# ScRNA-seq Expression of *APOC2* and *IFI27* Identifies Four Alveolar Macrophage Superclusters in Cystic Fibrosis and Healthy BALF

Xin Li[1], Fred W. Kolling[2], Daniel Aridgides[3], Diane Mellinger[3], Alix Ashare[1,3*], and Claudia V. Jakubzick[1*]

[1]Department of Microbiology and Immunology, Dartmouth Geisel School of Medicine, Hanover, NH [2]Department of Biomedical Data Science, Dartmouth Geisel School of Medicine, Hanover, NH [3]Department of Medicine, Dartmouth Hitchcock Medical Center, Lebanon, NH [*]Co-senior authors

Running title: Four Alveolar Macrophage Superclusters

Correspondences:

Claudia Jakubzick, PhD
Geisel School of Medicine at Dartmouth
Department of Microbiology and Immunology
626W Borwell
One Medical Center Drive
Lebanon, NH 03756
Phone: 303-898-3791
claudia.jakubzick@dartmouth.edu

Alix Ashare, MD, PhD
Geisel School of Medicine at Dartmouth
Dartmouth Hitchcock Medical Center
Department of Medicine
Department of Microbiology and Immunology
One Medical Center Drive
Lebanon, NH 03756
alix.ashare@hitchcock.org

Keywords: Cystic fibrosis, alveolar macrophages, monocytes, recruited macrophage, scRNA-seq, supercluster, macrophage taxonomy

## Abstract

**Rationale:** Alveolar macrophages (AMs) reside on the luminal surface of the airways and alveoli, ensuring proper gas exchange by ingesting cellular debris and pathogens, and regulating inflammatory responses. Recent studies have highlighted the heterogeneity of airspace macrophages from healthy bronchoalveolar lavage fluid and cystic fibrosis sputum. Understanding the heterogeneity and diverse roles played by AMs and recruited monocytes is critical for treating CF and other airway diseases.

**Objectives:** To identify and compare the cellular composition and functional diversity between CF and healthy airspace macrophages and monocytes.

**Methods:** We performed single-cell RNA sequencing (scRNA-seq) on 113,213 BAL cells from four healthy and three mild CF subjects. Transcriptional data and TotalSeq technology were used to confirm cell surface markers that distinguish resident AMs from recruited monocytes.

**Measurements and Main Results:**

Unbiased clustering identified four AM superclusters based on the expression of *APOC2* and *IFI27* genes. Each supercluster contains four shared subclusters named after their differentially expressed gene(s): IGF1.AMs, CCL18.AMs, CXCL5.AMs, and Cholesterol.AMs. AM Supercluster 4 contained two additional subclusters. Beyond the superclusters, several additional AM clusters were identified: Chemokine, Metallothionein, Interferon, and Cycling AMs. Interestingly, the Chemokine cluster further divides into four subclusters, each selectively expressing a unique combination of chemokines. Lastly, minimal differences were observed between mild CF and healthy individuals in airspace cellular composition.

**Conclusions:** ScRNA-seq identified four AM superclusters. Several fundamental questions require further investigation, including these clusters' locations, functions, transcription factors that regulate their distinct programming, and whether similar macrophage superclusters can be found in other organs.

## Introduction

Alveolar macrophages (AMs) are the primary phagocytes within the airspace of the lungs. In response to inflammatory stimuli, AMs secrete cytokines and chemokines to initiate the immune response. Subsequently, AMs and recruited monocytes contribute to the resolution of the inflammatory response to prevent excess damage (1-4). In addition, AMs play a major role in maintaining lipid homeostasis in the lung, which is essential for adequate gas exchange and alveolar epithelial integrity.

AMs may hold keys to understanding, preventing, and curing several diseases. They derive from prenatal monocytes, self-renew, and have shown remarkable plasticity by altering their transcriptome in response to the environmental changes (3, 5-8). Their transcriptional profile is altered in various inflammatory lung diseases, including asthma (2, 9), chronic obstructive pulmonary disease (10), cystic fibrosis (CF), and cancer (11-16). Therefore, enhancing our current understanding of the normal cellular composition and functional diversity of AM subtypes will help characterize disease-related transcriptional changes more precisely and assess targeting specifically a given AM subtype to restore immune balance.

Single-cell RNA sequencing (scRNA-seq) technology is critical for accomplishing these goals. AMs have traditionally been thought to be a uniform population of cells that can be activated by different disease states (5). However, recent scRNA-seq studies have revealed a rich diversity in AMs in bronchoalveolar lavage fluid (BALF) from healthy subjects, with multiple subpopulations that have yet to be characterized in CF and many other diseases (17). Another study investigated differences in immune population in sputum from CF subjects and found that sputum immune cells exhibited differential activation and maturity (18).

Our study was designed to investigate differences in the immune cell populations of BALF obtained from individuals with CF and healthy controls (HC). Using scRNA-seq, we found minor differences in cell populations between CF subjects with no significant lung disease and healthy controls, suggesting that normalizing the CF environment via treatments results in healthy-like airspace macrophages. Interestingly, unbiased analysis of this dataset identified four AM superclusters based on the expression of *APOC2* and *IFI27* genes, which have not been previously described. Each supercluster

contains four shared subclusters named after their differentially expressed genes (DEGs): IGF1.AMs, CCL18.AMs, CXCL5.AMs, and Cholesterol.AMs. Beyond the four AM superclusters, we found four additional clusters. The most striking was the Chemokine cluster, which further divides into four subclusters with each selectively expressing a distinct combination of chemokines. Hence, there appears to be a clear division of labor in leukocyte recruitment. Moreover, it also suggests that there are distinct transcription factors (based on our regulon analysis) that regulate the differentiation and programming of each Chemokine AM cluster. This pattern is analogous to what is found in T cell biology, where distinct transcription factors regulate the cell differentiation pathway to produce a defined combination of cytokines known as Th1, Th2, Th17, etc. Overall, our dataset opens many new areas for further investigation, including addressing how resident macrophage and monocyte subtypes are altered during disease and what factors regulate their programming.

## Methods

### Human Subjects

Four HC subjects and three subjects with CF underwent research bronchoscopy as previously described (19, 20). Briefly, after local anesthesia to the posterior pharynx and intravenous conscious sedation, a flexible fiberoptic bronchoscope was inserted transorally and passed through the vocal cords into the trachea. The bronchoscope was sequentially wedged into tertiary bronchi in the right upper lobe (RUL), right middle lobe (RML), and right lower lobe (RLL), saline was instilled, and BALF was collected. An aliquot of 10 mL of BALF from the RML was used for the studies outlined here. All subjects provided written informed consent, and this study was approved by the Institutional Review Board of Dartmouth-Hitchcock Medical Center (protocol #22781).

### Sample Processing

Sample processing and library construction were performed immediately after lavage. 1X phosphate-buffered saline (PBS) containing 0.5 mM EDTA and 0.1% bovine serum albumin (BSA) was added to the BALF. Samples were filtered through 100-micron cell strainers and centrifugated at 300 x *g* for five minutes at 4 °C, and supernatants were discarded. TotalSeq antibodies were added at concentrations recommended by the manufacturer and stained for 30 min on ice (21), details in online supplement. After staining, chilled Hank's balanced salt solution (HBSS) supplemented with 0.5% BSA was used to wash the sample. Samples were again filtered through 100-micron cell strainers and spun as above. Supernatants were aspirated, and pellets were resuspended with HBSS plus 0.5% BSA at an approximate concentration of $7.5 \times 10^5$ cells/mL. Cell quality and viability were assessed with a Cellometer K2 (Nexcelom Bioscience, Lawrence, MA). All samples had viability >80%. Single cells were then processed using the Chromium Next GEM Single Cell 3' Platform (10X Genomics, Pleasanton, CA). Approximately 30,000 cells were loaded on each channel with an average recovery rate of 24,000 cells. Libraries were sequenced on NextSeq 500/550 (Illumina, San Diego, CA) with an average sequencing depth of 50,000 reads/cell.

