

## HEMATOPOIESIS

# Identification of two distinct pathways of human myelopoiesis

Roy Drissen<sup>1\*</sup>, Supat Thongjuea<sup>2,3</sup>, Kim Theilgaard-Mönch<sup>4,5</sup>, Claus Nerlov<sup>1\*</sup>

Copyright © 2019  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim  
to original U.S.  
Government Works

Human myelopoiesis has been proposed to occur through oligopotent common myeloid progenitor (CMP) and lymphoid-primed multipotent progenitor (LMPP) populations. However, other studies have proposed direct commitment of multipotent cells to unilineage fates, without specific intermediary lineage cosegregation patterns. We here show that distinct human myeloid progenitor populations generate the neutrophil/monocyte and mast cell/basophil/eosinophil lineages as previously shown in mouse. Moreover, we find that neutrophil/monocyte potential selectively cosegregates with lymphoid lineage and mast cell/basophil/eosinophil potentials with megakaryocyte/erythroid potential early during lineage commitment. Furthermore, after this initial commitment step, mast cell/basophil/eosinophil and megakaryocyte/erythroid potentials colocalize at the single-cell level in restricted oligopotent progenitors. These results show that human myeloid lineages are generated through two distinct cellular pathways defined by complementary oligopotent cell populations.

## INTRODUCTION

Myeloid cell types carry out essential innate immune functions, with macrophages and neutrophils critical to antibacterial defense, eosinophils and basophils providing antiparasitic immunity, and mast cells acting as sentinels in the skin and gut and playing a key role in regulating allergy (1). Defects in myeloid cells lead to immune deficiencies, whereas excessive production can cause severe inflammation-induced tissue degeneration (2, 3), and myeloid cell types are critical regulators of inflammation (4) and the tumor microenvironment (5). Last, malignancies affecting the myeloid lineages include acute myeloid leukemia, myeloproliferative neoplasms, and myelodysplastic syndromes, all increasing in prevalence with population aging. Consequently, the ontogeny of myeloid cells has been intensively studied, with particular emphasis on the identification of progenitor populations committed to specific myeloid fates, as well as the transcription factors (TFs) (6) and cytokines (7) that regulate myeloid lineage commitment and differentiation. Initially, granulocyte and monocyte (GM) progenitors (GMPs) were identified in both murine (8) and human hematopoiesis (9) and proposed to constitute a single progenitor capable of forming all myeloid cell types. Subsequently, murine GMPs were found to contain eosinophil, basophil, and mast cell potential (10, 11). In contrast, in humans, eosinophil potential was located to the common myeloid progenitor (CMP) population (12), whereas the origin of human mast cell/basophil progenitors remains to be determined (13). We recently found that murine pre-GMs and GMPs could be subdivided using a *Gata1*-enhanced green fluorescent protein (EGFP) reporter, and that the *Gata1*-EGFP<sup>+</sup> and *Gata1*-EGFP<sup>-</sup> subpopulations contained mast cell/eosinophil and monocyte/neutrophil potentials, respectively (14). The separation of

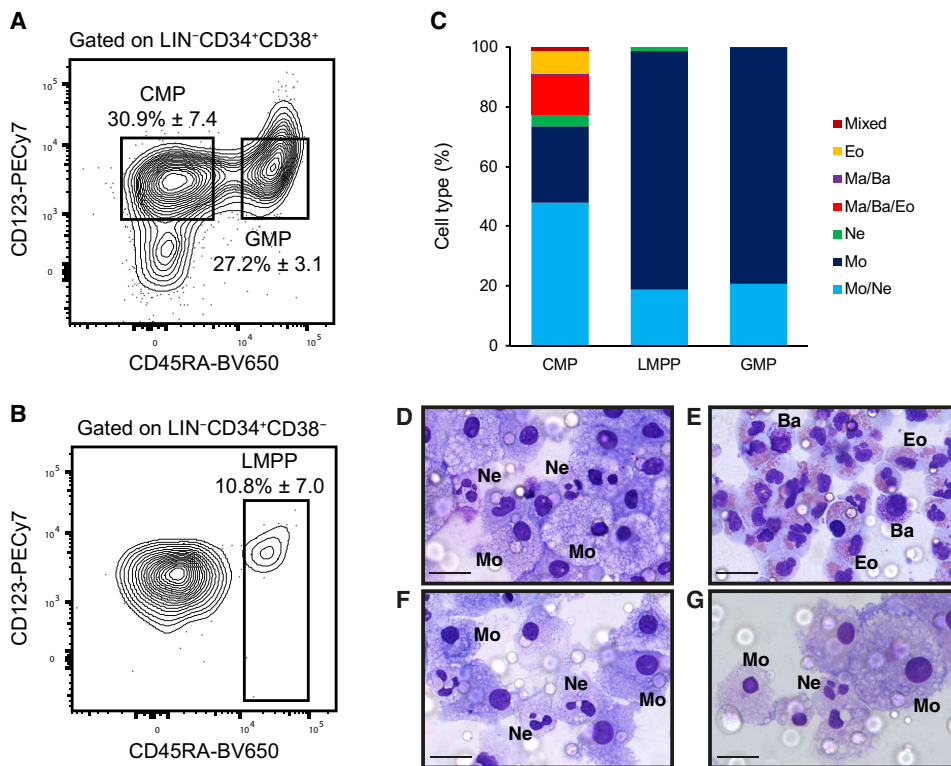
the two sets of myeloid lineage potentials preceded their separation from megakaryocyte/erythroid and lymphoid potentials, respectively. Because also human myeloid cell types can be classified into *GATA1*-expressing (mast cells, eosinophils, and basophils) and *GATA1*-nonexpressing cell types (neutrophils and monocytes/macrophages) (15, 16), we here investigate whether human myelopoiesis is similarly organized. We find that basophil/mast cell potential resides within the human CMP, and using single-cell RNA sequencing identify CD114 and CD131 as markers that define CMP subpopulations that contain neutrophil/monocyte and eosinophil/mast cell/basophil potential, respectively. Furthermore, neutrophil/monocyte potential cosegregates with lymphoid potential, whereas eosinophil/mast cell/basophil potential cosegregates with erythroid/megakaryocytic potential, with CD131<sup>+</sup> CMPs as oligopotent cells with combined eosinophil/mast cell/basophil and erythroid/megakaryocytic lineage potential. These findings lead to a revised model for human myelopoiesis and show that initial lineage commitment in human hematopoiesis involves the generation of lineage-restricted oligopotent progenitor populations containing defined subsets of lineage potentials.

## RESULTS

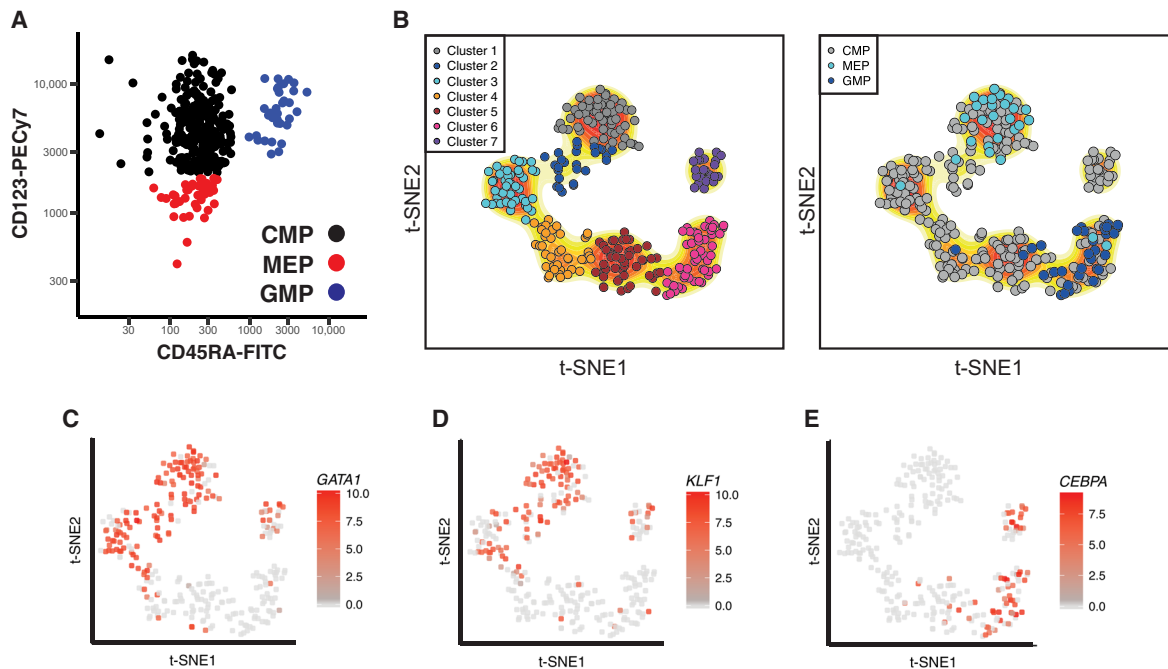
To localize human mast cell/basophil potentials, we cultured single CMPs (defined as LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD123<sup>+</sup>CD45RA<sup>-</sup>) and GMPs (LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD123<sup>+</sup>CD45RA<sup>+</sup>) (Fig. 1A) (9) in the presence of cytokines promoting the differentiation of all myeloid lineages (hSCF, hFLT3L, hIL-3, hIL-5; hIL-6, hGM-CSF, and hG-CSF; M-conditions) and analyzed the morphology of the cells generated. This showed that individual CMPs generated cultures containing eosinophils and basophils/mast cells (Fig. 1, C and E) or neutrophils and monocytes (Fig. 1, C and D), but only very rarely cells from both of these subsets (2 of 181 cultures; Fig. 1C). The majority of basophil/eosinophil-containing cultures (24 of 39 or 62%) were bilineage. In contrast, GMPs only generated neutrophils and monocytes under these conditions (Fig. 1, C and F). In the mouse, we found that lymphoid-primed multipotent progenitors (LMPPs) generate neutrophils and monocytes, but not eosinophils or mast cells (14). We therefore analyzed the myeloid lineages

<sup>1</sup>MRC Molecular Hematology Unit, MRC Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK. <sup>2</sup>MRC WIMM Centre for Computational Biology, MRC Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK. <sup>3</sup>NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK. <sup>4</sup>Department of Hematology and Finsen Laboratory, National University Hospital, Denmark. <sup>5</sup>Biotech Research and Innovation Centre and Program for Translational Hematology - The Novo Nordisk Foundation Center for Stem Cell Biology, University of Copenhagen, Copenhagen, Denmark.

