

Simultaneous Assessment of Kinship, Division Number, and Phenotype *via* Flow Cytometry for Hematopoietic Stem and Progenitor Cells

Alessandro Donada¹, Tamar Tak¹, Giulio Prevedello^{1,2,3,4}, Idan Milo¹, Ken R. Duffy², Leïla Perié¹

¹Quantitative Immuno-hematology, CNRS UMR168, Institut Curie ²Hamilton Institute, Maynooth University ³Institut Curie, PSL Research University, CNRS UMR 3348, Orsay ⁴Université Paris-Sud, Université Paris-Saclay, CNRS UMR 3348, Orsay

Corresponding Authors

Alessandro Donada

alessandro.donada@curie.fr

Leïla Perié

Leila.Perie@curie.fr

Citation

Donada, A., Tak, T., Prevedello, G., Milo, I., Duffy, K.R., Perié, L. Simultaneous Assessment of Kinship, Division Number, and Phenotype *via* Flow Cytometry for Hematopoietic Stem and Progenitor Cells. *J. Vis. Exp.* (193), e64918, doi:10.3791/64918 (2023).

Date Published

March 24, 2023

DOI

10.3791/64918

URL

joVE.com/video/64918

Abstract

Few techniques can assess phenotype and fate for the same cell simultaneously. Most of the current protocols used to characterize phenotype, although able to generate large datasets, necessitate the destruction of the cell of interest, making it impossible to assess its functional fate. Heterogeneous biological differentiating systems like hematopoiesis are therefore difficult to describe. Building on cell division tracking dyes, we further developed a protocol to simultaneously determine kinship, division number, and differentiation status for many single hematopoietic progenitors. This protocol allows the assessment of the *ex vivo* differentiation potential of murine and human hematopoietic progenitors, isolated from various biological sources. Moreover, as it is based on flow cytometry and a limited number of reagents, it can quickly generate a large amount of data, at the single-cell level, in a relatively inexpensive manner. We also provide the analytical pipeline for single-cell analysis, combined with a robust statistical framework. As this protocol allows the linking of cell division and differentiation at the single-cell level, it can be used to quantitatively assess symmetric and asymmetric fate commitment, the balance between self-renewal and differentiation, and the number of divisions for a given commitment fate. Altogether, this protocol can be used in experimental designs aiming to unravel the biological differences between hematopoietic progenitors, from a single-cell perspective.

Introduction

The past decade was marked by the worldwide spreading of single-cell approaches to cellular and molecular biology. Following in the steps of single-cell genomics^{1,2}, nowadays it is possible to study many components of a single cell (e.g.,

DNA, RNA, proteins), with novel single cell -omics techniques burgeoning every year. These techniques have shed light on old and new questions for the fields of immunology, neurobiology, oncology, and others, both using human

and model organism cells³. By highlighting the differences between individual cells, single cell -omics prompted the definition of a new model of hematopoiesis, centered on the heterogeneity of hematopoietic stem and progenitor cells (HSPCs) and moving away from the classical model of discrete homogeneous populations^{4,5}.

One of the few drawbacks of all the -omics techniques is the destruction of the cell of interest, precluding the possibility to assess its functionality. Conversely, other single-cell methods, such as single-cell transplantation assay and lineage tracing technologies, provide a readout of the functionality of the ancestor cell by assessing the fate of individual cells *in vivo*^{6,7}. Lineage tracing technologies involves labeling the cell of interest with a heritable genetic⁷ or a fluorescent label^{8,9}, allowing the fate of multiple single cells to be followed at the same time. However, the characterization of the starting cells is typically limited to a restricted number of parameters, such as the expression of a few surface proteins assessed by flow cytometry¹⁰. In addition, single-cell lineage tracing technologies require laborious detection of the cellular label, typically *via* DNA/RNA sequencing or imaging. This last point in particular limits the number of conditions that can be tested in a single experiment.

Another class of methods that are used to study the functionality of single cells are *ex vivo* cell culturing systems of single HSPCs. Easy to perform, these gold standard assays involve the sorting of individual cells into 96-wells cell culture vessels, and after culture, characterizing the cell progeny phenotype, typically by flow cytometry or morphological analysis. These assays have mostly been used to characterize the long-term differentiation of HSPCs into mature cells, typically after 2-3 weeks of culture^{11,12}.

Alternatively, they have been used to try to maintain and expand *ex vivo* HSPCs^{13,14,15,16,17,18}, with the promise of medical benefit for human stem cell transplantation¹⁹. Lastly, they have been used to study the early commitment of HSPCs using short-term culture²⁰, with the low number of cells generated in this culture being the main limiting factor. One drawback of these different kinds of *ex vivo* assays is that they only partially reflect the *in vivo* complexity; still, they are one of the rare ways to study human HSPC differentiation.

One missing piece of information from existing single-cell methods (single cell-omics, lineage tracing, and *ex vivo* culture) is the accurate detection of cell divisions, an essential parameter to consider when studying HSPC dynamics²¹. A simple way to assess the number of divisions *via* flow cytometry is the use of soluble "protein dyes", like the 5-(and 6)-carboxyfluorescein diacetate succinimidyl ester (CFSE)²². These division dyes diffuse inside the cytoplasm of the stained cells, and diluted by half and passed to the two daughter cells at every cell division, allowing to enumerate up to 10 divisions. Combining several division dyes, it is possible to seed multiple individual progenitors in the same well, as each individual dye allows separation of the different descendants. This is the principle behind the use of cell dyes for multiplex clonal and division-tracking that was first introduced for murine lymphocytes^{23,24}.

Here, we present the development of the MultiGen assay for use with murine and human HSPCs. It allows the testing of many single cells simultaneously for their properties of differentiation, division, and kinship *ex vivo*. This high-throughput, easy to perform, and inexpensive assay allows to measure the cellular phenotype, the number of divisions performed, and the cell kinship and clonal relationship with the other cells in the well, all at the same time. It can

be used to quantitatively assess symmetric and asymmetric fate commitment, the balance between self-renewal and differentiation, and the number of divisions necessary for a given commitment fate. The protocol requires a fluorescence activated cell sorter (FACS) and a flow cytometer with a plate reader, plus the equipment necessary to perform cell culture. In addition to the technical protocol for execution of the assay on human HSPCs, we also provide the detailed analysis framework, including the statistical testing necessary to assess cellular properties related to the concept of cell family²⁵. This protocol has already been used successfully to describe the murine HSPC compartment^{26,27}.

The following protocol uses magnetically enriched CD34⁺ cells as the starting material²⁸. In this way, it is possible to efficiently stain and isolate the human HSPCs from different blood sources (e.g., cord blood, bone marrow, peripheral blood). It is important not to discard the CD34⁻ fraction, as it will be used as part of the protocol to set different types of experimental controls. The mentioned cell amounts and volumes can be scaled up or down, according to the experimental workflow and necessities. Similarly, the protocol can be adapted to the study of different types of progenitors, simply by modifying the antibodies used for the cell sorting and flow cytometry steps.

Protocol

For the following protocol, de-identified umbilical cord blood was used as the HSPC source and collected in accordance with the guidelines defined by the Saint-Louis Hospital cord blood biobank (authorization AC-2016-2759) and with the Declaration of Helsinki.

NOTE: Before starting, ensure that all reagents and equipment necessary for this protocol are available, as listed

in the **Table of Materials** and mentioned in the protocol. Prepare the relevant reagents fresh and do not store them, unless explicitly mentioned.

1. Cell dye staining

NOTE: This section describes staining with four combinations of CFSE and violet dye (CTV) cell division dyes. Process all the tubes simultaneously, even if no cell dye solution is added. All steps are performed in sterile conditions to allow the following cell culture step. Time required: approximately 100 min.

1. Process the cord blood unit according to a magnetic sorting protocol²⁹. Ensure that two fractions are available: a large CD34⁻ fraction and a smaller CD34⁺ fraction. Spin both tubes for 5 min at 300 x g. Aspirate the supernatant without disturbing the pellet.
2. For the CD34⁺ fraction, resuspend it in 1 mL of Dulbecco's modified Eagle medium (DMEM) without fetal bovine serum (FBS). Count the cells using a hemocytometer; the cell density should not be higher than 3 x 10⁶ cells/mL. If this is the case, adapt the volume accordingly. For the CD34⁻ fraction, resuspend in DMEM w/o FBS, and adjust the volume to a maximum of 6 x 10⁶ cells/mL.
3. Aliquot 250 µL of the CD34⁺ fraction into four 15 mL polypropylene tubes. Label the tubes as follows: CD34⁺/CF (CFSE_only), CD34⁺/CV (CFSE_high CTV_low), CD34⁺/VC (CFSE_low CTV_high), and CD34⁺/VI (CTV_high). Aliquot 250 µL of the CD34⁻ fraction into another four 15 mL polypropylene tubes. Label the tubes as follows: CD34⁻/CF (CFSE_only), CD34⁻/

- CV (CFSE_{high} CTV_{low}), CD34⁻/VC (CFSE_{low} CTV_{high}), and CD34⁻/VI (CTV_{high}). The remaining cells from the CD34⁻ fraction can be discarded.
- Prepare two CFSE solutions, named CFSE_{high} and CFSE_{low}. For CFSE_{high} (10 μM), mix 1.1 mL of DMEM w/o FBS with 2.2 μL of CFSE stock (5 mM) solution. For CFSE_{low} (5 μM), mix 550 μL of DMEM w/o FBS and 0.55 μL of CFSE stock (5 mM) solution.
 - Add 250 μL of the CFSE_{high} solution to the CF and CV tubes, 250 μL of the CFSE_{low} solution to the VC tubes, and 250 μL of DMEM w/o FBS to the VI tube. To ensure an efficient mix of cell suspension and cell dye, incline the tube by almost 90 degrees, and deposit the CFSE solutions on the tube wall. Then, hold the tube vertically to mix the two solutions, and pipette three or four times to ensure a fast mixing of the CFSE solutions with the resuspended cells. Incubate at 37 °C for precisely 8 min.
 - After the incubation, add 5 mL of DMEM + 10% FBS. Keep the tubes at 37 °C for 5 min.
 - Spin the tubes for 5 min at 300 x g. Remove the supernatant *via* aspiration without disturbing the pellet, and wash the pellet with 5 mL of phosphate buffered saline 1x/ethylenediaminetetraacetic acid (PBS 1x/EDTA). Spin again for 5 min at 300 x g. Discard the supernatant without disturbing the pellet, and resuspend the cell pellet in 250 μL of 1x PBS /EDTA.
 - Prepare two CTV solutions, named CTV_{high} and CTV_{low}. For CTV_{high} (10 μM), mix 1.1 mL of PBS 1x/EDTA and 2.2 μL of CTV stock (5 mM). For CTV_{low} (5 μM), mix 550 μL of PBS 1x/EDTA with 0.55 μL of CTV stock (5 mM).
 - Add 250 μL of the CTV_{high} solution to the VC and VI tubes, 250 μL of the CTV_{low} solution to the CV tube,

and 250 μL of 1x PBS /EDTA to the CF tube. Use the same technique as described in step 1.5. Incubate at 37 °C for precisely 8 min.

