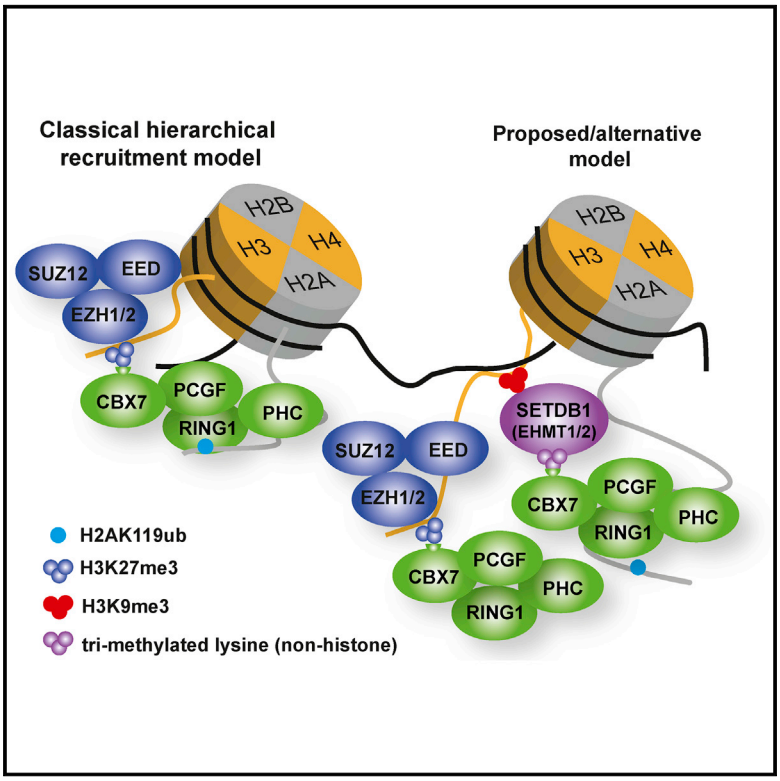
Article

Cell Reports

CBX7 Induces Self-Renewal of Human Normal and Malignant Hematopoietic Stem and Progenitor Cells by Canonical and Non-canonical Interactions

Graphical Abstract



Highlights

- CBX7 regulates self-renewal of human primitive normal and leukemic cells
- CBX7 binds SETDB1 and its inhibition induces differentiation of AML

Authors

Johannes Jung, Sonja C. Buisman, Ellen Weersing, ..., Raymond A. Poot, Leonid V. Bystrykh, Gerald de Haan

Correspondence

l.bystrykh@umcg.nl (L.V.B.), g.de.haan@umcg.nl (G.d.H.)

In Brief

Hematopoietic stem cells ensure production of mature blood cells during the lifetime of an individual. Excessive self-renewal of stem cells leads to leukemia. Jung et at. identify a mechanism that controls self-renewal of normal and leukemic stem cells and show how pharmacological molecules that inhibit this pathway repress leukemic cell growth.

Jung et al., 2019, Cell Reports *26*, 1906–1918
February 12, 2019
https://doi.org/10.1016/j.celrep.2019.01.050

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

CBX7 Induces Self-Renewal of Human Normal and Malignant Hematopoietic Stem and Progenitor Cells by Canonical and Non-canonical Interactions

Johannes Jung,[1] Sonja C. Buisman,[1] Ellen Weersing,[1] Albertina Dethmers-Ausema,[1] Erik Zwart,[1] Hein Schepers,[2] Mike R. Dekker,[3] Seka S. Lazare,[1] Franziska Hammerl,[1] Yulia Skokova,[4] Susanne M. Kooistra,[5] Karin Klauke,[1] Raymond A. Poot,[3] Leonid V. Bystrykh,[1,*] and Gerald de Haan[1,6,*]

[1]European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[2]Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[3]Department of Cell Biology, ErasmusMC, Rotterdam, the Netherlands
[4]University Hospital of Tübingen, Tübingen, Germany
[5]Department of Neuroscience, Section Medical Physiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[6]Lead Contact
*Correspondence: l.bystrykh@umcg.nl (L.V.B.), g.de.haan@umcg.nl (G.d.H.)
https://doi.org/10.1016/j.celrep.2019.01.050

SUMMARY

In this study, we demonstrate that, among all five CBX Polycomb proteins, only CBX7 possesses the ability to control self-renewal of human hematopoietic stem and progenitor cells (HSPCs). Xenotransplantation of *CBX7*-overexpressing HSPCs resulted in increased multi-lineage long-term engraftment and myelopoiesis. Gene expression and chromatin analyses revealed perturbations in genes involved in differentiation, DNA and chromatin maintenance, and cell cycle control. *CBX7* is upregulated in acute myeloid leukemia (AML), and its genetic or pharmacological repression in AML cells inhibited proliferation and induced differentiation. Mass spectrometry analysis revealed several non-histone protein interactions between CBX7 and the H3K9 methyltransferases SETDB1, EHMT1, and EHMT2. These CBX7-binding proteins possess a trimethylated lysine peptide motif highly similar to the canonical CBX7 target H3K27me3. Depletion of SETDB1 in AML cells phenocopied repression of CBX7. We identify CBX7 as an important regulator of self-renewal and uncover non-canonical crosstalk between distinct pathways, revealing therapeutic opportunities for leukemia.

INTRODUCTION

Hematopoietic stem cells (HSCs) are able to self-renew and differentiate into all mature blood cells to ensure peripheral blood cell homeostasis during an adult lifespan. In these primitive cells, the choice between self-renewal and differentiation must be well balanced to avoid either cytopenia or hyperproliferative conditions, such as leukemia. Self-renewal and differentiation are accompanied and controlled by a multitude of epigenetic changes of DNA and of histone proteins (Kamminga et al., 2006; Klauke et al., 2013; Rizo et al., 2008; Tadokoro et al., 2007). One important family of epigenetic regulators that is critical for stem cells is represented by the Polycomb group (PcG) genes.

PcG genes encode for chromatin-associated proteins, which assemble in various multimeric protein complexes and contribute to the regulation of gene expression patterns by post-translational modifications of histone tails (Bracken et al., 2006; Cao et al., 2005).

The two best-characterized PcG protein complexes are the canonical polycomb repressive complex 1 (PRC1) and PRC2. The canonical PRC1 is characterized by the presence of at least one of the five Polycomb chromobox proteins (CBX2, 4, 6, 7, and 8). Many functional and molecular studies have shown similar and overlapping binding patterns of PRC1- and PRC2-protein-containing complexes (Comet and Helin, 2014; Morey et al., 2012). Although the enzymatic activity of many individual epigenetic writers and erasers has been elucidated, our understanding of the biological role and the molecular dynamics of epigenetic protein complexes is still limited.

CBX proteins are characterized as chromodomain-containing proteins, recognizing trimethylated lysine 27 on histone H3 (H3K27me3), which is deposited by EZH1 and EZH2 (Fischle et al., 2003; Min et al., 2003). After recognition of H3K27me3 by the CBX proteins, the catalytic subunit of PRC1, RING1A and/or RING1B, ubiquitinates H2AK119 (Cao et al., 2005), leading to the repression of transcription through chromatin compaction and inhibition of RNA Polymerase 2 (Stock et al., 2007). Beyond this classical PRC2 and/or PRC1 recruitment model, evidence is emerging for a far more diverse and complicated composition and recruitment process. Most notably, it has become apparent that a plethora of distinct PRC1 complexes exist, some of which contain RYBP instead of CBX (Tavares et al., 2012). Furthermore, PRC1 can be present at genomic loci in the absence of any PRC2 activity (Kahn et al., 2016).

Notwithstanding our limited understanding of the complex protein-protein and protein-DNA interactions in which the PcG proteins are involved, it has become evident that PcG proteins are important regulators of self-renewal and differentiation of many types of pluripotent and adult stem cells (Morey et al., 2013). Indeed, deregulation of their expression or mutations in genes coding for PcG proteins can result in cancer development (Herrera-Merchan et al., 2012). We have previously shown that overexpression of the H3K27 methyltransferase Ezh2 in murine hematopoietic stem cells prevents their exhaustion in serial transplantation experiments (Kamminga et al., 2006). Furthermore, both EZH2 and BMI1 are important regulators of self-renewal of normal murine and human hematopoietic stem cells (Rizo et al., 2008). Interestingly, mutations in the *EZH2* gene were later found in patients with myelodysplastic syndromes and acute myeloid leukemia (Cancer Genome Atlas Research et al., 2013; Nikoloski et al., 2010)

More recently, we showed that Cbx7, but not Cbx2, -4, or -8, is a potent regulator of self-renewal of murine hematopoietic stem cells, and its enforced overexpression resulted in increased self-renewal and in phenotypically diverse leukemias (Klauke et al., 2013). In human cells, systematic short-hairpin-mediated repression of all CBX proteins in CD34+ cord blood cells resulted in decreased proliferation and colony-forming unit ability. In this experiment, knock down of CBX2 was shown to be most detrimental (van den Boom et al., 2013).

Collectively, these studies highlight the relevance of PcG proteins, and particularly CBX proteins, in maintaining blood cell homeostasis. As epigenetic changes are in principle reversible, elucidating the function of epigenetic writers, readers, and erasers in the context of healthy and malignant hematopoiesis is indispensable for identifying novel therapeutic targets.

Therefore, in the current study, we asked to what extent different CBX proteins are able to affect the balance between self-renewal and production of mature blood cells of normal human cord blood-derived primitive CD34+ hematopoietic stem and progenitor cells. We identify CBX7 as a potent inducer of self-renewal. Reversely, repression of CBX7 in acute myeloid leukemia (AML) cells results in their terminal differentiation. In addition, we identify evolutionary conserved non-histone interaction partners of CBX7. These interaction partners include multiple epigenetic enzymes, most notably SETDB1, EHMT1, and EHMT2, which are all H3K9 methyltransferases that carry a potential lysine site for trimethylation. These sites are in a conserved peptide context, which is similar to H3K9me3 and H3K27me3. Importantly, depletion of SETDB1, similar to CBX7, also induced differentiation of AML cells, suggesting that at least part of the self-renewal potential of CBX7 is dependent on its interaction with an H3K9 methyltransferase. H3K27me3 and H3K9me3 chromatin immunoprecipitation sequencing (ChIP-seq) and RNA sequencing (RNA-seq) experiments revealed direct and indirect CBX7 targets that comprise a complex network of both classical histone modifications and epigenetic interactions that collectively control the balance between self-renewal and differentiation in primitive human hematopoietic cells.

RESULTS

CBX7 Enhances Self-Renewal of Human CD34+ Cord Blood Cells *In Vitro* and *In Vivo*

To assess the role of the five different PRC1-CBX proteins on hematopoietic progenitor function, we overexpressed CBX2, 4, 6, 7, and 8 in CD34+ cord blood cells and performed colony-forming unit (CFU) assays. Whereas overexpression of CBX7 and CBX8 resulted in increased CFU frequencies, overexpression of CBX2 and CBX4 resulted in lower CFU frequencies in comparison to empty vector (EV) control. Overexpression of CBX6 had no discernable effect (Figure S1A). Although CBX8 overexpression resulted in a slightly higher CFU frequency in comparison to CBX7, replating of CBX7-overexpressing cells resulted in higher CFU frequency (Figure S1B). Overexpression of CBX8 in CD34+ HSPCs showed no proliferative advantage in a cytokine-driven suspension culture, whereas overexpression of CBX7 resulted in a strong proliferative advantage and cells could be kept in culture up to 10 weeks (Figure S1C). To determine the role of the five different CBX proteins in regulating the more primitive cell compartments, we performed cobblestone area-forming cell (CAFC) assays. CBX7 overexpression increased the CAFC day 35 frequency ~10 fold, whereas CBX8 overexpression resulted in a small increase in CAFC frequency (Figure S1D). In contrast, overexpression of CBX4 decreased the CAFC frequency dramatically (~50-fold), while overexpression of CBX2 and CBX6 had no effect. Reversely, short-hairpin-mediated knock down of CBX7 in CD34+ HSPCs with two distinct short hairpins resulted in a 3-fold reduced long-term culture initiating cell (LTC-IC) frequency (Figure S1E).

These *in vitro* phenotypes prompted us to analyze endogenous *CBX* expression levels in different primitive hematopoietic cell subsets by using previously published microarray experiments (Laurenti et al., 2013). *CBX7* expression decreased during differentiation from hematopoietic stem cells (HSC1 = Lin^- $CD34^+CD38^-CD45RA^-CD90^+CD49f^+$, HSC2 = Lin^- $CD34^+$ $CD38^-CD45RA^-CD90^-CD49f^+$) to more multipotent progenitor (MPP), common myeloid progenitor (CMP), and guanosine monophosphate (GMP) subsets (Figure S1F).

To assess consequences of CBX7 overexpression in *in vivo*, we transplanted 2×10^5 CD34+GFP+ cells in sub-lethally irradiated female NOD-SCID IL2rynull (NSG) mice. Mice transplanted with CBX7-overexpressing cord blood cells showed significantly higher engraftment of CD45+GFP+ cells in peripheral blood (PB) (Figure 1A) and in bone marrow (BM) (Figure 1B).

