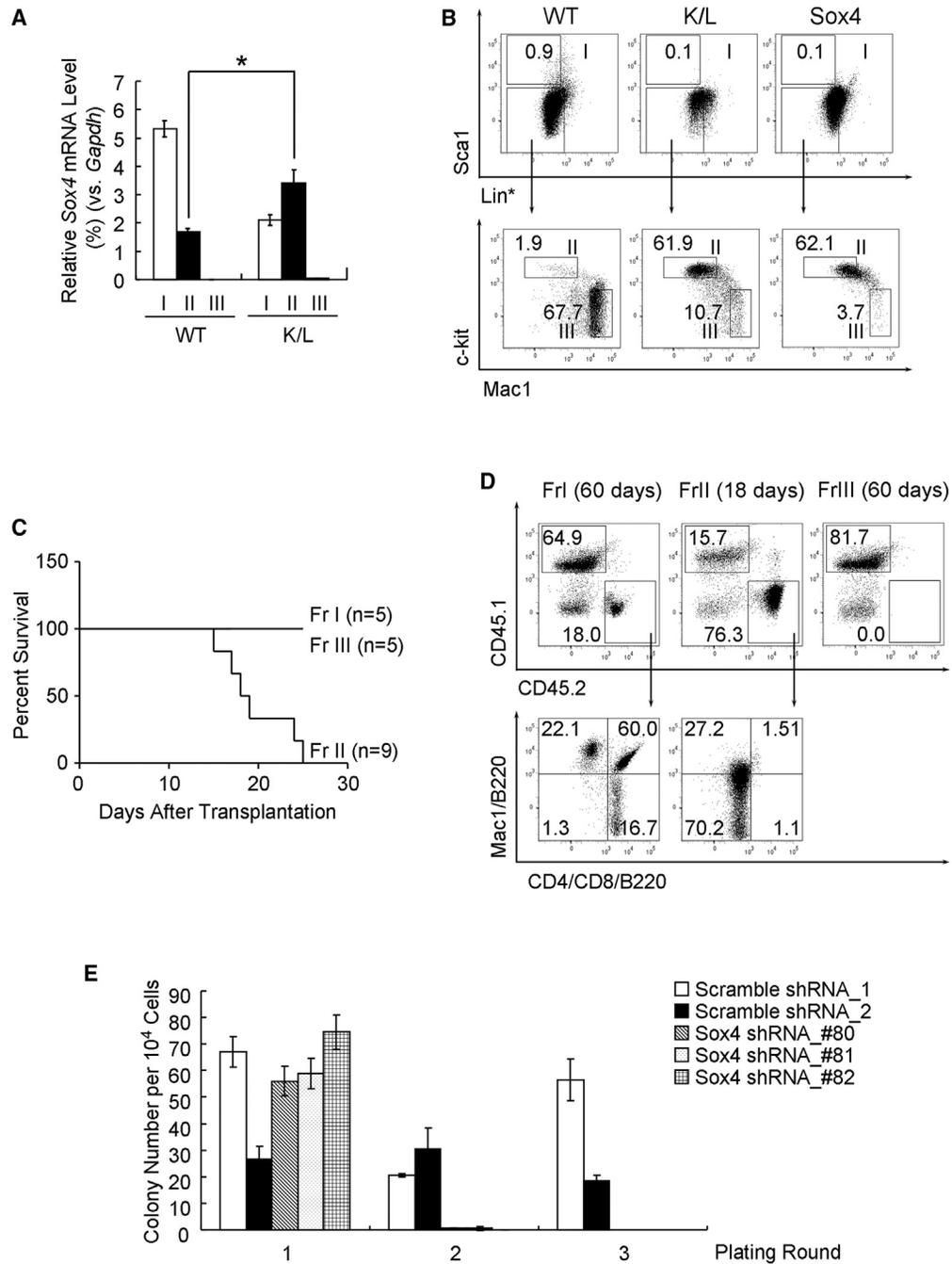
(D) Flow cytometry analysis of donor-derived myeloid cells in the bone marrow of mice receiving Cebpa KO LSK cells transduced with scrambled shRNA or mSox4 shRNA. The percentages of Mac1<sup>+</sup>Gr1<sup>+</sup> population within donor-derived (CD45.2<sup>+</sup>) GFP<sup>+</sup> and GFP<sup>-</sup> cells are indicated. Representative FACS plots of five independent experiments are shown.

(E) Quantitative analysis of donor-derived shRNA-transduced myeloid cell frequency (CD45.2<sup>+</sup>GFP<sup>+</sup>Mac1<sup>+</sup>Gr1<sup>+</sup>) in the bone marrow of recipient mice. Error bars indicate the mean  $\pm$  SEM of five independent experiments.  $p = 0.03$ .

See also Figure S4.

2006; Kvinlaug et al., 2011) (Figure 6A). This suggested the existence of a common leukemogenic gene expression signature characterizing Sox4 LIC and mutated C/EBP $\alpha$  LIC. Intriguingly, Cebpa-deficient HSCs also clustered together with LICs from K/L and Sox4 AML by sharing similar expression pattern of 686 genes, indicating underlying similarities among these three populations. We next investigated whether our results from the mouse models were of relevance to human AML. Previous unsupervised hierarchical gene expression clustering of leukemic blasts from 562 AML patients identified a homogenous group (cluster 4) in which both CEBPA-silenced and biallelic-mutated cases clustered together (Valk et al., 2004). Intriguingly, SOX4 expression was significantly increased within the very same group (Figure 6B), further confirming that SOX4 and mutated/silenced CEBPA are within one oncogenic pathway.