- After the incubation, add 5 mL of DMEM + 10% FBS. Keep at 37 °C for 5 min.
- Spin the tubes for 5 min at 300 x g, discard the supernatant without disturbing the pellet, and then wash the pellet with 5 mL of 1x PBS /EDTA. Spin again for 5 min at 300 x g.
- Discard the supernatant without disturbing the pellet, and resuspend the CD34⁻ fractions in 1x PBS /EDTA for a final concentration of 1.5 x 10⁶ cells/mL. Resuspend the CD34⁺ fractions in 40 μL of staining buffer, and transfer the cells into 1.5 mL tubes.

2. Antibody staining

NOTE: Antibody staining can be customized according to the experimental needs. Only the CD34⁺ fractions undergo antibody staining; the CD34⁻ fractions are used as a single staining control for the cell division dye combinations (CV, VC, CF, and VI fractions). The following panel is tailored for the detection of four types of HSPCs: hematopoietic stem cells (HSCs), multipotent progenitors (MPPs), lymphoid-primed multipotent progenitors (LMPPs), and hematopoietic progenitor cells (HPCs)¹². However, identification of HSCs and MPPs is presented. Time required: 75 min.

- Prepare the single staining for surface staining, using compensation beads. Mix negative beads and Immunoglobulin G (IgG) beads in a 1:1 ratio, for a total volume equivalent to 20 μL x the number of surface markers (e.g., 120 μL if the staining panel contains six antibodies).

2. Dispatch 20 μL of beads into individual 1.5 mL tubes for each marker. Add the volume corresponding to the dilution factor for each antibody in the corresponding tube (e.g., if the dilution factor is 1:20, add 1 μL).
3. To stain the CD34^+ cells, prepare a master mix of antibodies¹², based on **Table 1**. Mix the antibodies into a single 0.5 mL tube. Add 7 μL from the antibody master mix to each of the four CD34^+ conditions.
4. Incubate the compensation beads and the CD34^+ samples at 4° C for at least 30 min.

NOTE: The incubation time must be adapted to the technical details of the antibodies used for staining.
5. During incubation, prepare the 96-well round-bottom plate to be used for sorting, adding 100 μL of cell culture medium to each well using a multichannel pipette.

NOTE: Leave wells H8-H12 empty.
6. Label 5 mL polypropylene tubes for the surface staining controls (5, using beads), the cell division dyes controls (4, using the CD34^- fractions), and the CD34^+ samples (4).
7. At the end of the incubation, wash the cells and beads with 1 mL of staining buffer. Transfer the total volume to the 5 mL polypropylene tubes. Centrifuge the tubes for 5 min at 300 x g, then aspire the supernatant without disturbing the pellet.
8. Resuspend the cells in staining buffer, using approximately 500 μL each for the beads and CD34^+ cells, and 1 mL for the CD34^- tubes.

Table 1: Template to prepare the antibody mastermix for a cell sorting experiment. [Please click here to download this Table.](#)

3. Cell sorting

NOTE: The sorted cell numbers could vary according to the total amount of cells available. In the protocol, a minimum cell number for each control is provided. Time required (for a single plate): 100 min.

1. Open the template experiment or set a new experiment. Create a single specimen and multiple tubes, one per each condition.
2. Set the gating strategy detailed in **Figure 1**, creating six dot plot diagrams. First, visualize the cells on a FSC-A/SSC-A dot plot, and double click on the **polygon gating tool** to select a population with low side scatter (**Figure 1A**). In the following dot plot (FSC-A/FSC-H), right click on the **plot** and select the gate "**Cells**" from the dropdown menu by clicking on it. Use the same **gating tool** to select a tight population on the diagonal between the two axis (**Figure 1B**).
3. In the third dot plot (APC vs. FSH-H), display the population "**Single Cells**" and gate the cells negative for the expression of APC Lineage (Lin) (**Figure 1C**). In the fourth plot (CFSE vs. CTV), display the population "Lin⁻" and create four separated gates, one for each dye combination (**Figure 1D**).

NOTE: These gates must be tight, to select only a small fraction of homogeneously stained cells.
4. Use the fifth and sixth plot (APC-Cy7 vs. BV650 and PE-Cy7 vs. PE) to identify the progenitors of interest. Generously gate the **$\text{CD34}^+\text{CD38}^-$** population and the **$\text{CD34}^+\text{CD38}^+$** in the fifth plot (**Figure 1E**). Then, select the **$\text{CD34}^+\text{CD38}^-$** population in the sixth plot, and draw three gates, according to **Figure 1F**.

5. Run the single staining tubes containing the compensation beads, clicking on the **Acquire** button. Adjust the photomultiplier tube (PMT) voltages from the **Parameters** dropdown menu, particularly for the cell division dyes (between 10^4 and 10^5 on a biexponential scale).
6. Refine the compensation matrix accordingly to the panel used for sorting, using the **Compensation** tab. Record at least 5,000 events in the beads gate, clicking on the **Record** button.
7. Run the CD34⁻ fractions and check the compensation matrix again. **Record** at least 10,000 events in the single-cell gate.
8. Run the CD34⁺ fractions, recording at least 5,000 events in the single-cell gate. Adjust the gate for each dye combination, setting a tight gate to select a homogeneous population (**Figure 1D**). Similarly, adjust the gating for selecting HSCs and MPPs.
9. Once the analysis is completed and all tubes are recorded, insert the plate in the appropriate holder, after performing the standard Aria calibration for sorting on 96-well plates. Cooling the plate is recommended.
10. Prepare the plate sorting template according to the scheme presented in **Table 2**, using the experiment **sort layout**. The wells named "CD34⁻" contain 5,000-10,000 cells, sorted on the CF/CV/VC/VI gate. The "Bulk" wells contain at least 500 cells, sorted on the gate CD34⁺CD38⁻. Finally, the single-cell wells contain only one event per cell division dye combination per well, so four events per well in total.
NOTE: The "Bulk" populations can be adapted to specific progenitors' subset; do not sort less than 500 cells.
11. For the sorting, proceed in order, completing every cell division dye combination before moving to the next one. For example, start with sorting the CD34⁻ CF, in **yield** purity mode. Click on the acquire button, then on the sort button.
12. At the end of the CD34⁻ sorting, insert the CF CD34⁺ tube. Acquire, then click on the sort button, making sure to have ticked **0/16/0** as purity grade. Finally, sort the cells of interest, one cell per well, in **single-cell** purity, making sure to tick the **Index Sorting** option.
13. Move to the following cell division dye combination, repeating the same order. As a reference, **Table 2** provides an example of a sorted plate.
NOTE: The index sorting function generates individual files for each sorted condition.
14. At the end of the sorting, export the files as .fcs 3.0 files. Put the cells in a 37 °C, 5% CO₂ incubator. The cells are cultured for multiple days, according to the experimental design, for at least 24 h²⁶.

Table 2: Template for a cell sorting 96-well plate, based on the specific requirements for the successive flow cytometry analysis. [Please click here to download this Table.](#)

4. Cell sorting data analysis

NOTE: To validate the quality of cell sorting, the FACS data analysis is necessary before moving further. The main output of this step is the generation of a spreadsheet containing the marker intensities of each sorted individual cell.

1. Upload the .fcs 3.0 files in the analysis software.

2. Verify the compensation setting used during cell sorting, using the single staining files recorded before the actual sorting.
3. Set the revised gating strategy using the files corresponding to the different bulk. Copy and paste those gates on the index sorting files.
4. Check that the index sorted cells fell in the set gate. If there are some sorted cells that were incorrectly gated, they can be identified by exporting the plate coordinates recorded during index sorting, and removed later in the analysis.
5. **Export** the event from the index sorting files as compensated parameters. Export them as .csv files, ticking the options "**scale values**" and "**compensated parameters**". These files should be exported in a folder named "**Exported files**".
6. Combine all the files in a single .csv file, using the script in **Supplemental File 1**. Set the right path with the function "setwd". The output of this script is a spreadsheet containing all the differently gated events and the relative intensities for all the parameters.

5. After culture antibody staining

NOTE: Perform this part of the protocol in sterile conditions; several reagents are shared with the former steps, and need to stay sterile. For the flow cytometry analysis, use a flow cytometer with a plate reader. This allows performing the staining directly in the tissue culture plate, reducing the cell loss to a minimum by limiting the amount of pipetting and spinning. Prepare the surface marker single color staining using the compensation beads, except for wells A1-A4, which represent the single staining for the CF/CV/VC/VI colors and are already present in the 96-well plate. The bulk populations

sorted according to the cell dye help setting the gating strategy for the number of divisions and the general gating. Time required: 120 min.

