

Comparative single-cell lineage bias in human and murine hematopoietic stem cells

Isaac Shamie,^{1*} Meghan Bliss-Moreau,^{2,3*} Jamie Casey Lee,^{4*} Ronald Mathieu,² Harold M. Hoffman,⁴ Bob Geng,⁵ Nathan E. Lewis,^{1,6,7#} Yanfang Peipei Zhu^{4,8#} and Ben A. Croker^{2-4#}

¹Department of Bioengineering, UC San Diego, La Jolla, CA; ²Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA; ³Department of Pediatrics, Harvard Medical School, Boston, MA; ⁴Division of Rheumatology, Allergy & Immunology, Department of Pediatrics, School of Medicine, UC San Diego, La Jolla, CA; ⁵Rady Children's Hospital San Diego, San Diego, CA; ⁶Division of Host-Microbe Systems and Therapeutics, Department of Pediatrics, UC San Diego, La Jolla, CA; ⁷Center for Molecular Medicine, Complex Carbohydrate Research Center, Department of Biochemistry and Molecular Biology, University of Georgia, GA and ⁸Immunology Center of Georgia, Medical College of Georgia, Department of Biochemistry and Molecular Biology, Augusta University, GA, USA

^{*}IS, MB-M and JCL contributed equally as first authors.

[#]NEL, YPZ and BAC contributed equally as senior authors.

Correspondence: B.A. Croker
bcroker@health.ucsd.edu

Y.P. Zhu
pzhu@augusta.edu

N.E. Lewis
natelewis@uga.edu

Received: March 25, 2025.

Accepted: July 11, 2025.

Early view: July 17, 2025.

<https://doi.org/10.3324/haematol.2025.287897>

©2026 Ferrata Storti Foundation

Published under a CC BY-NC license



Abstract

The commitment of hematopoietic stem cells (HSC) to myeloid, erythroid, and lymphoid lineages is influenced by microenvironmental cues, and governed by cell-intrinsic and epigenetic characteristics that are unique to the HSC population. To investigate the nature of lineage commitment bias in human HSC, mitochondrial single-cell assay for transposase-accessible chromatin (ATAC)-sequencing was used to identify somatic mutations in mitochondrial DNA to act as natural genetic barcodes for tracking the *ex vivo* differentiation potential of HSC to mature cells. Clonal lineages of human CD34⁺ cells and their mature progeny were normally distributed across the hematopoietic lineage tree without evidence of significant skewing. To investigate commitment bias *in vivo*, mice were transplanted with limited numbers of long-term HSC (LT-HSC). Variation in the ratio of myeloid and lymphoid cells between donors was suggestive of a skewed output but was not altered by increasing numbers of LT-HSC. These data suggest that the variation in myeloid and lymphoid engraftment is a stochastic process dominated by the irradiated recipient niche with minor contributions from cell-intrinsic lineage biases of LT-HSC.

Introduction

Hematopoietic stem cells (HSC) are classically considered to have the capacity for complete regeneration of the hematopoietic compartment. More recent analyses indicate additional complexity and heterogeneity in the HSC compartment, with lineage-restricted or lineage-biased HSC considered a feature of mammalian hematopoiesis.¹⁻¹³ A partial differential equation model to study relationships between hematopoietic stem and progenitor cells (HSPC) emphasizes that myeloid bias cannot be accounted for solely by short-term HSC bias during inflammation but rather involves a combination of HSC and progenitor cell biases.¹⁴ Central to the concept of lineage bias is an assumption that cells used for studying HSC commitment are HSC and not multipotent progenitors or lineage-committed progenitors. Changes in differentiation of cells

downstream of the long-term HSC (LT-HSC) must also be evaluated when considering the potential lineage bias of a LT-HSC. Functional validation of these heterogeneous characteristics minimally requires demonstration of the capacity to transplant single HSC into lethally irradiated or hemoablated recipients.¹⁵ Retrospective analysis of single HSC injected into recipient mice and their progeny can validate the isolation procedure, and support data obtained from protocols for prospective isolation of these rare bone marrow cell populations. Next-generation sequencing and gene expression analyses have identified differentially expressed genes modulated specifically within the LT-HSC population, and enabled development of antibodies for purification and phenotypic analysis.¹⁶⁻¹⁹ However, successful engraftment of mice with single HSC still remains challenging. Many studies aiming to repopulate mice with HSC report variable reconstitution

using inconsistent definitions of engraftment including both survival and >0.1-1% repopulation in any lineage, suggesting that some HSC isolation protocols result in the purification of lineage-restricted progenitor cells and mature hematopoietic cells.

HSC tagging methodologies also provide a means to study the *in vivo* function of HSC. Clonal relationships between HSPC and mature hematopoietic lineages have been explored using inducible DNA “barcoding” methodologies during embryogenesis or postnatal life, genetically modified mice expressing fluorescent markers within the HSC population or upon the transplantation of virally transduced barcoded single HSC into lethally irradiated or hemoablated adult murine or non-human primate recipients.²⁰⁻²⁶ These methods have provided a wealth of data supporting the role of hematopoietic progenitor cells in driving steady-state hematopoiesis and the apparent lineage bias of LT-HSC, but also strong counterarguments.^{27,28} HSPC differentiation can also be tracked by mapping somatic mutations in the mitochondria, which act as natural barcodes for cells.²⁹ This modified single-cell assay for transposase-accessible chromatin (ATAC)-sequencing (mt-scATAC-seq) method captures both open nuclear chromatin and mitochondrial DNA sequences to identify somatic variants.

In this study, mt-scATAC-seq was used to exploit the somatic mutations in mitochondrial DNA as natural barcodes and thus track the *ex vivo* differentiation of human CD34⁺ HSC. Furthermore, mt-scATAC-seq was used to generate a dataset allowing the comparison of human CD34⁺ cell lineage differentiation *in vitro* with mouse LT-HSC differentiation *in vivo*. In this, we found that clonal lineages of human CD34⁺ cells and their mature progeny are normally distributed across the hematopoietic lineage, and no evidence of significant skewing of downstream cell types was observed. When single LT-HSC were transplanted into mice, we found limited variation in the ratio of myeloid and lymphoid cells between donors. Therefore, variation in myeloid and lymphoid engraftment is a stochastic process influenced by the irradiated recipient niche and not cell-intrinsic differences in LT-HSC.

Methods

Mitochondrial single-cell ATAC-sequencing library preparation

All donors provided written informed consent under protocols approved by an Institutional Review Board, and procedures for CD34⁺ cell collection were performed in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. mt-scATAC-seq libraries were generated by adapting the 10X Genomics protocol for single-cell ATAC-sequencing, according to modifications made by Lareau et al.²⁹ This modified protocol exploits a fixation step and a modified permeabilization

protocol to retain mitochondria in cells for subsequent capture and library preparation (Figure 1A). Whole cells were retained following an adapted protocol of 10X Genomics single-cell ATAC-sequencing. Cells were fixed in 1% formaldehyde and both digitonin and Tween 20 omitted in the lysis and wash buffers to generate a higher retention of mitochondrial DNA fragments per single cell. Cells were washed twice in phosphate-buffered saline and centrifuged for 5 min at 400xg and 4°C. Samples were multiplexed by taking the same number of cells per patient sample to get a total number of cells above 10⁵ to account for cell loss. Samples were incubated in lysis buffer for exactly 3 min on ice prior to washing. After centrifugation of the sample at 500xg for 5 min, the cell supernatant was discarded, and the cell pellet was resuspended in 1X Diluted Nuclei buffer. Cells were processed according to the Chromium Single Cell ATAC Solution user guide using the Chromium Next GEM Single Cell ATAC Library Kit, Chromium Next GEM Single Cell ATAC Gel Bead Kit, Chromium Next GEM Chip H Single Cell Kit, and Single Index Kit N Set A. Lastly, quality control tests were run on each library preparation using Agilent TapeStation High Sensitivity D1000 (Agilent) and Qubit dsDNA HS Assay kit (Invitrogen) prior to sequencing.

