nature cell biology

ARTICLES

Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation

Karin Klauke[1], Višnja Radulović[1], Mathilde Broekhuis[1], Ellen Weersing[1], Erik Zwart[1], Sandra Olthof[1], Martha Ritsema[1], Sophia Bruggeman[1], Xudong Wu[2], Kristian Helin[2], Leonid Bystrykh[1] and Gerald de Haan[1,3]

The balance between self-renewal and differentiation of adult stem cells is essential for tissue homeostasis. Here we show that in the haematopoietic system this process is governed by polycomb chromobox (Cbx) proteins. Cbx7 is specifically expressed in haematopoietic stem cells (HSCs), and its overexpression enhances self-renewal and induces leukaemia. This effect is dependent on integration into polycomb repressive complex-1 (PRC1) and requires H3K27me3 binding. In contrast, overexpression of Cbx2, Cbx4 or Cbx8 results in differentiation and exhaustion of HSCs. ChIP-sequencing analysis shows that Cbx7 and Cbx8 share most of their targets; we identified approximately 200 differential targets. Whereas genes targeted by Cbx8 are highly expressed in HSCs and become repressed in progenitors, Cbx7 targets show the opposite expression pattern. Thus, Cbx7 preserves HSC self-renewal by repressing progenitor-specific genes. Taken together, the presence of distinct Cbx proteins confers target selectivity to PRC1 and provides a molecular balance between self-renewal and differentiation of HSCs.

Mature blood cells have a limited lifespan and are continuously replenished by the activity of HSCs, which have the potential to undergo both self-renewal and differentiation. Identifying the molecular basis of this process will be of great value for a fundamental understanding of normal and malignant haematopoiesis and holds promise for stem cell-based therapies in the future.

Proteins that maintain or modify the transcriptional profile of stem cells by organizing chromatin structure can be considered as key epigenetic candidates. Polycomb group (PcG) proteins were first discovered as important developmental regulators in *Drosophila*[1]. According to the classical PcG model, they assemble into at least two complexes, PRC1 and PRC2, which collaborate to repress gene transcription by catalysing histone modifications[2,3]. PRC2 is comprised of three core components (Ezh, Suz12 and Eed), of which the SET-domain containing Ezh protein contributes to gene repression by tri-methylating histone H3 on Lys 9 and 27 (H3K9me3, H3K27me3). PRC1 contains Pcgf (for example Bmi1 and Mel18), Phc, Ring1 and Cbx, and catalyses mono-ubiquitylation of H2AK119.

Whereas in *Drosophila* each polycomb component is encoded by a single gene, the number of genes encoding for PcG proteins has increased in mammals. However, the relevance of this diversification is unclear. Yet, expression patterns of PcG family members vary between distinct tissues and between various cell types within the same tissue[4–7]. For example, in the haematopoietic system, Bmi1 is specifically expressed in immature haematopoietic cells, whereas the expression of the Bmi1 paralogue Mel18 is increased during differentiation[6]. This argues for the existence of cell-type- and differentiation-stage-specific PRC1 and PRC2 subcomplexes with unique molecular functions[8–11].

Analysis of the composition of the PRC1 complex showed that it contains a single representative of the Cbx family; either Cbx2, Cbx4, Cbx7 or Cbx8 (refs 12,13). The amino-terminal chromodomain of Cbx family members can bind H3K9me3 and H3K27me3 (refs 14,15), albeit with unequal affinities[16,17]. Therefore, the Cbx proteins are key components for targeting PRC1 to specific genomic loci to regulate gene expression. Whereas several PcG proteins, such as Ezh2 and Bmi1, have shown to be important for HSC self-renewal[6,18–22], no data are available for how different Cbx subunits contribute to HSC regulation.

By overexpressing individual Cbx family members in HSCs, we have interrogated the biological function of different Cbx-containing PRC1 complexes. We show that the decision of HSCs to self-renew or differentiate critically depends on the molecular composition of the PRC1 complex.

[1]European Institute for the Biology of Ageing (ERIBA), Section Ageing Biology and Stem Cells, University Medical Centre Groningen, University of Groningen, Groningen 9700 AD, The Netherlands. [2]Biotech Research and Innovation Centre (BRIC), Centre for Epigenetics and DanStem, University of Copenhagen, Copenhagen 2200, Denmark.
[3]Correspondence should be addressed to G.d.H. (e-mail: g.de.haan@umcg.nl)

Received 24 April 2012; accepted 1 February 2013; published online 17 March 2013; DOI: 10.1038/ncb2701

RESULTS

Distinct expression of Cbx family members during haematopoietic differentiation

First, we assessed the expression of Cbx family members in purified haematopoietic subpopulations. Cbx7 is highly expressed in long-term HSCs (LT-HSCs), and its transcript levels gradually decreases on lineage commitment (Fig. 1a,b). However, Cbx7 is also highly expressed in differentiated lymphoid lineages (Fig. 1a and Supplementary Fig. S2). The Cbx8 transcript is equally expressed in various primitive haematopoietic cells. However, at the protein level, Cbx8 levels increase on lineage commitment, indicating a possible role for post-transcriptional regulation[10,23]. Cbx4 is abundant in most haematopoietic cell populations whereas Cbx2 is rather marginally but ubiquitously expressed (Fig. 1a,b and Supplementary Fig. S2). Notably, of all Cbx family members, the Cbx7 transcript is most abundantly expressed in LT-HSCs. These data suggest that the PRC1 complex in HSCs preferentially contains Cbx7.

Cbx family members affect *in vitro* self-renewal potential

As Cbx family members compete for integration into PRC1 (refs 12,13), overexpression of individual Cbx family members can change the composition of PRC1 complexes in favour of inclusion of the overexpressed Cbx protein. We overexpressed individual Cbx family members in post-5FU-treated bone marrow (hereafter referred to as hematopoietic stem and progenitors cells (HSPCs)). In cytokine-driven cultures, Cbx7 overexpression results in a strong proliferative advantage (Fig. 1c) and substantially increases the fraction of cells in the S–G2/M phase, without affecting apoptosis (Supplementary Fig. S3a,b). In contrast, Cbx4 or Cbx8 overexpression leads to decreased proliferation (Fig. 1c and Supplementary Fig. S3a,b), whereas Cbx2 overexpression does not lead to any detectable changes in proliferation (Fig. 1c). In liquid cultures, Cbx7-overexpressing HSPCs retain an immature blast-like morphology (Supplementary Fig. S3c), whereas most Cbx8 cells differentiate into myeloid cells within 10 days (Supplementary Fig. S3c).

To measure self-renewal, cells were plated in methylcellulose-based colony assays, followed by whole-dish or single-colony replating. Overexpression of Cbx7 substantially increases the colony-forming ability. This effect is even more pronounced after replating (Fig. 1d). Most single Cbx7 colonies have a high replating potential; 15% of the primary colonies can form more than 200 secondary colonies (Fig. 1e). In contrast, overexpression of Cbx2, Cbx4 or Cbx8 results in lower numbers of secondary colonies (Fig. 1d,e), indicating diminished self-renewal.

The effect of Cbx overexpression was further studied in the cobblestone-area-forming cell (CAFC; Fig. 1f) assay. Strikingly, Cbx7 overexpression leads to a strong increase (~100-fold) in day-35 CAFC activity, which reflects the activity of the most primitive haematopoietic cells. In contrast, overexpression of Cbx4 or Cbx8 results in a significantly reduced number of CAFCs.

In summary, overexpression of individual Cbx proteins causes distinct effects on self-renewal *in vitro*. Cbx7 is unique among the Cbx members in its ability to stimulate self-renewal *in vitro*.

Cbx family members affect *in vivo* HSC repopulating potential

Next, we investigated the repopulating ability of HSPCs overexpressing Cbx2, Cbx4, Cbx7 or Cbx8 in a competitive transplantation setting. Mice transplanted with Cbx2-overexpressing HSPCs showed only B-cell reconstitution (Supplementary Fig. S4a), suggesting that Cbx2 has a role in lymphopoiesis, but not myelopoiesis. However, HSPCs that overexpress Cbx2, Cbx4 or Cbx8 all fail to contribute to long-term haematopoietic reconstitution (Fig. 2a), suggesting a rapid exhaustion of progenitors and HSCs. In contrast, cells that overexpress Cbx7 have a prominent competitive advantage over non-transduced cells, as revealed by high chimaerism levels in peripheral blood (Fig. 2a).

Out of 22 mice, 20 reconstituted with Cbx7-overexpressing HSPCs developed lethal leukaemias with variable latencies (Fig. 2b–d). Twelve mice were diagnosed with T-cell leukaemia. These mice showed increased white blood cell (WBC) counts (Supplementary Fig. S4d and Table S1), enlarged lymph nodes and splenomegaly (Fig. 2c and Supplementary Fig. S4d). Thymus and liver were variably affected (Supplementary Table S1). Cells were $CD3e^+$ (Fig. 2d), $CD4^+$ and $CD127^+$ (Supplementary Fig. S4b and Table S1). Interestingly, Cbx7 protein level was found to be elevated in human lymphoid tumours[24]. Previously, overexpression of Cbx7 in collaboration with c-Myc, in the lymphoid compartment has been shown to cause T-cell lymphomas[24]. Here, we neither observed upregulation of c-Myc nor of other major oncogenes involved in lymphomagenesis (Supplementary Fig. S5).

The second type of leukaemia developed with a short latency (~4 weeks for 3 out of 4 mice) and was classified as an immature leukaemia. Malignant cells did not express any of the lineage markers used for analysis (Fig. 2d and Supplementary Fig. S4c and Table S1). Peripheral WBC counts were increased (Supplementary Fig. S4d and Table S1) and cells showed a blast-like morphology (Supplementary Fig. S4d). Finally, 2 out of 22 mice developed erythroid leukaemias. Spleens were mildly enlarged and one mouse showed haemorrhagic lesions in lymph nodes (Fig. 2c). The numbers of peripheral WBCs and erythrocytes were decreased (Supplementary Fig. S4d and Table S1), whereas the number of reticulocytes was increased (Supplementary Fig. S4d). In addition, bone marrow and spleen samples revealed numerous erythroid precursors at variable stages of maturation (Supplementary Fig. S4d). Malignant GFP^+ cells expressed Ter119 (Fig. 2d and Supplementary Table S1). For two mice it was not possible to determine the leukemia subtype (mice number 2 and 20 in Supplementary Table S1). These results show that Cbx7 has a strong oncogenic potential *in vivo*, whereas Cbx2, Cbx4 and Cbx8 all induce HSC exhaustion.

These results show that Cbx7 has a strong oncogenic potential *in vivo*, whereas Cbx2, Cbx4 and Cbx8 all induce HSC exhaustion.

Cbx7 induces self-renewal in multipotent cells, but not in restricted progenitors

As post-5FU-treated bone marrow contains a heterogeneous pool of primitive cells, we investigated which specific cell population is responsible for Cbx7-induced self-renewal. We overexpressed Cbx7 and Cbx8 in purified LT-HSCs, short-term HSCs (ST-HSCs), multipotent progenitors (MPPs) or myeloid progenitors, and compared their effect on the proliferative response in these purified populations at the clonal level. Single transduced cells were seeded in stromal-free cytokine-supplemented cultures, and colony-forming ability was scored 14 days later (Fig. 3a). Overexpression of Cbx7 enhanced the proliferative capacity of LT-HSCs, ST-HSCs and MPPs, but not of myeloid progenitors (Fig. 3a). Morphological analysis (Fig. 3b) showed that many of the Cbx7 colonies consisted of undifferentiated cells. In contrast, Cbx8 induced differentiation

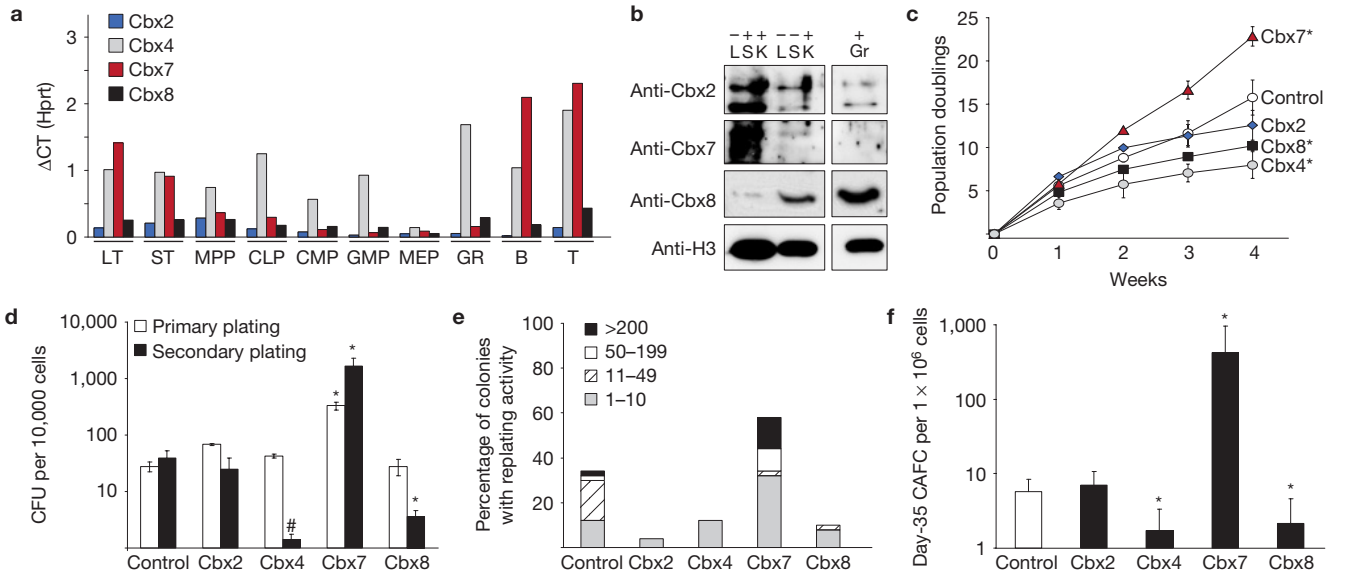

Figure 1 Enforced expression of Cbx genes reveals distinct effects on *in vitro* progenitor and HSC potential. (**a**) Gene expression of Cbx family members in different haematopoietic cell populations. Bars represent the mean of 2–3 technical replicates. See Supplementary Table S7 for raw data. GMP, granulocyte-macrophage progenitor; MEP, megakaryocyte-erythroid progenitor. (**b**) Protein abundance of Cbx2, Cbx7 and Cbx8 in Lin$^-$Sca1$^+$cKit$^+$ (LSK), Lin$^-$Sca$^-$cKit$^+$ (progenitor) and Gr1$^+$ cells. (**c**) Mean number of population doublings (±s.e.m.) for HSPCs overexpressing either control ($n=4$), Cbx2 ($n=3$), Cbx4 ($n=3$), Cbx7 ($n=4$) or Cbx8 ($n=4$). n represents independent experiments, $^*P \leq 0.05$. See Supplementary Table S7 for raw data. (**d**) Mean colony-forming ability (±s.e.m) of control and Cbx-overexpressing HSPCs (control $n=15$, Cbx2 $n=6$, Cbx4 $n=6$, Cbx7 $n=18$, Cbx8 $n=13$. n represents independent experiments. t-test $^*P<0.05$, $^\#P=0.06$). (**e**) Secondary colony formation on replating of individual primary colonies ($n=50$ primary replated colonies per sample). (**f**) Mean day-35 CAFC frequencies (±s.d.) of control and Cbx-overexpressing HSPCs (control $n=10$, Cbx2 $n=6$, Cbx4 $n=6$, Cbx7 $n=13$, Cbx8 $n=8$). n represents independent experiments. t-test $^*P<0.05$. Uncropped images of blots are shown in Supplementary Fig. S9a.

of LT-HSCs and ST-HSCs, because only a few colonies consisted of immature cells.

