Cell Reports, Volume *42*

Supplemental information

Systemic alterations in neutrophils and their precursors in early-stage chronic obstructive pulmonary disease

Theodore S. Kapellos, Kevin Baßler, Wataru Fujii, Christina Nalkurthi, Anna C. Schaar, Lorenzo Bonaguro, Tal Pecht, Izabela Galvao, Shobhit Agrawal, Adem Saglam, Erica Dudkin, Amit Frishberg, Elena de Domenico, Arik Horne, Chantal Donovan, Richard Y. Kim, David Gallego-Ortega, Tessa E. Gillett, Meshal Ansari, Jonas Schulte-Schrepping, Nina Offermann, Ignazio Antignano, Burcu Sivri, Wenying Lu, Mathew S. Eapen, Martina van Uelft, Collins Osei-Sarpong, Maarten van den Berge, Hylke C. Donker, Harry J.M. Groen, Sukhwinder S. Sohal, Johanna Klein, Tina Schreiber, Andreas Feißt, Ali Önder Yildirim, Herbert B. Schiller, Martijn C. Nawijn, Matthias Becker, Kristian Händler, Marc Beyer, Melania Capasso, Thomas Ulas, Jan Hasenauer, Carmen Pizarro, Fabian J. Theis, Philip M. Hansbro, Dirk Skowasch, and Joachim L. Schultze

SUPPLEMENTAL FIGURES

Fig. S1 – Related to Fig. 1. (A-B) Flow cytometry gating strategy for (A) myeloid and (B) lymphoid immune cells in the peripheral blood of control and COPD patients.

Fig. S2 – Related to Fig. 2. Analysis of peripheral blood compartment of control and COPD patients. (A) UMAP representation of 69,199 blood cells from 6 control and 8 COPD patients depicting the clustering result. (B) Dot plot of canonical gene markers for cell types found in human blood expressed in the single-cell RNA sequencing (scRNA-seq) clusters. Circle size represents the percentage of cells within a cluster that express a particular gene, circle color shows average scaled normalized expression within the cluster. (C) Heatmap of the proportion of cells within each cluster classified using the GenSigPro classifier labels [12]. (D) Network of the top 5 enriched Gene Ontology gene sets in control or COPD patients. Red arrows depict upregulated terms in COPD, blue arrows downregulated in COPD.

Fig. S3 – Related to Fig. 2. Characterization of peripheral blood neutrophil states in control and COPD patients. (A) Proportion of blood neutrophil states in each control and COPD patient. (B) Number of unique molecular identifiers (UMIs) for each blood neutrophil state grouped by disease status. (C) Number of genes for each blood neutrophil state grouped by disease status. (D-H) Gene set enrichment analysis of neutrophil state-specific markers using the Reactome gene sets. Enriched gene sets with less than 3 genes were discarded. (I) Trajectory inference for the identified blood neutrophil states. N1S/*LCN2*

neutrophils were selected as the start of the developmental pathway. Black lines connect developmentally similar populations. (J) Bar plots of blood neutrophil state frequencies in control and COPD patients. Data are represented as mean ± SD and analysis was carried out with a Wilcoxon test for non-normally distributed data, * $p<0.05$.

Fig. S4 – Related to Fig. 3. Characterization of peripheral blood neutrophil states in control and COPD patients (validation cohort). (A-B) UMAP representation of 9,269 whole blood cells from 3 control and 3 COPD patients depicting (A) cluster identity and (B) cluster annotation. (C) Dot plot of the top 5 differentially expressed (DE) genes for each blood cell cluster against the rest in 3 control and 3 COPD patients. Circle size represents the percentage of cells within a cluster that express a particular gene, circle color shows average scaled normalized gene expression within the cluster. (D) Proportion of blood neutrophil states in each control and COPD patient. (E) Enrichment of the top 20 unique genes for each Seq-well blood neutrophil state on Rhapsody blood neutrophil transcriptomes. (F) Table of the correspondence of blood neutrophil states between the test (Seq-well) and validation (BD Rhapsody) cohorts. (G) Violin plots of DE genes between control and COPD patients for blood neutrophil states.

Fig. S5 – Related to Fig. 4. Characterization of bronchoalveolar fluid neutrophil states in control and COPD patients. (A) Proportion of BALF neutrophil states in each control and COPD patient. (B) Number of unique molecular identifiers (UMIs) for each BALF neutrophil state grouped by disease status. (C) Number of genes for each BALF neutrophil state grouped by disease status. (D-F) Gene set enrichment analysis of

neutrophil state-specific markers using the Reactome gene sets. Enriched gene sets with less than 3 genes were discarded. (G) Enrichment of the top 20 unique genes for each BALF neutrophil state on the blood neutrophil transcriptomes from Fig. 2. (H) Heatmap of the top 5 marker genes for the blood neutrophil states from Fig. 2. Each column represents the scaled average normalized expression per patient.

