



OPEN

DATA DESCRIPTOR

Single-cell RNA-seq of primary bone marrow neutrophils from female and male adult mice

Minhoo Kim¹, Ryan J. Lu^{1,2} & Bérénice A. Benayoun^{1,3,4,5,6} ✉

Widespread sex-dimorphism is observed in the mammalian immune system. Consistently, studies have reported sex differences in the transcriptome of immune cells at the bulk level, including neutrophils. Neutrophils are the most abundant cell type in human blood, and they are key components of the innate immune system as they form a first line of defense against pathogens. Neutrophils are produced in the bone marrow, and differentiation and maturation produce distinct neutrophil subpopulations. Thus, single-cell resolution studies are crucial to decipher the biological significance of neutrophil heterogeneity. However, since neutrophils are very RNA-poor, single-cell profiling of these cells has been technically challenging. Here, we generated a single-cell RNA-seq dataset of primary neutrophils from adult female and male mouse bone marrow. After stringent quality control, we found that previously characterized neutrophil subpopulations can be detected in both sexes. Additionally, we confirmed that canonical sex-linked markers are differentially expressed between female and male cells across neutrophil subpopulations. This dataset provides a groundwork for comparative studies on the lifelong transcriptional sexual dimorphism of neutrophils.

Background & Summary

The mammalian immune system displays widespread sex dimorphism^{1–3}. In general, males are more susceptible to and have worse outcomes for severe infections, whereas females are more prone to autoimmune diseases^{3–6}. Consistently, transcriptome profiling studies have reported strong sex differences in the gene expression patterns of immune cells throughout life, including neutrophils^{7,8}. Neutrophils are key elements of the innate immune system that constitute the first line of defense in response to inflammatory stimuli⁹. Neutrophils protect the host by phagocytosis, production of antimicrobial granules, and release of neutrophil extracellular traps^{10–12}. Consequently, neutrophil dysfunction has been linked to the pathogenesis of various diseases, including atherosclerosis, macular degeneration, and cancer^{13–15}. Recently, our group performed a multi-omic study, including bulk RNA-seq analysis, of murine bone marrow neutrophils with respect to sex and age, and we observed significant sex differences in the neutrophil transcriptome and functional landscape throughout life⁸. Additionally, bulk RNA-seq analysis of murine spleen neutrophils by the ImmGen Consortium revealed transcriptional differences between female and male animals, and such sex differences were enhanced upon interferon stimulation⁷. Together, these findings suggest that neutrophils harbor sex dimorphic gene regulation.

Neutrophils are constantly produced in the bone marrow, and they migrate to the infected site through the circulatory system^{16,17}. Differentiation and maturation of neutrophils have been shown to produce distinct neutrophil subpopulations^{16,18,19}. Consequently, single-cell resolution approaches are essential to explore neutrophil heterogeneity and understand neutrophil biology. However, due to the low RNA and high RNase content of neutrophils, special precautions need to be taken to capture and robustly profile neutrophils in single-cell RNA-seq protocols^{20,21}. A recent landmark study by Xie *et al.* performed single-cell transcriptome profiling of neutrophils purified by flow cytometry from the bone marrow, peripheral blood and spleen of *female* mice²². The study identified eight distinct neutrophil subpopulations (G0–G4 and G5a–c) at different stages of maturation

¹Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA. ²Graduate Program in the Biology of Aging, University of Southern California, Los Angeles, CA, USA. ³Molecular and Computational Biology Department, USC Dornsife College of Letters, Arts and Sciences, Los Angeles, CA, USA. ⁴Biochemistry and Molecular Medicine Department, USC Keck School of Medicine, Los Angeles, CA, 90089, USA. ⁵USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA, USA. ⁶USC Stem Cell Initiative, Los Angeles, CA, USA. ✉e-mail: berenice.benayoun@usc.edu