Supervised clustering analysis identified a set of genes that were highly expressed in the wild-type HSC and LIC populations from Sox4 and K/L leukemias but exhibited decreased expression in committed progenitors (referred to as “self-renewal-associated LIC signature”) (Figure 6C; Table S2). We next compared this signature with a previously reported L-GMP self-renewal-associated signature (Krivtsov et al., 2006). This revealed that only a very limited set of genes crucial for acquisition of self-renewal properties in L-GMP (24 out of 420) are present in the LICs from Sox4 and K/L leukemias (24 out of 80) and the majority of this self-renewal-associated LIC signature (56 genes) is specific to the Sox4 and K/L leukemias (referred to as “specific C/EBP $\alpha$ -Sox4 LIC signature”) (Figures 6D and S6A; Tables S3, S4, and S5). Within the specific C/EBP $\alpha$ -Sox4 LIC signature, two key elements of TGF $\beta$ -FOXO signaling that have been implicated in maintaining LICs in chronic myeloid leukemia (Naka et al., 2010), Foxo3a and type II TGF $\beta$  receptor, are specifically upregulated (Table S4). In addition, a set of genes whose upregulation are functionally important for L-GMP transformation by MLL-AF9 and MOZ-TIF2 (Krivtsov et al., 2006; Kvinlaug et al., 2011) were either downregulated or unaltered in LICs, including TCF4, CDKN1C, HoxA5, HoxA10, and Lmo2 (Figure 6E; Table S5). These observations suggested a distinct transformation



**Figure 5. The LIC in Sox4-Induced AML Resembles the LIC in a Murine C/EBP $\alpha$  Mutant AML Model**

(A) qPCR analysis of Sox4 expression in donor-derived fractions I (Lin<sup>+</sup>Sca1<sup>+</sup>), II (Lin<sup>+</sup>Sca1<sup>-</sup>Mac1<sup>+</sup>c-kit<sup>+</sup>) and III (Lin<sup>+</sup>Sca1<sup>-</sup>Mac1<sup>+</sup>c-kit<sup>+/−</sup>) from wild-type control (WT) or leukemic K/L mice. Relative expression levels were determined as % *Gapdh* and presented as the mean  $\pm$  SEM of three independent experiments. \**p* < 0.01.

(B) Flow cytometry analysis of mice transplanted with bone marrow from wild-type mice, K/L mice, or with LSK cells overexpressing Sox4. Total bone marrow cells were subjected to lineage depletion excluding Mac1 and Gr1 (Lin<sup>+</sup>) followed by FACS analysis. Representative FACS plots of three independent experiments are shown.

(C) Kaplan-Meier survival analysis of mice that received fractions I, II, and III from Sox4 leukemic mice. Number of recipients for each fraction (n) is indicated.

(D) Flow cytometry analysis of mice that received fractions I, II, and III from Sox4 leukemic mice. Donor-derived cells (CD45.2<sup>+</sup>) in the bone marrow were analyzed at indicated time points. Representative FACS plots of three independent experiments are shown.

(E) Donor-derived fractions II from leukemic K/L mice were transduced with either scrambled shRNAs or Sox4 shRNAs and serially replated. The bar chart shows the mean  $\pm$  SEM of colony number of three independent experiments.

See also Figure S5.

programming resulting from the underlying initiating mutation of L-GMPs and the LICs from Sox4 and K/L leukemias. On the other hand, *Mef2C* (a cooperating oncogene in leukemogenesis) (Du et al., 2005; Krivtsov et al., 2006) and prostaglandinendoperoxide synthase 1 (*Ptgs1*) (a key enzyme in Wnt/ $\beta$ -catenin pathway-related prostaglandin synthesis) (Castellone et al., 2005; Goessling et al., 2009) are upregulated in the self-renewal-associated signature of both LICs and L-GMPs (Wang et al., 2010) (Table S3). These observations suggested that certain molecular pathways were commonly required for leukemic transformation regardless of the underlying initiating mutation of L-GMPs and LICs.

Gene set enrichment analysis (GSEA) further demonstrated that the specific C/EBP $\alpha$ -Sox4 LIC signature was highly enriched in gene expression profiles from human AML samples with biallelic *CEBPA* mutation or silent *CEBPA*, but not in those of human AML samples with wild-type *CEBPA* (Figures 6F and S6B), supporting the relevance of the LIC signature identified from the murine AML model in human AML.

#### Targeting SOX4 Rescues Myeloid Differentiation in Human CEBPA Mutant AML

To functionally evaluate the relevance of our findings in murine models to human leukemia, we first determined the ability of leukemogenic human *CEBPA* mutations to repress Sox4 mRNA expression. We analyzed Sox4 mRNA levels after transducing *Cebpa* KO LSK cells with viruses encoding either wild-type human CEBPA or a series of mutant proteins identified from human AML patients (Pabst et al., 2001). Interestingly, CEBPA mutant proteins 22N, 22C, and 10, which exhibit an increased ratio of 30 kDa to 42 kDa CEBPA and are defective in DNA binding, failed to repress Sox4, in contrast to wild-type human CEBPA (Figure 7A). Of note, CEBPA mutant 128, which exhibited the wild-type 30 kDa to 42 kDa ratio and maintained the ability to bind efficiently to C/EBP $\alpha$  binding sites, was still capable of repressing Sox4 expression at levels comparable to those of wild-type CEBPA (Figure 7A). In all of these experiments, CEBPA wild-type and mutant constructs were expressed at similar levels (Figure S7A). We next analyzed the effects of individual CEBPA mutant proteins on Sox4 promoter activity using reporter assays. Consistent with the findings of Sox4 mRNA expression analysis (Figure 7A), DNA-binding defective CEBPA mutants (22N, 22C, and 10) failed to repress Sox4 promoter activity (Figure 7B) when expressed at comparable protein levels (Figure S7B). Taken together, these data indicate that dysfunctional CEBPA proteins known to underlay the development of human AML fail to suppress SOX4.

