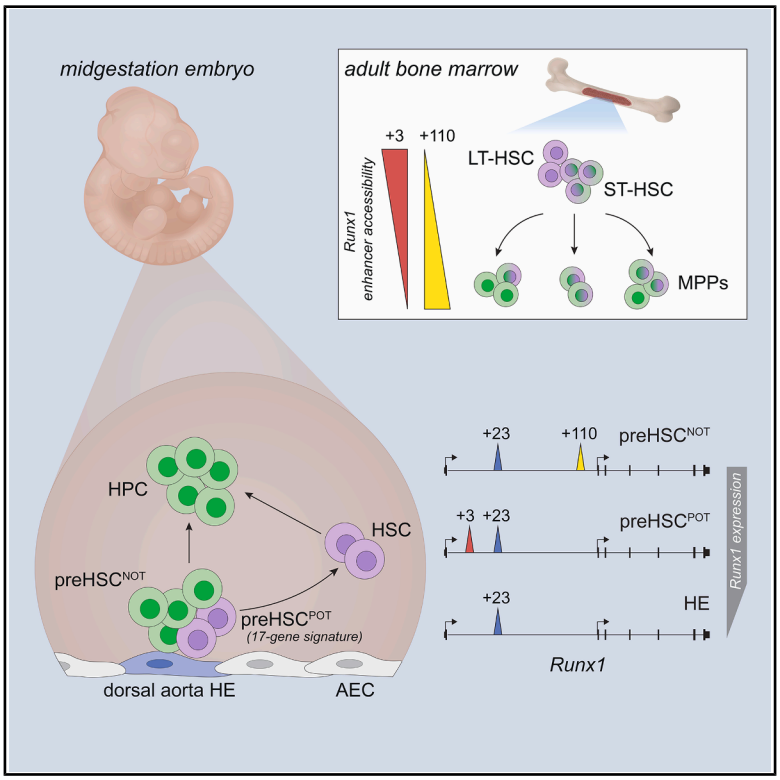
Cell Reports

Report

Resolving hematopoietic stem versus progenitor cell potential in the mouse dorsal aorta by differential Runx1 +110 enhancer activity

Graphical abstract



Highlights

- *Runx1* +110 enhancer activity segregates emerging dorsal aorta HPCs from HSCs
- Comparative transcriptomics revealed a molecular signature linked to HSC potential
- A single-cell multiome resource of transcriptomic and epigenomic states in EHT
- *Runx1* +3 enhancer is preferentially accessible in both embryonic and adult HSCs

Authors

Giorgio Anselmi, Vincent Frontera, Christina Rode, ..., Ditsa Levanon, Yoram Groner, Marella F.T.R. de Bruijn

Correspondence

giorgio.anselmi@icgeb.org (G.A.), marella.debruijn@imm.ox.ac.uk (M.F.T.R.d.B.)

In brief

During embryonic development, the specification of hematopoietic stem cells (HSCs) and progenitor cells (HPCs) critically depends on the transcription factor Runx1. Anselmi et al. describe a *Runx1* enhancer whose activity helps distinguish emerging HSCs from HPCs, enabling comparative transcriptomic and epigenomic profiling of the nascent HSC lineage.

Anselmi et al., 2026, Cell Reports *45*, 117737
August 25, 2026
https://doi.org/10.1016/j.celrep.2026.117737

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

Resolving hematopoietic stem versus progenitor cell potential in the mouse dorsal aorta by differential Runx1 +110 enhancer activity

Giorgio Anselmi,[1,4,9,*] Vincent Frontera,[1,5,9] Christina Rode,[1,6,9] Andrew Jarratt,[1,9] Naeema T. Mehmood,[1] Matthew Nicholls,[1] Luca Buttler,[1] Stella Antoniou,[1,7] John O'Rourke,[1] Emanuele Azzoni,[1,8] John Stamatoyannopoulos,[2] Ditsa Levanon,[3] Yoram Groner,[3] and Marella F.T.R. de Bruijn[1,10,*]

[1]MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DS, UK
[2]Altius Institute for Biomedical Sciences, 2211 Elliott Ave, Seattle, WA 98121, USA
[3]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel
[4]Present address: International Centre for Genetic Engineering and Biotechnology, ICGEB, Trieste, Italy
[5]Present address: QIAGEN, Aix-en-Provence, France
[6]Present address: QIAGEN, Aix-en-Provence, France
[7]Present address: Biobanking and BioMolecular Resources Infrastructure-European Research Infrastructure Consortium (BBMRI-ERIC), Graz, Austria
[8]Present address: School of Medicine and Surgery, University of Milano-Bicocca, Monza Italy, and Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy
[9]These authors contributed equally
[10]Lead contact
*Correspondence: giorgio.anselmi@icgeb.org (G.A.), marella.debruijn@imm.ox.ac.uk (M.F.T.R.d.B.)
https://doi.org/10.1016/j.celrep.2026.117737

SUMMARY

Hematopoietic stem cells (HSCs) are important in cell-based therapies for blood-related disorders. While progress has been made in generating HSCs by directed differentiation of pluripotent stem cells (PSCs), such cultures promote hematopoietic progenitor cells (HPCs) over HSCs. Thus, elucidating markers and factors associated with HSC versus HPC development is imperative. HSCs and HPCs originate from hemogenic endothelium (HE) in a Runx1-dependent process. Here, we characterize a *Runx1* enhancer (+110) that distinguishes emerging dorsal aorta HSCs from HPCs. Comparative transcriptomics reveal a 17-gene signature associated with *in vivo* long-term HSC potential, while single-cell multiome analysis demonstrates a clear epigenetic identity of dorsal aorta HE, preHSCs, and HPCs, providing a resource of stage-specific enhancer activity. Our study demonstrates the power of cell type-specific enhancer-reporter models to dissect cell fate decisions in development and provides new inroads to label and/or perturb HSC versus HPC potential *in vivo* and *in vitro*.

INTRODUCTION

Hematopoietic stem cells (HSCs) are responsible for lifelong maintenance of hematopoiesis. They are clinically important in cell-based therapies for blood disorders. In spite of recent breakthroughs using chemically defined cultures, it remains challenging to generate HSCs *de novo* in the laboratory.[1] To date, pluripotent stem cell (PSC)-derived hematopoietic differentiation cultures have generated many hematopoietic progenitor cells (HPCs) but only few HSCs.[2,3] During mouse embryonic development, the hematopoietic system is established in multiple asynchronous waves, tightly restricted in space and time.[4] The first two waves occur in the yolk sac (YS) between embryonic day 7.25 (E7.25) and E9.5 and generate primitive erythrocytes, megakaryocytes, and macrophages, followed by definitive-type lineage-restricted HPCs.[5–8] The third and last wave develops intra-embryonically in the midgestation aorta-gonad-mesonephros region (AGM) and is characterized by the generation of HSCs[9–12] via a stepwise process of maturation from hemogenic endothelium (HE), through E9.5 proHSCs and E10.5–E11.5 preHSC type 1 (preHSC1) and type 2 (preHSC2) intermediates.[13] This endothelial-to-hematopoietic transition (EHT)[14] is accompanied by the appearance of intra-aortic hematopoietic cell clusters (IAHCs) that bud from the ventral wall of the dorsal aorta (DA).[14] IAHCs are heterogeneous and contain (pre)HSCs as well as HPCs.[15,16] Between E10.5 and E11.5, the mouse aorta contains approximately 500–700 cluster cells, but only 1–2 are *bona fide* long-term multilineage reconstituting HSCs.[9,17,18] Arguably, a better understanding of the regulatory events driving the establishment of (pre)HSCs versus HPCs will aid the *de novo* generation of HSCs for therapeutic purposes.

To date, there is a paucity of reliable phenotypic markers discriminating between the emergence of HSC versus HPC

potential during AGM EHT. To obtain such tools, we probed the *Runx1* locus for cell type-specific *cis*-regulatory elements. The transcription factor (TF) Runx1 is a master regulator of developmental hematopoiesis, with loss of Runx1 leading to a block in EHT at the level of the HE and a complete absence of all HSCs and (pro)definitive HPCs.[19–25] *Runx1* is transcribed from two alternative promoters, neither of which conveys tissue-specific expression, suggesting that tissue/cell type specificity is mediated by distal enhancers.[26–28] Analysis of dynamic chromatin conformation changes *in vitro* suggested that the enhancers mediating hematopoietic-specific *Runx1* expression are restricted to the gene body.[29] We previously identified and characterized the *Runx1* +23 enhancer, which in transgenic enhancer-reporter mouse models mediates reporter expression to all cells undergoing EHT.[30–32] To date, no *Runx1 cis*-elements specific to either HE, HSCs, or HPCs have been identified.

Here, we report the specificity of the *Runx1* +110 enhancer in developmental hematopoiesis. We generated a transgenic mouse line carrying a GFP reporter transgene under the spatiotemporal control of the *Runx1* +110 enhancer. Phenotypic, functional, and transcriptomic analyses of GFP-expressing cells showed that the +110 enhancer recapitulates part of the endogenous *Runx1* expression pattern in YS and AGM HPCs while not being active in HE or functional HSCs and enabled the identification of a 17-gene signature associated with long-term reconstituting HSC (LT-HSC) potential. Finally, we generated a single-cell (sc) multiome dataset enriched for cells undergoing EHT, which provides a resource of stage-specific enhancers and identified an additional, complementary *Runx1* enhancer that is accessible in preHSCs.

RESULTS

The 110GFP enhancer reporter marks distinct subsets of Runx1-expressing hematopoietic cells at the onset of hematopoiesis

Within the *Runx1* gene body, we identified four enhancers which mediated *lacZ* reporter expression to the midgestation DA where *Runx1* is expressed in cells undergoing EHT[30,33] (Figure 1A). Analysis of E8.5 transient transgenic conceptusses suggested that, of these, the *Runx1* +110 enhancer displayed specificity for part of the *Runx1* expression pattern, labeling hematopoietic cells in the YS blood islands, but no DA HE (Figure 1B). To characterize the hematopoietic specificity of the *Runx1* +110 enhancer during developmental hematopoiesis, we generated a mouse line carrying a *hsp68Gfp* +110 transgene, referred to as 110GFP (Figure 1C). Whole-mount examination of E8–E14 110GFP transgenic embryos showed GFP expression at sites of hematopoietic cell emergence and colonization, namely the YS and fetal liver (FL; Figures 1D and S1A). In the E8 YS, 110GFP expression was detected in 24.4% ± 5.1% of Ter119$^+$ erythrocytes and 62.7% ± 7.6% of emerging VECad$^+$Ter119$^-$CD41$^+$CD45$^{-/+}$ HPCs. Among VECad$^+$Ter119$^-$CD41$^-$CD45$^-$ endothelial cells (ECs), in contrast, very little 110GFP expression was seen (Figures 1D and 1E; Figure S1B). Of note, the HPC gate is likely to include some 110GFP$^-$ ECs, as these express low levels of CD41 at E8.5[31]. At E9.5, 94.2% ± 2.2% of the erythromyeloid progenitors (EMPs) expressed 110GFP (Figure S1C). This was also borne out functionally, with most YS colony-forming unit in culture cells (CFU-Cs) marked by the 110GFP transgene; similarly, E11.5 FL CFU-Cs were mostly 110GFP$^+$ (Figure S1D).

In the midgestation DA, 110GFP expression was strikingly restricted to the Runx1$^+$ IAHCs budding into the vessel lumen but absent from the Runx1$^+$ HE and subaortic mesenchyme (Figure 1F). Most of the IAHCs (65%) contained a mix of 110GFP$^-$ and 110GFP$^+$ cells, while clusters containing either 110GFP$^-$ or 110GFP$^+$ cells were less prevalent (26% and 9%, respectively) and smaller (Figure S2). Flow cytometry analysis of the E10.5 110GFP AGM + vitelline and umbilical arteries (AVU) corroborated the restriction of 110GFP expression to hematopoietic cells, including erythrocytes (44.2% ± 10.1%) and CD41$^+$CD45$^+$ hematopoietic stem/progenitor cells (HSPCs) (51.7% ± 11.6%) (Figures 1G and 1H; Figures S1E and S1F). Functional analysis of midgestation 110GFP$^+$ AVU cells showed that, on average, 88%, of CFU-Cs were marked (Figure 1I). Altogether, these results demonstrate that 110GFP reporter expression is restricted to part of the hematopoietic cells at intra- and extra-embryonic sites of HSPC emergence.

