

High Dimensional Functionomic Analysis of Human Hematopoietic Stem and Progenitor Cells at a Single Cell Level

Thomas Luh¹, Kimberly Lucero¹, Wenji Ma², Jaeyop Lee¹,
Yu Jerry Zhou¹, Yufeng Shen² and Kang Liu^{1, *}

¹Department of Microbiology and Immunology, Columbia University Medical Center, New York, New York, USA; ²Department of Systems Biology and Department of Biomedical Informatics, Columbia University Medical Center, New York, New York, USA

*For correspondence: kl2529@cumc.columbia.edu

[Abstract] The ability to conduct investigation of cellular transcription, signaling, and function at the single-cell level has opened opportunities to examine heterogeneous populations at unprecedented resolutions. Although methods have been developed to evaluate high-dimensional transcriptomic and proteomic data (relating to cellular mRNA and protein), there has not been a method to evaluate corresponding high-dimensional functionomic data (relating to cellular functions) from single cells. Here, we present a protocol to quantitatively measure the differentiation potentials of single human hematopoietic stem and progenitor cells, and then cluster the cells according to these measurements. High dimensional functionomic analysis of cell potential allows cell function to be linked to molecular mechanisms within the same progenitor population.

Keywords: CD34⁺ hematopoietic stem and progenitor cells, Single cell culture, Differentiation, Quantitative clonal output, Barnes-Hut t-SNE, Dimension reduction

[Background] The development of techniques for single-cell measurements of cell transcription, signaling, and function at the single-cell level, alongside preexisting technologies such as flow cytometry, has allowed new lenses to examine complex, heterogeneous populations. Such methods generate large amounts of data, which can be interpreted with aid from dimensionality reduction algorithms, as illustrated on single-cell RNA-Seq using Mpath, Monocle, PCA, Wishbone, or diffusion map algorithms (Paul *et al.*, 2016; See *et al.*, 2017), and on CyTOF using tSNE or PhenoGraph (Amir el *et al.*, 2013; Levine *et al.*, 2015).

We developed this protocol to allow functional analysis and subsequent dimension reduction of large-scale culture of hematopoietic progenitors in single-cell environments. In this protocol, we describe a method to culture single cells of human CD34⁺ hematopoietic stem and progenitor cells (HSPCs) in a stromal cell culture with cytokines, to enumerate the clonal outcome (functionomics) of six different lineages (*i.e.*, granulocyte, monocyte, lymphocyte, CD141⁺ dendritic cell (conventional type 1 DC (cDC1)), CD1c⁺ dendritic cell (conventional type 2 DC (cDC2)), and plasmacytoid dendritic cells (pDC)) from each progenitor, and to cluster the progenitors according to function with a dimension reduction method. In our previous paper, we showed that this is feasible for a population of 2,247 progenitors; each progenitor can be individually plotted to form a two-dimensional ‘map’ of 2,247 data points (Lee *et*

al., 2017). Our protocol allows for the functional clustering of single cells. Such high dimensional functionomic analysis aids in linking cell function to molecular mechanism of any given cell population.

Materials and Reagents

1. 15 cm culture plates (Corning, Falcon®, catalog number: 353025)
2. 15 ml Falcon tubes (Corning, Falcon®, catalog number: 352097)
3. 96-well V-bottom plates (Thermo Fisher Scientific, Thermo Scientific™, catalog number: 249570)
4. Flat-bottom 96-well plates (Corning, Falcon®, catalog number: 353072)
5. 50 ml Falcon tubes (Corning, Falcon®, catalog number: 352098)
6. Serological pipettes (Fisher Scientific, catalog number: 13-678-11)
7. Opaque Eppendorf tubes (CELLTREAT Scientific Products, catalog number: 229437)
8. Eppendorf tubes (Fisher Scientific, catalog number: 05-408-129)
9. Pipette tips
10. Pasteur pipette (Fisher Scientific, catalog number: 13-678-8B)
11. 100 µm cell strainer (Corning, Falcon®, catalog number: 352360)
12. 0.20 µm filter (Corning, catalog number: 431229)
13. MS5 (Lee *et al.*, 2015a and 2015b)
14. OP9 (Lee *et al.*, 2015a and 2017)
15. dH₂O (Thermo Fisher Scientific, Invitrogen™, catalog number: 10977023)
16. Trypsin (Corning, catalog number: 25-052-CI)
17. Trypan blue (Thermo Fisher Scientific, Gibco™, catalog number: 15250061)
18. 70% ethanol
19. Ficoll-Paque Plus (GE Healthcare, catalog number: 17144002)
20. Anti-human CD34 MACS microbeads (Miltenyi Biotec, catalog number: 130-046-702)
21. Recombinant human Flt3L (Celldex Therapeutics, catalog number: CDX-301)
22. Recombinant human SCF (PeproTech, catalog number: 300-07)
23. Recombinant human GM-CSF (PeproTech, catalog number: 300-03)
24. Antibodies (details of clone name, fluorochrome, manufacturer and dilution are listed in Tables 1 and 3)