In the FBMD stromal co-cultures, day-35+ CAFC activity is restricted entirely to the most primitive LSK CD48$^-$ CD150$^+$ cells[25,26] (Fig. 4a). However, Cbx7 overexpression induced late cobblestone-forming activity also in ST-HSCs and MPPs, but not in myeloid progenitors (Fig. 4a). Cbx8 overexpression in ST-HSC and MPPs fails to induce such an effect.

We transplanted 750 Cbx7-overexpressing LT-HSCs into lethally irradiated recipients. Bone marrow analysis 12 months post-transplantation shows a significant increase in the frequency of GFP$^+$ ST-HSCs and MPPs, whereas the absolute number of lineage restricted progenitors (common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs)) and fully differentiated cells is reduced (Fig. 4b). In contrast, when Cbx8 is overexpressed, lower numbers of all GFP$^+$ haematopoietic cell subsets are observed, which indicates exhaustion of progenitors and HSCs *in vivo* (Fig. 4b). Importantly, when Cbx7 is overexpressed in myeloid progenitors or lymphoid progenitors, these cells do not contribute to haematopoietic reconstitution (Fig. 4c).

These results suggests that Cbx7 overexpression induces self-renewal of LT-HSCs, ST-HSCs and MPPs, and restricts their differentiation.

Loss of Cbx7 function reduces HSC self-renewal

To address whether the activity of HSCs and progenitors is dependent on Cbx7, we downregulated Cbx7 in HSPCs (Fig. 5a). Knockdown of Cbx7 results in a proliferative disadvantage in liquid culture (Fig. 5b), a reduction in their colony-forming ability (Fig. 5c) and a reduced frequency of CAFCs (data not shown). Cbx7 is therefore required for HSC and progenitor activity *in vitro* and other Cbx family members cannot compensate for its loss.

Next, we assessed whether the effect of Cbx7 overexpression on HSCs is dependent on its integration into PRC1. A mutant Cbx7 allele was generated in which the Pc-box was deleted[27]. This abrogates binding to Ring1b and Bmi1 (Supplementary Fig. S6b), and integration into PRC1 can therefore not occur. However, this mutant is still able to bind chromatin, although less efficiently (Supplementary Fig. S6c). In contrast to Cbx7^{WT}, overexpression of Cbx7$^{\Delta Pc}$ in HSPCs neither results in increased proliferative capacity in liquid culture (Supplementary Fig. S6a), nor to increased self-renewal as measured by the number of colony-forming units (CFUs; Fig. 6a). Furthermore, Cbx7$^{\Delta Pc}$ does not lead to an increased frequency of day-35 CAFC activity (Fig. 6b) and transplanted Cbx7$^{\Delta Pc}$ HSPCs fail to contribute to haematopoietic regeneration (Fig. 6c). We reason that Cbx7$^{\Delta Pc}$ acts in a dominant-negative manner by binding to target loci independently of PRC1, thereby blocking binding of endogenous PRC1 complexes.

Cbx proteins target PRC1 to chromatin by their chromodomain, which recognizes H3K9/27me3 (refs 14–17). To study whether Cbx7-induced self-renewal is dependent on recognizing this histone mark, a chromodomain-mutant (Cbx7^{AA}) was generated[27]. This amino-acid substitution abrogates binding to chromatin (Supplementary Fig. S6c), but it retains the ability to bind PRC1 components Bmi1 and Ring1b (Supplementary Fig. S6b). It therefore competes with endogenous Cbx proteins. Cbx7^{AA}, like Cbx7$^{\Delta Pc}$, acts in a dominant-negative manner because its overexpression results in a severe impairment of *in vitro* clonogenicity (Fig. 6a), self-renewal (Fig. 6b) and *in vivo* repopulating ability (Fig. 6c).

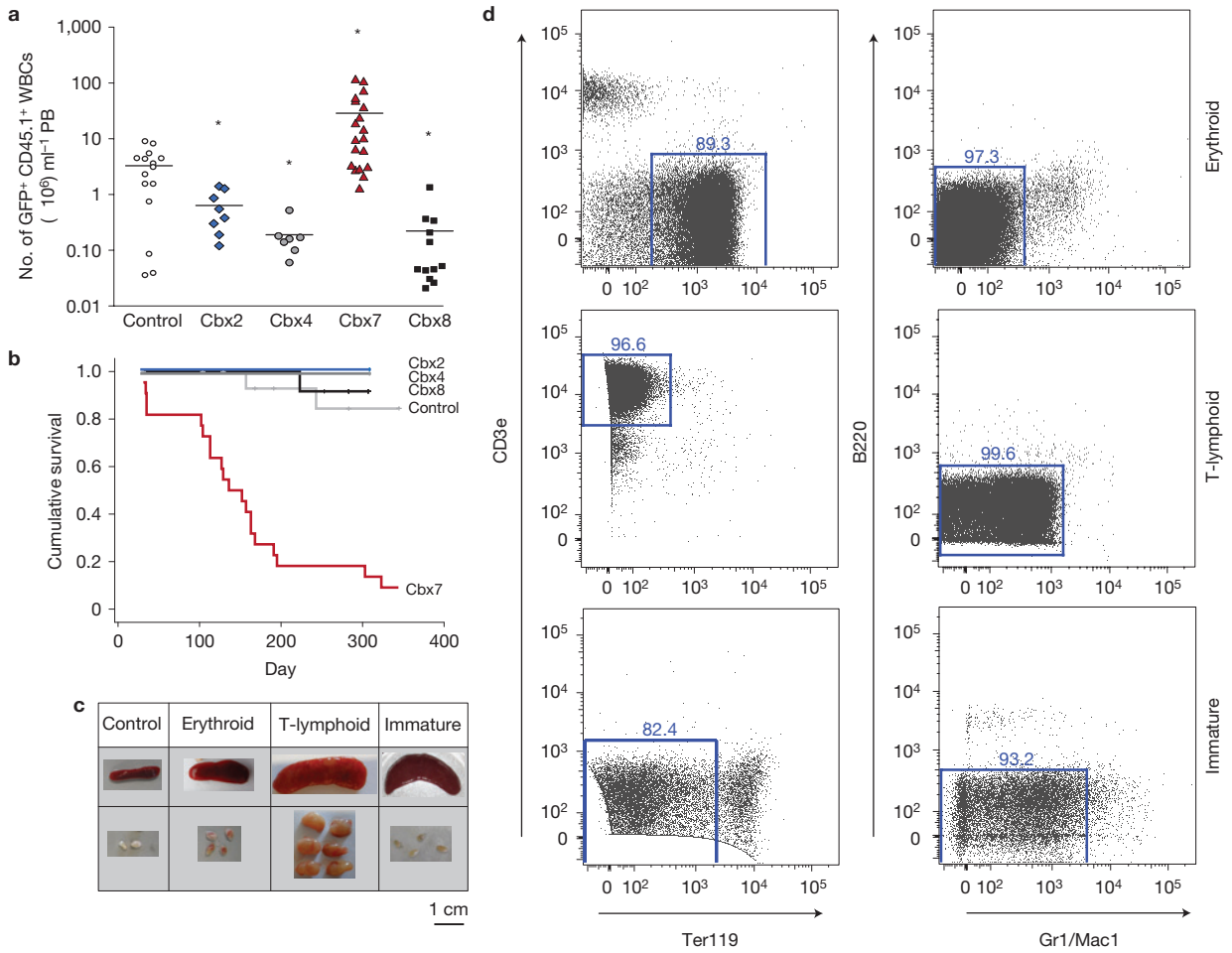

Figure 2 Cbx7 induces distinct types of leukaemia, whereas Cbx2, Cbx4 and Cbx8 induce a competitive disadvantage *in vivo*. Post-5FU-treated bone marrow cells from CD45.1 mice were isolated and transduced with control or Cbx-overexpressing retroviral vectors. Freshly transduced cells ($4.5–7.5 \times 10^6$) were then transplanted into lethally irradiated CD45.2 mice, without prior sorting for GFP. (**a**) Post-transplantation chimaerism levels of individual mice as measured by the absolute number of GFP$^+$ CD45.1$^+$ donor cells in peripheral blood (PB) at week 16 (control $n = 17$, Cbx2 $n = 8$, Cbx4 $n = 7$, Cbx8 $n = 12$), or on the day of sacrifice (Cbx7 $n = 22$). n represents individual mice. Line represents average chimaerism levels. t-test $^*P < 0.05$. (**b**) Kaplan–Meier survival curve for mice transplanted with control, Cbx2-, Cbx4-, Cbx7- or Cbx8-transduced bone marrow cells (control $n = 17$, Cbx2 $n = 8$, Cbx4 $n = 7$, Cbx7 $n = 22$, Cbx8 $n = 12$). n represents individual mice. (**c**) Spleen and lymph nodes from control, and Cbx7-induced leukaemic mice. (**d**) Representative FACS plots of GFP$^+$ spleen cells showing distinct dominant cell populations in Cbx7-induced leukaemias. Data from two independent transplantation experiments are combined. Gr1/Mac1: mixture of Gr1 and Mac1 antibodies with the same fluorochrome.

We verified whether perturbations of Cbx gene expression change the molecular composition of the PRC1 complex. Using Flag immunoprecipitation, we show that the core PRC1 component, Ring1b, interacts with both overexpressed Flag–Cbx7 and Flag–Cbx8 (Fig. 6d). Both Pcgf family members, Bmi1 and Mel18, bind equally to Cbx7 and Cbx8 (Fig. 4d), suggesting equal capabilities of overexpressed Cbx7 and Cbx8 to integrate into endogenous PRC1. As expected, binding between Cbx7 and Cbx8 was not detected (Fig. 6d). In addition, chromatin immunoprecipitation (ChIP) data show dissociation of endogenous Cbx8 from the *Ink4b–Ink4a–Arf* locus on overexpression of Cbx7, and vice versa (Supplementary Fig. S7a,b).

These data support the notion that Cbx7 is required for self-renewal and that the contrasting phenotypes after overexpression of Cbx proteins are caused by compositional rearrangements of PRC1.

Cbx7 and Cbx8 functional targets

As Cbx7 was the only Cbx family member inducing self-renewal, we reasoned that Cbx7 and Cbx8 bind to different genomic loci. Cbx8 was selected because its protein expression is reciprocal to Cbx7 and its overexpression showed a striking contrasting phenotype.

We evaluated whether Cbx7 and Cbx8 affect *Ink4b–Ink4a–Arf*, because studies have implicated that phenotypes of Bmi1 are, partially, dependent on this locus[28,29]. We show that both Cbx7 and Cbx8 bind throughout this locus (Fig. 7a,b and Supplementary Fig. S7a,b,d), which is enriched for H3K27me3 as evidenced by ChIP experiments (Supplementary Fig. S7c,d). No changes in expression of p15, p16 and p19 were observed (Supplementary Fig. S7e). Therefore, the contrasting HSC phenotypes induced by Cbx7 and Cbx8 are not explained by differential capabilities to repress this locus.

ChIP-sequencing (ChIP-seq) of Flag-precipitated DNA was used to identify genome-wide Cbx7- and Cbx8-binding sites in

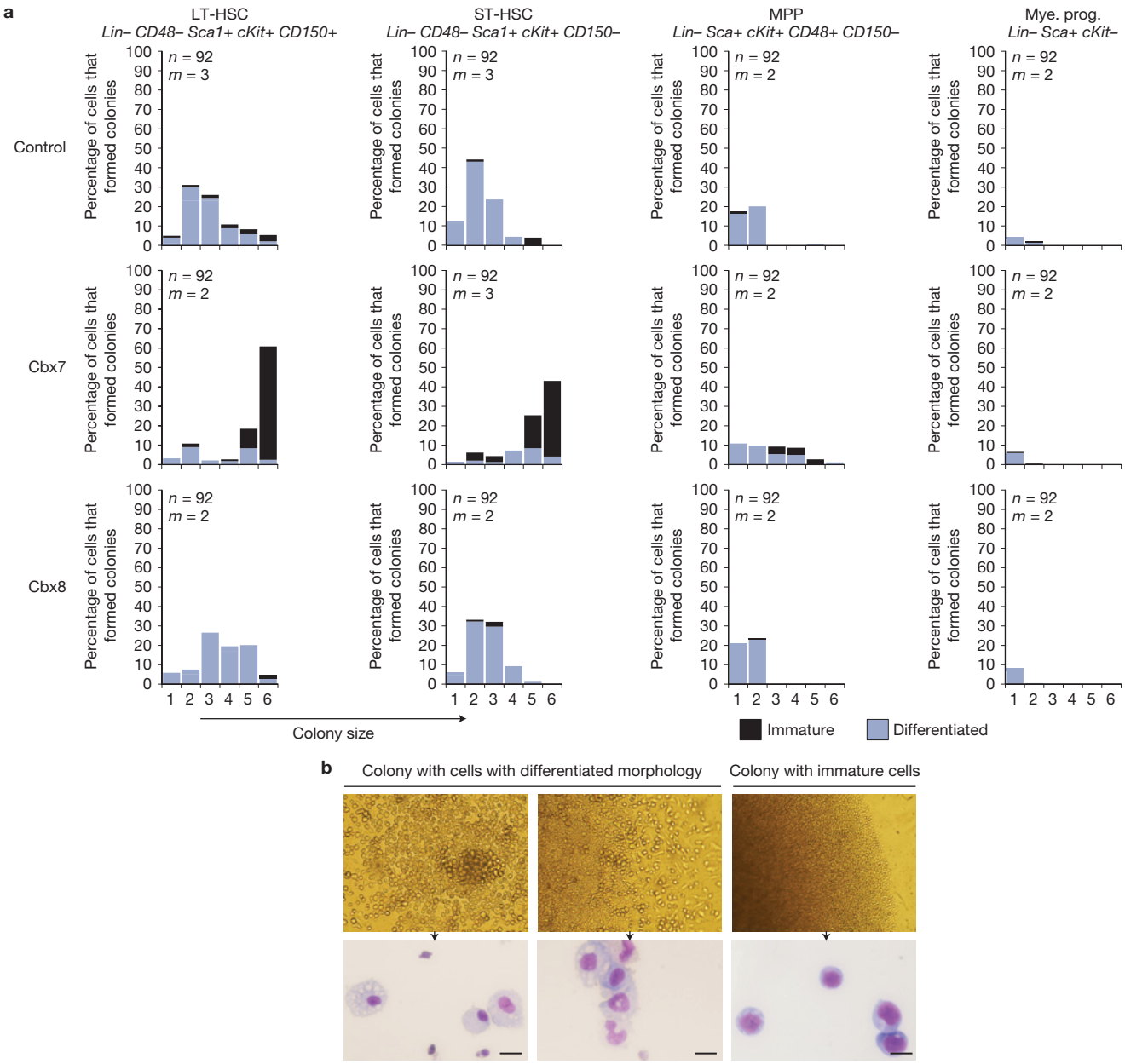

Figure 3 Cbx7 increases the proliferative capacity of purified multipotent haematopoietic subsets. (**a**) Clonal analysis of the proliferative capacity of indicated purified haematopoietic cells transduced with control, Cbx7 or Cbx8 retrovirus. Single transduced cells were seeded in cytokine-rich media and the colony size that these cells could produce was scored 14 days later. The figure represents the variability in colony-forming ability of single cells, scoring 92 cells (*n*) per independent experiment. Two or three independent experiments (*m*) were performed per sample. (**b**) The cell type of which the colonies consisted at day 14 was assessed using microscopical analysis of the colony itself and MGG-staining of cytospins from cells of these colonies. Scale bar, 10 μM.