The 110GFP enhancer reporter predominantly marks primitive macrophages, EryPs, and emerging Hoxa$^+$ HSPCs

To comprehensively identify the hematopoietic cells labeled by the 110GFP transgene, E10 110GFP$^+$ AVU cells were profiled by single-cell RNA sequencing (scRNA-seq) (Figure 2A). In parallel, a reference dataset capturing all E10 AVU hemogenic and hematopoietic cells was generated (from 23GFP transgenic embryos[31,32]). After quality control, the 110GFP$^+$ and 23GFP$^+$ datasets were integrated, and non-hematopoietic contaminants were removed (Figures S3A–S3D). Within the remaining 2,260 110GFP$^+$ and 4,246 23GFP$^+$ cells, thirteen clusters were identified, covering EHT and differentiating hematopoietic lineages (Figures 2B, S3E, and S3F; Table S3A). As expected, the 23GFP reporter labeled cells in 12 out of the 13 clusters, i.e., all but the primitive erythrocytes (EryPs, known to no longer express Runx1/23GFP at this time point).[22,30] In contrast, 110GFP labeled predominantly macrophages (pMac, Mac1, and Mac2) and EryPs but virtually none of the definitive erythrocytes (EryDs) and a few MEP, Mk, pMC, or pNeu progenitors (Figures 2B, and 2C, S3G, and S3H); of note, the latter may represent the few 110GFP- CFU-Cs in the E10.5/E11.5 AVU/YS/FL (Figure S1D). In line with the immunophenotypic and functional analyses (Figures 1F–1H), 110GFP also labeled HPCs (HPC1 and HPC2) and cells in a cluster with EHT characteristics (expressing endothelial genes together with *Runx1*, *Hlf*, and Myb) but very few if any ECs (Endo) (Figures 2B–2D, S3G, and S3H). Only the Endo, EHT, and HPC1 clusters expressed *Hoxa* genes (Figures 2D and 2E; Figure S3I), indicative of a recent intraembryonic origin.[35,36]

Re-clustering of the *Hoxa*$^+$ Endo, EHT, and HPC1 intraembryonic populations resulted in the identification of 4 main subclusters (Figure 2F). Based on differential endothelial and hematopoietic gene expressions, these were annotated as arterial endothelial cells (AECs; expressing *Cdh5*, *Sox17*, and *Gja4*), HE (expressing *Gja4*, *Gja5*, *Runx1* [low], *Vwf*, and *Procr*),

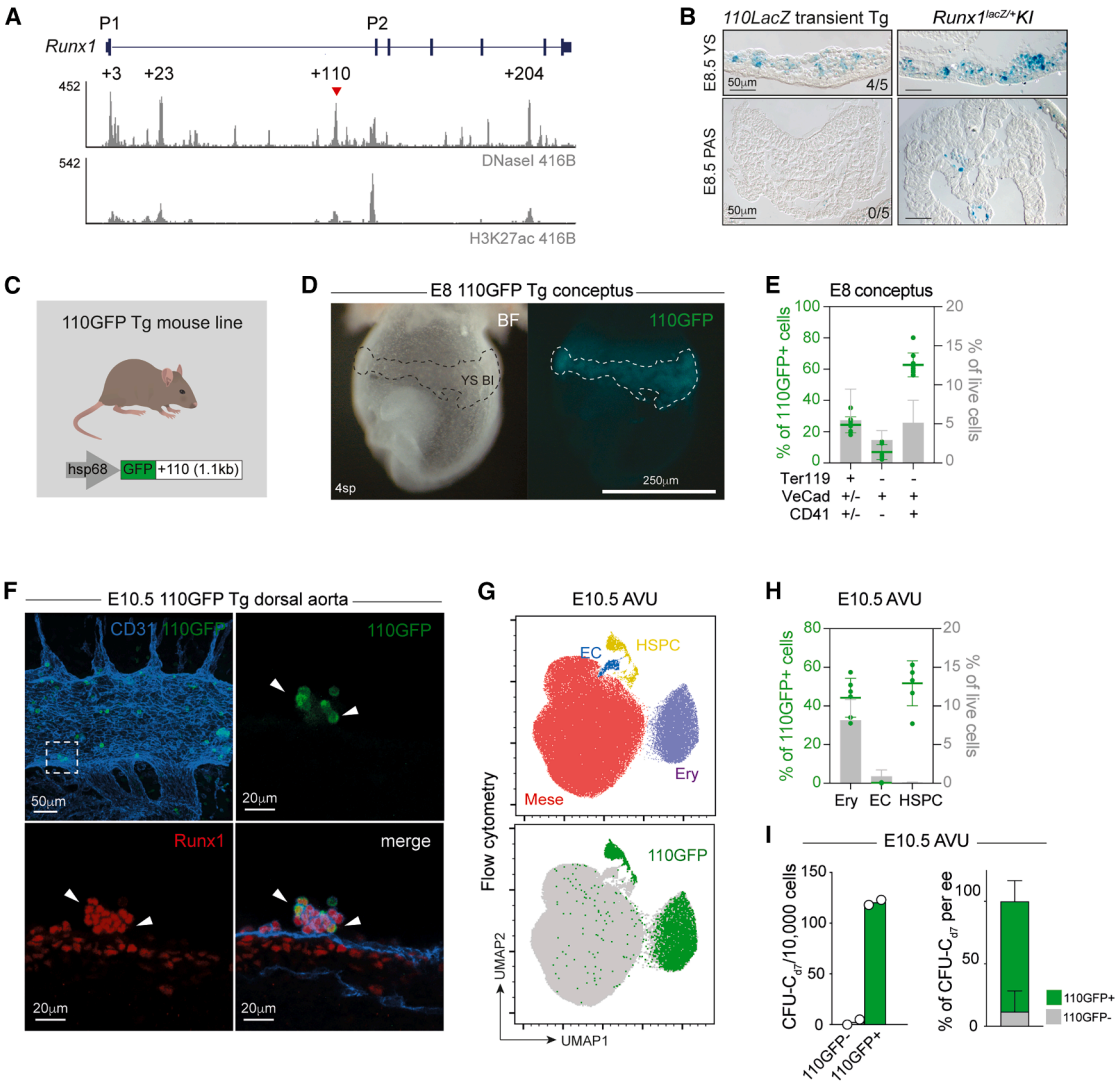


Figure 1. The 110GFP enhancer reporter marks distinct subsets of *Runx1*-expressing hematopoietic cells at the onset of hematopoiesis

(A) DNaseI-seq[34] and H3K27ac ChIP-seq[33] tracks of the *Runx1* locus in 416B cells. Red arrowhead indicates the +110 *Runx1* enhancer.

(B) X-gal staining (blue) in E8.5 YS and PAS from *Runx1* +110 enhancer-*lacZ* transient transgenic (Tg) embryos compared with Runx1$^{lacZ/+}$ KI. Fraction of embryos showing X-gal staining is reported. Scale bar = 50 μm.

(C) Schematic of the *hsp68Gfp+110* enhancer reporter construct used to generate the 110GFP Tg mouse by pronuclear injection.

(D) Whole-mount image of E8 110GFP Tg embryo. Dashed line highlights YS blood island. Scale bar = 250 μm.

(E) 110GFP expression in E8 concepti by flow cytometry (gating strategy in Figure S1B). Error bars represent mean ± SD.

(F) Whole-mount immunofluorescence of E10.5 110GFP Tg. Arrowheads indicate CD31$^+$Runx1$^+$ clusters. Scale bar = 50 μm (top left) and 20 μm (others).

(G) Flow cytometry analysis of E10 AVU of 110GFP Tg. Cell type-specific markers are presented in Figure S1D. GFP$^+$ cells (green) overlayed on GFP$^-$ cells (gray).

(H) 110GFP expression in E10.5 AVU cells, assessed by flow cytometry (gating strategy in Figure S1F). Error bars represent mean ± SD.

(I) CFU-C potential of GFP$^+$ and GFP$^-$ cells from E10.5 110GFP Tg AVU. Number of CFU-C$_{d7}$ per 10,000 cells (left) and frequency of CFU-C$_{d7}$ per embryo equivalent (ee) (right). Error bars represent mean ± SD.

HPCs (expressing *Runx1*, *Kit*, *Rac2*, and *Mpo*), and preHSCs (expressing *Vwf*, *Procr*, and *Gfi1*, along with *Runx1*, *Kit*, and *Hlf*) (Figures 2F and 2G; Table S3B). Comparison with published gene module scores[37–39] further supported this annotation (Figure S3J). While AECs and HE were only 23GFP$^+$, the preHSC and HPC clusters comprised both 110GFP$^+$ and 23GFP$^+$ cells (Figure 2F, bar graph), raising the question whether the HSC lineage is labeled by the *Runx1* +110 enhancer reporter.

Cells belonging to the functional HSC lineage are not marked by the 110GFP reporter

In the E9.5 para-aortic splanchnopleura with vitelline and umbilical arteries (PAS+VU), 110GFP expression was nearly absent from proHSCs (4.5% ± 1.1%, while marking 45.8% ± 3.9% of HPCs (Figures 3A and 3B; Figure S4A). In contrast, 37.2% ± 5.4% and 45% ± 5.9% of phenotypic E10.5 preHSC1 and E11.5 preHSC2/HSCs were marked by 110GFP, respectively

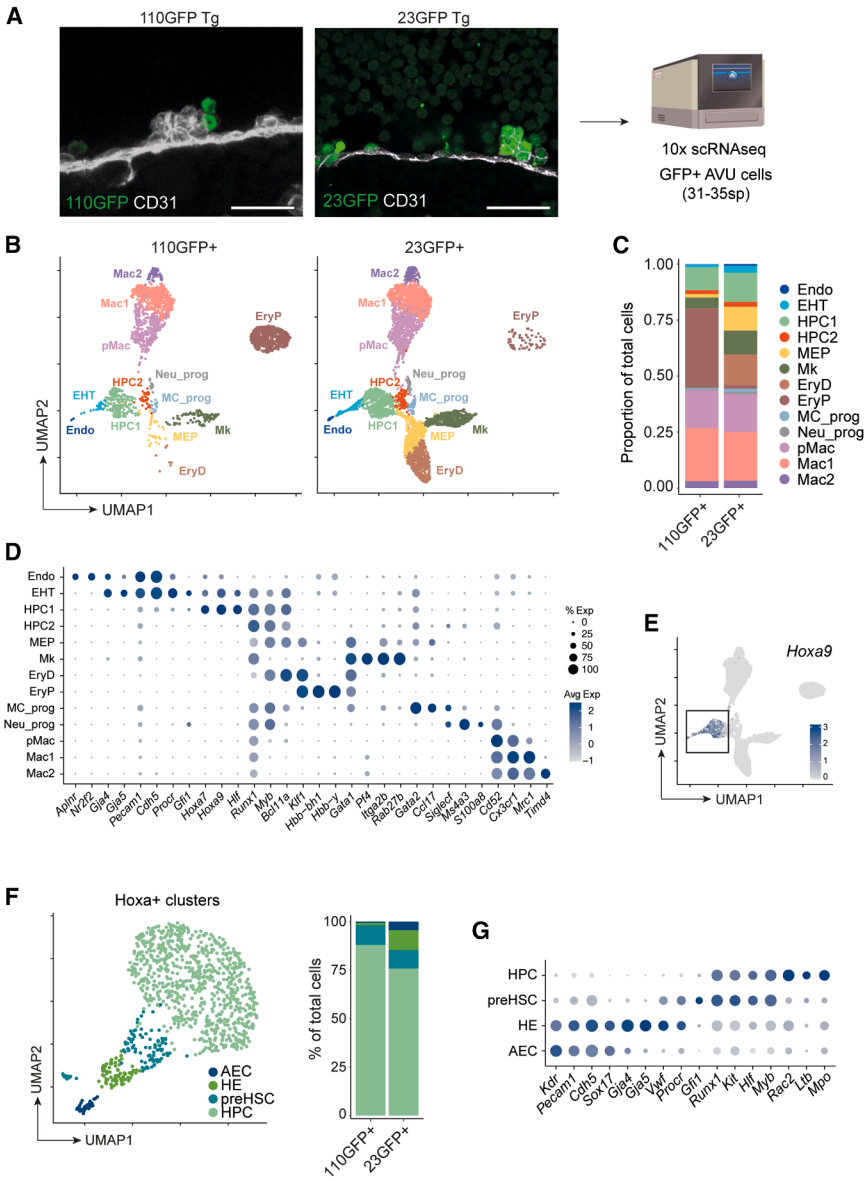


Figure 2. The 110GFP transgene marks primitive macrophages, EryPs, and emerging *Hoxa*$^{+}$ HSPCs

(A) Schematic of scRNA-seq analysis of E10 110GFP^{+} and 23GFP^{+} AVUs. Scale bar = 50 μm.

