Regulation of haematopoietic multipotency by EZH1

Linda T. Vo[1,2,3], **Melissa A. Kinney**[1,2], **Xin Liu**[4], **Yuannyu Zhang**[4,5], **Jessica Barragan**[1,2], **Patricia M. Sousa**[1,2], **Deepak K. Jha**[1,2], **Areum Han**[1,2], **Marcella Cesana**[1,2], **Zhen Shao**[5], **Trista E. North**[6], **Stuart H. Orkin**[2,3,7], **Sergei Doulatov**[8], **Jian Xu**[4], and **George Q. Daley**[1,2,3,*]

[1]Stem Cell Program, Boston Children's Hospital, Boston, MA

[2]Division of Hematology/Oncology, Boston Children's Hospital and Dana Farber Cancer Institute, Boston, MA

[3]Harvard Medical School, Boston, MA

[4]Children's Medical Center Research Institute, Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX

[5]Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

[6]Department of Pathology, Beth Israel-Deaconess Medical Center, Boston MA

[7]Howard Hughes Medical Institute

[8]Division of Hematology, Department of Medicine, University of Washington, Seattle, WA

SUMMARY

All haematopoietic lineages circulating in the blood of adult mammals derive from multipotent haematopoietic stem cells (HSCs)[1]. Haematopoiesis in the mammalian embryo stands in stark contrast, with lineage-restricted progenitors arising first, independently of HSCs, and HSCs emerging only later in gestation[2,3]. As best defined in the mouse, "primitive" progenitors first appear in the yolk sac (YS) at 7.5 days post-coitum (dpc)[2,3]. Subsequently, erythroid-myeloid progenitors (EMPs) that express fetal hemoglobin[4], as well as fetal lymphoid progenitors[5] develop in the YS and the embryo proper, but these cells lack HSC potential. Ultimately, "definitive" HSCs with long-term, multilineage potential and the capacity to engraft irradiated adults emerge at 10.5 dpc from arterial endothelium in the aorta-gonad-mesonephros (AGM) and other haemogenic

*Corresponding author George.Daley@childrens.harvard.edu.

vasculature[3]. The molecular mechanisms for reverse progression of haematopoietic ontogeny remain unexplained. We hypothesized that the definitive haematopoietic program might be actively repressed in early embryogenesis via epigenetic silencing[6], and that alleviating this repression would elicit multipotency in otherwise restricted haematopoietic progenitors. Here, we demonstrate that reduced expression of the Polycomb group protein EZH1 uncovers multi-lymphoid output from human pluripotent stem cells (PSCs) and precocious emergence of functional definitive HSCs at sites of primitive and/or EMP-biased haematopoiesis in vivo, identifying *EZH1* as a repressor of haematopoietic multipotency in the early mammalian embryo.

Differentiation of PSCs to hematopoietic lineages generates robust erythroid-myeloid lineage-restricted progenitors but not HSCs. This pattern bears striking similarities to early hematopoietic ontogeny. We hypothesized that the same epigenetic factors actively repress multipotency in embryogenesis and differentiation from PSCs. To identify these factors, we adopted a loss-of-function screen using lentivirally delivered shRNAs targeting 20 DNA and histone modifying factors (Extended Data 1a, Extended Table 1). Erythro-myeloid progenitors differentiated from human PSCs marked by CD34 and CD45 were expanded with five transcription factors (5F). They retained embryonic features, including lack of lymphoid potential[7], enabling us to screen for reactivation of lymphoid potential as a measure of multipotency. 5F cells were transduced with individual shRNAs and screened for T cell potential on OP9-DL1 stroma (Fig. 1a). Knockdown of 6 factors independently enhanced $CD4^+CD8^+$ T cell potential from 5F cells (Fig. 1b, Extended Data 1b).

Prospective validation revealed that only *EZH1* knockdown (shEZH1) elicited robust T (16.3 ± 7.4%) and B cell (22.5 ± 7.3%) potential (Fig. 1c–e), compared to shRNAs targeting a control luciferase gene (shLUC) (T cell 0.002 ± 0.002%; B cell 0.022 ± 0.006%) across multiple iPSC lines (Fig. 1f). EZH1-deficient cells retained erythro-myeloid potential by colony-forming assays (Fig. 1g) and flow cytometry (Fig. 1h, i). *EZH1* knockdown also promoted lymphoid potential independent of 5F, as evidenced by robust T cell differentiation from naive $CD34^+$ haemogenic endothelial (HE) cells (26.1± 16.5% shEZH1 vs. 2.3 ± 0.4% shLUC) (Extended Data 1c). Further characterization was prohibited due to the limited proliferation of PSC-HE. In contrast 5F cells expanded exponentially (Extended Data 1d) and showed increased $CD34^+$ progenitors with shEZH1 (78.8 ± 14.2% vs 29.3 ± 10.0%) (Extended Data 1e). Taken together, *EZH1* knockdown activates multipotency in restricted embryonic haematopoietic progenitors.

EZH1 is a component of the Polycomb Repressive Complex 2 (PRC2), which mediates epigenetic silencing of genes via methylation of lysine residue 27 of histone H3[8]. To dissect the role of PRC2 in repressing haematopoietic multipotency, we assessed T cell differentiation upon depletion of each PRC2 subunit. In addition to *EZH1, SUZ12* knockdown also enhanced T cell potential, albeit to a lesser extent. By contrast, knockdown of *EED* or *EZH2* had no effect on T cell potential and dual *EZH1* and *EZH2* knockdown phenocopied that of *EZH2* depletion (Fig. 2a, b). To determine if the catalytic SET domain was required, we overexpressed full-length murine *Ezh1* (mEzh1) or *Ezh1* without the SET domain (mEzh1 SET) (Fig. 3c). mEzh1 expression completely abrogated T cell potential in shEZH1 cells, whereas mEzh1 SET did not (Fig. 3c, d, Extended Data 2d–g). Furthermore,

overexpression of murine *Ezh2* failed to suppress T cell potential, despite the remarkable homology of the SET domains (Extended Data 2e, h, i). These data show that specific inhibition of *EZH1*, rather than antagonism of canonical PRC2, unlocks definitive lymphoid potential and the catalytic SET domain is required for this function.

To understand the molecular changes upon *EZH1* knockdown, we performed RNA-, ATAC- and ChIP-seq. Upregulated genes following *EZH1* knockdown were enriched for biological processes such as defense response (p=6.8×10^{-9}), immune response (p=1.2×10^{-7}) and T cell co-stimulation (p=0.03) (Fig. 3a, b). Human haematopoieteic gene signatures of multi-lymphoid progenitors[9], including HSC (stem), MLP (early lymphoid) and ProB, were highly enriched in shEZH1 cells, consistent with stem and lymphoid potential (Fig. 3c). We also performed RNA- and ATAC-seq on emergent HSPCs at 10.5 dpc[10,11,12] from the YS and AGM of wild-type (WT), *Ezh1*[+/−] and *Ezh1*[−/−] murine embryos (Fig. 4a). Interestingly, in WT embryos, expression of *Ezh1* was lower in AGM compared to YS, while *Ezh2* and *Eed* were higher in AGM (Fig. 4b). Significantly, *Ezh1*-deficiency in vivo also induced genes enriched for angiogenesis, haematopoietic/lymphoid development and immune system processes (Extended Data 3a–d).

Regions of increased chromatin accessibility (1610 ATAC peaks) in shEZH1 cells exhibited concomitantly increased gene expression upon *EZH1* knockdown and were associated with T cell development and lymphocyte activation pathways, as well as HSC, HSC/MLP, B and T cell signatures (Fig. 3d–e; Extended Data 3e–g). *EZH1* knockdown also increased accessibility to HSC/lymphoid transcription factors (TFs), such as *HLF*, *FOXO1* and *ARID5B*[13–15] (Fig. 3f). Downregulated peaks were enriched in alternative developmental processes and importantly, embryonic haematopoiesis (Fig. 3e, Extended Data 3e). In vivo, upregulated ATAC peaks in *Ezh1*-deficient AGM cells were enriched for immune response, T cell activation, lymphocyte differentiation pathways, as well as HSC and HSC/MLP signatures (Fig. 4a,c, d, Extended Data 3h, i); further, *EZH1* deficiency increased accessibility to target genes of master haematopoietic TFs, including *Runx1* (Extended Data 3j, k).

We hypothesized that these molecular changes upon *EZH1* knockdown were mediated by bivalent, or poised, chromatin domains, often implicated in control of developmentally-regulated genes[16]. Consistent with previous reports, EZH1 was broadly associated with repressive (H3K27me3), bivalent (H3K27me3 and H3K4me3) and active (H3K4me3) histone marks[17,18] (Fig. 3g, Extended Data 4a). While active genes were associated with housekeeping functions (Extended Fig. 4b), EZH1-bound bivalent and repressed genes were enriched for developmental and morphogenic processes (Extended Data 4c, d). *EZH1* knockdown increased expression of bivalent genes, which were associated with HSC and early lymphoid lineages (Extended Data 4e, f). These genes included targets of HSC TFs such as *RUNX1T1*, *SOX17* and NOTCH factors *HES1*, *HEY1* and *FOXC2*[19] (Fig. 3h). EZH1 directly bound promoters of HSC and ProB TFs including *HLF*, *PRDM16*, *LMO2*, *ETS1*, *MEIS1*, *RUNX1*, and *HOX* clusters (Extended Data 4e). We also observed a global reciprocal relationship between H3K27me3 and gene transcription (Fig. 3i, Extended Data 4g–k), with poised HSC genes exhibiting loss of H3K27me3 and increased expression upon

EZH1 knockdown (Extended Data 4h, i). 27/29 of these activated HSC genes are direct targets of *EZH1* including *HOPX*, *HLF*, *MEIS1* and *HES1* (p= 7.8×10^{-5}; Fig. 3j, k).