Fig. S6 – Related to Fig. 4. Imaging mass cytometry analysis of bronchial biopsies from control, normal lung function smokers and COPD patients. (A) Gating strategy for the annotation of neutrophils states, monocytes, macrophages, T, B, NK and epithelial cells in a representative lung biopsy. (B-D) Spatial allocation of detected cell types for each (B) control, (C) smokers and (D) COPD lung biopsy replicate. Axes represent x and y dimensions. (E) Neutrophil state relative proportions as percentages of $CD45^+$ cells in controls, smokers and COPD lung biopsies. Data are represented as mean ± SD and were analyzed with one-way ANOVA and Tukey's post-hoc corrections or the Dunn test for non-normally distributed data, * $p<0.05$, ** $p<0.01$. (F) Relative proportions of immune cell types as percentages of $CD45^+$ cells in control, smokers and COPD lung biopsies. Data are represented as mean ± SD and were analyzed with one-way ANOVA and Tukey's post-hoc corrections or the Dunn test for non-normally distributed data, * $p<0.05$, ** $p<0.01$.

Fig. S7 – Related to Fig. 5. Characterization of bronchoalveolar fluid neutrophil states from air and smoke-exposed mice. (A) Experimental design and sample processing. (B) Dot plot of top 5 differentially expressed (DE) genes for each identified BALF neutrophil

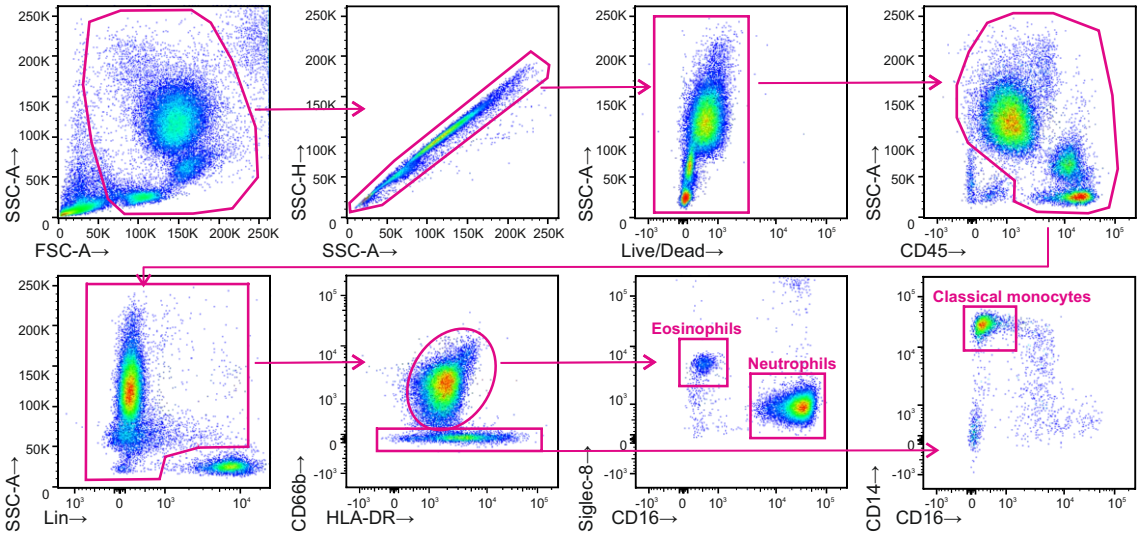
cluster against the rest. Circle size represents the percentage of cells within a cluster that express a particular gene, circle color shows average scaled normalized gene expression within the cluster. (C) Uniform Manifold Approximation and Projection (UMAP) representation of 18,406 bronchoalveolar lavage fluid (BALF) cells from 4 air and 4 cigarette smoke (CS)-exposed mice. (D) UMAP representation of 7,064 BALF neutrophils from 4 air and 4 CS-exposed mice. (E) Absolute neutrophil counts in the BALF of 4 air and 4 CS-exposed mice. (F) Bar plots of BALF neutrophil state frequencies in 4 air and 4 CS-exposed mice. (G) Heatmap of the top 20 unique murine BALF neutrophil state genes on the human BALF neutrophil states from Fig. 4. Murine genes were first converted to their human homologues. (H) Upset plot depicting the enriched gene sets (air vs smoke-exposed mice) between BALF neutrophil states. (I) UMAP representation of 1,677 lung neutrophils from 9 air and 15 CS-exposed mice for 2, 4 and 6 months (3 air and 5 CS-exposed animals per time point) from [45]. (J) Heatmap of the top 20 unique murine lung neutrophil state genes from [45] on this study's murine BALF neutrophil states. (K) Neutrophil state relative proportions as percentages of all neutrophils at each time point. All control animals were pooled together and were defined as time point 0. (L) UMAP representation colored by the pseudotime score of all neutrophils in the dataset. Arrow indicates the direction of the trajectory from time point 0 to 6 months. (M) Heatmap of the top 20 time-associated neutrophil genes for each cluster. Cells were ordered according to pseudotime.

