



OPEN ACCESS

EDITED BY

Selim Kuci,
University Hospital Frankfurt, Germany

REVIEWED BY

Armel Hervé Nwabo Kamdje,
University of Garoua, Cameroon
Naseem Ahamad,
The University of Texas Health Science Center
at San Antonio, United States
Frederic Deschaseaux,
Université Toulouse III Paul Sabatier, France

*CORRESPONDENCE

Magdalena Kucia,
✉ magdalena.kucia@wum.edu.pl

RECEIVED 10 March 2025

ACCEPTED 29 April 2025

PUBLISHED 15 May 2025

CITATION

Jarczak J, Kieszek P, Ratajczak MZ and Kucia M (2025) Optimizing single cell RNA sequencing of stem cells. A streamlined workflow for enhanced sensitivity and reproducibility in hematopoietic studies. The use of human umbilical cord blood-derived hematopoietic stem and progenitor cells. *Front. Cell Dev. Biol.* 13:1590889. doi: 10.3389/fcell.2025.1590889

COPYRIGHT

© 2025 Jarczak, Kieszek, Ratajczak and Kucia. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Optimizing single cell RNA sequencing of stem cells. A streamlined workflow for enhanced sensitivity and reproducibility in hematopoietic studies. The use of human umbilical cord blood-derived hematopoietic stem and progenitor cells

Justyna Jarczak ¹, Patrycja Kieszek ¹,
Mariusz Z. Ratajczak ^{1,2} and Magdalena Kucia ^{1*}

¹Laboratory of Regenerative Medicine, Preclinical Research and Technology Center, Medical University of Warsaw, Warsaw, Poland, ²Stem Cell Institute at Graham Brown Cancer Center, University of Louisville, Louisville, CO, United States

Background: Human hematopoietic stem/progenitor cells (HSPCs) are enriched in umbilical cord blood (UCB) among cell populations that express CD34 and CD133 (PROM1) antigens. These cells can be purified further and sorted by FACS as CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ cells. It has been postulated that the population of CD133⁺ HSPCs is enriched for more primitive stem cells. To address this issue at the molecular level, we performed single-cell RNA-sequencing (scRNA-seq) and analyzed the transcriptome of both cell types. We optimized the available protocols of scRNA-seq of HSPC and described our laboratory experiences with the limited number of cells obtained from human UCB.

Results: Herein, we report the results of scRNA-seq analysis paying special attention to the quality parameters of single cell libraries. We also present the similarities and differences in transcriptome between these cells (CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs) and their subpopulations identified and visualized as clusters using uniform manifold approximation and projection (UMAP), stressing the need for an integrated analysis of both datasets, which may be merged and treated as “pseudobulk.” We revealed that both populations do not differ significantly in gene expression, as evidenced by the very strong positive linear relationship between these cells ($R = 0.99$).

Conclusion: To obtain solid results that allow to draw conclusions that would have a biological translation, all parts of the scRNA-seq experiment are crucial and must be carried out with due care: cell sorting, single cell libraries preparation, quality control, and data analysis. The idea of working

with sorted material instead of the typical use of a full pellet of blood cells was right and confirmed the possibility of HSPC analysis, even with a limited number of cells.

KEYWORDS

HSPCs, single-cell RNA seq, clustering, marker detection, UCB, CD34⁺ cells, CD133⁺ cells, CD45⁺ cells

Background

Hematopoietic stem/progenitor cells (HSPCs) are specified into cells from all hematopoietic lineages, including erythrocytes, granulocytes, monocytes, lymphocytes, and megakaryocytes (Bhatia et al., 1997; McKinney-Freeman et al., 2012; Morrison et al., 1995; Andrews et al., 1989; Mikkola and Orkin, 2006). Different strategies are currently available to purify these cells; however, isolating them using a single marker with an FACS sorter is impossible (Wognum et al., 2003; Ng et al., 2009; Challen et al., 2009). Therefore, to enrich human bone marrow (BM), peripheral blood (PB), or umbilical cord blood (UCB)-derived HSPCs, a cocktail of antibodies for simultaneous positive and negative marker selection has to be employed (Rossi et al., 2011). HSPCs could be relatively quickly enriched after positive labeling with antibodies against CD34 and/or CD133 and CD45 antigens, along with the depletion of cells expressing lineage differentiation markers (Kato and Radbruch, 1993; Berardi et al., 1995; Notta et al., 2011; Bujko et al., 2019). Panels of cell surface markers including cell clustering, are employed to identify the intermediate stages of the hematopoietic specification. In this concept, HSPCs give rise to progressively lineage-restricted cell types, sometimes depicted as a “hematopoietic tree,” until adult blood cells are reached (Laurenti and Gottgens, 2018; Metcalf, 2007). This historical paradigm has changed in the past 5 years due to multiple studies that have reported the transcriptomes of thousands of individual hematopoietic cells in adult humans and mice, separated by cell surface markers (Calvanese and Mikkola, 2023). It has been suggested that a hematopoietic cell can be “locked” into a specific cell destiny by the stochastic production of lineage-specific transcription factors (TFs) over the noise threshold (Graf and Enver, 2009). It explains why HSPCs have been shown to co-express at low-level genes encoding essential transcription factors (TFs) linked to opposing lineages (Chambers et al., 2007; Hu et al., 1997; Lawrence et al., 2005; Miyamoto et al., 2002). This supports that there are subpopulations of cells in the HSPCs compartment that are permissive to pursue different cell fates before committing to a given lineage - a process known as priming (Ranzoni et al., 2021; Krause, 2002; Milner and Bigas, 1999).

Single-cell RNA-sequencing (scRNA-seq) is a method that analyzes the transcriptome at single-cell and single-base resolution. This method enables unraveling gene expression networks in rare cell types and demonstrates the heterogeneity in gene expression within temporally and spatially separated cell populations (Fast et al., 2021; Oguma et al., 2022; Zheng et al., 2022; Crosse et al., 2020; Farlik et al., 2016). This method (scRNAseq) is becoming an increasingly widely used tool in biological and biomedical research and has advanced the understanding of a range of biological processes and molecular mechanisms occurring in cells. It potentially may have a clinical application and may be valuable

in new diagnostic and therapeutic strategies (Su et al., 2022). In the case of hematopoietic stem cells, scRNA-seq allowed to decode both human (Popescu et al., 2019) and mice (Flohr et al., 2021) hematopoiesis. Furthermore, the first hematopoietic stem cells in human embryo were traced by scRNA-seq (Zeng et al., 2019). Somehow, to our surprise, no studies have yet been performed based on scRNA-seq to compare CD34⁺ and CD133⁺ HSPCs from UCB. Therefore, we purified from human UCB a population of CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HPSCs and subjected these cells to scRNA-seq analysis.

We report the transcriptomic profile of these cells, paying particular attention to the similarities and differences in mRNA expression. Based on this expression, we propose their classification into clusters of cell subpopulations. Importantly, we emphasized the hope that is placed in this type of research and the still existing need to explore the mysteries of the development and functioning of human hematopoietic stem/progenitor cells.

Materials and methods

Isolation of human CD34⁺ and CD133⁺ HSPCs

Human umbilical cord blood (hUCB) was obtained from a healthy newborn delivered at the Department of Obstetrics and Gynecology, Medical University of Warsaw (Warsaw Bioethics Committee permission number KB/3/2018) on 14th of September 2023. Participant gave written consent to participate in the study.

hUCB unit was diluted with phosphate-buffered saline (PBS) and carefully layered over Ficoll-Paque (GE Healthcare, Chicago, IL, United States) and centrifuged for 30 min at 400x g at 4°C. The mononuclear cells (MNCs) phase was collected, washed, and used for further analysis. Briefly, cells were stained with the following antibodies: hematopoietic lineage markers (Lin) cocktail of antibodies, each FITC-conjugated: CD235a (clone GA-R2 [HIR2]), anti-CD2 (clone RPA-2.10), anti-CD3 (clone UCHT1), anti-CD14 (clone M5E2), anti-CD16 (clone 3G8), anti-CD19 (clone HIB19), anti-CD24 (clone ML5), anti-CD56 (clone NCAM16.2) and anti-CD66b (clone G10F5) (all BD Biosciences, San Jose, CA, United States); PE-Cy7-conjugated anti-CD45 (clone HI30, BioLegend, San Diego, CA, United States), PE-conjugated anti-CD34 (clone 581, BioLegend, San Diego, CA, United States) and APC-conjugated anti-CD133 (clone CD133, MiltenyiBiotec, Gladbach, Germany). Antibodies were used in the manufacturer's recommended concentration. Cells were stained in the dark at 4°C for 30 min, then centrifuged and resuspended in RPMI-1640 medium containing 2% fetal bovine serum (FBS, Corning Inc., Corning, NY, United States). Cells were sorted

according to the strategy shown in [Supplementary Figure S1](#). Briefly, small events (2–15 μm in size) were included in the “lymphocyte-like” gate (P1) and then further analyzed for the expression of the Lin marker. Lin negative events were gated and subsequently analyzed for the expression of CD45 and CD34 or CD133 antigens ([Supplementary Figure S1](#)). Populations of CD34⁺ HSPCs (CD34⁺Lin[−]CD45⁺) and CD133⁺ HSPCs (CD133⁺Lin[−]CD45⁺) were sorted on the MoFlo Astrios EQ cell sorter (Beckman Coulter, Brea, CA, United States).

Single-cell sequencing

After sorting, cells were proceeded directly using Chromium X Controller (10X Genomics, United States) and Chromium Next GEM Chip G Single Cell Kit (10X Genomics, United States). Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1, and Single Index Kit T Set A (10X Genomics, United States) were used for library preparation according to manufacturer's guidelines. Libraries were then pooled and run on Illumina NextSeq 1000/2000 (Illumina, San Diego, CA, United States) in P2 flow cell chemistry (200 cycles) with paired-end sequencing mode (read 1–28 bp, read 2–90 bp, index 1–10 cycles, index 2–10 cycles), assuming 25,000 reads per single cell.

