Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection

Xuemei Xie1,2,3,4,8, Qiang Shi5,8, Peng Wu1, Xiaoyu Zhang1,2,3,4, Hiroto Kambara2,3,4, Jiayu Su2,3,4,5, Hongbo Yu6,7, Shin-Young Park2,3,4, Rongxia Guo1, Qian Ren1, Sudong Zhang1, Yuanfu Xu1, Leslie E. Silberstein2,3,4, Tao Cheng4,1, Fengxia Ma1✉, Cheng Li5✉, and Hongbo R. Luo2,3,4✉

The full neutrophil heterogeneity and differentiation landscape remains incompletely characterized. Here, we profiled >25,000 differentiating and mature mouse neutrophils using single-cell RNA sequencing to provide a comprehensive transcriptional landscape of neutrophil maturation, function and fate decision in their steady state and during bacterial infection. Eight neutrophil populations were defined by distinct molecular signatures. The three mature peripheral blood neutrophil subsets arise from distinct maturing bone marrow neutrophil subsets. Driven by both known and uncharacterized transcription factors, neutrophils gradually acquire microbicidal capability as they traverse the transcriptional landscape, representing an evolved mechanism for fine-tuned regulation of an effective but balanced neutrophil response. Bacterial infection reprograms the genetic architecture of neutrophil populations, alters dynamic transitions between subpopulations and primes neutrophils for augmented functionality without affecting overall heterogeneity. In summary, these data establish a reference model and general framework for studying neutrophil-related disease mechanisms, biomarkers and therapeutic targets at single-cell resolution.

Results

Mouse neutrophil atlas in the steady state. Gr1+ cells were isolated from BM, PB and spleen by fluorescence-activated cell sorting (FACS) (Fig. 1a). To capture the whole spectrum of neutrophil maturation and identify potential neutrophil populations with lower Gr1 antigen (mainly Ly6G) expression, we also included Gr1low and a few Gr1− cells in each sample. In addition, due to the low abundance of hematopoietic stem progenitor cells (HSPCs), to gain insights into granulopoiesis in its entirety, we included a sample enriched for c-Kit+ BM HSPCs mixed with BM Gr1+ cells at a 2:3 ratio to artificially create a BM c-Kit/Gr1 population. Using the mixed cell population minimized the sample-to-sample variation (batch effect).

After rigorous quality control (Extended Data Fig. 1a–d), we obtained 19,582 high-quality cells with an average of 1,241 genes per cell profiled, resulting in a total of 18,269 mouse genes detected in all cells (Extended Data Fig. 1e and Supplementary Table 1). Unbiased, graph-based clustering identified seven major cell populations (Extended Data Fig. 1f–h and Supplementary Table 2). To dissect neutrophil heterogeneity, we examined the neutrophil-related populations (myeloid progenitors and neutrophils). Unsupervised clustering partitioned differentiating and mature neutrophils into eight clusters (G0–4 and G5a–c; Fig. 1b). G0–4 mainly originated from BM and represented neutrophils differentiating in BM, while G5a–c mainly originated from peripheral tissue samples (Fig. 1c). There was substantial differential gene expression between the groups.

N neutrophils migrate from the circulating blood to infected tissues in response to inflammatory stimuli, where they protect the host by phagocytosing, killing and digesting bacterial and fungal pathogens. However, neutrophil populations are not homogenous. Differentiation and maturation produce distinct neutrophil subpopulations that may be pre-programmed with different functions. Discrete microenvironments can modify neutrophil function and behavior. In addition, rapid neutrophil aging, their short lifespan and mechanically induced cellular responses as they enter and exit capillaries also contribute to neutrophil heterogeneity. Neutrophil classification has traditionally relied on morphology, surface marker expression or gradient separation, which while simple and robust do not capture the full neutrophil compartment repertoire. Some neutrophil subpopulations overlap, making it difficult to gain insights into granulopoiesis in its entirety, we included a sample enriched for c-Kit+ BM HSPCs mixed with BM Gr1+ cells at a 2:3 ratio to artificially create a BM c-Kit/Gr1 population. Using the mixed cell population minimized the sample-to-sample variation (batch effect).

After rigorous quality control (Extended Data Fig. 1a–d), we obtained 19,582 high-quality cells with an average of 1,241 genes per cell profiled, resulting in a total of 18,269 mouse genes detected in all cells (Extended Data Fig. 1e and Supplementary Table 1). Unbiased, graph-based clustering identified seven major cell populations (Extended Data Fig. 1f–h and Supplementary Table 2). To dissect neutrophil heterogeneity, we examined the neutrophil-related populations (myeloid progenitors and neutrophils). Unsupervised clustering partitioned differentiating and mature neutrophils into eight clusters (G0–4 and G5a–c; Fig. 1b). G0–4 mainly originated from BM and represented neutrophils differentiating in BM, while G5a–c mainly originated from peripheral tissue samples (Fig. 1c). There was substantial differential gene expression between the groups.

1The State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, China. 2Department of Pathology, Harvard Medical School, Boston, MA, USA. 3Department of Laboratory Medicine, The Stem Cell Program, Boston Children’s Hospital, Boston, MA, USA. 4Dana-Farber/Harvard Cancer Center, Boston, MA, USA. 5School of Life Sciences, Center for Bioinformatics, Center for Statistical Science, Peking University, Beijing, China. 6Department of Pathology and Laboratory Medicine, VA Boston Healthcare System, Boston, MA, USA. 7Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA. 8These authors contributed equally: Xuemei Xie, Qiang Shi. ✉ e-mail: mafengxia@ihcams.ac.cn; cheng_li@pku.edu.cn; Hongbo.Luo@childrens.harvard.edu
Neutrophil differentiation and maturation trajectories.

According to known gene signatures\(^{4-7}\), we concluded that the G0 population mainly consisted of granulocyte monocyte progenitor (GMP) cells expressing typical genes such as \(Cd17\) (Kit), \(Cd34\) and Sox4 and neutrophil primary granule genes (Fig. 1e). We also conducted hierarchical clustering (Fig. 1g). Consistent with uniform manifold approximation and projection (UMAP) clustering, neutrophils in PB (G5a, G5b and G5c) were closely associated but more remote from BM G1–4 cells. Using Monocle\(^4\) to place differentiating neutrophil populations along possible granulopoiesis trajectories in pseudo-time, neutrophil differentiation and maturation occurred on a tightly organized trajectory, starting from G1 cells in BM and ending with G5 cells in PB and spleen (Fig. 1h). G1 to G2 cells underwent active proliferation, with cell division stopping abruptly thereafter (Fig. 1i). A cluster of G3 cells followed G2 expansion and expressed secondary granule genes such as \(Ltf\) and \(Npg\) (Fig. 1e). Neutrophil differentiation in BM concluded with a more mature G4 population highly expressing \(Mmp8\) (encoding a key granule protein for neutrophil-mediated host defenses) and \(Cxcl2\), which is important for neutrophil mobilization\(^{7}\) (Fig. 1d,e).

We measured the maturation score of each differentiating neutrophil population based on the expression of genes related to neutrophil differentiation and maturation (Supplementary Table 4). G5 cells, which accounted for the majority of neutrophils in the peripheral tissues, were the most mature neutrophils, while G4 cells showed the highest maturation score among BM maturing (G0–G4) neutrophils (Fig. 1j).

A sorting mechanism for generating heterogeneous neutrophil granules. A well-accepted mechanism explaining neutrophil granule heterogeneity is targeting by biosynthesis\(^{20,21}\) (that is, granule proteins synthesized at the same time in developing neutrophils will end up in the same granule without granule type-specific sorting). We examined the expression of various granule genes in differentiating neutrophils (Fig. 2a–c). Lactoferrin (LTF)-positive granules are often defined as specific (secondary) granules, while LTF-negative but gelatinase-positive granules are known as gelatinase (tertiary) granules. As expected, \(Camp\) (the gene for a major antibacterial-specific granule protein CAP-18)\(^{22}\) was expressed in G2 and G3 cells when specific granules were formed. Surprisingly, this gene, whose product is not in the tertiary granules and secondary vesicles, was also expressed in G4 neutrophils containing tertiary granules and secretory vesicles. Such inconsistency was also observed for \(Lcn2\) and \(Mmp8\). Granule proteins belong to a subset of proteins for which RNA expression dropped during differentiation while protein expression remained similar. This may be indicative of protein storage or sequestration in various granules\(^{33}\). Alternative targeting by timing of biosynthesis may not explain all granule heterogeneity, and some granule proteins may in fact be tagged to direct them to particular granules.

scRNA-seq-morphology correlates. Neutrophil maturation typically follows five main morphological stages: myeloblasts; promyelocytes; myelocytes; metamyelocytes; and band cells and segmented neutrophils\(^{8}\). We compared scRNA-seq-defined neutrophil populations with the classic morphology-defined neutrophils, which were isolated by FACs based on c-Kit and Ly6G expression\(^{43}\) and bulk sequenced (Fig. 2d). Most of the molecular signatures identified by scRNA-seq were also detected in bulk RNA-seq data (Fig. 2e and Supplementary Fig. 1a). To further dissect morphological heterogeneity, a regression-based deconvolution approach was applied on bulk expression profiles. Myeloblasts were a mixture of G0 and G1, promyelocytes were G1, myelocytes were G1/G2, metamyelocytes were G2, and band cells and segmented neutrophils were G3/G4 (Fig. 2f and Supplementary Fig. 1a–c).

Correlation with previously defined neutrophil subpopulations. Previous studies revealed a variety of distinct neutrophil subpopulations arising during differentiation and maturation. Olsson et al.\(^3\) performed scRNA-seq on stem/multipotent progenitor cells, common myeloid progenitor (CMP) cells, GMP cells and NK CD34\(^+\) cells (lin-c-Kit/CD34\(^−\)) that included granulocytic precursors. We calculated the fraction of each scRNA-seq-defined neutrophil subpopulation in these four samples and revealed that stem/multipotent progenitor cells and CMP cells exhibited a gene expression profile matching that of G0 cells. Importantly, G1 neutrophils started to appear in the GMP population. The NK CD34\(^+\) myeloid progenitor precursor population contained a substantial number (~20%) of G2 neutrophils (Extended Data Fig. 2a). We also correlated our scRNA-seq-defined neutrophil populations with the neutrophil subtypes reported by Evrard et al.\(^2\). Regression-based deconvolution analysis showed that the therein identified pre-neutrophils (preNeu), CXCR2\(^{−}\) immature neutrophils (immNeu) and CXCR2\(^{high}\) mature neutrophils (mNeu) were correlated with the G2, G3 and G4 neutrophils, respectively (Extended Data Fig. 2b). Next, we compared the C1 and C2 cells defined by Zhu et al.\(^{35}\) with the neutrophil