To explore whether CBX7 overexpression would be able to maintain human $CD34^+$ HSPCs in a more primitive state for a longer period *ex vivo*, we prolonged total *in vitro* culture time from 3 to 7 days and transplanted 1.5×10^6 GFP^+CD34^+ HSPCs. Again, NSG mice transplanted with *CBX7*-overexpressing HSPCs displayed significantly higher engraftment in PB (Figure 1C) and in bone marrow (Figure 1D). Mice transplanted with CBX7-overexpressing HSPCs showed a significantly increased percentage of CD33+ cells in bone marrow, suggesting that overexpression of CBX7 enhances myelopoiesis (Figures 1E and 1F). Similar results were obtained in the PB after 18 weeks (data not shown). Furthermore, these mice showed a significantly higher percentage of primitive

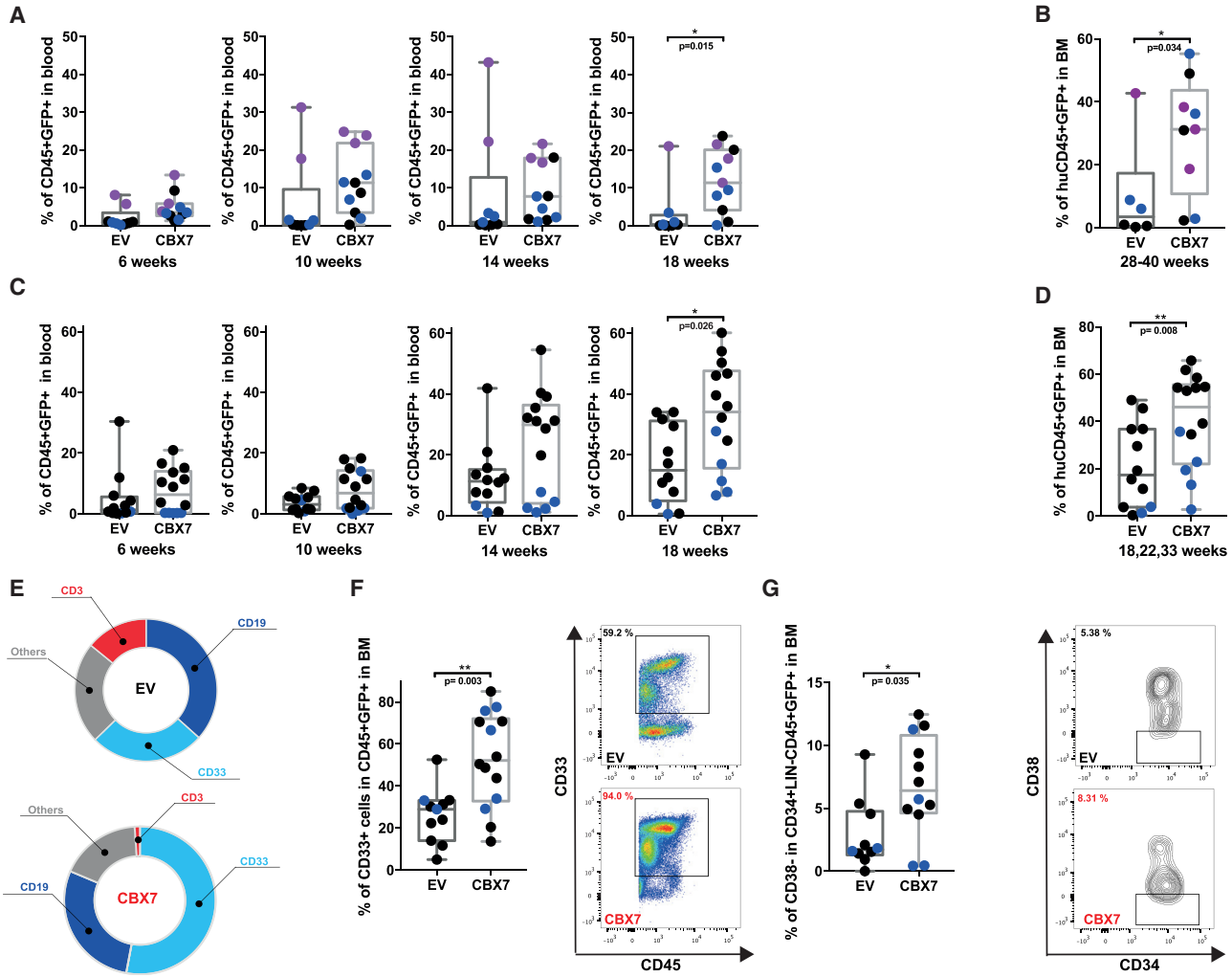


Figure 1. CBX7 Overexpression Induces Enhanced Long-Term Engraftment, Myelopoiesis, and Self-Renewal of Primitive CD34+CD38− HSPCs *In Vivo*

(A) Human chimerism levels in the peripheral blood of NSG mice upon transplantation of 200,000 CD34+GFP+CBX7-overexpressing or empty vector (EV) control cord blood cells.

(B) Primary recipients were sacrificed after 28–40 weeks, and human engraftment in bone marrow was evaluated.

(C) Human chimerism levels in the peripheral blood of NSG mice upon transplantation of 7 days *ex vivo*-cultured CBX7-overexpressing or empty vector control CD34+ cord blood cells.

(D) Human engraftment in bone marrow of mice shown in (C), combined analysis of mice sacrificed after 18, 22, and 33 weeks post-transplant.

(E) Relative engraftment of human CD19+, CD3+, and CD33+ cells within human CD45+GFP+ bone marrow cells of mice shown in (C), combined analysis of mice sacrificed after 18, 22, and 33 weeks post-transplantation.

(F) Human myeloid engraftment in bone marrow of mice shown in (C), combined analysis of mice sacrificed after 18, 22, and 33 weeks post-transplantation.

(G) Frequency of CD34+CD38− cells in huCD45+GFP+lin-CD34+ cells in bone marrow of mice shown in (C), 18 combined analysis of mice sacrificed after 18, 22, and 33 weeks post transplantation. Identically colored circles indicate 3 paired experimental and control samples that originate from the same cord. Statistical analysis was performed using a two-tailed Mann-Whitney test.

$CD38^-$ cells in the GFP^+ Lin^-CD34^+ compartment (Figure 1G), indicating that CBX7 controls *in vivo* proliferation or maintenance of human HSPCs.

Genome-wide Transcriptional Consequences of *CBX7* Overexpression in Human HSPCs

We next profiled the transcriptome of CBX7-overexpressing CD34+ HSPCs by using RNA-seq. Differential expression analysis showed a total of 1,463 genes significantly upregulated and 1,183 genes significantly downregulated. (Tables S1 and S2). To annotate CBX7-induced up- and downregulated genes, we performed Gene Ontology (GO) enrichment analysis. (Tables S3 and S4). This revealed that more than 100 genes were associated with "cell differentiation," particularly differentiation of hematopoietic cells. (Figure 2A; Tables S3 and S4). Furthermore, we found genes associated with cell cycle arrest

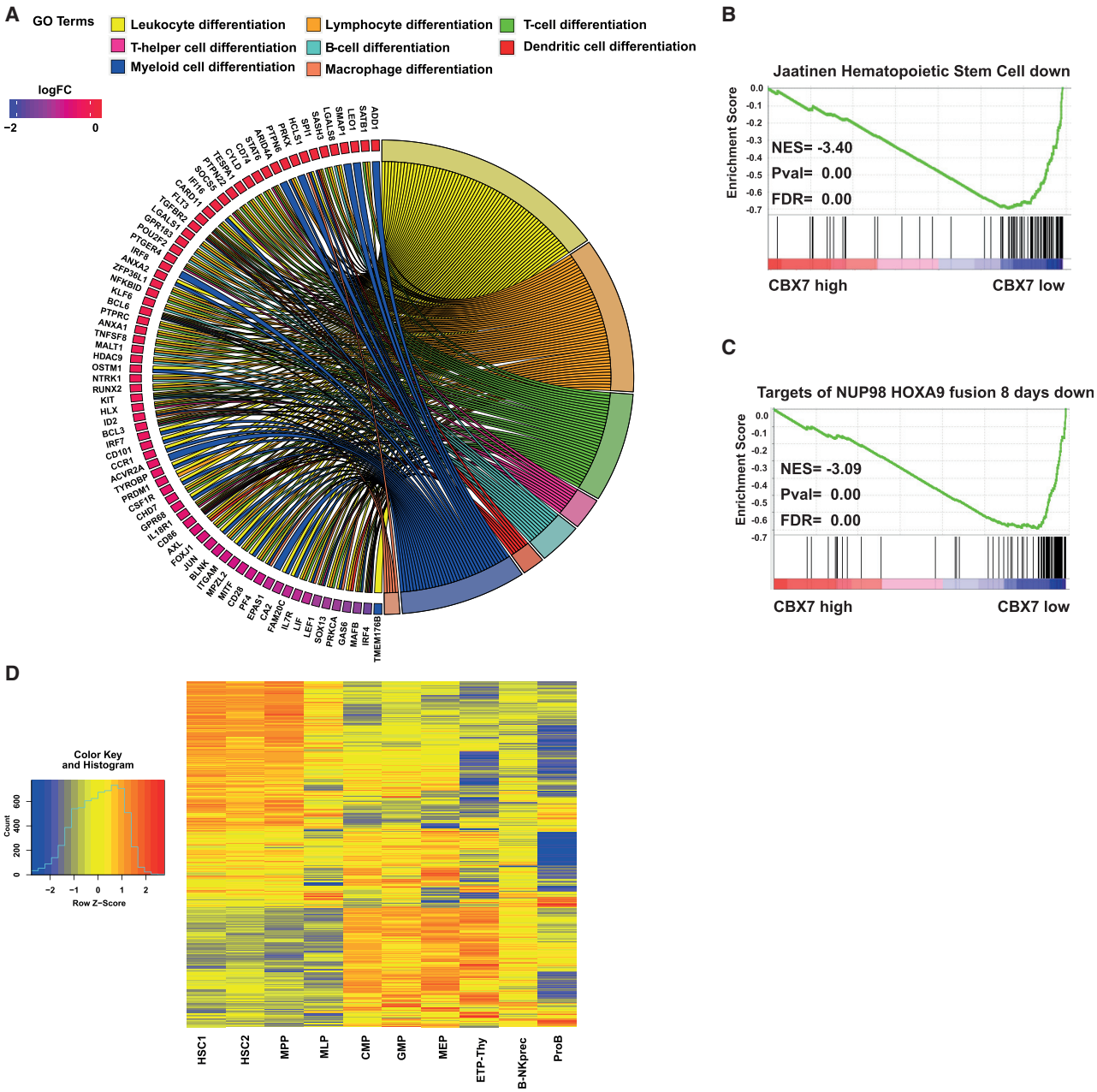


Figure 2. RNA-Seq Analysis of CBX7-Overexpressing CD34+ HSPC

Sets of differentially expressed genes were screened for Gene Ontology (GO) enrichment. GO categories were enriched for "differentiation," "cell cycle," "chromatin," and "DNA," shown using GO Chord plots. Preranked gene set enrichment analysis was performed for differentially expressed genes (FDR, <0.1) upon overexpression of CBX7 in comparison to empty vector control cells.

(A) GO Chord plot of genes repressed upon overexpression of CBX7 in comparison to control cells, associated with the GO terms "differentiation" of various hematopoietic cells.

(B and C) Gene Set Enrichment plot for genes downregulated upon overexpression of CBX7 reveal significant enrichment for (B) HSC genes and (C) NUP98-HOXA9 target genes ($p < 0.001$).

(D) Heatmap containing genes upregulated upon overexpression of CBX7 and their expression in multiple normal hematopoietic subsets according to previously published data from Laurenti et al., (2013).

to be repressed. In contrast, upregulated genes revealed transcripts related to "G1-S transition of mitotic cell cycle" (Figure S2A) and DNA replication (Figure S2B). These GO annotations are in good agreement with the *in vitro* and *in vivo* observation that overexpression of CBX7 leads to elevated self-renewal.

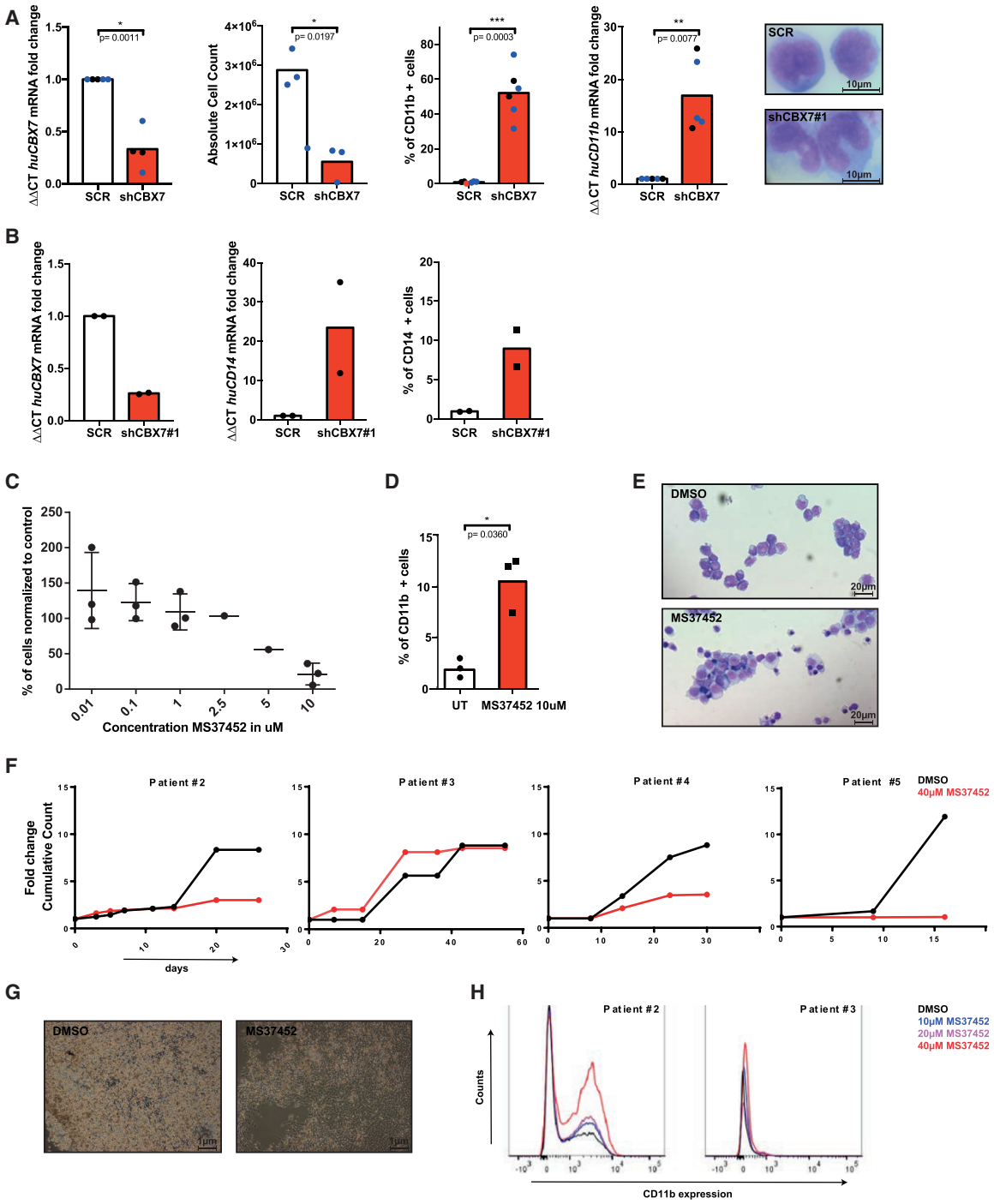


Figure 3. Genetic or Pharmacological Inhibition of CBX7 Induces Differentiation in AML cells

(A) Short-hairpin-mediated knock down of CBX7 mRNA (first panel, n = 5) in HL60 cells results in the loss of cell proliferation (second panel, n = 3), upregulation of CD11b mRNA and protein (third and fourth panels, n = 5) after six days in culture. Multiple cells showed signs of differentiation upon knock down of CBX7 (5th panel) (black = shCBX7#1, blue = shCBX7#2).