1. Before starting the protocol, mark the wells that contain at least one cell by checking the plate under an inverted microscope. This step allows to optimize the amount of antibody used for the staining and speeds up the procedure.
2. Prepare the antibody mastermix, according to **Table 3**. As there is a significant amount of pipetting, the table considers the technical error due to pipetting, including an extra 5% volume. The antibodies described in the table allow to characterize a range of HSPCs from human cord blood samples¹².
3. Centrifuge the plate for 5 min at 300 x g. Rapidly invert the plate under the hood and over a paper towel, to remove the supernatant.
4. Add 8 µL of staining buffer to wells A1-A4. Add 8 µL of the mix to the other wells.
5. Mix the negative beads and IgG compensation beads in a 1:1 ratio, for a total volume equivalent to 120 µL. Dispatch 20 µL in one 1.5 mL tube per each marker. Add the antibody volume corresponding to the dilution factor (e.g., if the dilution factor is 1:20, add 1 µL).
6. Incubate the plate and the single staining compensation controls at +4 °C, for at least 30 min.

NOTE: The incubation time must be adapted to the technical details of the antibodies used for the staining.

7. Wash the beads with 1 mL of staining buffer. Transfer the total volume to the 5 mL polypropylene tubes previously labeled. Centrifuge the tubes for 5 min at 300 x g, then remove the supernatant *via* aspiration.
8. Wash the cells in the plate by adding 100 µL of staining buffer per well using a multi-channel pipette. Centrifuge the plate at 300 x g for 5 min, then rapidly invert the plate under the hood and over a paper towel, to remove the supernatant.
9. Re-suspend the cells in 85 µL of staining buffer, using a multi-channel pipette.
10. Start the analysis on the flow cytometer (**Acquisition mode**), using the dedicated template and clicking on **custom**. This customized template takes into consideration the technical features of the 96-well round-bottom plate, particularly the dimensions of each well (diameter, depth and thickness). The probe must reach the bottom of the well, so set it at the exact centre of wells A1 and H12.
11. After selecting the **fluorophores** of interest from the list proposed by the software, set the **plate setup** following the plate template of Table 2, corrected for the number of wells containing at least one cell.
12. Select 100 µL as the **acquisition volume limit**. Tick the **agitation** option. Set the **acquisition rate** to 1 µL/s max, as lower speed improves the total volume analyzed per well.
13. Add the appropriate cleaning and washing solutions to wells H8-H12. The template in **Table 2** specifically leaves wells H8-H12 empty, as the flow cytometer needs to run a range of wash conditions at the end of the analysis.
NOTE: This step is adapted to the specifics of the flow cytometer used.

14. In the plots and gates section, first set the single-cell gate, using the FSC-A/SSC-A scatterplot then the FSC-H/FSC-A scatterplot. Create a **histogram** for each marker of interest.
15. Once the settings are confirmed, proceed to the **Analysis** section. Analyze the single staining first, **recording** no less than 5,000 events (optimal range: 5,000-15,000 events), both for the compensation beads and the CD34- stained fractions. Adjust the voltages if required.
16. Once the single stainings are all recorded, it is possible to start the actual acquisition, clicking on the **Acquisition** function.

Table 3: Antibody mastermix for a flow cytometry experiment, specifically for the identification of HSPCs from human cord blood. [Please click here to download this Table.](#)

6. Post-culture flow cytometry data analysis

NOTE: The data analysis described is specific for the software mentioned in the **Table of Materials**. The main output is the generation of a spreadsheet containing information for surface marker intensity, number of divisions, and kinship per each analyzed cell. Included in this part of the protocol is a script written in R, necessary to this workflow to generate the final analysis spreadsheet.

1. **Export** files from the flow cytometer as .fcs files. Upload them to the analysis software, grouping them together as "single staining", "bulk", and "single cell".
2. Prepare a compensation matrix using the single staining files, and apply it to the other two groups by **drag and drop**.

NOTE: If an automatic compensation tool is used, check the quality by hand before proceeding further.

3. To have a representative gating, concatenate the different bulk wells in a single file. This step rapidly highlights if two colors are overlapping (typically CV and VC) or other anomalies, and therefore need to be excluded. After clicking on the **concatenate populations** option, select all **uncompensated parameters** from the "parameters" menu, then click on **concatenate**.
4. Upload the concatenated file to the workspace, then apply the compensation matrix *via drag and drop*.
5. Prepare the gating strategy defined in **Figure 2** using the concatenated file. In the single-cell gate, display the events on a scatter plot with CFSE and CTV. Create a first gate called Labeled, including all four colors and excluding possible auto-fluorescence (**Figure 2C**). Then, gate each color individually.
6. Cells labeled with CV and VC need a transformed value, considering that the color is the result of the CFSE and CTV signals. The two coordinated signals are therefore rotated on a logarithmic scale by 45°, to allow the division dilution to proceed in parallel to the x-axis. This transformed value is derived manually, clicking on Tools and then **Derive Parameter**. Paste the following formula into the **formula** box:

$$\frac{\sqrt{2}}{2} * (\log (< Param name = "Comp - FL03 - A" />) + \log (< Param name = "Comp - FL17 - A" />))$$

NOTE: The equation²⁶ assumes that CFSE and CTV are parameters 03 and 17.
7. To correctly visualize this novel parameter named Derived Parameter, set a linear axis ranging ~3-7, clicking on the **Axis** parameter option and selecting **Customize Axis**.
8. Apply the gating to each color individually as a histogram plot: for CF and VI, set **CFSE-A** and **CTV-A** on the x-axis, respectively. For CV and VC, set the newly **Derived Parameter** on the x-axis. Set gates corresponding to each peak, as displayed in **Figure 3**.
9. Apply the gating to each individual single-cell well. Make sure to add the derived parameter to every analyzed well. Manually verify each color gate for each well, to detect events that are incorrectly assigned to a given peak. Examples of gating are displayed in **Figure 4**.
10. After the analysis is completed, and all wells have been verified, select all the CF/CV/VC/VI gates that contain at least one cell. **Export** them as .csv files, ticking the options "**scale values**" and "**compensated parameters**". These files are exported in a folder named "**Exported files**".
11. Combine all the files in a single .csv, using the R script in **Supplemental File 1**. Remember to set the right path with the function "setwd". The output of this script is a spreadsheet containing all the different gated events and the relative intensities for all the parameters.
12. Open the spreadsheet and rename the columns for each parameter, for example, using the following names: CFSE, CTV, CD90, CD123, CD45RA, CD34, CD38. These names will be used to identify the gating threshold to correctly assign to each cell their identity.
13. Add six columns named "Well", "Condition", "Color", "Generation", "Original_cell", and "Culture_time". These variables are the ones defined experimentally and are inferred from each row:

`export_A10_CD34 + PBS_CV_Peak 1.csv.1 = A10` (well), `CD34+` (Original_cell), `PBS` (Condition), `CV` (Color), `Peak_1` (Generation).

- Export the bulk wells to identify the threshold values for gating: export the compensated population of interest (e.g., `CD34+CD38-`) as .csv files, ticking the options "**scale values**" and "**compensated parameters**." Export these files in a folder named "**Exported files**".
- To find the threshold for `CD38`, identify the greatest numerical value for this parameter. Conversely, to find the threshold for `CD34`, identify the smallest numerical value for this parameter. Repeat this process for all the parameters of interest.

NOTE: For the analysis presented in the protocol, the marker `CD45RA` is used both to identify LMPPs in the `CD34+CD38-` gate and CMP/GMP in the `CD34+CD38+` gate. This means that two different threshold values need to be extracted for this marker.

- Copy and paste the threshold values in an excel file called "gating_matrix". This file is organized according to **Table 4**, and allows the analysis of multiple independent experiments. It is very important to name each column exactly with this scheme: `XXYYMMDD_xxh`, where `XX` stands for the two initials of the operator, `YY` the last two numbers for the year, `MM` for the month, `DD` for the day, and `xx` for the analysis time point.

Table 4: Gating matrix for the cell fate assignment, before the statistical analysis. The `CD45h` refers to the intensity of `CD45RA` for the HPC subset gating, while `CD45l` refers to the intensity of `CD45RA` for the `CD34+CD38-` subsets. [Please click here to download this Table.](#)

7. Statistical analysis

NOTE: The statistical testing of the data generated involves a custom-made analysis pipeline, coded using the programming language Python (**Supplemental File 2**, **Supplemental File 3**, and **Supplemental File 4**). The script is organized in three blocks: the first one for processing the spreadsheet, the second block to generate the heatmap for data visualization, and the last block to generate multiple histograms to analyze and test differentiation and division properties.

- Starting from the block "`0_process_data`" (**Supplemental File 2**), ensure that the `gating_matrix` and the data spreadsheet paths are correctly defined in the script.
- Define the "cell_cols" dictionary, to assign the relevant cell fates to each cell. In the specific case, the fates are HSCs, MPPs, LMPPs, common myeloid progenitors (CMPs), granulo-monocytic progenitors (GMPs), megakaryocytic-erythroid progenitors (MEPs), and `CD34-`.
- Using the threshold values defined from the bulk wells (step 6.16), define the "cell_class_exp_time" function. It is essential to be consistent in the column naming, to correctly define these thresholds, using the same name used to define each column in step 6.12.
- The cell phenotypes are defined in the script using a series of "if-else" statements, based on the thresholds detected during the flow cytometry analysis. **NOTE:** Different phenotypes can be displayed by modifying these statements to accommodate other marker combinations.