EZH2 also bound activated HSC genes, consistent with its ability to target the same regions[8] (Extended Data 4l); however, recent analysis of SET domain-swapping revealed context-specific sensitivity to an EZH2-specific inhibitor, further suggesting that while EZH1 and EZH2 can bind a common subset of HSC targets, these enzymes likely have distinct functions on chromatin[20]. Concordant with our observation that *SUZ12* knockdown partially phenocopies EZH1 loss (Fig. 2a, b), we observed specific enrichment of EZH1 and SUZ12 at activated HSC and ProB genes, consistent with non-canonical targets of EZH1-SUZ12 complex[17] (Extended Data 4m–q). Similarly, upregulated ATAC peaks in *Ezh1*-deficient AGM were also enriched for SUZ12 binding, but not EZH2, indicating a conserved role for non-canonical PRC2 regulation in vivo (Extended Data 4r). This data suggests that in addition to the canonical function of EZH1-PRC2 in mediating H3K27me3 changes at poised HSC loci, *EZH1* also regulates ProB genes through a complementary non-canonical EZH1-SUZ12 complex, highlighting an *EZH1*-specific function that is not phenocopied by *EZH2*.

The emergence of *bona fide* HSCs, defined by the capacity to repopulate irradiated adult recipients, marks the transition from embryonic to definitive haematopoiesis. We isolated AGM and YS from embryonic day (E)10.5 WT, *Ezh1*$^{+/-}$ and *Ezh1*$^{-/-}$ embryos and transplanted adult NOD/SCID-IL2Rγ^{null} (NSG) recipients (Fig. 4a). We detected peripheral blood reconstitution from WT AGM in 3/7 mice ($11.9 \pm 7.9\%$) at 4 weeks, but chimerism decreased by 16 weeks (2/7, $12.2 \pm 8.1\%$); this corresponds to 1 repopulating unit in ~10.4 ee, consistent with HSCs being exceedingly rare at E10.5[10,21]. By contrast, 5/8 mice transplanted with *Ezh1*$^{-/-}$ AGM cells were engrafted at 4 weeks ($39.2 \pm 9.4\%$) and stabilized at 16 weeks ($34.6 \pm 14.6\%$). Notably, *Ezh1*$^{+/-}$ AGM transplant recipients had the highest initial chimerism ($41.2 \pm 16.3\%$; 4/5), which increased by 16 weeks ($68.9 \pm 17.8\%$), and was predominantly multilineage (3/5) (Fig. 4e; Extended Data 5a, c). This corresponds to 1 repopulating unit in 3.6 *Ezh1*$^{-/-}$ and 2.2 *Ezh1*$^{+/-}$ ee, or a ~5-fold increase in HSC frequency over WT.

At E10.5, the YS is thought to contain few, if any, HSCs[21]. We detected low-level engraftment of WT YS cells in 5/9 recipients at 4 weeks ($3.4 \pm 0.7\%$), and in 3/9 mice at 16 weeks ($4.3 \pm 1.6\%$). Most *Ezh1*$^{-/-}$ ($4.5 \pm 0.9\%$, 6/7 engrafted) and all of *Ezh1*$^{+/-}$ YS-transplanted mice ($5.4 \pm 1.4\%$, 5/5 engrafted) showed stable long-term engraftment at 16 weeks. The number of repopulating units calculated was similar to that of the AGM (~1 in 12.3 ee WT; 1 in 2.6 *Ezh1*$^{-/-}$, 1 in <2 *Ezh1*$^{+/-}$). All engrafted mice were multilineage (Fig. 4f; Extended Data 5a, c). Importantly, up to 80% of peritoneal B cells in *EZH1*$^{+/-}$ AGM-engrafted mice were of the adult-like B-2 phenotype, as opposed to the embryonic B-1 cells (Extended Data 6a). Moreover, up to 95% of donor-derived CD45.2$^+$CD3$^+$ T cells expressed adult-type TCRβ, as opposed to embryonic TCR$\gamma\delta$, in *Ezh1*$^{-/-}$ and *Ezh1*$^{+/-}$ AGM and YS engrafted mice (Extended Fig. 6b). These data provide compelling evidence that *Ezh1* deficiency, and in particular haploinsufficiency, stimulates generation of definitive HSCs and adult-like lymphopoiesis.

The para-aortic splanchnopleura (PSP) at E9.5 lacks HSCs as determined by transplantation studies[3]. Transplantation of E9.5 WT PSP cells (Fig. 4a) failed to engraft adult recipients (0/5)[22]; in contrast, we detected chimerism in recipients of $Ezh1^{-/-}$ (3/3, 1.6 ± 0.3 %) and $Ezh1^{+/-}$ (4/6 mice, 3.6 ± 1.3%) PSP at 4 weeks (Fig. 4g; Extended Data 5b). By 16 weeks, chimerism increased in Ezh1$^{-/-}$ (3/3, 9.4 ± 5.1%) and $Ezh1^{+/-}$ (5/6, 13.1 ± 9.5%) recipients, and grafts were fully multilineage (Extended data 5c). Thus, $Ezh1$ deficiency stimulates precocious generation of bona fide HSCs during embryogenesis.

To assess self-renewal capacity of $Ezh1$-deficient HSCs, we performed secondary transplantation. No mice showed engraftment with E10.5 WT AGM (0/4) or YS (0/7). By contrast, 4/7 $Ezh1^{-/-}$ (4.4 ± 0.5%) and 9/9 $Ezh1^{+/-}$ (57.8 ± 10.2%) AGM-derived secondary recipients were engrafted (Fig. 4h; Extended Data 5d). Of note, while no $Ezh1^{-/-}$ YS recipients (0/10) were engrafted, we observed secondary chimerism from $Ezh1^{+/-}$ YS cells (5/7, 1.5 ± 0.3%), which increased by 16 weeks (6/7, 5.1 ± 1.9%) (Extended Data 5d, e). All engrafted secondary recipients were multilineage with no evidence of leukaemic transformation (Fig. 4h, Extended Data 5c, e). Taken together, these data indicate that genetic $Ezh1$ deficiency elicits precocious emergence of bona-fide HSCs in vivo.

It has long been a curiosity that haematopoietic ontogeny progresses in reverse order, with haematopoietic progenitors appearing first in embryonic development independently of HSCs[2,3]. We propose that $EZH1$ represses definitive loci in primitive blood progenitors differentiated from human PSCs and in murine embryos, which precludes precocious HSC emergence during gestation. $EZH1$ deficiency promotes multipotency in restricted blood progenitors and enables precocious emergence of HSCs. While PRC2 is a well-characterized HSC regulator, our data contribute compelling evidence for the distinct molecular functions of $EZH1$ and $EZH2$ and suggest a putative role for non-canonical PRC2, involving $EZH1$ and $SUZ12$. Homozygous loss of $Suz12$ in mice impairs HSC function and lymphopoiesis, but heterozygosity for $Suz12$ or Eed enhances HSC self-renewal[23,24]. Consistent with this, our data reinforce the concept that HSCs are exquisitely sensitive to PRC2 dosage, with partial reduction or elevation affecting function[23–26]. Interestingly, $Runx1$ haploinsufficiency also promotes premature HSC generation[27]. Our data unify these observations; EZH1 marks many TF binding sites, while $EZH1$ deficiency enhances accessibility to targets of key HSC TFs including $Runx1$ to promote HSC emergence (Extended Data 3j, k). We identify $Ezh1$ as a molecular regulator of lineage-restricted potential of the first blood progenitors in the mammalian embryo, which accounts in part for why early embryonic progenitors lack multipotency. Beyond developmental implications, our findings suggest that resolution of EZH1-marked domains may be essential for physiological specification of HSCs from PSCs, as a complementary approach to the synthetic reactivation of stem cell programs by HSC transcription factors[7,28].

METHODS

A step-by step protocol can be found at Protocol Exchange.

hPSC culture

All experiments were performed using MSC-iPS1[30], CD34-iPS and CD45-iPS, obtained from the Boston Children's Hospital Human Embryonic Stem Cell Core (hESC) and verified by immunohistochemistry for pluripotency markers, teratoma formation and karyotyping. All cells were routinely tested for mycoplasma contamination. Human iPS cells were maintained on mouse embryonic fibroblast (GlobalStem) feeders in DMEM/F12 + 20% KnockOut-Serum Replacement (Invitrogen), 1 mM L-glutamine, 1 mM NEAA, 0.1 mM β-mercaptoethanol, and 10 ng/mL bFGF. Media was changed daily, and cells were passaged 1:4 onto fresh feeders every 7 days using standard clump passaging with collagenase IV.

EB differentiation

EB differentiation was performed as previously described[31]. Briefly, hPSC colonies were scraped into non-adherent rotating 10 cm plates at the ratio of 2:1. The EB media was KO-DMEM + 20% FBS (Stem Cell Technologies), 1 mM L-glutamine, 1 mM NEAA, penicillin/streptomycin, 0.1 mM β-mercaptoethanol, 200 μg/mL h-transferrin, and 50 μg/mL ascorbic acid. After 24 hrs, media was changed by allowing EBs to settle by gravity, and replaced with EB media supplemented with growth factors: 50 ng/mL BMP4 (R&D Systems), 200 ng/mL SCF, 200 ng/mL FLT3, 50 ng/mL G- CSF, 20 ng/mL IL-6, 10 ng/mL IL-3 (all Peprotech). Media was changed on day 5, and day 10. EBs were dissociated on day 14 by digesting with collagenase B (Roche) for 2 hrs, followed by treatment with enzyme-free dissociation buffer (Gibco), and filtered through an 80 μm filter. Dissociated EBs were frozen in 10% DMSO, 40% FBS freezing solution.

Progenitor sorting

Dissociated EB cells were thawed following the Lonza Poietics protocol and resuspended at 1×10^6 per 100 μL staining buffer (PBS + 2% FBS). CD34+ cells were sorted from bulk EB culture using human CD34 microbeads (Miltenyi Biotec) and run through a magnetic column separator (MACS) as per manufacturer's instructions.

Lentiviral and shRNA library plasmids

5F lentiviral plasmids: HOXA9, ERG, RORA, SOX4, and MYB were cloned into pInducer-21 Dox-inducible lentiviral vector. shRNA library targeting 20 epigenetic modifiers[32] was obtained from the Broad Institute RNAi Consortium in pLKO.1 or pLKO.5 lentiviral vectors. Lentiviral particles were produced by transfecting 293T-17 cells (ATCC) with the lentiviral plasmids and 3rd-generation packaging plasmids. Virus was harvested 24 hours after transfection and concentrated by ultracentrifugation at 23,000 rpm for 3 hours. All viruses were titered by serial dilution on 293T cells.