(B) Short-hairpin-mediated knock down of CBX7 mRNA (1st panel) in OCI-AML3 cells results in upregulation of CD14 mRNA (second panel) and protein (third panel) after six days in culture (n = 2).

(C) Growth of OCI-AML3 cells treated with different concentrations of the CBX7 chromodomain inhibitor MS37452 after four days of culture (n = 3).

(D) Treatment of OCI-AML3 cells with MS37452 at a concentration of 10 μM results in increased expression of CD11b (n = 3).

(E) MS37452 induces monocyte and/or macrophage differentiation in OCI-AML3 cells. After treatment for 4 days with MS37452 at concentration of 10 μM, cytospin preparations were stained with May-Grünwald Giemsa stain. Magnification, 40×.

(legend continued on next page)

Complementary, we performed Gene Set Enrichment Analysis (GSEA) on a pre-ranked list of differentially expressed genes (false discovery rate [FDR], <0.1). Interestingly, GSEA revealed a strong negative correlation with a gene set containing genes with low abundance in hematopoietic stem cells, indicating that increased levels of *CBX7* result in repression of genes that are usually barely expressed in hematopoietic stem cells (Figure 2B). Furthermore, we identified two other sets with a high negative correlation, both containing genes downregulated upon overexpression of HOXA9 either with NUP98 or Meis1, suggesting that CBX7 targets overlap with targets of these fusion oncogenes (Figure 2C; Figure S2C). Additionally, we found a strong negative correlation with a gene set containing genes lower expressed in leukemic stem cells (CD34+CD38−) in comparison to leukemic blasts (CD34+CD38+), suggesting that genes downregulated by CBX7 overexpression are indeed expressed lower in leukemic stem cells than in leukemic blasts (Figure S2D).

To further characterize differentially expressed genes upon CBX7 overexpression, we compared these with steady-state transcriptomes of multiple subsets of hematopoietic cell types by using a previously published expression dataset as a cross reference (Laurenti et al., 2013). This analysis revealed that 378 transcripts that were higher expressed upon CBX7 overexpression were preferentially abundant in the more primitive cell compartments (HSC1, HSC2, and MPP versus multipotent lymphoid progenitor [MLP], CMP, GMP, megakaryocyte erythroid progenitor [MEP], ETP-Thy, B-NKprec, and ProB) (Figure 2D; Table S5). This suggests their involvement in maintaining elevated levels of self-renewal upon overexpression of CBX7 in hematopoietic stem cells.

CBX7 Expression Is Elevated in AML and Its Repression Results in Differentiation of AML Cells

Our data show that CBX7 is able to increase self-renewal of normal human hematopoietic stem and progenitor cells. To explore a putative role for CBX7 in the maintenance of AML cells, we first analyzed *CBX7* mRNA expression levels in AML patient samples in two previously published datasets. In the first dataset, containing 529 AML patient samples from patients treated at the Erasmus MC (Rotterdam, the Netherlands), CBX7 expression was significantly upregulated in comparison to peripheral-blood-mobilized CD34+ cells (Verhaak et al., 2009). The highest expression was observed in acute promyelocytic leukemia (APL), leukemias with a normal karyotype and *NPM1* mutated leukemia (Figure S3A). We additionally analyzed data from the Cancer Genome Atlas by Bloodspot (Bagger et al., 2016). Also, in this patient cohort, CBX7 was significantly higher expressed in multiple AML subtypes (Figure S3B).

To explore a functional role for high CBX7 expression in human leukemia, we assessed to what extent depletion of CBX7 would affect leukemic cell growth. As CBX7 is more abundantly expressed in APL (Figure S3A), we downregulated *CBX7* mRNA by using a short-hairpin approach in HL60 cells, which harbor a t(15;17) translocation. Knock down of *CBX7* was associated with a reduced abundance of CBX7 mRNA to ~40% of normal levels (Figure 3A, first panel) and lower absolute cell numbers after 6 days in culture (Figure 3A, second panel). Strikingly, downregulation of CBX7 resulted in a significant increase of CD11b expression, which is usually not expressed on primitive APL-blasts but rather on mature monocytes, macrophages, and granulocytes (Figure 3A, third panel). The changes of CD11b protein levels were associated with an increased expression of *CD11b* on mRNA level (Figure 3A, fourth panel), and morphological signs of cellular maturation upon May-Grünwald Giemsa staining (Figure 3A, fifth panel).

It has been reported that CBX7 can interact with mutated DNMT3A(R882) but not with wild-type DNMT3A in AML patient samples (Koya et al., 2016). Therefore, we decided to downregulate *CBX7* in OCI-AML3 cells, a cell line carrying *DNMT3A* R882 and mutant *NPM1*. Similar as in HL60 cells, upon knock down of *CBX7*, OCI-AML3 cells started to differentiate and upregulated the differentiation marker CD14 on the protein and mRNA level (Figure 3B). In summary, these experiments indicate that CBX7 is necessary for maintaining leukemic cells in an undifferentiated state, independent of DNMT3A.

We tested whether pharmacological inhibition of CBX7 would result in similar effects as short-hairpin-mediated repression. To this end, we cultured OCI-AML3 cells in the presence of increasing concentrations of the small molecule MS37452, which has been shown to bind to residues in the chromodomain of CBX7, thereby preventing binding of CBX7 to proteins harboring a trimethylated lysine residue. This loss of normal chromodomain function resulted in derepression of PRC target genes in prostate cancer cells (Ren et al., 2015). In OCI-AML3 cells, MS37452 resulted in the loss of cell growth in a time- and dose-dependent manner (Figure 3C). Furthermore, MS37452 treatment induced differentiation in leukemic cells, as evidenced by upregulation of the differentiation marker CD11b and by the strong increase of cells with a highly differentiated morphology (Figure 3D and E). We observed similar effects in THP1 cells, which carry a CDKN2A and RING1B deletion, suggesting that the inhibitory effect of MS37452 is not due to derepression of this locus (data not shown).

Finally, we tested whether pharmacological inhibition of CBX7 would similarly induce differentiation in primary, patient-derived, leukemic cells. To this end, we initiated stroma-associated cultures in which AML cells isolated from 4 different patients were seeded on MS5 stromal cells. When cell growth was observed, MS37452 was added to the cultures and cell growth was evaluated. In 3 out of 4 patient samples, MS37452 potently inhibited cell growth (Figure 3F). Figure 3G shows a micrograph of these cultures, clearly indicating that overall cell growth is severely impaired in the presence of the CBX7 inhibitor. In the sample of patient 3, no clear inhibitory effect of MS37452 was observed. This seemingly non-responsive patient sample displayed particular cytogenetics (46,XX, del(7)(q22q36)).

(F) Primary AML cells, derived from 4 patients, were cultured on MS5 stromal cells. When primary cells started to grow, MS37452 was added to the co-cultures and cell proliferation was monitored. In 3 out of 4 patients, cell growth was strongly inhibited by MS37452.
(G) Micrograph of primary AML cells of patient 2, co-cultured on MS5 stromal cells in the absence (left) of presence (right) of MS37452.
(H) MS37452 dose-dependently induces expression of CD11 in primary AML cells from patient 2, but not in refractory cells of patient 3.

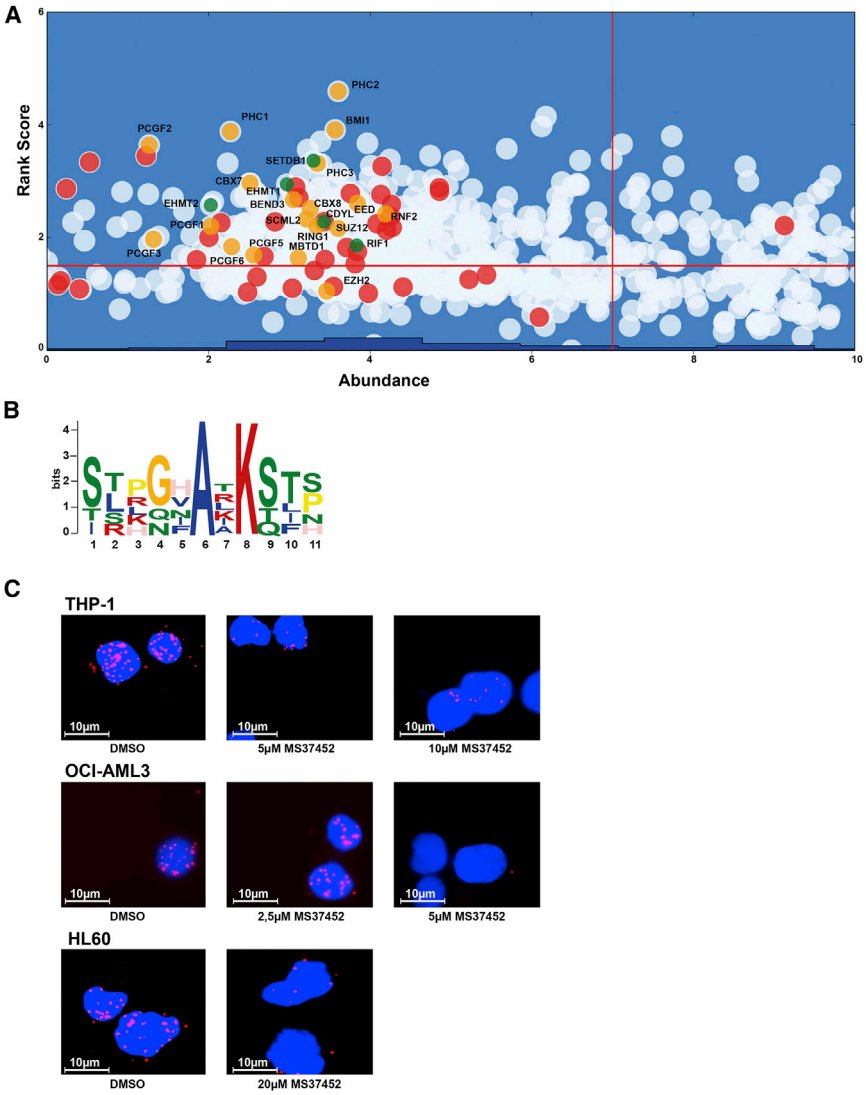


Figure 4. Mass Spectrometry Analysis Reveals Multiple H3K9 Methyltransferases as CBX7-Binding Partners

(A) Search for putative interaction partners of CBX proteins by label-free mass spectrometry. The plot depicts CBX7 interaction partners ranked on the basis of their cumulative rank score (derived from the frequency of spectral counts, corrected for GFP control samples), and the average abundance of these proteins in the human PaxDB database. The top-left corner represents priority candidates. Red symbols indicate known transcription factors, orange symbols refer to Polycomb proteins, and green symbols are proteins with known lysine trimethylation sites.

(B) Peptide motif of the CBX7 interaction partners harboring a lysine embedded in a motif highly similar to H3K9me3 and H3K27me3.

(C) Duolink proximity ligation assay (PLA) of endogenous SETDB1 and CBX7 interactions in THP-1, OCI-AML3, and HL60 cells in the absence (left panels) or presence of MS37452. Each PLA signal (Cy3, red) is indicative of one detected interaction event. Nuclei (blue) are stained with 4′,6-diamidino-2-phenylindole (DAPI).

Interestingly, EZH2 is located in this deleted region, suggesting that canonical CBX7 target loci may not be recognized. However, more samples would need to be tested to further provide insight into the differential sensitivity of primary AML cells. In addition, we determined the expression of the differentiation marker CD11b on MS37452-exposed AML cells (Figure 3H) and found increased CD11b expression of primary AML cells of patient 2 (responsive to MS37452) and no effect in patient 3 (non-responsive to MS37452). We observed that MS37452 dose-dependently resulted in increased expression of CD11b in primary AML cells over time (Figure 3H). Collectively, these data show that at least in some primary AML patient samples, inhibition of CBX7 has potent anti-leukemic effects.

