

SUPPLEMENTARY INFORMATION

DISTINCT ROUTES OF LINEAGE DEVELOPMENT RESHAPE THE HUMAN BLOOD HIERARCHY ACROSS ONTOGENY

Faiyaz Notta^{1,2,8†}, Sasan Zandi^{1,2†}, Naoya Takayama^{1,2}, Stephanie Dobson^{1,2}, Olga Gan¹, Gavin Wilson^{2,4}, Kerstin B Kaufmann^{1,2}, Jessica McLeod¹, Elisa Laurenti⁶, Cyrille F Dunant⁷, John D McPherson^{3,4}, Lincoln D Stein^{2,4}, Yigal Dror⁵, John E Dick^{1,2‡}

¹Princess Margaret Cancer Centre, University Health Network, ²Department of Molecular Genetics, and ³Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, ⁴Ontario Institute for Cancer Research, Toronto, Ontario Canada, ⁵The Hospital for Sick Children Research Institute, University of Toronto, Ontario, Canada, ⁶Wellcome Trust – Medical Research Council Cambridge Stem Cell Institute, Department of Haematology, University of Cambridge, Cambridge, United Kingdom, ⁷Ecole Polytechnique Fédérale de Lausanne, LMC, Station 12, Lausanne, CH-1015, Switzerland, ⁸Current address: Ontario Institute for Cancer Research.

TABLE OF CONTENTS

SUPPLEMENTARY TEXT	3
EXPERIMENTAL PROCEDURES	5
HUMAN CD34+ CORD BLOOD, BONE MARROW, AND FETAL LIVER	5
CELL SORTING	5
METHYLCELLULOSE AND MEGACULT™ ASSAY.....	6
SINGLE CELL ASSAY	6
FLUIDIGM ASSAY AND RNA SEQUENCING.....	7
IN VIVO ANALYSIS.....	8
REFERENCES.....	10

Supplementary Text

Colony assays resolve phenotypic heterogeneity within newly defined subpopulations of human myeloid progenitors

The unexpected cellular heterogeneity that we uncovered prompted us to determine whether these progenitor fractions also exhibited functional heterogeneity. We first applied standard methylcellulose (MC) colony assays to our newly defined cell populations. Median cloning efficiency across all subsets was 38% in FL, 46% in CB and 36% in BM (fig. S3 A-B). CD49f+ HSC derived colonies were exclusively myeloid of which nearly half were of high proliferative potential (CFU-HPP; fig. S3B), reflecting their stem cell potential. The percentage of CFU-HPP gradually decreased in HSC compartments of FL, CB and BM (CFU-HPP: FL - 38%, CB - 31%, BM - 22%; Fig. S4B left). These data indicate that CD49f uniformly enriches HSCs across development, and is consistent with ongoing *in vivo* studies (SZ, NT, FN, JED unpublished). HPP colonies persisted at the MPP stage (CFU-HPP: FL - 6%, CB - 7%, BM - 3%; fig. S4B middle), and surprisingly FL and CB CMP F1 cells still produced rare HPP colonies (CFU-HPP; FL: 4%, CB: 7%, BM: 0%; fig. S4B right). In BM, no subset within CD34+CD38+ population carried HPP potential. In general, FL colonies often generated 10.5 larger colonies compared to BM irrespective of the subset analyzed (fig. S4B). These results complement the work of Lansdorp et al. performed two decades ago with bulk CD34+ cells from FL, CB and BM (1). We extend these findings and establish that proliferative changes in development are not only restricted to the stem cell compartment, but occur at all hierarchical levels.