5F gene transfer and 5F culture

MACS separated CD34+ EB progenitors were seeded on retronectin-coated (10 μg/cm^2) 96 well plates at a density of $2\text{-}5 \times 10^4$ cells per well. The infection media was SFEM (StemCell) with 50 ng/mL SCF, 50 ng/mL FLT3, 50 ng/mL TPO, 50 ng/mL IL6, 10 ng/mL IL3 (all R&D Systems). Lentiviral infections were carried out in a total volume of 150 μL. The multiplicity of infection (MOI) each factor was: ERG MOI = 5, HOXA9 MOI = 5,

RORA MOI = 3, SOX4 MOI = 3, MYB MOI = 3, and MOI = 2 for shRNA. Virus was concentrated onto cells by centrifuging the plate at 2300 rpm for 30 min at RT. Infections were carried out for 24 hours. After gene transfer, 5F cells were cultured in SFEM with 50 ng/mL SCF, 50 ng/mL FLT3, 50 ng/mL TPO, 50 ng/mL IL6, and 10 ng/mL IL3 (all R&D Systems). Dox was added at 2 μg/mL (Sigma). Cultures were maintained at a density of <1 $\times 10^6$ cells/mL, and media were changed every 3-4 days.

T cell differentiation

After 14 days of respecification, 1×10^5 5F were plated in OP9-DL1 stromal co-culture[33]. Cells were cultured in α-MEM (Gibco), 1% penicillin/streptomycin, 20% FBS (Gemini), and 1 mM L-glutamine with 30 ng/mL SCF, 5 ng/mL FLT3, 5 ng/mL IL-7 (all R&D Systems) for 20 days with 2 μg/mL Dox followed by Dox removal. Cells were harvested by mechanical dissociation and filtered through a 40 μM filter and passaged onto fresh stroma every 5-7 days. T cell development was assessed after 35 days using CD45, CD7, CD3, CD4 and CD8.

B cell differentiation

After 14 days of respecification, 5×10^4 5F were plated into a single well of MS-5 stroma in a 6-well NUNC plate. Cells were cultured in Myelocult H5100 (Stem Cell Technologies) supplemented with 1% penicillin/streptomycin 50 ng/mL SCF (R&D), 10 ng/mL FLT3 (R&D), 25 ng/mL IL7 (R&D) and 25 ng/mL TPO (R&D) for 10 days with 2 μg/mL Dox followed by Dox removal.

Colony assays

After 14 days of respecification, 5×10^4 cells were plated into 3 mL of complete methylcellulose H3434 (StemCell Technologies) supplemented with 10 ng/mL IL6 (Peprotech), 10 ng/mL FLT3 (R&D), and 50 ng/mL TPO (R&D) without 2 μg/mL Dox. The mixture was distributed into two 60 mm dishes and maintained in a humidified chamber for 14 days.

Mouse transplantation

NOD/LtSz-scidIL2Rgnull (NSG) (Jackson Labs) mice were bred and housed at the Boston Children's Hospital animal care facility. Animal experiments were performed in accordance with institutional guidelines approved by Boston Children's Hospital Animal Care Committee. At least n=3 animals were used per cohort, based on previous transplantation studies. Mice were assigned randomly to groups and not blinded. Briefly, 8-12 week old mice were irradiated (275 rads) 24 hours before transplant. To ensure consistency between experiments, only female mice were used. Sublethally-irradiated adult NSG females were transplanted intravenously with 3.5 ee of whole E10.5 AGM. Mice were bled retroorbitally every 4 weeks to monitor donor chimerism up to 16 weeks post-transplantation. Twenty-four weeks after primary transplantation, primary recipients from each group were sacrificed and 4×10^6 whole bone marrow cells were transplanted into 1-3 secondary recipients. Cells were transplanted in a 200 μL volume using a 28.5 gauge insulin needle. Sulfatrim was

administered in drinking water to prevent infections after irradiation. Data points were combined from all independent experiments and outliers were not excluded.

Flow cytometry

The following antibodies were used for human cells: CD45 APC-Cy7 (557833, BD Biosciences), CD4 PE-Cy5 (IM2636U, Beckman Coulter Immunotech), CD8 BV421 (RPA-T8, BD Horizon), CD5 BV510 (UCHT2, BD Biosciences), TCRgd APC (555718, BD Biosciences), TCRab BV510 (T10B9.1A-31, BD Biosciences), CD3 PE-Cy7 (UCHT1, BD Pharmigen), CD7 PE (555361, BD Pharmigen), CD1a APC (559775, BD Pharmigen) for T cell staining. For B cell staining: CD45 PE-Cy5 (IM2652U, Beckman Coulter Immunotech), CD19 PE (4G7, BD), CD56 V450 (B159, BD Biosciences), CD11b APC-Cy7 (557754, BD Biosciences), For HSC/Progenitor sorting: CD34 PE-Cy7 (8G12, BD), CD45 (557833, BD Biosciences), CD38 PE-Cy7 (IM2651U, BD), DAPI. For myeloid and erythroid staining: CD11b APC-Cy7 (557754, BD Biosciences), GLYA PE-Cy7 (A71564, Beckman Coulter), CD71 PE (555537, BD Biosciences), CD45 PE-Cy5 (IM2652U, Beckman Coulter Immunotech). All stains were performed with $<1\times10^6$ cells per 100 μL staining buffer (PBS + 2% FBS) with 1:100 dilution of each antibody, 30 min at RT in dark. Compensation was performed by automated compensation with anti-mouse Igk and negative beads (BD). All acquisitions were performed on BD Fortessa or BD Aria cytometer.

The following antibodies were used for mouse cells: CD45.2 PE-Cy7 (104, eBioscience), CD45.1 FITC (A20, eBioscience), B220 PB (RA3-6B2, BD Biosciences), Ter119 PE-Cy5 (Ter 119, eBioscience), GR1 (RB6-8C5, BD Bioscience), CD3 APC (145-2C11, eBioscience), CD19 APC-Cy7 (1D3, BD Bioscience), MAC1 AF700 (M1/70, BD Bioscience) for engraftment analyses. For RNA seq sort: CD16/32 (93, Biolegend), Ter119 Biotin (Ter119, eBioscience), GR1 Biotin (RB6-8C, eBioscience), CD3 Biotin (17A2, eBioscience), CD5 Biotin (53-6.7, eBioscience), CD19 Biotin (eBio1D3, eBioscience), Streptavidin EF450 (eBioscience), CD45 PerCP-Cy5.5 (30-F11, eBioscience), CD144 EF660 (eBioBV13, eBioscience), CD117 APC-EF780 (2B8, eBioscience), CD41 PE-Cy7 (eBioMWReg30, eBioscience). For B cell staining: CD45.2 APC-CY7 (104, BioLegend), CD23 PE-Cy7 (B3B4, eBioscience), Ter119 PE-Cy5 (Ter 119, eBioscience), IgM EF660 (eB121-15F9, eBioscience), MAC1 A700 (M1/70, BD Bioscience), CD5 BV510 (53-7.3 BD Biosciences), IgM (11/41, eBioscience), B220 PE Cy5 (RA3-6B2, BD Biosciences). For T cell staining: CD45.2 APC-CY7 (104, BioLegend), TCRb PE-Cy5 (H57-597, BD Biosciences), CD8 APC-EF780 (53-6.7, eBioscience), CD4 APC (GK1.5, eBioscience), CD3 AF700 (17A2, BioLegend), TCRgd FITC (GL3, BD Biosciences). For HSPC sorting: CD16/32 (93, Biolegend), Ter119 Biotin (Ter119, eBioscience), Gr-1 Biotin (RB6-8C5, eBioscience), CD3 Biotin (17A2, eBioscience), CD5 Biotin (53-7.3, eBioscience), CD8 Biotin (53-6.7, eBioscience), CD19 Biotin (eBio1D3, eBioscience), Streptavidin eFluor450 (eBioscience), CD45 PerCP-Cy5.5 (30-F11, eBioscience), CD144 eFluor660 (eBioBV13, eBioscience), CD117 APC-eFluor 780 (2B8, eBioscience), CD41 PE-Cy7 (eBioMWReg30, eBioscience). All stains were performed with $<1\times10^6$ cells per 100 μL staining buffer (PBS + 2% FBS) with 1:100 dilution of each antibody, 30 min on ice in dark. Compensation was performed by automated compensation with anti-mouse Igk and negative beads (BD). All acquisitions were performed on BD Fortessa or BD Aria cytometer.

RNA-seq

Human cells were stained and sorted using the antibodies CD34 PE-Cy7 (8G12, BD) and CD38 PE-Cy7 (IM2651U, BD). RNA-seq libraries were prepared using the NEB Ultra (PolyA) kit as per manufacturer's protocol with 50 ng input RNA. Mouse cells were stained and sorted using the "HSPC stain" (see above). RNA-seq libraries were prepared using the Clontech SMARTer Universal Low Input kit as per manufacturer's protocol with 10 ng input RNA. Libraries were sequenced using the 200 cycle paired-end kit on the Illumina HiSeq2500 system. RNA-seq reads were analyzed with the Tuxedo Tools following a standard protocol[34]. Reads were mapped with TopHat version 2.1.0 and Bowtie2 version 2.2.4 with default parameters against build hg19 of the human genome, and build hg19 of the RefSeq human genome annotation. Samples were quantified with the Cufflinks package version 2.2.1. Differential expression was performed using Cuffdiff with default parameters.

ATAC-seq

ATAC-seq was performed as previously described[35]. $5 - 50 \times 10^3$ cells were used for each tagmentation using Tn5 transposases. The resulting DNA was isolated, quantified, and sequenced on an Illumina NextSeq500 system. The raw reads were aligned to the human genome assembly hg19 using Bowtie[36] with the default parameters, and only tags that uniquely mapped to the genome were used for further analysis. ATAC-seq peaks were identified using MACS[37].