CBX7 Interacts with Trimethylated Non-Polycomb Proteins

To further unravel the molecular mechanism by which CBX7 exerts its potent activity, and taking into account that PcG proteins are known to operate in large protein complexes, we decided to identify proteins directly interacting with CBX.

We performed label-free mass spectrometry analysis of benzoase-treated proteins that co-precipitated with FLAG-tagged CBX7, FLAG-tagged CBX8, and FLAG-tagged CBX4, using murine and human cells. A protein fraction that co-precipitated with FLAG-tagged GFP was used as a negative control. To prioritize candidates, we first removed proteins with low spectral counts (<10% of the cumulative spectral count) and then ranked proteins in relation to their spectral counts. We compared all mass spectrometry (MS) sets and screened for consistent binding partners of both murine and human CBX proteins. As expected, multiple members of PRC1 and PRC2 complexes were identified, including PCGF1, PCGF2, PCGF5, PCGF6, SCML2, PHC1, PHC2, PHC3, BMI1, RING2, RING1, EED, and SUZ12 (Figure 4A).

Because canonically CBX7 binds to the trimethylated lysine of H3K27 through its chromodomain, we hypothesized that the chromodomain could potentially associate with other trimethylated lysines in non-histone proteins when they contain a peptide context similar to H3K27. Therefore, we screened the list of CBX7 human and murine binding partners for proteins harboring a putative trimethylated lysine by using the PhosphoSitePlus database (Hornbeck et al., 2015). This screen revealed a list of 218 human and murine trimethylated proteins, corresponding to 335 known trimethylated human and murine peptides. We only considered proteins with high spectral counts (top 15% of the relative rank ordered proteins) and low protein abundance

(<100 ppm, https://pax-db.org/, average of all samples) and corrected for the binding of each candidate to GFP. This strict filtering narrowed our list down to four proteins (CDYL, SETDB1, EHMT1, and EHMT2). We then applied the same filtering for murine Cbx7-binding proteins and identified SETDB1, EHMT1, and EHMT2 as evolutionarily conserved binding partners. When using less strict filtering rules (top 70%) we also identified CDYL as a binding partner of murine Cbx7. Interestingly, this list of CBX7 interaction partners contains three H3K9 methyltransferases (EHMT1, EHMT2, and SETDB1) and one H3K9me3-associated protein (CDYL). In our mass spectrometry data, we were able to confirm the presence of EHMT1 and EHMT2 trimethylated peptides (Figures S4A and S4B). Next to our approach to use spectral counts for identifying putative interaction partners, we also calculated the relative enrichment in the exponentially modified protein abundance index (emPAI) in the CBX7 sample over the control sample (FLAGGFP). All our candidate proteins (SETDB1, EHMT1, EHMT2, and CDYL) showed a relative enrichment of at least 12. (Table S6).

We next assessed whether these human CBX7-binding partners contain a common signature. Indeed, the consensus-binding motif for CDYL, EHMT1, EHMT2, and SETDB1 was found to be S[LT]PGHA.Kme3ST[PS], which is highly similar to the peptide sequences to which human CBX7 is known to bind: A[RILFYV]Kme3[ST] (Kaustov et al., 2011) (Figure 4B). The similarity of these motifs suggests that CBX7 interacts with these non-histone proteins via the chromodomain. Our data show that human Polycomb CBX7 is able to interact with multiple H3K9 methylation-associated proteins that harbor a lysine that can be trimethylated, and we chose to focus on the CBX7-SETDB1 interaction. To validate the interaction of SETDB1 and CBX7 in FLAG-tagged CBX7-overexpressing K562 cells, we stained the immunoblots for CBX7 and SETDB1 (Figure S4C).

In addition, we performed proximity ligation assays (PLA) of endogenous SETDB1 and CBX7 in 3 different human acute myeloid and lymphoid cell lines (HL60, THP1, and OCI-AML3) to confirm that CBX7 and SETDB1 are indeed bona fide binding partners and found intense co-staining throughout the nucleus (Figure 4C, left panels). We next determined whether blocking the chromodomain of CBX7 by using MS37452 would result in loss of these CBX7-SETDB1 interactions. Indeed, MS37452 dose-dependently inhibited this interaction, as revealed by PLA analyses in the all leukemic cell lines (Figure 4C).

These data suggest that the chromobox of CBX7, which is essential and sufficient to bind lysine-trimethylated residues, is required for physiological binding to SETDB1. We identify a non-canonical (non-H3K27me3) role for CBX7 in crosstalking to epigenetic pathways governing H3K9 methylation.

Identification of CBX7 Target Loci and Their Association with H3K9me3 and H3K27me3

To identify genes directly controlled by CBX7 and to unravel their association with H3K27me3 and H3K9me3, we performed ChIP-seq experiments in primary HSPCs from different cord blood donors overexpressing either CBX7 or an empty vector. Before deep sequencing we tested each sample for enrichment at known target loci by qPCR (Figure 5A). Following sequencing, we only considered peaks that were present in at least two out of three independent ChIP-seq samples with an adjusted p value < 0.05. Upon overexpression, CBX7 peaks were strongly enriched at the transcription start site (TSS), confirming that CBX7 acted preferentially at core promoter regions. (Figure 5B).

We next searched genome wide for loci that were targeted by CBX7 and marked by H3K9me3 or H3K27me3. We observed that 23% of all CBX7 peaks were overlapping with H3K9me3, whereas 44% were overlapping with H3K27me3. Furthermore, 13% of all CBX7 peaks were associated with both H3K9me3 and H3K27me3 (Figure 5C), culminating in ~1/3 of all CBX7 peaks being associated with H3K9me3 ($p < 1.74e{-}244$).

Interestingly, in CBX7-overexpressing CD34+ cells ~20% of all TSSs marked with CBX7 were also marked with H3K9me3, providing further evidence of a joint gene regulatory function (Figure S5A). These molecular patterns are compatible with a model in which H3K9 methyltransferases act as binding partners of CBX7, at least for a subset of the genomic sites bound by CBX7.

To refine the list of direct targets of CBX7, we performed an integrative analysis of RNA-seq and ChIP-seq data and searched for genes repressed upon overexpression of CBX7, which were additionally marked by CBX7. Out of 1,183 repressed genes, 220 showed CBX7 peaks within 5 kb around their TSS and an additional 63 genes were marked within the gene body (Figure S5B shows some illustrative examples). All these 283 genes are likely to be primary targets of CBX7 in human HSPCs (Table S7).

The large majority of these primary targets (246 out of the 283) were also marked by H3K27me3 within 5 kb around the gene body, confirming the well-known interaction of CBX with the Polycomb repressive mark set by EZH2. Interestingly, and in agreement with our finding that CBX7 directly interacts with various H3K9 methyltranferases, 178 (i.e., 62%) of these direct CBX7 targets were also marked with H3K9me3 (Figure 5D).

Collectively, these molecular signatures reveal functional non-canonical cross talk between Polycomb CBX proteins and H3K9 methylation, as first suggested by the physical interaction of CBX7 and SETDB1, EHMT1, and EHMT2.

SETDB1 and CBX7 Share Functional Activity in AML Cells

As we identified the H3K9 methyltransferase SETDB1 as a CBX7-interacting protein and as we found that approximately one-third of the CBX7 genomic target loci were also covered by H3K9me3, we evaluated the function of SETDB1 in leukemic cells. Mutations in *SETDB1* are associated with the development of clonal hematopoiesis and SETDB1 is higher expressed in leukemic stem cells compared to leukemic blasts in AML patient samples (Eppert et al., 2011; Steensma et al., 2015); therefore, we set out to investigate whether short hairpin RNA (shRNA)-mediated knock down of SETDB1 in myeloid leukemic cells would phenocopy the effects observed upon CBX7 repression. We observed a loss of LTC-IC potential of normal CD34+ cells upon SETDB1 repression (data not shown). Similarly, SETDB1 knock down strongly impaired proliferation of HL60 cells (Figure 6A). In addition, SETDB1 knock down resulted in increased expression of CD11b, similar as CBX7 in HL60 cells (Figure 6B). Knock down of SETDB1 resulted in increased

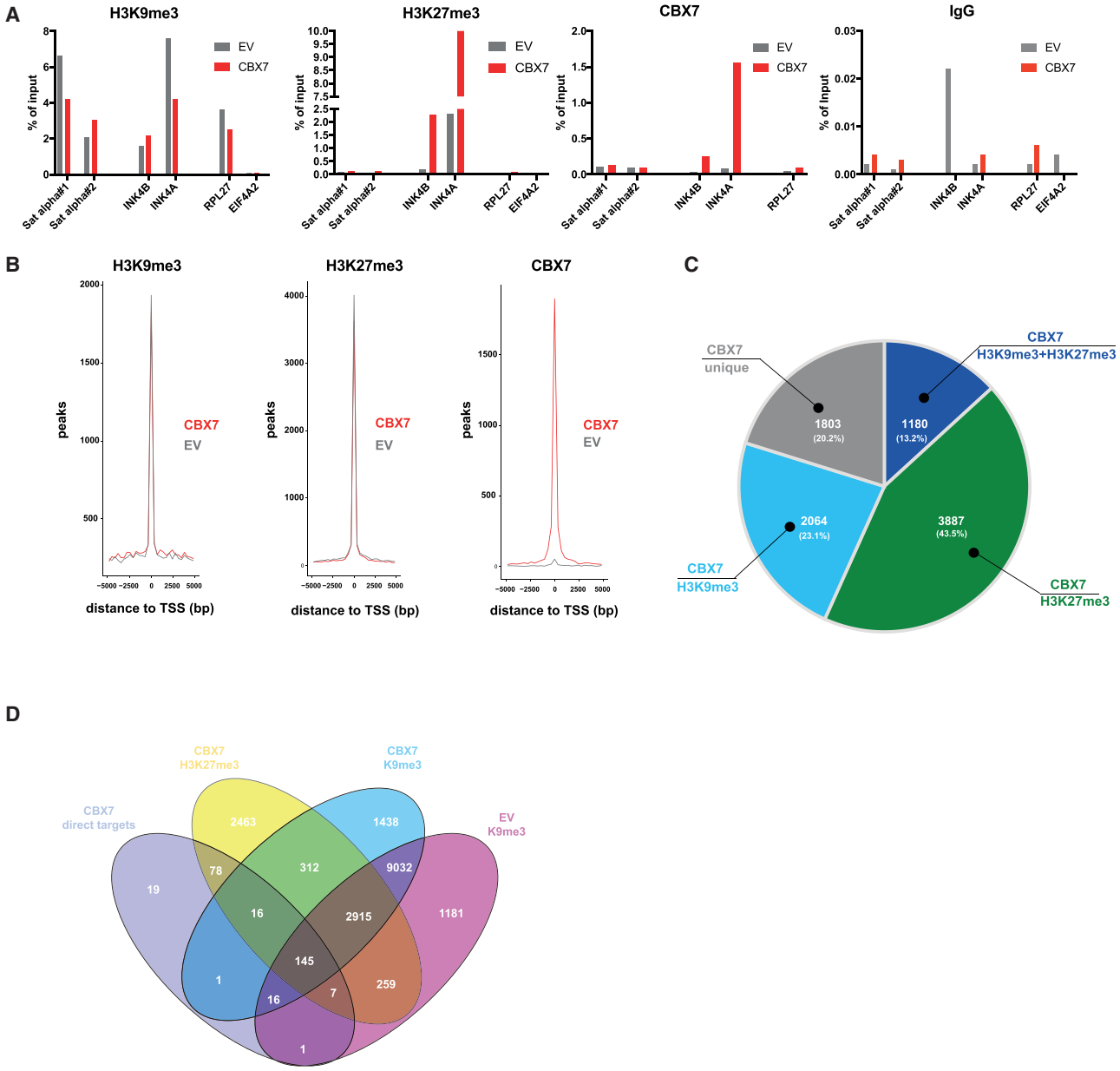


Figure 5. Identification of CBX7 Genome-wide Binding Sites in Primary Human CD34+ Cells and Their Association with H3K9me3 and H3K27me3

(A) ChIP-qPCR validation of selected positive and negative H3K9me3, H3K27me3, and CBX7 target loci, and IgG (control). (Data from one (out of 3) representative experiment are shown).

(B) Genome-wide distribution of H3K9me3, H3K27me3, and CBX7 peaks to nearest TSS in base pairs (bps).

(C) Pie chart showing absolute and relative numbers of genome-wide CBX7 peaks and their overlap with H3K9me3 and/or H3K27me3 peaks.

(D) Venn diagram showing overlap of genes marked with H3K27me3 (CBX7 H3K27me3) and H3K9me3 (CBX7 H3K9me3) in CBX7-overexpressing CD34+ HSPCs, H3K9me3 in control CD34+ HSPCs (empty vector H3K9me3) and direct targets of CBX7.

expression of *CBX7* (Figure 6C). Also, similar to CBX7, knock down of SETDB1 in OCI-AML3 cells induced increased the expression of differentiation markers (Figures 6D and 6E) and reduced proliferation (Figure 6F). Finally, the repression of SETDB1 resulted in the appearance of CD14+ cells 6 days after culturing (Figure 6G). The CD14 locus is a direct CBX7 target, as in primary CD34^{+} cells *CD14* is in the top three downregulated genes upon overexpression of *CBX7* and is marked with CBX7, H3K27me3, and H3K9me3 (Figure S4B). These experimental data indicate that CBX7 and SETDB1 jointly repress genes that are important for differentiation of leukemic cells toward mature myeloid cells.

