



CD66b⁻CD64^{dim}CD115⁻ cells in the human bone marrow represent neutrophil-committed progenitors

Federica Calzetti^{1,8}, Giulia Finotti^{1,8}, Nicola Tamassia^{1,8}, Francisco Bianchetto-Aguilera^{1,8}, Monica Castellucci², Stefania Canè³, Silvia Lonardi⁴, Chiara Cavallini⁵, Alessandro Matte⁶, Sara Gasperini¹, Ilaria Signoretto¹, Fabio Benedetti⁷, Massimiliano Bonifacio⁷, William Vermi⁴, Stefano Ugel³, Vincenzo Bronte³, Cristina Tecchio⁷, Patrizia Scapini¹ and Marco A. Cassatella¹✉

Here we report the identification of human CD66b⁻CD64^{dim}CD115⁻ neutrophil-committed progenitor cells (NCPs) within the SSC^{lo}CD45^{dim}CD34⁺ and CD34^{dim/-} subsets in the bone marrow. NCPs were either CD45RA⁺ or CD45RA⁻, and in vitro experiments showed that CD45RA acquisition was not mandatory for their maturation process. NCPs exclusively generated human CD66b⁺ neutrophils in both in vitro differentiation and in vivo adoptive transfer experiments. Single-cell RNA-sequencing analysis indicated NCPs fell into four clusters, characterized by different maturation stages and distributed along two differentiation routes. One of the clusters was characterized by an interferon-stimulated gene signature, consistent with the reported expansion of peripheral mature neutrophil subsets that express interferon-stimulated genes in diseased individuals. Finally, comparison of transcriptomic and phenotypic profiles indicated NCPs represented earlier neutrophil precursors than the previously described early neutrophil progenitors (eNePs), proNeus and COVID-19 proNeus. Altogether, our data shed light on the very early phases of neutrophil ontogeny.

Studies performed by single-cell sorting, transcriptional analysis, single-cell transplants and clonal tracking analysis^{1–9} have challenged the classical hierarchical model of hematopoiesis^{10,11}. It has been ascertained that hematopoiesis occurs as a continuous process along developmental trajectories following early erythroid/megakaryocyte/eosinophil/basophil and lympho/myeloid lineages to generate mature, terminally differentiated cells. Transcriptomic profiling has shown that intermediate compartments of progenitors, such as common myeloid progenitors (CMPs) or granulocyte-macrophage progenitors (GMPs)¹⁰, are largely composed of clusters of cells displaying lineage-selective commitment. However, even though single-cell RNA-sequencing (scRNA-seq) experiments reveal the heterogeneity of various cell subpopulations¹², flow-sorting and immunophenotyping approaches are still needed for their identification.

Conventional GMPs (cGMPs) represent the most restricted subset of myeloid progenitors in humans^{13,14}. Based on flow cytometry approaches¹⁵, cGMPs have been fractionated into three main compartments: granulocyte-monocyte-dendritic cell (DC) progenitors (GMDFPs), which generate CD66b⁺ granulocytes, CD14⁺ monocytes, conventional DCs (cDC1s, cDC2s) and plasmacytoid DCs (pDCs); monocyte-DC progenitors (MDPs), which generate DCs and CD14⁺ monocytes; and common dendritic progenitors (CDPs), which are exclusively committed to DCs. A different panel of markers which includes CD64 identified, within cGMPs, the common monocyte progenitors (cMoPs), which exclusively differentiate into pre-monocytes and subsequently CD14⁺ monocytes, and the

revised GMPs (rGMPs), which generate only CD66b⁺ granulocytes and CD14⁺ monocytes¹⁶. Although neutrophil precursors are included in cGMPs^{3,4,6,17–19}, little progress has been made to characterize the ontogeny of human neutrophils. A CD34⁺CD115⁻CD64⁺ subpopulation that generates granulocytes was identified in the fetal bone marrow (BM)²⁰, without discriminating between eosinophil, basophil or neutrophil potential. Similarly, the phenotypic and/or morphological features of the GMDFPs¹⁵ and rGMPs¹⁶-derived CD66b⁺ granulocytes were not pursued in detail. Various human neutrophil precursors, preNeus²¹, human Neutrophil Progenitors (hNePs)²², proNeus²³ and early neutrophil progenitors (eNePs)²⁴, have been described recently. However, they all express CD66b and CD15 and thus resemble the promyelocyte (PM), or more mature, neutrophil precursors. In contrast, the CD34⁺Ly6C⁺CD115⁻ proNeus in mice are found within GMPs²³, and represent the earliest identifiable precursors along the neutrophil maturation trajectory.

Here we report the identification, phenotypic characterization, single-cell transcriptomic analysis, and in vitro and in vivo functions of previously undescribed human uni-lineage CD34⁺ and CD34^{dim/-} NCPs, and show they are upstream of eNePs, proNeus and COVID-19 proNeus^{23–25}.

Results

CD123, CD115 and CD64 identify myeloid progenitors within cGMPs. To identify precursors of human neutrophils at earlier stages than PMs, we analyzed the SSC^{lo}CD45^{dim} cells present

¹Section of General Pathology, Department of Medicine, University of Verona, Verona, Italy. ²Centro Piattaforme Tecnologiche, University of Verona, Verona, Italy. ³Section of Immunology, Department of Medicine, University of Verona, Verona, Italy. ⁴Unit of Pathology, Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy. ⁵Interdepartmental Laboratory of Medical Research, Research Center LURM, University of Verona, Verona, Italy. ⁶Section of Internal Medicine B, Department of Medicine, University of Verona and AOUI Verona, Verona, Italy. ⁷Section of Hematology and Bone Marrow Transplant Unit, Department of Medicine, University of Verona, Verona, Italy. ⁸These authors contributed equally: Federica Calzetti, Giulia Finotti, Nicola Tamassia, Francisco Bianchetto-Aguilera. ✉e-mail: marco.cassatella@univr.it

in the low-density fraction of the BM (BM low-density cells (BM-LDCs)) (Fig. 1a and Extended Data Fig. 1a). SSC^{lo}CD45^{dim} cells were lineage negative (Extended Data Fig. 1b) and included CD34⁺CD38⁻ hematopoietic stem cells (HSCs) as well as CD34⁺ and CD34^{dim/-} myeloid-lymphoid progenitors^{26,27}, with the CD34^{dim/-} cells representing a transitional progenitor stage⁴. To analyze their potential ability to generate neutrophils, we stained BM-LDCs with an antibody panel including CD34, CD38, CD10, CD123, CD45RA, CD64 and CD115 (refs. ^{13,14,20,28}). We identified an SSC^{lo}Lin⁻CD45^{dim}CD10⁻CD38⁺ subset with variable expression of CD34 and CD45RA (Fig. 1a), including CD34⁺CD45RA⁻, CD34⁺CD45RA⁺, CD34^{dim/-}CD45RA⁺ and CD34^{dim/-}CD45RA⁻ cell subsets. The CD34⁺CD45RA⁺ cells were also CD135⁺, and thus represented cGMPs (not shown)¹⁴, while CD34⁺CD45RA⁻ cells (Fig. 1a) included more immature progenitors, such as CMPs and megakaryocyte-erythroid progenitors (MEPs)^{13,14}. Based on expression of CD123, CD115 and CD64, cGMPs were subdivided into five discrete cell populations (Fig. 1a): CDPs; CD64⁺CD115⁺ cells, which were also CLEC12A⁺ and resembled the monocyte-committed progenitor population²⁰, as well as cMoPs¹⁶; CD64⁻CD115⁺ cells, resembling MDPs¹⁵; CD64⁻CD115⁻ cells, resembling the conventional GMDPs¹⁵; and CD64^{dim}CD115⁻ cells, resembling previously described granulocyte progenitors²⁰.

Collectively, the CD123, CD115 and CD64 combination of markers discriminates five different myeloid progenitor subsets within cGMPs.

CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells are neutrophil precursors.

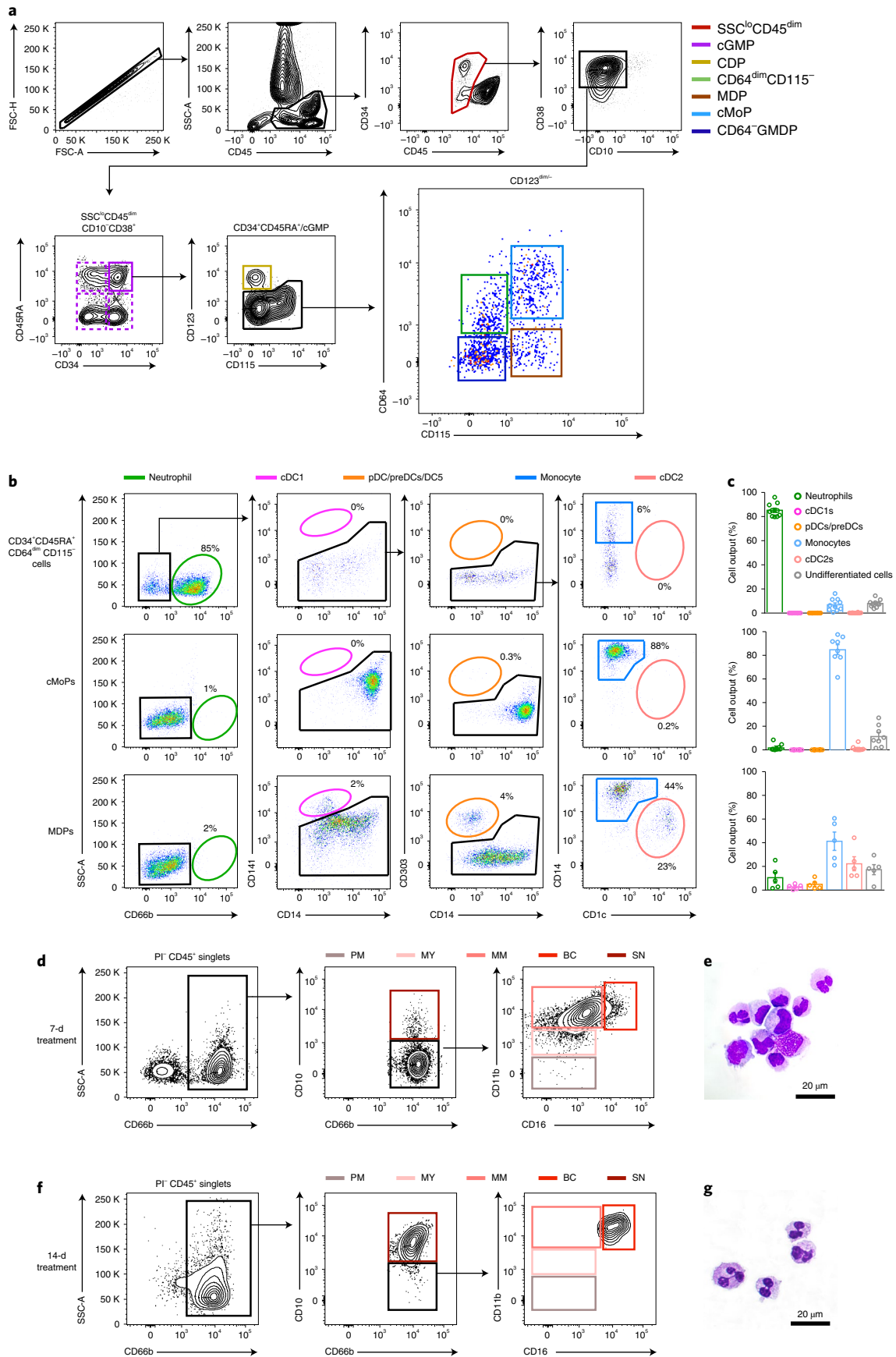
It was reported that neutrophils, but not basophils or eosinophils, derive from cGMPs¹⁷⁻¹⁹. To investigate whether CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells represented neutrophil-restricted progenitors, we cultured them with MS-5 cells²⁹ for 7 d in the presence of stem cell factor (SCF), FMS-like tyrosine kinase 3 ligand (Flt3L) and granulocyte-colony stimulating factor (G-CSF) (hereafter, SFGc). CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells gave rise mostly to CD66b⁺ cells, and very few CD14⁺ monocytes (Fig. 1b,c). By contrast, SFGc-cultured cMoPs differentiated mainly into CD14⁺ monocytes, and minimally to CD66b⁺ cells and DCs (Fig. 1b,c), confirming their commitment to monocytes¹⁶. Conversely, SFGc-treated MDPs gave rise to CD141⁺CD14⁻ cDC1s (which were also CLEC9A⁺), CD303⁺CD14⁻ pDCs/pre-DCs/DC5s (refs. ^{30,31}), CD14^{dim/-}CD1c⁺ cDC2s and CD14⁺ monocytes (Fig. 1b,c), consistent with their predicted heterogeneous composition^{15,32}. Moreover, SFGc fully maintained the viability of progenitor-generated cells after 7 d (Extended Data Fig. 2a) and promoted the expansion of CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells compared with cMoPs and MDPs (Extended Data Fig. 2b), consistent with the highest expression of CD114/G-CSFR in CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells (Extended Data Fig. 2c). Phenotypic (Fig. 1d) and morphologic (Fig. 1e) analysis of CD66b⁺ cells derived from SFGc-treated CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells indicated that they were mostly immature neutrophils, composed of CD11b⁻CD16⁻ PMs, CD11b^{dim/+}CD16⁻ myelocytes (MYs), CD11b⁺CD16⁺ metamyelocytes (MMs), CD11b⁺CD16⁺CD10⁻

band cells (BCs) and CD11b⁺CD16⁺⁺CD10⁺ segmented neutrophils (SNs) (Fig. 2d). Consistently, CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells incubated with SFGc for 14 d generated more mature SNs (Fig. 1f), as confirmed by nucleus morphology (Fig. 1g). When CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells were cultured with SCF, Flt3L and granulocyte-macrophage colony-stimulating factor (GM-CSF) (SFG), which favors monocyte or monocyte and DC differentiation from cMoPs or MDPs, respectively^{16,29}, the main progeny was, again, constituted by neutrophils (Extended Data Fig. 2d), confirming the uni-lineage commitment of CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells. Collectively, these experiments formally identified CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells as NCPs.

A CD64^{dim}CD115⁻ phenotype defines other CD34⁺ or CD34^{dim/-} NCPs. Because NCPs have been described within the CD45RA⁻ compartment^{3,19}, we tested whether other, not yet described, neutrophil-restricted progenitors could be identified within the SSC^{lo}CD45^{dim} region (Fig. 2a). A search for cells expressing the CD64^{dim}CD115⁻ phenotype in all the fractions delimited by the CD34 and CD45RA marker combination identified a total of four subsets of NCPs: CD34⁺CD45RA⁻ NCP1s, CD34⁺CD45RA⁺ NCP2s, CD34^{dim/-}CD45RA⁺ NCP3s and CD34^{dim/-}CD45RA⁻ NCP4s (Fig. 2a). Notably, we were able to identify all four NCPs in cord blood (CB) samples, spleen biopsies (not shown) and in reconstituting BM samples (day +21) from patients undergoing allogeneic hematopoietic stem cell transplant (alloHSCT) for hematological malignancies (Extended Data Fig. 3).

Next, we sorted CD34⁺ and CD34^{dim/-} NCPs to analyze their morphology and developmental potential. Morphologically, NCPs were found to display a high nuclear/cytoplasmic ratio (Fig. 2b) compatible with a progenitor identity. Moreover, granules were evident only in NCP3s and NCP4s, but at lower levels than in PMs (Fig. 2b). In addition, SFGc-treated NCP1s and NCP2s were found to display a comparable (Fig. 2c), but higher, proliferative potential than that exhibited by SFGc-treated NCP3s and NCP4s (Fig. 2c). Notably, NCPs were found to produce almost exclusively neutrophils at different stages of maturation, as determined by their phenotype and morphology (Fig. 2d). By the latter criteria, we found no eosinophils and/or basophils among the SFGc-generated cells (not shown), even if NCPs were incubated in SFG plus IL-3 and/or IL-5 (ref. ¹⁷). Moreover, NCPs generated no DCs when cultured in SFGc (not shown), while NCP2s and NCP3s exclusively generated a few CD14⁺ monocytes (not shown), which we subsequently found to derive from contaminating cMoP and pre-monocyte populations (not shown). In accordance with the results observed from the in vitro cultures, colony forming unit (CFU) assay experiments indicated that the colonies derived from NCPs were almost entirely granulocyte CFU (CFU-G), with a few macrophage CFU (CFU-M) arising from NCP2 and NCP3 only (Extended Data Fig. 4a). Of note, neutrophils generated by NCPs displayed respiratory burst ability (Extended Data Fig. 4b) and phagocytosis capacity (Extended Data Fig. 4c), both comparable to those displayed by peripheral mature neutrophils (PMNs). Finally, we found that NCP3s and NCP4s,

Fig. 1 | Identification of neutrophil precursors within cGMPs. **a**, Flow cytometry strategy for the identification of Lin⁻SSC^{lo}CD45^{dim}CD10⁻CD38⁺ BM-LDCs; CD34⁺CD45RA⁻, CD34^{dim/-}CD45RA⁺ and CD34^{dim/-}CD45RA⁻ subsets; and CD34⁺CD45RA⁺ cGMPs. CD123⁺CD115⁻ CDPs (yellow); CD123^{dim/-}CD64⁺CD115⁺ cMoPs (light blue), CD123^{dim/-}CD64⁻CD115⁺ MDPs (brown); CD123^{dim/-}CD64⁻CD115⁻ GMDPs (CD64⁻ GMDPs, dark blue) and CD123^{dim/-}CD64^{dim}CD115⁻ cells (CD64^{dim}CD115⁻, green) are displayed within cGMPs (one representative experiment out of 20 is shown). **b**, Flow cytometry strategy for the identification of CD66b⁺ neutrophils (green), CD141⁺CD14⁻ cDC1s (magenta), CD303⁺CD14⁻ pDCs/pre-DCs/DC5s (orange), CD14⁺CD1c⁻ monocytes (light blue) and CD14⁻CD1c⁺ cDC2s (pink) generated by CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells, cMoPs or MDPs incubated with SFGc for 7 d. **c**, Bar graphs reporting the percentages of the leukocyte types generated, as defined in **b**, by CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells ($n=10$), cMoPs ($n=8$) or MDPs ($n=5$). Data represent mean \pm s.e.m. **d,f**, Flow cytometry strategy for the identification of CD66b⁺CD10⁻CD11b^{dim/+}CD16⁻ MYs (pink), CD66b⁺CD10⁻CD11b⁺CD16⁺ MMs (light red), CD66b⁺CD10⁻CD11b⁺CD16⁺⁺CD10⁻ BCs (red) and CD66b⁺CD11b⁺CD16⁺⁺CD10⁺ SNs (dark red) derived from CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells cultured with SFGc for 7 d (**d**, $n=15$) and 14 d (**f**, $n=3$). **e,g**, Morphology of neutrophils derived from CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells after 7 d (**e**) and 14 d (**g**) of culture ($n=3$) K, thousand.