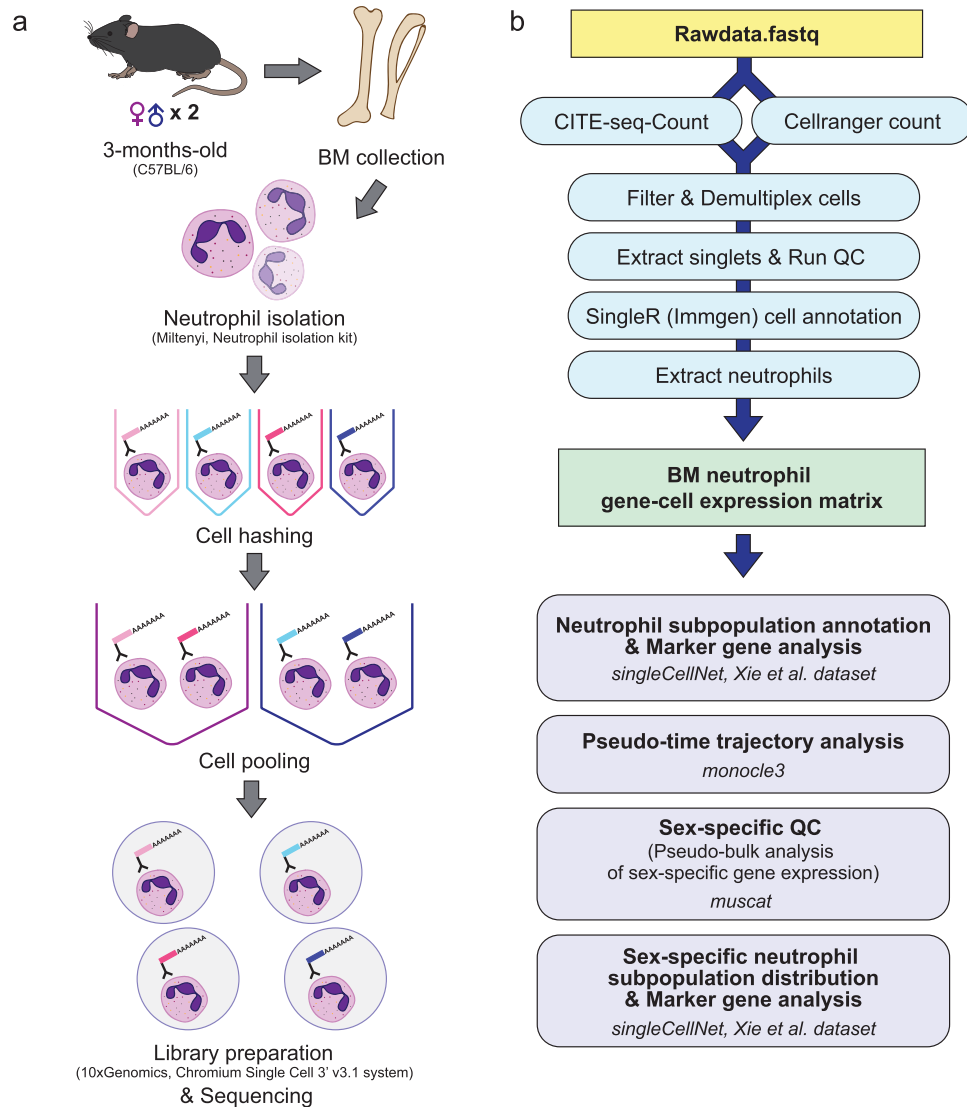


Fig. 1 Outline of sample preparation and data analysis workflow. **(a)** Outline of sample preparation workflow. Bone marrow was collected from two female and two male 3-month-old C57BL/6 mice. Isolated neutrophils were labeled with HTOs and pooled for sequencing library preparation and sequencing. **(b)** Outline of data analysis workflow. HTO counts and gene-barcode matrices were quantified using CITE-seq-Count³¹ and Cellranger count²⁹ functions, respectively. After demultiplexing, singlet neutrophils (annotated using SingleR³² and ImmGen³³ database) with gene count greater than 100 and mitochondrial gene count less than 25% were extracted to obtain a clean gene-cell expression matrix. Neutrophil subpopulation annotation and marker gene analyses were performed using singleCellNet³⁹ and the Xie *et al.* dataset³⁴. Sex-specific gene expression was quantified via pseudo-bulk analysis using muscat⁴¹. Pseudo-time trajectory analysis was performed using monocle3⁴⁰. HTO: Hash tag oligo. QC: Quality control.

in homeostatic state and proportions of each neutrophil subpopulation were different among bone marrow, peripheral blood and spleen²². In the bone marrow, relatively immature neutrophil subpopulations (G0-G4) were detected as more abundant, whereas more mature neutrophils (G5a-c) were found at higher frequency in the peripheral blood and spleen²². Additionally, Xie *et al.* leveraged the single-cell RNA-seq dataset to capture dynamic transitions between neutrophil subpopulations upon bacterial infection²². To gain insights on the regulatory networks and potential contribution of distinct neutrophil subpopulations to the observed sex differences in the neutrophil transcriptional landscape, single-cell resolution approaches are necessary. However, single-cell resolution datasets of neutrophil transcriptional landscapes that include both sexes are still lacking.

Here, we generated a single-cell RNA-seq dataset of primary neutrophils from 3-month-old female and male adult mouse bone marrow with biological replicates ($n = 2$ per sex) (Fig. 1a and Table 1). All the samples in our dataset were handled and processed together to eliminate potential batch effects, including the use of cell hashing²³. Single-cell multiplexing provides multiple benefits, including, but not limited to, mitigating batch effects and reducing costs²³. Through technical validation, as described below, we demonstrate technical quality of our dataset. After quality control, we annotated the neutrophil subpopulations by utilizing the annotation scheme

Sample	Sex	Age	HTO sequence
3m_F_1	Female	3-months-old	HTO1-ACCCACCAGTAAGAC
3m_M_1	Male		HTO2-GGTCGAGAGCATTCA
3m_F_2	Female		HTO3-CTTGCCGCATGTCAT
3m_M_2	Male		HTO4-AAAGCATTCTCACG

Table 1. Sample information.

from the study by Xie *et al.*²² - we identified comparable proportions of neutrophil subpopulations in our bone marrow neutrophil dataset as in the bone marrow neutrophil data from the original study. Additionally, we performed differential expression analysis of key marker genes for each neutrophil subpopulation between female and male bone marrow neutrophils. Overall, our dataset represents an important resource to investigate sex dimorphism in neutrophil biology.

Methods

Mouse husbandry. All animals were treated and housed in accordance with the Guide for Care and Use of Laboratory Animals. All experimental procedures were approved by the University of Southern California (USC)'s Institutional Animal Care and Use Committee (IACUC) and are in accordance with the institutional and national guidelines. Samples were derived from animals on approved IACUC protocol #20804. Adult female and male C57BL/6J mice (3-month-old animals) were obtained from Jackson Laboratories. Animals were acclimated at the animal facility at USC for 2 weeks before euthanasia. The animal facility at USC is on a 12-h light/dark cycle and animal housing rooms are maintained at 22 °C and 30% humidity. All animals were euthanized by CO₂ asphyxiation followed by cervical dislocation.

Isolation of primary neutrophils from the mouse bone marrow. Mouse bone marrow neutrophils were isolated using a previously described protocol that yields highly pure neutrophils^{8,24}. Specifically, hind limb long bones of each mouse were collected and kept on ice in D-PBS (Corning) supplemented with 1% antibiotic-antimycotic (Gibco) until further processing. Muscle and connective tissues were removed from the bones, and the bone marrow from the cleaned bones was collected into clean tubes by centrifugation²⁵. Red blood cells from the bone marrow were removed using the Red Blood Cell Lysis buffer (Miltenyi Biotec, no. 130-094-183), according to the manufacturer's instructions, albeit with no vortexing to avoid arbitrary neutrophil activation. The suspension was filtered on 70-µm mesh filters (Miltenyi Biotec, no. 130-110-916) to retain only single cells for downstream processing. Neutrophils were isolated from other bone marrow cells using MACS (Miltenyi Biotec, no. 130-097-658). Viability and yield were assessed using trypan blue exclusion and an automated COUNTESS cell counter (Thermo Fisher Scientific). Purified cells were immediately used for cell hashing.

Cell hashing, single-cell RNA-seq library preparation & sequencing. Cell hashing was performed using TotalSeqTM-A Antibodies (BioLegend) according to the manufacturer's instructions (TotalSeqTM-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit v3 or v3.1 (Single Index) Protocol) (Table 1). Specifically, 1 million purified neutrophils were first incubated with mouse Fc Blocking Reagent (Miltenyi Biotec, no. 130-092-575) at 4 °C for 10 minutes. Then, each blocked sample was incubated with 1 µg of specific TotalSeq Cell Hashing antibody (HTOs 1-4) at 4 °C for 30 minutes. After incubation, stained cells were washed 3 times, and assessed for viability and yield using the COUNTESS cell counter (Thermo Fisher Scientific). Female neutrophil samples (HTO 1, 3) and male neutrophil samples (HTO 2, 4) were pooled separately by sex of origin in equal volumes.

