EPIGENETICS

The Polycomb-associated factor PHF19 controls hematopoietic stem cell state and differentiation

Pedro Vizán[1]*, Arantxa Gutiérrez[1]*, Isabel Espejo[1], Marc García-Montolio[1], Martin Lange[1†], Ana Carretero[2], Atefeh Lafzi[3], Luisa de Andrés-Aguayo[1], Enrique Blanco[1], Roshana Thambyrajah[4], Thomas Graf[1], Holger Heyn[3], Anna Bigas[4,5], Luciano Di Croce[1,6,7‡]

Adult hematopoietic stem cells (HSCs) are rare multipotent cells in bone marrow that are responsible for generating all blood cell types. HSCs are a heterogeneous group of cells with high plasticity, in part, conferred by epigenetic mechanisms. PHF19, a subunit of the Polycomb repressive complex 2 (PRC2), is preferentially expressed in mouse hematopoietic precursors. Here, we now show that, in stark contrast to results published for other PRC2 subunits, genetic depletion of *Phf19* increases HSC identity and quiescence. While proliferation of HSCs is normally triggered by forced mobilization, defects in differentiation impede long-term correct blood production, eventually leading to aberrant hematopoiesis. At molecular level, PHF19 deletion triggers a redistribution of the histone repressive mark H3K27me3, which notably accumulates at blood lineage–specific genes. Our results provide novel insights into how epigenetic mechanisms determine HSC identity, control differentiation, and are key for proper hematopoiesis.

INTRODUCTION

Blood regenerates at a high level, providing an ideal platform for studying stem cell function. Hematopoietic stem cells (HSCs) in bone marrow (BM) are rare cells at the top of a hierarchical hematopoietic system; in this model, HSCs are defined by having long-term repopulation capacity and produce the more committed progenitors, which eventually produce all the differentiated cell types. Although well established in general terms, this model has been recently challenged in two ways: (i) HSCs can be already biased toward specific lineages, and (ii) new technologies (particularly single-cell transcriptomics) have demonstrated a high level of heterogeneity among HSCs, although it has been impossible to define discrete categories. Thus, it is becoming more accepted that boundaries between cell compartments, probably not only within HSCs but also in all undifferentiated progenitor compartments, are less strictly defined (*1*, *2*). In addition to being functionally defined as a group of cells capable of long-term repopulation capacity, HSCs can also be molecularly defined as cells that are able to maintain the highest level of plasticity. Strong gene regulation programs confer irreversibility during the differentiation process, and cell types/compartments can be functionally defined by their potency capacity after isolation using associated surface markers. In this context, epigenetic mechanisms are likely to play a central role in controlling and perpetuating the transcriptional programs important for the hematopoietic system. A number of epigenetic alterations have been linked to multiple hematopoietic diseases (*3*, *4*).

The Polycomb group (PcG) of proteins form multiprotein complexes that are major epigenetic regulators of gene expression. The Polycomb repressive complex 2 (PRC2) catalyzes the trimethylation of lysine-27 in the histone H3 N-terminal tail (H3K27me3), which is associated with chromatin compaction and gene repression (*5*, *6*). PRC2 comprises three core components [Enhancer of zeste homolog 2 (EZH2), suppressor of zeste 12 (SUZ12), and embryonic ectoderm development (EED)] and several substoichiometrical accessory factors that regulate its function (*7*, *8*). In the past two decades, PRC2 proteins have been studied in the context of hematopoiesis and their mutations have been linked to hematological disorders (*9*). Depletion of PRC2 core components in murine models lead to marked effects on hematopoiesis at either neonatal or fetal stage, demonstrating the crucial role of PRC2 in hematopoietic homeostasis (*10*–*12*). Three homologs of the *Drosophila melanogaster* Polycomb-like (*Pcl*) gene have been described as PRC2-associated factors: *Phf1/Pcl1*, *Mtf2/Pcl2*, and *Phf19/Pcl3*. PHF19 has been mainly characterized in mouse embryonic stem cells (mESCs) (*13*–*15*), where it plays a role in self-renewal and differentiation. *Phf19* expression declines upon mESC differentiation, although adult stem cell compartments within specific tissues still retain substantially high expression levels (*16*, *17*). Here, we report a murine model devoid of PHF19, which allowed us to investigate its role in hematopoiesis. We have uncovered how PHF19 controls HSCs identity and long-term contribution to blood production. Furthermore, we identified chromatin alterations in HSCs in the absence of PHF19 and unveiled the key role of H3K27me3 localization in HSCs differentiation and in preventing malignant hematopoiesis.

RESULTS

PHF19 plays a role in long-term hematopoietic reconstitution

In the hematopoietic system, *Phf19* expression is relatively elevated in undifferentiated progenitors and progressively decreases during differentiation (fig. S1A) (*18*, *19*). To study its role in adult homeostasis, we generated a mouse genetic model by deleting exons 4 and 5, which produces a stop codon in exon 6 (fig. S1, B to E). *Phf19*-KO (knockout) mice breed normally, and histopathological analysis of

[1]Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Dr. Aiguader 88, Barcelona 08003, Spain. [2]Department of Animal Health and Anatomy, School of Veterinary Medicine, Universitat Autònoma de Barcelona, Bellaterra, Spain. [3]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology, Baldiri Reixac 4, Barcelona 08028, Spain. [4]Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona 08003, Spain. [5]CIBERONC, Barcelona, Spain. [6]Universitat Pompeu Fabra (UPF), Barcelona, Spain. [7]ICREA, Pg. Lluis Companys 23, Barcelona 08003, Spain.
*These authors contributed equally to this work.
†Present address: Bayer AG, Pharmaceuticals Division, Research and Development, Berlin, Germany.
‡Corresponding author. Email: luciano.dicroce@crg.eu

young adult mice (8 to 12 weeks) did not exhibit any major alterations in the brain, lungs, heart, stomach, intestine, kidney, testis, ovary, or uterus. As reported for other Polycomb members (*20*, *21*), *Phf19*-KO mice presented anterior-to-posterior homeotic transformations, such as the presence of rib anlage or extra ribs in ~40% (6 of 14) of animals (fig. S1F; note that no alterations in *Phf19*-Flox mice were detected). Aged *Phf19*-KO mice (>60 weeks) presented a high penetrance of splenomegaly (fig. S1G), suggestive of a defect in the hematopoietic system.

We analyzed the hematopoietic hierarchy of fetal liver [embryonic day 14.5 (E14.5)] and of 8-week-old mouse BM. Lack of PHF19 led to an increase in embryonic phenotypically defined HSCs (Lineage$^-$ Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$), but in adults, we observed a reduction in HSCs, as well as in LSK (Lineage$^-$Sca-1$^+$c-Kit$^+$) progenitors (Fig. 1A). We also documented a significant reduction in cellularity (raw cell count) in BM (Fig. 1B). Furthermore, we investigated the adult HSC compartment by assessing the proportion of proliferative Ki67-positive cells (Fig. 1C) and of bromodeoxyuridine (BrdU) incorporation during 24 hours (Fig. 1D). Our data indicated a significant increase in quiescent cells in *Phf19*-KO–derived HSCs, which may account for the observed decrease in LSK, HSC cell populations, and total cellularity in adults.

To determine the role of *Phf19* in BM performance, we performed transplantation assays. Initially, we transplanted 0.5 million cells of whole BM (WBM) obtained from 8-week-old *Phf19*-Flox or *Phf19*-KO mice in competition with 0.5 million cells of CD45 congenic wild-type WBM cells. No differences were observed in the engraftment capacity or in blood contribution (fig. S1H). We then performed serial noncompetitive repopulation assays by transplanting 1 million cells of *Phf19*-Flox or *Phf19*-KO WBM and monitoring blood contribution after 3 months of each transplant (Fig. 1E and fig. S1I). Our analysis indicated a slight bias toward myeloid lineage in *Phf19*-KO transplanted mice, which was only statistically significant after the third transplant (data not shown). We observed a clear decrease in the contribution of *Phf19*-KO transplanted donors in serial transplants, pointing to a malfunction of HSCs. Colony-forming assays from lineage-specific depleted BM showed a similar behavior: No differences were found in the type or the total number of colonies after the first plating (fig. S1J), but after replating, *Phf19*-KO–derived cells were significantly impaired in their capacity to produce new colonies (Fig. 1F). Parallel experiments using progenitor LSK population led to similar results (fig. S1K).

Phf19-KO HSCs are mobilized upon stress but are not fully functional

To determine whether the decrease in BM HSC levels under homeostatic conditions (Fig. 1A) is the cause of the long-term lower efficiency in blood production (Fig. 1E), we analyzed the total proportion of LSK progenitors and HSCs upon different stress conditions that have been described to force HSC mobilization: aging, transplantation, and 5-fluorouracil (5-FU) injection. The hematopoietic compartment analysis showed that the reduced number of HSCs in *Phf19*-KO animals was compensated upon stress (Fig. 2A and fig. S2A), indicating that *Phf19*-KO HSCs are able to proliferate upon stress. Therefore, the reduced long-term capacity for replenishing the blood system might be related to a differentiation defect.

To assess cell division kinetics and differentiation capacity of PHF19-depleted HSCs, we next performed single-cell culture of phenotypically defined HSCs (*22*, *23*). Single Lineage$^-$Sca-1$^+$ c-Kit$^+$CD150$^+$CD48$^-$ HSCs were individually sorted on 96-well plates, and the number of cells was monitored for 4 days (Fig. 2B). Each individual well was scored according to its division dynamics: proliferating (i.e., sustained cell growth over the 4 days, rendering colony sizes from five to hundreds of cells), nonproliferating (i.e., not able to divide more than twice during 4 days), and proliferating and stop (i.e., division that eventually stopped or reduced the number of cells). Under culture conditions, most wells were scored as proliferating, both in *Phf19*-Flox and *Phf19*-KO plates, but a significant increase in the percentage of proliferating colonies in *Phf19*-KO wells was observed, which correlates with a reduction in nonproliferating wells (Fig. 2C). Notably, not only *Phf19*-KO wells had more proliferating colonies, but also these colonies individually contained a higher number of cells (fig. S2B). We kept colonies in culture for additional 5 days and then calculated colony size by fluorescence cytometer after propidium iodide (PI) staining. According to the division status during the first 4 days, the number of big colonies (>5000 cells) was higher in the *Phf19*-KO plates than in the *Phf19*-Flox plates (fig. S2C). Last, after 2 weeks, we measured the expression of lineage and HSCs markers. We observed an impairment of differentiation in *Phf19*-KO condition (Fig. 2D). Moreover, the remaining CD150$^+$CD48$^-$ population was higher in the *Phf19*-KO than in the *Phf19*-Flox cells (fig. S2D). Together, these data demonstrate that HSCs from *Phf19*-KO cells have higher proliferation but lower differentiation capacity compared to control cells. This is in agreement with our in vivo data, which indicate that *Phf19*-KO HSC cells are not able to fully differentiate in serial transplant experiments despite its increased capacity of generating progenitors upon stress.

