Title: Dysregulated granulopoiesis drives a sepsis endotype of immune suppression and high mortality

Authors: Andrew J. Kwok[1], Alice Allcock[1], Ricardo C. Ferreira[1], Madeleine Smee[1], Katie L. Burnham[3], Eddie Cano-Gamez[1,3], Yasemin-Xiomara Zurke[4], Oxford acute medicine/ED research[5], Stuart McKechnie[5], Alex Mentzer[1,2,5], Claudia Monaco[4], Irina Udalova[4], Charles J. Hinds[6], John A. Todd[1,2], Emma E. Davenport[3], Julian C. Knight[1,2,5*]

Affiliations: [1]Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [2]NIHR Oxford Biomedical Research Centre, Oxford, UK. [3]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK. [4]Kennedy Institute of Rheumatology, University of Oxford, Oxford, UK. [5]John Radcliffe Hospital, Oxford Universities Hospitals NHS Foundation Trust, UK. [6] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University, London, UK.

[*]Corresponding author. Email julian.knight@well.ox.ac.uk

Abstract: The maladaptive response to infection leading to the clinical syndrome of sepsis is highly heterogeneous, confounding immunotherapy trials. Here we establish the pathophysiology of a sepsis response endotype informative for immune state, disease severity, outcome and therapy. We applied whole blood single-cell multiomics (272,993 cells, n=39 individuals) and found patients with this endotype had increased *IL1R2*+ immature neutrophils, confirmed using CyTOF (n=41 samples) and RNA-sequencing (n=677 individuals), and upregulated STAT3 programs with plasma cytokine profiles of emergency granulopoiesis (EG). We characterised sepsis

hematopoietic stem cells (29,336 cells, n=27 samples) with multiomics, and demonstrated endotype-specific EG transcriptional skewing together with STAT3 and EG master regulator CEBPB epigenetic signatures. Our findings establish an immunohematopoietic basis for a deleterious sepsis endotype and nominate IL-6/G-CSF as potential therapeutic targets.

Single sentence summary: We define the pathophysiology of a poor outcome sepsis endotype, stratifying the maladaptive response for targeted therapy.

Main text: Addressing the widespread failure of clinical trials in sepsis requires better understanding and more effective therapeutic targeting of the highly heterogenous, maladaptive host response causing disease (*1*). Bulk transcriptomics and single-cell interrogation of peripheral blood mononuclear cells (PBMCs) have demonstrated maladaptive inflammation and immunosuppression involving myeloid and lymphoid cells, but granulocytes remain relatively under-characterised (*2-6*). Sepsis subphenotypes are recognised (*7*), and we previously described whole blood transcriptomic sepsis response signatures (SRS) associated with outcome and treatment response (*2, 8-10*). Here, we delineate specific neutrophil sub-populations in sepsis and provide evidence for heightened emergency granulopoiesis (EG) driving the poor outcome SRS1 endotype.

We first generated an unbiased single-cell atlas of peripheral blood leukocytes in sepsis. We profiled the transcriptome and expression of 30 cell-surface proteins in fresh whole blood for 26 sepsis patients and 13 controls (healthy controls [HC] and post-cardiac surgery sterile inflammation [CS]), performing broad and fine level annotations for 272,993 cells (Methods; Fig. 1A,B; Fig. S1,2; Tables S1-3). We identified multiple neutrophil clusters, including immature

subsets (Methods; Fig. 1B; Fig. S3A-E), and found that differentiation trajectories calculated using RNA velocity and partition-based graph abstraction followed the expected maturation sequence across subpopulations (Fig. S3F-H). Sampling neighborhoods of cells (Methods), we found proportionally more degranulating and *S100A8/9* high neutrophils in sepsis compared to HC, while all mononuclear cell (MNC) subsets except plasmablasts were reduced (Fig. 1C,D). Comparable differences were seen comparing CS against HC, suggesting these were non-specific features of inflammation (Fig. S4A,B). By contrast, higher abundance of the immature neutrophil populations was specific to sepsis (Fig. 1C-E; Fig. S4C).

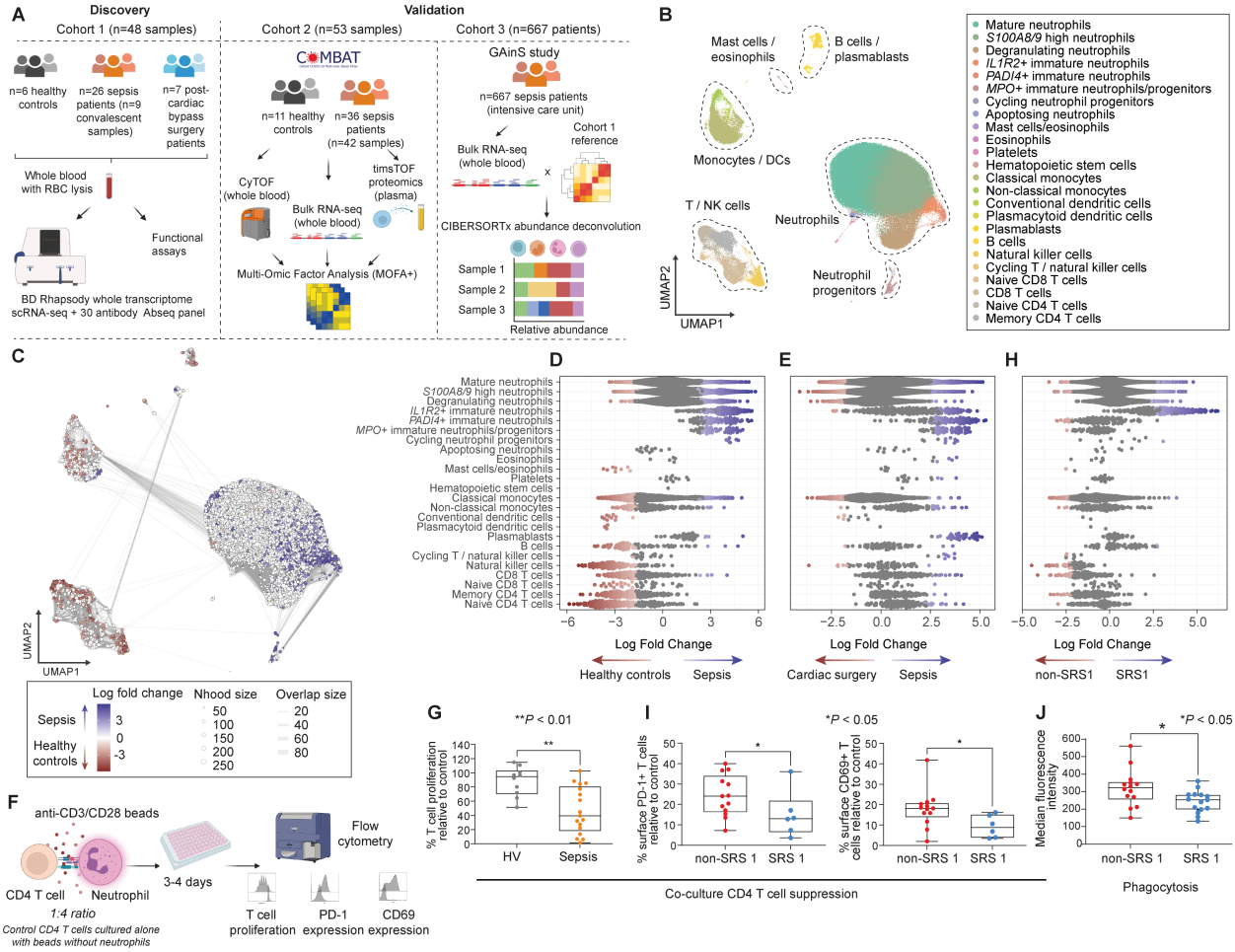

Figure 1. Whole blood single-cell census and neutrophil function in sepsis. (**A**) Cohort design. (**B-J**) Cohort 1. (**B**,**C**) UMAPs of (**B**) whole blood scRNA-seq/protein (272,993 cells) (**C**)

differential cell abundance for sepsis vs. HC, with sampled neighborhoods colored by statistical significance (spatial FDR < 0.05). (**D**,**E**) Beeswarm plot of differential cell abundance for sepsis vs. (**D**) HC (**E**) CS, with cluster labels of neighborhoods depicted. (**F**) Schematic of neutrophil-CD4 T cell co-culture. (**G**) Proliferative fraction of CD4 T cells after co-culture with neutrophils. (**H**) Sepsis differential cell abundance across SRS. (**I**) Surface CD69 (left) and PD1 (right) positive CD4 T cells after co-culture with sepsis neutrophils. (**J**) Neutrophil phagocytosis across SRS. Functional assays tested with two-sided Wilcoxon rank-sum tests.

It has been postulated that immature neutrophils are immunosuppressive (*11*). We found increased expression of genes for experimentally validated granulocytic myeloid-derived suppressor cells in both mature and immature neutrophils from septic patients compared to HCs (Fig. S5A,B). Functional co-culture validation showed sepsis neutrophils inhibited CD4 T cell proliferation (Fig. 1F,G) and activation (Fig. S5C). Increased arginase-1 activity and PD-1 PD-L1/2 pathways are proposed mechanisms for G-MDSC immunosuppression (*12, 13*), but we were unable to reverse suppression when inhibiting either pathway (Fig. S5D). Instead, we achieved partial reversal by inhibiting the prostaglandin EP2 receptor, which increased CD69 expression (Fig. S5D). Neither a cyclo-oxygenase or prostaglandin EP4 receptor inhibitor showed a similar effect, suggesting action at the level of pre-formed prostaglandin E2 (Fig. S5D).

We next investigated the cellular and molecular drivers of subphenotypes within sepsis. Whole blood transcriptomics identifies SRS1, or high quantitative SRS (SRSq) score patients, who have features of immunosuppression, more severe disease, and poor outcome (*2, 8-10*). We assigned SRS (Methods; Fig. S6A,B) and found expansion of *IL1R2*+ immature neutrophils and increased cycling neutrophil progenitors in SRS1, while MNCs including cMonos, NK, and memory T cells

were depleted (Fig. 1H; Fig. S6C). We observed large numbers of differentially expressed genes between SRS groups in mature, *S100A8/9* high, degranulating, and *IL1R2*+ immature neutrophils (Fig. S7A), with corresponding separation by SRS status on principal components analysis (PCA) for these populations (Fig. S7B). By contrast, there were minimal differences between SRS in the MNC subsets (Fig. S8), consistent with neutrophils driving SRS groupings. Functionally, there was no difference across SRS in neutrophil suppression of CD4 T cell proliferation (Fig. S9A), but we found that SRS1 neutrophils suppressed CD4 T cell activation more than non-SRS1 neutrophils (Fig. 1I). Moreover, SRS1 neutrophils displayed reduced phagocytosis (Fig. 1J), which we were unable to restore with GM-CSF (Fig. S9B). Taken together, our data were consistent with the hypothesis that SRS1 represents an immunosuppressed state.

To verify that neutrophils are the main drivers of differences in SRS, we re-analyzed whole blood bulk RNA-seq and single-cell mass cytometry (CyTOF) immunophenotyping data for 36 all-cause sepsis patients (42 samples) (cohort 2 (*14*); Fig. 1A; Table S1). After SRS assignment, we recapitulated higher early mortality in SRS1 (Fig. S10A,B). We identified eight neutrophil clusters in the CyTOF dataset (Methods; Fig. S10C-E; Fig. 2A) and found more immature and pro/pre-neutrophils in SRS1 (Fig. 2B; Fig. S10F), alongside greater density in the earlier stages of the pseudotime trajectory in SRS1 (Fig. S10G-I). PCA on cell proportions showed separation by neutrophil but not MNC compartments for SRS groups (Fig. 2C), confirming the centrality of neutrophils to SRS definitions. We then analyzed additional sepsis whole blood transcriptomic data (cohort 3, n=667 patients) (Fig. 1A) using cell type/state proportion deconvolution (Methods), and found significant *IL1R2*+ immature neutrophil and cycling neutrophil progenitor expansion in SRS1 (Fig. 2D).

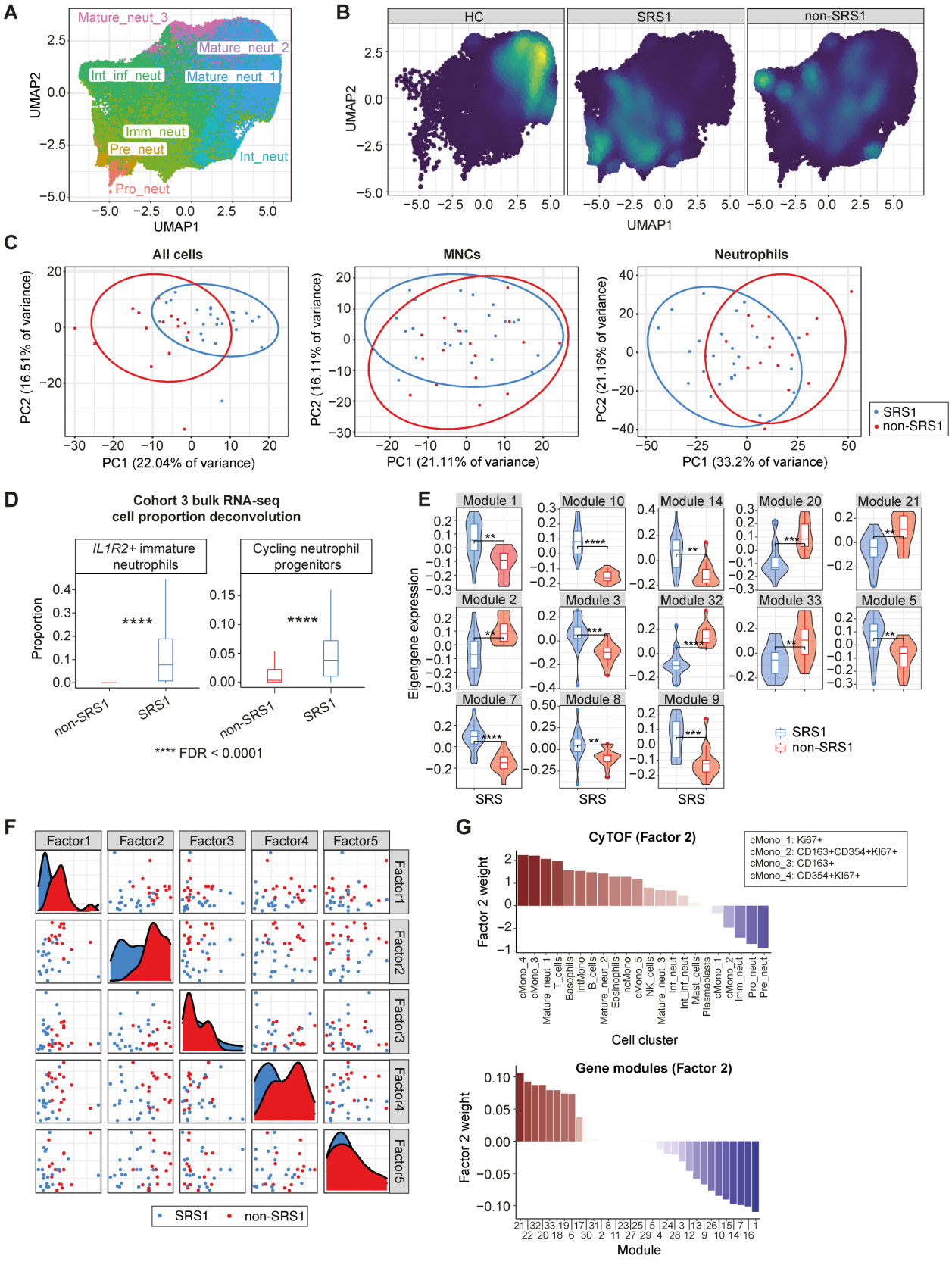
A
Mature_neut_3
Mature_neut_2
Int_inf_neut
Mature_neut_1
Imm_neut
Pre_neut
Int_neut
Pro_neut
UMAP2
UMAP1
B
HC
SRS1
non-SRS1
UMAP2
UMAP1
C
All cells
MNCs
Neutrophils
PC2 (16.51% of variance)
PC1 (22.04% of variance)
PC2 (16.11% of variance)
PC1 (21.11% of variance)
PC2 (21.16% of variance)
PC1 (33.2% of variance)
SRS1
non-SRS1
D
Cohort 3 bulk RNA-seq
cell proportion deconvolution
IL1R2+ immature neutrophils
Cycling neutrophil progenitors
Proportion
****
non-SRS1
SRS1
**** FDR < 0.0001
E
Module 1
Module 10
Module 14
Module 20
Module 21
Module 2
Module 3
Module 32
Module 33
Module 5
Module 7
Module 8
Module 9
Eigengene expression
SRS
SRS1
non-SRS1
F
Factor1
Factor2
Factor3
Factor4
Factor5
SRS1
non-SRS1
G
CyTOF (Factor 2)
cMono_1: Ki67+
cMono_2: CD163+CD354+KI67+
cMono_3: CD163+
cMono_4: CD354+KI67+
Factor 2 weight
cMono_4
cMono_3
Mature_neut_1
T_cells
Basophils
intMono
B_cells
Mature_neut_2
Eosinophils
ncMono
cMono_5
NK_cells
Mature_neut_3
Int_neut
Int_inf_neut
Mast_cells
Plasmablasts
cMono_1
cMono_2
Imm_neut
Pro_neut
Pre_neut
Cell cluster
Gene modules (Factor 2)
Factor 2 weight
Module

Figure 2. Validation of neutrophil immaturity underlying SRS1. (**A-C**) Cohort 2. (**A-B**) UMAP of (**A**) CyTOF neutrophil subsets (1,921,471 cells) showing (**B**) density-based distribution by SRS/HC status. (**C**) PCA of CyTOF neutrophil and MNC proportions. (**D**) Cell type/state proportions in cohort 3 sepsis patients (n=667) following whole blood bulk RNA-seq deconvolution (two-sided Wilcoxon rank-sum test). (**E-G**) Cohort 2. (**E**) Differentially expressed modules in whole blood bulk RNA-seq (FDR < 0.01, two-sided Wilcoxon rank-sum test). (**F**) Scatter plots of latent factor values (Methods) for sepsis patients (n=36). (**G**) Weights per cell cluster and gene module for latent factor two. FDR = false discovery rate.

Weighted gene co-expression network analysis (WGCNA) of cohort 2 bulk RNA-seq data identified 13 differentially expressed modules across SRS (FDR < 0.01) (Fig. 2E), with SRS1 upregulated modules enriched for signatures of differentiating neutrophils (*15*) (Fig. S11A). The WGCNA module most correlated with CyTOF immature neutrophil proportions (module 10; Fig. S11B) showed enrichment for a signature defining *IL1R2*+ immature neutrophils (Fig. S11C), replicating *IL1R2*+ immature neutrophil expansion in SRS1 (Fig. 1H). We further utilised factor analysis to integrate cell composition and bulk gene expression (with plasma protein measurements as a computational negative control) (Methods; Fig. S11D). Across SRS, factors one and two were the most divergent (Fig. 2F; Fig S11E,F), and connected immature neutrophil subsets with gene modules upregulated in SRS1 (Fig. 2E,G; Fig. S11E,G).