Bioinformatic analysis

Downstream analysis was performed using Seurat (version 5.0.1), preceded by the 10X Genomics Cell Ranger pipelines (CellRanger version 7.2.0, 10x Genomics, United States, Cell Ranger - Official 10x Genomics Support). Raw sequencing (BCL files) was demultiplexed and converted to fastq files using the bcl2fastq (version v2.20.0.422) within the 10X Genomics Cell Ranger mkfastq pipeline. Then, Cell Ranger count and aggregation pipelines were used for further processing. Cells with less than 200 and more than 2500 transcripts and those with more than 5% of mitochondria-related transcripts were excluded from the analysis. Sequencing reads were mapped to a human genome GRCh38 (version 2020-A) acquired from the 10 × Genomics website (<https://www.10xgenomics.com/support/software/cell-ranger/downloads#reference-downloads>). Gene expression measurements for each cell were normalized, and the normalized values were log-transformed (“LogNormalize” method) and then reduced to the most highly variable genes. Based on the information described in [Brennecke et al. \(2013\)](#), where authors calculated a subset of features that exhibit high cell-to-cell variation in the dataset, we used the default value of 2000 genes.

Non-linear dimensional reduction to visualize clusters was performed next with the use of uniform manifold approximation and d projection (UMAP) ([McInnes et al., 2018](#)) implemented in Seurat (version 5.0.1.) ([Satija et al., 2015](#); [Hao et al., 2024](#)). Cell clusters were recognized based on differentially expressed genes, both positive (upregulated genes) and negative (downregulated genes) markers, using an adjusted p-value (padj) (Seurat uses the Wilcoxon rank-sum test by default, padj values are after Benjamini–Hochberg correction). <0.05 and $\log_2\text{FC} > 1$. However, for this publication, we focused only on positive ones. These genes

were subsequently used for the functional analysis (Gene Ontology terms and KEGG pathways) and characterization of the identified samples and clusters. Functional analysis was performed using the Reactome Pathway Browser ([Orlic-Milacic et al., 2024](#)). Analysis was conducted in two ways: independently to thoroughly characterize the samples separately and integrated to visualize the differences between them.

Additionally, we selected hematopoietic lineage differentiation cell markers (CD2, CD3, CD4, CD11b (ITGAM), CD14, CD16 (FCGR3A), CD19, CD34, CD41 (ITGA2B), CD45 (PTPRC), CD66b (CEACAM8), CD68, GYPA, PROM1 (CD133), CD117 (c-KIT) to determine their expression in CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ and confirm the hematopoietic character of the studied cells.

Data analysis was done using R (R Core Team (2021) ([R Core Team, 2024](#)), while visualization was performed using the ggplot2 R package [version 3.4.4; ([Wickham, 2016](#))].

Data availability

The dataset(s) supporting the conclusions of this article are available in the Sequence Read Archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra>) repository and assigned unique persistent identifiers: PRJNA1128409 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA1128409>).

Results and discussion

Sorting of UCB-purified CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ HSPCs

We used multiparameter cell sorting, which enabled us to purify two subpopulations of HSPCs, CD34⁺Lin[−]CD45⁺, and CD133⁺Lin[−]CD45⁺, from hUCB MNCs isolated by Ficoll-Paque gradient centrifugation. Our sorting strategy included depleting Lineage marker-expressing cells, followed by cell isolation based on their simultaneous expression of CD45 and CD34 or CD133 antigen ([Abdelbaset-Ismail et al., 2016](#)). [Supplementary Figure S1](#) displays representative dot plots of our sorting strategy.

Briefly, UCB-derived MNCs were immunostained as described in Materials and Methods and then analyzed on a FACS machine. First, objects 2–15 μm in size were included in the “lymphocyte-like cells” gate (P1) ([Supplementary Figure S1A](#)). Next, we analyzed the expression of lineage markers and gated negative events (Lin[−]) ([Supplementary Figure S1B](#)). Subsequently, we evaluated the expression of CD45 and CD34 ([Supplementary Figure S1C](#)) and CD133 ([Supplementary Figure S1D](#)) and separately sorted the population of CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ HSPCs.

Additionally, we evaluated the number of CD133⁺Lin[−]CD45⁺ and CD34⁺Lin[−]CD45⁺ HSPCs and analyzed the co-expression of CD34 and CD133 antigens. The results were consistent with our earlier report where we analyzed a large number of hUCB units and calculated that CD34⁺Lin[−]CD45⁺ HSPCs comprised 0.337% (± 0.15) of gated events and CD133⁺Lin[−]CD45⁺ HSPC comprised 0.207% (± 0.08) of gated cells ([Bujko et al., 2023](#)).

Cells isolated from Ficoll-Paque were purified by a cell sorter (MoFlow Astrios) and subjected to droplet-based scRNA-seq (10x

Genomics) to investigate their transcriptomic profile and identify cell subpopulations.

Libraries quality control parameters and sequencing

First, we wanted to evaluate whether having a limited number of sorted cells, how it happens when working with HSPC, we are allowed to prepare libraries meeting the criteria for next-generation sequencing. According to the manufacturer's instructions, the average fragment length is ~300–600 bp and the expected insert size is ~400 bp. Quality control assessed with TapeStation 4150 (Agilent Technologies, United States) showed that all libraries had a fragment length of more than 300 bp and met the standards for sequencing (Supplementary Figures S2A,B). These were libraries prepared from isolated and sorted cell populations of HSPCs: CD133+lin-CD45⁺ (Supplementary Figure S2A), and CD34+lin-CD45⁺ (Supplementary Figure S2A). Therefore, libraries were next employed for quantitative measurements by KAPA Library Quantification Kit (Roche, Switzerland). The library concentrations were as follows: for HSPCs: CD133+lin-CD45⁺ = 173,67 nM and CD34+lin-CD45⁺ = 261,13 nM, with the mean value = 217.40 nM. Of note, sequencing parameters were as follows: Q30 = 93,99%; yield = 50,82 Gb; Aligned = 1,70 and error rate = 0,12. The mean number of reads was around 90 millions per library (Supplementary Figures S2C,D) as follows CD133+lin-CD45⁺ (Supplementary Figure S2C), and CD34+lin-CD45⁺ (Supplementary Figure S2D). Additionally, we captured 5862 CD133+Lin-CD45⁺ and 6811 CD34+Lin-CD45⁺ HSPCs that passed quality control, with a median of 2066 and 1613 genes detected per single cell, respectively (data not shown).

Single-cell RNA seq analysis of sorted human CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs

For CD34⁺Lin⁻CD45⁺, we detected as the most variable genes hemoglobin subunit gamma 2 and gamma 1 (HBG2, HBG1), hemoglobin subunit beta (HBB), hemoglobin subunit mu (HBM), and immunoglobulin kappa constant (IGKC) (Figure 1A). All these genes play a crucial role in hematopoiesis (Bao et al., 2019; Riether et al., 2015). They are involved in various physiological processes, including oxygen transport, hemoglobin assembly, DNA synthesis, mitosis, glycolysis, intracellular cAMP accumulation, and immune recognition. Importantly, we also detected differential expression of the pro-platelet basic protein (PPBP) (Figure 1A), which belongs to the CXC chemokine family and is involved in various cellular processes, including prostaglandin E2 secretion and synthesis of hyaluronic acid and sulfated glycosaminoglycan (Belperio et al., 2000; Broxmeyer and Kim, 1999). PPBP also stimulates the formation and secretion of plasminogen activators by synovial cells (Smith et al., 2015).

Similarly, for CD133⁺Lin⁻CD45⁺ HSPCs, the most differentially expressed genes were HBG1, HBA2, HBA1, HBB, HBM, HBG2, and IGKC (Figure 1B). We also noticed a differential expression of IGLC2, GNLY, and alpha hemoglobin stabilizing protein (AHSP),

which encodes a molecular chaperone that binds specifically to free alpha-globin involved in hemoglobin assembly (Weiss et al., 2005). This confirms that our UCB-purified HSPCs express a similar panel of genes involved in oxygen transfer.

Single-cell cluster analysis for the cell marker expression for sorted UCB CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs

We employed Seurat (version 5.0.1) to perform cluster analysis and differential gene expression within sorted HSPCs populations. Additionally, we used uniform manifold approximation and projection (UMAP) to investigate cellular heterogeneity, which allowed us to detect a total of 14 clusters for both CD34⁺Lin⁻CD45⁺ (Figure 2A) and CD133⁺Lin⁻CD45⁺ HSPCs libraries, respectively (Figure 2B). As expected, all clusters of CD34⁺Lin⁻CD45⁺ HSPCs except clusters 11 and 14 (Figures 3A,C) and clusters 6 and 13 of CD133+Lin-CD45⁺ HSPCs (Figures 3B,D) express at high-level PTPRC gene (CD45 marker). Within PTPRC (CD45) negative clusters in both sorted HSPCs populations, nearly all cells express glycophorin A (GYPA⁺) or tetraspanin (CD9⁺) (Figures 3C,D). While GYPA is a marker of erythroid commitment, CD9 has a diverse role in cellular processes, as it has also been shown to trigger platelet activation and aggregation (Huang et al., 2020; Worthington et al., 1990; Miao et al., 2001).

Two stem cell markers, CD34 and PROM1 (CD133), were observed in both populations (Figure 3). While, in CD34⁺Lin⁻CD45⁺ HPSC, the expression of the CD34 gene was observed in clusters no. 0, 1, 6, 7 (Figure 4A), in CD133⁺Lin⁻CD45⁺ HSPCs, the expression of the CD34 gene was observed in clusters no. 1, 4, and 11 (Figure 4C). Interestingly, in the population of CD34⁺Lin⁻CD45⁺ cells, we observed co-expression of mRNA for CD34 and PROM1 in clusters 0 and 1 only (Figures 4A,B). We suggest that these clusters are likely enriched with highly primitive HSPCs co-expressing CD34 and CD133 (PROM1). In contrast, within the CD133⁺Lin⁻CD45⁺ population, clusters 1, 4, and 11 exhibited simultaneous expression of PROM1 and CD34 mRNA (Figures 4C,D). Notably, in CD34⁺Lin⁻CD45⁺ cells, clusters 6 and 7 expressed CD34 mRNA but lacked PROM1 expression. Conversely, in CD133⁺Lin⁻CD45⁺ HSPCs, clusters 1, 4, and 11 showed co-expression of both PROM1 and CD34 transcripts. Importantly, the clusters with confirmed expression of two stem cell markers: CD34 and PROM1 (CD133) will be called later in the text as primitive or quiescent clusters.