ChIP-seq

ChIP experiments were performed as previously described[38] using the antibodies for H3K4me3 (04-745, Millipore) and H3K27me3 (07-449, Millipore) in 5F cells. For bioChIP analysis of EZH1 or EZH2 occupancy, FLAG-biotin(FB)-tagged EZH1 or EZH2 was stably expressed in 5F cells. The chromatin was isolated and immunoprecipiated by streptavidin Dynabeads (Life Technologies) as previously described[39]. ChIP-seq libraries were generated using NEBNext ChIP-seq Library Prep Master Mix following the manufacturer's protocol (New England Biolabs), and sequenced on an Illumina NextSeq500 system. ChIP-seq raw reads were aligned to the human genome assembly hg19 using Bowtie[36] with the default parameters, only tags that uniquely mapped to the genome were used for further analysis. ChIP-seq peaks were identified using MACS[37].

Bioinformatics and statistical analysis

All statistical calculations were performed using GraphPad Prism. Tests between two groups used two-tailed unpaired Student's t test. Data are presented as mean ± s.e.m. Where indicated, ANOVA was used, with p-values less than 0.05 considered significant. Gene set enrichment analysis (GSEA) and gene ontology (GO) were run according to default parameters in their native implementations. Statistical enrichment of gene lists was performed using Fisher's exact test.

Data Availability

All RNA-, ATAC-, and ChIP-seq data has been deposited to the Gene Expression Omnibus (GEO) database under the accession number GSE89418.
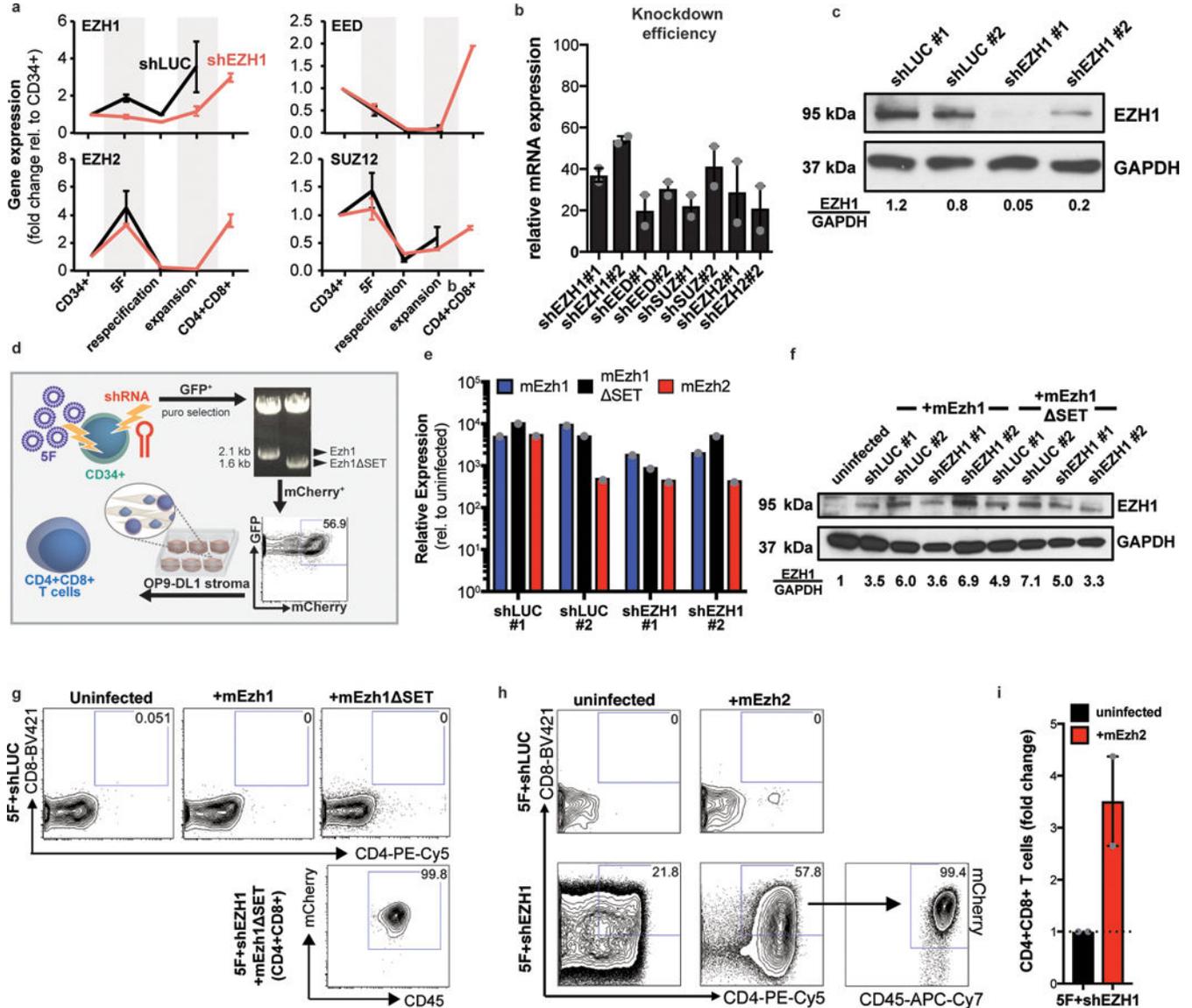
Extended Data
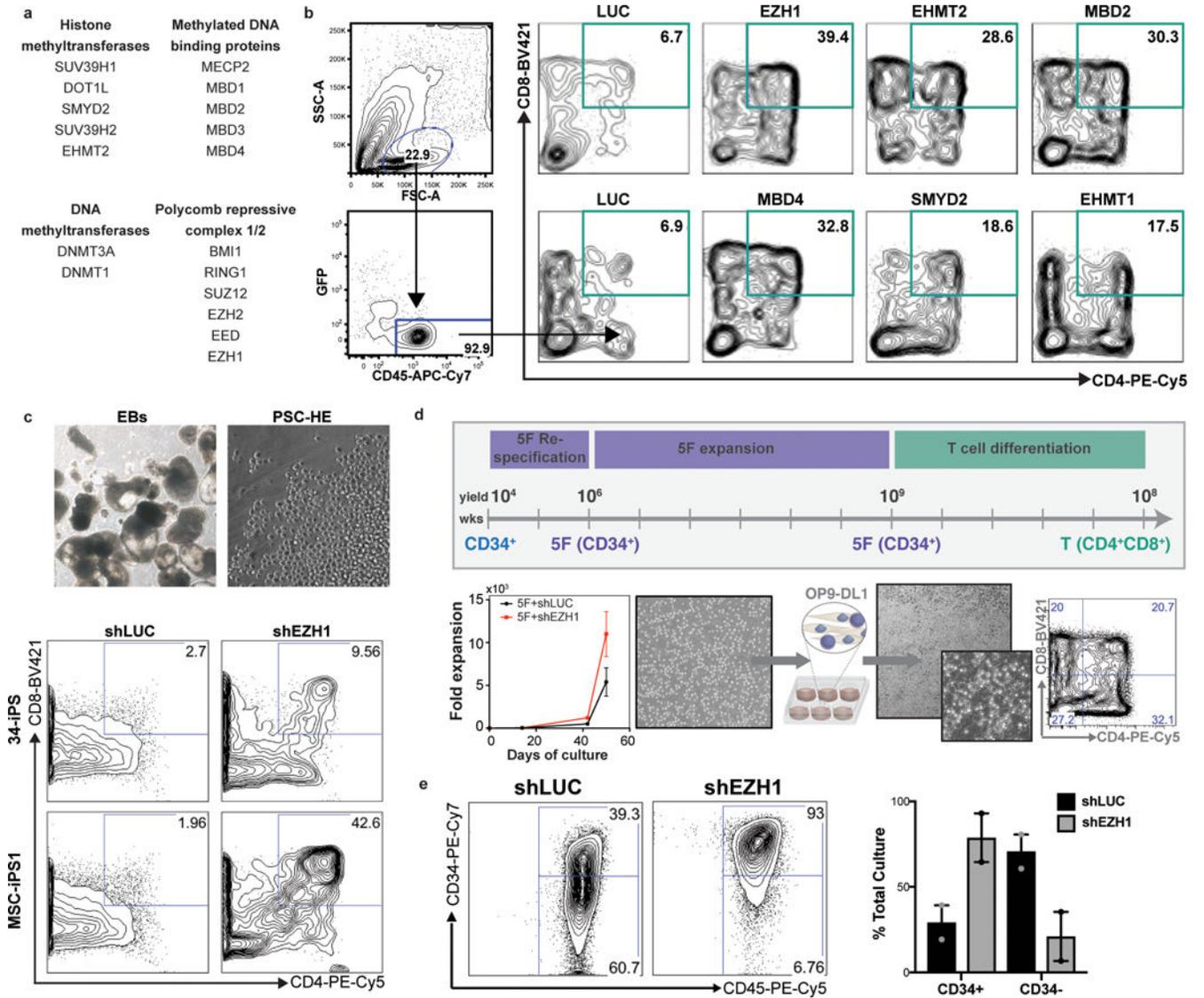
Extended Data 1. (related to Figure 1). EZH1 knockdown activates lymphoid potential from PSCs

(a) List of all candidate epigenetic modifiers in loss-of-function shRNA screen. (c) CD34+ cells were isolated after 9 days of EB differentiation (top left), transduced with shLUC or shEZH1 and cultured under endothelial-to-haematopoietic (EHT)-promoting conditions[40]. After 6 days, rounded haematopoietic cells (top right) were collected and co-cultured on OP9-DL1 stroma. Flow cytometric analysis of T cell potential in shLUC and shEZH1 cells without 5F is shown for two independent iPS lines in one experiment (n=2 biological replicates). (b) Representative flow plots of CD4+CD8+ T cell potential across top 6 candidates from 4 independent hairpins in two independent experiments (n=8). (d) Expansion and differentiation potential of 5F+shEZH1 cells after long-term in vitro culture. 5F+shEZH1 cells were maintained in +Dox cultures for 14 days respecification (~10^2-fold expansion), plus an additional 6 weeks (~10^3-fold expansion) and then plated into OP9-DL1

for T cell differentiation. Representative flow cytometric analyses of T cell potential of 5F +shLUC and 5F+shEZH1 cells after 13 weeks of expansion and differentiation in n=2 biological replicates. (e) Flow cytometric analysis (left) and quantification (right) of the proportion of $CD34^+$ and $CD34^-$ hematopoietic progenitors in +Dox suspension culture at day 25 in n=2 biological replicates.

