

## **Clonal analysis of fetal hematopoietic stem/progenitor cell subsets reveals how post-transplantation capabilities are distributed**

Olivia J Stonehouse<sup>1,2,6</sup>, Christine Biben<sup>1,2</sup>, Tom S Weber<sup>1,2</sup>, Alexandra Garnham<sup>1,2</sup>, Katie A Fennell<sup>3</sup>, Alison Farley<sup>1,2</sup>, Antoine F Terreaux<sup>1,2</sup>, Warren S Alexander<sup>1,2</sup>, Mark A Dawson<sup>3, 4, 5</sup>, Shalin H Naik<sup>1,2</sup>, and Samir Taoudi<sup>1,2,7\*</sup>

<sup>1</sup> The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia.

<sup>2</sup> Department of Medical Biology, The University of Melbourne, Australia.

<sup>3</sup> Peter MacCallum Cancer Centre, Melbourne, Australia.

<sup>4</sup> Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia.

<sup>5</sup> The University of Melbourne Centre for Cancer Research, The University of Melbourne, Melbourne, Victoria, Australia.

<sup>6</sup> Lowy Cancer Research Centre, University of New South Wales, Australia.

<sup>7</sup> School of Cellular and Molecular Medicine, University of Bristol, Bristol, United Kingdom.

\*Correspondence: samir.taoudi@bristol.ac.uk

Running title: Single-cell analysis of fetal stem/progenitors

## Abstract

It has been proposed that adult haematopoiesis is sustained by multipotent progenitor (MPP) clones that are specified during development. From an immunophenotypic perspective, it is known that hematopoietic stem cell (HSC) and MPPs are present in the fetal liver yet our understanding of how fetal MPPs functionally compare to those in the adult bone marrow is incomplete. Using acute-term transplantations, we found that at a population-level fetal immunophenotypic MPP classes exhibited similar lineage biases as adult cells, albeit with some difference in lymphoid output. Clonal assessment of fetal MPPs engraftment revealed that lineage biases largely resulted from differences in the pattern of single- or bi-lineage differentiation. Immunophenotypic long-term (LT)- and short-term (ST)-HSCs in the fetal liver were distinguished from MPPs according to propensity for clonal multi-lineage differentiation. We also discovered that a large cohort of long-term repopulating units (LT-RU) were within the immunophenotypic ST-HSC population, a significant portion of these were labelled using *Flt3*-cre. This finding has two implications: (1) use of the CD150+ LT-HSC immunophenotype alone will systematically underestimate the size and diversity of the fetal LT-RU pool; and, (2), given fetal LT-RUs with a ST-HSC immunophenotype have the functional attributes required to persist into adulthood.

## Introduction

Hematopoiesis ensures the continuous supply of mature blood lineages. Whether the adult hematopoietic hierarchy is stem cell-driven or is a process sustained by multipotent progenitor cells (MPPs) but underwritten by hematopoietic stem cells (HSCs) is a contentious issue (Busch et al., 2015; Patel et al., 2022; Pei et al., 2017; Pei et al., 2020; Rodriguez-Fraticelli et al., 2018; Sawai et al., 2016; Sawen et al., 2018; Schoedel et al., 2016; Sheikh et al., 2016; Sun et al., 2014).

Based on the ability to provide durable, high-level, and multilineage reconstitution following transplantation, HSCs are traditionally thought to be at the foundation of the hierarchy. Adult bone marrow (ABM) HSCs consist of subsets that can be classified according to durability of self-renewal and by hematopoietic lineage production (Dykstra et al., 2007; Eaves, 2015; Oguro et al., 2013; Pietras et al., 2015): myeloid-biased  $\alpha$ HSCs and lineage balanced  $\beta$ HSCs exhibit extensive self-renewal in serial transplantation assays;  $\gamma$ HSCs and  $\delta$ HSCs are both lymphoid-biased and fail to reconstitute hematopoiesis following serial transplantation (Dykstra et al., 2007; Muller-Sieburg et al., 2002). In contemporary studies, the most common convention used to resolve HSCs and MPPs is variations of the lineage marker (LIN)- SCA1+ KIT+ (LSK) and SLAM systems, known as the LSK-SLAM code (Kiel et al., 2005; Pietras et al., 2015). Long-term reconstituting HSCs (LT-HSCs) are recognized as LSK\_FLT3-CD150+CD48- cells, and short-term reconstituting HSCs (ST-HSCs) are LSK\_FLT3-CD150-CD48-(Kiel et al., 2005; Pietras et al., 2015). Most  $\alpha$ HSCs/ $\beta$ HSCs are CD150+, and  $\gamma$ HSCs/ $\delta$ HSCs are CD150- (Kent et al., 2009).

Downstream of HSCs are the MPPs. ABM MPPs are capable of acute term multilineage differentiation when transplanted into irradiated adult mice (Adolfsson et al., 2001; Kiel et al., 2005; Oguro et al., 2013; Pietras et al., 2015). Although the number of MPP subsets that can be immunophenotypically distinguished in the ABM is increasing (Mansson et al., 2007; Naik et al., 2013; Oguro et al., 2013; Pietras et al., 2015; Sommerkamp et al., 2021), there are three commonly studied MPPs subclasses (MPP2, MPP3, and MPP4). Largely based on studies in the adult, each of the MPP 2 – 4 subsets are capable of lymphoid, myelo-erythroid, and platelet production but exhibit clear lineage biases (Adolfsson et al., 2001; Adolfsson et al., 2005; Oguro et al., 2013; Pietras et al., 2015): MPP2 (LSK\_FLT3-CD150+CD48+) and MPP3 (LSK\_FLT3-CD150-CD48+) possess limited lymphoid potential; MPP2 exhibits erythroid and megakaryocyte/platelet bias;

MPP3 exhibits granulocytic bias; and, MPP4 (LSK\_FLT3+CD150-CD48+) exhibits a lymphoid bias. *In situ* barcoding studies suggest that long-term self-renewing MPPs are the drivers of native hematopoiesis because the barcodes present in mature lineages are found in the MPPs but not in LT-HSCs (Patel et al., 2022; Pei et al., 2017; Sun et al., 2014).

Understanding the biology of the HSPC lineages during embryogenesis is important because LT-HSCs (de Bruijn et al., 2000; Ema and Nakauchi, 2000; Ganuza et al., 2017; Gekas et al., 2005; Kumaravelu et al., 2002; Medvinsky and Dzierzak, 1996; Ottersbach and Dzierzak, 2005; Taoudi et al., 2008) and long-lived MPP clones (Patel et al., 2022) emerge during this period. Although the first long-term repopulating units (LT-RUs, a functional designation rather than an immunophenotypic description) express KIT (Sanchez et al., 1996) and SCA1 (de Bruijn et al., 2002) they lack CD150 expression (McKinney-Freeman et al., 2009). By E14.5, CD150 expression enables a high-frequency enrichment of LT-RUs in the fetal liver (Kim et al., 2006), the majority of these are of the  $\beta$ HSC (lineage balanced) class (Benz et al., 2012). Cells with the ST-HSC immunophenotype (CD150-) are present in the E14.5 FL (Kim et al., 2006; Patel et al., 2022), but their functional capacity has not been thoroughly investigated. Although it is known that long-term hematopoietic reconstitution is possible from CD150- cells in the E14.5 FL (Kent et al., 2009; Kim et al., 2006; Papathanasiou et al., 2009), how much of the LT-RU biomass they represent is not known. This is an important issue that requires resolution because of how ubiquitously CD150 expression is used as a hurdle criterion for the quantitative and qualitative analysis of the fetal LT-RU pool (for examples see (Che et al., 2022; Ganuza et al., 2022; Gao et al., 2022; Hall et al., 2022; Lee et al., 2022; Patel et al., 2022; Sakai et al., 2022; Van Deren et al., 2022; Young et al., 2021)).

It has been proposed that the first phase of HSC-derived blood production could be contributed to by developmentally restricted HSCs (drHSCs) (Beaudin et al., 2016). drHSCs derive along a *Flt3-Cre*-expressing ancestry, express the conventional LT-HSC immunophenotype (are LSK CD150+), and provide durable but strongly lymphoid-biased reconstitution following transplantation. drHSCs differ from *bona fide* LT-HSCs according to their ancestry, and the transient nature of their hematopoietic contribution under physiological developmental conditions

(Beaudin et al., 2016). One interpretation is that the drHSC subset of the fetal liver LT-HSCs could represent the prenatal equivalent of ST-HSCs (Patel et al., 2022).

MPP4-like features such as transcriptional lineage priming and lymphoid-bias have been reported in the E14.5 fetal liver (Benz et al., 2012; Kim et al., 2006; Mansson et al., 2007). Functional testing of the ABM MPP2 immunophenotype in cells from the E16.5 and E18.5 FL showed evidence of transient multilineage reconstitution (Hall et al., 2022). How functionally comparable the FL and ABM MPP immunophenotypes are remains incompletely understood.

To better understand the functional capability of fetal hematopoietic stem and progenitor cells, we used single cell RNA-Sequencing (scRNA-Seq), cellular barcoding, and transplantation experiments to investigate the E14.5 fetal liver LSK subsets. We found that functionally diverse MPP subtypes do exist within the embryo, and that the adult MPP markers broadly enrich for the same functional groups in the fetus. Clonal tracking following transplantation into adult mice revealed that although fetal MPPs were capable of multi-lineage outcomes in the majority of cases MPPs underwent single- or dual-lineage hematopoietic contribution. Most intriguingly, we discovered that two variants of *bona fide* LT-RUs (without overt lineage bias) co-exist in the fetal liver: one is within the LT-HSC immunophenotype and the other within the ST-HSC immunophenotype. Although the LT-RU biomass is distributed between both HSC immunophenotypes, the largest fraction is contained within the ST-HSC immunophenotype. Using *Flt3*-cre lineage tracking we found that ST-HSC immunophenotype was more significantly labelled than LT-HSCs. This suggests that multiple LT-RU forming pathways exist to establish the fetal HSC pool. Consequently, any strategy based on enriching LT-RUs without including those within the CD150-negative fraction will not only omit a large fraction of LT-RUs, but could be missing an important piece of the biological landscape of stem cell pool.

In combination with the insight that the FL ST-HSC population is developmentally distinguished from LT-HSCs according to *Flt3*-expressing ancestry and that embryonic *Flt3*-cre labelled cells contribute to native adult hematopoiesis (Patel et al., 2022), our findings suggest that LT-RUs present in the FL ST-HSC population could persist into adulthood and contribute to native hematopoiesis.

## Materials and Methods

### ***Mice***

Ly5.1, Ly5.1 x C57BL/6, *Rosa26\_floxed-EYFP* (Srinivas et al., 2001), UBC-GFP (Schaefer et al., 2001), and *Flt3-cre* (Benz et al., 2012) mice were maintained on a C57BL/6 background. For timed matings, noon of the day of a positive vaginal plug check was considered embryonic day (E) 0.5. Animal experiments were approved by the WEHI animal ethics committee.

### ***Fetal tissue collection***

The uterus/extra-embryonic tissues were washed with  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -containing DPBS (supplemented with 7% fetal calf serum, 100 units/ml penicillin-streptomycin). Embryos were dissected using watchmaker's forceps under a microscope (Leica M80) and staged according to morphological features according to Theiler's criteria. Fetal livers were dissected, mechanically dissociated in 1mL of  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free DPBS (supplemented with 7% fetal calf serum, 100 units/ml penicillin-streptomycin, 0.25mM EDTA), and filtered through a 70  $\mu\text{m}$  sieve to prepare single cell suspensions. The total number of cells in each sample was determined using a hemocytometer (BLAUBRAND, Neubauer).

### ***Adult tissue collection***

Femurs and tibias were dissected, washed, and crushed in  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free DPBS (supplemented with 7% fetal calf serum, 100 units/ml penicillin-streptomycin, 0.25mM EDTA) using a mortar and a pestle. Bone marrow suspensions were mechanically dissociated, and filtered through a 70  $\mu\text{m}$  sieve. Spleens and thymi were isolated and a single cell suspension were created by massaging each organ through a 70  $\mu\text{m}$  sieve. The total number of cells in each sample was determined using a Neubauer hemocytometer. Blood samples were collected via mandible or retro-orbital bleeds in EDTA-containing microvette tubes.

### ***Flow cytometry/Cell sorting***

KIT-expressing fetal liver or adult bone marrow cells were enriched using magnetic separation (Miltenyi Biotech) according to the manufacturer's instructions. KIT-enriched cells were stained for

flow analysis or sorting using the following immunophenotypes: LT-HSCs were defined as Lin-KIT<sup>+</sup>SCA1<sup>+</sup>FLT3-CD48-CD150<sup>+</sup>; ST-HSCs Lin-KIT<sup>+</sup>SCA1<sup>+</sup>FLT3-CD48-CD150<sup>-</sup>; MPP4s Lin-KIT<sup>+</sup>SCA1<sup>+</sup>FLT3<sup>+</sup>CD48<sup>+</sup>CD150<sup>-</sup>; MPP3s Lin-KIT<sup>+</sup>SCA1<sup>+</sup>FLT3-CD48<sup>+</sup>CD150<sup>-</sup>; MPP2s Lin-KIT<sup>+</sup>SCA1<sup>+</sup>FLT3-CD48<sup>+</sup>CD150<sup>+</sup>. Lineage cocktail contained; Ter119, B220, MAC1, CD3ε<sup>+</sup>, CD19, CD4, CD8, Ly6G, NK1.1. Anti-Mac1 antibody was omitted in the lineage cocktail for fetal liver preparations to avoid excluding some of the Lin-KIT<sup>+</sup>SCA1<sup>+</sup><sup>(Morrison et al., 1995)</sup>. Differentiated hematopoietic lineages were defined as: erythroid cells TER119<sup>+</sup>; B cells TER119<sup>-</sup> MAC1<sup>-</sup> B220<sup>+</sup> CD19<sup>+</sup>; Monocytes TER119<sup>-</sup> MAC1<sup>+</sup> Ly6G<sup>-</sup> SSC\_A low; Neutrophils TER119<sup>-</sup> MAC1<sup>+</sup> Ly6G<sup>+</sup>; Platelets TER119<sup>-</sup> CD41<sup>+</sup>; Natural killer cells TER119<sup>-</sup> MAC1<sup>-</sup> B220<sup>-</sup> CD3ε<sup>-</sup> NK1.1<sup>+</sup>; T cells TER119<sup>-</sup> MAC1<sup>-</sup> B220<sup>-</sup> NK1.1<sup>-</sup> CD3ε<sup>+</sup>. All antibody stains were performed on ice in the dark. Analysis or cell sorting were carried out on a BD LSRFortessa or a BD FACSAria (using an 85 μm nozzle), respectively. Dead cells were excluded using 7AAD. Samples were further analyzed using FlowJo software. Antibody details can be found in Table S1.