Four-step method for murine purification and transplantation of long-term hematopoietic stem cells

All animal experiments complied with the regulatory standards of, and were approved by, the Institutional Animal Care and Use Committee at Boston Children's Hospital. To evaluate long-term engraftment and differentiation of LT-HSC a four-step method for purification and transplantation of LT-HSC was used. First, bone marrow was pooled from the femoral and tibial bones of five to ten male and female 6- to 12-week-old C57BL/6J mice (expressing CD45.2) and red blood cells were lysed. Lineage depletion was performed using magnetic columns (Miltenyi) with a custom biotinylated lineage antibody cocktail (CD8a, CD4, CD11b, Gr1, Ter119, B220; BioLegend) with streptavidin magnetic beads (Miltenyi) (step one). Output cells were stained for fluorescence activated cell sorting with Lineage, CD34, CD48, CD150, CD117/cKit, Sca1 (BioLegend), and Fluorogold (Sigma). Bone marrow was bulk sorted (step two) for Lineage⁻ Sca-1⁺cKit⁺ (CD34⁻ LSK) cells on a BD FACS Aria II. The output of the first sorting was then plate-sorted for two, ten or 100 LT-HSC (LSK CD150⁺CD48⁻) into 96-well plates (step three). Two LT-HSC were sorted instead of targeting one LT-HSC to account for dead volume in the syringe during injection, thereby eliminating avoidable failures of engraftment. Cells were then cultured for 12-18 hours in Dulbecco's modified Eagle's medium/10% fetal bovine serum supplemented with 100 ng/mL murine stem cell factor (SCF), murine interleukin (IL)-3, murine IL-6 (BioLegend) at 37°C in 5% CO₂ (step four). The following day, cells were visualized on a microscope before transplantation. LT-HSC

were mixed with 2x10⁵ whole bone marrow from C57BL/6J CD45.1 wild-type mice and injected retro-orbitally into lethally irradiated (2 x 550R) recipient mice (CD45.1/CD45.2).

Results

Mitochondrial single-cell ATAC-seq defines clonal lineages in human CD34⁺ cells

To study differentiation of primary human CD34⁺ cells to committed myeloid and erythroid cells, CD34⁺ cells from two healthy donors were cultured for 72 h in the presence of recombinant human SCF, IL-3, IL-6, fms-like tyrosine 3 ligand (Flt3L), granulocyte colony-stimulating factor (G-CSF) and granulocyte-macrophage colony-stimulating factor (GM-CSF). Donor CD34⁺ cells were multiplexed for single-cell capture and library preparation before (input) or after culture (*Online Supplementary Table S1*, see Methods). Cells from an additional six healthy donors were analyzed to increase the power to delineate lineage rela-

tionships before and after cell culture. Using scATAC-seq, mitochondrial DNA somatic variants were identified and further used to generate clone-specific “natural” barcodes. These mitochondrial barcodes are inherited in clonal hierarchies to define clonality of CD34⁺ stem cells and their progeny. In parallel, the lower-coverage reads in nuclear open-chromatin regions were used for assigning cell lineages by assessing differential peaks in gene regulatory regions. Together these techniques allowed the fate tracking of human CD34⁺ HSPC to assess their lineage potential at a single-cell level. The modified scATAC-seq protocol adds a fixation step and an adapted permeabilization step designed to retain mitochondria in cells for subsequent capture and library preparation.²⁹ This method achieved uniform coverage across the mitochondrial genome and greater than 30-50x coverage in a minimum of 1,000 cells (Figure 1A, B) of the mitochondrial genome, which was suitable for variant calling and cellular “barcoding”. First, low-quality cells and mitochondrial alleles with low coverage and base-quality were

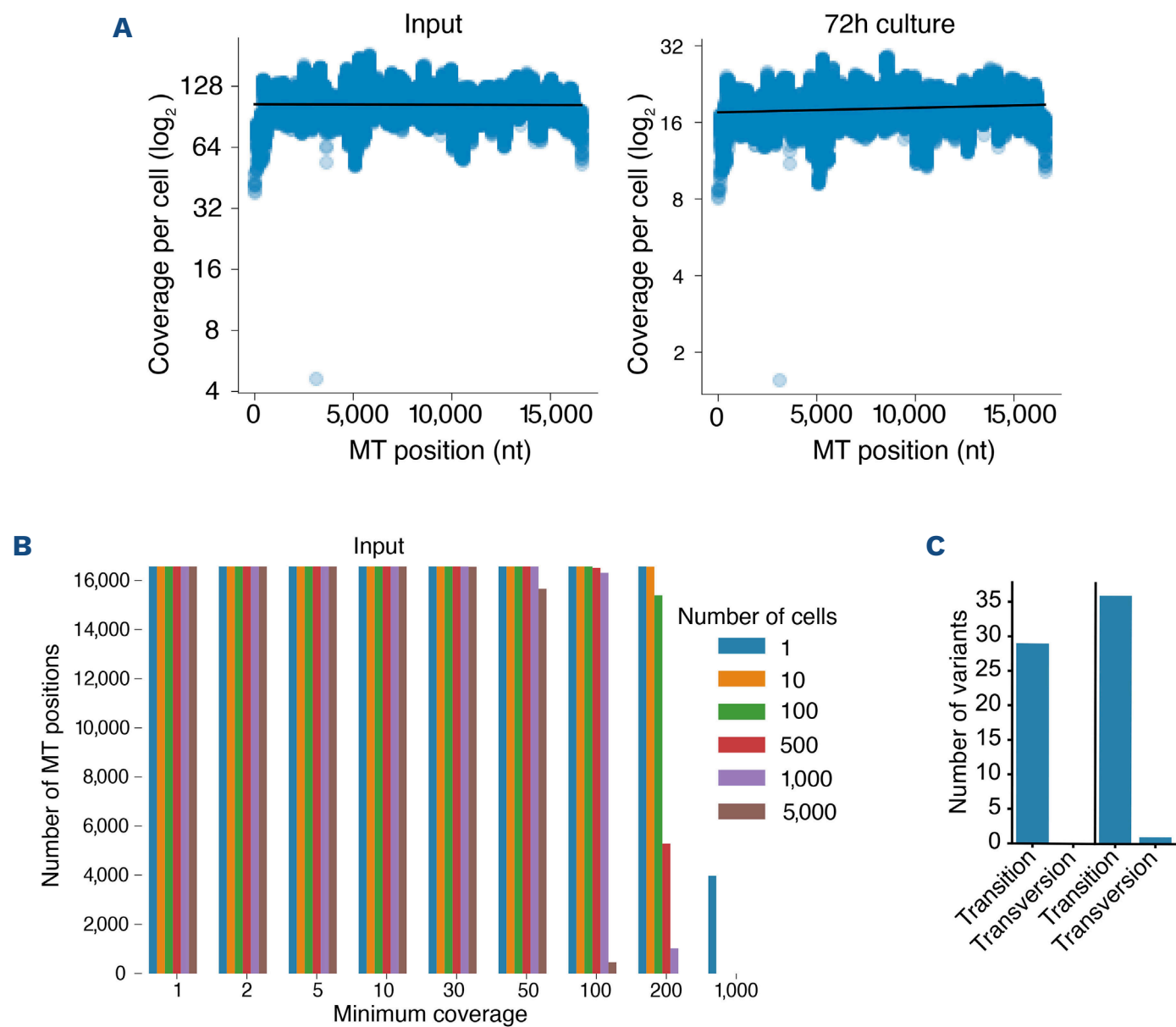


Figure 1. Mitochondrial single-cell ATAC-seq defines clonal lineages in mobilized human CD34⁺ cells. (A) Coverage across the mitochondrial genome in single cells across sequencing experiments. The black line is the mean at each position. (B) Number of mitochondrial positions covered across a range of cells and coverage thresholds across batches. (C) Mutation type for donor-specific variants. Transitions and transversions detected in donor 1 (left) and donor 2 (right). Donor-specific variants are defined as having greater than 0.8 variant allele frequency in over 90% of donor-assigned cells. MT: mitochondrial; nt: nucleotide.

filtered. Then additional variants were excluded using the Mitochondrial Genome Analysis Toolkit (MGATK), which removes variants with a low correlation of allelic reads across strands and a low variance-mean ratio (*Online Supplementary Figure S1A, B*, see Methods).²⁹ To de-multiplex each donor across conditions, the Vireo algorithm was applied on germline mitochondrial variants (*Online Supplementary Figure S1C-E*, see Methods). The donor predicted variant allele frequency (heteroplasmy) revealed 29 variants with high mean allele frequency (mean >0.7 in donors, <0.1 in others), highlighting many donor-specific variants that were primarily transition mutations (*Figure 1C*). Performance was assessed by varying the number of donors and calculating Vireo’s likelihood score, the evidence lower bound. The “elbow rule” confirmed that the true number of donors was the inflection point in which performance gain was reduced when additional possible donors were added to

the model (*Online Supplementary Figure S1D*). These data demonstrate that mt-scATAC-seq can be used to identify mitochondrial variants across the entire mitochondrial genome and to de-multiplex mixed samples.