PHF19 regulates HSC gene expression through epigenetic control of PRC2 activity

To gain insight into how PHF19 is able to control the quiescence state in homeostasis, we first performed RNA sequencing (RNA-seq) transcriptomic analysis on HSCs (Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$) isolated from young adult mice (table S1). Term enrichment analysis of common up-regulated genes of two independent replicates showed that nuclear receptors and, in particular, retinoic acid category were overrepresented. On the other hand, down-regulated genes were enriched for categories related to biosynthesis and energy production (fig. S3A). This is reminiscent of a recently characterized, highly unbiased, and quiescent HSC population, the so-called dormant HSCs (*24*).

We next did a gene set enrichment analysis (GSEA) for specific gene sets. HSC-specific signatures, a merge of differentially expressed genes provided by Gazit *et al.* (*25*) and Chambers *et al.* (*26*), were enriched in *Phf19*-KO (Fig. 3, A and C), indicating that HSC identity is enhanced upon PHF19 depletion. Furthermore, the Myc network, which is required for HSC differentiation (*24*, *27*), was down-regulated in *Phf19*-KO HSCs (Fig. 3, B and C). In addition, a decrease of enrichment in biosynthetic pathways (such as ribosome or mRNA processing) and an increase in positively regulated retinoic acid targets (*28*) were observed (Fig. 3C and fig. S3B), further supporting an intensification in HSC quiescence and dormancy (*24*, *29*). We also compared our data with an expression signature derived from HSCs depending on their divisional history (*30*). Up-regulated genes in HSCs that have undergone several divisions inversely correlate with gene expression of *Phf19*-KO HSCs and vice versa for down-regulated

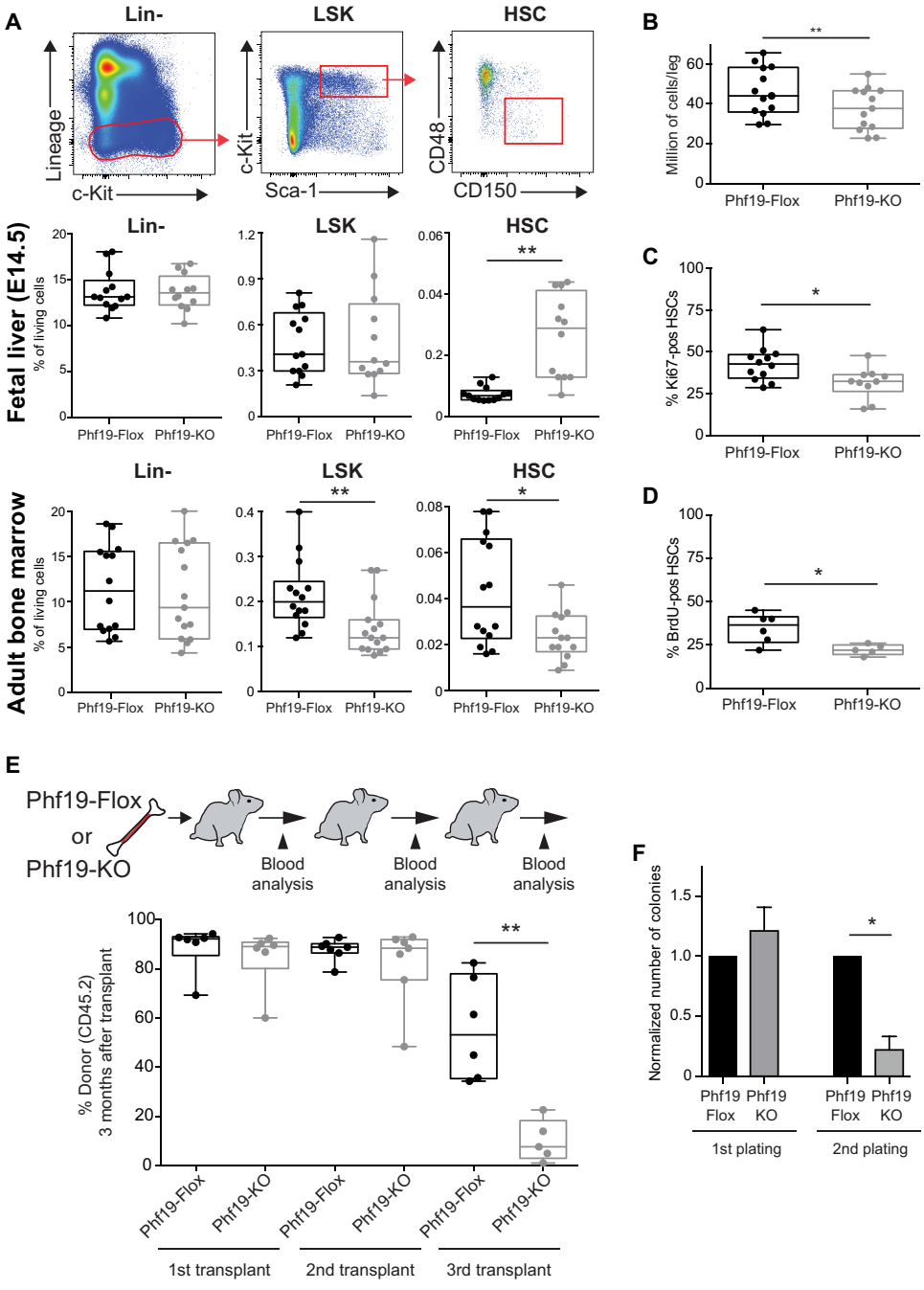

Fig. 1. Characterization of the hematopoietic system in a *Phf19*-depleted mouse model. (**A**) Top: BM gating strategy for phenotypically defining lineage negative (Lin$^-$), LSK, and HSCs (Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$). Middle: Quantification from 13 control floxed (*Phf19*-Flox) and 12 *Phf19*-KO E14.5 embryos, with percentages shown for the live cells measured as 4′,6-diamidino-2-phenylindole (DAPI) negative (box plot). Bottom: Quantification from 14 control floxed (*Phf19*-Flox) and 13 to 15 *Phf19*-KO mice, with percentages shown for the live cells measured as DAPI negative (box plot). (**B**) Raw number of WBM cells counted per leg in 13 independent experiments (with at least three animals per experiment) (box plot). (**C**) Percentage of Ki67$^+$ HSC cells in 12 mice for *Phf19*-Flox and 10 mice for *Phf19*-KO (box plot). (**D**) Percentage of BrdU$^+$ HSCs in six *Phf19*-Flox mice or five *Phf19*-KO mice at 24 hours after BrdU injection (box plot). (**E**) Percentage of donor-derived (CD45.2) cells in peripheral blood in serially transplanted recipients 3 months after each transplant (box plot). (**F**) Normalized number of colonies in the first and second lineage-negative plated cells for colony-forming assay, performed in three independent replicates (means + SEM). *$P < 0.05$ and **$P < 0.01$. Unpaired *t* test (A and C to E). Paired *t* test (B and F).

genes (fig. S3B). Moreover, since *Phf19*-KO HSCs are proliferative but defective in differentiation and long-term repopulation, we wondered whether they resembled aged HSCs (*31*). GSEA demonstrated that up-regulated genes in old animals are enriched in *Phf19*-KO condition; meanwhile, down-regulated genes are enriched in *Phf19*-Flox transcriptome (fig. S3B). Together, transcriptomics data corroborated that *Phf19*-depleted cells in homeostasis are in a more multipotent and quiescent state and have a decreased capacity upon

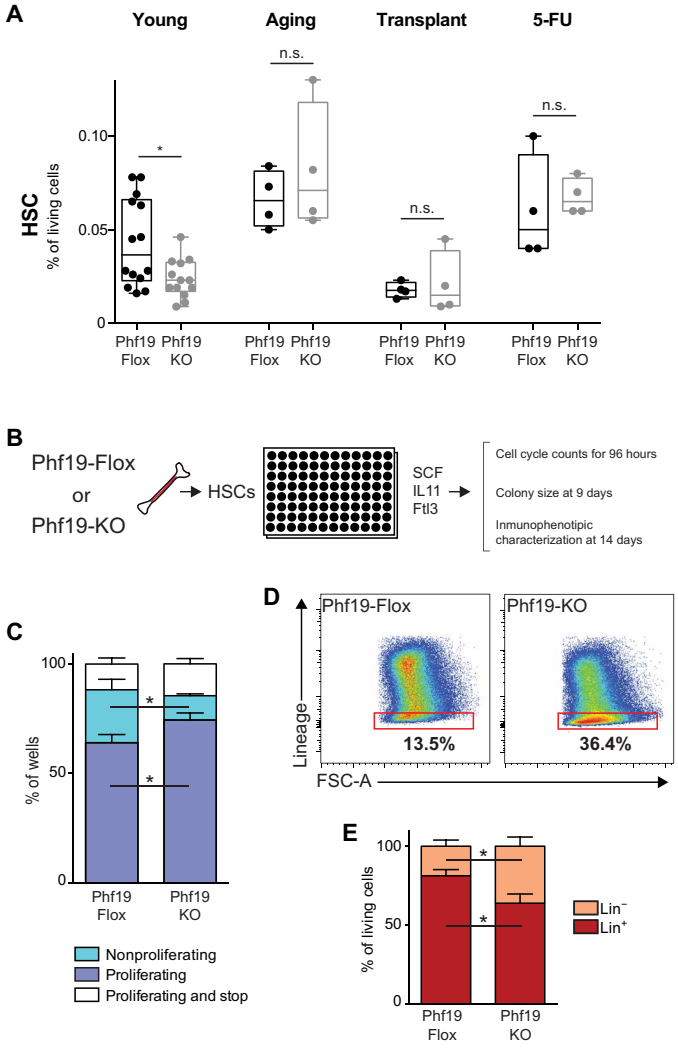

Fig. 2. Characterization of *Phf19*-KO HSCs after mobilization. (**A**) Quantification of phenotypically defined HSCs in young or aged (>60-week-old) mice after transplantation (at 6 months after transplant) and at 14 days after 5-FU injection. The percentage of the live cells measured as DAPI negative is shown (box plot). (**B**) Schematic representation of a single-cell HSC in vitro functional assay. (**C**) Proportion of wells of proliferating, nonproliferating, and proliferating and then stopped HSCs in three independent experiments (means + SEM). (**D**) Gating of lineage-negative marker versus FSC-A (forward scatter area) of pooled wells after 14 days in culture. (**E**) Percentage of lineage-negative and lineage-positive cells from three independent replicates (means + SEM). *$P < 0.05$; n.s., not significant. Paired *t* test (C and E).

induction of differentiation. Last, the role of Jumonji and AT-rich interaction domain containing 2 (JARID2) in the hematopoietic system has been recently investigated (*32*). A comparative analysis of transcriptomic data revealed a similar effect for JARID2 and PHF19 on HSCs gene regulation (fig. S3B).