Fig. S8 – Related to Fig. 6. Characterization of bone marrow neutrophil states from air and smoke-exposed mice. (A) Heatmap of mass cytometry (CyTOF) marker expression

in CD45^{+} bone marrow cells from air and cigarette smoke (CS)-exposed mice. Each column represents the average scaled normalized expression per cluster. (B-D) Expression levels of lineage-negative markers for neutrophils; (B) CD3, (C) B220 and (D) NK1.1. (E) Heatmap of the top 20 unique genes from the human bone marrow neutrophil populations from [47] on the bone marrow neutrophil states of this study. MPP; multipotent progenitor, CMP; common myeloid progenitor, GMP; granulocyte-monocyte progenitor. Black bars depict the correspondence of transcriptomic states to the phenotypically defined bone marrow neutrophil populations in Fig. 6B-D. (F) Dot plot of shared DE genes between human blood and BALF neutrophils and murine bone marrow early GMPs. Circle size represents the percentage of cells within a cluster that express a particular gene, circle color shows average scaled normalized gene expression within the cluster.

Figure S1 - related to Figure 1

A Myeloid panel



B Lymphoid panel

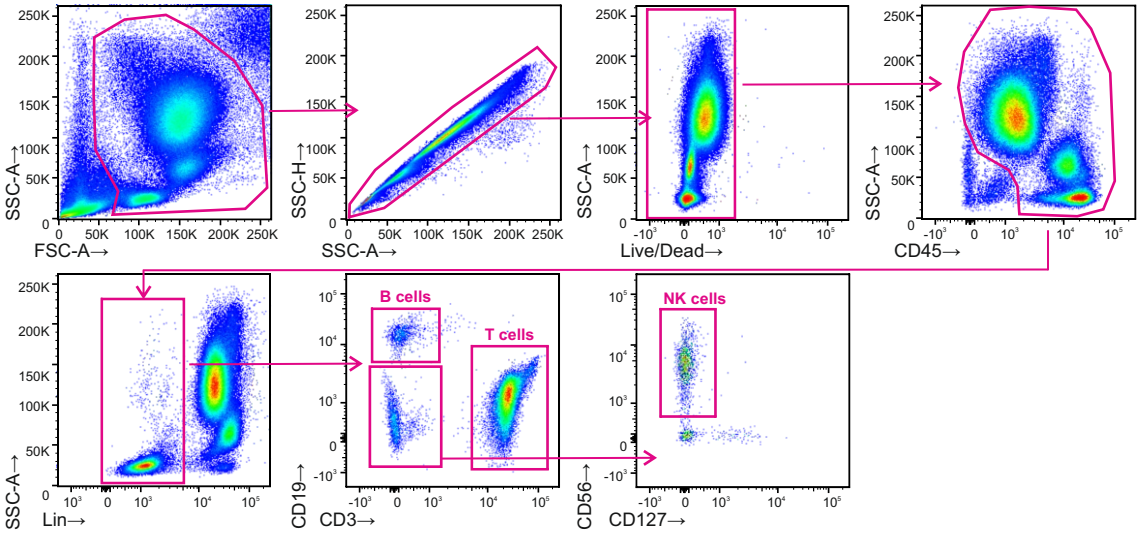


Figure S2 -related to Figure 2

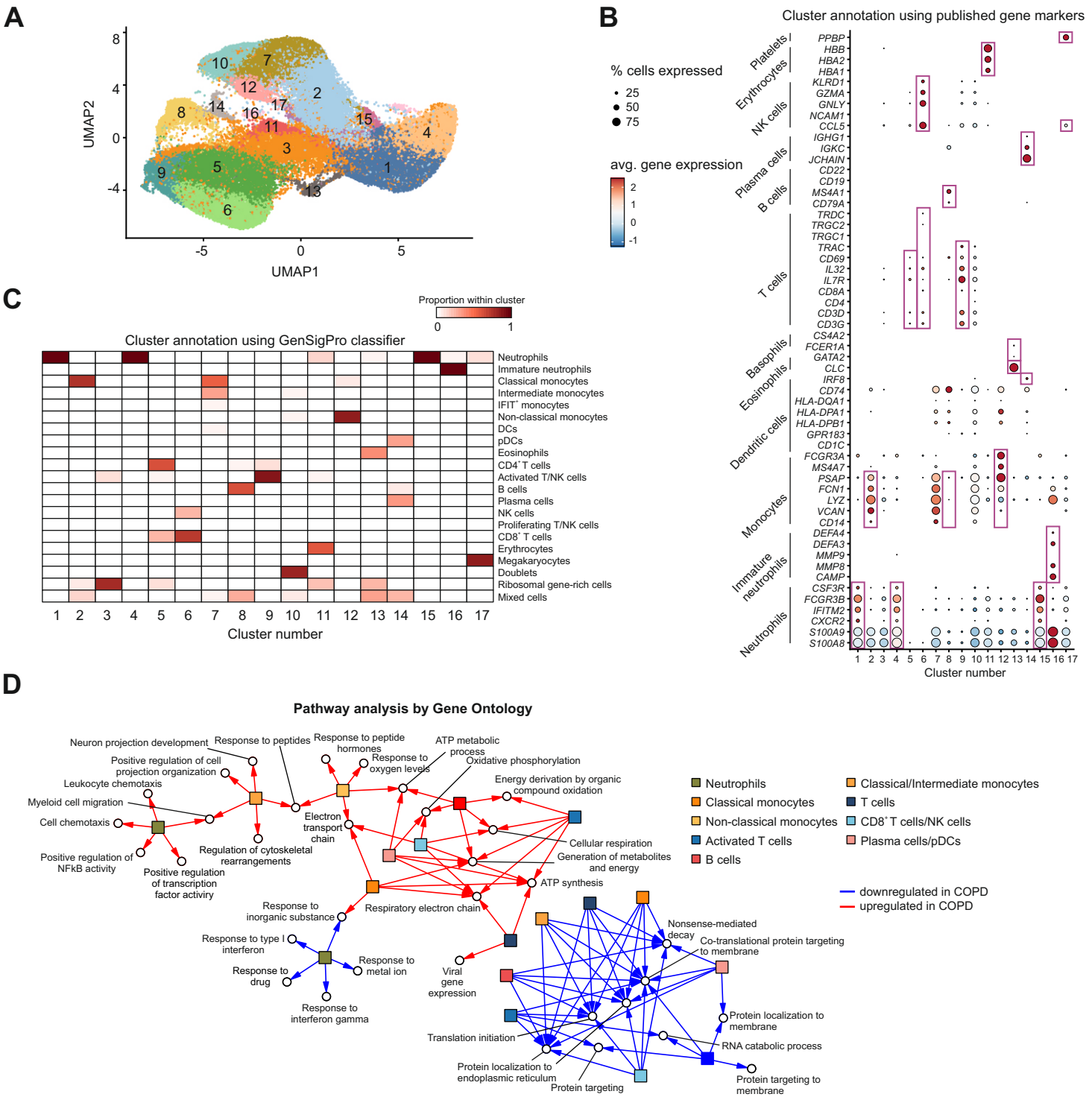


Figure S3 - related to Figure 2

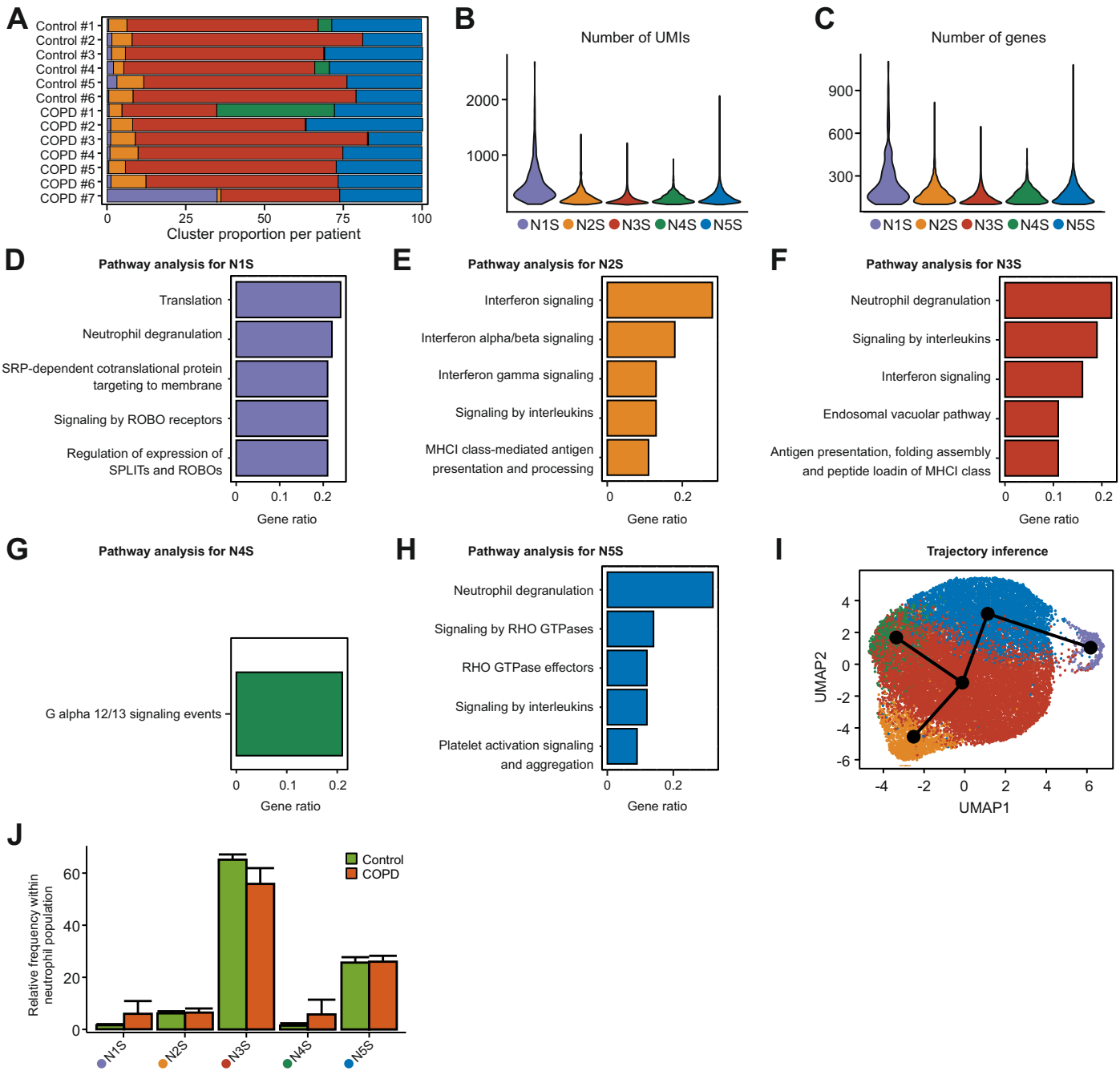


Figure S4 - related to Figure 3