Antibodies	Antigen	Company	Clone
AlexaFluor (AF)700	CD45	Biolegend	HI30
Q2dot-655	CD14	Invitrogen	Tuk4
BrilliantViolet (BV) 650	CD19	Biolegend	HIB19
BrilliantViolet (BV) 650	CD3	Biolegend	OKT
BrilliantViolet (BV) 650	CD56	Biolegend	HCD56
BrilliantViolet (BV) 650	CD16	Biolegend	3G8
APC-Cy7	CD34	Biolegend	581
PE	CD10	Biolegend	HI10a
PE	CLEC9a	Biolegend	8F9
FITC	CD45RA	Biolegend	HI100
BV421	CD38	Biolegend	HIT2
BV510	CD132	Biolegend	6H6
Live/Dead [®]		Life Technologies	
PerCP-Cy5.5	CD66b	Biolegend	G10F5
PE-Cy7	CD1c	Biolegend	L161
FITC	CD303	Biolegend	201A
APC-Cy7	CD56	Biolegend	HCD56
APC-Cy7	CD19	Biolegend	HIB19
APC	CD141	Miltenyi	AD5-14H12

25. MEMα with deoxyribonucleosides (Thermo Fisher Scientific, Gibco™, catalog number: 12571071)
26. Fetal bovine serum (FBS) (Thermo Fisher Scientific, Gibco™, catalog number: 10437028)
27. Pen-Strep (Thermo Fisher Scientific, Gibco™, catalog number: 15140122)
28. Mitomycin C (Sigma-Aldrich, catalog number: M4287-5X2MG)
29. Dulbecco's phosphate-buffered saline (DPBS) (GE Healthcare, catalog number: SH30378.02)
30. Ethylenediaminetetraacetic acid (EDTA) (Corning, catalog number: 46-034-CI)
31. Bovine serum albumin (BSA) (ThermoFisher, catalog number: BP1600-100)
32. FcR blocking buffer (Miltenyi Biotec, catalog number: 130-059-901)
33. CountBright Absolute Counting Beads (Thermo Fisher Scientific, Invitrogen™, catalog number: C36950)
34. Complete MEMα medium (see Recipes)
35. Mitomycin C stock (1 mg/ml) (see Recipes)
36. FACS buffer (see Recipes)
37. Anti-CD34 microbeads/FcR blocking mix (see Recipes)

Equipment

1. Pipettes
2. Incubator (Forma Scientific, model: 3354)
3. Centrifuge (Eppendorf, model: 5810 R, catalog number: 5811000320)

4. Vortex with universal holder (VWR, catalog number: 97043-562)
5. Microscope
6. Flow cytometer (BD, models: BD LSRII or LSRFortessa™)

Procedure

A. Culture of stromal cells (day -4 – day 0)

1. Passaging stromal cells (maximum of 13 times)
 - a. Grow MS5 and OP9 cells in complete MEM α medium. For 15 cm plates, seed 1.5 million cells in 20 ml medium.
 - b. For normal passage, microscopically verify that the cells are 80-90% confluent.
 - c. Remove medium from the culture plate by aspiration and wash cells with 10 ml sterile PBS to remove medium and FBS.
 - d. Remove the PBS, and then add 3 ml trypsin pre-warmed to 37 °C.
 - e. Incubate for 2-3 min at 37 °C, and microscopically verify that cells have started to detach.
 - f. Add 10 ml complete MEM α medium pre-warmed to 37 °C to stop trypsin activity.
 - g. Resuspend cells by pipetting up and down, and then transfer cells to 15 ml collecting tubes.
 - h. Centrifuge at 450 \times g (1,500 rpm) for 5 min at 4 °C.
 - i. Remove supernatant, and then resuspend the cells in 1 ml complete MEM α medium (Cell Concentrate).
 - j. In a new 15 cm plate, add 20 ml complete MEM α medium. Transfer 100 μ l cell concentrate to the new plate. This new plate will become confluent after 3-4 days.
 - k. Incubate at 37 °C, 5% CO₂ for 3-4 days, or until cells are 80-90% confluent.
 - l. Passage the cells again, to a maximum of 13 passages.
 - m. For MS5 and OP9 cells, cell yield from a 15 cm plate with 80-90% confluency is 15-20 million cells.
2. Treating stromal cells with Mitomycin-C
 - a. Verify that MS5/OP9 cells are 80-90% confluent.
 - b. Label collecting tubes ('MS5 + mitC', 'OP9 + mitC').
 - c. Add Mitomycin C stock (1 mg/ml, 100x) directly into the medium to achieve a final concentration of 10 μ g/ml. For example, aspirate medium and add back exactly 10 ml, and add 100 μ l Mitomycin C stock.
 - d. Mix by swirling gently for 3-5 sec.
 - e. Incubate at 37 °C for 3 h.
3. Plating stromal cells
 - a. Remove medium, and then add 10 ml PBS. Gently swirl the plate for 3-5 sec.
 - b. Remove the PBS, and then add 3 ml trypsin.
 - c. Incubate for 2-3 min at 37 °C, and use a microscope to verify that cells have started to detach.