Extended Data 2. (related to Figure 2). Ezh1, but not Ezh2, suppresses T cell potential and requires its catalytic domain

Ezh1, but not Ezh2, suppresses T cell potential and requires its catalytic domain(a) Quantitative PCR of PRC2 expression during the course of differentiation from hPSC-CD34+, respecification, expansion, OP9-DL1 co-culture and $CD4^+CD8^+$ T cells in n=2 biological replicates in one experiment. (b) Quantitative PCR of mRNA knockdown efficiency of individual shRNAs for PRC2 genes for n=2 replicates. (c) Western blot for EZH1 and GAPDH protein levels. (d) Scheme for rescue experiments. GFP+ 5F cells were

transduced with shRNAs and selected with puromycin. 5F+shRNA cells were then transduced with full-length murine Ezh1 ORF (mEzh1) or mEzh1 with the catalytic SET domain deleted (mEzh1 SET), marked by mCherry fluorescence. Triple-transduced (GFP[+], puro-resistant, mCherry[+]) cells were sorted and seeded onto OP9-DL1. T cells were analyzed by flow cytometry after 5 weeks of differentiation. (e) Expression of full-length murine Ezh1 (mEzh1), catalytic-deleted mEzh1 SET, or full-length Ezh2 in shLUC and shEZH1 cells by quantitative PCR. (f) Western blot validation of expression of mEzh1 or mEzh1 SET in shLUC and shEZH1 cells. (g) Representative flow cytometry plots of T cell potential for 5F+shLUC cells for rescue experiments detailed in (d) and from n=3 biological replicates (top). CD4+CD8+ T cells were verified for mCherry+ (bottom). See also Fig. 2c. All plots are gated on CD45+. (h) 5F+shRNA cells were transduced with full-length murine Ezh2 ORF (mEzh2) marked by mCherry fluorescence. Triple-transduced (GFP+, puro-resistant, mCherry+) cells were sorted and seeded onto OP9-DL1. T cells were analyzed by flow cytometry after 5 weeks of differentiation. Representative flow plots for n=2 biological replicates in one experiment. Quantitation for (h) is shown in (i) as mean ± SEM.
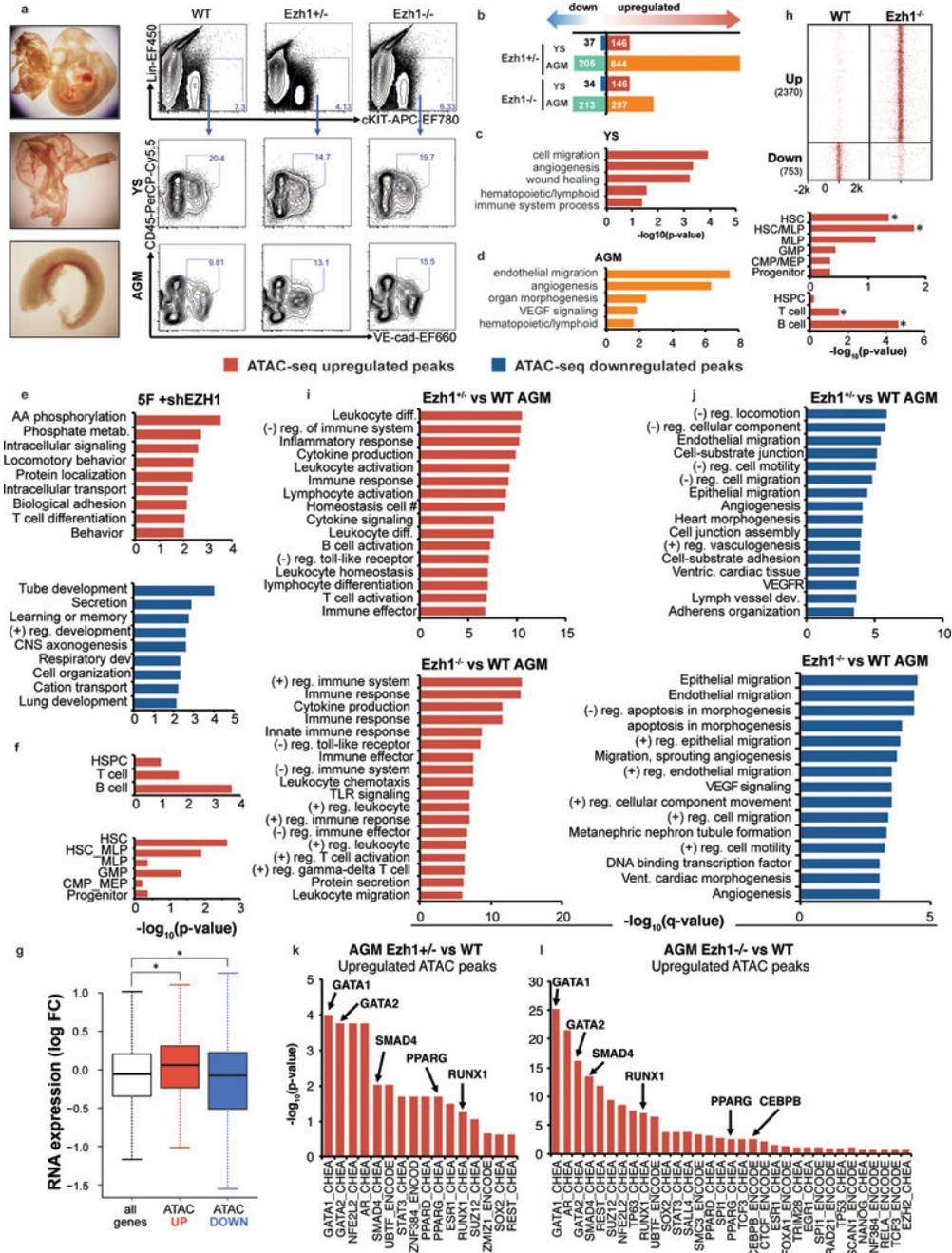
Extended Data 3. (related Figure 3). Ezh1 regulates hematopoietic and lymphoid programs in vitro and in vivo

(a) Representative images of E10.5 embryo (top), YS (middle) and AGM (bottom) from n>30 embryos. Lin⁻cKit⁺VE-Cadherin⁺CD45⁺CD41⁺ cells from E10.5 YS and AGM were FACS-sorted followed by RNA-seq analysis. (b) Genes upregulated and downregulated by >2-fold in *Ezh1*^+/− or *Ezh1*^−/− YS and AGM compared to WT. (c) and (d) GO term annotations of upregulated genes in *Ezh1*^+/− and *Ezh1*^−/− YS and AGM compared to WT. (e) GO analysis of enriched pathways of 1033 nearest neighbor genes associated with upregulated ATAC-seq peaks (top) GO analysis of nearest 1012 neighbor genes associated

with downregulated ATAC peaks (bottom). (f) Comparison of upregulated ATAC-seq peaks in 5F+shEZH1 cells with HSPC, B,T cell networks and (d) HSPC hierarchy signatures. (g) Box plot of expression of genes associated with upregulated (ATAC UP) and downregulated (ATAC DOWN) ATAC-seq peaks. *$p < 0.05$ via one-way ANOVA (h) ATAC-seq density map of $cKit^+VE$-cadherin$^+CD45^+$ HSPCs sorted from ~30 embryos of E10.5 WT and $Ezh1^{-/-}$ AGM (top) from one experiment. Significantly upregulated ATAC-seq peaks were compared to HSPC, T, B cell networks and signatures of the human HSPC hierarchy (bottom). (i) GO terms of enriched pathways of regions associated with significantly upregulated ATAC-seq peaks annotated by GREAT analysis in $Ezh1^{+/-}$ AGM (top) and $Ezh1^{-/-}$ AGM (bottom) compared to WT. (j) GO terms of enriched pathways of regions associated with significantly downregulated ATAC-seq peaks annotated by GREAT analysis in $Ezh1^{+/-}$ AGM (top) and $Ezh1^{-/-}$ AGM (bottom) compared to WT. (k) TF binding to genes with upregulated ATAC peaks in $Ezh1^{+/-}$ (left) and $Ezh1^{-/-}$ AGM (right) from (i) compared to WT AGM.
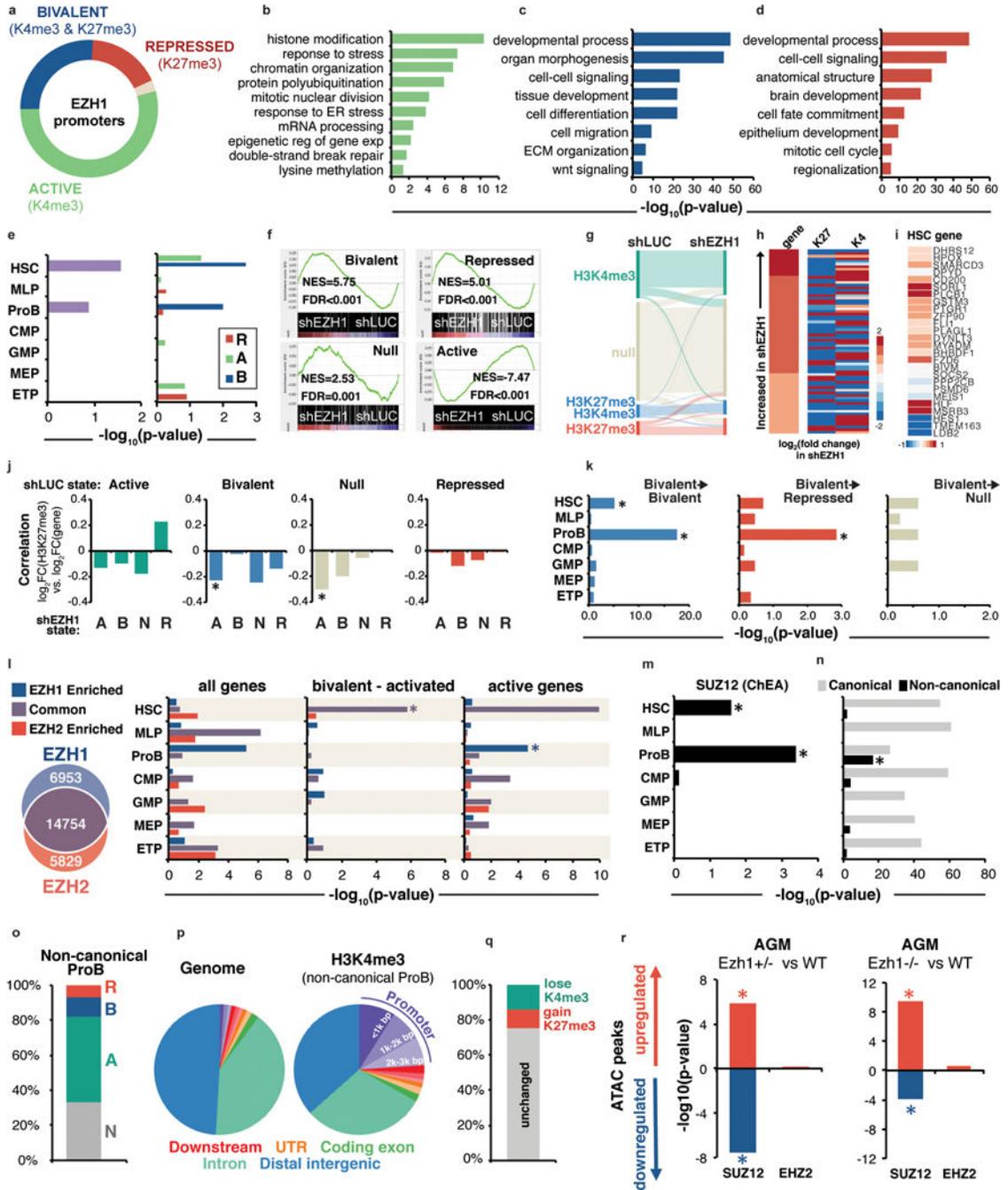
Extended Data 4. (related to Figure 3). Genome-wide chromatin occupancy reveals EZH1 enrichment at bivalent HSC genes and non-canonical active lymphoid genes