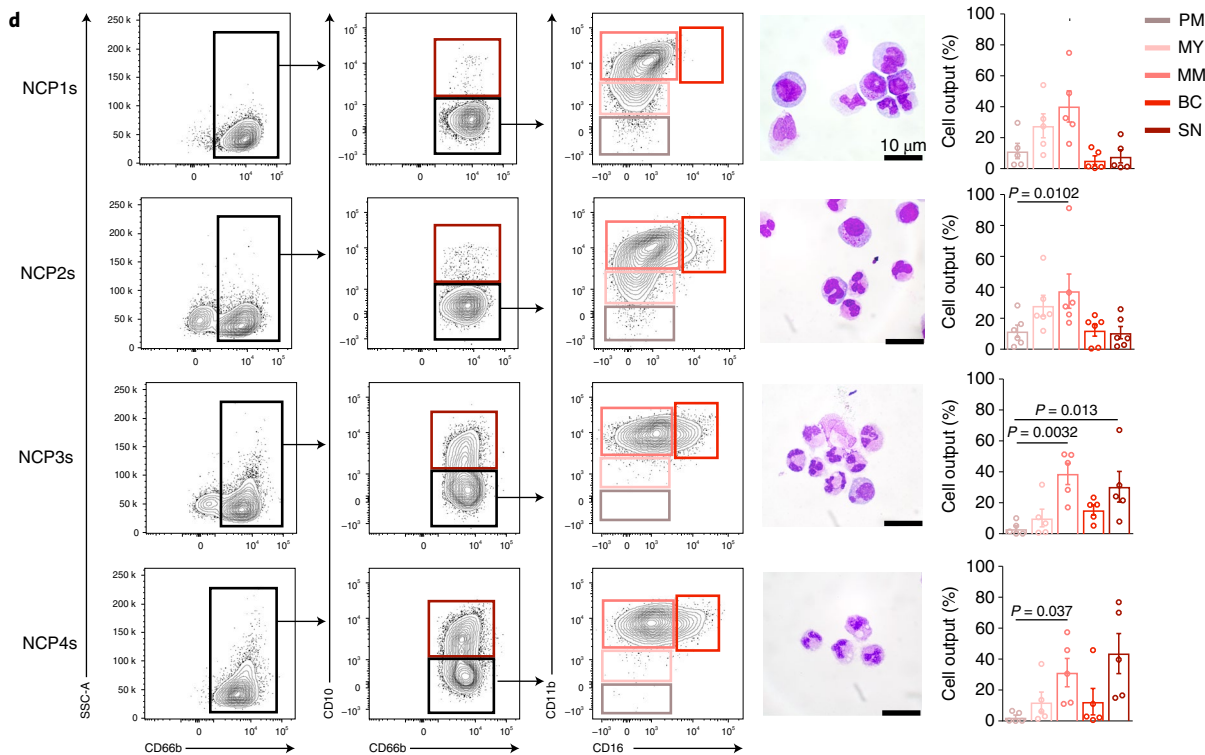
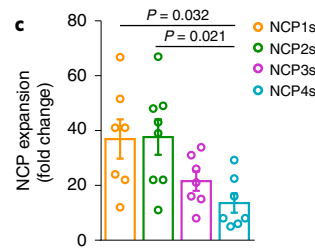
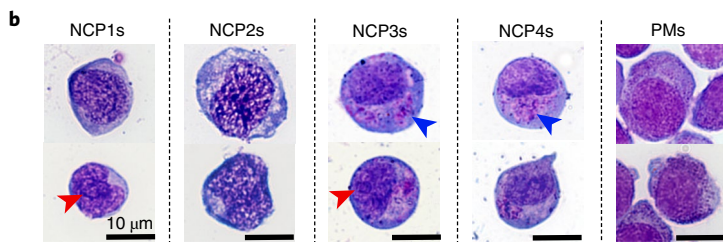
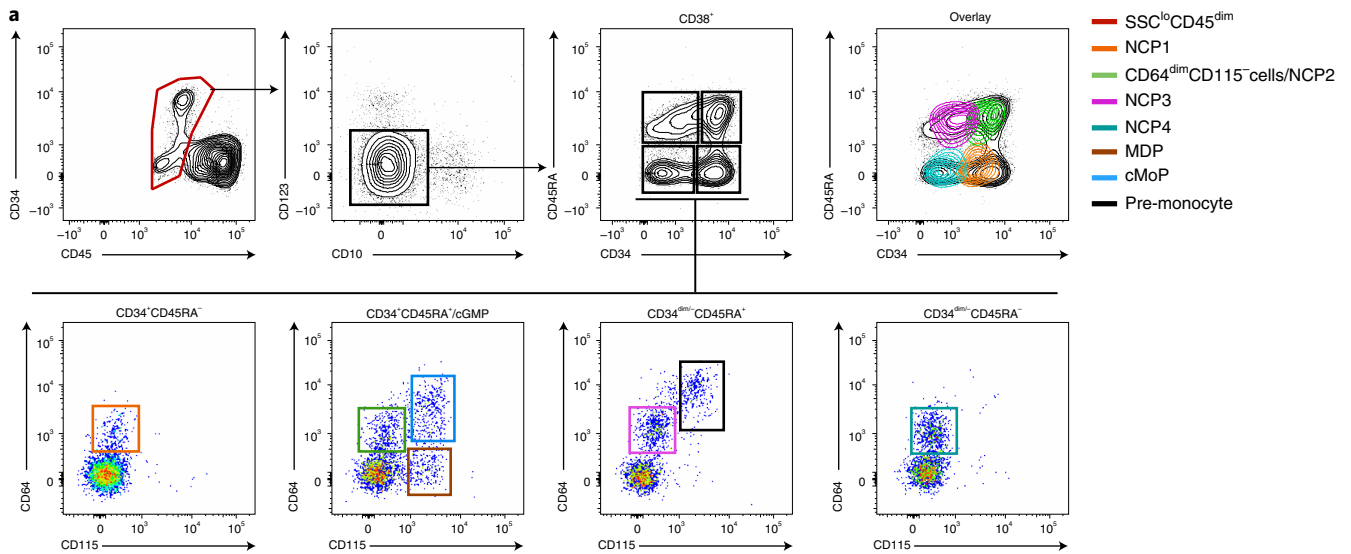


Fig. 2 | Identification of additional CD34⁺ and CD34^{dim/-} neutrophil-restricted progenitors. **a**, Flow cytometry strategy for the identification of CD64^{dim}CD115⁻CD34⁺CD45RA⁻ NCP1s (orange), CD64^{dim}CD115⁻CD34⁺CD45RA⁺ NCP2s (green), CD64^{dim}CD115⁻CD34^{dim/-}CD45RA⁺ NCP3s (magenta), CD64^{dim}CD115⁻CD34^{dim/-}CD45RA⁻ NCP4s (turquoise), cMoPs (light blue), MDPs (brown) and CD34^{dim/-}CD45RA⁺CD64⁺⁺CD115⁺ pre-monocytes (black) within SSC^{lo}CD45^{dim}CD38⁺CD10⁻CD123^{dim/-} BM-LDCs (one representative experiment out of 15 is shown). **b**, Morphology of purified NCP populations and PMs. Red arrows point to visible nucleoli, while blue arrows point to granules ($n=3$). **c**, Bar graphs depicting the fold expansion of purified NCPs treated with SFGc for 7 d (NCP1s, NCP3s and NCP4s, $n=7$; NCP2s, $n=8$). Data represent mean \pm s.e.m. Analysis was performed using one-way ANOVA and Tukey post hoc test. **d**, Flow cytometry strategy for identification of the phenotype of PMs (gray), MYs (pink), MMs (light red), BCs (red) and SNs (dark red) derived from NCPs cultured with SFGc for 7 d (left plots), as well as their morphology (middle panels). Right bar graphs report the percentages of the various neutrophil precursors. Data represent mean \pm s.e.m. (NCP1s, NCP3s and NCP4s, $n=5$; NCP2s, $n=6$). Analysis was performed by using Friedman test and two-sided Dunn post hoc test.

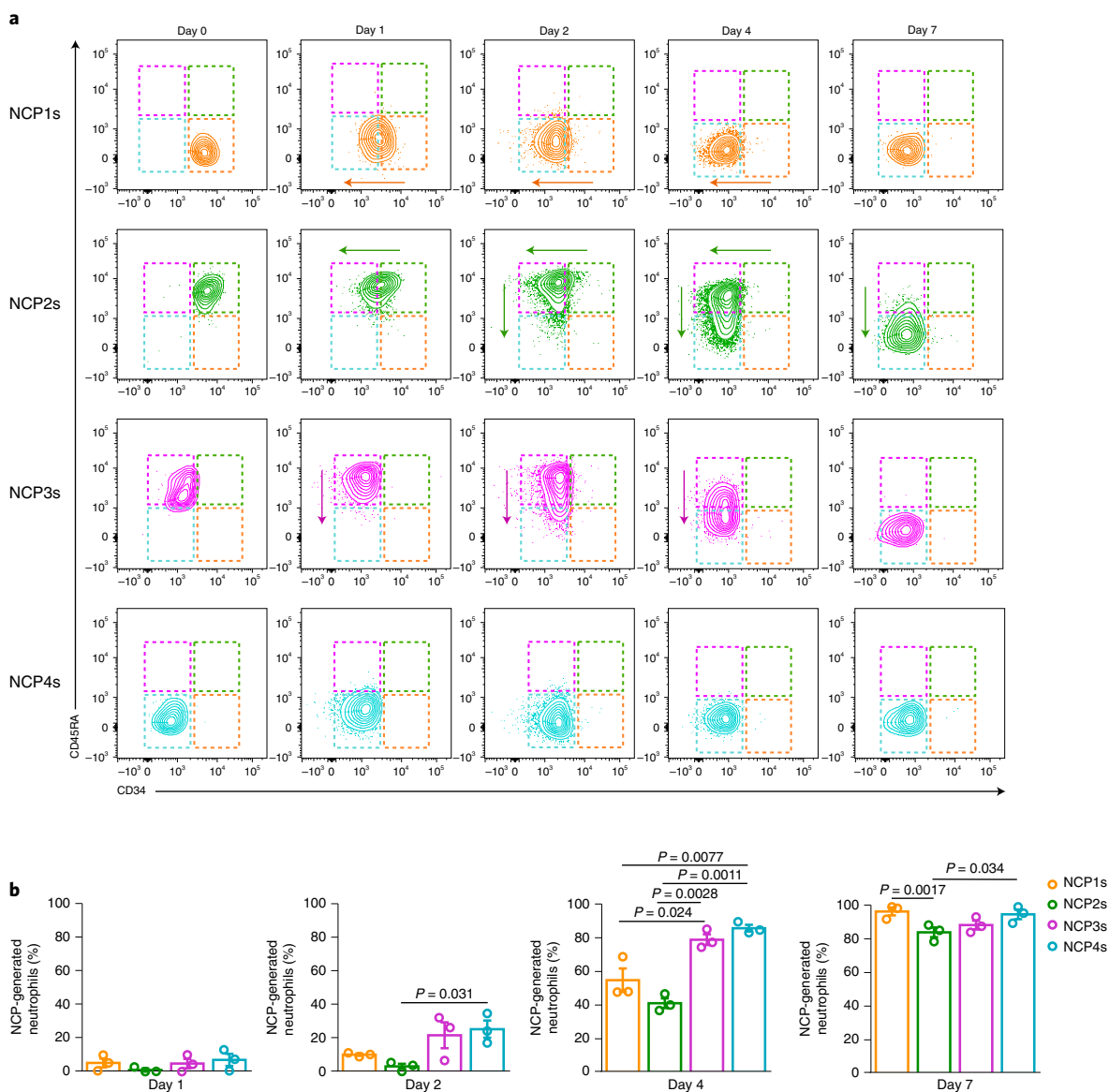


Fig. 3 | NCP1s and NCP2s independently differentiate into neutrophils. **a**, Differentiation of CD64^{dim}CD115⁻ NCPs based on the changes of CD34 and CD45RA expression during culture with SFGc for 1, 2, 4 and 7 d. Plots at day 0 show the phenotype of sorted NCPs, while plots relative to day 1 to day 7 show live, CD45⁺CD66b⁻CD14⁻ cells ($n=3$). **b**, Bar graphs depicting the percentages of live CD66b⁺ neutrophils generated by SFGc-treated NCPs. Data represent mean \pm s.e.m. ($n=3$). Analysis was performed using one-way ANOVA and Tukey post hoc test.

unlike NCP1s and NCP2s, expressed CD15, although at lower levels than PMs and mature SNs (Extended Data Fig. 4d). However, a fraction of NCP3s and NCP4s expressed CD15 at levels comparable to those expressed by cMoPs, pre-monocytes and mature monocytes

(Extended Data Fig. 4e), suggesting that CD15 alone does not specifically identify neutrophil precursors. Thus, these data exclude that NCPs correspond to SSC^{int}Lin⁻CD34⁻CD15^{int}CD11b⁻CD16⁻ early PMs³², which partially overlap with pre-monocytes. These results

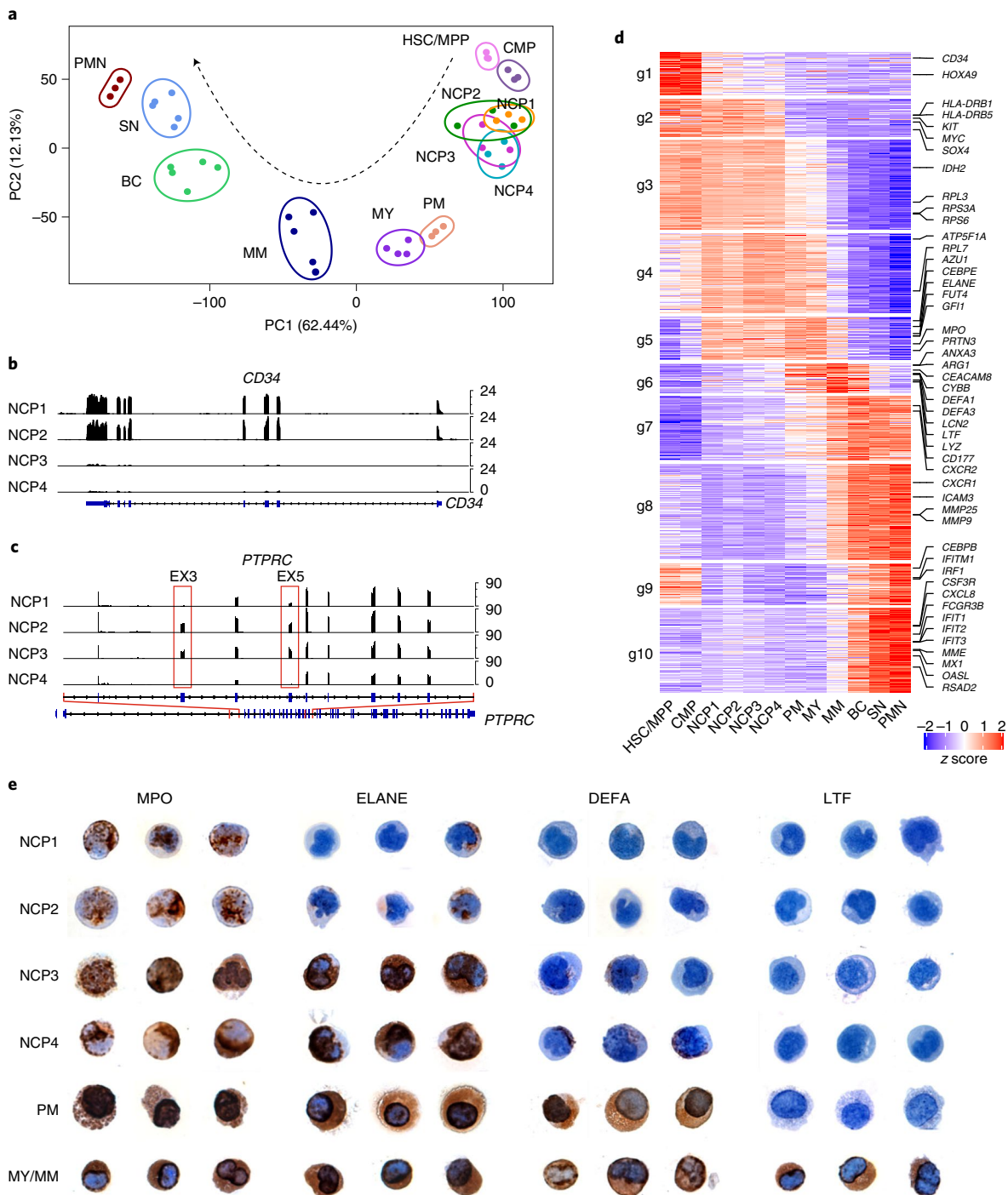


Fig. 4 | RNA-seq experiments confirm that NCPs represent very early precursors of neutrophils. a, PCA scatter plot based on the DEGs identified from bulk RNA-seq analyses of NCP1s (orange), NCP2s (green), NCP3s (magenta) and NCP4s (turquoise), as well as HSCs/MPPs (pink), CMPs (purple), PMs (salmon), MYs (blue violet), MMs (blue), BCs (light green), SNs (light blue) and PMNs (dark red) ($n = 3-5$). **b, c**, Schemes represent genome-browser views of RNA-seq signals at the *CD34* (**b**) and *PTPRC*/*CD45* (**c**) loci of NCPs. Exons 3 and 5 of *PTPRC*, whose transcription is essential for the expression of the *CD45RA* variants, are highlighted by red boxes in **c**. **d**, Heatmap displaying the expression patterns of the gene groups (g1-g10) resulting from the *k*-means analysis of DEGs identified among the various neutrophil-lineage cells and CMPs. The median gene expression levels of biological replicates were calculated, and data were represented as z score. **e**, Expression of antigenic MPO, ELANE, DEFA and LTF by immunocytochemical staining of NCPs, PMs and MYs/MMs. Three representative stained cells for each population are shown. Original magnification, $\times 600$.

indicated that $CD34^+$ and $CD34^{dim/-}$ NCPs were upstream of PMs in the neutrophil maturation pathway, were characterized by a $CD64^{dim}CD115^-$ phenotype and differentially expressed *CD45RA* (Supplementary Table 1 and Extended Data Fig. 5).

NCP1s and NCP2s independently differentiate into neutrophils. Because the acquisition of *CD45RA* conventionally represents a mandatory step for myeloid progenitor maturation that occurs at the GMP stage¹⁰, we investigated the maturation path of $CD34^+$

and CD34^{dim/-} NCPs in in vitro culture with SFGc. Assuming an NCP1 to NCP2 to NCP3 to NCP4 sequential progression, we found that NCP2s downregulated the expression of CD34, acquired an CD34^{dim/-}CD45RA⁺ phenotype and then progressively decreased the expression of CD45RA to become NCP4s (Fig. 3a). NCP3s matured into NCP4s by progressively decreasing CD45RA expression (Fig. 3a), while NCP4s remained CD34⁻ and CD45RA⁻ throughout the entire 7 d of culture (Fig. 3a). NCP1s downregulated the expression of CD34 without upregulating CD45RA, and directly matured into NCP4s (Fig. 3a). As expected, the frequency of NCPs progressively decreased between 4 d and 7 d, in line with their maturation into CD66b⁺ cells (Fig. 3b), and NCP3s and NCP4s differentiated through kinetics remarkably faster than those of NCP1s and NCP2s (Fig. 3b). As such, although NCP1s and NCP2s shared a CD34⁺CD64^{dim}CD115⁻ phenotype, they maintained the differential expression of CD45RA until their maturation into NCP4s.

Because NCP1s did not directly develop into NCP2s, the acquisition of CD45RA by the neutrophil precursors, which occurred exclusively along the NCP2 maturation path, could occur before the acquisition of CD64, implying that neutrophil precursors more immature than NCP1s and NCP2s would differ in their CD45RA expression. Hence, we sorted SSC^{lo}CD45^{dim}CD10⁻CD123^{dim/-}CD34⁺CD45RA⁻CD64⁻CD115⁻ cells (hereafter, CD45RA⁻CD64⁻ subset) and CD64⁻ GMDPs and cultured them with SFGc to test whether precursors contained within CD64⁻ GMDPs (Extended Data Fig. 6a) gave rise to NCP2s, as well as whether the CD45RA⁻CD64⁻ subset (Extended Data Fig. 6a) generated both NCP1s and CD64⁻ GMDPs, and, in turn, NCP2s. We noticed that the CD45RA⁻CD64⁻ subset concomitantly generated both CD34⁺CD45RA⁺ and CD34⁻CD45RA⁻ populations following two maturation routes (Extended Data Fig. 6b). Conversely, CD64⁻ GMDPs differentiated into CD34⁻CD45RA⁻ cells through an intermediate CD34^{dim/-}CD45RA⁺ stage, therefore never giving rise to CD34⁺CD45RA⁻ cells (Extended Data Fig. 6c). Analysis of the culture-derived cells indicated that the CD45RA⁻CD64⁻ subset restored both NCP1s and NCP2s (Extended Data Fig. 6b), while CD64⁻ GMDPs generated, as expected, only NCP2s (Extended Data Fig. 6c). Importantly, the CD45RA⁻CD64⁻ subset generated CD64⁻ GMDPs as early as day 2 (Extended Data Fig. 6b), while NCP2s emerged only at day 4, thus confirming that NCP2s derive from CD64⁻ GMDPs. In addition, both the CD45RA⁻CD64⁻ subset and the CD64⁻ GMDPs gave rise to both NCP3s and NCP4s as well as to PMNs and DCs/monocytes (not shown). These observations indicated that NCP1s and NCP2s differentiated through separate maturation routes and were derived from CD34⁺ progenitors that bifurcate into CD45RA⁺ and CD45RA⁻ NCPs before acquiring CD64 (Extended Data Fig. 6d).

NCPs precede PMs along the neutrophil maturation cascade.