primary HSPCs. Analysis of Cbx8 ChIP-seq relative to empty vector (control) revealed a total of 7,649 peaks, of which 7,091 peak positions could be annotated, mapping to 5,590 genes (Supplementary Information). Although overall sequencing depth was comparable (31.9×10^6 reads for Cbx7 versus 29×10^6 reads for Cbx8), substantially fewer peaks were identified for Cbx7; 3,621 peaks were called, of which 3,225 could be annotated, mapping to 2,768 genes (Supplementary Information). Cbx7- and Cbx8-enriched regions are predominantly found at the transcription start site (Fig. 7c,d).

Most identified targets were common for Cbx7 and Cbx8 (Fig. 7e). Cross-comparison of Cbx8 with Cbx7 (using the Cbx7 ChIP-seq signal as the background for Cbx8 peak calling) showed that of the 5,590 Cbx8 target genes, only 291 genes were not bound by Cbx7, and thus are unique for Cbx8 (Fig. 7e). Only 29 unique Cbx7 target genes were identified (Fig. 7e).

We assessed the quantitative differences of Cbx binding by counting the number of tags per peak on shared Cbx7 and Cbx8 loci (Fig. 7e and Supplementary Information). As we used the same antibody for ChIP-seq and the total read counts were comparable between

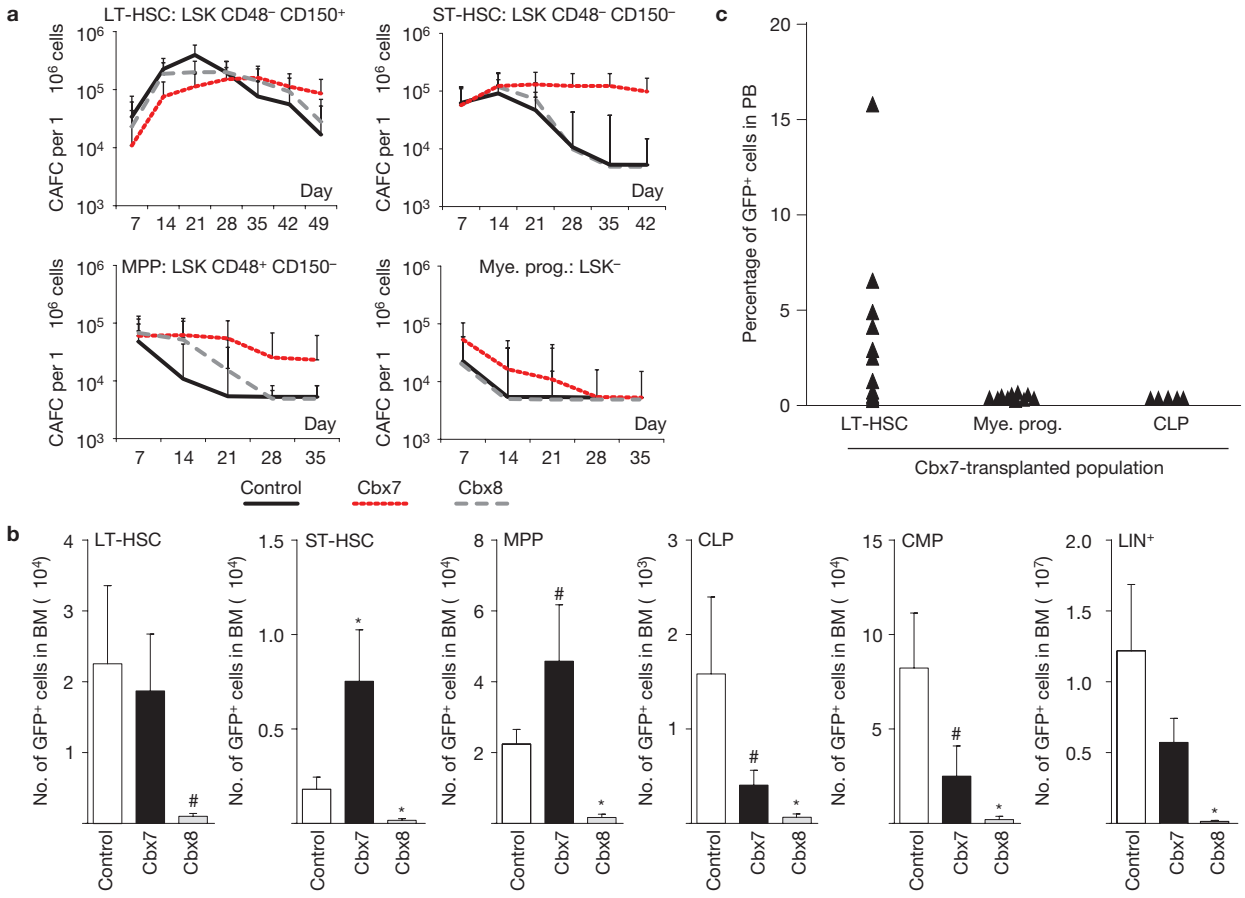

Figure 4 Cbx7 induces self-renewal of LT-HSCs, ST-HSCs and MPPs, but not of lineage restricted progenitors. (**a**) CAFC frequencies of purified LT-HSCs, ST-HSCs, MPPs or myeloid progenitors, transduced with the indicated vectors. Cells were plated in limiting dilution (1,3 and 10 cells per well), after which the average number of CAFCs (95% confidence interval) is calculated as described in ref. 40. (**b**) Phenotypical analysis of the bone marrow compartment of mice that were transplanted 12 months earlier with LT-HSCs that overexpress either control ($n = 5$), Cbx7 ($n = 5$) or Cbx8 ($n = 3$). The figure represents the average (±s.e.m.) number of absolute GFP$^+$ cells in the bone marrow (BM) of the indicated phenotype. n represents individual mice. t-test $^*P < 0.05$, $^\#P < 0.10$. See Supplementary Table S7 for raw data. (**c**) Chimaerism levels as indicated by the percentage of GFP$^+$ cells in the peripheral blood 24-weeks post transplantation. Mice were transplanted with either 750 Cbx7-overexpressing (GFP$^+$) LT-HSCs (LSK, CD48$^-$, CD150$^+$), myeloid progenitors (LSK$^-$) or CLPs (Lin$^-$ Sca^{mid} cKit^{mid} CD127$^+$). Symbols represent chimaerism levels of individual mice.

the Cbx7 and Cbx8, we reason that the number of tags per peak reflects the strength of occupancy on a particular genomic location, although we cannot exclude some extent of technical noise. For most overlapping loci, the number of tags per peak was consistently twice as high for Cbx8 as for Cbx7, suggesting that Cbx8 has a stronger overall affinity towards chromatin binding. This notion is further supported by the fact that Cbx8 has more targets than Cbx7 (5,590 versus 2,768, respectively). However, on some genomic positions (Fig. 7e), corresponding to 186 genes, the number of tags per peak was significantly higher for Cbx7.

Collectively, Cbx7 and Cbx8 share most of their targets. Only 291 unique Cbx8 targets and 29 unique Cbx7 targets were identified. Together with the 186 targets that showed quantitatively stronger binding by Cbx7, these genes are likely to account for the opposite phenotypes observed on HSCs (Fig. 7e and Supplementary Table S2 and Fig. S9a,b). Differential Cbx7 and Cbx8 targets contain oncogenes, differentiation-specific genes, cell-cycle regulators and transcription factors (Supplementary Table S3).

Cbx7 and Cbx8 targets are co-regulated during HSC differentiation

PcG-induced cell specification in the haematopoietic system is more likely orchestrated by changes in gene networks composed of signalling molecules, chromatin-modifying complexes and transcription factors, rather than by changes in the expression of a few individual genes. Therefore, we investigated whether any of the Cbx7 and Cbx8 targets exhibit correlated expression changes along HSC differentiation.

Expression data for the 215 Cbx7 and 291 Cbx8 targets were extracted from published transcriptome data in which four distinct haematopoietic cell stages (HSCs, progenitors, erythroid precursors and granulocytes) in two genetically distinct strains of mice were profiled (23 different micro-arrays, $n = 5/6$ for each cell population)[31]. From these data, gene networks were constructed[30] consisting of 33 Cbx7 and 85 Cbx8 targets that collectively showed correlated expression along haematopoietic differentiation (Fig. 7f and Supplementary Table S4).

Three highly connected subnetworks (modules) were identified. (Fig. 7f). In module 1, most genes that are coordinately repressed

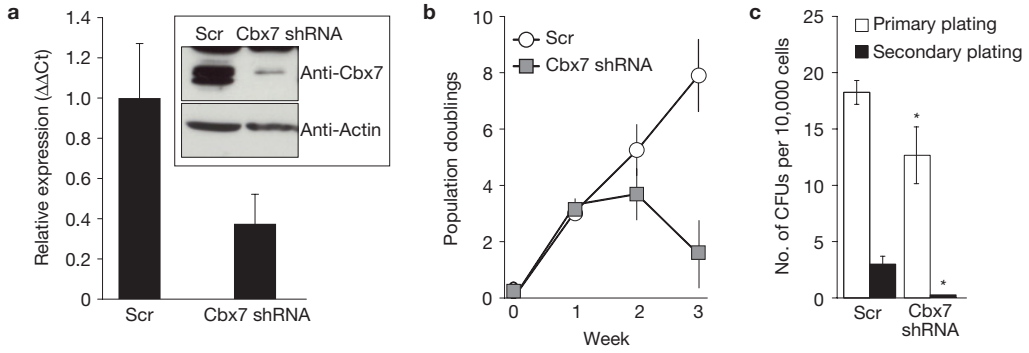

Figure 5 Cbx7 is required for HSPC self-renewal. (**a**) RT–qPCR analyses of Cbx7 messenger RNA expression in HSPCs transduced with Scrambled (Scr) and Cbx7 shRNA lentivirus. Expression was normalized to Hprt ($n = 3$ independent experiments, mean ± s.d.). Inset: western blot analysis of HSPCs transduced with control and Cbx7 shRNA lentivirus. Actin was used as a loading control. (**b**) Number of population doublings for HSPCs in which Cbx7 was downregulated (Cbx7 shRNA) or HSPCs transduced with Scr lentiviral vector ($n = 3$ independent experiments, mean ± s.d., t-test $P = 0.08$). (**c**) Colony-forming ability of control (Scr) and Cbx7-downregulating HSPCs ($n = 3$ independent experiments; mean ± s.d.; t-test $^*P < 0.05$). See Supplementary Table S7 for raw data.

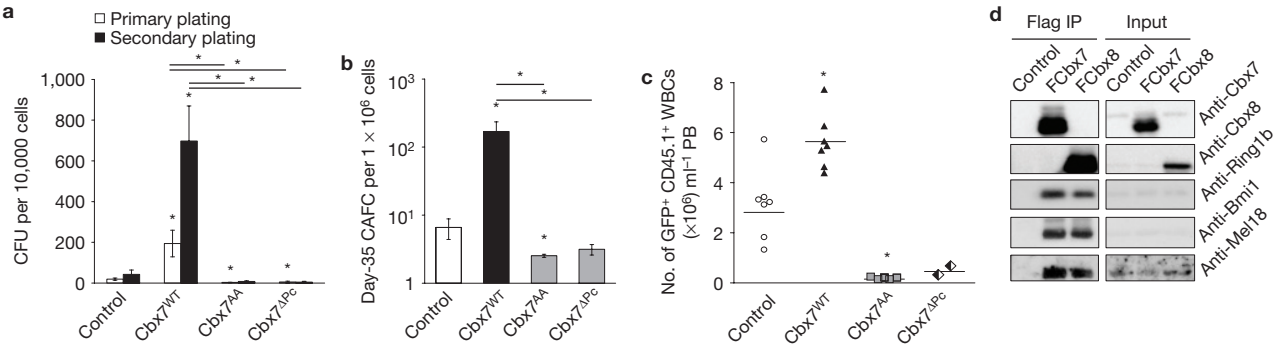

Figure 6 Cbx7-induced self-renewal requires PRC1 integration and H3K27me3 binding. (**a**) Primary and secondary colony-forming ability of (post-5-FU-treated) HSPCs transduced with control, Cbx7^{WT}, Cbx7 chromodomain-mutant (Cbx7^{AA}) and Cbx7 Pc-box deletion (Cbx7$^{\Delta Pc}$) vectors ($n = 4$ independent experiments, mean ± s.e.m., t-test $^*P < 0.05$). See Supplementary Table S7 for raw data. (**b**) Day-35 CAFC frequencies of control, Cbx7^{WT}, Cbx7^{AA} and Cbx7$^{\Delta Pc}$ bone marrow cells. $n = 4$ independent experiments, mean ± s.e.m., t-test $^*P < 0.05$. See Supplementary Table S7 for raw data. (**c**) Post-transplantation chimaerism levels as measured by the absolute number of GFP$^+$ CD45.1$^+$ donor cells in peripheral blood at week 12. Symbols represent chimaerism levels of individual mice (control $n = 7$, Cbx7^{WT} $n = 7$, Cbx7^{AA} $n = 5$; Cbx7$^{\Delta Pc}$ $n = 2$). Line represents average chimaerism levels. t-test $^*P < 0.01$. (**d**) Endogenous polycomb proteins bind ectopically expressed Flag–Cbx7 and Flag–Cbx8. Five per cent of input was used as a loading control. Uncropped images of blots are shown in Supplementary Fig. S9b.

in HSC and become upregulated in progenitors are Cbx7 targets (Fig. 7g and Supplementary Table S4). This is in agreement with the transcriptional repressive function of highly expressed Cbx7–PRC1 complexes in HSCs. Among them, *NPM1* is a regulator of the *ARF/p53* pathway, and loss-of-function mutations, deletions and translocations have been associated with acute myeloid leukaemia[32,33]. Within the same module, genes that show reciprocal expression changes during differentiation are mostly Cbx8 targets. These targets become repressed on the transition from HSCs into progenitors. *Bcl11a* and *Rlf* show many connections with other genes in this module, and are therefore putative major regulators. *Bcl11a* expression is found to be downregulated during differentiation, and its overexpression has been associated with leukaemogenesis[34,35].