*Data Preparation*: Raw sequencing reads were demultiplexed, mapped to the GRCh38 human reference genome, and gene-expression matrices were generated using CellRanger v6.1 (10X Genomics, Pleasanton, CA). The following analyses were conducted in R 4.1 (22) and Python 3.6. Seurat package v4.0 was used for downstream data analyses (23), and figures were produced using the package ggplot2. Following a standard workflow, the gene-expression matrix was filtered to discard cells with less than 200 genes, as well as genes that were expressed in less than 3 cells. Samples then underwent quality control to remove cells with either too many or too few expressed genes (average around 2000 and 7000) and cells with too many mtRNA (average around 10%), resulting in a total of 113,213 cells. Then, "SCTransform" was applied with the "glmGamPoi" method to normalize gene expression data (24, 25), and a CLR transformation was applied to normalize protein data within each cell. After individual preparation, all the samples were introduced into a combined Seurat object via “FindIntegrationAnchors” and “IntegrateData” functions (24). Then scaled values of variable genes were then subject to principal component analysis (PCA) for linear dimension reduction. A shared nearest neighbor network was created based on Euclidean distances between cells in multidimensional PC space (the first 50 PC were used) and a fixed number of neighbors per cell (30 neighbors). This was used to generate a two-dimensional Uniform Manifold Approximation and Projection (UMAP) for visualization as a harmonized atlas to dissect the cell populations in BALF of healthy and CF donors.

*Differentially Expressed Genes, Cell Type Identification*, *Cross-sample Cell Type Replicability*, *Gene Ontology, Regulon, and Pseudotime analysis,* and *Flow cytometry* details are provided in the supplement.

**Data Availability**

RNA sequencing data have been deposited in the Gene Expression Omnibus, accession code GSE193782. The data set can be visualized at https://XXXXXX.ucsc.edu.

## Results

*BAL Cell Populations Are Conserved across HC and CF*

To assess the cellular composition in the non-diseased airspace versus that with CF, four HCs and three subjects with CF underwent BAL (donors, Table 1). BALF samples were spun, washed, and immediately loaded onto the 10X platform for scRNA-seq. Sequenced cells were processed, normalized, and integrated using the Seurat package (details in Methods). Uniform manifold approximation and projection (UMAP) of all seven samples illustrates a similar cluster distribution across all subjects (Figure 1A). Using a curated gene list, we identified twelve major cell types (Complete list of genes: Table E1 and Figure 1B-1D) (17, 18, 26-35). As anticipated, the most abundant cell type is the alveolar macrophage, followed by monocytes, cycling myeloid cells, dendritic cell 2, FOLR2 and SPP1 interstitial macrophages, lymphocytes, epithelial cells, migratory dendritic cells, dendritic cell 1, plasmacytoid DCs, and cycling lymphoid cells (Figure 1C-1D). Next, MetaNeighbor analysis was used to confirm the classification and replicability of the identified clusters within our dataset and with two other recently published datasets, Mould et al., and Schupp et al. (17, 18, 36). Within our dataset, samples derived from HC and CF subjects demonstrated a strong correlation across all cell types (Figure 1E). Similarly, our HC sample macrophages, monocytes, and epithelial cells strongly correlated with healthy cell types from Mould et al., as did our CF sample cell types with CF sputum cell types (Figure 1F-1G). All in all, curated genes and robust alignment of our cells with previous studies support the identification of twelve cell types in healthy and CF BALF.

*Four AM superclusters and shared subclusters co-exist in the airspace of the lungs*

To define AM heterogeneity, we isolated and re-clustered the AM population. Unbiased clustering defined 8 AM clusters. Some clusters were previously described (37), such as the Metallothionein, Chemokine, IFN-reacting, and Cycling AM clusters. (Figure 2A-2D). In addition, we identified 4 novel AM superclusters based on the differential expression of two genes: *APOC2* and *IFI27* (Figure 2B-2E): *APOC*$^{+}$*IFI27*$^{+}$ AMs.S1, *APOC*$^{-}$*IFI27*$^{+}$ AMs.S2, *APOC*$^{+}$*IFI27*$^{-}$ AMs.S3, and *APOC*$^{-}$*IFI27*$^{-}$ AMs.S4. Besides *APOC2* and *IFI27*,

the four AM superclusters share the same top DEGs *IFI6*, *CTSC*, *RPS4X*, *RBP4*, *FBP1*, and *GSN*, which suggest anti-microbial and metabolic function (i.e., common macrophage phenotypes) (Figure 2D). Each AM supercluster contains 4 shared subclusters. The subclusters are named after their top DEG(s): IGF1, CCL18, CXCL5 and Cholesterol (based on the expression of cholesterol-biosynthesis-related genes) (Figure 3A-3C). AM Supercluster 4 (AMs.S4) contained two additional subclusters, GF and SNHG. Cells in the GF subcluster expressed the epithelial growth factor *AREG* and TGF-β superfamily ligands including, *GDF15,* whereas cells in SNHG expressed genes encoding small nucleolar RNA host genes (SNHGs) and genes encoding nuclear transcriptional regulators and transcription factors. The remaining cluster in each supercluster is referred to as a small-subcluster within the supercluster, noting a small 's' after AMs, for example, AMs.s#. This small-subcluster is the supercluster minus all four shared subclusters (Figure 3D).

In addition to the AM superclusters, four other AM clusters can be defined based on their expression of Metallothioneins (MT.AMs), Chemokines (CK.AMs), IFN-related genes (IFN.AMs), and Cycling genes (CK.AMs) (Figure 2B). The chemokine cluster can be further divided into 4 subpopulations, which we denote as CK.AMs.c1-4. The top 6 DEGs and curated genes illustrate the unique chemokine expression pattern within each subcluster (Figure 4D). CK.AMs highly express *CCL18*, *CCL23* and chemokine receptor *CXCR4* in CK.AMs.c1; *CCL3*, *CCL4*, *CXCL8* and *CCL20* in CK.AMs.c2; *CXCL9*, *CXCL10* and *CXCL11* (IFN-gamma inducible chemokines) in CK.AMs.c3, and *CXCL3* in CK.AM.c4 (Figure 4B-4D). The expression of unique combinations of chemokines in each CK.AMs suggests that different subtypes of AMs are programmed to chemoattract distinct leukocytes into the airspace, depending on the inflammatory stimuli they encounter. This is a novel finding since the levels of division of labor of resident AMs secreting chemokines have not been previously appreciated.

Next, given the number of AM subtypes identified, we performed Gene Ontology (GO) enrichment analysis to infer functionality based on their set of DEGs (Figure 4E). Predictably, the Chemokine AM subclusters contained GO terms regarding cell interaction and response to stimuli, together with other enriched GO terms consistent with the top DEGs in each AM subtype. For instance, the top 5 GO terms for

Metallothionein AMs pertain to cellular responses to metal ions (with the top GO term shown in Figure 4E and Figure E1). Likewise, the top 4 enriched terms for Cholesterol AMs contained "*cholesterol biosynthetic process*" and "*regulation of cholesterol biosynthetic process*" (Figure E1).

Lastly, to understand the transcription factors and co-factors driving each AM subtype, we performed single-cell level regulon analysis using the SCENIC (Single-Cell rEgulatory Network Inference and Clustering) package in R and its Python implementation pySCENIC (38). Co-expressed genes were used to identify and score regulons enriched in each cluster. For example, the proinflammatory regulator of the cytokine-induced apoptotic pathway, *BACH2*, is the top regulon identified for CK.AMs.c2, while the *IRF7* regulon is enriched in IFN-reacting AMs (Figure 4F and Figure E1). This analysis offers yet another vantage point for understanding the distinct AM subtypes and corroborates their defined DEGs and GO terms.

*Like AM Chemokine clusters, monocytes also display unique chemokine-expressing clusters*

We performed unbiased clustering for monocytes, which outlined 8 distinct populations (Figure 5A-5C). The largest cluster was denoted as conventional monocytes (Conv.Mono) that expresses classic myeloid cell markers shared with the other seven clusters (Figure 5D). An additional six populations can be defined as functional monocytes based on the expression of specific genes, including FCN1.Mono, CCL2.Mono, CXCL5.Mono, CXCL9.Mono, CCL13.Mono, and VEGFA.Mono (Figure 5D). To note, we also identified an intermediate monocyte cluster (Int.Mono) with distinct DEGs that encompasses cells transitioning from Conv.Mono to other functional monocyte phenotypes. Thus, monocytes appear to display a division of labor in orchestrating and recruiting leukocytes into the airspace, similar to what is observed for chemokine-expressing AMs.