Next, we used RNA-sequencing (RNA-seq) to profile NCPs, HSCs/multipotent progenitors (MPPs), CMPs, PMs, MYs, MMs, BCs, SNs

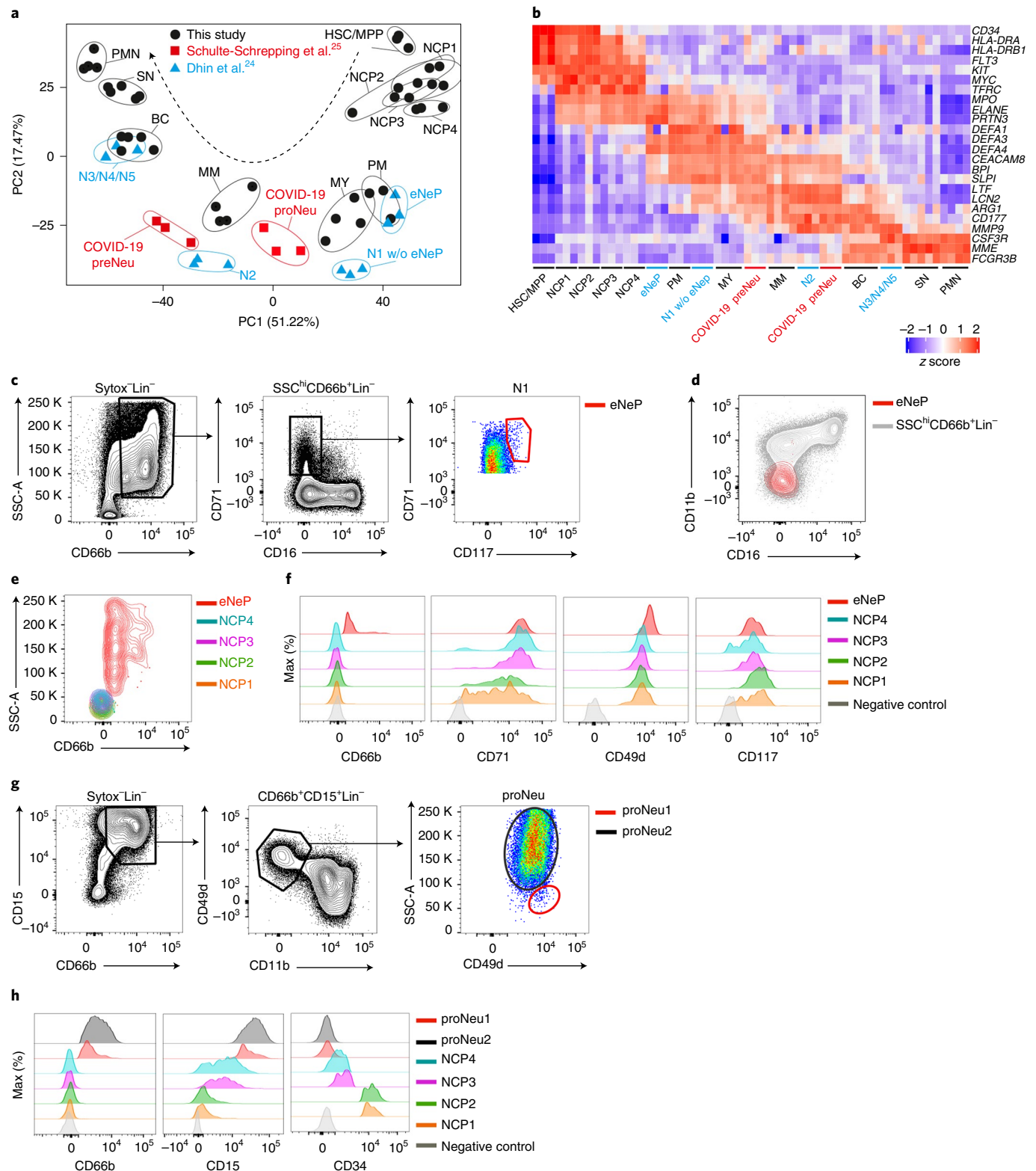
and PMNs. Both principal component analysis (PCA) (Fig. 4a) and hierarchical clustering analysis performed by optimal leaf ordering (Extended Data Fig. 7a) indicated that NCPs not only clustered by themselves but were also placed along the HSCs/MPPs and CMPs, PMs, MYs, MMs, BCs and SNs maturation trajectory, which ends in PMNs. Analysis of *CD34* and *PTPRC* (which encodes CD45) transcript expression (Fig. 4b,c) corroborated the flow cytometry data. Furthermore, by performing *k*-means clustering of differentially expressed genes (DEGs), ten main gene groups (g1–g10) were identified among all samples (Fig. 4d and Supplementary Table 2). Genes encoding markers of immature cells (such as *CD34*, *HOXA9*, *MYC*, *SOX4* and *KIT*), expressed only by HSCs/MPPs, CMPs and NCPs, were detected in g1 or g2 (Fig. 4d). Gene ontology (GO) analysis of the DEGs indicated the ‘MHC class II protein complex’ term was among the most significantly over-represented in g2 (Extended Data Fig. 7b), in line with the previously described expression of MHC class II in very immature neutrophil progenitors³³. Similarly, DEGs mainly involved in ribosome assembly, mitochondria formation and cell cycle regulation, characterizing g3, g4 and g5, respectively, were expressed in NCPs, PMs, MYs and MMs, but not in nonproliferating BCs and SNs (Fig. 4d, Extended Data Fig. 7b,c and Supplementary Table 2).

Importantly, g5 contained DEGs encoding messenger RNAs for the azurophilic granule (AG) proteins (such as *MPO*, *AZU1*, *PRTN3* or *ELANE*), which are transcribed first during granulopoiesis, as well as transcription factors known to be typically expressed in immature neutrophils, such as *GFI1* and *CEBPE* (ref. ³⁴) (Fig. 4d). A more focused analysis revealed that AG genes started to be expressed in CMPs, progressively increased in NCP1s/NCP2s and NCP3s/NCP4s, were maximally transcribed in PMs and MYs, and then gradually disappeared (Extended Data Fig. 7c). Defensin genes such as *DEFA1*, *DEFA1B*, *DEFA3* and *DEFA4*, which encode AG-stored proteins³⁵, were very poorly expressed in NCPs, and included in g6 (Fig. 4d and Supplementary Table 2). Moreover, g6, but also g7 and g8, included other DEGs absent in NCPs, and were enriched for mRNAs encoding specific granule (SG) proteins (for example, *LYZ*, *LTF*, *LCN2* and *CEACAM8/CD66b*), gelatinase granule (GG) proteins (for example, *MMP9*, *ARG1* and *CD177*), secretory vesicle (SV) proteins (for example, *MMP25* and *ANXA3*) and cell membrane (CM) proteins (for example, *CXCR1*, *CXCR2* and *ICAM3*) (Fig. 4d and Extended Data Fig. 7c), and thus associated to GO terms such as ‘neutrophil degranulation’ and ‘neutrophil activation’ (Extended Data Fig. 7b). Furthermore, g9 and g10 were enriched in interferon-stimulated genes (ISGs), as well as other genes associated to the GO terms ‘response to bacterium’ and ‘cytokine production’, present in BCs, SNs and PMNs (Fig. 4d and Extended Data Fig. 7b). Consistent with the literature³⁵, mRNAs encoding proteins involved in the production of reactive oxygen species, phagocytosis and chemotaxis (Extended Data Fig. 7c) were transcribed starting from the MM–BC stages and were highly expressed in PMNs.

Fig. 5 | NCPs stand at earlier stages than human eNePs, proNeu1/2s and COVID-19 proNeus. **a**, PCA scatter plot based on the top 500 most variable genes identified across: (1) bulk RNA-seq analyses of our HSCs/MPPs, CMPs, NCP1s, NCP2s, NCP3s, NCP4s, PMs, MYs, MMs, BCs, SNs and PMNs; (2) bulk RNA-seq analyses of eNePs, N1 without (w/o) eNeP, N2 and mixed N3/N4/N5 from the report by Dinh et al.²⁴ (cyan); (3) pseudo-bulk RNA-seq analyses of COVID-19 proNeus and preNeus from the report by Schulte-Schrepping et al.²⁵ (red). **b**, Heatmap displaying the expression grade of genes encoding for molecules distinctively associated to discrete neutrophil maturation stages by the cells represented in the PCA of panel **a**. Relative expression levels are represented as z score. **c**, Flow cytometry strategy for the identification of SSC^{hi}CD66b⁺Lin⁻CD71⁺CD117⁺ eNePs (red) within SSC^{hi}CD66b⁺Lin⁻N1s. **d**, Overlay plot of eNePs (red) with SSC^{hi}CD66b⁺Lin⁻ neutrophils (gray). Cells were analyzed based on CD11b and CD16 expression. **e**, Overlay plot of NCPs (orange, green, magenta and turquoise) and eNePs (red). Cells were analyzed based on SSC parameter and CD66b expression. **f**, Histogram showing the expression of CD71, CD117, CD49d and CD66b by NCPs (orange, green, magenta and turquoise) and eNePs (red), as compared with the fluorescence-negative control (light gray). **g**, Flow cytometry strategy for the identification of CD66b⁺CD15⁺Lin⁻CD11b⁻CD49d^{dim}SSC^{hi} proNeu1s (red) and CD66b⁺CD15⁺Lin⁻CD11b⁻CD49d^{dim}SSC^{hi} proNeu2s (black) within CD66b⁺CD15⁺Lin⁻ proNeus. **h**, Histogram overlays showing the expression of CD66b, CD15 and CD34 by NCPs (orange, green, magenta and turquoise), proNeu1s (red) and proNeu2s (black), as compared with fluorescence-negative control (light gray). For panels **c–h**, one representative experiment out of three is shown.

To get more insights into the specific transcriptomic differences among the NCPs, we performed DEG analysis by using the likelihood ratio test²⁶ and identified 1,114 DEGs. NCPs distinctly segregated from each other by PCA, indicating remarkable differences among their gene expression profiles (Extended Data Fig. 7d). However, while PC1 differences were mostly determined by cell cycle and AG genes, which were expressed more in NCP1s/NCP2s and NCP3s/NCP4s, respectively (Extended Data Fig. 7d), mRNAs that

mostly contributed to the PC2 variations, namely *PRG2*, *CLC*, *EPX* and *IL5RA*, found in NCP1s, or *IRF8*, *ANXA2*, *SAMHD1*, *SLAMF7*, *LYZ* and *F13A1*, found in NCP2s and NCP3s (Extended Data Fig. 7d), are typically expressed in progenitors of eosinophils or monocytes, respectively. Immunocytochemistry analysis of cytopins from NCPs, PMs and MYs/MMs indicated that NCPs were myeloperoxidase (MPO)⁺ and that NCP3/NCP4s, but not NCP1/NCP2s, consistently expressed neutrophil elastase (ELANE) (Fig. 4e).



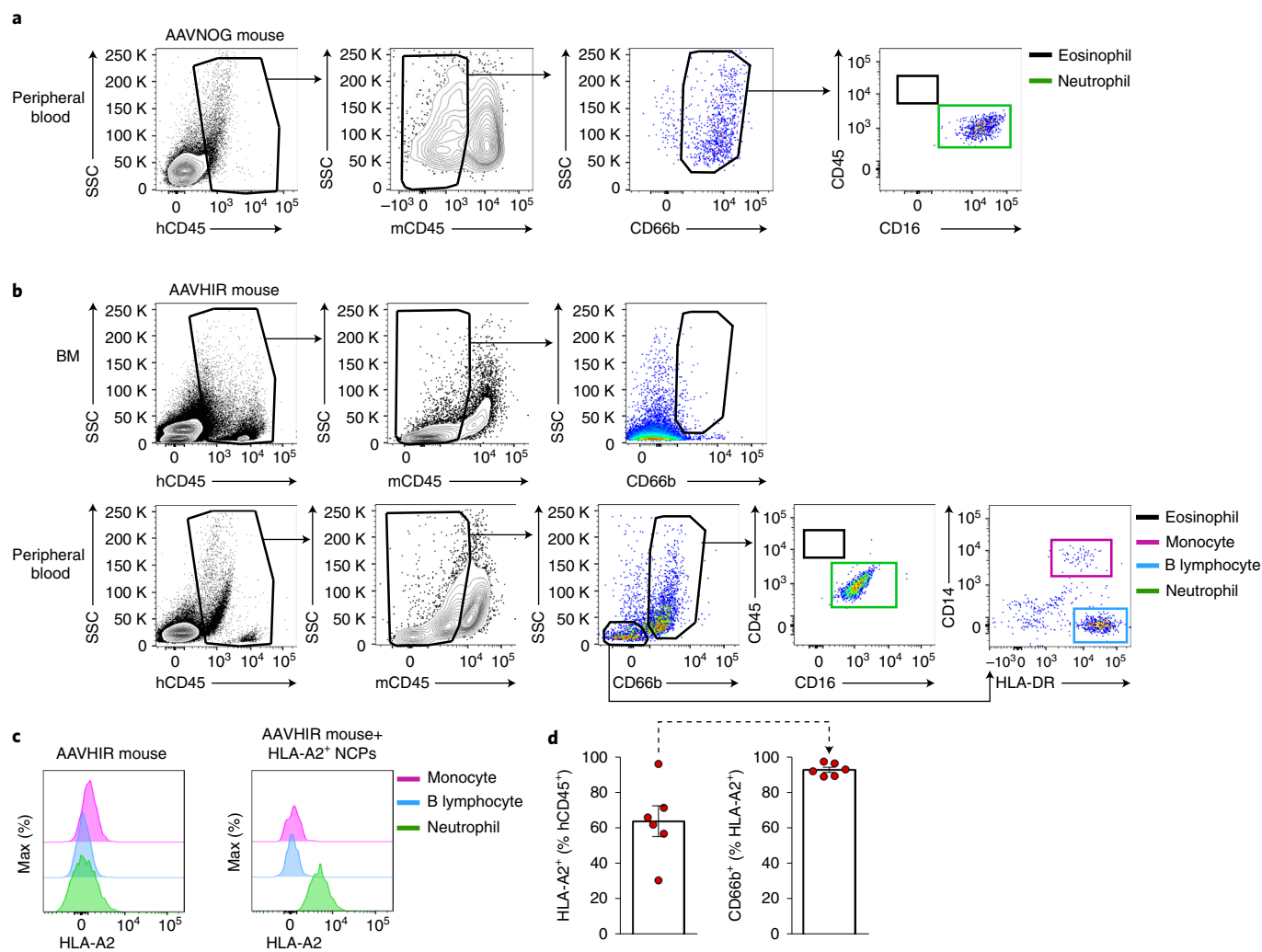


Fig. 6 | Generation of neutrophils by human NCPs in adoptive transfer experiments. a, Flow cytometry strategy for the identification of human $SSC^{\text{hi}}CD66b^+CD45^{\text{hi}}CD16^+$ neutrophils (green), by excluding $SSC^{\text{hi}}CD66b^+CD45^{\text{hi}}CD16^-$ eosinophils within the peripheral blood from AAVNOG mice, 11 d after intraocular injection of total NCPs. One representative mouse out of three with similar results is shown. **b**, Flow cytometry strategy for the identification of human $SSC^{\text{hi}}CD66b^+CD45^{\text{hi}}CD16^+$ neutrophils (green) in BM and peripheral blood of HLA-A2⁻ AAVHIR mouse, 14 d following intratibial injection of total HLA-A2⁺ NCPs. $CD14^+HLA-DR^+$ monocytes (magenta) and $CD14-HLA-DR^+$ B lymphocytes (light blue) are also detectable within the $SSC^{\text{lo}}CD66b^-$ region. One representative mouse out of six with similar results is shown. **c**, Histogram overlays showing the expression of HLA-A2 by monocytes (magenta), B lymphocytes (light blue) and neutrophils (green) identified within peripheral blood of HLA-A2⁻ AAVHIR mouse, before (left) and after (right) HLA-A2⁺ NCP injection ($n=6$). **d**, Bar graphs depicting the percentage of human HLA-A2⁺CD66b⁺ neutrophils (right) within total HLA-A2⁺ cells (left). Percentages were calculated by analyzing peripheral bloods of HLA-A2⁻ AAVHIR mice, 14 d following intratibial injection of HLA-A2⁺ NCPs. Data represent mean \pm s.e.m. ($n=6$).

Moreover, NCPs were defensin⁻, unlike PMs and MY/MM, with the latter also lactoferrin (LTF)⁺, as reported³⁷, and confirming the RNA-seq data (Fig. 4d and Supplementary Table 2). These observations indicated that NCPs were committed to the neutrophil lineage, and that they are placed upstream of PMs along the neutrophil maturation cascade.

NCPs are upstream of eNePs, human proNeu1/2s and COVID-19 proNeus. We next asked whether NCPs were upstream of the recently described Lin⁻CD66b⁺ eNePs²⁴, Lin⁻CD66b⁺ proNeu1/2s (ref. ²³) and other cells found in the blood of patients with COVID-19 (ref. ²⁵) defined as proNeus, based on the constitutive expression of mRNAs encoding markers for proNeu1/2s (including CD15, CD43, CD63, CD66b, CD81)²³. We compared publicly available RNA-seq datasets for eNePs and more mature precursors (N1 without eNeP, N2 and mixed N3/N4/N5)²⁴, as well as of scRNA-seq

datasets from COVID-19 proNeus and (more mature) preNeus²⁵, with the gene expression profiles of NCPs by PCA (Fig. 5a). Human proNeu1/2 transcriptomes were not available²³. eNePs clustered with PMs, while N1 without eNePs and N2 localized close to MYs and MMs, respectively (Fig. 5a). Moreover, in accordance with their more mature phenotypes, N3/N4/N5 grouped with BCs (Fig. 5a). Similarly, according to RNA-seq pseudo-bulk computing, COVID-19 proNeus and preNeus clustered before and after MMs, respectively (Fig. 5b). Examination of the mRNA accumulation encoding for molecules distinctively associated to discrete neutrophil maturation stages (Fig. 5b), indicated that only NCP1s and NCP2s expressed *CD34* mRNA, while all NCPs, but not eNePs or COVID-19 proNeus, accumulated mRNAs associated to stemness, such as *KIT*, *MYC* and MHC class II genes. Moreover, NCPs, eNePs and COVID-19 proNeus, as well as PMs and MYs, expressed AG genes (that is, *ELANE*, *PRTN3* and *MPO*), albeit at variable levels. By contrast,

only eNePs and COVID-19 proNeus expressed PM-associated genes, including defensins and *BPI* (that is, AG proteins), or *SLPI* and *CEACAM8/CD66b* (that is, SG proteins). Finally, and in agreement with PCA results (Fig. 5a), only COVID-19 proNeus, but not NCPs or eNePs, expressed SG and GG genes, such as *LCN2*, *LTF* and *ARG1*, at levels comparable to MYs (Fig. 5b).

Next, we reproduced the gating strategies used to identify eNePs²⁴ and human proNeu1/2s (ref.²³) to directly compare their phenotypes with those of NCPs. eNePs were defined as a CD117⁺ subset of the Lin⁻CD66b⁺CD71⁺ N1 cells (Fig. 5c). However, according to their CD11b, CD16 and CD66b expression (Fig. 5d–f), as well as side scatter (SSC) properties (Fig. 5e), eNePs corresponded mostly to a more immature fraction of PMs²⁴ (Fig. 5a). By contrast, NCPs were CD66b⁻ and exclusively located in the SSC^{lo} fraction of immature BM cells (Fig. 5e,f and Extended Data Fig. 5). In addition, NCPs were positive for markers not specifically associated to neutrophil precursors^{26,38–40}, namely CD71, CD117 and CD49d (Fig. 5f).

Because a detailed gating strategy for their identification and isolation, or a definitive phenotype, is not available for proNeus, we distinguished them as Lin⁻CD66b⁺CD15⁺CD11b⁻CD49d^{hi}SSC^{lo} proNeu1s and Lin⁻CD66b⁺CD15⁺CD11b⁻CD49d^{int}SSC^{hi} proNeu2s (ref.²³). proNeu1/2s (Fig. 5g) differed from NCPs in terms of expression of CD66b, CD15 and CD34 (Fig. 5h), but not CD64 and CD115 (not shown). Differently from NCPs, proNeu1s and proNeu2s expressed CD66b at mid/low and mid/high levels, respectively (Fig. 5h). Similarly, while NCP1s and NCP2s were CD15⁻ and NCP3/NCP4s were CD15^{dim/+}, proNeu1/2s had a higher expression of CD15 than NCPs (Fig. 5h). Furthermore, we found that proNeu1s were CD34⁻, similarly to proNeu2s, NCP3s and NCP4s (Fig. 5h), in contrast with the Infinity Flow approach identifying proNeu1s as CD34⁺ (ref.²⁴). This phenotypic analysis supported the notion that proNeu1s consisted of highly immature PMs, similarly to eNePs, while proNeu2s corresponded to PMs. Together, these data suggested that NCPs were more immature than eNePs and COVID-19 proNeus at the molecular level, as well as than eNePs and proNeu1/2s at the phenotypic level, while COVID-19 proNeus were more mature than originally assumed²⁵.

NCPs generate neutrophils in vivo. To investigate their contribution to granulopoiesis in vivo, NCPs were adoptively transferred into immunodeficient mice by different experimental approaches. In a first set of experiments, total NCPs were sorted and injected intravenously into AAVNOG mice. Human CD66b⁺ neutrophils were recovered in the peripheral blood (Fig. 6a), but not in the BM, of recipient AAVNOG mice 11 d after NCP transfer. In a second set of experiments, NCPs sorted from the BM of HLA-A2⁺ donors were transferred by intratibial injection into AAVHIR mice (that is, NOG mice engrafted with human HLA-A2-CD34⁺ cells, then treated with adeno-associated virus serotype 9 (AAV9) as before, and finally pre-conditioned with a nonmyeloablative dose of busulfan). CD66b⁺ neutrophils, although undetectable in the BM (Fig. 6b), represented the only human HLA-A2⁺ cells that were detected in the peripheral blood of recipient AAVHIR mice 14 d following NCP transfer (Fig. 6b–d). The remaining human cells consisted of HLA-A2⁻ monocytes and B lymphocytes (Fig. 6b,c), which likely survived the busulfan conditioning. Altogether, these in vivo adoptive transfer experiments indicated the neutrophil-specific commitment of NCPs and supported their biological role in neutropoiesis.