### ***Hematopoietic reconstitution assays***

Recipient mice were 8 – 10 weeks old Ly5.1 females. Lethal irradiation was delivered as a 1100 rads split dose (2 x 550 rads, three hour gap). Lethally irradiated recipients received 2 x 10<sup>5</sup> bone marrow filler cells (from Ly5.1 or Ly5.1/2 mice). Sub-lethal irradiation was delivered as a 600 rads dose (no filler cells were used). Short-term (four weeks) and long-term ( $\geq$  16 weeks) transplantations were performed using lethal irradiation. Acute-term transplantations (two weeks) were performed using a sublethal irradiation dose. For transplantation, cells were collected from the bone marrow of adult UBC-GFP mice, prepared in DPBS with 1% FCS and 100 units/ml penicillin-streptomycin, and delivered into recipients by intravenous injection. Bone marrow from UBC-GFP mice were used as donor cells. Cells were injected into recipients approximately three hours after irradiation. Recipients were maintained on neomycin for 14 – 21 days. For estimation of LT-RU frequency (limiting dilution analysis), reconstitution frequency from the transplantation of 10 – 100 purified cells was analyzed using the extreme limiting dilution method (Hu and Smyth, 2009).

### ***Single cell RNA-Seq***

Cells were index-sorted on a 384 well plate using an BD FACS AriaIII or Fusion. 9 C57BL/6 E14.5 fetal livers were pooled from two litters and sorted into a 384-well plates. Libraries were prepared using CELSeq2 (Hashimshony et al., 2016) with additional optimizations, as described in (Amann-Zalcenstein et al., 2020). E14.5 fetal livers were pooled from two litters and sorted on the same day across multiple replicate plates. Sorts were undertaken using aseptic protocols and using a 100  $\mu$ m nozzle. For flow alignment, 1 bead coated with Horse Radish Peroxidase was sorted into 1.2  $\mu$ L of horse-radish peroxidase substrate before and after sort controls. Colorimetric assessment confirmed proper alignment of the cell sorter. Single cells were sorted into plates containing a primer/lysis mix (dNTPs, ERCC and SUPERase Inhibitor). Plates were centrifuged for 1 min at 1,200  $\times g$  and immediately frozen down at  $-80^{\circ}\text{C}$  until further processing. For sequencing, libraries from a single plate (containing 375 single cells) were prepared with a CELSeq2 protocol (Amann-Zalcenstein et al., 2020). Sequencing reads were aligned and mapped to the mm10 mouse genome and ERCC spike-in sequences using Rsubread (Liao et al., 2019). GRCm38.88\_chr annotation was then added. The data was demultiplexed and reads overlapping genes were summarized into counts using scPipe v1.6.0 (Tian et al., 2018). All subsequent analyses were performed in R version 3.6.0. Seven cells were identified as outliers based on number of genes detected, total gene counts, ERCC percentage, mitochondrial gene percentage and ribosomal gene percentage, and were excluded leaving 368 cells for downstream analysis. Gender-related genes were removed to avoid potential gender biases in the analysis. Lowly expressed genes were also filtered out such that 15590 genes showing an average count greater than 1 in at least 10 cells were retained. Normalisation factors were computed by deconvolution using scran v1.12.1 (Lun et al., 2016b) and subsequently used to calculate log-transformed normalised expression values. Gene-specific variance of the biological and technical components of the data were then estimated. Cell cycle phases were predicted using the cyclone function in scran. Differential expression analyses both between cell types and between cell clusters were carried out using edgeR v3.26.6 (Chen et al., 2016b; Lun et al., 2016a). For each analysis, biological variation between samples were estimated using edgeR's estimateDisp function (Chen

et al., 2016b). Generalized log-linear models were then fitted to the count data, incorporating an adjustment for cell cycle phase. edgeR's likelihood ratio test pipeline was applied to identify differentially expressed genes between the cell populations (McCarthy et al., 2012). The false discovery rate (FDR) was controlled below 5% using the Benjamini Hochberg method. Gene Ontology and KEGG pathway analysis were performed using limma v3.40.6 (Ritchie et al., 2015). Dimension reduction using Principal Components Analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), and clustering was performed using Seurat v3.0.2 (Stuart et al., 2019). Code used for analysis will be made available on request.

### ***Lentiviral barcoding***

The mCherry-expressing SPLINTR library of DNA barcodes used contained ~70,000 unique barcodes (Fennell et al., 2022). E14.5 LSK subsets were collected from pooled littermates ( $n = 4$  independent experiments). A maximum of  $3 \times 10^4$  purified E14.5 FL LSK cells were transduced with SPLINTR virus in 96-well plates via spinfection (90 mins at  $1250 \times g$ ) at a multiplicity of infection of 0.01 - 0.02 (this provided 1 – 2% mCHERRY+ cells) to limit multiple integrations of the barcode per cell. Cells were washed and then injected intravenously into recipient mice. After two weeks mCHERRY+ barcoded cells were retrieved from the spleens of recipient mice two-weeks post transplantation. Spleens were stained with markers for the isolation of monocytes (TER119<sup>-</sup> MAC1<sup>+</sup>Ly6G<sup>-</sup> [low side-scatter]), neutrophils (TER119<sup>-</sup>MAC1<sup>+</sup>LY6G<sup>+</sup>), erythroblasts (TER119<sup>+</sup>CD71<sup>+</sup>), and B cells (TER119<sup>-</sup>MAC1<sup>-</sup>B220/CD19<sup>+</sup>). mCHERRY<sup>+</sup> cell populations were sorted and spun into pellets of no more than  $5 \times 10^5$  cells. Pellets were resuspended in 40  $\mu$ L Direct PCR Lysis Reagent (Cell) with 0.5 mg/mL Proteinase K. Cells underwent lysis at 55°C for 2 hrs, followed by 85°C for 30 mins and 95°C for 5 mins. Libraries were generated as outline in (Naik et al., 2013). Raw sequencing FASTQ files were processed using custom C++ code to count SPLINTR barcode reads in each sample, including exclusively barcodes that aligned to a previously generated reference library. All subsequent analyses were performed using the statistical computing language R. Counts below a threshold of 100 reads were set to zero. Barcodes below a threshold of 100 sequencing reads, or only detected in one PCR replicate were removed to minimise artefacts (Figure S1). Heatmaps of the barcode counts were generated using

the R package "heatmap 3". Finally, the fate (presence/absence in B/Ery/Mono/Neut) for each barcode was determined, and correlation between distribution of fates computed between different populations (iLT-HSC/iST-HSC/iMPP2/iMPP3/iMPP4) using the Pearson correlation coefficient. To quantify the contribution of each ancestral clone to a given lineage (lineage biomass), the proportion of reads attributed to each fate combination was computed per LSK subset.

### ***General statistical analysis***

Prism 9 (GraphPad) was used for data analysis and graph production. Data are represented as mean  $\pm$  standard deviation (SD), and analyzed using Student's t-test (two-way, unpaired). One-way ANOVA (using Sidak's P value adjustment) was used for multiple comparisons. 'n' was used to designate the number of independent experimental mice. ns, not statistically significant. \* = p<0.05, \*\* = p<0.01, \*\*\* = p<0.001, \*\*\*\* = p<0.0001.

## Results

### ***Quantitative and transcriptional investigation of fetal LSK subsets***

We used the adult LSK-SLAM immunophenotypes (Pietras et al., 2015) to investigate the development of HSPC (LSK) subsets in E11.5 - E14.5 fetal livers (Figure 1A – C). To avoid the presumption of functional equivalence to adult bone marrow HSPCs, we adopted the convention of referring to cell populations with the immunophenotypic ('i') prefix (e.g. iST-HSC), as has been implemented previously by others (Chen et al., 2016a; Dong et al., 2020). At E11.5 iST-HSC (LSK\_FLT3-CD150-CD48-), iMPP3 (LSK\_FLT3-CD150-CD48+), and iMPP4 (LSK\_FLT3+CD150-CD48+) populations were observed in the liver (Figure 1A – E). By E12.5 all LSK subtypes, including iLT-HSCs (LSK\_FLT3-CD150+CD48-) and MPP2 (LSK\_FLT3-CD150+CD48+), were detected (Figure 1A – E).

We next compared the transcriptomes of E14.5 FL LSK subsets to gain insight into potential functional differences. To this end, single cells from E14.5 FL LSK subsets were purified (as indicated in Figure 2A) using indexed flow cytometry (which recorded the immunophenotype of the purified cell). Index data was used to verify the immunophenotype of the purified cells. Unsupervised hierarchical clustering of single-cell RNA-Sequencing data indicated that E14.5 FL iMPP 2- 4 were transcriptionally distinct subsets (Figure 2B). Gene ontology (GO) term analysis using significantly differentially expressed genes between iMPP 2 – 4 revealed:

- Significant enrichment of pathways related to lymphoid differentiation/function in iMPP4s compared to iMPP2 and iMPP3 (Figures 2Ci and 2Di; Tables S2 and S3). This was in keeping with the lymphoid bias of ABM iMPP4 (Pietras et al., 2015).
- Significant enrichment of pathways associated with megakaryocytes and erythroid differentiation in iMPP2 compared to iMPP3 and iMPP4 (Figures 2Di and 2Ei; Tables S3 and S4). This was consistent with the megakaryocyte/erythroid bias of ABM iMPP2.
- Interestingly, in comparison with iMPP2 and iMPP4, iMPP3 presented with a more lineage-balanced transcriptional priming that spanned the erythro-myeloid, megakaryocytic, and lymphoid lineages (Figures 2Ci and 2Ei; Tables S2 and S4).

Visualization of the genes associated with enriched GO terms confirmed that the associated genes were broadly expressed within the relevant populations (Figures 2Cii – Eii and Table S5). This

indicated that adult MPP immunophenotypes effectively enriched for E14.5 LSKs with features of ABM-like lineage priming.

### ***Functional comparison of E14.5 FL and ABM immunophenotypic counterparts in vivo***

To compare the acute-term *in vivo* differentiation potential of E14.5 FL and ABM LSK subtypes, cells purified from UBC-GFP mice (Schaefer et al., 2001) were transplanted into GFP- sub-lethally irradiated recipients. To enable a like-for-like comparison the number of cells transplanted was determined by the availability of cells in the fetal liver, and the number of cells that would yield robust engraftment. From either E14.5 FL or ABM donors, 1200 iLT-HSCs, 4000 iST-HSCs, 7000 iMPP2s, 25000 iMPP3s, and 30000 iMPP4s were transplanted. Analysis of *in vivo* output was performed after two weeks. This time point was selected to ensure that the output of iMPPs (which rapidly exhaust (Pietras et al., 2015)) would be captured. Contribution to platelets, erythrocytes, myeloid, and lymphoid lineages was assessed in the peripheral blood, spleen, or bone marrow of recipient mice.

Inter-developmental stage comparison of the HSC subsets revealed largely equivalent outcomes (Figure 3A and B), the exception was greater erythroid output by FL iLT-HSCs (Figure 3A). Of note, the low T-cell observed from all of the HSC subsets two weeks after transplantation was consistent with observations from other studies (Forsberg et al., 2006; Papathanasiou et al., 2009; Pietras et al., 2015). iMPP 2 – 4 subsets also performed similarly between the developmental stages (Figure 3 C – E). The exceptions were enhanced B-cell output from FL iMPP3 (Figure 3D) and diminished T-cell output from FL iMPP4 (Figure 3E).

### ***Clonal tracking E14.5 FL LSK subsets following transplantation***

To understand the clonal nature of *in vivo* differentiation of FL iHSC and iMPP subsets we transplanted cells following lentiviral cellular barcoding. *Ex vivo* cellular barcoding involves the indelible labelling of individual cells with a high-diversity library of genetically heritable DNA sequences termed barcodes (Fennell et al., 2022; Gerrits et al., 2010; Lu et al., 2011; Naik et al., 2013; Schepers et al., 2008). Once LSK subsets are transduced with a barcode-carrying lentivirus they can be transplanted into irradiated recipients and their barcoded progeny can be purified by

flow cytometry. Clonal ancestry can be determined by retrieving barcode sequences from the genome of donor-derived cells.

Purified FL LSK subsets were transduced with a previously described mCHERRY-expressing lentiviral barcode library composed of  $\sim 7 \times 10^4$  unique barcodes (Fennell et al., 2022). A maximum of  $3 \times 10^4$  cells were cultured with a pre-titered library for 16 hrs to achieve 1 – 2 % transduction. This provided a >100-fold excess in barcode diversity, making repeat use of barcodes and multiple infection of individual cells unlikely (Kebschull and Zador, 2018; Naik et al., 2014) (Figure S1). Of note, the culture period did not affect *in vivo* output (Figure S2). Two weeks after transplantation, mCHERRY+ B-cells, erythroblasts, monocytes, and granulocytes were harvested from recipient mice. It has been previously shown that at two weeks post-transplantation the spleen not only contains the offspring of clones that also engraft the bone marrow but it also contains clones that have not engrafted the bone marrow (Naik et al., 2013). Accordingly, to ensure effective retrieval of barcoded offspring we focused on the retrieval of hematopoietic lineages from the spleen to determine the patterns of clonal outcomes for each of the E14.5 LSK subsets (Figure 4A).