Next, we performed gene ontology and regulon analysis on the monocyte clusters. The most notable GO term was for FCN1.Mono, whose genes were enriched for bacterial defense mechanisms suggesting the importance of this monocyte subtype during pulmonary bacterial infections (Figure 5E and Figure E2). The regulon analysis

illuminated the overrepresentation of zinc finger transcription factors suggesting a common dependency on this structural motif for the differentiation of monocyte subtypes into functionally distinct macrophages (Figure 5F and Figure E2).

*Pseudotime analysis suggests monocytes differentiate into AMs and IMs*

ScRNA-seq captures all cell types at a given time point, producing a series of expression profiles across cells as they transition from one state to another. To define the transitional processes of macrophage development, we performed pseudotime analysis using all the monocytes, AMs, and IMs in our dataset (39-42). The trajectory analysis revealed starting points of origin in the monocyte cluster (purple region and white open dots, Figure 6A), revealing two directions of differentiation: one toward IMs and the other toward cycling AMs and AMs.S4 (Figure 6A-6B). AMs.S4 appears to have its own starting branch point leading to all other AM clusters (Figure 6A-6B), which suggests that AMs.S4 may contribute to all other AM subtypes in the airspace.

To pinpoint the genes that change as a function of pseudotime trajectory, DEG analysis was applied. As monocytes differentiate into clusters of AMs and IMs, different genes are up-regulated sequentially, likely as a cause or consequence of changes in their cell state over time (Figure 6C). Genes for cell proliferation and myeloid cell function were among the top DEGs (Figure 6C). For example, *ISG15* is highly expressed during the later progression of pseudotime and is mainly upregulated in IFN-reacting AMs. Whereas *another among the top 10 DEGs, MT2A,* change along pseudotime and associate with Metallothionein AMs. Interestingly, AM superclusters AMs.S1 and AMs.S2 seem to evolve much later than their counterparts AMs.S3 and AMs.S4, even though they share 4 subclusters and most gene expression, except *IFI27* and *APOC2* (Figure 6C).

Although scRNA-seq has revolutionized our knowledge of macrophage heterogeneity, there are instances where investigators need to isolate bulk populations to perform other functional or morphological analyses. Therefore, flow cytometric antibodies that can be used to sort and enrich bulk AMs and recruited monocytes are in need. Using flow cytometry, we demonstrate that, in addition to high side-scatter for AMs and low side-scatter for monocytes, investigators can use antibodies against CD43

(SPN) and CD169 to define AMs and CD93, CD36, and CD14 to identify monocytes (Figure 6D) (43-47). ScRNA-seq and TotalSeq (for single-cell level protein expression assessment) also support the use of CD169 and CD43 for AMs compared to CD93, CD36, and CD14 for monocytes (Figure 6E-F). Thus, several forms of technology validate that higher expression of CD43 and CD169 in BAL cells define AMs over recruited monocytes.

*CF AMs have upregulated genes important for differentiation and inflammation*

Prior studies have demonstrated functional differences in CF compared to non-CF macrophages (48-50). To investigate the impact of aberrant CFTR function on cell heterogeneity and transcriptional differences in cells isolated from BALF, healthy and CF subjects with preserved lung function (Table 1), limited lung inflammation, and not currently on CFTR modulator treatment were enrolled. Two subjects had previously been on CFTR modulator treatment but stopped due to side effects, while the third subject declined CFTR modulator treatment due to overall good health. The subjects with CF in this study were stable on their medication regimens and reported no exacerbations within the prior 24 months. We observed differential expression of several genes within the total cell population of BALF from subjects with CF versus HCs (Figure 7A), including increased expression of *APOC1*, *APOC2*, *CCL18*, and *SOD2* in CF cells. Increased expression of *APOE* and *CCL18* in CF AMs compared to HC AMs is of particular interest given prior reports showed that CCL18 is associated with alternatively activated macrophages, and people with CF have a higher proportion of alternatively active macrophages (Figure 7B) (37, 51, 52). Among AM superclusters, CF AMs had increased expression of *APOC1* in AMs.s2, AMs.s3, and AMs.s4; *APOE* and *CCL18* in AMs.s4; and *S100A9* in AMs.s3 (Figure 7C). Similar to the bulk monocyte population, monocyte subclusters from subjects with CF did not demonstrate differential expression of inflammatory genes (Figure 7D).

## Discussion

AMs have multiple functional roles, including clearing surfactant components, cellular debris, inhaled particulates, and pathogens. Therefore, it is not surprising that in the airspace, there would be multiple AM and monocyte subtypes with different specializations. Single-cell RNA sequencing (scRNA-seq) technology has revolutionized our understanding of AMs, and we are only beginning to appreciate the depth and complexity of airway macrophages and monocytes in healthy and diseased lungs. Here, we used scRNA-seq to show that AMs form 4 superclusters based on the expression of two genes, *APOC2* and *IFI27.* Each supercluster contains subclusters defined by characteristic DEGs, including those for cytokines and chemokines. In addition to AM superclusters, we observed four other AM clusters. Of particular interest the Chemokine AM cluster, which divides further into four subclusters. Each of these chemokine clusters expressed a combination of chemokines that selectively chemoattract different leukocytes into the environment. Strikingly, there appears to be a defined niche for each AM subtype given the tightly regulated frequency of each AM subtype across all samples.

Although our unbiased AM analysis in Figure 2B resulted in 8 clusters, four AM superclusters, and four other clusters (i.e., Chemokine, Metallothionein, Interferon, and Cycling AMs), we questioned whether the non-supercluster AMs express all four combinations of *APOC2* and *IFI27*. Indeed, we observed a quadrant distribution of APOC2 and IFI27 for Chemokine, Metallothionein, and Interferon AMs (Figure E3). This suggests that non-supercluster AMs clusters might be a part of the four AM superclusters. And the reason this was not observed in the analysis of Figure 2B is because of the overpowering expression of other functional genes in the non-supercluster AMs. Based on this observation, a hypothesis might be that discrete compartments in the lung dictate the expression of APOC2 and IFI27 and within these compartments, a family of AMs, all the AM subtypes, exist to ensure proper physiological functions (Figure E4).

Why do the superclusters and subclusters exist? In addition to the hypothesis that each AM supercluster exists in a discrete compartment of the lung, AM superclusters might regulate their environment in a tissue-specific niche and functionally

specific manner. For example, it would be scientifically and clinically interesting to know whether certain CK.AM clusters regulate the leukocyte environment in the airspace during and after infection with respiratory viruses like SARS-CoV-2.

Chemokines are chemoattractant cytokines, and this type of cytokine profiling resembles what has been observed for decades in T cell biology. Each T cell type, such as Th1 and Th2, selectively produces a characteristic set of cytokines. It is known that the differentiation of T cells is driven by distinct transcription factors like T-bet and Gata-3. Likewise, each AM profiled in this study outlined a distinct set of transcription factors, observed by our regulon analysis. Therefore, an important future goal in studying the AM system will be to investigate which transcription factors regulate the individual Chemokine AM subtypes and beyond. Overall, the findings presented here open the door to identify new therapeutic targets for inflammatory lung disease, while significantly enhancing our understanding of the complexity that exists with airspace macrophages.

Many lines of evidence point to impaired innate immune cell function in the CF lung (48-50, 53, 54). Our study was originally designed to investigate the cellular and transcriptional differences in cells extracted from BALF between a cohort of CF subjects with normal lung function compared to healthy controls. Such differences may contribute to the phenotypic differences described in prior studies (48-50, 53, 54). Although there were minimal differences between CF and healthy controls, we demonstrate that all four AM superclusters have increased expression of genes important for regulating of macrophage differentiation and inflammation in CF versus healthy control subjects. This suggests that AMs in the CF lung have the potential to generate a more robust immune response, through greater priming of inflammatory cytokine and chemokine production. The CF subjects in our study had preserved lung function and limited inflammation, by design, as evidenced by lack of neutrophilic infiltration in BALF. This suggests that our findings may be reflective of CFTR mutations and mild disease compared to severe CF subjects. Future studies should include subjects with more significant lung disease to determine if this change in the lung environment has an additional impact on CF immune cell gene expression. Additionally, comparisons of CF subjects with and without lung disease could help identify factors in airspace cells that predispose one to disease progression.