(a) Breakdown of EZH1 binding at promoter regions and associated histone marks. GO term analysis of EZH1-bound active (b), bivalent (c) and repressed (d) genes. (e) Distribution of EZH1-bound genes across the hematopoietic hierarchy (left) and their associated histone marks (right). R = repressed (H3K27me3-marked), A = active (H3K4me3-marked), B = bivalent (H3K4me3 and H3K27me3-marked). (f) GSEA analysis of EZH1-bound genes correlated with RNA-seq data upon *EZH1* knockdown. (g) Sankey diagram showing genome-wide changes in histone methylation status upon *EZH1* knockdown. (h)

Upregulated genes exhibit reciprocal decreases in H3K27me3 levels, as quantified by EpiChIP software. (i) Activated (formerly bivalent) HSC genes exhibit increased gene expression upon EZH1 knockdown and loss of H3K27me3. (j) Correlations between changes in H3K27me3 and gene expression levels upon *EZH1* knockdown, subdivided by subgroups corresponding to methylation changes. (k) breakdown of bivalent-bivalent (left), bivalent-repressed (center), and bivalent-null (right) genes upon *EZH1* knockdown across the hematopoietic hierarchy. (l) Overlap of EZH1 and EZH2 enriched peaks and the distribution of all EZH1 enriched, EZH2 enriched or common genes across the hierarchy (left), or specifically bivalent genes that become activated upon *EZH1* knockdown (middle) and active genes, marked by H3K4me3 in shLUC (right). (m) *SUZ12* binding (from the ChEA database) across the hematopoietic hierarchy. (n) Canonical and non-canonical targets, previously identified by Xu et al. *Mol Cell* (2015) across the hematopoietic hierarchy. (o) Breakdown of histone marks on non-canonical ProB genes and (p) the genome-wide distribution from CEAS analysis. (q) Changes in actively marked, non-canonical ProB genes (green bar, panel o), upon *EZH1* knockdown. (r) SUZ12 and EZH2 binding (ChEA database) at ATAC peaks in $Ezh1^{+/-}$ and $Ezh1^{-/-}$ AGM. *$p < 0.05$ via one-way ANOVA.
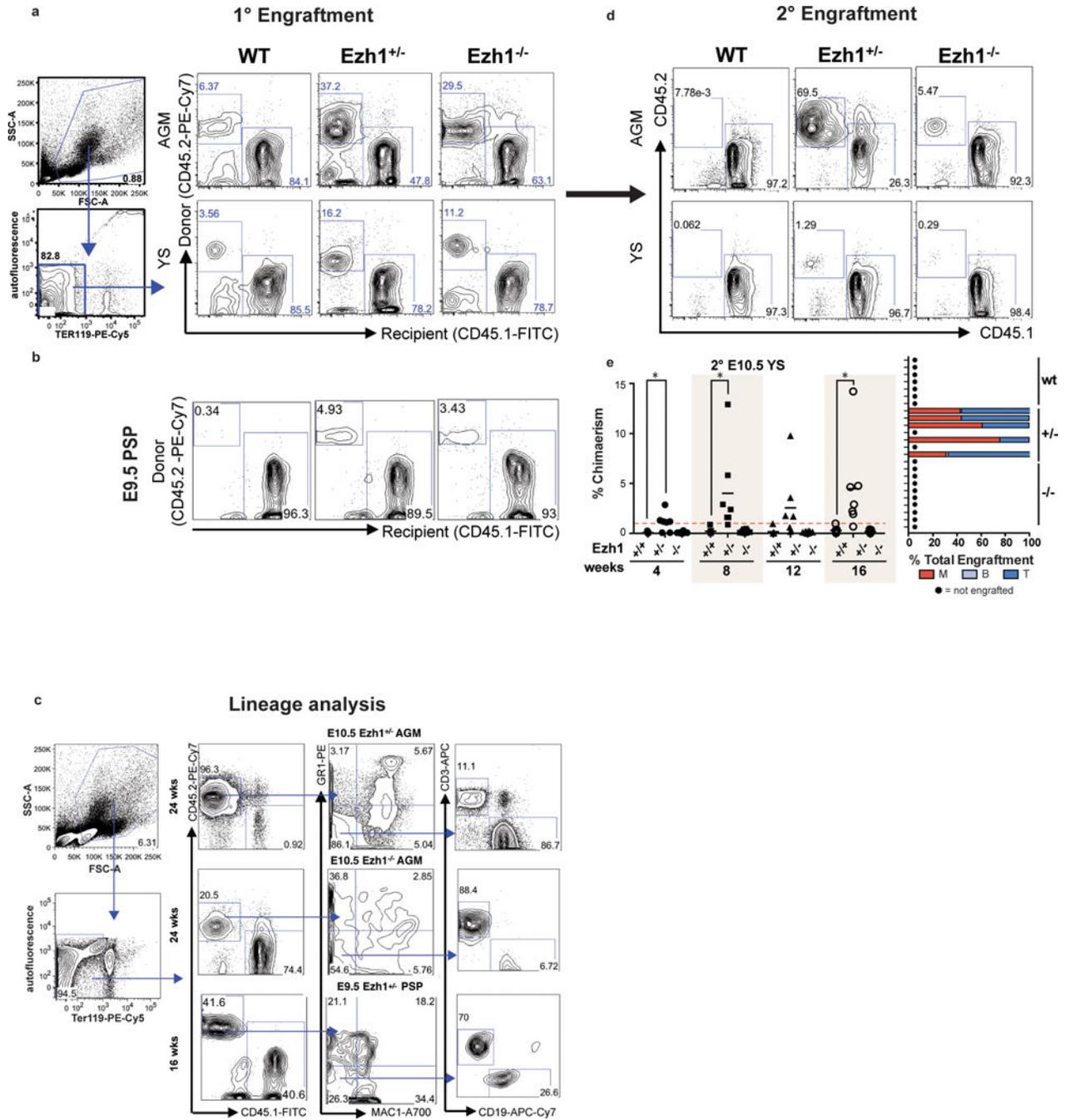
Extended Data 5. (related to Figure 4). Ezh1 deficiency enhances embryonic HSPC engraftment
(a) Whole E10.5 AGM and YS were transplanted intravenously into sublethally irradiated NSG adult females. Chimaerism was monitored via retroorbital bleeding every 4 weeks. Representative flow plots are shown for analysis after 4 weeks in n ≥ 3 mice. (b) Whole E9.5 PSP was transplanted intravenously into sublethally irradiated NSG adult females. Chimaerism was monitored via retroorbital bleeding every 4 weeks. Representative flow plots are shown for analysis after 8 weeks in n ≥ 3 mice. (c) Representative flow plots of lineage analysis in E10.5 AGM *Ezh1*[+/−] and *Ezh1*[−/−] primary transplant recipients after 24