We proceeded to investigate which pathways and mediators were contributing to the observed differences in SRS neutrophil subsets. In cohort 1, consensus non-negative matrix factorisation (Methods) showed correlation of SRSq score with gene expression programs in mature neutrophils, *S100A8/9* high neutrophils, *IL1R2*+ and *PADI4*+ immature neutrophils (Fig. 3A; Fig. S12A-E;

Table S3). These programs were enriched for pathways including IL-6 JAK-STAT3 signalling, TNF, MTORC1, hypoxia and complement (Fig. 3B); and targets of transcription factors (TFs) including RUNX1, SPI1 and CBFB (roles in myelopoiesis) (*16*), CEBPB and STAT3 (key to emergency granulopoiesis [EG]) (*17, 18*), as well as CEBPA (role in steady state granulopoiesis, SSG) (Fig. 3C). This was supported by enrichment of STAT3 and SPI1 targets in module 10 from cohort 2 (immature neutrophil correlated module) (Fig. 3D). Plasma levels of myelopoietic cytokines M-CSF and GM-CSF were not differential across SRS, whereas G-CSF and IL-6, which canonically signal through STAT3 and are known to be key contributors to EG (*18, 19*), were elevated in SRS1 (Fig. 3E). Previous reports linked G-CSF priming of neutrophils in cancer to increased neutrophil NETosis (*20*), and we similarly found SRS1 neutrophils underwent more NETosis than non-SRS1 neutrophils (Fig. 3F,G).

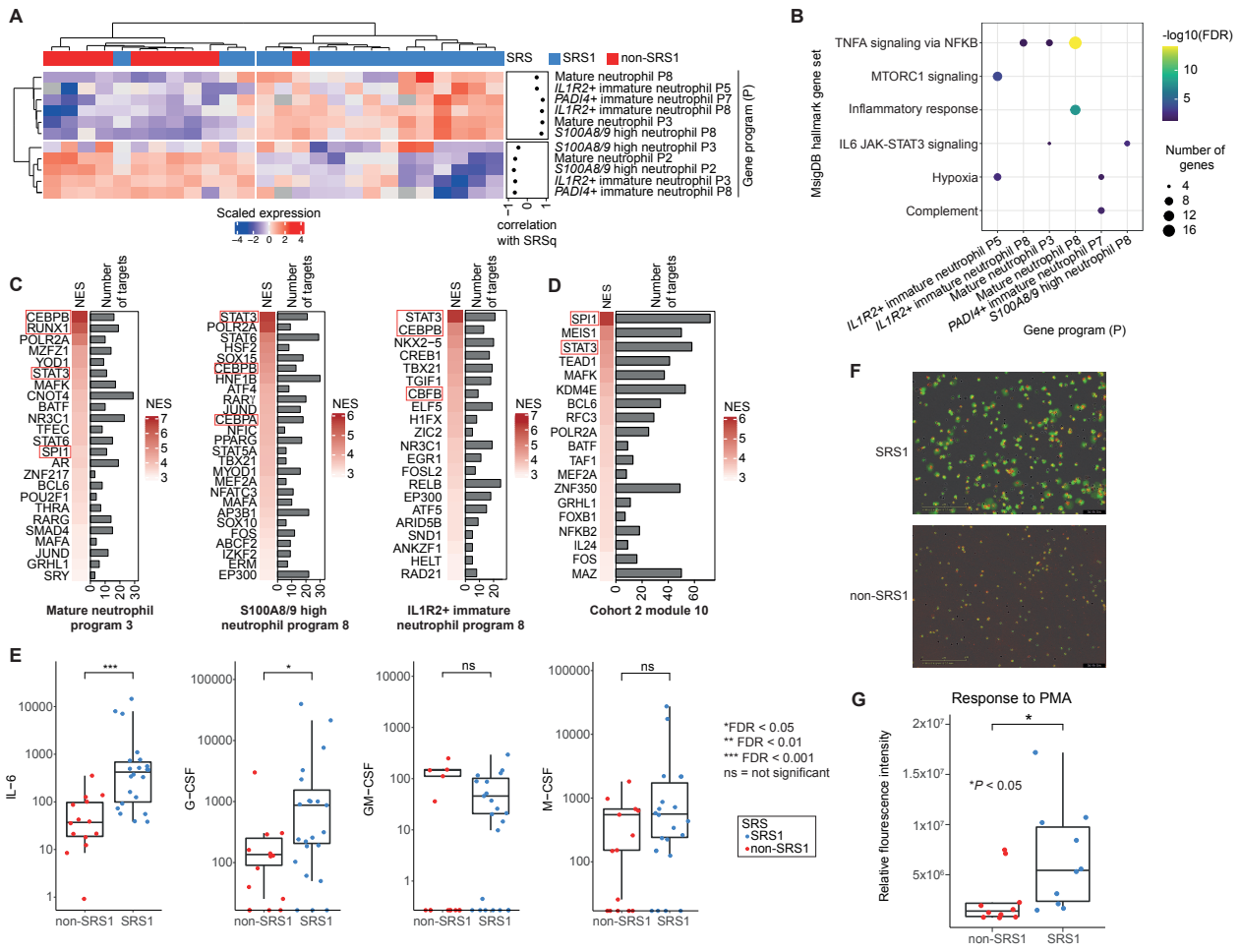

Figure 3. Gene programs and mediators in SRS1 neutrophils. (**A-C**) Neutrophil gene programs in cohort 1. (**A**) Gene expression heatmap for programs significantly correlated with SRSq (Spearman *Rho*, FDR < 0.05). (**B**) Pathway enrichment for programs with significant positive correlations with SRSq. (**C**) Transcription factor (TF) prediction analysis (top 50 genes per program) for three neutrophil programs positively correlated with SRSq. (**D-E**) Cohort 2. (**D**) TF prediction analysis for top 1% genes showing highest correlation with whole blood bulk RNA-seq module 10 eigengene. (**E**) Plasma myelopoietic cytokine levels by SRS (two-sided Wilcoxon rank-sum test). (**F**) Representative images of neutrophils undergoing NETosis from SRS1 and non-SRS1 patients (four hours post-PMA stimulation, fluorescence green denotes cells undergoing

NETosis). (**G**) NETosis in neutrophils across SRS (two-sided Wilcoxon rank-sum test). FDR = false discovery rate.

We therefore hypothesized that SRS1 was a state of dysregulated granulopoiesis, specifically involving heightened EG. To test this, we characterized CD34+ hematopoietic stem cells (HSCs) enriched from PBMCs (n=27 samples), assaying RNA and chromatin (assay for transposase accessible chromatin-sequencing; ATAC-seq) landscapes of each cell (sample set 4; Fig. 4A; Methods). We analyzed 29,336 HSCs (Fig. S13 and S14; Methods), using both modalities to identify seven clusters (Fig. 4B; Methods). We observed shifts in HSC distribution in sepsis vs. HC with cluster C6 most enriched in HCs (Fig. S15A,B), and across SRS, with cluster C5 enriched in SRS1 (Fig. 4B; Fig. S15C). We next investigated the relationship between cluster C5 and granulopoiesis. Overall, 184 chromatin peaks were differentially open in C5 (FDR < 0.05, FC > 1.5) (Table S4; Fig. S16A), and enriched for reported HSC/multipotent progenitor and common myeloid progenitor chromatin profiles (Fig. S16B) suggesting a myelopoietic shift. This was supported by identification of CEBPB and STAT3 motifs within C5 peaks on enrichment analysis (Fig. S16C), and peak set overlap with CEBPB/STAT3 occupancy in public ChIP-seq datasets (Fig. 4C; Fig. S16D). Comparing C5 with other clusters, 251 genes were differentially upregulated (Table S4; Fig. S17A). We then leveraged annotated HSC and progenitor scRNA-seq data (Methods) to derive mononuclear phagocyte and granulocyte defining myelopoietic gene signatures (Fig. S17B). We saw no difference across SRS in expression of the former, but found higher expression of the latter in SRS1 HSCs (Fig. 4D). Beyond SSG, we identified EG gene sets from a mouse model of sepsis where granulocyte progenitors were isolated (Fig. S17C), and scored our HSCs for genes upregulated in sepsis vs. control animals of the same granulocyte progenitor type, finding higher expression of three EG gene sets in SRS1 vs. non-SRS1 (Fig. 4D).

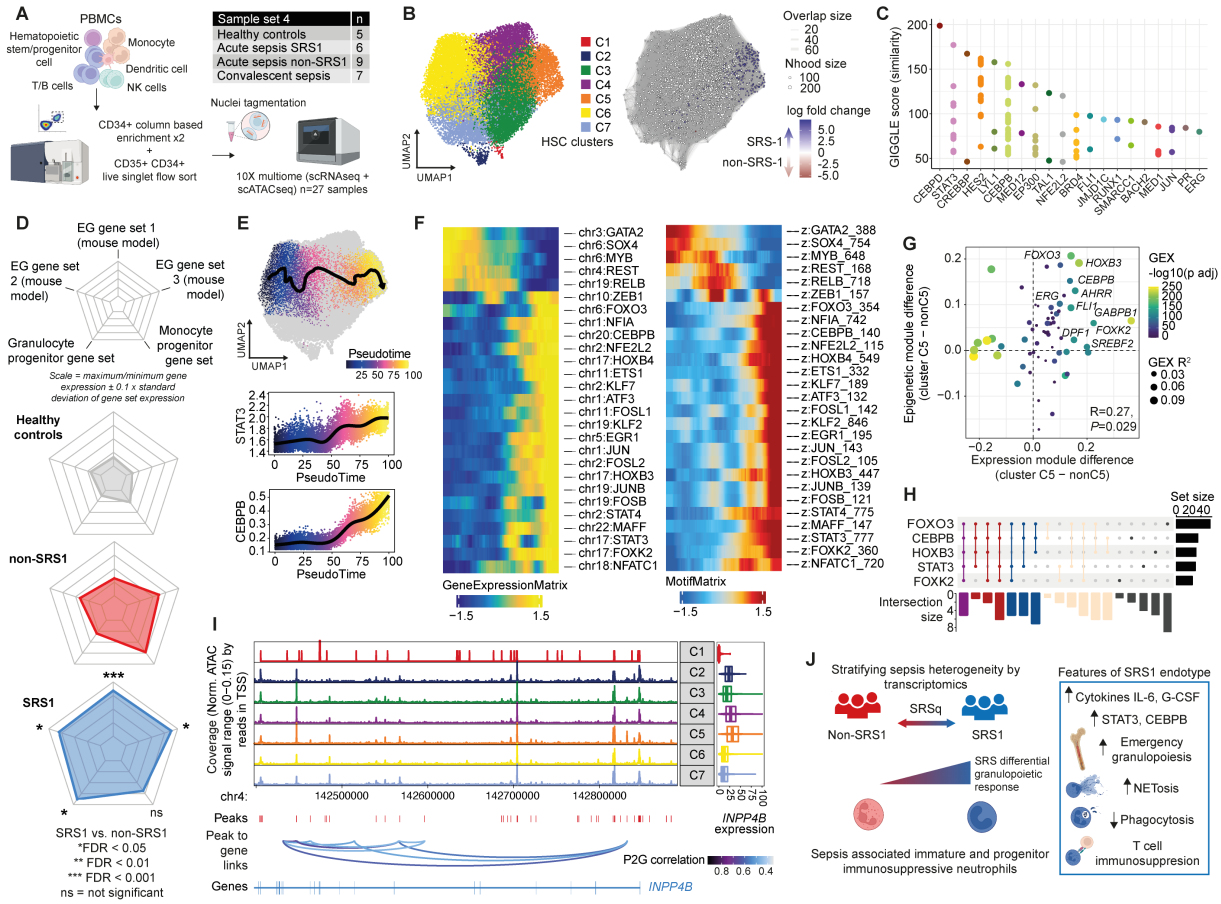

Figure 4. Heightened emergency granulopoiesis in SRS1. (**A**) HSC approach (sample set 4). (**B**) UMAP of RNA and ATAC defined HSC clusters (29,336 cells) (left) and differential cell abundance across SRS (right) with sampled neighborhoods colored by statistically significant enrichment (spatial FDR < 0.05). **(C)** ChIP-seq overlap analysis (GIGGLE score, Methods) for cluster C5 differentially open peaks. (**D**) Gene set enrichment of HSCs for signatures of human granulopoiesis and monopoiesis, and sepsis mouse model emergency granulopoiesis (EG). **(E)** UMAP of HSCs with supervised pseudotime trajectory (left) and STAT3 and CEBPB gene expression with pseudotime (right). **(F)** Genes (left) and transcription factor (TF) motifs (right) which correlate along pseudotime. **(G)** TF modules identified from integrative RNA and chromatin modeling (Methods). **(H)** UpSet plot of genes regulated by EG TFs. **(I)** *INPP4B* locus showing

aggregated scATAC profiles and gene expression per HSC cluster, peaks called and peak-to-gene links. **(J)** Summary of sepsis endotypes and features defining the poor outcome SRS1 state.

To gain further insight into the TF regulatory dynamics in SRS1 EG, we integrated both data modalities. We first established a supervised pseudotemporal trajectory from cluster C6 (HC-enriched) to C5 (SRS1 EG). *STAT3* and *CEBPB* gene expression increased along the trajectory as anticipated (Fig. 4E). We then extracted genes and chromatin TF motifs which changed along the trajectory (Fig. S18A) and retained TFs which were correlated across modalities to identify a set of TFs potentially regulating EG (Fig. 4F). Next, we modeled gene expression as a function of TF expression and regulatory region accessibility to identify 65 TF modules (Methods; Table S5). On mixed model testing for differential TF module activity across C5 vs. other clusters, the CEBPB (EG) module ranked highly in both significance and variance explained (Fig. 4G), while the CEBPA (SSG) module was lower in C5 (Fig. S18B), providing further evidence for heightened EG in SRS1. Finally, we intersected the top TF modules by variance explained with TFs identified from the correlational trajectory analysis, identifying FOXO3, FOXK2 and HOXB3 as potential novel regulators of EG (Fig. 4G,H; Fig. S18C,D). Overlapping genes in these TF modules with CEBPB and STAT3 modules, we found *INPP4B, FKBP5, NEAT1, SETBP1* and *DLGAP2* to be commonly regulated (Fig. 4H,I; Fig. S18E-H), implicating these TFs and their putative target genes for future work investigating SRS1 based EG.

Our data reinforce the specificity of immature and progenitor neutrophils to sepsis, definitively link their expansion to a deleterious SRS1 endotype, and show disruption of neutrophil function in sepsis and SRS1. While evidence for altered myelopoiesis in sepsis has been described in mice (*21, 22*) and modeled in vitro (*23*), here we are able to demonstrate this in patients, in the form of

amplified EG in SRS1. Previous clinical trials of G-CSF in sepsis failed to reduce mortality (*24-26*), while trials of GM-CSF are ongoing (*27*). However, an improved understanding of the mechanisms by which these drugs influence neutrophils and granulopoiesis, and which patients should be targeted, will be key to success. The addition of granulopoietic stimuli may not enhance the functional capacity of circulating immature and immunosuppressive cells, and the effects on an already altered hematopoietic system are unknown. In view of increased levels of G-CSF and IL-6, heightened EG, and immunosuppressive neutrophils in SRS1 (Fig. 4J), inhibition of these two cytokines in SRS1 patients represents a potential immunotherapeutic strategy to boost immune function.

References and notes:

1. T. van der Poll, M. Shankar-Hari, W. J. Wiersinga, The immunology of sepsis. *Immunity* **54**, 2450-2464 (2021).
2. E. E. Davenport *et al.*, Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med* **4**, 259-271 (2016).
3. B. P. Scicluna *et al.*, Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. *Lancet Respir Med* **5**, 816-826 (2017).
4. T. E. Sweeney *et al.*, Unsupervised Analysis of Transcriptomics in Bacterial Sepsis Across Multiple Datasets Reveals Three Robust Clusters. *Crit Care Med* **46**, 915-925 (2018).
5. M. L. Washburn *et al.*, T Cell- and Monocyte-Specific RNA-Sequencing Analysis in Septic and Nonseptic Critically Ill Patients and in Patients with Cancer. *J Immunol* **203**, 1897-1908 (2019).
6. M. Reyes *et al.*, An immune-cell signature of bacterial sepsis. *Nat Med* **26**, 333-340 (2020).
7. K. Reddy *et al.*, Subphenotypes in critical care: translation into clinical practice. *Lancet Respir Med* **8**, 631-643 (2020).
8. K. L. Burnham *et al.*, Shared and Distinct Aspects of the Sepsis Transcriptomic Response to Fecal Peritonitis and Pneumonia. *Am J Respir Crit Care Med* **196**, 328-339 (2017).
9. D. B. Antcliffe *et al.*, Transcriptomic Signatures in Sepsis and a Differential Response to Steroids. From the VANISH Randomized Trial. *Am J Respir Crit Care Med* **199**, 980-986 (2019).
10. E. Cano-Gamez *et al.*, An immune dysfunction score for stratification of patients with acute infection based on whole blood gene expression. *medRxiv*, 2022.2003.2017.22272427 (2022).

11. F. Veglia, M. Perego, D. Gabrilovich, Myeloid-derived suppressor cells coming of age. *Nat Immunol* **19**, 108-119 (2018).
12. F. Uhel *et al.*, Early Expansion of Circulating Granulocytic Myeloid-derived Suppressor Cells Predicts Development of Nosocomial Infections in Patients with Sepsis. *Am J Respir Crit Care Med* **196**, 315-327 (2017).
13. A. Meghraoui-Kheddar *et al.*, Two New Neutrophil Subsets Define a Discriminating Sepsis Signature. *Am J Respir Crit Care Med* **205**, 46-59 (2022).
14. C. O.-M.-o. B. A. C. E. a. julian.knight@well.ox.ac.uk, C. O.-M.-o. B. A. Consortium, A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* **185**, 916-938 e958 (2022).
15. A. Giladi *et al.*, Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat Cell Biol* **20**, 836-846 (2018).
16. R. Avellino, R. Delwel, Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation. *Blood* **129**, 2083-2091 (2017).
17. M. G. Manz, S. Boettcher, Emergency granulopoiesis. *Nat Rev Immunol* **14**, 302-314 (2014).
18. H. Zhang *et al.*, STAT3 controls myeloid progenitor growth during emergency granulopoiesis. *Blood* **116**, 2462-2471 (2010).
19. F. Walker *et al.*, IL6/sIL6R complex contributes to emergency granulopoietic responses in G-CSF- and GM-CSF-deficient mice. *Blood* **111**, 3978-3985 (2008).
20. M. Demers *et al.*, Cancers predispose neutrophils to release extracellular DNA traps that contribute to cancer-associated thrombosis. *Proc Natl Acad Sci U S A* **109**, 13076-13081 (2012).
21. G. F. Weber *et al.*, Interleukin-3 amplifies acute inflammation and is a potential therapeutic target in sepsis. *Science* **347**, 1260-1265 (2015).
22. I. Kwok *et al.*, Combinatorial Single-Cell Analyses of Granulocyte-Monocyte Progenitor Heterogeneity Reveals an Early Uni-potent Neutrophil Progenitor. *Immunity* **53**, 303-318 e305 (2020).
23. M. Reyes *et al.*, Plasma from patients with bacterial sepsis or severe COVID-19 induces suppressive myeloid cell production from hematopoietic progenitors in vitro. *Sci Transl Med* **13**, (2021).
24. L. Bo, F. Wang, J. Zhu, J. Li, X. Deng, Granulocyte-colony stimulating factor (G-CSF) and granulocyte-macrophage colony stimulating factor (GM-CSF) for sepsis: a meta-analysis. *Crit Care* **15**, R58 (2011).
25. R. K. Root *et al.*, Multicenter, double-blind, placebo-controlled study of the use of filgrastim in patients hospitalized with pneumonia and severe sepsis. *Crit Care Med* **31**, 367-373 (2003).
26. D. P. Stephens *et al.*, Randomized, double-blind, placebo-controlled trial of granulocyte colony-stimulating factor in patients with septic shock. *Crit Care Med* **36**, 448-454 (2008).
27. E. M. Pinder *et al.*, Randomised controlled trial of GM-CSF in critically ill patients with impaired neutrophil phagocytosis. *Thorax* **73**, 918-925 (2018).
28. B. Kaminow, D. Yunusov, A. Dobin, STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. *bioRxiv*, 2021.2005.2005.442755 (2021).