(caption: Fig. 1 scRNA-seq analysis of steady-state BM, PB and spleen neutrophils. a, Schematic of the study design. b, UMAP of 12,285 neutrophils from BM, PB and spleen, colored by cell type (top left), sample origin (top right) and inferred cluster identity (bottom). The contaminating population (cont) mainly consisted of low-quality cells and was therefore discarded from further analysis. Similarly, unless otherwise stated, GM cells (low-quality G3-like cells with a low UMI count per cell and per gene and a high percentage of mitochondrial UMI counts) were excluded from all other downstream analyses. c, Proportions of the nine neutrophil clusters in four samples. d, Heatmap showing the row-scaled expression of the 20 highest DEGs (Bonferroni-corrected \(P\) values < 0.05; Student’s \(t\)-test) per cluster for all neutrophils except cells from the GM population. e, Dot plot showing the scaled expression of selected signature genes for each cluster, colored by the average expression of each gene in each cluster scaled across all clusters. Dot size represents the percentage of cells in each cluster with more than one read of the corresponding gene. f, Gene Ontology analysis of DEGs for each cluster. Selected Gene Ontology terms with Benjamini–Hochberg-corrected \(P\) values < 0.05 (one-sided Fisher’s exact test) are shown and colored by gene ratio. g, Unsupervised hierarchical clustering of the eight clusters based on the average gene expression of cells in each cluster. For each neutrophil subpopulation, the proportion of cells from indicated BM, PB or spleen samples is shown. h, Monocle trajectories of neutrophils colored by sample origin (left) and cluster identity (right). Each dot represents a single cell. Cell orders are inferred from the expression of the most variable genes across all cells. Trajectory directions were determined by biological prior. i, Heatmap showing the row-scaled expression of cell cycle-related genes for GO-G4 neutrophils. j, Violin plot of maturation scores for each cluster. For the box plot within each violin plot, middle lines indicate median values, boxes range from the 25th to 75th percentiles, and upper/lower whiskers extend from the hinge to the largest/smallest value no further than 1.5 times the interquartile range (IQR) from the hinge. Colored areas indicate density distribution of data. MP, myeloid progenitor; rRNA, ribosomal RNA; SP, spleen; WT, wild type.)
subpopulations identified in our study. The result suggested C1 to be a mixture of G0/G1, and C2 to be G2/G3 (Extended Data Fig. 2c,d). Notably, in an earlier study, Kim et al.24 also defined a population of proliferative late-stage neutrophil precursors (NeuP) in BM characterized by a lin-c-Kit+CD11b-Ly6G−Ly6B−CD115−Gfi1+ signature that should be located in the G1/G2 population. It was later found that this NeuP population was highly heterogeneous and contained other myeloid progenitors (G0 cells). Furthermore, we examined the correlation of scRNA-seq-defined neutrophil populations with the stage I and stage II neutrophils defined by Giladi et al.13 and revealed that stage I neutrophils were mainly G2–G4 cells while stage II neutrophils were a mixture of G4 and G5 cells (Extended Data Fig. 2e,f). Finally, a population of committed neutrophil progenitors (proNeu) was proposed. Recently, Muench et al.29 and Kwok et al.30 independently characterized this population. We compared the neutrophil subpopulations identified in our study with the proNeu population and found that proNeu cells perfectly correlated with G1 neutrophils (Extended Data Fig. 2g,h).
PB and spleen contain three distinct neutrophil subpopulations. Three major neutrophil subpopulations were identified in PB and spleen with 172 DEGs (Supplementary Table 3 and Fig. 2g). Similar to G4 BM cells, G5a cells highly expressed Mmp8 and Slc10a8 (Fig. 2g), as well as genes related to neutrophil migration and inflammatory responses (Fig. 2h). Interestingly, a group of G5b neutrophils expressed a set of interferon-stimulated genes (ISGs), such as Ifit3 and Isg15 (Extended Data Fig. 3a). Trajectory analysis showed that G5a and G5b neutrophils gradually developed into G5c neutrophils (Extended Data Fig. 3b), with the latter showing the highest aging score (Fig. 2i,j). By applying a two-component Gaussian mixture model, we further identified 15% of G5c neutrophils as aged—a significantly higher proportion than in G5b or G5a populations (Fig. 2). Although G5c cells appeared to more aged, the mitochondrial unique molecular identifier (UMI) percentage was not elevated in G5c cells, indicating continued viability in PB and spleen (Extended Data Fig. 3c). Noticeably, the initial cell clustering was performed from a mixture of c-Kit+ BM HSPCs and Gr1+ neutrophils from BM, spleen and PB. Clustering of circulating (PB) neutrophils only gave rise to the same three G5 clusters (Supplementary Fig. 2a,b).

Experimental isolation of neutrophil subsets. scRNA-seq defined BM differentiating neutrophil populations (G0–G4) tightly correlated with the morphology-defined neutrophil subpopulations (Fig. 2). On the basis of this correlation, we were able to identify and sort G1 and G2 neutrophils in BM (Fig. 3a). To separate G3 and G4, we added another surface marker, Cxcr2, which was more highly expressed on G4 cells (Fig. 3b). The proportions of G1–G4 neutrophils in BM measured by FACS analysis (Fig. 3c) fit well with those calculated based on the scRNA-seq data (Fig. 1c).

We isolated PB G5b cells based on their expression of IFIT1 (Fig. 3d). G5c cells are relatively aged neutrophils with high surface expression of Cxcr4 and thus were isolated as IFIT1−Cxcr4a neutrophils by FACS sorting, while G5a cells were identified as IFIT1−Cxcr4a neutrophils (Fig. 3d). The identity of the sorted G5a, G5b and G5c neutrophils was confirmed by measuring the expression of distinct signature genes (Fig. 3e,f). Using this approach, we calculated the percentage of each G5 subpopulation in PB and spleen. The result was consistent with that derived from the scRNA-seq data (Fig. 3g).

The pathogen clearance machinery is continuously and gradually built during neutrophil differentiation, maturation and aging. Phagocytosis, chemotaxis and neutrophil activation scores increased drastically during the early stages of granulopoiesis, peaked at G3 and remained relatively stable thereafter (Extended Data Fig. 4a–c). Similarly, the NADPH oxidase score increased during G0 to G1 and G2 transition, peaked at G3 and then decreased by 20% in mature neutrophils (Extended Data Fig. 4d). However, the dynamics of the oxidase complex subunits varied through neutrophil differentiation (Extended Data Fig. 4e). Sequential subunit expression ensures maximum stimulation-triggered NADPH oxidase activation at the later stages of neutrophil maturation and minimum activation in immature neutrophils in BM. Notably, genes related to mitochondria-mediated reactive oxygen species (ROS) production were significantly downregulated during neutrophil maturation, further supporting that neutrophil ROS production is mainly mediated by phagocytic NADPH oxidase (Extended Data Fig. 4f).

Mature neutrophils derive energy mainly from glycolysis81. However, metabolism-related genes (Extended Data Fig. 4g), including those related to glycolysis (Extended Data Fig. 4h), were downregulated in mature neutrophils. Similarly, genes related to glucose transportation were also not upregulated in mature neutrophils (Extended Data Fig. 4i). These data suggest that glycolysis-dominant metabolism in neutrophils is likely to be driven by post-transcriptional or/and post-translational mechanisms rather than transcriptional upregulation of related genes.

Organ-specific transcriptome features. Compared with early-stage maturing BM neutrophils, most neutrophils in PB and spleen were mature G5 neutrophils with similar gene expression patterns (Extended Data Fig. 5a). Kyoto Encyclopedia of Genes and Genomes enrichment analysis revealed that PB neutrophils were enriched for genes related to malaria and African trypanosomiasis, while spleen neutrophils were more enriched for genes related to leishmaniasis and Yersinia infection, suggesting that PB and spleen neutrophils may play distinct roles in combating different infections (Extended Data Fig. 5a,b). Spleen neutrophils also expressed T cell differentiation, interleukin-17 signaling, tumor necrosis factor signaling, and antigen processing and presentation-related genes, suggesting a role for splenic neutrophils in adaptive immunity. Under homeostatic conditions, more than 98% of neutrophils (about 200 million) stored in BM32,33. Although the percentage of G5 neutrophils in BM was significantly lower (Fig. 1c and Extended Data Fig. 5c), their absolute number was comparable or even higher than it was in PB or spleen (Extended Data Fig. 5d). Compared with their counterparts in PB and spleen, BM G5a and G5c neutrophils preferentially expressed genes related to cell migration, chemotaxis, adhesion and antimicrobial peptides (Extended Data Fig. 5e–h). PB and spleen G5 neutrophils displayed higher maturation and apoptosis scores than BM G5 neutrophils (Extended Data Fig. 5i,j). These results
suggest that although the identity of each neutrophil subpopulation is determined by programed expression of distinct signature genes, the organ-specific microenvironment also plays a significant role in driving transcription in each subpopulation.

Unexpected complexity in neutrophil mobilization. Next, we traced cell fate and reconstructed cell lineage direction using the recently developed RNA velocity approach (Fig. 4a and Supplementary Fig. 3). Consistent with Monocle (Fig. 1h), BM maturation (from G2 to G4) followed a single main branch without significant division, with G3 bearing long vectors and indicating a strong tendency to progress to G4 (Fig. 4a and Supplementary Fig. 3). G5c cells were firmly at the end of neutrophil maturation and differentiation, showing the highest apoptosis scores (Fig. 4b) and proportion of apoptotic cells (~20%; Fig. 4c) among the most mature G5 population. There was also significant apoptosis in G5a and G5b cells (Fig. 4c), suggesting that death programs can be independent of maturation. Next, we performed velocity analysis using only G3–G5 subsets and the cells in different compartments were presented separately to accurately assess their relationship with other cell types.
**Fig. 3 | Analysis of neutrophil subpopulations by flow cytometry.** a–c, Separation of BM maturating G1–G4 neutrophils. a, FACS and staining strategy. The experiments were conducted as described in Fig. 2d. Data are representative of two independent experiments. b, Expression pattern of the gene Cxcr2 projected on the UMAP plot. Only BM cells are shown. c, Percentage and absolute number of different neutrophil populations in BM. Data represent means ± s.d. (n = 4 mice) of two independent experiments. d–f, Separation of G5a, G5b and G5c neutrophils by flow cytometry. d, FACS and staining strategy for PB G5a (IFIT1−CXCR4lo), G5b (IFIT1+) and G5c (IFIT1−CXCR4hi) neutrophils. Fluorescence minus one (FMO) controls (top) were used to control for spillover-related contribution to background in each channel. e, Left: UMAP plot of G5a, G5b and G5c neutrophils colored by cluster identity. Right: expression patterns of the marker genes of G5a (Lyz2 and S100a8), G5b (Ifit1 and Isg15) and G5c (Cxcr4 and Gm2a) projected on the UMAP plot. f, Relative mRNA expression of the six marker genes in sorted PB G5a, G5b and G5c neutrophils as measured by real-time qPCR. Data represent means ± s.d. (n = 15 mice) of two independent experiments. g, Comparison of FACS and scRNA-seq-based analyses of G5 subpopulations in PB and spleen. The scRNA-seq-based percentages of each G5 subpopulation were derived from Fig. 1c. FACS-based measurement was conducted as described in d. Data represent means ± s.d. (n = 3–6 mice) of two independent experiments.
Fig. 4 | The trajectory and transcriptional control of neutrophil maturation. a–d. The origin and inter-relationship of neutrophil subpopulations. a. Velocity analysis revealing the origin and inter-relationship of neutrophil subpopulations. Velocity fields were projected onto the UMAP plot. b. Violin plot of apoptosis scores (GO: 0043065) for G5 clusters. c. Proportion of apoptotic cells in each cluster, identified by a two-component Gaussian mixture model. d. As in a, but only showing G3 neutrophils originating from PB. e–j. The formation of neutrophil subpopulations is driven by both known transcription factors and a large set of uncharacterized ones. e. Heatmap showing the row-scaled gene expression of transcription factors known to be involved in granulopoiesis and neutrophil function. f. UMAP of the regulon activity matrix of neutrophils and 7,209 non-neutrophils under normal conditions. g. Percentage of neutrophils and other non-neutrophil populations. Only regulons with at least one absolute t-value >100 are visualized. Previously uncharacterized neutrophil-specific transcription factors are marked in red, with binding motifs of these transcription factors shown on the right. h. Activities of the four newly identified neutrophil-specific regulons. j. As in h, but with t-values representing activity change between the current developmental stage and the previous one. Only regulons with at least one absolute t-value >40 are visualized. Regulons are hierarchically clustered based on activation pattern (red and orange: early activated; yellow: G3 inactivated; green: late activated; blue: globally activated). Mono, monocyte; DC, dendritic cell.
populations and/or cells in other compartments (Supplementary Fig. 4). Interestingly, the trajectory of a significant number of G3 neutrophils was towards the peripheral G5a population, suggesting mobilization of G3 cells to PB or tissue without first undergoing full G4 maturation (Supplementary Fig. 4). The BM G4 population split into (1) the peripheral G5a population and (2) the ISG-related G5b population without entering the G5a stage. Thus, although G5a and G5b were most similar (Fig. 1g), they are two separate and independent PB neutrophil populations, with G5a derived from BM G3 and G4 cells and G5b derived solely from G4 cells. G5a to G5b conversion was rarely detected in PB (Fig. 4a and Supplementary Fig. 4).

Both the G3 and G4 populations are differentiating neutrophils that mainly exist in BM (Extended Data Fig. 5d). A small number of immature neutrophils also circulate, which are thought to be derived from accidental release of cells closest to maturation35–37. We detected significantly more G3 cells in the periphery of healthy mice (5% of PB and 6% of spleen neutrophils; Fig. 1c) compared with G4 cells. Thus, during BM granulopoiesis, more G3 neutrophils escape from the BM niche, migrate into PB and travel to other organs. PB and BM G3 cells consistently overlapped on velocity analysis, with some falling into the PB G5a cluster (Fig. 4d and Supplementary Fig. 4). Furthermore, PB G3 cells directly differentiated into G5a without going through G4, consistent with the low number of G4 cells in PB and spleen (Fig. 4d and Supplementary Fig. 4).