To determine if targeting SOX4 could revert human CEBPA biallelic mutant AML, we next studied the effect of SOX4 shRNA knockdown in such cells in vivo. Two CEBPA mutant AML patient samples were transduced with lentiviruses carrying a GFP marker and either SOX4 shRNA or a scrambled shRNA (Figure S7C). Transduced cells were then transplanted into sublethally irradiated NSG mice. Flow cytometry of bone marrow cells 5 weeks after transplantation revealed pronounced neutrophilic differentiation as evidenced by CD15 surface marker expression in the GFP<sup>+</sup> population of SOX4 shRNA-transduced human patient cells (Figure 7C). Furthermore, cytology of sorted GFP<sup>+</sup> human cells (CD45<sup>+</sup>GFP<sup>+</sup>) showed the presence of mature

neutrophils in SOX4 shRNA-transduced cells only (Figure 7D). Taken together, downregulation of SOX4 in human CEBPA mutant AML samples partially restores granulocytic differentiation in vivo, thus providing strong experimental evidence that enhanced SOX4 is a major contributor to the leukemic phenotype in human leukemia with dysfunctional CEBPA.

## DISCUSSION

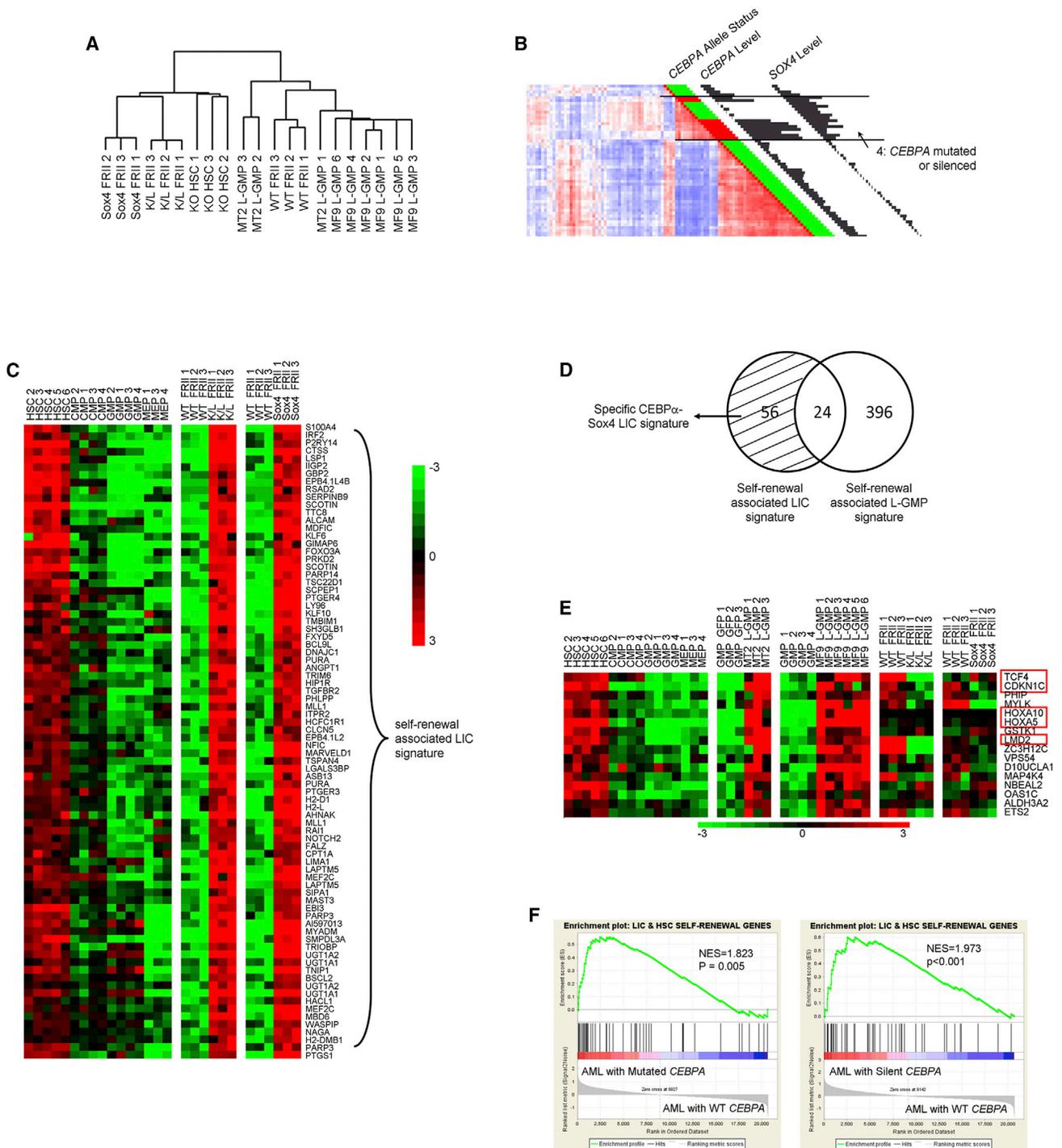
### Sox4 Is a Central Mediator of AML with Defective C/EBP $\alpha$

In this study, we demonstrate that Sox4 contributes to the leukemic phenotype of C/EBP $\alpha$  mutant AML in murine models as well as in human AML. C/EBP $\alpha$  is both a transcription activator and repressor, depending on the target genes and cell context. As a tumor suppressor, C/EBP $\alpha$  modulates cell growth through repression of oncogenes *Myc*, *Mycn*, and *Bmi1* (Johansen et al., 2001; Ye et al., 2013; Zhang et al., 2004). Here, we identified and validated Sox4 as a C/EBP $\alpha$  downstream target.