5. Specify the experimental specific conditions, using the function "cond_rule" (e.g., different experimental treatments). For the dataset provided, the conditions are named "GT" and "Diff". Describe the two different cell culture media used to culture the cells. This information will be used by the block "1_dot_plot" (**Supplemental File 3**) for plotting the heatmap.
6. In the block "2_bar_plot" (**Supplemental File 4**), define the dictionary "class_dct", including the discrete cell fates of interest. For the dataset provided, the cell fates of interest are the same described for the "cell_cols" dictionary.
7. Define "conds" (conditions), "or_cells" (original cell), "sym_labs" (symmetry labels), and "times" (the experimental time point). These are reiterating filters necessary for plotting. "conds" take the conditions defined in "cond_rule" again, "or_cells" are HSCs and MPPs, and "sym_labs" describe the type of divisions.
8. In the block "2_bar_plot", it is possible to plot cells that have progressed up to division 6.
NOTE: The dataset provided only includes cells up to division 4, so an error message pops up, but this does not prevent the script from working.
9. The figures generated by the script can be retrieved in the folder named "figures" as pdf files. The files named "Test" represent the different statistical tests performed for the corresponding histogram.

Representative Results

FACS sorting

The sorting gating strategies presented in this protocol are based on widely accepted strategies^{12,30,31}. For the gating strategy presented in **Figure 1**, the starting material is cord blood progenitors previously purified *via* CD34⁺ magnetic enrichment, which explains the negligible percentage of Lineage positive cells. It is essential to use tight gates for the four intracellular dye combinations (e.g., the CTV in the figure), to improve the peaks' resolution during the following analysis and to gate the correct cell population (**Figure 1D**). In the case displayed in the figure, the gates select for the largest and better defined population. The presence of multiple, close populations for each cell division dye combination is, in our experience, not representative of biological differences. Instead, it could indicate a) a non-optimal staining procedure, or b) a large heterogeneity (especially in size) in the starting pool of cells. This is not unexpected when starting from cord blood or other complex biological sources (e.g., bone marrow aspirates, peripheral blood). If the gate is not tightly defined, the progressive dilution of the different dye combinations can lead to merging of the later peaks, specifically for the conditions CV and VC (**Figure 2D**). Another negative consequence of a suboptimal gating is the inability to efficiently distinguish different peaks after cell culture, as a heterogeneous starting population can lead to shallow peaks.

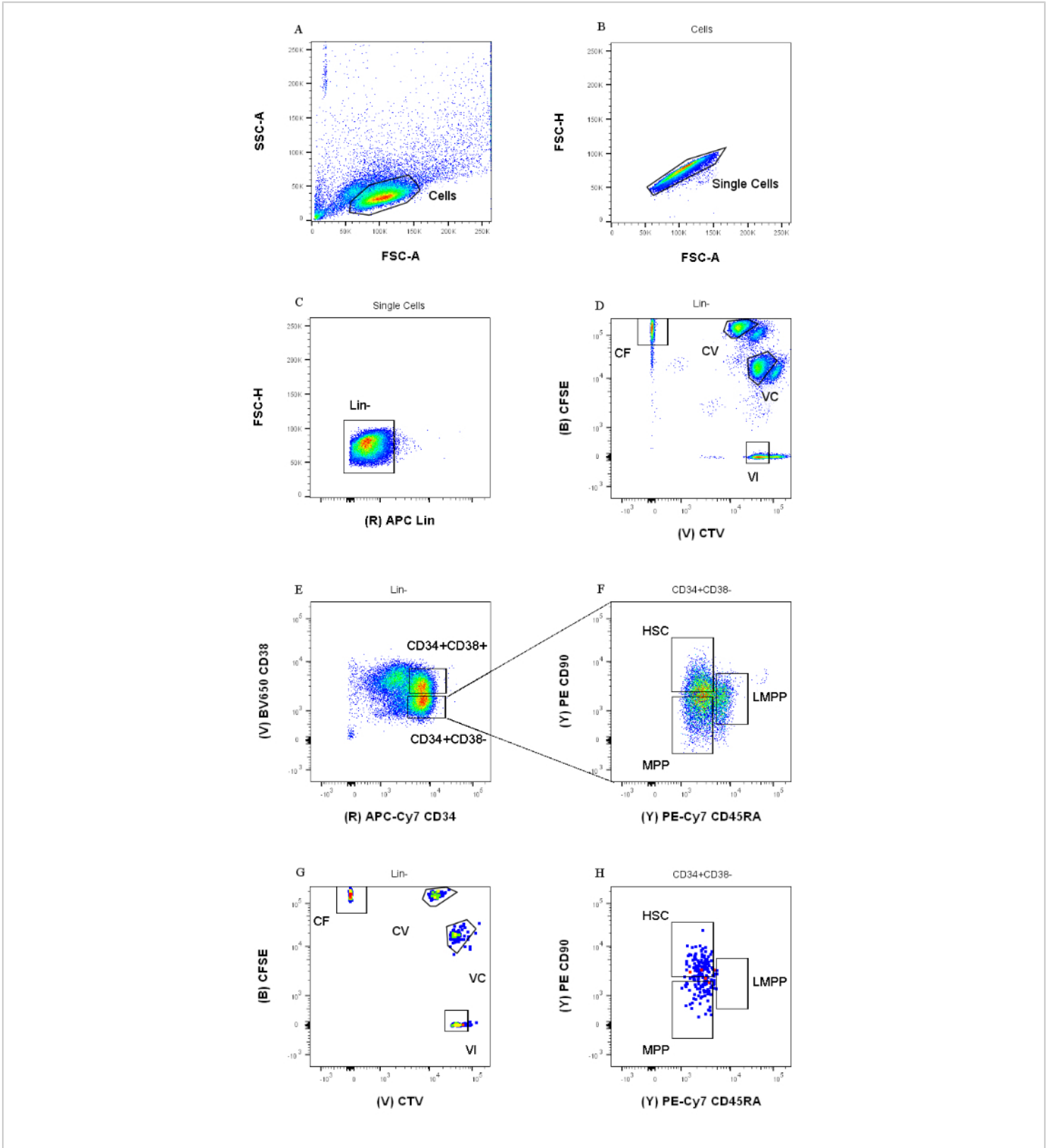


Figure 1: Gating strategy for cell sorting. (A) FSC-A versus SSC-A, to exclude debris and contaminating cells. (B) FSC-A versus FSC-H, to exclude doublets and cell clumps. (C) Lin versus FSC-H, to exclude cells that are Lin+. (D) CTV

versus CFSE, to univocally identify the cells stained with the dye combinations CF, CV, VC, and VI. The gates should be strict enough to include a homogeneous population. (E) CD34 versus CD38, to separate the restricted progenitors $CD34^+CD38^+$ (also called HPCs) from the multipotent compartment $CD34^+CD38^-$. (F) CD45RA versus CD90, from the $CD34^+CD38^-$ population, to separate between the most immature progenitors enriched in the HSC ($CD90^+CD45RA^-$), the LMPP ($CD90^{mid}CD45RA^+$), and the more committed MPP ($CD90^-CD45RA^-$). (G) Index sorted events, represented here for their cell dye combination staining and (H) the expression of surface markers CD90 and CD45RA. [Please click here to view a larger version of this figure.](#)

Flow cytometry analysis after cell culture

The data in **Figure 2** is representative of human cord blood HSCs, kept in culture for 72 h, in the presence of multiple cytokines able to support a range of myeloid progenitors and precursors. Panels **2A** to **2D** represent the gating necessary to establish the kinship of each of the individual cells, while panels **2E** to **2G** allow cellular phenotyping. The reduced presence of MEPs in the figure is probably the consequence of the culture conditions used for this representative experiment (**Figure 2F**). Using different cytokines and culture conditions alter the relative percentage of each subset, similarly to selecting different starting cells for the experiment.