**Acknowledgement**

This work was supported by National Institutes of Health grants R01 HL115334, R01 HL135001, and R35 HL155458 (CJ). 10x Genomics and Illumina sequencing workflows were performed in collaboration with the Genomics Shared Resource and Single Cell Genomics Cores at Dartmouth, which receives support from NCI CCSG grant P30 CA023108, NIGMS P20 GM130454, and S10 OD025235. XL, AA, and CJ prepared the manuscript; XL, FK and CJ performed the experiments; DA and AA enrolled all subjects, collected samples, and curated clinical data; DM processed clinical samples; XL performed the scRNA-seq and other bioinformatics analyses; all authors provided intellectual input, critical feedback, discussed results, and designed experiments. Thank you to Dr. Tor Wager and William T. King for critical feedback.

**Competing interests**

All authors declare that there are no competing interests.

**Table 1. Subject Characteristics**

| **Characteristic** | **CF (n = 3)** | **HC (n = 4)** |
|---|---|---|
| **Sex, female % (n)** | 33% (1) | 50% (2) |
| **Average age, years (SD)** | 28 ± 4.6 | 29 ± 6.1 |
| **FEV1, percent predicted (SD)** | 89.8 ± 11 | |
| ***Pa* colonization % (n)** | 67% (2) | |

Values are means ± standard deviation (SD); n = number of subjects; FEV1 = forced expiratory volume in 1 second; Pa = Pseudomonas aeruginosa; CF = cystic fibrosis; HC = healthy control.

## Figure legends

**Figure.1 Twelve BAL cell types identified in healthy control (HC) and CF samples.**
(A) ScRNA-seq data shows integrated UMAP with HC and CF, along with individual sample distribution. (B) UMAP demonstrating the twelve major cell types: alveolar macrophages (AMs), myeloid and lymphoid cycling cells (Cyc.Mye and Cyc.Lym), FOLR2 and SPP1 interstitial macrophages, monocytes (Mono), Dendritic cell 1 (DC1), Dendritic cell 2 (DC2), Migratory DC (Mig.DC), plasmacytoid DCs (pDCs), lymphocytes (Lym), and epithelial cells (Epi). (C) Percentage bar graph shows the distribution of cell types in each sample. (D) Dot plot shows the expression of curated genes in each individual cell type. (E) Correlation heat map of cell types against all samples in in-house dataset (Li). (F) Correlation heat map of HC cell types in in-house and Mould dataset. The lower quadrant highlights the corresponding cell types. (G) Correlation heat map of HC and CF cell types in in-house and Schupp dataset. The lower quadrant highlights the corresponding cell types.

**Figure.2 Unbiased analysis of scRNA-seq data illustrates four AM superclusters based on IFI27 and APOC2 expression.**
(A) UMAP of re-clustered AMs shows sample source distribution. (B) UMAP shows distinct AM clusters: AM Supercluster 1 (AMs.S1), AMs.S2, AMs.S3 AMs.S4, Metallothionein (MT.AMs), Chemokine (CK.AMs), IFN-reacting (IFN.AMs) and Cycling AM (Cyc.AMs) clusters. (C) Percentage of individual AM clusters for each subject. (D) Dot plot shows the expression of the top 6 DEGs in individual AM clusters, duplicates removed. For example, AMs.S4's top 6 DEGs are already present in the first three AM superclusters; therefore, AMs.S4 is seen on the Y axis but not the X axis. (E) Feature plot shows the expression of *IFI27* and *APOC2*, which define the four AM superclusters.

**Figure.3 AM superclusters contain 4 shared subclusters and a small-subcluster.**
(A) UMAP illustrates the shared AM subclusters. (B) Percentage of individual AM subclusters for each subject. (C) Featured plots show four representative genes for each shared AM subcluster. (D) Dot plot illustrates the expression of the top 6 DEGs in each individual AM subcluster, duplicates removed.

**Figure.4 Chemokine AM subcluster express a unique combination of chemokines.**
(A) UMAP illustrates distinct chemokine-expressing AM subclusters: CK.AMs.c1, CK.AMs.c2, CK.AMs.c3, and CK.AMs.c4. (B) Percentage of individual chemokine-expressing AM subclusters for each subject. (C) Featured plots show four representative gene expression for individual chemokine-expressing AM subcluster. (D) Dot plot illustrates the expression of the top 6 DEGs (duplicates removed) and curated chemokine genes in individual chemokine-expressing AM subcluster. (E) Heat map shows the top 5 gene ontology terms (biological process) for each AM subtype. Top 1 term is labeled. (F) Heat map shows the top 5 regulons for each AM subtype. Top 1 regulon is labeled.

**Figure.5 BAL monocytes contain 8 distinct clusters.**
(A) UMAP of monocyte shows sample source distribution. (B) UMAP shows distinct monocyte clusters: conventional monocyte cluster (Conv.Mono) and intermediate monocyte cluster (Int.Mono) and others. (C) Percentage bar graph outlines the distribution of individual monocyte clusters in each sample. (E) Dot plot shows the expression of the top 6 DEGs in individual monocytes clusters, duplicates removed. (F) Heat map shows the top 5 gene ontology terms (biological process) for each monocyte subtype. Top 2 terms are labeled. (G) Heat map shows the top 5 regulons for each monocyte subtype. Top 2 regulons are labeled.

**Figure 6. Pseudotime analysis of monocytes and macrophages in BALF.**
(A) Pseudotime plot shows the evolvement of monocytes to macrophages with the root as white dots and the principal points as solid black dots. (B) UMAP shows distinct monocyte and macrophage clusters included for Pseudotime analysis. (C) Heat map shows the top 100 pseudotime DEGs across the monocyte-to-macrophage branch. Top 10 DEGs are labelled. (D) Flow cytometry plot illustrates the high expression of CD43 and CD169 on AMs and CD93, CD36 and CD14 on monocytes. (E) Featured plots show representative gene expression and TotalSeq detection in AMs and monocytes. (F) Radar plots illustrate the comparison of expression and protein detection levels of representative genes and TotalSeq antibodies.

**Figure.7 Gene expression analysis between HC and CF samples**
(A) Volcano plot demonstrates the DEGs in either HC or CF samples. Genes with a P value smaller than 0.05 and log2 fold change larger than 25% are shown. Top 15 DEGs are labeled. (B) Dot Chart shows the top 3 DEGs in either HC or CF samples for the twelve major cell types. (C) Dot Chart shows the top 3 DEGs in either HC or CF samples for individual AM subtypes. (D) Dot Chart shows the top 3 DEGs in either HC or CF samples for individual monocyte subtypes.