Single-cell RNA-seq libraries were generated using the Chromium Next GEM Single-cell 3' v3.1 assay (10xGenomics) according to the manufacturer's instructions (10xGenomics User Guide Chromium Next GEM Single-cell 3' Reagent Kits v3.1 (CG000204, Rev D))²⁶ with modifications to accommodate for low RNA content of neutrophils. Briefly, cell numbers equivalent to a target cell recovery of 5,000 cells per sample after sequencing were loaded onto a 10x Genomics Chromium Next GEM G Chip with the reverse transcription enzyme master mix. We used the Chromium Next GEM Single-cell 3' Reagent Kits v3.1 to perform mRNA capture, barcoding and reverse transcription within the GEMs. Single-cell RNA-seq libraries were generated using the Single-cell 3' Reagent Kit. To allow for library complexity despite low input RNA, an additional 2 cycles were added to the cDNA amplification step (for a total of 14 cycles). HTO Additive primer v2 (Integrated DNA Technologies) was added to the cDNA amplification reaction to amplify the HTO molecules. The amplified HTOs were recovered in the cleanup step for the cDNA amplification, and HTO libraries were built using the 2x NEBNext PCR Master Mix and Truseq indexing oligos (Integrated DNA Technologies).

Single-cell RNA-seq and HTO libraries were quantified and quality controlled using a Qubit fluorometer and the 4200 TapeStation system (Agilent Technologies) with a High Sensitivity DNA ScreenTape. Paired-end reads (26 + 8 + 98 bp) were generated on the Illumina NextSeq550 platform at the USC Genome Core.

Bioinformatic analysis. For sequencing data analysis, we followed the best-practice recommendations for single-cell RNA-seq analysis (Fig. 1b). Our pipeline was tested and validated in R²⁷ versions 3.6.3 and 4.1.2 and Seurat²⁸ versions 3.2.2 and 4.1.0. We report here the results using R 4.1.2 and Seurat 4.1.0. Each step of our bioinformatic analysis pipeline is discussed in detail below.

Sample	Reads	Q30 Bases in Barcode	Q30 Bases in RNA Read	Q30 Bases in UMI	Confident Mapping to Genome	Confident Mapping to Transcriptome
Female Pool	199,687,860	96.9%	90.6%	95.8%	94.7%	87.3%
Male Pool	215,735,119	97.1%	90.9%	96.0%	95.0%	88.0%

Table 2. Detailed QC report of 10x Genomics sequencing files (Cell Ranger).

Sample	Estimated Number of Cells	Mean Reads per Cell	Median Genes per Cell	Median UMI per Cell	Fraction Reads in Cells	Sequencing Saturation
Female Pool	2,848	70,115	1,691	5,712	96.8%	87.9%
Male Pool	3,630	59,431	1,544	4,755	97.3%	87.7%

Table 3. Sequencing statistics of 10x Genomics libraries (Cell Ranger).

Sample	Reads	Mean Quality Score	Matched Hashtags
Female HTO	1,722,815	27.2	79%
Male HTO	2,042,585	27.4	82%

Table 4. QC and information for HTO libraries.

Sample	QC Neutrophils	Median Genes per Neutrophil	Median UMI per Neutrophil	Median HTO per Neutrophil	Mean Mitochondrial Reads
3m_F_1	1,396	1,662	5,552	331	<0.5%
3m_F_2	1,205	1,737	5,980	208	<0.5%
3m_M_1	1,802	1,541	4,716	273	<0.5%
3m_M_2	1,622	1,540	4,765	260	<0.5%

Table 5. Final per sample information after HTO demultiplexing, singlet and neutrophil filtering.

Filter and demultiplex cells. Raw reads were aligned to the mouse genome (mm10) and processed using the “cellranger count” (6.0.2) pipeline²⁹ (`--expect-cells = 5000 --include-introns`) to generate gene-barcode matrices (Tables 2 and 3). We used the `--include-introns` option to account for the limited number of genes expressed in neutrophils³⁰, as recommended by 10xGenomics. HTO quantification was performed using CITE-seq-Count (v. 1.4.5)³¹ with the following parameters: `-cbf 1 -cbl 16 -umif 17 -umil 26 -cells 5000` (Table 4). Gene-barcode matrices and HTO counts were loaded into R and cells that were detected by both RNA and HTO were filtered for downstream analysis. Additionally, genes that were not detected in at least 20 cells were excluded to eliminate possible random noise. After setting up a Seurat²⁸ object using the gene-barcode matrices, HTO counts were added to the Seurat object as an independent assay and normalized using the centered log-ratio (CLR) transformation. Cells were demultiplexed using the `MULTIseqDemux()` function into four samples according to their HTO barcode of origin.

Extract singlets and run QC. Singlets were retained for downstream analyses based on the `MULTIseqDemux()` annotations. For quality control, cells that have unique feature counts less than 100 or mitochondrial gene count greater than 25% were excluded (`subset = nFeature_RNA > 100 & percent.mito < 25`) (Table 5). After removing the unwanted cells, the dataset was normalized using the `NormalizeData()` function with the following parameters: `normalization.method = “LogNormalize”, scale.factor = 10000`.

Dimensional reduction and clustering. Prior to dimensional reduction, the dataset was scaled using the `SCTransform()` function with `vars.to.regress = c(“nFeature_RNA”, “percent.mito”)`. The first 15 principal components (PCs) were used for Uniform Manifold Approximation and Projection (UMAP) (`RunUMAP()`) and clustering (`FindNeighbors()` and `FindClusters()`, `resolution = 0.3`).

SingleR cell annotation. To eliminate possible contaminants (non-neutrophil cells) from the dataset, quality controlled singlets were annotated using SingleR v.1.8.1³². The ImmGen expression dataset from `celldex` was used as the reference for cell annotation³³. 6,025 cells out of 6,073 cells from the dataset were annotated as neutrophils (~99.21% purity). Cells annotated as non-neutrophil cells were excluded for downstream analysis.