Module 2 represents a subnetwork (Supplementary Fig. S8c) of genes specific for myeloid differentiation that almost exclusively consists of Cbx8 targets. These genes show repression on the transition from HSCs into progenitors and become highly upregulated during myeloid differentiation (Supplementary Fig. S8c and Table S4). This myeloid network is negatively regulated by a single Cbx8 target: *Gfi1b*, which is a well-known transcription factor that plays a critical role in the control of erythropoiesis and megakaryopoiesis[36–38]. Strikingly, two Cbx7 targets, *Sfpq* and *Ube3c*, also negatively regulate the Cbx8 myeloid specific gene network and show reciprocal gene expression patterns (Supplementary Fig. S8c).

Finally, module 3 contains genes differentially expressed in erythroid cells. Genes that become downregulated during differentiation from stem to progenitor to erythroid cells represent predominantly Cbx8 targets, whereas genes with a reciprocal expression pattern are enriched for Cbx7 targets (Supplementary Fig. S8d and Table S4).

Our approach shows that Cbx7 and Cbx8 targets are negatively co-regulated during HSC specification. Generally, genes bound only by Cbx8 are higher expressed in HSCs and become repressed during differentiation into progenitors, indicating that Cbx8 selectively targets HSC-specific genes. In turn, Cbx7 targets show the opposite expression pattern. Therefore, Cbx7 might induce HSC self-renewal by repressing progenitor-specific genes.

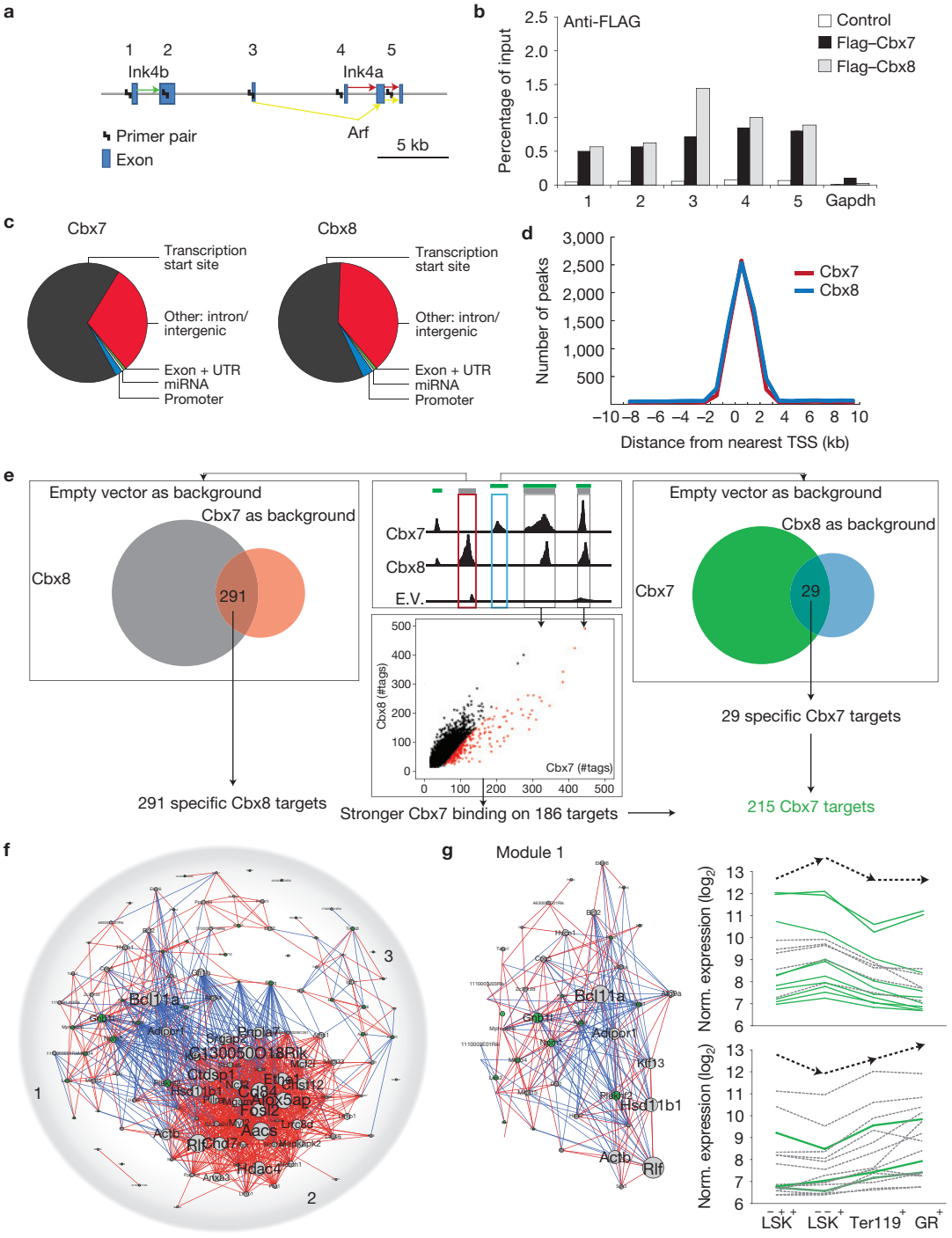

Figure 7 Cbx7 and Cbx8 targets are coordinately regulated during HSC differentiation and show reciprocal expression patterns. (**a**) Schematic representation of the mouse *Ink4b–Ink4a–Arf* locus and primer pairs used for ChIP analysis in **b**. kb, kilobases. (**b**) Binding of Cbx7 and Cbx8 to the *Ink4b–Ink4a–Arf* locus assessed by Flag ChIP. Bars represent mean of 2 independent experiments. See Supplementary Table S7 for raw data. (**c**) Diagram illustrating the overall distribution of Cbx7- and Cbx8-binding sites on TSS (±1 kb), promoter (−1 to −5 kb), microRNA, exon+UTR, intron and intergenic regions. (**d**) Localization of Cbx7 and Cbx8 ChIP-seq peaks relative to the nearest TSS. (**e**) Schematic representation illustrating data analysis of Flag ChIP-seq. E.V., empty vector. (**f**) Graphic representation of the gene network of Cbx7 and Cbx8 targets that show correlated gene expression changes during HSC differentiation. Expression of the 215 Cbx7 and 291 Cbx8 target genes was extracted from transcriptome data of four distinct haematopoietic cell populations (HSCs ($L^-S^+K^+$), progenitors ($L^-S^-K^+$), erythroid cells ($Ter119^+$) and granulocytes ($Gr1^+$)). Green nodes represent Cbx7 targets; grey nodes represent Cbx8 targets. Node size corresponds to the number of significant correlations of gene expression of a particular gene with other genes. Lines between nodes represent positive (red) or negative (blue) correlations of gene expression during haematopoietic differentiation. (**g**) Module 1 within the network represents one of the strongest connections within the total network. Right panels: genes (represented by nodes) are clustered on the basis of similarities in gene expression profiles and plotted in two graphs that show an inverse expression pattern along haematopoietic differentiation ($L^-S^+K^+$ HSCs, $L^-S^-K^+$ progenitors, erythroid cells ($Ter119^+$), granulocytes (Gr^+)). Green lines represent expression changes of Cbx7 targets, grey dotted lines of Cbx8 targets. Raw expression data can be found in Supplementary Table S4.

Table 1 GO analysis of genes targeted by Cbx7 in both ESCs and HSCs.

GO ID	Molecular function	Total genes in GO	Matched genes (of 600 total)	Corrected *P*-value Benjamini FDR	Gene families*
GO:0003700	Sequence-specific DNA-binding transcription factor activity	706	133 (18,84%)	4.10×10^{-78}	Dlx(3), Fox(12), Hox(14), Lhx(6), Nr(5), Pitx(3), Six(3), Sox(3), Tbx(3)
GO:0008270	Zinc ion binding	1,334	59 (4,42%)	2.24×10^{-4}	Lhx(6), Prdm(4), Sp(3), Zic(3)
GO:0003682	Chromatin binding	249	19 (7,63%)	5.06×10^{-4}	Dlx(2), Gata(2)
GO:0004714	Transmembrane receptor protein tyrosine kinase activity	49	9 (18,37%)	1.27×10^{-4}	Fgf(2), Eph(2)
GO:0005102	Receptor binding	236	17 (7,2%)	2.60×10^{-3}	Wnt(6)
GO:0003707	Steroid hormone receptor activity	47	7 (14,89%)	4.75×10^{-3}	Nr(5)
GO:0017147	Wnt-protein binding	26	5 (19,23%)	8.86×10^{-3}	Fzd(3)

*Numbers in parentheses represent the number of identified gene family members. GO, gene ontology. FDR, false discovery rate.

DISCUSSION

During evolution, the number of genes encoding polycomb proteins has increased, which has resulted in structural and functional diversification of PRCs. Here, we studied the role of different PRC1-associated Cbx family members in HSC regulation. Cbx family members show distinct haematopoietic cell-stage-specific expression patterns, which allows for the formation of variant PRC1 complexes during haematopoietic differentiation. As Cbx7 is highest expressed in the most immature haematopoietic populations, the PRC1 complex in HSCs preferentially contains Cbx7. PRC1 complexes containing Cbx7 induce self-renewal of HSCs by repressing the expression of progenitor-specific genes. Other Cbx proteins can compete with Cbx7, resulting in Cbx2-, Cbx4- or Cbx8-containing PRC1 complexes that target HSC-specific genes and thereby induce entrance into the differentiation pathway. We suggest that stoichiometric fluctuations between PRC1 complexes containing different Cbx family members play a key role in balancing self-renewal and differentiation of HSCs (Supplementary Fig. S1).

Diversification of polycomb genes might also account for tissue-specific stem cell functioning. Whereas we show that Cbx7 is the most important Cbx family member in HSCs, in epidermal stem cells Cbx4 plays this essential role[39]. It remains to be determined which of the Cbx family members are required for other types of adult stem cells. Recently, Cbx7 has been identified as an embryonic stem cell (ESC) pluripotency factor[8,10]. So far, genes that have been associated with ESC pluripotency, including *Oct4* and *Nanog*, have been very distinct from genes required for self-renewal in adult stem cells. To our knowledge, Cbx7 is the first gene to be identified as being important for self-renewal both in ESCs and adult stem cells, which indicates that the molecular mechanism by which it exerts its activity might be conserved during ontogenesis.

In the search for such a mechanism, we compared targets that we identified in HSCs with Cbx7 targets previously identified in ESCs (ref. 8). This ESC–HSC cross-comparison revealed 600 common targets that represent approximately 25% of all Cbx7 targets in ESCs (2,349; ref. 8) and HSCs (2,768; Supplementary Table S5). Interestingly, we identified 133 transcription factors, many of which are members of the same family (for example, 12 *Fox* genes, 14 *Hox* genes and 3 *Sox* genes; Table 1 and Supplementary Table S6). In addition, Cbx7 ECS–HSC targets include *Fgf* and *Eph* receptors and genes involved in Wnt signalling. We reason that Cbx7-regulated control of these genes is a core element of self-renewal.

Our data support the notion that the balance between self-renewal and differentiation, which constitutes a key hallmark of HSCs, is regulated by the molecular composition of the PRC1 complex. Therefore, genetic or pharmacologic perturbation of the composition of PRC1 may provide a strategy to induce or repress self-renewal in normal or aberrant haematopoiesis. □

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

We thank H. Moes, G. Mesander, H. de Bruin and R. J. van der Lei for expert flow cytometry assistance, the entire staff of the Central Animal Facility at the UMCG, K. Magnussen from the National High-throughput Sequencing Centre of the University of Copenhagen, B. Dethmers-Ausema, R. Bron, F. Feringa, K. van der Laan, V. Stojanovska and K. Wakker for laboratory assistance, J. Engelbert for bioinformatical assistance, and B. Dykstra, M. Niemantsverdriet, R. van Os and H. Schepers for valuable scientific discussions. We also acknowledge financial support from the Netherlands Organization for Scientific Research (VICI 918-76-601 and ALW to G.d.H), the Netherlands Institute for Regenerative Medicine, the Dutch Cancer Society (grant 2007-3729, and UMCG-2011-5277 to S.B.) and the European Community (EuroSystem, 200720). The work in the Helin laboratory was supported by the Danish National Research Foundation, the Novo Nordisk Foundation and the Danish Cancer Society.

AUTHOR CONTRIBUTIONS

K.K., L.B., K.H. and G.d.H. initiated research and developed the concept of the paper. K.K., L.B., and G.d.H. designed research; K.K. and V.R. performed experiments with contributions from M.B., E.W., S.O., M.R., S.B. and X.W; E.Z. performed bioinformatics analyses with contributions from L.B. and K.K.; K.K. and L.B. analysed and interpreted data; and K.K. wrote the manuscript with contributions from L.B., K.H. and G.d.H.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at www.nature.com/doifinder/10.1038/ncb2701
Reprints and permissions information is available online at www.nature.com/reprints

1. Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276,** 565–570 (1978).
2. Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10,** 697–708 (2009).