Single CD34+ cells generate clonal populations of variable size identified by mitochondrial variants

Clone assignment is critical to track the fate of CD34+ stem cells at the single-cell level. Somatic mitochondrial variants were used to facilitate clone assignment in cultures of human CD34+ cells (*Figure 2A, Online Supplementary Figure S1A*), using a community-based k-nearest neighbors (KNN) clustering algorithm on heteroplasmy across single cells (*Online Supplementary Figure S2*).²⁹ After removing clones with fewer than five cells, 26-50 clones per donor were detected per capture (*Online Supplementary Table S1*). Clone size was variable between donors, with clone size in

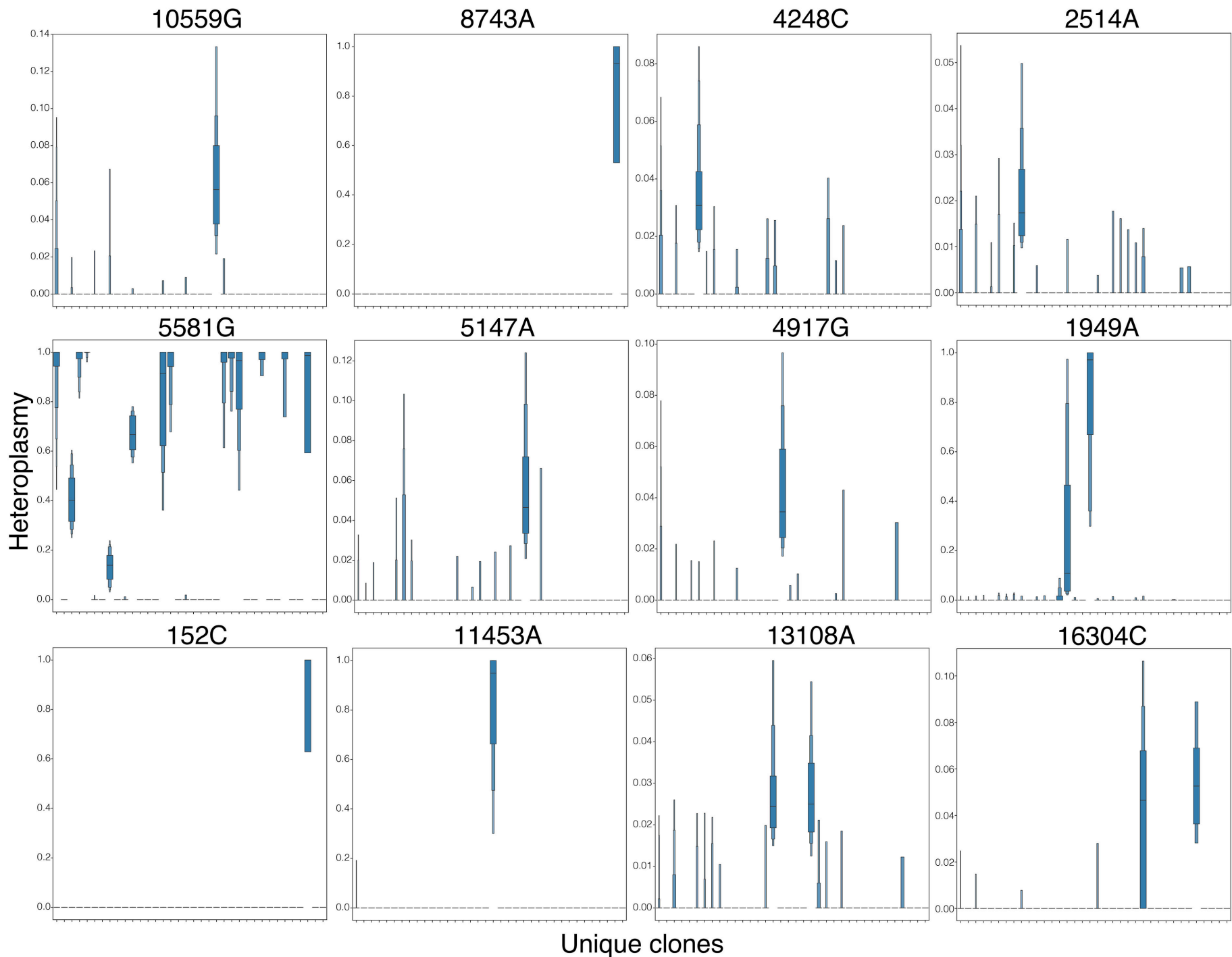


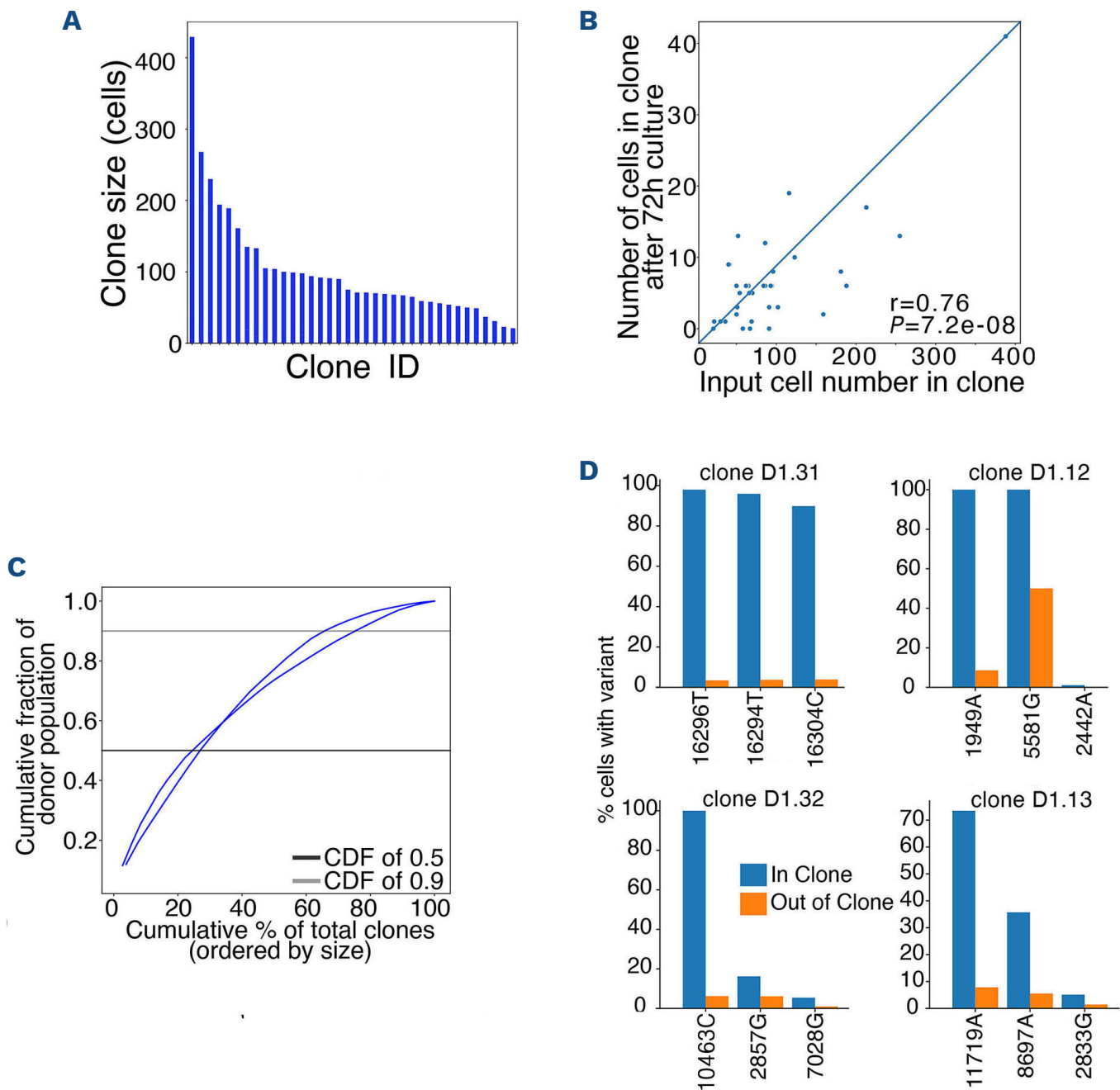
Figure 2. Mitochondrial heteroplasmy in detected clones. Distribution of somatic mitochondrial barcodes (y axis) across detected clones (x axis) in donor 1. Each variant is labeled as position followed by the alternative variant. Each boxplot is the heteroplasmy distribution of that allele in all cells within that clone.

the input sample correlating with the clone size after 72 h of culture (Figure 3A, B). Clones were identified based on combinations of unique mitochondrial variants (Figure 3D, E). Examination of the cumulative fraction of donor cells based on the percentage of total clones, ordered by size, indicated that the larger clones did not dominate the pool in these healthy donors (Figure 3C). Clone assignment was based on heteroplasmy (fraction of reads with the variant) rather than binary variant calls because the mitochondrial genome is heteroplasmic. Indeed, cells in clones were distinguished across a range of heteroplasmy (Figures 2A and 3E). Some variants were shared across clones, such as variant 5581G (Figure 3D, E), and these are predicted to have arisen from a common ancestral stem cell. Different variant-calling and clone-detection workflows were then compared and we observed a high concordance of cell-pair clonal relationships across parameters (see Methods). For example, we varied the number of neighbors (k) in the KNN algorithm and noticed a high concordance of clonal assignment with higher values of k. A lower performance was noted when the number of nearest neighbors in the KNN algorithm was low,

as it resulted in larger, more sparse clusters, leading to lower consistency in clones detected when running the algorithm on a subsample of the population (*Online Supplementary Figure S2*). While clonal detection using nucleotide barcoding is sensitive to the number of cells sampled and read depth, these smaller clones may represent a low fraction of total hematopoiesis.^{26,37} Taken together, these data demonstrate that mt-scATAC-seq can be used to identify clonal lineages in human CD34⁺ cells using mitochondrial variants acting as natural barcodes. Additionally, the proliferative potential of CD34⁺ cells in ex vivo culture conditions is not related to the size of the clone within the CD34⁺ pool (Figure 3B, C).