- d. Add 10 ml complete MEM α medium to stop trypsin.
- e. Resuspend cells by pipetting up and down, and then transfer cells to labeled harvesting tubes.
- f. Centrifuge at 450 $\times g$ for 5 min at 4 °C.
- g. Remove medium, and then resuspend in 1 ml complete MEM α medium.
- h. Count cells, and calculate the number and volume of cells needed for the final plating. Each well in a flat-bottom 96-well plate, should contain 6,250 OP9 cells and 37,500 MS5 cells in 50 μ l complete MEM α medium. To prepare for one 96-well plate, resuspend 0.625×10^6 OP9 cells and 3.75×10^6 MS5 cells in 5 ml complete MEM α medium, mix well, and then distribute 50 μ l into each well.
- i. Incubate cells at 37 °C overnight.
- j. Verify that cells are securely attached to plate by shaking gently while viewing under a light microscope. With secure attachment, the cells should not drift around with the fluid movement.
- k. Within 1-3 days, sort and plate progenitors on these cells.

B. Purification of CD34⁺ hematopoietic stem and progenitor cells from human cord blood (day 0)

1. Isolation of cord blood mononuclear cells using Ficoll-Paque

Note: This step should be handled at room temperature.

- a. Add 25 ml PBS into multiple 50 ml tubes.
- b. Sterilize the outside of the cord blood sample container with ethanol. Samples usually contain about 100 ml of blood.
- c. In 15 ml portions, evenly distribute the cord blood sample into the prepared 50 ml Falcon tubes containing PBS.
- d. Rinse the blood sample container with some additional room temperature PBS, and use the rinse to 'top off' the last Falcon tube (such that each tube has a fluid volume of 40 ml). On average, use 7 Falcon tubes per bag of cord blood.
- e. Tightly seal each of the tubes, and gently invert them 3-5 times to resuspend any sedimented cells.
- f. Underlay the cell suspension with 10.5 ml of Ficoll-Paque with a 10 ml Serological pipette, in the following manner:
 - i. Load 10.5 ml into the pipette.
 - ii. Without releasing any fluid, gently bring the pipette to the bottom of the tube.
 - iii. For the first 2 ml, very slowly release Ficoll-Paque into the bottom of the tube. The goal is to have a pure layer of Ficoll-Paque at the bottom of the tube.
 - iv. For the next 8 ml, the Ficoll-Paque release can be slightly sped up, but use caution to avoid turbulence that may disrupt the Ficoll-Paque-blood boundary.
 - v. For the last 0.5 ml, slow down the Ficoll-Paque release, and observe the meniscus in the pipette. Use caution to prevent introducing air bubbles into Ficoll-Paque.

- g. Centrifuge at $450 \times g$ (1,500 rpm) for 25 min at room temperature. Use lower acceleration (level 3 of 9) and no brakes for deceleration. Orient the opaque sections of the test tube face towards the central motor of the centrifuge, so that cell deposits will collect on a transparent section of the test tube, away from the motor, and so can be easily visualized.
 - h. After centrifugation, there should be four layers:
 - i. The topmost layer (yellow/clear, ~35 ml) is a PBS/serum mix.
 - ii. The second layer from top (cloudy, ~1 ml) is a leukocyte-containing buffy coat.
 - iii. The third layer from top (clear, ~10 ml) contains Ficoll-Paque.
 - iv. The bottommost layer (red, ~5 ml) contains RBCs.
 - v. Additionally, at the buffy coat layer, a smeared pellet of cells (cloudy, ~1 mm) should be deposited along the test tube wall.
 - i. Transfer the collected cells into buffy collection tube:
 - i. Prepare one 15 ml buffy collection tube for each 50 ml tube from the previous steps. Add 2 ml PBS in each tube.
 - ii. Remove the PBS/serum layer until only 10 ml of this layer remains, using caution to avoid turbulence by aspirating from the top of the layer. (*i.e.*, if the buffy coat line is at the 30 ml line, then retain 40 ml total fluid remains.)
 - iii. With a 10 ml Serological pipette, transfer the buffy coat layer to a buffy collection tube, and use caution to minimize fluid transfer. To do so, hold the pipette tip directly above the buffy coat, and slowly aspirate, then eject fluid to the buffy collection tube. Leave ~5 ml of PBS/serum for the next step.
 - iv. Resuspend the smeared pellet of cells in the Ficoll-Paque by gently rocking the tube, using caution to avoid mixing the fluid layers. The rocking motion allows the liquid-liquid interface between the Ficoll-Paque and the PBS/serum mix to rub the pellet off the wall.
 - v. Transfer as much Ficoll-Paque layer to the buffy collection tube as possible, while completely avoiding the RBC layer. Avoiding the PBS/serum layer is ideal, but not critical.
 - j. Centrifuge at $450 \times g$ (1,500 rpm) for 10 min at 4 °C.
 - k. Remove supernatant, and resuspend cells in 1 ml FACS buffer.
 - l. Collect cell suspensions from all collection tubes into a single 50 ml Falcon tube.
 - m. Wash the cells by adding FACS buffer until the total fluid volume is 50 ml.
 - n. Count the cells. Each bag of cord blood yields ~300-1,000 million cells.
 - o. Centrifuge at $450 \times g$ for 5 min at 4 °C.
 - p. Remove supernatant, deriving a mononuclear cell pellet.
2. Enrichment of CD34⁺ cells
- a. Prepare anti-CD34 microbeads/FcR blocking mix.
 - b. In a 15 ml conical tube, dilute the cells in 2 μ l anti-CD34 microbeads/FcR blocking mix per 1 million nucleated cells.
 - c. Incubate at 4 °C for 30-40 min.