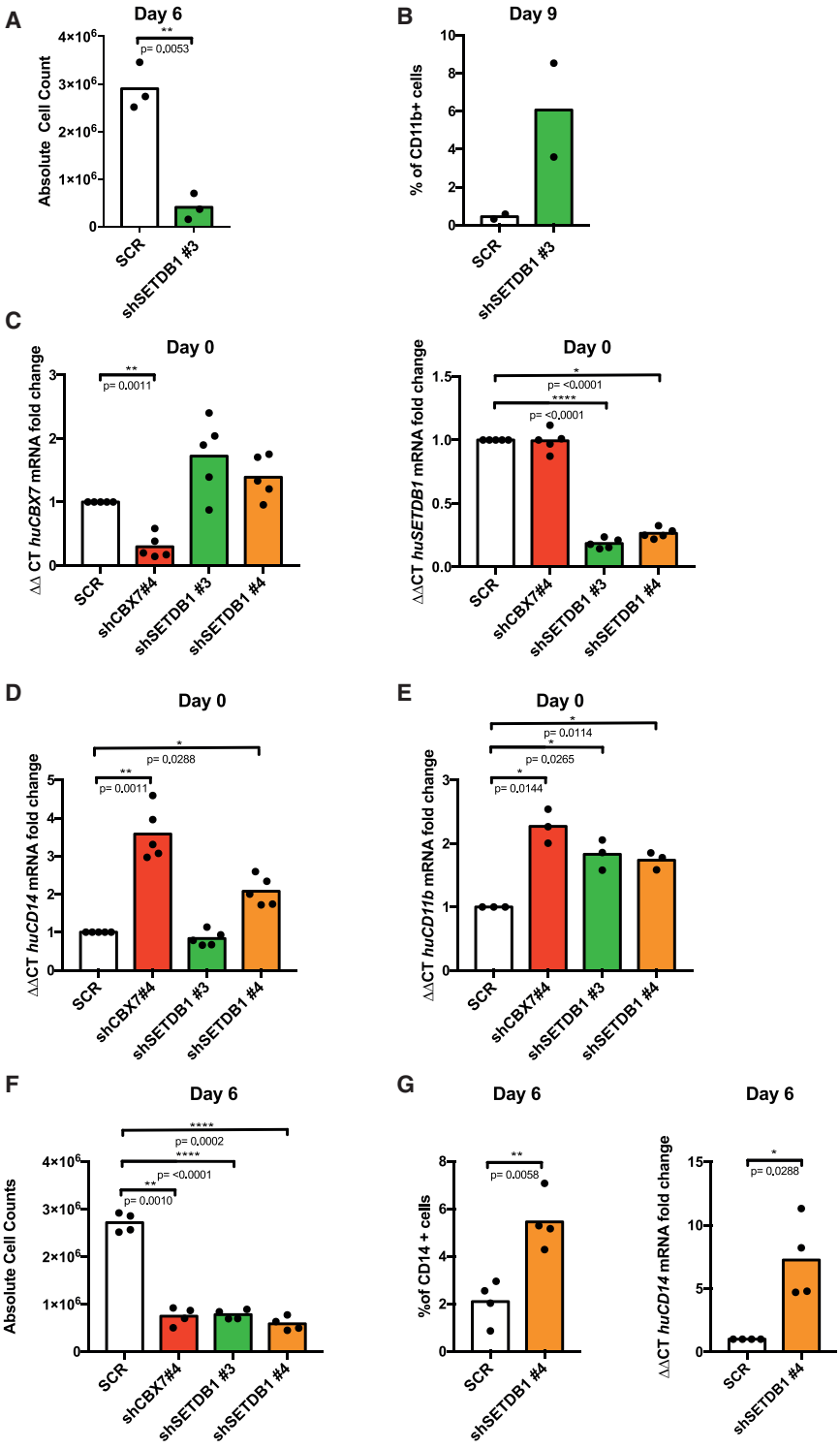


Figure 6. Repression of SETDB1 Phenocopies Repression of CBX7 in Leukemic Cells

(A) Absolute cell numbers after 6 days of culturing 150,000 HL60 cells upon short-hairpin-mediated knock down of SETDB1 (n = 3).

(B) Percentage of CD11b+ HL60 cells after 9 days in culture upon knock down of SETDB1 (n = 2).

(C) Fold change of CBX7 mRNA (left) and hu-SETDB1 mRNA (right) expression in OCI-AML3 cells 24 h after transduction with multiple short hairpins targeting CBX7 or SETDB1 (n = 5).

(D and E) Fold change of CD11b (D) and CD14 (E) mRNA expression in OCI-AML3 cells 24 h after transduction with multiple short hairpins targeting CBX7 or SETDB1 (n = 3–5).

(F) Absolute cell numbers after 6 days of culturing 150,000 OCI-AML3 cells upon short-hairpin-mediated knock down of SETDB1 or CBX7 (n = 4).

(G) Percentage of CD14+ cells (left) and fold change of CD14 mRNA expression (right) upon knock down of SETDB1 in OCI-AML3 cells six days after sort (n = 4). All statistical analyses were performed using paired t test, two-tailed.

DISCUSSION

In this study, we identify CBX7 as a regulator of self-renewal in normal and leukemic hematopoietic cells. We describe the complex molecular architecture of CBX7-induced cell proliferation and discover a biologically relevant, non-canonical role for CBX7 as a binding partner of multiple H3K9 methyltransferases, including SETDB1. Because transplantation of CBX7-overexpressing CD34+ cells resulted in an increased frequency of myeloid CD33+ cells and primitive CD34+CD38- cells in the bone marrow and CBX7 overexpression resulted in repression of differentiation-associated genes, we explored the role of CBX7 in AML. We show that genetic repression and pharmacological inhibition of *CBX7* in AML cells impairs their proliferation and results in derepression of differentiation-associated genes.

Polycomb CBX proteins are key components of the PRC1 complex, where their function is believed to be essential for the recruitment of PRC1 to H3K27me3-modified genomic loci. Thus, the chromobox domain contained in all CBX proteins is able to recognize H3K27me3 modifications deposited by EZH1 and/or EZH2 as part of the PRC2 complex, which contributes to the repression of target genes. Whereas the *Drosophila* genome contains a single *cbx* gene, during evolution amplification of CBX homologs has occurred in mammals. CBX2, −4, −6, −7, and −8 have all been described to be part of the PRC1 complex, and it is likely that various assemblies of PRC1 have distinct biological targets. In this project, we investigated the role of all five PRC1-CBX proteins in regulating human CD34+ HSPCs. We show that CBX7 is uniquely able to

enhance cell growth of primitive hematopoietic cell subsets. Additionally, transplantation of CBX7-overexpressing CD34+ cells resulted in enhanced long-term engraftment, multi-lineage differentiation potential, and an increased frequency of myeloid CD33+ cells and primitive CD34+CD38− cells in the bone marrow. These results are reminiscent of data of mouse Cbx7, which we reported earlier (Klauke et al., 2013) and established CBX7 as an important evolutionary conserved regulator of self-renewal of human CD34+ HSPCs.

Overexpression of CBX7 resulted in the repression of genes associated with differentiation and led to an upregulation of genes involved in cell cycle and DNA replication. ChIP-seq analysis showed that ~1/3 of the repressed differentiation-associated genes were direct CBX7 targets. Furthermore, many genes, which were upregulated upon overexpression of CBX7, are preferentially expressed by primitive hematopoietic cell subsets, and, thus, are likely to contribute to the maintenance of the primitive phenotype.

Our data indicate that CBX7 regulates the self-renewal activity of primitive cells. As we show that CBX7 represses genes important for differentiation, we hypothesized that CBX7 may also play a role in AML, where self-renewal is enhanced and, conversely, differentiation is repressed. We show that knock down of *CBX7* in leukemic cell lines and primary patient samples affects their proliferation and results in derepression of genes that are normally expressed on differentiated cells.

The molecular mechanism by which CBX7 represses differentiation-inducing genes remains to be elucidated, but our studies strongly suggest that the interplay between the canonical, H3K27me3-mediated, and a non-canonical, H3K9-mediated pathway plays an important role. Whereas *in vitro* the *Drosophila* Polycomb Cbx protein can only recognize H3K27me3 but not H3K9me3, biochemical studies have revealed that multiple mammalian CBX homologs can also bind to H3K9me3 in cell-free systems, each with different binding affinities (Bernstein et al., 2006; Kaustov et al., 2011). So far, no H3K9 methyltransferases were described to interact with CBX proteins *in vivo*. As CBX proteins interact with trimethylated lysine residues on histone proteins via their chromodomain, we hypothesized that CBX proteins might also interact with non-histone proteins harboring a trimethylated lysine embedded in a motif highly similar to histone proteins. Indeed, mass spectrometry analysis revealed multiple of such candidates. Interestingly, all four evolutionarily conserved CBX interacting proteins (EHMT1 [also known as GLP], EHMT2 [also known as G9A], SETDB1, and CDYL) have been shown to physically interact and are strongly associated with H3K9 methylation (Fritsch et al., 2010).

We focused our further studies on the interaction between CBX7 and SETDB1. SETDB1 is an H3K9 methyltransferase that is best known for its role in repressing the expression of endogenous retroviral elements in the genome (Collins et al., 2015). Interestingly, both SETDB1 and CBX7 have been identified as regulators of embryonic stem cell states (Bilodeau et al., 2009), but the role of SETDB1 in hematopoiesis has only recently emerged. Interestingly, mutations in *SETDB1* have been associated with clonal hematopoiesis in elderly individuals (Steensma et al., 2015). Recently, it has been shown that deletion of Setdb1 in murine hematopoietic stem cells results in bone marrow failure (Koide et al., 2016).

Biochemical studies have revealed that the chromodomain of CBX7 has high affinity for a trimethylated 24-amino acid peptide, representing exactly the consensus amino acid sequence of SETDB1 (amino acids 1,157 to 1,181). In fact, the affinity of CBX7 for this sequence is higher for peptides representing the amino acid sequence of H3K27me3 or H3K9me3 (Kaustov et al., 2011). Three lysine residues of SETDB1 have been identified that can be trimethylated (K490, K1170, and K1178), and all could serve as putative binding sites for CBX7 (Hornbeck et al., 2015).

In accordance with the direct *in vivo* interaction between CBX7 and SETDB1, nearly one-third of all CBX7 target loci were simultaneously covered with H3K9me3. In addition, the fact that 62 out of the 95 differentially expressed direct CBX7 target genes associated with differentiation were marked by both CBX7 and H3K9me3, strongly suggesting that self-renewal of human HSPCs is dependent on CBX7-mediated joint repression of target loci by methylation of both H3K27 and H3K9.

Reversely, we demonstrate that proliferation is decreased in leukemic cells when either CBX7 or SETDB1 is downregulated or when CBX7 is pharmacologically inhibited. Similarly, it has been shown that exposure of murine AML blasts to the EHMT2 inhibitor UNC0638 leads to inhibition of growth and induction of myeloid differentiation (Lehnertz et al., 2014).

As for the exact molecular mechanism by which CBX7, SETDB1, and H3K9me3 interact, we hypothesize that such interactions are locus specific and dependent on the composition of the protein complex involved. We speculate that regulation follows a stepwise program, where trimethylated SETDB1 initially converts H3K9me or H3K9me2 into H3K9me3, resulting in attraction of PRC1 by binding of CBX7 to SETDB1. An alternative, not mutually exclusive, possibility is that CBX7 first recognizes trimethylated SETDB1, by which it is then recruited to H3K9me2 loci to ensure further chromatin compaction. These recruitment models would be independent of H3K27me3/PRC2. At loci where both H3K27me3 and H3K9me3 histone marks are present, CBX7 could be recruited to both. It is interesting to note that one of the CBX7-binding proteins we identified, CDYL, can bind to EZH2 as well as to SETDB1 (Escamilla-Del-Arenal et al., 2013; Fritsch et al., 2010; Zhang et al., 2011), allowing for multiple alternative Polycomb and H3K9 methyltransferase interactions. At this point, we do not formally know whether trimethylation of SETDB1 (and EHMT1 and EHMT2) in leukemic cells is required for recognition by CBX7, and if so, how trimethylation of these H3K9 methyltransferases is regulated. Further elucidation of the daunting complexities by which seemingly independent epigenetic pathways converge will allow the understanding of the molecular machinery by which self-renewal is ensured. Disruption of these self-renewal pathways is likely to offer therapeutic opportunities in leukemia.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING

- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Human cord blood samples
  - Primary leukemic cells
- METHOD DETAILS
  - Cloning of retroviral vector constructs
  - Cloning of flag-tagged huSETDB1 cDNA in pRRLA
  - Cloning of short-hairpins in lentiviral expression vectors
  - Transduction of 32D cells
  - Production of a stable retroviral producer cell line (PG13)
  - Production of lentiviral supernatant
  - Retroviral virus production and transduction of CD34+ cells
  - Lentiviral transduction of CD34+ cells
  - Lentiviral transduction of HL60 and OCI-AML3 cells
  - MS37452 treatment of OCI-AML3 cells
  - FACS analysis of HL60 and OCI-AML3 cells
  - Sort of GFP+CD34+ cells (MoFlo Astrios and MoFloXDP)
  - CFU assay
  - Cobblestone area-forming cell assay
  - Long-term culture initiating cell assay
  - Suspension culture experiment with CBX7, CBX8 and empty vector overexpressing cells
  - Xenotransplantation of transduced CD34+ cells 24 hours after last transduction round
  - Xenotransplantation of transduced CD34+ cells after one week *in vitro* culture
  - Bleeding of xenotransplanted NSG mice
  - Bone marrow analysis of NSG mice
  - RNA-Seq of CD34+ cells
  - Chip-Seq of transduced CD34+ cells
  - Mass spectrometry of pull-downs of FLAG-tagged huCBX7, huCBX8, huCBX4 and GFP in K562 cells and FLAG-tagged muCbx7 and GFP in 32d-cells
  - Detection of CBX7 and SETDB1 interaction in HL60 cells by DUOLINK *in situ* proximity ligation assay (PLA)
  - Purification of FLAG-tagged protein
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and eight tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2019.01.050.