Both known and uncharacterized transcription factors drive neutrophil subpopulations. Next, we sought to characterize transcription factor dynamics across neutrophil differentiation and maturation35–37 (Fig. 4e). Genes related to stem cell maintenance and early lineage commitment, such as Gata2, Irf8 and Runx1, were highly expressed in the G0 population. Genes highly expressed in G1 included Gfi1 and Cepha, which play essential roles in neutrophil development, strongly suggesting that specific neutrophil lineage commitment occurs during G1. To assess specific global gene regulatory networks associated with neutrophil maturity, we applied single-cell regulatory network inference and clustering (SCENIC) analysis38 (Fig. 4f). There was high consistency between Seurat clusters and SCENIC clusters (Fig. 4g). To further dissect the regulatory differences between neutrophils and other cell types, we compared regulon activities from each neutrophil group versus all non-neutrophil populations using a generalized linear model (GLM)39. This identified 19 neutrophil-specific networks, including previously reported transcription factors such as Cebpe, Spi1 and Klf5 (Fig. 4h and Supplementary Table 5). Importantly, this analysis also identified four new regulons, Nfil3, Max, Mlx and Xbp1, which are closely related to the expression of neutrophil-specific genes (Fig. 4i). Next, we examined the regulatory events responsible for transitioning between consecutive neutrophil differentiation stages (Fig. 4j). Coarse-grained clustering revealed at least five regulon groups with distinct activation patterns, including two early activated, one late activated, one globally activated and one specifically inactivated after G2. While many transcription factor networks such as Cebpe, Ets1, Klf5, Rad21 and E2f2 contributed to neutrophil commitment, changes in Xbp1 and Mlx networks were specifically associated with G0/G1 transition. Additionally, the dramatic loss of regulatory networks such as E2f1, Nefl and Rb1 indicated a potential functional change between G2 and G3.

**Bacterial infection primes neutrophils for augmented functionality without affecting overall heterogeneity.** Next, we investigated how bacterial infection affected neutrophil subpopulations, including in the liver and peritoneal cavity (Extended Data Fig. 6a–f). At the same sequencing depth, the gene number and total UMIs both increased in neutrophils isolated from BM, spleen and PB of *E. coli*-challenged mice compared with control mice, indicating elevated transcriptional activity during bacterial infection (Extended Data Fig. 6g,h). Leveraging the fact that cells in each subpopulation from unchallenged mice retain core signature genes beyond perturbations, we were able to identify every population in challenged mice using a well-accepted data integration method40 (Fig. 5a,b) validated independently through unsupervised dimension reduction at transcriptome levels. The expression of signature molecules (Fig. 5c), NADPH oxidase components (Extended Data Fig. 6i) and granular proteins (Extended Data Fig. 6j) remained remarkably consistent after *E. coli* challenge. Thus, the identity of each neutrophil population was maintained during acute bacterial infection and the signature genes still successfully determined neutrophil identity under inflammatory conditions (Fig. 5c). However, infection up- and downregulated numerous genes in each neutrophil subpopulation (Fig. 5d). In differential gene expression analysis (Extended Data Fig. 7a and Supplementary Table 6), DEGs in G0 and G1 cells were also preferentially involved in regulating immune effector processes and ROS metabolism, respectively, suggesting that immune adaptation to bacterial infection could occur as early as within early progenitor cells (Fig. 5e and Extended Data Fig. 7b). In relatively mature G4 and G5 neutrophils, bacterial infection triggered significant upregulation of cytokine production and secretion genes (Fig. 5e and Extended Data Fig. 7b). Finally, in bacteria-challenged hosts, neutrophil functions related to bactericidal activities including synthesis of granular proteins (Fig. 5f), NADPH oxidase complex (Fig. 5g), phagocytosis and chemotaxis (Fig. 5h) were all upregulated. Thus, during bacterial infection, core neutrophil subpopulations are maintained but genes related to pathogen clearance are upregulated at each stage of neutrophil maturation to maximize host defenses.

We also examined whether bacterial infection had a universal impact on transcriptional regulatory networks across neutrophil populations. Overall, there was a coherent drift in gene regulatory network activities in each subpopulation after bacterial challenge (Extended Data Fig. 8a), perhaps driven by upregulation of defense response-associated transcription factor networks such as Irf7 and downregulation of metabolic transcription factor networks such...

---

**Fig. 5 | Bacterial infection primes neutrophils for augmented functionality without affecting their overall heterogeneity.** a. Comparison of control and *E. coli*-challenged neutrophils originating from BM, PB and spleen. All neutrophils under both control and *E. coli*-challenged conditions (13,687 cells) were projected together by UMAP but are displayed separately by experimental condition. b. Comparison of neutrophil composition between control and *E. coli* challenge in BM, PB and spleen before and after *E. coli* challenge. Each cell in the *E. coli*-challenged dataset was annotated based on the transcriptomic similarity between this cell and cells in the reference dataset. Neutrophils sharing similar transcriptomic profiles were placed into the same cluster. c. Dot plot showing the scaled expression of signature genes for each cluster before and after *E. coli* challenge, colored by the average expression of each gene in each cluster scaled across all clusters. Dot size represents the percentage of cells in each cluster with more than one read of the corresponding gene. d. Heatmap showing the log([fold-change]) in gene expression of the representative cluster-based DEGs between control and *E. coli*-challenged neutrophils. The asterisks mean log([fold-change]) > 1 in corresponding cells. e. Gene Ontology analysis of cluster-based DEGs between control and *E. coli*-challenged neutrophils. Selected Gene Ontology terms with Benjamini-Hochberg-corrected *P* values < 0.05 (one-sided Fisher’s exact test) are shown. f–h, Comparisons of functional scores (granule scores (f), NADPH oxidase complex scores (g) and chemotaxis and phagocytosis scores (h)) between control and *E. coli*-challenged neutrophils for each cluster. Significance was determined by Student’s *t*-test. NS, not significant (*P > 0.05); *P < 0.05; **P < 0.01; ***P < 0.0001. MHC, major histocompatibility complex.
as Foxp1 and Ctcf (Extended Data Fig. 8b). Interestingly, we also identified transcription factor networks (for example, Fos and Atf4) showing different responses in immature (upregulated) and mature (downregulated) populations (Extended Data Fig. 8b). These networks were gradually activated from G1 to G5 under normal conditions (Extended Data Fig. 8c and Supplementary Table 5), and their...
Fig. 6 | The ISG-expressing neutrophil population is present in both humans and mice and expands during bacterial infection. a. Quantitative image analysis of the spatial distribution of G5b in whole spleen sections. Left: LSC image analysis of the whole spleen. The first column shows low-magnification images of axial spleen cryosections immunostained for DAPI (blue), S100a8 (green) and IFIT1 (red). The second column shows representative images of spleens from the selected region shown by a red box in the first column. The third column shows localization diagrams of S100a8 and spleen shown in the first column. Cells gated positive based on the fluorescence intensity in the S100a8 channel. Each dot represents a single cell. The fourth column shows localization diagrams of S100a8*IFIT1* (G5b) cells gated from the cells shown in the third column. Each dot represents a single cell. Right: LSC images (top) and confocal images (bottom) of representative G5b neutrophils. b. Quantification and relative frequency of spleen Ly6G+ cells in the whole spleen (left; measured by FACS) and spleen G5b cells (right; S100a8*IFIT1*) measured by LSC at different time points after E. coli infection. Results are means ± s.d. of three independent experiments (n = 3 mice for each time point; 3–5 slides were scanned and quantified for each mouse). c. Heatmap showing the log_{10}(fold-change) in expression of 49 ISGs before and after E. coli challenge for each cluster. Genes are marked with an asterisk if their expression changed significantly, as identified by Student’s t-test (Bonferroni-corrected Pvalue < 0.05). d. UMAP of neutrophils from human PB, colored by cluster identity. The fraction of cells in each cluster is displayed on the right. e. Heatmap showing the row-scaled expression of the five highest DEGs (Bonferroni-corrected Pvalues < 0.05; Student’s t-test) per cluster for all human G5 (hG5) neutrophils. f. Heatmap showing the expression of 37 ISGs for the three human neutrophil clusters. Genes marked in red are conserved across mice and humans.Ctl, control.
**Fig. 7** Bacterial infection accelerates G1 cell division and post-mitotic maturation without altering overall neutrophil differentiation programs.

**a.** Monocle trajectories of *E. coli*-challenged neutrophils colored by sample origin (left) and cluster identity (right). Each dot represents a single cell. Cell orders are inferred from the expression of the most variable genes across all cells. The trajectory direction was determined by biological prior. Correlation matrices of t-values for regulon activity change during each group transition event under normal conditions (top) or after *E. coli* challenge (bottom). For each group transition event after challenge, the direct comparison with all normal transition events is shown (bottom).

**b.** Comparisons of proliferation score (**c**) and S-phase and G2/M-phase score (**d**) between control and *E. coli*-challenged neutrophils for each of the seven clusters.

**e.** In vivo EdU incorporation assay. **e.** Top: schematic. Bottom: gating strategy of the three neutrophil subpopulations: immature (Ly6G^hi^CXCR4^hi^; black), intermediate (Ly6G^hi^CXCR4^hi^; blue) and mature neutrophils (Ly6G^hi^CXCR4^lo^; red).

**f.** In vivo EdU proliferation assay of neutrophil subsets in BM, PB and spleen at sequential time points with or without *E. coli* challenge. Data are represented as percentages of EdU^+^ cells in the corresponding gated subpopulation. Results are the mean ± s.d. of three independent experiments (n = 3–5 mice for each time point). HI *E. coli*, heat-inactivated *E. coli*; LV, liver; PC, peritoneal cavity.
**Fig. 8** | **Bacterial infection reprograms the structure of the neutrophil population and the dynamic transition between each subpopulation.**

**a.** Comparison of the organ distributions of each neutrophil subpopulation before and after *E. coli* challenge, measured by cell number. **b.** Comparison of neutrophil dynamics (with the velocity field projected on the UMAP plot) before and after *E. coli* challenge. **c.** Neutrophil proportion (left) and cell number (right) in the peritoneal cavity, measured at different time points after *E. coli* challenge. Results are the means ± s.d. of three independent experiments (*n* = 3 mice for each time point). **d.** Left: UMAP of *E. coli*-challenged neutrophils from BM, PB, spleen and peritoneal cavity, colored by cluster identity. Right: peritoneal cavity cells are highlighted in the UMAP plot. The proportions of each neutrophil cluster in the PC are shown below. **e.** Comparison of apoptosis scores and necroptosis scores between control and *E. coli*-challenged neutrophils for G3–G5 clusters. Statistical significance was determined by Student’s *t*-test. **** *P* ≤ 0.0001. **f.** Dynamic transition between neutrophil subpopulations under steady-state and bacterial infection conditions. We cataloged differentiating and mature mouse neutrophils in an unbiased manner using scRNA-seq. Based on the correlation analyses (Supplementary Fig. 3), the G0, G1, G2, G3 and G4 clusters characterized here were aligned to BM GMP, proNeu, preNeu, immNeu and mNeu cells, respectively. The names proNeu, preNeu, immNeu and mNeu were adopted from ref. 6. G5a, G5b and G5c (the major neutrophil populations in PB) represented the most mature neutrophils with typical polymorphonuclear morphology and were named PMNa, PMNb and PMNc, respectively. Under homeostatic conditions, the PMNa cells in PB can arise from both mNeu and immNeu, while PMNb cells mainly arise from BM mNeu cells. The transformation from immNeu to PMNa cells was suppressed during infection and immNeu cells cells in infected hosts predominantly differentiated to mNeu cells.
The ISG-related G5b neutrophil population exists in both humans and mice and expands during infection. The percentage of G5b cells increased significantly in E. coli-challenged hosts (Fig. 5b). We investigated the distribution of G5b neutrophils in spleen using a laser scanning cytometer, co-staining spleen tissue sections with an anti-S100a8 antibody and an anti-IFIT1 antibody to identify G5b neutrophils (Fig. 6a). Under normal conditions, S100a8+ neutrophils were uniformly distributed in the red pulp, while G5b (S100a8+IFIT1+) cells were preferentially subcapsular. After E. coli challenge, the overall number of neutrophils in spleen increased significantly, as did the percentage and number of G5b (Fig. 6b), which were still preferentially subcapsular, their specialized location further demonstrating the uniqueness of this subpopulation. Although multiple ISGs (for example, Ifitm1) were upregulated in basically all neutrophil subpopulations after bacterial stimulation (Fig. 6c), many ISGs such as Igf15 and Oas2, which are specifically expressed in G5b, were not upregulated, suggesting that ISG-related G5b expansion was not due to bacteria-induced ISG expression.

Next, we examined whether this G5b population was also present in human blood by scRNA-seq of sorted human PB neutrophils13,19,26, we separated relatively immature (Ly6G lowCXCR4hi), these cells post-mitotically in BM and PB (Fig. 7e). Based on the transcription regulatory networks in control and challenged samples (Fig. 7b), while both G1 and G2 cells are proliferative, the proliferation score increased only in the G1 population (Fig. 7c), as did genes related to G2/M-phase progression5, while genes related to S-phase progression were paradoxically reduced in G2 cells during acute infection (Fig. 7d).

During bacterial infection, the G3 and G4 pool must increase and post-mitotic maturation without altering overall neutrophil differentiation. Neutrophil populations significantly expand during bacterial infection14,15. However, the neutrophil differentiation and maturation trajectory was largely maintained in E. coli-challenged mice (Fig. 7a). The overall stability of the neutrophil differentiation program after bacterial infection was also demonstrated by correlation of SCENIC transcription regulatory networks in control and challenged samples (Fig. 7b). While both G1 and G2 cells are proliferative, the proliferation score increased only in the G1 population (Fig. 7c), as did genes related to G2/M-phase progression5, while genes related to S-phase progression were paradoxically reduced in G2 cells during acute infection (Fig. 7d).