(B) UMAP and unsupervised clustering of 110GFP and 23GFP cells from E10 AVU. Endo, endothelial cells; EHT, endothelial-to-hematopoietic transition; HPCs, hematopoietic progenitor cells; MEP, megakaryocyte and erythrocyte progenitor; Mk, megakaryocytes; EryD, definitive erythrocytes; EryP, primitive erythrocytes; MC_prog, mast cell progenitor; Neu_prog, neutrophil progenitor; pMac, macrophage precursors; Mac, macrophage.

(C) Frequency of clusters within each sample.

(D) Marker genes defining cell clusters in (B).

(E) Expression of *Hoxa9* in 110GFP^{+} and 23GFP^{+} integrated cells.

(F) UMAP and unsupervised clustering of subsets expressing *Hoxa* genes (left) and cluster frequency within each sample (right).

(G) Marker genes defining *Hoxa*$^{+}$ clusters. AEC, arterial endothelial cells; HE, hemogenic endothelium; preHSC, HSC precursors; HPC, hematopoietic progenitor cell.

(Figures 3A and 3B). Of note, (pre)HSCs and HPCs of the E10.5/11.5 AVUs share the same phenotype,[13] thus requiring functional tests to distinguish between them. Transplantation assays (direct for HSCs, after OP9 co-aggregates for preHSCs[13,40]) clearly showed that functional preHSC and HSC potential segregated entirely to the 110GFP^{-} AVU cells (Figures 3C and 3D). Notably, phenotypically defined LT- and ST (short term)-HSCs in adult bone marrow (BM) also lacked 110GFP expression (Figure S4B). Interestingly, while 15.9% ± 9.2% of phenotypic HSCs in the E14 FL did express 110GFP (Figure S4C), this

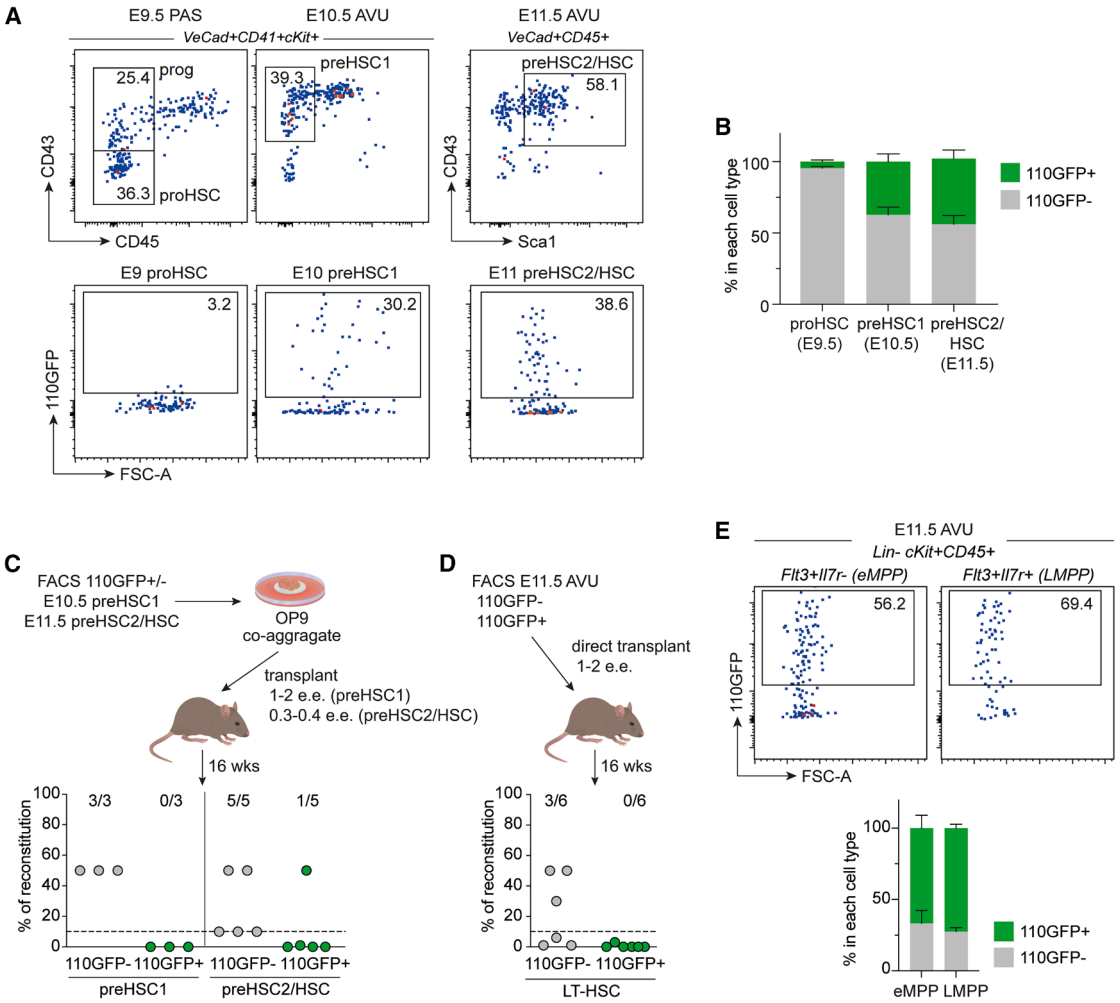


Figure 3. Cells belonging to the functional HSC lineage are not marked by 110GFP

(A) Flow cytometry analysis of E9.5 pro-HSCs, E10.5 preHSC1, and E11.5 preHSC2/HSC.

(B) Quantification of 110GFP expression in cell types from (A). Error bars represent mean ± SD.

(C) Repopulation activity of 110GFP$^+$ or 110GFP$^-$ pre-HSC1 and pre-HSC2/HSC after OP9 co-aggregate cultures 16 weeks post-transplant (engraftment threshold: 10%). The ratio of reconstituted/injected mice is reported.

(D) Repopulation activity of 110GFP$^+$ and 110GFP$^-$ cells from E11.5 AVU 16 weeks post-transplant (engraftment threshold: 10%). The ratio of reconstituted/injected mice is reported.

(E) 110GFP expression in eMPPs and LMPPs from E11.5 AVU, assessed by flow cytometry. Error bars represent mean ± SD.

may be due to the presence of hematopoietic cells devoid of LT-HSC potential.[41]

Recently, AGM-derived preHSCs were suggested to give rise to Flt3$^+$cKit$^+$CD45$^+$Il7r$^-$embryonic multipotent progenitors (eMPPs) that contribute to adult MMP3/4 and a significant fraction of multilineage adult blood.[42] Flow cytometry analysis of E11.5 AVUs showed that 66.7% ± 9% of phenotypic eMPPs expressed 110GFP, which, together with 110GFP expression among E14 FL and adult BM MMP3 and MMP4 (Figure 3E; Figures S4B–S4D), suggests that 110GFP labels functional eMPPs. Examining this directly requires lineage tracing experiments, which is beyond the scope of this study. In addition to eMMPs, 72.5% + 2.8% of lymphoid-primed multipotent progenitors (LMPPs) expressed 110GFP (Figures 3E and S4D). In summary, our data showed that while part of the phenotypic HPCs, eMPPs, and LMPPs of the midgestation AVU expressed 110GFP, cells belonging to the HSC lineage, i.e., proHSCs and functional preHSC1/2 and HSCs, were not marked by 110GFP.

Emergence of HSC potential in 110GFP$^-$ cells is marked by a 17-gene signature

To identify markers that can distinguish cells with preHSC potential among the emerging hematopoietic cells in the midgestation AVU, we analyzed known HSC-associated genes in the *Hoxa*$^+$ "preHSC" UMAP (uniform manifold approximation and projection) cluster (Figure 2F) and observed distinct expression patterns between 23GFP$^+$ versus 110GFP$^+$ cells. For example, 23GFP$^+$ but not 110GFP$^+$ "preHSCs" expressed *Tcf15*,

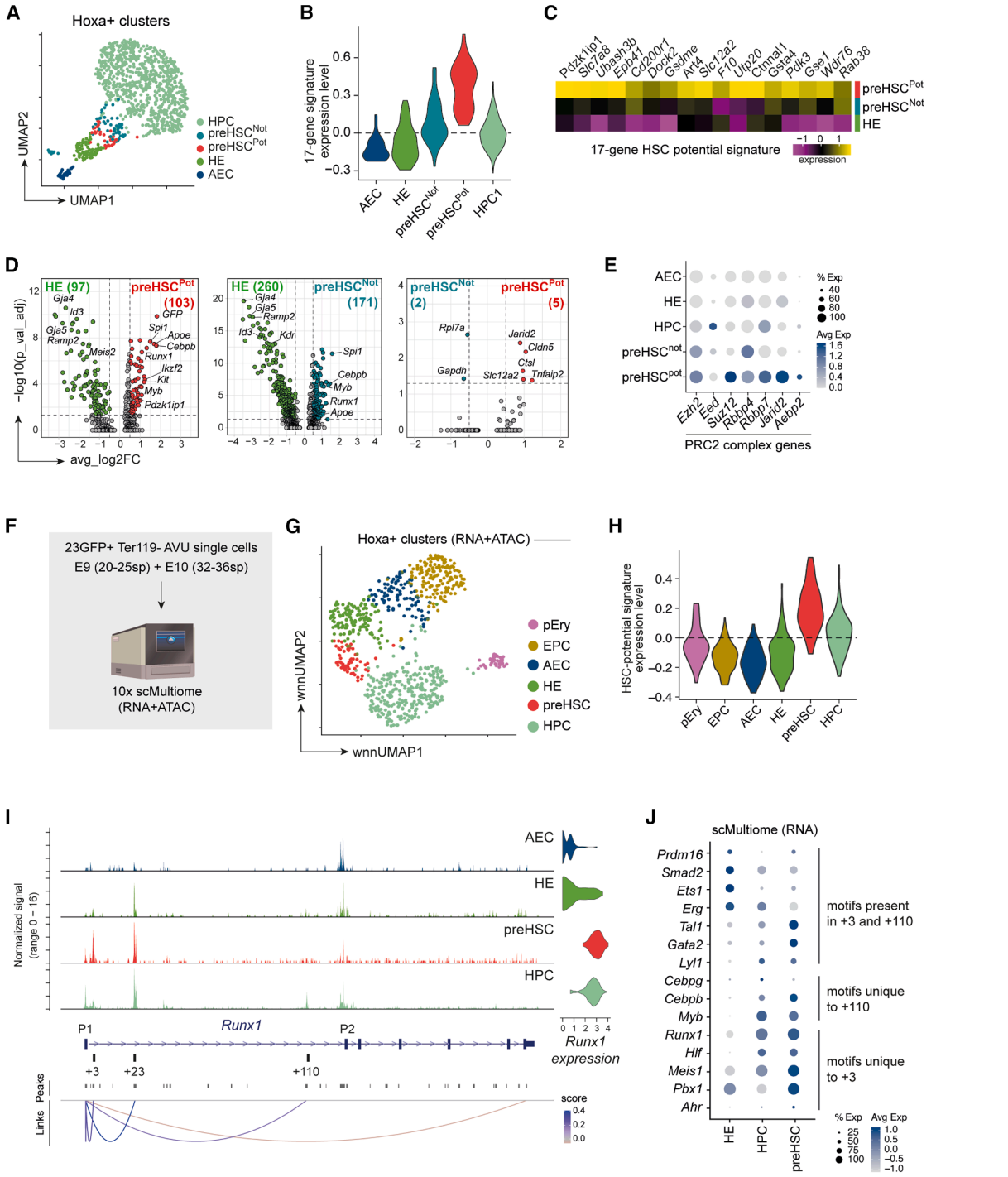


Figure 4. Emergence of HSC potential in 110GFP$^-$ cells is marked by a 17-gene signature and the activation of the *Runx1* +3 enhancer

(A) Refined annotation of *Hoxa*$^+$ clusters including functionally defined preHSCs (preHSC^{Pot} cluster from Figure S5C in red).

(B) Expression of the HSC potential signature in *Hoxa*$^+$ clusters.