The spectrum of lineage colonies was consistent across phenotypically matched subsets in FL, CB and BM (fig. S2A), and advocate that segregation of cellular subsets based on functional surface markers is able to stably discriminate comparable progenitors in development. Subsets defined by absence of CD71 and BAH-1 such as HSCs and F1 subsets of MPPs, CMPs and MEPs were highly enriched for CFU-My activity (fig. S3A). We reasoned that the absence of lineage commitment markers in the CMP compartment (CMP F1) would identify true human CMPs. However, greater than 70% of colonies from FL and CB CMP F1 were My only, and the rest were BFU-E (fig. S3A). In BM, CMP F1 cells were exclusively My (CFU-My, FL: 70%, CB: 76%, BM: 99%; fig. S3A). The increased proportion of CFU-My in BM correlates well with increased myelopoiesis observed in aging (2). These data propose that previously defined CMPs consist mostly of unilineage My and Er progenitors. Although resolving Er progenitors with CD71 was expected, we found that BFU-E derived from MPP

F2/3 were often larger than analogous F2/3s defined in the CMPs and MEPs (fig. S4A). This indicates that Er progenitors derived from CD34+CD38⁻ versus CD34+CD38⁺ are functionally distinct.

Mk activity predominates in CD34+CD38⁻ stem cell enriched compartment

Human Mk progenitors are poorly defined in the CD34⁺ hierarchy. The shortcomings of previous studies were partly due to use of standard colony assays, which rely on morphological based assessment for Mks. Alternate cytokine requirements of Mk indicate standard colony assays do not efficiently support differentiation along this lineage. We employed the highly specific collagen-based MegacultTM assay that detects colony-forming unit-megakaryocyte (CFU-Mk) based on GPIIb/IIIa (CD41) staining (fig S3C-D). Using this assay, we found a striking enrichment of CFU-Mk in the CD34+CD38⁻ stem cell compartment (fig. S3C). Of the 11 subsets we fractionated from FL, CB and BM, MPP F2 was particularly enriched for CFU-Mk potential (figs S3C). Although CFU-Mk activity was found in the CD34+CD38⁺ progenitor compartment, it was minimal in comparison (fig. S3C). Particularly important was the finding that MEP fractions were not strikingly enriched for CFU-Mk activity (fig. S3C). These results were consistent across FL, CB and BM. It was clear that the bulk of Mk potential resides within the CD34+CD38⁻ compartment across development, rather than in the tradition view of MEP only arising from CMP. In conjunction with MC colony data that demonstrated concurrent enrichment for BFU-E from these subsets, it is formally possible that true MEPs reside within the stem cell compartment and not within classically defined MEPs within the CD34+CD38⁺ progenitor compartment. But absolute evidence that both Mk and Er colonies are derived from a common progenitor required single cell analysis.

Single cell transcription profiling of newly identified populations

Virtually, all single cells from CMP F1 and GMPs expressed at least one of three myeloid lineage factors such as *CEBPa*, *MPO* and *CSF2RA* (Fig. 5A) *CEBPa* is a master regulator of steady state myelopoiesis and activates the promoter of several critical myeloid genes to lock down myeloid development. Approximately 10% of FL and BM CMP F1 cells expressed *CEBPa* alongside *GATA1* or *EPOR*, suggesting molecular factors of My and Er fates co-exist in some cells within CMP F1 (Fig. 5A). This coincides with the small percentage of mixed colony potential from this subset in the single cell assay (Fig. 3B). 20% of BM CMP F1 also co-expressed *CEBPa* with *EPOR* or *GATA1*. However unlike FL, BM CMP F1 cells did not co-express *EPOR* and *GATA1*. Moreover the expression of *CEBPa* and *MPO* was dramatically

higher in BM CMP F1 suggesting a dominant My program is already established in CMP F1 cells that co-express My and Er factors (Fig 5A). Dramatically higher levels of *EPO* or *GATA1* observed in fully committed erythroid committed subsets, such as MEP F3 suggests the established E/Mk program that is also supported by the single cell functional assay (Fig. 5A).