Previous studies reported that Sox4 is upregulated in L-GMPs from murine MOZ-TIF2, AML1-ETO, and NUP98-HOXA9 AML models in an immediate leukemia initiation signature (Kvinlaug et al., 2011). However, among 16 groups of AML patients classified on the basis of molecular signatures (Valk et al., 2004), SOX4 is only upregulated in AML subtypes carrying biallelic mutated and silent *CEBPA* or internal tandem duplication mutation of FLT3 (FLT3-ITD). Previous observations that FLT3-ITD blocks C/EBP $\alpha$  through transcriptional repression or posttranslational modification (Radomska et al., 2006; Zheng et al., 2004), and that CEBPA biallelic mutant AML patients have a lower coincidence of FLT3-ITD mutations (Green et al., 2010), suggest that aberrant levels of SOX4 in FLT3-ITD subtypes is likely due to CEBPA deficiency.

Previous studies based on knockin mice demonstrated that biallelic *Cebpa* mutations induce leukemia by expanding stem/progenitor cells with development of a transformed progenitor pool blocked in myeloid differentiation (Bereshchenko et al., 2009). Here, we demonstrate that dysregulated Sox4 contributes to these leukemia phenotypes. Furthermore, the expansion of a highly proliferative premalignant LSK compartment caused by DNA-binding defective *Cebpa* mutations is alleviated in dKO mice, suggesting that enhanced Sox4 levels are crucial to maintain this transformable pool. The exclusive increase of Sox4 level in C/EBP $\alpha$  mutant LICs also suggests its potential role in LIC function. Unfortunately, complete loss of LIC leukemogenic potential after even a very short period of liquid culture precludes addressing this issue by viral-mediated RNAi knockdown followed by transplantation (Claus Nerlov and Jorg Cammenga, personal communication [and confirmed by us]). However, serial-replating assays, a surrogate for self-renewal potential that has been validated with transplantation assays in a number of studies (Huntly et al., 2004; Lavau et al., 1997; Moran-Crusio et al., 2011), suggest a causative role for Sox4 in the acquisition of LIC self-renewal.

Although myelopoiesis in Sox4 KO mice appears normal, Sox4 can inhibit myeloid differentiation in 32DCI3 cells (Boyd et al., 2006) and primary stem/progenitor cells treated with cytokines, consistent with the progressive decrease in Sox4 expression



**Figure 6. Gene Expression Signatures of Sox4 and K/L LICs Are Distinct from MLL-AF9 and MOZ-TIF2 L-GMPs**

(A) Hierarchical clustering of fraction II population from leukemic K/L and Sox4 recipients and their wild-type counterparts (WT FRII), L-GMP population from leukemic MLL-AF9 and MOZ-TIF2 recipients, and HSCs from *Cebpa* KO mice (KO HSC).

(B) Pairwise correlations between *CEBPA* and *SOX4* in gene expression profiles of human AML patients. The bar and histograms next to each sample indicate *CEBPA* allele status (mutations are shown in red and wild-type are shown in green) and *CEBPA* and *SOX4* mRNA expression levels. The black lines designate cluster 4 AML patients with either mutated or silent *CEBPA* allele(s).

(C) Heatmap shows the 80-gene self-renewal-associated LIC signature that is specifically upregulated in the wild-type HSCs (versus committed progenitors) as well as Sox4 and K/L LICs (versus WT FRII) (fold change cutoff = 1.5; p value cutoff = 0.05).

(D) The Venn diagram shows the overlap between the “self-renewal-associated LIC signature” (80 genes) and the “self-renewal-associated L-GMP signature” (420 genes) in a single two-class comparison. Shadow area indicates the “specific C/EBP $\alpha$ -Sox4 LIC signature” (56 genes).

(legend continued on next page)

during normal myeloid lineage commitment and terminal differentiation. Our finding that suppression of *Sox4* partially restored myeloid differentiation of dKO LSK and human CEBPA mutant AML samples further demonstrates that enhanced *Sox4* expression blocks myeloid differentiation in this AML subtype. Collectively, we demonstrate a central role of *Sox4* in C/EBP $\alpha$  mutant induced leukemogenesis.

Our findings in murine C/EBP $\alpha$  mutant AML suggest that abrogation of SOX4 may also affect the function of LICs from human CEBPA mutated or silent AML. However, a lack of phenotypic definition of human CEBPA mutated/silent LICs omitted further investigations, since LICs cannot be universally defined for all AML samples on the basis of the CD34<sup>+</sup>CD38<sup>-</sup> phenotype (Eppert et al., 2011; Sarry et al., 2011; Taussig et al., 2008, 2010).