\*Corresponding author. Email: claus.nerlov@imm.ox.ac.uk (C.N.); roy.drissen@imm.ox.ac.uk (R.D.)



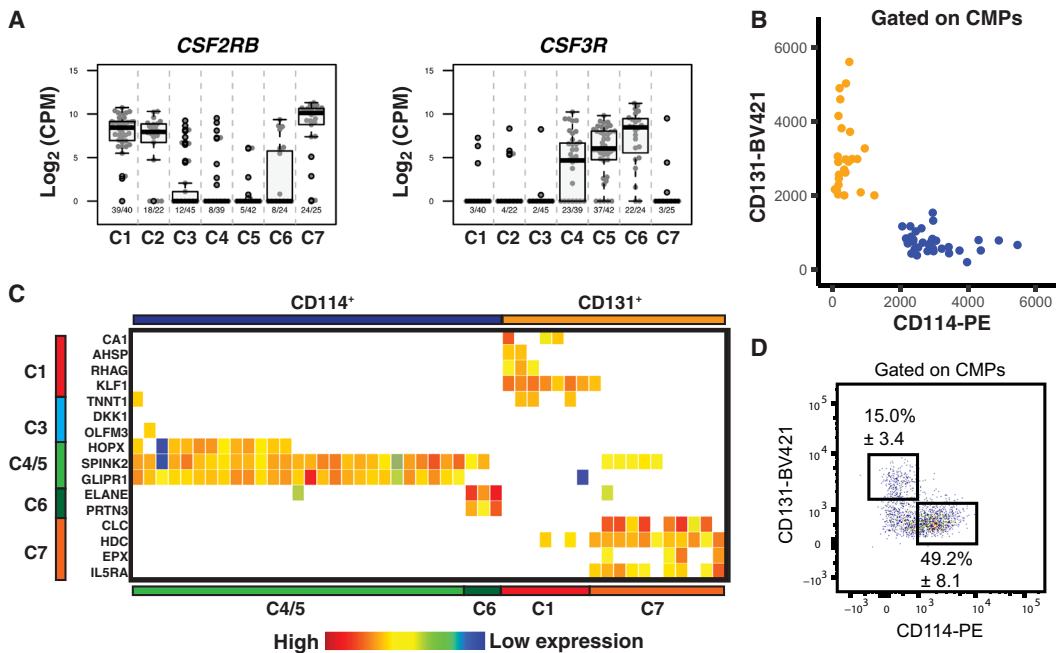
**Fig. 1. Myeloid potential in predefined hematopoietic progenitor populations.** (A and B) Gating strategy of populations in adult human bone marrow samples. Displayed are live, lineage-negative singlets, positive or negative for CD34 and CD38 as indicated above the plots. CD10 was included in the lineage cocktail. Data are representative of seven donors in 16 experiments, and the numbers show gated cells as percentage  $\pm$  SD of the parental gate. PECy7, phycoerythrin cyanine 7. (C) Histogram showing cell type determined by morphology of cytopins from single-cell cultures of CMP ( $n = 181$ ), LMPP ( $n = 75$ ), and GMP ( $n = 77$ ) as percentage of total number of cultures analyzed. Mo, monocyte; Ne, neutrophil; Eo, eosinophil; Ma/Ba, mast cell/basophil; Mixed, neutrophil/monocyte in combination with eosinophil/basophil/mast cell morphology. Data are summary of six independent experiments (table S1A). (D to G) Morphology of single-cell cultures from CMP (D and E), GMP (F), or LMPP (G), representative of the data in (C). Scale bars, 25  $\mu$ m.



**Fig. 2. Heterogeneity of human CMPs revealed by single-cell RNA sequencing.** (A) Plot indicating the CD45RA and CD123 expression measures by flow cytometry of the CMP, GMP, and MEP cells used for single-cell RNA sequencing. FITC, fluorescein isothiocyanate. (B) Left: t-SNE plot of 237 CMP, 32 GMP, and 27 MEP cells showing seven clusters as indicated by coloring. The contour plot (red gradient color in the background) indicates the kernel smoothing density of cells. Right: The same t-SNE plot showing the individual cell types: CMP (gray), MEP (light blue), and GMP (dark blue). (C) *GATA1*, (D) *KLF1*, and (E) *CEBPA* expression [as  $\log_2$  (CPM)] superimposed on the t-SNE plot from (B).

generated by LMPPs (defined as  $LIN^-CD34^+CD38^-CD45RA^+$  (Fig. 1B) (17)) and observed only neutrophils and monocytes (Fig. 1, C and G).

Human mast cells/basophil progenitors therefore reside within the phenotypically defined CMP compartment, as previously reported for eosinophil progenitors (12). In addition, the observed combined



**Fig. 3. Characterization of progenitor populations defined by CD114 and CD131.** (A) Gene expression level of *CSF2RB* and *CSF3R* in single CMP cells in identified clusters. Y axis represents the expression level in the  $\log_2$  (CPM) scale. X axis represents the cluster identification. Numbers below each plot show the number of expressing cells and the total number of cells within the respective cluster. Box plot shows median and quartile values, and whiskers show outlier values within 1.5 times of the interquartile range. (B) Plot indicating the CD114 and CD131 flow cytometry signals of index-sorted single CMP cells used for gene expression analysis. PE, phycoerythrin. (C) Heat map showing gene expression of index-sorted CD131<sup>+</sup> and CD114<sup>+</sup> CMPs (indicated on top) as shown in (B). The genes were chosen as markers for the indicated clusters (left). Cells were clustered manually, and the gene expression was used to annotate the cells to one of the clusters (bottom). (D) Flow cytometry plot showing CD114 and CD131 signals of CMP cells and gates used for sorting CMP subpopulations. Numbers are cells within each gate as average percentage  $\pm$  SD of parental gate. Data are representative of 11 independent experiments using six different donors.

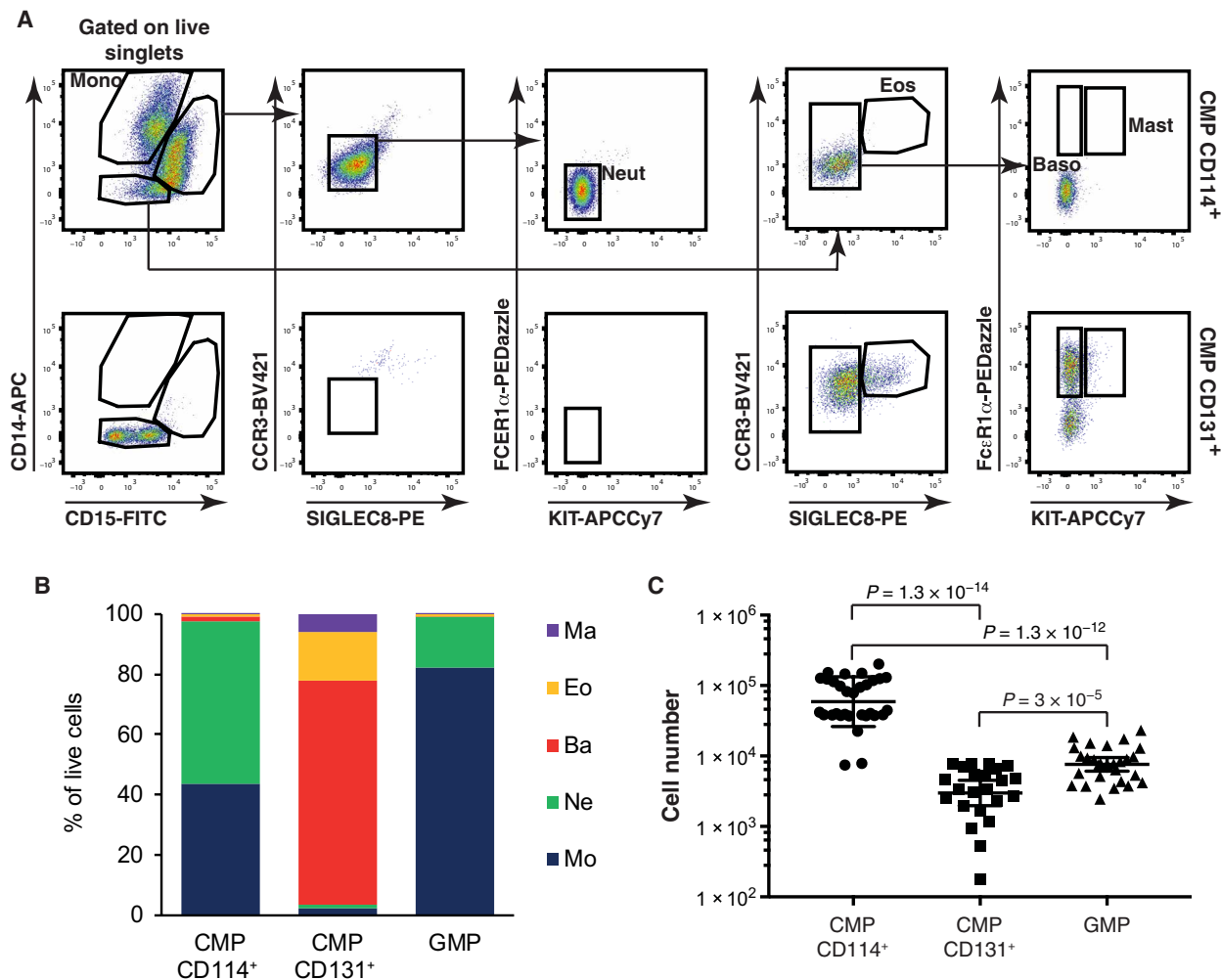
eosinophil and mast cell/basophil output indicated the presence of a common progenitor for these lineages, distinct from progenitors with monocyte/neutrophil potential, within the CMP compartment. To deconvolute these functionally distinct progenitor subsets, we performed Smart-seq2-based single-cell RNA sequencing (18) of index-sorted CMPs [237 cells after final quality control (QC)], including GMPs (32 cells) and megakaryocyte/erythroid progenitors (MEPs) (27 cells) as reference populations (Fig. 2A). T-distributed stochastic neighbor embedding (t-SNE) analysis (19) followed by kernel density estimation and *k*-mean clustering analysis identified seven cellular clusters (Fig. 2B). MEPs and GMPs primarily mapped to cluster 1 (C1) and C6, respectively, indicating that these represent megakaryocyte/erythroid- and neutrophil/monocyte-committed CMP subpopulations. Consistent with this, *GATA1* and *KLF1* were highly expressed in C1 to C3 and C7, but not C4 to C6 (Fig. 2, C and D), with *CEBPA* expressed in C4 to C6 and C7 (Fig. 2E). We previously observed coexpression of *Cebpa*, *Klf1*, and *Gata1* in murine progenitors with mast cell/eosinophil potential, whereas neutrophil/monocyte progenitors expressed neither *Gata1* nor *Klf1*. Furthermore, *Cebpa* expression was lost in committed erythroid progenitors, where *Klf1* and *Gata1* expression were maintained (14). We therefore hypothesized that C4 to C6 were candidate neutrophil/monocyte progenitors, whereas C7 was a candidate eosinophil/basophil/mast cell progenitor population.