ACKNOWLEDGMENTS

The authors wish to thank Geert Mesander, Johan Teunis, and Theo Bijma of the Central Flowcytometry Unit of the University Medical Center Groningen and Dick H. Deckers and Jeroen Demmers of the Mass Spectrometry Facility of Erasmus MC for their excellent support. This study was supported by a grant from the Dutch Cancer Society (RUG 2014-7178), the Netherlands Organization for Scientific Research (Mouse Clinic for Cancer and Aging), and a personal fellowship of the Deutsche Krebshilfe to J.J.

AUTHOR CONTRIBUTIONS

J.J., S.C.B., E.W., B.D.-A., F.H., M.R.D., and J.S. performed research; H.S., S.S.L., S.M.K., K.K., R.A.P., L.V.B., and G.d.H. designed the research; J.J., S.C.B., E.Z., L.V.B., R.A.P., and G.d.H. analyzed data; J.J., S.C.B., R.A.P., L.V.B., and G.d.H. wrote the paper.

DECLARATION OF INTERESTS

The authors have no financial conflict of interest to disclose.

Received: February 14, 2018
Revised: December 19, 2018
Accepted: January 14, 2019
Published February 12, 2019

REFERENCES

Bagger, F.O., Sasivarevic, D., Sohi, S.H., Laursen, L.G., Pundhir, S., Sønderby, C.K., Winther, O., Rapin, N., and Porse, B.T. (2016). BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. Nucleic Acids Res. *44*, D917–D924.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol. Cell. Biol. *26*, 2560–2569.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

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

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell *20*, 845–854.

Collins, P.L., Kyle, K.E., Egawa, T., Shinkai, Y., and Oltz, E.M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. Proc. Natl. Acad. Sci. USA *112*, 8367–8372.

Comet, I., and Helin, K. (2014). Revolution in the Polycomb hierarchy. Nat. Struct. Mol. Biol. *21*, 573–575.

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

Escamilla-Del-Arenal, M., da Rocha, S.T., Spruijt, C.G., Masui, O., Renaud, O., Smits, A.H., Margueron, R., Vermeulen, M., and Heard, E. (2013). Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. Mol. Cell. Biol. *33*, 5005–5020.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870–1881.

Fritsch, L., Robin, P., Mathieu, J.R.R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. Mol. Cell *37*, 46–56.

Herrera-Merchan, A., Arranz, L., Ligos, J.M., de Molina, A., Dominguez, O., and Gonzalez, S. (2012). Retraction: Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. Nat. Commun. *8*, 14005.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. *43*, D512–D520.

Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods *347*, 70–78.

Kahn, T.G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., Reim, I., Pirrotta, V., and Schwartz, Y.B. (2016). Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. Nucleic Acids Res. *44*, 10132–10149.

Kamminga, L.M., Bystrykh, L.V., de Boer, A., Houwer, S., Douma, J., Weersing, E., Dontje, B., and de Haan, G. (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood *107*, 2170–2179.

Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G.A., Li, Z., Vedadi, M., Schapira, M., et al. (2011). Recognition and specificity determinants of the human cbx chromodomains. J. Biol. Chem. *286*, 521–529.

Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., et al. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat. Cell Biol. *15*, 353–362.

Koide, S., Oshima, M., Takubo, K., Yamazaki, S., Nitta, E., Saraya, A., Aoyama, K., Kato, Y., Miyagi, S., Nakajima-Takagi, Y., et al. (2016). Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. Blood *128*, 638–649.

Koya, J., Kataoka, K., Sato, T., Bando, M., Kato, Y., Tsuruta-Kishino, T., Kobayashi, H., Narukawa, K., Miyoshi, H., Shirahige, K., and Kurokawa, M. (2016). DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation. Nat. Commun. *7*, 10924.

Laurenti, E., Doulatov, S., Zandi, S., Plumb, I., Chen, J., April, C., Fan, J.B., and Dick, J.E. (2013). The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat. Immunol. *14*, 756–763.

Lehnertz, B., Pabst, C., Su, L., Miller, M., Liu, F., Yi, L., Zhang, R., Krosl, J., Yung, E., Kirschner, J., et al. (2014). The methyltransferase G9a regulates HoxA9-dependent transcription in AML. Genes Dev. *28*, 317–327.

Cancer Genome Atlas Research Network; Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. *17*, 1823–1828.

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S.A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell *10*, 47–62.

Morey, L., Aloia, L., Cozzuto, L., Benitah, S.A., and Di Croce, L. (2013). RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. Cell Rep. *3*, 60–69.

Nikoloski, G., Langemeijer, S.M.C., Kuiper, R.P., Knops, R., Massop, M., Tönnissen, E.R.L.T.M., van der Heijden, A., Scheele, T.N., Vandenberghe, P., de Witte, T., et al. (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat. Genet. *42*, 665–667.

Ren, C., Morohashi, K., Plotnikov, A.N., Jakoncic, J., Smith, S.G., Li, J., Zeng, L., Rodriguez, Y., Stojanoff, V., Walsh, M., and Zhou, M.M. (2015). Small-molecule modulators of methyl-lysine binding for the CBX7 chromodomain. Chem. Biol. *22*, 161–168.

Rizo, A., Dontje, B., Vellenga, E., de Haan, G., and Schuringa, J.J. (2008). Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. Blood *111*, 2621–2630.

Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P., and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood *126*, 9–16.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat. Cell Biol. *9*, 1428–1435.

Tadokoro, Y., Ema, H., Okano, M., Li, E., and Nakauchi, H. (2007). De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. J. Exp. Med. *204*, 715–722.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell *148*, 664–678.

van den Boom, V., Rozenveld-Geugien, M., Bonardi, F., Malanga, D., van Gosliga, D., Heijink, A.M., Viglietto, G., Morrone, G., Fusetti, F., Vellenga, E., and Schuringa, J.J. (2013). Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells. Blood *121*, 2452–2461.

Verhaak, R.G., Wouters, B.J., Erpelinck, C.A., Abbas, S., Beverloo, H.B., Lugthart, S., Löwenberg, B., Delwel, R., and Valk, P.J. (2009). Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica *94*, 131–134.

Zhang, Y., Yang, X., Gui, B., Xie, G., Zhang, D., Shang, Y., and Liang, J. (2011). Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. J. Biol. Chem. *286*, 42414–42425.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Human BD Fc block	BD Bioscience	564220
CD16/ CD32 Mouse BD Fc block	BD Bioscience	553142
Mouse anti human CD3 APC-Cy7 (SK7)	BD Bioscience	557832
Mouse anti human CD4 PE-Cy7 (SK3)	BD Bioscience	557852
Mouse anti human CD33 BV421(WM53)	BD Bioscience	562854
Mouse anti human CD45 APC (HI30)	BD Bioscience	560973
Mouse anti human CD19 PE (HIB19)	BD Bioscience	555413
BD Horizon Brilliant Stain buffer	BD Bioscience	563794
Mouse anti human CD38 PE (HIT2)	BD Bioscience	555460
Mouse anti human CD90 AF700 (5E10)	BioLegend	328120
Mouse anti human CD34 APC (581)	BD Bioscience	555824
Mouse anti human CD45 RA PE-Cy7 (L48)	BD Bioscience	337186
Mouse anti human CD45 BV421 (HI30)	BioLegend	304032
Mouse anti human CD2 PE-Cy5 (RPA-2.10)	BioLegend	300210
Mouse anti human CD3 PE-Cy5 (UCHTI)	BioLegend	300410
Mouse anti human CD4 PE-Cy5 (RPA-T4)	BioLegend	300510
Mouse anti human CD7 PE-Cy5 (6B7)	BioLegend	343110
Mouse anti human CD8 PE-Cy5 (RPA-T8)	BioLegend	301010
Mouse anti human CD19 PE-Cy5 (HIB19)	BioLegend	302210
Mouse anti human CD20 PE-Cy5 (2H7)	BioLegend	302308
Mouse anti human CD235a PE-Cy5 (HIR2)	BioLegend	306606
Mouse anti human CD11b PE-Cy5 (ICRF44)	BioLegend	301308
Mouse anti human CD14 PE-Cy5 (TuK4), TRI-COLOR	ThermoFisher Scientific	MHCD1406
Mouse anti human CD56 PE-Cy5 (MEM-188)	BioLegend	304608
Mouse Anti-Human Alexa Fluor 700 CD14 Clone M5E2 (RUO)	BD Bioscience	557923
Mouse Anti-Human BV421 CD11b/MAC-1 (RUO)	BD Bioscience	562632
Human anti CD34 PE-Cy7 (8G12)	BD Bioscience	348811
H3K9me3 polyclonal antibody- Premium	Diagenode	C15410193
Anti H3K27me3	Merck	07-449
Anti CBX7	Merck	07-981
Pierce Protein A/G Magnetic beads	Thermo Scientific	#88803
Monoclonal ANTI-FLAG M2 antibody	Sigma	F3165-2 mg
Polyclonal rabbit anti human/mouse CBX7 p15	Santa Cruz Biotechnology	SC 70-232
SETDB1 Antibody (5H6A12)	Pierce Protein	MA5-15722
Biological Samples		
Human Cord blood sampes	Department Obstetrics, Isala Hospital Zwolle	N/A
Primary AML samples	Department of Hematology, UMCG	N/A
Chemicals, Peptides, and Recombinant Proteins		
IMDM 2% FCS	StemCell Technologies	# 07700
MethoCult H4435 Enriched	StemCell Technologies	#04435 and 04445
Myelocult H5100	StemCell Technologies	#05100 and 05150
StemSpan SFEM	StemCell Technologies	# 09650

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hydrocortisone	StemCell Technologies	#07904
Human recombinant TPO	RnD	288-TP-200
Human recombinant FLT3L	RnD	308-FK-025
Human recombinant SCF	RnD	255-SC-050
Human recombinant IL3	Sigma	I1646
Mouse recombinant IL3	RnDSystems	403-ml-50
RetroNectin 2,5 mg	Westburg/ Takara	T100/B
Hexadimethrine bromide	Sigma	h-9268
FuGene HD transfection reagent	Promega	E2312
Formaldehyde, 37% formaldehyde solution	Santa Cruz Biotechnology	Sc-203049
SDS Solution, 20%	Fisher Scientific	BP 1311-1
cOmplete, EDTA-free Protease Inhibitor Cocktail	Merck	#000000011873580001
QIAquick PCR Purification Kit	Quiagen	#28104
Novagen Benzoase Nuclease, Purity 99%	Merck Millipore	70664
SnakeSkin Dialysis Tubing, 7K MWCO, 22mm	Thermo Scientific	68700
3xFlag peptide	Sigma	F3290-4MG
CryoStor CS10	StemCell Technologies	#07930
Protease Inhibitor Cocktail	Sigma	P8340
Anti-Flag M2 Magnetic beads	Sigma	M8823
NT Protein Labeling kit Red-NHS	Nanotemper	MO-L001
MS37452	Sigma	SML1405
Deposited Data		
RNA Seq data	This paper	https://www.ebi.ac.uk/ena/data/view/PRJEB22831
ChIP Seq data	This paper	https://www.ebi.ac.uk/ena/data/view/PRJEB22344
Experimental Models: Cell Lines		
MS5 cells	DSMZ	ACC 441
HL60 cells	ATCC DSMZ	CCL-240 ACC 3
OCI-AML3 cells	DSMZ	ACC 582
32D cells		Kind gift from Ivo Touw
PG13 cells	ATCC	CRL-10686
Phoenix-ECO cells	ATCC	CRL-3214
K562 cells	ATCC	CCL-243
293FT	ThermoFisher Scientific	R70007
Experimental Models: Organisms/Strains		
NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ	Mice were purchased from Charles River Laboratory (L'Arbesle Cedex, France) and bred in house.	N/A
Oligonucleotides		
See Table S8		N/A
Software and Algorithms		
Graphpad Prism (v5-7)	Graphpad Prism	https://www.graphpad.com
ELDA	(Hu and Smyth, 2009)	http://bioinf.wehi.edu.au/software/elda/
FlowJo	Version X.0.7	
Other		
CD34+ MicrobeadKit	Miltenyi Biotec	130-056-702
Lymphoprep	Stem cell technologies	7861
LS Columns	Miltenyi Biotec	130-042-401

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gerald de Haan (g.de.haan@umcg.nl).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

Mouse experiments were performed in line with international and national guidelines. All experiments were approved by the Institutional Animal Care and Use Committee of the University of Groningen (IACUC-RUG).

For all xenotransplantation studies, we performed single cord transplantations of freshly isolated CD34+ cord blood cells. Female 11-22 weeks old NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice were radiated three hours before transplantation with 1.8 Gy. In each experiment age of mice was balanced with maximum 2 weeks of difference between the experimental and control group. No antibiotic prophylaxis after radiation was given.

Human cord blood samples

CD34+ cord blood isolation

Cord blood was obtained from healthy full-term pregnancies after informed consent in accordance with the Declaration of Helsinki from the Obstetrics Department at the Isala Hospital in Zwolle, the Netherlands. Initially, cord blood volume and cell counts were measured and then diluted 1:1 with PBS+ 2mM EDTA+0.5% BSA. Maximum 30 mL of diluted cord blood was carefully layered on 15 mL of Lymphoprep in a 50ml falcon tube and centrifuged for 20 minutes, 800 g, without brakes. Middle layer containing mononuclear cells was harvested and diluted 1:1 with PBS 2mM EDTA 0.5% BSA and then centrifuged for 5 minutes at 800 g. Cell pellets were collected and washed with PBS 2mM EDTA 0.5% BSA and centrifuged for 10 minutes at 200 g. Immunomagnetic labeling and separation were performed according to the manufacturer's manual of the CD34 MicroBead Kit, human (Miltenyi Biotec). Cells were either used immediately for experiments or frozen in Cryostor CS10.