- d. Add FACS buffer up to the 15 ml line.
- e. Centrifuge at $450 \times g$ (1,500 rpm) for 5 min at 4 °C.
- f. Remove supernatant.
- g. Resuspend cells in FACS buffer (1 ml per 100 million cord blood mononuclear cells).
- h. Prepare an elution tube: 15 ml Falcon tube + 2 ml FACS buffer.
- i. Perform positive selection for CD34⁺ cells with an LS MACS column:
 - i. Place a 100 µm cell strainer on top of the column.
 - ii. Equilibrate the column by filtering 5 ml FACS buffer through the strainer; discard the flow-through.
 - iii. Gently vortex the cells to ensure suspension.
 - iv. Filter the cells through the LS MACS column; retain the flow-through.
 - v. Optional: wash the column once by passing 4 ml FACS buffer through it; discard this wash. Then, re-load the flow-through from the previous step on the column to increase cell yield.
 - vi. Wash the column by passing 4 ml FACS buffer through it; discard the wash. Repeat the wash a second time, and discard the wash again.
 - vii. Remove the MACS column from its magnetic holder, and place it on top of the elution tube.
 - viii. Add 4 ml of FACS buffer to the column, and allow gravity elution to occur.
 - ix. Add 4 ml of FACS buffer, taking caution to prevent the column from drying out, and use a plunger to push the FACS buffer through.
- j. Count the cells. Each bag of cord blood yields 300-1,000 $\times 10^6$ CD34⁺ cells.
- k. Centrifuge at $450 \times g$ for 5 min at 4 °C.
- l. Remove supernatant.
- m. (Optional) If necessary, cells can be stored overnight on ice in 4 °C, or frozen in 10% DMSO + FBS media; however, proceeding with fresh cells without an intermediate freezing step is recommended to avoid affecting clonal output and cell viability.

C. Culture of purified cells in single-cell environment (day 0-7)

1. Single-cell sorting and plating
 - a. Stain CD34⁺ HSPCs with antibody mix. Use 10 µl antibody mix for 1×10^6 cells according to the following dilution (Table 1):

Table 1. Antibody preparation for purification of progenitors from human cord blood

Fluorochrome	Antigen	Company	Clone	Dilution factor	Volume per 1×10^6 cells (μ l)
AlexaFluor (AF) 700	CD45	Biolegend	HI30	200	0.05
Qdot-655	CD14	Invitrogen	Tuk4	800	0.01
Brilliant Violet (BV) 650	CD19	Biolegend	HIB19	200	0.05
BV650	CD3	Biolegend	OKT	200	0.05
BV650	CD56	Biolegend	HCD56	200	0.05
BV650	CD16	Biolegend	3G8	200	0.05
APC-Cy7	CD34	Biolegend	581	200	0.05
PE	CD10	Biolegend	HI10a	100	0.10
FITC	CD45RA	Biolegend	HI100	80	0.13
BV421	CD38	Biolegend	HIT2	80	0.13
BV510	CD123	Biolegend	6H6	80	0.13
Additional PBS					9.21
Final volume					10.00

- b. Incubate cells on ice for 40 min, and using flow cytometry, sort cells according to following gating (Table 2, Figure 1):