To allow further characterization of these subpopulations, we mined the RNA sequencing data for differentially expressed genes encoding surface markers (table S2). Ten candidates were identified

(Fig. 3A and fig. S1A), and the encoded surface markers were tested for their ability to subdivide the CMP population (fig. S1, B and C). We observed that Fc fragment of IgE receptor Ia (Fc $\epsilon$ R1 $\alpha$ ), CD114, and CD131 each labeled a subfraction of CMPs. Of these, CD131 and CD114 were of particular interest, because genes that encode them (*CSF2RB* and *CSF3R*, respectively) were expressed in *GATA1*-expressing (C1 to C3 and C7) and *GATA1*-nonexpressing (C4 to C6) clusters, respectively (Fig. 3A). The CD133 marker (encoded by *PROM1*) previously used to separate myeloid progenitor subsets (20) showed an expression pattern similar to CD114 by both gene expression (fig. S1A) and flow cytometry (fig. S2A), but was less specific for C4 to C6. We depicted CD131 and CD114 expression on MEP, GMP, and more immature populations. CD131 expression is not observed in hematopoietic stem cells and MPP (HSC/MPP) or LMPP but is sustained on MEPs. CD114 is expressed at low levels in both the HSC/MPP and LMPP subsets but up-regulated in CD114<sup>+</sup> CMPs and sustained in GMPs (fig. S2B).

To test the ability of these markers to prospectively separate the identified clusters, signature genes selectively expressed in each cluster were identified (fig. S3A and table S2). We found that C4 and C5 were separated only by distinct cell cycle status (fig. S3, B and C), and common signature genes were therefore identified for these two clusters. C2 was too weak to generate signature genes.

To test whether CD114 and CD131 surface expression identified the predicted clusters, single CMPs were index-sorted on CD114 and CD131 (Fig. 3B), and expression of the above identified cluster signature genes was analyzed by microfluidics-based quantitative reverse transcription polymerase chain reaction (qRT-PCR). As predicted, C1 and C7 signature genes were expressed in CD114<sup>-</sup>CD131<sup>+</sup> CMPs (henceforth CD131<sup>+</sup> CMPs) and C4/5 and C6 signatures in CD114<sup>+</sup>CD131<sup>-</sup> CMPs (henceforth CD114<sup>+</sup> CMPs) (Fig. 3C), validating the ability of these markers to prospectively separate clusters with molecular characteristics of neutrophil/monocyte progenitors (C4/5 + C6) and eosinophil/mast cell progenitors (C7). The CD114<sup>-</sup>CD131<sup>-</sup> CMPs were a mixed population containing predominantly cells expressing C4/5 signature genes, as well as some C1/3 and C7 signature gene-expressing cells. The CD114<sup>-</sup>CD131<sup>-</sup> double-negative CMP fraction therefore did not contain distinct cell types, but rather cells with functional characteristics of CD114<sup>+</sup> CMPs or CD131<sup>+</sup> CMPs that could not be separated using these markers.

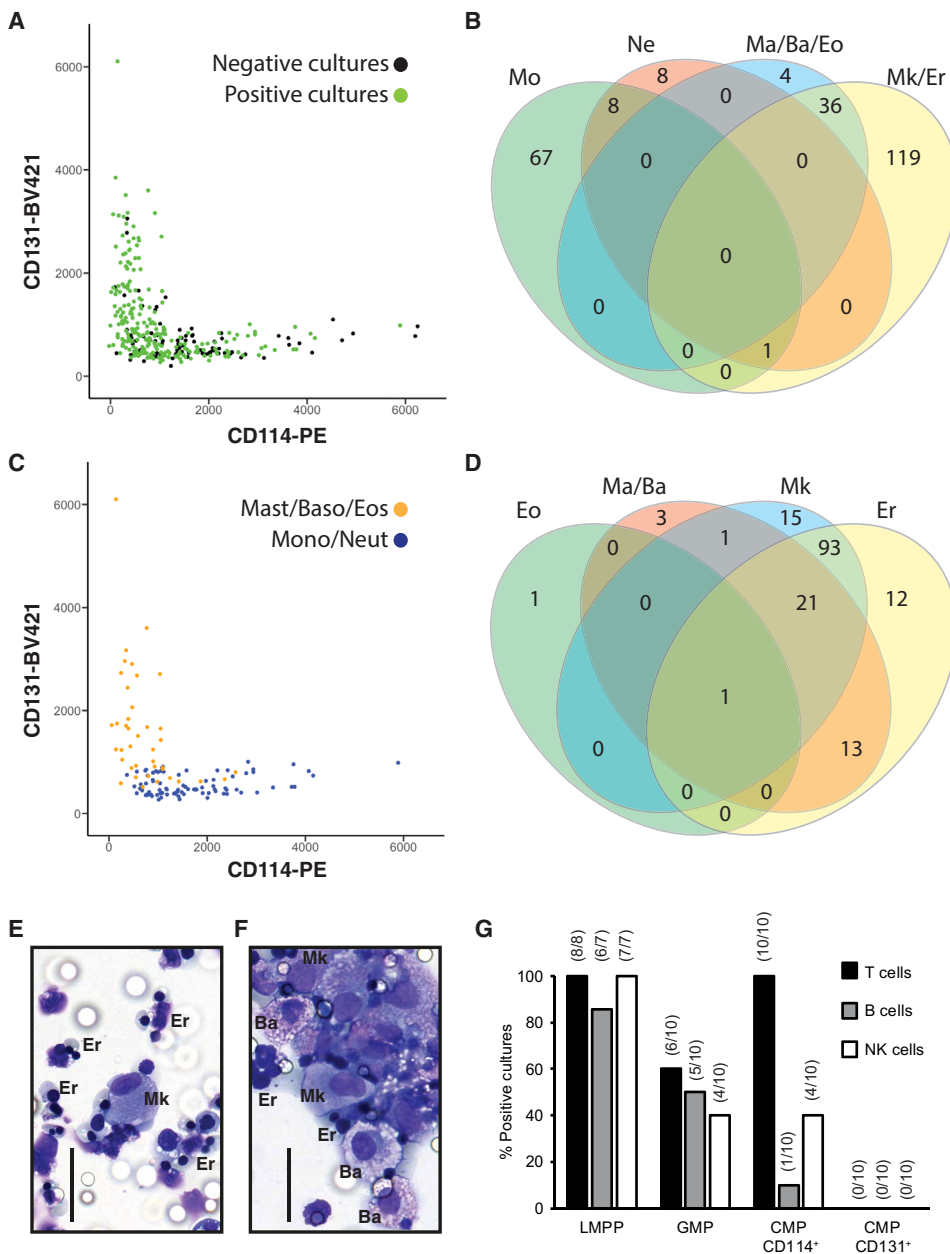


**Fig. 4. Measurement of lineage output from CD114<sup>+</sup> and CD131<sup>+</sup> CMPs.** (A) Gating strategy for flow cytometry analysis of bulk cultured progenitors. Examples of cultures of CD114<sup>+</sup> CMPs (top) and CD131<sup>+</sup> CMPs (bottom) are shown. Arrows in the top panel indicate sequential gating. APC, allophycocyanin. (B) Quantification of cell types identified as in (A) after culture of indicated progenitor populations, presented as frequency of total defined cells. Data are averaged from five (GMP) or six (CMP CD114<sup>+</sup> and CMP CD131<sup>+</sup>) donors, and 2 to 10 cultures were grown for each population. (C) Number of live cells retrieved from cultures in (B) measured by flow cytometry. *P* values indicate significance of differences in cell numbers (Mann-Whitney *U* test).