Neutrophil subpopulation annotation and marker gene analysis. Neutrophil heterogeneity has been reported by multiple studies^{16,18,19}. A recent work by Xie *et al.* characterized neutrophil subpopulations (G0-G4 and G5a-c) in the murine bone marrow via single-cell RNA-seq analysis of flow cytometry²² (hereafter referred to as the “Xie *et al.* dataset”³⁴). To annotate the neutrophil subpopulations in our dataset, we leveraged the gene-barcode matrix

Library	BioProject	BioSample
Female Neutrophils	PRJNA796634 ⁴²	SAMN24905300 ⁴³
Male Neutrophils	PRJNA796634 ⁴²	SAMN24905301 ⁴⁴
Female HTO	PRJNA796634 ⁴²	SAMN24905302 ⁴⁵
Male HTO	PRJNA796634 ⁴²	SAMN24905303 ⁴⁶

Table 6. Raw sequencing data accession.

(GSE137539³⁴) and cell annotation data (GSM4081545³⁵, GSM4081546³⁶, GSM4081547³⁷ and GSM4081548³⁸) from the Xie *et al.* dataset provided in the Gene Expression Omnibus (GEO) database. singleCellNet v.0.1.0 was used to classify neutrophil subpopulations within our dataset³⁹. After extracting the common genes found in both our dataset and the Xie *et al.* dataset, we trained a classifier using the `scn_train()` function with the following parameters: $nTopGenes = 100$, $nTopGenePairs = 50$, $nRand = 50$, $nTrees = 1000$. `scn_predict()` and `get_cate()` functions were used to classify cells within our dataset and assign annotation to each cell, respectively. Marker genes for neutrophil subpopulations identified from the Xie *et al.* dataset (17 genes) were used to analyze marker gene expression in our dataset²². Dot plots of marker gene expression levels were generated using the `DotPlot()` function from Seurat²⁸.

Pseudo-time trajectory analysis. Single-cell trajectory was constructed using monocle3 v.1.0.0⁴⁰. After learning the principal graph (`learn_graph()`), cells were ordered using `order_cells(root_cells = "G2")`. "G2" subpopulation was used as root cells as they were the least matured neutrophil subpopulation detected in our dataset²².

Pseudo-bulk analysis of sex-specific gene expression. As a quality control, we assessed the expression of sex-specific genes, *Xist* and *Ddx3y*, in each neutrophil subpopulation using pseudo-bulk analysis with muscat v.1.5.2⁴¹. Each neutrophil subpopulation single-cell data was aggregated to pseudo-bulk data using the `aggregateData()` function with the `fun = "sum"` option. Differential state was assessed using the `pbDS()` function with the following parameters: `method = "edgeR"`, `min_cells = 8` (edgeR v. 3.36.0).

Dimensional reduction of pseudo-bulk analysis data. To assess the potential global sex differences in the transcriptional landscapes between female and male neutrophils (globally and for each subpopulation), muscat v.1.5.2⁴¹ was used to perform Multidimensional Scaling (MDS, `pbMDS()`) of the aggregated pseudo-bulk data (output of `aggregateData()`, described above).

Data Records

Sequencing data have been submitted to the Sequence Read Archive accessible through BioProject PRJNA796634⁴² (BioSamples SAMN24905300⁴³, SAMN24905301⁴⁴, SAMN24905302⁴⁵ and SAMN24905303⁴⁶; Table 6). We used publicly available neutrophil single-cell RNA-seq annotation data from the GEO (GSE137539³⁴, 8-to-10-week-old female mouse bone marrow neutrophil samples) to annotate neutrophil subpopulations within our dataset. The final annotated Seurat object⁴⁷ has been made available on Figshare for use and analysis (<https://doi.org/10.6084/m9.figshare.19623978>).

Technical Validation

Quality control of the single-cell RNA-seq dataset. We utilized cell hashing²³ to multiplex our neutrophil single-cell RNA-seq samples. Biological replicates ($n = 2$ animals per sex) of each sex were pooled separately for sequencing library preparation (Fig. 1a). Hash tag oligo (HTO) quantification was performed using CITE-seq-Count³¹ (Fig. 1b). As shown in Fig. 2a, we confirmed clear enrichment of each HTO (HTO1, HTO2, HTO3 and HTO4) after demultiplexing using the `MULTIseqDemux()` function from Seurat²⁸, indicating successful cell hashing of neutrophils. Additionally, comparable Unique Molecular Identifier (UMI) counts were observed among four samples prior to quality control (Fig. 2b). To exclude potential low-quality data and noise, doublets and negative cells and cells with gene count less than 100 or mitochondrial gene count greater than 25% were excluded from the dataset (Fig. 2c). Additionally, to eliminate non-neutrophil cells from the dataset, we used SingleR³² and the ImmGen³³ database to annotate the cells (Fig. 2d). Cells annotated as neutrophils (6,025 cells out of 6,073 cells from the post-filter dataset; ~99.21%) were extracted for downstream analysis (Table 5). We used the filtered neutrophil singlets to perform dimensional reduction via Uniform Manifold Approximation and Projection (UMAP) using the first 15 principal components (Fig. 3a). We did not find any noticeable sample origin-related clustering and observed a homogeneous distribution of cells derived from female and male animals (Fig. 3a).

Annotation of neutrophil subpopulations. Recently, Xie *et al.* identified eight distinct neutrophil subpopulations (G0-4 and G5a-c) through a single-cell RNA-seq analysis of the bone marrow, peripheral blood, and spleen neutrophils from female mice²². In the same study, each subpopulation was identified as the following neutrophil subsets in different maturation states: G0 – granulocyte-monocyte progenitors, G1 – committed neutrophil progenitors, G2 – pre-neutrophils, G3 – immature neutrophils and G4, G5a-c – mature neutrophils²². To identify and annotate neutrophil subpopulations in our dataset, we used singleCellNet³⁹ and the annotation information from the Xie *et al.* dataset. As shown in Fig. 3b, we identified six neutrophil subpopulations, G2-4 and G5a-c, within our dataset. Importantly, G0 and G1 subpopulations, granulocyte-monocyte progenitors and committed neutrophil progenitors, respectively, were not detected in our dataset, most likely due to depletion