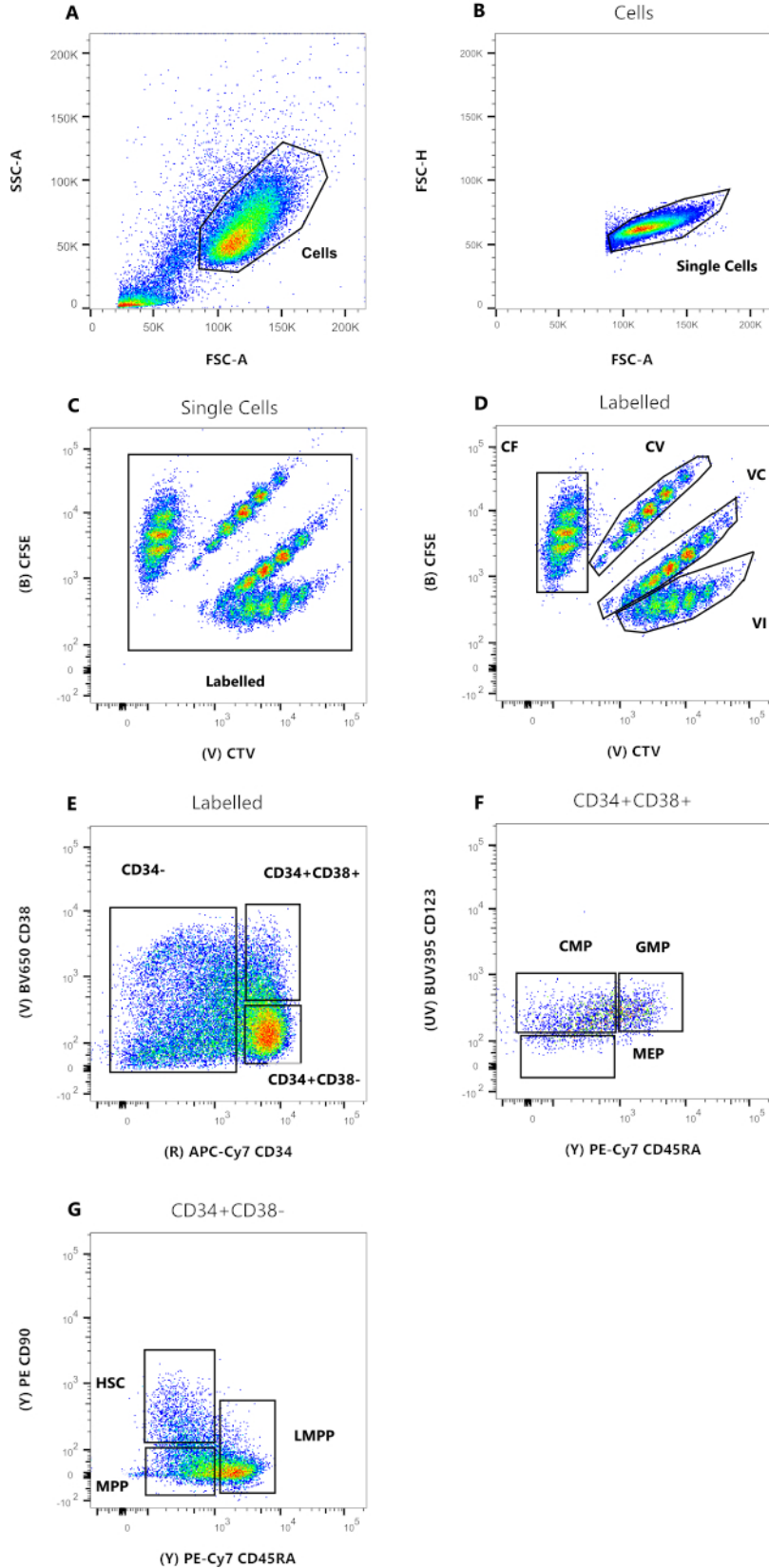


Figure 2: Gating strategy for the flow cytometry analysis. (A) FSC-A versus SSC-A, to exclude debris and contaminating cells. (B) FSC-A versus FSC-H, to exclude doublets and cell clumps. (C) CTV versus CFSE, the gate Labeled allows to exclude any auto-fluorescent event that could affect the data resolution. (D) CTV versus CFSE. It is extremely important to stringently gate the four populations, based on the cell division dye dilutions. (E) CD34 versus CD38, to distinguish between committed precursors ($CD34^-$), restricted progenitors (HPC) ($CD34^+CD38^+$), and immature progenitors ($CD34^+CD38^-$). (F) CD45RA versus CD123, to distinguish three types of restricted progenitors: CMP ($CD123^+CD45RA^-$), MEP ($CD123^-CD45RA^-$), and GMP ($CD123^+CD45RA^+$). (G) CD45RA versus CD90, from the $CD34^+CD38^-$, to identify HSCs, LMPPs, and MPPs. [Please click here to view a larger version of this figure.](#)

The peak definition and assignment steps (**Figure 3** and **Figure 4**) are crucial aspects of the protocol and require the definition of strict gates. For the peak definition (**Figure 3**), at least 1,000 events are necessary for a reliable identification. In this sense, it could be beneficial to isolate more cells during the cell sorting step for the "Bulk" wells. **Figure 4** describes four examples of single wells containing multiple families. This figure clarifies the importance of **Figure 2D** and **Figure 3** gating, especially for the identification of each family and each peak. **Figure 4A** illustrates a straightforward example, as all the cells in the gate CF are very close to each other and can be easily assigned to a single peak. **Figure 4C** shows

another example of a family distributed univocally on two well separated peaks, as it is clearly displayed in the histogram of **Figure 4D**. **Figure 4E,G** reveal the importance of strict gating based on a large number of events; they both display few events that are close, but outside the dye combination gates. Those events could be incorrectly included in the VI and CF gates, based exclusively on the single-well analysis. Finally, **Figure 4F,H** display two different examples of families spread on multiple peaks, with one example of two similar intensity peaks (**Figure 4F**) and one with two unequal intensity peaks (**Figure 4H**).

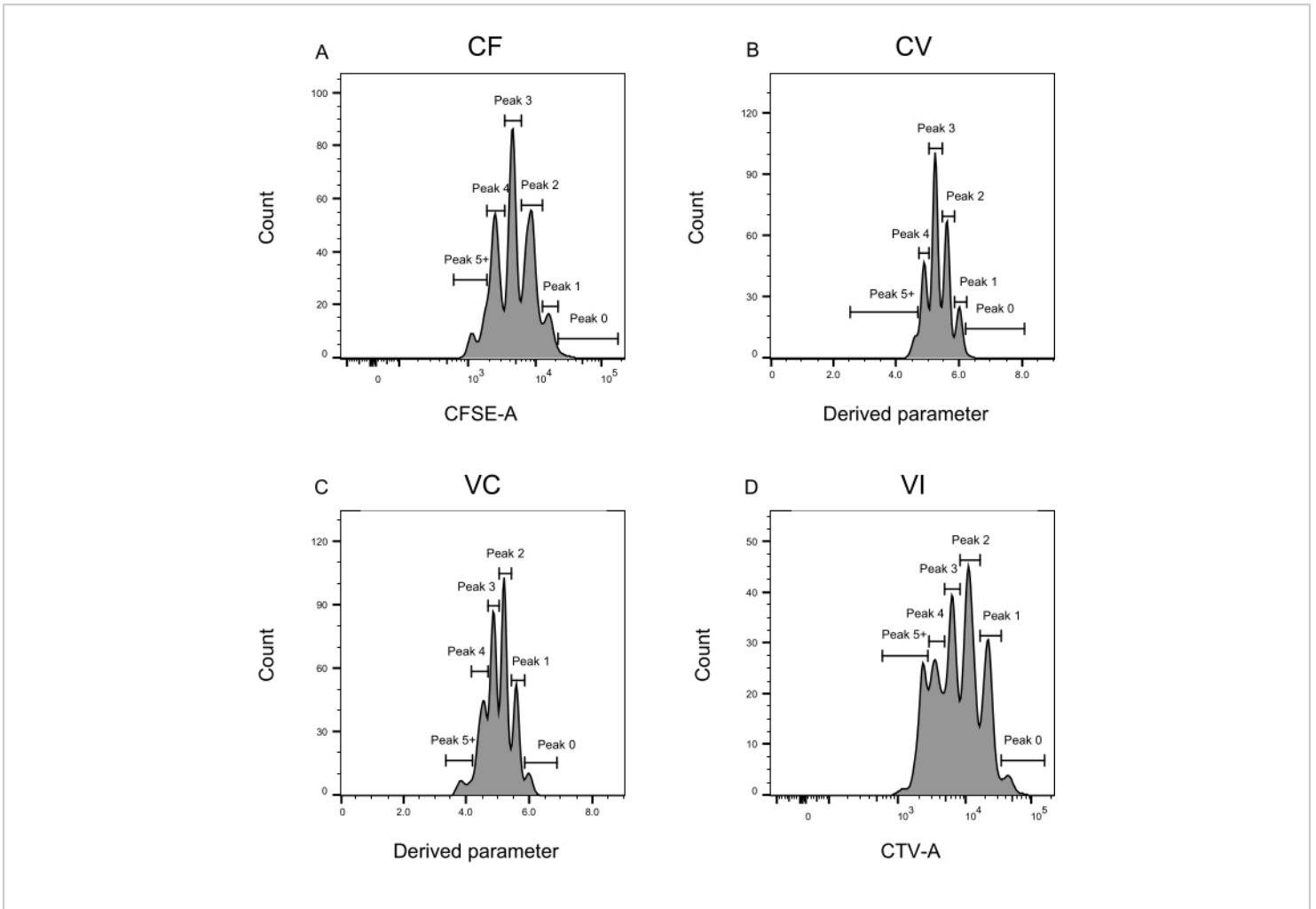


Figure 3: Peak definition for the flow cytometry analysis. (A-D) Peaks should be defined registering at least 500 events, to assure a good representation for every individual peak. **(A)** Histogram for the CFSE-A intensity. Several peaks can be identified, each one corresponding to a different population of dividing cells. **(B,C)** Histograms for the intensity of the derived parameter, representing the CFSE-CTV mixture, CV **(B)** and VC **(C)**, respectively. **(D)** Histogram for the CTV-A intensity.

[Please click here to view a larger version of this figure.](#)