Mitochondrial single-cell ATAC-sequencing identifies lineage commitment of CD34⁺ cells in culture

After successfully mapping the clonal lineage of human CD34⁺ cells using the mitochondrial open-chromatin reads captured by mt-scATAC-seq, the nuclear genomic open-chromatin reads captured from the same experiments were then used to identify active genomic loci for determining cell lineage. Coupling mitochondrial and nu-



Continued on following page.

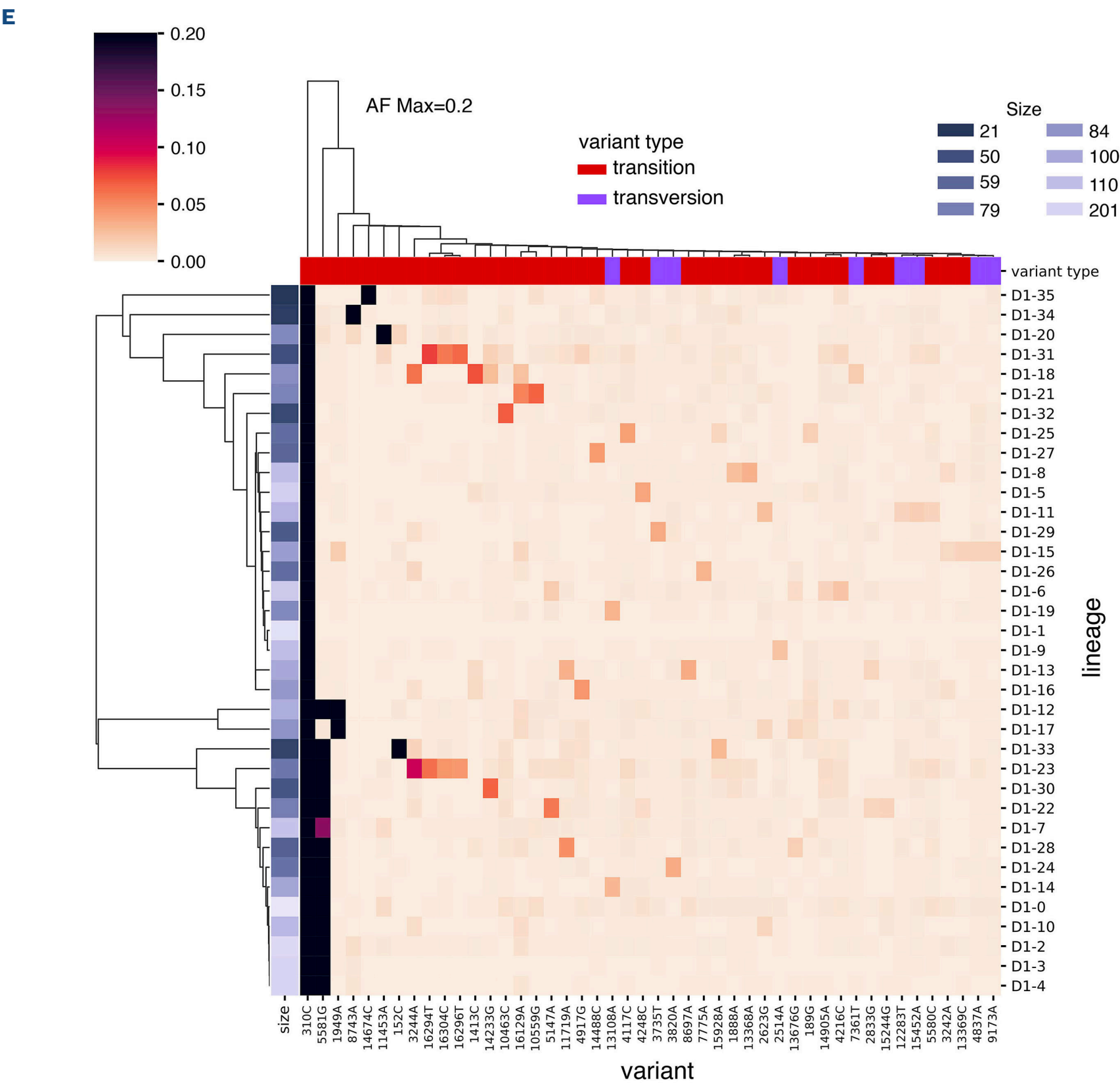


Figure 3. Single CD34⁺ stem cells expand into clones of variable sizes identified by mitochondrial variants. (A) Number of cells in each clone. (B) Scatterplot of number of cells across input and cultured cells. Pearson correlation r and P value are shown. (C) Cumulative distribution of the number of cells captured with increasing number of clones, sorted by largest to smallest, across both donors. Each donor line is colored by the conditions used. (D) Barcodes detected in representative clones. The percentages of cells with the barcode in the clone (blue) and outside of the clone (orange) are shown for the top distinguishing barcodes for each clone. (E) Average heteroplasmy of each variant in each clone. Maximum cut-off at 0.2. Variant types shown for each variant and number of cells for each clone. CDF: cumulative distribution function; AF: allele frequency.

clear chromatin reads enabled the differentiation potential of each clonal population of CD34⁺ cells to be mapped. Nuclear open-chromatin reads were processed using conventional scATAC-seq tools, enabling dimensionality reduction and cell clustering. Quality control of experiments showed a comparable number of detected peaks across experiments (*Online Supplementary Figure S5A-C*). Data from input CD34⁺ cells and cultured CD34⁺ cells were

embedded and clustered using uniform manifold approximation and projection (UMAP) and Louvain clustering. Twelve populations were identified using analysis of gene activity scores in exonic and promoter regions critical for each cell lineage (Figure 4A, Methods). As peaks were also found in intergenic and other non-coding regions (*Online Supplementary Figure S5D*), transcription factor motifs were also identified in open-chromatin peaks us-

ing ChromVar.³⁵ The gene activity scores of many lineage markers were differentially regulated between clusters supporting the differentiation of the human CD34⁺ cells to unique blood cell lineages, and also supporting the lineage-fate relationship of each cluster of these *ex vivo* cultured CD34⁺ human HSPC (*Online Supplementary Figure S4*). Transcription factor activity scores also supported the assignment of clusters towards the erythroid (GATA1-TAL1) and myeloid (SPI1 and CEBPE) lineages (*Online Supplementary Figure S4D*). To improve the resolution of the lineage assignment by gene activity and transcription factor activity, cells were integrated from all samples (*Online Supplementary Figure S4*). Subsequent analysis showed a higher resolution of clusters, including two clusters in the neutrophil developmental lineage (*Online Supplementary Figure S4B*). These data support the

dual use of mt-scATAC-seq for assigning clonal lineages with mitochondrial natural barcodes as well as nuclear accessible chromatin read signatures for assigning cell lineages.

Lineage bias of human CD34⁺ cells using mitochondrial single-cell ATAC-sequencing

The differentiation of CD34⁺ cells into diverse lineages in these assays indicated that an assessment of lineage bias could be performed on these clonal populations. When overlaying cells from clonal lineages and quantitatively analyzing their distribution, lineage-restricted cells were clustered with negligible evidence of lineage bias (Figures 4B, C and 5A-C, *Online Supplementary Figure S3E*). To test whether clones were biased to specific HSPC clusters, a hypergeometric test was applied, which