During bacterial infection, the G3 and G4 pool must increase to produce more mature neutrophils. Nevertheless, the proportions of G3 and G4 cells were not increased in the BM of E. coli-challenged hosts (Fig. 5b), indicating that post-mitotic maturation may be accelerated in BM. To test this hypothesis, we labeled -challenged hosts (Fig. 5b), indicating that post-mitotic maturation may be accelerated in BM. To test this hypothesis, we labeled E. coli-challenged mice, including those in BM, PB, spleen and peritoneal cavity, revealed irregular multiple directions and short vectors in peritoneal cavity G5 populations, indicating a rather inactive transition among these populations (Supplementary Fig. 7). Interestingly, although G5c cells accounted for only about 25% of PB neutrophils before bacterial challenge and <10% of PB neutrophils after bacterial challenge, >45% of peritoneal cavity neutrophils in challenged mice were G5c cells (Fig. 8d), suggesting that these cells may possess higher transendothelial migration capability than G5a or G5b cells. G5c neutrophils showed the highest aging score compared with other G5 cells (Fig. 8i,j) with no difference detected among BM, PB and spleen G5c cells in unchallenged mice (Supplementary Fig. 8a). During bacterial infection, the aging score and the proportion of aged cells in the peritoneal cavity G5c population became significantly higher compared with in the PB and spleen G5c populations (Supplementary Fig. 8b). These results are in agreement with a previous study showing that in the acute inflammatory response during endotoxicemia aged neutrophils stop returning to BM and instead rapidly migrate to sites of inflammation46,47. Infection delays neutrophil death46. Paradoxically, genes related to apoptosis or necroptosis were significantly upregulated in every neutrophil subpopulation in E. coli-challenged mice (Fig. 8e), indicating that the delayed neutrophil death in infected hosts is mainly determined by the activation of apoptotic factors and pathways rather than the level of the related proteins.

Discussion

In this study, we used single-cell transcriptome profiling to reveal neutrophil heterogeneity and orchestrated maturation during bacterial infection reprograms the neutrophil population structure and dynamic transitions between subpopulations. In BM, the proportion of G1 cells increased during bacterial infection, indicating elevated proliferation of myeloid progenitors (Fig. 5b and Fig. 8a). The percentage of BM G2 cells remained the same, suggesting balanced influx from G1 cells and transformation from G2 to G3 cells. Velocity analysis showed that the obvious transformation from G3 to G5a cells under homeostatic conditions was suppressed during infection, and G3 cells in infected hosts predominantly differentiated to G4 cells (Fig. 8b and Supplementary Fig. 5). G4 cells decreased from 38 to 30% in BM but significantly increased in PB and spleen (Fig. 5b and Fig. 8a). PB G4 cells were mainly derived from BM G3 cells (Fig. 8b). Additionally, infection significantly suppressed the G5a and G5b to G5c transition, leading to a smaller G5c population in E. coli-challenged hosts (Fig. 5b and Fig. 8b). In the velocity analysis, cells from BM and PB were analyzed together to reveal the origin and inter-relationship of neutrophil subpopulations in each compartment. Regardless of whether we investigated all (G0–G5) (Fig. 8b and Supplementary Fig. 5) or only post-mitotic (G3–G5) (Supplementary Fig. 6) neutrophils, a similar pattern of cell transition and fate direction was observed. In PB and spleen, the conversion between G5a and G5b neutrophils was rarely detected and G5c clearly localized at the end of the neutrophil maturation trajectory (Supplementary Fig. 6). A significant number of G5b neutrophils also existed in BM (Extended Data Fig. 5d). A bidirectional transition was observed in the BM G5b population with BM G4 and PB/spleen G5b at each end. This switched to an overly unidirectional transition from BM G5b to PB/spleen G5b during bacterial infection (Supplementary Fig. 6).

To assess neutrophil heterogeneity at the site of infection, neutrophils were extracted from the inflamed peritoneal cavity (Fig. 8c,d). Velocity analysis on all G3–G5 cells in E. coli-challenged mice, including those in BM, PB, spleen and peritoneal cavity, revealed irregular multiple directions and short vectors in peritoneal cavity G5 populations, indicating a rather inactive transition among these populations (Supplementary Fig. 7). Interestingly, although G5c cells accounted for only about 25% of PB neutrophils before bacterial challenge and <10% of PB neutrophils after bacterial challenge, >45% of peritoneal cavity neutrophils in challenged mice were G5c cells (Fig. 8d), suggesting that these cells may possess higher transendothelial migration capability than G5a or G5b cells. G5c neutrophils showed the highest aging score compared with other G5 cells (Fig. 8i,j) with no difference detected among BM, PB and spleen G5c cells in unchallenged mice (Supplementary Fig. 8a). During bacterial infection, the aging score and the proportion of aged cells in the peritoneal cavity G5c population became significantly higher compared with in the PB and spleen G5c populations (Supplementary Fig. 8b). These results are in agreement with a previous study showing that in the acute inflammatory response during endotoxicemia aged neutrophils stop returning to BM and instead rapidly migrate to sites of inflammation46,47. Infection delays neutrophil death46. Paradoxically, genes related to apoptosis or necroptosis were significantly upregulated in every neutrophil subpopulation in E. coli-challenged mice (Fig. 8e), indicating that the delayed neutrophil death in infected hosts is mainly determined by the activation of apoptotic factors and pathways rather than the level of the related proteins.
homeostasis and bacterial infection. The scRNA-seq-defined neutrophil populations characterized here were correlated with classical morphology-based and various previously reported distinct BM neutrophil subpopulations arising during differentiation and maturation (Extended Data Fig. 10). In an effort to unify the naming scheme for cellular clusters in neutrophil development, a set of recently used terms for differentiating neutrophils were adopted. Based on correlation analyses (Extended Data Fig. 2), the G0, G1, G2, G3, and G4 clusters aligned to BM GMP, proNeu, preNeu, immNeu and mNeu, respectively (Fig. 8f).

PB contained three main neutrophil subsets (G5a, G5b and G5c), which represented the most mature neutrophils with typical polymorphonuclear morphology and were named PMNa, PMNb and PMNc, respectively (Fig. 8f). They are three transcriptionally distinct mature neutrophil subpopulations that may be pre-programmed with different functions. First, PMNb were discrete and definable ISG-expressing neutrophils. They were more similar to PMNa than PMNc neutrophils, and the majority of PMNb neutrophils directly developed from BM mNeu (Fig. 8f).

Interferon and interferon-related pathways are implicated in both viral and non-viral infections and play a critical role in host defenses. ISG-related PMNb neutrophils may be primed to combat invading pathogens even before infection occurs. Interestingly, a group of ISG-expressing tumor-infiltrating neutrophils were recently identified in human and mouse lung cancers, their transcriptome was significantly different from that of PMNb neutrophils, indicating significant neutrophil reprogramming in the tumor microenvironment. Second, the difference between PMNa, PMNb and PMNc is unlikely to be due to mechanically induced cellular responses caused by transendothelial migration. Most neutrophils in PB have experienced the same mobilization from BM to the circulation and have never crossed capillaries. Third, the difference between PMNa, PMNb and PMNc is not a result of neutrophil activation. These distinct subpopulations existed in unchallenged mice, and their identity was stable and largely maintained during bacterial infection. Finally, these distinct subpopulations are not a direct result of neutrophil aging or death. Although PMNb clearly localized at the end of the neutrophil maturation trajectory and had a higher aging score than PMNa and PMNc (Fig. 2j), only 15% of PMNb neutrophils could be defined as aged neutrophils. Aged neutrophils also existed in the PMNa and PMNb populations, albeit at lower percentages. Additionally, PMNc marker genes were significantly enriched for ribosome biogenesis, cytoplasmic translation, post-transcriptional regulation and lipopolysaccharide-mediated signaling pathways by Gene Ontology analysis, indicating the high functionality of this population.

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/s41590-020-0736-z.

Received: 15 October 2019; Accepted: 11 June 2020; Published online: 27 July 2020

References
1. Nicolas-Avila, J. A., Adrover, J. M. & Hidalgo, A. Neutrophils in homeostasis, immunity, and cancer. *Immunity* **46**, 15–28 (2017).
2. Nauseef, W. M. & Borregaard, N. Neutrophils at work. *Nat. Immunol.* **15**, 602–611 (2014).
3. Silvestre-Roig, C., Hidalgo, A. & Soehnlein, O. Neutrophil heterogeneity: implications for homeostasis and pathogenesis. *BLOOD* **127**, 2173–2181 (2016).
4. Ley, K. et al. Neutrophils: new insights and open questions. *Sci. Immunol.* 3, eaat4579 (2018).
5. Scapini, P., Marini, O., Tecchio, C. & Cassatella, M. A. Human neutrophils in the saga of cellular heterogeneity: insights and open questions. *Immunol. Rev.* **273**, 48–60 (2016).
6. Ng, L. G., Ostuni, R. & Hidalgo, A. Heterogeneity of neutrophils. *Nat. Rev. Immunol.* **19**, 255–265 (2019).
7. Adrover, J. M., Nicolas-Avila, J. A. & Hidalgo, A. Aging: a temporal dimension for neutrophils. *Trends Immunol.* **37**, 334–345 (2016).
8. Yvan-Charvet, L. & Ng, L. G. Granulopoiesis and neutrophil homeostasis: a metabolic, daily balancing act. *Trends Immunol.* **40**, 598–612 (2019).
9. Doerschuk, C. M. Leukocyte trafficking in alveoli and airway passages. *Respir. Res.* **1**, 136–140 (2000).
10. Wang, Q. & Doerschuk, C. M. The signaling pathways induced by neutrophil–endothelial cell adhesion. *Antioxid. Redox Signal.* **4**, 1447–1452 (2002).
11. Adlung, L. & Amit, I. From the Human Cell Atlas to dynamic immune maps in human disease. *Nat. Rev. Immunol.* **18**, 597–598 (2018).
12. Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A. & Teichmann, S. A. Single-cell transcriptomics to explore the immune system in health and disease. *Science* **358**, 58–63 (2017).
13. Giladi, A. et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat. Cell Biol.* **20**, 836–846 (2018).
14. Nestorowa, S. et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20–e31 (2016).
15. Pelsen, L. et al. Human hematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).
16. Paul, F. et al. Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell 163*, 1663–1677 (2015).
17. Karamitos, D. et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat. Immunol.* **19**, 85–97 (2018).
18. Qiu, X. et al. Reversed gradient embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
19. Eash, K. J., Greenbaum, A. M., Gopalan, P. K. & Link, D. C. CXCR2 and CXCR4 antagonistically regulate neutrophil trafficking from murine bone marrow. *J. Clin. Invest.* **120**, 2423–2431 (2010).
20. Cowland, J. B. & Borregaard, N. Granulopoiesis and granules of human neutrophils. *Immunol. Rev.* **237**, 11–28 (2016).
21. Borregaard, N., Sorensen, O. E. & Theilgaard-Møch, K. Neutrophil granules: a library of innate immunity proteins. *Trends Immunol.* **28**, 340–345 (2007).
22. Sorensen, O., Arnliots, K., Cowland, J. B., Rainton, D. F. & Borregaard, N. The human antibacterial cathelicidin, hCAP-18, is synthesized in myelocytes and metamyelocytes and localized to specific granules in neutrophils. *Blood* **90**, 2796–2803 (1997).
23. Hoogendijk, A. J. et al. Dynamic transcriptome–proteome correlation networks reveal human myeloid differentiation and neutrophil-specific programming. *Cell Rep.* **29**, 2505–2519.e4 (2019).
24. Saltake, S. et al. C/EBPP is involved in the amplification of early granulocyte precursors during candidemia-induced “emergency” granulopoiesis. *J. Immunol.* **189**, 4546–4555 (2012).
25. Olsson, A. et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
26. Evrad, M. et al. Developmental analysis of bone marrow neutrophils reveals populations specialized in expansion, trafficking, and effector functions. *Immunity* **48**, 364–379.e8 (2018).
27. 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**, 2339–2341.e8 (2018).
28. Kim, M.-H. et al. A late-lineage murine neutrophil precursor population exhibits dynamic changes during demand-adapted granulopoiesis. *Sci. Rep.* **7**, 39804 (2017).
29. Muench, D. E. et al. Mouse models of neutropenia reveal progenitor-stage-specific defects. *Nature* **2**, 109–114 (2020).
30. Kwok, I. et al. Combinatorial single-cell analyses of granulocyte–monocyte progenitor heterogeneity reveals an early uni-potent neutrophil progenitor. *Immunity* https://doi.org/10.1016/j.immuni.2020.06.005 (2020).
31. Borregaard, N. & Herlin, T. Energy metabolism of human neutrophils during phagocytosis. *J. Clin. Invest.* **70**, 550–557 (1982).
32. Chervenick, P. A., Boggs, D. R., Marsh, J. C., Cartwright, G. E. & Wintrobe, M. M. Quantitative studies of blood and bone marrow neutrophils in normal mice. *Am. J. Physiol.* **215**, 353–360 (1968).
33. Colvin, G. A. et al. Murine marrow cellularity and the concept of stem cell competition: geographic and quantitative determinants in stem cell biology. *Leukemia* **18**, 575–583 (2004).
34. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
35. Broxmeyer, H. E. Chemokines in hematopoiesis. *Curr. Opin. Hematol.* **15**, 49–58 (2008).
36. Furze, R. C. & Rankin, S. M. Neutrophil mobilization and clearance in the bone marrow. *Immunology* **125**, 281–288 (2008).
37. Suratt, B. T. et al. Neutrophil maturation and activation determine anatomic site of clearance from circulation. *Am. J. Physiol. Lung Cell Mol. Physiol.* **281**, 1913–1921 (2001).
38. Theilgaard-Monch, K. et al. The transcriptional program of terminal granulocytic differentiation. *Blood* **105**, 1785–1796 (2005).
39. Monticelli, S. & Natoli, G. Transcriptional determination and functional specificity of myeloid cells: making sense of diversity. *Nat. Rev. Immunol.* **17**, 595–607 (2017).
40. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
41. Lambrechts, D. et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289 (2018).
42. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
43. Kwak, H. J. et al. Myeloid cell-derived reactive oxygen species externally regulate the proliferation of myeloid progenitors in emergency granulopoiesis. *Immunity* **42**, 159–171 (2015).
44. Manz, M. G. & Boettcher, S. Emergency granulopoiesis. *Nat. Rev. Immunol.* **14**, 302–314 (2014).
45. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
46. Uhl, B. et al. Aged neutrophils contribute to the first line of defense in the acute inflammatory response. *Blood* **128**, 2327–2337 (2016).
47. Kołaczkowska, E. The older the faster: aged neutrophils in inflammation. *Blood* **128**, 2280–2282 (2016).
48. Luo, H. R. & Loison, F. Constitutive neutrophil apoptosis: mechanisms and regulation. *Am. J. Hematol.* **83**, 288–295 (2008).
49. Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* **32**, 513–545 (2014).
50. 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.e10 (2019).