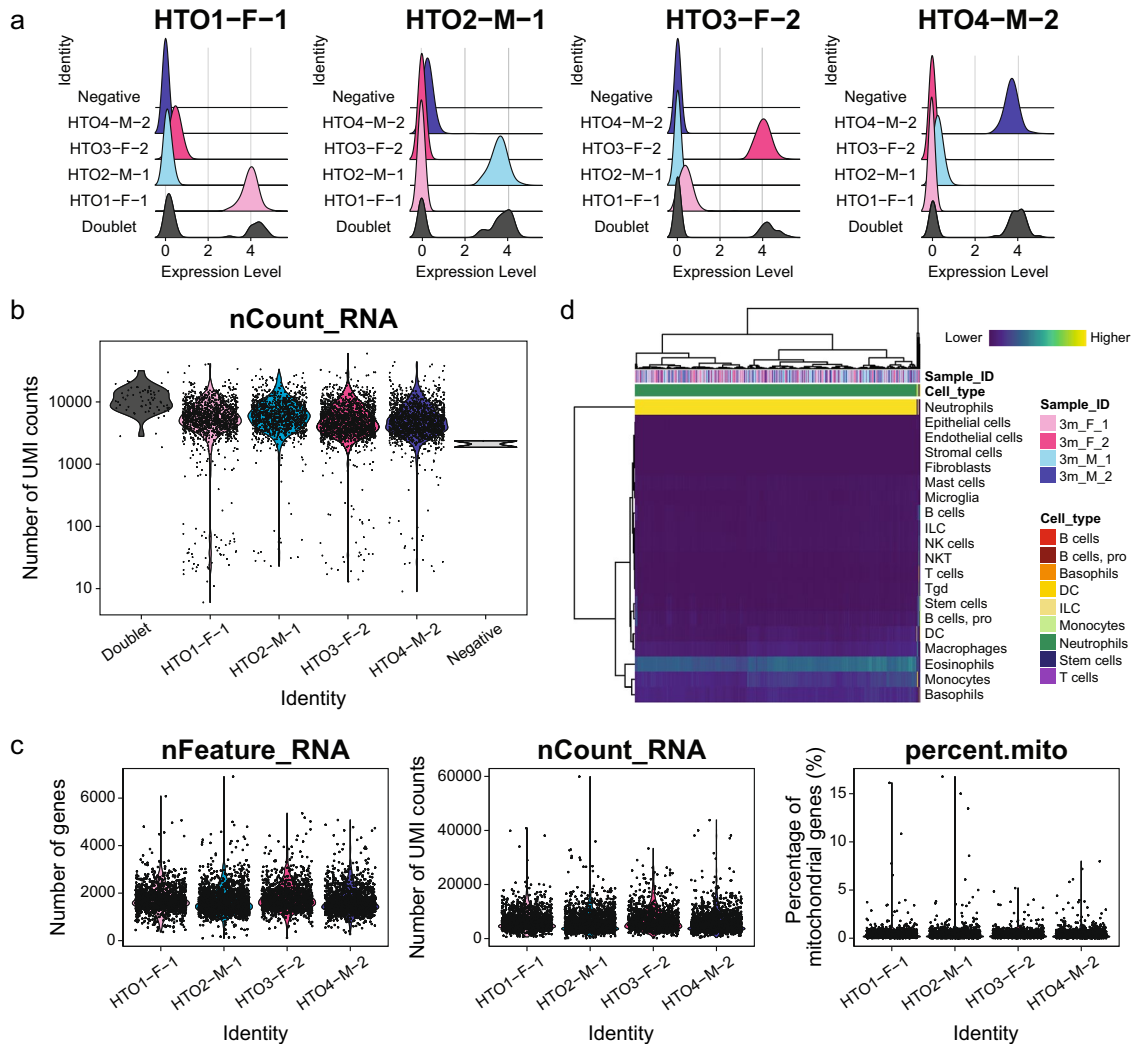


Fig. 2 Single-cell RNA-seq dataset quality assessment. **(a)** Ridgeplots of HTO sample enrichment after demultiplexing. **(b)** Violin plot of UMI counts (nCount_RNA) from each sample after demultiplexing. **(c)** Violin plots of gene counts (nFeature_RNA, left panel), UMI counts (nCount_RNA, middle panel) and percentage of mitochondrial gene counts (percent.mito, right panel) after quality control filtering. **(d)** Heatmap of cell annotation scores and cell annotation via SingleR³² and ImmGen³³ database. HTO: Hash tag oligo. UMI: Unique molecular identifier.

of progenitors by negative selection during the MACS neutrophil isolation step of our workflow⁸. To note, G0 and G1 subpopulations were detected only in small proportions in the bone marrow data of the Xie *et al.* dataset²². We also assessed the expression levels of marker genes associated with each neutrophil subpopulation. In the Xie *et al.* dataset, 18 marker genes were identified for subpopulations G2-4 and G5a-c. We observed similar relative expression levels of the marker genes among the neutrophil subpopulations in our dataset as was shown for the Xie *et al.* dataset, except for *Gm5483*, which was not detected in our dataset, and the expression of marker genes with respect to each neutrophil subpopulation was not grossly affected by the biological sex (Fig. 3c and Supplementary Figure 1a). Additionally, we used monocle3⁴⁰ to construct a single-cell trajectory along pseudo-time to validate the temporal relationships among the neutrophil subpopulations along the maturation process. As shown in Fig. 3d, we confirmed neutrophil maturation along the constructed trajectory in pseudo-time, starting from G2 through G5a-c, as was described for the Xie *et al.* dataset²².

Expression analysis of sex-linked genes. To quality check for sex-specificity of our dataset, we assessed the expression of sex-chromosome linked genes, i.e. female-specific *Xist* and male-specific *Ddx3y* (Fig. 4a,b). Through sample-wise single-cell expression analysis and subpopulation-wise pseudo-bulk analysis of *Xist* and *Ddx3y* expression, we confirmed the expression of sex-specific genes in their respective samples (Fig. 4a,b). Additionally, our sex-wise comparisons confirmed that all six neutrophil subpopulations (G2-4 and G5a-c) were represented across the female and male samples in similar proportions (Fig. 4c).