E14.5 FL iLT-HSC clones ( $n = 41$ ) and FL iST-HSC clones ( $n = 43$ ) contributed to similar clonal outputs (Figure 4B and 4C). This ranged from complete multi-lineage outcomes that spanned the lymphoid, myeloid, and erythroid lineages (~30% of clones), various combinations of restricted multi-outcomes, erythro-myeloid restriction, and single-lineage restriction (Figure 4B and 4C). This was in stark contrast to the outcome of MPPs. Clonal tracking of MPPs revealed that:

- Of iMPP2 clones ( $n = 15$ ), some were capable of restricted multi-outcomes but 86% of clones underwent single-lineage myeloid or erythroid production (Figure 4D).
- Of iMPP3 clones ( $n = 57$ ), 68% were single-lineage restricted; this included production of B-cells, monocytes, or erythroblasts. 30% of clones underwent restricted multi-outcomes (myeloid lineage and B-cell outcome), and only 1 clone underwent complete multi-outcome (Figure 4E).
- Although similar in the performance to iMPP3, iMPP4 clones ( $n = 168$ ) were more prone to B-cell restricted outcome (Figure 4F).

Although all iMPPs (2 – 4) exhibited some degree of multi-lineage outcomes, contribution to all tested lineages was a feature largely restricted to iLT-HSCs and iST-HSCs.

Correlation between distribution of outcomes between the LSK subsets revealed a marked difference between iLT-HSCs and iST-HSCs compared to the iMPPs (Figure 4G). The iHSC versus iMPP segregation was also evident when contribution of clones to lineage biomass was considered (Figure 4H). These data revealed that E14.5 iLT-HSCs and iST-HSCs contained cells with the capacity for complete multi-lineage outcomes, and that the lineage-biased outcomes observed from MPPs at the population-level were likely driven by a mix of clones that underwent with single-lineage restricted outcomes.

### ***Comparison of stemness between E14.5 FL iLT-HSCs and iST-HSCs***

From the barcoding experiments we noted that FL iLT-HSCs and iST-HSCs performed similarly. ABM iST-HSCs diverged from ABM iLT-HSCs according to differences in durability of multilineage reconstitution, and by the inability to self-renew (Morita et al., 2010; Oguro et al., 2013; Pietras et al., 2015). To further investigate FL iST-HSCs, we next compared the function of E14.5 FL iLT-HSCs and iST-HSCs in long-term transplantation assays. To distinguish immunophenotypic classification from stem cell function we will continue to refer to HSC immunophenotypes as either iLT-HSC or iST-HSC but the provision of durable multilineage reconstitution will be referred to as the output of a long-term repopulating unit (LT-RU).

We first compared the performance of E14.5 FL iLT-HSCs and iST-HSCs following transplantation into lethally irradiated recipients (the gold-standard HSC assay). Analysis of erythroid, platelet, and leucocyte reconstitution at 5, 16, and 40 weeks after transplantation of 1000 cells per recipient indicated that *bona fide* LT-RUs were present in both populations (Figure 5A). Analysis of the LSK compartment of recipients (after 40 weeks) revealed that LT-RUs from both populations generated all ABM LSK subsets, including iLT-HSCs (Figure 5B).

We next performed limiting dilution experiments to determine the frequency of LT-RUs in the FL iLT-HSC and iST-HSC populations. For this, we transplanted 10, 50, 100, or 1000 purified cells into lethally irradiated recipients; mice were analyzed after 16 weeks (Figure 5C). Using the extreme limiting dilution method (Hu and Smyth, 2009), we found that LT-RUs were

present at a frequency of 1/19 iLT-HSCs (which is in keeping with a previous estimate (Kim et al., 2006)) and 1/34 iST-HSCs (Figure 5D). Considering the absolute number of iLT-HSCs and iST-HSCs per E14.5 FL (Figure 1B), we estimated that the iST-HSC population contained approximately twice the number of LT-RUs than the iLT-HSC population (Figure 5E).

To investigate the self-renewal capacity of LT-RUs in the iST-HSC population (herein referred to as iST:LT-RU, 1/10 or 1/100 doses of ABM from primary recipients (that had received 100 donor cells) were transplanted into lethally irradiated secondary recipients. We found no significant difference in secondary reconstitution between iLT:LT-RU and iST:LT-RU (Figure 5F). Thus, the E14.5 FL ST-HSC population contains large number of *bona fide* LT-RUs that did not exhibit notable lineage biases.

### ***Refinement of the E14.5 FL ST-HSC LT-RU immunophenotype***

We next investigated our E14.5 FL LSK scRNA-Seq dataset to understand if a unifying LT-RU transcriptional signature could be identified. Based on similarity matrix analysis, 7 transcriptional clusters were identified; most iLT-HSCs and iST-HSCs segregated into clusters 4 and 7 (Figure 6A). Derivation of signature genes for these clusters indicated that cluster 4 likely contained LT-RUs: *Mecom*, *Hlf*, and *Mllt3* are strongly associated with LT-RU function (Calvanese et al., 2019; Kataoka et al., 2011; Komorowska et al., 2017; Wahlestedt et al., 2017) and were enriched in cluster 4 (Figure 6A, Figure S3, Table S6, and Table S7). In contrast, erythroid-associated genes (e.g. *Epor*, *Klf1*, and *Redrum*) were expressed at higher level in cluster 7 (Table S6 and Table S7).

Intra-cluster differential gene expression analysis between within iLT-HSCs and iST-HSCs revealed that cells within clusters were transcriptionally very similar (Figure 6B, Table S8, and Table S9), while many genes were differentially expressed between clusters (Figures 6C, Table S6, and Table S7). Although very similar, five genes were differentially expressed between iLT-HSCs and iST-HSCs within cluster 4: *Gm42418*, *Igf2*, *Igf1r*, *Mir6236*, and *Satb1* were more highly expressed in iST-HSCs (Table S8). Previous studies have shown that: *in vitro* treatment of FL and ABM iLT-HSCs with IGF2 results in LT-RU expansion (Zhang and Lodish, 2004); level of IGF2 receptor expression in FL and ABM cells correlates with greater hematopoietic reconstitution (Zhang and Lodish, 2004); enforced expression of *Igf2* in ABM iLT-HSCs enhances hematopoietic

reconstitution (Thomas et al., 2016); and, high-levels of *Satb1* expression correlates with improved hematopoietic reconstitution (Doi et al., 2018). This suggests that some degree of functional difference between iLT:LT-RUs and iST:LT-RUs might exist.

Given that LT-RU frequency in the iST-HSC population was half that observed in the iLT-HSC population, to directly compare iST:LT-RUs with iLT:LT-RU we first aimed to improve iST:LT-RU enrichment. To this end, differential gene expression analysis was performed between iST-HSCs in cluster 4 and cluster 7. Amongst the significantly upregulated genes in the cluster 4, *Ptprc* (which encodes CD45) was found to be strongly up-regulated (Figure 6Di and Table S10). Other genes included *Procr* (which encodes EPCR) and *Esam* (Figure 6Di and Table S10), both of which are known to be expressed by E14.5 iLT-HSCs and enrich for LT-RU activity (Balazs et al., 2006; Kent et al., 2009; Yokota et al., 2009). CD45 is expressed by LT-RUs once they have emerged in the E11.5 AGM region (Taoudi et al., 2008; Taoudi et al., 2005), in the E12.5 yolk sac (Taoudi et al., 2005), and in the E13.5 – E15.5 FL (Kent et al., 2009; Taoudi et al., 2005). Thus, we considered CD45 expression to be a promising marker for improved FL iST:LT-RU enrichment. Using flow cytometry we found that  $85\% \pm 10\%$  of iLT-HSCs and  $56\% \pm 12\%$  of iST-HSCs were CD45+ (Figure 6Dii).

We next compared the ability of iLT-HSCs, CD45+ iST-HSCs, and CD45low/- iST-HSCs to provide acute- and long-term reconstitution. To this end, 100 cells of each immunophenotype were transplanted into lethally irradiated recipients. In keeping with the enhanced reconstitution observed with IGF2-signalling and *Satb1* expression (Doi et al., 2018; Thomas et al., 2016; Zhang and Lodish, 2004), after two weeks CD45 low/- iST-HSCs cells failed to contribute to reconstitution but CD45+ iST-HSCs cells provided robust reconstitution that was significantly greater than that of iLT-HSCs (Figure 6E). By 16 weeks, CD45 low/- iST-HSCs did not provide any reconstitution, but iLT-HSCs and CD45+ iST-HSCs provided equivalent hematopoietic contribution (Figure 6F). Secondary transplantation demonstrated the long-term self-renewal capacity of CD45+ iST-HSC LT-RU reconstitution (Figure 6G). As all LT-RUs were contained within the 56% of iST-HSCs that expressed CD45 (Figure 6Dii), the LT-RU frequency in the CD45+ iST-HSC population in  $\sim 1/19$  cells, a frequency equivalent to LT-RUs in the iLT-HSC population (Figure 5D).

***The iST:LT-RU variant does not contain developmentally-restricted (lymphoid-biased) HSCs but is labelled by Flt3-cre***

Developmentally-restricted HSCs (drHSCs) (Beaudin et al., 2016) are a component of the E14.5 FL LSK. Similar to LT-HSCs, drHSCs express CD150 and can self-renew in secondary transplantation assays. However, under physiological conditions drHSCs exhaust between the neonatal period and adulthood. Two features distinguish drHSCs from LT-HSCs: (1) from eight weeks post-transplantation, drHSC-derived reconstitution is lymphoid-biased (Beaudin et al., 2016). This bias is transferred upon secondary transplantation (Beaudin et al., 2016); and, (2) drHSCs emerge via a *Flt3*-expressing ancestry. Use of a *Flt3*-cre expressing mouse line elegantly distinguishes LT-HSCs from drHSCs in E14.5 FL LSK CD150+ compartment (Beaudin et al., 2016). Our data showed that iST:LT-RUs were CD150- and provided lineage-balanced long-term reconstitution, thus were not a component of the drHSC axis.

A recent study demonstrated cells that non LT-HSCs that derive via a *Flt3*-expressing ancestry contribute to native hematopoiesis in adulthood (Patel et al., 2022); these cells were termed long-lived embryonic MPPs. To investigate if iST:LT-RUs were a component of this long-lived embryo-derived axis of native adult hematopoiesis, we inter-crossed *Flt3*-cre mice with floxed-STOP *Rosa26-EYFP* mice. In keeping with the findings of Patel *et al* (Patel et al., 2022), we found that CD45+ iST-HSCs were significantly more effectively labelled than iLT-HSCs (Figure 7A). To compare reconstitution potential, 100 YFP+ or YFP- iST-HSC[CD45+] cells were transplanted into lethally irradiated mice. After 4 weeks YFP+ and YFP- cells contributed to monocyte and B-cell reconstitution (Figure 7B); after 16 weeks robust monocyte, B-cell, and T-cell reconstitution was observed (Figure 7C). Thus, iST:LT-RUs of *Flt3*-expressing ancestry were capable of long-term lineage-balanced hematopoietic reconstitution.

## Discussion

Transcriptional analysis of iMPP2 – 4 from the E14.5 FL revealed that these populations were transcriptionally diverse along axes that were largely predictable based on insight from their adult counterparts (Pietras et al., 2015): iMPP2 was distinguished by the expression of megakaryocytic and erythroid genes, iMPP4 was distinguished by the expression of lymphopoiesis genes, and iMPP3 expressed a balance of megakaryocytic, erythroid, and lymphoid genes. Functional analysis (via population-level transplantation) demonstrated that iMPP2 had a platelet (megakaryocyte), erythrocyte, and monocyte production bias; iMPP4 exhibited a lymphoid bias (relative to iMPP2) and, iMPP3 was statistically similar to iMPP4. Clonal investigation revealed that the FL iMPPs were a heterogeneous mix cells that underwent single- or dual-lineage outcomes, while multi-lineage outcomes were largely reserved for the iLT- and iST-HSC subsets. The subdued clonal capacity of FL iMPPs is consistent with observation made from ABM iMPPs following transplantation (Naik et al., 2013) and during native hematopoiesis (Rodriguez-Fraticelli et al., 2018).

The most distinguishing feature of FL iMPPs was the capacity of iMPP3 to produce B-cells, this outcome largely derived from clones that were capable of monocyte and B-cell production, whereas iMPP4 achieved B-cell production predominantly via uni-outcome clones. E14.5 FL LT-HSCs and ST-HSCs were strikingly distinct from the iMPPs. This was most evident in the lentiviral barcoding experiments where complete multi-lineage outcomes were only repeatedly detected in the HSC subsets.

We made the surprising discovery that in acute-term (two weeks), short-term (four weeks), and long-term (> 16 weeks) transplantation assays the E14.5 iST-HSC population was functionally very similar to E14.5 iLT-HSCs, including clonal output. E14.5 long-term repopulating units (LT-RUs) were not only distributed between the iLT-HSC and iST-HSC immunophenotypes, but most LT-RUs were concealed within the CD45-expressing fraction of the iST-HSC population. ABM iST-HSCs have the capacity for short-term multi-lineage reconstitution but were generally less able to sustain this in the long-term, and ABM iST-HSCs generally could not reconstitute secondary recipients (thus did not self-renew) (Morita et al., 2010; Oguro et al., 2013; Pietras et al., 2015). In contrast to this, we found that LT-RUs within the E14.5 ST-HSC population (iST:LT-RUs) were

functionally similar to LT-RUs within the E14.5 LT-HSC population (iLT:LT-RUs), this included: durable multilineage reconstitution for > 40 weeks post-transplantation and following secondary transplantation; and clonal multi-lineage production which enabled the production of platelet, erythroid, monocyte, and B-cell lineages following transplantation.

The E14.5 FL iLT-HSC population contains a sub-class of CD150+ LT-RUs known as developmentally-restricted HSCs (drHSCs) which are lymphoid-biased HSCs that contribute to the first HSC-derived lineages of the fetus (Beaudin et al., 2016). Using the *Flt3*-cre:YFP mouse line (which labels drHSCs (Beaudin et al., 2016)) we found that the FL iST-HSCs were significantly more labelled than iLT-HSCs, and that *bona fide* LT-RUs within the iST-HSC had derived via a *Flt3*-expressing ancestry. This suggests that multiple LT-RU forming pathways exist to establish the E14.5 fetal stem cell pool: one that derives via a *Flt3*-expressing ancestry that yielded YFP+ LT-RUs in the iST-HSC population and one that yielded LT-RUs in iLT-HSC that were predominantly YFP-. Whether these pathways originate from different anatomical sites, or from different precursors within the same stem-cell forming organ will be a subject of future investigation.