To efficiently separate the *GATA1*-expressing and *GATA1*-nonexpressing clusters, we therefore focused on CD114<sup>+</sup> and CD131<sup>+</sup> CMPs. These two populations were sorted (Fig. 3D), bulk cells cultured under M-conditions and analyzed for the generation of myeloid cell types by flow cytometry (Fig. 4A). We observed that CD114<sup>+</sup> CMPs generated predominantly neutrophils and monocytes (>98%; Fig. 4B), whereas CD131<sup>+</sup> CMPs generated predominantly eosinophils, mast cells, and basophils (>92%; Fig. 4B), demonstrating that CD114 and CD131 expression can be used for the prospective separation of these two sets of myeloid lineage potentials. Furthermore, in this assay, GMPs generated essentially only neutrophils and monocytes (Fig. 4B), and these cultures were less proliferative than the CD114<sup>+</sup> CMPs (Fig. 4C).

To colocalize lineage potentials at the single-cell level, we index-sorted CMPs followed by single-cell culture under conditions supporting megakaryocyte/erythroid lineage output [myeloid/megakaryocyte/erythroid (MME) conditions; see Materials and Methods]. To stringently define lineage readout, all cultures were subjected to both morphological analysis and microfluidics-based expression analysis

of key lineage-specific genes, and only readouts supported by both criteria were considered positive, except for megakaryocytes, which could be reliably identified by morphology alone. At least one positive lineage readout was obtained from 243 cultures (Fig. 5A). We observed that megakaryocyte/erythroid lineage output was highly dissociated from neutrophil/monocyte output ( $P = 2 \times 10^{-14}$ ,  $\chi^2$  test), but not from eosinophil/basophil/mast cell lineage output ( $P = 0.22$ ,  $\chi^2$  test) (Fig. 5B). Mapping single cells to CD114 and CD131 expression confirmed that CD131<sup>+</sup> CMPs were restricted to mast cell/basophil/eosinophil output, whereas CD114<sup>+</sup> CMPs generated neutrophils and monocytes (Fig. 5C). Separate analysis of all *GATA1*<sup>+</sup> lineages showed that the majority of cultures were bi- or oligolineage under MME conditions (129 of 160 or 81%; Fig. 5D), including megakaryocyte/erythroid (Fig. 5E) and megakaryocyte/erythroid/basophil (Fig. 5F) readouts. However, compared with the data obtained under M-conditions (Fig. 1), we noticed that eosinophil lineage output was reduced. We therefore repeated the analysis, using the same donor bone marrows, under M-conditions. Here, neutrophil and eosinophil readouts, and consequently bilineage myeloid readouts, were improved

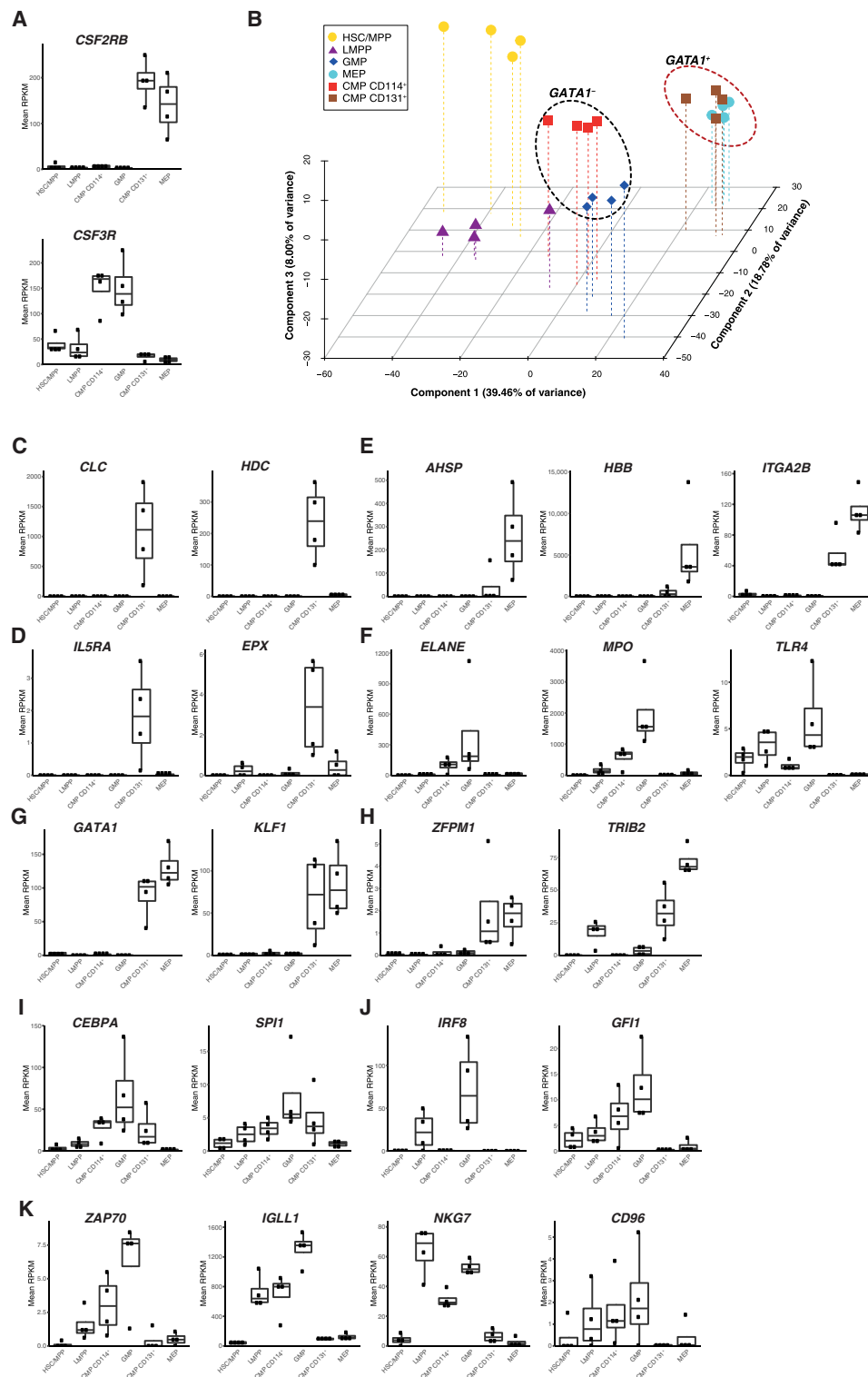


**Fig. 5. Lineage affiliation of prospectively isolated myeloid progenitors.** (A) CMP cells were index-sorted and cultured under MME conditions. The CD131 and CD114 expression of cells with at least one positive lineage readout indicated. The data are cumulative from two independent experiments (table S1B). (B) Venn diagram of lineage output from cells in (A) with positive lineage readout defined by the presence of both appropriate morphology and lineage-specific gene expression. (C) CD114 and CD131 intensity of isolated CMPs that gave rise to cultures containing mast cells, basophils or eosinophils (orange), or monocytes or neutrophils (blue). (D) Venn diagram as in (B) showing the overlap between positive eosinophil, mast cell/basophil, megakaryocyte, and erythrocyte lineage readouts. (E and F) Cytospins of single CMP cultures showing (E) erythroid (Er) and megakaryocyte (Mk) morphology and (F) erythroid, megakaryocyte, and basophil (Ba) morphology. (G) Lymphoid potential of indicated progenitor cell populations (30 cells per culture) grown in B/NK cell conditions (MS-5 stromal cells) or T cell conditions (OP9-hDL1 stromal cells) shown as frequency of cultures producing B, NK, or T cells. Data are from two donors in two independent experiments. Numbers above bars indicate number of cultures with a positive readout/number of cultures.

in M-condition compared with MME-condition cultures (fig. S4, A and B). Under both conditions, a significant dissociation of neutrophil/monocyte and mast cell/basophil/eosinophil readouts was observed (M-conditions:  $P = 0.0014$ ; MME conditions:  $P = 1.2 \times 10^{-7}$ ,  $\chi^2$  test).

Although this dissociation was absolute under MME conditions, under M-conditions, we observed rare cultures where mast cell/basophil/eosinophil potential was combined with neutrophil potential (but not monocyte potential) (fig. S4, B to D), an observation similar to that previously made in murine pre-GMs (14). Therefore, optimal measurement of myeloid and myeloid-erythroid copotentials may require distinct culture conditions, and suboptimal conditions lead to an underestimate of the potency of the cells analyzed. Last, to determine the relationship of the two myeloid lineage subsets to lymphoid lineages, we measured lymphoid lineage output of LMPPs, GMPs, CD114<sup>+</sup> CMPs, and CD131<sup>+</sup> CMPs in murine marrow stromal cell line (MS-5) [for B and natural killer (NK) cell readout] and osteo-petrotic 9-human delta-like 1 (OP9-hDL1) cocultures (for T cell readout). We observed B, T, and NK cell output from populations with neutrophil/monocyte potential (LMPP, CD114<sup>+</sup> CMP, and GMP), but not from CD131<sup>+</sup> CMPs (Fig. 5G and fig. S4E).