Next, we analyzed the remaining cell markers related to hematopoietic lineage specification: (CD2, CD3, CD4, CD11b (ITGAM), CD14, CD16 (FCGR3A), CD19, CD34, CD41 (ITGA2B), CD45 (PTPRC), CD66b (CEACAM8), CD68, GYPA, PROM 1 (CD133), CD117 (c-KIT) (Figures 3C,D; Figures 5A,B). The clusters with confirmed expression of lymphocyte (T and B), monocytes, macrophages, NK cells and granulocytes markers (e. g. CD2, CD3, CD19, CD14, FCGR3A, CD68) were called 'fate decision' clusters. In sorted CD34⁺Lin⁻CD45⁺ HSPCs, T lymphocyte markers (CD2, CD3) were detected in clusters no. 5 and 13 and CD4 in clusters 2, 3, 4, and 5. In contrast, for CD133⁺Lin⁻CD45⁺ HSPCs, CD2, and CD3 were detected in cluster no. 3 and 7, and CD4 in clusters 1, 2, 3,

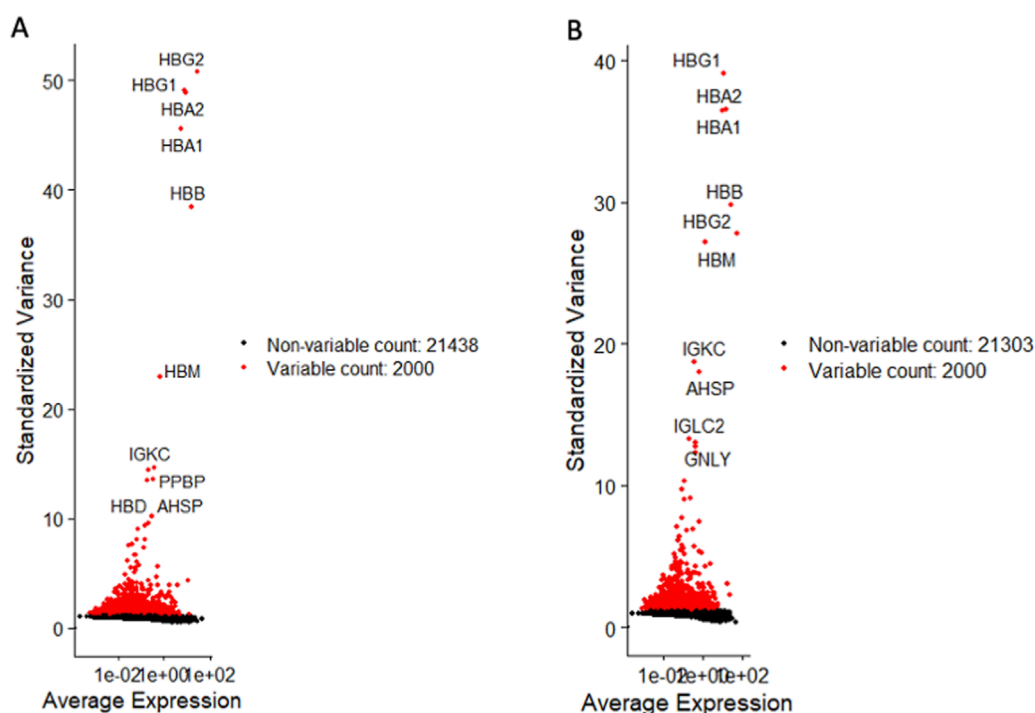


FIGURE 1
Identification of highly variable genes in both HSPC datasets. Volcano plots display differentially variable genes of CD34⁺Lin⁻CD45⁻ (A) and CD133⁺Lin⁻CD45⁻ (B). Variable genes are shown in red, while non-variable genes are shown in black.

and 4. B lymphocyte marker (CD19) was expressed in cluster no. 13 for CD34⁺Lin⁻CD45⁺ HSPC and cluster 8 for CD133⁺Lin⁻CD45⁺ HSPC. The expression of mRNA for monocytes, macrophages, NK cells, and granulocytes (CD14, FCGR3A, and CD68) was present in clusters 2, 3, and 4 (CD14), clusters 3, 4, and 10 (FCGR3A), clusters 2, 3 and 10 (CD68) of CD34⁺Lin⁻CD45⁺ HSPCs. In the case of CD133⁺Lin⁻CD45⁺ HSPCs, CD2, and CD3 were expressed in cluster 3 and CD4 in clusters 0, 2, and 3 CD19 was detected in clusters 8, CD14 in clusters 0 and 2, FCGR3A in clusters 0, 5, 7, and CD68 in clusters 0 and 2. The erythrocyte marker glycophorin A (GYPA) was found in cluster 11 for the CD34⁺Lin⁻CD45⁺ dataset and cluster 6 for the CD133⁺Lin⁻CD45⁺ HSPCs dataset. At the same time, mRNA of megakaryopoietic marker - ITGA2B (CD41) detected in clusters 14 of CD34⁺Lin⁻CD45⁺ HSPC and 13 of CD133⁺Lin⁻CD45⁺ HSPC (Figures 5A,B; Figures 3C,D). We also identified in both studied datasets the expression of mRNA for c-KIT (CD117) that, to a large extent, correlated with the expression of mRNA for CD34 and CD133 (Figures 5A,B).

In addition to markers for hematopoietic specification of sorted cells toward erythropoiesis, myelopoiesis, and thrombopoiesis, we also identified changes in the expression of genes involved in innate immunity responses. We observed, for example, the downregulation of S100A8 and S100A9 in numerous clusters of CD34⁺Lin⁻CD45⁺ cells (clusters 1, 5, 6, 7, 9, 10, 11, and 13) (Supplementary Tables S1–S9). However, at the same time, the expression of these markers was upregulated in clusters 2 and 3 (Supplementary Tables S10, S11). S100A8 and S100A9 belong to the S100 protein family, also known as myeloid-related proteins (MRPs), and are primarily expressed in myeloid

lineage cells, including neutrophils, monocytes, and macrophages (Xia et al., 2017; Takizawa et al., 2012). They belong to the alarmin family of molecules.

We observed a similar situation in the expression of S100A8 and S100A9 in CD133⁺Lin⁻CD45⁺ cells. Among the top 50 downregulated genes in many clusters, we found other proteins from the family S100 proteins, such as S100A6, S100A10, S100A11, and S100A12. Decreased expression of S100A12 was observed in clusters 1, 3, 4, 7, 8, 9 and 11 (Supplementary Tables S12–S18). This protein, also known as Calgranulin C or EN-RAGE (Extracellular Newly identified RAGE-binding protein), is primarily involved in inflammatory responses and host defense mechanisms, and is expressed and secreted by neutrophils, monocytes, and macrophages in response to inflammatory stimuli (Cheng et al., 2019).

The summary of progenitor fractions and their lineage commitment in the single-cell clusters of CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs is presented in Table 1.

Transcriptomic profiles of selected clusters of UCB CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs

Transcriptomic profiles of quiescence/primitive clusters (expressing PROM1, CD34 genes)

Next, we analyzed the transcriptomic profile of selected clusters in CD34⁺Lin⁻CD45⁺ HSPCs. Firstly, we focused on the transcriptomic differences between clusters expressing mRNA