Taken together with previous studies, our findings indicate that three LT-RU populations must co-exist in the E14.5 fetal liver: conventional CD150+ iLT:LT-RUs (Kim et al., 2006), CD150+ drHSCs (Beaudin et al., 2016), and CD150- iST:LT-RUs. Although other studies have previously shown that LT-RUs were present in the CD150- fraction of the E14.5 FL (Kent et al., 2009; Kim et al., 2006; Papathanasiou et al., 2009) they did not quantify the LT-RU biomass distribution between the CD150+ and CD150- fractions of LSK FLT3- CD48- cells. Thus, the significance of CD150- LT-RU went unappreciated. We found that approximately 70 % of E14.5 FL LT-RUs are within the ST-HSC immunophenotype (LSK\_FLT3-CD150-CD48-, Figure 5E). Based on our findings (analysis of our scRNA-Seq data and functional tests), the use of LT-RU-associated markers such as EPCR (Balazs et al., 2006; Kent et al., 2009), ESAM (Yokota et al., 2009), or CD45 ((Kent et al., 2009; Taoudi et al., 2008; Taoudi et al., 2005) and Figure 6E – G) in LSK\_FLT3- CD48- cells would provide a single unifying LT-RU enriching immunophenotype. However, this approach would censor important information such as differences in the developmental pathways of iLT-HSCs and iST-HSCs as indicated by *Flt3*-expressing ancestry ((Patel et al., 2022) and Figure 7).

When LT-RUs first emerge in the E11.5 AGM region they do not express CD150 (McKinney-Freeman et al., 2009), nor do LT-RUs express CD150 in the E12.5 placenta (McKinney-Freeman et al., 2009). Although E14.5 LT-RUs can be purified at high frequency according to CD150 expression (Kim et al., 2006), our data demonstrates that the E14.5 FL LT-RUs biomass is a mix of CD150- and CD150+ LT-RUs. We were not able to test whether during continued physiological development E14.5 FL iST:LT-RUs would initiate CD150 expression and so become iLT:LT-RUs. Although it could be presumed that early LT-RUs (e.g, those produced in the AGM region) all upregulate CD150 expression during fetal development, this is currently untested. Of note, findings from a previous study suggest that FL iLT-HSCs and iST-HSCs are ontologically distinct lineages that largely maintain segregation and persist into adulthood: using a *Flt3*-creERT2 mouse line, Patel *et al* (Patel et al., 2022) demonstrated that after inducing labelling at E10.5 or at E14.5, iST-HSCs and iMPPs in the E15.5 FL could be labelled without labelling iLT-HSCs. The authors noted that *Flt3*-creERT2 labelled lineages from the fetus persisted into adulthood and contributed to ongoing hematopoiesis (Patel et al., 2022). Although it was concluded that embryonic MPPs drove the persistence of life-long *Flt3*-cre derived hematopoiesis, given our findings it is possible that LT-RUs present in the iST-HSC population could persist into adulthood and contribute to native hematopoiesis.

### ***Acknowledgements***

**Acknowledgments.** This work was supported by: the Australian Research Council (ARC) Stem Cells Australia program; NH&MRC project (1128993, 1129012, 2011770), program (1113577), and Investigator (1173342) grants; Independent Research Institutes Infrastructure Support Scheme grant 361646 from the NH&MRC, the Australian Cancer Research Fund, and Victorian State Government Operational Infrastructure Support. S.T was supported by a fellowship from the Lorenzo and Pamela Galli Charitable Trust and a Royal Society Wolfson Fellowship (233002).

### **Author contributions.**

O.J.S designed experiments, performed experiments, analyzed data, and wrote the manuscript first draft. C.B designed experiments, performed experiments, and analyzed data. A.G analyzed scRNA-Seq data. K.A.F, M.A.D, and S.H.N generated essential reagents and designed experiments. T.S.W analyzed lentiviral barcoding data. S.T, A.F, A.F.T, and W.S.A provided essential intellectual input. S.T. conceived the study, designed experiments, and analyzed data. All authors contributed to writing the manuscript.

**Competing Interests.** The authors declare no competing interests exist.

**Materials and correspondence.** Correspondence should be addressed to Samir Taoudi (samir.taoudi@bristol.ac.uk).

## References

Adolfsson, J., O.J. Borge, D. Bryder, K. Theilgaard-Monch, I. Astrand-Grundstrom, E. Sitnicka, Y. Sasaki, and S.E. Jacobsen. 2001. Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* 15:659-669.

Adolfsson, J., R. Mansson, N. Buza-Vidas, A. Hultquist, K. Liuba, C.T. Jensen, D. Bryder, L. Yang, O.J. Borge, L.A. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, and S.E. Jacobsen. 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* 121:295-306.

Amann-Zalcenstein, D., L. Tian, J. Schreuder, S. Tomei, D.S. Lin, K.A. Fairfax, J.E. Bolden, M.D. McKenzie, A. Jarratt, A. Hilton, J.T. Jackson, L. Di Rago, M.P. McCormack, C.A. de Graaf, O. Stonehouse, S. Taoudi, W.S. Alexander, S.L. Nutt, M.E. Ritchie, A.P. Ng, and S.H. Naik. 2020. A new lymphoid-primed progenitor marked by Dach1 downregulation identified with single cell multi-omics. *Nat Immunol* 21:1574-1584.

Balazs, A.B., A.J. Fabian, C.T. Esmon, and R.C. Mulligan. 2006. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood* 107:2317-2321.

Beaudin, A.E., S.W. Boyer, J. Perez-Cunningham, G.E. Hernandez, S.C. Derderian, C. Jujjavarapu, E. Aaserude, T. MacKenzie, and E.C. Forsberg. 2016. A transient developmental hematopoietic stem cell gives rise to innate-like B and T cells. *Cell Stem Cell* 19:768-783.

Benz, C., M.R. Copley, D.G. Kent, S. Wohrer, A. Cortes, N. Aghaeepour, E. Ma, H. Mader, K. Rowe, C. Day, D. Treloar, R.R. Brinkman, and C.J. Eaves. 2012. Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell* 10:273-283.

Busch, K., K. Klapproth, M. Barile, M. Flossdorf, T. Holland-Letz, S.M. Schlenner, M. Reth, T. Hofer, and H.R. Rodewald. 2015. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* 518:542-546.

Calvanese, V., A.T. Nguyen, T.J. Bolan, A. Vavilina, T. Su, L.K. Lee, Y. Wang, F.D. Lay, M. Magnusson, G.M. Crooks, S.K. Kurdistani, and H.K.A. Mikkola. 2019. MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature* 576:281-286.

Che, J.L.C., D. Bode, I. Kucinski, A.H. Cull, F. Bain, H.J. Becker, M. Jassinskaja, M. Barile, G. Boyd, M. Belmonte, A.G.X. Zeng, K.J. Igarashi, J. Rubio-Lara, M.S. Shepherd, A. Clay, J.E. Dick, A.C. Wilkinson, H. Nakauchi, S. Yamazaki, B. Gottgens, and D.G. Kent.

2022. Identification and characterization of in vitro expanded hematopoietic stem cells. *EMBO Rep* 23:e55502.

Chen, J.Y., M. Miyanishi, S.K. Wang, S. Yamazaki, R. Sinha, K.S. Kao, J. Seita, D. Sahoo, H. Nakauchi, and I.L. Weissman. 2016a. Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* 530:223-227.

Chen, Y., A.T. Lun, and G.K. Smyth. 2016b. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* 5:1438.

de Bruijn, M.F., X. Ma, C. Robin, K. Ottersbach, M.J. Sanchez, and E. Dzierzak. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* 16:673-683.

de Bruijn, M.F., N.A. Speck, M.C. Peeters, and E. Dzierzak. 2000. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J* 19:2465-2474.

Doi, Y., T. Yokota, Y. Satoh, D. Okuzaki, M. Tokunaga, T. Ishibashi, T. Sudo, T. Ueda, Y. Shingai, M. Ichii, A. Tanimura, S. Ezoe, H. Shibayama, T. Kohwi-Shigematsu, J. Takeda, K. Oritani, and Y. Kanakura. 2018. Variable SATB1 Levels Regulate Hematopoietic Stem Cell Heterogeneity with Distinct Lineage Fate. *Cell Rep* 23:3223-3235.

Dong, F., S. Hao, S. Zhang, C. Zhu, H. Cheng, Z. Yang, F.K. Hamey, X. Wang, A. Gao, F. Wang, Y. Gao, J. Dong, C. Wang, J. Wang, Y. Lan, B. Liu, H. Ema, F. Tang, B. Gottgens, P. Zhu, and T. Cheng. 2020. Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis. *Nat Cell Biol* 22:630-639.

Dykstra, B., D. Kent, M. Bowie, L. McCaffrey, M. Hamilton, K. Lyons, S.J. Lee, R. Brinkman, and C. Eaves. 2007. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell* 1:218-229.

Eaves, C.J. 2015. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* 125:2605-2613.

Ema, H., and H. Nakauchi. 2000. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* 95:2284-2288.

Fennell, K.A., D. Vassiliadis, E.Y.N. Lam, L.G. Martelotto, J.J. Balic, S. Hollizeck, T.S. Weber, T. Semple, Q. Wang, D.C. Miles, L. MacPherson, Y.C. Chan, A.A. Guirguis, L.M. Kats, E.S. Wong, S.J. Dawson, S.H. Naik, and M.A. Dawson. 2022. Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature* 601:125-131.

Forsberg, E.C., T. Serwold, S. Kogan, I.L. Weissman, and E. Passegue. 2006. New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. *Cell* 126:415-426.

Ganuza, M., T. Hall, D. Finkelstein, A. Chabot, G. Kang, and S. McKinney-Freeman. 2017. Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. *Nature cell biology* 19:1153-1163.

Ganuza, M., T. Hall, J. Myers, C. Nevitt, R. Sanchez-Lanzas, A. Chabot, J. Ding, E. Kooienga, C. Caprio, D. Finkelstein, G. Kang, E. Obeng, and S. McKinney-Freeman. 2022. Murine foetal liver supports limited detectable expansion of life-long haematopoietic progenitors. *Nat Cell Biol* 24:1475-1486.

Gao, S., Q. Shi, Y. Zhang, G. Liang, Z. Kang, B. Huang, D. Ma, L. Wang, J. Jiao, X. Fang, C.R. Xu, L. Liu, X. Xu, B. Gottgens, C. Li, and F. Liu. 2022. Identification of HSC/MPP expansion units in fetal liver by single-cell spatiotemporal transcriptomics. *Cell Res* 32:38-53.

Gekas, C., F. Dieterlen-Lievre, S.H. Orkin, and H.K. Mikkola. 2005. The placenta is a niche for hematopoietic stem cells. *Dev Cell* 8:365-375.

Gerrits, A., B. Dykstra, O.J. Kalmykowa, K. Klauke, E. Verovskaya, M.J. Broekhuis, G. de Haan, and L.V. Bystrykh. 2010. Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* 115:2610-2618.

Hall, T.D., H. Kim, M. Dabbah, J.A. Myers, J.C. Crawford, A. Morales-Hernandez, C.E. Caprio, P. Sriram, E. Kooienga, M. Derecka, E.A. Obeng, P.G. Thomas, and S. McKinney-Freeman. 2022. Murine fetal bone marrow does not support functional hematopoietic stem and progenitor cells until birth. *Nat Commun* 13:5403.

Hashimshony, T., N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K.J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, and I. Yanai. 2016. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol* 17:77.

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

Kataoka, K., T. Sato, A. Yoshimi, S. Goyama, T. Tsuruta, H. Kobayashi, M. Shimabe, S. Arai, M. Nakagawa, Y. Imai, K. Kumano, K. Kumagai, N. Kubota, T. Kadokawa, and M. Kurokawa. 2011. Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. *J Exp Med* 208:2403-2416.

Kebschull, J.M., and A.M. Zador. 2018. Cellular barcoding: lineage tracing, screening and beyond. *Nat Methods* 15:871-879.

Kent, D.G., M.R. Copley, C. Benz, S. Wohrer, B.J. Dykstra, E. Ma, J. Cheyne, Y. Zhao, M.B. Bowie, Y. Zhao, M. Gasparetto, A. Delaney, C. Smith, M. Marra, and C.J. Eaves. 2009. Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood* 113:6342-6350.

Kiel, M.J., O.H. Yilmaz, T. Iwashita, O.H. Yilmaz, C. Terhorst, and S.J. Morrison. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121:1109-1121.

Kim, I., S. He, O.H. Yilmaz, M.J. Kiel, and S.J. Morrison. 2006. Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood* 108:737-744.

Komorowska, K., A. Doyle, M. Wahlestedt, A. Subramaniam, S. Debnath, J. Chen, S. Soneji, B. Van Handel, H.K.A. Mikkola, K. Miharada, D. Bryder, J. Larsson, and M. Magnusson. 2017. Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration. *Cell Rep* 21:3514-3523.

Kumaravelu, P., L. Hook, A.M. Morrison, J. Ure, S. Zhao, S. Zuyev, J. Ansell, and A. Medvinsky. 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.

Lee, Y., E. DiMaulo-Milk, J. Leslie, and L. Ding. 2022. Hematopoietic stem cells temporally transition to thrombopoietin dependence in the fetal liver. *Sci Adv* 8:eabm7688.

Liao, Y., G.K. Smyth, and W. Shi. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic acids research* 47:e47-e47.

Lu, R., N.F. Neff, S.R. Quake, and I.L. Weissman. 2011. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat Biotechnol* 29:928-933.

Lun, A.T., Y. Chen, and G.K. Smyth. 2016a. It's DE-licious: A Recipe for Differential Expression Analyses of RNA-seq Experiments Using Quasi-Likelihood Methods in edgeR. *Methods Mol Biol* 1418:391-416.