29. A. T. L. Lun *et al.*, EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* **20**, 63 (2019).
30. A. Gayoso *et al.*, Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat Methods* **18**, 272-282 (2021).
31. S. Sinha *et al.*, Dexamethasone modulates immature neutrophils and interferon programming in severe COVID-19. *Nat Med* **28**, 201-211 (2022).
32. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**, 1408-1414 (2020).
33. E. Dann, N. C. Henderson, S. A. Teichmann, M. D. Morgan, J. C. Marioni, Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat Biotechnol*, (2021).
34. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
35. D. Kotliar *et al.*, Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *Elife* **8**, (2019).
36. R. Janky *et al.*, iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol* **10**, e1003731 (2014).
37. D. Bayik *et al.*, Myeloid-Derived Suppressor Cell Subsets Drive Glioblastoma Growth in a Sex-Specific Manner. *Cancer Discov* **10**, 1210-1225 (2020).
38. H. Alshetaiwi *et al.*, Defining the emergence of myeloid-derived suppressor cells in breast cancer using single-cell transcriptomics. *Sci Immunol* **5**, (2020).
39. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
40. S. B. Hay, K. Ferchen, K. Chetal, H. L. Grimes, N. Salomonis, The Human Cell Atlas bone marrow single-cell interactive web portal. *Exp Hematol* **68**, 51-61 (2018).
41. I. Korsunsky *et al.*, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296 (2019).
42. M. Ogishi *et al.*, Multibatch Cytometry Data Integration for Optimal Immunophenotyping. *J Immunol* **206**, 206-213 (2021).
43. S. Van Gassen *et al.*, FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636-645 (2015).
44. M. Nowicka *et al.*, CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res* **6**, 748 (2017).
45. K. E. Diggins, A. R. Greenplate, N. Leelatian, C. E. Wogsland, J. M. Irish, Characterizing cell subsets using marker enrichment modeling. *Nat Methods* **14**, 275-278 (2017).
46. K. Street *et al.*, Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
47. R. Argelaguet *et al.*, MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* **21**, 111 (2020).
48. A. M. Newman *et al.*, Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773-782 (2019).
49. Y. Huang, D. J. McCarthy, O. Stegle, Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol* **20**, 273 (2019).
50. J. M. Granja *et al.*, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403-411 (2021).

51. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
52. J. M. Granja *et al.*, Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol* **37**, 1458-1465 (2019).
53. R. Zheng *et al.*, Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res* **47**, D729-D735 (2019).
54. R. M. Layer *et al.*, GIGGLE: a search engine for large-scale integrated genome analysis. *Nat Methods* **15**, 123-126 (2018).
55. J. S. Fleck *et al.*, Inferring and perturbing cell fate regulomes in human cerebral organoids. *bioRxiv*, 2021.2008.2024.457460 (2021).

Acknowledgments: We thank the patients and healthy donors who participated for their generous contribution to this study. We also thank the clinical staff, particular Andrew Brent and Alex Novak, and Oxford University NHS Hospital Trusts recruiting hospitals involved in patient recruitment and sample collection. We thank the Oxford Genomics Centre for sequencing support, in particular Angela Lee for project management and Isar Nassiri for bioinformatic support. We thank the Wellcome Centre for Human Genetics and Jenner Institute flow facilities for assistance in sorting experiments. We are grateful to Mosutafa Attar, Fabiola Curion, Kasia Kedzierska, Kaspar Märtens, Yuxin Mi, Patrick McLean, Giuseppe Scozzafava and members of the Knight lab for support and critical discussion. We also thank the COMBAT consortium for access to data for re-analysis. Schematic representations of experimental strategies and SRS models were produced with Biorender.

Funding: We acknowledge the support of the National Institute for Health Research through the Comprehensive Clinical Research Network for patient recruitment. This research was funded by the Medical Research Council (MR/V002503/1) (J.C.K., E.D.), Wellcome Trust Investigator Award (204969/Z/16/Z) (J.C.K), Chinese Academy of Medical Sciences (CAMS) Innovation 537 Fund for Medical Science (2018-I2M-2-002 (J.C.K.), Croucher Foundation (A.J.K.), Wellcome Trust Grants (090532/Z/09/Z and 203141/Z/16/Z) to core facilities Wellcome Centre for Human

Genetics, Wellcome Trust core funding to the Wellcome Sanger Institute (Grant number WT206194), and NIHR Oxford Biomedical Research Centre (J.C.K.). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Authors contributions: Conceptualization: A.J.K. and J.C.K. Data curation: A.J.K. Formal Analysis: A.J.K., A.A., M.S. and Y-X.Z. Funding acquisition: J.C.K. and E.D. Investigation: A.J.K., A.A., M.S. and R.C.F. Methodology: A.J.K., E.C-G. and K.L.B. Project administration: J.C.K., C.M., I.U., J.A.T. and E.D. Resources: A.J.K., J.C.K., A.M., S.M. and C.J.H. Supervision: J.C.K. Visualization: A.J.K. and J.C.K. Writing – original draft: A.K. and J.C.K. Writing – review & editing: All authors read, provided input on, and approved the manuscript.

Competing interests: All other authors declare no competing interests.

Data and materials availability: Raw and derived data for whole blood single-cell sequencing and hematopoietic stem cell single-cell sequencing datasets are deposited on the European Genome-phenome Archive (EGA) (EGAS00001006283**)** and Zenodo. For sequence level raw datasets deposited at EGA, access is managed by a Data Access Committee. Code will be made available upon reasonable request.

Supplementary Materials

Materials and methods

Study ethics, patient cohorts and sample sets

Cohort 1

Volunteers self-reporting as healthy (healthy controls, HCs) and with no history of infection in the past 14 days were consented and recruited into the Genetic diversity and gene expression in white blood cells study [South Central Oxford REC B, reference 06/Q1605/55]. Blood samples were collected on one occasion.

Sepsis patient samples were collected from patients older than 18 years of age admitted to Oxford University Hospitals NHS Foundation Trust, UK. Patients were recruited either from the intensive care unit (ICU) if they had symptoms and signs of established sepsis (suspected infection with an acute change in total Sequential Organ Failure Assessment (SOFA) score of ≥2 points), or from the Emergency Department and medical wards if they had a change in quick SOFA score by ≥2 points and a NEWS2 score of ≥7 or intensive care review requested. Exclusion criteria were as previous in the UK Genomic Advances in Sepsis (GAinS) study (NCT00121196) (*2*) and included: patients or consultees unwilling or unable to give consent or advice declaration; advanced directive to withhold or withdraw life sustaining treatment; admission for palliative care only; pregnancy and 6 weeks post-partum; or severe acquired immunodeficiency including systemic high dose steroid therapy (prednisolone 0.5mg/kg/day for 14 days or equivalent), HIV infection, known regular therapy with immunosuppressive agents e.g. azathioprine, or neutrophil counts <1000/mL

due to any cause including metastatic disease and hematological malignancies or chemotherapy but excluding severe sepsis and solid organ/bone marrow transplant recipient receiving immunosuppressive therapy.

Post cardiac surgery samples were collected from patients older than 18 years of age admitted to Oxford University Hospitals NHS Foundation Trust, UK. Patients were deemed eligible if they were 1) undergoing cardiac bypass surgery, 2) required post-operative intensive care unit stay and 3) did not have an infection prior to surgery. Exclusion criteria were identical to those for sepsis patients.

Sepsis and post cardiac surgery patients were consented into the Sepsis Immunomics Study [South Central Oxford REC C, reference:19/SC/0296] between May and November 2021. Samples were collected on days 1, 3 or 5 of either hospital or intensive care admission. Written informed consent was obtained from adults or advice declarations from personal/nominated consultees for patients lacking capacity, with retrospective consent obtained from the patient once capacity was regained.

Cohort 2

HCs and sepsis patients were recruited into the same studies as in cohort 1 under the same criteria and ethics and were analyzed previously as part of a COVID-19 centric bioresource (*14*).

The primary endpoint for the SRS groups was chosen to be 14-day mortality from sampling based on previous SRS analyses that early mortality was significantly different between the SRS groups (*2, 8*).

Cohort 3

Adult patients (≥18 years old) were recruited through the UK Genomic Advances in Sepsis (GAinS) study (NCT00121196) upon admission to intensive care with sepsis due to community-acquired pneumonia (CAP) or faecal peritonitis (FP); https://ukccggains.com/). Patients were recruited from 34 UK intensive care units (ICUs), with diagnoses, inclusion and exclusion criteria as described by Davenport et al (*2*). Ethics approval was granted nationally and locally (RECs 05/MRE00/38, 08/H0505/78, and 06/Q1605/55), with informed consent obtained from all patients or their legal representative.

Sample set 4

Sample set 4 (Fig. 4A) included samples from sepsis patients and HCs recruited under the same studies as in cohorts 1 and 2, with one and eight acute sepsis samples overlapping with cohorts 1 and 2 respectively, while six acute sepsis samples were not analyzed in either cohort. Of the eight convalescent sepsis samples, four were from patients in each of cohorts 1 and 2. Sample set 4 also included seven age and sex matched HCs, of which five were the same HCs as in cohort 2.

Fresh whole blood sample processing for single-cell sequencing (cohort 1)

Blood samples were drawn with EDTA Vacutainer tubes (BD Biosciences) and processed within one hour of collection. One mL of whole blood was lysed with nine mL of 1X eBioscience red blood cell (RBC) lysis buffer (Thermofisher) for 10 minutes, quenched with 35 mL of Hanks' Balanced Salt Solution (Lonza) and centrifuged at 300g for 10 minutes. The RBC lysis was repeated with three mL lysis buffer for three minutes and quenched and centrifuged again. Cells were then counted and 200,000 cells were transferred for antibody staining and single cell capture.

Whole blood single cell RNA and cell surface protein profiling (cohort 1)

Single cell RNA and cell surface protein sequencing was performed with the BD Rhapsody platform (633731 and 633733), using the whole transcriptome amplification assay (633801) and 30 Rhapsody Abseq antibodies (Table S2). One µL of each Abseq antibody was added to 200,000 total white cells and topped up with 45 µL of BD staining buffer for 30 minutes. Samples were washed three times in 100 µL BD Rhapsody sample buffer + 200 µL BD staining buffer at 400g (5 minute centrifugation), before counting and diluting for single cell capture.

The concentration of cells was diluted to target a capture of 6000 cells per sample before loading onto the BD Rhapsody cartridge. Reverse transcription, complementary DNA (cDNA) amplification (633773) and library construction (633801) were performed following manufacturers' recommendations in six batches. Libraries were sequenced on the NovaSeq6000 (Illumina).

Whole blood single cell multi-omic analysis (cohort 1)

Analysis was performed with v4.0.0 R and python 3.8.6.

Preprocessing and quality control

Gene expression data were aligned to the GRCh38 reference genome using STARsolo (*28*) (v2.7.9a) with settings:

```
--soloType CB_UMI_Complex \
--soloUMIlen 8 \
--soloCBposition 0_0_0_8 0_21_0_29 0_43_0_51 \
--soloUMIposition 0_52_0_59 \
--soloFeatures Gene Velocyto \
--soloMultiMappers EM \
--soloCBmatchWLtype 1MM \
```

Unfiltered files i.e. containing all 884,736 cell barcodes on the whitelist were used downstream for an additional round of cell calling by the emptyDrops function within DropletUtils (v1.10.3) (*29*) with a UMI threshold of 100 and FDR of 0.5%. Spliced and unspliced counts matrices were also produced with the Velocyto option in –soloFeatures.

Abseq reads were trimmed with Trimmomatic (v0.39) to 48 bases i.e. 12bp UMI + 36bp Abseq nucleotide sequence. The reads were then aligned to an artificial reference of the Abseq nucleotide sequences also using STARsolo with settings:

```
--alignIntronMax 1 \
--alignMatesGapMax 1 \
--winAnchorMultimapNmax 1 \
--chimMainSegmentMultNmax 1 \
--outFilterMultimapNmax 1 \
--seedMultimapNmax 1 \
```

Cells expressing fewer than 100 genes, more than 4000 genes, > 10% mitochondrial reads, > 2% haemoglobin reads and with a log10(UMI/gene) < 0.6 were removed (Fig. S1D-G). Gene filtering was then performed with genes expressed in < 10 cells or a total count of < 3 removed.

Scrublet (v.0.2.3) and doublet detect (v3.0) were then both used on default settings to generate doublet scores. Automatic thresholds were applied to remove multiplets.

Normalization, dimensionality reduction, visualisation and clustering

Data were normalized (scanpy v1.7.2: normalize_total), log + 1 corrected (scanpy: log1p) and 4000 HVGs identified using the Seurat vst algorithm (scanpy: highly_variable_genes, flavor='seurat_v3').

Multi-modal dimensionality reduction (DR) was performed with TotalVI (scvi-tools v0.10.0) (*30*) on default settings with all 30 proteins and 4000 HVGs with each individual sample set as a batch (as all samples were processed separately). The 20 latent dimensions identified were used for downstream UMAP visualisation and graph construction for clustering and differential cell abundance analysis.

Unsupervised clustering was performed by first building a shared nearest neighbours graph with *k* = 30 nearest neighbours in Seurat (v4.0). Choice of *k* parameter for number of nearest neighbours was made following inspection of Clustree (v0.4.20) analysis of a range of values for *k* and the effect on cluster splitting. Graph building was followed by clustering with the smart local moving algorithm in the Seurat FindClusters function (algorithm = 3).

Evaluation of clustering resolution was done by combining three metrics including assessment of cluster neighbourhood purity, cluster average silhouette width and a 30 iteration bootstrap to determine cluster stability with respect to sampling noise (bluster v1.0), and one subjective element i.e. inspection of clusters and their top defining markers to match clustering results with known biology and understand any potential value in merging vs. splitting. Combining these four elements, the default clustering resolution of 0.8 was chosen, with the splitting of one immature neutrophil cluster into two (*MPO*+ immature neutrophils/progenitors and *PADI4*+ immature neutrophils), based on the clustering resolution of 1.1.

Cell type and state annotation and RNA velocity/trajectory analysis

Cell types and states as defined at the above chosen resolutions were then merged for annotation of major known immune cell types at a broad level, and kept at the clustering resolution of choice for annotation at a fine granularity. Protein markers enabled identification of major lineages which were merged for the broad level annotations. Correct lineage assignment was confirmed by corroborating marker driven annotation with SingleR (v1.6.1) automatic assignments (Fig. S2B). Fine annotation was then performed based on inspection of biologically known/potentially meaningful gene markers to subdivide T cell and neutrophil populations (Table S3). For neutrophil clusters, maturity annotation was assisted by CD10 expression (a marker of neutrophil maturity) on UMAP projection (Fig. S3A) which was consistent with the low CD10 neutrophil clusters expressing marker genes associated with immaturity, including *CD24* and *ARG1 (31)* (Fig. S3A,C), as well as genes encoding neutrophil granule proteins such as *MPO, ELANE, AZU1, CTSG,* and *RNASE2* (Fig. S3B). We annotated the clusters according to gene markers that particularly distinguished the neutrophil state, including a cluster of neutrophils with marked *IL1R2* expression (Fig. S3C) and a population undergoing cell cycle (Fig.S3D).

The unspliced and spliced counts from STARsolo were used for RNA velocity analysis of the neutrophils (without degranulating or apoptosing neutrophils) with scVelo (v0.2.3, default settings of stochastic model) (*32*). Moments were estimated using TotalVI reduced dimensions, and partition based graph abstraction was run based on the annotated neutrophil clusters on default settings.

Differential abundance (DA) analysis

Cell type/state DA changes across conditions were identified by sampling neighborhoods of cells from a *k*-nearest neighbors graph and looking for enrichment of either condition in each neighborhood as implemented in MiloR (*33*). The 20 batch corrected latent dimensions identified via multi-modal DR with TotalVI were used for MiloR (v0.99.19) kNN graph construction (k = 30). The proportion of cells sampled for neighbourhood indexing was 0.1. Differential abundance (DA) testing was then performed with generalised linear models including age and sex as covariates. Significantly different neighbourhoods were determined by a spatial corrected FDR < 0.05.

Differential gene expression (DGE) analysis

For pseudobulk analysis, gene expression was aggregated per individual and per cell type/state into a pseudobulk count matrix i.e. raw UMI counts per gene were summed across all cells belonging to the same cluster and to the same individual. Genes were filtered per pseudobulk based on minimum expression of n counts in at least X samples, where X was the smallest comparator group, and n was defined for each pseudobulk based on histograms of logged count distributions. Sample and cell type/state based pseudobulks were then normalized by the EdgeR trimmed mean of M-values (TMM) method (v3.30.3) (*34*). The normalized count matrix was log transformed for principal components analysis (PCA).

Pseudobulk DGE was performed with a generalised linear model as implemented in EdgeR, with the contrasts of interest specified and age, sex and sequencing batch included in the model.

Consensus non-negative matrix factorization (cNMF) of neutrophil gene expression

First, 1000 highly variable genes for each neutrophil state were selected with the SelectIntegrationFeatures function in Seurat. Each of the neutrophil states (except degranulating neutrophils, apoptosing neutrophils and cycling neutrophil progenitors) were then analyzed with cNMF (v1.2) as previously described (100 factorizations) (*35*) for between four to 15 programs, with final program number for each neutrophil state chosen based on a tradeoff of stability vs. error for a range of program numbers factorized out as recommended by the original authors. Mean program usage per sample for each neutrophil state was correlated with SRSq scores per sample with FDR adjustment for all programs except programs which picked up outlier individual cell-specific expression signatures (*IL1R2*+ immature neutrophil program 1, *PADI4*+ immature neutrophil program 5, *MPO*+ immature neutrophil/progenitor program 7; Fig. S12). The top 50 genes per program were taken for downstream pathway and transcription factor prediction analysis.