3. Morey, L. & Helin, K. Polycomb group protein-mediated repression of transcription. *Trends Biochem. Sci.* **35,** 323–332 (2010).
4. Lessard, J., Baban, S. & Sauvageau, G. Stage-specific expression of polycomb group genes in human BM cells. *Blood* **91,** 1216–1224 (1998).
5. Gunster, M. J. *et al.* Differential expression of human Polycomb group proteins in various tissues and cell types. *J. Cell. Biochem. Suppl.* **36,** 129–143 (2001).
6. Iwama, A. *et al.* Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1. *Immunity* **21,** 843–851 (2004).
7. Kato, Y., Koseki, H., Vidal, M., Nakauchi, H. & Iwama, A. Unique composition of polycomb repressive complex 1 in hematopoietic stem cells. *Int. J. Hematol.* **85,** 179–181 (2007).
8. Morey, L. *et al.* Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10,** 47–62 (2012).
9. Tavares, L. *et al.* RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148,** 664–678 (2012).
10. O'Loghlen, A. *et al.* MicroRNA regulation of Cbx7 mediates a switch of Polycomb orthologs during ESC differentiation. *Cell Stem Cell* **10,** 33–46 (2012).
11. Gao, Z. *et al.* PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45,** 344–356 (2012).
12. Maertens, G. N. *et al.* Several distinct polycomb complexes regulate and co-localize on the INK4a tumor suppressor locus. *PLoS One* **4,** e6380 (2009).
13. Vandamme, J., Volkel, P., Rosnoblet, C., Le Faou, P. & Angrand, P. O. Interaction proteomics analysis of polycomb proteins defines distinct PRC1 complexes in mammalian cells. *Mol. Cell Proteomics* **10,** M110.002642 (2011).
14. Min, J., Zhang, Y. & Xu, R. M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* **17,** 1823–1828 (2003).
15. Fischle, W. *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **17,** 1870–1881 (2003).
16. Bernstein, E. *et al.* Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell Biol.* **26,** 2560–2569 (2006).
17. Kaustov, L. *et al.* Recognition and specificity determinants of the human cbx chromodomains. *J. Biol. Chem.* **286,** 521–529 (2011).
18. Van Lohuizen, M. *et al.* Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. *Cell* **65,** 737–752 (1991).
19. Kamminga, L. M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107,** 2170–2179 (2006).
20. Lessard, J. *et al.* Functional antagonism of the Polycomb-Group genes eed and Bmi1 in hemopoietic cell proliferation. *Genes Dev.* **13,** 2691–2703 (1999).
21. Park, I. K. *et al.* Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423,** 302–305 (2003).
22. Rizo, A., Dontje, B., Vellenga, E., de Haan, G. & Schuringa, J. J. Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. *Blood* **111,** 2621–2630 (2008).
23. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proc. Natl Acad. Sci. USA* **105,** 10762–10767 (2008).
24. Scott, C. L. *et al.* Role of the chromobox protein CBX7 in lymphomagenesis. *Proc. Natl Acad. Sci. USA* **104,** 5389–5394 (2007).
25. Ploemacher, R. E., van der Sluijs, J. P., van Beurden, C. A., Baert, M. R. & Chan, P. L. Use of limiting-dilution type long-term marrow cultures in frequency analysis of marrow-repopulating and spleen colony-forming hematopoietic stem cells in the mouse. *Blood* **78,** 2527–2533 (1991).
26. De Haan, G., Nijhof, W. & Van Zant, G. Mouse strain-dependent changes in frequency and proliferation of hematopoietic stem cells during aging: correlation between lifespan and cycling activity. *Blood* **89,** 1543–1550 (1997).
27. Gil, J., Bernard, D., Martinez, D. & Beach, D. Polycomb CBX7 has a unifying role in cellular lifespan. *Nat. Cell Biol.* **6,** 67–72 (2004).
28. Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A. & van Lohuizen, M. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397,** 164–168 (1999).
29. Oguro, H. *et al.* Differential impact of Ink4a and Arf on hematopoietic stem cells and their BM microenvironment in Bmi1-deficient mice. *J. Exp. Med.* **203,** 2247–2253 (2006).
30. Voy, B. H. *et al.* Extracting gene networks for low-dose radiation using graph theoretical algorithms. *PLoS Comput. Biol.* **2,** e89 (2006).
31. Walasek, M. A. *et al.* The combination of valproic acid and lithium delays hematopoietic stem/progenitor cell differentiation. *Blood* **119,** 3050–3059 (2012).
32. Colombo, E., Alcalay, M. & Pelicci, P. G. Nucleophosmin and its complex network: a possible therapeutic target in hematological diseases. *Oncogene* **30,** 2595–2609 (2011).
33. Falini, B. *et al.* Acute myeloid leukemia with mutated nucleophosmin (NPM1): any hope for a targeted therapy? *Blood Rev.* **25,** 247–254 (2011).
34. Yin, B. *et al.* A retroviral mutagenesis screen reveals strong cooperation between Bcl11a overexpression and loss of the Nf1 tumor suppressor gene. *Blood* **113,** 1075–1085 (2009).
35. Deambrogi, C. *et al.* Analysis of the REL, BCL11A, and MYCN proto-oncogenes belonging to the 2p amplicon in chronic lymphocytic leukemia. *Am. J. Hematol.* **85,** 541–544 (2010).
36. Osawa, M. *et al.* Erythroid expansion mediated by the Gfi-1B zinc finger protein: role in normal hematopoiesis. *Blood* **100,** 2769–2777 (2002).
37. Laurent, B. *et al.* High-mobility group protein HMGB2 regulates human erythroid differentiation through trans-activation of GFI1B transcription. *Blood* **115,** 687–695 (2010).
38. Van der Meer, L. T., Jansen, J. H. & van der Reijden, B. A. Gfi1 and Gfi1b: key regulators of hematopoiesis. *Leukemia* **24,** 1834–1843 (2010).
39. Luis, N. M. *et al.* Regulation of human epidermal stem cell proliferation and senescence requires polycomb- dependent and -independent functions of Cbx4. *Cell Stem Cell* **9,** 233–246 (2011).
40. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. *Int. AAAI Conf. on Weblogs and Social Media* (2009).

METHODS

Animals. Female 10–16-week-old C57BL/6 CD45.2 (Harlan) or CD45.1 (bred at the Central Animal Facility) mice were used. All experiments are approved by the Animal Care Committee.

Cell purification. LT-HSCs (Lin^- $Sca1^+$ $c\text{-}Kit^+$ $CD48^-$ $CD34^-$ $CD150^+$ $EPCR^+$), ST-HSCs (Lin^- $Sca1^+$ $c\text{-}Kit^+$ $CD48^-$ $CD34^+$ $CD150^-$) and multipotent progenitor cells (Lin^- $Sca1^+$ $c\text{-}Kit^+$ $CD150^-$ $CD34^+$) were sorted from pooled bone marrow cells as described previously[41].

CLPs (Lin^- $CD127^+$ $c\text{-}Kit^{mid}$ Sca^{mid}), CMPs (Lin^- $CD127^-$ $c\text{-}Kit^+$ $Sca1^-$ $CD16/32^{mid}$ $CD34^{mid}$), granulocyte-macrophage progenitors (Lin^- $CD127^-$ $c\text{-}Kit^+$ $Sca1^-$ $CD16/32^{high}$ $CD34^{high}$) and megakaryocyte-erythroid progenitors (Lin^-, $CD127^-$, $c\text{-}Kit^+$, $Sca1^-$, $CD16/32^{low}$, $CD34^{low}$) were isolated by staining with Alexa-Fluor-conjugated lineage antibodies (B220 clone RA3-6B2, Biolegend 135010, 1:500; Gr1 clone RB6-8C5, Biolegend 108422, 1:75; Mac1 clone M1/70, Biolegend 101222, 1:100; Ter119 clone TER119, Biolegend 116220, 1:100; CD3 clone 17A2, Biolegend 100216, 1:150), CD127–PE (clone A7R34, Biolegend 135010, 1:250), Sca1–Pacific Blue (clone D7, Biolegend 108120, 1:150), cKit–APC (clone 2B8, BD Biosciences 561074, 1:75), CD16/32–PeCy7 (clone 93, eBioscience 25-0161-82, 1:75) and CD34–FITC (clone RAM34, BD Biosciences, 1:100). Sorting gates were set according to ref. 42. Granulocytes ($Gr1^+$), T cells ($CD3\varepsilon^+$) and B cells ($B220^+$) were isolated using B220–Pacific blue (clone RA3-6B2, Biolegend 103227, 1:1000), CD3ε–APC (clone 145-2C11, eBiosciences 17-0031, 1:100) and Ly-6G(Gr1)–AlexaFluor700 (1:100).

Retroviral vectors. The retroviral SF91–IRES–GFP vector (gift from C. Baum, Hannover Medical School, Germany) was used as the backbone for cloning complementary DNAs upstream of IRES. Cbx2, Cbx4 and Cbx7 cDNA clones were ordered from OpenBiosystems (accession numbers: Cbx2 BC035199, Cbx4 BC117801, Cbx7 BC021398). Myc–Flag–Cbx8 was cloned from PINCO (ref. 43) into SF91.

N-terminal Flag-tagged Cbx7 was generated by PCR using a forward primer containing the Flag sequence: forward: 5′ GCAGCGGCCGCATGGACTACAAGGACGACGATGACAAGATGGAGCTGTCAGCCATAGGCGAGCAGGTG-3′ and reverse: 5′-GGTCGACAGGGGAAGCCGCTATTCACAG-3′, and cloned into TOPO-blunt. The Cbx7 chromodomain mutant ($Cbx7^{AA}$: K31W32 to A31A32) was cloned using PCR on TOPO–Flag–Cbx7 using forward: 5′-CAAAGTTGAATATCTGGTGGCGGCGAAGGATGGCCCCCCA-3′ and reverse: 5′-CTTGGGGGGCCACCTTTCGCCGCCACCAGATATTCAACTT-3′.

$Cbx7^{\Delta Pc}$ (ΔTVTFREAQAAEGF) was cloned by insertion of Xba1 restriction sites flanking the Pc-box using F:5′-GCTCTAGATTCCGAGACCGCAACGAGAAG-3′ and R:5′-GCTCTAGAGACGGAGTTGGCGGTGATG-3′. The Pc-box was then deleted by Xba1 digestion and self-ligation. Subcloning from TOPO into SF91 was performed by Not1–Sal1 digestion and ligation. All constructs were verified by Sanger sequencing.

Retroviral overexpression in primary bone marrow. Post-5FU-treated primary bone marrow cells from female C57BL/6 (CD45.2) mice were isolated from tibia, femora and pelvic bones and transduced with the indicated vectors as described[44]. $EGFP^+$ cells were sorted using a MoFlo flow cytometer (BeckmanCoulter).

Downregulation of Cbx7 in primary bone marrow. Lentiviral construct containing a shRNA against Cbx7 (5′-CCGGCCTCAAGTGAAGTTACCGTGACTCGAGTCACGGTAACTTCACTTGAGGTTTTTG-3′; ref. 8) was generated by cloning annealed oligonucleotides into the AgeI and EcoRI cloning sites of pLKO.1 (Addgene plasmid 10878; ref. 45). As a control, a pLKO.1 vector containing a scrambled RNAi (5′-CCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGG-3′; ref. 46) was used (Addgene plasmid 1864).

HEK293T cells (T75) at 20% confluency were transfected with 3 μg of pLKO.1 short hairpin RNA (shRNA; Cbx7 or scrambled), 0.7 μg of pCMV–VSV-G and 0.3 μg of pCMV-8.91 plasmids in DMEM . After 18 h, the medium was changed to StemSpan (StemCell Technologies). Virus supernatant was added to HSPCs 18 h later. Cells were selected with 3 μg ml^{-1} puromycin 48 h after infection. Seventy-two hours later cells were plated in assays. Liquid cultures were maintained under selective antibiotics.

***In vitro* assays.** Post-5FU-treated GFP^+ HSPCs cells were collected in StemSpan supplemented with 10% FBS, 300 ng ml^{-1} rSCF (Amgen), 20 ng ml^{-1} rmIL11 (R&D Systems), 1 ng ml^{-1} Flt3 ligand (Amgen), and penicillin and streptomycin. Cytospots were generated by spinning at 70*g* in 200 μl FCS (Shandon Cytospin3) 10-days after culturing. Ten-day cultured cells were lysed, fixed and permeabilized according to the manufacturer's protocol (BD Phosflow, BD Biosciences) and stained with 1.2 μg ml^{-1} DAPI. Apoptosis was assessed by staining 10-day cultured cells with 10 μg ml^{-1} 7AAD (Sigma) and 50 μl ml^{-1} Annexin-V (IQ-Products) in calcium buffer. Cells were analysed by flow cytometry (LSRII, BD Biosciences).

Post-5FU-transduced 7-day cultured GFP^+ cells were plated in methylcellulose supplemented with 100 ng ml^{-1} mSCF and 20 ng ml^{-1} rmGM-CSF (R&D systems). After a 6-day incubation period, CFU-GM colonies were scored. Cells were collected by scraping and washed 3× in PBS. All cells were collected in fresh methylcellulose medium, vortexed and replated. For single-colony replating, 50 single colonies were picked, made single-cell by vortexing in new methylcellulose and plated in a 35 mm dish. Colonies were scored 7 days later.

CAFC assays were performed as previously described[47]. Directly after transduction, GFP^+ cells were plated in limiting dilution on a FBMD stromal cell layer and CAFCs were scored weekly.

Bone marrow transplantation. Freshly transduced HSPCs ($4.5–7.5 \times 10^6$; CD45.1) were transplanted into lethally irradiated (9.0 Gy) CD45.2 mice, without prior GFP sorting. Transduction percentages were 20–40%. The relative contribution of transduced cells (GFP^+) compared with identical treated non-transduced cells (GFP^-) was determined by blood withdrawal from the retro-orbital plexus. Red blood cells were subjected to lysis and WBCs were stained with CD45.1–PE (clone A20, Biolegend 110708, 1:200), B220–Pacific blue (clone RA3-6B2, Biolegend 103227, 1:1,000), CD3ε–APC (clone 145-2C11, eBiosciences 17-0031, 1:100), Ter119–Pe/Cy7 (clone TER-119, Biolegend 116222, 1:200), CD11b–AlexaFluor 700 (clone M1/70, Biolegend 101222, 1:100), and Ly-6G(Gr1)–AlexaFluor 700 (clone RB6-8C5, Biolegend 108422, 1:100).

At clear signs of morbidity, blood was taken from the orbital plexus. The mouse was euthanized and bone marrow cells were isolated by crushing. Single-cell suspensions were generated from different organs (spleen, lymph nodes, liver, thymus). Cell numbers were counted. Blood smears and cytospins were made and stained by MGG. Immunophenotypic analysis was performed as above. In a few cases, spleen cells were also stained with CD3–AF700 (clone 17A2, Biolegend 100216, 1:150), CD4–APC (clone RM4-5, Biolegend 100516, 1:200), and CD127–PE (clone A7R34, Biolegend 135010, 1:250) or CD41–PE (clone MWReg30, Biolegend 133906, 1:100) and CD16/32–PeCy7 (clone 93, eBioscience 25-0161-82, 1:75).

Transduction and transplantation of purified haematopoietic cells. LT-HSCs, ST-HSCs, MPPs and myeloid progenitors were FACS-purified as explained above, and pre-stimulated for 24 h in StemSpan with supplements. CLPs were pre-stimulated in StemSpan supplemented with 10% FBS, 100 ng ml^{-1} SCF, 50 ng ml^{-1} Flt3, 5 ng ml^{-1} IL7, and penicillin and streptomycin. At 24 h after pre-stimulation, cells were transferred to retronectin-coated wells and virus-supernatant produced in StemSpan was added. GFP^+ cells were sorted 24 h later.