Table 2. Characterization of progenitor populations in human cord blood

Progenitor	Markers
Hematopoietic stem cell (HSC)	CD34+CD38-CD45RA-CD90+
Multipotent progenitor (MPP)	CD34+CD38-CD45RA-CD90-
Lymphoid-primed multipotent progenitor (LMPP)	CD34+CD38-CD45RA+CD10-
Multilymphoid progenitor (MLP)	CD34+CD38-CD45RA+CD10+
B and NK cell progenitor (BNKP)	CD34+CD38+CD45RA+CD123int/-CD115-CD10+
Common myeloid progenitor (CMP)	CD34+CD38+CD45RA-CD10-CD123int
Granulocyte/Monocyte/DC progenitor (GMDP)	CD34+CD38+CD45RA+CD10-CD123int
Monocyte/DC progenitor (MDP)	CD34+CD38+CD45RA+CD123intCD115+
Common DC progenitor (CDP)	CD34+CD38+CD45RA+CD123hiCD115-
Megaerythrokaryocyte progenitor (MEP)	CD34+CD38+CD45RA-CD10-CD123-

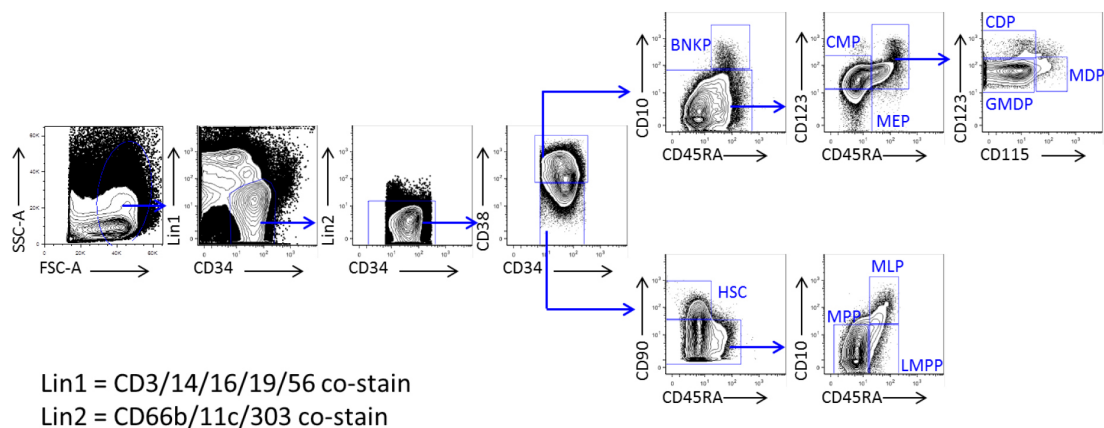


Figure 1. Visual gating strategy of progenitor populations in human cord blood. Flow cytometry plots show separation of CD34⁺ cord blood cells into ten populations, HSC, MPP, LMPP, MLP, BNKP, CMP, MEP, GMDP, MDP and CDP.

- c. Sort single cells into each well of stromal-cell-containing 96-well plates with 50 µl complete MEMα medium per well.
2. Cell culture over time
 - a. Prepare a 2x mixture of cytokines in complete medium:
 - i. 100 ng/ml Flt3L (final 1x concentration 50 ng/ml).
 - ii. 20 ng/ml SCF (final 1x concentration 10 ng/ml).
 - iii. 10 ng/ml GM-CSF (final 1x concentration 5 ng/ml) (not required to culture cells; stimulates and inhibits proliferation of various cell subtypes).
 - b. Add the 2x mixture of cytokines to seeded stromal cell wells. (Add 50 µl for 96-well plate)
 - c. Replenish media with cytokine every 7 days.
 - i. On day 7, for each well, prepare 50 µl medium containing cytokines at 3x concentration. Gently add 50 µl fresh media with 3x cytokines into the well by pipetting onto the wall. *Note: A higher concentration of cytokines is prepared because the final media volume will be increased, while the volume of media-cytokines mixture remains the same. Media is not removed from the wells yet, due to the possibility of disturbing cell attachment.*
 - ii. On day 14, prepare 50 µl medium containing cytokine at 3x concentration as in Step C2c i. Gently remove 50 µl of the old medium from each well, and then add 50 µl fresh media into the well by pipetting onto the wall.
- D. Harvest and analyze cells (day 7~21)
 1. Read cells via flow cytometry
 - a. Prepare antibody mixture in FACS buffer using the following dilution (Table 3).

Table 3. Antibody preparation for analysis of progeny from human cord blood

Fluorochrome	Antigen	Company	Clone	Dilution factor	Volume per sample (24-well plate) (µl)	Volume per sample (96-well plate) (µl)
Live/Dead ®		Life Technologies		1000	0.01	0.01
AlexaFluor (AF) 700	CD45	Biologend	HI30	200	0.05	0.02
Qdot-655	CD14	Invitrogen	Tuk4	800	0.01	0.01
PE	CLEC9a	Biologend	8 F9	500	0.02	0.01
PerCP-Cy5.5	CD66b	Biologend	G10F5	100	0.10	0.04
PE-Cy7	CD1c	Biologend	L161	100	0.10	0.04
FITC	CD303	Biologend	201A	100	0.10	0.04
APC-Cy7	CD56	Biologend	HCD56	100	0.10	0.04
APC-Cy7	CD19	Biologend	H1B19	80	0.13	0.05
BV510	CD123	Biologend	6H6	80	0.13	0.05
APC	CD141	Miltenyi	ADS-14H12	80	0.13	0.05
Additional PBS					9.13	3.65
Final volume					10.00	4.00

- b. Add PBS such that each well is filled, using a multichannel pipet. (250 µl on a 96-well plate)
- c. Pipet up and down 10-20 times to resuspend cells, and transfer 150 µl to 96-well V-bottom plates for staining.
- d. Centrifuge at 450 x g for 5 min at 4 °C.
- e. Discard the supernatant by microtip aspiration, using caution to keep the fluid level above the slanted portion of the well.