Transcription factor (TF) prediction analysis

Gene programs positively correlated with SRSq were used for TF prediction analysis with the Cytoscape (v3.9.1) plugin iRegulon (v1.3) (*36*). The top 50 genes of each of the gene programs were used and the prediction analysis was run on default settings (i.e. 10K motif collection, 20kb around TSS putative regulatory region, 20kb centred around TSS (7 species) Motif rankings database). TFs with a normalized enrichment score (NES) threshold of >3.0 were extracted and the NES plotted per TF. The same TF analysis was performed for the top 1% genes correlating with module 10 eigengene from cohort 2 RNA-seq WGCNA analysis (below).

IL1R2+ neutrophil defining gene set

A gene set specific to *IL1R2+* neutrophils was identified by taking the *IL1R2+* neutrophil data from sepsis samples and looking for gene markers of these cells against all other cells with the FindMarkers function in Seurat. The resulting markers were then filtered for an adjusted *P*-value of < 0.01 and an average log2FC > 1 to obtain markers specific to the cells.

Gene set enrichment scoring for single cells

Gene sets for granulocytic myeloid derived suppressor cells (G-MDSC) (*37, 38*) were scored in the dataset using the AddModuleScore function in Seurat, followed by two sided Wilcoxon rank-sum tests to test for differences across different groups of cells.

Pathway enrichment analysis

Pathway enrichment analysis was performed against MSigDB Hallmark pathways via the ClusterProfiler R package (v4.1.4) (*39*) with over-representation analysis via Fisher's exact tests as implemented in the enricher function.

Sepsis endotype (sepsis response signature) assignment/scoring

To assign SRS endotypes for sepsis samples, gene expression was aggregated per individual for all cells to produce a per sample pseudobulk count matrix i.e. raw UMI counts per gene were summed across all cells belonging to the same individual. Genes were then filtered per pseudobulk based on minimum expression of 5 counts in at least *X* samples, where *X* was the smallest sample group by comparator group i.e. six HCs. Sample pseudobulks were then normalized by the EdgeR

trimmed mean of M-values (TMM) method. The relevant genes (*DYRK2*, *CCNB1IP1*, *TDRD9*, *ZAP70*, *ARL14EP*, *MDC1*, and *ADGRE3*) from the normalized count matrix were then used for sepsis sample SRS assignment and SRSq score calculation by SepstratifieR (v0.0.0.9) (*10*).

Neutrophil functional assays

Isolation of neutrophils from whole blood

Neutrophils were isolated from 5-10 mL whole blood from EDTA Vacutainer tubes (BD Biosciences) using EasySep HLA Chimerism Whole Blood CD66b positive selection kit (StemCell) following manufacturer's instructions. Cells were washed with HBSS (Lonza) and resuspended in complete media (RPMI-1640 10% FCS, 100 IU/mL L-glutamine, 100 IU/mL penicillin-streptomycin) for downstream assays.

Phagocytosis Assay

Neutrophils were incubated in five mL polypropylene test tubes for 20 minutes in complete media with or without pHrodo Green E.Coli Bioparticles (Invitrogen) at a concentration of 1 in 10 for assessing phagocytosis. Cells were washed and stained for surface markers for 30 minutes with 7-AAD, CD66b-AF700 (G10F5) and Siglec-8-APC (7C9) antibodies from Biolegend and fixed using Fix/Lyse solution (Invitrogen) for 10 minutes before acquisition using BD LSRFortessa X-20 analyzer. The median fluorescent intensity (MFI) for single 7AAD- CD66b+ Siglec-8+ gated cells was determined by subtraction of the MFI of the *fluorescence minus one* (FMO) control sample that was not incubated with any pHrodo bioparticles.

Neutrophil - allogeneic CD4 T cell co-culture

Cryopreserved CD4+ T cells isolated from healthy donor leukocyte cones were thawed and stained with eBioscience Cell Proliferation Dye eFluor450 (Invitrogen) at a final concentration of 10 μM for 20 minutes following manufacturer's recommendations. T cells were then immediately co-cultured with the isolated neutrophils in 96 well U bottom plates at 37°C 5% CO_2 at a 4:1 neutrophil to T cell ratio in complete media supplemented with 50 IU/mL recombinant human IL-2 (Biolegend) and with anti-CD3/CD28 dynabeads (Gibco). The volume of anti-CD3/CD28 beads added was equivalent to a 1:1 bead to T cell ratio and the total cell density was maintained at 200,000 cells per well. As controls, T cells were also plated alone without neutrophils in every run, with and without anti-CD3/CD28 beads. To reverse T cell inhibition, we tested co-cultures with 1 mM L-arginine (Sigma-Aldrich), 1 μM arginase-1 inhibitor CB-1158 (Fisher Scientific), anti-CD274 (eBioscience clone MIH1), anti-CD273 (PD-L2) (eBioscience clone MIH18), 20 μg/mL PGE_2 (Sigma-Aldrich), 10 μM indomethacin (Sigma-Aldrich), 15 μM EP_2 inhibitor TG-6-10-1 (MedChemExpress) or 15 μM EP_4 inhibitor GW627368 (MedChemExpress).

After 3-4 days of co-culture, cells were stained for surface markers for 30 minutes at room temperature followed by staining with Annexin-V and 7-AAD for 15 minutes to determine the percentage of live non-apoptosing cells for each sample. Samples were analyzed using the BD LSRFortessa X-20. All antibodies were purchased from Biolegend unless otherwise stated: CD3-APC (UCHT1), CD4-BUV395 (BD Biosciences, SK3), CD66b-AF700 (G10F5), PD1-PE (NAT105), CD69-PECy7 (FN50), 7-AAD, Annexin V- FITC. T cells were gated from a live single cell population as CD66b- CD4+ CD3+ Annexin-V- 7AAD- cells. Proliferation analysis by dye dilution was established using the non-bead stimulated T cells cultured alone without neutrophils

as the baseline for determining the proliferative fraction of cells in each sample. The proliferative fraction and percentage of cells expressing PD-1 and CD69 were calculated as a percentage relative to the anti-CD3/28 bead stimulated no co-culture T cells control to account for donor variation.

Flow cytometry data analysis

All flow cytometry data was analyzed using Flowjo (v10) and Graphpad Prism (v9).

NETosis assay

Neutrophils were plated at 20,000 cells/well in low auto-fluorescence F-12K Nut Mix (Gibco) media (Hams-F12) using flat-bottomed 96-well plates which had been coated for 1 hour with 0.01% poly-L-ornithine solution (Sigma). Cells were allowed to attach for up to 1hr prior to treatment with 100nM phorbol 12-myrisate 13 acetate (PMA), as per Incucyte assay recommendations, or left untreated as controls. All cells were also treated with 250 nM IncuCyte Cytotox Green Dye, used to measure NETosis-induced cell death. The final volume in each well was made up to 100 μL using Hams-F12 media. The plated cells were then transferred to the IncuCyte Live-Cell Analysis System incubator and scans were scheduled for every 20 minutes for up to 4 hours, using standard scan type, objective 20X with green fluorescence channel selection.

Bulk RNA-seq analysis (cohort 2)

Analysis was performed with v4.0.0 R. The edgeR trimmed mean of M-value normalized and filtered count matrix was obtained from the COMBAT consortium (*14*), with 23,063 genes from 42 samples analyzed.

Exploratory analysis and sepsis endotype (sepsis response signature) assignment/scoring

For cohort 2 sepsis samples, the relevant genes (*DYRK2*, *CCNB1IP1*, *TDRD9*, *ZAP70*, *ARL14EP*, *MDC1*, and *ADGRE3*) were extracted from the normalized count matrix and used for SRS assignment and SRSq score calculation by SepstratifieR (v0.0.0.9) (*10*). PCA was performed with the top 15% most variable features transformed at log2(counts per million + 1).

Weighted gene co-expression network analysis

WGCNA (v1.70-3) was conducted as per the vignettes from (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/) with the log normalized count matrix. The soft thresholding power of 4 was first chosen by taking the lowest power which reached a scale free topology model fit of 0.8. This was then used to construct a signed-hybrid network with biweight midcorrelation and a deepsplit parameter of 2 with a minimum module size of 30. Modules were summarised by eigengene expression (the value of the first principal component of the module) and modules with eigengene correlation of <0.2 were merged for a total of 33 modules identified. Differential testing of modular eigengene expression across SRS groups was conducted with two-sided Wilcoxon rank-sum tests, adjusting for multiple testing with the Benjamini-Hochberg method.

Pathway analysis

Pathway enrichment analysis was performed against neutrophil gene sets from the Human Cell Atlas (*40*) and signatures of differentiating and differentiated neutrophils from mouse single cell transcriptomics (*15*) via the ClusterProfiler R package with over representation analysis via

Fisher's exact tests as implemented in enricher function. Gene set enrichment analysis for cohort 1 *IL1R2*+ neutrophil defining genes was performed via the GSEA function for the WGCNA module most correlated with immature neutrophil proportions.

CyTOF analysis (cohort 2)

Pre-processing

Normalized, debarcoded, and bead and doublet cleaned FCS files were obtained from the COMBAT consortium (*14*). To avoid biases due to highly varying cell numbers per sample, a maximum of 50,000 cells per sample was taken. As not all batches contained control samples for batch correction, a control sample free correction method Harmony (v1.0) *(41)* was adapted from scRNA-seq methodology for use given previous demonstration of utility in CyTOF data (*42*).

Clustering, trajectory inference and differential abundance analyses

The batch corrected matrix was fed into clustering analysis via FlowSOM (*43*) as implemented in the CATALYST package (v1.14.0) (*44*). This was first done with all cells and all 41 markers (12x12 starting grid) with metacluster merging by consensus clustering to 50 metaclusters followed by manual annotation and merging. As the 41 marker panel was designed to be myeloid focused i.e. to capture both monocyte and neutrophil subsets, we annotated five monocyte clusters (Fig. S9C, all neutrophil clusters merged) before subsetting the neutrophils and reclustering to 30 metaclusters with 17 selected markers. Neutrophil clusters were then manually annotated and merged based on median marker expression into eight subsets, with marker enrichment modeling scores (*45*) for analytical validation of annotation (Fig. S9E), leaving a final total of 22 clusters at the finest level of annotation (14 non-neutrophil and eight neutrophil).

For trajectory analysis, neutrophils were downsampled to 15,000 cells per comparator group (SRS1, non-SRS1 and healthy volunteers) for diffusion map dimension reduction as implemented in the CATALYST package followed by principal curve fitting on the minimum spanning tree with default settings on the slingshot function in the Slingshot package (v1.8.0) *(46)*.

The number of cells identified in each of the 22 clusters was divided by the total number of cells in the patient to evaluate the proportion of the cluster. PCA of cluster proportions was performed after quantile normalizing the cell.

Cluster DA analysis was performed using a generalised linear mixed model using the glmmPQL function from the MASS package (v7.3-51.6) to account for repeated sampling in six patients, with batch included as a covariate in the model and patient included as a random effect. An extra random effect term for each individual sample was included to model the overdispersion in proportions seen in high dimensional cytometry data (*44*). Adjustment for multiple testing of each cluster was performed with the Benjamini-Hochberg method.

Multi-modal analysis by MOFA+ (cohort 2)

To integrate cohort 2 bulk RNA-seq and CyTOF data, we used MOFA+ (v1.0.1) (*47*) to identify latent spaces of shared variation across datasets and to look for a direct relationship between pro/pre- and immature neutrophils and SRS1 upregulated gene modules. As a negative control, we included a plasma proteomic dataset from the same patient samples, reasoning that the latent factors linking cell clusters to gene expression profiles should not show an overly strong

contribution from the plasma proteomic data. Concentrations of 105 plasma proteins as determined by the timsTOF mass spectrometer were obtained from the COMBAT consortium for this (*14*).

For each sample, the eigengenes of the 33 identified modules from WGCNA, quantile normalized proportion of the 22 cell clusters from CyTOF clustering and annotation, and log2 transformed and median centered intensity values for 105 plasma proteins were utilised, with six samples missing plasma protein data and one sample missing CyTOF data (Fig. S11D). Views were left unscaled, number of factors was set at 10, and all other model options were left with default values. Training convergence was done in slow mode and other options were left with default values. Differences in factor values between SRS groups was tested with a two-sided Wilcoxon rank-sum test, adjusting for multiple testing with the Benjamini-Hochberg method.

Cytokine analysis (cohort 2)

Concentrations of 51 plasma analytes as determined by the Luminex assay were obtained from the COMBAT consortium (*14*). Differential testing of cytokine concentrations across SRS groups was conducted with two-sided Wilcoxon rank-sum tests, adjusting for multiple testing with the Benjamini-Hochberg method.

Cell type/state deconvolution of 864 sepsis bulk RNA-seq samples (cohort 3)

Bulk RNA sequencing data were obtained for 864 samples (667 patients) from the GAinS study. Briefly, whole blood samples were taken on the first, third, and/or fifth day following ICU admission and leukocytes were isolated using the LeukoLOCK system (Thermo Scientific). RNA was extracted using the manufacturer's protocol (Ambion), and libraries prepared using NEB Ultra

II Library Prep kits (Illumina) were sequenced on a NovaSeq6000 (Illumina). Reads were aligned to the GRCh38 reference genome using STAR (v2.7.3) and gene expression quantified using featureCounts. Finally, the count matrix was normalized using TMM from edgeR and log-transformed.

A signature matrix for CIBERSORTx (*48*) was constructed using our cohort 1 single cell dataset, with each finely annotated cell type and state except apoptosing cells downsampled to 100 cells per type/state. The signature matrix was then created via the Create Signature Matrix analysis module with a parameters min. expression = 0.25, replicates = 100 and sampling = 0.5.

Cell fractions of the 864 sample whole blood bulk RNA-seq dataset were then estimated with the Impute Cell Fractions analysis module using the single cell reference matrix, with batch correction S-mode enabled and quantile normalization disabled. For DA testing of cell types/states across SRS, only the 667 samples from unique individuals were used for two-sided Wilcoxon rank-sum tests per cell type/state with Benjamini-Hochberg correction for multiple testing.

PBMC isolation and cryopreservation (sample set 4)

PBMCs were isolated from sepsis (acute and convalescent) and HC whole blood using density gradient centrifugation with Leucosep tubes (Greiner) and lymphoprep (StemCell). After centrifugation, PBMCs were washed in cold HBSS (Lonza) (300g, 10 minutes at 4˚C), RBC lysed (four mL 1X RBC lysis buffer (Invitrogen) for four minutes on ice before quenching with cold HBSS), washed again and frozen down in FCS (Sigma) containing a final 10% DMSO (Cell

Signalling Technology). Cryovials were placed in a cryogenic storage container and stored at -80˚C for 24-72h before being moved to liquid nitrogen for long-term storage.

CD34+ haematopoietic stem and progenitor cell (HSPC) enrichment (sample set 4)

PBMCs from 15 sepsis patient samples, 7 age and sex matched HCs and 8 convalescent sepsis samples were used for CD34+ HSPC isolation. Individuals were recruited under the same study ethics as above in cohorts 1 and 2. Three convalescent sepsis samples were removed after preprocessing and demultiplexing ("HSPC single cell multi-omic analysis") but prior to downstream analysis due to clinical inappropriateness from retrospective evaluation of clinical information (samples judged too far from acute episode [over six months past hospital discharge]). The 30 samples were processed in six batches, with each batch containing at least one sample from each comparator group.

Cryopreserved PBMCs were thawed at 37°C and immediately washed twice, first with RPMI containing 25 U/mL benzonase (88700, Thermofisher), followed by magnetic activated cell sorting (MACS) buffer (phosphate buffered saline [PBS] + 0.5% BSA + 2mM EDTA). Cells from each sample were resuspended in 300 µL MACS buffer and 100 µL each of Fc block and CD34 microbeads (130-100-453, Miltenyi Biotec) were added and incubated for 30 minutes at 4°C.

Cells were washed once with MACS buffer (10 mL) at room temperature, followed by MACS column sorting twice i.e. the flow through from one column was passed through a second column for enhanced yield. Cells from both columns were eluted and centrifuged at 4°C in PBS + 2% BSA once (300g, 10 minutes) and stained at 4°C for FACS sorting with one µL each of

CD34-PE (clone AC136, Miltenyi Biotec) and CD45-FITC (clone HI30, Biolegend) for 20 minutes with DRAQ7 (live/dead) dye (424001, BioLegend) diluted 1:833 from the stock solution added in the final five minutes. Each sample was then FACS sorted for live, singlet, CD34+ CD45+ HSPCs.

HSPC single cell multi-omics (RNA + ATAC-seq; sample set 4)

Single-cell RNA and assay for transposase-accessible chromatin (ATAC) sequencing on CD34+ HSPCs was performed by isolating nuclei of flow sorted HSPCs. Buffers were made according to recommendations by the 10X Demonstrated Protocol (CG000365 Rev B) i.e. 1X nuclei buffer with 1mM DTT and 1 U/μL RNase inhibitor (3335399001, Sigma); wash and lysis buffers with 10mM Tris-HCl (pH 7.4), 10mM NaCl, 3mM $MgCl_2$, 1% BSA (130-091-376, Miltenyi Biotec), 0.1% Tween-20 (1662404, Bio-Rad), 1mM DTT and 1 U/μL RNase inhibitor, and in the lysis buffer additionally 0.1% Nonidet P40 Substitute (74385, Sigma) and 0.01% Digitonin (BN2006, Thermofisher).

An equal number of cells per sample was first pooled for each batch. Cells were pelleted in an microcentrifuge tube (400g at 8°C for four minutes) and after removal of supernatant, 40 μL of 1X lysis buffer was added to the cell pellet with three pipette mixes immediately. 60 μL of chilled wash buffer was then added immediately without mixing. Cells were spun at 500g for eight minutes at 4°C, after which the supernatant was aspirated and 45 μL of chilled 1X nuclei buffer was added for another wash. All supernatant was then aspirated and nuclei were resuspended in 5.5 μL 1X nuclei buffer for transposition immediately.

5 µL of the suspension was then used for nuclei transposition and gel bead in emulsion generation followed by RNA and ATAC library preparation (1000285, 10X) as per the 10X manufacturer's protocol. The RNA libraries were sequenced on the NovaSeq6000 (Illumina) and ATAC libraries sequenced on the NextSeq500 (Illumina) as per the 10X official requirement.

HSPC single cell multi-omic analysis (sample set 4)

Analysis was performed with v4.1.0 R.

Preprocessing and demultiplexing

Raw FASTQ files of scRNA-seq and scATAC-seq were aligned to the GRCh38 reference genome using 10X-arc (v2).

Genetic demultiplexing was performed for each batch as previously described (*14*). In brief, the GATK variant calling pipeline (https://github.com/gatk-workflows/gatk4-rnaseq-germline-snps-indels) (v4.1.7.0) was applied to the bulk RNA-seq data to identify sequence variants for the genetic demultiplexing of the single cell data. Vireo (v0.3.2) was then used for demultiplexing and heterotypic doublet calling (*49*).

Arrow file creation with ArchR software and doublet removal

Preprocessed HSPC multi-omic data were input into ArchR (v1.0.2) (*50*) to create ArrowFiles for analysis. In brief, fragment files from each 10X channel were reorganised to include sample

information from vireo demultiplexing, and an arrow file was created per sample with the MakeArrowFile function with mintss=4 and minFrags=1000 from ArchR resulting in 27 arrow files with three convalescent samples not included downstream (as above in "HSPC single cell multi-omics").