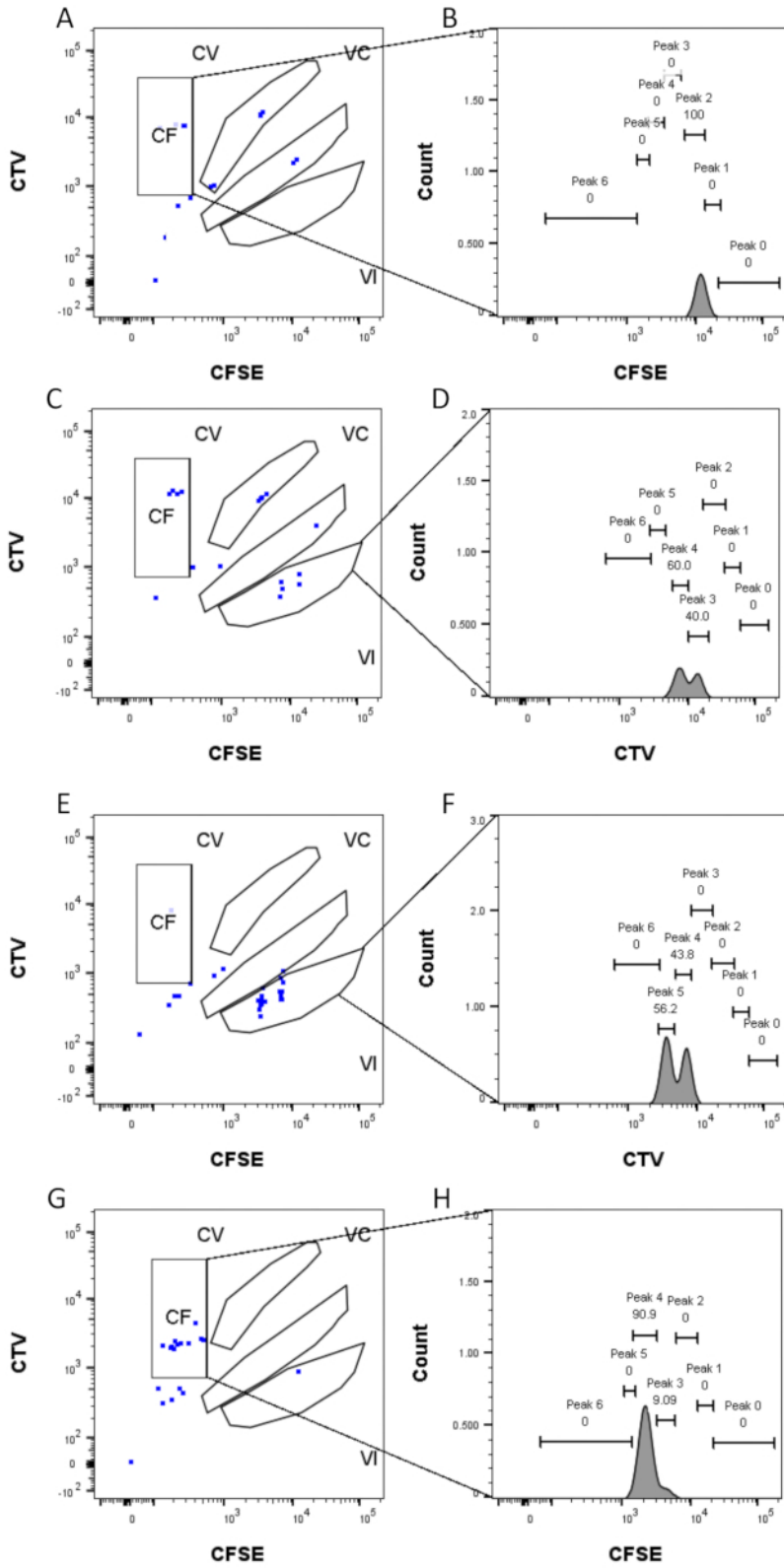


Figure 4: Peak assignment. (A,B) One peak only can be detected for this well, in the CF gate. (C,D) Two peaks of almost equal intensity can be detected in this well, in the VI gate. The peaks are well-resolved. (E,F) Two peaks of comparable intensity can be detected in this well, in the VI gate. Only the events in the gate have been considered, based on the strategy set using the bulk wells. (G-H) Two peaks of unequal intensity can be detected in this well, in the gate CF. [Please click here to view a larger version of this figure.](#)

Data representation and statistical testing

Figure 5 shows different types of data representation of two separate experiments, both performed after 72 h of cell culture. HSCs and MPPs have been cultured in two different cell culture media, supposed to alter the cell division and differentiation properties. These media are named "Diff" (Differentiation)³² and "GT"³³; the first promote myeloid and erythroid differentiation, as it contains erythropoietin (EPO) and granulo-monocytic colony stimulating factor (GM-CSF), while the second has been developed in the context of gene therapy clinical trials, with the goal to maintain and amplify a high percentage of HSPCs. **Figure 5A** is a representative heatmap for the condition "Diff", representing a variety of cell families, both in cell fates and divisions. In this heatmap, each row represents an individual family, each square an individual cell, and the columns are grouping all cells that are in the same generation (e.g., cells in generation 2 divided at least twice). It is possible to distinguish highly homogeneous families, composed of a single cell type and displaying the same number of divisions (e.g., family #63), and heterogeneous families, including three cell types over two generations (e.g., family #84). Because the cellular recovery rate for this analysis is approximately 70%, complete families, which are defined by having all their cells recovered in possibly different generations (e.g., a family of one cell in generation 1 and two cells in generation 2), are rarely observed (displaying a hashtag next to their ID number in **Figure 5A**). There are multiple explanations accounting for the incomplete detection, which could be

technical (staining issue, cell loss due to the protocol) or biological (cell death and/or apoptosis). Technical limitations can be overcome using an analyzer designed to reduce the dead volume associated with the individual sample, and by performing the cell staining directly in the cell culture plate to reduce the volume pipetting. Conversely, orthogonal methods to determine the amount of cell death (e.g., *via* live-cell imaging experiments) can help to distinguish the technical and biological factors resulting in incomplete detection.

Figure 5Bi shows how to visualize the effect of the culture condition on the cell type composition, as if one had performed a bulk assay. Here, the Diff condition promotes a larger number of fates, and a higher percentage of CD34⁺ cells (defined as all the cell types except CD34⁻). The confidence intervals are calculated in the script *via* basic bootstrap, with 250,000 bootstrapped datasets³⁴. It is worth noting that all the other histograms in **Figure 5** display confidence intervals calculated in the same way. **Table 5** recapitulates all the information about the number of families and number of cells in each generation.

Figure 5Bii graphically represents the output of the statistical testing performed in the script "2_bar_plot". A formal description of the statistical framework is available²⁶. In brief, this framework enables statistical hypothesis testing while assuming that cells from the same family are dependent (an assumption that is itself testable), contrary to classical statistics that would require independence between all observed cells. In the specific case presented in the figure,

the statistical test challenges the hypothesis that the cell fate choices of MPPs, measured as the frequencies of the different cell types present in the culture, are independent of the cell culture conditions used. First, the G-test statistic is used to evaluate the discrepancy between cell type frequencies from different cell media (for the example in **Bii**, this statistic is represented with the red bar). Then, a randomization of the data is performed *via* permutation, swapping entire families of cells between the two cell culture conditions. This is to preserve the dependence between family-related cells, while keeping the number of families in each set consistent with the original data. The G-test statistic is calculated from the randomized data set. The blue values represented in **5Bii** are the G-test statistic for 250,000 permutations. Finally, the *p* value is computed to assess the extent to which the original dataset deviates from the distribution of the permuted ones. In the example, the original statistic largely deviates from the distribution, resulting in a small *p* value and thus rejecting the hypothesis that the cell fate of MPPs is independent of the culture conditions.

Figure 5C represents the percentage of cell families per maximum generation, to explore how different conditions change cell division per cell family. This data plot shows that at 72 h, the cells cultured in the Diff condition complete a larger number of divisions than cells in the GT condition. Represented is the number of maximum generations per each family, so one family that displays cells in generations 1 and 2 is considered as generation 2. The same statistical framework used for **Figure 5B** can be used to statistically

test the independence between cellular division and culture condition.

Figure 5D explores the type of symmetry/asymmetry of the first division for the different ancestor types (HSCs or MPPs). For the complete cell families in generation 1-the only generation where it is possible to definitively establish the two daughter cells as sister cells-four different types of symmetry/asymmetry can be defined: the label "Sym Undiff" describes families where both daughters retain the phenotype of the cell of origin. "Sym Diff" means that both daughters have the same phenotype, and it is different from the cell of origin. "Asym Undiff" means that one daughter only retains the phenotype of the cell of origin. Finally, "Asym Diff" describes families where both daughters have different phenotypes, and none of them are the same as the cell of origin. To gain statistical power in assessing these symmetric/asymmetric fates, performing the MultiGen analysis at early time points is desirable, in order to observe more families whose offspring are found in generation 1.

Finally, **Figure 5E** represents the percentages of cell types as a function of the number of divisions, to gain insights about the differentiation pattern progression across divisions. For example, the data displayed in the figure suggest that cells progress to the CD34⁻ state, with over 50% of detected cells in this class after only three divisions. Moreover, it is possible to infer that MPPs do not favor self-renewal division, as a small percentage of cells retain the original phenotype. Some of these conclusions can then be tested using the statistical framework presented in the previous figures.