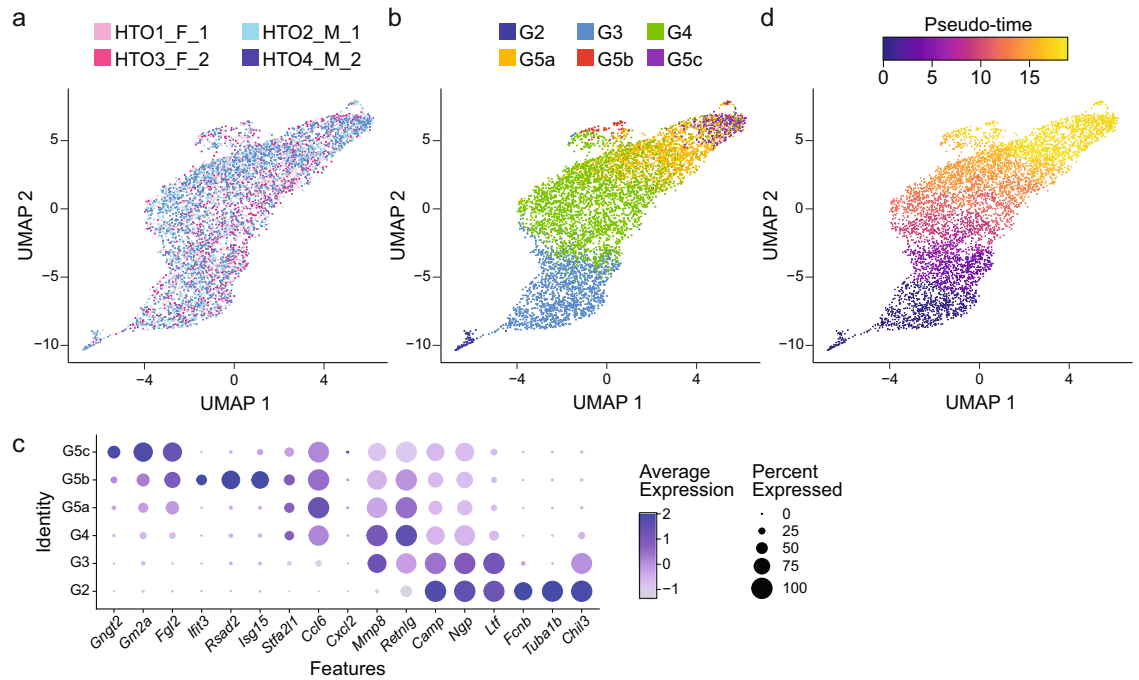


Fig. 3 Neutrophil subpopulation annotation and marker gene analysis. (a), (b) and (d) Two-dimensional cell clustering via UMAP using the first 15 principal components. Cells are labelled by (a) HTO labels, (b) neutrophil subpopulations and (d) monocle3⁴⁰ pseudo-time trajectory scores. (c) Dot plot of scaled expression levels of neutrophil subpopulation marker genes derived from the Xie *et al.* dataset³⁴. UMAP: Uniform Manifold Approximation and Projection.

Pseudo-bulk-level analysis of sex-dimorphic neutrophil transcriptomes. More generally, to determine whether the transcriptomes of neutrophils and the underlying neutrophil subpopulations displayed sex-dimorphic expression patterns, we further analyzed the pseudo-bulk expression data. Using Multi Dimensional Scaling (MDS) as a dimensional reduction approach, we observed clear distinction between female and male samples at the global level (Fig. 4d). Further, when we compared the transcriptomes of each neutrophil subpopulation, we found similar clustering of female vs. male samples in all subpopulations, except for G5b (Fig. 4e and Supplementary Figure 1b), consistent with the notion that sex-dimorphic gene regulation occurs across neutrophil subpopulations. This observation suggests that the majority of neutrophil subpopulations present with sex-dimorphic transcriptional landscapes.

Together, through the technical validations described above, we show that our single-cell RNA-seq dataset of bone marrow neutrophils from adult female and male mice is of high-quality. Additionally, we observed distinct neutrophil transcriptional landscapes between females and males. This dataset will serve as a valuable resource to interrogate sex differences in neutrophil landscapes at the single-cell level and to investigate the regulatory mechanisms underlying sex-dimorphism in neutrophil biology.

Usage Notes

Here, we generated a single-cell RNA-seq dataset of murine primary bone marrow neutrophils from adult female and male animals. From our neutrophil subpopulation marker gene expression analysis, we observed comparable expression levels of the marker genes among female and male samples (Supplementary Figure 1a). Thus, cell identities of neutrophil subpopulations seem to be preserved between the two sexes. On the other hand, our pseudo-bulk-level analysis revealed distinct transcriptional landscapes of neutrophils between female and male subjects (Fig. 4d), which is consistent with previous studies that showed sex-dimorphic transcriptomes of neutrophils via bulk RNA-seq analyses^{7,8}. In addition, we also provide evidence that sex-dimorphic gene expression also occurs in different neutrophil subpopulations (Fig. 4e and Supplementary Figure 1b), and is likely to not stem solely from sex-dimorphism in underlying patterns of heterogeneity. A potential limitation of this dataset includes the relatively small sample size ($n = 2$ per sex), which only allows for the discovery of large sex differences. Thus, future studies with larger numbers of animals will be needed for the detection of more subtle sex-dimorphic effects. Overall, the dataset described here will be an invaluable resource to start assessing the sex differences in the gene expression profiles across neutrophil subpopulations. This dataset may also be leveraged to identify differences in RNA velocity and/or transcription factor regulon activity as a function of sex and stage of neutrophil maturation (for example, using SCENIC⁴⁸). Importantly, this dataset represents data from sexually mature C56BL/6J animals in good health and should be treated as such when inferring potential biological