To molecularly characterize the identified CMP subpopulations, we performed Smart-seq2-based RNA sequencing of bulk cell populations from four independent bone marrow donors. In addition to CD114<sup>+</sup> CMPs and CD131<sup>+</sup> CMPs, we also profiled GMPs, MEPs, LMPPs, and HSCs/MPPs (defined as LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD45RA<sup>-</sup>) (fig. S5). We validated that the purified CMP subpopulations expressed *CSF3R* and *CSF2RB*, respectively (Fig. 6A). In principal components analysis (PCA), CD131<sup>+</sup> CMPs clustered with MEPs, whereas CD114<sup>+</sup> CMPs clustered with GMPs and were adjacent to LMPPs, consistent with these populations defining distinct developmental pathways (Fig. 6B). Similarly, the expression of mast cell/basophil (Fig. 6C), eosinophil (Fig. 6D), and megakaryocyte/erythroid (Fig. 6E) genes was observed in CD131<sup>+</sup> CMPs and MEPs, whereas neutrophil/monocyte genes were expressed in CD114<sup>+</sup> CMPs and GMPs (Fig. 6F). Comparison of CD114<sup>+</sup> and CD131<sup>+</sup> CMPs using gene set enrichment analysis [GSEA (21)] showed that the former were enriched for neutrophil, monocyte, and lymphoid (B, T, and NK cell) gene expression (fig. S6, A and B), whereas basophil and eosinophil genes were enriched in the latter (fig. S6C). The gene expression signatures of the CMP subfractions therefore correspond to their lineage potentials. Hematopoietic lineage specification is controlled by the combinatorial action



**Fig. 6. Gene expression in prospectively isolated progenitor populations.** (A) The indicated cell populations were purified from four independent bone marrow donors and RNA-sequenced. The box-and-whisker plots show the expression of *CSF2RB* and *CSF3R*. Boxes show mean and central quartiles; whiskers show data range. Individual data points are overlaid on the plots. (B) The three-dimensional plot shows PCA using the first three components derived from 1868 genes [ $P < 0.05$  (ANOVA) and  $CV \geq 0.3$ ]; encircled areas indicate clusters containing myeloid-erythroid progenitor cells with or without *GATA1* expression as indicated. (C to K) Gene expression levels measured by RNA sequencing of the indicated genes, shown as in (A).

of key TFs (6). CD131<sup>+</sup> CMPs express TFs associated with both erythroid/megakaryocytic differentiation (*KLF1* and *GATA1*; Fig. 6G) and with specification of *GATA1*<sup>+</sup> myeloid cell types (*CEBPA* and *SPI1*; Fig. 6I). Expression of myeloid TFs (*SPI1* and *CEBPA*) was decreased upon MEP specification, concomitant with expression of *ZFPM1* and its target gene *TRIB2* (22), a negative regulator of CCAAT/enhancer binding proteins (C/EBPs) (Fig. 6H) (23, 24), and paralleled by down-regulation of basophil and eosinophil programs (fig. S6D), consistent with the loss of myeloid lineage potentials in this population (9). In contrast, CD114<sup>+</sup> CMPs and GMPs expressed genes involved in monocyte/neutrophil specification (*GFI1* and *IRF8*; Fig. 6J), as well as general myeloid lineage specification (*SPI1* and *CEBPA*), but lack expression of *GATA1* (Fig. 6G). GSEA showed that neutrophil and monocyte/macrophage genes were expressed at higher levels in GMPs compared with CD114<sup>+</sup> CMPs (fig. S6E), consistent with their lower expression of the TFs driving these differentiation programs (*IRF8*, *GFI1*, *CEBPA*, and *SPI1*), and indicating that CD114<sup>+</sup> CMPs represent an earlier, committed stage of neutrophil/monocyte development than GMPs. In line with the lymphoid potentials found in the populations, T cell (*ZAP70*), B cell (*IGLL1*), NK cell (*NKG7*), and pan-lymphoid (*CD96*) representative genes are all preferentially expressed in the lymphoid-competent progenitor subsets (LMPPs, CD114<sup>+</sup> CMPs, and GMPs), when compared with CD131<sup>+</sup> CMPs and MEPs (Fig. 6K).

## DISCUSSION

We have here used a combination of single-cell RNA sequencing and single-cell functional readouts to assess the lineage potentials of individual human bone marrow progenitor cells. We have assessed all myeloid cell fates and combined morphological, flow cytometric, and gene expression analysis to obtain coherent and reliable identification of the cell types generated. We observe that when assessed *in vitro* at the single-cell level, CD34<sup>+</sup>CD38<sup>+</sup> human progenitor cells have either neutrophil/monocyte or eosinophil/basophil/mast cell potential, but virtually never a combination

of these. We also find that restriction to neutrophil/monocyte myeloid fate occurs in LMPPs, where lymphoid lineage readout was previously found to co-occur with neutrophil and/or monocyte readout in the vast majority of single cells analyzed (25). Conversely, most of the CD131<sup>+</sup> CMPs that generate basophil/mast cell/eosinophils also produce megakaryocyte or erythroid lineage cells, but not neutrophils, monocytes, or lymphoid cells (this study). Together with previous findings that GMPs (9) and LMPPs (17) lack megakaryocyte/erythroid lineage potential, these results support a hierarchical model of adult human bone marrow hematopoiesis where *GATA1*<sup>-</sup> (neutrophils and monocytes) and *GATA1*<sup>+</sup> myeloid lineages (mast cells/basophils/eosinophils) separate before they segregate from lymphoid and megakaryocyte/erythroid lineages, respectively (fig. S7).

This model has several key features in common with our current understanding of murine hematopoiesis (14). In particular, in neither mouse nor human has a myeloid lineage-committed progenitor containing all myeloid lineage potentials at the single-cell level been identified. Instead, progenitors restricted to neutrophil/monocyte differentiation can be prospectively isolated in both species [human CD114<sup>+</sup> CMP and GMP, respectively, corresponding to murine *Gata1*-pre-GM and GMP (14)], as can progenitors restricted to mast cell/basophil and eosinophil myeloid differentiation (CD131<sup>+</sup> CMP and *Gata1*<sup>+</sup> GMP). Single-cell RNA sequencing of human myeloid progenitor cells has been used to computationally generate trajectories leading to eosinophil and mast cell/basophil progenitor formation, variably predicting that these cell types cosegregate with neutrophil/monocyte (26) or erythroid lineage progenitors (27). Our data would be largely compatible with the latter model. Our model is not readily compatible with the previously observed segregation of neutrophil (CD133<sup>+</sup>) and basophil/eosinophil (CD133<sup>-/lo</sup>) progenitors, where monocytes were generated from both populations (20). The use of two markers (CD114 and CD131) and the greater selectivity of CD114 expression for neutrophil/monocyte-restricted progenitors likely explain why we obtain strict segregation of neutrophil-monocyte from basophil-mast cell-eosinophil, as well as megakaryocyte/erythroid potential.

Recent studies have proposed direct commitment of multipotent HSCs/MPPs to unilineage fate in adult bone hematopoiesis, on the basis of the inability to detect oligopotent cells with combined myeloid and erythroid/megakaryocytic lineage potential functionally (28) or computationally based on coexpression of lineage-affiliated genes in single cells (26). However, we here observe combined megakaryocyte/erythroid and myeloid readout from >15% of human bone marrow CMPs. One possible explanation for this discrepancy is that our assay conditions and analysis were optimized for the detection of basophils/mast cells and eosinophils, which are not fully detected by the CD11b antibody used by Notta *et al.* (28) to identify myeloid cells (fig. S8). Although this remains to be clarified, it underscores the importance of evaluating the output of myeloid lineages individually rather than as a whole, because they derive from two distinct progenitor pathways and therefore cannot be used as proxies for each other.

Our results establish the existence of two complementary oligopotent progenitors in human hematopoiesis: one containing neutrophil/monocyte as well as lymphoid lineage potentials [the previously described LMPP (17, 29)] and one capable of generating megakaryocytes, erythroid cells, and the *GATA1*-expressing myeloid cell types (basophils/mast cells and eosinophils), here desig-

nated as EMPP (erythroid/megakaryocyte-primed MPP; corresponding to the CD131<sup>+</sup> CMP). As discussed above, in both of these populations, myeloid potentials colocalize with other potentials at the single-cell level, demonstrating true oligopotentiality. These observations are therefore consistent with an early lineage bifurcation that generates *GATA1*<sup>+</sup> and *GATA1*<sup>-</sup> progenitor domains (fig. S7), with EMPPs and LMPPs, respectively, defining the entry points, similarly to what has been proposed for murine hematopoiesis (14). However, our data do not exclude that direct commitment to individual lineages also occurs, a notion supported by the identification of murine HSCs that are fate restricted to platelet lineage output (30–32). Last, LMPPs and GMPs have been identified as target cells for transformation of the neutrophil/monocyte lineages (17), and it will now be relevant to investigate the role of the herein identified eosinophil and basophil/mast cell lineage-restricted progenitor populations in sustaining malignancies of these lineages.

## MATERIALS AND METHODS

### Study design

The aim of this study was to determine heterogeneity within the CMP population of healthy adult human bone marrow cells and to prospectively isolate subpopulations for investigating their lineage potentials and thus providing evidence for independent pathways toward different myeloid cell types. Progenitor cells were analyzed by gene expression and fluorescence-activated cell sorting (FACS), and cell potentials were tested by *in vitro* cultures. Sample size, replicates, number of experiments, statistical analysis, and donor sample information are specified in figure legends and in this section.

### Human bone marrow cells

Bone marrow samples were from AllCells or taken from healthy male volunteers at the age of 21 to 29 years old, who provided written informed consent in accordance with local guidelines established by and with the approval of the local Ethics Committee of the Cities of Copenhagen and Frederiksberg. Mononuclear cells were isolated using Ficoll density gradient. Cryopreserved mononuclear cells were thawed and processed for flow cytometry as previously described (33). Where possible and relevant, experiments were repeated with cells from at least two donors (table S3).