Single GFP^+ cells were seeded into liquid cultures as described in ref. 41. After 14 days, wells were scored for the presence of a clone and the clone size was assessed. Colony size-1 represents colonies containing 1–30 cells, 2 31–100 cells, 3 101–1,000 cells, 4 ~5,000 cells, 5 ~15,000 cells and size 6 > 50,000 cells. Purified subpopulations were also plated in CAFC assays (1, 3 and 10 cells per well).

Transduced $CD45.1^+GFP^+$ cells (750) were transplanted in CD45.2 recipients plus 0.75×10^6 $CD45.2^+$ fresh bone marrow cells. After 12 months, bone marrow from mice that showed >1% chimaerism was analysed.

Immunoprecipitation. Immunoprecipitation using M2 Flag–agarose beads (Sigma) was carried out according to a standard protocol in IP150 buffer (50 mM Tris–HCl at pH 7.5, 150 mM NaCl, 5% glycerol, 0.2% Igepal and protein inhibitor cocktail (Roche)) in a myeloid cell line (32D).

ChIP assays of transduced HSPCs were in essence performed as described previously[48]. Flag–ChIP using M2 Flag–agarose beads was performed in IP150 buffer and complexes were washed 4× in 500 mM wash buffer (1% Triton X-100, 0.1% SDS, 500 mM NaCl, 2 mM EDTA at pH 8.0 and 20 mM Tris–HCl at pH 8.0) and the final wash step was performed in TE buffer. All other ChIPs were performed in IP buffer (66.7 mM Tris–HCl, 100 mM NaCl, 5 mM EDTA, 0.2% NaN3, 1.67% Triton-X-100 and 0.33% SDS) and complexes were washed in 500 mM wash buffer followed by LiCl buffer (20 mM Tris–HCl at pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP40 and 0.5% Na-deoxycholate) and a final wash step in TE. One microgram of rabbit-anti-CBX8 (LAST; ref. 49), rabbit-anti-Cbx7 (Millipore, 07-981) and rabbit-anti-H3K27me3 (Upstate, 07-449) or IgG from rabbit serum (Sigma, I8140) was used per immunoprecipitation. Prot A and Prot G fast-flow Sepharose (GE Healthcare) were used for pulldown. Primer sequences are available on request.

Western blotting. SDS–PAGE was performed according to standard protocols. Membranes were incubated with rabbit-anti-Cbx2 (1:1,000, Abcam, ab80044), rabbit-anti-CBX8 (1:1,000, LAST), rabbit-anti-Cbx7 (1:100, Santa Cruz, P-15), rabbit-anti-Ring1b (1:1,000, NAST; ref. 50), goat-anti-Mel18 (1:250, Abcam, ab5267), mouse-anti-Bmi1 (1:1,000, AF27; ref. 51), mouse-anti-H3 (Cell Signaling 3638, 1:300) and rabbit-anti-actin (Cell Signaling 4970, 1:5,000).

DOI: 10.1038/ncb2701

ChIP-sequencing. DNA samples obtained from the ChIP assays were adaptor-ligated and amplified (NEB, E6040) and analysed by Illumina Genome Analyser II. Duplicate reads were removed to avoid PCR bias. Using the Bowtie alignment algorithm for the Illumina sequence format (version 1.1.2) on the Galaxy server[52–54], the sequence data sets were mapped to the mouse genome (NCBI37/mm9). Peak detection was performed using MACS1.0.1 (refs 55,56; genome size 1.87×10^9, tag size 36 base pairs, cutoff peak detection 1×10^{-5}). Chromosomal positions were annotated to the RefSeq/ENSEMBL database. To define the distribution of peaks across the genome, we used the Bioconductor (version 2.9) package ChIPpeakAnno[57] and the built in data set TSS.mouse.NCBI37. For other features a custom annotation set was created by the BioConductor/R package BiomaRt (ref. 58).

Enriched genomic regions for Cbx7 and Cbx8 were determined using the empty vector (control) ChIP signal as a negative control. In addition, a cross-comparison was performed where the Cbx7 signal served as a control for Cbx8 peak calling, and vice versa. This comparative analysis determines regions where the ChIP signals are enriched relative to control ChIP (empty vector), or to the alternative Cbx protein, by a conditional binomial model (75-base-pair window, 0.1 FDR cutoff). All statistics and plotting were performed in R.

Gene network analysis. Gene expression network analysis of Cbx7 and Cbx8 targets was performed as described previously[31] using gene expression data of 4 primary cell subsets (LSK, $L^-S^-K^+$, $Gr1^+$ and $Ter119^+$). Probes were defined as differentially expressed during differentiation if $(V_{max} - V_{min})/\text{s.d.} \geq 4$ ($P < 0.05$), assuming normal distribution (3σ). Only transcripts that show significant differential expression were subjected to relevance network analysis as described previously[30]. Pearson correlation (r) between each pair of transcripts was measured and transcripts with $-0.8 < r > 0.8$ were marked as connected. Networks were visualized in Gephi software (version 0.8β; ref. 40) and spatially arranged by built-in Yifan-Hu and Fruchterman-Reingold algorithms. A built-in modularity function was used to partition the network into significant tight connected communities (modules). Significant differential clustering between Cbx7 and Cbx8 targets was tested using binomial distribution.

Statistical analysis. Statistical significant differences of growth curves were using the compareGrowthCurves function of the statmod software package (R). Differences between the distribution of colony sizes were tested by the χ^2 test (data transformed to $n = 50$). CAFCs by limiting dilution were calculated by maximum likelihood as described previously[59]. Other statistical analyses were performed by two-tailed t-tests assuming equal variances by SPSS or Excel.

Accession numbers. Primary accessions: ChIP-seq data files are available at the Gene Expression Omnibus (GEO) under GSE36658. Referenced accession: GSE35292.

41. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & de Haan, G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J. Exp. Med.* **208,** 2691–2703 (2011).
42. Challen, G. A., Boles, N., Lin, K. K. & Goodell, M. A. Mouse hematopoietic stem cell identification and analysis. *Cytometry A* **75,** 14–24 (2009).
43. Dietrich, N. *et al.* Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. *EMBO J.* **26,** 1637–1648 (2007).
44. Gerrits, A. *et al.* Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* **115,** 2610–2618 (2010).
45. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
46. Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307,** 1098–1101 (2005).
47. Van Os, R. P., Dethmers-Ausema, B. & de Haan, G. *In vitro* assays for cobblestone area-forming cells, LTC-IC, and CFU-C. *Methods Mol. Biol.* **430,** 143–157 (2008).
48. Frank, S. R., Schroeder, M., Fernandez, P., Taubert, S. & Amati, B. Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. *Genes Dev.* **15,** 2069–2082 (2001).
49. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20,** 1123–1136 (2006).
50. Dietrich, N. *et al.* REST-mediated recruitment of polycomb repressor complexes in mammalian cells. *PLoS Genet.* **8,** e1002494 (2012).
51. Bracken, A. P. *et al.* The Polycomb group proteins bind throughout the INK4a-ARF locus and are disassociated in senescent cells. *Genes Dev.* **21,** 525–530 (2007).
52. Giardine, B. *et al.* Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15,** 1451–1455 (2005).
53. Blankenberg, D. *et al.* Manipulation of FASTQ data with Galaxy. *Bioinformatics* **26,** 1783–1785 (2010).
54. Goecks, J., Nekrutenko, A. & Taylor, J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11,** R86 (2010).
55. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
56. Feng, J., Liu, T. & Zhang, Y. Using MACS to identify peaks from ChIP-Seq data. *Curr. Protoc. Bioinform.* (2011) Chapter 2, Unit 2 14.
57. Zhu, L. J. *et al.* ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinform.* **11,** 237 (2010).
58. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4,** 1184–1191 (2009).
59. de St Groth, Fazekas The evaluation of limiting dilution assays. *J. Immunol. Methods* **49,** R11–R23 (1982).

DOI: 10.1038/ncb2701

nature cell biology

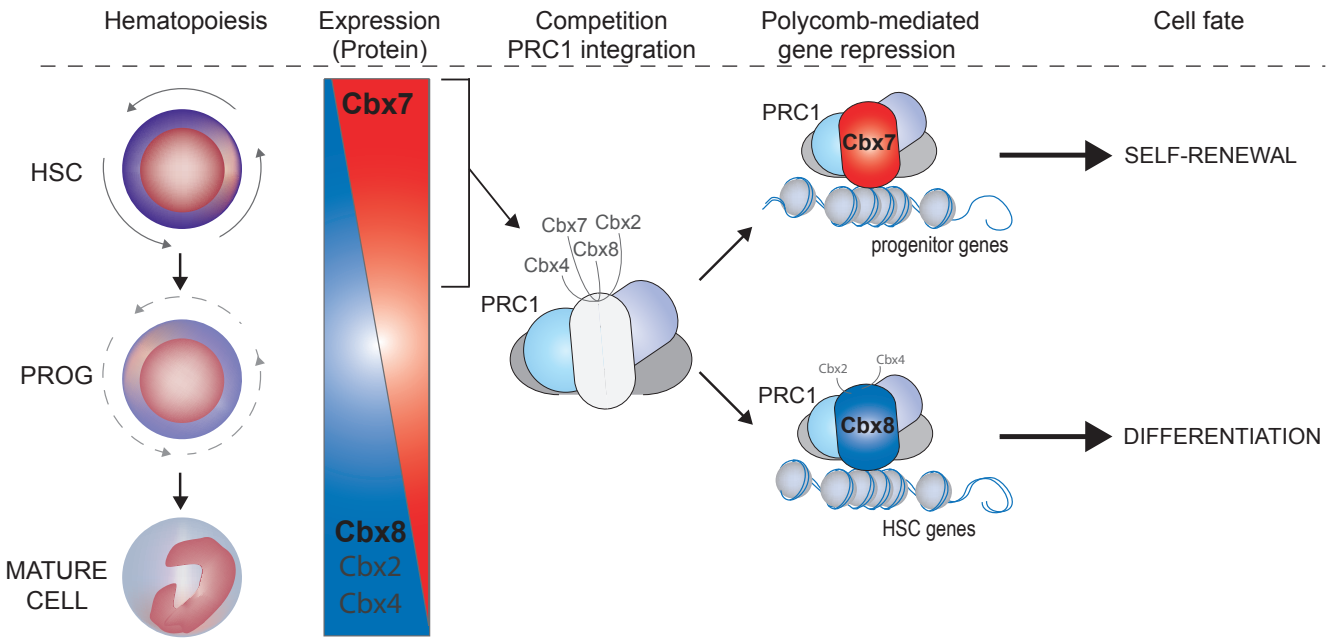

Figure S1 PRC1-Cbx control of HSC fate decisions. Model summarizing the findings in this paper. Polycomb Cbx orthologs compete for PRC1 integration and balance HSC self-renewal and differentiation. Cbx7-containing PRC1 complexes induce self-renewal of HSCs by repressing the expression of progenitor-specific genes. Other Cbx proteins can outcompete Cbx7 from PRC1, resulting in Cbx2-, Cbx4-, or Cbx8-containing PRC1 complexes that target HSC-specific genes and thereby induce entrance into the differentiation pathway.

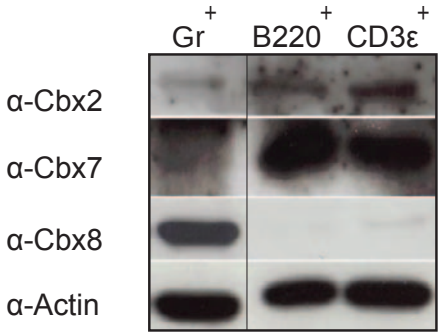

Figure S2 Protein abundance of Cbx orthologs in differentiated hematopoietic cells. Western blot analysis of protein abundance of Cbx2, Cbx7 and Cbx8 in granulocytes (Gr1+), B cells (B220+) and T- cells (CD3e+). Actin was used as loading control. Also see Figure S9.

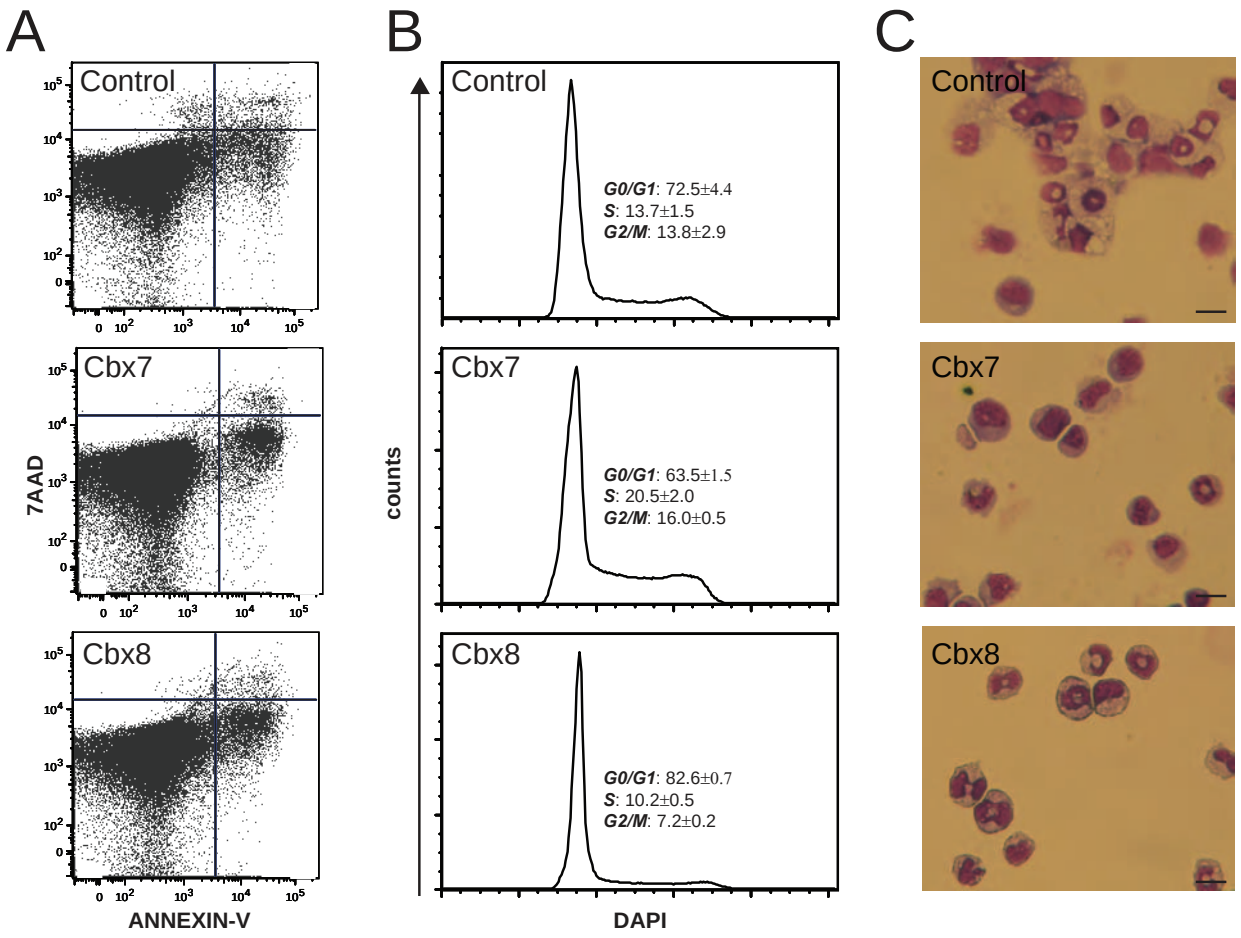

Figure S3 Effect of Cbx7 and Cbx8 overexpression on the cell-cycle. Post 5FU treated bone marrow cells were transduced with Control or Cbx-overexpressing retroviral vectors. GFP+ HSPCs were plated in cytokine-supplemented media and at day 10 cells were used for apoptosis assays, cell-cycle analysis and cytospins. **A)** Apoptotic frequency of control, Cbx7 and Cbx8 bone marrow cells using FACS analysis of annexin-V and 7AAD positive cells. One representative FACS plots is shown from 3 independent experiments. **B)** Cell cycle analysis by measurement of DNA content in control, Cbx7 and Cbx8 HSPCs (n=2 independent experiments, mean±s.d.). **C)** May-Grünwald Giemsa staining of control, Cbx7 and Cbx8 cells. One representative image is shown. Scale bar 10μM. Also see Table S7 for raw data.