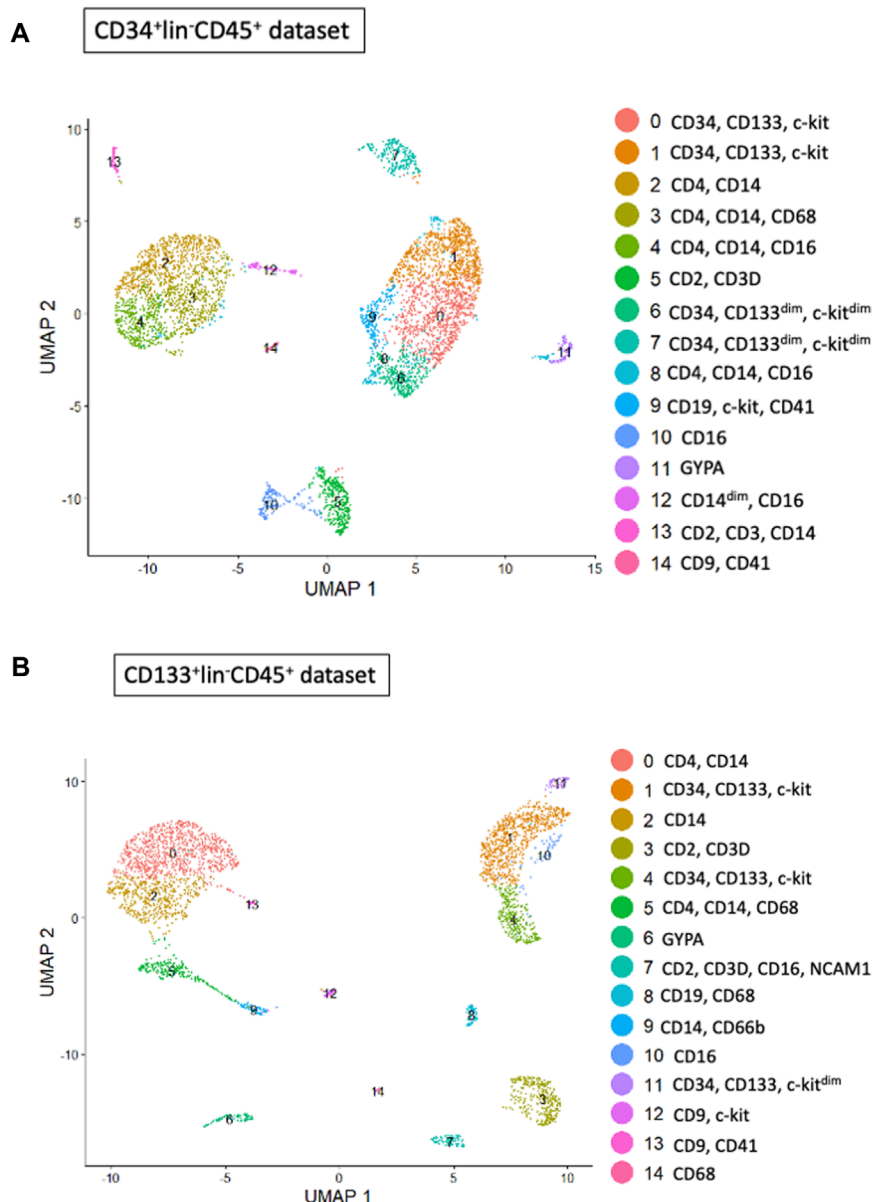


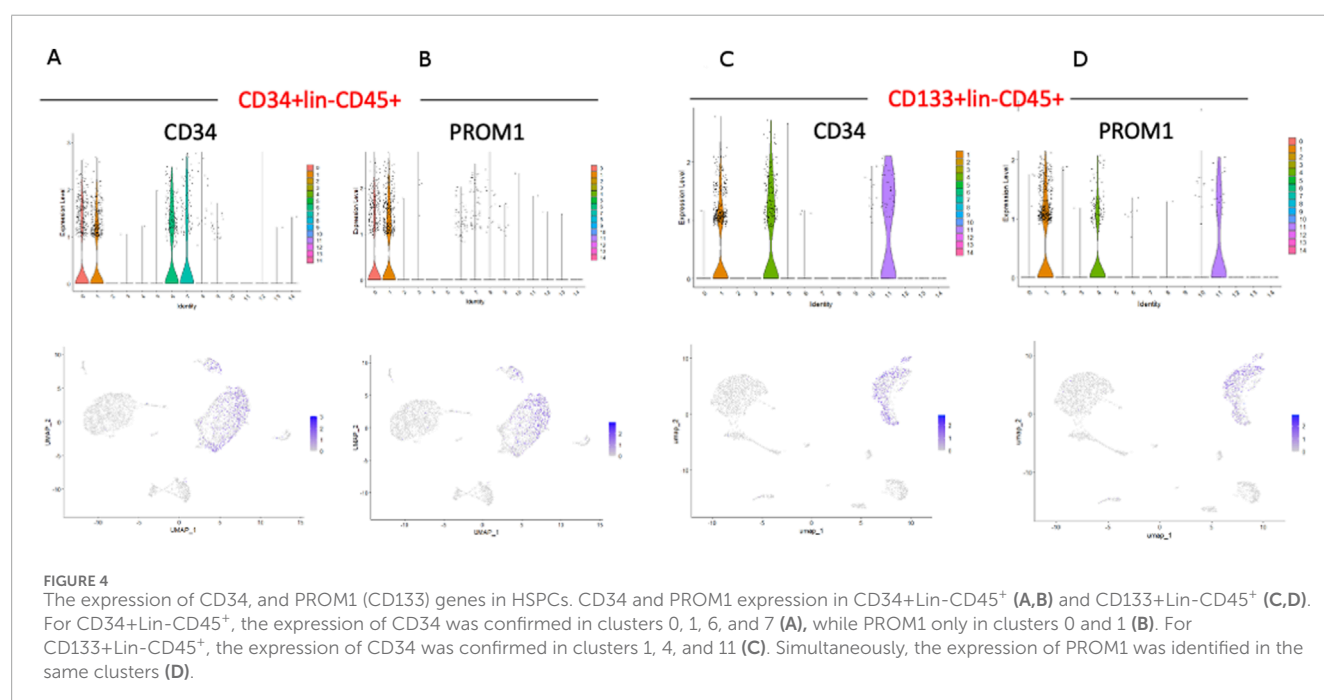
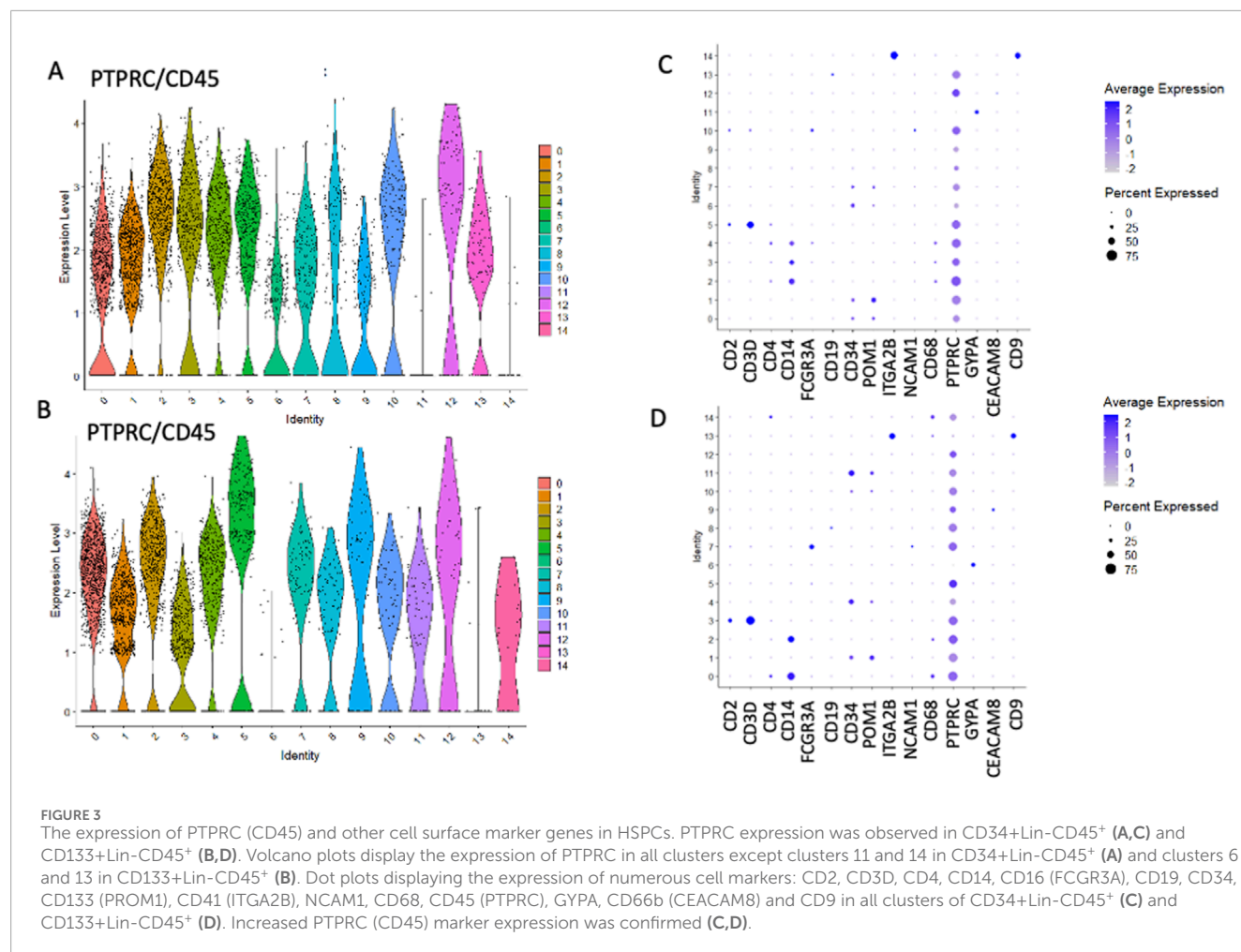
FIGURE 2

Cell subpopulations (clusters) found in both HSPC datasets. Clusters within CD34⁺Lin⁻CD45⁺ (A) and CD133⁺Lin⁻CD45⁺ (B) visualized by UMAP implemented in Seurat (version 5.0.1). The datasets are first analyzed without integration. Human umbilical cord blood isolated HSPCs plot as a total of 14 clusters for CD34⁺ (A) and 14 for CD133⁺ (B) HSPCs libraries.

CD133 (PROM1) and CD34 *versus* clusters expressing CD34 only. This comparison was possible only for the CD34⁺Lin⁻CD45⁺ HSPC population because, in these cells, we can observe these two subpopulations positive for both genes encoding PROM1 and CD34. Therefore, we compared clusters 0 and 1 (expressing mRNA for CD34 and PROM1 and clusters 6 and 7 (expressing mRNA for CD34 only) (Supplementary Table S19).

Among the TOP 50 upregulated genes, the expression of two genes linked with developmental processes was increased almost 3 times in CD34 and PROM1 expressing clusters compared to the clusters with the only expression of CD34. The first one was SOX5, which is responsible for the embryonic development regulation ($\log_2FC = 2.77$; $padj = 2.32E-05$) (Lai et al., 2008), and

the second one, MECOM, responsible for embryonic development, cell differentiation, and regulation of transcription ($\log_2FC = 2.71$; $padj = 9.49E-28$) (Supplementary Table S19) (Goyama et al., 2008). Interestingly, there were several genes downregulated at the level from -5 to -10 log fold change, e.g., PTGDS ($\log_2FC = -10.62$; $padj = 2.02E-05$) regulating inflammation; DNMT ($\log_2FC = -9.42$; $padj = 5.42E-21$), which is crucial in the process of generating a diversity of antigen receptors for effective immune response; IGLC1 ($\log_2FC = -7.78$; $padj = 0.036$) and IGLC2 ($\log_2FC = -7.08$; $padj = 5.33E-07$) triggering complement by antibody-mediated activation (Supplementary Table S19). The list of TOP 50 up and downregulated genes is presented in Supplementary Table S19. As mentioned above, such comparison was not possible for