### Flow cytometry

For flow cytometry and cell sorting, a BD LSRFortessa, LSR II, FACSAria II, FACSAria III, and FACSAria Fusion (BD Biosciences) were used. FlowJo analysis software was used for subsequent data analysis. All antibody stainings were preceded by incubation with human FcR Blocking Reagent (Miltenyi Biotec, 130-059-901) at a 1:10 dilution. For antibody stainings that included purified CD131, cells were first stained with purified CD131, followed by BV421 anti-mouse immunoglobulin G1 (IgG1). The cells were then resuspended in purified mouse IgG (500 µg/ml) (PMP01X, Bio-Rad), and after 5 to 10 min, an equal volume containing relevant antibodies was added. In all flow experiments, 7-aminoactinomycin D (7-AAD) (40037, Biotium) was used at a final concentration of 1 µg/ml to exclude dead cells. Antibodies, suppliers, and dilutions used are listed in table S4. Populations were defined as follows: Lineage cocktail: CD2, CD3, CD4, CD7, CD8, CD10, CD11b, CD14, CD19, CD20, CD56, and CD235a. HSC/MPP: LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>, LMPP: LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>+</sup>, CMP: LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD45RA<sup>-</sup>CD123<sup>+</sup>,

MEP: LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD45RA<sup>-</sup>CD123<sup>-</sup>, and GMP: LIN<sup>-</sup>CD34<sup>+</sup>CD38<sup>+</sup>CD45RA<sup>+</sup>CD123<sup>+</sup>. Cultured cells: monocytes: CD14<sup>+</sup>CD15<sup>-</sup>, neutrophils: CD14<sup>+</sup>CD15<sup>+/lo</sup>Siglec8<sup>-</sup>CCR3<sup>-</sup>CD117<sup>-</sup>FcεR1α<sup>-</sup>, mast cells: CD14<sup>-</sup>CD15<sup>-/lo</sup>Siglec8<sup>-</sup>CD117<sup>+</sup>FcεR1α<sup>+</sup>, basophils: CD14<sup>-</sup>CD15<sup>-/lo</sup>Siglec8<sup>-</sup>CD117<sup>-</sup>FcεR1α<sup>+</sup>, eosinophils: CD14<sup>-</sup>CD15<sup>-/lo</sup>Siglec8<sup>+</sup>CCR3<sup>+</sup>, B cells: CD45<sup>+</sup>CD14<sup>-</sup>CD15<sup>-</sup>CD19<sup>+</sup>CD56<sup>-</sup>, NK cells: CD45<sup>+</sup>CD14<sup>-</sup>CD15<sup>-</sup>CD19<sup>-</sup>CD56<sup>+</sup>, and T cells: CD45<sup>+</sup>CD7<sup>hi</sup>.

### Generation of complementary DNA libraries using Smart-seq2 protocol

Single cells or bulk cells (100 cells) were isolated by FACS into 96-well plates (Thermo Fisher Scientific, single cells) or Eppendorf tubes (bulk) containing 4 μl of a lysis mix, consisting of 0.2% Triton X-100, 4 U of ribonuclease (RNase) inhibitor (Takara), 2.5 μM oligo-dT<sub>30</sub>VN (Biomers), and 2.5 mM deoxyribonucleotide triphosphate (dNTP) mix (Fermentas). This was stored at -80°C for up to 1 week. For the reverse transcriptase, 6 μl of the following mix was added: 2 μl of Superscript II first-strand buffer, 0.5 μl of dithiothreitol (100 mM), 2 μl of betaine (5 M), 0.1 μl of MgCl<sub>2</sub> (1 M), 0.25 μl of RNase inhibitor (40 U/μl), 0.1 μl of template switching oligonucleotide (TSO) (100 μM), 0.25 μl of Superscript II reverse transcriptase (200 U/μl), and 0.8 μl (single cells) or 0.4 μl (bulk) of water. After reverse transcriptase, 15 μl was added containing 12.5 μl of KAPA HiFi HS Ready Mix (2×) and 0.125 μl of ISPCR primers (10 μM). The thermal conditions for reverse transcriptase and preamplification were according to the original Smart-seq2 protocol (18). The number of cycles used for PCR amplification was 22 for single cells and 16 for bulk samples. After PCR amplification, complementary DNA (cDNA) libraries were purified using AMPure XP magnetic beads according to the manufacturer's instructions. After purification, the libraries were resuspended in 17.5 μl of elution buffer (Qiagen) and stored at -20°C. Quality and concentration of the cDNA libraries generated was assessed using High-Sensitivity Bioanalyzer (Agilent).

### Illumina library preparation and sequencing

From single cells, 1.25 μl of cDNA was used, and from bulk samples, 0.7 ng of cDNA was tagged using the Nextera XT DNA Sample Preparation kit (Illumina) according to the manufacturer's instructions, except that one-fourth of the volumes indicated were used. Purification of the product was done with a 1:1 ratio of AMPure XP beads, with a final elution in 17.5 μl in resuspension buffer provided by the Nextera kit. Samples were loaded on a High Sensitivity DNA chip (Agilent Technologies) to check the size and quality of the indexed library, and the concentration was measured with a Quant-iT PicoGreen double-stranded DNA Assay Kit (Invitrogen) on a CLARIOstar (BMG LABTECH) or with Qubit high-sensitivity DNA kit (Invitrogen). Libraries were pooled to a final concentration of 4 nM and were sequenced with Illumina NextSeq 500 (76-base pair single-end read) after preparation according to the manufacturer's instructions. For the bulk samples, gene expression data were accumulated from two sequence runs.

### Single-cell RNA sequencing analysis

Short reads were aligned to the human (GRCh37/hg19) genome using TopHat (v2.0.13) (34). The mapping parameter "-g 1" was used to allow one alignment to the reference for a given read. A total of 20 cells with <500,000 mapped reads, with the percentage of mapping to the mitochondrial chromosome >10% or <2000 detected genes, were excluded from further analysis. A total of 296 cells (237 CMPs,

32 GMPs, and 27 MEPs) fulfilled these criteria. The featureCounts (35) software was used to count reads on the basis of the RefSeq gene model. Counts per million (CPM) values were calculated using a script in R. We normalized the CPM values into log<sub>2</sub> (CPM) scale and set up the limit of detection at 1 CPM. Log<sub>2</sub> scale of genes expressed at <1 CPM was set to zero. We selected 4731 predicted variable genes based on a simple noise model using the Lowess model of the average gene expression and the coefficient of variation (CV) (36). We then performed PCA. The top 100 genes, without the cell cycle-related genes, with the highest absolute correlation coefficient (PCA component loadings, one of the first three components) were used for the t-SNE analysis. The Rtsne package, a Barnes-Hut implementation, was used to perform the t-SNE analysis. We used the kernel density estimation function in the kernel smoothing package to perform the kernel smoothing density estimation using the basis of t-SNE analysis results (from dimensions 1 and 2) to visualize high-density regions of data points. We estimated seven high-density regions. We then used the centers of each high-density region as the input for the *k*-means clustering analysis (*k* = 7) to assign the cell clusters. Differentially expressed genes analysis was performed using nonparametric Wilcoxon test for the expression level and Fisher's exact test for the expressing cell frequency. *P* values generated from both tests were then combined using Fisher's method and were adjusted using Benjamini-Hochberg. Differentially expressed genes were selected on the basis of the absolute log<sub>2</sub> fold change of >2 and the adjusted *P* < 0.05. Core cell cycle genes from S and G<sub>2</sub>M phases were previously described (37).

### Bulk RNA sequencing analysis

The same mapping and gene expression quantification procedures were performed as described in the single-cell analysis. Reads per kilobase of transcript per million mapped reads (RPKM) values were calculated using a script in R. A total of 1868 genes were selected for PCA on the basis of the high variation of gene expression across populations [analysis of variance (ANOVA) with adjusted *P* < 0.05 and CV equal to or larger than 0.3]. Differentially expressed gene analysis was performed on the read count using DESeq2 (38).

### Gene set enrichment analysis

Gene set enrichment was performed using GSEA v3.0 (<http://software.broadinstitute.org/gsea/index.jsp>) and human cell type-specific gene sets for neutrophils, monocytes, macrophages, basophils, and eosinophils (<http://nano.princeton.edu/standard/>). Human gene sets for NK, B, and T cells were generated by taking the top 100 genes identified as specific for these cell types by Novershtern *et al.* (39) (table S5). The monocyte gene set was >500 genes and was therefore divided into two sets and tested individually with similar results. A representative analysis is shown.

### Myeloid cell cultures

To test myeloid potential, single cells were cultured in Terasaki plates in 20 μl of Iscove's Modified Dulbecco's Medium (IMDM) with L-glutamine (Gibco), 20% HyClone Defined Fetal Bovine Serum (SH30070.03, GE Healthcare), penicillin-streptomycin (Invitrogen), and 0.1 μM β-mercaptoethanol (Sigma-Aldrich), supplemented with hSCF (20 ng/ml), hFlt3L (20 ng/ml), hIl-3 (20 ng/ml), hIl-5 (50 ng/ml), hIl-6 (20 ng/ml), hGM-CSF (50 ng/ml), and hG-CSF (20 ng/ml). For bulk cultures, 50 to 150 cells were cultured in 400 μl of the same culture medium in 48-well plates. For combined erythroid,



megakaryocyte, and myeloid readout, single cells were cultured in round-bottom 96-well plates in 50  $\mu$ l of StemSpan (STEMCELL Technologies) with hSCF (20 ng/ml), hFlt3L (20 ng/ml), hIL-3 (20 ng/ml), hIL-5 (50 ng/ml), hIL-6 (20 ng/ml), hGM-CSF (50 ng/ml), hG-CSF (20 ng/ml), hLDL (40  $\mu$ g/ml; Sigma-Aldrich), erythropoietin (0.5 U/ml), and thrombopoietin (100 ng/ml). Cells were cultured at 37°C, 5% CO<sub>2</sub>. Cytokines, suppliers, and concentrations used are found in table S6. Cytospins were prepared with a Shandon Cytospin at 1000 rpm with low acceleration, followed by May-Grünwald-Giemsa stain (VWR).