weeks, and in E9.5 PSP *Ezh1*[+/−] primary transplant recipient after 16 weeks (n 3 mice per group). (d) Primary recipients in (a) were sacrificed after 24 weeks post-transplantation and 4×10^6 whole bone marrow was transplanted into sublethally irradiated adult NSG females intravenously. Chimaerism was monitored via retroorbital bleeding. Representative flow plots of E10.5 AGM and YS secondary transplants after 4 weeks in n 3 mice. (e) Secondary transplantation of E10.5 YS primary recipients in (Fig. 4f). (Right) Lineage distribution of E10.5 YS secondary recipients. Data is pooled across three independent experiments. *p<0.05, ** p<0.01 by unpaired two-sided t-test; see supplementary information for exact p values per time point. N.E. = not engrafted.
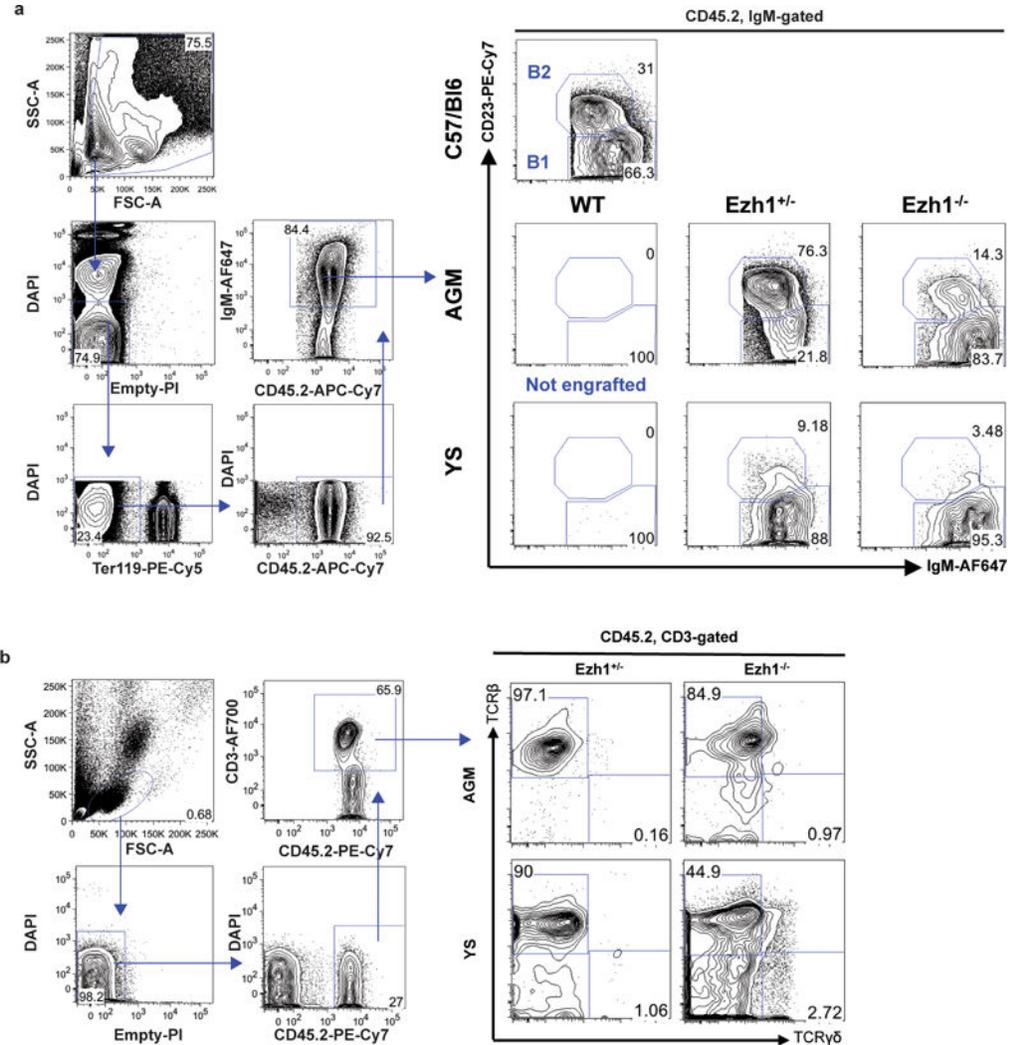
Extended Data 6. (related to Figure 4). Ezh1-deficient embryonic HSPCs contribute to adult-type lymphopoiesis in vivo

(b) Flow analysis of B1 and B2 progenitors in the peritoneal cavity of engrafted primary recipients (n=1 mouse per group). (b) Flow analysis of TCRβ and TCRγδ frequencies of donor-derived peripheral CD3+ T cells from engrafted primary recipients (n=1 mouse per group).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008; 132:631–644. [PubMed: 18295580]

2. Medvinsky A, Rybtsov S, Taoudi S. Embryonic origin of the adult hematopoietic system: advances and questions. Development. 2011; 138:1017–1031. [PubMed: 21343360]

3. Dzierzak E, Speck NA. Of lineage and legacy: the development of mammalian hematopoietic stem cells. Nat Immunol. 2008; 9:129–136. [PubMed: 18204427]

4. McGrath KE, et al. Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. Cell Rep. 2015; 11:1892–1904. [PubMed: 26095363]

5. Böiers C, et al. Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell. 2013; 13:535–548. [PubMed: 24054998]

6. Cedar H, Bergman Y. Epigenetics of haematopoietic cell development. Nat Rev Immunol. 2011; 11:478–488. [PubMed: 21660052]

7. Doulatov S, et al. Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. Cell Stem Cell. 2013; 13:459–470. [PubMed: 24094326]

8. Shen X, et al. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell. 2008; 32

9. Laurenti E, et al. The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat Immunol. 2013; 14:756–763. [PubMed: 23708252]

10. North TE, et al. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity. 2002; 16:661–672. [PubMed: 12049718]

11. Rybtsov S, et al. Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J Exp Med. 2011; 208:1305–1315. [PubMed: 21624936]

12. Boisset JC, et al. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature. 2010; 464:116–120. [PubMed: 20154729]

13. Riddell J, et al. Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. Cell. 2014; 157:549–564. [PubMed: 24766805]

14. Papaemmanuil E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet. 2009; 41:1006–1010. [PubMed: 19684604]

15. Tothova Z, et al. FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. Cell. 2007; 128:325–339. [PubMed: 17254970]

16. Bernstein BE, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006; 125:315–326. [PubMed: 16630819]

17. Xu J, et al. Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. Mol Cell. 2015; 57:304–316. [PubMed: 25578878]

18. Margueron R, et al. Ezh1 and Ezh2 maintain repressive chromatin through distinct mechanisms. Mol Cell. 2008; 32

19. Jang IH, et al. Notch1 acts via Foxc2 to promote definitive hematopoiesis via effects on hemogenic endothelium. Blood. 2015; 125

20. Hohmann AF, et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. Nat Chem Biol. 2016; 12:672–679. [PubMed: 27376689]

21. Müller AM, Medvinsky A, Strouboulis J, Grosveld F, Dzierzak E. Development of hematopoietic stem cell activity in the mouse embryo. Immunity. 1994; 1:291–301. [PubMed: 7889417]

22. Yoder M, Hiatt K. Engraftment of embryonic hematopoietic cells in conditioned newborn recipients. Blood. 1997; 89:2176–2183. [PubMed: 9058742]

23. Lee SC, et al. Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. Blood. 2015; 126:167–175. [PubMed: 26036803]

24. Xie H, et al. Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. Cell Stem Cell. 2014; 14:68–80. [PubMed: 24239285]

25. Majewski IJ, et al. Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. PloS Biol. 2008; 6

26. Herrera-Merchan A, et al. Ectopic expression of the histone methyltransferase Ezh2 in hematopoietic stem cells causes myeloproliferative disease. Nat Commun. 2012; 3

27. Cai Z, et al. Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity. 2000; 13:423–431. [PubMed: 11070161]

28. Sugimura R, et al. Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature. 2017; 545:432–438. [PubMed: 28514439]

29. Cahan P, et al. CellNet: network biology applied to stem cell engineering. Cell. 2014; 158:903–915. [PubMed: 25126793]

30. Park IH, et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature. 2008; 451:141–146. [PubMed: 18157115]

31. Chadwick K, et al. Cytokines and BMP-4 promote hematopoietic differentiation of human embryonic stem cells. Blood. 2003; 102:906–915. [PubMed: 12702499]

32. Onder TT, et al. Chromatin-modifying enzymes as modulators of reprogramming. Nature. 2012; 483:598–602. [PubMed: 22388813]

33. Holmes R, Zúñiga-Pflücker JC. The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. Cold Spring Harb Protoc. 2009; 2009 pdb.prot5156.

34. Trapnell C, et al. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013; 31:46–53. [PubMed: 23222703]

35. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods. 2013; 10:1213–1218. DOI: 10.1038/nmeth.2688 [PubMed: 24097267]

36. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. [PubMed: 19261174]

37. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]

38. Huang J, et al. Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. Dev Cell. 2016; 11:9–23.

39. Kim J, Cantor AB, Orkin SH, Wang J. Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. Nat Protoc. 2009; 4:506–517. [PubMed: 19325547]
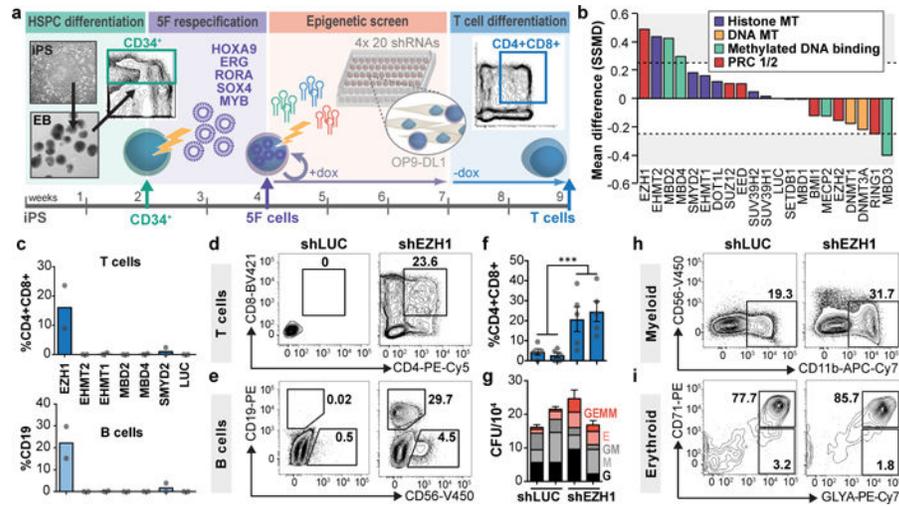
40. Ditadi A, et al. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nat Cell Biol. 2015; 17