Lun, A.T., D.J. McCarthy, and J.C. Marioni. 2016b. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research* 5:

Mansson, R., A. Hultquist, S. Luc, L. Yang, K. Anderson, S. Kharazi, S. Al-Hashmi, K. Liuba, L. Thoren, J. Adolfsson, N. Buza-Vidas, H. Qian, S. Soneji, T. Enver, M. Sigvardsson, and S.E.W. Jacobsen. 2007. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* 26:407-419.

McCarthy, D.J., Y. Chen, and G.K. Smyth. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40:4288-4297.

McKinney-Freeman, S.L., O. Naveiras, F. Yates, S. Loewer, M. Philitas, M. Curran, P.J. Park, and G.Q. Daley. 2009. Surface antigen phenotypes of hematopoietic stem cells from embryos and murine embryonic stem cells. *Blood* 114:268-278.

Medvinsky, A., and E. Dzierzak. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* 86:897-906.

Morita, Y., H. Ema, and H. Nakauchi. 2010. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med* 207:1173-1182.

Morrison, S.J., H.D. Hemmati, A.M. Wandycz, and I.L. Weissman. 1995. The purification and characterization of fetal liver hematopoietic stem cells. *Proc Natl Acad Sci U S A* 92:10302-10306.

Muller-Sieburg, C.E., R.H. Cho, M. Thoman, B. Adkins, and H.B. Sieburg. 2002. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood* 100:1302-1309.

Naik, S.H., L. Perie, E. Swart, C. Gerlach, N. van Rooij, R.J. de Boer, and T.N. Schumacher. 2013. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* 496:229-232.

Naik, S.H., T.N. Schumacher, and L. Perie. 2014. Cellular barcoding: a technical appraisal. *Exp Hematol* 42:598-608.

Oguro, H., L. Ding, and S.J. Morrison. 2013. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* 13:102-116.

Ottersbach, K., and E. Dzierzak. 2005. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Dev Cell* 8:377-387.

Papathanasiou, P., J.L. Attema, H. Karsunky, J. Xu, S.T. Smale, and I.L. Weissman. 2009. Evaluation of the long-term reconstituting subset of hematopoietic stem cells with CD150. *Stem Cells* 27:2498-2508.

Patel, S.H., C. Christodoulou, C. Weinreb, Q. Yu, E.L. da Rocha, B.J. Pepe-Mooney, S. Bowling, L. Li, F.G. Osorio, G.Q. Daley, and F.D. Camargo. 2022. Lifelong multilineage contribution by embryonic-born blood progenitors. *Nature* 606:747-753.

Pei, W., T.B. Feyerabend, J. Rossler, X. Wang, D. Postrach, K. Busch, I. Rode, K. Klapproth, N. Dietlein, C. Quedenau, W. Chen, S. Sauer, S. Wolf, T. Hofer, and H.R. Rodewald. 2017. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* 548:456-460.

Pei, W., F. Shang, X. Wang, A.K. Fanti, A. Greco, K. Busch, K. Klapproth, Q. Zhang, C. Quedenau, S. Sauer, T.B. Feyerabend, T. Hofer, and H.R. Rodewald. 2020. Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding. *Cell Stem Cell* 27:383-395 e388.

Pietras, E.M., D. Reynaud, Y.A. Kang, D. Carlin, F.J. Calero-Nieto, A.D. Leavitt, J.M. Stuart, B. Gottgens, and E. Passegue. 2015. Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell* 17:35-46.

Ritchie, M.E., B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, and G.K. Smyth. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 43:e47-e47.

Rodriguez-Fraticelli, A.E., S.L. Wolock, C.S. Weinreb, R. Panero, S.H. Patel, M. Jankovic, J. Sun, R.A. Calogero, A.M. Klein, and F.D. Camargo. 2018. Clonal analysis of lineage fate in native haematopoiesis. *Nature* 553:212-216.

Sakai, S.S., A. Hasegawa, R. Ishimura, N. Tamura, S. Kageyama, S. Komatsu-Hirota, M. Abe, Y. Ling, S. Okuda, M. Funayama, M. Kikkawa, Y. Miura, K. Sakimura, I. Narita, S. Waguri, R. Shimizu, and M. Komatsu. 2022. Loss of Atg2b and Gskip Impairs the Maintenance of the Hematopoietic Stem Cell Pool Size. *Mol Cell Biol* 42:e0002421.

Sanchez, M.J., A. Holmes, C. Miles, and E. Dzierzak. 1996. Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity* 5:513-525.

Sawai, C.M., S. Babovic, S. Upadhyaya, D. Knapp, Y. Lavin, C.M. Lau, A. Goloborodko, J. Feng, J. Fujisaki, L. Ding, L.A. Mirny, M. Merad, C.J. Eaves, and B. Reizis. 2016. Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity* 45:597-609.

Sawen, P., M. Eldeeb, E. Erlandsson, T.A. Kristiansen, C. Laterza, Z. Kokaia, G. Karlsson, J. Yuan, S. Soneji, P.K. Mandal, D.J. Rossi, and D. Bryder. 2018. Murine HSCs contribute actively to native hematopoiesis but with reduced differentiation capacity upon aging. *Elife* 7:

Schaefer, B.C., M.L. Schaefer, J.W. Kappler, P. Marrack, and R.M. Kedl. 2001. Observation of antigen-dependent CD8+ T-cell/ dendritic cell interactions in vivo. *Cell Immunol* 214:110-122.

Schepers, K., E. Swart, J.W. van Heijst, C. Gerlach, M. Castrucci, D. Sie, M. Heimerikx, A. Velds, R.M. Kerkhoven, and R. Arens. 2008. Dissecting T cell lineage relationships by cellular barcoding. *The Journal of experimental medicine* 205:2309-2318.

Schoedel, K.B., M.N.F. Morcos, T. Zerjatke, I. Roeder, T. Grinenko, D. Voehringer, J.R. Gothert, C. Waskow, A. Roers, and A. Gerbaulet. 2016. The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. *Blood* 128:2285-2296.

Sheikh, B.N., Y. Yang, J. Schreuder, S.K. Nilsson, R. Bilardi, S. Carotta, H.M. McRae, D. Metcalf, A.K. Voss, and T. Thomas. 2016. MOZ (KAT6A) is essential for the

maintenance of classically defined adult hematopoietic stem cells. *Blood* 128:2307-2318.

Sommerkamp, P., M.C. Romero-Mulero, A. Narr, L. Ladel, L. Hustin, K. Schonberger, S. Renders, S. Altamura, P. Zeisberger, K. Jacklein, D. Klimmeck, A. Rodriguez-Fraticelli, F.D. Camargo, L. Perie, A. Trumpp, and N. Cabezas-Wallscheid. 2021. Mouse multipotent progenitor 5 cells are located at the interphase between hematopoietic stem and progenitor cells. *Blood* 137:3218-3224.

Srinivas, S., T. Watanabe, C.-S. Lin, C.M. William, Y. Tanabe, T.M. Jessell, and F. Costantini. 2001. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC developmental biology* 1:1-8.

Stuart, T., A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert, and R. Satija. 2019. Comprehensive Integration of Single-Cell Data. *Cell* 177:1888-1902 e1821.

Sun, J., A. Ramos, B. Chapman, J.B. Johnnidis, L. Le, Y.J. Ho, A. Klein, O. Hofmann, and F.D. Camargo. 2014. Clonal dynamics of native haematopoiesis. *Nature* 514:322-327.

Taoudi, S., C. Gonneau, K. Moore, J.M. Sheridan, C.C. Blackburn, E. Taylor, and A. Medvinsky. 2008. Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ pre-definitive HSCs. *Cell Stem Cell* 3:99-108.

Taoudi, S., A.M. Morrison, H. Inoue, R. Gribi, J. Ure, and A. Medvinsky. 2005. Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. *Development* 132:4179-4191.

Thomas, D.D., A.G. Sommer, A.B. Balazs, I. Beerman, G.J. Murphy, D. Rossi, and G. Mostoslavsky. 2016. Insulin-like growth factor 2 modulates murine hematopoietic stem cell maintenance through upregulation of p57. *Exp Hematol* 44:422-433 e421.

Tian, L., S. Su, X. Dong, D. Amann-Zalcenstein, C. Biben, A. Seidi, D.J. Hilton, S.H. Naik, and M.E. Ritchie. 2018. scPipe: A flexible R/Bioconductor preprocessing pipeline for single-cell RNA-sequencing data. *PLoS Comput Biol* 14:e1006361.

Van Deren, D.A., S. De, B. Xu, K.M. Eschenbacher, S. Zhang, and M.R. Capecchi. 2022. Defining the Hoxb8 cell lineage during murine definitive hematopoiesis. *Development* 149:

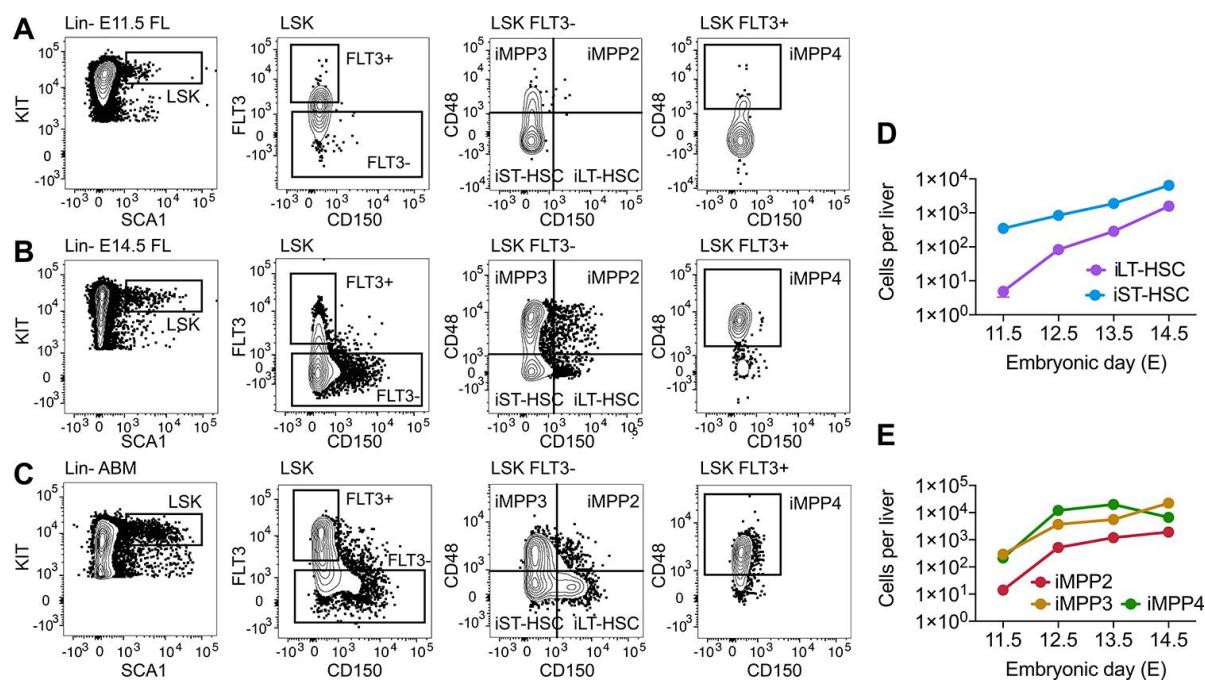
Wahlestedt, M., V. Ladopoulos, I. Hidalgo, M. Sanchez Castillo, R. Hannah, P. Sawen, H. Wan, M. Dudenhofer-Pfeifer, M. Magnusson, G.L. Nordahl, B. Gottgens, and D. Bryder. 2017. Critical Modulation of Hematopoietic Lineage Fate by Hepatic Leukemia Factor. *Cell Rep* 21:2251-2263.

Yokota, T., K. Oritani, S. Butz, K. Kokame, P.W. Kincade, T. Miyata, D. Vestweber, and Y. Kanakura. 2009. The endothelial antigen ESAM marks primitive hematopoietic progenitors throughout life in mice. *Blood* 113:2914-2923.

Young, I.C., B. Wu, J. Andricovich, S.T. Chuang, R. Li, A. Tzatsos, R.C. Wu, and M.Y. Wu. 2021. Differentiation of fetal hematopoietic stem cells requires ARID4B to restrict autocrine KITLG/KIT-Src signaling. *Cell Rep* 37:110036.

Zhang, C.C., and H.F. Lodish. 2004. Insulin-like growth factor 2 expressed in a novel fetal liver cell population is a growth factor for hematopoietic stem cells. *Blood* 103:2513-2521.