Experimental Procedures

Human CD34+ Cord Blood, Bone Marrow, and Fetal Liver

Human cord blood samples were obtained from Trillium Hospital (Mississauga, Ontario, Canada with informed consent in accordance to guidelines approved by University Health Network (UHN) Research Ethics Board. Cord blood was processed 24 - 48h post-delivery. Mononuclear cells were enriched using Ficoll-Paque (GE Healthcare Life Sciences, cat. no. 17-1440-02) followed by red-blood lysis by ammonium chloride (Stem cell Technoloiges, cat. no. 07850). CD34+ selection was performed using CD34 Microbead kit (Miltenyi Biotech, cat. no. 130-046-703). Cells were viably stored at -80°C or -150°C. In most cases, individual donors were processed for experiments. CD34+ adult bone marrow was purchased from Lonza (cat. no. 2M-101C). Only bone marrow obtained from donors between 25 and 35 years of age were used in this study. CD34+ human fetal liver cells were purchased from AllCells (AllCells LLC, Alameda, CA, US, cat. no. FL-CD34-001F or FL-CD34-002F/~20-22wk gestation).

Cell Sorting

Viable cells were thawed via dropwise addition of IMDM based (IMDM + 20% fetal calf serum (FCS) + DNaseI). Final concentration of DNase I (Roche Applied Science, 10104159001) in IMDM solution was 200ug/mL. Post-thaw, cells were spun at low RPM (~1000) for 20min at 4°C. After the spin, thawing solution was removed and cells were resuspended in 100uL of PBS+5% FCS for exposure to antibodies. The following antibodies were used for cell sorting: CD7 PB (BD bioscience, clone M-T701), CD10 A700 (BD bioscience, clone HI10a, custom conjugation), CD34 APC-Cy7 (BD bioscience, clone 581, custom conjugation), CD38 PC7 (BD bioscience, clone HB7, cat. no. 335790), CD110(cMPL) PE (BD bioscience, clone BAH-1, custom conjugation), CD71 NC650 (ebioscience, clone OKT9, custom conjugation), CD135(FLT3) biotin (BD bioscience, clone 4G8, custom conjugation), CD45RA FITC (BD bioscience, clone HI100, cat. no. 555488), Thy1/CD90 APC (BD bioscience, clone 5E10, cat.

no. 559869), CD49f PE-Cy5 BD bioscience, clone GoH3, cat. no. 551129), CD19 PE (BD bioscience, clone SJ25C1, cat. no. 340364), Streptavidin Qdot605 conjugate (Life Technologies, cat. no. Q10101MP). Cell sorting was performed on the BD FACSAria III.

Methylcellulose and MegaCult™ Assay

Cell populations were sorted directly into complete methylcellulose (cytokines: SCF, GM-CSF, IL-3, G-CSF, EPO; cat. no. H4034, Stem cell technologies). Complete methylcellulose medium was additionally supplemented with FLT3 ligand (20ng/mL final concentration), IL-6 (50ng/mL final concentration), LDL (4ug/mL final concentration, cat. no. 02698, Stem cell technologies). 300 cells were deposited by cell sorting into 3mL of methylcellulose, mixed and 1mL (100 cells) was plated into 35 mm dishes in duplicates. Colonies were allowed to differentiate for 14 days and were morphologically assessed for colonies. In the case for fetal liver, colonies were often picked and analyzed for flow cytometry as BFU-E did not hemoglobinize well. For cell counting, plates were resuspended in PBS/FCS and subsequently counted on the Vi-cell XR (Beckman Coulter). For Megacult, 200 sorted cells in technical duplicates were plated in a collagen-based material in double chamber culture slides and cultured for 10 days (MegaCult®-C Medium with Cytokines, Cat no: 04901/04951, Stem cell technologies). Staining for GPIIb/IIIa antibody and scoring of CFU-Mk colonies were performed according to manufacturer's protocol.