- f. Transfer the remaining 100 μ l of cells to the 96-well V-bottom plates.
- g. Centrifuge at 450 $\times g$ for 4 min at 4 $^{\circ}$ C.
- h. Discard the supernatant, as before.
- i. Gently vortex the plate to resuspend the cells, using caution to avoid any fluid spillage. To do so, reduce the vortex speed to minimum, grasp the plate firmly, and firmly place it on the vortex mixer. Slowly increase the vortex speed, and gently reduce pressure on the vortex mixer so that the vortex mixer can shake. Then, decrease the vortex speed and remove the plate. As needed, this process may be repeated in each of the plate's four quadrants. It is advised to practice this technique with a test plate and tap water to ensure no sample loss. (An alternative resuspension method is to gently pipet 10-20 times.)
- j. Add 4 μ l of antibodies.
- k. Gently vortex the plate to mix cells with the antibody, as before.
- l. Centrifuge with a quick spin (50 $\times g$, 5 sec) to bring down the cells and staining mix to the bottom of the well.
- m. Incubate in the dark for 40 min at 4 $^{\circ}$ C.
- n. Add 200 μ l PBS to wash.
- o. Centrifuge at 450 $\times g$ for 5 min at 4 $^{\circ}$ C. Discard the supernatant.
- p. Resuspend in 50 μ l FACS buffer.
- q. Vortex the counting beads stock solution for 30 sec, and make a 1:10 dilution of counting beads in FACS buffer.
- r. Add 10 μ l of dilute counting beads to each well (~1,000 beads per well).
- s. Acquire cells on a flow cytometer (BD-LSRII or LSR Fortessa), and gate cells according to the following gating (Table 4) (Figure 2 and Lee *et al.*, 2017).

Table 4. Characterization of progeny populations in human cord blood

Progeny	Markers
Granulocyte (G)	CD45+CD66b+
CD1c DC (C)	CD45+CD66b-CLEC9a+CD141+
CD141 DC (A)	CD45+CD66b-CLEC9a-CD1c+
Monocyte (M)	CD45+CD66b-CLEC9a-CD1c-CD14+
pDC (P)	CD45+CD66b-CLEC9a-CD1c-CD14-CD123+CD303+
Lymphocyte (L)	CD45+CD66b-CLEC9a-CD1c-CD14-CD123-CD19/56+

2. Proceed to data analysis

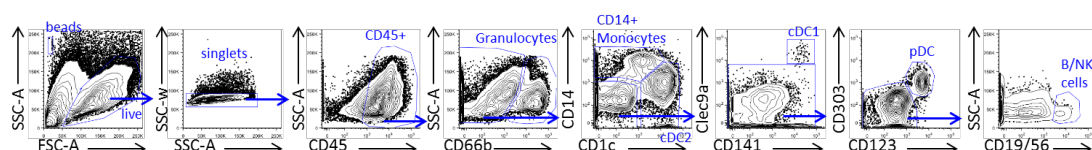


Figure 2. Visual gating strategy for analysis of progeny populations in human cord blood.

Flow cytometry plots show phenotype of six cell types differentiated from CD34⁺ cells cultured

in MP+FSG, granulocytes, monocytes, cDC1, cDC2, pDC and B/NK (lymphoid) cell.

Data analysis

A. Data preparation

1. Export flow cytometry data and save in a .csv file, such that:
 - a. Each data column contains one 'dimension' of data.
 - b. Each row contains one sample's data.
 - c. One column 'Batch' exists (case sensitive); all samples from one patient should have identical batch IDs.
 - d. One column 'Progenitor' exists (case sensitive).
 - e. All data columns are consecutive, each containing the number of relevant reads in each sample-progeny data point. (Figure 3)

Batch	Progenitor	Sample ID	CD45	G	M	C	A	P	L	Junk
B106	HSC	9qsf	3630	1	25	12	0	5	98	25
B106	GMDP	beus	81	0	0	7	0	0	29	44
B106	CDP	9teg	6	0	0	3	3	0	0	57
A176	LMPP	66pt	183	1	6	116	58	2	0	8
A176	CMP	7bqy	998	978	18	0	0	0	0	3
A302	MPP	wvcs	511	7	21	92	345	30	8	83
C152	LMPP	fvh8	226	0	0	9	7	10	185	33