### Lymphoid cultures

For NK and B cell potential, cultures of 30 or 5 cells were grown in gelatin-coated 24-well plates seeded with MS-5 feeder cells in  $\alpha$ -minimum essential medium ( $\alpha$ -MEM), GlutaMAX Supplement (32561, Gibco/Life Technologies) supplemented with 10% HyClone Defined Fetal Bovine Serum (SH30070.03, GE Healthcare), 1% penicillin-streptomycin, 1% L-glutamine, hSCF (20 ng/ml), hFLT3L (10 ng/ml), hIL-2 (10 ng/ml), and hIL-15 (10 ng/ml). Half of the medium was changed every week. Cultures were analyzed by flow cytometry after 3 weeks. Cultures with more than eight cells in the respective gates were considered positive for NK or B cells. For T cell potential, cultures of 30 or 5 sorted cells were grown in gelatin-coated 24-well plates seeded with OP9-hDL1 feeder cells in freshly prepared  $\alpha$ -MEM medium (Gibco/Thermo Fisher Scientific, 12000-063) with 20% HyClone Defined Fetal Bovine Serum, 1% penicillin-streptomycin, 1% L-glutamine, hSCF (10 ng/ml), hFLT3L (5 ng/ml), and hIL-7 (5 ng/ml). Every week, cells were transferred to new plates with feeder cells and fresh medium. Cultures were analyzed by flow cytometry after 5 weeks. Cultures with more than 50 cells in the respective gates were considered positive T cells.

### Quantitative PCR

For single cells, the CellsDirect One-Step qRT-PCR kit (Life Technologies, 11753-100) was used according to the manufacturer's protocol for preparation and amplification of cDNA. The BioMark 192.24 Dynamic Array platform (Fluidigm) and TaqMan assays (table S7) were used according to the manufacturers' instructions.  $\Delta$ Ct values were zero-centered for each gene by subtraction of the mean value of all positive cells for the gene. These normalized values were used to generate a heat map using the web-based tool Morpheus (<https://software.broadinstitute.org/morpheus/>). For gene expression of cultured cells, medium of the cultures was removed and the cells were resuspended in 15  $\mu$ l of CellDirect 2 $\times$  Reaction Mix containing SUPERase In RNase Inhibitor (0.2 U/ $\mu$ l; AM2694) and placed at -80°C. Preparation of cDNA and preamplification were done with 2.5  $\mu$ l of the lysed cells, using 22 preamplification cycles in a total volume of 5  $\mu$ l. This was diluted 50 times for further gene expression using the BioMark 192.24 Dynamic Array platform (Fluidigm) and TaqMan assays according to the manufacturers' instructions.  $\Delta$ Ct values relative to *HPRT* [ $\text{Ct}(HPRT) - \text{Ct}(\text{Gene})$ ] were zero-centered for each gene by subtraction of the mean value of all positive cells for the gene. Single-cell-derived cultures were divided with one-half used for cytospin and one-half for quantitative PCR. A positive lineage readout was defined as both morphologically differentiated cells and signature gene expression being observed, with positive readout defined as  $\text{Ct}(IRF8) \leq \text{Ct}(HPRT)$  for monocytes,  $\text{Ct}(CSF3R) \leq \text{Ct}(HPRT)$  for neutrophils,  $\text{Ct}(HDC) \leq \text{Ct}(HPRT)$  for mast cells/basophils,  $\text{Ct}(IL5RA) \leq \text{Ct}(HPRT)$ , and  $\text{Ct}(EPX) \leq \text{Ct}(HPRT)$  for eosinophils;

and  $\text{Ct}(KLF1) \leq \text{Ct}(HPRT) + 4$ ,  $\text{Ct}(AHSP) \leq \text{Ct}(HPRT) + 4$ , and  $\text{Ct}(CA) \leq \text{Ct}(HPRT) + 4$  for erythroid cells.

### Statistical analysis

Significance of differences in distribution of lineage potentials in single cells was analyzed using the  $\chi^2$  test. For the identification of differentially expressed genes from single-cell RNA sequencing data, a combination of the Wilcoxon and Fisher's exact test was used; combined significance was calculated using Fisher's method and *P* values were adjusted using Benjamini-Hochberg.

### SUPPLEMENTARY MATERIALS

[immunology.sciencemag.org/cgi/content/full/4/35/eaau7148/DC1](http://immunology.sciencemag.org/cgi/content/full/4/35/eaau7148/DC1)

Fig. S1. Candidate cell surface markers for separation of myeloid progenitor subsets.

Fig. S2. Expression of CD131, CD114, and CD133 on hematopoietic stem and progenitor subsets.

Fig. S3. Single-cell heat map of clustered CMPs.

Fig. S4. Influence of culture conditions on lineage readout.

Fig. S5. Gating strategy for isolation of human stem and progenitor cell subsets.

Fig. S6. Myeloid and lymphoid potential of progenitor populations is reflected in their gene expression.

Fig. S7. Proposed model of the human hematopoietic hierarchy.

Fig. S8. Expression of CD11b on myeloid cell types generated in vitro.

Table S1. Cloning efficiencies of myeloid progenitors.

Table S2. Genes differentially expressed between CMP clusters.

Table S3. Donor samples.

Table S4. Antibodies used for flow cytometry and FACS.

Table S5. Gene sets used for GSEA.

Table S6. Cytokines used for progenitor assays.

Table S7. TaqMan probes for qRT-PCR.

### REFERENCES AND NOTES

1. L. Koenderman, W. Buurman, M. R. Daha, The innate immune response. *Immunol. Lett.* **162**, 95–102 (2014).
2. M. G. Manz, S. Boettcher, Emergency granulopoiesis. *Nat. Rev. Immunol.* **14**, 302–314 (2014).
3. H. Sadrzadeh, O. Abdel-Wahab, A. T. Fathi, Molecular alterations underlying eosinophilic and mast cell malignancies. *Discov. Med.* **12**, 481–493 (2011).
4. A. Ortega-Gómez, M. Perretti, O. Soehnlein, Resolution of inflammation: An integrated view. *EMBO Mol. Med.* **5**, 661–674 (2013).
5. D. I. Gabrilovich, S. Ostrand-Rosenberg, V. Bronte, Coordinated regulation of myeloid cells by tumours. *Nat. Rev. Immunol.* **12**, 253–268 (2012).
6. S. H. Orkin, L. I. Zon, Hematopoiesis: An evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
7. D. Metcalf, Lineage commitment and maturation in hematopoietic cells: The case for extrinsic regulation. *Blood* **92**, 345–347; discussion 352 (1998).
8. K. Akashi, D. Traver, T. Miyamoto, I. L. Weissman, A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193–197 (2000).
9. M. G. Manz, T. Miyamoto, K. Akashi, I. L. Weissman, Prospective isolation of human clonogenic common myeloid progenitors. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11872–11877 (2002).
10. H. Iwasaki, S.-i. Mizuno, R. Mayfield, H. Shigematsu, Y. Arinobu, B. Seed, M. F. Gurish, K. Takatsu, K. Akashi, Identification of eosinophil lineage-committed progenitors in the murine bone marrow. *J. Exp. Med.* **201**, 1891–1897 (2005).
11. Y. Arinobu, H. Iwasaki, M. F. Gurish, S.-i. Mizuno, H. Shigematsu, H. Ozawa, D. G. Tenen, K. F. Austen, K. Akashi, Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18105–18110 (2005).
12. Y. Mori, H. Iwasaki, K. Kohno, G. Yoshimoto, Y. Kikushige, A. Okeda, N. Uike, H. Niuro, K. Takenaka, K. Nagafuji, T. Miyamoto, M. Harada, K. Takatsu, K. Akashi, Identification of the human eosinophil lineage-committed progenitor: Revision of phenotypic definition of the human common myeloid progenitor. *J. Exp. Med.* **206**, 183–193 (2009).
13. O. Schmetzer, P. Valentin, M. K. Church, M. Maurer, F. Siebenhaar, Murine and human mast cell progenitors. *Eur. J. Pharmacol.* **778**, 2–10 (2016).
14. R. Drissen, N. Buza-Vidas, P. Woll, S. Thongjuea, A. Gambardella, A. Giustacchini, E. Mancini, A. Zriwil, M. Lutteropp, A. Grover, A. Mead, E. Sitnicka, S. E. W. Jacobsen, C. Nerlov, Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.* **17**, 666–676 (2016).