PHF19 is an associated PRC2 factor (*7*) and controls the trimethylation on lysine-27 of histone H3 (H3K27me3) on chromatin (*8*). We thus performed chromatin immunoprecipitation sequencing (ChIP-seq) experiments on Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$ population, and unexpectedly, we found a global increase in H3K27me3 levels, even when spike-in control for quantitative measurements between conditions was used (fig. S3C). We also corroborated that such increase is global and not only detectable inside the peaks (fig. S3D). This acquired H3K27me3 deposition pattern resembled the one observed in aged HSCs (*31*) since genes marked with H3K27me3 in aged HSCs intensively accumulate H3K27me3 upon depletion of PHF19, and conversely, the H3K27me3 genes in young HSCs are also more intensely marked with H3K27me3 in *Phf19*-Flox HSCs (fig. S3E). At molecular level, Polycomb complexes regulate differentiation by restricting lineage-determining gene expression (*5*). We thus reasoned that in *Phf19*-KO cells, this mechanism could be implicated in the increased HSC identity observed at a functional and transcriptomic level. As expected, the genomic accumulation of H3K27me3 levels inversely correlated with expression data (fig. S3F). However, despite the global increase in H3K27me3, the HSC gene sets did not show a significant increase in H3K27me3 levels (Fig. 3D, left). Significant increase was observed for genes that are characteristic of differentiated lineages (Fig. 3D, middle). Moreover, transcription factors implicated in HSC differentiation into erythroid, granulocyte, lymphoid, or megakaryocyte (*20*) also showed a significant increase in H3K27me3 (Fig. 3D, right). Increased H3K27me3 deposition in lineage-specific transcription factors could be observed for essential genes—such as *Cebpa* (for myeloid) (*33*); *Pax5* and *Irf5* (for B cells) (*34*, *35*); Tcf7 (for T cells) (*36*); several members of Notch pathway [*Dll1* (depicted), *Dll4*, *Jag2*, or *Hes1*], which are well-known regulators of both T and B cells differentiation (*37*); and Epo (for erythroid cells) (*38*)—and could not be observed in HSC fingerprint genes (*Vdr* depicted as an example; fig. S3G). To further validate how the depletion of PHF19 influences chromatin organization, we performed assay for transposase-accessible chromatin and sequencing (ATAC-seq) and verified that the increase in H3K27me3-containing regions was accompanied by a loss of chromatin accessibility (fig. S3H). This decrease could be visualized in most of the master differentiation transcription factors mentioned (Fig. 3E).

Differentiation program genes are poorly expressed in HSCs. Therefore, to evaluate the impact of chromatin changes on gene expression, we designed two strategies. First, we performed single-cell RNA-seq. Our first observation was that *Phf19*-KO HSCs were transcriptionally more similar between each other than *Phf19*-Flox HSCs, both using all expressed genes (fig. S3I) and using a HSC signature specifically defined from single-cell data (Fig. 3F) (*23*). Phenotypically defined HSC comprises a heterogeneous population, which contains cells primed for differentiation. We reasoned that we could detect transcripts of differentiation genes in individual HSC cells. We thus computed the number of cells expressing differentiation gene set (Fig. 3D, middle) (*25*, *26*). As expected, some, but not all, of the cells expressed differentiation genes. Interestingly, we observed we observed a significant decrease in the number of cells that expressed these genes in the *Phf19*-KO condition (Fig. 3G). Second, upon forced in vitro differentiation (as depicted in Fig. 2B), we monitored by quantitative polymerase chain reaction (qPCR) the expression of transcription factors that presented an accumulation of H3K27me3 mark within their promoter regions. As represented in Fig. 3H, all genes showed a decrease in mRNA expression, with the only exception of *Cebpa*. Since differentiation is not completely abolished in *Phf19*-KO cultured HSCs, we hypothesized that myeloid lineage could be enhanced (see next section), which is in agreement with the increased expression levels of *Cebpa*. Together, our data indicate that the absence of PHF19 produces an epigenetic rewiring of HSCs, leading to an enhancement of stem identity likely instructed by an aberrant repression of differentiation programs.

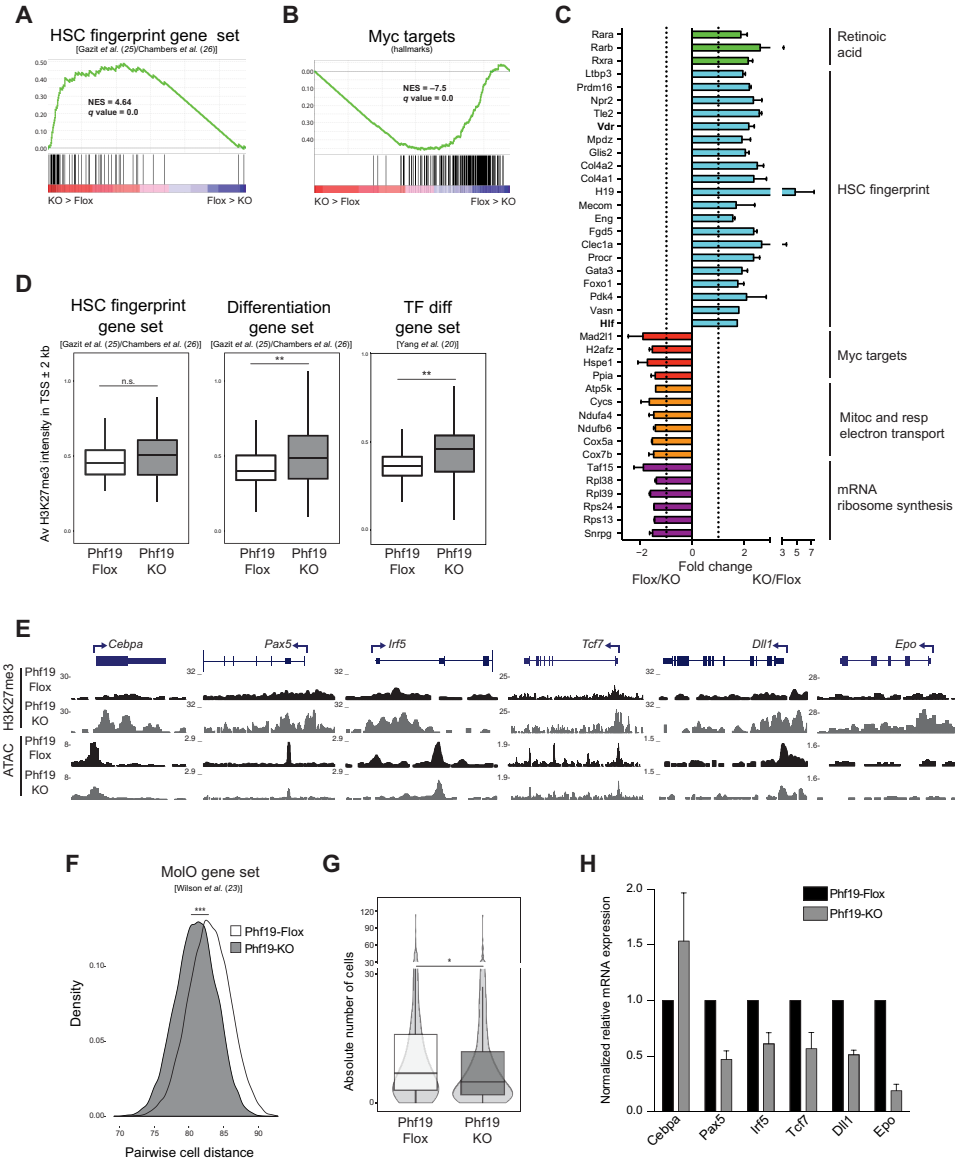

Fig. 3. HSC gene expression control and epigenetic changes associated with *Phf19* depletion. (**A**) GSEA showing positive enrichment in *Phf19*-KO transcriptome for HSC gene set [Gazit *et al.* (*25*) and Chambers *et al.* (*26*)]. (**B**) GSEA showing negative enrichment in *Phf19*-KO transcriptome for Myc target gene set. (**C**) Normalized fold change expression from two independent RNA sequencing (RNA-seq) experiments in relevant genes (means + SD). (**D**) ChIP-seq levels of H3K27me3 in the transcription start site (TSS) ± 2 kb of specific gene sets: the HSC gene set, the differentiation gene set, and for transcription factors associated with hematopoietic differentiation processes. (**E**) University of California Santa Cruz (UCSC) genome browser screenshots for H3K27me3 and ATAC of hematopoiesis master regulators in *Phf19*-Flox and *Phf19*-KO. (**F**) Density plot of pairwise distances of cells based on single cell–based HSC (MolO) signature (*23*) genes for *Phf19*-Flox and *Phf19*-KO cells. (**G**) Distribution shown by violin plot and box plot of absolute number of cells that expressed differentiation genes for 174 *Phf19*-Flox and *Phf19*-KO cells. (**H**) mRNA expression levels of cultured HSCs after 9 days under growth/differentiation conditions depicted in Fig. 2B, relative to *Rplp0* and normalized for *Phf19*-Flox of two replicates (means + SEM). **P* < 0.05, ***P* < 0.01, and ****P* < 0.001. Paired *t* test (D). Wilcoxon rank-sum test (F and G). NES, normalized enrichment score.