**Figure 1**



**Figure 1. Quantification of LSK subsets in the fetal liver between E11.5 – E14.5.**

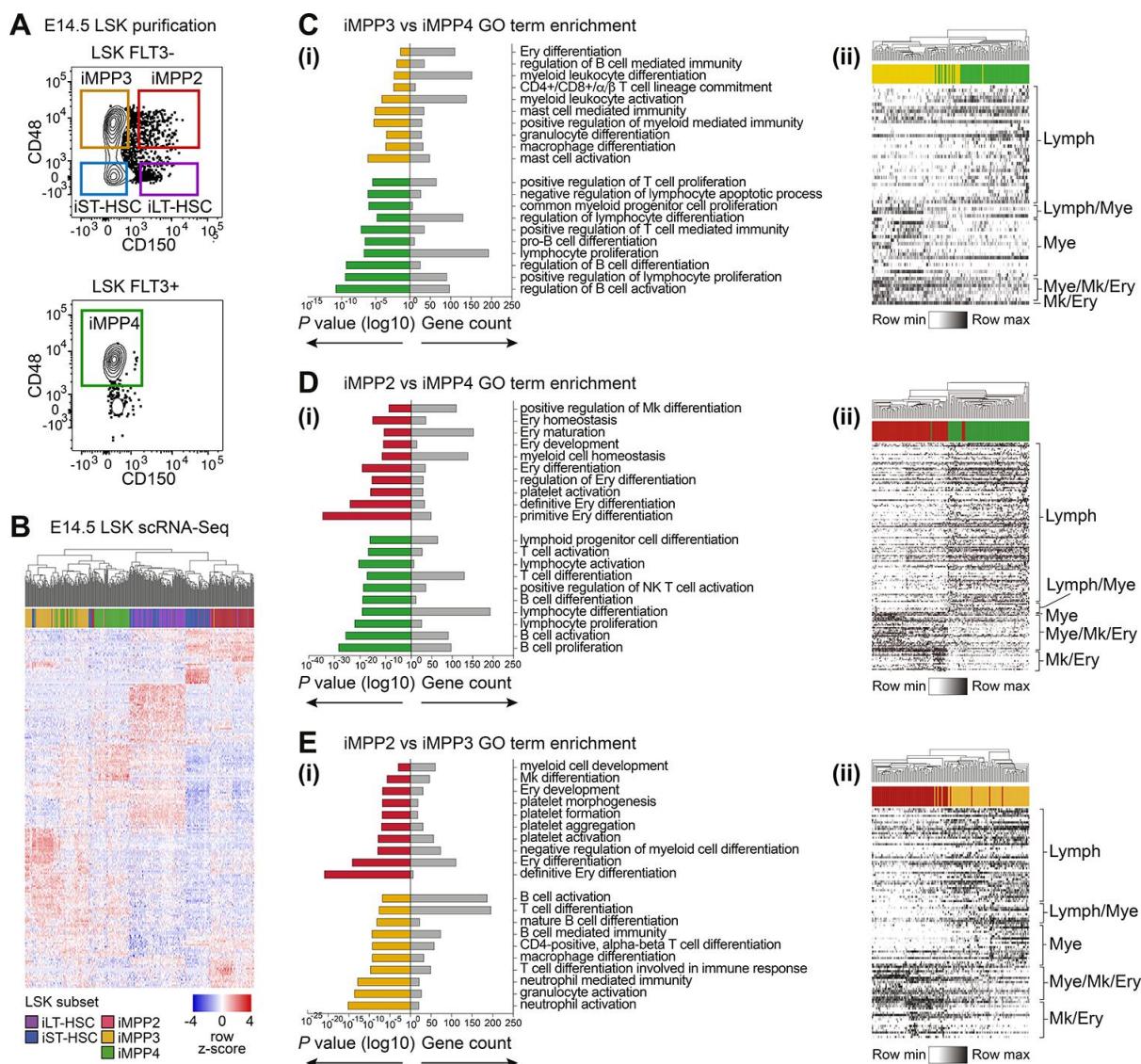
**(A – C)** Representative flow cytometry plots showing the identification of LSK subsets in the E11.5 fetal liver (FL) **(A)**, E14.5 fetal liver **(B)** and adult bone marrow **(C)**.

**(D)** Quantification iLT- and iST-HSCs in the fetal liver between E11.5 – E14.5. E11.5  $n = 7$  individual embryos; E12.5, E13.5, and E14.5  $n = 6$  individual embryos per developmental stage.

**(E)** Quantification of iMPP2, iMPP3, and iMPP4 in the fetal liver between E11.5 – E14.5. E11.5  $n = 7$  individual embryos; E12.5, E13.5, and E14.5  $n = 6$  individual embryos per developmental stage.

iMPP2, immunophenotypic multipotent progenitor 2. iMPP3, immunophenotypic multipotent progenitor 3. iMPP4, immunophenotypic multipotent progenitor 4. iLT-HSC, immunophenotypic long-term hematopoietic stem cell (iLT-HSC). iST-HSC, immunophenotypic short-term hematopoietic stem cell.

**Figure 2**



**Figure 2. Transcriptional investigation of iMPP subsets in the E14.5 fetal liver.**

**(A)** Representative example of the gating strategy used to purify LSK subsets from the E14.5 fetal liver.

**(B)** Heatmap of the top 100 most variable genes for each E14.5 FL LSK subtype.

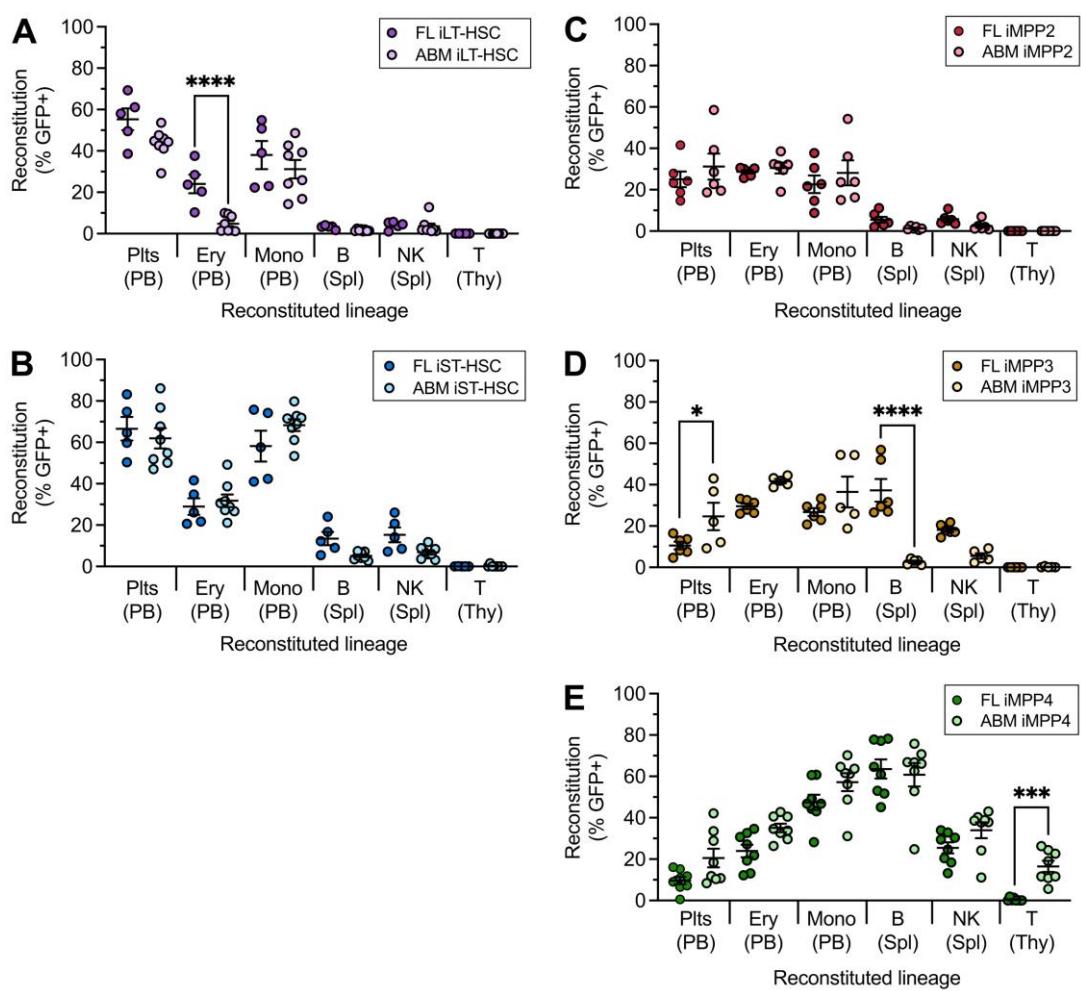
**(C)** Gene Ontogeny (GO) term enrichment analysis for significantly differentially expressed genes for iMPP3 versus iMPP4 (i). (ii) Heatmap of genes present in significantly enriched GO terms.

**(D)** Gene Ontogeny (GO) term enrichment analysis for significantly differentially expressed genes for iMPP2 versus iMPP4 (i). (ii) Heatmap of genes present in significantly enriched GO terms.

**(E)** Gene Ontogeny (GO) term enrichment analysis for significantly differentially expressed genes for iMPP2 versus iMPP3 (i). (ii) Heatmap of genes present in significantly enriched GO terms.

Fetal liver (FL). Lineage-KIT<sup>+</sup>SCA1<sup>+</sup> (LSK). iMPP2, immunophenotypic multipotent progenitor 2. iMPP3, immunophenotypic multipotent progenitor 3. iMPP4, immunophenotypic multipotent progenitor 4. iLT-HSC, immunophenotypic long-term hematopoietic stem cell (iLT-HSC). iST-HSC, immunophenotypic short-term hematopoietic stem cell. Lymphoid (Lymph). Lymphoid and myeloid (Lymph/Mye). Mye (Myeloid). Mk/Ery (Megakaryocyte and erythroid). Myeloid, megakaryocyte and erythroid (Mye/Mk/Ery).

**Figure 3**



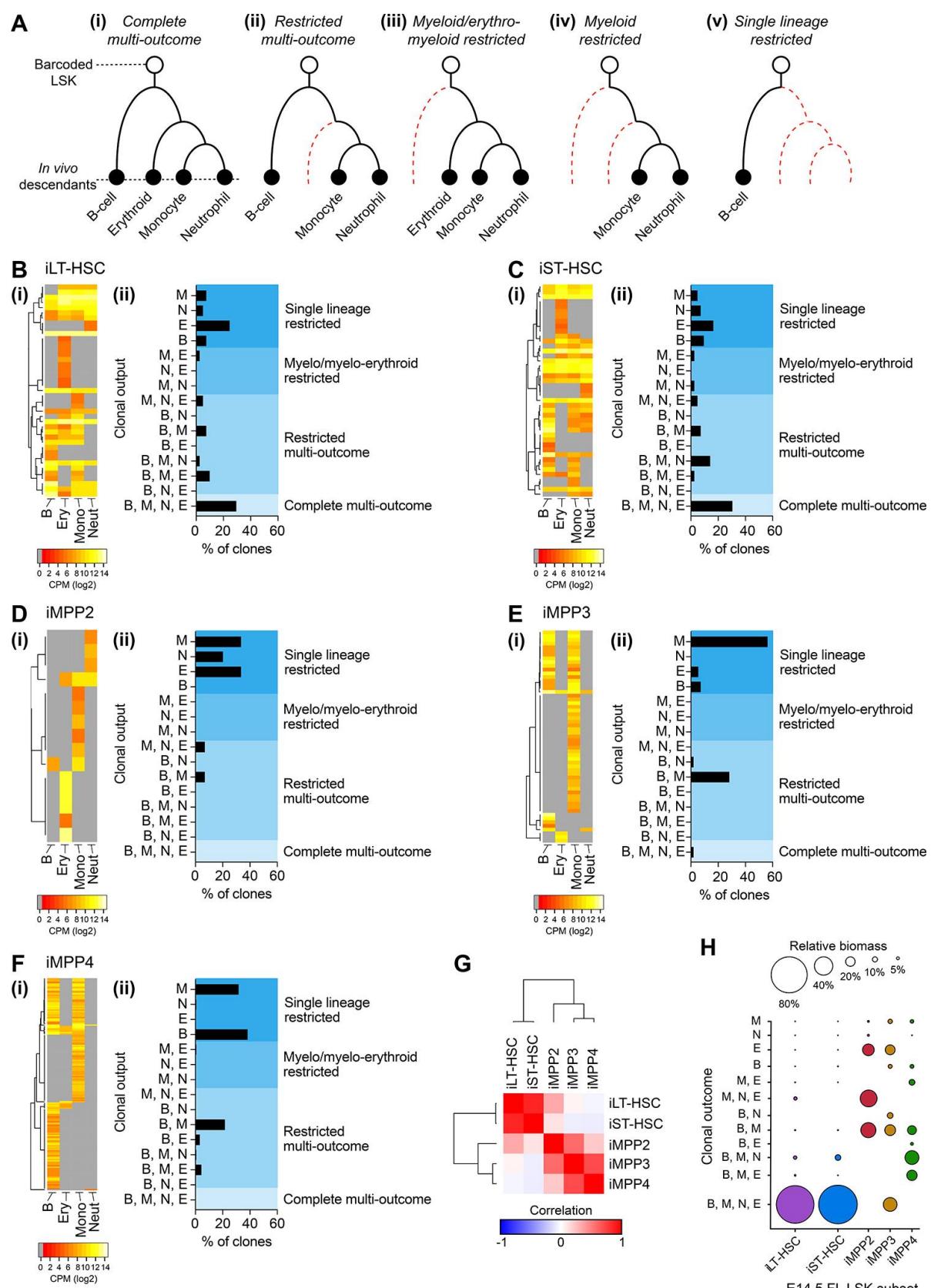
**Figure 3. Functional comparison of E14.5 FL LSK subsets with ABM counterparts in acute-term hematopoietic reconstitution assays.**

Comparison of multi-lineage reconstitution between E14.5 FL and ABM iLT-HSCs (A), iST-HSCs (B), iMPP2 (C), iMPP3 (D), and iMPP4 (E) two weeks after transplantation. Data were analyzed using one-way ANOVA. The Šídák test was used to correct for multiple comparisons. ns, not statistically significant. \*,  $p < 0.05$ . \*\*,  $p < 0.005$ , \*\*\*,  $p < 0.0005$ , \*\*\*\*,  $p < 0.00005$ .

ABM iLT-HSC,  $n = 8$  independent recipient mice. ABM iST-HSC,  $n = 8$  independent recipient mice. E14.5 iLT-HSC,  $n = 4$  independent recipient mice. E14.5 iST-HSC  $n = 5$  independent recipient mice. ABM iMPP2,  $n = 6$  independent recipient mice. ABM iMPP3,  $n = 5$  independent recipient mice. ABM iMPP4,  $n = 8$  independent recipient mice. E14.5 iMPP2,  $n = 6$  independent recipient mice. E14.5 iMPP3,  $n = 6$  independent recipient mice. E14.5 iMPP4,  $n = 7$  independent recipient mice.

iMPP2, immunophenotypic multipotent progenitor 2. iMPP3, immunophenotypic multipotent progenitor 3. iMPP4, immunophenotypic multipotent progenitor 4. iLT-HSC, immunophenotypic long-term hematopoietic stem cell (iLT-HSC). iST-HSC, immunophenotypic short-term hematopoietic stem cell. Plt, Platelet. Ery, erythroid. Mono, Monocyte. B, B-cell. T, T-cell. NK, natural killer cell. Parentheses, analyzed tissue. PB, peripheral blood. Spl, spleen. Thy, thymus.