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. 2020
Methods

Mouse strains. Female C57BL/6 mice were purchased from The Jackson Laboratory. Eight- to ten-week-old mice were used in all of the experiments. All animal experiments were conducted in accordance with the Animal Welfare Guidelines of the Children’s Hospital Boston. The Children’s Hospital Animal Care and Use Committee approved and monitored all of the procedures.

Mouse peritonitis model. Wild-type mice were intraperitoneally injected with 1×10^9 E. coli (ATCC 19138) in 300 µl phosphate-buffered saline (PBS). At different time points after injection, mice were anesthetized with isoflurane. retro-orbital blood was collected, and then mice were sacrificed by euthanizing with CO2 (ref. 1). Cells from different organs such as BM, spleen, liver and peritoneal exudate were collected as detailed below.

Mouse neutrophil isolation. Neutrophils display circadian oscillations in number and phenotype, and neutrophil aging is an intrinsically driven bona fide circadian process45. Thus, all samples in this study were prepared from mice sacrificed at the same time in the morning (08:00). Eight-week-old female mice were used for scRNA-seq analysis. PB (600–800 µl) was collected by retro-orbital bleeding and diluted with 3 ml Hank’s balanced salt solution (HBSS) containing 15 mM EDTA 5. Cells were centrifuged for 10 min at 500 g. Red blood cells were lysed by resuspension in 5 ml ammonium-chloride-potassium (ACK) lysis buffer (Thermo Fisher Scientific) for 5 min at room temperature. Next, 10 ml RPMI+2% fetal bovine serum were added to stop lysis followed by centrifugation at 500g for 5 min. Cells were washed twice with 10 ml HBSS+2 mM EDTA +1% fetal bovine serum (FBS) to remove remaining red blood cells. BM cells were first enriched by positive selection using c-kit (CD117) microbeads (Miltenyi Biotec) and further purified by FACS sorting c-kit+ cells. To isolate spleen neutrophils, spleens were dissected, placed in 3 ml PBS + 1 mM EDTA and then gently disaggregated through a 70 µm cell strainer using a 1 ml syringe plunger. Whole spleen cells were collected, centrifuged and resuspended in 1 ml PBS + 1 mM EDTA. The red blood cells were then lysed with 5 ml ACK lysis buffer for 2 min at room temperature. After centrifugation, cells were washed twice with PBS + 1 mM EDTA and resuspended in 500 µl PBS + 1% BSA. BM and blood neutrophils were harvested by sorting with the following fluorescent-conjugated antibodies: APC-conjugated anti-CD11b; APC/ 

EDU incorporation assay. EdU—a thymidine analog—can track cells post-mitotically in BM and PB (Fig. 7c). EdU is incorporated into DNA in the S phase of the cell cycle, and the half-life of EdU is only about 30 min, so incorporation can only occur in the first 1–2 h after EdU intraperitoneal injection. After 1 h of intraperitoneal injection with 0.5 mg EdU, mice were injected with E. coli as above to induce peritonitis. Mice were sacrificed at designated time points, and BM, blood and spleen cells were harvested followed by staining with the following fluorescent-conjugated antibodies: APC-conjugated anti-CD11b; APC/ 

Spleen cryosection preparation. Spleens were fixed in 1% formaldehyde (StatLab) for 4–8 h, dehydrated in 30% sucrose solution for 72 h, and snap frozen in O.C.T. (Sakura Finetek Japan). Single-cell-thick (5 µm) spleen cryosections were obtained using a Leica Cryostat and the CryoJane tape transfer system (Leica Microsystems). For immunofluorescence staining, slides were rehydrated in PBS for 10 min followed by rinsing in PBST (PBS + 0.1% Tween 20). Blocking was performed with PBS + 10% donkey serum for 20 min. The diluted primary rat anti-S100a8 (Thermo Fisher Scientific; 335308) and rabbit anti-iFT1 (Abscam; ab263256) antibodies were added and incubated for 1 h at room temperature. After three washes with PBST, Alexa Fluor 488-conjugated donkey anti-rat antibody (Jackson Immunoresearch; 141697) and Cy3-conjugated donkey anti-rabbit antibody (Jackson Immunoresearch; 143460) were added and incubated for 30 min at room temperature. Slides were washed 3x with PBST and then stained with DAPI (0.5 µM) for 3 min. Slides were rinsed in PB and were covered with mounting solution (Vectorshield; Vector Laboratories).

Laser scanning cytometry (LSC). LSC is an emerging technology that images and quantitatively analyzes cellular and subcellular features within tissues, re-interrogating identified cell subpopulation(s) for in situ characterization of the molecular and cellular points associated with those cells. LSC was performed with an iCyS Research Imaging Cytometer four-laser system (Throlabs)43. Each section was first scanned with a 10x objective using the 405 nm laser to generate low-resolution images of the LSC-imaged nuclei and obtain a general view of the spleen. Subsequently, the sections were divided into small regions and scanned with a dry objective lens to create high-resolution field images. Data were analyzed using iCyS Cytometric Analysis Software (Throlabs).

ConfoCal imaging. Sections with a thickness of 20 µm were prepared and stained as described above. Confocal images were obtained using the Zeiss LSM 710 laser scanning confocal microscope (Carl Zeiss AG). Data were analyzed using Imaris Software (Oxford Instruments).

Intracellular protein staining for FACS analysis. PB and spleen cell suspensions were prepared as described above. After being washed with PBS twice, cells were blocked with rat anti-mouse CD16/CD32 antibody on ice for 10 min. APC-conjugated anti-CD45, APC/cy7-conjugated anti-CD11b, PE-conjugated anti-Ly6G and BV711-conjugated anti-CD44 antibodies were added and incubated for 1 h at room temperature. Cells were then incubated for 20 min at 4°C in the dark. Cells were washed with PBS twice, fixed and permeabilized with 1 ml PBS containing 4% paraformaldehyde (Electron Microscopy Sciences) and 0.1% saponin (Sigma–Aldrich) at 4°C for 30 min, and pelleted by centrifugation at 3,000g for 3 min at 4°C. After being washed with 1 ml anti-CD34. Myeloblasts, promyelocytes, metamyelocytes, myelocytes, mature band cells and segmented neutrophils were sorted with a MoFlo cell sorter (Beckman Coulter). Total RNA was extracted from those populations using the Qagen RNase Mini Kit (Qagen). RNA quality was evaluated spectrophotometrically, and the quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies). All samples showed RNA integrity of >7.5. RNA-seq libraries were prepared using the KAPA mRNA HyperPrep Kit (Illumina). Once prepared, indexed cDNA libraries were pooled in equimolar amounts and sequenced with paired-end reads on an Illumina HiSeq 2500.

Wright–Giemsa staining and examination of morphology-defined neutrophil populations. The first recognizable cells of neutrophil lineage in BM are myeloblasts, which are characterized by a high nuclear-to-cytoplasmic ratio and dispersed chromatin. Myeloblasts then irreversibly differentiate into promyelocytes, which are characterized by a round nucleus and azurophil granules, followed by myelocytes characterized by a round nucleus and specific granules. Metamyelocytes are characterized by nuclear indentations (kidney-shaped nucleus) and the emergence of secretory vesicles. Finally, metamyelocytes are divided into band cells with a band-shaped nucleus and segmented cells (segmented neutrophil; also known as polymorphonuclear cells). Cells were sorted (Fig. 2d) and concentrated onto microscope slides by cytopinning. Slides were dried and stained using the Diff–Quick Stain Set (Siemens). Stained slides were rinsed under running tap water and air-dried for 10 min. Images were obtained under a microscope with a 63x objective.
Wash Buffer (PBS containing 0.2% BSA and 0.1% saponin), cells were blocked for 30 min with mixing buffer (PBS containing 5% goat serum and 5% BSA) and then stained with anti-IFIT1/p56 antibody (Sigma–Aldrich) for 30 min at 4 °C in 200 μl staining buffer (PBS containing 5% goat serum, 5% BSA and 0.1% saponin). Cells were washed twice with 1 ml Wash buffer followed by incubation with the secondary goat anti-rabbit-Alexa Fluor 488 antibody (Invitrogen) in staining buffer for 30 min. Cells were then washed with 1 ml Wash buffer and resuspended in 1 ml PBS containing 0.5% BSA. The stained cells were either sorted with a FACSaria III cell sorter (BD Biosciences) or analyzed on an Attune NxT Flow Cytometer (Thermo Fisher Scientific). Fixation, washing, staining and sorting were performed at a concentration of 5–10×10^6 cells per ml.

RNA purification, cDNA synthesis and preamplification and quantitative PCR (qPCR). Total RNA was extracted from sorted fixed cells using a RecoverAll Total Nucleic Acid Isolation Kit (Thermo Fisher Scientific) according to the manufacturer's recommendations. The pre-amplified cDNA was subjected to qPCR in which the amplified product was detected using TB Green Premix Ex Taq (Tli RNase H Plus) (Takara Bio) on a CFX96 Real-Time PCR Detection System (Bio-Rad). ΔΔCt was calculated using GAPDH as a normalizer.

scRNA-seq data processing. The quality of sequencing reads was evaluated using FastQC and MultiQC. Cell Ranger version 2.2.0 was used to align the sequencing reads (fastq) to the mm10 mouse transcriptome and quantify the expression of transcripts in each cell. The pipeline resulted in a gene expression matrix for each sample, which records the number of UMIs for each gene associated with each cell barcode. For human data, sequenced reads were aligned to the hg38 human transcriptome, then the expression of transcripts in each cell was quantified using the BD Rhapsody Whole Transcriptome Assay Analysis Pipeline. Unless otherwise stated, all downstream analyses were implemented using R version 3.5.2 and the BD Rhapsody Whole Transcriptome Assay Analysis Pipeline. Unless otherwise stated, all downstream analyses were implemented using R version 3.5.2 and the BD Rhapsody Whole Transcriptome Assay Analysis Pipeline.

Dimension reduction. Dimension reduction was performed at three stages of the analysis: the selection of variable genes; PCA; and UMAP. The FindVariableGenes function (y.cutoff = 1.2 for total control cells; y.cutoff = 1.2 for control neutrophils; y.cutoff = 0.7 for E. coli–challenged total cells) was applied to select highly variable genes covering most of the biological information contained in the whole transcriptome. Then, the variable genes were used for PCA implemented with the RunPCA function. Next, we selected principal components 1–20 (for total cells) or 1–15 (for neutrophils) as input to perform the RunUMAP function to obtain bidimensional coordinates for each cell.