Lack of PHF19 causes long-term aberrant hematopoiesis

To study the consequences of the increased proliferation and impairment of differentiation observed in *Phf19*-KO HSCs, we performed WBM transplant from aged mice, which already presented defects in hematopoiesis (as revealed by splenomegaly; fig. S1G), to irradiated young recipient mice. We again detected a peripheral blood lineage myeloid bias after serial transplant (fig. S4A), and impairment of blood contribution by *Phf19*-KO donors was accelerated and already evident in the first and second rounds of transplant (fig. S4B). Moreover, a high incidence of splenomegaly was observed in mice euthanized for WBM extraction for the second transplant (Fig. 4A). Moreover, survival of *Phf19*-KO versus *Phf19*-Flox transplanted mice was significantly reduced (Fig. 4B).

Flow cytometry characterization of BM from survivor mice showed the presence of an uncharacterized population (fig. S4C), suggesting anomalous hematopoiesis. We next analyzed spleens of the same young transplant survivor mice (four of six with splenomegaly in *Phf19*-KO and one of six with splenomegaly in *Phf19*-Flox)

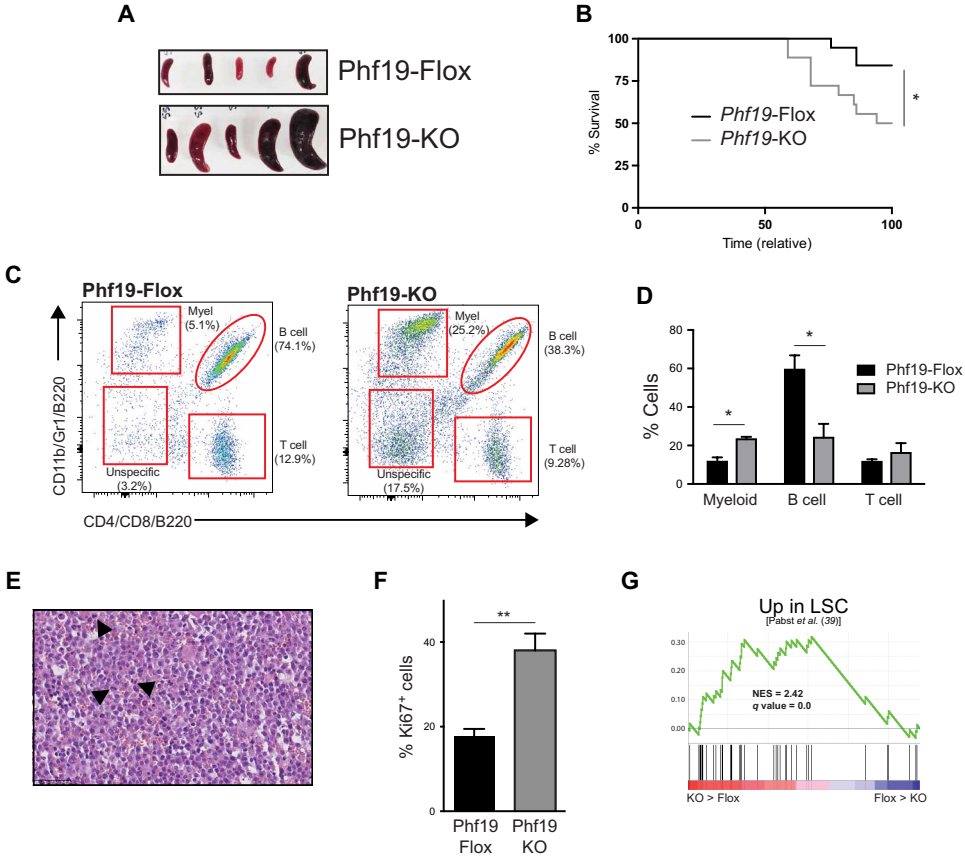

Fig. 4. Lack of Phf19 causes aberrant hematopoiesis over the long term. (**A**) Picture depicting splenomegaly at 6 months after transplantation with WBM from aged mice. Photo credit: Pedro Vizán, CRG. (**B**) Survival curve of all mice (*Phf19*-Flox versus *Phf19*-KO) after their last transplant. Time normalized for each experiment to final sacrifice time of all animals. (**C**) Representative gating of whole spleen analyzed for myeloid, B cell, and T cell proportions (donor CD45.2-positive cells present in the spleen). (**D**) Quantification of donor CD45.2-positive myeloid, B cell, and T cell proportions in five *Phf19*-Flox and three *Phf19*-KO mice. In *Phf19*-KO conditions, the sum does not reach 100% because of the unspecific population observed in the panel for hematoxylin and eosin staining (means + SEM). (**E**) Representative hematoxylin and eosin staining of *Phf19*-KO spleens. Arrowheads indicate mitotic figures. (**F**) Percentage of Ki67-positive cells in two spleens per condition (more than four white pulp areas per spleen) (means + SEM). (**G**) GSEA showing positive enrichment in *Phf19*-KO transcriptome for leukemic stem cell (LSC) differentially up-regulated genes (*37*). $^{*}P < 0.05$ and $^{**}P < 0.01$. Log-rank (Mantel-Cox) test (B). Unpaired *t* test (D and F).

by flow cytometry. When we evaluated donor CD45.2 cell contributions, we observed a decrease in B cell content together with an increase in myeloid infiltration (Fig. 4, C and D), as well as the presence of an unusual, unspecific splenic cell population (Fig. 4C). Further characterization of the spleens by histological analysis revealed that the tissue architecture was compromised in *Phf19*-KO transplanted mice, with almost indistinguishable white and red pulps (fig. S4D), together with the presence of a population of large pleomorphic cells with irregularly shaped nuclei (Fig. 4E). The cytological features of these cells are compatible with myeloid dysplasia. Many mitotic figures were present, and immunohistochemistry for the proliferation marker Ki67 corroborated a higher rate of proliferation in *Phf19*-KO spleens than in *Phf19*-Flox spleens (Fig. 4F). Moreover, the macrophage marker CD68, which is found mostly in the red pulp (fig. S4E, left), was diffusely spread in all areas of the *Phf19*-KO samples (fig. S4E, right), in accordance with our flow cytometry observation of unspecific and increased myeloid cells.

In light of these results, we reevaluated our transcriptomic data. Term enrichment of up-regulated genes in *Phf19*-KO condition also showed the pathways in cancer category [Kyoto Encyclopedia of Genes and Genomes (KEGG) mouse 2019, $P = 9.16 \times 10^{-7}$]. Furthermore, a set of differentially up-regulated genes found in a human leukemic stem cell (LSC) population (*39*) was enriched in *Phf19*-KO versus *Phf19*-Flox (Fig. 4G). It has also been observed that LSC transcriptional programs are partially shared with stem cell transcription programs, and a reduced gene catalog of 17 genes was reported to correlate with poor prognosis and relapse for acute myeloid leukemia (*40*). We manually examined these genes in our transcriptomic data. Thirteen of the original 17 genes were expressed [reads per kilobase million (RPKM) > 0.1] in our dataset, and 10 (~77%) were up-regulated. These observations are quite notable since these specific alterations in the transcriptome are already evident in young mice. Together, our results indicate that PHF19 is necessary to maintain a correct hematopoietic balance and that its depletion induces an aberrant state, which primes HSCs for malignant progression.

DISCUSSION

Here, we have shown that PHF19 contributes to maintain the correct levels of H3K27me3 within the HSC compartment, influencing their long-term function and differentiation predisposition. Blood production has been classically explained by a hierarchical model in

which HSCs have the capacity to self-renew and differentiate into a population of more committed progenitors, which are still able to self-renew and produce all blood types of the organism. In the past few years, this model is being reinterpreted as a continuous change between cellular states rather than a stepwise process. Within the HSC compartment, variability is given by characteristics that are not initially related to their differentiation potential but rather with their cellular state, such as biosynthetic rate or energy metabolism. Thus, HSCs have been further characterized in dormant and active with a continuum among the two states (*24*). Our results indicate that PHF19 controls the balance between these two conditions.

Epigenetics is one of the main factors that govern cell plasticity, as well as the transition towards new cell states. In particular, the PcG of proteins, which main function is transcriptional repression, have been reported to play a role in HSC performance (*9*). Loss-of-function studies of PRC2 core components demonstrated their essential role in hematopoiesis (*10*–*12*), but the associated lethality has hampered an in-depth characterization of transcriptional pathways and H3K27me3 (re)distribution implicated in HSC identity. Similarly, metal response element binding transcription factor 2 (MTF2) deletion impairs fetal erythropoiesis (*41*), causing a prebirth death that has also impeded further studying of the role of H3K27me3 in adult HSCs. Several studies have unveiled that PCL proteins and JARID2 give rise to two biochemically distinct versions of the PRC2 complex: PRC2.1 and PRC2.2 (*6*). It has been recently reported that JARID2 depletion in hematopoietic precursors (*32*) affects the expression of self-renewal genes in HSCs and multiple progenitor precursors. Transcriptional comparison shows that JARID2-depleted HSCs resemble *Phf19*-KO HSCs (fig. S3B) despite a reduction in H3K27me3 levels. We believe that PRC2 variants control self-renewal and differentiation by complementary mechanisms. Notably, PRC2.2 may be more active than PRC2.1, as previously suggested (*42*). Thus, PRC2.2 reduction might cause direct derepression of a set of HSC genes, and conversely, PRC2.1 reduction might allow PRC2.2 to increment H3K27me3 deposition at differentiation genes, producing a more homogeneous HSC-like transcriptome. These two distinct molecular mechanisms would paradoxically produce a similar phenotype: increase of HSC/dormancy features and impairment of proper differentiation. Together, this highlights the importance of the proper balance between PRC2 variations and their specific role with respect to the control of adult stem cell functionality. Unfortunately, nowadays, there are no available data of PRC2 component localization at genome-wide level, which impedes a thorough analysis of the impact of distinct PRC2 versions in HSC gene regulation.