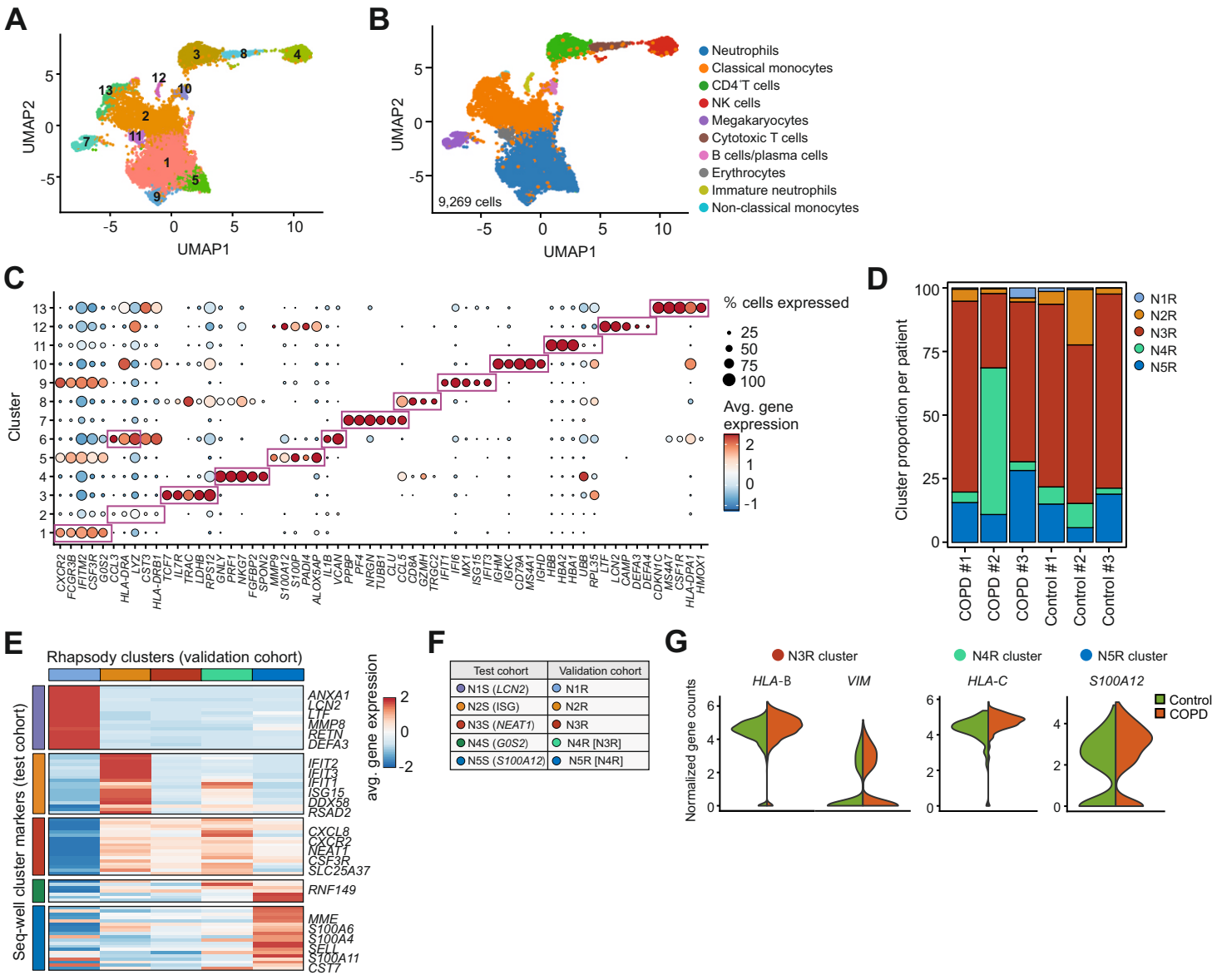


Figure S5 - related to Figure 4

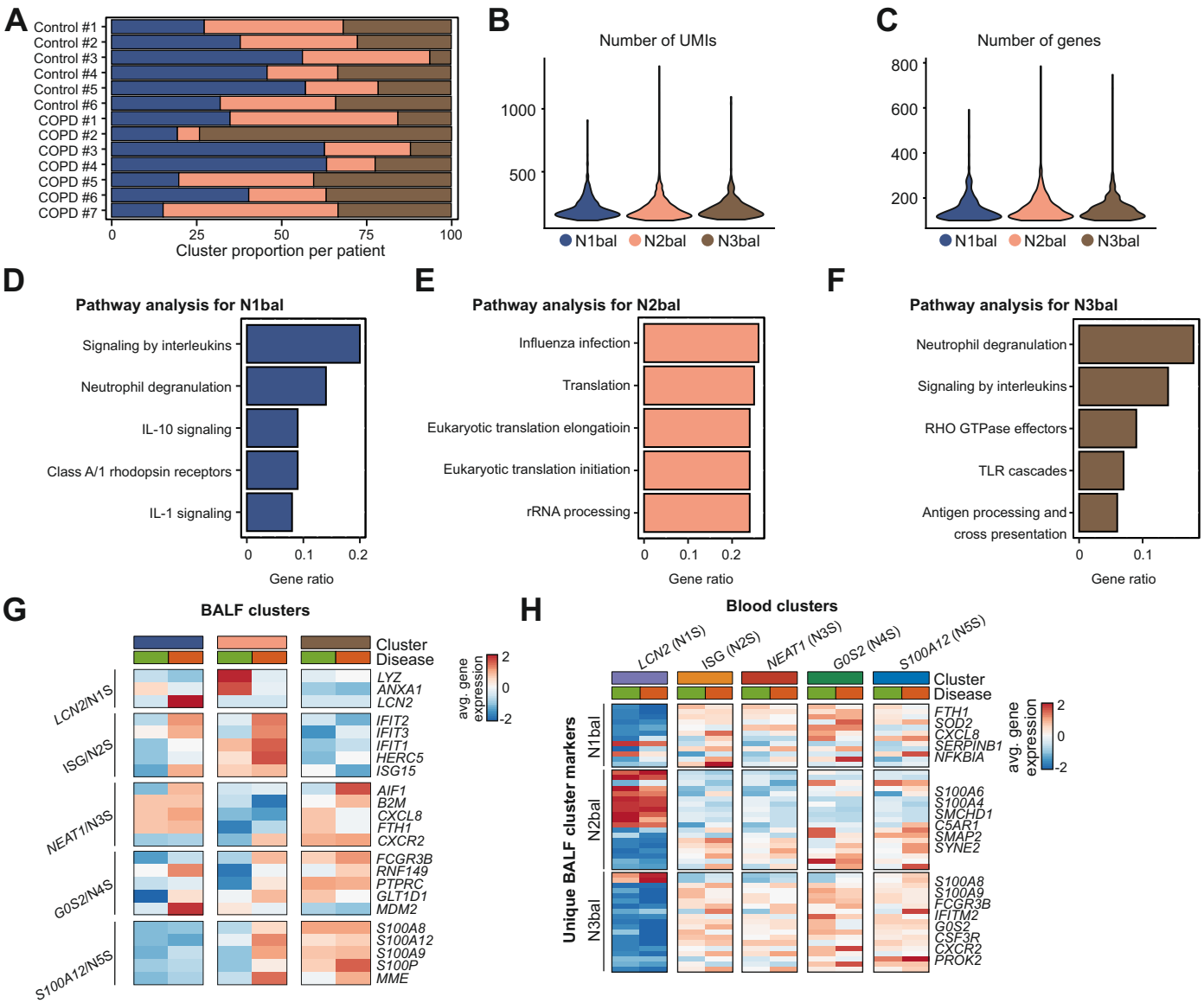


Figure S6 - related to Figure 4

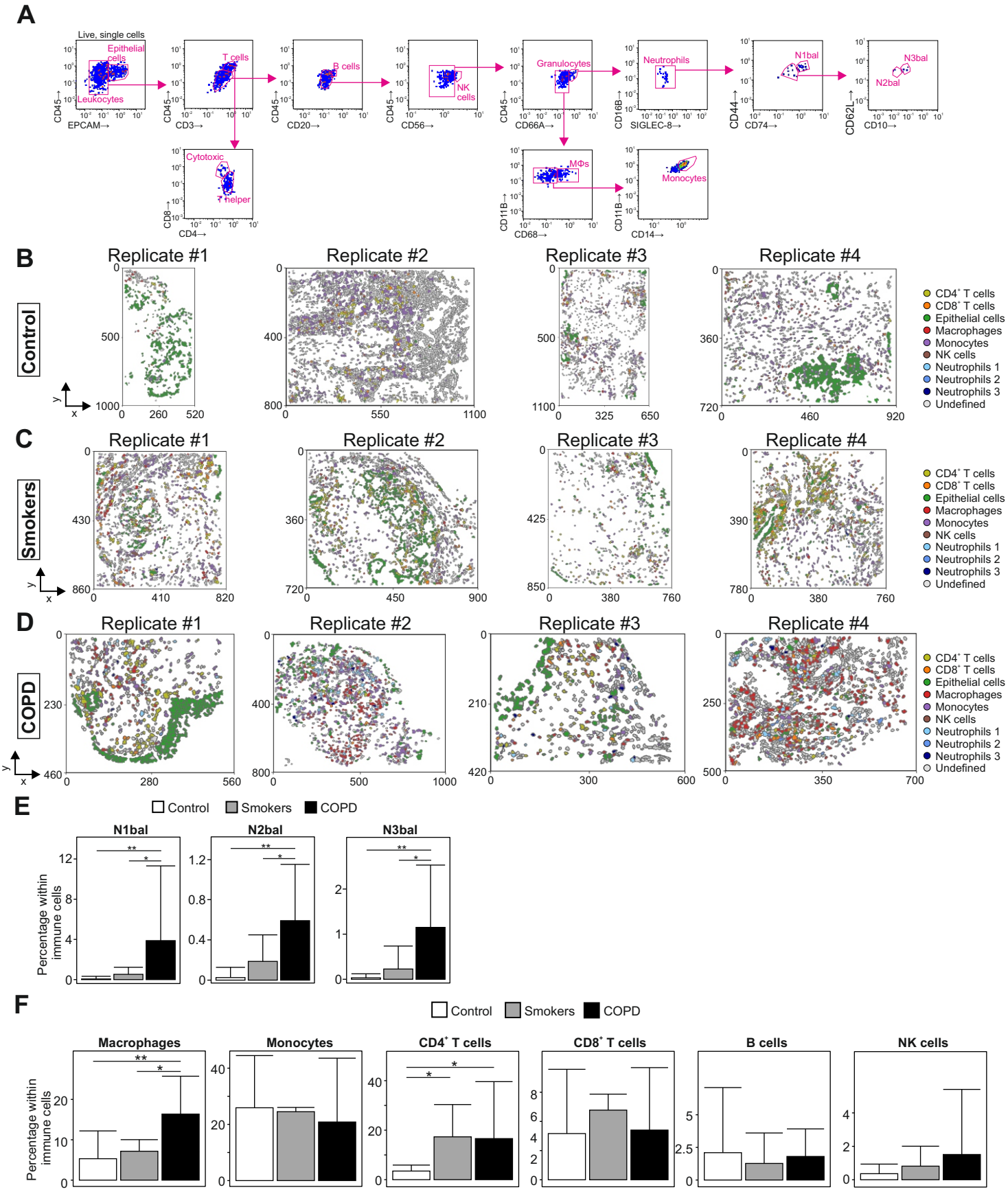


Figure S7 - related to Figure 5

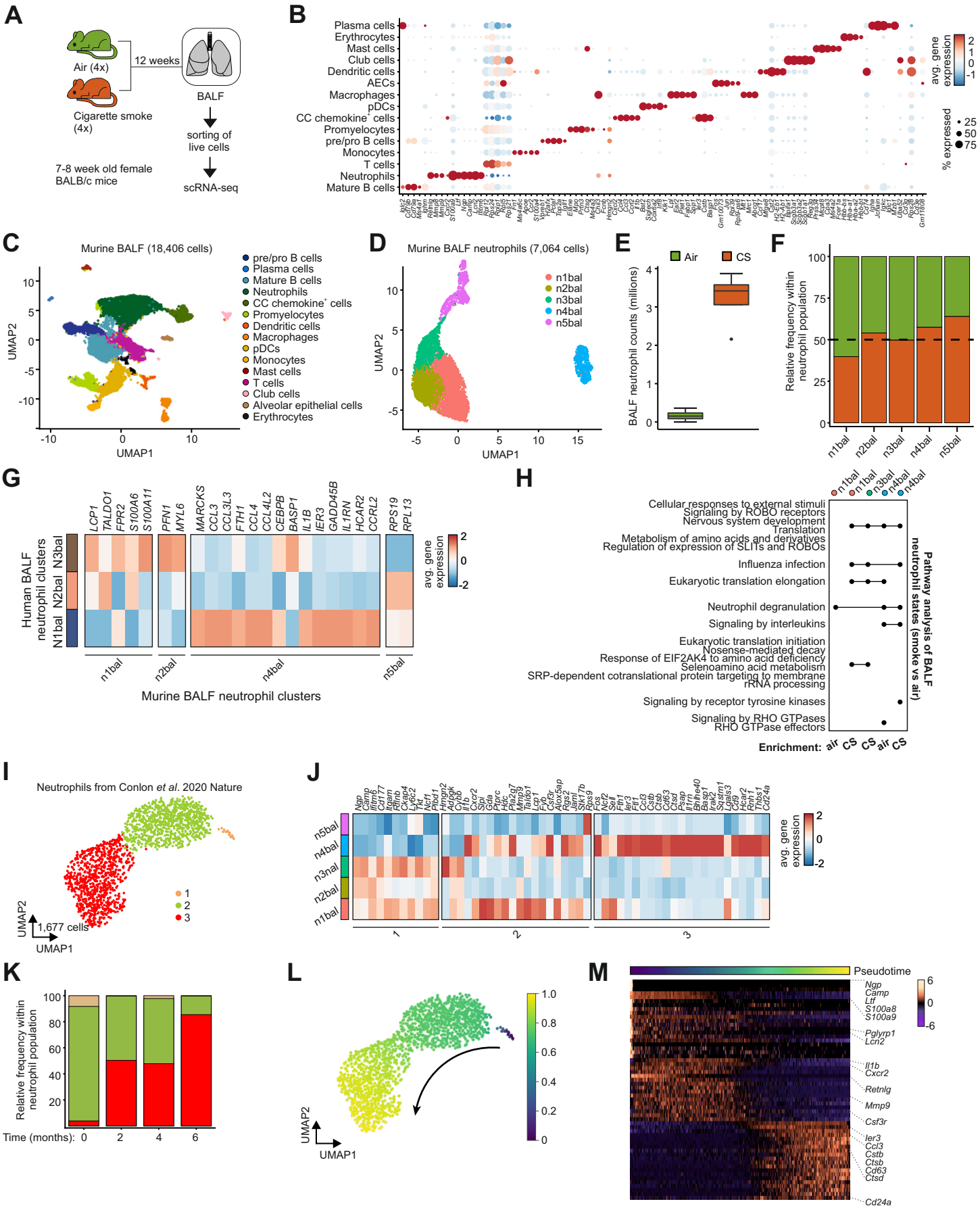


Figure S8 - related to Figure 6