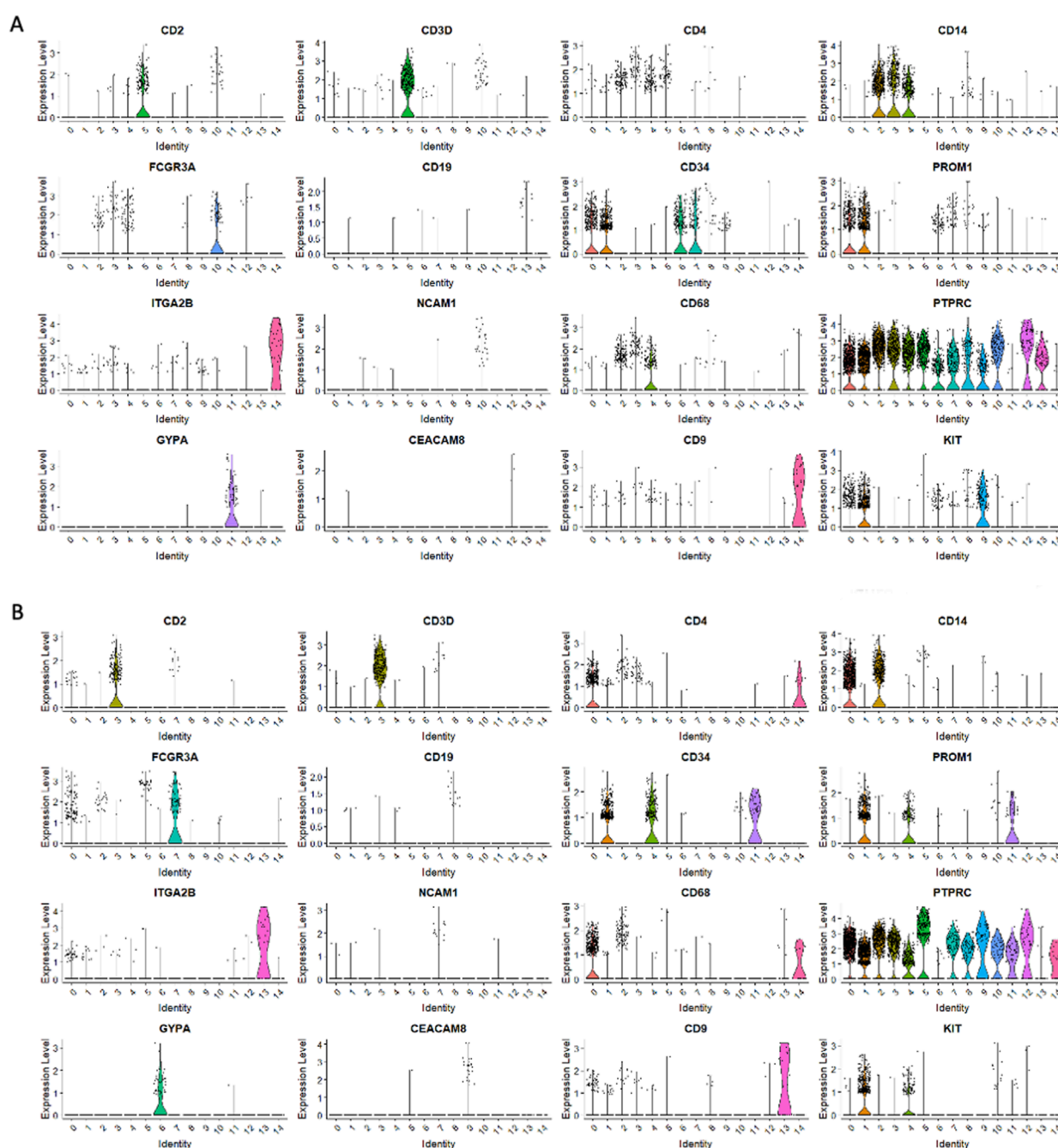


FIGURE 5

Expression of cell markers in CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs. Violin plots displaying the expression of numerous cell markers: CD2, CD3D, CD4, CD14, CD16 (FCGR3A), CD19, CD34, CD133 (PROM1), CD41 (ITGA2B), NCAM1, CD68, CD45 (PTPRC), GYPA, CD66b (CEACAM8), CD9 and KIT (CD117) in all clusters of CD34⁺Lin⁻CD45⁺ (A) and CD133⁺Lin⁻CD45⁺ (B).

CD133+Lin-CD45+HSPC because, within this population, we can observe clusters expressing both markers: CD34 and PROM.

Transcriptomic profiles of quiescent *versus* “fate decision” clusters

Next, we compared datasets of primitive clusters of CD34+Lin-CD45⁺ and CD133+Lin-CD45⁺ HSPCs, that express mRNA for

both CD34 and CD133 markers, with the selected clusters showing already lineage specification. As first, we compared CD34+Lin-CD45⁺ HSPCs datasets, clusters no. 0, 1, 6, and 7 (expressing CD34, PROM1 and KIT (CD117)) with clusters no. 2, 3, 4 (expressing CD4, CD14, CD68) (Figure 5A; Supplementary Table S20). We noticed increase in expression of several genes in clusters 0, 1, 6 and 7 of connected with developmental processes (DPPA4, log2FC = 8.03;

padj = 5.42E-56), cell differentiation and proliferation (MYCT1, log2FC = 8.99; padj = 7.52E-36; MPDZ, log2FC = 8.85; padj = 5.20E-30; MAGI1, log2FC = 8.27; padj = 5.18E-16 as well as KIF7, log2FC = 8.16; padj = 2.71E-19) (Supplementary Table S20). DPPA4 gene is present in embryonic stem cells and is likely to play a role in maintaining their pluripotency (Abu-Dawud et al., 2022; Li et al., 2021; Somanath et al., 2018). This somehow surprising increase in expression of this gene in a primitive subpopulation of CD34+Lin-CD45⁺ HSPCs requires further studies. Among TOP 50 upregulated genes, we also found: CFH (log2FC = 9.02; padj = 8.89E-34), playing an important role in regulation of the complement cascade as well as MAPK12 (log2FC = 7.96; padj = 7.54E-16), involved in signaling pathways related to stress, inflammation, and apoptosis (Supplementary Table S20). This supports a potential interaction of CD34+Lin-CD45⁺ HSPC with the immune system. Of note the expression of CD34 gene was highly upregulated in this dataset (log2FC = 8.44; padj = 7.54E-16) (Supplementary Table S20). In contrast, among TOP 50 downregulated genes, we found those connected with monocytes, macrophages and lymphocyte specification including CX3CR1 (log2FC = -10.75; padj = 2.16E-81), FCGR3A (log2FC = -10.60; padj = 2.79E-35), CD300E (NKG2D) (log2FC = -10.04; padj = 6.74E-130), TLR8 (log2FC = -9.75; padj = 5.79E-38), and CLEC4D (log2FC = -9.50; padj = 1.21E-25) (Supplementary Table S20).

The same comparison has been performed for CD133+Lin-CD45⁺ HSPCs clusters 1, 4 and 11 (expressing CD34, PROM1 and KIT *versus* clusters no. 0, 2 and 14 (expressing CD4, CD14, CD68) (Figure 5B; Supplementary Table S21). Similarly, to the population of CD34+Lin-CD45⁺ HSPCs, we found genes connected with developmental processes including MEST (log2FC = 9.08; padj = 1.07E-56) as well as those connected to cell differentiation: LRP6 (log2FC = 8.80; padj = 2.99E-37) and FZD6 (log2FC = 8.50; padj = 1.85E-34) (Supplementary Table S21). These are both frizzled receptors and key mediators of Wnt signaling pathways. Moreover, in primitive subpopulations of CD133+Lin-CD45⁺ HSPCs we identified genes with increased expression, involved in chromatin modification and regulation of gene expression such as HIST1H2BF (Histone H2B type 1-C/E/F/G/I) (log2FC = 8.85; padj = 6.00E-42); cell signaling - FGD5 (log2FC = 8.73; padj = 1.23E-41) and RNF150 (log2FC = 8.69; padj = 1.32E-32), as well as those involved in intracellular transport: TCTEX1D1 (log2FC = 8.51; padj = 1.57E-33) (Supplementary Table S21).

Transcriptomic comparison of both cell populations

If we compare the TOP 50 upregulated genes from CD34+Lin-CD45⁺ HSPC with CD133+Lin-CD45⁺ HSPCs datasets, we will find four common genes (CD34, SCN3A, COL6A2, and SOCS2-AS1) with increased expression in primitive subpopulations *versus* clusters showing a higher level of differentiation (Supplementary Tables S20, S21). Among these genes, SOCS2-AS1 as long non-coding RNA (lncRNA) and antisense transcript of the SOCS2 gene regulates its expression through various mechanisms such as RNA interference or chromatin modification. As reported, the SOCS2 gene acts as a negative regulator of signaling pathways in controlling cell proliferation, differentiation,

and functioning of immune cells (Cramer et al., 2022; Rico-Bautista et al., 2006). Analysis of TOP50 downregulated genes revealed mostly genes connected with monocytes, macrophages and lymphocytes functions including CLEC4E (log2FC = -10.44; padj = 9.46E-68), CCL3 (log2FC = -9.37; padj = 1.42E-18) and CD86 (log2FC = -9.18; padj = 3.57E-90), TLR8 (log2FC = -8.94; padj = 1.01E-28) and CLEC4D (log2FC = -8.64; padj = 1.89E-17) (Supplementary Table S21). Moreover, mRNA for genes CSTA, VCAN, LGALS2, PID1, SDC2, SERPINB2, TLR8, CLEC4D, LAMB3, and CYP1B (Supplementary Tables S20, S21) was downregulated in primitive subpopulations of CD34+Lin-CD45⁺ and CD133+Lin-CD45⁺ HSPCs. Interestingly, among these genes VCAN, LGALS2 and SDC2 are involved in cell-cell communication.

Unsupervised clustering to analyze transcriptional states in CD34+Lin-CD45⁺ and CD133+Lin-CD45⁺ HSPCs

Analysis of differentially expressed genes in CD34+Lin-CD45⁺ dataset revealed upregulation of several genes in CD34+PROM1+c-KIT⁺ cells (clusters 0 and 1) including C1QTNF4 (log2FC = 1.35; padj = 3.48E-159), AVP, SNHG29, AREG, SPINK2, CALN1, NKAIN2 and FAM30A (Supplementary Table S22). These genes are involved in cell metabolism (C1QTNF4, SPINK2), immunity (C1QTNF4, SPINK2), and several cellular processes (SNHG29, AREG, FAM30A). Disruption of their expression is associated with leukemia and the development of lymphoma. Top downregulated genes detected for clusters 0 and 1, including LYZ, VCAN, TYROBP, NAMPT, and SAT1, are related to immune responses and cellular processes.

Next, by employing Reactome and the gene ontology, we identified several pathways highly up (Figure 6A) and downregulated (Figure 6B) in a cluster of CD34+Lin-CD45⁺ that express PROM1 and c-KIT. These top 50 upregulated genes belong to circadian clock regulation, cellular responses to stress, transcription, translation, and metabolism of RNA and proteins. We also confirmed our previously published data that SARS-CoV2 might infect HSPCs because among the top 50 highly expressed genes in scRNA-Seq of CD34+PROM1+c-Kit⁺ dataset revealed those related to SARS-CoV2 infection (Kucia et al., 2021; Ropa et al., 2020). In parallel, in the same dataset, we observed downregulation of genes annotated to inflammation, cytokine signaling, Toll-like receptor signaling, caspase 3 dependent cell death - pyroptosis, platelet activation, and hemostasis (Figure 6B).

Finally, we compared the expression of top upregulated (Figure 6C) and downregulated (Figure 6D) genes between a subpopulation of CD34+PROM1 *versus* CD34+PROM1low/negative among CD34+Lin-CD45⁺ HSPCs. Reactome and GO analysis revealed upregulation of genes involved in transcription and signaling by Rho GTPase, while among the top downregulated genes were the ones associated with complement cascade activation and NOTCH and interferon signaling.