As we have demonstrated in this study, *Phf19* deletion does not affect breeding or life span. Although we observed changes in the number of HSCs (higher proportion in fetal liver and lower proportion in young adults), these alterations do not have fatal consequences during development or in adult homeostasis under animal facility conditions. However, under challenging conditions, such as aging or serial transplants, we have unveiled the role for PHF19 in controlling HSC functions. In *Phf19*-KO animals, dormant HSC characteristics are enhanced under resting conditions, such as high quiescence, or low metabolic and biosynthetic activities. Upon induced mobilization, proliferation is triggered, but differentiation is impaired. This double-edged behavior could be linked to epigenetic states. The main role of PcG is to maintain cell identity by restricting lineage-specific gene transcription. In *Phf19*-KO HSCs, we observed an increase in H3K27me3 mark and chromatin compaction at genes necessary for differentiation. At the same time, once differentiation is triggered, this compaction would reduce the accessibility of specific transcription factors.

In summary, we have demonstrated that PHF19 depletion increases the dormancy-like HSC state by a mechanism that, at the same time, impairs differentiation and reduces durability upon serial transplants. Several studies have used label dilution models to quantify HSC divisional history, which has allowed a link between low mitotic rate and greatest repopulation ability (*30*, *43*). The transcriptional program associated with low divisional history resembles the *Phf19*-KO expression data (fig. S3B). Thus, although we observed a higher number of HSCs in fetal liver, which could be indicative of an increased proliferation, adult HSCs are more quiescent. This could be explained by an early exhaustion of highly proliferative HSCs during fetal and early life stages or by the fact that label dilution models tend to misinterpret label retention and are simply labeling HSCs with high repopulation potential, irrespective of mitotic history (*44*). Beyond these strong functional and transcriptional similarities, the quiescent features in *Phf19*-KO do not yield in fully functional dormant HSCs. In this regard, it is remarkable that *Phf19*-KO HSCs resemble aged HSCs (fig. S3B), which have been observed to accumulate during aging, but are less functional at generating new blood cells (*45*). Notably, the H3K27me3 increase observed upon PHF19 depletion is in line with the same mark distribution reported for aged HSCs (fig. S3E) (*31*). As for the *Phf19*-KO mice, the consequences in aged HSCs are a myeloid bias and a higher propensity to hematological disorders. Moreover, quiescence is the most significant feature in cancer stem cells and the leading cause of relapse. This study underlines the complex balance between the silencing of differentiation programs and the maintenance of elevated levels of plasticity to sustain stem cell identity.

METHODS

Mice

All animal procedures were approved by the ethical committee of the Barcelona Biomedical Research Park and by the animal experimentation commission of the Catalan government (Generalitat de Catalunya). All animals used for this study had a C57BL/6 background. LoxP sites were introduced by homologous recombination around exons 4 and 5, the deletion of which generates a stop codon in exon 6. $Phf19^{fl/fl}$ animals were crossed with Sox2Cre (mice were provided by A.R. Nebreda, Institute for Research in Biomedicine, Barcelona). The CRE activity in the Sox2CRE strain has been of maternal inheritance, irrespective of whether they inherit the transgene (*46*). Therefore, after backcrossing with new male floxed animals, animals with the knocked-out allele without the SoxCre transgene were selected. Once generated, mice were maintained as heterozygosis ($Phf19^{-/fl}$) and as homozygosis ($Phf19^{-/-}$). In both cases, offspring were viable and bred normally. Equal numbers of male and female mice were used; no gender biases were noted. For BM transplants experiments, 8- to 12-week-old congenic CD45.1 and CD45.1/CD45.2 mice (bred in-house) were used.

BM extraction and splenocytes isolation

For fetal liver analysis, the tissue was harvested from E14.5 embryos and the dissected liver was crushed with the end of a 1-ml syringe through a 40-μm cell strainer into Iscove's modified Dulbecco's medium and 10% fetal bovine serum. For cell BM analysis of individual

mice, clean femurs and tibias were flushed using a 23-gauge needle and a 10-cm^3 syringe with cold dissection medium [1× phosphate-buffered saline (PBS), 2% fetal calf serum (FCS), and 2 mM EDTA] before cell surface staining (see next section). For pooled mice analyses, the clean bones were crushed with mortar and pestle in cold dissection medium (1× PBS, 2% FCS, and 2 mM EDTA). For sorting and colony-forming unit assays (see below), lineage depletion was performed using the mouse Lineage Cell Depletion Kit (Miltenyi Biotec) by magnetic-activated cell sorting (MACS) technique according to the manufacturer's instructions. Splenocytes were freshly dissociated by mincing the tissue to obtain a single-cell suspension.

Flow cytometry

Before flow cytometry analysis, fetal liver cell suspension, blood, BM, and splenocytes, erythrocytes were lysed with 1× red blood cell lysis buffer (eBioscience) and washed with 1× PBS. For surface staining, single-cell suspensions in fluorescence-activated cell sorting (FACS) buffer (1× PBS, 2% FCS, and 2 mM EDTA) were blocked for 10 min with anti-CD16/CD32 (mouse Fc block, BD Pharmingen). Samples were stained in FACS buffer for 30 min at 4°C with fluorochrome-conjugated antibodies and washed before analysis. Blood and splenocytes were incubated with the following fluorochrome-conjugated monoclonal antibodies for lineage analysis: phycoerythrin (PE) anti-mouse CD4, PE anti-mouse CD8, PE anti-mouse B220, PECy7 anti-mouse B220, allophycocyanin (APC) anti-mouse CD45.1, fluorescein isothiocyanate (FITC) anti-mouse CD45.2, PECy7 anti-mouse CD11b, and PECy7 anti-mouse GR-1. Fetal liver and BM were analyzed using the lineage marker (Lin) mixture (BD Biosciences) that include the following V450-conjugated antibodies: CD3ε (145-2C11); CD11b (M1/70) (this antibody was excluded in fetal liver analyses); CD45R/B220 (RA3-6B2); mouse erythroid cells Ly-76 (Ter119), Ly6G, and Ly-6C (RB6-8C5); and the following fluorochrome-conjugated antibodies for HSC and precursors (LSK): PECy7 anti-mouse Sca I, APC-Fluo780 anti-mouse cKit, APC anti-mouse CD150, and PE anti-mouse CD48. Cell viability was evaluated with 4′,6-diamidino-2-phenylindole (DAPI) at a final concentration of 1:10,000.

For intracellular staining (Ki67 or BrdU), BM cells were resuspended in FACS buffer and stained for 10 min at 4°C with anti-CD16/CD32 (mouse FC block, BD Pharmingen) for 30 min at 4°C with a cocktail of lineage marker biotin-conjugated monoclonal antibodies against CD2, CD3, CD11b, CD14, CD15, CD16, CD19, CD56, CD123, and CD235a (Miltenyi Biotec) and, lastly, with fluorochrome-conjugated antibodies specific for HSCs (as detailed above). Cells were then washed and incubated with the secondary antibody PE-CF594 streptavidin for 10 min at 4°C. For Ki67 immunostaining, surface-stained BM cells were washed and fixed with reagent A (Invitrogen) for 15 min at room temperature and then washed and permeabilized in reagent B in the presence of FITC anti-mouse Ki67 for 20 min at room temperature. After two washes, cells were resuspended in 500 μl of 1× PBS containing ribonuclease (RNase; 5 μg/ml) and DAPI (2 μg/ml). For BrdU (see next section for BrdU in vivo injection), surface-stained BM cells were washed, and the BrdU incorporation assay was performed using the BrdU Kit from BD Biosciences.

BrdU incorporation assay

Subject mice were intraperitoneally injected twice (at 12-hour intervals) with BrdU (150 μl of 10 mg/ml). At 12 hours after the second injection, mice were euthanized, and BM cells were isolated and stained for cytometry as described above.

BM transplantation

For competitive transplantation assay, 0.5×10^6 million nucleated BM cells from femurs and tibias (see next section) of 8- to 12-week-old CD45.2 mice were mixed with 0.5×10^6 million CD45.1 BM cells and transplanted by retro-orbital injection into 8- to 12-week-old CD45.1/CD45.2 recipient mice that were lethally irradiated using a split dose of 8 gray (Gy) (two doses of 4 Gy administered at least 4 hours apart). CD45.2 contribution to peripheral blood was assessed 3 months after transplantation. For both young and old serial transplantation assays, 1×10^6 nucleated BM cells from 8- to 12-week-old (young) or 80- to 85-week-old (old) CD45.2 mice were transplanted by retro-orbital injection into 8- to 12-week-old lethally irradiated (two split doses of 4 Gy, 4 hours apart) CD45.1/CD45.2 recipient mice. The CD45.2 contribution to peripheral blood was assessed at 3 months after transplantation. For serial transplants, animals were euthanized after 6 months, and pooled 1×10^6 nucleated BM cells were used to transplant new 8- to 12-week-old lethally irradiated CD45.1/CD45.2 recipient mice. Recipient mice were treated for 4 weeks after transplantation with Enrovet (enrofloxacin) at 600 μg/ml in their drinking water.

Colony-forming assay

Lineage-depleted or LSK-sorted BM cells were plated in duplicate in 35-mm tissue culture dishes (STEMCELL Technologies) in methylcellulose-based medium MethoCultTM GF M3434 (STEMCELL Technologies) at 2000 cells per plate. Plates were incubated 10 to 14 days at 37°C, 100% humidity, and 5% CO_2. Colonies were scored using an inverted microscope and resuspended and counted in 1× PBS. Cells at 100,000 per plate were used for replating.

Single-cell HSC culture

Lineage-depleted BM cells were stained as previously described for HSC. Single-cell sorting (FACSAria, BD) was performed in 96-well plates containing 100 μl of filtered HSC expansion medium [StemSpan serum-free expansion medium, murine stem cell factor (300 ng/ml), murine interleukin-11 (20 ng/ml), and human Flt3L (1 ng/ml)]. The plates were incubated at 37°C with 5% CO_2, and medium was replaced every 7 days. Cell divisions were visually scored for 96 hours. At day 9, single wells were stained with PI (1 μg/ml) and counted using FACSCanto II High Throughput Sampler. In addition, an independent plate was harvested and pooled for RNA extraction. At day 14, cells were harvested, pooled, and stained with BM surface markers as previously described.