(C) HSC potential signature (3-way comparison of HE, preHSC^{Not}, and preHSC^{Pot}). Rows represent pseudobulk average expression in each cluster.

(D) DEGs in pairwise comparisons of HE, preHSC^{Pot}, and preHSC^{Not}.

(E) Average expression of the PRC2 complex genes in *Hoxa*$^+$ clusters defined in (A).

(F) Schematic of scMultiome analysis of Ter119-23GFP$^+$ cells from E9 PAS and E10 AVU.

(G) UMAP of joint RNA and ATAC measurements, using weighted nearest neighbor (WNN) integration for scMultiome *Hoxa*$^+$ clusters.

(H) Expression of the HSC potential signature in scMultiome *Hoxa*$^+$ clusters.

(legend continued on next page)

Pdzk1ip1, and *Id3* (Figure S5A), previously associated with adult BM LT-HSC activity,[43–45] in line with 23GFP marking the HSC lineage throughout development[31,32] (Figure S5B). Conversely *Cd27*, which is expressed at intermediate levels in HSCs,[39] was highly expressed in 110GFP$^+$ cells (Figure S5A). On the basis of differential expression of HSC-associated genes, we could divide the ''preHSC'' UMAP cluster into two distinct subsets, where one harbors functional preHSCs (110GFP$^-$ preHSC^{Pot}) and the other does not (110GFP$^+$ preHSC^{Not}) (Figures S5C and 4A). While, at the immunophenotypic level, preHSC^{Pot} and preHSC^{Not} were indistinguishable (CITE-seq, Figure S5D), direct comparison and differential gene expression testing of preHSC^{Pot}, preHSC^{Not}, and HE identified a 17-gene signature that correlated strongly with the preHSC potential (Figures 4B and 4C; Table S3C; STAR Methods). Importantly, the 17-gene signature was also enriched in published scRNA-seq datasets of functionally validated mouse AGM/FL/BM preHSCs/HSCs and among a population that included human AGM HSCs[46] where 10 out of 17 genes were expressed, indicating its broader applicability in distinguishing HSC lineage potential (Figure S6).

To further explore candidate factors associated with the establishment of HSC potential in the DA, we performed pairwise comparisons between preHSC^{Pot}, preHSC^{Not}, and HE. Both preHSC^{Pot} and preHSC^{Not} were clearly distinct from HE, with significantly upregulated hematopoiesis-associated genes (e.g., *Runx1*, *Spi1*, *Cebpb*, *Kit*, and *Myb*) and downregulated HE-associated arterial markers (e.g., *Gja4*, *Gja5*, *Id3*, and *Ramp2)* (Figure 4D; Tables S3D–S3G). In contrast, only a few differentially expressed genes (DEGs) could be identified between preHSC^{Pot} and preHSC^{Not} (Tables S3H and S3I). Interestingly, the most significantly upregulated gene in preHSC^{Pot} was *Jarid2* (Figure 4D). Jarid2 is part of the Polycomb repressive complex 2 (PRC2), and, in line with this, an increased expression of PRC2 core (*Eed*, *Suz12*, *Ezh2*, and *Rbbp4* or *Rbbp7)* and accessory (*Aebp2)* subunits was observed in preHSC^{Pot} cells, suggesting a potential PRC2-mediated role of Jarid2 in HSC development (Figure 4E).

Identification of *Runx1* enhancer activity during AVU EHT that is complementary to the +110 enhancer

Having established that *Runx1* +110 enhancer activity inversely correlates with HSC emergence, we asked whether there are other *Runx1* enhancers associated specifically with the HSC lineage. For this, we generated a scMultiome dataset (RNA + ATAC) on 23GFP$^+$ cells undergoing EHT in the E9.5 PAS+VU and E10.5 AVU (Figures 4F and S7; Table S4A). Re-clustering of *Hoxa*-expressing multiome clusters identified six distinct populations (Figures 4G and S8A–S8C). Four of these corresponded to AEC, HE, preHSC, and HPC clusters, also seen in E10 scRNA-seq (Figures 4F and 2F). The remaining two were endothelial progenitor cells (EPCs) and a small cluster of erythrocyte progenitors (pEry) (Figures 4F and S8D). The preHSC multiome cluster was enriched for our 17-gene signature, indicating the presence of preHSC^{Pot} (Figure 4H). Genome-wide analysis of the accessible peaks in HE, preHSC, and HPC clusters identified 151, 543, and 896 unique open chromatin peaks, respectively, which contained shared as well as cell type-specific TF motifs (Figures S8E and S8F; Tables S4B and S4C). HPC-specific motifs were enriched for hematopoietic TFs (in line with the more differentiated state of these cells), while the preHSC cluster was enriched for HOX motifs and the motifs associated with TGFb/BMP signaling (SMAD), neural crest development (GSX and ISL1), and chromatin architecture (CTCF and DLX) (Figure S8G).

Within the *Runx1* locus, we detected two differentially accessible regions, corresponding to the +110 and the +3 enhancer, enriched in HPCs versus preHSCs, respectively (Figures 4I and S8H; Table S4B). Interestingly, in adult BM, the +3 enhancer also showed the highest accessibility in LT/ST-HSCs,[47] while the +110 enhancer became accessible upon differentiation from ST-HSCs (Figure S9A). The differential accessibility of the +3, +110, and +23 (accessible throughout hematopoiesis; Figures 4I, S8H, and S9A) enhancers was reflected in their differential activity in luciferase assays (Figure S9B). Importantly, the +3 enhancer mediated reporter gene expression to sites of hematopoietic cell emergence and colonization in E8.5 and E10 transient transgenic embryos (Figures S9C and S9D),[33] which support its role in developmental hematopoiesis.

Comparing candidate regulatory modules in the +3 and +110 enhancers showed 13 and 11 conserved TF motifs unique to the +3 and +110 enhancers, respectively, and 11 shared motifs (Tables S5A and S5B; STAR Methods). Shared motifs included broad hematopoietic TF motifs such as ETS, E-box, and GATA (Figures S10 and S11), mutation of which negatively affected enhancer activity.[33] Motifs unique to the +110 enhancer included CEBP and MYB, in line with 110GFP labeling myeloid cells. Motifs unique to the +3 enhancer, in contrast, included RUNX, HLF, MEIS, PBX, and AHR (Figure S10). Interestingly, upstream TF for these motifs showed differential expression between HE, preHSCs, and HPCs in mouse, with broadly comparable differential expression seen in humans (Figures 4J, S10B, S10C, S11B, and S11C).

DISCUSSION

During embryonic development, the first HSCs are generated alongside HPCs in the major arteries of the embryo, the DA and vitelline and umbilical arteries. Here, we showed that *Runx1* +110 enhancer activity inversely correlates with the emergence of the HSC lineage and demonstrated its use in elucidating possible mechanisms involved in HSC specification *in vivo*.

In developmental hematopoiesis, 110GFP reporter expression was observed in YS blood cells, in the FL, and in IAHCs in the midgestation AVUs where it labeled part of the phenotypic identified HPCs, eMPPs, LMPPs, and preHSCs. However, while

(I) ATAC-seq tracks and peak-gene links for the Runx1 locus in *Hoxa*$^+$ clusters. AEC, arterial endothelial cell; HE, hemogenic endothelium; preHSC, HSC precursor; HPC, hematopoietic progenitor cell.
(J) Gene expression of TFs whose motifs were found in conserved phylogenetic footprints of the +3 and +110 enhancers.

marking CFU-Cs, 110GFP labeled neither functional preHSCs (preHSC^{Pot}) nor *bona fide* HSCs. Its complete absence from the functional HSC lineage and heterogeneous activity pattern among hematopoietic progenitors was corroborated in an ATAC-seq study of adult BM[47] and reflected in our single-cell transcriptome analysis of 110GFP^{+} AVU cells. We exploited the absence of 110GFP labeling on DA preHSC^{Pot} to identify a 17-gene transcriptomic signature associated with the emergence of LT-HSC potential in the mouse embryo. This signature was also enriched in functionally validated scRNA-seq datasets of mouse DA preHSC/HSC, FL, and BM HSCs,[37,39] as well as in human AGM HSCs.[46] Of note, markers known to be associated with emerging HSCs, such as *Hlf*, *Mllt3*, and *Mecom*,[46] were not part of our signature, as we found them to be expressed also in HE and/or HPCs. Instead, our signature included *Pdzk1ip1* (aka Map17) and *Ctnnal1* (aka α-catulin) that have been previously associated with adult BM HSCs.[44,48–52] Thus, the 110GFP reporter and the 17-gene signature are valuable negative and positive indicators of HSC potential, respectively. Moreover, the extensive sequence conservation between the mouse *Runx1* +110 enhancer and its human equivalent (+140), along with broadly comparable expression of candidate upstream regulators, supports further investigation of the potential of the human *RUNX1* +140 enhancer reporter to identify nascent transplantable induced PSC (iPSC)-derived HSC (iHSCs). In addition, it will be of interest to test expression of the 17-gene signature in iHSCs, that, along with the enhancer reporter, could be used to optimize iHSC generation.

The heterogeneous expression of 110GFP among IAHCs (including preHSCs) in the midgestation AVU raises questions about the relationship between the HSC lineage, emerging 110GFP^{+} non-HSC progenitors, and HE. As clusters of $\geq$5 cells were reported to be polyclonal, while smaller ones (<5 cells) were found to be partly clonal,[53] we explored the composition of the small IAHCs. Among these, 110GFP^{-}, 110GFP^{+}, and mixed phenotypes were observed, precluding unambiguous inference of lineage relationships. However, as 110GFP^{+} preHSC^{Not} were detected in the AVU one full day before *bona fide* HSCs, it is unlikely that they are the progeny of HSCs. Whether preHSC^{Not} derive directly from HE, proHSCs, or preHSC^{Pot}, we cannot, at present, distinguish. In the YS where no HSCs were generated, nearly all emerging CFU-C and EMP were 110GFP^{+}, indicating that 110GFP^{+} progenitors can emerge directly from HE.

The transcriptional regulation of *Runx1* is complex, and its expression levels require precise spatiotemporal control for normal hematopoietic development and adult hematopoietic homeostasis.[24,54,55] To assess which enhancers other than +110 and +23 may be active during EHT, we analyzed chromatin accessibility in primary E9.5-E10.5 23GFP^{+} DA EHT cells (scMultiome). This dataset showed that distinct epigenetic profiles characterize consecutive stages of EHT progression, representing a valuable resource of stage-specific candidate enhancers. Within the *Runx1* locus, the specificity of the +110 enhancer for HPCs and that of the +23 enhancer for all cells undergoing EHT were corroborated. Interestingly, accessibility of another *Runx1* enhancer element, the +3, peaked in AVU preHSCs in a pattern complementary to the +110. This, combined with labeling of hemogenic sites in +3 enhancer-reporter transient transgenic mouse embryos, warrants further examination of +3 enhancer activity in the emerging HSC lineage.

Conserved TF motifs in the *Runx1* +3 and +110 enhancers only partially overlapped. The +110 enhancer harbors GATA and CEBP motifs, while the +3 enhancer harbors GATA, MEIS, and RUNX motifs, with corresponding upstream TFs showing dynamic expression between HE, HPCs, and preHSCs in mouse and humans. Interestingly, a recent study by Frömel et al.[56] on the design principles of cell state-specific enhancers in hematopoiesis reported that Gata1/2, Cebpa, Meis1, and Runx1 are involved in complex interplays that could lead to cell type-specific activation, neutralization, or repression depending on the cell lineage. This suggests that the specificity of the +3 and +110 enhancers is likely determined by a complex interplay between these factors on the *Runx1* enhancer elements. Unraveling the full regulatory complexity of *Runx1* is relevant to better understand *RUNX1*-associated hematopoietic disorders.

Collectively, our data support a model in which phenotypically homogeneous preHSCs arise from HE in the embryo major arteries. A subset of them, the preHSC^{Not} (where *Runx1* +23 and +110 enhancer are active), will only contribute to HPCs and, possibly, eMPPs/LMPPs. In contrast, the preHSC^{Pot} (in which the +23 and +3 enhancers are active) will mature into engraftable HSCs. Future studies into the emergence of preHSC^{Pot} and preHSC^{Not} and their neighboring cells will shed light on the molecular mechanisms determining these potentials and feed into designing and monitoring robust *in vitro* culture conditions supporting *de novo* HSC generation and/or expansion.