15. L. I. Zon, Y. Yamaguchi, K. Yee, E. A. Albee, A. Kimura, J. C. Bennett, S. H. Orkin, S. J. Ackerman, Expression of mRNA for the GATA-binding proteins in human eosinophils and basophils: Potential role in gene transcription. *Blood* **81**, 3234–3241 (1993).
16. M. Babina, Y. Schülke, L. Kirchof, S. Guhl, R. Franke, S. Böhm, T. Zuberbier, B. M. Henz, A. F. Gombart, The transcription factor profile of human mast cells in comparison with monocytes and granulocytes. *Cell. Mol. Life Sci.* **62**, 214–226 (2005).
17. N. Goardon, E. Marchi, A. Atzberger, L. Quek, A. Schuh, S. Soneji, P. Woll, A. Mead, K. A. Alford, R. Rout, S. Chaudhury, A. Gilkes, S. Knapper, K. Beldjord, S. Begum, S. Rose, N. Geddes, M. Griffiths, G. Standen, A. Sternberg, J. Cavenagh, H. Hunter, D. Bowen, S. Killick, L. Robinson, A. Price, E. Macintyre, P. Virgo, A. Burnett, C. Craddock, T. Enver, S. E. W. Jacobsen, C. Porcher, P. Vyas, Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell* **19**, 138–152 (2011).
18. S. Picelli, Å. K. Björklund, O. R. Faridani, S. Sagasser, G. Winberg, R. Sandberg, Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
19. L. van der Maaten, G. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
20. A. Görgens, S. Radtke, M. Mollmann, M. T. Cross, J. Dürig, P. A. Horn, B. Giebel, Revision of the human hematopoietic tree: Granulocyte subtypes derive from distinct hematopoietic lineages. *Cell Rep.* **3**, 1539–1552 (2013).
21. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
22. E. Mancini, A. Sanjuan-Pla, L. Luciani, S. Moore, A. Grover, A. Zay, K. D. Rasmussen, S. Luc, D. Bilbao, D. O'Carroll, S. E. Jacobsen, C. Nerlov, FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *EMBO J.* **31**, 351–365 (2012).
23. K. Keeshan, Y. He, B. J. Wouters, O. Shestova, L. Xu, H. Sai, C. G. Rodriguez, I. Maillard, J. W. Tobias, P. Valk, M. Carroll, J. C. Aster, R. Delwel, W. S. Pear, Tribbles homolog 2 inactivates C/EBP $\alpha$  and causes acute myelogenous leukemia. *Cancer Cell* **10**, 401–411 (2006).
24. T. Naiki, E. Saijoui, Y. Miyaoka, K. Sekine, A. Miyajima, TRB2, a mouse Tribbles ortholog, suppresses adipocyte differentiation by inhibiting AKT and C/EBP $\beta$ . *J. Biol. Chem.* **282**, 24075–24082 (2007).
25. D. Karamitros, B. Stoilova, Z. Aboukhalil, F. Hamey, A. Reinisch, M. Samitsch, L. Quek, G. Otto, E. Repapi, J. Doondeea, B. Usukhbayar, J. Calvo, S. Taylor, N. Goardon, E. Six, F. Pflumio, C. Porcher, R. Majeti, B. Göttgens, P. Vyas, Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat. Immunol.* **19**, 85–97 (2018).
26. L. Velten, S. F. Haas, S. Raffel, S. Blaszkiewicz, S. Islam, B. P. Hennig, C. Hirche, C. Lutz, E. C. Buss, D. Nowak, T. Boch, W.-K. Hofmann, A. D. Ho, W. Huber, A. Trumpp, M. A. G. Essers, L. M. Steinmetz, Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).
27. S. Zheng, E. Papalexii, A. Butler, W. Stephenson, R. Satija, Molecular transitions in early progenitors during human cord blood hematopoiesis. *Mol. Syst. Biol.* **14**, e8041 (2018).
28. F. Notta, S. Zandi, N. Takayama, S. Dobson, O. I. Gan, G. Wilson, K. B. Kaufmann, J. McLeod, E. Laurenti, C. F. Dunant, J. D. McPherson, L. D. Stein, Y. Dror, J. E. Dick, Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016).
29. J. Adolfsson, R. Månsson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. W. Yang, O. J. Borge, L. A. Thoren, K. L. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, S. E. W. Jacobsen, Identification of Flt3+ lympho-myeloid stem cells lacking erythromegakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295–306 (2005).
30. A. Grover, A. Sanjuan-Pla, S. Thongjuea, J. Carrelha, A. Giustacchini, A. Gambardella, I. Macaulay, E. Mancini, T. C. Luis, A. Mead, S. E. W. Jacobsen, C. Nerlov, Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat. Commun.* **7**, 11075 (2016).
31. J. Carrelha, Y. Meng, L. M. Kettyle, T. C. Luis, R. Norfo, V. Alcolea, H. Boukarabila, F. Grasso, A. Gambardella, A. Grover, K. Högstrand, A. M. Lord, A. Sanjuan-Pla, P. S. Woll, C. Nerlov, S. E. W. Jacobsen, Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* **554**, 106–111 (2018).
32. A. E. Rodriguez-Fraticelli, S. L. Wolock, C. S. Weinreb, R. Panero, S. H. Patel, M. Jankovic, J. Sun, R. A. Calogero, A. M. Klein, F. D. Camargo, Clonal analysis of lineage fate in native haematopoiesis. *Nature* **553**, 212–216 (2018).
33. P. S. Woll, U. Kjällquist, O. Chowdhury, H. Doolittle, D. C. Wedge, S. Thongjuea, R. Erlandsson, M. Ngara, K. Anderson, Q. Deng, A. J. Mead, L. Stenson, A. Giustacchini, S. Duarte, E. Giannoulatou, S. Taylor, M. Karimi, C. Scharenberg, T. Mortera-Blanco, I. C. Macaulay, S.-A. Clark, I. Dybedal, D. Josefsen, P. Fenaux, P. Hokland, M. S. Holm, M. Cazzola, L. Malcovati, S. Tauro, D. Bowen, J. Boultood, A. Pellagatti, J. E. Pimanda, A. Unnikrishnan, P. Vyas, G. Göhring, B. Schlegelberger, M. Tobiaasson, G. Kvalheim, S. N. Constantinescu, C. Nerlov, L. Nilsson, P. J. Campbell, R. Sandberg, E. Papaemmanuil, E. Hellström-Lindberg, S. Linnarsson, S. E. W. Jacobsen, Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. *Cancer Cell* **25**, 794–808 (2014).
34. D. Kim, G. Perete, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
35. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
36. A. Giustacchini, S. Thongjuea, N. Barkas, P. S. Woll, B. J. Pavinelli, C. A. G. Booth, P. Sopp, R. Norfo, A. Rodriguez-Meira, N. Ashley, L. Jamieson, P. Vyas, K. Anderson, Å. Segerstolpe, H. Qian, U. Olsson-Strömberg, S. Mustjoki, R. Sandberg, S. E. W. Jacobsen, A. J. Mead, Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat. Med.* **23**, 692–702 (2017).
37. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth II, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A.-C. Villani, C. M. Johannesson, A. Y. Andreev, E. M. Van Allen, M. Bertagnoli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
38. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
39. N. Novershtern, A. Subramanian, L. N. Lawton, R. H. Mak, W. N. Haining, M. E. McConkey, N. Habib, N. Yosef, C. Y. Chang, T. Shay, G. M. Frampton, A. C. B. Drake, I. Leskov, B. Nilsson, F. Pfeffer, D. Dombkowski, J. W. Evans, T. Liefeld, J. S. Smutko, J. Chen, N. Friedman, R. A. Young, T. R. Golub, A. Regev, B. L. Ebert, Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).

**Acknowledgments:** We thank F. Pflumio and E. Six for providing OP9-hDL1 stromal cells; B. Stoilova and D. Karamitros for advice on lymphoid culture experiments; and P. Vyas, I. Roberts, and the Nerlov laboratory for helpful discussions. The assistance of the WIMM flow cytometry facility, WIMM single cell facility, and WIMM sequencing facility is gratefully acknowledged. **Funding:** This work was supported by the MRC (G0701761, G0900892, and MC\_UU\_12009/7 to C.N.) and the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The WIMM Single Cell Core Facility is supported by the MRC MHU (MC\_UU\_12009), the Oxford Single Cell Biology Consortium (MR/M00919X/1) and the WT-ISSF (097813/Z/11/B#) funding, and the WIMM Strategic Alliance award nos. G0902418 and MC\_UU\_12025]. The WIMM FACS Core Facility is supported by the MRC HIU, MRC MHU (MC\_UU\_12009), NIHR Oxford BRC and the John Fell Fund (131/030 and 101/517), the EPA fund (CF182 and CF170), and WIMM Strategic Alliance awards (G0902418 and MC\_UU\_12025). **Author contributions:** R.D. and C.N. designed the experiments. R.D. performed the experiments. S.T. performed bioinformatics analysis. K.T.-M. provided bone marrow samples. R.D., S.T., and C.N. wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials. RNA sequencing data are available from the Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) under the accession number GSE113046.

Submitted 9 July 2018  
Accepted 27 March 2019  
Published 24 May 2019  
10.1126/sciimmunol.aau7148

**Citation:** R. Drissen, S. Thongjuea, K. Theilgaard-Mönch, C. Nerlov, Identification of two distinct pathways of human myelopoiesis. *Sci. Immunol.* **4**, eaau7148 (2019).

## Identification of two distinct pathways of human myelopoiesis

Roy Drissen, Supat Thongjuea, Kim Theilgaard-Mönch and Claus Nerlov

*Sci. Immunol.* 4, eaau7148.  
DOI: 10.1126/sciimmunol.aau7148

### Goodbye CMPs

Advances in single-cell analyses continue to profoundly change our understanding of hematopoiesis. Here, using single-cell RNA sequencing in conjunction with functional studies, Drissen *et al.* studied the differentiation potentials of human common myeloid progenitor (CMP) cells. They report that CMPs that have been defined on the basis of cell surface markers are a mixture of at least two cell types with the potential to give rise to either mast cells/basophils/eosinophils or neutrophils/monocytes. In other words, they report that CMPs having the potential to give rise to all myeloid cell types do not exist. Furthermore, they found that the cells in the former group retain the potential to give rise to megakaryocytes and erythrocytes, whereas cells in the latter groups retain lymphoid potential.

#### ARTICLE TOOLS

<http://immunology.sciencemag.org/content/4/35/eaau7148>

#### SUPPLEMENTARY MATERIALS

<http://immunology.sciencemag.org/content/suppl/2019/05/20/4.35.eaau7148.DC1>

#### REFERENCES

This article cites 39 articles, 13 of which you can access for free  
<http://immunology.sciencemag.org/content/4/35/eaau7148#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Immunology* (ISSN 2470-9468) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Immunology* is a registered trademark of AAAS.

Copyright © 2019 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works