Unsupervised clustering and annotation. We performed the FindClusters function (resolution: 0.3, 0.6 and 0.2 for control total cells, neutrophils and E. coli–challenged total cells, respectively) to cluster cells using the Louvain algorithm based on the same principal components as for the RunUMAP function. Clusters G1–G5 were neutrophils at different maturation stages. G1 and G2 were early-stage neutrophils with a higher expression of Elane, Mpo, Ftbh and Camp (Fig. 1d,e). Neutrophils are terminally differentiated. The transition from a proliferative cell to terminal differentiation was accomplished by a dramatic change in the expression of the important cell cycle regulatory proteins, so we next performed a single-cell-resolution analysis of cell cycle activation during neutrophil differentiation based on the expression of G1S- and G2/M-phase-specific genes (Fig. 1f). Cells in the G0 to G2 stages underwent active proliferation, while cell division stopped abruptly thereafter. CDC28 protein kinase regulatory subunit 2 (CDSK2), Mki67 and Cdc20 were all strongly downregulated at the messenger RNA (mRNA) level.

Identification of DEGs. We used the FindMarkers or FindAllMarkers function with test.use = “m”, logfc.threshold = log1(1.5) based on normalized data to identify DEGs. Pvalue adjustment was performed using Bonferroni correction based on the total number of genes in the dataset. DEGs with adjusted Pvalues < 0.05 were filtered out. Gene Ontology analysis was performed by using the R package clusterProfiler. In the experiment described in Extended Data Fig. 7, we conducted differential gene expression analysis in each neutrophil subpopulation using the non-parametric Wilcoxon rank-sum test and identified DEGs with an average expression fold-change > 2.

Developmental trajectory inference. Pseudo-time was generated with Monocle version 2 (ref. 21) to infer the potential lineage differentiation trajectory. The newCellDataSet function (lowerDetectionLimit = 0.5; expressionFamily = negbinomial.size) was used to build the object based on the above highly variable genes identified by Seurat version 2.3.4.

Bulk RNA-seq analysis. The quality of sequencing reads was evaluated using FastQC and MultiQC. Adapter sequences and low-quality score bases were trimmed using trimmomatic/0.36. The resulting reads were then mapped to the mouse reference sequence (GRCh38/mm10; Ensembl release 81) and counted using STAR/2.5.2b alignment software. Gene differential expression analysis was performed using the R package EdgeR.

Scoring of biological processes. Individual cells were scored for their expression of gene signatures representing certain biological functions. For all signatures except neutrophil aging, functional scores were defined as the average normalized expression of corresponding genes. Aging score was defined as the weighted average of Z scores of age-related genes, where the Z scores were calculated by scaling the normalized expression of a gene across all cells. Gene weights were set to either 1 or –1 to reflect positive or negative relationships. The neutrophil maturation signature was derived by identifying the top 50 DEGs (as listed in Supplementary Table 4) with the highest fold-changes and adjusted Pvalues < 0.05 between the mature cluster (G4) and immature clusters (G0–G3). Granule signatures were from ref. 17. Other functional signatures were derived from the Gene Ontology database11, with the full gene list provided in Supplementary Table 4. For instance, to access the phagocytosis function at the transcript level, we determined a phagocytosis score by calculating the average expression of genes in the Gene Ontology term ‘phagocytosis, engulfment’ (GO:0006911). The apoptosis score was measured by the upregulation of the integrated proapoptotic pathway (Fig. 4b). To further dissect apoptotic heterogeneity in G5 populations independent of transcriptome-based sub-clustering, we fit a two-component Gaussian mixture model to the apoptotic heterogeneity score of all G5 cells using the R package mixtools version 1.1.0 (ref. 22). We then clustered the samples on the mean of the mixture component with the highest G5 proportion and assigned each cell to one of the two groups based on its posterior (Fig. 4c).

Age-related genes were summarized from the previous literature (Fig. 2). Aging is a main mechanism that accounts for neutrophil heterogeneity7,63: aged neutrophils are smaller with fewer granules and granular multi-lobed nuclei and produce more neutrophil extracellular traps (NETs). Related to function, aged neutrophils express less of the adhesion molecule t-selectin (CD62L; encoded by Sell) and more CD11b (n; encoded by Itgam), lymphocyte function-associated antigen-1 (CD11a/i2), CD49d (integrin α6; encoded by Itgad), TLR4, ICAM-1, CD44 and CD11c (encoded by Itgax). Additionally, aged neutrophils express more surface CXCR4 and less CXCR2, which regulates their release from and return to BM. CXCR4 also play a role in the exiting aged, senescent neutrophils at BM sites. Anti-CXCR4 antibodies or CXCR4 antagonists impede neutrophil homing to BM64. Finally, aged neutrophils exhibit increased expression of CD24 (a glycosylphosphatidylinositol-linked glycoprotein that induces apoptosis when crosslinked) and reduced expression of CD47 (the ‘don’t eat me’ signal that inhibits efferocytosis—a process leading to clearance of dead neutrophils).

ROS-mediated pathogen killing is a major host defense mechanism. In neutrophils, ROSs are mainly produced by the phagocytic NADPH oxidase (aka the NOX2 complex). During cell activation, cytosolic components of the NADPH oxidase NCF2 (p67phox), Rac1 and/or Rac2, NCF4 (p40-phox) and NCF1 (p47phox) are recruited to the membrane to form a complex with membrane bound CYBA (p22-phox) and CYBB (p40-subunit beta). We evaluated the NADPH oxidase score based on the expression of the seven NADPH oxidase-related genes (Supplementary Fig. 6d).

Comparison of scRNA-seq-defined populations with morphology-defined neutrophil subpopulations. To benchmark single-cell transcriptomic
neutrophil classification against existing morphological classification schemes, we deconvoluted bulk RNA-seq profiles based on the expression of scRNA-seq-identified group-specific signatures. This approach was similar to other existing deconvolution methods such as CIBERSORT\(^\text{17}\), but we used a linear regression model with the constraint of non-negative coefficients (that is, the non-negative least-squares problem) instead of the linear support vector regression in CIBERSORT. Although we manually chose 20 genes with the highest fold-changes as signatures for each single-cell group, we noted that the deconvolution in this case was robust to the choice of signatures. The regression model was built using the R package rtlns (version 1.4.1). Bulk profiles were quantile normalized.

At different morphology-defined neutrophil differentiation stages, neutrophils produce different granules containing distinct enzymes and antimicrobial compounds. Thus, we also examined the expression of various granule genes in differentiating neutrophils. Genes related to primary (azurophilic) granules such as *Mpo* started to be expressed in some G0 cells, peaked in G1 cells and then rapidly decreased in G2 cells (Fig. 2a–b). Myeloperoxidase (MPO)-negative granules can be divided into granules containing LTF but no gelatinase (MMP9), granules that contain both and granules that contain gelatinase but no LTF\(^\text{11}\). We found sequential production of these granules in maturing neutrophils, with LTF-containing granules emerging in G2 cells, LTF and gelatinase-containing granules emerging in G3 cells, and gelatinase-containing granules (LTF low) emerging in G4 cells (Fig. 2a–c). Of the proteins that localize exclusively to secretory vesicles such as *PFP1* (encoded by *Frp1*) and *VAMP2* (encoded by *Vamp2*), their cognate mRNA profiles peaked in G4 cells in BM and continued to be expressed in PB neutrophils.

**SCENIC analysis.** SCENIC is a computational tool that infers regulatory modules or regulons by analyzing the co-expression of transcription factors and their putative target genes characterized by enrichment of corresponding transcription factor binding sites in regulatory regions. SCENIC was performed on all control and *E. coli*-challenged samples using the Python package pySCENIC (version 0.9.11)\(^\text{40}\) with default parameters. We scaled the network challenge samples using the Python package E. coli-performed on all control and neutrophil classification against existing morphological classification P-values and corresponding to different biological questions such as neutrophil cluster transition putative target genes characterized by enrichment of corresponding transcription velocities of each cell were estimated using the gene.relative.velocity.estimates approach as was used for comparing scRNA-seq-defined populations with morphology-defined neutrophil subpopulations (see above). We deconvoluted bulk RNA-seq profiles of BM GMPs, preNeu cells, immature neutrophils and mature neutrophils, as well as PB neutrophils based on their expression of scRNA-seq-identified group-specific signatures. The 20 highest DEGs of each single-cell group (G0–G5) were selected as signatures for deconvolution.

Finally, Giladi et al.\(^\text{13}\) also defined two BM neutrophil subpopulations. c-Kit\(^+\) stage I neutrophils express a set of genes used to approximate the neutrophil differentiation axis, while stage II neutrophils display a mature neutrophil signature defined by genes upregulated in the most terminally differentiated neutrophils. The initial increase in the expression of the maturation signature is independent of PU.1, but further neutrophil maturation and activation of stage II genes is completely blocked in PU.1 knockout\(^\text{11}\). To reveal the correlation of these neutrophil subtypes with scRNA-seq-defined neutrophil populations, we applied the same regression-based deconvolution approach as was used for comparing scRNA-seq-defined populations with morphology-defined neutrophil subpopulations (see above). We deconvoluted bulk RNA-seq profiles of BM GMPs, preNeu cells, immature neutrophils and mature neutrophils, as well as PB neutrophils based on their expression of scRNA-seq-identified group-specific signatures. The 20 highest DEGs of each single-cell group (G0–G5) were selected as signatures for deconvolution.

**Cell label transfer.** Total cells were partitioned into distinct cell types annotated by the expression of known marker genes. Neutrophils in their steady state were partitioned into eight clusters based on gene expression profiles annotated according to their development order. *E. coli*-challenged neutrophils were annotated using a well-accepted method\(^\text{12}\). Briefly, we first identified pairwise correspondences (as anchors) between single cells across datasets (before and after *E. coli* challenge) to quantify the batch effect. Each cell in the *E. coli*-challenged dataset was then annotated based on the transcriptomic similarity between this cell and cells in the reference dataset. Specifically, cells would receive corresponding labels with the highest similarity scores, whereas cells with the highest similarity score lower than 0.5 were defined as unassigned. The unassigned cells accounted for c<010\% of the total cell population, distributed randomly on the UMAP plot, and thus were excluded from further investigation. In this way, each neutrophil from the new stimulated dataset was assigned a cluster name, and neutrophils sharing similar transcriptomic profiles were placed into the same cluster. Hence, each cell in the bacterial infection state was assigned to one of the nine cluster labels. This transfer procedure was implemented using the FindTransferAnchors (dists = [115], and TransferData (dists = [115]) functions in Seurat version 3.0.2 (ref.\(^\text{29}\)) with the combination of top 100 DEGs of each cluster.

**Correlation of scRNA-seq-defined neutrophil populations with previously reported neutrophil subpopulations.** Previous studies revealed a variety of distinct BM neutrophil subpopulations arising during differentiation and maturation. Using scRNA-seq coupled with a new analytical tool, iterative clustering and guide-gene selection and clonogenic assays, Olsson et al.\(^\text{25}\) analyzed discrete genomic states and the transitional intermediates that span myelopoiesis. They performed scRNA-seq on stem/multipotent progenitor cells, CMP cells, GMP cells and LK CD34\(^+\) cells (lin−*c-Kit*−*CD34*+) that included granulocytic precursors. We calculated the fraction of each scRNA-seq-defined cluster cell in the four samples per cell cluster in the four samples per cell cluster using the cell label transfer method described above. The cluster identity of each cell was inferred based on the transcriptomic similarity between this cell and the reference clusters (G0–G5) defined in the current study. This same method was used to analyze the C1 and C2 neutrophil clusters reported by Zhu et al.\(^\text{12}\). Recently, a proliferative unipotent neutrophil precursor that suppresses T cell activation and promotes tumor growth was identified in the mouse BM that generates neutrophils after intra-BM adoptive transfer. scRNA-seq analysis of BM Lin−*c-Kit*−*Ly6a*−*Ly6c*− cells revealed two populations: an early-stage c-Kit*Gfi1*−*Cebpα*Ly6g*−* progenitor with stem cell morphology (C1) and a late-stage *c-Kit*−*Gfi1*−*Cebpα*Ly6g*−* precursor with more mature phenotypic features similar to transient neutrophil precursors (C2)\(^\text{13}\). Further analysis showed that cluster C1 is the early-stage committed unipotent neutrophil progenitor. Interestingly, the late-stage progenitors were mostly similar to the preNeu population identified by Evrard et al.\(^\text{12}\). In the current study, the raw data related to C1 and C2 cells were retrieved and reanalyzed. We annotated each cell cluster in the four samples per cell cluster in the four samples per cell cluster using the Neutrophil Classification Reference (NCBR)\(^\text{27}\) (principal components 1–12; resolution parameter set at 0.03). Using mass cytometry (CyTOF) and cell cycle–based analysis, Evrard et al.\(^\text{12}\) identified three neutrophil subsets within BM: committed c-Kit*−*Gfi1*−*proliferative neutrophil precursors expressing primary and secondary granule proteins (preNeu); CXCR2*−*non-proliferating immature neutrophils highly expressing secondary granule proteins (immature neutrophils); and CXCR2*−*mature neutrophils highly expressing gelatinase granule proteins (mature neutrophils). To reveal the correlation of these neutrophil subtypes with scRNA-seq-defined neutrophil populations, we applied the same regression-based deconvolution approach as was used for comparing scRNA-seq-defined populations with morphology-defined neutrophil subpopulations (see above). We deconvoluted bulk RNA-seq profiles of BM GMPs, preNeu cells, immature neutrophils and mature neutrophils, as well as PB neutrophils based on their expression of scRNA-seq-identified group-specific signatures. The 20 highest DEGs of each single-cell group (G0–G5) were selected as signatures for deconvolution.