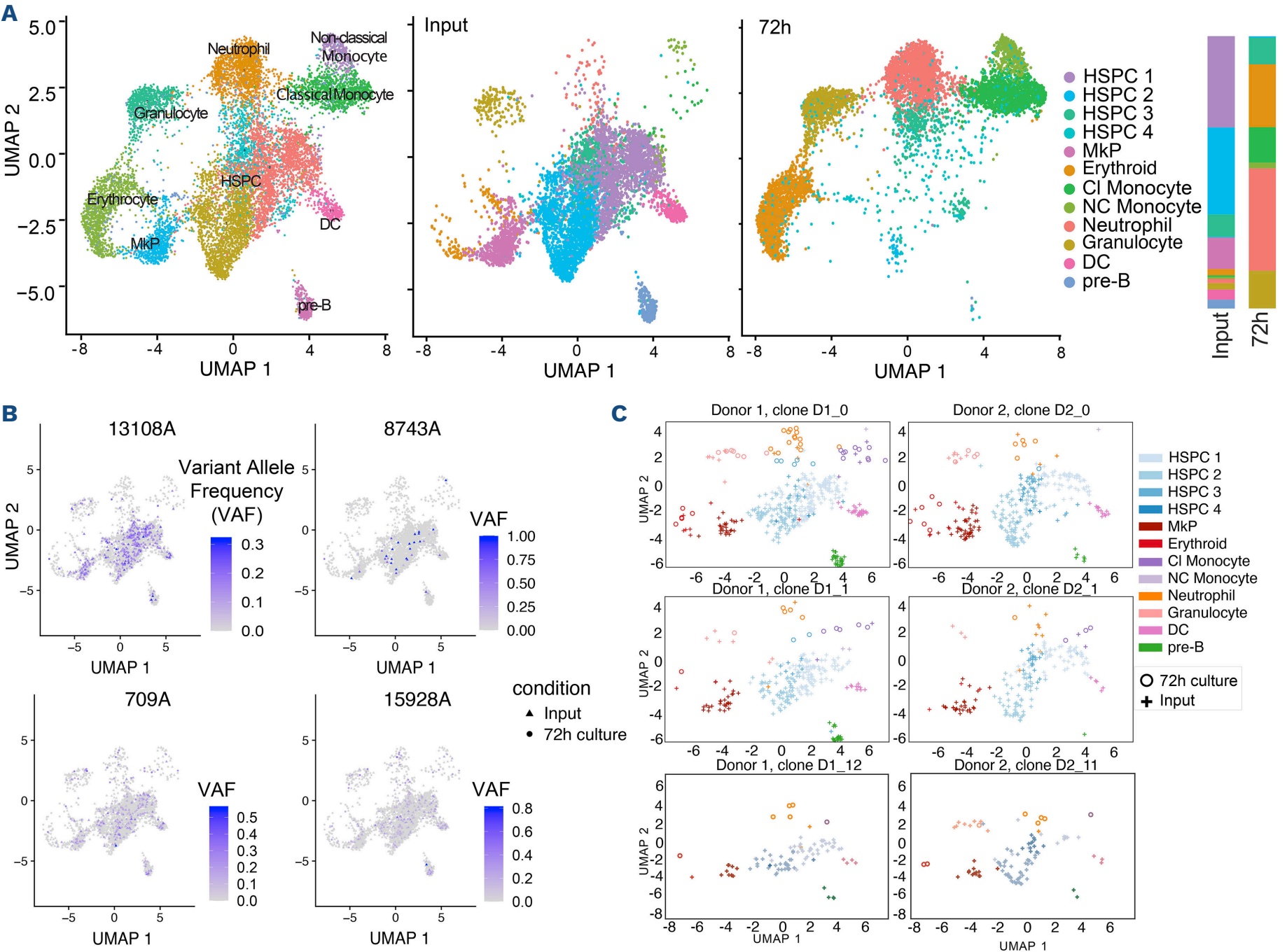


Figure 4. Mitochondrial single-cell ATAC-sequencing reveals minimum lineage-bias in the *ex vivo* differentiation of human CD34⁺ cells. (A) Uniform manifold approximation and projection (UMAP) of cells in both input and culture colored by cluster. Left: both conditions; middle: input; right: after 72 hours of culture. (B) Mitochondrial variants across UMAP for selected variants in donor 1. (C) Lineage fates of cells in input. Cells in a representative clone in each donor embedded in the UMAP. MkP: megakaryocyte progenitor; HSPC: hematopoietic stem and progenitor cell; pre-B: pre-B cell; DC: dendritic cell; CI: classical; NC: non-classical; VAF: variant allele frequency.

detected few clone-cluster pairs with evidence of bias (Figures 4C and 5A-C). The clones D1-12, D2-11, D2-17, and D2-22 were overrepresented in lymphoid, granulocyte, and HSPC clusters (Benjamini-Hochberg adjusted P value <0.1 , see Methods); however, these contained fewer cells

overall. Furthermore, the lymphoid lineage cells were not found in the ex vivo-differentiated condition (Figure 6, *Online Supplementary Figure S4*). The lineage bias was also examined after pooling data from all donors using the normalized entropy metric (see Methods). Using this

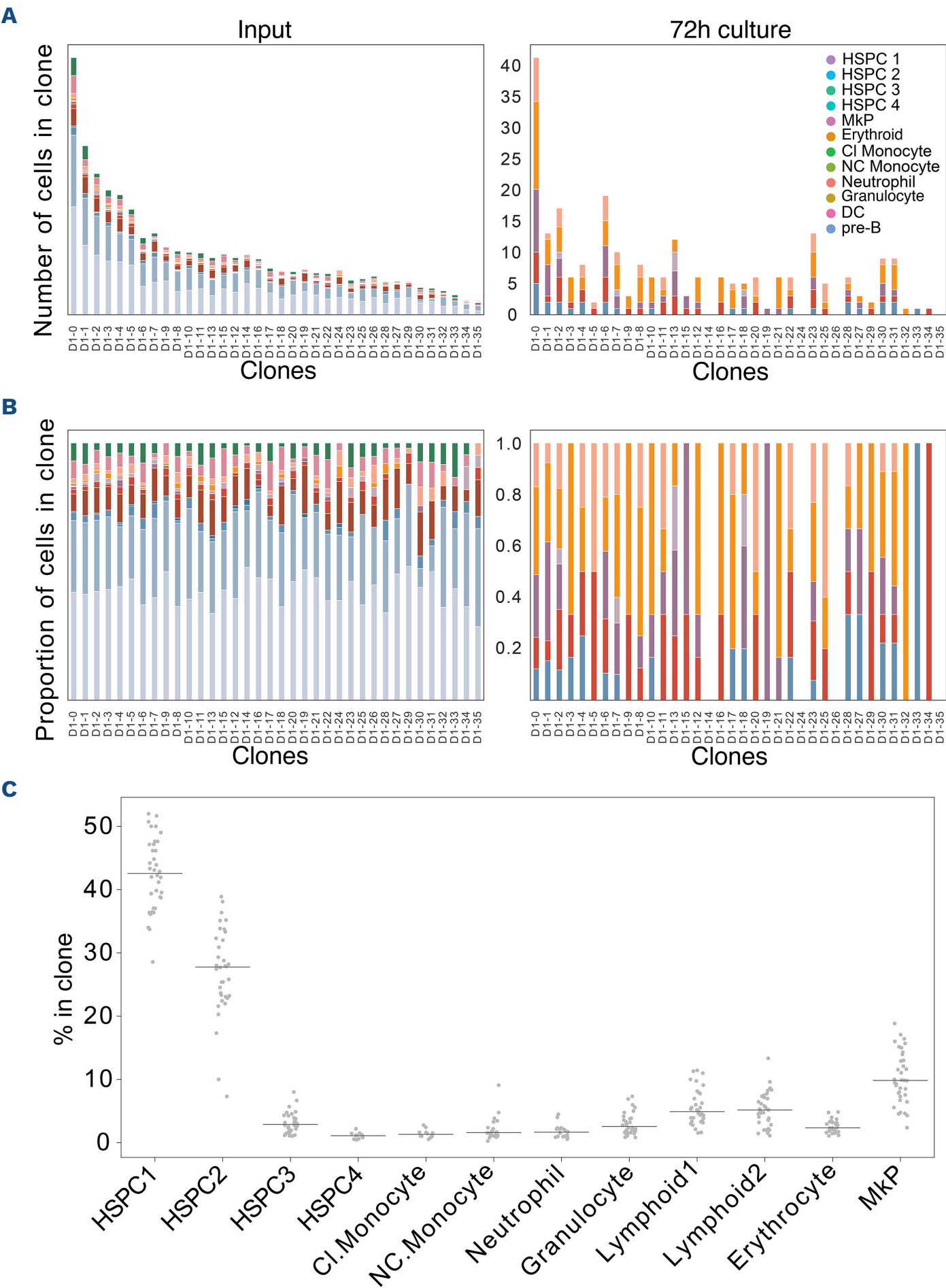


Figure 5. Tracking individual clonal differentiation of human CD34⁺ cells by mitochondrial single-cell ATAC-sequencing. (A) Raw cell counts in each clone, colored by lineage cluster. (B) Percent of hematopoietic cell type clusters across each clone for donors. The color legend is same as in (A). (C) Percent of lineage in a clone, across all clones for donor 1 and donor 2. HSPC: hematopoietic stem and progenitor cell; MkP: megakaryocyte progenitor; Cl: classical; NC: non-classical; DC: dendritic cell; pre-B: pre-B cell.

algorithm, if a clone is biased to one clone, it will display a lower entropy. Analysis of clones revealed little variability in clone entropy with smaller clones correlating with lower entropy (*Online Supplementary Figure S3F*). These data suggest that stochastic changes in lineage commitment are involved, rather than biased differentiation of LT-HSC.