### Sox4 and C/EBP $\alpha$ Mutant Leukemias Share a Unique Gene Expression Signature

The overlap between gene expression patterns of Sox4 LICs, K/L LICs, and *Cebpa* KO HSCs suggests the existence of a transformation signature specific to Sox4-mediated C/EBP $\alpha$ -defective LICs that is distinct from that of L-GMPs from MLL-AF9 and MOZ-TIF2 AML (Krivtsov et al., 2006; Kvinlaug et al., 2011). This suggests that Sox4 and K/L LICs use different transformation pathways from MLL-AF9 and MOZ-TIF2 L-GMPs, a notion further strengthened by very limited overlap between their self-renewal signatures. Intriguingly, gene expression profile comparisons demonstrated increased Foxo3a and type II TGF $\beta$  receptor levels only in the specific C/EBP $\alpha$ -Sox4 LIC signature and in human AML samples with mutated and silent CEBPA. Although the exact role of this signaling cascade in AML remains unknown, its relevance to LICs in CML has been demonstrated (Naka et al., 2010). In a BCR-ABL-induced CML model, pharmaceutical inhibition of TGF $\beta$ -FOXO signaling in combination with a tyrosine kinase inhibitor decreased the colony-forming capacity of LICs and depleted CML in vivo (Naka et al., 2010). It will be of interest to investigate whether the TGF $\beta$ -FOXO signaling pathway is specifically activated and required for maintaining Sox4-mediated C/EBP $\alpha$  defective LICs.

Self-renewal associated signatures of both Sox4/C/EBP $\alpha$  LICs and MOZ-TIF2/MLL-F9 L-GMPs share a central enzyme of the prostaglandin E (PGE) biosynthetic pathway, Ptg1. In addition, our gene expression profiling analysis revealed that not only the synthase for PGE (Ptgs1) but also the receptors for PGE (PTGERs) exhibit increased expression in both LICs and L-GMPs compared to their normal counterparts. This suggests that this pathway or its related regulatory network might serve as a crucial node for leukemic transformation in general. Previous studies have suggested that PGE activated the Wnt signaling pathway by relieving the inhibitory phosphorylation of  $\beta$ -catenin (Castellone et al., 2005; Goessling et al., 2009). Investigating whether the Wnt/ $\beta$ -catenin pathway is commonly

required for leukemogenic transformation, regardless of the underlying initiating mutation, will be of future interest.

The findings of our mouse model are directly translatable to human AML. SOX4 levels are strongly enhanced in human AML patients carrying mutated or silent CEBPA, and a set of self-renewal-associated genes that are specifically upregulated in Sox4 and C/EBP $\alpha$  LICs are also highly enriched in such human AML samples. Furthermore, human AML cells with CEBPA deficiency recapitulate the response of murine cells to Sox4 knockdown and are driven into differentiation. Considering the selective toxicity of Sox4 ablation to leukemic cells, but not to normal HSCs (Schilham et al., 1996), it may therefore be a promising target for effective and leukemia subtype-specific therapies.

## EXPERIMENTAL PROCEDURES

### Serial Replating Assay

Cells were plated in Methocult M4343 media (Stem Cell Technologies), and the colony number was counted and subjected to replating every 10 days.

### Mouse Strains

*Cebpa* conditional knockout mice, Sox4<sup>loxP/+</sup> heterozygous mice, and K/L mice have been described previously (Bereshchenko et al., 2009; Penzo-Méndez et al., 2007; Zhang et al., 2004). Sox4 conditional knockout mice were generated by breeding Sox4<sup>loxP/+</sup> mice to Sox4<sup>loxP/loxP</sup> homozygous mice and intercrossing Sox4<sup>loxP/loxP</sup> mice with Mx1-Cre mice. Sox4/*Cebpa* conditional double-knockout mice were generated by intercrossing the *Cebpa*<sup>loxP/loxP</sup> mice carrying the Mx1-Cre transgene with Sox4<sup>loxP/loxP</sup> mice.

### Chromatin Immunoprecipitation

ChIP assays were performed on stem/progenitor cells (Lin<sup>-</sup>c-kit<sup>+</sup>) or myeloid cells (Mac1<sup>+</sup>Gr1<sup>+</sup>) as described previously (Alberich-Jordà et al., 2012). Immunoprecipitation was performed with rabbit antibodies against C/EBP $\alpha$  (sc-61X, Santa Cruz) or normal rabbit immunoglobulin G (12-370, Millipore). Specific regions were quantified by qPCR with SYBR reagent (Bio-Rad).

### Bone Marrow Transplantation

A total of  $1 \times 10^4$  LSKs or  $2 \times 10^5$  viral-transduced cells were transplanted into lethally irradiated congenic B6.SJL-Ptprca mice by retro-orbital injection with  $2 \times 10^5$  supporting bone marrow cells from unirradiated B6.SJL-Ptprca mice. For transplantation of cells from leukemic mice,  $2 \times 10^3$  cells of fraction I, II, or III were transplanted into sublethally irradiated mice. Mice were monitored by peripheral blood analysis between 2 and 6 months posttransplantation and were considered moribund when they were severely anemic.

Human AML samples were transduced with lentiviruses for 6 hr and immediately transplanted into sublethally irradiated (150 rads) NSG (NOD.Cg-Prkdc<sup>scid</sup> IL2rg<sup>tm1Wjl</sup>/szJ) mice. Donor reconstitution of recipient hematopoiesis was analyzed 5 weeks posttransplantation.