## References

1. Fujii W, Kapellos TS, Bassler K, Handler K, Holsten L, Knoll R, Warnat-Herresthal S, Oestreich M, Hinkley ER, Hasenauer J, Pizarro C, Thiele C, Aschenbrenner AC, Ulas T, Skowasch D, Schultze JL. Alveolar macrophage transcriptomic profiling in COPD shows major lipid metabolism changes. *ERJ Open Res* 2021; 7.
2. Hetzel M, Ackermann M, Lachmann N. Beyond "Big Eaters": The Versatile Role of Alveolar Macrophages in Health and Disease. *Int J Mol Sci* 2021; 22.
3. Janssen WJ, Barthel L, Muldrow A, Oberley-Deegan RE, Kearns MT, Jakubzick C, Henson PM. Fas Determines Differential Fates of Resident and Recruited Macrophages During Resolution of Acute Lung Injury. *Am J Respir Crit Care Med* 2011.
4. Mould KJ, Barthel L, Mohning MP, Thomas SM, McCubbrey AL, Danhorn T, Leach SM, Fingerlin TE, O'Connor BP, Reisz JA, D'Alessandro A, Bratton DL, Jakubzick CV, Janssen WJ. Cell Origin Dictates Programming of Resident versus Recruited Macrophages during Acute Lung Injury. *Am J Respir Cell Mol Biol* 2017; 57: 294-306.
5. Mould KJ, Jackson ND, Henson PM, Seibold M, Janssen WJ. Single cell RNA sequencing identifies unique inflammatory airspace macrophage subsets. *JCI Insight* 2019; 4.
6. Yona S, Kim KW, Wolf Y, Mildner A, Varol D, Breker M, Strauss-Ayali D, Viukov S, Guilliams M, Misharin A, Hume DA, Perlman H, Malissen B, Zelzer E, Jung S. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* 2013; 38: 79-91.
7. Jakubzick C, Gautier EL, Gibbings SL, Sojka DK, Schlitzer A, Johnson TE, Ivanov S, Duan Q, Bala S, Condon T, van Rooijen N, Grainger JR, Belkaid Y, Ma'ayan A, Riches DW, Yokoyama WM, Ginhoux F, Henson PM, Randolph GJ. Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity* 2013; 39: 599-610.
8. Misharin AV, Morales-Nebreda L, Reyfman PA, Cuda CM, Walter JM, McQuattie-Pimentel AC, Chen CI, Anekalla KR, Joshi N, Williams KJN, Abdala-Valencia H, Yacoub TJ, Chi M, Chiu S, Gonzalez-Gonzalez FJ, Gates K, Lam AP, Nicholson TT, Homan PJ, Soberanes S, Dominguez S, Morgan VK, Saber R, Shaffer A, Hinchcliff M, Marshall SA, Bharat A, Berdnikovs S, Bhorade SM, Bartom ET, Morimoto RI, Balch WE, Sznajder JI, Chandel NS, Mutlu GM, Jain M, Gottardi CJ, Singer BD, Ridge KM, Bagheri N, Shilatifard A, Budinger GRS, Perlman H. Monocyte-derived alveolar macrophages drive lung fibrosis and persist in the lung over the life span. *J Exp Med* 2017; 214: 2387-2404.
9. Fricker M, Gibson PG. Macrophage dysfunction in the pathogenesis and treatment of asthma. *Eur Respir J* 2017; 50.
10. O'Beirne SL, Kikkers SA, Oromendia C, Salit J, Rostmai MR, Ballman KV, Kaner RJ, Crystal RG, Cloonan SM. Alveolar Macrophage Immunometabolism and Lung Function Impairment in Smoking and Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2020; 201: 735-739.
11. Deriy LV, Gomez EA, Zhang G, Beacham DW, Hopson JA, Gallan AJ, Shevchenko PD, Bindokas VP, Nelson DJ. Disease-causing mutations in the cystic fibrosis transmembrane conductance regulator determine the functional responses of alveolar macrophages. *J Biol Chem* 2009; 284: 35926-35938.
12. Bonfield TL. Macrophage Dysfunction in Cystic Fibrosis: A Therapeutic Target to Enhance Self-Immunity. *Am J Respir Crit Care Med* 2015; 192: 1406-1407.

13. Hey J, Paulsen M, Toth R, Weichenhan D, Butz S, Schatterny J, Liebers R, Lutsik P, Plass C, Mall MA. Epigenetic reprogramming of airway macrophages promotes polarization and inflammation in muco-obstructive lung disease. *Nat Commun* 2021; 12: 6520.
14. Bruscia EM, Bonfield TL. Cystic Fibrosis Lung Immunity: The Role of the Macrophage. *J Innate Immun* 2016; 8: 550-563.
15. Li X, Rawat K, Jakubzick CV. Targeting resident macrophages in cancer. *Nat Immunol* 2021; 22: 1078-1079.
16. María Casanova-Acebes ED, Andrew Leader, Jessica LeBerichel, Jovan Nikolic, Blanca M. Morales, Markus Brown, Christie Chang, Leanna Troncoso, Steven T. Chen, Ana Sastre-Perona, Matthew D. Park, Alexandra Tabachnikova, Maxime Dhainaut, Pauline Hamon, Barbara Maier, Catherine M. Sawai, Esperanza Agulló-Pascual, Markus Schober, Brian Brown, Boris Reizis, Thomas Marron, Ephraim Kenigsberg, Christine Moussion, Philippe Benaroch, Julio A. Aguirre-Ghiso, and Miriam Merad. Tissue-resident macrophages provide a pro-tumorigenic niche to early NSCLC cells. *Nature* 2021.
17. Mould KJ, Moore CM, McManus SA, McCubbrey AL, McClendon JD, Griesmer CL, Henson PM, Janssen WJ. Airspace Macrophages and Monocytes Exist in Transcriptionally Distinct Subsets in Healthy Adults. *Am J Respir Crit Care Med* 2021; 203: 946-956.
18. Schupp JC, Khanal S, Gomez JL, Sauler M, Adams TS, Chupp GL, Yan X, Poli S, Zhao Y, Montgomery RR, Rosas IO, Dela Cruz CS, Bruscia EM, Egan ME, Kaminski N, Britto CJ. Single-Cell Transcriptional Archetypes of Airway Inflammation in Cystic Fibrosis. *Am J Respir Crit Care Med* 2020; 202: 1419-1429.
19. Hogan DA, Willger SD, Dolben EL, Hampton TH, Stanton BA, Morrison HG, Sogin ML, Czum J, Ashare A. Analysis of Lung Microbiota in Bronchoalveolar Lavage, Protected Brush and Sputum Samples from Subjects with Mild-To-Moderate Cystic Fibrosis Lung Disease. *PLoS One* 2016; 11: e0149998.
20. Aridgides DS, Mellinger DL, Armstrong DA, Hazlett HF, Dessaint JA, Hampton TH, Atkins GT, Carroll JL, Ashare A. Functional and metabolic impairment in cigarette smoke-exposed macrophages is tied to oxidative stress. *Sci Rep* 2019; 9: 9624.
21. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017; 14: 865-868.
22. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL
. *http://wwwR-projectorg* 2014.
23. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. *Cell* 2021; 184: 3573-3587 e3529.
24. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell* 2019; 177: 1888-1902 e1821.
25. Ahlmann-Eltze C, Huber W. glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* 2021; 36: 5701-5702.

26. Gibbings SL, Thomas SM, Atif SM, McCubbrey AL, Desch AN, Danhorn T, Leach SM, Bratton DL, Henson PM, Janssen WJ, Jakubzick CV. Three Unique Interstitial Macrophages in the Murine Lung at Steady State. *Am J Respir Cell Mol Biol* 2017; 57: 66-76.
27. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S, Jardine L, Dixon D, Stephenson E, Nilsson E, Grundberg I, McDonald D, Filby A, Li W, De Jager PL, Rozenblatt-Rosen O, Lane AA, Haniffa M, Regev A, Hacohen N. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* 2017; 356.
28. Collin M, Bigley V. Human dendritic cell subsets: an update. *Immunology* 2018; 154: 3-20.
29. Collin M, McGovern N, Haniffa M. Human dendritic cell subsets. *Immunology* 2013; 140: 22-30.
30. Cai Y, Dai Y, Wang Y, Yang Q, Guo J, Wei C, Chen W, Huang H, Zhu J, Zhang C, Zheng W, Wen Z, Liu H, Zhang M, Xing S, Jin Q, Feng CG, Chen X. Single-cell transcriptomics of blood reveals a natural killer cell subset depletion in tuberculosis. *EBioMedicine* 2020; 53: 102686.
31. Brown CC, Gudjonson H, Pritykin Y, Deep D, Lavallee VP, Mendoza A, Fromme R, Mazutis L, Ariyan C, Leslie C, Pe'er D, Rudensky AY. Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity. *Cell* 2019; 179: 846-863 e824.
32. Gibbings SL, Goyal R, Desch AN, Leach SM, Prabagar M, Atif SM, Bratton DL, Janssen W, Jakubzick CV. Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. *Blood* 2015; 126: 1357-1366.
33. Gautier EL, Shay T, Miller J, Greter M, Jakubzick C, Ivanov S, Helft J, Chow A, Elpek KG, Gordonov S, Mazloom AR, Ma'ayan A, Chua WJ, Hansen TH, Turley SJ, Merad M, Randolph GJ, Immunological Genome C. Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat Immunol* 2012; 13: 1118-1128.
34. Jin X, Meng L, Yin Z, Yu H, Zhang L, Liang W, Wang S, Liu G, Zhang L. Characterization of dendritic cell subtypes in human cord blood by single-cell sequencing. *Biophysics Reports* 2019; 5: 199-208.
35. Leach SM, Gibbings SL, Tewari AD, Atif SM, Vestal B, Danhorn T, Janssen WJ, Wager TD, Jakubzick CV. Human and Mouse Transcriptome Profiling Identifies Cross-Species Homology in Pulmonary and Lymph Node Mononuclear Phagocytes. *Cell Rep* 2020; 33: 108337.
36. Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* 2018; 9: 884.
37. Gardai SJ, McPhillips KA, Frasch SC, Janssen WJ, Starefeldt A, Murphy-Ullrich JE, Bratton DL, Oldenborg PA, Michalak M, Henson PM. Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte. *Cell* 2005; 123: 321-334.
38. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, van den Oord J, Atak ZK, Wouters J, Aerts S. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017; 14: 1083-1086.

39. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 2019; 566: 496-502.
40. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017; 14: 979-982.
41. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with Census. *Nat Methods* 2017; 14: 309-315.
42. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014; 32: 381-386.
43. Bharat A, Bhorade SM, Morales-Nebreda L, McQuattie-Pimentel AC, Soberanes S, Ridge K, DeCamp MM, Mestan KK, Perlman H, Budinger GR, Misharin AV. Flow Cytometry Reveals Similarities Between Lung Macrophages in Humans and Mice. *Am J Respir Cell Mol Biol* 2016; 54: 147-149.
44. Yu YA, Hotten DF, Malakhau Y, Volker E, Ghio AJ, Noble PW, Kraft M, Hollingsworth JW, Gunn MD, Tighe RM. Flow Cytometric Analysis of Myeloid Cells in Human Blood, Bronchoalveolar Lavage, and Lung Tissues. *Am J Respir Cell Mol Biol* 2015.
45. Yu YA, Tighe RM. Isolation and Characterization of Human Lung Myeloid Cells. *Methods Mol Biol* 2018; 1809: 111-119.
46. Desch AN, Gibbings SL, Goyal R, Kolde R, Bednarek J, Bruno T, Slansky JE, Jacobelli J, Mason R, Ito Y, Messier E, Randolph GJ, Prabagar M, Atif SM, Segura E, Xavier RJ, Bratton DL, Janssen WJ, Henson PM, Jakubzick CV. Flow Cytometric Analysis of Mononuclear Phagocytes in Non-diseased Human Lung and Lung-draining Lymph Nodes. *Am J Respir Crit Care Med* 2015.
47. Gibbings SL, Jakubzick CV. A Consistent Method to Identify and Isolate Mononuclear Phagocytes from Human Lung and Lymph Nodes. *Methods Mol Biol* 2018; 1799: 381-395.
48. Hazlett HF, Hampton TH, Aridgides DS, Armstrong DA, Dessaint JA, Mellinger DL, Nymon AB, Ashare A. Altered iron metabolism in cystic fibrosis macrophages: the impact of CFTR modulators and implications for Pseudomonas aeruginosa survival. *Sci Rep* 2020; 10: 10935.
49. Bessich JL, Nymon AB, Moulton LA, Dorman D, Ashare A. Low levels of insulin-like growth factor-1 contribute to alveolar macrophage dysfunction in cystic fibrosis. *J Immunol* 2013; 191: 378-385.
50. Chen Y, Armstrong DA, Salas LA, Hazlett HF, Nymon AB, Dessaint JA, Aridgides DS, Mellinger DL, Liu X, Christensen BC, Ashare A. Genome-wide DNA methylation profiling shows a distinct epigenetic signature associated with lung macrophages in cystic fibrosis. *Clin Epigenetics* 2018; 10: 152.
51. Murphy BS, Bush HM, Sundareshan V, Davis C, Hagadone J, Cory TJ, Hoy H, Hayes D, Jr., Anstead MI, Feola DJ. Characterization of macrophage activation states in patients with cystic fibrosis. *J Cyst Fibros* 2010; 9: 314-322.
52. Ruytinx P, Proost P, Van Damme J, Struyf S. Chemokine-Induced Macrophage Polarization in Inflammatory Conditions. *Front Immunol* 2018; 9: 1930.
53. Di A, Brown ME, Deriy LV, Li C, Szeto FL, Chen Y, Huang P, Tong J, Naren AP, Bindokas V, Palfrey HC, Nelson DJ. CFTR regulates phagosome acidification in macrophages and alters bactericidal activity. *Nat Cell Biol* 2006; 8: 933-944.

54. Paemka L, McCullagh BN, Abou Alaiwa MH, Stoltz DA, Dong Q, Randak CO, Gray RD, McCray PB, Jr. Monocyte derived macrophages from CF pigs exhibit increased inflammatory responses at birth. *J Cyst Fibros* 2017; 16: 471-474.
55. Armstrong DA, Lee MK, Hazlett HF, Dessaint JA, Mellinger DL, Aridgides DS, Hendricks GM, Abdalla MAK, Christensen BC, Ashare A. Extracellular Vesicles from Pseudomonas aeruginosa Suppress MHC-Related Molecules in Human Lung Macrophages. *Immunohorizons* 2020; 4: 508-519.