**Figure 4**



**Figure 4. Clonal multilineage outcomes are restricted to the E14.5 iLT-HSCs and iST-HSCs.**

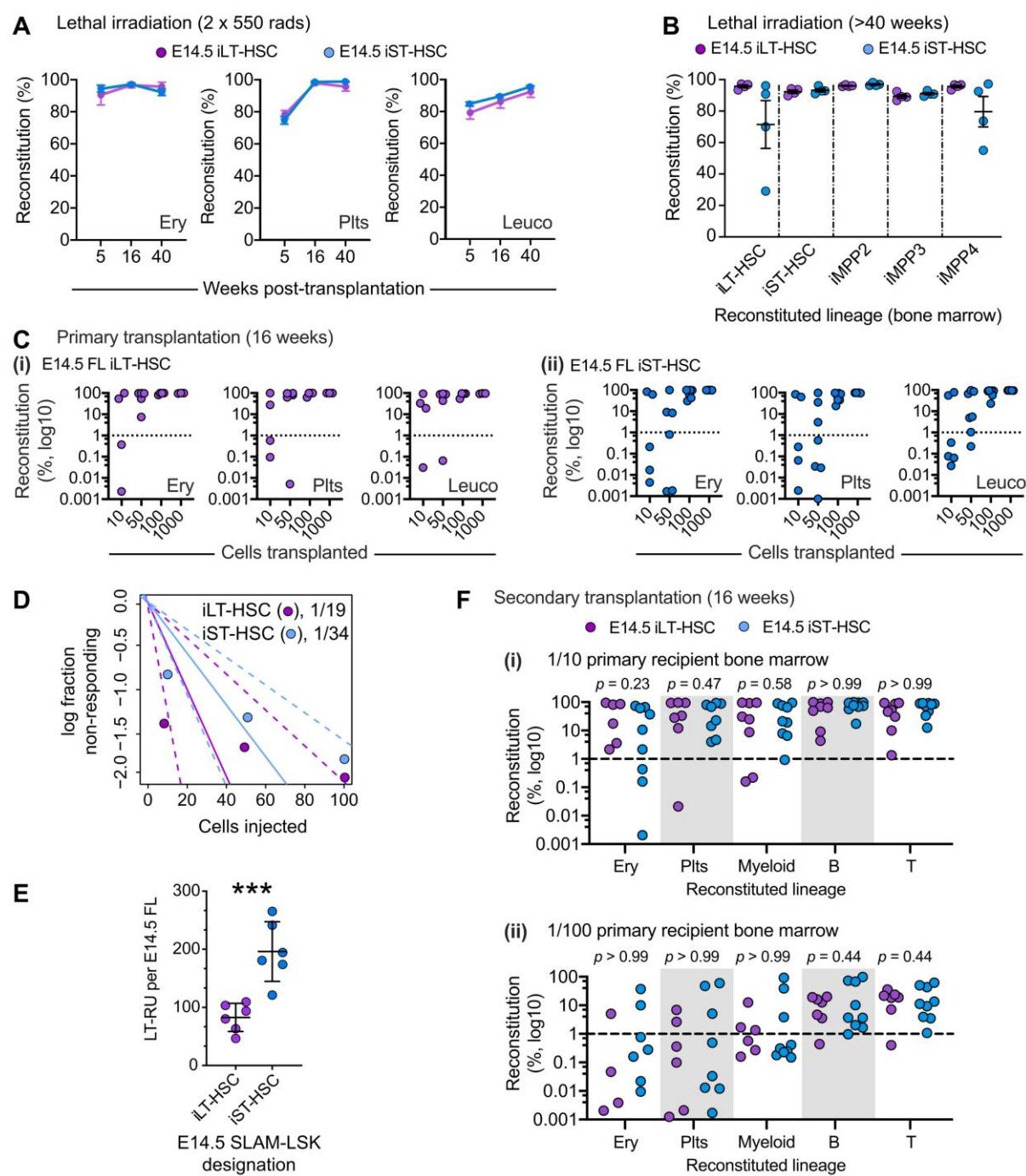
**(A)** Clonal output from barcoded E14.5 LSK subsets was determined by retrieving barcodes in donor-derived hematopoietic lineages two weeks after transplantation. Based on the lineages with shared barcodes, clones were classified as having either: **(i)** Complete multi-lineage outcome (complete multi-outcome) when the same barcode was retrieved from each lineage; **(ii)** Restricted multi-outcome when contribution to B-cells and to at least one of the erythro-myeloid lineages; **(iii)** Erythro-myeloid restricted when contribution to only erythroblasts, monocytes, and granulocytes was observed; **(iv)** Myeloid restricted when contribution to only monocytes, and granulocytes was observed; **(v)** Single-lineage restricted when contribution to only one of the lineages was observed.

**(B - F)** Heatmaps of barcodes reads **(i)** and histograms summarising the frequencies of clonal outputs **(ii)** for iLT-HSC (**B**,  $n = 41$  clones), iST-HSC (**C**,  $n = 43$  clones), iMPP2 (**D**,  $n = 15$  clones), iMPP3 (**E**,  $n = 57$  clones), and iMPP4 (**F**,  $n = 168$  clones). M, monocyte. N, neutrophil. E, erythroid cell. B, B-cell. CPM, barcode sequence counts per million reads.

**(G)** Similarity matrix analysis of global clonal outcomes of E14.5 LSK subsets.

**(H)** Bubble plot summary of the proportion of reads associated with specific clonal outcomes (relative biomass) for the E14.5 LSK subtypes. Rows are the clonal outcomes (lineages produced). Size of the bubble represents the relative biomass contribution. M, monocyte. N, neutrophil. E, erythroid cell. B, B-cell.

**Figure 5**



**Figure 5. Large numbers of LT-RUs are concealed in the E14.5 iST-HSC immunophenotype.**

**(A)** 1,000 E14.5 FL iLT-HSCs (purple) or iST-HSCs (blue) were transplanted into lethally irradiated (2 x 550 rads) recipients. Peripheral blood platelets (Plt), erythrocytes (Ery), and leucocytes (leuco) were analyzed at 5, 16, and 40 weeks post transplantation. iLT-HSCs,  $n = 10$  recipient mice. iST-HSCs,  $n = 11$  recipient mice.

**(B)** Reconstitution of ABM LSK subsets following transplantation with 1,000 E14.5 FL iLT-HSCs or iST-HSCs in lethally irradiated recipients >40 weeks after transplantation. iLT-HSCs,  $n = 4$  recipient mice. iST-HSCs,  $n = 4$  recipient mice.

**(C)** Titration of E14.5 iLT-HSCs **(i)** or iST-HSCs **(ii)**. 10, 50, 100 or 1000 cells were injected into lethally irradiated recipients. Peripheral blood platelets (Plt), erythrocytes (Ery), and leucocytes (leuco) were analyzed 16 weeks after transplantation. Dashed line represents the 1% reconstitution threshold.

**(D)** Limiting dilution analysis was used to estimate long-term repopulating unit (LT-RU) frequency in the E14.5 iLT-HSC and iST-HSC populations. LT-RU frequency was 1/19 of iLT-HSCs and 1/34 of iST-HSCs. Dotted lines represent 95% confidence interval for each population.

**(E)** Estimation of absolute numbers of LT-RUs in the E14.5 iLT-HSC and iST-HSC populations. iLT-HSC  $n = 6$  individual embryos. iST-HSC  $n = 6$  individual embryos.

**(F)** Secondary transplantations were performed using 1/10 **(i)** or 1/100 **(ii)** doses of primary recipients that had received either a 100-cell dose of E14.5 iLT-HSCs or of iST-HSCs. Peripheral blood platelets (Plt), erythrocytes (Ery), and leucocytes (leuco) were analyzed 16 weeks after transplantation. For (i) iLT-HSC secondary recipients,  $n = 6$  secondary recipients from 3 independent primary recipients; iST-HSC secondary recipients,  $n = 9$  secondary recipients from 3 independent primary recipients. For (ii) iLT-HSC secondary recipients,  $n = 6$  secondary recipients from 3 independent primary recipients; iST-HSC secondary recipients,  $n = 8$  secondary recipients from 3 independent primary recipients. Dashed line represents the 1% reconstitution threshold.  $p$  values are derived from contingency analysis (using Fisher's exact test) of reconstitution outcomes (above versus below the 1% reconstitution threshold) from iLT-HSCs and iST-HSCs for each lineage.

**Figure 6**

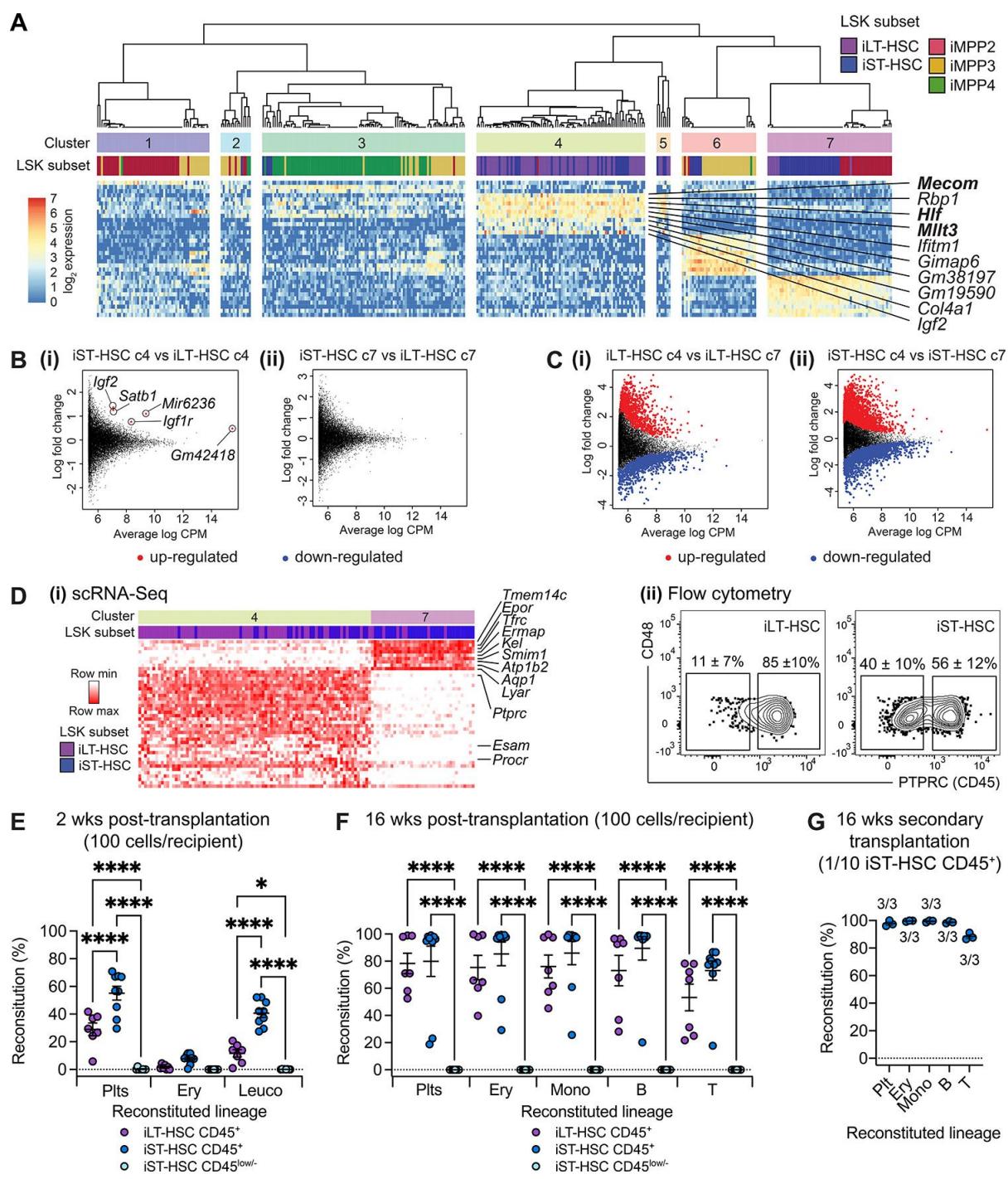


Figure 6. Use of single-cell transcriptomics to define a more refined iST:LT-RU immunophenotype.

**(A)** E14.5 FL LSK subsets can be segregated into seven transcriptional clusters based on single-cell RNA-Sequencing data. Heatmap represents marker genes for each transcriptional cluster. Note the enrichment of HSC-associated genes in cluster 4 which are composed of cells from both iLT-HSC and iST-HSC populations.

**(B)** Intra-cluster mean average (MA) plots of iLT-HSCs versus iST-HSCs in cluster 4 **(i)** and iLT-HSCs versus iST-HSCs in cluster 7 **(ii)**.

**(C)** Inter-cluster mean average (MA) plots of iLT-HSCs in cluster 4 versus cluster 7 **(i)** and iST-HSCs in cluster 4 versus cluster 7 **(ii)**.