Real-time PCR

RNA was isolated from cultured HSCs with the RNeasy Plus Mini Kit (Qiagen), and cDNA was generated from 1 μg of RNA with the qScript cDNA Synthesis Kit (Quantabio), both following the manufacturer's instructions. Real-time PCR reactions were performed using SYBR Green I PCR Master Mix (Roche) and the Roche LightCycler 480.

Population RNA-seq

Two RNA replicates were obtained from 5000 to 10,000 HSCs ($Lineage^-Sca\text{-}1^+c\text{-}Kit^+CD150^+CD48^-$) sorted from flow cytometry as described above. RNA extraction was performed using the RNeasy Micro Kit (Qiagen). Library preparation was performed with

3 to 5 ng of total RNA using the SMARTer Ultra Low RNA Kit (Clontech) and sequenced [50–base pair (bp) single end] on an Illumina HiSeq 3000. The RNA-seq samples were mapped against the mm9 mouse genome assembly using TopHat (*47*) with the option –g 1 to discard those reads that could not be uniquely mapped in just one region. Cufflinks (*48*) was run to quantify the expression in RPKMs of each annotated transcript in RefSeq (*49*). We applied the following thresholds to find differentially expressed genes between each pair of *Phf19*-Flox versus *Phf19*-KO samples [RPKM > 0.5 and KO/Flox FC ≥ 1.5 or Flox/KO FC ≥ 1.3]. We intersected the up- or down-regulated genes in common between duplicates to end up with the final sets of up-regulated (901) and down-regulated (556) genes, respectively. Reports of functional enrichments of Gene Ontology and other genomic libraries were generated using the Enrichr tool (*44*). For specific GSEA, we considered as expressed genes those with RPKM > 0.1. Next, we rank-ordered them according to fold change and analyzed using the preranked tool of GSEA (*50*).

Chromatin immunoprecipitation sequencing

Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$ cells (10,000) were sorted by flow cytometry as described above and freshly cross-linked with 1% formaldehyde. Low Cell ChIP-seq kit (Active Motif) was used following the manufacturer's instructions for histone H3 lysine-27 trimethylated ChIP grade antibody (39155, Active Motif). For ChIP experiments with spike-in control, a ratio of 1:2 (*Drosophila*:mouse) S2 cells was added prior cell lysis. Total chromatin was sheared with a Covaris sonication system for 20 min at intensity 8 and 20% duty cycle. ChIP-seq libraries were prepared with total recovered chromatin, using the Next Gen DNA Library Kit and the Next Gen Indexing Kit (Active Motif) as per the manufacturer's instructions. Libraries were sequenced (50-bp single end) on a HiSeq 3000 platform (Illumina). ChIP-seq samples were mapped against the mm9 mouse genome assembly using Bowtie with the option –m 1 to discard those reads that could not be uniquely mapped to just one region (*51*). ChIP-seq samples normalized by spike-in were mapped against a synthetic genome constituted by the mouse and the fruit fly chromosomes (mm9 and dm3) using Bowtie with the option –m 1 to discard reads that did not map uniquely to one region. MACS was run with the default parameters but with the shift size adjusted to 100 bp to perform the peak calling against the corresponding control sample (*52*). Each set of target genes was retrieved by matching the ChIP-seq peaks in the region 2.5 kbp upstream of the transcription start site (TSS) until the end of the transcripts as annotated in RefSeq annotations (*49*). The heat maps displaying the density of ChIP-seq reads around the TSS of each target gene set were generated by counting the number of reads in this region for each individual gene and normalizing this value with the total number of mapped reads of the sample. Genes on each ChIP heat map were ranked by the logarithm of the average number of reads in the same genomic region. Box plots showing the ChIP-seq level distribution for H3K27me3 for particular sets of genes were calculated by determining the average value at transcript level on the region ±2 kbp around each TSS. The values of the samples including spike-in were corrected by the number of fly reads mapped of the sequencing experiment. The University of California Santa Cruz (UCSC) genome browser was used to generate the screenshots depicted in the manuscript (*53*). To confirm the general increase of H3K27me3 ChIP-seq levels from Phf19-Flox to Phf19-KO conditions, we segmented the mouse genome into bins of 100 bp. Next, we counted within each bin the number of normalized reads of H3K27me3 corrected by spike-in at each sample. To discard regions of the genome virtually with no signal, we focused on the bins in which, at least in one of the conditions, a minimum amount of ChIP signal was detected (threshold: 5).

Omni-ATAC-seq

Omni-ATAC-seq libraries were generated as previously described (*54*). Briefly, 10,000 Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$ cells per condition were washed in cold PBS and resuspended in 50 μl of cold lysis buffer 1 [10 mM tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% (v/v) Igepal CA-630, 0.01% digitonin, and 0.1% Tween 20]. Samples were incubated on ice for 3 min and washed out with 1 ml of cold lysis buffer 2 [10 mM tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.1% Tween 20]. Samples were centrifuged for 10 min at 4°C, and the nuclei pellet was resuspended in the transposition reaction mix [25 μl of 2× transposition reaction buffer from Nextera kit, 2.5 μl of Nextera Tn5 transposase from Nextera kit, 16.5 μl of PBS, 0.5 μl of digitonin (1%), 0.5 μl of Tween 20 (10%), and 5 μl of nuclease-free water] and incubated at 37°C for 1 hour. Samples were purified using the Qiagen MinElute PCR Purification Kit. Transposed DNA was eluted in 10 μl of elution buffer and subjected to 5 cycles of PCR before amplification using barcoded primers and NEBNext High-Fidelity PCR Master Mix. We determined the optimal total number of PCR cycles for each Omni-ATAC-seq library using qPCR. Omni-ATAC-seq libraries were purified using 1.8× volumes of AMPure XP beads to remove fragments below 100 bp and sequenced (50-bp single end) on an Illumina HiSeq 3000. The ATAC-seq samples were mapped against the mm9 mouse genome assembly using Bowtie with the option –m 1 to discard those reads that could not be uniquely mapped to just one region (*51*). Mitochondrial reads were removed from each resulting map, and down-sampling was applied to obtain the same number of mapped fragments per sample. Box plots showing the ATAC-seq level distribution particular sets of genes were calculated by determining the average value at transcript level on the region ±2 kbp around each TSS. The UCSC genome browser was used to generate the screenshots depicted in the manuscript (*53*).

Single-cell RNA-seq

Library preparation (MARSseq)

Single-cell libraries from polyadenylate-tailed RNA were constructed, applying massively parallel single-cell RNA-seq (MARSseq) (*55*, *56*). Briefly, single HSC (Lineage$^-$Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$) cells were FACS-sorted into 384-well plates, containing lysis buffer [0.2% Triton (Sigma-Aldrich) and RNase inhibitor (Invitrogen)] and reverse transcription (RT) primers. The RT primers contained the single-cell barcodes and unique molecular identifiers (UMIs) for subsequent demultiplexing and correction for amplification biases, respectively. Single-cell lysates were denatured and immediately placed on ice. The RT reaction mix containing SuperScript III reverse transcriptase (Invitrogen) was added to each sample. In the RT reaction, spike-in artificial transcripts (External RNA Controls Consortium, Ambion) were included at a dilution of $1:16 \times 10^6$ per cell. After RT, the cDNA was pooled using an automated pipeline (epMotion, Eppendorf). Unbound primers were eliminated by incubating the cDNA with exonuclease I [New England Biolabs (NEB)]. A second pooling was performed through cleanup with solid-phase reversible immobilization (SPRI) magnetic beads (Beckman Coulter). Subsequently, pooled cDNAs were converted into double-stranded DNA

with the Second Strand Synthesis enzyme (NEB), followed by cleanup and linear amplification by T7 in vitro transcription overnight. Afterward, the DNA template was removed by Turbo deoxyribonuclease I (Ambion), and the RNA was purified with SPRI beads. Amplified RNA was chemically fragmented with Zn^{2+} (Ambion) and then purified with SPRI beads. The fragmented RNA was ligated with ligation primers containing a pool barcode and partial Illumina Read1 sequencing adapter using T4 RNA ligase I (NEB). Ligated products were reversed transcribed using the AffinityScript RT enzyme (Agilent Technologies) and a primer complementary to the ligated adapter, partial Read1. The cDNA was purified with SPRI beads. Libraries were completed through a PCR step using the KAPA HiFi HotStart ReadyMix (Kapa Biosystems) and a forward primer that contains Illumina P5-Read1 sequence and a reverse primer containing the P7-Read2 sequence. The final library was purified with SPRI beads to remove excess primers. Library concentration and molecular size were determined with High Sensitivity DNA Chip (Agilent Technologies). The libraries consist of 192 single-cell pools. Multiplexed pools (*2*) were run in one Illumina HiSeq 2500 Rapid two-lane flow cell following the manufacturer's protocol. Primary data analysis was carried out with the standard Illumina pipeline.

Data preprocessing

Sequencing was carried out as paired-end reads, wherein the first read contains the transcript sequence and the second read contains the cell barcode and UMIs. Data analyses were performed as described previously (*56*). Briefly, quality check of the generated reads was performed with the FastQC quality control suite. Samples that reached the quality standards were then processed to deconvolute the reads to single-cell level by demultiplexing according to the cell and pool barcodes. Reads were filtered to remove polyT sequences. Sequencing reads were mapped to the mouse reference genome (Gencode release M11, assembly GRCm38) with the RNA pipeline of the GEMTools 1.7.0 suite using default parameters (6% of mismatches, minimum of 80% matched bases, and minimum quality threshold of 26). Cells with less than 60% of reads mapping to the reference genome or more than 2×10^6 total reads were discarded. Gene quantification was performed using UMI-corrected transcript information to correct for amplification biases, collapsing read counts for reads mapping on a gene with the same UMI (allowing an edit distance up to two nucleotides in UMI comparisons). Only unambiguously mapped reads were considered. Thresholds were set not only to reduce technical noise but also to conserve the sensitivity to identify low-frequency outlier cell populations and to capture differences between fresh and cryopreserved cells.