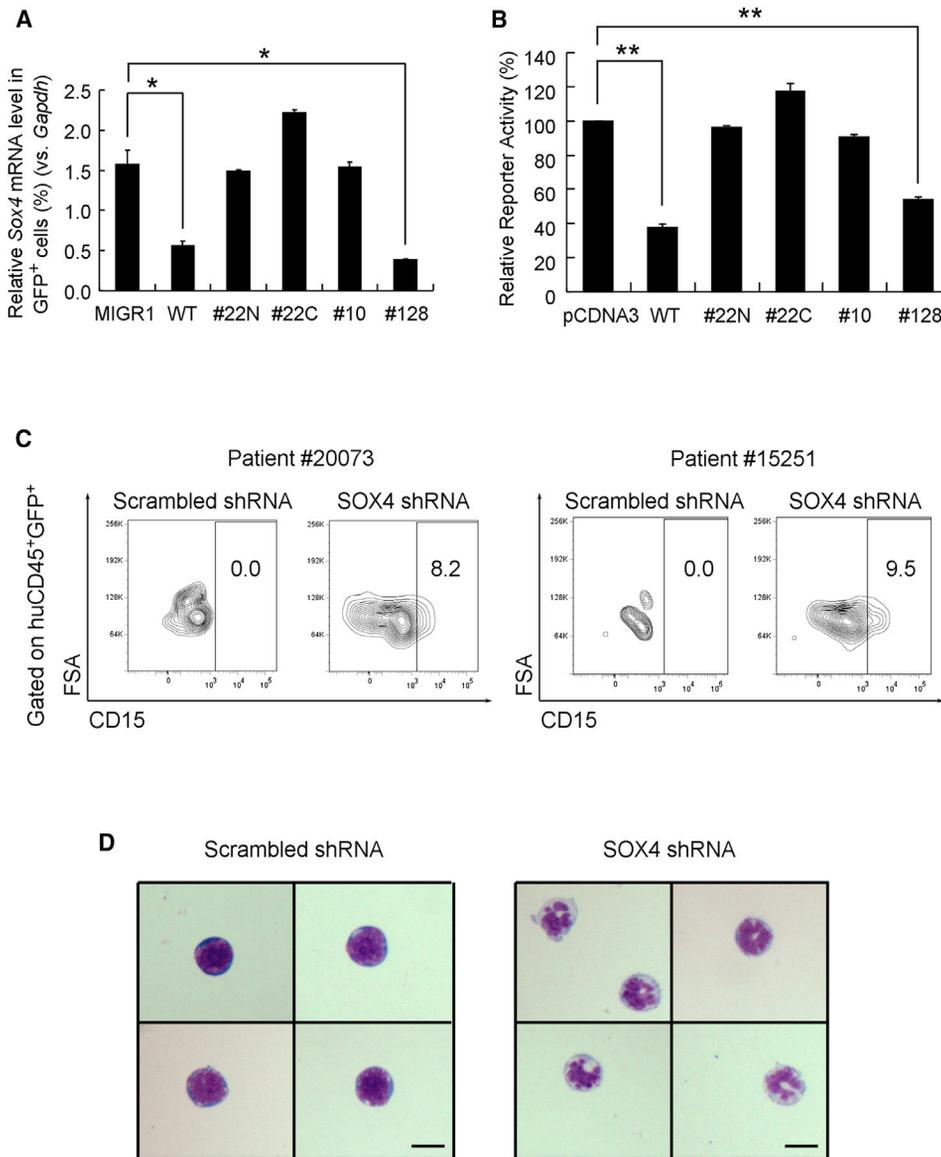
### Microarray Analysis

Total RNA was purified using a QIAGEN RNeasy Plus Micro kit and cDNA was amplified by the Nugen Ovation Pico WTA System. The cDNA library was labeled and hybridized to Affymetrix Mouse430 2.0 chips. Robust multiarray average was run for Cel files generated from our lab (GSE45430) or public databases (GSE3725 and GSE24797). Integrated data were normalized using the cross-correlation method and further log<sub>2</sub>-transformed. Subtraction of the

(E) Supervised analysis of gene expression profiles from wild-type HSCs, CMPs, GMPs, and MEPs, as well as those from MLL-AF9 L-GMP, MOZ-TIF2 L-GMP, Sox4 LICs (Sox4 FRII), K/L LICs (K/L FRII), and their control counterparts, identified a set of self-renewal-associated genes that show increased expression in the L-GMPs, but not in the LICs (fold change cutoff = 1.5; p value cutoff = 0.05).

(F) Gene set enrichment analysis demonstrated significant enrichment of the murine "specific C/EBP $\alpha$ -Sox4 LIC signature" in human AML with mutated CEBPA (left panel, p = 0.005) or with silent CEBPA (right panel, p < 0.001) as compared to other human AML with wild-type CEBPA.

See also Figure S6, Tables S2–S4, and Table S5.



**Figure 7. Abnormal Expression of SOX4 Is Correlated with Mutations of CEBPA in Human AML**

(A) Expression of *Sox4* levels in *Cebpa* KO LSK cells transduced with empty virus (MIGR1) or viruses expressing wild-type human CEBPA or its mutants identified from AML patients (22N, 22C, 10, and 128). Relative gene expression levels were determined by qPCR as % *Gapdh*.

(B) HEK293 cells were transfected with murine *Sox4* promoter reporter and either empty vector (pCDNA3) or constructs expressing wild-type CEBPA (WT) or CEBPA mutants. Luciferase activities were normalized to Renilla activities and are presented as percentage of the empty vector.

(C) Primary human AML patient cells #20073 and #15251 were transduced with lentiviruses carrying scrambled shRNA or human SOX4 shRNA and transplanted into NSG mice. Five weeks later, flow cytometry analysis was performed on the human GFP+ population in the bone marrow of transplanted mice. Differentiated myeloid cells were identified by CD15+.

(D) Morphologically distinct cells were identified by Wright-Giemsa staining of donor-derived GFP+ cells from the bone marrow of recipient mice. Scale bar, 5  $\mu$ m. Error bars indicate the mean  $\pm$  SEM of three independent experiments. \* $p < 0.01$ ; \*\* $p < 0.001$ . See also Figure S7.

mean of the means was carried out for each gene for each sample set. Hierarchical clustering was executed to generate the clustering tree. Human data of AML with mutated or silent *CEBPA* and normal *CEBPA* were taken from GSE14468 and normalized as described above. Gene set enrichment analysis was performed as previously described (Ye et al., 2013).