Analysis of differentially expressed genes - up and downregulated in subpopulations of CD133+Lin-CD45⁺ HSPCs

TABLE 1 The summary of progenitor fractions and their lineage commitment in the single-cell clusters of CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs.

Lineage commitment	Cluster name	CD34 ⁺ Lin ⁻ CD45 ⁺ clusters	CD133 ⁺ Lin ⁻ CD45 ⁺ clusters	Key markers
Primitive HSPCs	Quiescent/primitive	0, 1, 6, 7	1, 4, 11	CD34, PROM1
Early progenitors		0, 1, 6, 9	1, 4	c-KIT
T-cell lineage		2, 3, 4, 5	0, 2, 3	CD2, CD3, CD4
B cell lineage		13	8	CD19
Myeloid lineage	Fate decision	2, 3, 4, 10	0, 2, 7, 14	CD14, FCGR3A, CD68
Erythroid lineage		11	6	GYPA
Megakaryocytic lineage		14	13	ITGA2B

dataset (Figures 6E,F) revealed high expression of C1QA, CLEC10A, GPRC5A, SASH1, KRBOX1, RNF150, PRSS51, HSPA12A, PDE10A, MMP16, AC015908.2, and ARHGAP20 involved in ubiquitin-dependent protein catabolic processes, proteolysis, ATP binding activity, signal transduction due to regulating the availability of intracellular cyclic nucleotides, and NLRP-dependent pyroptosis. Reactome and GO analysis of the top 50 upregulated genes for PROM1⁺CD34⁺ clusters in CD133⁺Lin⁻CD45⁺ HSPCs dataset identified mRNA encoding RAC1 and CDC42 GTPase, genes involved in metabolism, hormone receptors, and lipoproteins remodeling (Figure 6E). The top 50 downregulated genes in a primitive subpopulation of PROM1⁺CD34⁺ cells in the CD133⁺Lin⁻CD45⁺ HSPCs dataset were associated with differentiation, TLRs signaling, antigen processing and presentation, neutrophil degranulation, and cytokine-related immune responses (Figure 6F).

Integrated analysis of both datasets

Finally, we performed an integrated analysis of CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs datasets merged and treated as “pseudobulk”. We compared their pseudobulk profiles and observed that CD34⁺ and CD133⁺ HSPCs do not differ significantly in gene expression, as evidenced by the strong positive linear relationship between the cell populations ($R = 0.99$). We observed several genes expressed at similar levels in both cell populations (Supplementary Figure S3). Therefore, next, we visualized clusters using uniform manifold approximation and d projection (UMAP) implemented in Seurat (version 5.0.1) (Satija et al., 2015; Hao et al., 2024). Initially, the analysis was done with the datasets without integration. Visual inspection of the result confirmed that clusters are defined by the cell population type. The UMAP plot illustrates clustering of cells, with CD34⁺Lin⁻CD45⁺ (red) and CD133⁺Lin⁻CD45⁺ (blue) forming distinct but partially overlapping populations, indicating transcriptional similarities (Supplementary Figure S3). However, the gene expression profiles

for both populations overlap, and we do not observe unique subpopulations within CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs (Supplementary Figures S3A,C). This has been confirmed after integration because cells from both cell populations cluster together (Supplementary Figures S3A,B). Additionally, when we compared CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺ HSPCs side-by-side, we did not see differences in the distribution of subpopulations (Supplementary Figure S3C).

Finally, we analyzed in integrated datasets the expression of cell marker genes to highlight the differences in the expression of primitive hematopoietic stem cell markers: CD34, PROM1, PECAM1 (CD31), c-KIT (CD117); lymphocyte markers: CD2, CD4; monocytes and macrophages markers: CD14; erythrocyte marker: GYPA and megakaryocytes marker: ITGA2B. We also checked the expression of PTPRC, a marker common for leukocytes, present at the surfaces of all immune system cells except for erythrocytes and platelets. We observed that there are two main cluster clouds in both cell populations. While CD34, PROM1, and c-KIT expression characterize the first, the second one expresses CD4 and CD14. It is consistent with the results from non-integrated analysis. The aggregation analysis of both datasets shows high similarity by identifying 23 cell clusters in each one that align together (Supplementary Figures S3A–C). The very high correlation ($R = 0.99$) between the transcriptomes of CD34⁺ and CD133⁺ HSPCs in our pseudobulk analysis highlights a strong transcriptional overlap between these two commonly used markers. This finding supports the notion that CD34 and CD133 mark substantially overlapping populations of primitive hematopoietic progenitors. While both subsets exhibit similar expression of genes associated with stemness and early differentiation, our clustering analysis reveals functionally relevant substructures with differential gene expression, particularly in genes related to signaling, lineage priming, and immune regulation. Importantly, the observed correlation validates the use of either CD34 or CD133 as viable purification targets for the enrichment of multipotent HSPCs in research and clinical settings. At the same time, the subtle transcriptomic differences underscore the potential benefit of dual-marker strategies

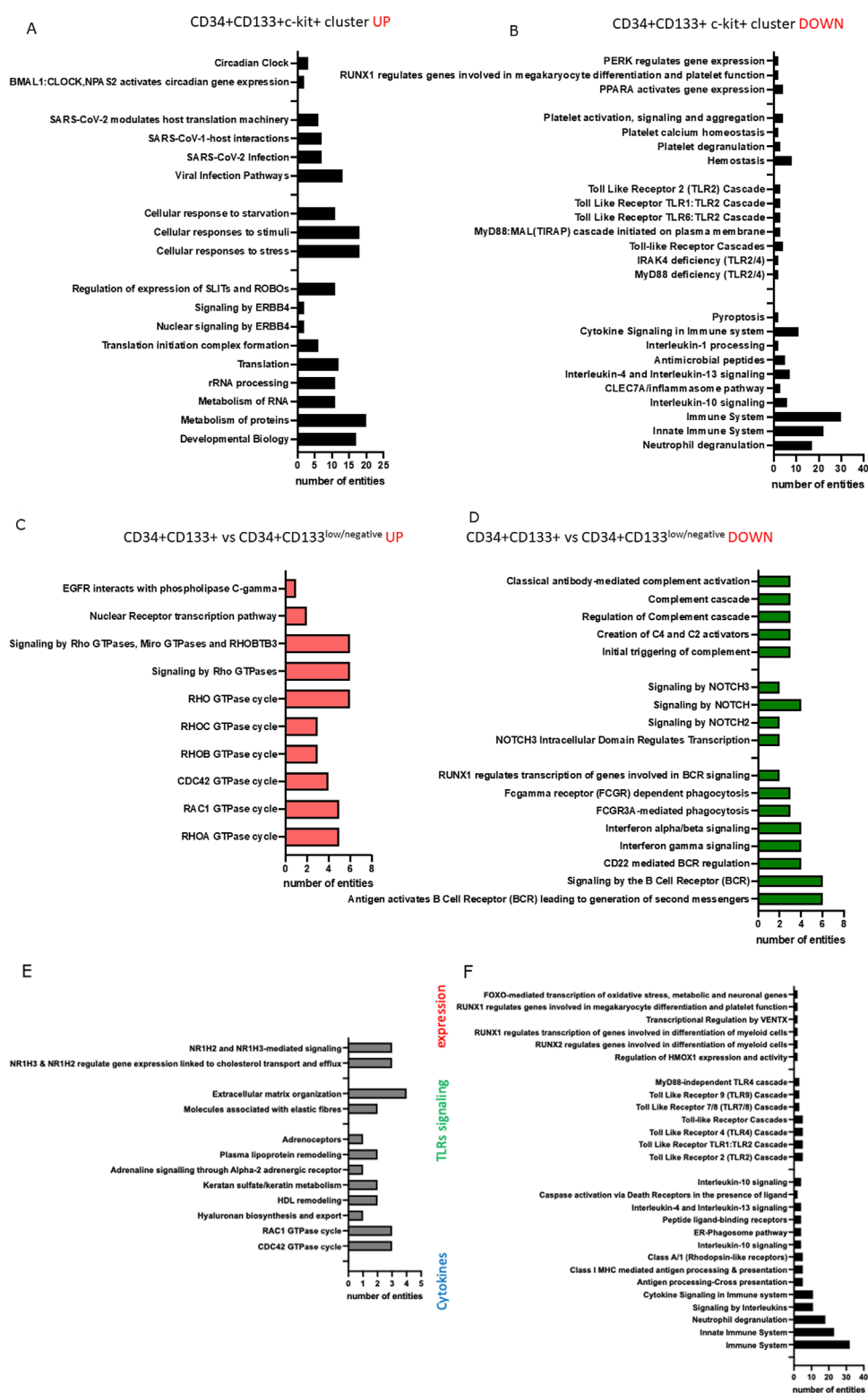


FIGURE 6

Reactome pathways identified based on the top 50 upregulated (A) and top 50 downregulated (B) genes were found as differentially expressed between clusters 0 and 1 (expressing CD34, PROM1, and KIT), and clusters 6 and 7 (expressing only CD34) of CD34⁺Lin⁻CD45⁺. Reactome pathways were identified based on the top 50 upregulated (C) and top 50 downregulated (D) genes found in clusters 1 and 4 (expressing CD34, PROM1, and KIT) of CD133+Lin⁻CD45⁺. Reactome pathways were identified based on the top 50 upregulated (E) and top 50 downregulated (F) genes in a primitive subpopulation of PROM1+CD34⁺ cells of the CD133+Lin⁻CD45⁺ HSPCs dataset.

(CD34⁺CD133⁺) for capturing the most primitive, quiescent subpopulations.

There are several different strategies to purify and identify HSPCs (Bhatia et al., 1997; Kato and Radbruch, 1993; Bujko et al., 2019). Hematopoietic stem cells (HSCs) and multipotent hematopoietic progenitors are routinely isolated using various markers. The most common is FACS or immunomagnetic sorting using HSPCs expressed markers. Additionally, HSPCs could be purified based on minimal accumulation of metabolic fluorochromes such as pyronin Y (mRNA marker), Rh123 (efflux pump substrate), and Hoe33342 (DNA marker). It has been proposed that applying SLAM family markers, CD150, CD48, CD229, and CD244, can distinguish HSCs and more restricted progenitors. In this paper, we employed FACS to sort two populations of HSPCs based on the expression of CD34, CD133, and CD45 antigens and the lack of expression of lineage differentiation markers. The first one, the CD34 antigen, is a member of a family of single-pass transmembrane sialomucin proteins expressed on human UCB and bone marrow lymphohematopoietic stem cells, progenitor cells, endothelial cells, mast cells, and a subpopulation of dendritic cells (Kato and Radbruch, 1993). The other one, CD133 antigen (PROM1 - prominin-1), is a member of pentaspan transmembrane glycoproteins, which specifically localize to cellular protrusions expressed on HSPCs, endothelial progenitors, embryonic stem cells, some types of neural cells and is associated as a marker with several tumors (Yin et al., 1997; Weigmann et al., 1997; Katoh and Katoh, 2007; Rappa et al., 2008). As reported, human HSPCs co-express CD34 and CD133 antigens, and 30%–60% of CD34⁺ cells co-express CD133 antigens.