In addition to heterotypic doublet (doublets from different samples) removal from vireo demultiplexing, homotypic doublets were removed with the ArchR removeDoublet function for each sample individually (default settings).

The gene expression matrix was added to each arrow file by reading in the unfiltered 10X-arc outputs and matching the barcodes from this with the barcodes called as cells from ArchR.

RNA and ATAC quality control

Quality control was then performed on both scRNA and scATAC data (Fig. S12). Cells expressing fewer than 100 genes, more than 6000 genes or 25000 UMIs, and with a log10(UMI/gene) < 0.8 were removed. Cells with a TSS enrichment < 7 and < 1000 unique fragments were also filtered out. Gene filtering was then performed with genes expressed in < 10 cells or a total count of < 3 removed. This left 46,782 cells with a median TSS enrichment of 14.6 and median fragment count of 13,821 (Fig. S12D,E) and 26,660 genes for downstream analysis.

Cell type/state annotations and non-HSC filtering

We sought to specifically isolate HSCs for analysis and therefore annotated cell type/state identity through a multi-step data driven strategy involving RNA based reference mapping and de novo joint RNA chromatin clustering.

After QC, we assigned identity of HSPCs and contaminating cells by mapping our scRNA-seq data to two healthy donor reference bone marrow mononuclear cell (BMMC) scRNA-seq datasets from Hao et al (*51*) (reference dataset 1) (Fig. S14A) and Granja et al (*52*) (reference dataset 2) (Fig. S14C) with Seurat (v4.0). We used the supervised PCA mapping for reference dataset 1 as recommended by the Seurat authors as the reference dataset was annotated after WNN analysis of CITE-seq data, while we used PCA mapping for reference dataset 2. Cells with non-HSPC labels from either reference were filtered out (i.e. only cells labelled as HSPCs from both reference mappings were retained), leaving 46,156 HSPCs.

HSPCs were then taken downstream for multi-modalDR in ArchR with the default iterativeLSI function for scRNA (gene expression matrix) and scATAC (tile matrix) resulting in 30 dimensions each for scRNA and scATAC. Both scATAC and scRNA dimensions were corrected for batch with Harmony (default settings in ArchR) and combined with the addCombinedDims function. The cells were then clustered on the combined dimensions at resolution 0.4 with the addClusters function (smart local moving algorithm) (Fig. S14E). HSPC identity was then finalised by assigning the majority RNA mapping identity from reference dataset 1 per cluster (Fig. S14F). Progenitor cells were thus filtered out and only 29,336 HSCs retained for downstream analysis.

HSC peak calling, re-multi-modal DR, batch correction and clustering

Each individual sample was pseudobulked together per sample (i.e. all HSCs per sample) for peak calling with MACS2 and iterative peak overlapping / removal within ArchR with minimum cells set as the number of cells in the sample with the fewest cells, and minimum replicates set to 5 (lowest sample group size i.e. convalescent sepsis samples).

Before de novo HSC clustering, we re-performed multi-modal DR and batch correction on the gene expression matrix and peak matrix by repeating on default settings iterativeLSI and Harmony in ArchR respectively and addCombinedDims. HSCs were then re-clustered based on these reduced dimensions with the smart local moving algorithm at resolution 0.2 (a low resolution chosen to avoid overclustering and challenging interpretability given the purity of the cell type already).

Differential abundance analysis

HSC cluster DA changes across conditions were identified by sampling neighborhoods of cells from a *k*-nearest neighbors graph and looking for enrichment of either condition in each neighborhood as implemented in MiloR (v1.2.0). The 60 batch corrected reduced dimensions (30 each of scRNA and scATAC) were used in MiloR kNN graph construction ($k = 20$). The proportion of cells sampled for neighbourhood indexing was 0.3. DA testing was performed with generalised linear models including age and sex as covariates. Significantly different neighbourhoods were determined by a spatial corrected FDR < 0.05.

Differential chromatin accessibility and differential gene expression analysis

To identify the defining genomic regions and gene expression patterns of the HSC clusters, we utilised the getMarkerFeatures function in ArchR for either the peak matrix for scATAC or the gene expression matrix for scRNA, with peaks/genes called as significantly differentially up or downregulated at thresholds of FDR < 0.05 and FC > 1.5.

Bulk ATAC/TF motif enrichment analysis and ChIP-seq dataset overlapping

Differentially open regions for cluster C5 were taken for bulk ATAC/TF motif enrichment analyses and overlapping with publically available ChIP-seq datasets. The former was performed in ArchR with the peakAnnoEnrichment after addArchRAnnotations (collection="ATAC") and addMotifAnnotations. The latter was done by overlapping the peaks with ENCODE TF binding site datasets with the same functions in ArchR as with motif enrichment but with ENCODE TFBS annotations rather than motif annotations.

For further validation of overlaps with ChIP-seq data, we input our 184 differentially open peaks of cluster C5 into the cistromeDB toolkit (http://dbtoolkit.cistrome.org/) (*53*) using the top 1k peaks according to peak enrichment for peak number of Cistrome sample to use. The Giggle score denotes a composite significance and effect size scoring (*54*).

Steady-state and emergency granulopoiesis (EG) gene set enrichment scoring for single cells

To obtain gene sets defining steady state granulopoiesis and EG, we obtained re-analyzed scRNA-seq data of HCA HSPCs, and granulocyte progenitor RNA-seq data from cecal and ligation puncture (CLP) mouse models of sepsis respectively.

For the HCA HSPC scRNA-seq data, we utilised the annotations from Hay et al (*40*) to contrast all monocyte-dendritic cell progenitors (MDPs) with granulocytic progenitors to obtain MNP and granulocytic defining myelopoietic gene sets.

For the CLP sepsis mouse model data, we obtained a count matrix from GSE151682 (*22*) for all granulocyte monocyte progenitors, proNeu1, proNeu2 and preNeu, and normalized the data with edgeR TMM, then performed a QLF ANOVA equivalent test for control granulocyte monocyte progenitors (GMPs) vs. CLP GMPs, control pro-neutrophil (proN) 1 vs. CLP proN1 and control GMP vs. CLP proN1. We took the CLP upregulated genes (FDR < 0.05 and FC > 2) for each contrast as separate gene sets for EG.

We then scored our individual cells with these gene sets using the AddModuleScore function in Seurat. We examined differential gene set activity in all HSCs across SRS by testing a full linear mixed model of gene set~SRS + age + sex + batch + 1|sample for each gene set against a reduced model of gene set~age + sex + batch + 1|sample with an ANOVA test and FDR multiple testing correction.

Integrative trajectory analysis

A supervised pseudotemporal trajectory in sepsis and HC HSCs was constructed with the addTrajectory function in ArchR on the combined batch corrected LSI dimensions with clusters C6 (most enriched in HCs, Fig. S15B) and C5 (most enriched in SRS1, Fig. 4B) as the backbone. The trajectories for either gene expression or TF motifs were then obtained with the getTrajectory

function in ArchR (log2 normalization for gene expression but not TF motifs), and the trajectories were correlated with the correlateTrajectories function (maxDist = 500kb).

Gene regulatory network (GRN) construction and differential TF module testing

We built GRNs through the Pando framework (*55*) for sepsis and HC HSCs using 4000 HVGs from the FindVariableFeatures function in Seurat (default settings). GRNs were inferred with the infer_grn function in Pando with the regulatory region defined as 500kb +/- the transcription start site, a threshold of 0.05 for TF_cor, the Pando custom motif database and the GREAT option for selecting regulatory regions near genes. Thresholds of a minimum of 10 genes per TF module, R^2 of 0.1 per model and maximum of 200 regions per TF module were used to retain modules with reasonable interpretability, leaving 65 TF modules for analysis.

Module gene expression activity was calculated using the AddModuleScore function in Seurat for the genes positively regulated by each TF. Module epigenetic activity was determined by computing deviations in accessibility from the background (addBgdPeaks in ArchR) for the regions positively regulated by each TF with the addDeviationsMatrix function in ArchR. Module gene expression and epigenetic activity differences were calculated by subtracting the median module gene expressions and the z scores for epigenetic deviations respectively between clusters of interest.

Differential module gene expression activity across HSC clusters was tested per cell with a linear mixed model of module gene expression~cluster + age + sex + batch + 1|sample. Differential module gene expression activity across clusters was tested against a reduced model of module

gene expression ~age + sex + batch + 1|sample with an ANOVA test and FDR multiple testing correction. A mixed model marginal R^2 value per TF module was obtained fitting just module gene expression ~cluster + 1|sample.

Putative TF target gene chromatin accessibility plotting

For *INPP4B, FKBP5, NEAT1, SETBP1 and DLGAP2* which were commonly regulated by the TFs STAT3, CEBPB, FOXO3, FOXK2 and HOXB3, the local chromatin accessibility at these genes was plotted with the ArchR function plotBrowserTrack to show peak gene correlations at a minimum Pearson $R > 0.35$ and $FDR < 0.05$.

Supplementary figures

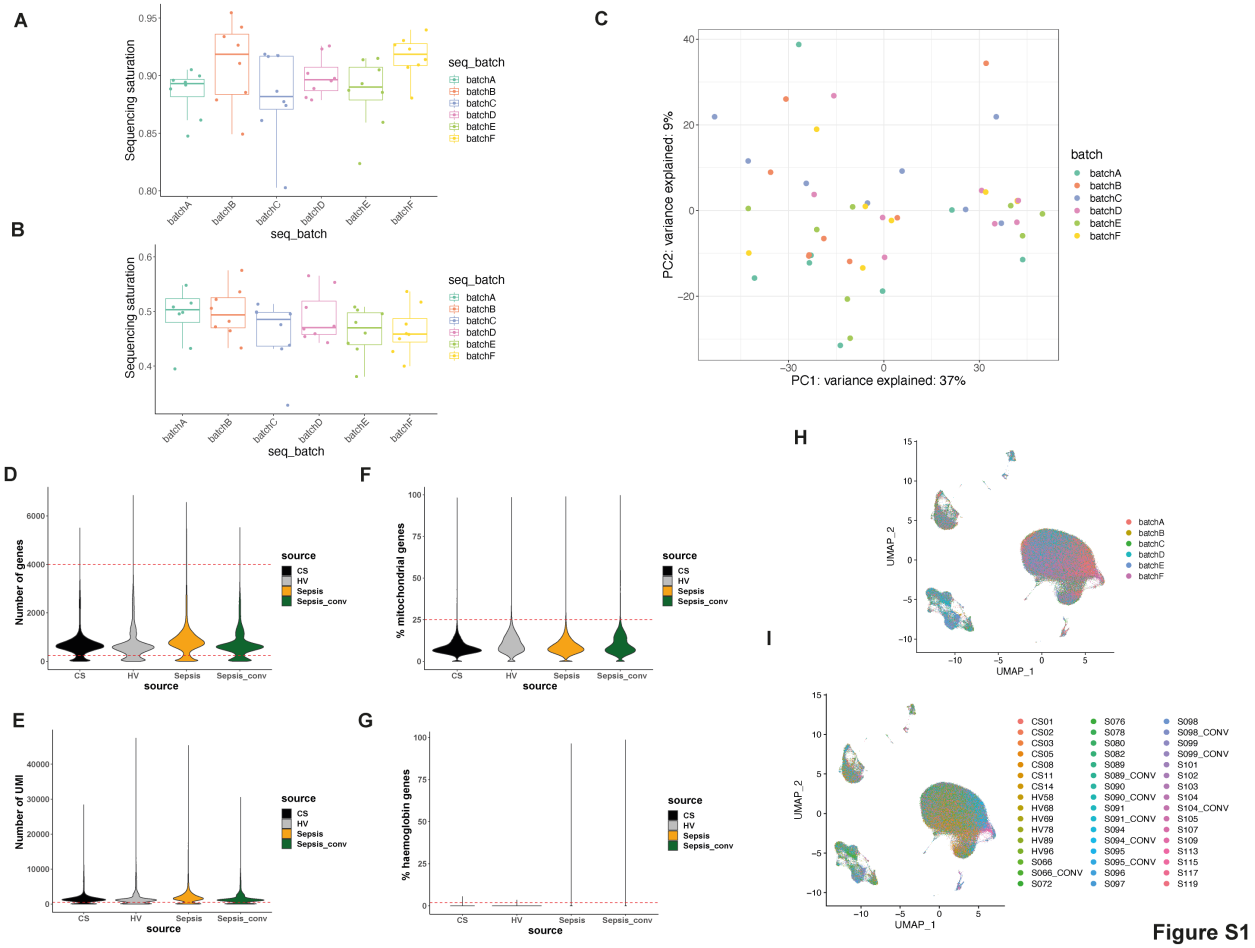

Figure S1

Figure S1. Whole blood single cell RNA and cell surface protein sequencing quality control (QC). **(A-B)** Sequencing saturation for **(A)** single cell RNA and **(B)** protein libraries per sequencing batch (seq batch). **(C)** First two principal components (PCs) from principal components analysis (PCA) of all 48 samples of cohort 1 pseudobulked colored by sequencing batch. **(D-G)** QC violin plots for various filtering thresholds (red dotted lines) for **(D)** number of genes detected per cell **(E)** number of UMIs per cell **(F)** percentage of mitochondrial genes detected per cell and **(G)** percentage of hemoglobin genes detected per cell. **(H-I)** UMAP of 272,993 cells after QC filtering colored by **(H)** sequencing batch or **(I)** sample identifier.

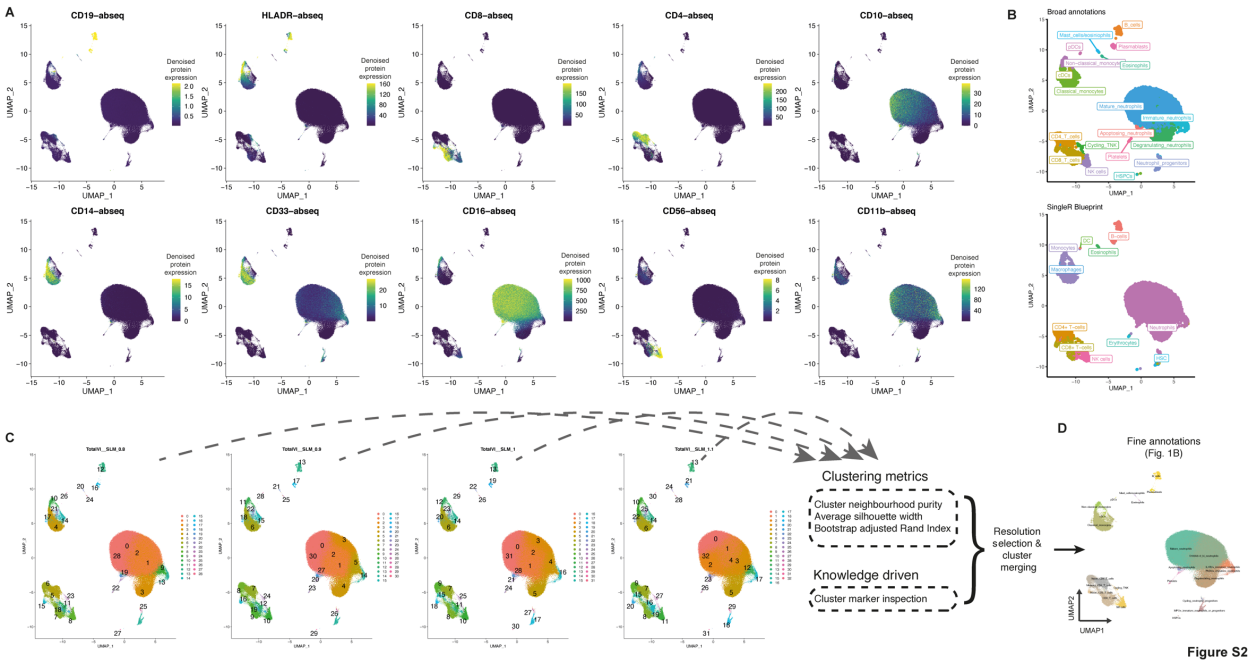

Figure S2. Annotation of and clustering cohort 1 single cell RNA and cell surface protein sequencing dataset. (A) UMAP of 272,993 cells with TotalVI denoised surface protein marker expression overlayed for selected lineage defining markers. **(B)** Comparison of broad annotations (Methods "Whole blood single cell multi-omic analysis (cohort 1)") with data driven algorithmic labelling of cell state/identity by SingleR based on reference bulk RNA-seq profiles of pure cell populations from the Blueprint Consortium. **(C)** Clustering results at varying resolutions (0.8-1.1) with smart local moving algorithm for fine annotation strategy. **(D)** UMAP of fine level annotations.

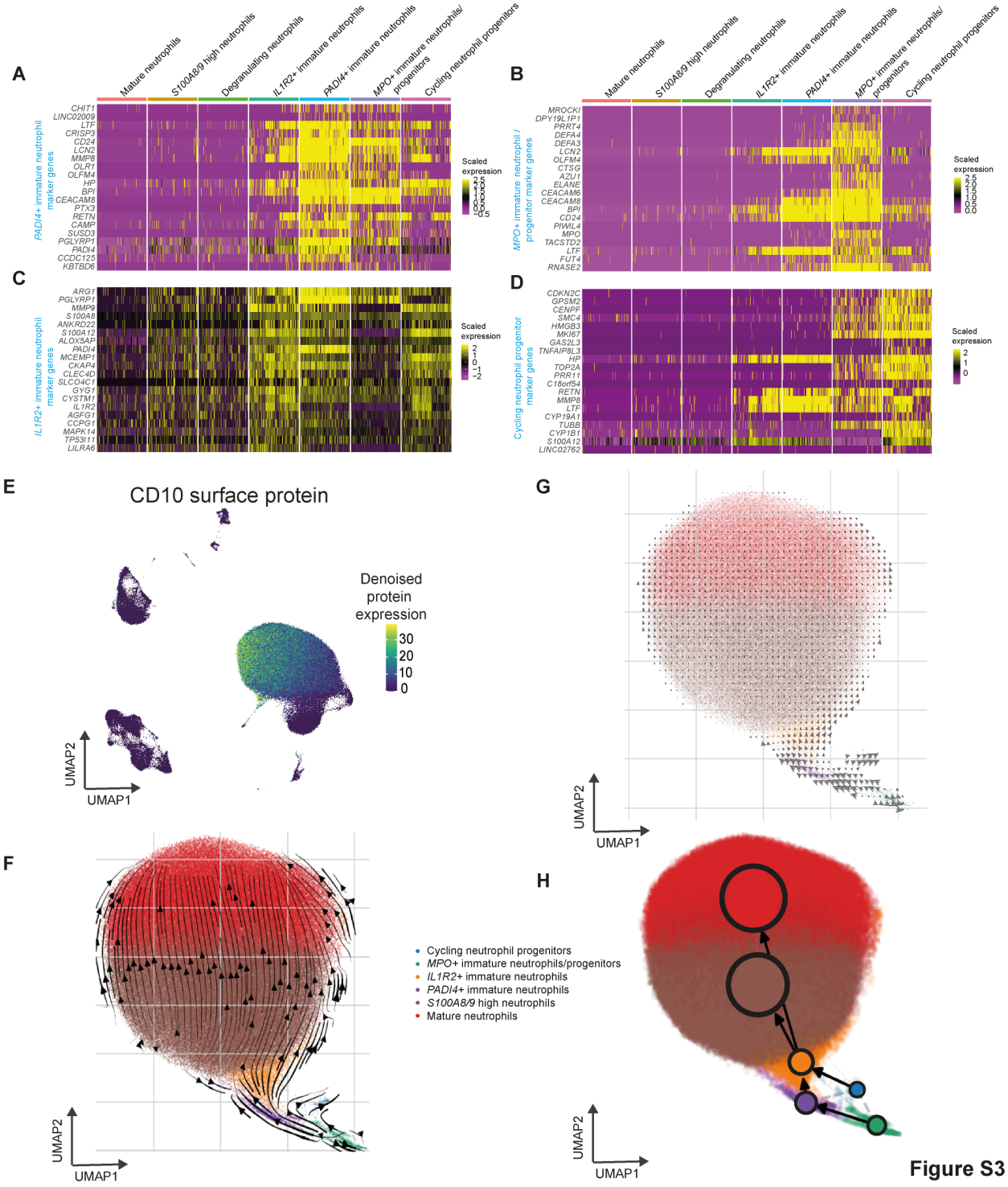

Figure S3. Immature neutrophil markers and neutrophil RNA velocity analysis. (A-D) Heatmaps of gene markers for **(A)** *PADI4+* immature neutrophils **(B)** *IL1R2+* immature neutrophils **(C)** *MPO+* immature neutrophils/progenitors and **(D)** cycling neutrophil progenitors.