Cell lines

Phoenix-ECO cells were culture in DMEM +1% P/S + 10% heat-inactivated FCS. PG13 cells were culture in DMEM +1% P/S + 10% heat-inactivated FCS. HEK293FT cells were culture in DMEM +1% P/S + 10% heat-inactivated FCS. HL60 cells were cultured in RPMI+1%P/S+20% heat-inactivated FCS. OCI-AML3 cells were cultured in RPMI+1%P/S+10% heat-inactivated FCS. K562 cells were cultured in RPMI+1%P/S+10% heat-inactivated FCS.

Primary leukemic cells

Primary AML cells were provided by dr. JJ Schuringa from the Department of Hematology, University Medical Center Groningen after informed consent.

METHOD DETAILS

Cloning of retroviral vector constructs

The consensus cDNA of CBX2,4,6,7 and 8 and FLAG-tagged versions of the cDNA were inserted in the retroviral vector backbone of SF91-IRES-GFP (Klauke et al., 2013) upstream of IRES by PCR based cloning using Not1 and Sal1 restriction sites. Primers used for PCR based cloning are listed above. FLAG-tagged GFP vector was cloned by vector-PCR of SF91-FLAG tagged muCbx7 (Klauke et al., 2013) with MluI restriction site containing primers and subsequent ligation.

Cloning of flag-tagged huSETDB1 cDNA in pRRLA

A FLAG-tagged versions of SETDB1 cDNA was inserted in the lentiviral vector backbone of pRRLA IRES-GFP upstream of IRES by PCR based cloning using Mlu1 and Xba1 restriction sites. Primers used for PCR based cloning are listed above.

Cloning of short-hairpins in lentiviral expression vectors

Corresponding oligos for SCR, shCBX7, shSETDB1#3 and shSETDB1#4 were annealed and cloned into the empty pLKO.1_mCherry vector upon digestion with Age1 and EcoR1.

Transduction of 32D cells

Initially, 300,000 Phoenix-ECO cells/well (of a six-well plate) were seeded in DMEM+1%P/S+10%FCS. On day 2 cells were transfected with 1 μg of plasmid with the help of FuGene® in a 1:3 ratio. 24 hours after transfection medium was changed to RPMI+10%FCS+1%P/S. On day 4 non-treated six- well plates were coated with RetroNectin® according to the manufacturer's manual. Viral supernatant was harvested and filtered through a sterile syringe filter with a 0.45 μm pore size hydrophilic PVDF membrane. Then 2 mL of viral supernatant, 300000 32D cells and hexadimethrine bromide (2 μg/ml) and muIL3 (10 ng/ml) were

added/well. Six-well plates were centrifuged for 45 minutes at room temperature for 45 minutes at 400 G. 24 hours after transduction virus-supernatant was replaced by RPMI+10%FCS+1%P/S+ muIL3 (10 ng/ml).

Production of a stable retroviral producer cell line (PG13)

Initially, 300,000 Phoenix-ECO cells/well (of a six-well plate) were seeded in DMEM+1%P/S+10%FCS. On day 2 cells were transfected with 1 μg of plasmid with the help of FuGene® in a 1:3 ratio. 24 hours after transfection medium was refreshed and 10,000 PG13 cells were plated out/ well in a tissue culture treated six-well plate. 48 hours after transfection viral supernatant was harvested and used to transduce the retroviral packaging cell line PG13 with the help of Hexadimethrine bromide (2 μg/ml). One day after transduction medium of transduced PG13 was changed to DMEM+1%P/S+10%FCS and cultured at 37°C/5% CO2.

Production of lentiviral supernatant

2.75 10^6 293FT cells were plated in gelatin coated cell-culture treated 10 cm dishes in DMEM+10%FCS+1%P/S and incubated overnight at 37°C/5%CO_2. On the next day cells were transfected with 3 μg of the pLKO.1 or pRRLA vector, 3 μg of the packaging plasmid pCMV8.91 and 0.7 μg of envelope plasmid VSV-G and 21 μl of FuGene®. On the next day medium was changed to either StemSpan SFEM or RPMI. Two days after transfection the virus was collected, filtered through a sterile syringe filter with a 0.45 μm pore size hydrophilic PVDF membrane and used either immediately for transduction or was frozen.

Retroviral virus production and transduction of CD34+ cells

24 hours before the first transduction round CD34+ cells were prestimulated in StemSpan SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml at 37°C and 5% CO2. Medium of transduced PG13 cells was changed to StemSpan SFEM. On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin® according to the manufacturer's manual.

Then viral supernatant of virus-producing PG13 cells was harvested and filtered through a sterile syringe filter with a 0.45 μm pore size hydrophilic PVDF membrane. Between 500,000 and 1,000,000 CD34+ cells were transduced with 2 mL of viral supernatant in the presence of SCF 100 ng/ml, FLT3L 100 ng/ml, TPO 100 ng/ml and Hexadimethrine bromide to a final concentration of 2 μg/ml. Six-well plates were centrifuged at 400 g for 1 hour at room temperature. Transduction was repeated two times in 8-12 hour time intervals. After last transduction round medium was changed to StemSpan SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml.

Lentiviral transduction of CD34+ cells

CD34+ cells were cultured in StemSpan SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml 24 hours before first transduction round at 37°C and 5% CO_2. On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin according to the manufacturer's manual. Lentiviral supernatant was thawn on ice. Between 500,000 and 1,000,000 CD34+ cells were transduced with 2 mL of viral supernatant in the presence of SCF 100 ng/ml, FLT3L 100 ng/ml, TPO 100 ng/ml and Hexadimethrine bromide 2 μg/ml. Six-well plates were centrifuged at 400G for 1 hour at room temperature. Transduction was repeated once in 8-12 hour time intervals. After last transduction round medium was changed back to StemSpan SFEM containing SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml.

Lentiviral transduction of HL60 and OCI-AML3 cells

On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin according to the manufacturer's manual. Between 300,000 and 500,000 cells were transduced in 2 mL of viral supernatant containing Hexadimethrine bromide 2 μg/ml. Six-well plates were centrifuged at 400G for 1 hour at room temperature. Transduction was repeated once in 8-12 hour time intervals. After last transduction round medium was changed back to RPMI+1%P/S+10% (OCI-AML3) or 20% of FCS (HL60). At several time points cells were counted manually with a hemacytometer.

MS37452 treatment of OCI-AML3 cells

Initially, 500.000 OCI-AML3 cells/well (of a six-well plate) were seeded in RPMI+1%P/S+10% heat-inactivated FCS supplemented with MS37452 (dissolved in DMSO at a concentration of 50μM) at different concentrations. After four days cells were counted manually using a hemacytometer.

FACS analysis of HL60 and OCI-AML3 cells

Samples were incubated with Human BD Fc block to prevent unspecific binding at 4°C in the dark. After blocking, 5 μl mouse anti huCD11b BV421 and/or 5 μl mouse anti huCD14 Alexa Fluor 700 antibodies were added and samples were incubated for 20-25 minutes at 4°C in the dark. Afterward, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II.

Sort of GFP+CD34+ cells (MoFlo Astrios and MoFloXDP)

24 hours after the last transduction round cells were harvested, washed and resuspended in PBS+BSA 0.2%. Cells were incubated with Human BD Fc block for 15 minutes according to the manufacturer's manual to prevent unspecific binding. After blocking, 5 μl of

mouse anti huCD34 PE-Cy7 was added and incubated at 4 degrees for 20 minutes. Cells were washed with PBS+BSA 0.2% and resuspended in PBS+BSA 0.2% with the viability dye (PI).

CFU assay

All experiments with CD34+ cells were performed with single (not pooled) cords (except Chip-Seq experiments). 5,000 CD34+GFP+ cells were sorted in a FACS tube containing 1 mL of IMDM 2% FBS. 0.3 mL of the sorted cells in IMDM 2%FCS were added to a pre-aliquoted 3 mL MethoCult tube. Afterward the tube was vortexed for at least 4 s and then let stand for a minimum of 5 minutes. For dispensing the MethoCult mixture into 35 mm culture dishes a 3 mL syringe with a 16 gauge Blunt-End Needle was used to add 1.1 mL per dish. Dishes were cultured at 37°C / 5% CO2 conditions. Colonies were counted and typed after 14 days. For replating cells from primary dishes were harvested, centrifuged and counted with a hemocytometer. For CBX7 5,000 cells and for CBX8 30,000 cells were plated out as described above. Control cells were plated out at same cell numbers as experimental groups. Figure 1 A+B show single data points of each experiment. Each single experiment was performed in duplicates and the average of the technical replicates was plotted.

Cobblestone area-forming cell assay

96-well flat-bottom plates were pre-coated with 0.1% gelatin. Two days before sort 10,000 MS5 cells were plated in 200 μl of Myelocult H5100 supplemented with 10^{-6} M hydrocortisone +1% P/S. On the next day cells were radiated with 30 Gy. One day post radiation CD34+GFP+ cells were sorted directly into 96-well plates at limiting dilution and cultured for 5 weeks with weekly performed half-medium changes. Cobblestones were analyzed with a phase-contrast microscope. Frequency of each experiment was calculated with ELDA software. (Hu and Smyth, 2009) Figure 1D shows single data points of each experiment. (The y axis indicates the number of cells that need to be plated for a CAFC to develop.)

Long-term culture initiating cell assay

96-well flat-bottom plates were pre-coated with 0.1% gelatin. Two days before sort 10,000 MS5 cells were plated out in 200 μl of Myelocult H5100 supplemented with 10^{-6} M hydrocortisone.

Cells were sorted directly into 96-well plates at limiting dilution and cultured for 5 weeks with weekly performed half-medium changes. After 5 weeks medium was replaced by Methocult H4335 and incubated for two further weeks at 37°C/5% CO2. Colony-formation was accessed by phase-contrast microscopy. Frequency of each experiment was calculated with ELDA software.

(The y axis indicates the number of cells that need to be plated for a LTC-IC to develop)

Suspension culture experiment with CBX7, CBX8 and empty vector overexpressing cells

100,000 CD34+GFP+ cells were sorted in a six-well plate containing 2 mL of Myelocult H5100, 1% P/S, 10^{-6} M hydrocortisone, TPO 100 ng/ml, IL3 50 ng/ml, SCF 100 ng/ml, FLT3L 100 ng/ml. Cells were cultured at 37°C/5% CO2. Cells were counted manually with a hemacytometer every week and 100,000 cells were re-seeded under the same conditions.

Xenotransplantation of transduced CD34+ cells 24 hours after last transduction round

Isolation of CD34+ cells and transduction was performed as described above. 24 hours after transduction the percentage of CD34+GFP+ cells was determined after Fc blocking and staining with CD34+PE-Cy7 as described above. Cells were harvested and counted manually with a hemocytometer and trypan blue and resuspended in PBS. In total equivalents of 200,000 CD34+GFP+ cells were transplanted per mouse via retro-orbital injection. A small aliquot was kept in culture for determining the exact number of transplanted CD34+GFP+ cells 24 hours later. Because the transduction efficiency of experimental group was always lower than in the control group only absolute percentages of GFP engraftment are illustrated.

Xenotransplantation of transduced CD34+ cells after one week *in vitro* culture

Isolation of CD34+ cells was performed as described above. After isolation, cells were cultured in StemSpan SFEM exposed to FLT3L, TPO and SCF each with 100 ng/ml. Then three transduction rounds in a 24 hours time intervals were performed. 24 hours after the last transduction round cells were cultured for further 96 hours, so that cells were after isolation for one week *in vitro* cultured before transplantation.

On the day of transplantation percentage of CD34+GFP+ cells was determined after Fc blocking and staining with CD34+PE-Cy7 as described above. Cells were harvested and counted manually with a hemocytometer and trypan blue and resuspended in PBS. In total equivalents of 1.5 million of CD34+GFP+ cells were transplanted per mouse via retro-orbital injection under general anesthesia. Because transduction efficiency of experimental group was always lower than in the control group only absolute percentages of GFP engraftment are illustrated.

Bleeding of xenotransplanted NSG mice

Beginning 6-weeks after transplantation chimerism in PB was determined in 4-week intervals. Blood samples were taken under general anesthesia via retro-orbital bleeding. Blood was lysed with ammonium chloride, washed two times with PBS+BSA 0.2% and resuspended in 50 μl of PBS+BSA 0.2%. Samples were incubated with Human BD Fc block and CD16/ CD32 Mouse BD Fc block

to prevent unspecific binding. After incubation for 10 minutes at room temperature antibody master mix and BV stain buffer was added and samples were incubated for 20-25 minutes at 4°C in the dark. Afterward, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II.

Bone marrow analysis of NSG mice

Mice were sacrificed and dissected under general anesthesia after the end of the experiment or reaching the human endpoint of the experiment. Bones (Femur, tibia, fibula and pelvis) were collected and cleaned. Bones were crushed in the presence of PBS+0.2% PBS with a mortar and pestle. The obtained cell suspension was filtered through a 40 μm filter. Remaining erythrocytes were lysed with ammonium chloride. Cells were then pelleted by centrifugation and resuspended in PBS+BSA 0.2%. For preventing unspecific binding samples were incubated with Human BD Fc block and CD16/ CD32 Mouse BD Fc block.