Single cell Assay

MS-5 stroma (kindly provided by Dr. K. Itoh, Japan) was seeded into 96-well flat bottom plates (Nunc) at a density of 3×10^3 to 5×10^3 cells in Myelocult medium (H5100, Stem cell technologies) per well. MS-5 cells were seed in 100uL of medium an allowed to adhere for 24-48h prior to cell sorting. Before cell sorting, media was removed with multichannel pipet and 200uL of serum free media (StemPro34 SFM with nutrient, cat. no. 10639, Life Technologies) supplemented with (SCF - 100ng/mL, FLT3 - 20ng/mL, TPO - 100ng/mL, EPO - 3units/mL, IL-6 - 50ng/mL, IL-3 - 10ng/mL, IL-11 - 50ng/mL, GM-CSF - 20ng/mL, LDL - 4ug/mL, 2-ME, L-Glutamine, Pen-strep). Single cells were cultured for 2 weeks without media change. For HSC fraction, where some clones did not emerge till after 2 weeks, a half media change was performed. Wells content was harvested and subjected to flow cytometry for analysis (LSRII, BD biosciences). The following antibodies were used to assess myeloid cell lineages:

(CD15 PB [clone MMA, cat. no. 642917, BD Biosciences], CD45 FITC [BD Biosciences, clone 2D1], GlyA PE [clone 11E4B-7-6, cat. no. IM2211U, Beckman Coulter], CD41 PC5 [clone P2, cat. no. PN 6607116, Beckman Coulter], CD14 PC7 [clone RM052, cat. no. A22331, Beckman Coulter], CD42b APC [clone HIP1, cat. no. 551061, BD Biosciences], CD11b APC7 [clone xxx, cat. no. xx, BD Biosciences]). Generally, greater than 10 cells were required to call a positive colony.

Fluidigm Assay and RNA sequencing

For RNA sequencing, HSC and progenitor subsets from CB were sorted (2000-4000 cells) and RNA was isolated using the mirVana RNA isolation kit (Ambion). RNA was subjected to amplification using SMARTer™ Ultra Low RNA Kit for Illumina Sequencing (Clontech, Cat. No. 634826). Library preparation of cDNA was performed using the Nextera protocol (Illumina, Nextera DNA Sample Preparation Kit, Cat No. FC-121-1031). Each population was subjected to at least half lane of sequencing on Illumina HiSeq2000.

For the fluidigm protocol, single cells were sorted into eppendorf 96 well skirted plates with 5ul of cell lysis buffer containing 10% NP40 (Fisher Scientific, Cat# 28324), RNaseOut (life technology, Cat# 10777-019) and 5x VILO reaction mix (LT, Cat# 11754-050). To make cDNA from single cells, 0.15ul of 10x Supermix VILO (LT, Cat# 11754-050) and 0.12ul T4 Gene 32 protein (NEB, cat# M0300S) added to the mix and incubated in 25°C, 5min; 50°C, 30min; 55°C, 25min; 60°C, 5min and 70°C 10min. TaqMan preamp Master Mix (life technology, Cat# 4391128) and up to 48 multiplexed inventoried TaqMan assays (final dilution 0.05x) were used to amplify 48 target cDNA. The cDNA target primers used in this study are shown below. Zero, 2, 4 and 8 cells were included as control. RT-PCR pre-amplification cycling conditions were: 20x(95 °C, 5 sec; 60 °C, 4 min). Amplified material then cleaned up using 1.2ul of Exo I 20U/ul and 0.6ul of Exo I (NEB, Cat#M0293S) for 30 min in 37°C. cDNA diluted 5x with water or TE buffer. qPCR was done with the same inventoried TaqMan assays using Fluidigm 48.48 Gene expression Dynamic Arrays (Fluidigm) chips and read in BioMark HDsystem according to manufacturer protocol for single cell QPCR. The data were exported into Microsoft Excel for downstream analysis. Data were analyzed using modified SINGuLAR™ Analysis Toolset and pheatmap package used to generate heatmaps in R. We set the cutoff level for expression at Ct>28. Samples with low GAPDH and PGK1 expression (Ct higher than 20) were considered as outliers and excluded from the analysis.