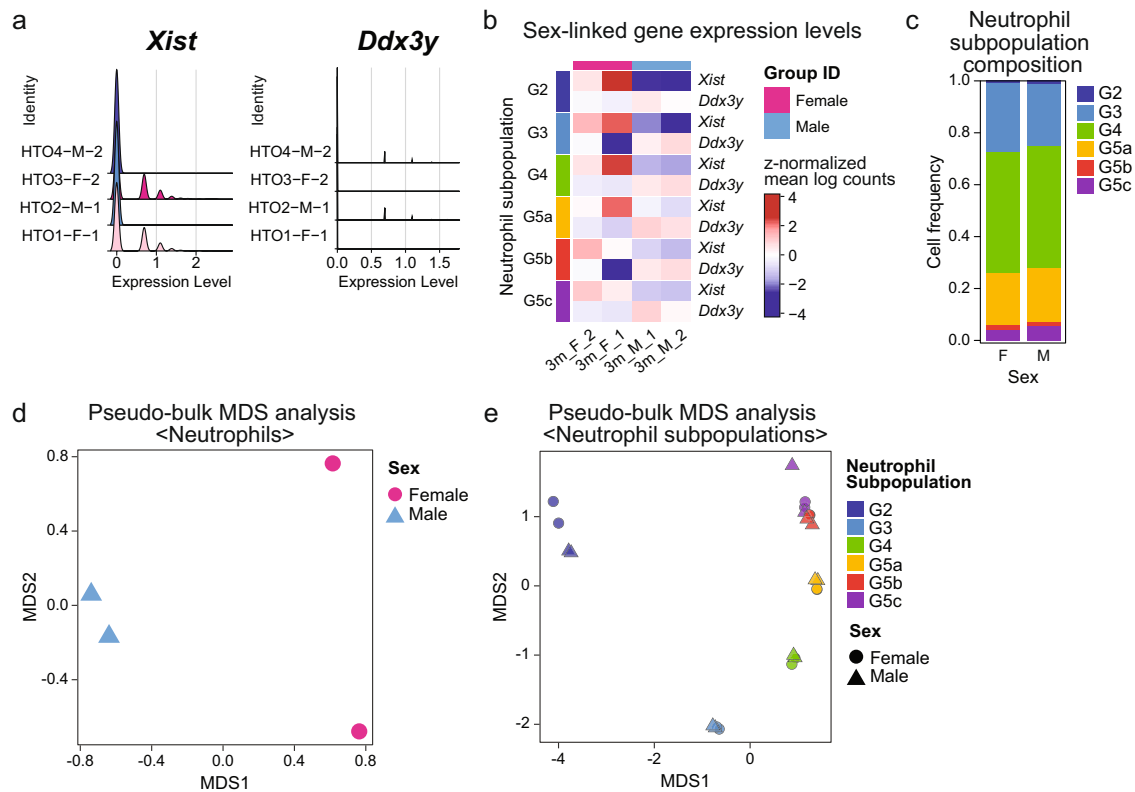


Fig. 4 Sex-specific gene expression and neutrophil subpopulation distribution analysis. **(a)** Ridge plots of expression levels of *Xist* (female-specific) and *Ddx3y* (male-specific). **(b)** Heatmap of expression levels of sex-specific genes in each neutrophil subpopulations quantified by muscat⁴¹ pseudo-bulk analysis. **(c)** Percent stacked barplot of neutrophil subpopulation distribution in female and male samples. **(d)** and **(e)** MDS plots of female vs. male neutrophil pseudo-bulk data for all neutrophils **(d)** and separated by neutrophil subpopulations **(e)**. F: Female. M: Male. MDS: Multidimensional scaling.

responses as a function of genetic background, in response to infection or with respect to animal organismal age. Future work including profiling of neutrophils from other mouse strains (both inbred and outbred), other ages and/or health states will be important to determine the conservation of sex-dimorphic transcriptional patterns of neutrophils across biological contexts.

Code availability

All analytical code used for processing and technical validation is available on the Benayoun Laboratory GitHub repository (https://github.com/BenayounLaboratory/Neutrophil_scRNAseq_2022). The provided R code was run and tested using R 4.1.2²⁷.

Received: 25 April 2022; Accepted: 8 July 2022;

Published online: 23 July 2022

References

- Clocchiatti, A., Cora, E., Zhang, Y. & Dotto, G. P. Sexual dimorphism in cancer. *Nat Rev Cancer* **16**, 330–339, <https://doi.org/10.1038/nrc.2016.30> (2016).
- Klein, S. L., Jedlicka, A. & Pekosz, A. The Xs and Y of immune responses to viral vaccines. *Lancet Infect Dis* **10**, 338–349, [https://doi.org/10.1016/S1473-3099\(10\)70049-9](https://doi.org/10.1016/S1473-3099(10)70049-9) (2010).
- Whitacre, C. C., Reingold, S. C. & O’Looney, P. A. A gender gap in autoimmunity. *Science* **283**, 1277–1278, <https://doi.org/10.1126/science.283.5406.1277> (1999).
- Guerra-Silveira, F. & Abad-Franch, F. Sex bias in infectious disease epidemiology: patterns and processes. *PLoS One* **8**, e62390, <https://doi.org/10.1371/journal.pone.0062390> (2013).
- McClelland, E. E. & Smith, J. M. Gender specific differences in the immune response to infection. *Arch Immunol Ther Exp (Warsz)* **59**, 203–213, <https://doi.org/10.1007/s00005-011-0124-3> (2011).
- Beeson, P. B. Age and sex associations of 40 autoimmune diseases. *Am J Med* **96**, 457–462, [https://doi.org/10.1016/0002-9343\(94\)90173-2](https://doi.org/10.1016/0002-9343(94)90173-2) (1994).
- Gal-Oz, S. T. et al. ImmGen report: sexual dimorphism in the immune system transcriptome. *Nat Commun* **10**, 4295, <https://doi.org/10.1038/s41467-019-12348-6> (2019).
- Lu, R. J. et al. Multi-omic profiling of primary mouse neutrophils predicts a pattern of sex and age-related functional regulation. *Nat Aging* **1**, 715–733, <https://doi.org/10.1038/s43587-021-00086-8> (2021).
- Furze, R. C. & Rankin, S. M. Neutrophil mobilization and clearance in the bone marrow. *Immunology* **125**, 281–288, <https://doi.org/10.1111/j.1365-2567.2008.02950.x> (2008).