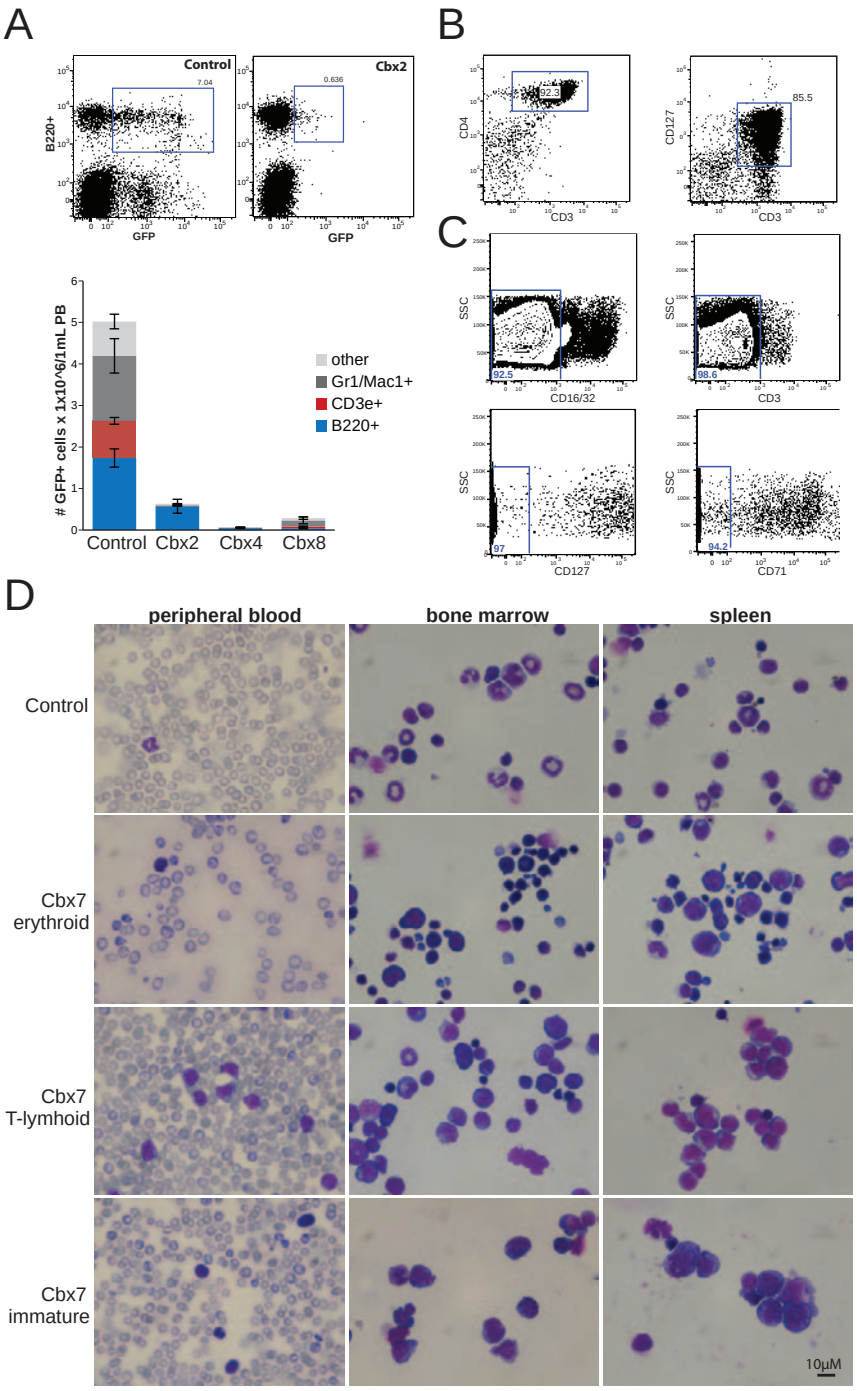

Figure S4 Cbx7-induced leukemia subtypes. **A)** Analysis of lineage contribution of transplanted HSPCs transduced with Control, Cbx2, Cbx4 and Cbx8 retrovirus, 16 weeks post-transplantation. *n*=7-9 individual mice, mean±s.e.m. **B)** FACS plot example of spleen cells from a mouse (mouse9 suppl table 1) with (CD3+, CD4+, CD127+) T-cell leukemia. Cells within CD45.1+, GFP+ fraction are shown. **C)** Staining of spleen cells of mouse 14 (suppl table 1) with an immature leukemia type with CD3, CD127, CD71 and CD16/32. Cells within CD45.1+, GFP+ fraction are shown. **D)** May-Grunwald-Giemsa staining of cells in peripheral blood, bone marrow, and spleen of control mice and Cbx7 mice with indicated leukemia subtypes. One representative image is shown for every condition. Scale bar 10μM.

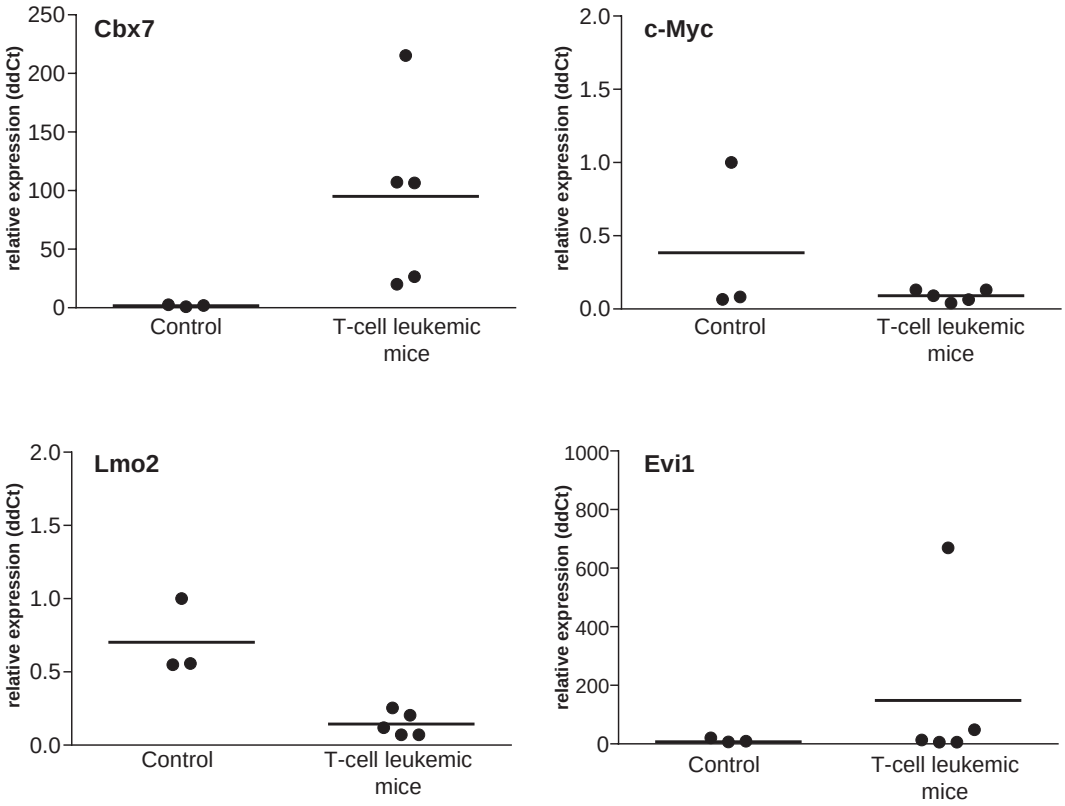

Figure S5 Oncogene expression in Cbx7-induced lymphoid leukemias. Relative gene expression of Cbx7, c-Myc, Evi-1, Lmo2 in spleen cells from individual Control (n=3) and T-cell leukemic mice (n=5). Gapdh was used for normalization.

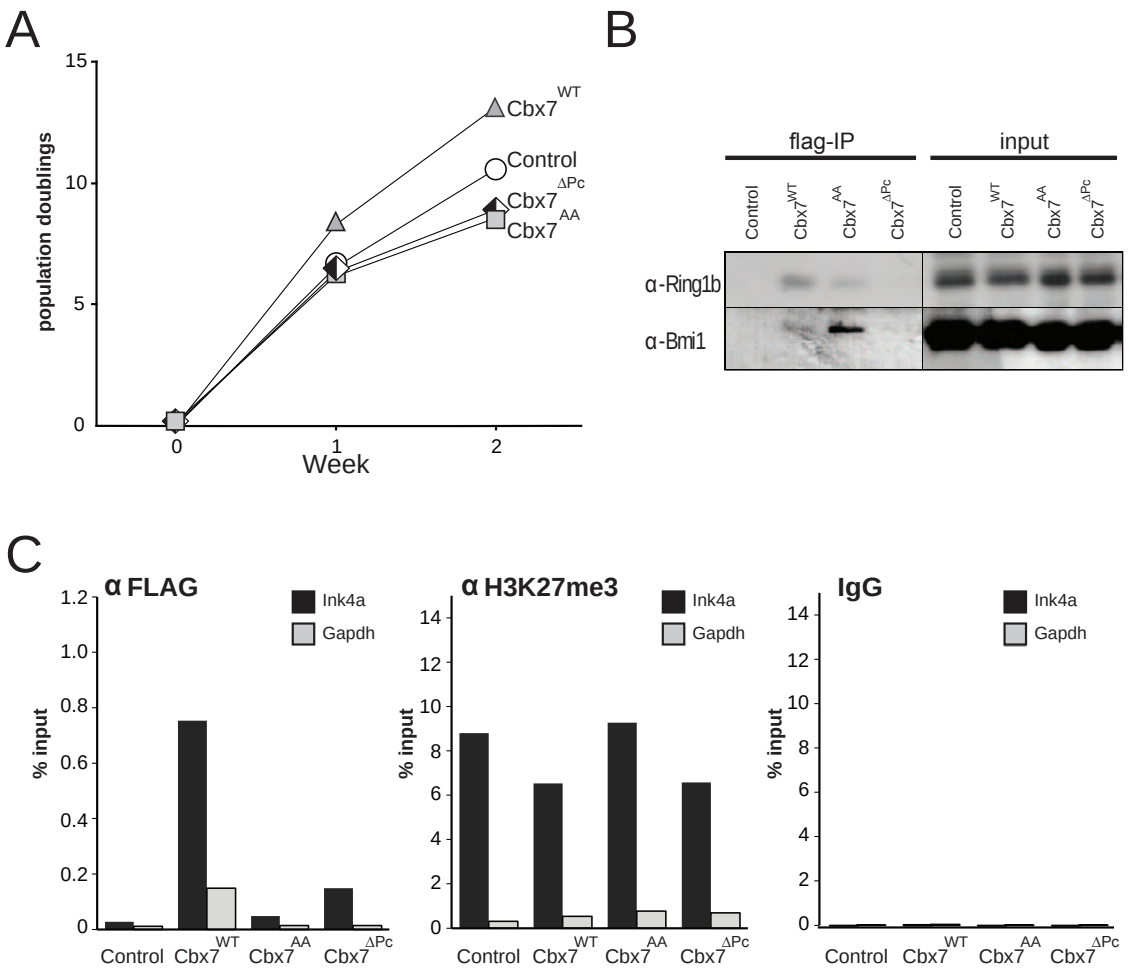

Figure S6 Cbx7-induced HSC self-renewal is Polycomb dependent. **A)** Population doublings of HSPCs overexpressing Cbx7^{WT}, flag-Cbx7^{AA}, and flag-Cbx7$^{\Delta Pc}$. Bars represent mean from 2 independent experiments. **B)** Flag-immunoprecipitation of Ring1b and Bmi1 with flag-Cbx7^{WT}, flag-Cbx7^{AA}, and flag-Cbx7$^{\Delta Pc}$ in 32D cells. One representative experiment is shown from two independent experiments. Also see Figure S9. **C)** Binding of Cbx7^{WT}, flag-Cbx7^{AA}, and flag-Cbx7$^{\Delta Pc}$ to the transcription start site of Ink4a. HSPCs were transduced with indicated vectors. Chromatin-immunoprecipitation was then performed using Flag-M2 agarose beads and H3K27me3 antibodies coupled to sepharose beads. IgG was used as negative control. Pulled down DNA was purified and binding to Ink4a was assessed using qPCR. Bars represent mean from 2 independent experiments. Also see Table S7 for raw data.