Limitations of the study

Formally establishing the developmental origin of 110GFP^{+} preHSC^{Not} and their relationship with other HSPCs/HE in the DA requires lineage tracing studies using cell type-specific and temporally controlled Cre lines and/or barcodes, which is outside of the scope of the present study. Irrespective of this, we believe that the 110GFP marker remains a unique tool to separate HSC versus HPC potential when they first emerge from HE.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Marella F. T. R. de Bruijn (marella.debruijn@imm.ox.ac.uk).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact, with a completed material(s) transfer agreement.

Data and code availability

scRNA-seq and scMultiome data are available under the ArrayExpress accession numbers E-MTAB-15202 and E-MTAB-15412, respectively. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank current and past members of the de Bruijn laboratory for their practical help, advice, and discussions, Simone Riva for computational advice, and Pik-Shan Li and Jackie Sloane-Stanley for generation of mouse transgenics. This work was supported by programs in the MRC Molecular Haematology Unit

Core award to M.D.B. (MC_U137970202, MC_UU_12009/2, MC_UU_00016/2, MC_UU_00029/5) and a Weizmann - UK Joint Research Program ("Making Connections") supported by the Weizmann friends in the UK to M.D.B., D.L., and Y.G. We would like to acknowledge Kevin Clark, Sally Clark, and Paul Sopp in the flow cytometry facility at the MRC WIMM for providing cell sorting services. The facility is supported by the MRC TIDU; MRC MHU (MC_UU_12009); NIHR Oxford BRC; Kay Kendall Leukaemia Fund (KKL1057), John Fell Fund (131/030 and 101/517); the EPA fund (CF182 and CF170); and the MRC WIMM Strategic Alliance awards G0902418 and MC_UU_12025. We thank Neil Ashley and Maria Greco for help on 10× Chromium sample processing. The WIMM Single Cell Core Facility was supported by the MRC MHU (MC_UU_12009), the Oxford Single Cell Biology Consortium, (MR/M00919X/1), and the WT-ISSF (097813/Z/11/B#). The facility was supported by WIMM Strategic Alliance awards G0902418 and MC_UU_12025. We would like to acknowledge Tim Rostron in the Sequencing facility at the MRC WIMM for providing sequencing services. The facility is supported by the MRC TIDU and by the EPA fund (CF268). We thank Christopher Lagerholm and Jana Koth for their imaging assistance. The Wolfson Imaging Center Oxford is supported by the MRC via the WIMM Strategic Alliance (G0902418), the MHU (MC_UU_12009), the TIDU (MC_UU_12010), the Wolfson Foundation (18272), and the Wellcome Trust (Micron 107457/Z/15Z) grants.

AUTHOR CONTRIBUTIONS

Conceptualization, G.A. and M.F.T.R.d..B.; methodology, G.A., V.F., C.R., A.J., N.T.M., L.B., M.N., J.O., S.A., E.A., and M.F.T.R.d..B.; formal analysis, G.A., V.F., and C.R.; investigation, G.A., V.F., C.R., D.L., Y.G., and M.F.T.R.d..B.; resources, G.A., C.R., A.J., J.S., and M.F.T.R.d..B.; data curation, GA; writing – original draft, G.A., C.R., and M.F.T.R.d..B.; writing – review & editing, G.A., C.R., M.N., E.A., D.L., Y.G., and M.F.T.R.d..B.; visualization, G.A., C.R., and M.N.; supervision, G.A. and M.F.T.R.d..B.; funding acquisition, D.L., Y.G., and M.F.T.R.d..B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
- METHOD DETAILS
  - Timed matings and embryo collection
  - Analysis of enhancer-lacZ transient transgenic embryos
  - Dissections and generation of cell suspensions
  - Flow cytometry and cell sorting
  - Immunofluorescence analysis and imaging
  - Hematopoietic progenitor assay
  - Long-term HSC reconstitution analysis
  - OP9 maintenance and co-aggregate cultures
  - Single-cell RNA-seq cell isolation and library preparation
  - Single-cell RNA-seq analysis
  - HSC potential gene signature
  - Single-cell multiome cell isolation and library preparation
  - Single-cell multiome analysis
  - Motif analysis of Runx1 +3 and +110 enhancers
  - Luciferase reporter assay
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplementary data related to this article can be found online at https://doi.org/10.1016/j.celrep.2026.117737.

Received: July 15, 2025
Revised: May 19, 2026
Accepted: July 7, 2026

REFERENCES

1. Wilkinson, A.C., and de Bruijn, M.F.T.R. (2024). A developmental route to hematopoietic stem cells. Nat. Biotechnol. *43*, 1242–1243. https://doi.org/10.1038/s41587-024-02398-7.
2. Piau, O., Brunet-Manquat, M., L'Homme, B., Petit, L., Birebent, B., Linard, C., Moeckes, L., Zuliani, T., Lapillonne, H., Benderitter, M., et al. (2023). Generation of transgene-free hematopoietic stem cells from human induced pluripotent stem cells. Cell Stem Cell *30*, 1610–1623.e7. https://doi.org/10.1016/j.stem.2023.11.002.
3. Ng, E.S., Sarila, G., Li, J.Y., Edirisinghe, H.S., Saxena, R., Sun, S., Bruveris, F.F., Labonne, T., Sleebs, N., Maytum, A., et al. (2024). Long-term engrafting multilineage hematopoietic cells differentiated from human induced pluripotent stem cells. Nat. Biotechnol. *43*, 1274–1287. https://doi.org/10.1038/s41587-024-02360-7.
4. de Bruijn, M., and Palis, J. (2024). Kaufman's Atlas of Mouse Development Supplement. Gillian Morriss-Kay and Shankar Srinivas) Ch. *9*, 145–157.
5. Palis, J., Chan, R.J., Koniski, A., Patel, R., Starr, M., and Yoder, M.C. (2001). Spatial and temporal emergence of high proliferative potential hematopoietic precursors during murine embryogenesis. Proc. Natl. Acad. Sci. USA *98*, 4528–4533. https://doi.org/10.1073/pnas.071002398.
6. McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep. *11*, 1892–1904. https://doi.org/10.1016/j.celrep.2015.05.036.
7. Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J.C.A., Azzoni, E., Woll, P.S., Mead, A.J., Hultquist, A., Swiers, G., et al. (2013). Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell *13*, 535–548. https://doi.org/10.1016/j.stem.2013.08.012.
8. Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073–5084. https://doi.org/10.1242/dev.126.22.5073.
9. Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzakt, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301. https://doi.org/10.1016/1074-7613(94)90081-7.
10. Medvinsky, A., and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. Cell *86*, 897–906. https://doi.org/10.1016/s0092-8674(00)80165-8.
11. de Bruijn, M.F., Speck, N.A., Peeters, M.C., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. EMBO J. *19*, 2465–2474. https://doi.org/10.1093/emboj/19.11.2465.
12. Taoudi, S., Gonneau, C., Moore, K., Sheridan, J.M., Blackburn, C.C., Taylor, E., and Medvinsky, A. (2008). Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ predefinitive HSCs. Cell Stem Cell *3*, 99–108. https://doi.org/10.1016/j.stem.2008.06.004.
13. Rybtsov, S., Sobiesiak, M., Taoudi, S., Souilhol, C., Senserrich, J., Liakhovitskaia, A., Ivanovs, A., Frampton, J., Zhao, S., and Medvinsky, A. (2011). Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J. Exp. Med. *208*, 1305–1315. https://doi.org/10.1084/jem.20102419.
14. Weijts, B., and Robin, C. (2024). Capturing embryonic hematopoiesis in temporal and spatial dimensions. Exp. Hematol. *136*, 104257. https://doi.org/10.1016/j.exphem.2024.104257.

15. Vink, C.S., and Dzierzak, E. (2023). The (intra-aortic) hematopoietic cluster cocktail: what is in the mix? Exp. Hematol. *118*, 1–11. https://doi.org/10.1016/j.exphem.2022.12.002.

16. Ganuza, M., Clements, W., and McKinney-Freeman, S. (2022). Specification of hematopoietic stem cells in mammalian embryos: a rare or frequent event? Blood *140*, 309–320. https://doi.org/10.1182/blood.2020009839.

17. Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., and Medvinsky, A. (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. Development *129*, 4891–4899. https://doi.org/10.1242/dev.129.21.4891.

18. Yokomizo, T., and Dzierzak, E. (2010). Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. Development *137*, 3651–3661. https://doi.org/10.1242/dev.051094.

19. Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887–891. https://doi.org/10.1038/nature07619.

20. Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321–330. https://doi.org/10.1016/s0092-8674(00)80986-1.

21. Wang, Q., Stacy, T., Miller, J.D., Lewis, A.F., Gu, T.L., Huang, X., Bushweller, J.H., Bories, J.C., Alt, F.W., Ryan, G., et al. (1996). The CBFbeta subunit is essential for CBFalpha2 (AML1) function in vivo. Cell *87*, 697–708. https://doi.org/10.1016/s0092-8674(00)81389-6.

22. North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marín-Padilla, M., and Speck, N.A. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development *126*, 2563–2575. https://doi.org/10.1242/dev.126.11.2563.

23. Cai, Z., de Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, J.R., and Dzierzak, E. (2000). Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity *13*, 423–431. https://doi.org/10.1016/s1074-7613(00)00042-x.

24. de Bruijn, M., and Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. Blood *129*, 2061–2069. https://doi.org/10.1182/blood-2016-12-689109.

25. Gao, L., Tober, J., Gao, P., Chen, C., Tan, K., and Speck, N.A. (2018). RUNX1 and the endothelial origin of blood. Exp. Hematol. *68*, 2–9. https://doi.org/10.1016/j.exphem.2018.10.009.

26. Ghozi, M.C., Bernstein, Y., Negreanu, V., Levanon, D., and Groner, Y. (1996). Expression of the human acute myeloid leukemia gene AML1 is regulated by two promoter regions. Proc. Natl. Acad. Sci. USA *93*, 1935–1940. https://doi.org/10.1073/pnas.93.5.1935.

27. Bee, T., Ashley, E.L.K., Bickley, S.R.B., Jarratt, A., Li, P.S., Sloane-Stanley, J., Göttgens, B., and de Bruijn, M.F.T.R. (2009). The mouse Runx1 +23 hematopoietic stem cell enhancer confers hematopoietic specificity to both Runx1 promoters. Blood *113*, 5121–5124. https://doi.org/10.1182/blood-2008-12-193003.

28. Bee, T., Liddiard, K., Swiers, G., Bickley, S.R.B., Vink, C.S., Jarratt, A., Hughes, J.R., Medvinsky, A., and de Bruijn, M.F.T.R. (2009). Alternative Runx1 promoter usage in mouse developmental hematopoiesis. Blood Cells Mol. Dis. *43*, 35–42. https://doi.org/10.1016/j.bcmd.2009.03.011.

29. Owens, D.D.G., Anselmi, G., Oudelaar, A.M., Downes, D.J., Cavallo, A., Harman, J.R., Schwessinger, R., Bucakci, A., Greder, L., de Ornellas, S., et al. (2022). Dynamic Runx1 chromatin boundaries affect gene expression in hematopoietic development. Nat. Commun. *13*, 773. https://doi.org/10.1038/s41467-022-28376-8.

30. Nottingham, W.T., Jarratt, A., Burgess, M., Speck, C.L., Cheng, J.F., Prabhakar, S., Rubin, E.M., Li, P.S., Sloane-Stanley, J., Kong-A-San, J., and de Bruijn, M.F.T.R. (2007). Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood *110*, 4188–4197. https://doi.org/10.1182/blood-2007-07-100883.

31. Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D., et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nat. Commun. *4*, 2924. https://doi.org/10.1038/ncomms3924.

32. Bee, T., Swiers, G., Muroi, S., Pozner, A., Nottingham, W., Santos, A.C., Li, P.S., Taniuchi, I., and de Bruijn, M.F. (2010). Nonredundant roles for Runx1 alternative promoters reflect their activity at discrete stages of developmental hematopoiesis. Blood *115*, 3042–3050. https://doi.org/10.1182/blood-2009-08-238626.

33. Schütte, J., Wang, H., Antoniou, S., Jarratt, A., Wilson, N.K., Riepsaame, J., Calero-Nieto, F.J., Moignard, V., Basilico, S., Kinston, S.J., et al. (2016). An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability. eLife *5*, e11469. https://doi.org/10.7554/eLife.11469.

34. Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R.S., Stehling-Sun, S., Sabo, P.J., Byron, R., Humbert, R., et al. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. Science *346*, 1007–1012. https://doi.org/10.1126/science.1246426.