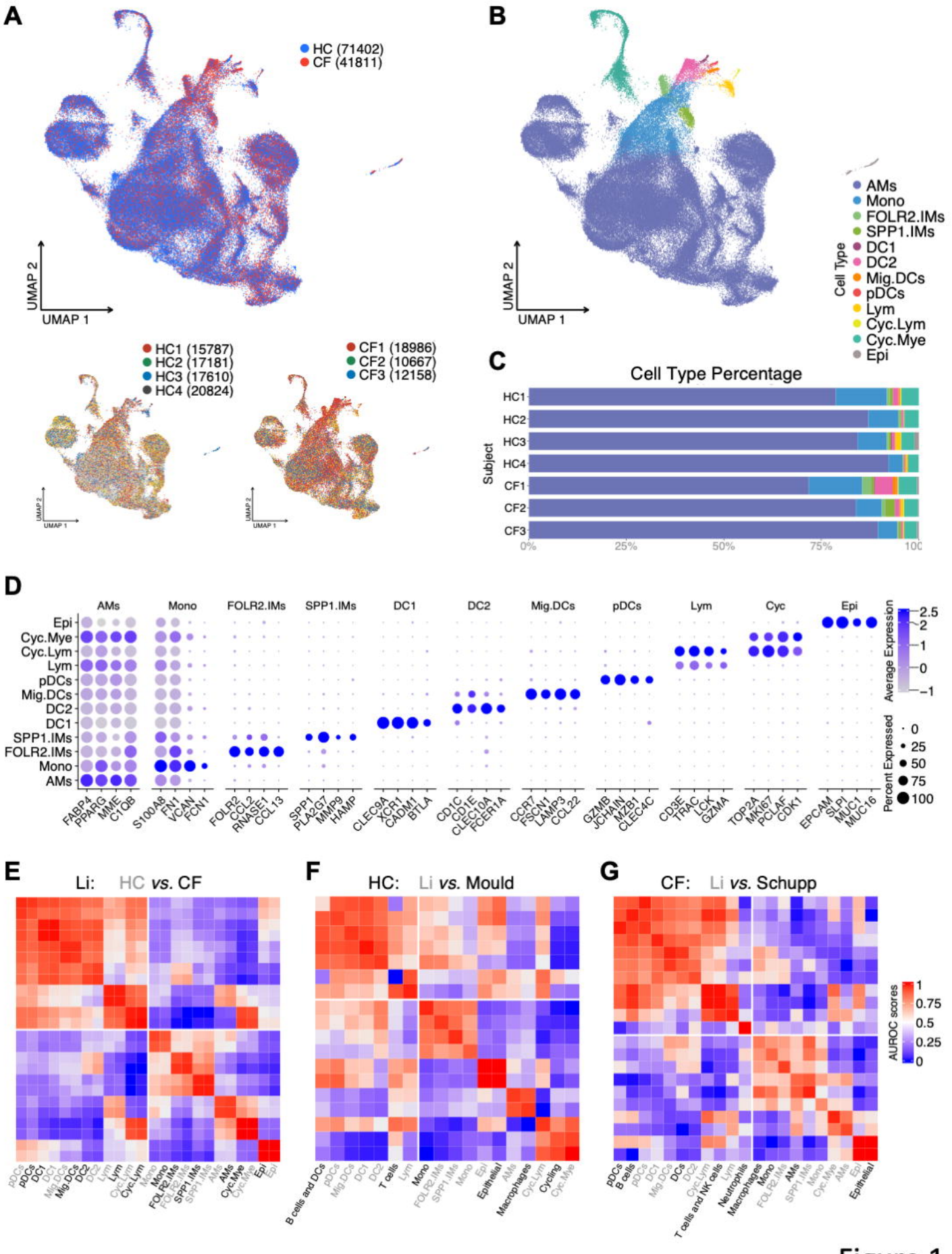

A
HC (71402)
CF (41811)
UMAP 2
UMAP 1
HC1 (15787)
HC2 (17181)
HC3 (17610)
HC4 (20824)
CF1 (18986)
CF2 (10667)
CF3 (12158)
B
AMs
Mono
FOLR2.IMs
SPP1.IMs
DC1
DC2
Mig.DCs
pDCs
Lym
Cyc.Lym
Cyc.Mye
Epi
Cell Type
C
Cell Type Percentage
Subject
HC1
HC2
HC3
HC4
CF1
CF2
CF3
0%
25%
50%
75%
100
D
AMs
Mono
FOLR2.IMs
SPP1.IMs
DC1
DC2
Mig.DCs
pDCs
Lym
Cyc
Epi
FABP4 PPARG MME C1QB
S100A8 FN1 VCAN FCN1
FOLR2 CCL2 RNASE1 CCL13
SPP1 PLA2G7 MMP9 HAMP
CLEC9A XCR1 CADM1 BTLA
CD1C CD1E CLEC10A FCER1A
CCR7 FSCN1 LAMP3 CCL22
GZMB JCHAIN MZB1 CLEC4C
CD3E TRAC LCK GZMA
TOP2A MKI67 PCLAF CDK1
EPCAM SLPI MUC1 MUC16
Average Expression
2.5
2
1
0
−1
Percent Expressed
0
25
50
75
100
E
Li: HC vs. CF
F
HC: Li vs. Mould
G
CF: Li vs. Schupp
AUROC scores
1
0.75
0.5
0.25
0

Figure 1

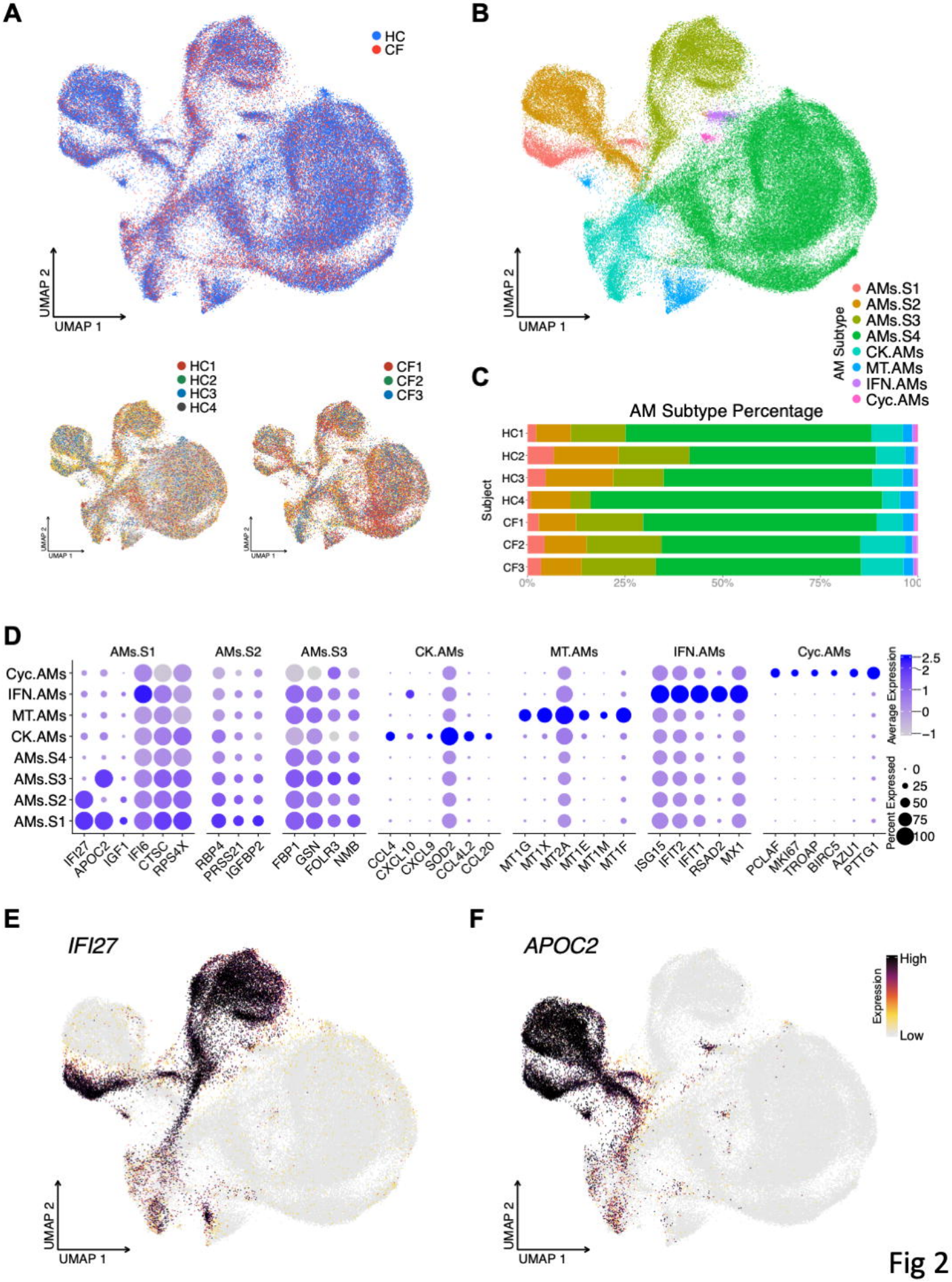

A
HC
CF
UMAP 2
UMAP 1
HC1
HC2
HC3
HC4
CF1
CF2
CF3
B
AMs.S1
AMs.S2
AMs.S3
AMs.S4
CK.AMs
MT.AMs
IFN.AMs
Cyc.AMs
AM Subtype
C
AM Subtype Percentage
Subject
0%
25%
50%
75%
100
D
AMs.S1
AMs.S2
AMs.S3
CK.AMs
MT.AMs
IFN.AMs
Cyc.AMs
IFI27
APOC2
IGF1
IFI6
CTSC
RPS4X
RBP4
PRSS21
IGFBP2
FBP1
GSN
FOLR3
NMB
CCL4
CXCL10
CXCL9
SOD2
CCL4L2
CCL20
MT1G
MT1X
MT2A
MT1E
MT1M
MT1F
ISG15
IFIT2
IFIT1
RSAD2
MX1
PCLAF
MKI67
TROAP
BIRC5
AZU1
PTTG1
Average Expression
Percent Expressed
E
IFI27
F
APOC2
High
Low
Expression