(E) UMAP of all neutrophils with denoised CD10 protein marker expression overlay. **(F)** UMAP of neutrophils (excluding degranulating neutrophils) with RNA velocity stream directions plotted. **(G)** UMAP as per **(F)** with size of velocity vector throughout vector field depicted. **(H)** Partition based graph abstraction analysis of RNA velocity predicted cellular trajectory transitions.

5

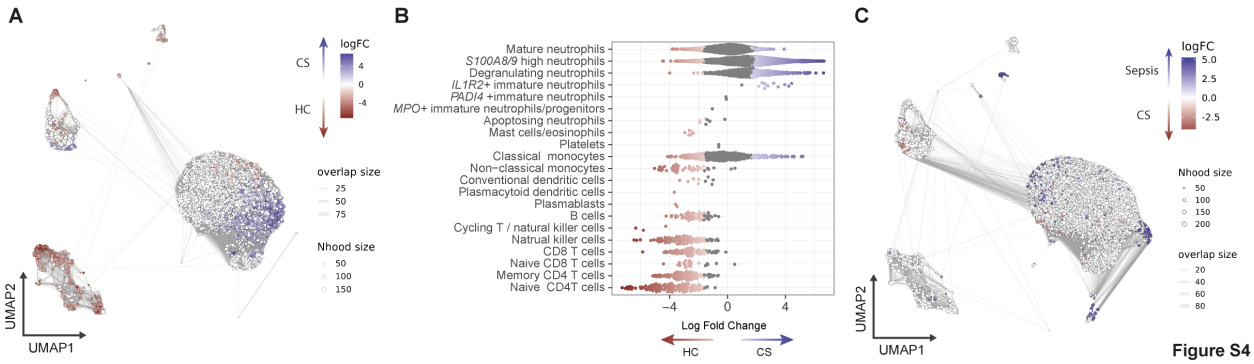

Figure S4

Figure S4. Healthy control (HC) – cardiac surgery (CS) and CS-sepsis differential abundance (DA) analysis. (A) DA analysis using graph neighborhood-based method to detect enrichment of neighborhoods between CS and HC samples, with UMAP of sampled neighborhoods and statistically significant enrichment colored (spatial FDR < 0.05 with generalised linear modeling). **(B)** Beeswarm plot of CS-HC DA analysis with cluster labels of neighborhoods depicted. **(C)** UMAP of DA analysis between CS and sepsis.

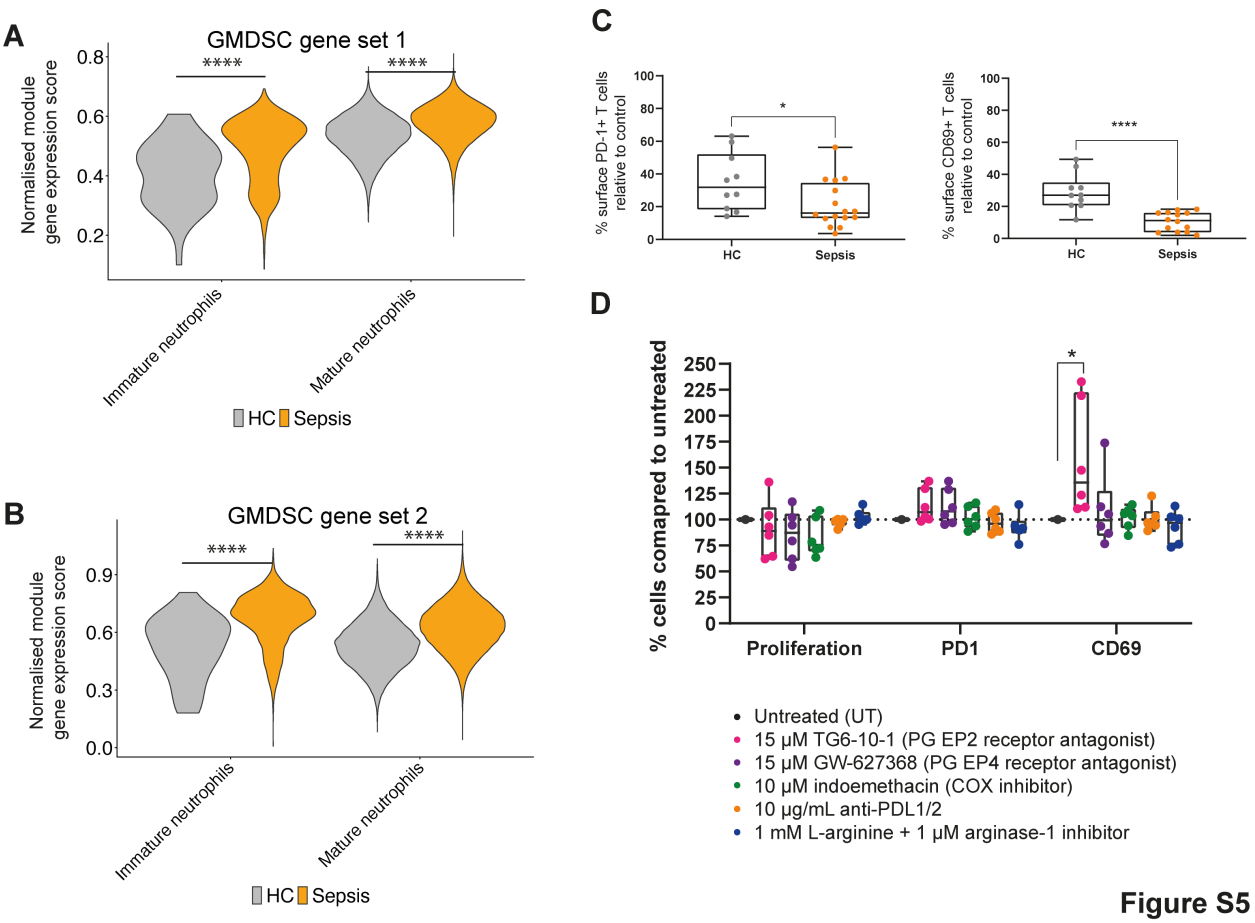

Figure S5

Figure S5. Functional interrogation of sepsis neutrophil immunosuppression. (A-B) Gene set scoring of mature and immature neutrophils in sepsis and HC samples by gene signatures of granulocytic myeloid derived suppressor cells. **(C)** Boxplots of surface PD-1 (left) and CD69 (right) positive anti-CD3/28 bead stimulated CD4 T cells after co-culture with neutrophils at a 4 neutrophil:1 T cell ratio, compared to positive controls of CD4 T cells cultured with anti-CD3/28 beads alone. **(D)** Percentage of CD4 T cells proliferating, expressing PD-1 or expressing CD69 in sepsis neutrophil-allogeneic T cell co-cultures with various treatments compared to UT control. All *P*-values were calculated with two-sided Wilcoxon rank-sum tests. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001. *GMDSC = granulocytic myeloid derived suppressor cells, UT = untreated, PG = prostaglandin, COX = cyclo-oxygenase.*

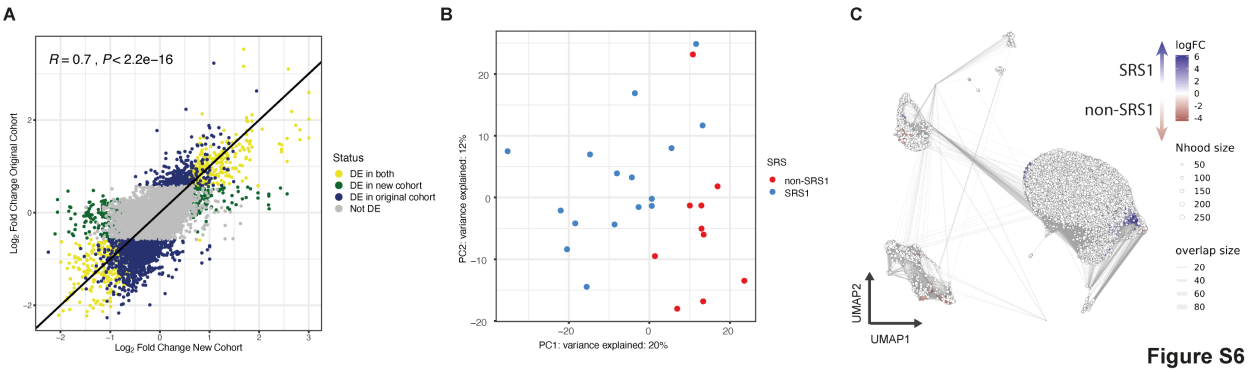

Figure S6. Cohort 1 SRS assignment and DA analysis. (A) Correlation of fold change in gene expression between SRS1 vs. non-SRS1 in cohort 1 and cohort in which SRS was originally derived (*2*). **(B)** First two PCs from PCA of pseudobulked single cell expression profiles of sepsis samples (all cells per sample). **(C)** UMAP of sampled neighborhoods from DA analysis across SRS with statistically significant enrichment colored (spatial FDR < 0.05 with generalised linear modeling).

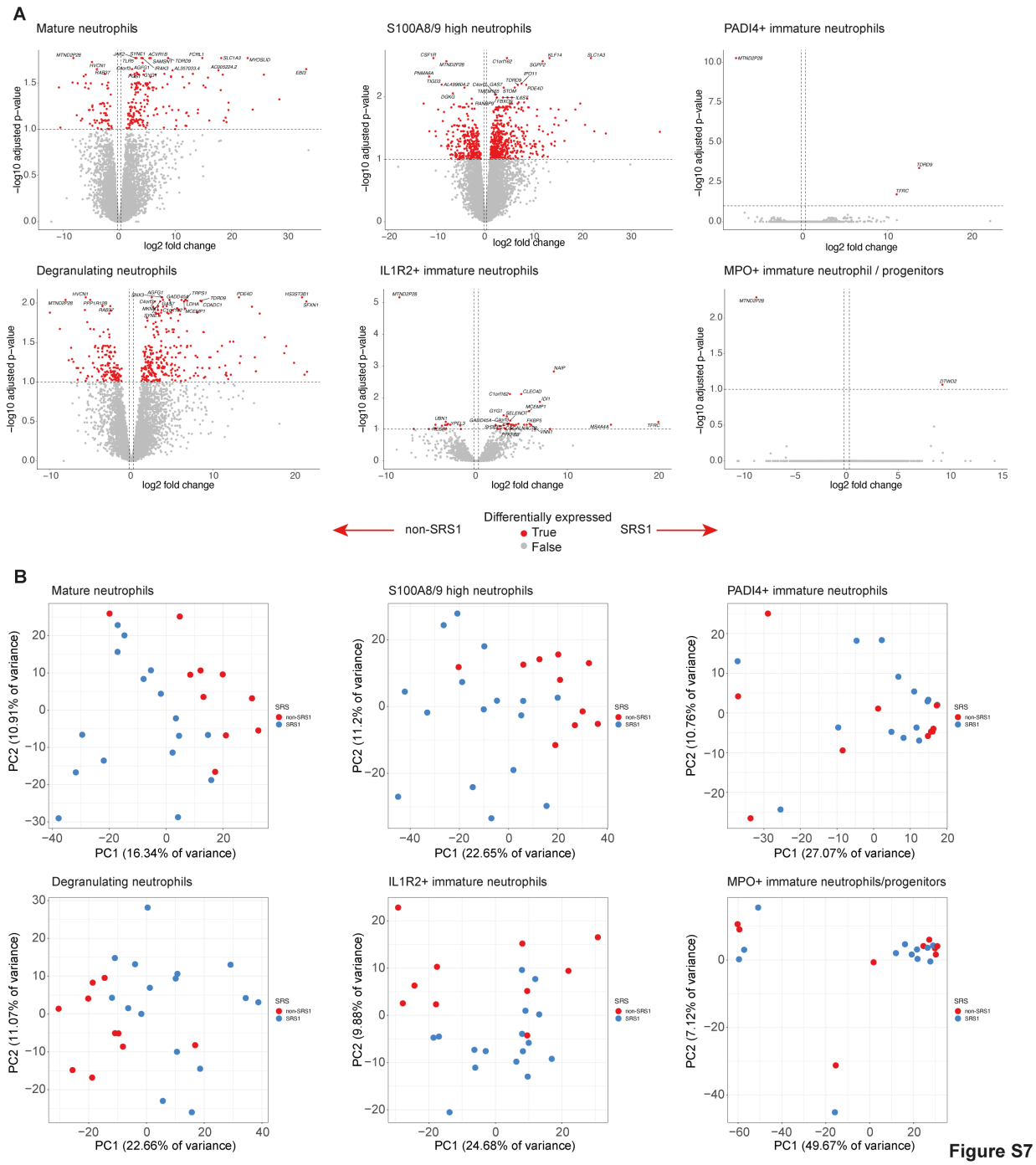

Figure S7

Figure S7. SRS neutrophil pseudobulk gene expression analysis. (A) Volcano plots of DGE analysis between SRS groups for pseudobulked neutrophil states (red denoting genes with fold change > 1.2 and FDR < 0.1) (positive fold change denoting upregulation in SRS1). **(B)** First two PCs from PCAs of pseudobulked neutrophil states colored by SRS groups.

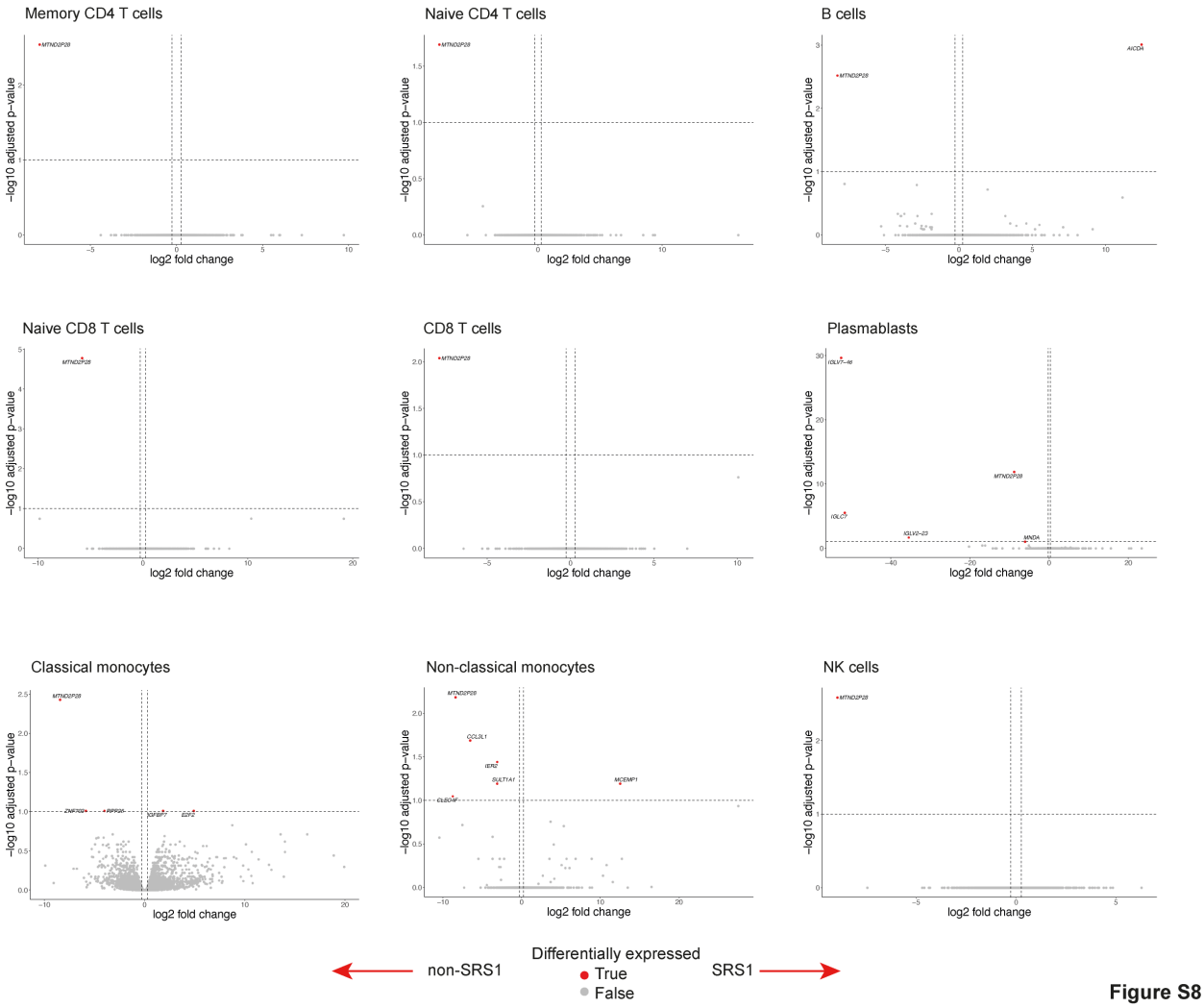

Figure S8

Figure S8. SRS mononuclear cell pseudobulk gene expression analysis. Volcano plots of DGE analysis between SRS groups for pseudobulked mononuclear cells (red denoting genes with fold change > 1.2 and FDR < 0.1) (positive fold change denoting upregulation in SRS1).