Skeletal staining

Skeletal whole mounts were performed as previously described (*57*). Briefly, the completely eviscerated and skinned animals were fixed in 96% ethanol for 5 days and transferred to acetone for 2 days. Staining was performed in 0.005% Alizarin Red S, 0.015% Alcian blue 8GS in 5% acetic acid, 5% H_2O, and 90% ethanol for 5 days at room temperature. Samples were washed in H_2O and cleared for 2 days in 1% KOH, followed by clearing steps in 0.8% KOH and 20% glycerol, in 0.5% KOH and 50% glycerol, and in 0.2% KOH and 80% glycerol for 1 week each. Cleared skeletons were stored in 100% glycerol.

Spleen histopathology

Paraffin-embedded tissue sections (3 μm thick) were air dried and further dried at 60°C overnight and then dewaxed. Hematoxylin/eosin standard protocols were done using a CoverStainer (Dako-Agilent) following the manufacturer's procedures. Immunohistochemistry was performed using an Autostainer Plus (Dako-Agilent). Before immunohistochemistry, sections were dewaxed for Ki67 as part of the antigen retrieval process using the low-pH EnVision FLEX Target Retrieval Solutions (Dako, Burlington). For CD68, samples were dewaxed and rehydrated, and thereafter, antigen retrieval was performed using an autoclave for 20 min at 121°C with citrate buffer (pH 6). Samples were blocked with Peroxidase-Blocking Solution (S2023, Dako-Agilent), 5% normal goat serum, and 2.5% bovine serum albumin. Primary antibodies rabbit polyclonal anti-Ki67 (Abcam) and rabbit polyclonal anti-CD68 (orb47985, Biorbyt) were diluted at 1:2000 and 1:4000, respectively, with EnVision FLEX Antibody Diluent (Dako, Agilent) and incubated for 60 min. The secondary antibody used was a BrightVision poly–horseradish peroxidase anti-rabbit immunoglobulin G (IgG) biotin-free, ready to use (Immunologic, DPVR-110HRP). Antigen-antibody complexes were visualized with 3-3′-diaminobenzidine (Dako). Sections were counterstained with hematoxylin (Dako, S202084) and mounted with Toluene-Free Mounting Medium (CS705, Dako) using the Dako CoverStainer. Specificity of staining was confirmed with a rabbit IgG polyclonal isotype control (Abcam). Bright-field images were acquired with a NanoZoomer-2.0 HT C9600 digital scanner (Hamamatsu) equipped with a 20× objective. All images were visualized with a gamma correction set to 1.8 in the image control panel of the NDP.view2 U123888-01 software (Hamamatsu Photonics, France).

Quantification and statistical analysis

Animal data are presented as box plots produced using Prism 6 software (GraphPad, San Diego, CA). Quantified data are presented as means ± SEM. Statistical tests between groups are detailed in the figure legends. Student's paired or unpaired *t* test, depending on the experimental design, was broadly used in all figures. In addition, Wilcoxon rank-sum test for single-cell data (Fig. 3, F and G, and fig. S3I) and log-rank (Mantel-Cox) test (Fig. 4B) were used. Significance is depicted throughout the figures as $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$.

Data and software availability

Raw data and processed information of the ChIP-seq, RNA-seq, and ATAC-seq experiments generated here were deposited in the National Center for Biotechnology Information Gene Expression Omnibus (*58*) repository under the accession number GSE135283.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/32/eabb2745/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. S. Haas, A. Trumpp, M. D. Milsom, Causes and consequences of hematopoietic stem cell heterogeneity. *Cell Stem Cell* **22**, 627–638 (2018).
2. E. Laurenti, B. Gottgens, From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418–426 (2018).
3. P. Ntziachristos, O. Abdel-Wahab, I. Aifantis, Emerging concepts of epigenetic dysregulation in hematological malignancies. *Nat. Immunol.* **17**, 1016–1024 (2016).
4. C. Plass, S. M. Pfister, A. M. Lindroth, O. Bogatyrova, R. Claus, P. Lichter, Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat. Rev. Genet.* **14**, 765–780 (2013).
5. L. Di Croce, K. Helin, Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147–1155 (2013).

6. B. Schuettengruber, H. M. Bourbon, L. Di Croce, G. Cavalli, Genome regulation by Polycomb and Trithorax: 70 years and counting. *Cell* **171**, 34–57 (2017).
7. P. Vizan, M. Beringer, C. Ballare, L. Di Croce, Role of PRC2-associated factors in stem cells and disease. *FEBS J.* **282**, 1723–1735 (2015).
8. A. Laugesen, J. W. Hojfeldt, K. Helin, Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. *Mol. Cell* **74**, 8–18 (2019).
9. V. Di Carlo, I. Mocavini, L. Di Croce, Polycomb complexes in normal and malignant hematopoiesis. *J. Cell Biol.* **218**, 55–69 (2019).
10. H. Xie, J. Xu, J. H. Hsu, M. Nguyen, Y. Fujiwara, C. Peng, S. H. Orkin, Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. *Cell Stem Cell* **14**, 68–80 (2014).
11. S. C. W. Lee, S. Miller, C. Hyland, M. Kauppi, M. Lebois, L. D. Rago, D. Metcalf, S. A. Kinkel, E. C. Josefsson, M. E. Blewitt, I. J. Majewski, W. S. Alexander, Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. *Blood* **126**, 167–175 (2015).
12. W. Yu, F. Zhang, S. Wang, Y. Fu, J. Chen, X. Liang, H. Le, W. T. Pu, B. Zhang, Depletion of polycomb repressive complex 2 core component EED impairs fetal hematopoiesis. *Cell Death Dis.* **8**, e2744 (2017).
13. C. Ballaré, M. Lange, A. Lapinaite, G. M. Martin, L. Morey, G. Pascual, R. Liefke, B. Simon, Y. Shi, O. Gozani, T. Carlomagno, S. A. Benitah, L. D. Croce, Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat. Struct. Mol. Biol.* **19**, 1257–1265 (2012).
14. J. Hunkapiller, Y. Shen, A. Diaz, G. Cagney, D. M. Cleary, M. Ramalho-Santos, N. Krogan, B. Ren, J. S. Song, J. F. Reiter, Polycomb-like 3 promotes polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal. *PLOS Genet.* **8**, e1002576 (2012).
15. G. L. Brien, G. Gambero, D. J. O'Connell, E. Jerman, S. A. Turner, C. M. Egan, E. J. Dunne, M. C. Jurgens, K. Wynne, L. Piao, A. J. Lohan, N. Ferguson, X. Shi, K. M. Sinha, B. J. Loftus, G. Cagney, A. P. Bracken, Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat. Struct. Mol. Biol.* **19**, 1273–1281 (2012).
16. M. D. Doynova, J. F. Markworth, D. Cameron-Smith, M. H. Vickers, J. M. O'Sullivan, Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro. *Skelet. Muscle* **7**, 5 (2017).
17. A. K. San Roman, A. Tovaglieri, D. T. Breault, R. A. Shivdasani, Distinct processes and transcriptional targets underlie CDX2 requirements in intestinal stem cells and differentiated villus cells. *Stem Cell Rep.* **5**, 673–681 (2015).
18. F. O. Bagger, S. Kinalis, N. Rapin, BloodSpot: A database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res.* **47**, D881–D885 (2019).
19. D. Lara-Astiaso, A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman, I. Amit, Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).
20. J. Yang, Y. Tanaka, M. Seay, Z. Li, J. Jin, L. X. Garmire, X. Zhu, A. Taylor, W. Li, G. Euskirchen, S. Halene, Y. Kluger, M. P. Snyder, I.-H. Park, X. Pan, S. M. Weissman, Single cell transcriptomics reveals unanticipated features of early hematopoietic precursors. *Nucleic Acids Res.* **45**, 1281–1296 (2017).
21. S. Wang, F. He, W. Xiong, S. Gu, H. Liu, T. Zhang, X. Yu, Y. Chen, S. Wang, F. He, W. Xiong, S. Gu, H. Liu, T. Zhang, X. Yu, Y. Chen, Polycomblike-2-deficient mice exhibit normal left-right asymmetry. *Dev. Dyn.* **236**, 853–861 (2007).
22. I. Beerman, J. Seita, M. A. Inlay, I. L. Weissman, D. J. Rossi, Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37–50 (2014).
23. N. K. Wilson, D. G. Kent, F. Buettner, M. Shehata, I. C. Macaulay, F. J. Calero-Nieto, M. S. Castillo, C. A. Oedekoven, E. Diamanti, R. Schulte, C. P. Ponting, T. Voet, C. Caldas, J. Stingl, A. R. Green, F. J. Theis, B. Göttgens, Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* **16**, 712–724 (2015).
24. N. Cabezas-Wallscheid, F. Buettner, P. Sommerkamp, D. Klimmeck, L. Ladel, F. B. Thalheimer, D. Pastor-Flores, L. P. Roma, S. Renders, P. Zeisberger, A. Przybylla, K. Schönberger, R. Scognamiglio, S. Altamura, C. M. Florian, M. Fawaz, D. Vonficht, M. Tesio, P. Collier, D. Pavlinic, H. Geiger, T. Schroeder, V. Benes, T. P. Dick, M. A. Rieger, O. Stegle, A. Trumpp, Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell* **169**, 807–823.e19 (2017).
25. R. Gazit, P. K. Mandal, W. Ebina, A. Ben-Zvi, C. Nombela-Arrieta, L. E. Silberstein, D. J. Rossi, Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *J. Exp. Med.* **211**, 1315–1331 (2014).
26. S. M. Chambers, N. C. Boles, K.-Y. K. Lin, M. P. Tierney, T. V. Bowman, S. B. Bradfute, A. J. Chen, A. A. Merchant, O. Sirin, D. C. Weksberg, M. G. Merchant, C. J. Fisk, C. A. Shaw, M. A. Goodell, Hematopoietic fingerprints: An expression database of stem cells and their progeny. *Cell Stem Cell* **1**, 578–591 (2007).
27. A. Wilson, M. J. Murphy, T. Oskarsson, K. Kaloulis, M. D. Bettess, G. M. Oser, A.-C. Pasche, C. Knabenhans, H. R. MacDonald, A. Trumpp, c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18**, 2747–2763 (2004).
28. J. E. Balmer, R. Blomhoff, Gene expression regulation by retinoic acid. *J. Lipid Res.* **43**, 1773–1808 (2002).
29. F. K. B. Lauridsen, T. L. Jensen, N. Rapin, D. Aslan, A. S. Wilhelmson, S. Pundhir, M. Rehn, F. Paul, A. Giladi, M. S. Hasemann, P. Serup, I. Amit, B. T. Porse, Differences in cell cycle status underlie transcriptional heterogeneity in the HSC compartment. *Cell Rep.* **24**, 766–780 (2018).
30. J. M. Bernitz, K. Rapp, M. G. Daniel, D. Shcherbinin, Y. Yuan, A. Gomes, A. Waghray, R. Brosh, A. Lachmann, A. Ma'ayan, D. Papatsenko, K. A. Moore, Memory of divisional history directs the continuous process of primitive hematopoietic lineage commitment. *Stem Cell Rep.* **14**, 561–574 (2020).
31. D. Sun, M. Luo, M. Jeong, B. Rodriguez, Z. Xia, R. Hannah, H. Wang, T. Le, K. F. Faull, R. Chen, H. Gu, C. Bock, A. Meissner, B. Göttgens, G. J. Darlington, W. Li, M. A. Goodell, Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* **14**, 673–688 (2014).
32. H. Celik, W. K. Koh, A. C. Kramer, E. L. Ostrander, C. Mallaney, D. A. C. Fisher, J. Xiang, W. C. Wilson, A. Martens, A. Kothari, G. Fishberger, E. Tycksen, D. Karpova, E. J. Duncavage, Y. Lee, S. T. Oh, G. A. Challen, JARID2 functions as a tumor suppressor in myeloid neoplasms by repressing self-renewal in hematopoietic progenitor cells. *Cancer Cell* **34**, 741–756.e8 (2018).
33. F. Rosenbauer, D. G. Tenen, Transcription factors in myeloid development: Balancing differentiation with transformation. *Nat. Rev. Immunol.* **7**, 105–117 (2007).
34. C. Pridans, M. L. Holmes, M. Polli, J. M. Wettenhall, A. Dakic, L. M. Corcoran, G. K. Smyth, S. L. Nutt, Identification of Pax5 target genes in early B cell differentiation. *J. Immunol.* **180**, 1719–1728 (2008).
35. C. Lien, C.-M. Fang, D. Huso, F. Livak, R. Lu, P. M. Pitha, Critical role of IRF-5 in regulation of B-cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4664–4668 (2010).
36. H. Y. Kueh, E. V. Rothenberg, Regulatory gene network circuits underlying T cell development from multipotent progenitors. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **4**, 79–102 (2012).
37. A. Bigas, L. Espinosa, Hematopoietic stem cells: to be or Notch to be. *Blood* **119**, 3226–3235 (2012).
38. E. Dzierzak, S. Philipsen, Erythropoiesis: Development and differentiation. *Cold Spring Harb. Perspect. Med.* **3**, a011601 (2013).
39. C. Pabst, A. Bergeron, V.-P. Lavallée, J. Yeh, P. Gendron, G. L. Norddahl, J. Krosl, I. Boivin, E. Deneault, J. Simard, S. Imren, G. Boucher, K. Eppert, T. Herold, S. K. Bohlander, K. Humphries, S. Lemieux, J. Hébert, G. Sauvageau, F. Barabé, GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. *Blood* **127**, 2018–2027 (2016).
40. S. W. K. Ng, A. Mitchell, J. A. Kennedy, W. C. Chen, J. M. Leod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A. D. Schimmer, A. C. Schuh, K. W. Yee, L. Bullinger, T. Herold, D. Görlich, T. Büchner, W. Hiddemann, W. E. Berdel, B. Wörmann, M. Cheok, C. Preudhomme, H. Dombret, K. Metzeler, C. Buske, B. Löwenberg, P. J. M. Valk, P. W. Zandstra, M. D. Minden, J. E. Dick, J. C. Y. Wang, A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437 (2016).
41. J. L. M. Rothberg, H. B. Maganti, H. Jrade, C. J. Porter, G. A. Palidwor, C. Cafariello, H. L. Battaion, S. T. Khan, T. J. Perkins, R. F. Paulson, C. Y. Ito, W. L. Stanford, Mtf2-PRC2 control of canonical Wnt signaling is required for definitive erythropoiesis. *Cell Discov.* **4**, 21 (2018).
42. M. Beringer, P. Pisano, V. D. Carlo, E. Blanco, P. Chammas, P. Vizán, A. Gutiérrez, S. Aranda, B. Payer, M. Wierer, L. D. Croce, EPOP functionally links Elongin and Polycomb in pluripotent stem cells. *Mol. Cell* **64**, 645–658 (2016).
43. J. Qiu, D. Papatsenko, X. Niu, C. Schaniel, K. Moore, Divisional history and hematopoietic stem cell function during homeostasis. *Stem Cell Rep.* **2**, 473–490 (2014).
44. M. N. F. Morcos, T. Zerjatke, I. Glauche, C. M. Munz, Y. Ge, A. Petzold, S. Reinhardt, A. Dahl, N. S. Anstee, R. Bogeska, M. D. Milsom, P. Säwén, H. Wan, D. Bryder, A. Roers, A. Gerbaulet, Continuous mitotic activity of primitive hematopoietic stem cells in adult mice. *J. Exp. Med.* **217**, e20191284 (2020).
45. H. Geiger, G. de Haan, M. C. Florian, The ageing haematopoietic stem cell compartment. *Nat. Rev. Immunol.* **13**, 376–389 (2013).
46. S. Hayashi, T. Tenzen, A. P. McMahon, Maternal inheritance of Cre activity in a Sox2Cre deleter strain. *Genesis* **37**, 51–53 (2003).
47. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
48. C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
49. N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. M. Veigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Niell, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully,