To further define the relationship between clone size, cell type, and lineage bias potential in a CD34⁺ clonal lineage, the lineage proportions in each clone were compared across cell types. This was also measured across all donors using an entropy measurement in each lineage and converting the clonal counts into probability distributions for each lineage. The lineage proportions in each clone varied widely across cell type (Figures 4 and 5, *Online Supplementary Figure S3*). In contrast, the variation across

clones within each lineage was smaller within each lineage compared to between the lineages (Figure 5C, *Online Supplementary Figure S3F*). While most clones showed no skewing across clusters, it was necessary to determine whether the resolution of clone detection influenced these data. To do this, individual mitochondrial variant barcodes across the HSPC clusters were examined and no significant biases across the lineage clusters in donor 1 and donor 2 were found (Figures 4B, C and 5A-C). Together, these data suggest that detectable HSPC clones have multi-potent capacity contributing to hematopoiesis in humans without substantial lineage bias.

To validate the results of mt-scATAC-seq, flow cytometry of CD34⁺ cell cultures was performed in parallel. Dimensionality reduction and automated clustering were performed on human CD34⁺ cells prior to culture and after

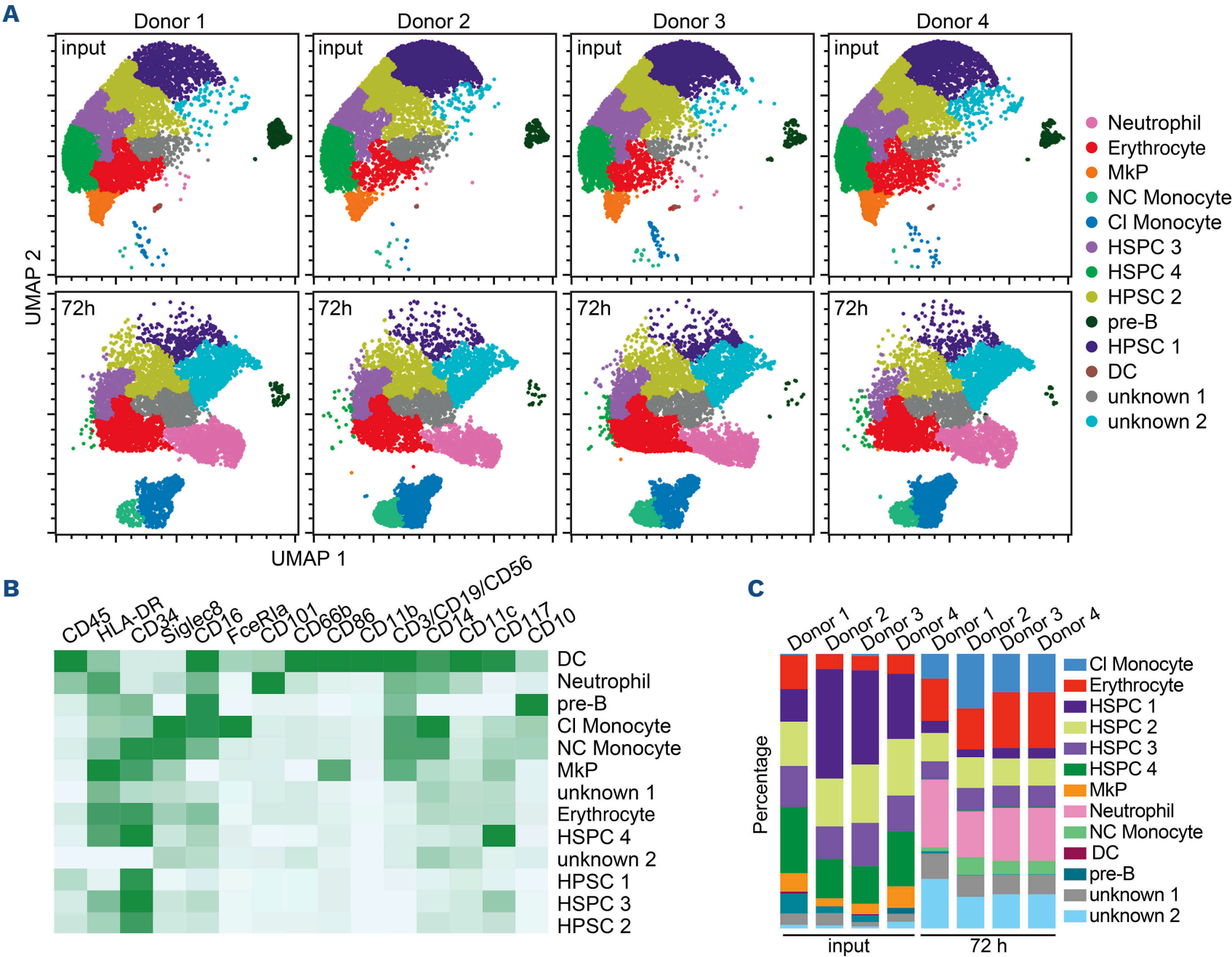
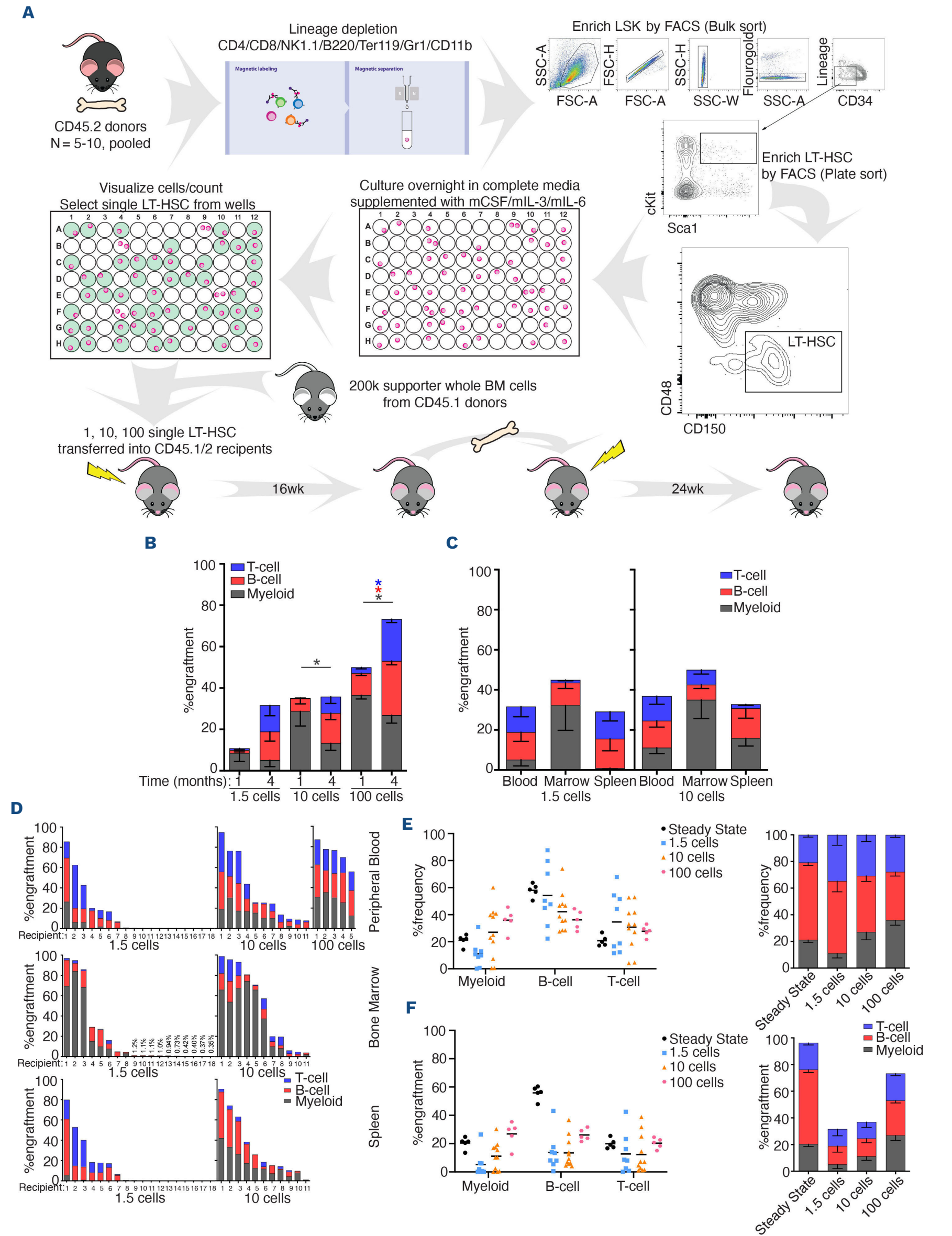


Figure 6. Flow cytometry identifies variable cell lineages in human CD34⁺ stem cells at steady state and in ex vivo culture. (A) Uniform manifold approximation and projection combined with FlowSOM clustering identified 13 cell clusters based on median fluorescence intensity (MFI) expression of CD45, HLA-DR, CD34, Siglec8, CD16, FcεRIα, CD101, CD66b, CD86, CD11b, CD3, CD19, CD56, CD14, CD11c, CD117, CD10. (B) Heatmap showing MFI expression levels of indicated markers in each cell cluster. (C) Bar graphs showing the proportion of each cluster in each donor before and after 72 hours of culture. UMAP: uniform manifold approximation and projection; MkP: megakaryocyte progenitor; NC: non-classical; Cl: classical; HSPC: hematopoietic stem and progenitor cell; pre-B: pre-B cell; DC: dendritic cell.