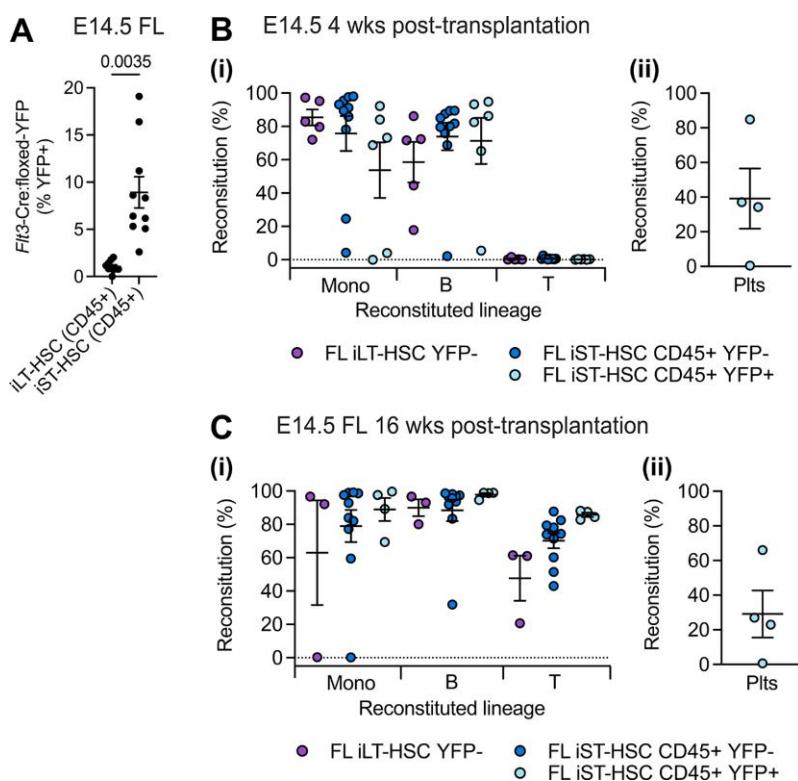
**(D)** **(i)** Heatmap of gene differentially expressed between iST-HSCs in cluster 4 versus cluster 7.

**(ii)** Representative flow cytometry profiles of PTPRC (CD45) expression by E14.5 FL iLT-HSCs and iST-HSCs.  $n = 14$  independent embryos. Inset values represent mean  $\pm$  SEM.

**(E – F)** Hematopoietic reconstitution after transplantation of 100 iLT-HSC CD45 $^{+}$  cells, 100 iST-HSC CD45 $^{+}$  cells, or 100 iST-HSC CD45 $^{\text{low/-}}$  cells into lethally irradiated mice. Peripheral blood Plts, Ery, and Leuco lineages were analyzed 2 **(E)** or 16 **(F)** weeks after transplantation. Data were analyzed using one-way ANOVA. The Šídák test was used to correct for multiple comparisons. ns, not statistically significant. \*,  $p < 0.05$ . \*\*\*\*,  $p < 0.00005$ . iLT-HSC CD45 $^{+}$ ,  $n = 7$  independent recipient mice. iST-HSC CD45 $^{+}$ ,  $n = 9$  independent recipient mice. iST-HSC CD45 $^{\text{low/-}}$ ,  $n = 8$  independent recipient mice. Dashed line represents the 1% reconstitution threshold.

**(G)** Secondary transplantations were performed using 1/10 doses of primary recipients that had received iST-HSC CD45 $^{+}$  cells. Primary recipients had originally received a 100-cell dose of iST-HSC CD45 $^{+}$  cells. Peripheral blood Plt, Ery, and Leuco lineages were analyzed 16 weeks after transplantation.  $n = 3$  secondary recipients from 2 independent primary recipients. Inset values represent the frequency of reconstitution secondary recipients. Dashed line represents the 1% reconstitution threshold. Plt, platelets. Ery, erythrocytes. Leuco, leucocytes.

**Figure 7**



**Figure 7. E14.5 iST-HSCs contain LT-RUs of *Flt3*-cre ancestry that have lineage balanced *in vivo* outputs.**

**(A)** Quantification of *Flt3*-cre driven recombination of the floxed-YFP locus in E14.5 FL iLT-HSC and iST-HSC CD45+ populations.  $n = 10$  independent mice. Data were analyzed using the unpaired two-tailed student T-test. Exact p value is shown.

**(B and C)** E14.5 FL HSCs were fractioned according to *Flt3*-cre:floxed-YFP labelling. 100 cells YFP+ or YFP- cells from population were injected into lethally irradiated recipients and analyzed after 4 **(B)** and 16 **(C)** weeks. Due to the paucity of iLT-HSC YFP+ cells, this fraction was not transplanted. **(i)** Contribution to peripheral blood monocytes (Mono), B-cells (B), and T-cells (T). **(ii)** Reconstitution of peripheral blood platelet (Plts). Data were analyzed using one-way ANOVA. The Šídák test was used to correct for multiple comparisons. Only comparisons where  $p < 0.05$  are displayed. Exact p values are shown. iST-HSC CD45+ YFP-,  $n = 10$  independent recipient mice. iST-HSC CD45+ YFP+,  $n = 6$  independent recipient mice.

## Stonehouse *et al* Supplementary information

### Supplementary methods

**Table S1.** List of antibodies used.

**Table S2.** Differential gene expression between E14.5 FL iMPP4 and iMPP3.

**Table S3.** Differential gene expression between E14.5 FL iMPP4 and iMPP2.

**Table S4.** Differential gene expression between E14.5 FL iMPP2 and iMPP3.

**Table S5.** Genes associated with iMPP-derived GO terms.

**Table S6.** Inter-cluster 4 and 7 differential gene expression between E14.5 FL iLT-HSCs.

**Table S7.** Inter-cluster 4 and 7 differential gene expression between E14.5 FL iST-HSCs.

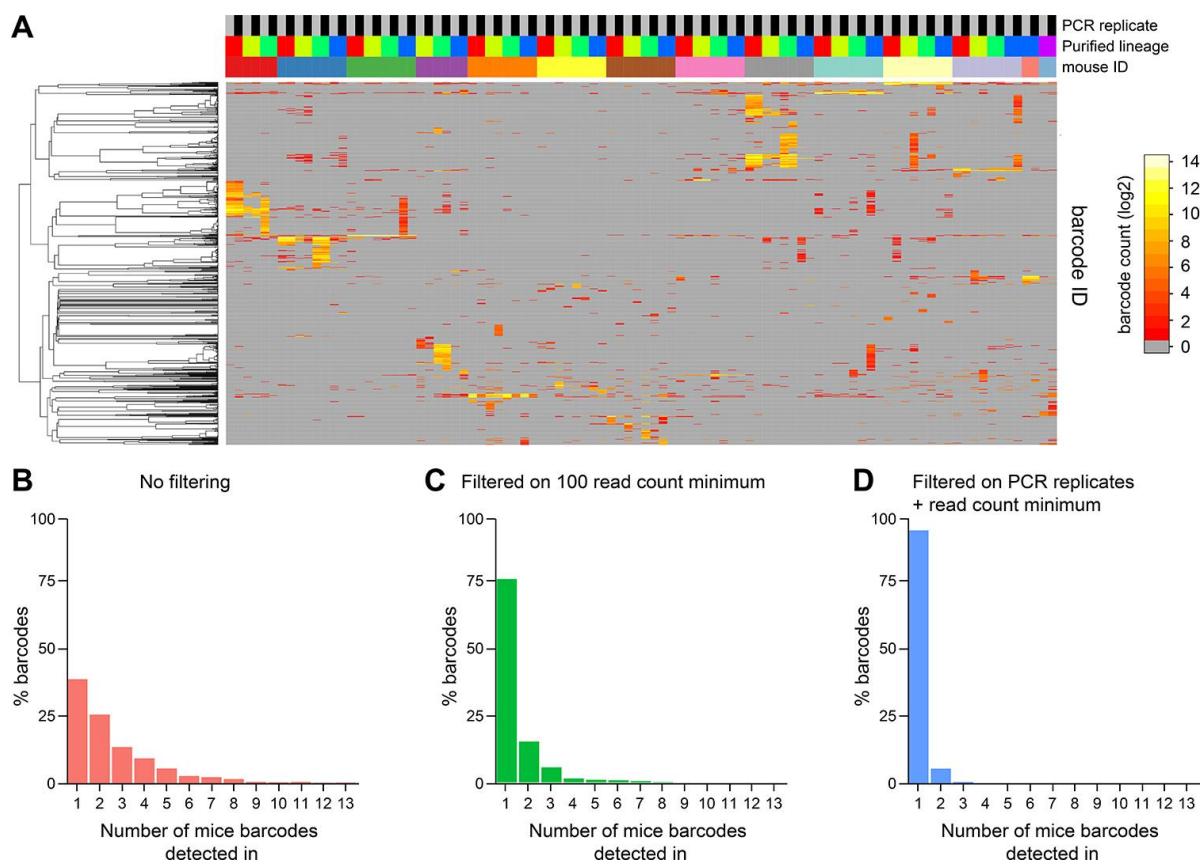
**Table S8.** Intra-cluster 4 differential gene expression between E14.5 FL iLT-HSCs and iST-HSCs.

**Table S9.** Intra-cluster 7 differential gene expression between E14.5 FL iLT-HSCs and iST-HSCs.

**Table S10.** Inter-cluster 4 and 7 differential gene expression - genes with plasma membrane associated products.

## Supplementary Figures

**Figure S1**

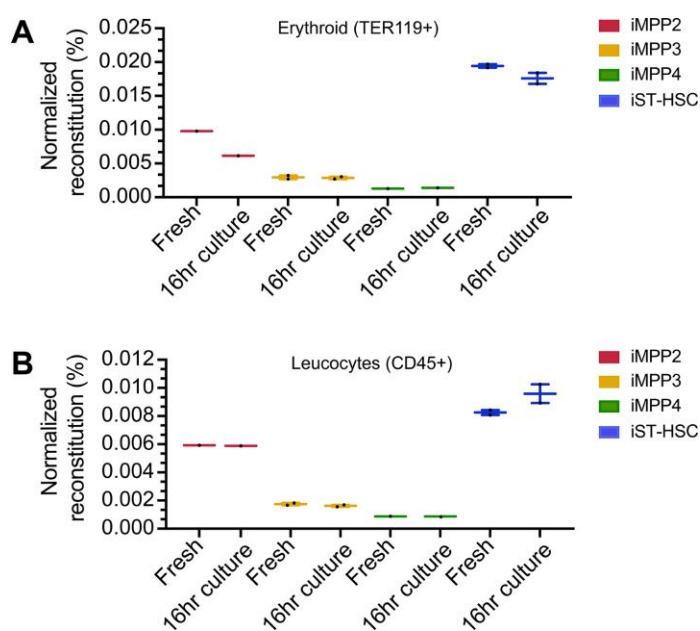


**Figure S1. Filtering lentiviral barcoding sequences.**

**(A)** Heatmap of all barcode sequences detected before filtering in the experimental cohort of 14 animals. Mouse ID are color-coded. Purified lineage: Red: B cells, Yellow: Erythrocytes, Green: Monocytes, Blue: Neutrophils, Purple: blank. 2 PCR replicates (grey/black) were performed for each sample.

**(B – D)** Barcode sharing between all cohorts of experimental mice with no data filtering **(B)**, after removal of low read counts ("min\_read") **(C)**, and after PCR replicate filtering ("PCR rep") and "min\_read" filtering **(D)**. Data analyzed in this study were based on "PCR rep" + "min\_read" filtering. This shows that >95% bona fide barcodes are detected in one mouse only, suggesting that repeated usage of barcodes is unlikely.

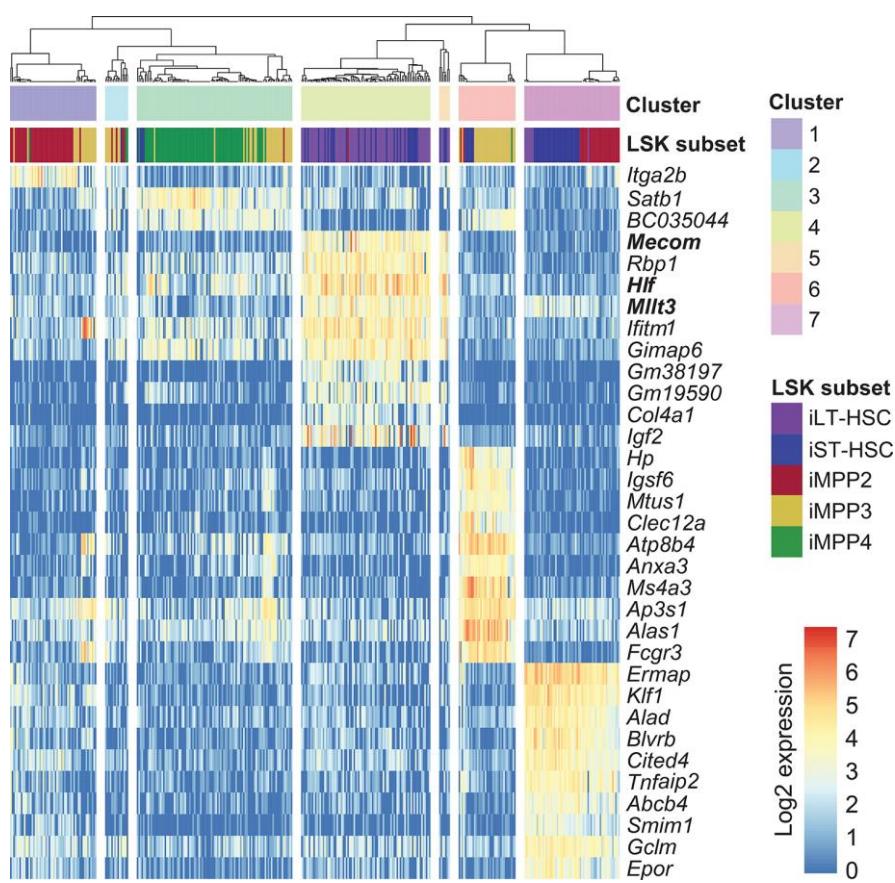
**Figure S2**



**Figure S2. 16 hrs of *ex vivo* culture did not affect LSK subset function.**

Normalized reconstitution per cell two weeks after transplantation of freshly isolated E14.5 FL LSK subsets or E14.5 FL LSK subsets following 16 hrs of *in vitro* culture. **(A)** Reconstitution in the erythroid lineage. **(B)** Reconstitution in leucocytes. iMPP2,  $n = 1$  recipient. iMPP3,  $n = 2$  independent recipients. iMPP4,  $n = 1$  recipient. iST-HSC,  $n = 2$  independent recipients.

**Figure S3**



**Figure S3. Signature genes of E14.5 FL LSK transcriptional clusters.** Heatmap of E14.5 FL LSK transcriptional cluster signature genes (linked to Figure 6A).