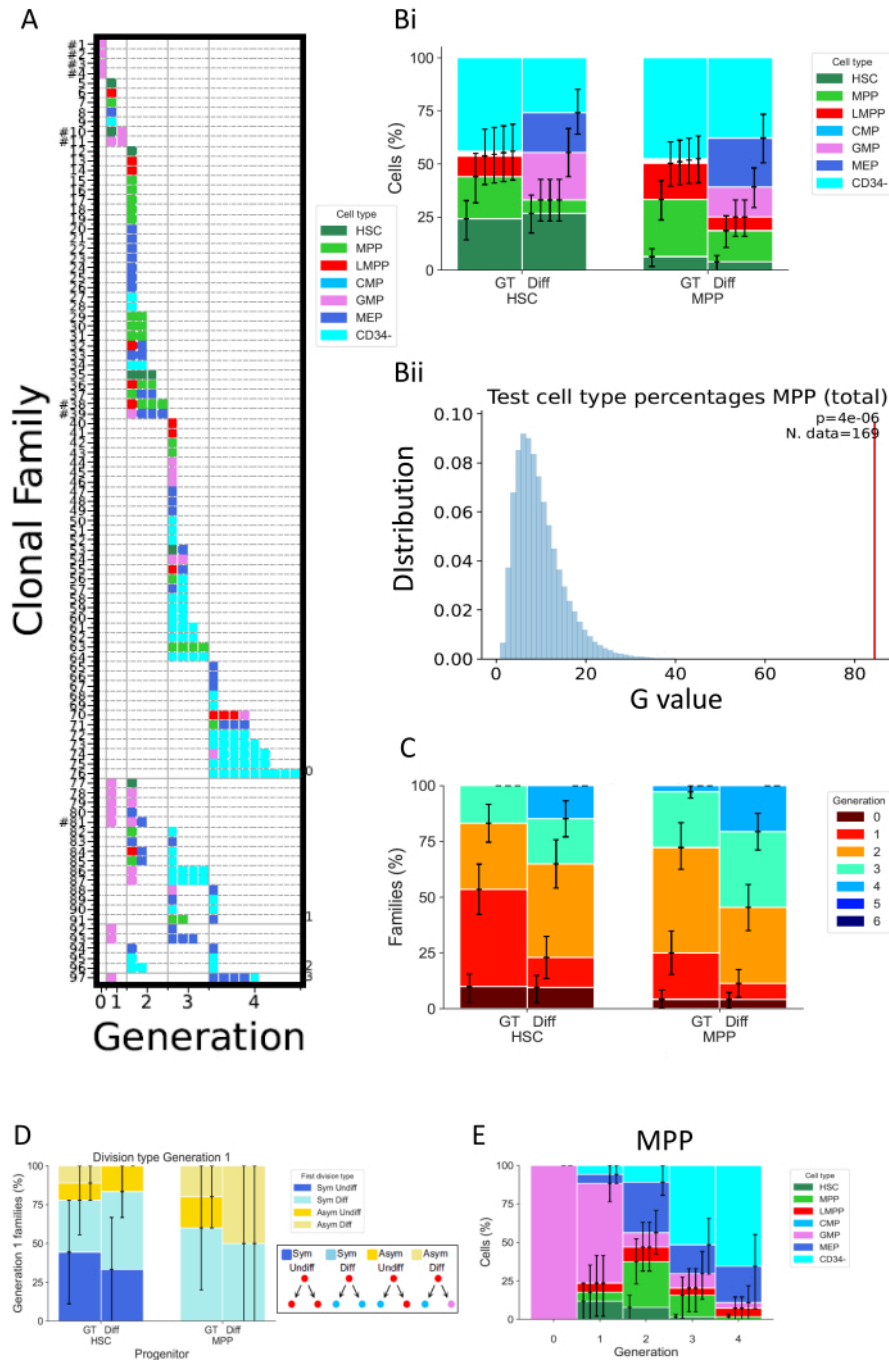


Figure 5: Example of data representation for one 72 h experiment using cord blood HSPCs. (A) Heatmaps for a selected dataset (HSC, in "Diff" medium, after 72 h of culture). The plots represent all individual cells (squares) according to their kinship (rows), number of divisions performed (columns, called generation), and phenotype (colors). **(Bi)** Histogram comparing proportions of the cell types of the cell progenies of HSCs and MPPs, between the condition GT and the condition Diff. **(Bii)** The graph represents the statistical tests performed in the "2_bar_plot" script for MPPs at 72h of culture, comparing

between the "Diff" and the "GT" cytokine cocktails. The experimental value is displayed in red, and the values generated *via* 250,000 permutations in blue. The *p*. value of the G-test is indicated in the top right corner with the number of families used for the test. **(C)** Histogram comparing the percentage of families (314 families in total) in each generation (color coded), for HSCs and MPPs per culture condition. The confidence intervals are calculated with 250,000 bootstrapped datasets. **(D)** Histogram representing the type of symmetry/asymmetry between the fate of the daughter cells for the families with two cells in generation 1: Sym Undiff (both daughters retain the phenotype of the cell of origin), Sym Diff (both daughters have the same phenotype, and it is different from the cell of origin), Asym Undiff (only one daughter retains the phenotype of the cell of origin), and Asym Diff (both daughters have different phenotypes and none of them resemble the cell of origin). **(E)** Histograms of the contribution of cell types classified per generation for MPPs cultured with the "Diff" cocktail; *n* = 204 cells and 97 families. [Please click here to view a larger version of this figure.](#)

Table 5: Description of the number of families and cells analyzed per each experimental condition (cell of origin and cell culture medium). [Please click here to download this Table.](#)

Supplemental File 1: [Please click here to download this File.](#)

Supplemental File 2: [Please click here to download this File.](#)

Supplemental File 3: [Please click here to download this File.](#)

Supplemental File 4: [Please click here to download this File.](#)

Discussion

The MultiGen assay is a high-throughput, easy to perform, and unexpensive assay, that has been instrumental to study lymphocyte^{23,24,35} and murine hematopoietic cells^{26,27}. Here, we present a new development of the approach that allows to decipher *ex vivo* the early phase of human HSPC commitment, at the single-cell level using short term culture (**Figure 6**). Single-cell *ex vivo* culture systems are typically used to assess the long-term fate of HSPCs into mature cells, but some fates occur earlier than others³⁶, potentially biasing the analysis toward fewer fates. In addition, these culture systems usually miss information about divisions

during fate commitment. The first commitment steps have been shown to occur as early as at the start of the culture, sometimes without division^{26,37}, making short term culture and tracking division essential to study early fate commitment. By simultaneously following the fate, division, and kinship, this assay allows to understand the role of the first division and fate decision in human HSPCs. Using the assay, it is possible to infer after how many divisions the commitment process occurs, the balance between self-renewal and differentiation for those early progenitors, and how those properties are inherited across generations. To our knowledge, this is the only assay that allows these types of measurements for human HSPCs, at single-cell resolution. In addition, using different combinations of cell division dyes, we increased the throughput of the analysis, making this assay a valuable tool for generating large datasets quickly. The dye combinations allow to follow several families in the same wells, increasing the number of cells available for analysis in short-term culture. The number of combinations could potentially be increased even more, *via* the addition of other dyes (e.g., yellow dye) or modifying the ratio of CFSE and CTV. However this reduces the number of other parameters that can be analyzed.

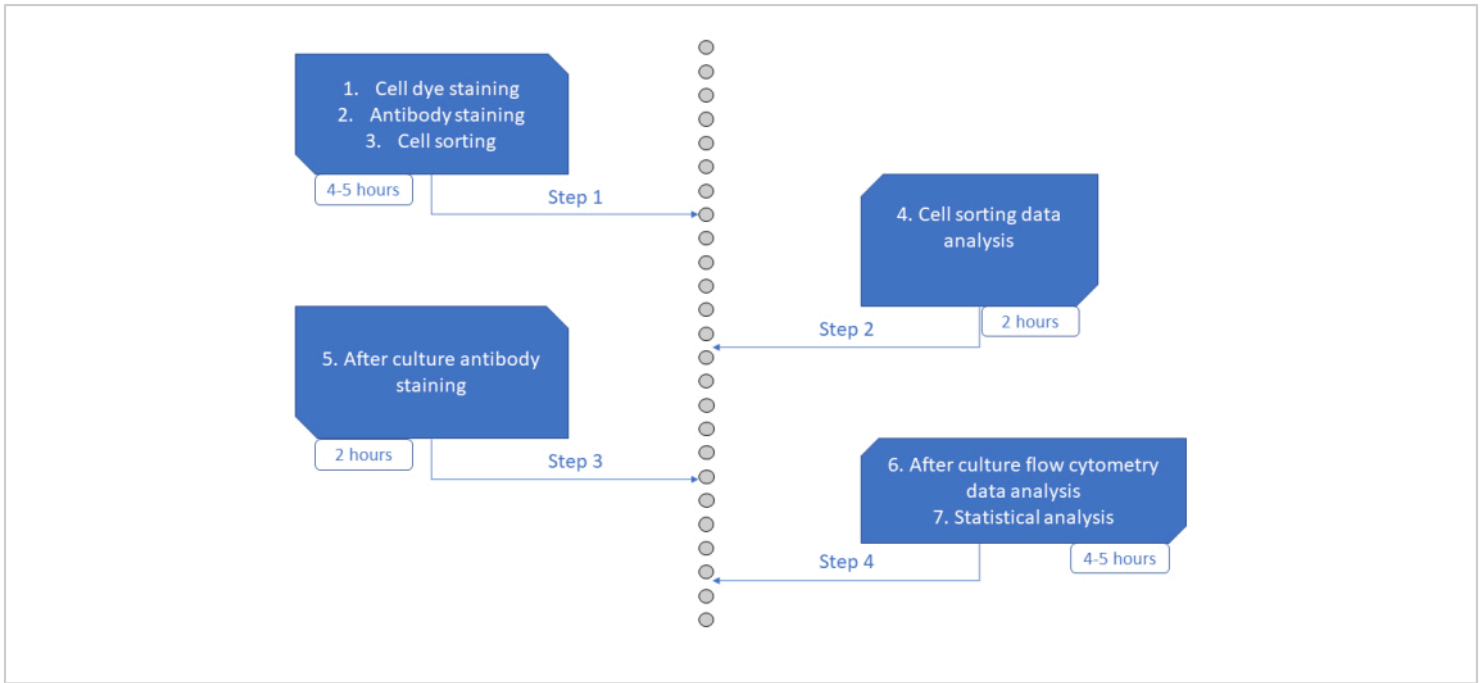


Figure 6: Schematic representation of the protocol. [Please click here to view a larger version of this figure.](#)

To perform the analysis successfully, due to the large number of wells and the reduced number of cells to analyze, it is necessary to run the flow cytometry analysis on an analyzer equipped with a plate reader. The new generation of bench analyzers is particularly adapted to this assay, as most of them have a smaller dead volume to reduce the percentage of cell loss. This in turn guarantees a higher efficiency in recovering the entirety of each well, prompting an efficiency estimated in the 70% range²⁶. Estimating the cell loss during the flow cytometry acquisition is crucial for the analysis of each individual family. For example, assuming no cell death and counting the number of divisions, it is possible to estimate the number of cells per each family. Nevertheless, it is desirable to run some confirmatory experiments, particularly in the estimation of cell death in the tested culture conditions and measuring the recovery rate experimentally using a defined numbers of cells.

One of the crucial steps of this protocol is the peak assignment. As already mentioned, a good quality peak distribution is strongly dependent from the isolation of very narrow peaks at cell sorting. Nevertheless, it is still difficult to assign the correct number of divisions based uniquely on the distribution. As cell sorting and flow cytometry analysis are performed on two different machines, it is not possible to directly compare the intensity of each signal, so it could be difficult to know if the first peak observed on the right end of the histogram is peak 0 or peak 1. In this regard, few solutions are possible; one way is to perform an orthogonal experiment to accurately measure the number of divisions performed by these cells (e.g., live-cell imaging). Another possibility is to simply count the number of cells in the well under an inverted brightfield microscope, before running the flow cytometry analysis. This will infer an average number of divisions (assuming no cell death). Finally, a post-hoc solution for peak assignment is the detection of an unusual number

of "impossible families"; these families are composed of a larger than possible number of cells per generation (e.g., five cells in generation 2, or two cells in generation 1 and one cell in generation 2). The possibility to exclude impossible families is coded in the statistical analysis step, and flags the impossible family. If the occurrence of these errors is too high, it is reasonable to assume that the peak assignment needs to be revised.