After incubation for 10 minutes at room temperature, antibody master mix and BV stain (only for lineage-specific staining) buffer was added and samples were incubated for 20-25 minutes at 4°C in the dark. Afterward, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II (Lineage staining) and LSRII (progenitor staining).

RNA-Seq of CD34+ cells

100,000 GFP+ CD34+ cells were sorted into lysis buffer 4 days post-transduction. RNA was extracted using the Nucleospin RNA XS Kit, with the addition of a second elution step to increase yield. RNA quality and quantity was assessed using the Bioanalyzer RNA total Pico Assay. RNA samples with an RNA-integrity > 8 were processed for RNA-seq library preparation using the SMARTer Stranded Total RNA-seq Kit. Briefly, 10 ng of total RNA was reverse transcribed using random primers and amplified via PCR during which barcoded Illumina adapters were added. After amplification, ribosomal RNA and mitochondrial cDNA were removed by annealing specific R-probes, resulting in cleavage of ribosomal and mitochondrial cDNA in the presence of ZapR. After cleavage of ribosomal and mitochondrial cDNA, the remaining cDNA was amplified again during another round of PCR. The nM concentration of RNA-seq libraries were quantified based on library size (Bioanalyzer) and cDNA concentration (Qubit) and normalized to 2nM prior to pooling. RNA-sequencing was performed on an Illumina HiSeq 2500 machine, three single-end runs with a read length of 63-64nt, resulting in fastq samples consisting of 26 to 64 million reads. Sample mapping was done with STAR (version 2.5.1b-2.5.2b), a custom genome index was build using Genecodegenes.org release 24 (GRCh38.p5) Ensembl 83, December 2015. STAR outputs read counts per gene, these where filtered by removing ribosomal and transfer RNA. Differential expression (DE) analyses by EdgeR (version 3.16.2), with upper quartile normalization. DE genes per condition ranked by p value and adjusted for multiple testing using a Benjamini- Hochberg method. The final DE gene lists were filtered by FDR < 0.05.

The RNA-seq data are deposited at ENA (PRJEB22831).

Gene Annotation and GO search was done using String database (https://string-db.org). Corresponding annotated and GO files were downloaded from the site and further analyzed with custom scripts. GOChord plots were done in R (GOplot packages).

RNA expression data from Laurenti et al., (2013), were obtained from GEO (GSE42141). Expression heatmap was done using R using heatmap.2 function in gplot library. Data were clustered by correlation and z-transformed. Gene Set-Enrichment Analysis was performed on a preranked list of all differentially expressed genes (FDR < 0.1). The number of permutations was set to 1000, with exclusion of filtersets < 10. (Furthermore, we applied the following filters scoring_sheme weighted, norm meandiv, make_sets true, gui false, set_max 500, set_min 15, npmerm 1000)

Chip-Seq of transduced CD34+ cells

Frozen human CD34+ enriched cord blood cells were thawn, pooled (Batch 1 = 21 cords, Batch 2 = 21 cords, Batch 3 = 31 cords) and cultured in StemSpan SFEM with TPO, SCF and FLT3L (each 100 ng/ml) for prestimulation. After 24 – 48 hours three transduction rounds with retroviral supernatant for either EV or CBX7 at 24-hour time intervals were performed. Transduction was performed as described above. After the final transduction round, cells were further expanded in StemSpan SFEM with TPO, SCF and FLT3L (each 100 ng/ml). One week after thawing, cells were stained and CD34+GFP+ cells sorted as described above. Sorted cells were washed in ice-cold PBS+BSA 0,2%, centrifuged (450G, 5 min, 4°C) and resuspended in 1% cold formaldehyde for fixation. Tubes were incubated on a rotator at 4°C for 10 minutes. Fixation was stopped by adding glycine to a final concentration of 0.125M. After addition of glycine cells were incubated on a rotator at 4°C for 5 minutes, washed two times with cold PBS, transferred to a low-adherent tube and resuspended in SDS buffer (NaCl 100mM, Tris-Cl pH8.1 50 mM, NaN_3 0.2%, 0.5% SDS) + cOmplete protease inhibitors (1 tablet/ 50 mL SDS buffer). Samples were snapfrozen on dry ice and then transferred to −80°C for storage. The following table describes the number of transduced CD34+ cells/ experiment and per antibody in million of cells.

	H3K9me3	*H3K27me3*	*CBX7*	*IgG*
1st Experiment	1.4	1.4	2	0.2
2nd Experiment	0.65	0.8	0.8	0.125
3rd Experiment	1	1	2	0.2

For chromatin-immunoprecipitation, samples were thawn, centrifuged for 5 minutes at 900 g at room temperature and the supernatant was discarded. Pellets were resuspended in 500 μl of IP buffer (30 mL SDS buffer + 15 mL Triton dilution buffer (100mM Tris-Cl pH8.6 + 100mM NaCl + 5mM EDTA ph 8.0 + 0.2% NaN_3 + 5%Triton X-100 + cOmplete protease inhibitors) and sonicated to an average length of 400-500 bps (Bioruptor 30 s on/ 30 s off / cycle, high, in total 3 cycles). 5% of each sample was reversed crosslinked in TE with 1% SDS and 200 mM NaCl overnight at 65°C. On the next day decrosslinked DNA was isolated with the QIAquick PCR purification kit (QIAGEN) and appropriate fragment length was confirmed via agarose gel electrophoresis.

Protein A/G magnetic beads were washed three times with cold PBS and once with cold IP buffer. Crosslinked and fragmented samples were thawn and centrifuged for 30 min at 17,000 g at 4°C. Samples were precleared by rotating at 4°C with 7.5 μl of washed beads for 1 hour. Samples were then distributed into several low adherence tubes for incubation with 5 μg of antibody and incubated overnight at 4°C on a rotating platform. The next day, samples were incubated with 20 μl of washed beads for 4 hours on a rotating platform at 4°C. Using a magnetic stand, the supernatant was removed and beads were washed 4x in a low salt buffer (150 mM) and once with TE. After the last washing step samples were incubated in 1% SDS, 200 nM NaCl overnight at 65°C with 1100 RPM to reverse crosslinks. DNA was isolated with QIAquick PCR purification kit and enrichment of positive and negative loci was confirmed via qPCR.

ChIP libraries for sequencing were prepared with the Microplex Library Preparation Kit V2 (Diagenode, C05010012) according to the manufacturers protocol. The concentration of individual Chip-seq samples was determined based on library size (Agilent, Bioanalyzer 2100) and DNA concentration (Thermo Fisher, Qubit) and diluted to 2nM prior to pooling.

ChIP-Seq was performed on an Illumina NextSeq 500 machine, paired-end 79-80 bp, a custom genome index was build using Genecodegenes.org release 25 (GRCh38.p7) Ensembl 85, July 2016. BAM mem (version 0.7.15) produced BAM files were generated and processed with Samtools (version 1.3.1). MACS2 (version 2.1.0) with the settings -f BAMPE –-nomodel –-broad –broad-cutoff -g 2.7e9 –kep-dup 1. The output BED files where analyzed using bedtools (version 2.26.0) functions intersect, closest and Deeptools (version 2.4.2).

The Chip-seq data are deposited at ENA (PRJEB22344).

Mass spectrometry of pull-downs of FLAG-tagged huCBX7, huCBX8, huCBX4 and GFP in K562 cells and FLAG-tagged muCbx7 and GFP in 32d-cells

32D cells and K562 cells were transduced as described above, sorted and expanded. Cells were harvested and washed. For each experiment 2-4 mL of cell pellets were used. Cell pellets were resuspended in 4-5 pellet volumes of ice-cold buffer A (10mM HEPES ph 7.6, 1.5mM MgCl2 and 10mM KCl, 0.5 mM DTT+ complete protease inhibitor), lysed for 10 minutes on ice and centrifuged for 10 min at 3000 rpm. Pellets were resuspended in 2 pellet volumes of buffer A. The cell suspension was homogenized with a Dounce Homogenizer with 10 strokes (pestle A). The homogenized suspension was centrifuged for 10 minutes with 3000 rpm. Supernatant was removed and pellets were centrifuged for 1 minute with 3000 rpm. Supernatant was removed and pellets were resuspended in 1.5 pellet volumes of buffer C (20 mM HEPES ph 7.6, 20% glycerol, 420 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM DTT, complete protease inhibitor). Nuclei suspension was homogenized with Dounce homogenizer with 10 strokes (pestle B). The suspension was then rotated on a rotor suspension at 4°C for 30 minutes and afterward centrifuged for 15 min with maximal speed. Nuclear extracts were dialyzed with a SnakeSkin Dialysis tube to buffer D (20 mM HEPES [pH 7.6], 0.2 mM EDTA, 1.5 mM $MgCl_2$, 100 mM KCl, 20% glycerol, 0,5 mM DTT, 0.2 mM PMSF, 9.5 mg/l sodium Metabisulfite).

60 μL of anti-FLAG M2 agarose beads (Sigma) equilibrated and washed in buffer C-100 (20 mM HEPES ph 7.6, 20% glycerol, 1.5 mM MgCl2, 100 mM KCl, 0.2 mM EDTA, 0,02% NP40, 0.5 mM DTT, complete proteaseinhibitor) were added to 1.5 mL of nuclear extract in low-adherence microcentrifuge tubes and incubated for 3 hr at 4°C in the presence of 225 units of Benzonase (Novagen) on a rotator. Afterward beads were washed 5-times with buffer C-100. Bound proteins were eluted four times with 60 μl of 3xFLAG-peptide solution (buffer c-100 + 0.2 mg/ml 3xFLAG peptide) for 15 minutes at 4°C. Efficiency of co-immunoprecipitation and elution was checked via western blot with staining against FLAG. Most efficient elutions were pooled, TCA precipitated, and proteins separated by polacrylamide gel electrophoresis stained with (Invitrogen). MS was done in a label free format. For each identified peptide and protein spectral counts were calculated using commercial PEAK studio software, using standard filtering settings. Further, data were merged for all 10 slices into a single table, the sum of all spectral counts (per slice) was taken as a measure of the protein amount in the sample. Data were grouped into three experiments (Human, mouse 1 and mouse 2), sorted by spectral count abundance. The last 10% of the least abundant proteins were removed from each pull down list. The rest of the proteins were ranked in relative scale (0 for the most abundant to 0.9 as the least abundant). The cumulative abundance rank index was calculated for every candidate protein by subtracting each abundance rank index from the control rank index (GFP).

The list of proteins with K-Me3 modifications was downloaded from Phosphosite database. Information for proteins abundance was downloaded from PaxDB site average abundance across all human samples was used for this analysis.

Cross comparison of the gene/protein lists from different databases was done using custom scripts. All illustrations were prepared in R and Python using standard graphic packages.

For identification of trimethylated interaction partners of CBX7 we only considered proteins which were in the top 20% of the relative ranked ordered CBX7-binding proteins and which have an abundance score below 100 (ppm) over all cell types and whose relative ranked score in the experimental sample is higher than in the control sample.

To confirm interaction of CBX7 and SETDB1 in K562 cells we performed Co-IP in FLAG-tagged CBX7 overexpressing K562 cells as described above and performed SDS-PAGE. After protein transfer membrane was stained against CBX7 and SETDB1.

Detection of CBX7 and SETDB1 interaction in HL60 cells by DUOLINK *in situ* proximity ligation assay (PLA)

5×10^4 HL60 cells were fixed in ice-cold methanol for 5 minutes on cytospin slides. Interaction of the endogenous CBX7 and SETDB1 proteins in HL60 cells was assessed using the Duolink *in situ* Proximity Ligation Assay (PLA) (Olink Bioscience, Uppsala, Sweden), as described by the manufacturer. Used antibodies:

- Polyclonal rabbit anti human/mouse CBX7 p15 Santa Cruz Biotechnology, SC 70-232
- anti- SETDB1 Antibody (5H6A12) Pierce Protein, MA5-15722

Purification of FLAG-tagged protein

Retroviral supernatant for overexpressing flag-tagged huCBX7 and lentiviral supernatant for overexpressing flag-tagged huSETDB1 was produced as described above. 293FT cells were transduced with two transduction rounds and expanded.

FLAG-tagged SETDB1 and CBX7 was purified using anti-Flag M2 magnetic beads (Sigma). Frozen pellets were resuspended in 10 vol (v/v) lysis buffer (50 mM Tris-HCl pH 7.5, 600 mM NaCl, 2 mM EDTA, 1.0% NP40 (v/v), 1:1000 μl Protease Inhibitor Mix (Sigma-Aldrich). Lysates were incubated for 30 min on ice while being inverted (3 times) every 5 min. After incubation the lysate was forced 10 times through a 26 g needle. Debris was pelleted by centrifugation (20000 g, 30 min, 4°C) and the supernatant transferred to a new tube. 1 original pellet vol. (v/v) anti-Flag M2 magnetic beads (Sigma-Aldrich) was washed 4 times with 1 mL TBS (50 mM Tris-Cl, pH 7.6; 150 mM NaCl) and 1 time with lysis buffer. The supernatant and the washed beads were combined and antigen capture was performed for 4 hr at 4°C using a head-over-tail rotator (HOT). After incubation, samples were placed on a magnetic stand for 1 min. The supernatant was removed and the bead-bound protein washed 3 times with 1ml lysis buffer (5 min HOT, 1 min magnetic stand) and 3 times with 1 mL wash buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM EDTA, 0.1% NP40 (v/v)). The FLAG-tagged protein was eluted by adding 150 μl elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 3xFlag-Peptide 1mg/ml) and incubation for 5 hr at 4°C. 20 μl 85% glycerol were added to the eluted protein. Proteins were stored at −20°C.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details for each experiment are indicated in the legend of each figure, where appropriate. Statistical analyses of mass spectrometry data is provided in Table S6. Software used is provided in the Key Resource Table.