#### Statistical Analysis

Survival analysis was performed using the Kaplan-Meier method. Differences between survival distributions were analyzed using the log-rank

test. All other statistical analyses were performed using the unpaired Student's t test. Statistical computations were performed using GraphPad Prism.

#### Study Approval

The animal study was approved by the Institutional Animal Care and Use Committee (Protocol # 201-2011) and the study involving human subjects was approved by the Committee on Clinical Investigations of Beth Israel Deaconess Medical Center. All experiments conform to the relevant regulatory

standards. Patients' informed consent was obtained in accordance with the Declaration of Helsinki.

#### ACCESSION NUMBERS

Gene expression data for LICs from Sox4 and C/EBP $\alpha$  mutant leukemic mice and their normal counterparts and ChIP-seq data of genome binding/occupancy profiling of C/EBP $\alpha$  in primary macrophage cells have been deposited in the Gene Expression Omnibus with the accession numbers GSE45430 and GSE50565, respectively.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ccr.2013.09.018>.

#### ACKNOWLEDGMENTS

This study was supported by National Institutes of Health grants HL 56745 and CA118316 and Harvard Stem Cell Institute grant DP-0086-10-00. M.A. was supported by Navrat grant LK21307 from the Czech Ministry of Youth, Health and Sports. G.A. was supported by a Collegio Ghislieri fellowship. P.S. was supported by the Austrian Research Foundation and European Union. A.D.R. was supported by the Societa' Italiana di Ematologia Sperimentale borsa di studio "Dr. Tito Bastianello." R.S.W. was supported by the Jose Carreras Leukemia Foundation (FJJC F11/01). E.L. was supported by FAMRI CIA. V.L. was supported by National Institutes of Health grant NIH/NIAMS R01 AR54153. C.B. and M.H. were supported by grants from Leukaemia Lymphoma Research. We thank all members of the Tenen laboratory and Kristian Reckzeh for helpful discussions, the Harvard Stem Cell Institute/Beth Israel Deaconess Medical Center Flow Cytometry Core for their expertise, Angie Tan Lay Keng and Lee Ming Hui from the NUS-Duke genomic facility in Singapore for their expertise in microarray analysis, and Archibald Perkins and Joseph R. Nevins for providing us with experimental reagents.

Received: April 1, 2013

Revised: July 12, 2013

Accepted: September 25, 2013

Published: October 31, 2013

#### REFERENCES

Alberich-Jordà, M., Wouters, B., Balastik, M., Shapiro-Koss, C., Zhang, H., Di Ruscio, A., Radoska, H.S., Ebralidze, A.K., Amabile, G., Ye, M., et al. (2012). C/EBP $\gamma$  deregulation results in differentiation arrest in acute myeloid leukemia. *J. Clin. Invest.* **122**, 4490–4504.

Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Månsson, R., Luc, S., Grover, A., Jacobsen, S.E., Bryder, D., and Nerlov, C. (2009). Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBP $\alpha$  mutant AML. *Cancer Cell* **16**, 390–400.

Boyd, K.E., Xiao, Y.Y., Fan, K., Poholek, A., Copeland, N.G., Jenkins, N.A., and Perkins, A.S. (2006). Sox4 cooperates with Ev1 in AKXD-23 myeloid tumors via transactivation of proviral LTR. *Blood* **107**, 733–741.

Castellone, M.D., Teramoto, H., Williams, B.O., Druey, K.M., and Gutkind, J.S. (2005). Prostaglandin E2 promotes colon cancer cell growth through a Gs-axin-beta-catenin signaling axis. *Science* **310**, 1504–1510.

Deneault, E., Cellot, S., Faubert, A., Laverdure, J.P., Fréchette, M., Chagraoui, J., Mayotte, N., Sauvageau, M., Ting, S.B., and Sauvageau, G. (2009). A functional screen to identify novel effectors of hematopoietic stem cell activity. *Cell* **137**, 369–379.

Du, Y., Spence, S.E., Jenkins, N.A., and Copeland, N.G. (2005). Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* **106**, 2498–2505.

Eppert, K., Takenaka, K., Lechman, E.R., Waldron, L., Nilsson, B., van Galen, P., Metzler, K.H., Poepl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17**, 1086–1093.

Goessling, W., North, T.E., Loewer, S., Lord, A.M., Lee, S., Stoick-Cooper, C.L., Weidinger, G., Puder, M., Daley, G.Q., Moon, R.T., and Zon, L.I. (2009). Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell* **136**, 1136–1147.

Gombart, A.F., Hofmann, W.K., Kawano, S., Takeuchi, S., Krug, U., Kwok, S.H., Larsen, R.J., Asou, H., Miller, C.W., Hoelzer, D., and Koeffler, H.P. (2002). Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein alpha in myelodysplastic syndromes and acute myeloid leukemias. *Blood* **99**, 1332–1340.

Green, C.L., Koo, K.K., Hills, R.K., Burnett, A.K., Linch, D.C., and Gale, R.E. (2010). Prognostic significance of CEBPA mutations in a large cohort of younger adult patients with acute myeloid leukemia: impact of double CEBPA mutations and the interaction with FLT3 and NPM1 mutations. *J. Clin. Oncol.* **28**, 2739–2747.

Huntly, B.J., Shigematsu, H., Deguchi, K., Lee, B.H., Mizuno, S., Duclos, N., Rowan, R., Amaral, S., Curley, D., Williams, I.R., et al. (2004). MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell* **6**, 587–596.

Jafarnejad, S.M., Ardekani, G.S., Ghaffari, M., and Li, G. (2013). Pleiotropic function of SRY-related HMG box transcription factor 4 in regulation of tumorigenesis. *Cell. Mol. Life Sci.* **70**, 2677–2696.