35. Ng, E.S., Azzola, L., Bruveris, F.F., Calvanese, V., Phipson, B., Vlahos, K., Hirst, C., Jokubaitis, V.J., Yu, Q.C., Maksimovic, J., et al. (2016). Differentiation of human embryonic stem cells to HOXA(+) hemogenic vasculature that resembles the aorta-gonad-mesonephros. Nat. Biotechnol. *34*, 1168–1179. https://doi.org/10.1038/nbt.3702.

36. Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M., et al. (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat. Cell Biol. *18*, 595–606. https://doi.org/10.1038/ncb3354.

37. Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., et al. (2016). Tracing haematopoietic stem cell formation at single-cell resolution. Nature *533*, 487–492. https://doi.org/10.1038/nature17997.

38. Zhu, Q., Gao, P., Tober, J., Bennett, L., Chen, C., Uzun, Y., Li, Y., Howell, E.D., Mumau, M., Yu, W., et al. (2020). Developmental trajectory of prehematopoietic stem cell formation from endothelium. Blood *136*, 845–856. https://doi.org/10.1182/blood.2020004801.

39. Vink, C.S., Calero-Nieto, F.J., Wang, X., Maglitto, A., Mariani, S.A., Jawaid, W., Göttgens, B., and Dzierzak, E. (2020). Iterative Single-Cell Analyses Define the Transcriptome of the First Functional Hematopoietic Stem Cells. Cell Rep. *31*, 107627. https://doi.org/10.1016/j.celrep.2020.107627.

40. Rybtsov, S., Batsivari, A., Bilotkach, K., Paruzina, D., Senserrich, J., Nerushev, O., and Medvinsky, A. (2014). Tracing the origin of the HSC hierarchy reveals an SCF-dependent, IL-3-independent CD43(-) embryonic precursor. Stem Cell Rep. *3*, 489–501. https://doi.org/10.1016/j.stemcr.2014.07.009.

41. Stonehouse, O.J., Biben, C., Weber, T.S., Garnham, A., Fennell, K.A., Farley, A., Terreaux, A.F., Alexander, W.S., Dawson, M.A., Naik, S.H., and Taoudi, S. (2024). Clonal analysis of fetal hematopoietic stem/progenitor cells reveals how post-transplantation capabilities are distributed. Stem Cell Rep. *19*, 1189–1204. https://doi.org/10.1016/j.stemcr.2024.07.003.

42. Patel, S.H., Christodoulou, C., Weinreb, C., Yu, Q., da Rocha, E.L., Pepe-Mooney, B.J., Bowling, S., Li, L., Osorio, F.G., Daley, G.Q., and Camargo, F.D. (2022). Lifelong multilineage contribution by embryonic-born blood progenitors. Nature *606*, 747–753. https://doi.org/10.1038/s41586-022-04804-z.

43. Rodriguez-Fraticelli, A.E., Weinreb, C., Wang, S.W., Migueles, R.P., Jankovic, M., Usart, M., Klein, A.M., Lowell, S., and Camargo, F.D. (2020). Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. Nature *583*, 585–589. https://doi.org/10.1038/s41586-020-2503-6.

44. Sawai, C.M., Babovic, S., Upadhaya, S., Knapp, D.J.H.F., Lavin, Y., Lau, C.M., Goloborodko, A., Feng, J., Fujisaki, J., Ding, L., et al. (2016).

Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. Immunity *45*, 597–609. https://doi.org/10.1016/j.immuni.2016.08.007.

45. van Galen, P., Kreso, A., Wienholds, E., Laurenti, E., Eppert, K., Lechman, E.R., Mbong, N., Hermans, K., Dobson, S., April, C., et al. (2014). Reduced lymphoid lineage priming promotes human hematopoietic stem cell expansion. Cell Stem Cell *14*, 94–106. https://doi.org/10.1016/j.stem.2013.11.021.

46. Calvanese, V., Capellera-Garcia, S., Ma, F., Fares, I., Liebscher, S., Ng, E.S., Ekstrand, S., Aguadé-Gorgorió, J., Vavilina, A., Lefaudeux, D., et al. (2022). Mapping human haematopoietic stem cells from haemogenic endothelium to birth. Nature *604*, 534–540. https://doi.org/10.1038/s41586-022-04571-x.

47. Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., et al. (2019). The cis-Regulatory Atlas of the Mouse Immune System. Cell *176*, 897–912.e20. https://doi.org/10.1016/j.cell.2018.12.036.

48. Forsberg, E.C., Passegué, E., Prohaska, S.S., Wagers, A.J., Koeva, M., Stuart, J.M., and Weissman, I.L. (2010). Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. PLoS One *5*, e8785. https://doi.org/10.1371/journal.pone.0008785.

49. Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S., et al. (2014). Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. Cell Stem Cell *15*, 507–522. https://doi.org/10.1016/j.stem.2014.07.005.

50. Wilson, N.K., Kent, D., Buettner, F., Shehata, M., Macaulay, I., Calero-Nieto, F., Sánchez Castillo, M., Oedekoven, C., Diamanti, E., Schulte, R., et al. (2015). Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. Cell Stem Cell *16*, 712–724. https://doi.org/10.1016/j.stem.2015.04.004.

51. Acar, M., Kocherlakota, K.S., Murphy, M.M., Peyer, J.G., Oguro, H., Inra, C.N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K., and Morrison, S.J. (2015). Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. Nature *526*, 126–130. https://doi.org/10.1038/nature15250.

52. Kokkaliaris, K.D., Kunz, L., Cabezas-Wallscheid, N., Christodoulou, C., Renders, S., Camargo, F., Trumpp, A., Scadden, D.T., and Schroeder, T. (2020). Adult blood stem cell localization reflects the abundance of reported bone marrow niche cell types and their combinations. Blood *136*, 2296–2307. https://doi.org/10.1182/blood.2020006574.

53. Zape, J.P., Lizama, C.O., Cautivo, K.M., and Zovein, A.C. (2017). Cell cycle dynamics and complement expression distinguishes mature haematopoietic subsets arising from hemogenic endothelium. Cell Cycle *16*, 1835–1847. https://doi.org/10.1080/15384101.2017.1361569.

54. Levanon, D., and Groner, Y. (2004). Structure and regulated expression of mammalian RUNX genes. Oncogene *23*, 4211–4219. https://doi.org/10.1038/sj.onc.1207670.

55. Bellissimo, D.C., and Speck, N.A. (2017). RUNX1 Mutations in Inherited and Sporadic Leukemia. Front. Cell Dev. Biol. *5*, 111. https://doi.org/10.3389/fcell.2017.00111.

56. Frömel, R., Rühle, J., Bernal Martinez, A., Szu-Tu, C., Pacheco Pastor, F., Martinez-Corral, R., and Velten, L. (2025). Design principles of cell-state-specific enhancers in hematopoiesis. Cell *188*, 3202–3218.e21. https://doi.org/10.1016/j.cell.2025.04.017.

57. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., III, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

58. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296. https://doi.org/10.1186/s13059-019-1874-1.

59. Wolock, S.L., Lopez, R., and Klein, A.M.S. (2019). Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. *8*, 281–291.e289. https://doi.org/10.1016/j.cels.2018.11.005.

60. Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience *9*, giaa151. https://doi.org/10.1093/gigascience/giaa151.

61. van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell *174*, 716–729.e27. https://doi.org/10.1016/j.cell.2018.05.061.

62. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

63. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341. https://doi.org/10.1038/s41592-021-01282-5.

64. Gao, C.H., Chen, C., Akyol, T., Dusa, A., Yu, G., Cao, B., and Cai, P. (2024). ggVennDiagram: Intuitive Venn diagram software extended. iMeta *3*, e177. https://doi.org/10.1002/imt2.177.

65. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978. https://doi.org/10.1038/nmeth.4401.

66. Gu, Z. (2022). Complex heatmap visualization. iMeta *1*, e43. https://doi.org/10.1002/imt2.43.

67. Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018. https://doi.org/10.1093/bioinformatics/btr064.

68. Aaron R. Quinlan, Ira M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics, Volume 26, Issue 6, March 2010, Pages 841-842, https://doi.org/10.1093/bioinformatics/btq033

69. Kothary, R., Clapoff, S., Darling, S., Perry, M.D., Moran, L.A., and Rossant, J. (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. Development *105*, 707–714. https://doi.org/10.1242/dev.105.4.707.

70. Luis, T.C., Luc, S., Mizukami, T., Boukarabila, H., Thongjuea, S., Woll, P.S., Azzoni, E., Giustacchini, A., Lutteropp, M., Bouriez-Jones, T., et al. (2016). Initial seeding of the embryonic thymus by immune-restricted lympho-myeloid progenitors. Nat. Immunol. *17*, 1424–1435. https://doi.org/10.1038/ni.3576.

71. Yokomizo, T., Yamada-Inagawa, T., Yzaguirre, A.D., Chen, M.J., Speck, N.A., and Dzierzak, E. (2012). Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. Nat. Protoc. *7*, 421–431. https://doi.org/10.1038/nprot.2011.441.

72. Sheridan, J.M., Taoudi, S., Medvinsky, A., and Blackburn, C.C. (2009). A novel method for the generation of reaggregated organotypic cultures that permits juxtaposition of defined cell populations. Genesis *47*, 346–351. https://doi.org/10.1002/dvg.20505.

73. Azzoni, E., Frontera, V., Anselmi, G., Rode, C., James, C., Deltcheva, E.M., Demian, A.S., Brown, J., Barone, C., Patelli, A., et al. (2021). The onset of circulation triggers a metabolic switch required for endothelial to hematopoietic transition. Cell Rep. *37*, 110103. https://doi.org/10.1016/j.celrep.2021.110103.

74. Mass, E., Ballesteros, I., Farlik, M., Halbritter, F., Günther, P., Crozet, L., Jacome-Galarza, C.E., Händler, K., Klughammer, J., Kobayashi, Y., et al. (2016). Specification of tissue-resident macrophages during organogenesis. Science *353*, aaf4238. https://doi.org/10.1126/science.aaf4238.

75. Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Sharipov, R.N., Fedorova, A.D., Rumynskiy, E.I., Medvedeva, Y.A., Magana-Mora, A., Bajic, V.B., Papatsenko, D.A., et al. (2018). HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. *46*, D252–D259. https://doi.org/10.1093/nar/gkx1106.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD41-PE	BD Pharmigen	Cat # 558040; RRID: AB_397004
CD41-eFluor450	eBioscience	Cat # 48-0411-82; RRID: AB_1582238
CD41-PE-Cy7	eBioscience	Cat # 25-0411-80; RRID: AB_1234972
CD45-APCeFluor780	eBioscience	Cat # 47-0451-82; RRID: AB_1548781
CD45-PE	BD Pharmigen	Cat # 553081; RRID: AB_394611
CD45-BV510	Biolegend	Cat # 103138; RRID: AB_2563061
SAV-AF405	Invitrogen	Cat # S32351; RRID: N/A
SAV-APC	Invitrogen	Cat # SA1005; RRID: N/A
SAV-PE-Cy7	Invitrogen	Cat # SA1012; RRID: N/A
Ter119-BV501	Biolegend	Cat # 116237; RRID: AB_2561661
Ter119-PE-Cy7	BD Pharmigen	Cat # 557853; RRID: AB_396898
Ter119-APC-Cy7	BD Pharmigen	Cat # 560509; RRID: AB_1645230
Ter119-biotin	BD Pharmigen	Cat # 553672; RRID: AB_394985
CD144-APC	eBioscience	Cat # 17-1441-82; RRID: AB_10598508
CD117-APC	BD Pharmigen	Cat # 553356; RRID: AB_398536
CD117-APC-R700	BD Pharmigen	Cat # 565476; RRID: N/A
CD43-PE	eBioscience	Cat # 12-0431-82; RRID: AB_465659
Sca1-PE-Cy7	eBioscience	Cat # 25-5981-81; RRID: AB_469668
CD135-BV421	BD Biosciences	Cat # 562898; RRID: AB_2737876
CD127-PE	eBioscience	Cat # 12-1271-83; RRID: AB_465845
CD48-PE	eBioscience	Cat # 12-0481-82; RRID: AB_465694
CD150-PE-Cy7	eBioscience	Cat # 25-1502-82; RRID: AB_10805742
CD304-BV421	Biolegend	Cat # 145209; RRID: AB_2562358
CD31-BV711	Biolegend	Cat # 102449; RRID: AB_2860594
CD44-BV785	Biolegend	Cat # 103041; RRID: AB_11218802
CD146-PE-Cy7	Biolegend	Cat # 134713; RRID: AB_2563108
CD73-APC	Biolegend	Cat # 127209; RRID: AB_11219400
CD3-biotin	BD Pharmigen	Cat # 553060; RRID: AB_394593
B220-biotin	BD Pharmigen	Cat # 553086; RRID: AB_394615
Gr1-biotin	BD Pharmigen	Cat # 553125; RRID: AB_394641
CD11b-biotin	BD Pharmigen	Cat # 553309; RRID: AB_394773
CD41-TotalSeqB	Biolegend	Cat # 133941; RRID: AB_2876468
CD117-TotalSeqB	Biolegend	Cat # 105849; RRID: AB_2813938
CD44-TotalSeqB	Biolegend	Cat # 103071; RRID: AB_2832299
CD45-TotalSeqB	Biolegend	Cat # 103161; RRID: AB_2813931
CD127-TotalSeqB	Biolegend	Cat # 135055; RRID: AB_2860677
CD115-TotalSeqB	Biolegend	Cat # 135543; RRID: AB_2832487
Ter119-TotalSeqB	Biolegend	Cat # 116251; RRID: AB_2832398
CD304-TotalSeqB	Biolegend	Cat # 145219; RRID: AB_2876495
F4/80-TotalSeqB	Biolegend	Cat # 123155; RRID: AB_2819847
CD73-TotalSeqB	Biolegend	Cat # 127235; RRID: AB_2860664
goat anti-mouse/rat/human CD31	R&D Systems	Cat # AF3628; RRID: AB_2161028
rat anti-mouse CD31	BD Pharmgen	Cat # 553370; RRID: AB_396660
rabbit anti-mouse/rat/human Runx1	Abcam	Cat # ab92336; RRID: AB_2049267