NCPs fall into multiple scRNA-seq cell clusters. To unequivocally dissect their transcriptional heterogeneity, we performed scRNA-seq in sorted NCP1s, NCP2s, NCP3s, NCP4s and cMoPs. Dimensionality reduction by Uniform Manifold Approximation and Projection (UMAP)⁴¹ on the 17,902 cells sequenced indicated that NCPs clearly segregate from cMoPs, while NCP1s mostly

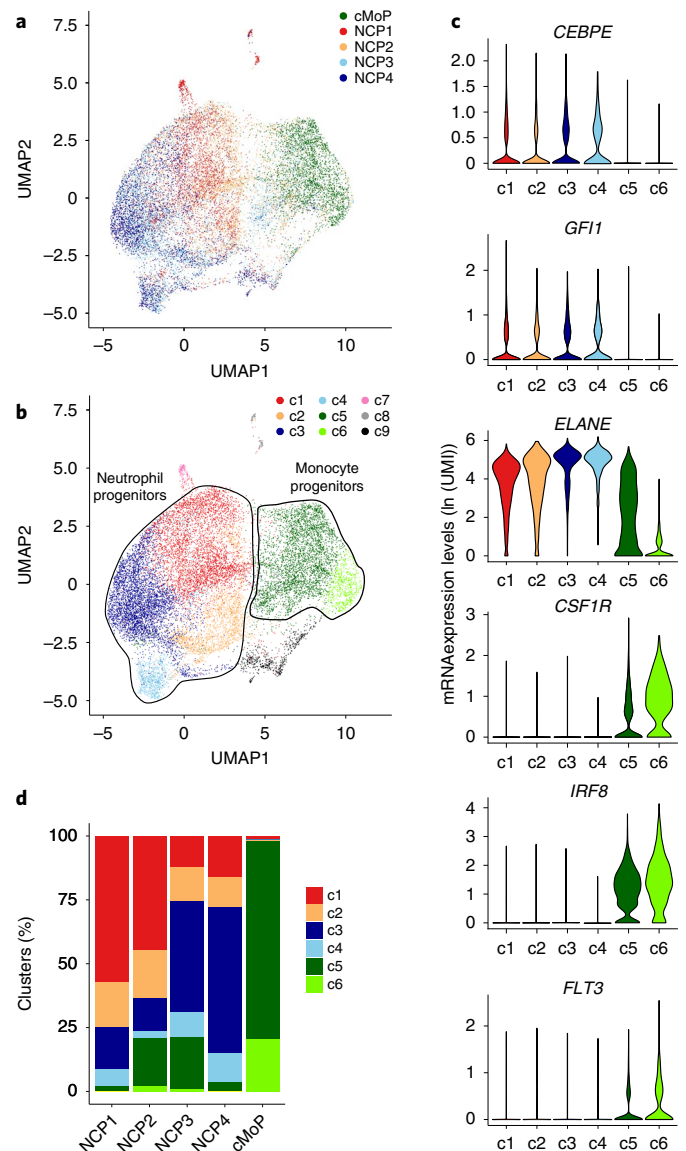


Fig. 7 | scRNA-seq experiments of NCPs and cMoPs reveal that they consist of multiple cell clusters. **a**, UMAP visualization of scRNA-seq profiles of NCP1s (red), NCP2s (orange), NCP3s (light blue), NCP4s (blue) and, for comparison, cMoPs (green), isolated from three different BMs. **b**, Identification of the clusters (c1–c9) in the UMAP shown in **a**, by Louvain clustering analysis. **c**, Violin plots showing the mRNA expression levels (as ln(UMI)) of neutrophil-lineage-specific genes (that is, *CEBPE*, *GFI1* and *ELANE*) and monocyte lineage-specific genes (that is, *CSF1R*, *IRF8* and *FLT3*), across c1–c6 cells. Cluster colors are the same of those shown in Fig. 7b. **d**, Stacked bar graph shows the relative abundances of c1–c6 cells in NCP1s, NCP2s, NCP3s, NCP4s and cMoPs.

overlap with NCP2s and NCP3s with NCP4s (Fig. 7a and Extended Data Fig. 8a). Unbiased, graph-based clustering by Seurat⁴² identified nine discrete cell clusters (Fig. 7b), of which three (c7, c8 and c9) were excluded from all subsequent analyses because they consisted of eosinophil precursors (c7, identified by *EPX* mRNA expression), megakaryocyte/erythroid precursors (c8, identified by *PF4* and *HBB* mRNA expression) or cells with an unclear ontogeny and high expression of apoptosis-related genes (c9) (Fig. 7b). Of the six remaining cell clusters, four were unequivocally attributable to the neutrophil lineage (c1, c2, c3 and c4), and two to the monocyte lineage (c5 and c6) (Fig. 7b). Such segregation was also confirmed

by hierarchical clustering analysis of their DEGs (Extended Data Fig. 8b), as genes typical of the neutrophil (such as *CEBPE*, *GFI1* or *ELANE*) and monocyte (such as *IRF8*, *CSF1R* or *FLT3*) lineages were exclusively expressed in c1–c4 and c5–c6, respectively (Fig. 7c). Moreover, distribution analysis of scRNA-seq cell clusters in NCPs and cMoPs evidenced that c1–c4 were completely absent in cMoPs (Fig. 7c), indicating that CD115 expression efficiently discriminates NCPs. By contrast, c5 cells were found to contaminate NCP2s and NCP3s, but not NCP1s and NCP4s (Fig. 7d), therefore explaining the PCA results on bulk RNA-seq (Extended Data Fig. 7d).

The distribution of c1–c4 cells in NCP1s and NCP4s is peculiar to that of NCP2s and NCP3s, respectively (Fig. 8a). Moreover, while NCP1s and NCP2s were enriched mainly in c1 cells, NCP3s and NCP4s were preferentially accumulated in c3 and c4 cells (Fig. 8a). By contrast, c2 cells were equally distributed among the four NCPs (Fig. 8a). Calculation of the pseudotime value for every cell by *Destiny*⁴³ (Fig. 8b), to define their maturation state, showed that c1 contained the more immature cells, while c3 and c4 contained relatively more mature cells (Fig. 8c). Accordingly, *CD34* was preferentially expressed in c1 and c2 cells (Fig. 8d), consistent with their relatively more abundant presence in NCP1s and NCP2s (Fig. 8a). c2 cells, instead, displayed variable pseudotime values (Fig. 8c), in line with their presence in all NCPs (Fig. 8a). Seurat analysis of the DEGs in c1–c4 cells resulted in the identification of 136 genes (Fig. 8e and Supplementary Table 3). c1 cells had high expression of typical genes of immature proliferating cells (such as *CD34*, *TOP2A*, *TUBB* and *HIST1H4C*) (Fig. 8e and Extended Data Fig. 8c,d), consistent with their lowest pseudotime values (Fig. 8c). c1 cells also expressed genes associated to the ‘MHC class II protein complex’ GO term (Extended Data Fig. 8e), consistent with the g2 genes from bulk RNA-seq data (Extended Data Fig. 7a). By contrast, c3 and c4 cells had high expression of AG genes (Fig. 8e and Extended Data Fig. 8c,d), and displayed GO terms mostly enriched for ‘neutrophil activation’ and ‘neutrophil degranulation’ (Extended Data Fig. 8e), in accordance with their elevated pseudotime values (Fig. 8c), and their prevalent correspondence with NCP3s and NCP4s (Fig. 8a). Notably, c4 cells specifically expressed high mRNA levels of *BEX1* (Fig. 8e and Extended Data Fig. 8c), while c2 cells had high expression of ISGs, such as *ISG15*, *IFI6*, *IFIT3* and many more (Fig. 8e and Extended Data Fig. 8c), in line with the enrichment of the GO term ‘defense response to virus’ in this cluster (Extended Data Fig. 8e). No SG or GG genes were present in scRNA-seq datasets (Extended Data Fig. 8d), confirming the RNA-seq data. Finally, using *Destiny*⁴³ to assess their potential developmental trajectories, we found that c1–c4 cells distributed along three branches. The c1 branch included the majority of c1 cells (Fig. 8f,g and Extended Data Fig. 8f), together with a minor quote of the c2, c3 and c4 cells, and branched into two different trajectories (Fig. 8f): one, defined as the ‘conventional trajectory’, was mainly characterized by cells expressing increased AG mRNA (c3 and c4 cells), but also included *BEX1*⁺ cells (c4 cells); the other, defined as the ‘ISG trajectory’ (Fig. 8f), was characterized by ISG^{hi} cells (c2 cells) (Fig. 8h). As such, scRNA-seq clustering analysis identified four clusters of neutrophil precursor cells at different stages of maturation, distributed along two maturation routes, and indicated that

NCP1s and NCP2s on one hand, and NCP3s and NCP4s on the other hand, were identical in terms of cluster distribution.

Discussion

In this study, we identified the phenotypes of four different CD64^{dim}CD115⁻ NCPs within the SSC^{lo}Lin⁻CD45^{dim} region of human BM. Based on their differential CD34 and CD45RA expression, NCPs were subdivided into CD34⁺CD45RA⁻ NCP1s, CD34⁺CD45RA⁺ NCP2s, CD34^{dim/-}CD45RA⁺ NCP3s and CD34^{dim/-}CD45RA⁻ NCP4s. Within cGMPs, we also defined CD64⁺CD115⁺ cMoPs (refs. 16,20), CD64⁻CD115⁺ MDPs and CD64⁻CD115⁻ GMDPs (ref. 15). Because GMDPs generate neutrophils, but not eosinophils and basophils^{4,17–19}, we propose to rename GMDPs as NMDPs (that is, neutrophil-monocyte-DC progenitors).

Our experiments uncovered that, while NCP1s originated from CD45RA⁻ progenitors and directly matured into NCP4s, NCP2s developed from NMDPs and matured into NCP4s through NCP3s. Hence, our data not only confirmed that the transition of neutrophil precursors from HSCs to cGMPs was associated to their CD45RA upregulation¹⁰, but also uncovered the existence of a second developmental route, independent from CD45RA upregulation, which might be explained by the necessity to increase the number of peripheral neutrophils during emergency granulopoiesis. Alternatively, it could be that CD45RA may serve either for the localization of CD34⁺ and CD34^{dim/-} NCPs in specific BM niches, or for a yet unknown function under pathological conditions.

Using transcriptomic dataset and phenotypic analysis comparisons, our data showed that NCP neutrophil precursors represent an earlier stage of differentiation compared with the recently described eNePs²⁴, human proNeu1/2s (ref. 23) and COVID-19 proNeus²⁵. In fact, clustering analysis of DEGs among CD34⁺ and CD34^{dim/-} NCPs evidenced a remarkable enrichment of genes related to cell proliferation by NCP1s and NCP2s, as well as of AG-encoding genes by NCP3s and NCP4s.

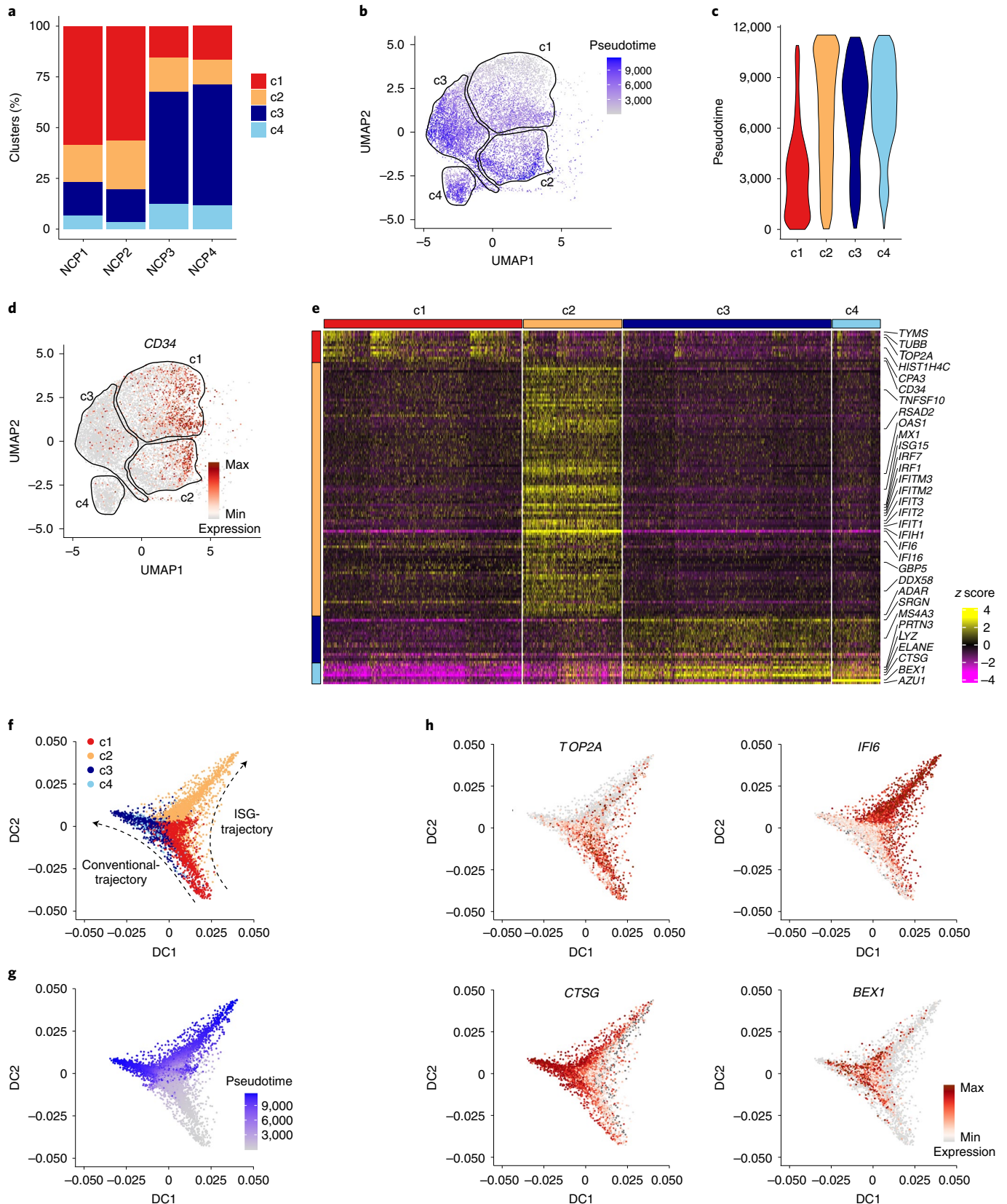
scRNA-seq experiments uncovered that the CD34⁺ and CD34^{dim/-} NCPs consisted of four cell clusters (c1–c4) distributed at substantially identical ratios in NCP1s and NCP2s, as well as in NCP3s and NCP4s. One of them (c2) was enriched in ISGs, consistent with the recent identification of PMN subsets expressing ISGs during health^{44,45} or pathological conditions^{25,46–49}. Treatment of patients with severe COVID-19 with dexamethasone was reported to reduce the peripheral ISG⁺ neutrophil subset on the one hand, but to expand the subset of immature immunosuppressive cells on the other hand⁴⁸, suggesting neutrophil polarization could be modulated by pharmacological interventions.

Clusters of neutrophil precursors displaying levels of maturation similar to c1–c4 cells have been identified by scRNA-seq of HSPCs from human BM, even though the sorting strategy for their isolation was not described. For instance, four clusters of neutrophil precursors (N0–N3) were found in CMPs (N0) and cGMPs (N1, N2 and N3)³. N1, N2 and N3 might correspond in part to the c1 cells given their expression of CD34 protein and AG mRNA. However, N3 could be ascribed to a monocyte progenitor due to its expression of *LYZ*, *SAMHD1*, *CSF1R*, *ANXA2*, *KLF4* and *IRF8* mRNA³, which we found to be peculiar to cMoPs. Similarly, it remains uncertain whether N0, which only expresses MPO among the AG genes³,

Fig. 8 | Characterization of the scRNA-seq cell clusters composing NCPs. **a**, Stacked bar graph shows the relative abundance of c1–c4 cells in NCP1s, NCP2s, NCP3s and NCP4s, after exclusion of cells from monocyte progenitors (c5 and c6). **b, c**, UMAP (**b**) and violin (**c**) plots restricted to c1–c4 cells showing the pseudotime value for every cell, as calculated by the *Destiny* algorithm. **d**, Expression patterns of *CD34* mRNA projected on a UMAP plot restricted to c1–c4 cells. **e**, Heatmap showing scaled expression of discriminative gene sets (Bonferroni-corrected $P < 0.05$; two-sided Wilcoxon Rank Sum test) by c1–c4 cells, as defined in Fig. 7b. The most specific genes of each cluster are depicted on the right y axis. The color scheme is based on z score distribution from –4 (purple) to 4 (yellow). **f, g**, Trajectory plots showing the distribution of c1–c4 cells (**f**), and their pseudotime values (**g**). In **f**, the arrows indicate both the ‘conventional’, and the ‘ISG’, developmental trajectories, along the neutrophil lineage. **h**, mRNA expression patterns of genes characteristic of c1 (that is, *TOP2A*), c2 (that is, *IFI6*), c3 (that is, *CTSG*) and c4 (that is, *BEX1*) cells, projected on the trajectory plots from **f**.

corresponds to a neutrophil progenitor, because MPO can also be expressed by cMoPs⁹. However, none of these studies^{3,19} detected specific ISG- or BEX1-containing clusters (as we did), likely because of the much lower number of neutrophil precursors investigated compared with our scRNA-seq set.

In summary, we identified CD34⁺ and CD34^{dim/-} NCPs within human BMs, all of them characterized by an SSC^{lo}Lin⁻CD45^{dim}CD64^{dim}CD115⁻ phenotype. As such, CD34⁺ and CD34^{dim/-} NCPs were easily sortable. We expect that future work will uncover specific membrane markers that would allow the sorting



of the cell clusters within the CD34⁺ and CD34^{dim/-} NCPs, to allow an improved characterization and to evaluate if and how they expand under pathological conditions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41590-022-01189-z>.

Received: 31 May 2021; Accepted: 15 March 2022;

Published online: 28 April 2022

References

- Notta, F. et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016).
- Paul, F. et al. Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **163**, 1663–1677 (2015).
- Velten, L. et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).
- Pellin, D. et al. A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. *Nat. Commun.* **10**, 2395 (2019).
- Yanez, A. et al. Granulocyte-monocyte progenitors and monocyte-dendritic cell progenitors independently produce functionally distinct monocytes. *Immunity* **47**, 890–902.e4 (2017).
- Zheng, S. et al. Molecular transitions in early progenitors during human cord blood hematopoiesis. *Mol. Syst. Biol.* **14**, e8041 (2018).
- Karamitros, D. et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat. Immunol.* **19**, 85–97 (2018).
- Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, eaaw3381 (2020).
- Muench, D. E. et al. Mouse models of neutropenia reveal progenitor-stage-specific defects. *Nature* **582**, 109–114 (2020).
- Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: a human perspective. *Cell Stem Cell* **10**, 120–136 (2012).
- Watcham, S., Kucinski, I. & Gottgens, B. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood* **133**, 1415–1426 (2019).
- Jacobsen, S. E. W. & Nerlov, C. Haematopoiesis in the era of advanced single-cell technologies. *Nat. Cell Biol.* **21**, 2–28 (2019).
- Manz, M. G., Miyamoto, T., Akashi, K. & Weissman, I. L. Prospective isolation of human clonogenic common myeloid progenitors. *Proc. Natl Acad. Sci. USA* **99**, 11872–11877 (2002).
- Doulatov, S. et al. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nat. Immunol.* **11**, 585–593 (2010).
- Lee, J. et al. Restricted dendritic cell and monocyte progenitors in human cord blood and bone marrow. *J. Exp. Med.* **212**, 385–399 (2015).
- Kawamura, S. et al. Identification of a human clonogenic progenitor with strict monocyte differentiation potential: a counterpart of mouse cMoPs. *Immunity* **46**, 835–848.e4 (2017).
- Mori, Y. et al. 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).
- Gorgens, A. et al. Revision of the human hematopoietic tree: granulocyte subtypes derive from distinct hematopoietic lineages. *Cell Rep.* **3**, 1539–1552 (2013).
- Drissen, R., Thongjuea, S., Theilgaard-Monch, K. & Nerlov, C. Identification of two distinct pathways of human myelopoiesis. *Sci. Immunol.* **4**, eaau7148 (2019).
- Olweus, J., Thompson, P. A. & Lund-Johansen, F. Granulocytic and monocytic differentiation of CD34^{hi} cells is associated with distinct changes in the expression of the PU.1-regulated molecules, CD64 and macrophage colony-stimulating factor receptor. *Blood* **88**, 3741–3754 (1996).
- Evrard, M. et al. Developmental analysis of bone marrow neutrophils reveals populations specialized in expansion, trafficking, and effector functions. *Immunity* **48**, 364–379.e8 (2018).
- Zhu, Y. P. et al. Identification of an early unipotent neutrophil progenitor with pro-tumoral activity in mouse and human bone marrow. *Cell Rep.* **24**, 2329–2341.e8 (2018).
- Kwok, I. et al. Combinatorial single-cell analyses of granulocyte-monocyte progenitor heterogeneity reveals an early uni-potent neutrophil progenitor. *Immunity* **53**, 303–318.e5 (2020).
- Dinh, H. Q. et al. Coexpression of CD71 and CD117 identifies an early unipotent neutrophil progenitor population in human bone marrow. *Immunity* **53**, 319–334.e6 (2020).
- Schulte-Schrepping, J. et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* **182**, 1419–1440.e23 (2020).
- van Lochem, E. G. et al. Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: reference patterns for age-related changes and disease-induced shifts. *Cytom. B Clin. Cytom.* **60**, 1–13 (2004).
- Gorczyca, W. et al. Immunophenotypic pattern of myeloid populations by flow cytometry analysis. *Methods Cell. Biol.* **103**, 221–266 (2011).
- Olweus, J. et al. Dendritic cell ontogeny: a human dendritic cell lineage of myeloid origin. *Proc. Natl Acad. Sci. USA* **94**, 12551–12556 (1997).
- Breton, G., Lee, J., Liu, K. & Nussenzweig, M. C. Defining human dendritic cell progenitors by multiparametric flow cytometry. *Nat. Protoc.* **10**, 1407–1422 (2015).
- Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
- See, P. et al. Mapping the human DC lineage through the integration of high-dimensional techniques. *Science* **356**, eaag3009 (2017).
- Lee, J. et al. Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors. *Nat. Immunol.* **18**, 877–888 (2017).
- Olweus, J., Lund-Johansen, F. & Terstappen, L. W. Expression of cell surface markers during differentiation of CD34⁺, CD38^{-lo} fetal and adult bone marrow cells. *Immunomethods* **5**, 179–188 (1994).
- Cowland, J. B. & Borregaard, N. Granulopoiesis and granules of human neutrophils. *Immunol. Rev.* **273**, 11–28 (2016).
- Cassatella, M. A., Ostberg, N. K., Tamassia, N. & Soehnlein, O. Biological roles of neutrophil-derived granule proteins and cytokines. *Trends Immunol.* **40**, 648–664 (2019).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- Mora-Jensen, H. et al. Technical advance: immunophenotypical characterization of human neutrophil differentiation. *J. Leukoc. Biol.* **90**, 629–634 (2011).
- Elghetany, M. T. Surface antigen changes during normal neutrophilic development: a critical review. *Blood Cells Mol. Dis.* **28**, 260–274 (2002).
- Olweus, J., Lund-Johansen, F. & Terstappen, L. W. CD64/Fc gamma RI is a granulocyte-monocytic lineage marker on CD34⁺ hematopoietic progenitor cells. *Blood* **85**, 2402–2413 (1995).
- Gross, S. et al. Characterization and phenotypic analysis of differentiating CD34⁺ human bone marrow cells in liquid culture. *Eur. J. Haematol.* **59**, 318–326 (1997).
- Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2018).
- Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
- Angerer, P. et al. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016).
- Xie, X. et al. Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. *Nat. Immunol.* **21**, 1119–1133 (2020).
- Gupta, S. et al. Sex differences in neutrophil biology modulate response to type I interferons and immunometabolism. *Proc. Natl Acad. Sci. USA* **117**, 16481–16491 (2020).
- Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334.e1310 (2019).
- Mistry, P. et al. Transcriptomic, epigenetic, and functional analyses implicate neutrophil diversity in the pathogenesis of systemic lupus erythematosus. *Proc. Natl Acad. Sci. USA* **116**, 25222–25228 (2019).
- Sinha, S. et al. Dexamethasone modulates immature neutrophils and interferon programming in severe COVID-19. *Nat. Med.* **28**, 201–211 (2022).
- Bost, P. et al. Deciphering the state of immune silence in fatal COVID-19 patients. *Nat. Commun.* **12**, 1428 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Isolation of cells from human BM and peripheral blood samples. BM samples were obtained from healthy donors ($n = 35$; Supplementary Table 4) selected for donation purposes according to the Italian Bone Marrow Donor Registry criteria and following informed, written consent. Fresh BM samples (1.5/2 ml) were collected under aseptic conditions in heparinized sterile syringes, processed using endotoxin-free polypropylene tubes (Greiner Bio-One) and subjected to density gradient centrifugation onto Ficoll-Paque (GE Healthcare). BM-LDCs were then collected, and either immediately processed for flow cytometry analysis, or suspended in α MEM growth medium (Corning) supplemented with 10% low-endotoxin FBS (<0.5 endotoxin unit (EU) ml^{-1}), in the presence of penicillin/streptomycin (from now on termed 'culture medium')²⁹, to be finally distributed in tissue-culture plates (Corning) and preincubated for 20 h at 37 °C in CO₂ (ref. ²⁰). Occasionally, fresh aliquots of BM-LDCs were frozen in 90% FBS (Sigma-Aldrich) plus 10% dimethylsulfoxide (Sigma-Aldrich) and stored in liquid nitrogen, since results obtained by using thawed or unfrozen BM-LDC samples were found fully comparable. Neutrophils from healthy donors were purified by the EasySep Human Neutrophil Enrichment Kit (Stem Cell Technologies; $>99.7\%$ purity) after centrifuging buffy coats onto Ficoll-Paque gradient density³⁰. BM samples, obtained from three patients undergoing alloHSCT for hematological malignancies, were collected at day +21 after alloHSCT (when hematopoiesis is actively reconstituting from donor-derived HSCs). At this time-point, all three patients were in complete remission and had achieved full donor chimerism, according to variable number tandem repeats analysis (available from the authors). The study has been cleared by the Ethics Committee of the Azienda Ospedaliera Universitaria Integrata di Verona (Italy) (CMRI/55742).