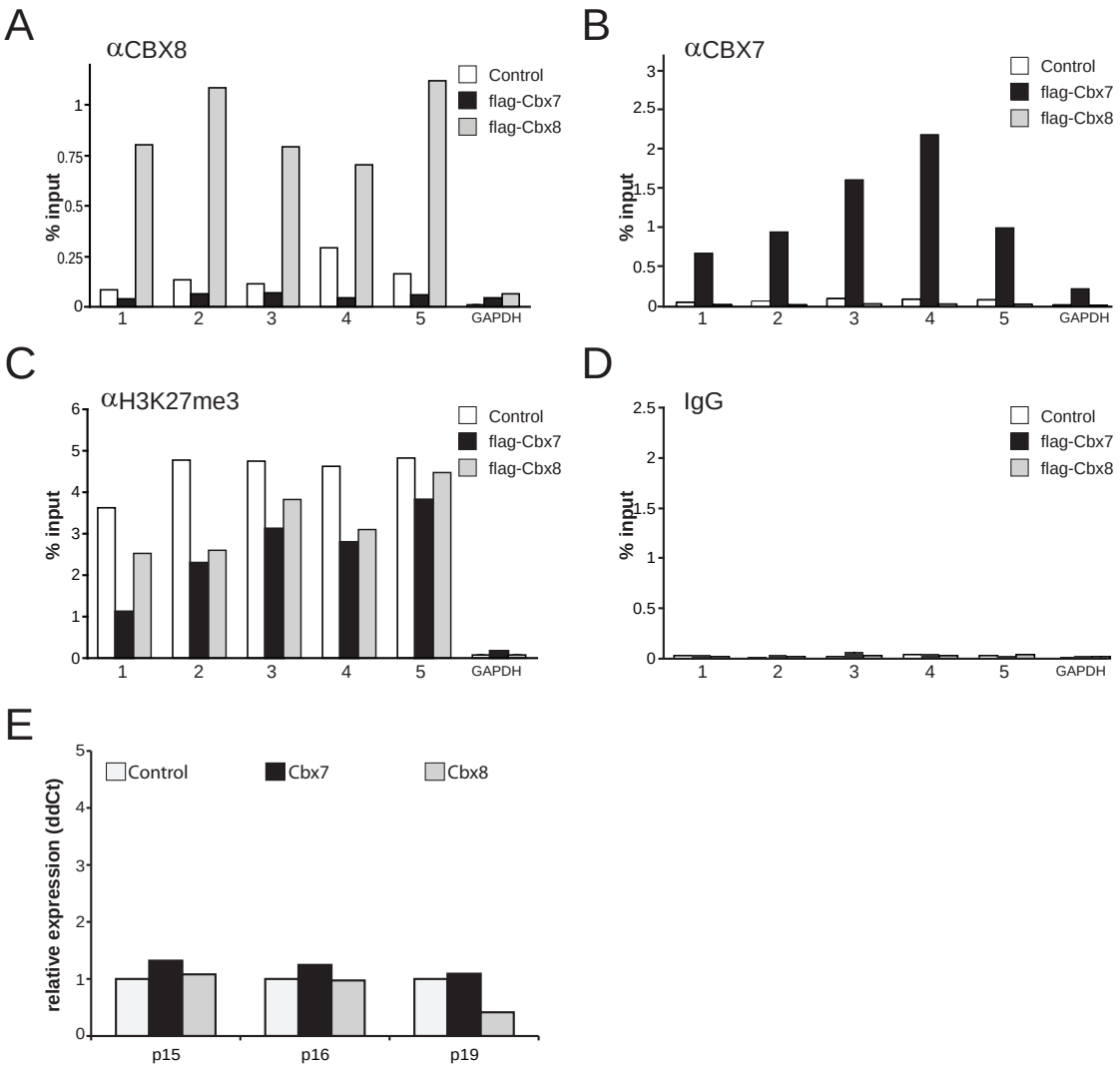

Figure S7 Cbx7 and Cbx8 bind the senescence locus. **A-D)** Post-5FU treated bone marrow cells transduced with Control or Cbx-overexpressing retroviral vectors were sorted based on GFP expression and cultured for 7 days in cytokine-supplemented media. Next, ChIP was performed in control, Cbx7 and Cbx8 bone marrow cells using indicated antibodies. Numbers correspond to the *Ink4b-Ink4a-Arf* primer pairs used for qPCR, as shown in Figure 4a. Bars represent mean from 2 independent experiments. **E)** Relative expression of p15, p16 and p19 after Cbx7 and Cbx8 overexpression in bone marrow compared to control, as determined by rt-PCR. Expression was calculated according to ddCT of Gapdh. Bars represent mean from 2 independent experiments. Also see Table S7 for raw data.

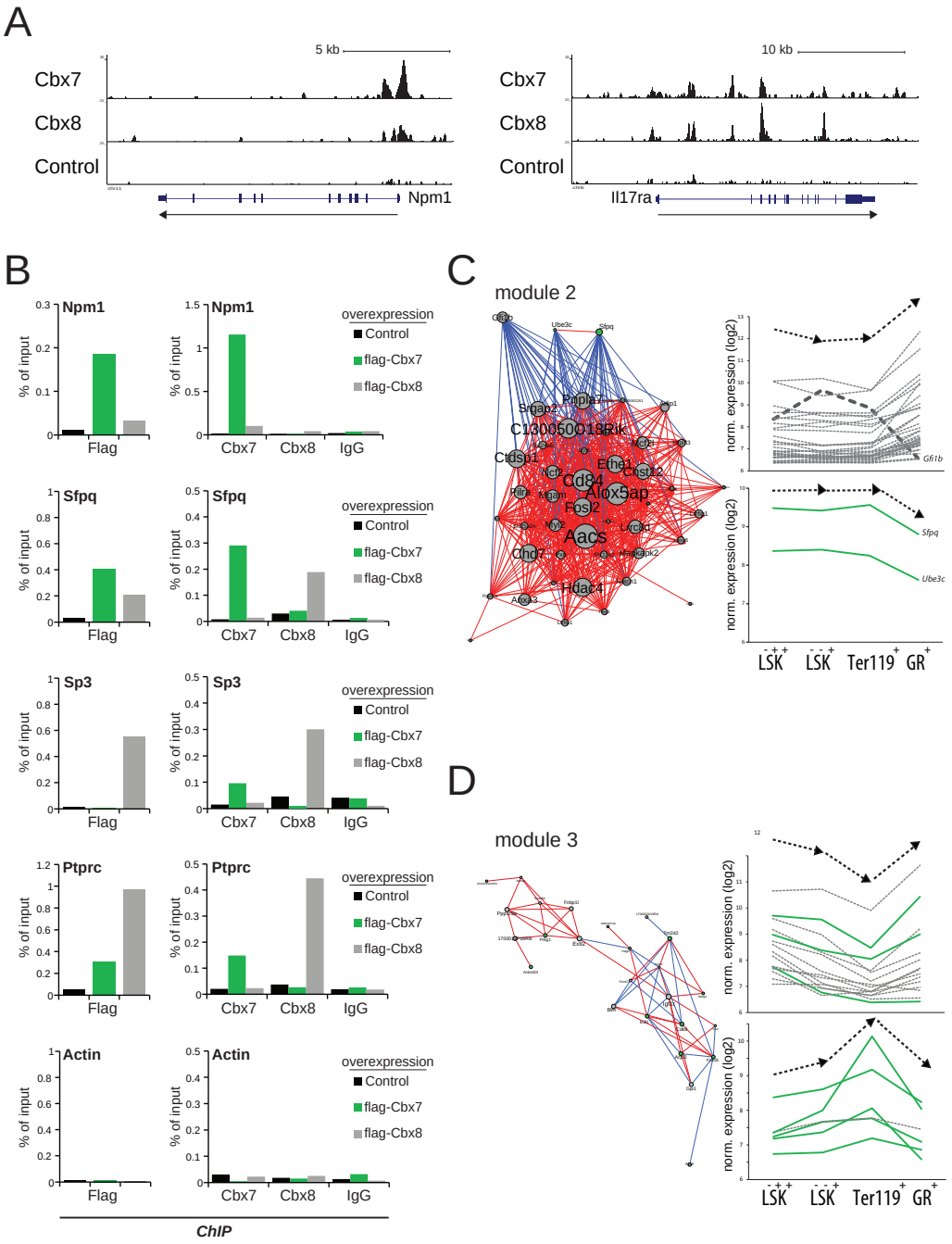

Figure S8 Cbx7 and Cbx8 targets and gene network. **A)** Representative examples of ChIP-seq results in Control, Cbx7-overexpressed and Cbx8-overexpressed bone marrow cells. The y-axis of the binding profiles indicates tag counts. **B)** Validation of ChIP-seq targets by flag-ChIP, Cbx7-ChIP and Cbx8-ChIP followed by qpcr. Flag-ChIP and Cbx-ChIPs represent independent experiments. Npm1 and Sfpq are Cbx7 targets (either not bound by Cbx8, or stronger bound by Cbx7 than by Cbx8), Sp3 and Ptprc are specific Cbx8 targets. Actin was used as a negative control locus. Also see Supplementary Table 7 for raw data. **C)** Module 2 represents a myeloid specific sub-network within the network. **D)** Module 3 represents an erythroid specific sub-network. Right panels of **C)** and **D)**: Genes (represented by nodes) are clustered on basis of similarities in gene expression profiles and plotted in two graphs that show an inverse expression pattern along hematopoietic differentiation (L-S+K+ HSCs, L-S-K+ progenitors, erythroid cells (Ter119+), granulocytes (Gr+)). Green lines represent expression changes of Cbx7 targets, grey dotted lines of Cbx8 targets. Raw expression data can be found in Table S4.

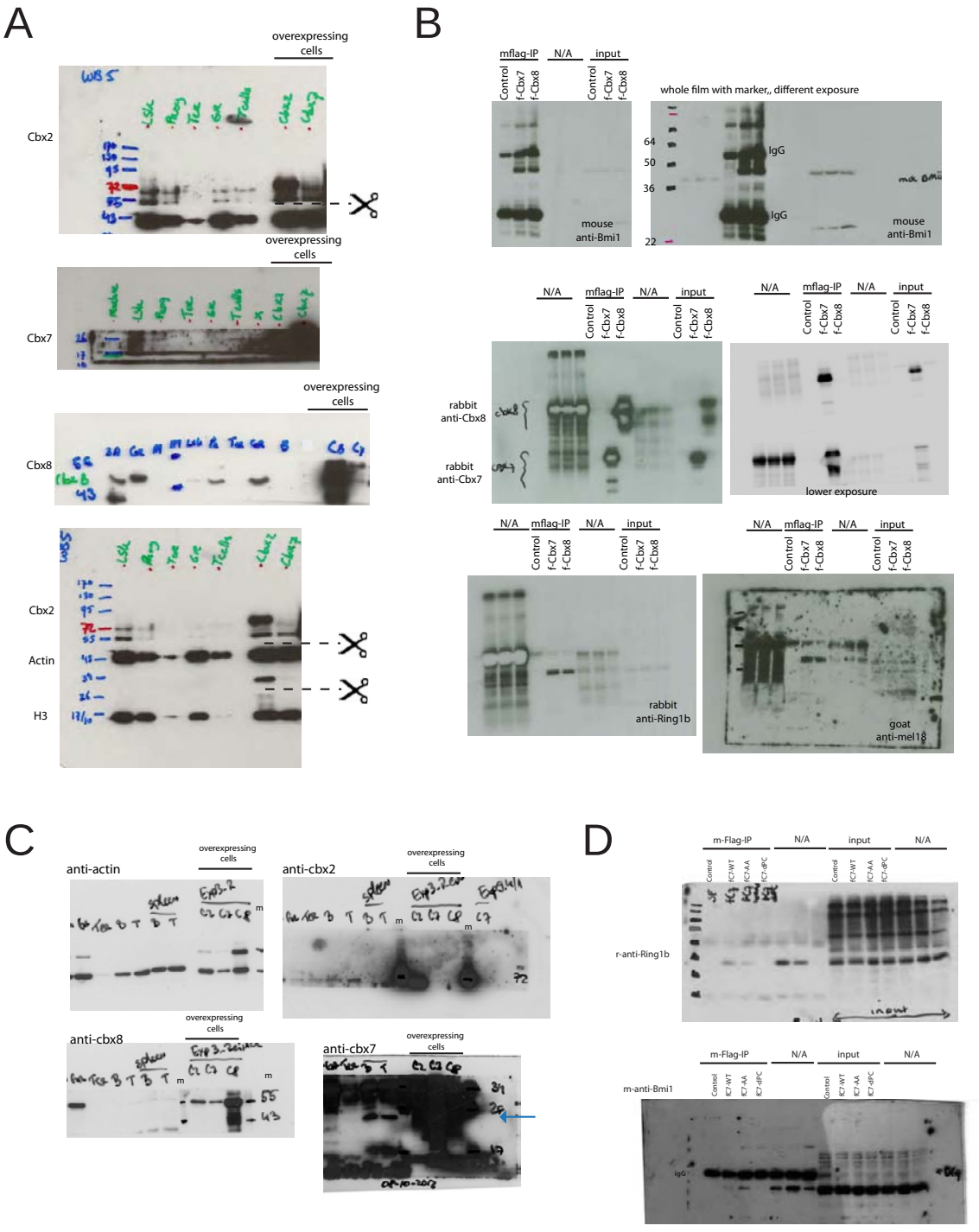

Figure S9 Uncropped western blots. Full scans of cropped westernblots of **(A)** Figure 1b, **(B)** Figure 6d, **(C)** Figure S2, and **(D)** Figure S6b.

Supplementary Tables

Supplementary Table 1: Cbx7-induced leukemic mice.
This table show cell counts and FACS analysis of cells from different tissues from all individual Cbx7-induced leukemic mice and control mice.
FTP= femur, tibia and pelvic bones, WBC= white blood cells, RBC= red blood cells

Supplementary Table 2: Differential Cbx7 and Cbx8 targets
This table shows the final lists of differential Cbx7 and Cbx8 target genes.

Supplementary Table 3: Gene-ontology analysis of differential Cbx7 and Cbx8 targets
Classification of differential Cbx7 and Cbx8 targets in GO-categories.

Supplementary Table 4: Cbx7 and Cbx8 targets show reciprocal expression patterns during HSC differentiation
Summary of all genes presented in modules 1, 2, and 3 of the gene network. Connectivity represents the amount of connections within the network (edges). In addition, gene expression data in two genetically distinct mice,C57Bl/6 (B6) and DBA (D2), of these genes in different hematopoietic cell populations are shown in worksheet 2,3 and 4. Differential clustering between Cbx7 and Cbx8 targets was tested using binomial distribution. Stem= Stem cells (Lin^-, Sca^+, $cKit^+$), Pro= progenitors (Lin^-, Sca^+, $cKit^-$), Ter=erythroid cells ($Ter119^+$), Gr=Granulocytes ($Gr1^+$). E=equal, D=down, U=up.

Supplementary Table 5: Cbx7 targets in HSCs overlap with Cbx7 targets in ESCs
Overlapping and unique Cbx7 targets in ESCs (Morey et al., 2012) and HSCs.

Supplementary Table 6: Gene-ontology analysis of common Cbx7 targets in ESCs and HSCs.
Full list of significant overrepresented GO-categories of common Cbx7 ESC and HSC targets (sheet 1). In addition, genes that were found that fall into these categories are listed in sheet 2.

Supplementary Table 7: raw data
Raw data for experiments with small *n* (*n*<5). Raw data for each figure panel are listed in separate worksheets.

Supplementary Data 1: ChIP-seq analysis of Cbx7 and Cbx8 binding sites.
ChIP-seq identified targets for Cbx7 and Cbx8 by peakcalling using empty vector as background (worksheet 1 and 2), and by cross background peakcalling (worksheet 3 and 4). In the final worksheet, overlapping Cbx7 and Cbx8 peak positions are listed.