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
chicken anti-GFP	Abcam	Cat # ab13970; RRID: AB_300798
donkey anti-chicken IgG (H + L)	Jackson immuno research	Cat # 703-545-155; RRID: AB_2340375
donkey anti-rabbit IgG (H + L)	Invitrogen	Cat # A31572; RRID: AB_162543
donkey anti-goat IgG (H + L)	Invitrogen	Cat # A21447; RRID: AB_141844
Goat anti-rat IgG (H + L)	Southern Biotech	Cat # 3050-04; RRID: AB_2795829
TruStain FcX™ PLUS	Biolegend	Cat# 156603; RRID: N/A
Biological samples		
Mouse embryos	This paper	N/A
Mouse adult bone marrow	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Recombinant IL-3	Peprotech	Cat # AF-213-13
Recombinant SCF	Peprotech	Cat # 250-03
Recombinant Flt3 ligand	Peprotech	Cat # 250-31L
Hoechst 33258	Merck	Cat# 861405
7-aminoactinomycin D (7-AAD)	ThermoFisher Scientific	Cat# A1310
Hyclone FCS	Fisher Scientific	Cat# SH30070.03HI
Collagenase I	Merck	Cat# C0130
Hot StartTaq Master Mix	QIAGEN	Cat# 203445
Iscove's Modified Dulbecco's Medium (IMDM)	ThermoFisher Scientific	Cat# 12440053
Critical commercial assays		
MethoCult GF M3434	STEMCELL Technologies	Cat# 03434
Chromium Next GEM Single Cell 3′ v3.1 Kit	10*x* Genomics	Cat# 1000269
Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Kit	10*x* Genomics	Cat# 1000283
Dual Luciferase Reporter Assay System	Promega	Cat# E1910
Deposited data		
Single-cell RNA-seq analysis of embryonic AGM+VU from Runx1 enhancer-reporter transgenic mice	This paper	ArrayExpress: E-MTAB-15202
Single-nuclei Multiome (RNA+ATAC) analysis of embryonic AGM+VU from Runx1 +23 enhancer-reporter transgenic mouse	This paper	ArrayExpress: E-MTAB-15412
Experimental models: Cell lines		
OP9	Laboratory of Alexander Medvinsky	N/A
Experimental models: Organisms/strains		
Mouse: 23GFP	Our laboratory	Bee et al., 2010; Swiers et al., 2013
Mouse: 110GFP	This paper	N/A
+3LacZ transient transgenic	Our laboratory	Schutte et al., 2016
+110LacZ transient transgenic	Our laboratory	Schutte et al., 2016
Oligonucleotides		
+110_F: GACTCGAGTGCACACAAACACACACACC	IDT	This paper
+110_R: GAGTCGACCTTCTCTCAGCTTGCCCATC	IDT	This paper
LacZ_F: GCAGATGCACGGTTACGATG	IDT	Nottingham et al., 2007
LacZ_R: GTGGCAACATGGAAATCGCTG	IDT	Nottingham et al., 2007
GFP_F: GACGTGAACGGCCACAAGTTCA	IDT	Bee et al., 2010
GFP_R: GTGGCGGATCTTGAAGTTCACC	IDT	Bee et al., 2010
Software and algorithms		
FACS Diva	BD	RRID: SCR_001456
FlowJo	BD	RRID: SCR_008520

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
IMARIS	Bitplane	RRID: SCR_007370
GraphPad Prism	GraphPad Software	RRID: SCR_002798
Photoshop	Adobe	RRID: SCR_014199
Illustrator	Adobe	RRID: SCR_010279
Zeiss Zen	Zeiss	https://www.zeiss.com/microscopy/int/products/microscope-software/zen.html
R version 4.1.3 and 4.3.2	The R Foundation	https://www.r-project.org
Seurat version 4.0.4 and 5.0.1	Hao et al.,[57]	https://satijalab.org/seurat/
SCTransform	Hafemeister et al.[58]	N/A
CellRanger version 5.0.0	10x Genomics	https://support.10xgenomics.com
Scrublet	Wolock et al.[59]	https://github.com/swolock/scrublet
SoupX	Young et al.[60]	https://github.com/constantAmateur/SoupX
MAGIC	Van Dijk et al.[61]	https://github.com/KrishnaswamyLab/MAGIC/
Cell Ranger ARC version 2.0.2	10x Genomics	https://support.10xgenomics.com
MACS2	Zhang et al.[62]	N/A
Signac version 1.12.0	Stuart et al.[63]	https://stuartlab.org/signac/
ggVennDiagram version 1.5.7	Gao et al.[64]	https://gaospecial.github.io/ggVennDiagram/
chromVAR	Schep et al.,[65]	https://github.com/GreenleafLab/chromVAR
ComplexHeatmap package version 2.28.0	Gu et al.,[66]	https://github.com/jokergoo/ComplexHeatmap
FIMO (MEME suite version 5.5.3)	Grant et al.,[67]	https://meme-suite.org/meme/tools/fimo
Bedtools version 2.29.2	Quinlan et al.,[68]	N/A
Other		
BD FACSAria™ Fusion II	BD Biosciences	N/A
BD LSRFortessa™ cell analyzer	BD Biosciences	N/A
NucleoCounter® NC-3000™	Chemometec	N/A
Membrane Filter, Pores 0.8 μm, 25 mm Diameter	Merck	Cat# AAWP02500

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

Genomic fragments spanning the *Runx1* enhancers of interest were generated by PCR (primer sequences in Table S1). Fragments were cloned downstream of a *LacZ or Gfp* gene in a reporter vector driven by the hsp68 core promoter.[69] Mouse transient transgenic embryos and transgenic mouse lines were generated by pronuclear injection of (CBA x C57BL/6)F1 zygotes following standard procedures. Transient transgenic embryos carrying *Runx1* enhancer-*LacZ* reporter construct were identified by *LacZ* specific PCR (primer sequences in Table S1) on genomic DNA isolated from yolk sac or ectoplacental cone biopsies. Transgenic founder mice carrying *Runx1* enhancer-*Gfp* reporter construct were identified by genomic PCR for *Gfp* on DNA isolated from tail or ear biopsies (primers in Table S1). Four transgenic 110GFP F0 founder mice were generated. Transgenic reporter gene expression was analyzed in F1 embryos to identify those with GFP expression in a *Runx1*-specific pattern, and in line with the pattern observed in the corresponding enhancer-*LacZ* reporter transient transgenic embryos. A stable 110GFP transgenic line carrying <15 copies of the transgene was generated. All mice were housed with free access to food and water in an enhanced environment. All procedures were in compliance with United Kingdom Home Office regulations and the Oxford University Clinical Medicine Animal Welfare and Ethical Review Committee.

METHOD DETAILS

Timed matings and embryo collection

For timed pregnancies, (CBA x C57BL/6)F1 females (5–16 weeks old) were mated overnight with transgenic males (maintained on a mixed CBA x C57BL/6 background). Embryos were collected in $CaCl_2$ and $MgCl_2$-containing PBS (Gibco) supplemented with 10% FCS (Gibco), 50 U/mL penicillin, 50 mg/mL streptomycin (Cambrex Corporation). Embryos were staged by counting somite pairs (E8 to E10.5) or tail somite pairs (E11.5).[70] 110GFP transgenic embryos were identified by fluorescence illumination (X-Cite 120; Improvision) on a Leica MZFLIII microscope.

Analysis of enhancer-lacZ transient transgenic embryos

Transient transgenic embryos were analyzed as described previously[30]. Briefly, embryos were fixed in PBS containing 1% formaldehyde, 0.2% glutaraldehyde, 2mM $MgCl_2$, 5mM EGTA (pH 8.0), and 0.02% NP-40, washed in PBS 0.02% NP-40, and stained in Xgal staining solution (5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$ 3 H_2O, 2mM $MgCl_2$, 0.01% sodium deoxycholate, 0.02% NP-40, and 1 mg/mL Xgal in PBS). Embryos were washed in PBS, postfixed, incubated in 15% sucrose in PBS, and embedded in Tissue-Tek OCT compound (Sakura, Siemens Medical Solutions Diagnostics). Transverse sections (8-10 μm) were cut on a Leica CM3050S Cryostat (Leica Microsystems), collected on SuperFrost Plus microscope slides (VWR International) and coverslipped using Kaiser glycerol gelatin (VWR International). Embryo sections were examined, and pictures taken on a Nikon Eclipse E600 brightfield microscope with either a Nikon PlanFluor 40*x*/0.75 or 60*x*/1.40 objective and using the Nikon Digital Camera DXM 1200c setup.

Dissections and generation of cell suspensions

Cell suspensions from dissected embryos were prepared as described.[31] Briefly, E8.5 concepti, E9.5 caudal parts (CP) and yolk sac (YS) and E10.5/E11.5 AGM with vitelline and umbilical arteries attached (AVU), were harvested and dissected in $CaCl_2$ and $MgCl_2$-containing PBS supplemented with 10% FCS (Gibco),
50U/mL penicillin, and 50μg/mL streptomycin (Cambrex Corporation) using forceps (Watchmakers) and 25G needles (detailed sample information available in Table S6). Tissues were pooled and incubated in 0.12% (w/v) of collagenase (Type I, Sigma) in PBS without $CaCl_2$ and $MgCl_2$ supplemented with 10% FCS (Gibco),
50U/mL penicillin, and 50μg/mL streptomycin (Cambrex Corporation), dissociated by pipetting and washed. E14 fetal liver was dissociated by pipetting and washed. Viable cells were counted using Neubauer hemocytometer and trypan blue (Invitrogen) or NucleoCounter NC-3000TM automated cell counter (Chemometec).

Flow cytometry and cell sorting

FACS analysis was performed as previously described.[31] In brief, embryonic cells were isolated using a 100 μm nozzle on a BD FACSAria Fusion II instrument. Analysis was performed on CyAn ADP (Beckman Coulter) or BD LSRFortessa cell analyzer. Antibody staining was carried out in PBS (without $CaCl_2$ and $MgCl_2$) supplemented with 10% FCS (Gibco),
50U/mL penicillin, and 50μg/mL streptomycin (Cambrex Corporation). After sorting, cells were collected in tubes containing 100% FCS or sorted directly into OP9 co-culture media. Antibodies (Table S2) were titrated to determine their optimal concentration of use and gates were defined using unstained, single stained and fluorescence-minus-one stained cells. Dead cells were excluded based on Hoechst 33258 (Invitrogen) or 7AAD (eBioscience) uptake. Data were acquired and analyzed using FACS Diva (BD) and FlowJo software.