Figure 3. Example of first few rows of a valid .csv file viewed in Excel

B. Data normalization in R

1. Open cPlot.R in RStudio, then Source the file. (Code > Source) (The file is available online at https://github.com/kangliulab/cplot/blob/master/cPlot_v0.7.R.)
2. Change working directory to the location of data files. (Session > Set Working Directory > Choose Directory)

Note: Do not use a working directory in a synced folder (OneDrive, Dropbox, etc.), as this may cause some problems with file outputs.
3. Normalize data using one of the following functions. To use, enter the function into the console, replacing inFile with the file name, firstData with the first data column (leftmost column is #1), and dataCols with the number of data columns. Keyword may be omitted if desired; if used, it adds the keyword to the start of each row to help differentiate between files. When replacing words, keep any quotation marks that are denoted below.
 - a. `addCB ("inFile", firstData, dataCols, "keyword")` is used for samples that must be normalized both by subject (*i.e.*, batch) and by progeny (*i.e.*, lineage).
 - b. `addPatient ("inFile", firstData, dataCols, "keyword")` is used for samples that must be normalized only by subject (*i.e.*, batch).
 - c. `addPreNorm ("inFile", firstData, dataCols, "keyword")` is used for samples that have already been normalized.

- d. Examples:
 - i. `addCB("data-file-v123_for-test-use.csv", 5, 6).`
 - ii. `addCB("data-file-v123_for-test-use.csv", 5, 6, "v123").`
4. Generate t-SNE maps using the below function. To use, enter it into the console. You may choose to omit all numbers, which causes the script to use the default numbers (denoted below). To replace numbers, indicate both the variable to be changed and its new value.
 - a. `generateManyMaps(dataCols = 6, perVector = c(50), thetaVector = c(0.10), iterVector = c(1000), etaVector = c(200), n = 1, seedSet = FALSE, randSeed = 6)` allows multiple parameters to be simultaneously tested. It can also be used to generate multiple maps with a single set of parameters. Note that it will try every possible combination of parameters, and will try each combination *n* times. `dataCols` should be equal to your number of data columns, `perVector` selects perplexity values, `thetaVector` selects theta values, `iterVector` selects iterations to run, `etaVector` selects eta values, `n` selects number of repetitions per combination of parameters, `seedSet` forces each plot to have the same random seed, and `randSeed` selects the random seed of the first plot (or, the random seed of all plots if `seedSet` is TRUE).
 - b. Examples:
 - i. `generateManyMaps().`
 - ii. `generateManyMaps(n = 10, seedSet = TRUE).`
 - iii. `generateManyMaps(perVector = c(40, 50, 60), etaVector = c(150, 200, 250), n = 3).`
5. Assign cell lineages using each of the below functions in sequential order.
 - a. `pickMap("inFile")` allows a particular t-SNE map to be imported for further analysis. `inFile` should represent a data file that was created by `generateManyMaps().`
 - b. `assignLineages()` automatically calculates an appropriate lineage track for each data point. It can only be used after `pickMap()` has been used. It can take a long time (3-8 h) to run, especially for larger data files!
 - c. `compileAllData("inFile")` generates a .csv with raw data, normalized data, tSNE coordinates, and assigned lineages. It assumes that data has already been imported into RStudio. If RStudio was closed sometime between importing the data (in step 3) and this step, then you should re-import the data before running `compileAllData().` `inFile` should represent a data file that was created by `assignLineages().`
 - d. Examples:
 - i. `pickMap("data1_p50_t0.1_s6_i1000.csv").`
 - ii. `assignLineages().`
 - iii. `compileAllData("2assignedLineages_FC0.7.csv").`

Above, we describe an outlined approach to our data. As described in our previous paper, each progenitor's output is imagined as a 6D vector. A preliminary data cleanup step involves neglecting low cell counts: below 2 cells for CDPs and below 6 cells for all other cell types; the neglect occurs

by changing the reading to 0 cells. Progenitors with a total of 0 cell output after data cleanup should then be purged. Data normalization of cell type yield follows the DESeq procedure, assuming the geometric mean of total clonal output for a single progenitor phenotype across different donors should be similar. A second normalization of cell type yield was also done via DESeq, to ensure similar geometric mean of each progeny type yield across all progenitors.

Barnes-Hut t-SNE (Van Der Maaten, 2014) was performed to visualize the inferred pair-wise similarity data in 2 dimensions. Potency similarity between any pair of cells is inferred by their proximity in this space using Gaussian diffusion kernel.

To assign lineages to each progenitor, we begin by defining six 'backbones'; one for each type of progeny output. The backbone of each progeny type consists of any progenitors with > 70% commitment degree to that progeny; commitment degree is defined as (yield of one progeny type)/(total yield of all progeny types). Then, distance of each progenitor to any given track is defined as the 6D Euclidean distance between a progenitor of interest and all backbone progenitors of that track. The progenitor is assigned to the lineage corresponding with the track that it has the lowest distance to.