**Figure 1. In vitro screen for epigenetic modifiers that restrict definitive lymphoid potential**
(a) Scheme for human PSC differentiation into haematopoietic progenitors. CD34+ cells were transduced with HOXA9, ERG, RORA, SOX4, and MYB (5F). 5F cells were then transduced with individual shRNAs (×4 each) targeting each epigenetic modifier and seeded onto OP9-DL1 stroma to induce T cell differentiation. (b) Strictly standardized mean difference (SSMD) of CD4+CD8+ T cell frequencies across all 4 shRNAs targeting each epigenetic modifier in 5F cells in n=2 independent experiments using two different iPSC lines, CD45-iPS and MSC-iPS1. (c) Prospective analysis of T cell and B cell frequencies from 5F+shRNA targeting top candidates (n=2 biological replicates). (d) Flow analysis of CD4+CD8+ T cell development of 5F cells with shRNAs targeting luciferase (shLUC) or EZH1 (shEZH1) after 5 weeks differentiation on OP9-DL1. (e) Flow analysis of CD19+ B cell potential. (f) Quantitation (mean ± SEM) of T cell potential of 5F+shEZH1 cells compared to 5F+shLUC cells pooled across 2 hairpins and 5 independent experiments (n=10) using multiple iPSC lines (CD34-iPS, CD45-iPS, MSC-iPS1). Source data files show individual values obtained for each hairpin. ***p=0.001 by unpaired two-tailed t-test (g) Quantitation of colony-forming potential in n=3 independent experiments. (h) Flow analysis of myeloid (CD11b+) and (i) erythroid (CD71+GLYA+) potential. Experiments replicated at least twice.
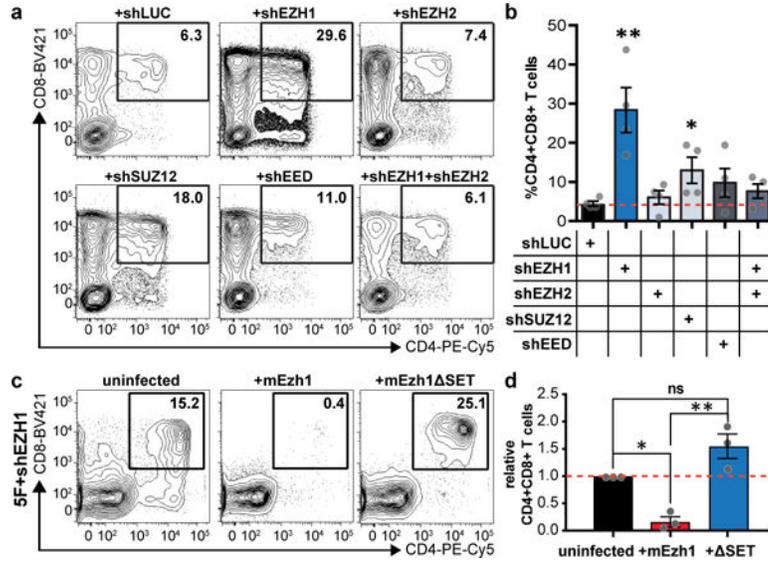
Figure 2. Repression of canonical PRC2 subunits does not activate lymphoid potential
(a) Representative flow plots of T cell potential of 5F cells with shRNAs targeting individual components of PRC2. (b) Quantitation (mean ± SEM) of T cell potential of 5F+shRNA targeting the indicated subunit in (a) shown as using two hairpins across two independent experiments (n=4) *p=0.0457, **p=0.0061 by unpaired two-tailed t-test. (c) Representative flow analysis of T cell potential in 5F+shEZH1 cells with mEzh1 or mEzh1 SET co-expressed. (d) Quantitation of flow analysis in (c) shown as mean ± SEM of n=3 biological replicates. *p=0.0146, **p=0.0011 by one-way ANOVA. All plots are gated on CD45[+]. Data is pooled across two independent experiments.
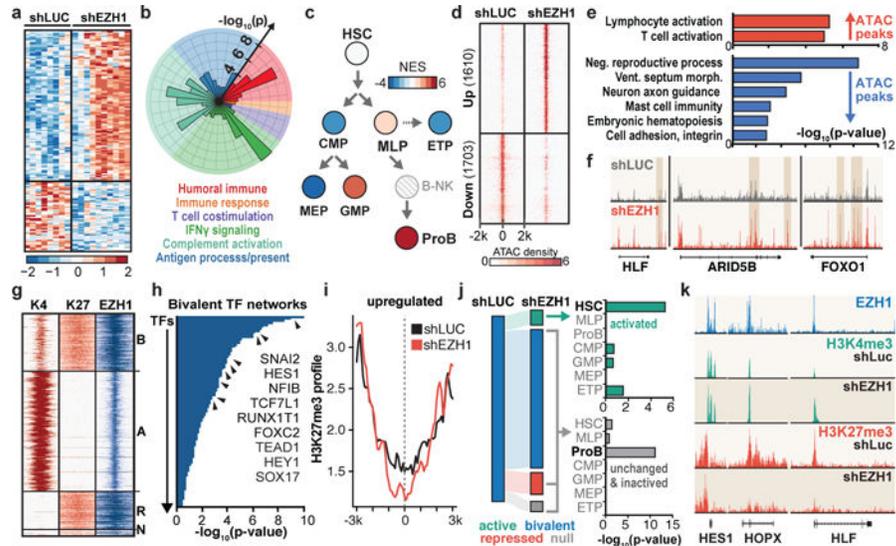
Figure 3. EZH1 directly binds to and modulates expression and chromatin accessibility of HSC and lymphoid genes

(a) Heatmap of upregulated (104) and downregulated (49) genes (>2-fold; Benjamini-Hochberg corrected t-test, p<0.1) from RNA seq analysis of CD34$^+$CD38$^-$ HSPCs 5F+shEZH1 (n=10 biological replicates) compared to 5F+shLUC cells (n=8 biological replicates). (b) GO analysis of biological processes associated with significantly upregulated genes in (a), subdivided by GO hierarchical categories and p-values labeled along radius. (c) Enrichment of human HSC and progenitor signatures by GSEA in 5F+shEZH1 compared with 5F+shLUC cells, overlaid on the map of human HSPC hierarchy. (d) Density map of upregulated and downregulated ATAC peaks by MAnorm in 5F+shEZH1 cells compared to 5F+shLUC from n=2 biological replicates. (e) GO terms of enriched biological processes of ATAC-seq peaks in (d) by GREAT analysis. (f) Tracks of representative genes that acquire a significant ATAC peak upon EZH1 knockdown. (g) ChIP-seq density map of EZH1 peaks within bivalent (B), repressed (R), active (A) or null (N) promoter groups from n=2 biological replicates. K4 = H3K4me4, K27 = H3K27me3 (h) Waterfall plot of CellNet[29] predicted regulators of EZH1-bound bivalent gene networks. (i) Sitepro quantitative analysis of H3K27me3 levels at all upregulated genes around the transcription start site upon EZH1 knockdown, relative to shLUC in n=2 biological replicates. (j) Sankey diagram illustrating histone methylation changes of all bivalent genes in shLUC cells and after EZH1 knockdown n=2 biological replicates (left). Genes that lose H3K27me3 (become activated) are specifically enriched in HSC signature, whereas bivalent genes that are unchanged or repressed are enriched in ProB signature (right) by Fisher's exact test. (k) ChIP-seq tracks of EZH1, H3K4me3 and H3K27me3 at representative HSC promoter regions in shLUC and shEZH1 cells. Experiments replicated at least twice.
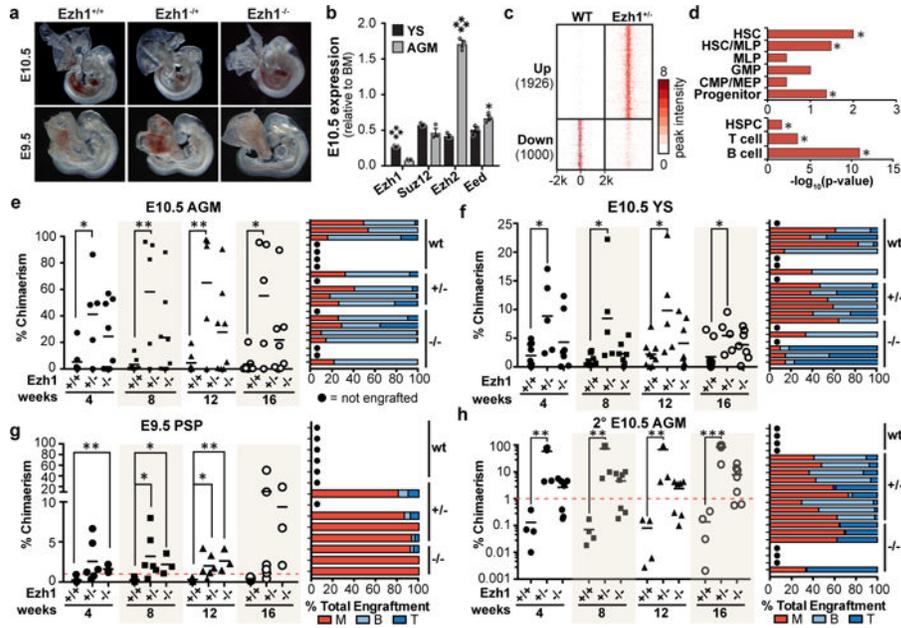
Figure 4. Ezh 1 deficiency increases lymphoid potential and engraftment of embryonic HPSCs
(a) Representative images of E9.5 and E10.5 embryos (n>50 embryos). (b) Quantitative PCR of each PRC2 subunit in E10.5 WT YS and AGM as mean ± SEM of n=3 biological replicates. *p=0.0439,****p<0.0001 by unpaired two-tailed t-test (c) ATAC-seq density map of cKit$^+$VE-cadherin$^+$CD45$^+$ HSPCs sorted from 30 pooled embryos of E10.5 WT and Ezh1$^{+/-}$ AGM. (d) Significantly upregulated ATAC-seq peaks were compared to HSPC, T, B cell networks and signatures of the human HSPC hierarchy. *p<0.05 by Fisher's exact test. (e) Engraftment of E10.5 AGM (3.5ee) in sublethally-irradiated adult NSG females. Donor chimerism marked by CD45.2$^+$ was measured in peripheral blood every 4 weeks up to 16 weeks post-transplantation. Each dot represents a single transplant recipient. (Right) Lineage distribution of engrafted mice showing T cell (T), B cell (B), and myeloid (M) contribution. (f) Engraftment of E10.5 YS (5ee). (Right) Lineage distribution of engrafted mice. (g) Engraftment of E9.5 PSP (10ee). (Right) Lineage distribution of engrafted mice. (h) Serial transplantation of whole BM from primary recipients of E10.5 AGM cells in (e). Secondary transplant was carried out after 24 weeks of primary transplant. (Right) Lineage distribution of engrafted mice. n ≥ 3 mice per group; *p<0.05, ** p<0.01, ***p<0.0001 by unpaired two-tailed t-test. See supplemental information for exact p values per time point. Data is pooled across four independent experiments in (e, f), four independent experiments in (g), three independent experiments in (h); experiment in (c) performed once.