10. Lee, W. L., Harrison, R. E. & Grinstein, S. Phagocytosis by neutrophils. *Microbes Infect* **5**, 1299–1306, <https://doi.org/10.1016/j.micinf.2003.09.014> (2003).
11. Lacy, P. Mechanisms of degranulation in neutrophils. *Allergy Asthma Clin Immunol* **2**, 98–108, <https://doi.org/10.1186/1710-1492-2-3-98> (2006).
12. Sollberger, G., Tilley, D. O. & Zychlinsky, A. Neutrophil Extracellular Traps: The Biology of Chromatin Externalization. *Dev Cell* **44**, 542–553, <https://doi.org/10.1016/j.devcel.2018.01.019> (2018).
13. Ionita, M. G. *et al.* High neutrophil numbers in human carotid atherosclerotic plaques are associated with characteristics of rupture-prone lesions. *Arterioscler Thromb Vasc Biol* **30**, 1842–1848, <https://doi.org/10.1161/ATVBAHA.110.209296> (2010).
14. Ghosh, S. *et al.* Neutrophils homing into the retina trigger pathology in early age-related macular degeneration. *Commun Biol* **2**, 348, <https://doi.org/10.1038/s42003-019-0588-y> (2019).
15. Treffers, L. W., Hiemstra, I. H., Kuijpers, T. W., van den Berg, T. K. & Matlung, H. L. Neutrophils in cancer. *Immunol Rev* **273**, 312–328, <https://doi.org/10.1111/imr.12444> (2016).
16. Nicolas-Avila, J. A., Adrover, J. M. & Hidalgo, A. Neutrophils in Homeostasis, Immunity, and Cancer. *Immunity* **46**, 15–28, <https://doi.org/10.1016/j.immuni.2016.12.012> (2017).
17. Nauseef, W. M. & Borregaard, N. Neutrophils at work. *Nat Immunol* **15**, 602–611, <https://doi.org/10.1038/ni.2921> (2014).
18. Silvestre-Roig, C., Hidalgo, A. & Soehnlein, O. Neutrophil heterogeneity: implications for homeostasis and pathogenesis. *Blood* **127**, 2173–2181, <https://doi.org/10.1182/blood-2016-01-688887> (2016).
19. Ng, L. G., Ostuni, R. & Hidalgo, A. Heterogeneity of neutrophils. *Nat Rev Immunol* **19**, 255–265, <https://doi.org/10.1038/s41577-019-0141-8> (2019).
20. Can I process neutrophils (or other granulocytes) using 10x Single Cell applications? (2022).
21. Neutrophil Analysis in 10x Genomics Single Cell Gene Expression Assays (2021).
22. Xie, X. *et al.* Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. *Nat Immunol* **21**, 1119–1133, <https://doi.org/10.1038/s41590-020-0736-z> (2020).
23. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* **19**, 224, <https://doi.org/10.1186/s13059-018-1603-1> (2018).
24. McGill, C. J., Lu, R. J. & Benayoun, B. A. Protocol for analysis of mouse neutrophil NETosis by flow cytometry. *STAR Protoc* **2**, 100948, <https://doi.org/10.1016/j.xpro.2021.100948> (2021).
25. Amend, S. R., Valkenburg, K. C. & Pienta, K. J. Murine Hind Limb Long Bone Dissection and Bone Marrow Isolation. *J Vis Exp*, <https://doi.org/10.3791/53936> (2016).
26. Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry) (2019).
27. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2021).
28. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529, <https://doi.org/10.1016/j.cell.2021.04.048> (2021).
29. Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 14049, <https://doi.org/10.1038/ncomms14049> (2017).
30. Hay, S. B., Ferchen, K., Chetal, K., Grimes, H. L. & Salomonis, N. The Human Cell Atlas bone marrow single-cell interactive web portal. *Exp Hematol* **68**, 51–61, <https://doi.org/10.1016/j.exphem.2018.09.004> (2018).
31. Roelli, P., Flynn, B. & Gui, G. Hoohm/CITE-seq-Count: 1.4.2. *Zenodo* <https://doi.org/10.5281/zenodo.2590196> (2019).
32. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**, 163–172, <https://doi.org/10.1038/s41590-018-0276-y> (2019).
33. Heng, T. S. & Painter, M. W., Immunological Genome Project, C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091–1094, <https://doi.org/10.1038/ni1008-1091> (2008).
34. *Gene Expression Omnibus at NCBI*. <https://identifiers.org/geo:GSE137539> (2020).
35. *Gene Expression Omnibus at NCBI*. <https://identifiers.org/geo:GSM4081545> (2020).
36. *Gene Expression Omnibus at NCBI*. <https://identifiers.org/geo:GSM4081546> (2020).
37. *Gene Expression Omnibus at NCBI*. <https://identifiers.org/geo:GSM4081547> (2020).
38. *Gene Expression Omnibus at NCBI*. <https://identifiers.org/geo:GSM4081548> (2020).
39. Tan, Y. & Cahan, P. SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst* **9**, 207–213 e202, <https://doi.org/10.1016/j.cels.2019.06.004> (2019).
40. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386, <https://doi.org/10.1038/nbt.2859> (2014).
41. Crowell, H. L. *et al.* muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat Commun* **11**, 6077, <https://doi.org/10.1038/s41467-020-19894-4> (2020).
42. *BioProject at NCBI*. <https://identifiers.org/bioproject:PRJNA796634> (2022).
43. *BioSample at NCBI*. <http://www.ncbi.nlm.nih.gov/biosample?term={SAMN24905300}> (2022).
44. *BioSample at NCBI*. <http://www.ncbi.nlm.nih.gov/biosample?term={SAMN24905301}> (2022).
45. *BioSample at NCBI*. <http://www.ncbi.nlm.nih.gov/biosample?term={SAMN24905302}> (2022).
46. *BioSample at NCBI*. <http://www.ncbi.nlm.nih.gov/biosample?term={SAMN24905303}> (2022).
47. Kim, M., Lu, R. J. & Benayoun, B. A. Annotated Seurat object of scRNA-seq from female and male murine neutrophils. *figshare* <https://doi.org/10.6084/m9.figshare.19623978> (2022).
48. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083–1086, <https://doi.org/10.1038/nmeth.4463> (2017).

Acknowledgements

We would like to thank Suchi Patel at the USC Genome Core for her help with sequencing of the transcriptomic libraries on the Illumina NextSeq550 platform. We would also like to thank Cassandra McGill, Juan Bravo and Dr. Bryan Teefy for providing helpful feedback on the manuscript. This work was supported by GCRLE-2020 post-doctoral fellowship from the Global Consortium for Reproductive Longevity and Equality at the Buck Institute, made possible by the Bia-Echo Foundation to M.K., NIA T32 AG052374 predoctoral fellowship and a Diana Jacobs Kalman/AFAR Scholarships for Research in the Biology of Aging to R.J.L., Pew Biomedical Scholar award #00034120, an innovator grant from the Rose Hills foundation, and the Kathleen Gilmore Biology of Aging research award to B.A.B.

Author contributions

Conceptualization, R.J.L. and B.A.B.; neutrophil isolation, R.J.L. and B.A.B.; 10x Genomics and HTO library preparation, B.A.B.; data analysis M.K. and B.A.B.; writing - original draft, M.K.; writing - review and editing, M.K., R.J.L. and B.A.B.; funding acquisition, B.A.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-022-01544-7>.

Correspondence and requests for materials should be addressed to B.A.B.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022