Flow cytometry and FACS experiments. For flow cytometry experiments, BM-LDCs, sorted progenitor subpopulations or culture-derived cells were counted, resuspended in 50 μl of PBS buffer (Corning) plus 2% FBS and 2 mM EDTA (Sigma-Aldrich) (from now on termed 'staining buffer'), and subsequently incubated for 10 min in the presence of 5% human serum (Sigma-Aldrich). Cells were then stained for 30 min on ice by fluorochrome-conjugated monoclonal antibodies (mAbs) for either 8-color MACSQuant10 Analyzer (Miltenyi Biotec), or 14-color LSRFortessa X-20 (BD) and MACSQuant16 Analyzer (Miltenyi Biotec) flow-cytometers, as listed in Supplementary Tables 5 and 6. Viability of both fresh and 20-h-preincubated BM-LDC samples was assessed by either propidium iodide (PI) (Sigma-Aldrich) or alternatively Sytox Blue (Invitrogen) exclusion, and reproducibly found $>95\%$. The mean fluorescence intensity (MFI) relative to selected markers was obtained by subtracting either the MFI of the correspondent isotype control mAbs, or the cell autofluorescence (fluorescence minus one, FMO), to their fluorescence value. FlowJo software v.10.7.1 was used for data analysis. For FACS of the myeloid progenitors, 50–100 $\times 10^6$ BM-LDCs were resuspended at 100 $\times 10^6$ per ml and labeled for 45 min at 4 °C (in the dark) with the mAbs reported in Supplementary Table 7. Anti-CD14 and anti-IL5R α mAbs were always included in the sorting panel to exclude, respectively, mature monocytes and eosinophil-committed progenitors. Conversely, we omitted the use of antiCD38 and antiCD117 antibodies since NCPs were found as CD38⁺ and CD117⁺. Cells were then washed and resuspended at 30 $\times 10^6$ per ml in staining buffer, to be ultimately filtered through a 0.35- μm nylon mesh. Cells were finally sorted by using a FACSAria Fusion (BD) cell sorter equipped with 85- μm nozzle, immediately centrifuged, resuspended in α MEM medium, counted and used for experiments. Alternatively, sorted cells were lysed in RLT buffer (Qiagen) for RNA extraction³¹. Sorted cell populations displayed a $>95\%$ purity, as verified by flow cytometry analysis.

In vitro differentiation assay. To evaluate the differentiation potential of sorted BM progenitors, the latter cells were cultured on the MS-5 stromal cell line (ACC 441, Leibniz-Institut DSMZ), according to protocols previously described²⁹. Briefly, the day before the coculture experiments, MS-5 cells cultured in α MEM medium at 95% confluence were incubated with 10 $\mu\text{g ml}^{-1}$ mitomycin C (Sigma-Aldrich) for 3 h at 37 °C. Then, after treatment with 0.05% trypsin/EDTA (Corning), MS-5 cells were collected and resuspended in α MEM medium at 0.25 $\times 10^6$ per ml, to be finally seeded in round-bottom 96-well tissue-culture plates for 24 h. Thus, 1–3 $\times 10^3$ of the various sorted BM progenitors were resuspended in 100 μl of α MEM medium, seeded on top of MS-5 cells and incubated with 20 international units (IU) ml^{-1} Flt3L (Miltenyi Biotec), 10 IU ml^{-1} SCF (Miltenyi Biotec) and either 100 IU ml^{-1} GM-CSF (Peprotech)²⁹ or, as in previous studies^{52,53}, 6,500 IU ml^{-1} G-CSF (Myelostim, Italfarmaco Spa), to compose, respectively, the SFG²⁹ or SFGc cocktail. Cells were then collected from the MS-5 cells starting at day 1, and up to day 14, of the culture, either for flow cytometry staining, or for functional studies. Expansion of purified progenitor populations is reported as mean fold-change of live, CD45⁺ output cells from the number of cells plated as input at time 0.

Morphological analysis of neutrophil precursors and culture-derived cell populations. Cytospins prepared from either sorted neutrophil precursors or culture-derived cell populations were stained by the May-Grunwald Giemsa procedure. Slides were analyzed by a Leica DM 6000 B microscope, equipped with a Leica DFC 300FX Digital Color Camera (Leica Microsystems).

O₂⁻ production and phagocytosis assay. Either progenitor-derived neutrophils collected from NCPs cultured for 7 d, or freshly isolated healthy donor neutrophils, were washed and resuspended at 0.25 $\times 10^6$ per ml in HBSS supplemented with 10% FBS, containing 1 mM CaCl₂ and 5 mM glucose. O₂⁻ production in response to 20 ng ml^{-1} PMA (Sigma-Aldrich) was assessed by the Cytochrome C reduction assay⁵⁴. To assess phagocytosis capacity, 0.025 $\times 10^6$ cells per 100 μl were incubated in Eppendorf tubes at 37 °C, in the absence/presence of 20 $\mu\text{g ml}^{-1}$ unopsonized zymosan particles. Cells were intermittently resuspended and, after 30 min, treated with an excess of cold HBSS and centrifuged to stop phagocytosis. Cytospin preparations of these samples were stained by the May-Grunwald Giemsa procedure, and slides analyzed by a Leica DM 6000 B microscope, equipped with a Leica DFC 300FX Digital Color Camera (Leica Microsystems).

Immunocytochemistry. Sorted cells were spotted on polarized slides and immediately fixed by 95% ethanol. Before the immunostaining procedure, slides were placed for 10 min in methanol containing 0.3% H₂O₂ for endogenous peroxidase quenching. Slides were rehydrated for 3 min in ethanol at 95 °C, and then washed in tap and distilled water. Slides were then washed in pH 7.4/7.6 Tris-HCl for 5 min before incubation for 2 h at room temperature with antibodies towards MPO (Rb polyclonal, dilution 1:5,000, A0398 Agilent Technologies), ELANE (1:200 dilution, clone NP57, Agilent Technologies), α -Defensins (DEFA; dilution 1:70, clone H-2, Santa Cruz Biotechnology) and LTF (Rb polyclonal, dilution 1:500, hpa059976, Sigma-Aldrich). Reactions were revealed by using the EnVision Detection System HRP, followed by diaminobenzidine for 5 min. Subsequently, slides were counterstained by Meyer's hematoxylin for 10 min and coverslipped. Immunostained slides were photographed using a DP-73 Olympus digital camera, mounted on an Olympus BX60 microscope ($\times 600$ magnification, square side: 20 μm).

CFU assay. The CFU assay was performed by using the 'complete Stem MACS HSC-CFU medium' (containing 20 ng ml^{-1} G-CSF, 20 ng ml^{-1} GM-CSF, 20 ng ml^{-1} IL-3, 20 ng ml^{-1} IL-6 and 3 U ml^{-1} Epo, from Miltenyi Biotec), according to the manufacturer's protocol.

Adoptive transfer experiments in mice. NOD.Cg-Prkdc^{scid}Il2rg^{tm1Sug}/JicTac (NOG) mice were purchased from Taconic Biosciences and maintained under specific-pathogen-free conditions at the Verona University animal facility. All animal experiments were approved by the local Ethics Committee (<https://www.univr.it/it/cirsal>), and conducted according to the guidelines of the Federation of European Laboratory Animal Science Associations. Animal experiments were performed according to national (protocol number 12722, approved by the Ministerial Decree Number 14/2012-B of 18 January 2012, and protocol number BR15/08, approved by the Ministerial Decree Number 925/2015-PR of 28 August 2015), as well as European, laws and regulations. All animals were housed in conditions of 22 °C and 45–65% humidity with 12-h light cycle. For adoptive transfer into NOG mice, animals (4-week-old females) were intramuscularly injected with 10⁸ multiplicity of infection (MOI) of recombinant AAV9 vectors coding for human CSF2, CSF3, IL-3 and CXCL8, to favor the expansion of the myeloid compartment. In brief, cytokine complementary DNAs were purchased from GeneScript and AAV9 vectors were produced by Prof. M. Giacca's laboratory (International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste and King's College, London) by co-transfection of HEK293T cells (ACC 635, Leibniz-Institut DSMZ). After 2 weeks, mice were γ -irradiated (1.2 Gy), and 6 h later engrafted with a total of 10⁵ pooled NCPs (to have enough cells) via intraocular injection. After 11 d, mice were euthanized, and BM and blood analyzed by FACS (3 out of 3 mice were successfully engrafted). For adoptive transfer into Human Immune Reconstituted (HIR) mice, animals (4-week-old females) were γ -irradiated (1.2 Gy) and, after 6 h, engrafted with 10⁵ hHLA-A2-CD34⁺ cells (catalog no.: 4M-105, Lonza) via tail vein injection (intravenously). AAV9 vectors coding for human CSF2, CSF3, IL-3 and CXCL8 were intramuscularly administered at week 4 after engraftment, as already detailed. Hematopoietic engraftment was considered adequate when 25% of human CD45-expressing cells could be detected in peripheral blood of mice by FACS, according to the formula: frequency of hCD45⁺ cells/(frequency of hCD45⁺ + frequency of mCD45⁺ cells) $\times 100$. At week 12 post-engraftment, mice were treated with a single nonmyeloablative dose of busulfan (10 mg kg^{-1}) and, 24 h later, animals received an intratibial injection of a total of 1.5 $\times 10^4$ pooled HLA-A2⁺ NCPs. After 14 d, mice were euthanized, and BM and blood were analyzed by FACS (6 out of 6 mice were successfully engrafted by HLA-A2⁺ NCPs).

RNA-seq. Total RNA was extracted by the RNeasy Mini Kit (Qiagen) after cell lysis⁵⁵. Libraries for transcriptome analysis were prepared using the Smart-seq2 protocol⁵⁶. Briefly, 2 ng of total RNA was copied into a first-strand cDNA by reverse transcription and template-switching, using oligo (dT) primers (5'-AAGCAGTGGTATCAACGCAGAGTACT30VN-3', from Thermo Fisher) and a locked nucleic acid (LNA)-containing template-switching oligo (5'-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3', from Exicon). The resulting cDNA was pre-amplified, purified and tagged with Tn5 transposase. cDNA fragments generated after tagmentation were gap-repaired, enriched by

PCR and purified to create the final cDNA library. Libraries were sequenced on the Illumina NextSeq 500 in single-read mode (1 × 75 cycles) at the Centro Piattaforme Tecnologiche of the University of Verona.

scRNA-seq. Sorted cells were labeled by using the BD Single-Cell Multiplexing Kit (BD Biosciences), strictly following the manufacturer's protocol. Such a procedure allowed us to combine five samples (NCP1s, NCP2s, NCP3s, NCP4s and cMoPs) into a single pool. Each sample was washed twice in FACS buffer and resuspended in cold BD Sample Buffer (BD Biosciences). Only samples with high viability (>85%, as evaluated by the BD Rhapsody scanner) were used for sequencing. Samples from the same donor were pooled in equal amounts to achieve approximately 15,000 cells in 620 µl, and loaded onto a BD Rhapsody cartridge for an incubation of 20 min at room temperature. Then, Cell Capture Beads (BD Biosciences) were added to the cartridge and incubated at room temperature for 3 min, and thereafter cartridges were washed. Cells were then lysed, and the released mRNAs were captured by Cell Capture Beads. mRNAs were then retrieved, to be washed before performing reverse transcription and treatment with Exonuclease I. cDNA libraries were prepared by using the BD Rhapsody Whole Transcriptome Analysis Amplification kit and BD Single-Cell Multiplexing kit (BD Biosciences). Size-distribution (quality) of final libraries was assessed by Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape and quantified using a Qubit Fluorometer using the Qubit dsDNA HS Kit (Thermo Fisher, no. Q32854). Sequencing was performed in paired-end mode (2 × 75 cycles) on a NextSeq 500 System (Illumina). This procedure was utilized for three different BMs, whose data were integrated as outlined below.

RNA-seq computational analysis. Computational analysis of transcriptome datasets generated by Smart-seq2 was performed by using the bioinformatic pipeline, as previously described³⁷. Briefly, after quality filtering, according to the Illumina pipeline, removal of contaminant adapters and base quality trimming were performed using Trim Galore! (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) script. Gene counts were normalized among various samples using DESeq2 (ref. ³⁶), and only genes coding for protein and long noncoding RNA were retained for downstream analysis. DEGs were identified using DESeq2, by using as selection parameter adjusted *P* value lower than 0.01 and likelihood ratio test³⁶. Batch effects were removed using the limma package's 'removeBatchEffect' function before performing PCA. PCA was performed on DEGs by using Bioconductor/R package *pcaExplorer* v.2.10.0.

Computational inference of developmental path. Hierarchical clustering was performed using the Euclidean distance and Ward aggregation as criteria. To find a suitable linear order within the hierarchical clustering dendrogram (HSCs/MPPs, CMPs, NCPs, PMs, MYs, MMs, BCs, SCs and mature blood neutrophils), we used the optimal leaf ordering seriation method of R package *seriation*, v.1.2-9 (ref. ⁵⁸).

Seven Bridges processing for scRNA-seq data. After demultiplexing of bcl files by using *Bcl2fastq2* v.2.20 from Illumina and assessment of reads quality, paired-end scRNA-seq reads were then filtered for valid cell barcodes using the barcode whitelist provided by BD. Then, sequenced reads were aligned to the hg38 human transcriptome and the expression of transcripts in each cell was quantified via the standard Rhapsody analysis pipeline (BD Biosciences) on Seven Bridges (<https://www.sevenbridges.com>), following manufacturer's recommendations.

Seurat workflow for scRNA-seq data analysis. The R package Seurat⁵⁹ was utilized for all downstream analysis. For each single-cell dataset, the number of detected genes, the number of unique molecular identifiers (UMIs), as well as the fraction of UMIs corresponding to mitochondrial features, which altogether reflect the transcriptome quality of each cell, were calculated. Only cells that transcribed at least 200 genes, and only genes that were expressed in at least 10 cells, were included in the analysis. In sum, 15,969 cells and 17,398 genes were obtained. Next, cells with more than 25% mitochondrial transcripts and with a number of genes above the 99th percentile or below the 1st percentile were removed for downstream analysis. Therefore, after such a rigorous quality control, a total of 15,057 cells were analyzed (4,465, 6,958 and 3,634 from BMs of, respectively, donor 1, donor 2 and donor 3). To examine cell cycle variation in our data, by using *CellCycleScoring* function, we assigned a score to each cell, based on its expression of G2/M and S phase markers. To remove batch effects across data from different donors, we performed dataset integration using *SCTransform* integration workflow (<https://satijalab.org/seurat/v3.2/integration.html>). Normalization and detection of highly variable genes were performed using the difference between the S and G2M cell cycle scores and the percentage of mitochondrial UMIs as covariates. To identify the integration of anchor genes among the three datasets from different BMs, the *FindIntegrationAnchors*() function was used applying default parameters. Using Seurat's *IntegrateData*(), samples were combined into one object. These 'integrated' batch-corrected values were then set as the 'default assay', and gene expression values were scaled before running PCA. The dimensional reduction of the integrated dataset was computed by summarizing the first 50 principal components and visualized in a two-dimensional UMAP representation. Clustering was conducted using the *FindNeighbors*() and *FindClusters*() functions using the same 50 principal components, and a resolution parameter set to 0.3. Differential

expression tests were performed using the *FindAllMarkers*() function. DEGs were identified using the nonparametric Wilcoxon rank sum test, based on normalized data. *P* value adjustment was performed using Bonferroni correction based on the total number of genes in the dataset. Genes with >0.25 log-fold changes, at least 25% expressed in tested groups and Bonferroni-corrected *P* < 0.01 were considered as significant DEGs. Euclidean distance between clusters was calculated on DEGs and hierarchical clustering was performed in R, using the 'ward.D' linkage clustering. The average gene expression of clusters was calculated using the function *AverageExpression*().

GO of DEGs from bulk RNA-seq and scRNA-seq. GO analysis was performed on DEGs from bulk RNA-seq and scRNA-seq for the cellular component, biological process and molecular function ontology domains by using the Bioconductor/R package *clusterProfiler* (v.3.14.3)⁶⁰.

Analytic comparison among transcriptomes obtained from NCPs, eNePs and COVID-19 proNeus. RNA-seq datasets from eNePs and other more mature neutrophil precursors (that is, N1 without eNeP, N2 and mixed N3/N4/N5)²⁴ were analyzed exactly as our bulk RNA-seq analyses. For COVID-19 proNeus and preNeus²⁵, cells categorized as 'pre-Neutrophils' and 'pro-Neutrophils' (that is, clusters 6 and 8, respectively, from Fig. 7b of Schulte-Schrepping et al. paper²⁵) were retrieved from scRNA-seq datasets and then normalized to count per million, by using the *NormalizeData*() function of Seurat R packages and applying the following settings: 'normalization.method = RC' and 'scale.factor = 1e6'. Subsequently, pseudo-bulk RNA-seq analyses of proNeus and preNeus from selected donors ('BN-12', 'BN-18', 'BN-19') were computed by evaluating the mean expression for each gene. Each of the two datasets^{24,25} was merged with ours, and only genes present in all three datasets were included in the analysis and transformed to log₂ pseudo-count values for downstream assays.

Trajectory analysis and pseudotime calculation. Trajectory analysis was performed by using the *destiny* algorithm v.3.01 (ref. ⁴³). In brief, the neutrophil precursor space was subsetted and the diffusion map was calculated based on the top 2,000 variable genes with a sum of at least 10 counts over all cells. Based on the diffusion map, a diffusion pseudotime was calculated to infer a transition probability between the different cell states of the neutrophils.

Statistical analysis. Data are expressed as mean ± s.d. or s.e.m. of the number of indicated experiments. Where applicable, normality distribution was estimated using the D'Agostino–Pearson or Shapiro–Wilk normality test. Statistical evaluation for normally distributed data was performed by using one-way or two-way analysis of variance (ANOVA), followed by Tukey's or Bonferroni's post hoc test, respectively. Non-normally distributed data were assessed with Mann–Whitney test, or, for multiple group comparison, with Kruskal–Wallis or Friedman test followed by Dunn's multiple comparison test. The tests used are indicated in the respective figure legends. Values of *P* < 0.05 were considered statistically significant. Statistical analysis was performed by GraphPad Prism v.7.0 software. No randomization was performed in experiments with BM cells since this was not relevant. Samples were allocated into experimental groups based on gating strategy. In animal experiments, mice were randomly assigned to each group. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications^{24,44,45}.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Raw datasets have been submitted to the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) and are available under the accession number GSE164687. Bulk RNA-seq datasets of eNePs and HSCs/CMPs were downloaded from GSE157103 and GSE113182, respectively. COVID-19 proNeus and preNeus scRNA-seq datasets were downloaded from <https://www.fastgenomics.org>. Source data are provided with this paper.