**Statistical analyses.** For most experiments, comparisons were made using a two-tailed, unequal paired Student’s t-test. The values shown in each figure represent mean ± s.d. P<0.05 was considered statistically significant. All statistical analyses and graphics were made using GraphPad Prism (GraphPad) and R (The R Project for Statistical Computing).

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.
References
51. Sakai, J. et al. Reactive oxygen species-induced actin glutathionylation controls actin dynamics in neutrophils. Immunity 37, 1037–1049 (2012).
52. Adrover, J. M. et al. A neutrophil timer coordinates immune defense and vascular protection. Immunity 50, 390–402.e10 (2019).
53. Hou, Q. et al. Inhibition of IP6K1 suppresses neutrophil-mediated pulmonary damage in bacterial pneumonia. Sci. Transl. Med. 10, eaal4045 (2018).
54. Loison, F. et al. Proteinase 3-dependent caspase-3 cleavage modulates neutrophil death and inflammation. J. Clin. Invest. 124, 4445–4458 (2014).
55. Karatepe, K. et al. Proteinase 3 limits the number of hematopoietic stem and progenitor cells in murine bone marrow. Stem Cell Rep. 11, 1092–1105 (2018).
56. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).
57. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. 37, 38–48 (2018).
58. Kowalczyk, M. S. et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res. 25, 1860–1872 (2015).
59. Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature 539, 309–313 (2016).
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).
61. Consortium, G. O. The Gene Ontology resource: 20 years and still GOing strong. Nucleic Acids Res. 47, D350–D358 (2018).
62. Benaglia, T., Chauveau, D., Hunter, D. & Young, D. mixtools: An R package for analyzing finite mixture models. J. Stat. Softw. https://doi.org/10.18637/jss.v032.i06 (2009).
63. Zhang, D. et al. Neutrophil ageing is regulated by the microbiome. Nature 525, 528–532 (2015).
64. Suratt, B. T. et al. Role of the CXCR4/SDF-1 chemokine axis in circulating neutrophil homeostasis. Blood 104, 565–571 (2004).
65. Martin, C. et al. Chemokines acting via CXCR2 and CXCR4 control the release of neutrophils from the bone marrow and their return following senescence. Immunity 19, 583–593 (2003).
66. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods 12, 453–457 (2015).
67. Mullen, K. M. & van Stokkum, I. H. M. nnls: The Lawson–Hanson algorithm for non-negative least squares (NNLS) v1.4 (2012); https://cran.r-project.org/web/packages/nnls/index.html
68. Kjeldsen, L., Bainton, D. F., Sengeløv, H. & Borregaard, N. Structural and functional heterogeneity among peroxidase-negative granules in human neutrophils: identification of a distinct gelatinase-containing granule subset by combined immunocytochemistry and subcellular fractionation. Blood 82, 3183–3191 (1993).
69. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849 (2016).

Acknowledgements
The authors thank A. Hidalgo, L. G. Ng, J. Manis and L. Chai for helpful discussions and suggestions. E.M., Y.X. and T.C. are supported by grants from the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2017-12M-1-015, 2016-12M-1-017 and 2016-12M-1-003), the Non-profit Central Research Institute Fund of the Chinese Academy of Medical Sciences (2018RC31002, 2018PT32034 and 2017PT31033), the National Natural Science Foundation of China (81976107 and 81421002) and the Natural Science Foundation of Tianjin City (18JCJYJC25700). Q.S. and C.L. were supported by the National Natural Science Foundation of China (31871266), Chinese National Key Projects of Research and Development (2016YFA0100103) and NSFC Key Research Grant 71532001. H.R.L. is supported by National Institutes of Health grants (1 R01 AI142642, 1 R01 AI145274, 1 R01 AI141386, R01HL92020 and F01 HL095489) and a grant from FAMRI (CIA 123080). Part of the data analysis was performed on the High Performance Computing Platform of the Center for Life Sciences, Peking University.

Author contributions
H.R.L., C.L. and F.M. conceptualized the study. H.R.L., I.E.S. and C.L. designed the experiments. X.X., P.W., X.Z. and S.Z. acquired samples. Q.S., X.X., P.W., J.S. and X.Z. performed the RNA-seq data analysis. H.R.L., C.L., T.C., Y.X. and I.E.S. provided resources. H.R.L., C.L., Y.X., T.C. and F.M. supervised all of the work. H.R.L., X.X., J.S. and Q.S. prepared the original manuscript. H.R.L., C.L., X.X., J.S., Q.S. and F.M. revised the manuscript. All coauthors read, reviewed and approved the manuscript.

Competing interests
The authors declare no competing interests.
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Sample preparation, quality controls, and related parameters and results related to scRNA-seq analysis.

a, Fluorescence-activated cell sorting (FACS) strategy for scRNA-seq sample preparation. b, Summary of sample information. c, Cell viability percentages immediately before cells were loaded into the 10X Chromium Controller. d, Representative GEM formation after the 10X Chromium Controller under the microscope. e, Violin plots of the number of genes, number of UMIs, mitochondria count percentage, and UMI per gene of all QC-passed cells in different organs. f, Uniform manifold approximation and projection (UMAP) of 19,582 cells from the bone marrow (BM), peripheral blood (PB), and spleen (SP) colored by sample origin and cell type, respectively. Expression of unique genes specifically distinguished each cluster and associated them with neutrophils (Neu) (S100a8 and S100a9), myeloid progenitors (MP) (Cd34, Kit, Mpo and Elane), hematopoietic stem progenitor cells (HSPC, not including MP) (Cd34, Kit, Mpo- and Elane-), monocytes (Mono) (S100a4 and Ccl9/MIP-1γ), B cells (Cd79a and Cd79b), T cells (Cd3d and Ccl5), and dendritic cells (DC) (Siglech), respectively. Cont: contaminated cells. g, Heatmap showing the five highest differentially expressed genes (DEGs) per cell type for all QC-passed cells. h, As in e but using only neutrophils in different organs. i, Comparison of Gr1+ BM neutrophil populations in our data with Ly6g+ BM neutrophil populations in Dr. Ido Amit’s data. Cluster labels are transferred from our data to Dr. Ido Amit’s data13 (Methods). Left: UMAPs of 3591 Gr1+ neutrophils and 2304 Ly6g+ neutrophils colored by data set or cluster identity. Right: Neutrophil compositions in our data and Dr. Ido Amit’s data. j, Violin plots of the number of genes and number of UMIs of our Gr1+ neutrophils and Dr. Ido Amit’s Ly6g+ neutrophils.
Extended Data Fig. 2 | See next page for caption.
Extended Data Fig. 2 | scRNA-seq defined neutrophil populations correlated with previously reported neutrophil subpopulations. a, Correlation of scRNA-seq defined neutrophil populations with the indicated four samples characterized by Olsson et al. Coefficient matrix showing deconvolution results of bulk profiles of indicated neutrophil subpopulations. The 20 highest DEGs per single-cell group (G0-G5) were selected as signatures for deconvolution. Each column is normalized by column sums. b, Correlation of scRNA-seq defined neutrophil populations with the neutrophil subtypes reported by Evrard et al. Coefficient matrix showing deconvolution results of bulk profiles of indicated neutrophil subpopulations. The 20 highest DEGs per single-cell group (G0-G5) were selected as signatures for deconvolution. Each column is normalized by column sums. c, Top: t-distributed stochastic neighbor embedding (t-SNE) plot of the C1 and C2 cells characterized by Zhu et al. The raw data was retrieved from GEO website and reanalyzed. Bottom: t-SNE plot of C1 and C2 cells colored based on scRNA-seq defined clusters (G0-G5). The cluster identity of each cell was determined as described in (A). d, Heatmap showing row-scaled expression of the 10 highest DEGs of C1 and C2 in indicated scRNA-seq defined clusters. Signature genes Ly6g, Cebpα and Cebpe were also included in the map. e, The Stage I (x-axis) and Stage II (y-axis) score of each single cell in the reference sample of current study (Fig.1b). The scRNA-seq defined neutrophil identity of each cell is indicated. f, Violin plots of Stage I and Stage II score for each scRNA-seq defined neutrophil population. g, Correlation between scRNA-seq-defined neutrophil populations and the neutrophil subpopulations reported by Muench et al. Left: The fraction of indicated scRNA-seq defined clusters (G0-G5) in samples characterized by Muench et al. Right: The fraction of indicated scRNA-seq defined clusters (G0-G5) in each neutrophil subtype. Each row is normalized by row sums. The cluster identity of each cell was determined as described in (a). h, Correlation between scRNA-seq-defined neutrophil populations and the neutrophil subpopulations reported by Kwok et al. Shown are the fraction of indicated scRNA-seq defined clusters in each neutrophil subtype. Each row is normalized by row sums.
Extended Data Fig. 3 | Three major neutrophil subpopulations, including an ISG-expressing G5b population, were identified in the PB and spleen. 

a, Heatmap showing row-scaled expression of 47 interferon-stimulated genes (ISGs) for each averaged cluster. 

b, Monocle trajectories of neutrophil population G5a, G5b, and G5c. Each dot represents a single cell. 

c, Violin plots of the number of genes, number of UMI, mitochondria count percentage, and UMI per gene of neutrophils in each cluster.
Extended Data Fig. 4 | Characterization of neutrophil subpopulations. a–d, Violin plot of phagocytosis score (GO:0006911), chemotaxis score (GO:0030593), neutrophil activation score (GO:0042119), and NADPH oxidase score for each cluster. e, Heatmap showing relative expression of seven genes of the NADPH oxidase complex for all neutrophils. f, As in (a–d) but displaying mitochondria-mediated ROS production score (reactive oxygen species biosynthetic process, GO:1903409) for each cluster. g, Violin plots of metabolic scores for each cluster. Glycolysis (Reactome Pathway Database #R-MMU-70171); Oxidative phosphorylation (GO:000619); Electron transport chain (GO:0022900); Tricarboxylic acid cycle (GO:0006099). h–i, Heatmaps showing relative expression of glycolysis-related genes and glucose transport-related genes.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Organ-specific transcriptome features. a, Heatmap showing row-scaled expression of the ten highest DEGs per organ for each averaged organ profile. b, As in a, but for each G5 subpopulation between PB and SP. KEGG analysis of DEGs for each G5 in these two organs. Left: selected KEGG terms with Benjamini-Hochberg-corrected P-values < 0.05 are shown. c, The percentages of each neutrophil subpopulation in the BM and PB calculated based on the scRNA-seq data (Fig. 1c). d, The absolute numbers of each neutrophil subpopulation in the BM and PB calculated based on the percentage in (c) and predicted total BM neutrophil count. e-j, The transcriptome feature of the three G5 populations in the BM, PB, and SP. e, Heatmap showing row-scaled expression of DEGs across organs in each G5 cluster. (f-h), GO analysis of DEGs across organs. Selected GO terms with Benjamini-Hochberg-corrected P-values < 0.05 (one-sided Fisher’s exact test) are shown. i, Violin plots of maturation score and apoptosis score for each G5 neutrophil subpopulation across organs. j, Proportions of apoptotic cells in each G5 neutrophil subpopulation across organs.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Single cell RNA-seq analysis of neutrophils in E. coli-challenged mice. a, Number of white blood cells and the proportion of neutrophils in mice before and after E. coli challenge evaluated by a hematology analyzer (Mindray BC-5000 Vet). Results are the mean ± SD of three independent experiments. b, Experimental scheme of the sample collection process after E. coli challenge. c, Summary of sample information. Organ distribution of neutrophils is shown on the right. d, Cell viability percentages immediately before cells were loaded into the 10X Chromium Controller. e, UMAPs of all 24,943 cells from BM, PB, and SP from E. coli-challenged mice colored by sample origin and cell type, respectively. f, Heatmap showing row-scaled expression of the five highest DEGs for all QC-passed cells colored by cell type. g, Comparisons of the number of genes, number of UMIs, mitochondria count percentage, and UMI per gene of all QC-passed cells in each organ before and after E. coli challenge. h, As in g but only of all neutrophils in each organ. i-j, Heatmaps showing expression of 7 genes of the NADPH oxidase complex (i) and neutrophil granule-related genes (j) for all neutrophils.
Extended Data Fig. 7 | Differentially expressed genes in each neutrophil subpopulation in E. coli-challenged mice. a, MA plots displaying genes that are up-(red) or downregulated (blue) after E. coli challenge for each cluster. Dashed lines denote fold change thresholds used when identifying DEGs. b, Gene ontology (GO) analysis of DEGs before and after E. coli challenge for each cluster. Selected GO terms with Benjamini-Hochberg-corrected P-values < 0.05 (one-sided Fisher’s exact test) are shown.
Extended Data Fig. 8 | Alteration of transcription networks in E. coli-challenged mice. **a**, UMAP of the regulon activity matrix of 32,888 cells (11,992 normal neutrophils, 13,687 challenged neutrophils, and 7209 other cells under normal conditions) colored by Seurat cluster identity (top) or experimental condition (bottom, only neutrophils). **b**, Heatmap of the t-values of regulon activity derived from a generalized linear model for the difference between cells from one challenged neutrophil subpopulation and cells from the corresponding normal subpopulation. Only regulons with at least one absolute t-value greater than 18 are visualized. Regulons are hierarchically clustered based on challenge-response pattern (purple: upregulated, yellow: first up- then downregulated, green: downregulated). **c**, Heatmap showing activity change of regulons identified in **(b)** during normal group transitions.
Extended DataFig. 9  |  Single-cell RNA-seq analysis of human peripheral blood neutrophils. a, Overview of study design and the gating strategy for isolating human PB neutrophils. b, UMAP plots of neutrophils from three healthy donors (D1, D2, or D3) colored by cluster identity. c, The combined UMAP plot of the three donors. d, Dot plot showing scaled expression of selected signature genes for each cluster colored by average expression of each gene in each cluster scaled across all clusters. Dot size represents the percentage of cells in each cluster with more than one read of the corresponding gene. The analysis was conducted using cells from all three human donors. e, Row-scaled expression of the ten highest differentially expressed genes (Bonferroni-corrected P values < 0.05, Student's t-test) in each neutrophil cluster. D1+D2+D3, the analysis was conducted using cells from all three human donors. f, Row-scaled expression of 37 interferon-stimulated genes in each neutrophil cluster. The analysis was conducted using cells from all three human donors.
Extended Data Fig. 10 | Single-cell transcriptome profiling reveals eight neutrophil subpopulations defined by distinct molecular signatures. Summary of dynamic change of morphology, gene expression (Ly6g and c-kit), TF expression (Cebpe and Cebpb) and granules (Azurophil, Specific, Gelatinase granules and Secretory Vesicles) between each subpopulation. Comparison between our scRNA transcriptome profiles with other published neutrophil populations. 