A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. D. Cuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
50. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
51. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
52. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, D. Haussler, The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
54. M. R. Corces, A. E. Trevino, E. G. Hamilton, P. G. Greenside, N. A. Sinnott-Armstrong, S. Vesuna, A. T. Satpathy, A. J. Rubin, K. S. Montine, B. Wu, A. Kathiria, S. W. Cho, M. R. Mumbach, A. C. Carter, M. Kasowski, L. A. Orloff, V. I. Risca, A. Kundaje, P. A. Khavari, T. J. Montine, W. J. Greenleaf, H. Y. Chang, An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
55. D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul, N. Elefant, F. Paul, I. Zaretsky, A. Mildner, N. Cohen, S. Jung, A. Tanay, I. Amit, Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).
56. A. Guillaumet-Adkins, G. Rodríguez-Esteban, E. Mereu, M. Mendez-Lago, D. A. Jaitin, A. Villanueva, A. Vidal, A. Martinez-Marti, E. Felip, A. Vivancos, H. Keren-Shaul, S. Heath, M. Gut, I. Amit, I. Gut, H. Heyn, Single-cell transcriptome conservation in cryopreserved cells and tissues. *Genome Biol.* **18**, 45 (2017).
57. N. M. van der Lugt, J. Domen, K. Linders, M. van Roon, E. Robanus-Maandag, H. te Riele, M. van der Valk, J. Deschamps, M. Sofroniew, M. van Lohuizen, Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev.* **8**, 757–769 (1994).
58. T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: Archive for functional genomics data sets—Update. *Nucleic Acids Res.* **41**, D991–D995 (2013).

Acknowledgments: We thank all the members of the Di Croce laboratory for helpful discussions and V. A. Raker for scientific editing. **Funding:** The work in the Di Croce laboratory is supported by grants from the Spanish of Economy, Industry, and Competitiveness (MEIC) (BFU2016-75008-P), "Fundación Vencer El Cancer" (VEC), the European Regional Development Fund (FEDER), Fundació "La Marató de TV3," and from AGAUR. The laboratory of A.B. is supported by SAF2016-75613-R from the Ministerio de Ciencia, Innovación y Universidades. H.H. is a Miguel Servet (CP14/00229) researcher funded by the Spanish Institute of Health Carlos III (ISCIII) and received funding from the European Union's Horizon 2020 research and innovation programme (MSCA-ITN-2015-675752) and the Ministerio de Ciencia, Innovación y Universidades (SAF2017-89109-P; AEI/FEDER, UE). P.V. was supported by "Fundación Científica de la Asociación Española Contra el Cáncer." I.E. was supported by La Caixa INPhINIT program. M.G.-M. was supported by the FPI Program, and R.T. was supported by Generalitat de Catalunya (Beatriu de Pinós program; 2016 BP 00021). We acknowledge support from the Spanish MEIC through the Instituto de Salud Carlos III and the EMBL partnership, Centro de Excelencia Severo Ochoa, CERCA Programme/Generalitat de Catalunya, the Generalitat de Catalunya through Departament de Salut and Departament d'Empresa i Coneixement, and the Co-financing by the Spanish Ministry of Economy, Industry and Competitiveness (MEIC) with funds from the European Regional Development Fund (ERDF) corresponding to the 2014-2020 Smart Growth Operating Program. **Author contributions:** P.V. and A.G. conceived and planned this project and collected, analyzed, and interpreted data. I.E. and M.G.-M. collected, analyzed, and interpreted data. M.L. and L.D.C. conceived and planned this project. A.C. performed skeletal staining and interpreted data. A.L. and H.H. performed and analyzed single-cell transcriptomic data. L.d.A.-A. and T.G. conceived and planned transplant experiments and provided insightful advice during the execution of the project. E.B. performed the bioinformatic analyses. R.T. and A.B. performed and analyzed fetal liver analyses and provided insightful advice during the execution of the project. P.V. and L.D.C. wrote the manuscript, with the assistance and final approval of all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 12 February 2020
Accepted 19 June 2020
Published 5 August 2020
10.1126/sciadv.abb2745

Citation: P. Vizán, A. Gutiérrez, I. Espejo, M. García-Montolio, M. Lange, A. Carretero, A. Lafzi, L. de Andrés-Aguayo, E. Blanco, R. Thambyrajah, T. Graf, H. Heyn, A. Bigas, L. Di Croce, The Polycomb-associated factor PHF19 controls hematopoietic stem cell state and differentiation. *Sci. Adv.* **6**, eabb2745 (2020).