Code availability

R scripts for data processing are available through <https://github.com/fbianchetto/Neutrophil-Committed-Progenitors>.

References

- Calzetti, F. et al. The importance of being "pure" neutrophils. *J. Allergy Clin. Immunol.* **139**, 352–355.e6 (2017).
- Zimmermann, M. et al. IFN α enhances the production of IL-6 by human neutrophils activated via TLR8. *Sci. Rep.* **6**, 19674 (2016).
- Tura, O. et al. Optimal ex vivo expansion of neutrophils from PBSC CD34⁺ cells by a combination of SCF, Flt3-L and G-CSF and its inhibition by further addition of TPO. *J. Transl. Med.* **5**, 53 (2007).
- Jie, Z. et al. Large-scale ex vivo generation of human neutrophils from cord blood CD34⁺ cells. *PLoS ONE* **12**, e0180832 (2017).

54. Serra, M. C., Calzetti, F., Ceska, M. & Cassatella, M. A. Effect of substance P on superoxide anion and IL-8 production by human PMNL. *Immunology* **82**, 63–69 (1994).
55. Tamassia, N. et al. The MyD88-independent pathway is not mobilized in human neutrophils stimulated via TLR4. *J. Immunol.* **178**, 7344–7356 (2007).
56. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
57. Bianchetto-Aguilera, F. et al. Deciphering the fate of slan⁺-monocytes in human tonsils by gene expression profiling. *FASEB J.* **34**, 9269–9284 (2020).
58. Hahsler, M., Hornik, K. & Buchta, C. Getting things in order: an introduction to the R package seriation. *J. Stat. Softw.* **25**, 1–34 (2008).
59. Butler, A. et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
60. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

Acknowledgements

This work was supported by grants to M.A.C. (from AIRC grant no. IG-20339; MIUR-PRIN grant no. 2017714E75_004; and Fondazione Cariverona) and to N.T. (grant no. GR-2016-02361263). We thank P. Mazzi for preparing cytopins, E. Pietronigro for slide acquisition, S. Pasqualato (European Institute of Oncology, Milan, Italy) for providing Tn5 transposase and A. Carraro (Centro Trapianti Fegato, AOUI Verona, Italy) and V. Bergamini (UOC Ostetricia e Ginecologia B, AOUI Verona, Italy) for providing spleen and cord blood samples. We thank G. Zini (University Cattolica, Hematology Department, Rome) for her professional consultancy about NCP morphology. Centro

Piattaforme Tecnologiche (CPT) of the University of Verona has been instrumental for the access to the flow cytometry/cell analysis and genomic/transcriptomic platforms.

Author contributions

F.C. and G.F. performed and analyzed flow cytometry, cell sorting and in vitro differentiation experiments. C.C. and A.M. were involved in cell sorting experiments. N.T., M.C. and S.G. prepared samples for RNA-seq and scRNA-seq. N.T. and F.B.-A. analyzed RNA-seq and scRNA-seq data. F.C. and I.S. performed CFU assay. S.L. and W.V. performed and analyzed ICC. F.C., G.F., S.C., S.U., A.M. and V.B. were involved in adoptive transfer experiments. F.B., M.B. and C.T. provided patient samples. F.C., G.F., N.T., P.S. and M.A.C. designed and wrote the manuscript. M.A.C. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

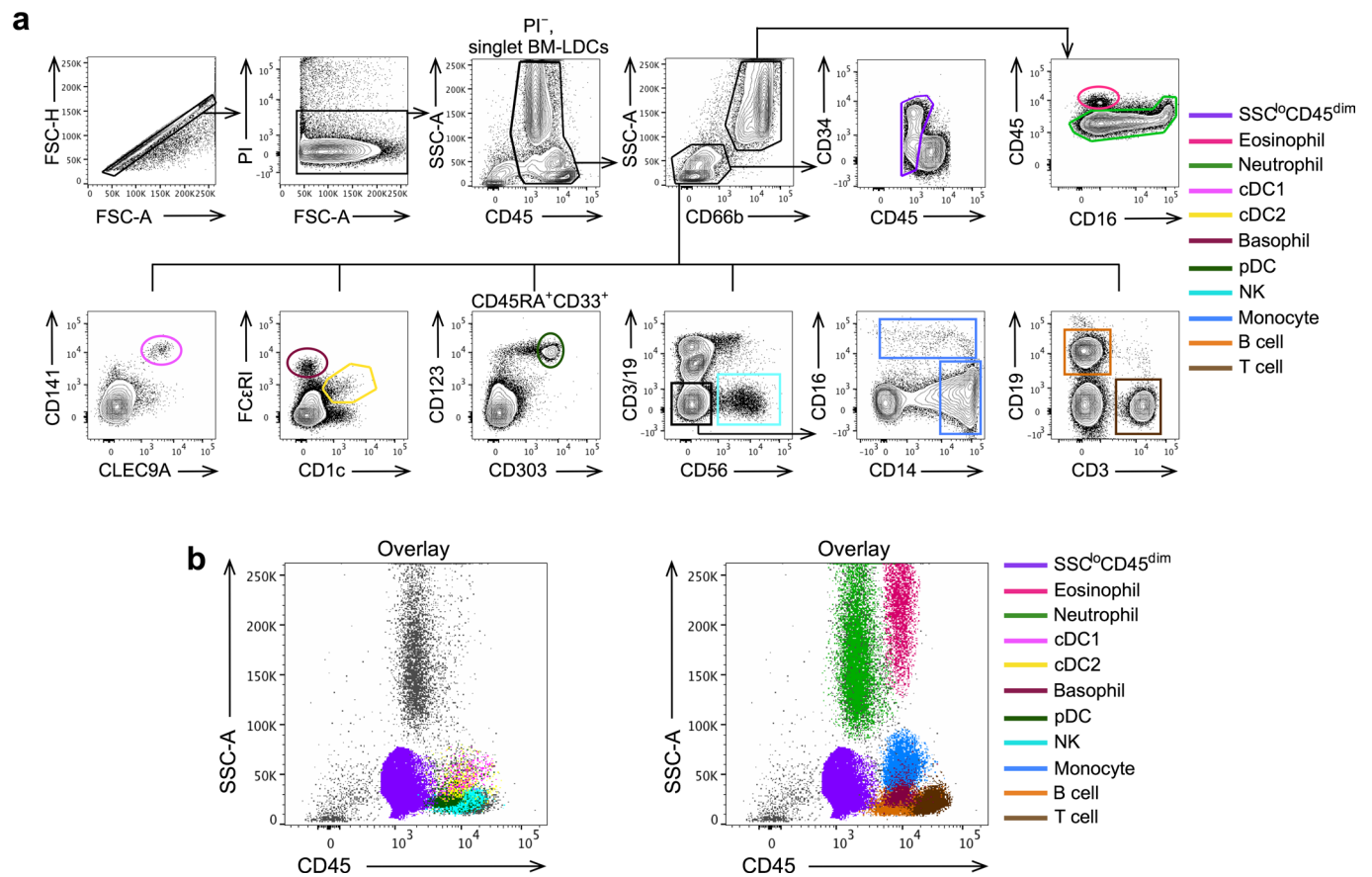
Extended data is available for this paper at <https://doi.org/10.1038/s41590-022-01189-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41590-022-01189-z>.

Correspondence and requests for materials should be addressed to Marco A. Cassatella.

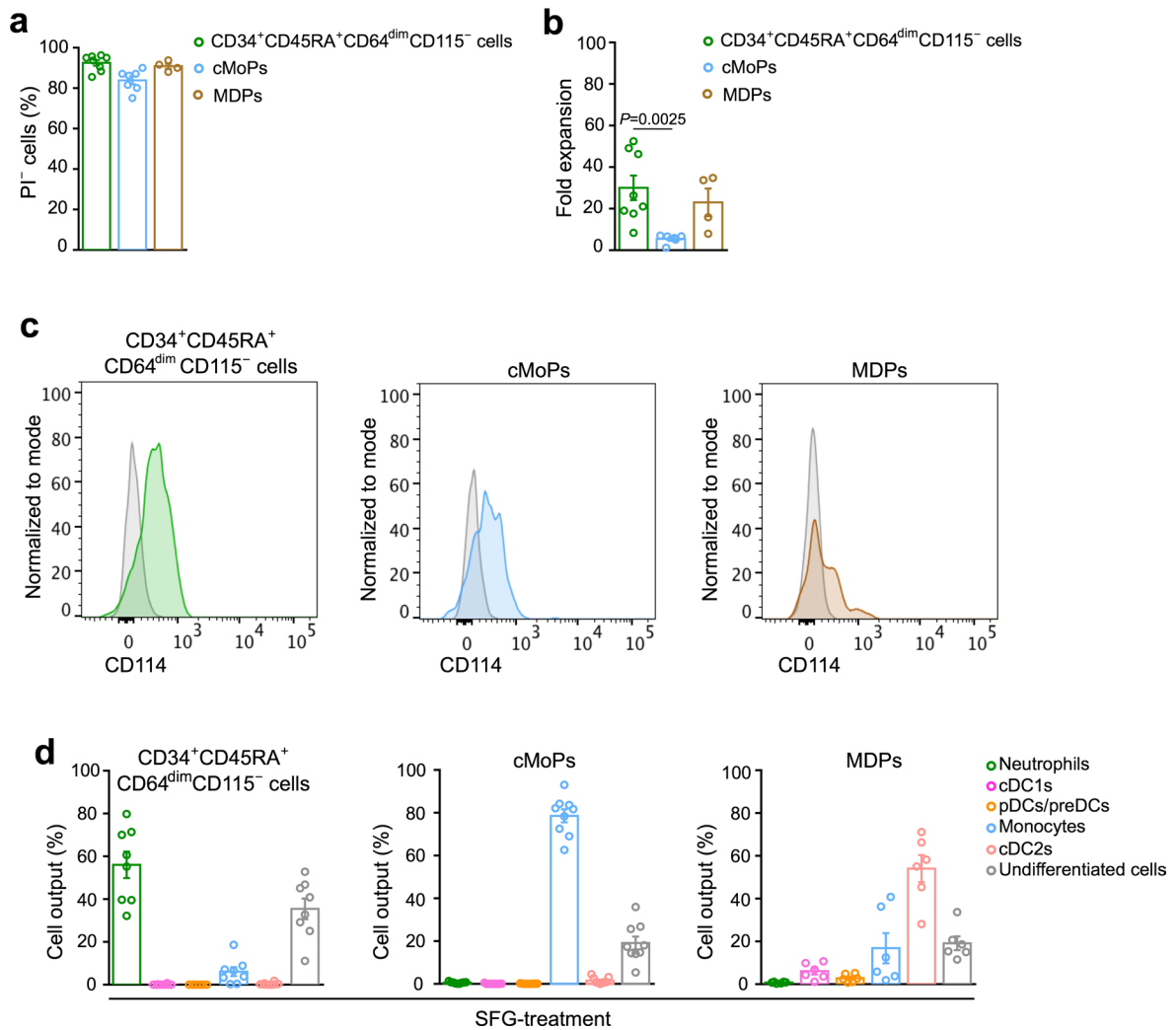
Peer review information *Nature Immunology* thanks Pierre Guermonprez and Hongbo Luo for their contribution to the peer review of this work. Ioana Visan was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

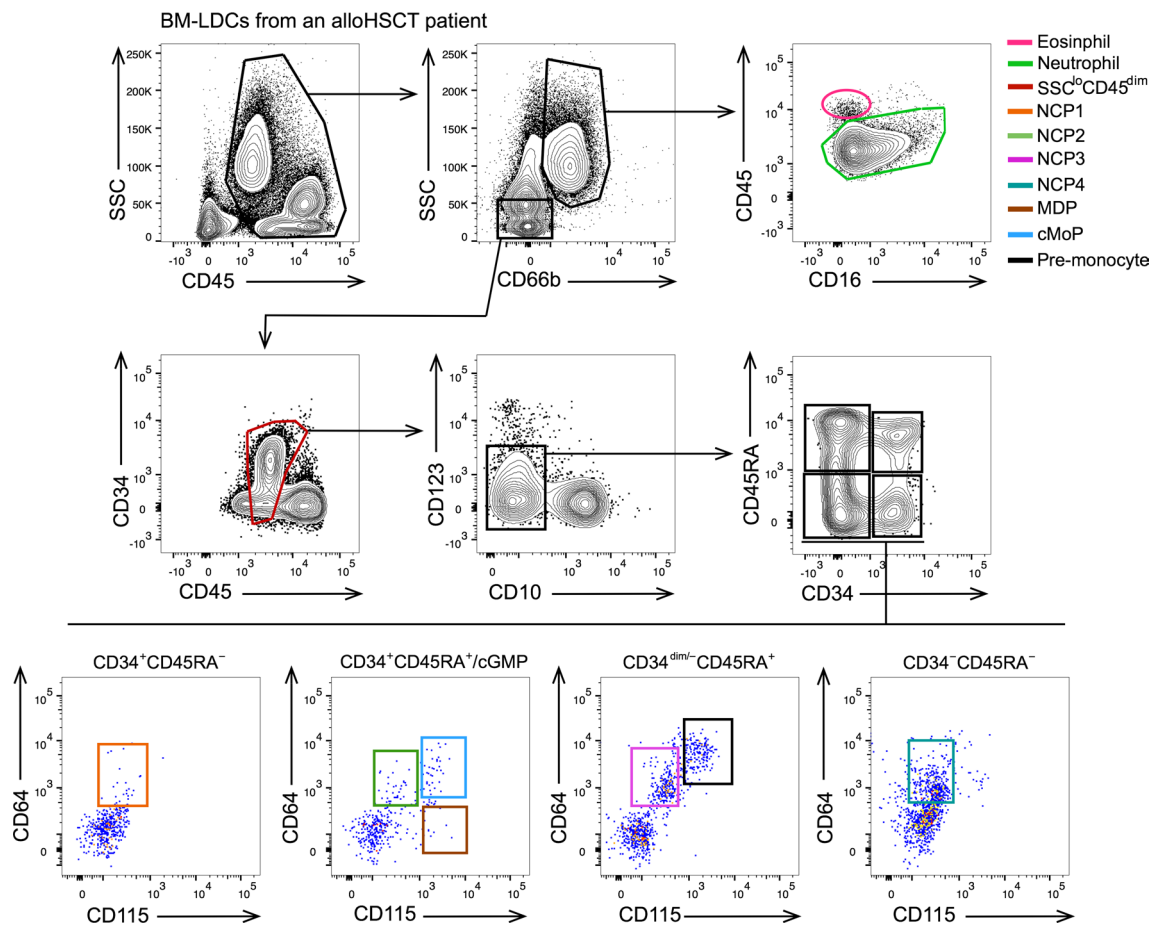


Extended Data Fig. 1 | Gating strategy to identify lineage-positive cells and immature $SSC^{lo}CD45^{dim}$ myeloid/lymphoid progenitors within BM-LDCs.

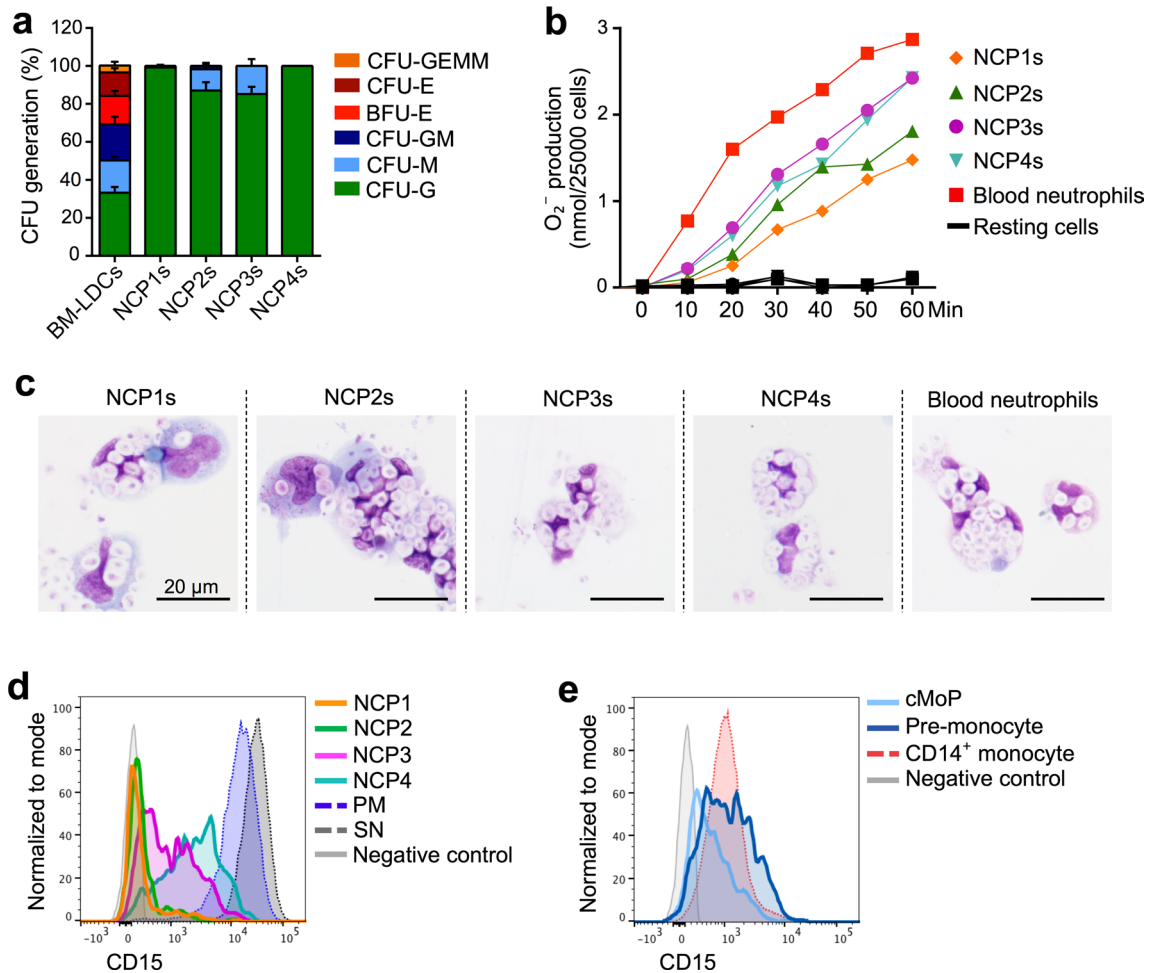
a, Flow cytometry workflow for the identification of mature leukocyte populations and $SSC^{lo}CD45^{dim}$ immature myeloid/lymphoid progenitors within BM-LDCs. Mature cell populations were identified as: $SSC^{hi}CD66b^{+}CD16^{-}CD45^{hi}$ eosinophils (pink); $SSC^{hi}CD66b^{+}CD16^{-}/^{+}CD45^{+}$ neutrophils (green); $SSC^{lo}CD141^{+}CLEC9A^{+}$ cDC1s (magenta); $SSC^{lo}FcεRI^{+}CD1c^{-}$ basophils (dark red); $SSC^{lo}FcεRI^{dim}CD1c^{+}$ cDC2s (yellow); $SSC^{lo}CD45RA^{+}CD33^{+}CD123^{+}CD303^{+}$ pDCs (green); $SSC^{lo}CD3^{-}CD19^{-}CD56^{+}$ NK cells (light blue); $SSC^{lo}CD3^{-}CD19^{-}CD56^{-}CD14^{+}/^{-}CD14^{+}/^{-}CD16^{+}/^{-}$ total monocytes (light blue); $SSC^{lo}CD3^{+}CD19^{-}$ T cells (brown); $SSC^{lo}CD3^{-}CD19^{+}$ B cells (orange). Immature myeloid/lymphoid progenitors were identified as $SSC^{lo}CD45^{dim}$ cells (purple). **b**, Dot plot overlays of lineage-positive/mature cells and $SSC^{lo}CD45^{dim}$ immature progenitors.