Continued on following page.

Figure 7. Long-term hematopoietic stem cell reconstitution varies by recipient, tissue, and time after engraftment. (A) Four-step method for purification and transplantation of long-term hematopoietic stem cell (LT-HSC) to evaluate long-term engraftment and differentiation. Step 1: bone marrow from 5–10 CD45.2 wild-type mice was pooled and lineage depletion was performed using magnetic columns with a custom lineage antibody cocktail. Output cells were stained for flow cytometry sorting. Step 2: CD34⁺ Lin[−]Sca⁺Kit⁺ (CD34⁺ LSK) cells were bulk sorted. Step 3: LT-HSC (LSK CD150⁺CD48[−]) were plate sorted into 96-well plates. Step 4: cells were cultured for 12–18 hours in complete media supplemented with murine colony-stimulating factor, murine interleukin-3 and murine interleukin-6. LT-HSC were mixed with whole bone marrow from CD45.1 wild-type mice and injected into lethally irradiated recipient mice. (B) Reconstitution of peripheral blood at 1 and 4 months after transplant of lethally irradiated mice with 1.5, 10, or 100 LT-HSC. * $P < 0.05$, 1 vs. 4 months for myeloid, T-cell, or B-cell comparisons. (C) Reconstitution of peripheral blood, bone marrow, and spleen at 4 months after transplant of lethally irradiated mice with 1.5 or 10 LT-HSC. (D) Reconstitution of peripheral blood, bone marrow, and spleen at 4 months after transplant of lethally irradiated mice with 1.5, 10, or 100 LT-HSC. Individual mice shown. Twelve to 18 mice engrafted with more than 1% donor LT-HSC-derived cells in myeloid and lymphoid compartments in the bone marrow. (E, F) Comparison of the lineage reconstitution in the peripheral blood of successfully engrafted lethally irradiated mice receiving 1.5, 10, or 100 LT-HSC compared to non-irradiated wild-type mice at steady state. Data are represented as the frequency (E) or percent engraftment (F) when compared to non-irradiated wild-type mice at steady state. FACS: fluorescence-activated cell sorting; SSC: side scatter; FSC: forward scatter; mCSF: murine colony-stimulating factor; mIL-3: murine interleukin-3; mIL-6: murine interleukin-6; BM: bone marrow.

72 h of culture using antibodies specific to human HSPC markers (CD34/CD117; c-Kit), lymphoid lineage markers (CD3/CD19/CD56), granulocyte lineage markers (CD66b/FcεRIα/Siglec8), monocyte lineage markers (CD14/CD16/CD86/CD11c), and additional developmental and maturation markers (CD45/CD10/CD101/CD11b/HLA-DR). Analysis of these flow cytometry data using dimensionality reduction identified 13 clusters including four HSPC subsets, one granulocyte-like cluster, one classical monocyte cluster, one non-classical monocyte cluster, one pre-B cluster, one dendritic cell cluster, one additional CD34⁺ cluster and three CD34[−] clusters (Figure 6A, B). Consistent with mt-scATAC-seq findings, the pre-B and dendritic cell clusters were no longer represented in the cell population after 72 h of culture, suggesting that these two populations may be contaminants of CD34⁺ cell sorting and die in the *ex vivo* culture. The frequency of these clusters identified by mt-scATAC-seq and flow cytometry are consistent, supporting the validity of the mt-scATAC-seq approach and the capacity of the CD34⁺ cells to differentiate into diverse cell types in the presence of SCF, IL-3, IL-6, Flt3L, G-CSF and GM-CSF (Figure 6C). These flow cytometry data support the *ex vivo* mt-scATAC-seq mitochondrial barcoding system to study short-term differentiation of human CD34⁺ cells to lineage-committed cells.

Assessing lineage bias in differentiation of single mouse hematopoietic stem cells *in vivo*

To compare the differentiation of clonal human CD34⁺ HSC, as assessed by mt-scATAC-seq, and mouse LT-HSC, a four-step method for mouse LT-HSC purification and transplantation was optimized and used to evaluate long-term engraftment and differentiation of single LT-HSC (Figure 7A). To confirm that the phenotype of single cells was consistent with HSC, reconstitution in primary and secondary lethally irradiated recipients was examined by multicolor flow cytometry panels that allowed for the identification of donor white blood cells (CD45.2) from recipient (CD45.1/CD45.2) and support (CD45.1) cells

(Online Supplementary Figure S7). The conditions used to isolate HSC yielded >1% engraftment of 12/18 recipient mice 4 months following transplant with a single HSC, and 100% of recipient mice transplanted with ten HSC or 100 HSC (Figure 7B–D). Changes in developmental-related reconstitution and ratios of myeloid and lymphoid cells were evident between 1 month and 4 months after reconstitution (Figure 7B) and also between blood, bone marrow, and spleen (Figure 7C). Flow cytometric analysis of donor myeloid and lymphoid populations in bone marrow, spleen, and blood, and across time in the peripheral blood, indicated that single LT-HSC were capable of high efficiency, long-term multi-lineage repopulation (Figure 7D–F). These data demonstrate that this sorting strategy results in high efficiency engraftment and reconstitution of the hematopoietic system.

When compared to wild-type steady-state (not transplanted) mice, flow cytometric analysis of hematopoietic tissues from irradiated single-cell HSC recipient animals revealed substantial variability in reconstitution of cells from the myeloid and lymphoid lineages, pointing to potential lineage bias of LT-HSC used for reconstitution (Figure 7D). However, this variability in lineage reconstitution was also evident when ten or 100 HSC were used to reconstitute animals (Figure 7E, F). In these latter cases, any bias would be predicted to be masked by the average response of the larger numbers of LT-HSC population in these animals. Taken together, the variable reconstitution in the myeloid and lymphoid lineages suggests that other factors, such as the inflammatory milieu and the hematopoietic niche of the lethally irradiated host, must also contribute substantially to the behavior of donor LT-HSC.

Discussion

Lineage bias in the HSC pool has been widely investigated *in vivo* but relies on a number of key assumptions: (i) the donor cell populations of HSC are pure and do not

contain any multipotent progenitors or lineage-committed progenitors; (ii) *in vivo* barcoding and other labeling methodologies are labeling HSC with high efficiency; (iii) that variation in the repopulation of myeloid and lymphoid cells in recipient animals reconstituted with single HSC is not influenced by irradiated host cells or support bone marrow donor cells such as Th1 cells, Foxp3⁺ regulatory T cells and NK cells; (iv) that skewed differentiation is not a consequence of pressures exerted on multipotent progenitors and/or lineage-committed progenitors; (v) that transplantation of typically >50 HSC can be relied on to investigate the cell-intrinsic lineage-bias of individual HSC; (vi) that the dynamic changes in myeloid:lymphoid:erythroid ratios following transplantation are acknowledged; and (vii) that dwindling proportions of donor hematopoietic cells over time are consistent with the behavior/function of HSC at steady state.³⁸⁻⁴²

In this study, significant variation was demonstrated in the ratio of myeloid:lymphoid cells in mice reconstituted with a single HSC. This could be interpreted as evidence of lineage bias in donor HSC; however, the myeloid:lymphoid ratio in mice reconstituted with ten HSC showed similar variation in the myeloid:lymphoid ratio. Critically, mice fully reconstituted using 100 HSC are not equivalent to the steady state, or mice reconstituted with one HSC or ten HSC. If lineage bias exists between HSC, then the variation in the overall myeloid:lymphoid ratio should reduce with increasing numbers of donor HSC as the average output of lineage-biased HSC approaches that of mice reconstituted with whole bone marrow or compared to mice at steady state. However, a reduction in the variation of myeloid:lymphoid ratio was not observed in mice receiving ten HSC despite an increase in successful engraftment and efficiency of reconstitution in these experiments. These findings suggest that environmental and inflammatory stimuli in the irradiated host are dominant factors that alter the differentiation of multipotent progenitors and lineage-restricted hematopoietic progenitor cells, while arguing for a more modest role of lineage bias in HSC capable of reconstituting mouse hematopoiesis. These findings are consistent with recent fate-mapping studies of fetal liver cells which concluded that the expansion potential of HSC is not pre-determined but rather influenced by the niche.²⁷ Similarly, human gene therapy studies also concluded that HSC distributed their clonal progeny without lineage bias.²⁸ These data support mathematical population dynamic modeling of myeloid-biased HSPC during inflammation.¹⁴

The methodology also provides an opportunity to enhance the study of mouse and human HSC *ex vivo* and *in vivo*. With improved, standardized, and concurrently validated

methods for HSC isolation and HSC lineage tracing, the study of genetic and biochemical regulators of HSC proliferation and differentiation can be explored using diverse methodology including clonal culture assays, single-cell transcriptomics and epigenomics, and *in vivo* differentiation and function.⁴⁰ The effects of cytokines and microbial products on gene expression, cell function, and donor cell engraftment can also be explored with higher resolution and confidence. Barcoding and transplantation techniques provide excellent clone-detection methods; however, the cellular and immune response to these editing techniques can introduce confounding effects that require consideration in experimental and therapeutic design,⁴³ a factor that is avoided using naturally occurring mitochondrial DNA barcodes.^{29,44-46} The rapid advances in novel conditioning regimens and genetically modified recipient mice with an enhanced capacity for engraftment demand improvements in reproducible methods for the isolation of pure HSC populations.^{41,42} Comparison of the engraftment efficiency of single HSC in lethally irradiated mice and in non-hemoablative murine models is also needed. Perhaps most relevant for the field of hematopoiesis would be a unified approach with unambiguous functional definitions and not simply molecular definitions when reporting the outcome of experiments that investigate HSC and hematopoietic progenitor cells.