We performed single-cell RNA-sequencing (RNA-seq) analysis of the transcriptome between UCB-derived CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ cells enriched for HSPCs and report similarities and differences in transcriptome between these cells at single cell level and, described mRNA clusters detected by RNA-seq. We detected some differences in transcriptome between subpopulations of these cells identified by scRNA-seq visualized as clusters using uniform manifold approximation and d projection (UMAP). Nevertheless, integrated analysis of CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ datasets, which were merged and treated as “pseudobulk” revealed that both populations do not differ significantly in gene expression, as evidenced by the very strong positive linear relationship between these cells. Our results show that the expression profiles for both populations overlap, and we do not observe unique subpopulations within CD34⁺Lin[−]CD45⁺ and CD133⁺Lin[−]CD45⁺ HSPCs.

To achieve reliable results that enable biologically meaningful conclusions, every stage of the scRNA-seq experiment must be meticulously executed, including cell sorting, single-cell library preparation, quality control, and data analysis. The decision to work with sorted material, rather than the conventional approach of using a full pellet of blood cells, proved to be valid and

demonstrated that HSPC analysis is feasible even with a limited cell number.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/sra/PRJNA1128409>.

Ethics statement

The studies involving humans were approved by Warsaw Bioethics Committee (permission number KB/50/2022). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JK: Formal Analysis, Investigation, Methodology, Software, Writing – original draft. PK: Data curation, Methodology, Writing – original draft. MZR: Supervision, Writing – review and editing. MK: Conceptualization, Project administration, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Science Centre, Poland OPUS grant UMO-2021/41/B/NZ3/01589 to MZR and National Science Centre, Poland OPUS grant UMO-2022/45/B/NZ3/00476 to MK.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2025.1590889/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

Cell sorting strategy for the isolation of HSPCs from hUCB. Immunostained cells were first visualized by dot plot showing forward scatter (FSC) vs. side scatter (SSC) signals, where small events ranging from 2 to 15 μm were gated (P1) (A) and

further analyzed for the expression of Lineage markers. Lineage negative events were gated (Lin⁻) (B) and analyzed for the expression of CD45 and CD34 (C) or CD133 (D) antigens. The populations of CD34⁺Lin⁻CD45⁺ HSPCs (C) and CD133⁺Lin⁻CD45⁺ HSPCs (D) were separately sorted on the MoFlo Astrios EQ cell sorter. MNCs isolation and staining were described in the Materials and Methods section. Representative dot plots saved during the sample acquisition are shown.

SUPPLEMENTARY FIGURE S2

Quality control parameters of final libraries. Left panel: Electropherograms of final libraries: CD133+lin-CD45⁺ (A) and CD34+lin-CD45⁺ (B) prepared with the use of Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1, and Single Index Kit T Set A (10X Genomics, United States), assessed with the D5000 ScreenTape assay. DNA shows a peak size of around 400 bp, which matches the expected size range between 400 and 600 bp. Right panel: FastQC reports showing per base sequences quality in all samples: CD133+lin-CD45⁺ (C) and CD34+lin-CD45⁺ (D).

SUPPLEMENTARY FIGURE S3

Integration analysis of HSPC libraries. Integrating two scRNA Seq datasets (CD34⁺Lin⁻CD45⁺ and CD133⁺Lin⁻CD45⁺) visualized the lack of differences between two subtypes of HSPCs (A) and identified 23 common cell clusters within both datasets (B,C). Slight differences result from the different number of cells obtained for both libraries (C).

References

- Abdelbaset-Ismail, A., Suszynska, M., Borkowska, S., Adamiak, M., Ratajczak, J., Kucia, M., et al. (2016). Human haematopoietic stem/progenitor cells express several functional sex hormone receptors. *J. Cell. Mol. Med.* 20 (1), 134–146. doi:10.1111/jcmm.12712
- Abu-Dawud, R., Mistry, B. V., Almahdali, R., Matin, M. M., Alaiya, A., Alghamdi, F., et al. (2022). Dynamic cytosolic foci of DPPA4 in human pluripotent stem cells. *Tissue Cell* 78, 101893. doi:10.1016/j.tice.2022.101893
- Andrews, R. G., Singer, J. W., and Bernstein, I. D. (1989). Precursors of colony-forming cells in humans can be distinguished from colony-forming cells by expression of the CD33 and CD34 antigens and light scatter properties. *J. Exp. Med.* 169 (5), 1721–1731. doi:10.1084/jem.169.5.1721
- Bao, E. L., Cheng, A. N., and Sankaran, V. G. (2019). The genetics of human hematopoiesis and its disruption in disease. *EMBO Mol. Med.* 11 (8), e10316. doi:10.15252/emmm.201910316
- Belperio, J. A., Keane, M. P., Arenberg, D. A., Addison, C. L., Ehlert, J. E., Burdick, M. D., et al. (2000). CXCL chemokines in angiogenesis. *J. Leukoc. Biol.* 68 (1), 1–8. doi:10.1189/jlb.68.1.1
- Berardi, A. C., Wang, A., Levine, J. D., Lopez, P., and Scadden, D. T. (1995). Functional isolation and characterization of human hematopoietic stem cells. *Science* 267 (5194), 104–108. doi:10.1126/science.7528940
- Bhatia, M., Wang, J. C., Kapp, U., Bonnet, D., and Dick, J. E. (1997). Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* 94 (10), 5320–5325. doi:10.1073/pnas.94.10.5320
- Brennecke, P., Anders, S., Kim, J. K., Kolodziejczyk, A. A., Zhang, X., Proserpio, V., et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* 10 (11), 1093–1095. doi:10.1038/nmeth.2645
- Broxmeyer, H. E., and Kim, C. H. (1999). Regulation of hematopoiesis in a sea of chemokine family members with a plethora of redundant activities. *Exp. Hematol.* 27 (7), 1113–1123. doi:10.1016/s0301-472x(99)00045-4
- Bujko, K., Ciechanowicz, A. K., Kucia, M., and Ratajczak, M. Z. (2023). Molecular analysis and comparison of CD34(+) and CD133(+) very small embryonic-like stem cells purified from umbilical cord blood. *Cytom. A* 103 (9), 703–711. doi:10.1002/cyto.a.24767
- Bujko, K., Kucia, M., Ratajczak, J., and Ratajczak, M. Z. (2019). Hematopoietic stem and progenitor cells (HSPCs). *Adv. Exp. Med. Biol.* 1201, 49–77. doi:10.1007/978-3-030-31206-0_3
- Calvanese, V., and Mikkola, H. K. A. (2023). The genesis of human hematopoietic stem cells. *Blood* 142 (6), 519–532. doi:10.1182/blood.2022017934
- Challen, G. A., Boles, N., Lin, K. K., and Goodell, M. A. (2009). Mouse hematopoietic stem cell identification and analysis. *Cytom. A* 75 (1), 14–24. doi:10.1002/cyto.a.20674
- Chambers, S. M., Boles, N. C., Lin, K. Y., Tierney, M. P., Bowman, T. V., Bradfute, S. B., et al. (2007). Hematopoietic fingerprints: an expression database of stem cells and their progeny. *Cell. Stem Cell* 1 (5), 578–591. doi:10.1016/j.stem.2007.10.003
- Cheng, P., Eksioğlu, E. A., Chen, X., Kandell, W., Le Trinh, T., Cen, L., et al. (2019). S100A9-induced overexpression of PD-1/PD-L1 contributes to ineffective hematopoiesis in myelodysplastic syndromes. *Leukemia* 33 (8), 2034–2046. doi:10.1038/s41375-019-0397-9
- Cramer, A., Galvao, I., Venturini de Sa, N., Gaio, P., Fernanda de Melo Oliveira, N., Rates Gonzaga Santos, M., et al. (2022). Role of Suppressor of cytokine signaling 2 during the development and resolution of an experimental arthritis. *Cell. Immunol.* 372, 104476. doi:10.1016/j.cellimm.2021.104476
- Crosse, E. I., Gordon-Keylock, S., Rybtsov, S., Binagui-Casas, A., Felchle, H., Nnadi, N. C., et al. (2020). Multi-layered spatial transcriptomics identify secretory factors Promoting human hematopoietic stem cell development. *Cell. Stem Cell* 27 (5), 822–839. doi:10.1016/j.stem.2020.08.004
- Farlik, M., Halbritter, F., Muller, F., Choudry, F. A., Ebert, P., Klughammer, J., et al. (2016). DNA Methylation Dynamics of human hematopoietic stem cell differentiation. *Cell. Stem Cell* 19 (6), 808–822. doi:10.1016/j.stem.2016.10.019
- Fast, E. M., Sporrij, A., Manning, M., Rocha, E. L., Yang, S., Zhou, Y., et al. (2021). External signals regulate continuous transcriptional states in hematopoietic stem cells. *Elife* 10, e66512. doi:10.7554/eLife.66512
- Flohr, S. A., Yang, D., Kim, K., Lazare, S., Skinder, N., Zwart, E., et al. (2021). A comprehensive transcriptome signature of murine hematopoietic stem cell aging. *Blood* 138 (6), 439–451. doi:10.1182/blood.202009729
- Goyama, S., Yamamoto, G., Shimabe, M., Sato, T., Ichikawa, M., Ogawa, S., et al. (2008). Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell. Stem Cell* 3 (2), 207–220. doi:10.1016/j.stem.2008.06.002
- Graf, T., and Enver, T. (2009). Forcing cells to change lineages. *Nature* 462 (7273), 587–594. doi:10.1038/nature08533
- Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., et al. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 42 (2), 293–304. doi:10.1038/s41587-023-01767-y
- Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., et al. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. *Genes. Dev.* 11 (6), 774–785. doi:10.1101/gad.11.6.774
- Huang, P., Zhao, Y., Zhong, J., Zhang, X., Liu, Q., Qiu, X., et al. (2020). Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells. *Proc. Natl. Acad. Sci. U. S. A.* 117 (23), 12868–12876. doi:10.1073/pnas.1915085117
- Kato, K., and Radbruch, A. (1993). Isolation and characterization of CD34+ hematopoietic stem cells from human peripheral blood by high-gradient magnetic cell sorting. *Cytometry* 14 (4), 384–392. doi:10.1002/cyto.990140407
- Katoh, Y., and Katoh, M. (2007). Comparative genomics on PROM1 gene encoding stem cell marker CD133. *Int. J. Mol. Med.* 19 (6), 967–970. doi:10.3892/ijmm.19.6.967
- Krause, D. S. (2002). Regulation of hematopoietic stem cell fate. *Oncogene* 21 (21), 3262–3269. doi:10.1038/sj.onc.1205316