| Reference  | Neutrophil Subpopulation |
|------------|--------------------------|
| Olsson et al., 2016 | LSK/CMP, GMP, LK CD34* |
| Kim et al., 2017 | Neu precursors |
| Zhu et al., 2018 | Early-stage progenitors (C1), Late-stage progenitors (C2) |
| Giladi et al., 2018 | Stage I Neu, Stage II Neu |
| Evrard et al., 2018 | CMP/GMP, proNeu, preNeu, Immature-Neu, Mature-Neu |
| Muench et al., 2020 | CMP/GMP |
| Kwok et al., 2020 | CMP/GMP, proNeu, preNeu, Immature-Neu, Mature-Neu |
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 Authors & Referees 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

- 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 mouse samples, sc-RNA libraries were constructed using 10X genomics Chromium Single Cell 3’ Library & Gel Bead Kit v2. For human samples, sc-RNA libraries were constructed using Targeted mRNA and AbSeq Reagent. Sequencing data were collected through NOVA-seq6000 from illumina platform. For bulk RNA sequencing, total RNA was extracted from indicated populations using the Qiagen RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA quality was evaluated spectrophotometrically, and the quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). All samples showed RNA integrity >7.5. RNA-seq libraries were prepared using the KAPA mRNA HyperPrep Kit (Illumina). Once prepared, indexed cDNA libraries were pooled in equimolar amounts and sequenced with paired-end reads on an Illumina HiSeq2500.

Data analysis

Cell Ranger v2.2.0 was used to align the sequencing reads (fastqs) to the mm10 mouse transcriptome. After gene expression matrix for each sample was created, further analysis was performed using R package Seurat v2.3.4 & v3.0.2, Monocle v2 and Velocity. For human data, sequenced reads were aligned to the hg38 human transcriptome, then the expression of transcripts in each cell was quantified using BD™ Rhapsody Whole Transcriptome Assay Analysis Pipeline. After gene expression matrix was created, further analysis was performed using R package Seurat v2.3.4. For bulk RNA-sequencing analysis, the quality of sequencing reads was evaluated using FastQC and MultiQC. Adaptor sequences and low-quality score bases were trimmed using trimmomatic/0.36. The resulting reads were then mapped to the mouse reference sequence (GRCm38/mm10, Ensemble release 81) and counted using STAR2.5.2b alignment software. Gene differential expression analysis was performed using the R package EdgeR.

Flow cytometry data were analyzed using Flowjo V10.5.

Laser scanning cytometry data were analyzed using iCys Cytometric Analysis Software (Thorlabs).

Confocal microscopy data were analyzed using iCams Software V8.
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

All sequencing data have been deposited at NCBI GEO depository and are accessible with the accession number GSE137540 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE137540).

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-lst.pdf

Life sciences study design

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

**Sample size**

Sample size was determined based on the knowledge on good sample size to ensure adequate data for reliable assessments. Sample sizes are always indicated in figure legends or methods section. Sample size for scRNA-seq analysis is defined by the number of the mice required to obtain sufficient amount of cells for 10X genomics chromium platform. The four bone marrow samples were collected from two mice; the two peripheral blood samples were collected from six mice; the spleen, liver, and peritoneal samples were collected from two mice. The three human peripheral blood samples were collected from three individual healthy donors and processed separately.

**Data exclusions**

No data was excluded from the analysis.

**Replication**

Apart from sc-RNA sequencing, all other experiments were performed at least three times independently and successfully reproduced. Reproducibility of the experiments and significances of the results are shown in details in figure legends or method.

**Randomization**

No randomization was used in this study, because all experimental mice were purchased from the Jackson laboratory and kept under the same environment.

**Blinding**

The investigators were blinded to group allocation during data collection and analysis. The unbiased clustering was conducted using indicated computer software.

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**

| n/a | Involved in the study |
|-----|-----------------------|
| ☑   | Antibodies            |
| ☑   | Eukaryotic cell lines |
| ☑   | Palaeontology         |
| ☑   | Animals and other organisms |
| ☑   | Human research participants |
| ☑   | Clinical data         |

**Methods**

| n/a | Involved in the study |
|-----|-----------------------|
| ☑   | ChIP-seq              |
| ☑   | Flow cytometry        |
| ☑   | MRI-based neuroimaging |

**Antibodies**

Antibodies used

Mouse single cell suspensions were stained with fluorophore-conjugated antibodies (APC/CY7-conjugated anti-Gr1, FITC-conjugated-anti-CD45). Mouse bulk-RNA-seq samples were incubated with biotin-conjugated anti-CD4, biotin-conjugated anti-CD8a, biotin-conjugated anti-ter119, and biotin-conjugated anti-CD20/LD16R antibodies for 20 min, washed with PBS, and then stained with PE/cy7-conjugated streptavidin, APC conjugated anti-e-kit, PE-conjugated anti-Ly6G, and FITC-conjugated anti-CD34 antibodies. For human sc-RNA-seq, the sample was stained with Percp-cy5.5-conjugated anti-human CD33 antibody.

For EDU incorporation assay, cells were stained with APC-conjugated anti-CD11b, APC/cy7-conjugated anti-Ly6G, and PE-
conjugated anti-CXCR4 antibodies.

FITC-conjugated anti-CD34 antibody was purchased from BD bioscience. All other flow antibodies were from Biolegend.

For laser scanning cytometry, samples were incubated with rat anti-510Daβ [Thermo Fisher Scientific #335806] and rabbit anti-iFT1 (Abcam, Cambridge, UK #ab236256) antibodies, followed by secondary staining with Alexa Fluor 488-conjugated donkey anti-rabbit antibody [Jackson Immunoresearch, West Grove, PA #141697] and CyTM3-conjugated donkey anti-rabbit antibody [Jackson Immunoresearch #143460].

For intracellular protein staining, samples were incubated with anti-iFT1/p56 antibody [Sigma #AB117], and then stained with secondary goat anti-rabbit-Alexa Fluor 488 antibody [Invitrogen # A-11008].

Validation

All antibodies used for flow cytometry and cell sorting are well-established lineage marker antibodies. Staining patterns were consistent with the manufacturer product information as well as published data.

Animals and other organisms

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

Laboratory animals  Female C57BL/6 mice were purchased from the Jackson Laboratory. Eight to ten week-old mice were used in all experiments.

Wild animals  No wild animal was used in this study.

Field-collected samples  N/A

Ethics oversight  All animal experiments were conducted in accordance with the Animal Welfare Guidelines of the Children’s Hospital Boston. The Children’s Hospital Animal Care and Use Committee approved and monitored all procedures.

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  Donor #1: a healthy 32-year-old male. Donor #2: a healthy 47-year-old female. Donor #3: a healthy 39-year-old male.

Recruitment  Healthy people undergoing routine medical check-up at participating site were asked whether they were interested in participating in this study. No incentives were provided and there is no bias in selection.

Ethics oversight  The Ethics Committee of Tianjin Blood Disease Hospital approved the study protocol. All participating blood donors provided written informed consent for sample collection and data analysis.

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  For mouse scRNA-seq analysis, all samples were prepared from mice sacrificed at the same time in the morning (8:00 AM). Eight-week-old female mice were used in this study.

Peripheral blood (600-800 μl) was collected by retro-orbital bleeding and diluted with 3 ml HBSS containing 15 mM EDTA. Cells were centrifuged for 10 min at 500 x g. Red blood cells were lysed by resuspension in 5 ml ACK (ammonium-chloride-potassium) lysis buffer (Thermo Fisher Scientific, Waltham, MA) for 5 min at RT. 10 ml RPMI + 2% FBS were added to stop lysis followed by centrifugation at 500 x g for 5 min. Cells were washed twice with 10 ml HBSS + 2 mM EDTA + 1%BSA before being re-suspended in 500 μl PBS + 1% BSA.

For bone marrow neutrophil isolation, whole bone marrow cells were flushed from the femur, tibia, and ilia leg bones with 5 ml HBSS + 2 mM EDTA + 1% BSA and filtered through a 70 μm cell strainer. Cells were centrifuged for 10 min at 500 x g. Red blood cells were lysed with 1 ml ACK lysis buffer for 2 min at room temperature (RT) and washed twice with HBSS + 2 mM EDTA + 1% BSA and re-suspended in 200 μl PBS + 1% BSA. c-kit-positive bone marrow cells were first enriched by positive selection using c-
kit (CD117) microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany) and further purified by FACS sorting c-kit-positive cells.

To isolate spleen neutrophils, spleens were dissected, placed in 3 ml PBS + 1mM EDTA, and then gently disaggregated through a 70 μm cell strainer using a 1 ml syringe plunger. Whole spleen cells were collected, centrifugated, and resuspended in 1 ml PBS + 1mM EDTA. The red blood cells were then lysed with 5 ml ACK lysis buffer for 2 min at room temperature. After centrifugation, cells were washed twice with PBS + 1 mM EDTA and re-suspended in 500 μl PBS + 1% BSA. Finally, peritoneal cavity exudate cells were harvested by three successive washes with 10 ml HBSS + 15 mM EDTA + 1% BSA. After centrifugation, cells were washed twice with the same solution and re-suspended in 100 μl PBS + 1% BSA.

Human sample collection
Peripheral blood (10 ml) was collected from healthy donors into heparin anticoagulant tubes. An equal volume (10 ml) of 6% hydroxyethyl solution was added into the heparinized blood and mixed gently several times for adequate mixing. The blood was kept at RT for 20-30 min before pipetting the supernatant into a 50 ml Falcon tube followed by centrifugation at 290 x g for 5 mins without braking. Cells were washed twice and lysed with ACK to completely remove red blood cells. Samples were stained with Percp-cy5.5-conjugated anti-human CD33 antibody for 20 min and DAPI was added to cells prior to FACS sorting.

Instrument
FACSaria III cell sorter (BD Biosciences, Franklin Lakes, NJ) and Attune Nxt Flow Cytometer (ThermoFisher).

Software
8D FACSDIVA

Cell population abundance
Abundance of the cell population in the sorted samples were indicated in Extended Fig.1b and Extended Fig.6 a-c.

Gating strategy
Intact cells were gated according to the FSC-A and SSC-A. Doublets were excluded by FSC-H and FSC-A. Dead cells were excluded based on DAPI staining. Gating strategies for specific cell populations were illustrated in Extended Data Fig.1a, Figure 3a, Figure 3d and Extended Data Fig.9a.

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