In this protocol, we included a few examples of data representation and analysis for the assay, as this has become an essential step in the generation and interpretation of large datasets³⁸. The first example is the heatmap showing the totality of all analyzed cells, organized per family. This is an efficient tool to explore the general properties of the data and potential conclusions: are families composed by multiple cell types or do they tend to be homogeneous in composition? Are families spread over multiple generations, or are they mostly dividing the same number of times? This exploratory analysis needs to then be complemented with more specific plots and statistical testing. It can be used to quantitatively assess symmetric and asymmetric fate commitment, differentiation without division, the balance between self-renewal and differentiation, and the number of divisions for a given commitment fate. It is fundamental, during the experimental planning, to set the cell culture length accordingly to the type of question asked; for example, for the first two questions (symmetric/asymmetric balance and differentiation without division), planning very short cultures steps allows the isolation of a large number of families that have performed only one or no divisions at all²⁶. Conversely, longer experiments allow the exploration of the number of divisions required for a specific cell commitment, as they sample families at different stages of differentiation. Nevertheless, this method is not designed for long-term

cultures (2-3 weeks), as cell dye dilution is not able to accurately track more than seven or eight divisions²². As a consequence, this tool is mostly adapted to study the early commitment of hematopoietic progenitors, and is not designed to make robust conclusions on the long term differentiation properties of these cells.

The statistical framework was developed specifically for the analysis of this type of data and based on the concept of permutations²⁶. This was necessary because of the observation of a familial dependence on the cell type distribution and on the number of divisions performed. In other words, cells that are part of the same family are also more likely to display similar phenotypes and divide the same number of times. While an in-depth analysis is beyond the scope of this work, the provided set of statistical tests should be sufficient when assessing different conditions.

In conclusion, this protocol constitutes a valuable tool for assessing the cellular dynamics of hematopoietic stem and progenitor cells *ex vivo*, in a rapid and inexpensive way. Due to its flexibility and versatility regarding time point, culture conditions, and type of HSPCs analyzed, it allows to test a variety of experimental conditions. As a flow cytometry-based assay, it can be implemented in most labs, and it does not require extensive prior knowledge, making it a good candidate for screenings and pilot experiments.

Disclosures

The authors declare no conflict of interest relevant to this work. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Acknowledgments

We would like to thank the members of the Institut Curie Flow Facility for their help with setting up the flow cytometry experiments. We also want to acknowledge the contributions of the other members of the Team Perié, during multiple discussions. We thank Dr. Julia Marchingo and Prof. Phil Hodgkin (Walter and Eliza Hall Institute of Medical Research) for sharing their protocol of the multiplexing of cell division dyes on lymphocytes. We thank the Saint Louis hospital cord blood biobank for providing the biological resources necessary for the development of this protocol. The study was supported by an ATIP-Avenir grant from CNRS and Bettencourt-Schueller Foundation (to L.P.), grants from the *Labex CelTisPhyBio* (ANR-10-LBX-0038) (to L.P. and A.D.), Idex Paris-Science-Lettres Program (ANR-10-IDEX-0001-02 PSL) (to L.P.), the Canceropole INCA Emergence (2021-1-EMERG-54b-ICR-1, to L.P.), and the ITMO MIIC grant (21CM044, to L.P.). As well as funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program ERC StG 758170-Microbar (to L.P.), A.D. was supported by a fellowship from Fondation de France.

References

- Ginhoux, F., Yalin, A., Dutertre, C. A., Amit, I. Single-cell immunology: Past, present, and future. *Immunity*. **55** (3), 393-404 (2022).
- Ke, M., Elshenawy, B., Sheldon, H., Arora, A., Buffa, F. M. Single cell RNA-sequencing: A powerful yet still challenging technology to study cellular heterogeneity. *Bioessays*. **44** (11), e2200084 (2022).
- Regev, A. et al. The human cell atlas. *Elife*. **6**, e27041 (2017).
- Laurenti, E., Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. **553** (7689), 418-426 (2018).
- Haas, S., Trumpp, A., Milsom, M. D. Causes and consequences of hematopoietic stem cell heterogeneity. *Cell Stem Cell*. **22** (5), 627-638 (2018).
- Loughran, S. J., Haas, S., Wilkinson, A. C., Klein, A. M., Brand, M. Lineage commitment of hematopoietic stem cells and progenitors: insights from recent single cell and lineage tracing technologies. *Experimental Hematology*. **88**, 1-6 (2020).
- Perié, L., Duffy, K. R. Retracing the in vivo haematopoietic tree using single-cell methods. *FEBS Letters*. **590** (22), 4068-4083 (2016).
- Yu, V. W. C. et al. Epigenetic memory underlies cell-autonomous heterogeneous behavior of hematopoietic stem cells. *Cell*. **167** (5), 1310-1322.e17 (2016).
- Ganuza, M. et al. Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. *Nature Cell Biology*. **19** (10), 1153-1163 (2017).
- Naik, S. H., Schumacher, T. N., Perié, L. Cellular barcoding: A technical appraisal. *Experimental Hematology*. **42** (8), 598-608 (2014).
- Quek, L. et al. Genetically distinct leukemic stem cells in human CD34 – acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. *The Journal of Experimental Medicine*. **213** (8), 1513-1535 (2016).
- Karamitros, D. et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nature Immunology*. **19** (1), 85-97 (2018).

13. Boitano, A. E. et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science*. **329** (5997), 1345-1348 (2010).
14. Delaney, C. et al. Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution. *Nature Medicine*. **16** (2), 232-236 (2010).
15. Fares, I. et al. Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science*. **345** (6203), 1509-1512 (2014).
16. Guo, B., Huang, X., Lee, M. R., Lee, S. A., Broxmeyer, H. E. Antagonism of PPAR- γ 3 signaling expands human hematopoietic stem and progenitor cells by enhancing glycolysis. *Nature Medicine*. **24** (3), 360-367 (2018).
17. Vannini, N. et al. The NAD-booster nicotinamide riboside potently stimulates hematopoiesis through increased mitochondrial clearance. *Cell Stem Cell*. **24** (3), 405-418.e7 (2019).
18. Gupta, R. et al. Nov/CCN3 enhances cord blood engraftment by rapidly recruiting latent human stem cell activity. *Cell Stem Cell*. **26** (4), 527-541.e8 (2020).
19. Horwitz, M. E. et al. Omidubicel vs standard myeloablative umbilical cord blood transplantation: results of a phase 3 randomized study. *Blood*. **138** (16), 1429-1440 (2021).
20. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D., Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science*. **367** (6479), eaaw3381 (2020).
21. Loeffler, D., Schroeder, T. Understanding cell fate control by continuous single-cell quantification. *Blood*. **133** (13), 1406-1414 (2019).
22. Tario, J. D. et al. Optimized staining and proliferation modeling methods for cell division monitoring using cell tracking dyes. *Journal of Visualized Experiments*. (70), e4287 (2012).
23. Marchingo, J. M. et al. T-cell stimuli independently sum to regulate an inherited clonal division fate. *Nature Communications*. **7**, 13540 (2016).
24. Horton, M. B. et al. Multiplexed division tracking dyes for proliferation-based clonal lineage tracing. *Journal of Immunology*. **201** (3), 1097-1103 (2018).
25. Lehmann, E. L., Romano, J. P., Casella, G. *Testing statistical hypotheses*. 784. Springer, New York. (2005).
26. Tak, T. et al. HSPCs display within-family homogeneity in differentiation and proliferation despite population heterogeneity. *Elife*. **10**, 360624 (2021).
27. Sommerkamp, P. et al. Mouse multipotent progenitor 5 cells are located at the interphase between hematopoietic stem and progenitor cells. *Blood*. **137** (23), 3218-3224 (2021).
28. Kato, K., Radbruch, A. Isolation and characterization of CD34+ hematopoietic stem cells from human peripheral blood by high-gradient magnetic cell sorting. *Cytometry*. **14** (4), 384-392 (1993).
29. Miltenyi, S., Müller, W., Weichel, W., Radbruch, A. High gradient magnetic cell separation with MACS. *Cytometry*. **11** (2), 231-238 (1990).
30. Doulatov, S. et al. Revised map of the human progenitor hierarchy shows the origin of macrophages

- and dendritic cells in early lymphoid development. *Nature Immunology*. **11** (7), 585-593 (2010).
31. Goardon, N. et al. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell*. **19** (1), 138-152 (2011).
 32. Laurenti, E. et al. CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell*. **16** (3), 302-313 (2015).
 33. Aiuti, A. et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science*. **341** (6148), 1233151 (2013).
 34. Davison, A. C., Hinkley, D. V. *Bootstrap Methods and their Application*. Cambridge University Press. (1997).
 35. Horton, M. B. et al. Lineage tracing reveals B cell antibody class switching is stochastic, cell-autonomous, and tuneable. *Immunity*. **55** (10), 1843-1855.e6 (2022).
 36. Notta, F. et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. **351** (6269), aab2116 (2016).
 37. Grinenko, T. et al. Hematopoietic stem cells can differentiate into restricted myeloid progenitors before cell division in mice. *Nature Communications*. **9** (1), 1898 (2018).
 38. Saeys, Y., Van Gassen, S., Lambrecht, B. N. Computational flow cytometry: Helping to make sense of high-dimensional immunology data. *Nature Reviews Immunology*. **16** (7), 449-462 (2016).