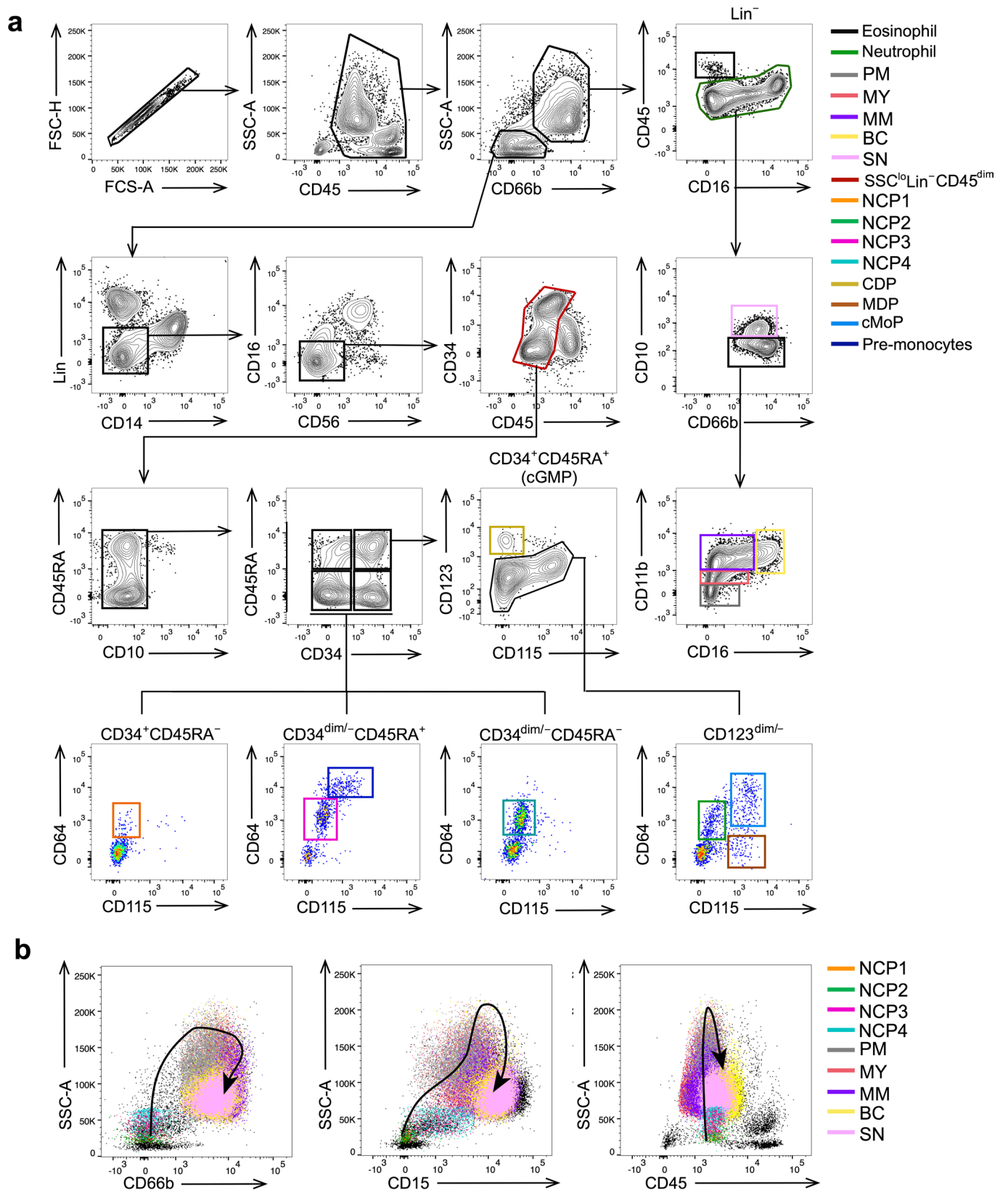
Extended Data Fig. 2 | Biological features of CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells, cMoPs and MDPs prior to, and after treatment with, either SFGc, or SFG. **a, Bar graphs depicting the percentage of PI⁻, live cells generated by CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells (n = 8), cMoPs (n = 7) and MDPs (n = 4) treated with SFGc for 7 days. Data represent means \pm s.e.m. **b**, Bar graphs showing the fold expansion of CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells (n = 8), cMoPs (n = 6) and MDPs (n = 4) treated with SFGc for 7 days. Data represent means \pm s.e.m. Analysis was performed by using Kruskal-Wallis test, Dunn's post-hoc test. **c**, Representative histogram overlays showing the expression of CD114 by CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells (green), cMoPs (light blue) and MDP (brown) populations. Gray profile corresponds to isotype control (n = 3). **d** Bar graphs reporting the percentages of neutrophils (green), cDC1s (magenta), pDCs/preDCs/DC5s (orange), monocytes (light blue) and cDC2s (pink) generated by CD34⁺CD45RA⁺CD64^{dim}CD115⁻ cells (n = 8), cMoPs (n = 9) and MDPs (n = 6) treated with SFG for 7 days. Data represent means \pm s.e.m.**



Extended Data Fig. 3 | Reconstitution of NCPs and mature neutrophils in BMs of allogeneic HSC-transplanted patients. Flow cytometry strategy for the identification of reconstituting neutrophils (green), eosinophils (pink), NCP1s (orange), NCP2s (green), NCP3s (magenta), NCP4s (turquoise), cMoPs (light blue), MDPs (brown) and pre-monocytes (black) within BM-LDCs of reconstituting BM (at day +21) from a patient undergoing allogeneic hematopoietic stem cell transplantation (alloHSCt) (n=3).



Extended Data Fig. 4 | Functionality of neutrophils derived from NCPs, as well as CD15 expression by both neutrophil and monocyte precursors among BM-LDCs and their mature counterparts. **a**, Histogram plot showing the percentages and the type of colonies generated after 14 days of NCP and BM-LDC culture ($n=3$). Data represent means \pm s.e.m. **b**, Representative plot showing the time-course of O_2^- production by neutrophils derived from NCPs cultured for 7 days with SFGc (orange, green, magenta and turquoise) as compared to blood neutrophils (red) ($n=3$). **c**, Representative cytopins showing the phagocytosis of zymosan particles by neutrophils derived from NCPs cultured for 7 days with SFGc ($n=3$). **d,e**, Histogram overlays showing the expression of CD15 by NCPs (orange, green, magenta and turquoise), PMs (purple), SNs (black), cMoPs (light blue), pre-monocytes (blue) and CD14⁺ monocytes (red) as compared to the fluorescence-negative control (light gray) ($n=5$).

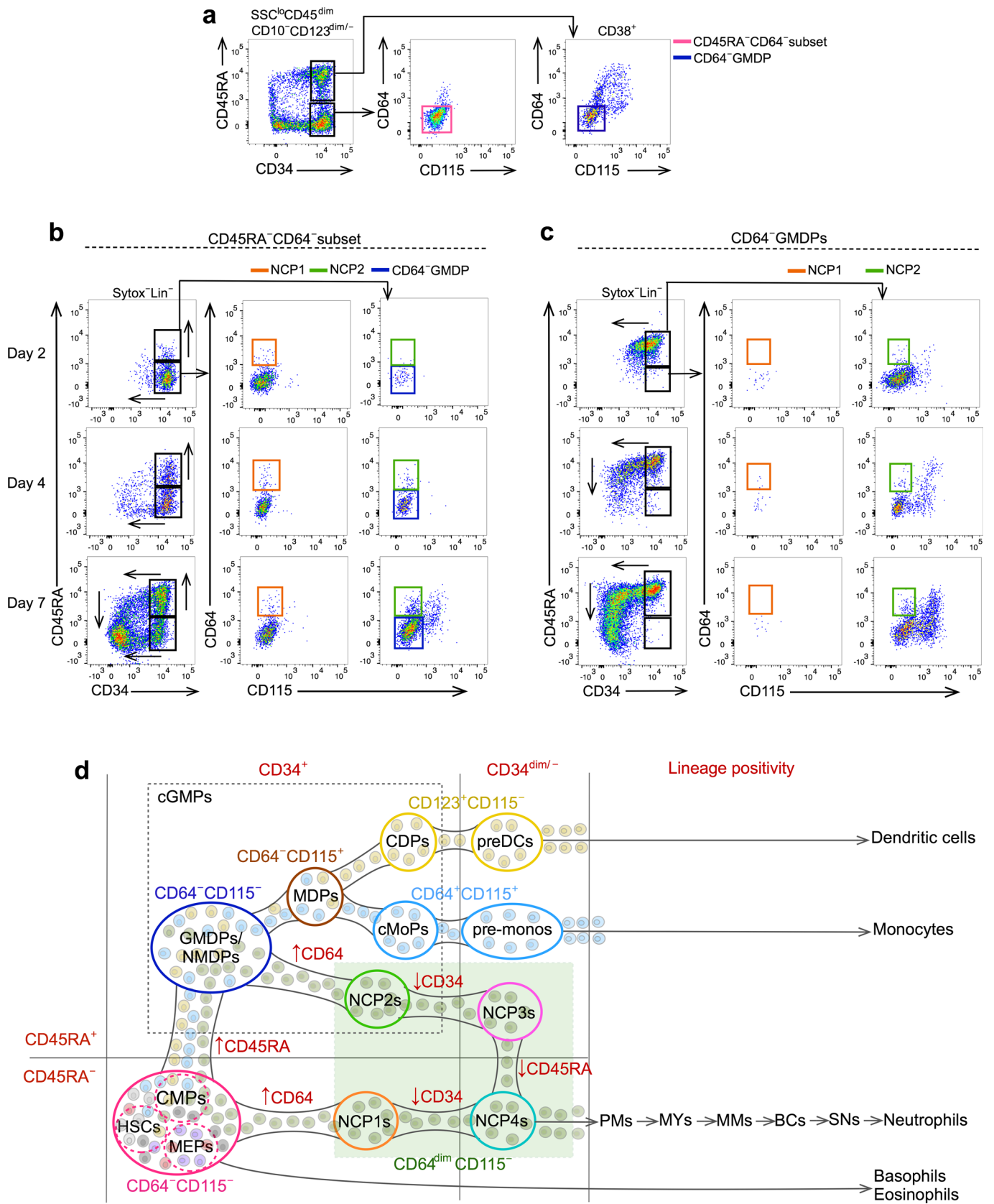


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | 14-color flow cytometry antibody panel to identify NCPs, monocyte/DC progenitors and immature neutrophil populations

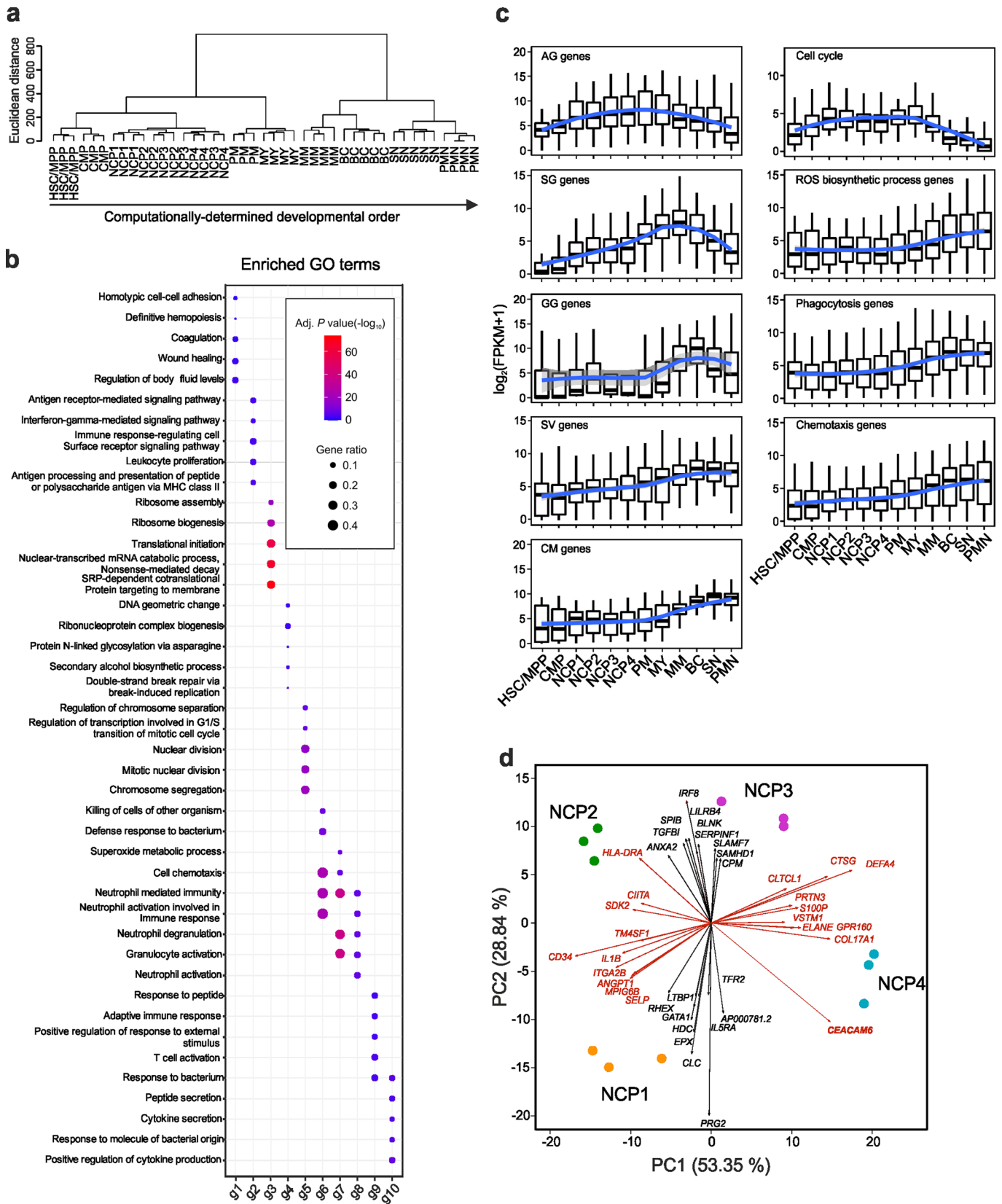
in human BMs. a. Flow cytometry workflow (Supplementary Table 6) showing the gating strategy to identify myeloid progenitors within the $SSC^{lo}CD66b^{-}Lin(CD3/CD19/CD1c/CD141)^{-}CD14^{-}CD16^{-}CD56^{-}CD45^{dim}$ region of BM-LDCs (dark red): $CD34^{+}CD45RA^{+}$ cGMPs, $CD34^{+}CD45RA^{+}CD123^{+}CD115^{-}$ CDPs (yellow), $CD34^{+}CD45RA^{+}CD123^{dim/-}CD64^{dim}CD115^{-}$ NCP2s (green), $CD34^{+}CD45RA^{+}CD123^{dim/-}CD64^{+}CD115^{+}$ cMoPs (light blue), $CD34^{+}CD45RA^{+}CD123^{dim/-}CD64^{-}CD115^{+}$ MDPs (brown), $CD34^{+}CD45RA^{-}CD123^{dim/-}CD64^{dim}CD115^{-}$ NCP1s (orange), $CD34^{dim/-}CD45RA^{+}CD123^{dim/-}CD64^{dim}CD115^{-}$ NCP3s (magenta), $CD34^{dim/-}CD45RA^{+}CD123^{dim/-}CD64^{++}CD115^{+}$ pre-monocytes (blue), $CD34^{dim/-}CD45RA^{-}CD123^{dim/-}CD64^{dim}CD115^{-}$ NCP4s (turquoise). The same flow cytometry workflow identify $SSC^{hi}CD66b^{+}CD16^{-}CD45^{hi}$ eosinophils (black) as well as all the neutrophil precursors within the $SSC^{hi}CD66b^{+}Lin(CD3/CD19/CD1c/CD141)^{-}$ region (green): $CD66b^{+}CD10^{-}CD11b^{-}CD16^{-}$ PMs (grey), $CD66b^{+}CD10^{-}CD11b^{dim/+}CD16^{-}$ MYs (light red), $CD66b^{+}CD10^{-}CD11b^{+}CD16^{+}$ MMs (purple), $CD66b^{+}CD10^{-}CD11b^{+}CD16^{++}CD10^{-}$ BCs (light yellow), and $CD66b^{+}CD11b^{+}CD16^{++}CD10^{+}$ SNs (pink).

b. Dot plot overlays depicting the phenotype variation (black arrows) of the populations composing the neutrophil maturation cascade in terms of SSC parameter and CD66b, CD15 and CD45 marker modulation.



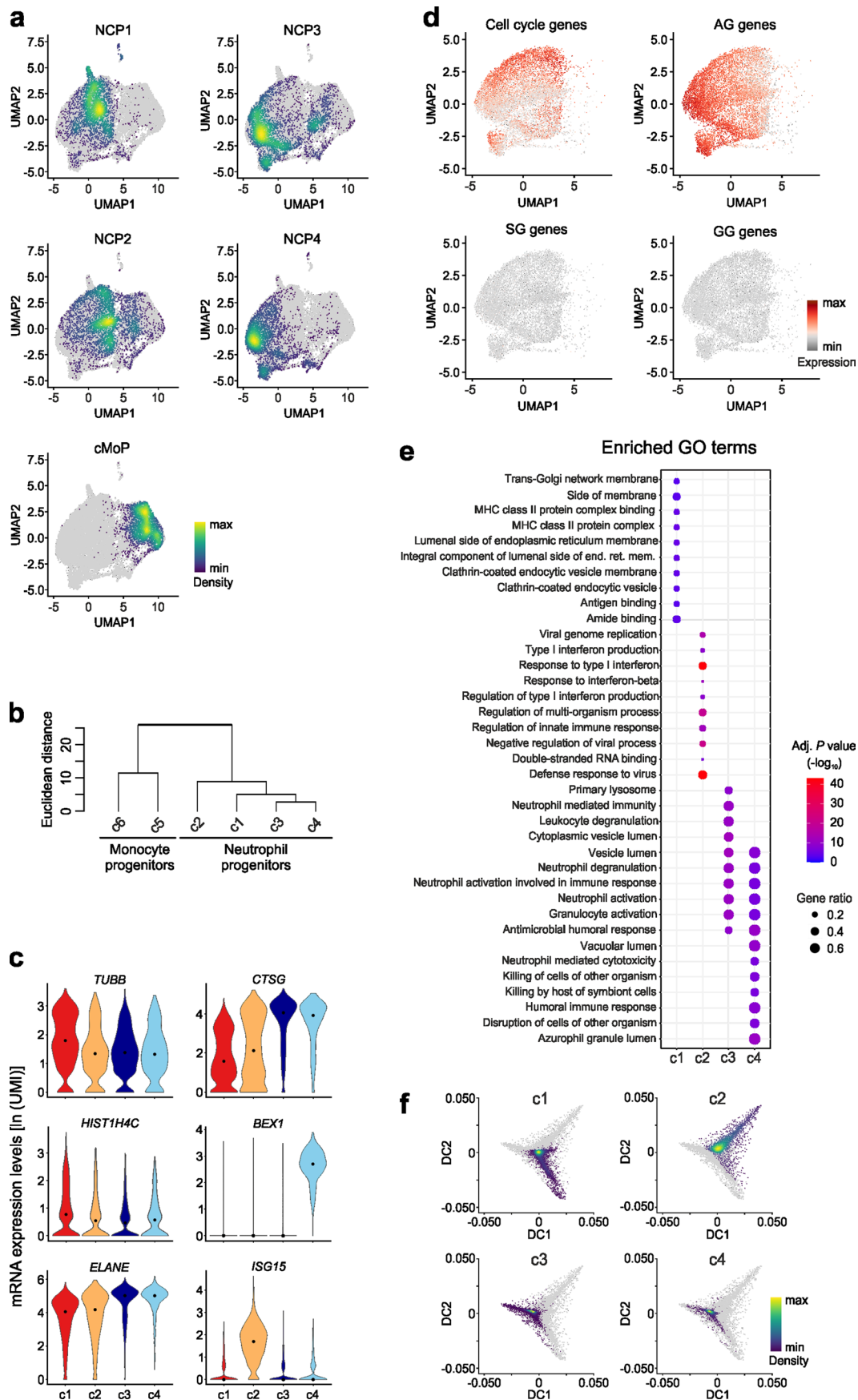
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | NCP1s and NCP2s derive from phenotypically distinct CD34⁺ progenitors. **a**, Flow cytometry strategy for the identification of SSC^{lo}CD45^{dim}CD10⁻CD123^{dim/-}CD34⁺CD45RA⁻CD64⁻ subset (CD45RA⁻CD64⁻ subset, pink) and SSC^{lo}CD45^{dim}CD10⁻CD123^{dim/-}CD38⁺CD34⁺CD45RA⁺CD64⁻CD115⁻GMDPs (CD64⁻GMDPs, dark blue). **b,c**, Plots showing the differentiation potential of CD45RA⁻CD64⁻ subset (**b**) and CD64⁻GMDPs (**c**) based on the changes of CD34 and CD45RA, as well as CD64 and CD115, expression by the generated cells during culture with SFGc for 2, 4 or 7 days (n=3). **d** Hypothetic model of myeloid cell ontogeny according to our results. The scheme shows that the acquisition of CD45RA represents a very premature event (that occurs prior to that of CD64) along the maturation trajectories of early progenitors of neutrophils, monocytes and DCs, occurring from the multilineage CD34⁺CD45RA⁻CD64⁻CD115⁻progenitor subset (including HSCs, CMPs, MEPs) into transitional multilineage progenitor pools. The latter pools include GMDPs (also named as neutrophil-, monocyte- and DC-committed progenitors, NMDPs), MDPs and CDPs, that subsequently mature into uni-lineage precursors, including NCP2s/NCP3s and cMoPs/pre-monocytes. The scheme also shows that the acquisition of CD45RA does not occur in those progenitors present within CD34⁺CD45RA⁻CD64⁻CD115⁻ subset that directly upregulates CD64 expression and generates NCP1s. As shown in the scheme, NCP1s directly differentiate into NCP4s, while NCP2s originate NCP4s via NCP3s.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Characterization of NCP transcriptomes as determined by RNA-seq. **a**, Developmental path of NCP1s, NCP2s, NCP3s and NCP4s, as well as HSCs/MPPs, CMPs, PMs, MYs, MMs, BCs, SNs and mature neutrophils (PMN), computationally determined from bulk RNA-seq datasets by using the optimal leaf ordering (OLO) algorithm. **b**, GO terms enriched by genes associated with the ten gene groups (g1-g10) identified by K-means analysis, as shown in Fig. 4d. The top five GO terms with Benjamini-Hochberg-corrected P values <0.05 (one-sided Fisher's exact test) are shown for every gene group. 'Gene ratio' indicate the fraction of DEGs present in the given GO term. **(c)**, Box plots showing the distribution of mRNA expression levels [as $\log_2(\text{FPKM} + 1)$] for genes associated to cell cycle, AG, SG, GG, SV and GM, as well as ROS biosynthetic process, phagocytosis, and chemotaxis. The box plot shows the median with the lower and upper quartiles representing a 25th to 75th percentile range and whiskers extending to $1.5 \times$ interquartile range (IQR). LOESS fitting of the data with relative confidence interval is represented by a blue line with a shadow area. **d**, PCA biplots based on the DEGs identified by LRT among bulk RNA-seq of NCP1s (orange), NCP2s (green), NCP3s (magenta) and NCP4s (turquoise). The graph lists the ten most relevant genes contributing to sample variations (indicated by vectors) for both PC1 and PC2, under both positive and negative directions. Vector lengths correlate with the weight of the given gene within the components.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Additional characterization of NCP and cMoP scRNA-seqs. **a**, Density plots of NCP1s, NCP2s, NCP3s, NCP4s and cMoPs overlaid on the UMAP of Fig. 7a. Density of cells in each plot is depicted according to the indications of the color bar. **b**, Hierarchical clustering dendrogram based on the DEGs identified among the neutrophilic and monocytic cell clusters shown in Fig. 7b. The vertical axis of the dendrogram represents the dissimilarity between clusters (that is, Euclidean distance). **c**, Violin plots showing the mRNA expression levels [as $\ln(\text{UMI})$] of selected genes across the four neutrophil clusters (c1-c4) chosen among the top defining genes indicated in Fig. 8e. Clusters are colored according to Fig. 7b. **d**, Expression patterns of cell-cycle, AG, SG, and GG genes projected on the UMAP plot restricted to neutrophil progenitor clusters (c1-c4). **e**, Gene Ontology analysis of DEGs for c2-c4. For every cluster (x-axis), the top ten Gene Ontology terms with Benjamini-Hochberg-corrected P values <0.05 (one-sided Fisher's exact test) are shown. For cluster 1 no enrichment of biological processes GO term was identified. **f**, Trajectory plots of c1-c4 cells as defined in Fig. 7b. In each plot is depicted the density of cells according to the color bar placed at the right bottom corner of the panel.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

For bulk RNA sequencing, total RNA was extracted by the RNeasy Mini Kit (Qiagen, Venlo, Limburg, Netherlands) after cell lysis. To completely remove any possible contaminating DNA, an on-column DNase digestion with the RNase-free DNase set (Qiagen) was performed during total RNA isolation. Quality control of the total RNA was performed using Agilent TapeStation 2200 (Agilent Technologies). RNA integrity (RIN) was routinely found to be optimal (RIN > 7.0). Libraries for transcriptome analysis were prepared using the Smart-seq2 protocol. Briefly, 2 ng total RNA were copied into a first strand cDNA by reverse transcription and template-switching, using oligo (dT) primers and a LNA-containing template-switching oligo (TSO). The resulting cDNA was pre-amplified, purified, and tagged with Tn5 transposase. cDNA fragments generated after tagmentation were gap-repaired, enriched by PCR and purified to create the final cDNA library. Libraries were sequenced on the Illumina NextSeq 500 in single-read mode (1x75 cycles).