Notes

Although this culture protocol only reports differentiation potential of granulocyte, monocyte, lymphocyte, CD141⁺ dendritic cell (cDC1), CD1c⁺ dendritic cell (cDC2), and plasmacytoid dendritic cells (pDC) of hematopoietic lineage, the analysis method can be applied to any culture system that nurtures different lineages as far as that the cellular output can be quantified.

Recipes

1. Complete MEM α medium
MEM α
10% FBS
1x Pen/Strep
2. Mitomycin C stock (100x, 1 mg/ml)
 - a. Suspend 2 mg Mitomycin C powder in 2 ml dH₂O
 - b. Vortex solution
 - c. Filter solution with a 0.20 μ m filter
 - d. Keep stock for up to 3 months at -20 °C
3. FACS buffer
DPBS
2 mM EDTA
0.5% BSA
4. Anti-CD34 microbeads/FcR blocking mix

Prepare 2 μ l of the mix per 1 million total nucleated cells
15% FcR blocking buffer
25% anti-CD34 MACS beads
60% FACS buffer

Acknowledgments

This work was supported by the Empire State Stem Cell Fund through the New York State Department of Health (C029562 to K.L.) and The US National Institute of Health (AI101251 and OD023291 to K.L.). Research reported in this manuscript was performed partly by the Columbia Center for Translational Immunology Flow Cytometry Core, supported in part by the Office of the Director of the US National Institutes of Health (S10RR027050 and S10OD020056). This protocol was adapted from previous work Lee *et al.*, 2017, Lee *et al.*, 2015a. Authors declare no conflicts of interest or competing interests. The recombinant human Flt3L was a gift from Celldex Therapeutics.

References

1. Amir el, A. D., Davis, K. L., Tadmor, M. D., Simonds, E. F., Levine, J. H., Bendall, S. C., Shenfeld, D. K., Krishnaswamy, S., Nolan, G. P. and Pe'er, D. (2013). [viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia](#). *Nat Biotechnol* 31(6): 545-552.
2. Levine, J. H., Simonds, E. F., Bendall, S. C., Davis, K. L., Amir el, A. D., Tadmor, M. D., Litvin, O., Fienberg, H. G., Jager, A., Zunder, E. R., Finck, R., Gedman, A. L., Radtke, I., Downing, J. R., Pe'er, D. and Nolan, G. P. (2015). [Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis](#). *Cell* 162(1): 184-197.
3. Lee, J., Zhou, Y. J., Ma, W., Zhang, W., Aljoufi, A., Luh, T., Lucero, K., Liang, D., Thomsen, M., Bhagat, G., Shen, Y. and Liu, K. (2017). [Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors](#). *Nat Immunol* 18(8): 877-888.
4. Lee, J., Breton, G., Aljoufi, A., Zhou, Y. J., Puhr, S., Nussenzweig, M. C. and Liu, K. (2015a). [Clonal analysis of human dendritic cell progenitor using a stromal cell culture](#). *J Immunol Methods* 425: 21-26.
5. Lee, J., Breton, G., Oliveira, T. Y., Zhou, Y. J., Aljoufi, A., Puhr, S., Cameron, M. J., Sekaly, R. P., Nussenzweig, M. C. and Liu, K. (2015b). [Restricted dendritic cell and monocyte progenitors in human cord blood and bone marrow](#). *J Exp Med* 212(3): 385-399.
6. Paul, F., Arkin, Y., Giladi, A., Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., David, E., Cohen, N., Lauridsen, F. K. B., Haas, S., Schlitzer, A., Mildner, A., Ginhoux, F., Jung, S., Trumpp, A., Porse, B. T., Tanay, A. and Amit, I. (2016). [Transcriptional heterogeneity and lineage commitment in myeloid progenitors](#). *Cell* 164(1-2):

325.

7. See, P., Dutertre, C. A., Chen, J., Gunther, P., McGovern, N., Irac, S. E., Gunawan, M., Beyer, M., Handler, K., Duan, K., Sumatoh, H. R. B., Ruffin, N., Jouve, M., Gea-Mallorqui, E., Hennekam, R. C. M., Lim, T., Yip, C. C., Wen, M., Malleret, B., Low, I., Shadan, N. B., Fen, C. F. S., Tay, A., Lum, J., Zolezzi, F., Larbi, A., Poidinger, M., Chan, J. K. Y., Chen, Q., Renia, L., Haniffa, M., Benaroch, P., Schlitzer, A., Schultze, J. L., Newell, E. W. and Ginhoux, F. (2017). [Mapping the human DC lineage through the integration of high-dimensional techniques](#). *Science* 356(6342).
8. Van Der Maaten, L. (2014). [Accelerating t-sne using tree-based algorithms](#). *J Mach Learn Res* 15: 3221-3245.