- Kucia, M., Ratajczak, J., Bujko, K., Adamiak, M., Ciechanowicz, A., Chumak, V., et al. (2021). An evidence that SARS-Cov-2/COVID-19 spike protein (SP) damages hematopoietic stem/progenitor cells in the mechanism of pyroptosis in Nlrp3 inflammasome-dependent manner. *Leukemia* 35 (10), 3026–3029. doi:10.1038/s41375-021-01332-z
- Lai, T., Jabaudon, D., Molyneaux, B. J., Azim, E., Arlotta, P., Menezes, J. R., et al. (2008). SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron* 57 (2), 232–247. doi:10.1016/j.neuron.2007.12.023
- Laurenti, E., and Gottgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature* 553 (7689), 418–426. doi:10.1038/nature25022
- Lawrence, H. J., Christensen, J., Fong, S., Hu, Y. L., Weissman, I., Sauvageau, G., et al. (2005). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood* 106 (12), 3988–3994. doi:10.1182/blood-2005-05-2003
- Li, H., Long, C., Xiang, J., Liang, P., Li, X., and Zuo, Y. (2021). Dppa2/4 as a trigger of signaling pathways to promote zygote genome activation by binding to CG-rich region. *Brief. Bioinform* 22 (4), bbaa342. doi:10.1093/bib/bbaa342
- McInnes, L., Healy, J., Saul, N., and Grobberger, L. (2018). UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* 3, 861. doi:10.21105/joss.00861
- McKinney-Freeman, S., Cahan, P., Li, H., Lacadie, S. A., Huang, H. T., Curran, M., et al. (2012). The transcriptional landscape of hematopoietic stem cell ontogeny. *Cell Stem Cell* 11 (5), 701–714. doi:10.1016/j.stem.2012.07.018
- Metcalfe, D. (2007). Concise review: hematopoietic stem cells and tissue stem cells: current concepts and unanswered questions. *Stem Cells* 25 (10), 2390–2395. doi:10.1634/stemcells.2007-0544
- Miao, W. M., Vasile, E., Lane, W. S., and Lawler, J. (2001). CD36 associates with CD9 and integrins on human blood platelets. *Blood* 97 (6), 1689–1696. doi:10.1182/blood.v97.6.1689
- Mikkola, H. K., and Orkin, S. H. (2006). The journey of developing hematopoietic stem cells. *Development* 133 (19), 3733–3744. doi:10.1242/dev.02568
- Milner, L. A., and Bigas, A. (1999). Notch as a mediator of cell fate determination in hematopoiesis: evidence and speculation. *Blood* 93 (8), 2431–2448. doi:10.1182/blood.v93.8.2431.408k33_2431_2448
- Miyamoto, Y., Hieda, M., Harreman, M. T., Fukumoto, M., Saiwaki, T., Hodel, A. E., et al. (2002). Importin alpha can migrate into the nucleus in an importin beta- and Ran-independent manner. *EMBO J.* 21 (21), 5833–5842. doi:10.1093/emboj/cdf569
- Morrison, S. J., Uchida, N., and Weissman, I. L. (1995). The biology of hematopoietic stem cells. *Annu. Rev. Cell. Dev. Biol.* 11, 35–71. doi:10.1146/annurev.cb.11.110195.000343
- Ng, Y. Y., Baert, M. R., de Haas, E. F., Pike-Overzet, K., and Staal, F. J. (2009). Isolation of human and mouse hematopoietic stem cells. *Methods Mol. Biol.* 506, 13–21. doi:10.1007/978-1-59745-409-4_2
- Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J. E. (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* 333 (6039), 218–221. doi:10.1126/science.1201219
- Oguma, Y., Kuroda, Y., Wakao, S., Kushida, Y., and Dezawa, M. (2022). Single-cell RNA sequencing reveals different signatures of mesenchymal stromal cell pluripotent-like and multipotent populations. *iScience* 25 (11), 105395. doi:10.1016/j.isci.2022.105395
- Orlic-Milacic, M., Rothfels, K., Matthews, L., Wright, A., Jassal, B., Shamovsky, V., et al. (2024). Pathway-based, reaction-specific annotation of disease variants for elucidation of molecular phenotypes. *Database (Oxford)* 2024, baa031. doi:10.1093/database/baa031
- Popescu, D. M., Botting, R. A., Stephenson, E., Green, K., Webb, S., Jardine, L., et al. (2019). Decoding human fetal liver haematopoiesis. *Nature* 574 (7778), 365–371. doi:10.1038/s41586-019-1652-y
- Ranzoni, A. M., Tangherloni, A., Berest, I., Riva, S. G., Myers, B., Strzelecka, P. M., et al. (2021). Integrative single-cell RNA-seq and ATAC-seq analysis of human developmental hematopoiesis. *Cell Stem Cell* 28 (3), 472–487.e7. doi:10.1016/j.stem.2020.11.015
- R Core Team (2024). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>.
- Rappa, G., Fodstad, O., and Lorico, A. (2008). The stem cell-associated antigen CD133 Prominin-1 is a molecular therapeutic target for metastatic melanoma. *Stem Cells* 26 (12), 3008–3017. doi:10.1634/stemcells.2008-0601
- Rico-Bautista, E., Flores-Morales, A., and Fernandez-Perez, L. (2006). Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. *Cytokine Growth Factor Rev.* 17 (6), 431–439. doi:10.1016/j.cytogfr.2006.09.008
- Riether, C., Schurch, C. M., and Ochsenbein, A. F. (2015). Regulation of hematopoietic and leukemic stem cells by the immune system. *Cell Death Differ.* 22 (2), 187–198. doi:10.1038/cdd.2014.89
- Ropa, J., Cooper, S., Capitano, M. L., Van't Hof, W., and Broxmeyer, H. E. (2020). Human hematopoietic stem, progenitor, and immune cells Respond *ex vivo* to SARS-CoV-2 spike protein. *Stem Cell. Rev. Rep.* 17, 253–265. doi:10.1007/s12015-020-10056-z
- Rossi, L., Challen, G. A., Sirin, O., Lin, K. K., and Goodell, M. A. (2011). Hematopoietic stem cell characterization and isolation. *Methods Mol. Biol.* 750, 47–59. doi:10.1007/978-1-61779-145-1_3
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33 (5), 495–502. doi:10.1038/nbt.3192
- Smith, N. L., Bromley, M. J., Denning, D. W., Simpson, A., and Bowyer, P. (2015). Elevated levels of the neutrophil chemoattractant pro-platelet basic protein in macrophages from individuals with chronic and allergic aspergillosis. *J. Infect. Dis.* 211 (4), 651–660. doi:10.1093/infdis/jiu490
- Somanath, P., Bush, K. M., and Knoepfler, P. S. (2018). ERBB3-Binding protein 1 (EBP1) is a novel developmental pluripotency-associated-4 (DPPA4) Cofactor in human pluripotent cells. *Stem Cells* 36 (5), 671–682. doi:10.1002/stem.2776
- Su, M., Pan, T., Chen, Q. Z., Zhou, W. W., Gong, Y., Xu, G., et al. (2022). Data analysis guidelines for single-cell RNA-seq in biomedical studies and clinical applications. *Mil. Med. Res.* 9 (1), 68. doi:10.1186/s40779-022-00434-8
- Takizawa, H., Boettcher, S., and Manz, M. G. (2012). Demand-adapted regulation of early hematopoiesis in infection and inflammation. *Blood* 119 (13), 2991–3002. doi:10.1182/blood-2011-12-380113
- Weigmann, A., Corbeil, D., Hellwig, A., and Huttner, W. B. (1997). Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* 94 (23), 12425–12430. doi:10.1073/pnas.94.23.12425
- Weiss, M. J., Zhou, S., Feng, L., Gell, D. A., Mackay, J. P., Shi, Y., et al. (2005). Role of alpha-hemoglobin-stabilizing protein in normal erythropoiesis and beta-thalassemia. *Ann. N. Y. Acad. Sci.* 1054, 103–117. doi:10.1196/annals.1345.013
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis* (Springer-Verlag New York).
- Wognum, A. W., Eaves, A. C., and Thomas, T. E. (2003). Identification and isolation of hematopoietic stem cells. *Arch. Med. Res.* 34 (6), 461–475. doi:10.1016/j.arcmed.2003.09.008
- Worthington, R. E., Carroll, R. C., and Boucheix, C. (1990). Platelet activation by CD9 monoclonal antibodies is mediated by the Fc gamma II receptor. *Br. J. Haematol.* 74 (2), 216–222. doi:10.1111/j.1365-2141.1990.tb02568.x
- Xia, C., Braunstein, Z., Toomey, A. C., Zhong, J., and Rao, X. (2017). S100 proteins as an important regulator of macrophage inflammation. *Front. Immunol.* 8, 1908. doi:10.3389/fimmu.2017.01908
- Yin, A. H., Miraglia, S., Zanjani, E. D., Almeida-Porada, G., Ogawa, M., Leary, A. G., et al. (1997). AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 90 (12), 5002–5012. doi:10.1182/blood.v90.12.5002.5002_5012
- Zeng, Y., He, J., Bai, Z., Li, Z., Gong, Y., Liu, C., et al. (2019). Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. *Cell Res.* 29 (11), 881–894. doi:10.1038/s41422-019-0228-6
- Zheng, Z., He, H., Tang, X. T., Zhang, H., Gou, F., Yang, H., et al. (2022). Uncovering the emergence of HSCs in the human fetal bone marrow by single-cell RNA-seq analysis. *Cell Stem Cell* 29 (11), 1562–1579.e7. doi:10.1016/j.stem.2022.10.005