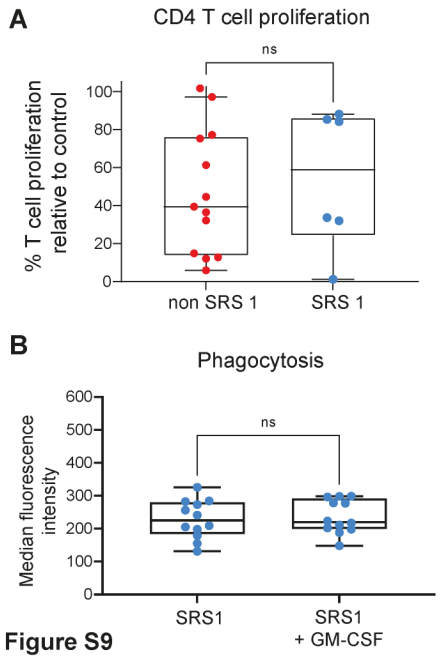

Figure S9. Functional interrogation of neutrophils across SRS. (A) Boxplot of proliferative fraction of anti-CD3/28 bead stimulated CD4 T cells after co-culture with neutrophils at a 4 neutrophil:1 T cell ratio, compared to positive controls of CD4 T cells cultured with anti-CD3/28 beads alone. **(B)** Neutrophil phagocytosis with and without GM-CSF stimulation. *P*-values were calculated with two-sided Wilcoxon rank-sum tests. *ns = not significant.*

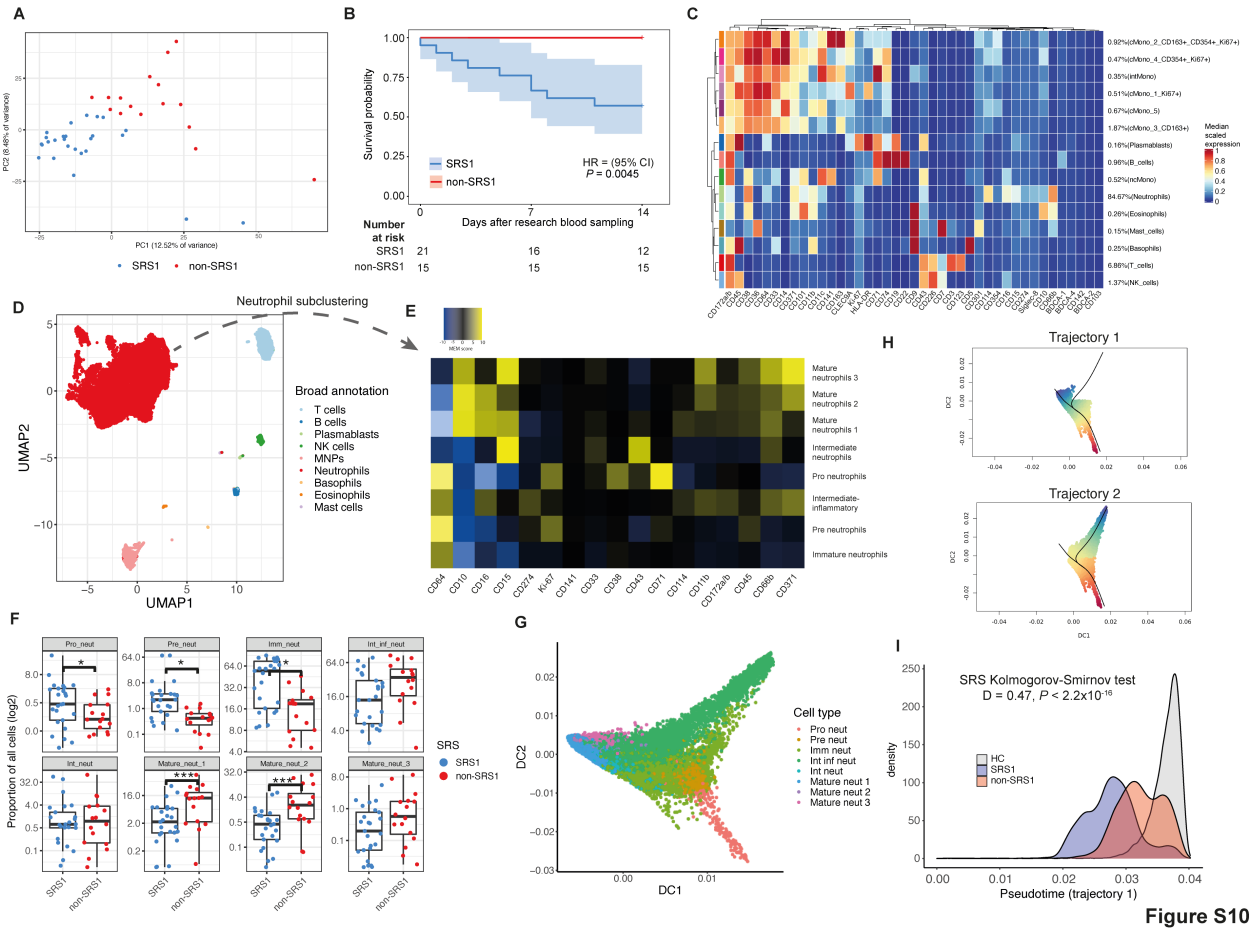

Figure S10. Cohort 2 SRS and CyTOF analysis. (A) First two PCs from PCA of n = 42 sepsis patient samples (36 individuals) for peripheral blood bulk RNA-seq (23,063 genes). **(B)** Kaplan-Meier survival by SRS (log-rank test *P*-value, 95% confidence intervals). **(C)** Heatmap of 15 immune cell clusters (with neutrophil clusters merged; Methods) from clustering on 41 protein markers for n = 41 sepsis (36 individuals) and 11 healthy control samples. **(D)** UMAP of nine broad immune cell types (2,287,410 cells) from CyTOF immunophenotyping. **(E)** Marker enrichment modeling (MEM) score heatmap of eight neutrophil clusters (1,921,471 cells), identified from neutrophil subclustering on 17 selected protein markers. **(F)** Neutrophil cluster frequency as a proportion of all cells with proportional differences across SRS analysed by generalised linear mixed models (n = 41 samples, 36 individuals). **(G)** Diffusion map dimension reduction of neutrophils. **(H)** Diffusion components (DCs) 1 and 2 with black lines depicting two

trajectories identified by principal curve fitting. Color reflects pseudotime value of cell with red indicating early pseudotime and blue late pseudotime. **(I)** Density distribution of neutrophils over pseudotime with Kolmogorov-Smirnov test for SRS1 vs. non-SRS1 distributions. *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001. *cMono = classical monocyte, ncMono = non-classical monocyte, NK cell = natural killer cell, pro-neut = pro-neutrophil, Pre-neut = pre-neutrophil, imm neut = immature neutrophil, int neut = intermediate neutrophil, int inf neut = intermediate inflammatory neutrophil, mature neut = mature neutrophil, DC = diffusion component.*

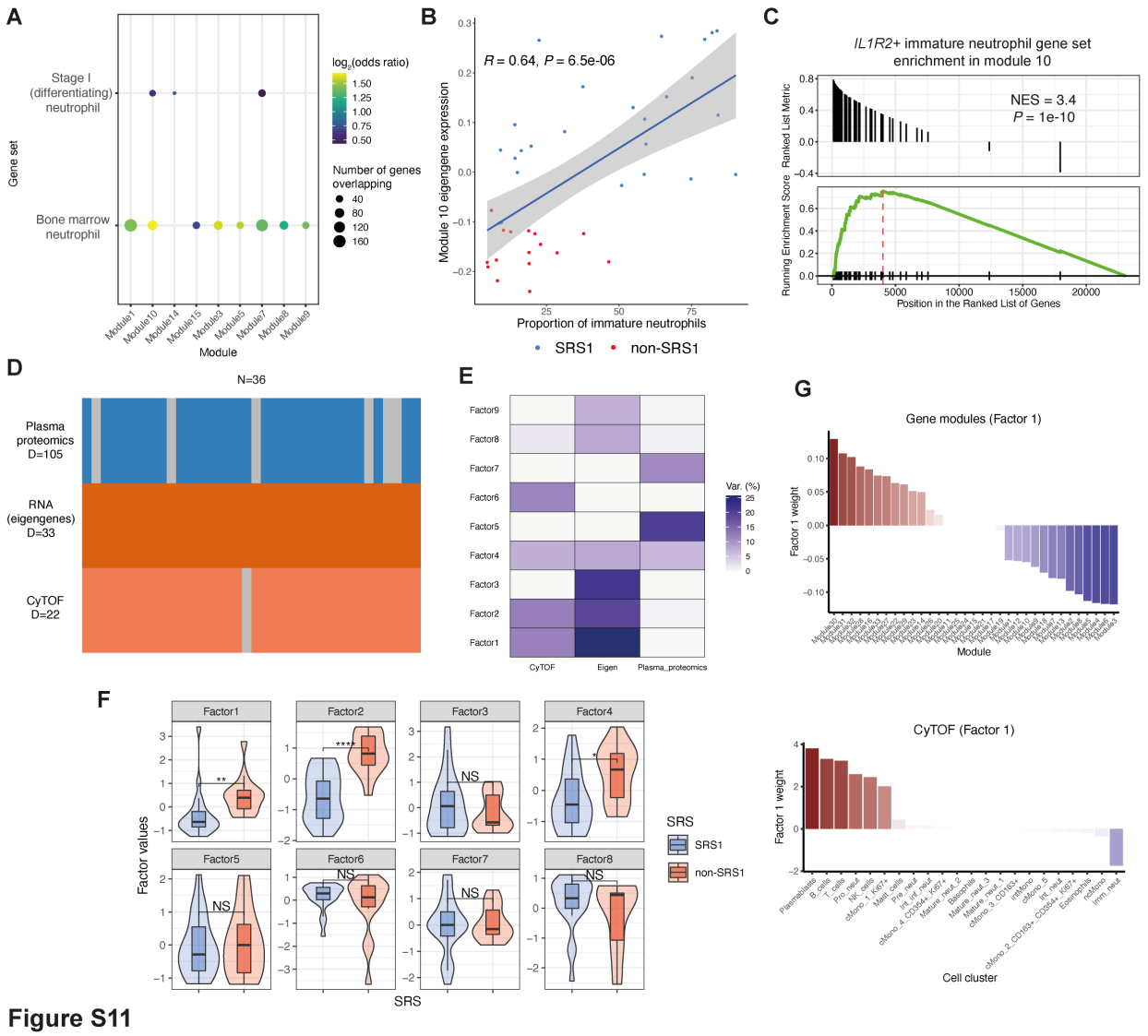

Figure S11

Figure S11. Cohort 2 validation of SRS1 *IL1R2*+ immature neutrophil expansion and factor analysis. (A) Enrichment of neutrophil related gene sets of weighted gene co-expression network analysis (WGCNA) modules upregulated in SRS1. **(B)** Correlation of module 10 eigengene expression against proportion of immature neutrophils from CyTOF per sample of cohort 2. **(C)** Leading edge analysis for enrichment of cohort 1 derived *IL1R2*+ immature neutrophil defining signature in WGCNA module 10. **(D)** Input data into MOFA+ model consisting of three modalities (105 plasma proteins, 33 module eigengenes from WGCNA and 22 cell clusters from CyTOF) for n = 36 sepsis samples with one and six missing samples for CyTOF and plasma proteomics

respectively. **(E)** Variance decomposition showing percentage variance explained for individual modalities by each latent factor. **(F)** Factor values by SRS groups (two-sided Wilcoxon rank sum-test). **(G)** Weights per gene module (top) and cell cluster (bottom) for latent factor one. *FDR < 0.05, **FDR < 0.01, *** FDR < 0.001, **** FDR < 0.0001. *FDR = false discovery rate, NES = normalized enrichment score*, *cMono = classical monocyte, ncMono = non-classical monocyte, NK cell = natural killer cell, pro-neut = pro-neutrophil, Pre-neut = pre-neutrophil, imm neut = immature neutrophil, int neut = intermediate neutrophil, int inf neut = intermediate inflammatory neutrophil, mature neut = mature neutrophil.*

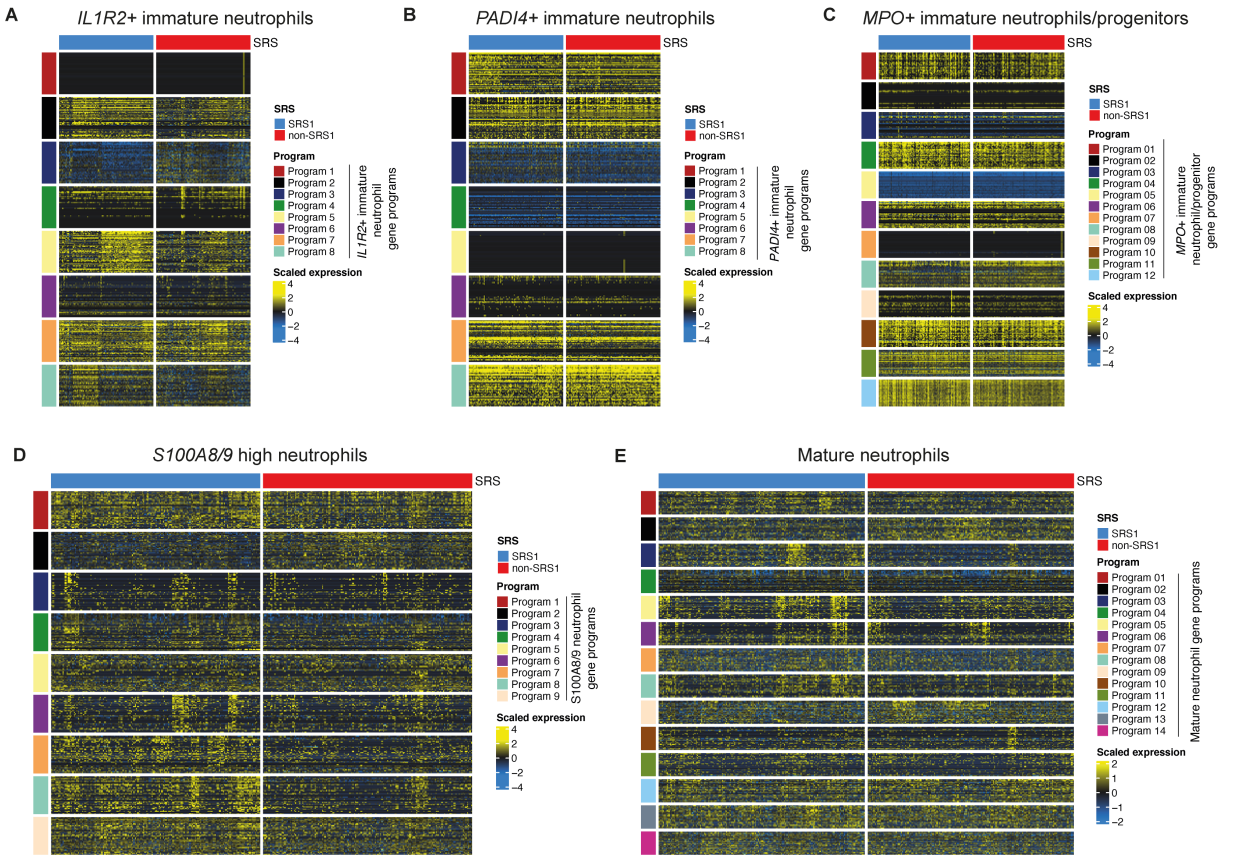

Figure S12

Figure S12. SRS neutrophil consensus non-negative matrix factorization (cNMF) analysis. **(A-E)** Heatmaps of gene programs from cNMF of **(A)** *IL1R2*+ immature neutrophils **(B)** *PADI4*+ immature neutrophils **(C)** *MPO*+ immature neutrophils/progenitors **(D)** *S100A8/9* high neutrophils and **(E)** mature neutrophils. Each heatmap shows 150 downsampled SRS1 and non-SRS1 cells. Each program neutrophil state was analyzed independently i.e. programs are not shared between each neutrophil state.

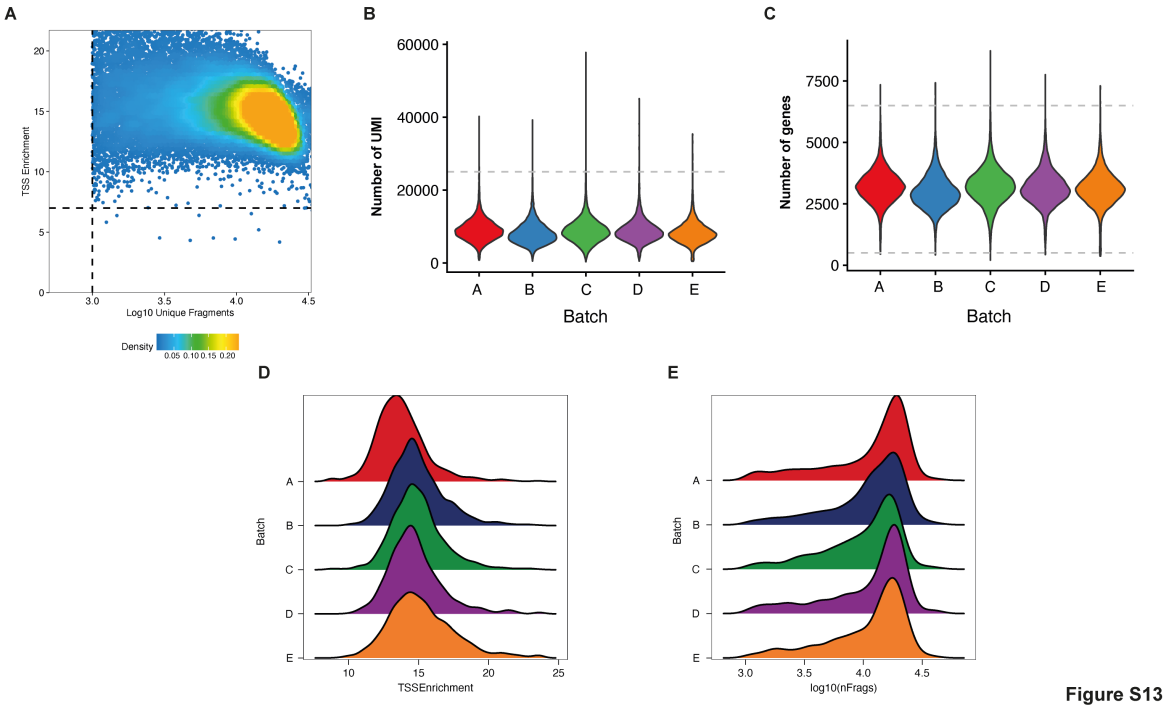

Figure S13

Figure S13. CD34+ hematopoietic stem and progenitor cell single cell RNA and ATAC-seq quality control (QC). (A) Summary scatterplot of TSS enrichment vs. unique fragment profiles pre-QC with filtering thresholds shown. **(B-C)** QC violin plots with RNA based filtering thresholds (grey dotted lines) for **(B)** number of UMI per cell and **(C)** number of genes detected per cell. **(D-E)** Ridge plots per batch for ATAC-seq data post QC for **(D)** TSS enrichment and **(E)** number of unique fragments detected.