Disclosures

No conflicts of interest to disclose.

Contributions

IS, MB-M, JCL, RM, HMM, BG, NEL, YPZ and BAC conceived the study. IS, MB-M, JCL, RM, NEL, YPZ and BAC were responsible for the methodology. IS, MB-M, JCL, NEL, YPZ and BAC performed the investigations. IS, MB-M, JCL, NEL, YPZ and BAC were responsible for visualization. HMM, BG, NEL and BAC acquired funding for the study. MB-M, HMM, BG, NEL, YPZ and BAC provided supervision. IS, MB-M, JCL, RM, HMM, BG, NEL, YPZ and BAC wrote the manuscript.

Funding

This work was supported by NIH grant R01AI155869, HL152958, R35GM119850, the V Foundation for Cancer Research, and a T32 award (5T32HL007574-36 to MB-M).

Data-sharing statement

All mouse data generated during this study are included in the manuscript. Human data needed to evaluate the conclusions in the paper are present in the paper, in the Online Supplementary Materials, or available in genetic sequencing databases.

References

- Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A*. 2011;108(50):20012-20017.
- Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232-236.
- Carrelha J, Meng Y, Kettle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*. 2018;554(7690):106-111.
- Grover A, Sanjuan-Pla A, Thongjuea S, et al. Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat Commun*. 2016;7:11075.
- Pinho S, Marchand T, Yang E, Wei Q, Nerlov C, Frenette PS. Lineage-biased hematopoietic stem cells are regulated by distinct niches. *Dev Cell*. 2018;44(5):634-641.e4.
- Lu R, Czechowicz A, Seita J, Jiang D, Weissman IL. Clonal-level lineage commitment pathways of hematopoietic stem cells in vivo. *Proc Natl Acad Sci U S A*. 2019;116(4):1447-1456.
- Busch K, Klapproth K, Barile M, et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*. 2015;518(7540):542-546.
- Sarrazin S, Mossadegh-Keller N, Fukao T, et al. MafB restricts M-CSF-dependent myeloid commitment divisions of hematopoietic stem cells. *Cell*. 2009;138(2):300-313.
- Muller-Sieburg CE, Cho RH, Karlsson L, Huang J-F, Sieburg HB. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood*. 2004;103(11):4111-4118.
- Kirschner K, Chandra T, Kiselev V, et al. Proliferation drives aging-related functional decline in a subpopulation of the hematopoietic stem cell compartment. *Cell Rep*. 2017;19(8):1503-1511.
- Hoggatt J, Mohammad KS, Singh P, Pelus LM. Prostaglandin E2 enhances long-term repopulation but does not permanently alter inherent stem cell competitiveness. *Blood*. 2013;122(17):2997-3000.
- Rundberg Nilsson A, Soneji S, Adolfsson S, Bryder D, Pronk CJ. Human and murine hematopoietic stem cell aging is associated with functional impairments and intrinsic megakaryocytic/erythroid bias. *PLoS One*. 2016;11(7):e0158369.
- Frisch BJ, Hoffman CM, Latchney SE, et al. Aged marrow macrophages expand platelet-biased hematopoietic stem cells via interleukin1B. *JCI Insight*. 2019;5(10):e124213.
- Singh A, Chia JJ, Rao DS, Hoffmann A. Population dynamics modeling reveals that myeloid bias involves both HSC differentiation and progenitor proliferation biases. *Blood*. 2025;145(12):1293-1308.
- Purton LE, Scadden DT. Limiting factors in murine hematopoietic stem cell assays. *Cell Stem Cell*. 2007;1(3):263-270.
- Kiel MJ, Yilmaz OH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell*. 2005;121(7):1109-1121.
- Rector K, Liu Y, Van Zant G. Comprehensive hematopoietic stem cell isolation methods. *Methods Mol Biol*. 2013;976:1-15.
- Balazs AB, Fabian AJ, Esmon CT, Mulligan RC. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood*. 2006;107(6):2317-2321.
- Christensen JL, Weissman IL. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A*. 2001;98(25):14541-14546.
- Naik SH, Perié L, Swart E, et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature*. 2013;496(7444):229-232.
- Lin DS, Tian L, Tomei S, et al. Single-cell analyses reveal the clonal and molecular aetiology of Flt3L-induced emergency dendritic cell development. *Nat Cell Biol*. 2021;23(3):219-231.
- Sun J, Ramos A, Chapman B, et al. Clonal dynamics of native haematopoiesis. *Nature*. 2014;514(7522):322-327.
- Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, et al. Clonal analysis of lineage fate in native haematopoiesis. *Nature*. 2018;553(7687):212-216.
- Spencer Chapman M, Ranzoni AM, Myers B, et al. Lineage tracing of human development through somatic mutations. *Nature*. 2021;595(7865):85-90.
- Kim S, Kim N, Presson AP, et al. Dynamics of HSPC repopulation in nonhuman primates revealed by a decade-long clonal-tracking study. *Cell Stem Cell*. 2014;14(4):473-485.
- Koelle SJ, Espinoza DA, Wu C, et al. Quantitative stability of hematopoietic stem and progenitor cell clonal output in rhesus macaques receiving transplants. *Blood*. 2017;129(11):1448-1457.
- Bowling S, Sritharan D, Osorio FG, et al. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. *Cell*. 2020;181(6):1410-1422.e27.
- Scala S, Basso-Ricci L, Dionisio F, et al. Dynamics of genetically engineered hematopoietic stem and progenitor cells after autologous transplantation in humans. *Nat Med*. 2018;24(11):1683-1690.
- Lareau CA, Ludwig LS, Muus C, et al. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat Biotechnol*. 2021;39(4):451-461.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013 May 26. doi:arXiv:1303.3997v2 [preprint, not peer-reviewed].
- Huang Y, McCarthy DJ, Stegle O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol*. 2019;20(1):273.
- Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*. 2019;9(1):5233.
- Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods*. 2021;18(11):1333-1341.
- Waltman L, van Eck NJ. A smart local moving algorithm for large-scale modularity-based community detection. *Eur Phys J*. 2013;86:471.
- Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14(10):975-978.
- Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-272.
- Bystrykh LV, Verovskaya E, Zwart E, Broekhuis M, de Haan G. Counting stem cells: methodological constraints. *Nat Methods*. 2012;9(6):567-574.
- van Galen P, Kreso A, Wienholds E, et al. Reduced lymphoid lineage priming promotes human hematopoietic stem cell

- expansion. *Cell Stem Cell*. 2014;14(1):94-106.
39. Gekas C, Graf T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood*. 2013;121(22):4463-4472.
40. Chapple RH, Tseng Y-J, Hu T, et al. Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Adv*. 2018;2(11):1220-1228.
41. Palchaudhuri R, Saez B, Hoggatt J, et al. Non-genotoxic conditioning for hematopoietic stem cell transplantation using a hematopoietic-cell-specific internalizing immunotoxin. *Nat Biotechnol*. 2016;34(7):738-745.
42. Czechowicz A, Palchaudhuri R, Scheck A, et al. Selective hematopoietic stem cell ablation using CD117-antibody-drug-conjugates enables safe and effective transplantation with immunity preservation. *Nat Commun*. 2019;10(1):617.
43. Ansari AM, Ahmed AK, Matsangos AE, et al. Cellular GFP toxicity and immunogenicity: potential confounders in in vivo cell tracking experiments. *Stem Cell Rev Rep*. 2016;12(5):553-559.
44. Cromer MK, Vaidyanathan S, Ryan DE, et al. Global transcriptional response to CRISPR/Cas9-AAV6-based genome editing in CD34+ hematopoietic stem and progenitor cells. *Mol Ther*. 2018;26(10):2431-2442.
45. Ludwig LS, Lareau CA, Ulirsch JC, et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell*. 2019;176(6):1325-1339.e22.
46. Miller TE, Lareau CA, Verga JA, et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat Biotechnol*. 2022;40(7):1030-1034.