Immunofluorescence analysis and imaging

Sections for immunofluorescence analysis were processed as previously described.[31] In brief, embryos were fixed in 4% paraformaldehyde, washed in PBS and soaked in 15% (w/v) sucrose. Samples were snap-frozen in Tissue-Tek OCT compound (Sakura, Siemens Medical Solutions Diagnostics), sectioned at 10μm (Leica CM3050s) and collected onto glass slides (Superfrost plus, VWR). Sections were block permeated and stained in 5% FCS in 0.4% Triton-X/PBS (primary and secondary antibody list in Table S2). Slides were mounted with Vectashield hard-set mounting medium containing DAPI (Vector Laboratories). Image acquisition was performed using a Zeiss 510 and Zeiss 780 Inverted confocal microscopes. 25*x* (LD LCI PA 25*x*/0.8 DIC), 40*x* (LD C-Apochromat 40*x*/1.1) and 63*x* (Plan Apochromat 63*x*/1.4) objectives were used with oil immersion. Images were processed using Adobe Photoshop. Wholemount immunofluorescence was performed according to Yokomizo et al.[71] In brief, the head, tail, limbs and YS of E10.5 embryos were removed and embryos fixed in 4% paraformaldehyde and washed in PBS. For long-term storage, 50% and 100% MeOH washes were performed and embryos stored in 100% MeOH at −20°C. Prior to staining, embryos were washed in 50% MeOH and PBS and incubated in individual Eppendorf tubes in PBS 5–10% FCS 0.4% Triton-X on a shaking rack. Primary and secondary staining (Table S2) were performed in 100-200μL of PBS 5–10% FCS 0.4% Triton-X overnight at 4°C on a shaking rack. In between staining and after secondary antibody incubation, wholemounts were washed in the same buffer. Embryos were dehydrated in MeOH, cleared in BABB (1:2 benzyl alcohol:benzyl benzoate) and mounted into Fastwells (Grace Bio-Labs). Image acquisition was performed overnight using a Zeiss LSM 780 Upright confocal microscope with a 25× oil immersion objective (LD LCI PA 25×/0.8 DIC) and the ZEN 2011 LSM (black) software. Z-stacks and tiles of wholemount images were then 3D reconstructed in IMARIS (Bitplane) or OMERO software. Images were analyzed with IMARIS (Bitplane) and Adobe Photoshop was used for final image processing.

Hematopoietic progenitor assay

Colony-forming unit-culture (CFU-C) assays were performed using Methocult M3434 (Stem Cell Technologies). Cells were plated in duplicate in 35mm culture dishes according to manufacturer's instructions. Cultures were grown at 37°C with 5% CO2 and colonies counted after 7 days (detailed sample information in Table S6).

Long-term HSC reconstitution analysis

Male mice (C57BL/6xCBA)/F1 up to 3 months old were sub-lethally irradiated with a split dose of 9Gy from a 137Cs source. Donor cells were obtained from either E11.5 AVU or from OP9 co-aggregates and injected together with 2×10^5 spleen carrier cells into the

tail veins of irradiated recipients (detailed sample information available in Table S6). The level of reconstitution in peripheral blood at 4 and 16 weeks post-transplant was determined by Gfp-specific PCR (Table S1) and quantitative analysis (Image Lab, Bio-Rad) of genomic DNA extracted from peripheral blood using the DNeasy Blood & Tissue kit (Qiagen) according to manufacturer's instructions. Percentage reconstitution was calculated based on comparison to standards containing 0%, 1%, 10%, 50%, and 100% of donor genomic DNA.

OP9 maintenance and co-aggregate cultures

OP9 cells were maintained as previously described.[28] Sorted cell populations from pools of E10.5/E11.5 AVU (detailed sample information available in Table S6) were assayed as previously described.[12,72,73] In brief, a cell mixture containing AGM cells (1ee) and 10^5 OP9 was suspended in Iscove's modified Dulbecco's medium (IMDM, Gibco), aspirated in a 200mL pipette tip sealed with paraffin, and centrifuged at 400g for 5 min at RT. Co-aggregates were transferred onto a floating 0.8μm nitrocellulose membranes (Millipore) and cultured at the liquid–gas interface in IMDM (Gibco) supplemented with 100ng/mL IL-3, 100ng/mL SCF, 100ng/mL Flt3 ligand (Peprotech), L-glutamine and penicillin/streptomycin (Gibco). After 5 days, aggregates were dissociated and analyzed as previously described.[12]

Single-cell RNA-seq cell isolation and library preparation

AVU from pooled E10 23GFP Tg (9 embryos, 32–35 sp) and 110GFP Tg (5 embryos, 32–34 sp) were dissected and processed into single cell suspension as described above (detailed sample information in Table S6). Cells were incubated in TruStain FcX PLUS blocking reagent (Biolegend) and TotalSeqB antibody cocktail (Biolegend) (Table S2), as per manufacturer's instructions. 23GFP+ and 110GFP+ cells were FACS-sorted using a BD FACSAria Fusion sorter (100μm nozzle) and cell viability was assessed (99% and 95% live cells for 23GFP and 110GFP, respectively) before loading each sample onto independent channels of a Chromium chip for single cell partitioning and barcoding on Chromium Controller (10*x* Genomics). Single-cell cDNA synthesis and library preparation were performed using a Chromium Next GEM Single Cell 3′ v3.1 Dual Index Kit (10*x* Genomics) as per manufacturer's instructions and sequenced on Illumina NovaSeq 6000.

Single-cell RNA-seq analysis

Raw sequencing data were processed using Cell Ranger Software (version 5.0.0, 10*x* Genomics) and a customised version of the GRCm39 mouse genome assembly where the sequence for the Gfp gene was added using the cellranger mkref pipeline (10*x* Genomics). After sequencing, 5,751 cells were recovered for 23GFP AVU and 2,997 for 110GFP AVU. Ambient RNA, mostly consisting in low expression of globin genes, was removed using SoupX package[60] (contamination fraction was selected manually as 10% for 23GFP AVU and 20% for 110GFP AVU). Low quality cells and empty droplets were filtered out based on the number of detected genes, UMI and percentage of mitochondrial genes (<5%) and doublets were identified and excluded using Scrublet.[59] Downstream analysis was performed using Seurat package (version 4.0.4)[57] in R (version 4.1.3). Each sample was normalised and scaled independently using SCTransform[58] (including cell cycle-associated genes, UMI counts and mitochondrial genes regression) and samples were integrated using Seurat RPCA approach (k = 5). When Seurat subset function was used to remove contaminating clusters or select specific populations in the original dataset, SCTransform and RPCA integration were performed again on the selected subsets using the same parameters. To improve cluster annotation, MAGIC[61] was used to impute missing values and reduce dropout effects. Clusters were annotated based on differentially expressed gene (DEG) analysis (using Seurat FindAllMarkers, min.pct = 0.25, logfc.threshold = 0.25) and the expression of known cell type-specific marker genes. The FindSubCluster function was used when marker gene expression suggested further heterogeneity within a defined cluster. Finally, cluster annotation was confirmed by visualising the expression of gene signatures from publicly available datasets including embryonic hematopoietic tissues[37–39,42,74] using the AddModuleScore function (ctrl = 100).

HSC potential gene signature

To identify marker genes specifically expressed in HSC^{Pot} when compared to HE and HSC^{Not}, DEG analysis was performed using Seurat FindAllMarkers function (3-way comparison) and only genes that showed a minimum difference of 25% in the fraction of detection (min.diff.pct = 0.25) and an average log2 fold-change of at least 0.5 (logfc.threshold = 0.5) between groups were tested.

Single-cell multiome cell isolation and library preparation

AVU from pooled E9 (33 embryos, 20–25 sp) and E10 (24 embryos, 32–36 sp) 23GFP Tg were dissected, processed into single cell suspension and stained for FACS sorting, as described above (detailed sample information available in Table S6). Ter119-23GFP+ cells were FACS-sorted using a BD FACSAria Fusion sorter (100μm nozzle) and nuclei were isolated as 10*x* Genomics demonstrated protocols (CG000365). Nuclei quality was assessed, and nuclei were counted before loading each sample onto independent channels of a Chromium chip for single cell partitioning and barcoding on Chromium Controller (10*x* Genomics). Single-cell cDNA synthesis and library preparation were performed using a Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Kit (10*x* Genomics) as per manufacturer's instructions and sequenced on Illumina NovaSeq 6000.

Single-cell multiome analysis

Raw sequencing data were processed using Cell Ranger ARC (version 2.0.2, 10*x* Genomics) and GRCm38 mouse genome assembly. Downstream analysis was performed using Signac (version 1.12.0)[63] and Seurat packages (version 5.0.1)[57] in R (version 4.3.2). Per-cell quality control metrics were computed on the merged samples for both RNA (1000 < detected genes <7000, 1000 < UMI< 70000 and mitochondrial genes <20%) and ATAC (2.5 < TSS enrichment score <10, Nucleosome signal <1.5 and 1500 < ATAC fragments <25000) assays, and low-quality nuclei were filtered out. Gene expression data of the 7349 nuclei that passed QC were normalised and scaled using SCTransform[58] (including cell cycle-associated genes, UMI counts and mitochondrial genes regression). Clusters were identified using the FindClusters function (resolution = 1) in Seurat and dimension reduction was performed using RunUMAP function (dims = 1:50). Clusters were annotated based on differentially expressed gene (DEG) analysis (using Seurat FindAllMarkers, min.pct = 0.25, logfc.threshold = 0.25) and the expression of known cell type-specific marker genes. ATAC-seq peaks were identified using MACS2[62] and the CallPeaks function in Signac (additional.args = "-q 0.01"). DNA accessibility data were processed by performing latent semantic indexing (LSI) in Signac and dimension reduction was performed using the RunUMAP function (reduction = LSI, dims = 2:50). Joint UMAP visualisation of RNA and ATAC assays was computed using the weighted nearest neighbor method (WNN) in Seurat. When Seurat subset function was used to select Hoxa+ (hoxa7, hoxa9 and hoxa10) hematopoietic clusters in the original dataset, both RNA and ATAC processing was performed again using the same parameters. Peak-to-gene linkage on analysis on Hoxa+ clusters was performed using the LinkPeaks function in Signac. HSC-potential signature was calculated as described above. Differentially accessible peaks were calculated using FindMarkers function (with default arguments apart from test.use = "LR"; latent.vars = "nCount_peaks", "TSS.enrichment", "percent.mt", "sample"; min.pct = 0.01), by comparing each cluster of interest (i.e., HE, preHSC and HPC) to the others (1 vs. 2 comparison). The resulting sets of peaks were further filtered for the ones showing a minimum difference of 10% in the fraction of detection (pct.1 – pct.2 > 0.1) and an average log2 fold-change >1. Over-represented motifs in each cluster-specific set of peaks were calculated using Signac FindMotifs function and results were visualised using ggVennDiagram (version 1.5.7).[64] Per-cell motif score was calculated using chromVAR[65] and the activity of the top 10 enriched motifs for each cluster of interest were visualise using ComplexHeatmap package (version 2.28.0).[66]

Motif analysis of Runx1 +3 and +110 enhancers

Genomic regions corresponding to the ATAC peaks of the Runx1 +3 and +110 enhancers were searched for known transcription factor motifs using FIMO[67] from the MEME suite (version 5.5.3) and the HOCOMOCOv11_core_MOUSE database.[75] Mean conservation score for each motif was calculated by intersecting their sequence with base wise conservation scores (phyloP) of 59 vertebrate genomes with mouse GRCm38 from UCSC, using Bedtools (version 2.29.2). Significant hits resulting from FIMO analysis were then filtered by i) mean conservation score >0; ii) presence within the highly conserved region of each enhancer (i.e., conserved in mouse, human, dog and opossum) and iii) expression in the cell type of interest.

Luciferase reporter assay

Luciferase assays were performed as described[30] with minor modifications. Briefly, 10^7 416B cells and 5×10^6 HPC7 cells were electroporated with test plasmids (10μg and 5μg, respectively) at 220V and 960μFD. After 24 h, relative luciferase activity was measured using the Dual Luciferase Reporter Kit (Promega) and a FLUOstar Optima luminometer (BMG Labtech, Aylesbury, United Kingdom). Transfection efficiency correction was performed using the ratio of firefly over renilla luciferase activity.

QUANTIFICATION AND STATISTICAL ANALYSIS

Results are presented as average±SD. Exact *n* value and detailed sample information are available in Table S6. Unpaired, two-tailed Student's *t* test with Welch's correction was used to determine the level of significance. The following *p*-values were considered statistically significant and were highlighted by asterisks: * $p < 0.05$; ** $p < 0.01$.