For sc-RNA samples, sorted cells were labelled by using the BD Single-Cell Multiplexing Kit (BD Biosciences), strictly following the manufacturer's protocol (BD Biosciences). Such a procedure allowed us to combine 5 samples (NCP1s, NCP2s, NCP3s, NCP4s and cMoPs) into a single pool. Each sample was washed twice in FACS buffer and resuspended in cold BD Sample Buffer (BD Biosciences). Only samples with high viability (> 85 %, as evaluated by the BD Rhapsody scanner) were used for sequencing. Samples from the same donor were pooled in equal amounts to achieve approximately 15000 cells in 620 μ l, and loaded onto a BD Rhapsody cartridge for an incubation of 20 min at room T. Then, Cell Capture Beads (BD Biosciences) were added to the cartridge, incubated at room T for 3 min, and thereafter Cartridges washed. Cell were then lysed, and the released mRNA captured by Cell Capture Beads. mRNAs were then retrieved, to be washed prior to performing reverse transcription and treatment with Exonuclease I. cDNA Libraries were prepared by using the BD Rhapsody Whole Transcriptome Analysis (WTA), Amplification kit and BD Single-Cell Multiplexing kit (BD Biosciences). Size-distribution (Quality) of final libraries was assessed by Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape and quantified using a Qubit Fluorometer using the Qubit dsDNA HS Kit (ThermoFisher, #Q32854). Sequencing was performed in paired-end mode (2x75 cycles) on a NextSeq 500 System (Illumina). This procedure was utilized for three different BMs.

Data analysis

Sequencing Analysis:

For sc-RNA samples, after demultiplexing of bcl files by using Bcl2fastq2 V2.20 from Illumina and assessment of reads quality, paired-end scRNA-seq reads were then filtered for valid cell barcodes using the barcode whitelist provided by BD bioscience. Then, sequenced reads were

aligned to the hg38 human transcriptome and the expression of transcripts in each cell was quantified via the standard Rhapsody analysis pipeline (BD Biosciences) on Seven Bridges (<https://www.sevenbridges.com>), following manufacturer's recommendations. After gene expression matrix was created, integration and further analysis were performed using R package Seurat v3.2.2. Pseudo-time trajectories were determined using R package destiny v3.0.1

For bulk RNA sequencing, computational analysis of transcriptome datasets generated by Smart-seq2 has been performed by using the following bioinformatic pipeline. Briefly, after quality filtering, according to the Illumina pipeline, removal of contaminant adapters and base quality trimming were performed using Trim Galore! v0.6.4_dev (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) script. Gene counts were normalized among various samples using DESeq2 v1.26.0, and only genes coding for protein and long non-coding RNA (lncRNA) were retained for downstream analysis. Only genes expressed above 1 FPKM, in at least one sample, were considered as "expressed" genes and thus included for downstream analysis. Differentially expressed genes (DEGs) were identified using DESeq2, by using as selection parameter adjusted P-value lower than 0.01 and likelihood ratio test (LRT). Batch effects were removed using the limma v3.42.2 package's "removeBatchEffect" function before performing principal component analysis (PCA). PCA was performed on DEGs by using Bioconductor/R package pcaExplorer v.2.10.0.

Gene ontology (GO) analysis:

GO analysis was performed on DEGs from bulk RNA-seq and scRNA-seq for the cellular component, biological process and molecular function ontology domains by using the Bioconductor/R package clusterProfiler (version 3.14.3). Over-representation analysis was performed using the Benjamini-Hochberg (BH) procedure, with P value Cutoff = 0.01 and Q-value Cutoff = 0.05. Redundant GO terms were removed by using the "simplify" function in the clusterProfiler package, using the Wang similarity measure and a similarity cutoff of 0.5. The most significant terms were then plotted in R.

R scripts for data processing are available through <https://github.com/fbianchetto/Neutrophil-myeloblast>

Flow Cytometry:

Flow cytometry quantification was performed using FlowJo v10.7.1

Statistical analysis was performed using GraphPad Prism 7

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw and processed datasets have been submitted to the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) and are available under the accession number GSE164687.

Public Datasets Used:

GSE113182: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113182>

GSE153263: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153263>

seurat_COVID19_Neutrophils_cohort2_rhapsody_jonas_FG_2020-08-18.rds: <https://beta.fastgenomics.org/datasets/detail-dataset-ee4b1a0f339140ad82f861aea35076f1>

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

The sample size was chosen based on past experience in the lab and on the standard practices in the field with a minimum of n=3 and was in individual experiments increased to be sufficient to demonstrate statistically significant differences in comparisons between experimental groups by statistical tests described in the text. Further increase in sample size was mainly limited by the availability of human donor material (alloHSCT Bone Marrows). No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications (PMID:32719519; PMID: 32601182; PMID: 31160568)

Data exclusions

No data were excluded

Replication

Experiments were performed at least three times independently. All attempts at replication were successful within a reasonable range of variability due to inter-individuals diversity. The number of replicates for every experiments is indicated in the figure legends

Randomization

No randomisation was performed in experiments with BM cells because not relevant. Samples were allocated into experimental groups based on gating strategy.
In each animal experiment, mice were randomly assigned to each group.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

- | n/a | Included in the study |
|-------------------------------------|-----------------------------------------------------------------|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

- | n/a | Included in the study |
|-------------------------------------|----------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

CD14 (clone REA599)-Miltenyi Biotec; Cat# 130-110-524; RRID:AB_2655055; 1:50
 CD34 (clone AC136) -Miltenyi Biotec ; Cat# 130-113-182; RRID:AB_2726008; 1:50
 CD33 (clone WM53) -Biolegend; Cat# 303416; RRID:AB_2561690; 1:20
 CD10 (clone H110a) -Biolegend; Cat# 312218; RRID:AB_2561833; 1:20
 CD125 (clone A14) -BD; Cat# 743927; RRID:AB_2741855; 1:20
 CD15 (clone W6D3) -Biolegend; Cat# 323040; RRID:AB_2566520; 1:20
 CD45 (clone HI30) -Biolegend; Cat# 304036; RRID:AB_2561940; 1:20
 CD123 (clone 6H6) -Biolegend; Cat# 306022; RRID:AB_2562068; 1:20
 CD15 (clone W8D3) -BD; Cat# 562979; RRID:AB_2744292; 1:20
 CD45RA (clone HI100) -Biolegend; Cat# 304136; RRID:AB_2563653; 1:20
 CD14 (clone M5E2) -BD; Cat# 563698; RRID:AB_2744287; 1:20
 CD115 (clone 9-4D2-1E4) -Biolegend; Cat# 347312; RRID:AB_2566758; 1:20
 CD66b (clone G10F5) -Biolegend; Cat# 305104; RRID:AB_314496; 1:20
 CD14 (clone REA599) -Miltenyi Biotec; Cat# 130-110-518; RRID:AB_2655049; 1:50
 CD141 (clone AD5-14H12) -Miltenyi Biotec; Cat# 130-113-317; RRID:AB_2726094; 1:50
 CD1c (clone AD5 8E7) -Miltenyi Biotec; Cat# 130-113-301; RRID:AB_2726080; 1:50
 CD303 (clone AC144)-Miltenyi Biotec; Cat# 130-113-192; RRID:AB_2726017; 1:50
 CD64 (clone 10.1) -Biolegend; Cat# 305008; RRID:AB_314492; 1:20
 CLEC9A (clone 8F9) -Miltenyi Biotec; Cat# 130-097-368; RRID:AB_2657789; 1:50
 FCεR1 (clone CRA1) -Miltenyi Biotec; Cat# 130-100-162; RRID:AB_2660603; 1:50
 CD10 (clone HI10a) -Biolegend; Cat# 312204; RRID:AB_314915; 1:20
 CD114 (clone LMM741) -Biolegend; Cat# 346106; RRID:AB_2083867; 1:20
 CD123 (clone AC145) -Miltenyi Biotec; Cat# 130-113-326; RRID:AB_2726103; 1:50
 CD1c (clone AD5-8E7) -Miltenyi Biotec; Cat# 130-113-302; RRID:AB_2726081; 1:20
 CD45 (clone HI30) -BD; Cat# 562279; RRID:AB_11154577; 1:20
 CD45 (clone 5B1) -Miltenyi Biotec; Cat# 130-113-120; RRID:AB_2725948; 1:50
 CD16 (clone 3G8) -Biolegend; Cat# 302027; RRID:AB_893262; 1:20
 CD45RA (clone HI100) -Biolegend; Cat# 304122; RRID:AB_893357; 1:20
 CD11b (clone REA713) -Miltenyi Biotec; Cat# 130-110-555; RRID:AB_2654673; 1:50
 CD123 (clone AC145) -Miltenyi Biotec; Cat# 130-113-327; RRID:AB_2726104; 1:50
 CD14 (clone REA599) -Miltenyi Biotec; Cat# 130-110-521; RRID:AB_2655061; 1:50
 CD141 (clone AD5-14H12) -Miltenyi Biotec; Cat# 130-113-319; RRID:AB_2726096; 1:50
 CD19 (clone REA675) -Miltenyi Biotec; Cat# 130-113-647; RRID:AB_2726200; 1:50
 CD1c (clone AD5-8E7) -Miltenyi Biotec; Cat# 130-113-303; RRID:AB_2726082; 1:50
 CD3 (clone REA613) -Miltenyi Biotec; Cat# 130-113-140; RRID:AB_2725968; 1:50
 CD10 (clone HI10a) -Biolegend; Cat# 312210; RRID:AB_314921; 1:20
 CD14 (clone REA399) -Miltenyi Biotec; Cat# 130-110-520; RRID:AB_2655053; 1:50
 CD19 (clone LT19) -Miltenyi Biotec; Cat# 130-113-165; RRID:AB_2725993; 1:50
 CD303 (clone AC144) -Miltenyi Biotec; Cat# 130-113-190; RRID:AB_2726015; 1:50
 CD34 (clone AC136) -Miltenyi Biotec; Cat# 130-113-176; RRID:AB_2726003; 1:50
 CD66b (clone G10F5) -Biolegend; Cat# 305118; RRID:AB_2566607; 1:20
 CD66b (clone G10F5) -Biolegend; Cat# 305114; RRID:AB_2566038; 1:20
 CD11b (clone ICRF44) -Biolegend; Cat# 301342; RRID:AB_2563395; 1:20
 CD16 (clone 3G8) -Biolegend; Cat# 302018; RRID:AB_314218; 1:20
 CD38 (clone HIT2) -Biolegend; Cat# 303534; RRID:AB_2561605; 1:20
 CD45RA (clone HI100) -Biolegend; Cat# 304128; RRID:AB_10708880; 1:20
 HLA-DR (clone L243) -Biolegend; Cat# 307618; RRID:AB_493586; 1:20
 CD3 (clone UCHT1) -Biolegend; Cat#300436; RRID:AB_2562124; 1:20

CD19 (clone HIB19) -Biolegend; Cat#302236; RRID:AB_2563606; 1:20
 HLA-A2 (clone BB7.2)- Biolegend; Cat#343316; RRID:AB_2561573; 1:20
 mouse CD45.1 (A20)- Thermo Fisher Scientific; Cat#47-0453-82; RRID:AB_1582228; 1:20
 CD56 (clone REA196) -Miltenyi Biotec; Cat#130-114-550; RRID:AB_2726696; 1:50
 CD123 (clone REA918)-Miltenyi Biotec; Cat#130-115-272; RRID:AB_2726975; 1:50
 CD49d (clone MZ18-24A9)-Miltenyi Biotec; Cat#130-125-206; RRID:AB_2857751; 1:50
 CD71 (clone CV1G4) -Biolegend; Cat#334110; RRID:AB_2563117; 1:20
 CD117 (clone A3C6E2) -Miltenyi Biotec; Cat#130-113-541; RRID:AB_2733335; 1:50
 MPO (Rabbit polyclonal) -Agilent Technologies; Cat# A0398; RRID:AB_2335676; 1:5000
 ELANE (clone NP57)-Agilent Technologies; Cat# M0752; RRID:AB_630032; 1:200
 α -Defensins (clone H-2) - Santa Cruz Biotechnology; Cat# sc-390796; 1:70
 LTF (Rabbit polyclonal) -Sigma-Aldrich; Cat# hpa059976; RRID:AB_2684185; 1:500

Validation

All the antibodies are well recognized clones in publications, commercially available and validated by manufacturer (<https://www.biolegend.com/en-us/quality/quality-control>; <https://www.miltenyibiotec.com/US-en/products/macsc-antibodies/Antibody-production-development-and-quality-control.html>; <https://www.thermofisher.com/it/en/home/life-science/antibodies/invitrogen-antibody-validation.html>; <https://wwwbdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents>). See RRIDs for respective websites for each antibody. In addition these antibodies are routinely used in our laboratory with good reproducibility.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Mouse MS-5 cells from German Collection of Microorganisms and Cell Cultures GmbH, DSMZ Cat# ACC-441
 HEK293T cells from German Collection of Microorganisms and Cell Cultures GmbH, DSMZ Cat# ACC 635

Authentication

None of the cell lines used have been authenticated.

Mycoplasma contamination

All the cell lines were negative for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Female NOD.Cg-Prkdcscidll2rgtm1Sug/JicTac (NOG) mice, 4 weeks old, were used for this study. All the animals were housed in conditions of 22°C and 45-65% humidity with 12 hours light cycle.

Wild animals

This study did not involve wild animals.

Field-collected samples

This study did not involve field-collected samples.

Ethics oversight

NOD.Cg-Prkdcscidll2rgtm1Sug/JicTac (NOG) mice were purchased from Taconic Biosciences. All mice were maintained under specific pathogen-free conditions in the animal facility of the University of Verona. Animal experiments were performed according to national (protocol number 12722 approved by the Ministerial Decree Number 14/2012-B of January 18, 2012 and protocol number BR15/08 approved by the Ministerial Decree Number 925/2015-PR of August 28, 2015) and European laws and regulations. All animal experiments were approved by Verona University Ethical Committee (<https://www.univr.it/it/cirsal>) and conducted according to the guidelines of Federation of European Laboratory Animal Science Association (FELASA).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Buffy-coats are from anonymous, random gender, random blood group adult donors. BM samples were obtained from healthy donors selected for BM donation purposes according to the Italian Bone Marrow Donor Registry (IBMDR) criteria. BM samples were obtained also from patients undergoing allogeneic hematopoietic stem cell transplantation (alloHSCT) for hematological malignancies. To dissect the different steps of myeloid differentiation, BM samples were collected at day +21 after alloHSCT, when patients have achieved full donor chimerism according to variable number tandem repeats (VNTR) analysis.
 BM samples were obtained from healthy donors (n=35, 51% male). Age and gender for every donor are indicated in Supplementary Table 4. Mean age of the donors is 29

Recruitment

Blood from healthy donors was obtained from the Transfusion Center, University hospital, Verona (Italy). Bone marrows from healthy donors and alloHSCT patients were obtained from the Bone Marrow Transplant Unit, University hospital, Verona (Italy).

Ethics oversight

Human samples were obtained following informed written consent by HDs and alloHSCT patients. The study has been cleared by the Ethic Committee of the Azienda Ospedaliera Universitaria Integrata di Verona (Italy) (CMRI/55742).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

BM samples were collected under aseptic conditions in heparinized sterile syringe, processed using endotoxin-free polypropylene tubes (Greiner bio-one), and subjected to density gradient centrifugation onto Ficoll-Paque (GE Healthcare). For fluorescence-activated cell sorting of the myeloid progenitors, 50-100*10⁶ BM-LDCs were resuspended at 100*10⁶/ml and labeled for 45 min at 4°. Cells were then washed and resuspended at 30*10⁶/ml in staining buffer, to be ultimately filtered through a 0.35 µM nylon mesh. BM-LDCs, sorted progenitor subpopulations or culture-derived cells were counted, resuspended in 50 µl PBS buffer plus 2 % FBS and 2 mM EDTA (Sigma-Aldrich), and subsequently processed for flow cytometry. Cells were incubated for 10 min in the presence of 5 % human serum. Cells were then stained for 30 min on ice by fluorochrome-conjugated monoclonal antibodies.

Instrument

Flow cytometry data collection was performed with MACSQuant10 Analyzer (Miltenyi Biotec), LSRFortessa X-20 (BD) or MACSQuant16 Analyzer (Miltenyi Biotec) flow-cytometers. Fluorescence-activated cell sorting was performed by using the FACSria Fusion (BD) cell sorter.

Software

Flow cytometry data were analyzed using FlowJo v10.7.1.

Cell population abundance

Sorted cell populations displayed a > 95 % purity, as verified by flow cytometry analysis and dead cell exclusion. Usually, immature progenitors or peripheral blood populations were sorted in a range from 2*10³ to 1*10⁶ cells.

Gating strategy

Mature cell populations were identified by sequentially gating on singlets, then on PI- cells, and finally on CD45+ cell. Both lineage-positive cells and immature progenitors were identified by gating on either SS_{low}CD66b⁻, or SS_{Chi}CD66b⁺, fraction, as it follows: SS_{Chi}CD66b⁺CD16-CD45^{hi} eosinophils; ii) SS_{Chi}CD66b⁺CD16-/CD45⁺ neutrophils (from promyelocytes to segmented neutrophils); iii) SS_{low}CD141+CLEC9A+ cDC1 DCs; iv) SS_{low}FceRI+CD1c- basophils; v) SS_{low}FceRI_{dim}CD1c+ cDC2 DCs; vi) SS_{low}CD45RA+CD33+CD123+CD303+ pDCs; vii) SS_{low}CD3-CD19-CD56+ NK cells; viii) SS_{low}CD3-CD19-CD56-CD14+CD16-/CD14_{dim}-CD16+ total monocytes; ix) SS_{low}CD3+CD19- T cells; x) SS_{low}CD3-CD19+ B cells; and xi) SS_{low}CD45_{dim} cells including CD34+CD38- HSCs, CD34+ and CD34_{dim}- myeloid/lymphoid progenitors. For immature progenitors within BM-LDCs, analysis was performed on live, SS_{low}CD66b⁻ cells, CD3/CD19/CD1c/CD141/CD16/CD56(Lin)-negative cells. Within the latter cells, the CD45_{dim} region defines the SS_{low}Lin-CD45_{dim} progenitor pool, from which we could subsequently identify each progenitor population by using an antibody panel comprising key markers conventionally used to detect myeloid and lymphoid progenitors, namely CD34, CD38, CD10, CD123, CD45RA, CD64 and CD115. Fluorescence Minus One (FMO) or isotype control mAbs were used to determine background staining.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.