Johansen, L.M., Iwama, A., Lodie, T.A., Sasaki, K., Felsner, D.W., Golub, T.R., and Tenen, D.G. (2001). c-Myc is a critical target for c/EBP $\alpha$  in granulopoiesis. *Mol. Cell. Biol.* **21**, 3789–3806.

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.

Kustikova, O.S., Geiger, H., Li, Z., Brugman, M.H., Chambers, S.M., Shaw, C.A., Pike-Overzet, K., de Ridder, D., Staal, F.J., von Keudell, G., et al. (2007). Retroviral vector insertion sites associated with dominant hematopoietic clones mark "stemness" pathways. *Blood* **109**, 1897–1907.

Kuwahara, M., Yamashita, M., Shinoda, K., Tofukuji, S., Onodera, A., Shinnakasu, R., Motohashi, S., Hosokawa, H., Tumes, D., Iwamura, C., et al. (2012). The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF- $\beta$  and suppresses T(H)2 differentiation. *Nat. Immunol.* **13**, 778–786.

Kvinlaug, B.T., Chan, W.I., Bullinger, L., Ramaswami, M., Sears, C., Foster, D., Lasic, S.E., Okabe, R., Benner, A., Lee, B.H., et al. (2011). Common and overlapping oncogenic pathways contribute to the evolution of acute myeloid leukemias. *Cancer Res.* **71**, 4117–4129.

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.

Lin, F.T., MacDougald, O.A., Diehl, A.M., and Lane, M.D. (1993). A 30-kDa alternative translation product of the CCAAT/enhancer binding protein alpha message: transcriptional activator lacking antimitotic activity. *Proc. Natl. Acad. Sci. USA* **90**, 9606–9610.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24.

Naka, K., Hoshii, T., Muraguchi, T., Tadokoro, Y., Ooshio, T., Kondo, Y., Nakao, S., Motoyama, N., and Hirao, A. (2010). TGF-beta-FOXO signalling maintains leukaemia-initiating cells in chronic myeloid leukaemia. *Nature* **463**, 676–680.

Nerlov, C. (2004). C/EBP $\alpha$  mutations in acute myeloid leukaemias. *Nat. Rev. Cancer* **4**, 394–400.

Pabst, T., Mueller, B.U., Zhang, P., Radoska, H.S., Narravula, S., Schnittger, S., Behre, G., Hiddemann, W., and Tenen, D.G. (2001). Dominant-negative

- mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat. Genet.* **27**, 263–270.
- Penzo-Méndez, A., Dy, P., Pallavi, B., and Lefebvre, V. (2007). Generation of mice harboring a Sox4 conditional null allele. *Genesis* **45**, 776–780.
- Pietras, E.M., Warr, M.R., and Passegué, E. (2011). Cell cycle regulation in hematopoietic stem cells. *J. Cell Biol.* **195**, 709–720.
- Porse, B.T., Xu, X., Lindberg, B., Wewer, U.M., Friis-Hansen, L., Nerlov, C., and Pedersen, T.A. (2001). E2F repression by C/EBPalpha is required for adipogenesis and granulopoiesis in vivo. *Cell* **107**, 247–258.
- Radomska, H.S., Bassères, D.S., Zheng, R., Zhang, P., Dayaram, T., Yamamoto, Y., Sternberg, D.W., Lokker, N., Giese, N.A., Bohlander, S.K., et al. (2006). Block of C/EBP alpha function by phosphorylation in acute myeloid leukemia with FLT3 activating mutations. *J. Exp. Med.* **203**, 371–381.
- Sarry, J.E., Murphy, K., Perry, R., Sanchez, P.V., Secreto, A., Keefer, C., Swider, C.R., Strzelecki, A.C., Cavellier, C., Récher, C., et al. (2011). Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2R $\gamma$ c-deficient mice. *J. Clin. Invest.* **121**, 384–395.
- Schilham, M.W., Oosterwegel, M.A., Moerer, P., Ya, J., de Boer, P.A., van de Wetering, M., Verbeek, S., Lamers, W.H., Kruisbeek, A.M., Cumano, A., and Clevers, H. (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* **380**, 711–714.
- 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.
- Taussig, D.C., Vargaftig, J., Miraki-Moud, F., Griessinger, E., Sharrock, K., Luke, T., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S.G., et al. (2010). Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. *Blood* **115**, 1976–1984.
- Valk, P.J., Verhaak, R.G., Beijnen, M.A., Erpelinck, C.A., Barjesteh van Waalwijk van Doorn-Khosrovani, S., Boer, J.M., Beverloo, H.B., Moorhouse, M.J., van der Spek, P.J., Löwenberg, B., and Delwel, R. (2004). Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* **350**, 1617–1628.
- Wang, Y., Krivtsov, A.V., Sinha, A.U., North, T.E., Goessling, W., Feng, Z., Zon, L.I., and Armstrong, S.A. (2010). The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science* **327**, 1650–1653.
- Ye, M., Zhang, H., Amabile, G., Yang, H., Staber, P.B., Zhang, P., Levantini, E., Alberich-Jordà, M., Zhang, J., Kawasaki, A., and Tenen, D.G. (2013). C/EBP $\alpha$  controls acquisition and maintenance of adult hematopoietic stem cell quiescence. *Nat. Cell Biol.* **15**, 385–394.
- Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A., et al. (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity* **21**, 853–863.
- Zheng, R., Friedman, A.D., Levis, M., Li, L., Weir, E.G., and Small, D. (2004). Internal tandem duplication mutation of FLT3 blocks myeloid differentiation through suppression of C/EBPalpha expression. *Blood* **103**, 1883–1890.