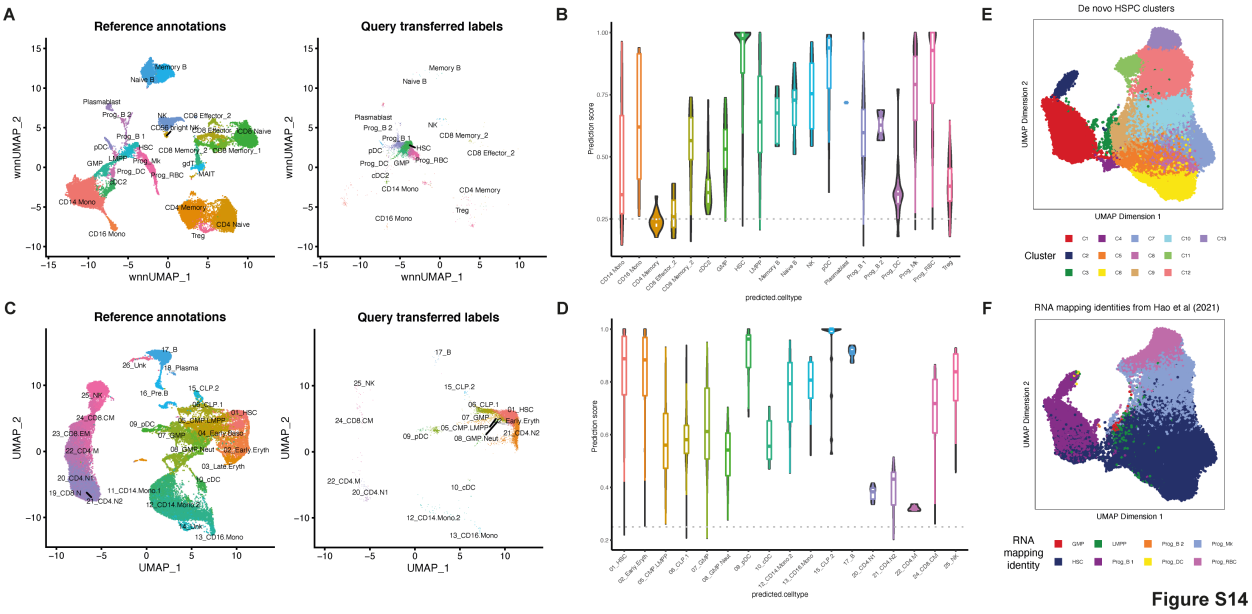

Figure S14

Figure S14. Hematopoietic stem and progenitor cell (HSPC) RNA based identity mapping and hematopoietic stem cell filtering. (A) HSPC (right) mapping to Seurat bone marrow mononuclear cell (BMMC) dataset (*51*) (left) and **(B)** mapping QC scores. **(C)** HSPC (right) mapping to healthy donor BMMC dataset from Granja et al (*52*) (left) and **(D)** mapping QC scores. **(E)** UMAP of Harmony batch corrected RNA and chromatin reduced dimensions after non-HSCs were filtered out (n = 46,156 HSPCs; Methods). **(F)** UMAP colored by RNA prediction identities from HSPC mapping to Seurat BMMC dataset. *c/pDC = conventional/plasmacytoid dendritic cell, GMP = granulocyte monocyte progenitor, LMPP = lymphoid-primed multipotential progenitors, NK = natural killer, Prog = progenitor, Mk = megakaryocyte, RBC = red blood cell, Eryth = erythrocyte, CMP = common myeloid progenitor, Mono = monocyte, CD4 N1/2 = CD4 naïve 1/2, CD4 M = memory CD4, CD8 CM = central memory CD8.*

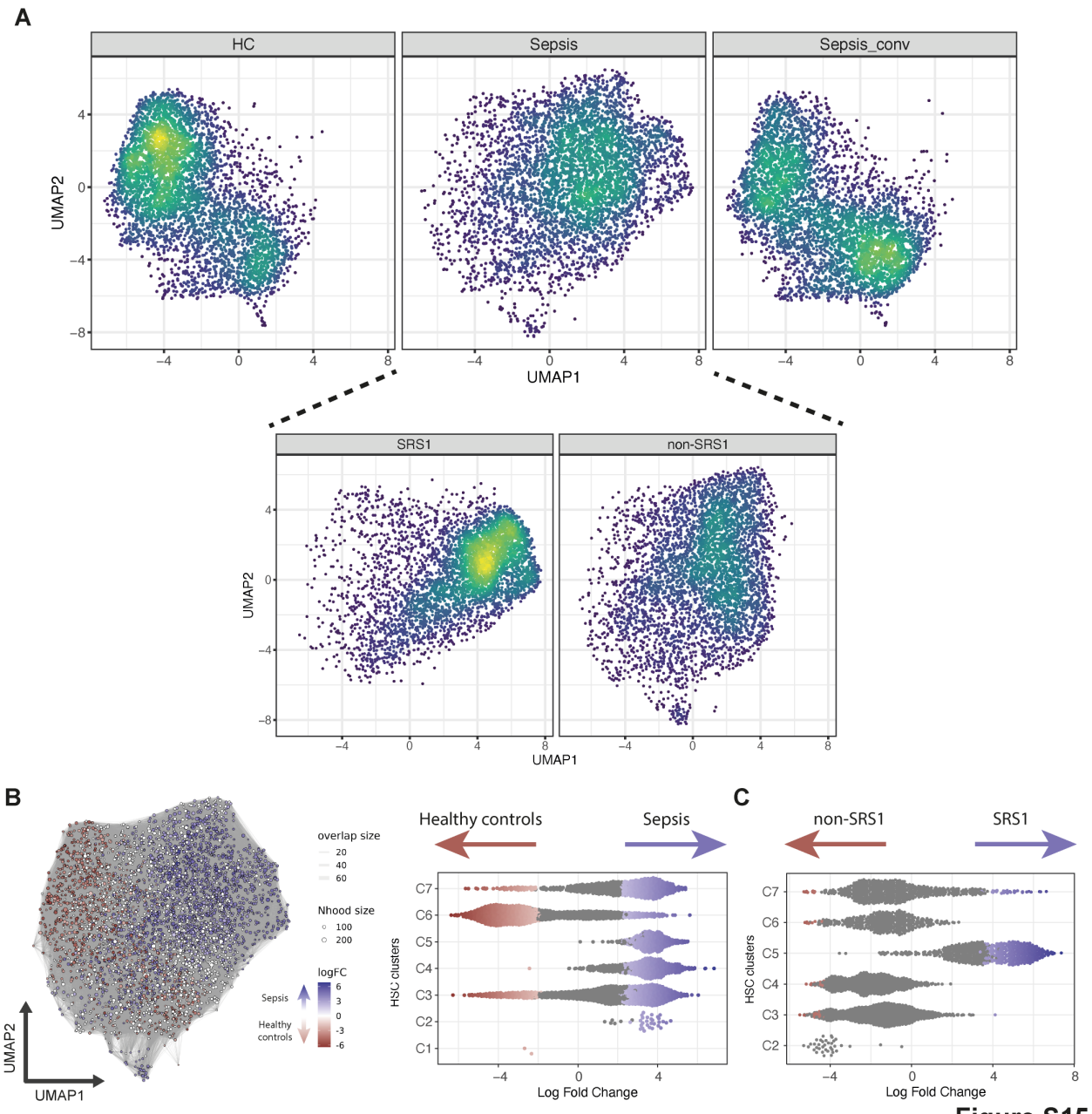

Figure S15

Figure S15. Hematopoietic stem cell (HSC) differential abundance (DA) analysis. (A) Density-based UMAP of HSC distributions by comparator group. **(B)** DA for healthy control-sepsis with UMAP showing sampled neighborhoods colored by statistically significant enrichment (spatial FDR < 0.05) (left) and beeswarm plot with HSC cluster labels (right). **(C)** Beeswarm plot for DA across SRS.

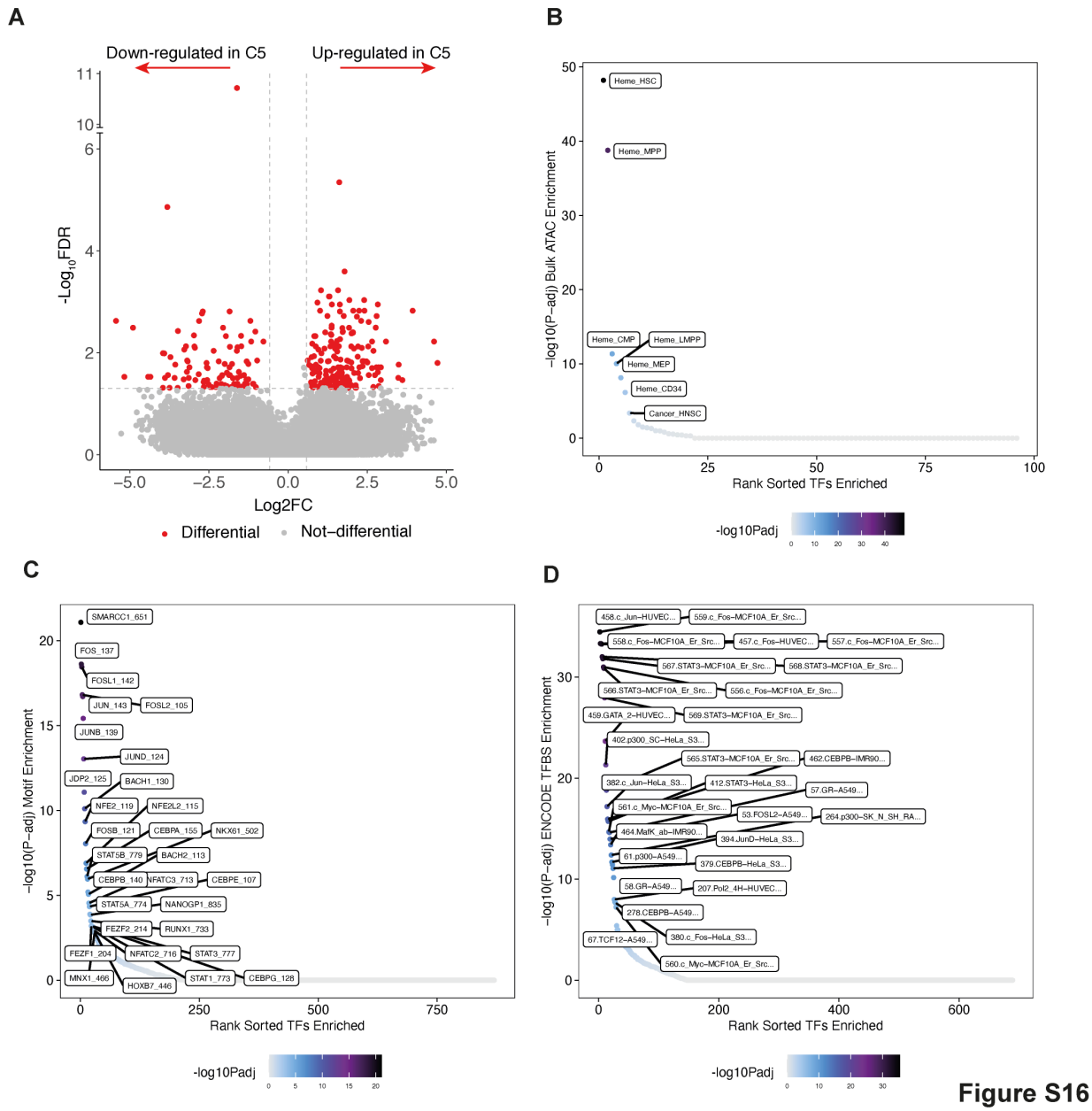

Figure S16

Figure S16. Hematopoietic stem cell (HSC) cluster C5 differentially open peak set overlap analysis. (A) Volcano plot of differentially accessible (DA) peaks (fold change > 1.5, FDR < 0.05) for HSC cluster C5 vs. all other clusters (two-sided Wilcoxon rank-sum test). **(B)** Enrichment of bulk ATAC-seq profiles with 184 peaks differentially open in HSC cluster C5. **(C)** Enrichment of transcription factor (TF) motifs of C5 DA peaks. **(D)** Enrichment of ENCODE TF binding site (TFBS) data for C5 DA peaks.

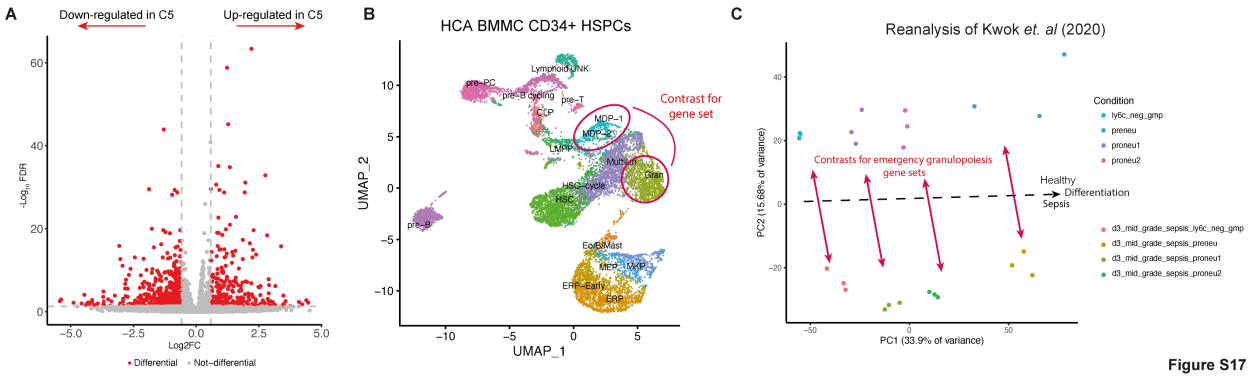

Figure S17. Hematopoietic stem cell (HSC) gene expression and granulopoiesis gene set analysis. (A) Volcano plot of differentially expressed genes at fold change > 1.5 and FDR < 0.05 of cluster C5 of HSCs vs. all other HSC clusters (two-sided Wilcoxon rank-sum test). **(B)** UMAP of CD34+ HSPCs from annotated Human Cell Atlas BMMC scRNA-seq data (*40*). **(C)** PCA of bulk RNA-seq profiles from granulocyte/myeloid progenitors of septic and healthy mice from Kwok et al (*22*).

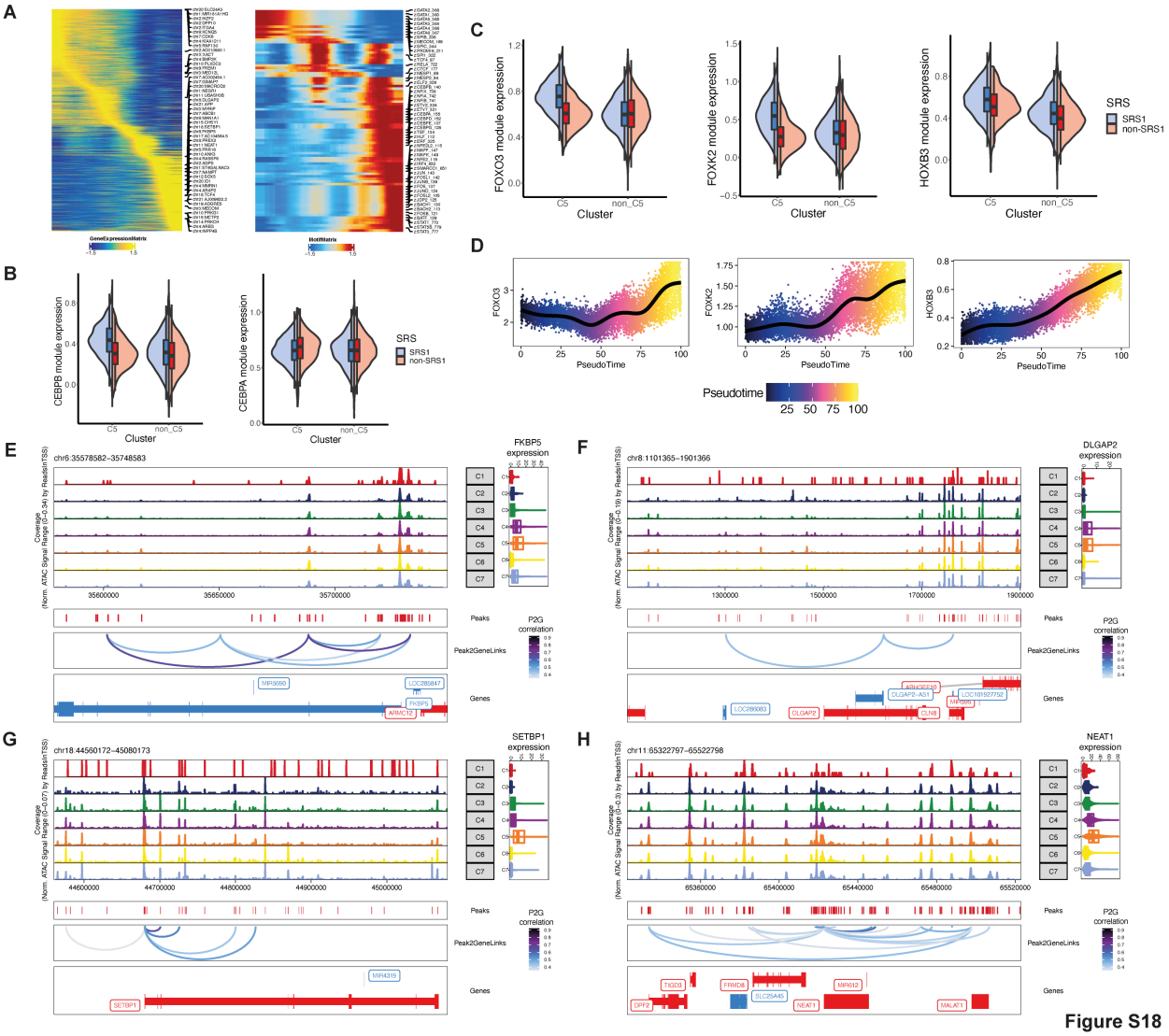

Figure S18

Figure S18. Hematopoietic stem cell (HSC) trajectory analysis, transcription factor (TF) modules from integrative scRNA scATAC analysis and putative emergency granulopoiesis TF regulators and target genes. (A) Heatmaps of gene expression (left) and TF motifs (right) across pseudotime. **(B)** Expression of CEBPB (left) and CEBPA (right) modules. **(C)** Expression of FOXO3 (left), FOXK2 (middle) and HOXB3 (right) modules. **(D)** FOXO3 (left), FOXK2 (middle) and HOXB3 (right) gene expression with pseudotime. **(E-H)** Browser tracks for **(E)** *FKBP5* **(F)** *DLGAP2* **(G)** *SETBP1* and **(H)** *NEAT1*. Top, aggregated scATAC profiles per HSC

cluster. Right, expression of relevant gene. Middle, peaks called and peak-to-gene links. Bottom, gene locus.

Supplementary tables

Table S1. Clinical data for cohorts 1 and 2.

Table S2. Antibodies used for cell surface protein profiling for cohort 1 on BD Rhapsody platform.

Table S3. Gene markers for each annotated cell type/state in cohort 1 single cell profiling and top 50 genes of each program for neutrophil states from cohort 1 neutrophil consensus non-negative matrix factorization.

Table S4. Differentially accessible peaks and differentially expressed genes (fold change > 1.5, FDR < 0.05) for hematopoietic stem cell cluster C5.

Table S5. Transcription factors (TF) and regulated genes ("TF modules") identified from gene regulatory network construction by modeling gene expression as a function of TF expression and regulatory region accessibility.