Assay ID	Gene Name	Cat#	Inventoried
Hs99999903_m1	ACTB	4331182	yes
Hs00269972_s1	CEBPA	4331182	yes
Hs00538896_m1	CSF2RA	4331182	Yes
Hs00237052_m1	CXCR4	4331182	yes
Hs00395519_m1	EBF1	4351372	yes
Hs00181092_m1	EPOR	4331182	yes
Hs00174690_m1	FLT3	4331182	yes
Hs99999905_m1	GAPDH	4331182	yes
Hs01085823_m1	GATA1	4331182	yes
Hs01119304_m1	MPL	4351372	yes
Hs00924296_m1	MPO	4331182	yes
Hs00172003_m1	PAX5	4331182	yes
Hs99999906_m1	PGK1	4331182	yes
Hs01548149_m1	PU.1	4351372	yes
Hs01109452_m1	VWF	4351372	yes

In Vivo Analysis

Only CB cells were used for in vivo analysis due to the number of cells required to generate engraftment. CB populations were sorted, recounted and resuspended in a volume of 25 μ L/mouse for transplant. Eight to twelve week old female NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG) mice were sublethally irradiated at 225cGy, 12-24 hours prior to transplant (3). Intrafemoral injections were performed as previously described (4). Briefly, mice were anesthetized with isoflurane, the right knee was secured in a bent position and a 27 gauge needle was used to drill into the right femur. Cells were then injected using a 28.5 gauge 1/2 cc syringe. Following 2-4 weeks mice were euthanized by cervical dislocation and peripheral blood (PB), right femur (RF) and bone marrow (left and right tibias and left femur, BM) were collected. BM bones were crushed in 1mL of IMDM + 5% FBS using a 2 oz glass mortar and pestle (Fisher Scientific). Crushed bones were subsequently washed with 20 mL of media, filtered through a 40 μ M cell strainer (BD Falcon), centrifuged and the resulting marrow cells resuspended in 1mL of IMDM + 5% FBS. Cells from RF and BM were counted using Vicell XR (Beckman Coulter). PB was collected in tubes containing heparin (Sigma cat. no H3149-10KU) and lysed with ammonium chloride (StemCell Technologies) prior to antibody staining. Cell aliquots were stained in 96-well round bottom plates (BD Falcon) and the human grafts were analyzed by flow cytometry (LSRII Becton Dickinson) for RF/BM: 15 PB (BD biosciences, clone MMA, cat. no. 642917), 45 FITC (BD bioscience, clone 2D1, cat. no. 347463), 19 PE (BD biosciences, clone 4G7, cat. no. 349209), GlyA PE (Coulter, clone 11E4B-7-6, cat. no. IM2211U), 45 PC5 (Coulter, clone J.33, cat. no. IM2653U), 14 PC7 (Coulter, clone RM052, cat. no. A22331), 33 APC (BD biosciences, clone p67.6, cat.no. 340474), 71 APC (BD bioscience, clone M-A712, cat. no. 551374), 34 APC7 (BD bioscience, clone 581, custom conjugation) and PB: CD45 FITC (BD

bioscience), CD19 PE (BD biosciences, clone 4G7, cat. no. 349209), CD45 PC5 (Coulter, clone J.33, cat. no. IM2653U), CD41 PC7 (Coulter, clone P2, cat. no. 6607115), CD42b PE or APC (BD biosciences, clone HIP1), CD3 APC (BD biosciences, clone UCHT1, cat. no. 555335), CD11b APC7 (BD biosciences, clone ICRF44, cat. no. 557754).

REFERENCES

1. P. M. Lansdorp, W. Dragowska, H. Mayani, Ontogeny-related changes in proliferative potential of human hematopoietic cells. *J Exp Med* 178, 787 (Sep 1, 1993).
2. W. W. Pang *et al.*, Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proceedings of the National Academy of Sciences of the United States of America* 108, 20012 (Dec 13, 2011).
3. F. Notta, S. Doulatov, J. E. Dick, Engraftment of human hematopoietic stem cells is more efficient in female NOD/SCID/IL-2Rgcnnull recipients. *Blood*, (Mar 5, 2010).
4. F. Mazurier, M. Doedens, O. I. Gan, J. E. Dick, Rapid myeloerythroid repopulation after intrafemoral transplantation of NOD-SCID mice reveals a new class of human stem cells. *Nat Med* 9, 959 (Jul, 2003).