Fig 2

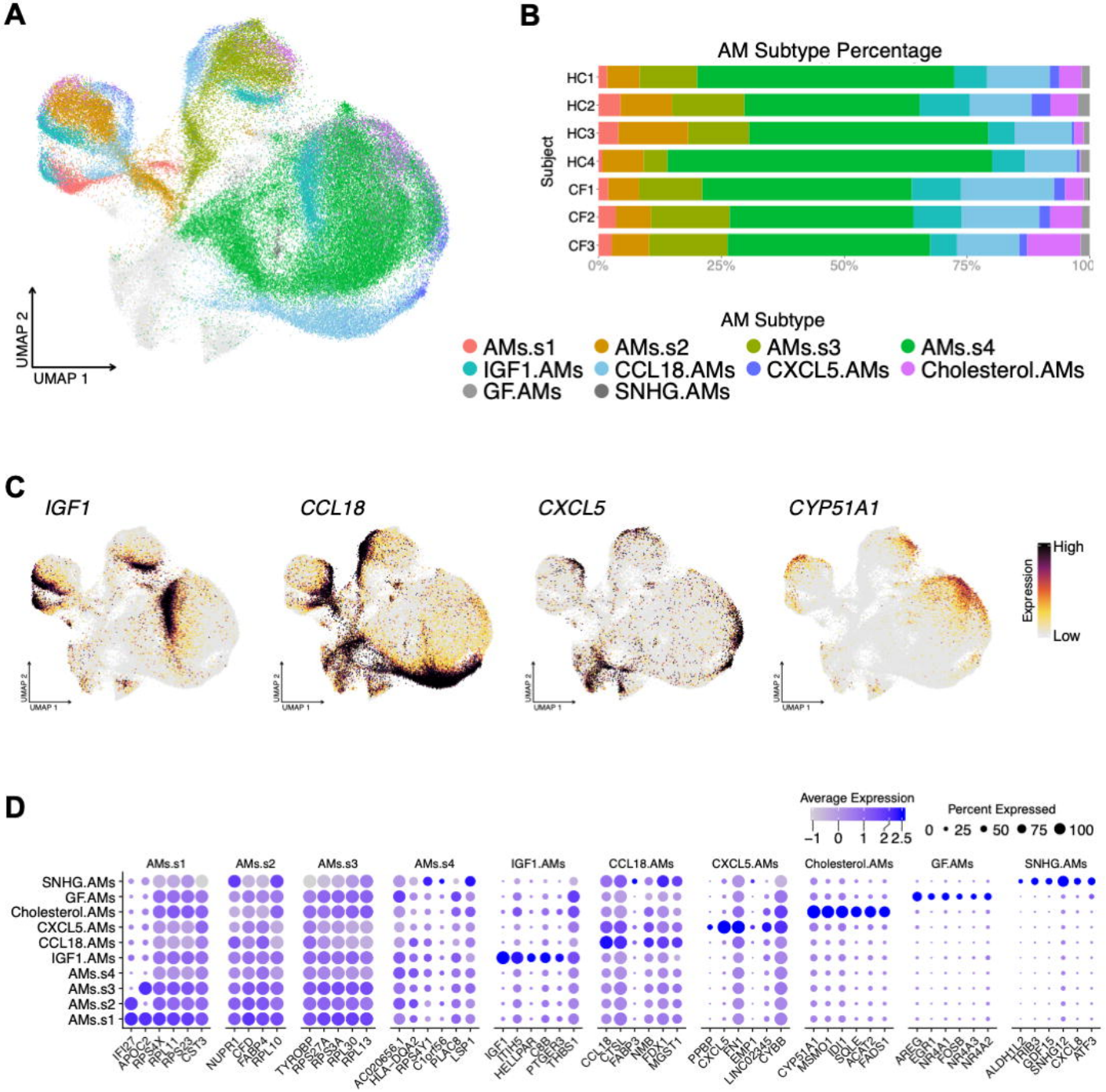

A
UMAP 2
UMAP 1
B
AM Subtype Percentage
Subject
HC1
HC2
HC3
HC4
CF1
CF2
CF3
0%
25%
50%
75%
100
AM Subtype
AMs.s1
AMs.s2
AMs.s3
AMs.s4
IGF1.AMs
CCL18.AMs
CXCL5.AMs
Cholesterol.AMs
GF.AMs
SNHG.AMs
C
IGF1
CCL18
CXCL5
CYP51A1
High
Low
Expression
D
Average Expression
−1 0 1 2 2.5
Percent Expressed
0 25 50 75 100
AMs.s1
AMs.s2
AMs.s3
AMs.s4
IGF1.AMs
CCL18.AMs
CXCL5.AMs
Cholesterol.AMs
GF.AMs
SNHG.AMs
IFI27 APOC2 RPS4X RPL11 RPS23 CST3
NUPR1 CFD FABP4 RPL10
TYROBP RPS21A RPS3A RPL30 RPL13
AC020656.1 HLA-DQA2 RPS4Y1 C1orf56 PLAC8 LSP1
IGF1 ITIH5 HELLPAR CBB PTGER3 THBS1
CCL18 CTSL FABP3 NMB FDX1 MGST1
PPBP CXCL5 ENO1 EMP1 LINC02345 CYBB
CYP51A1 MSMO1 IDI1 SQLE ACAT2 FADS1
AREG EGR1 NR4A1 FOSB NR4A3 NR4A2
ALDH1L2 TRIB3 GDF15 SNHG12 CXCL8 ATF3

Figure 3

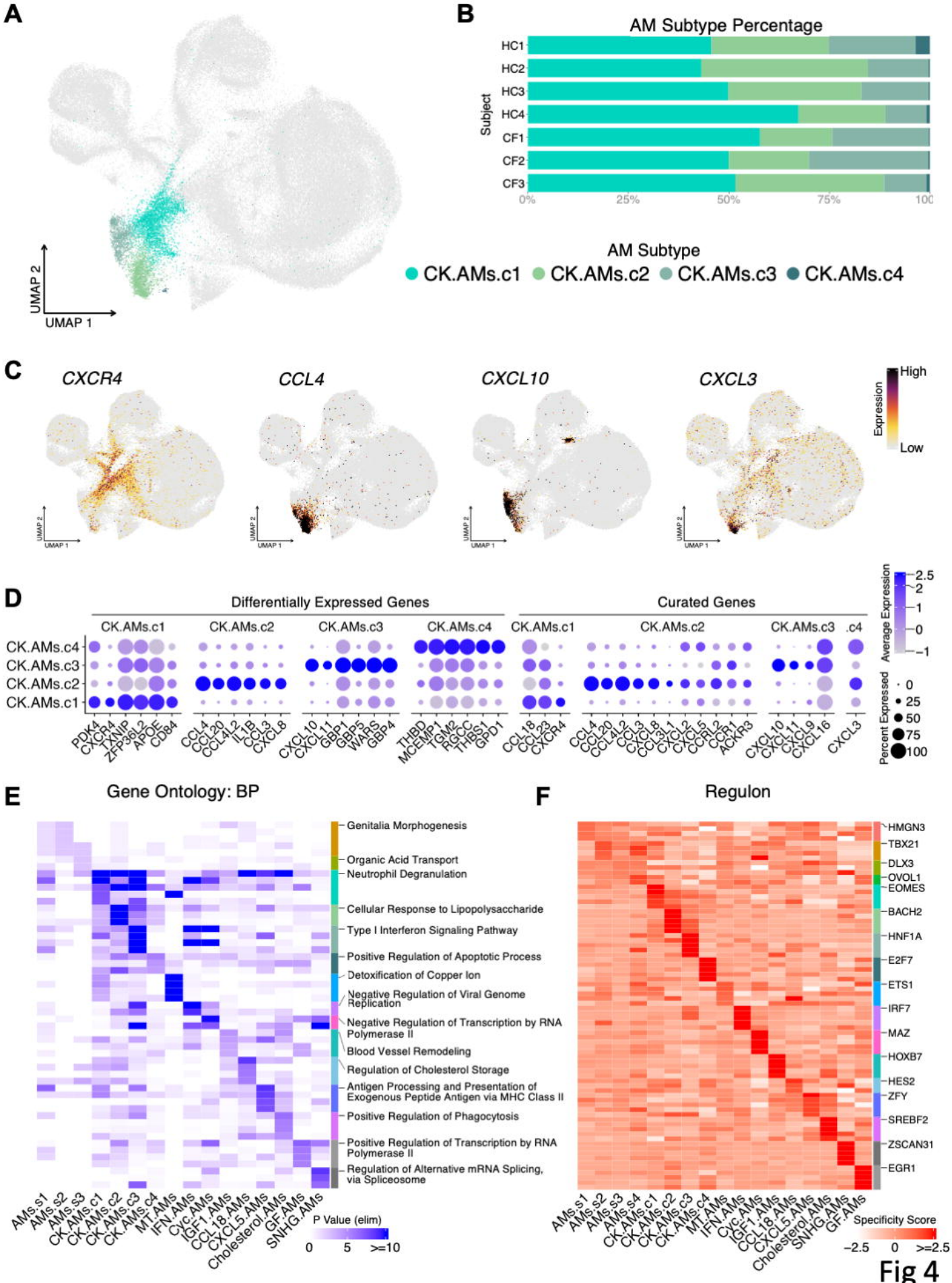
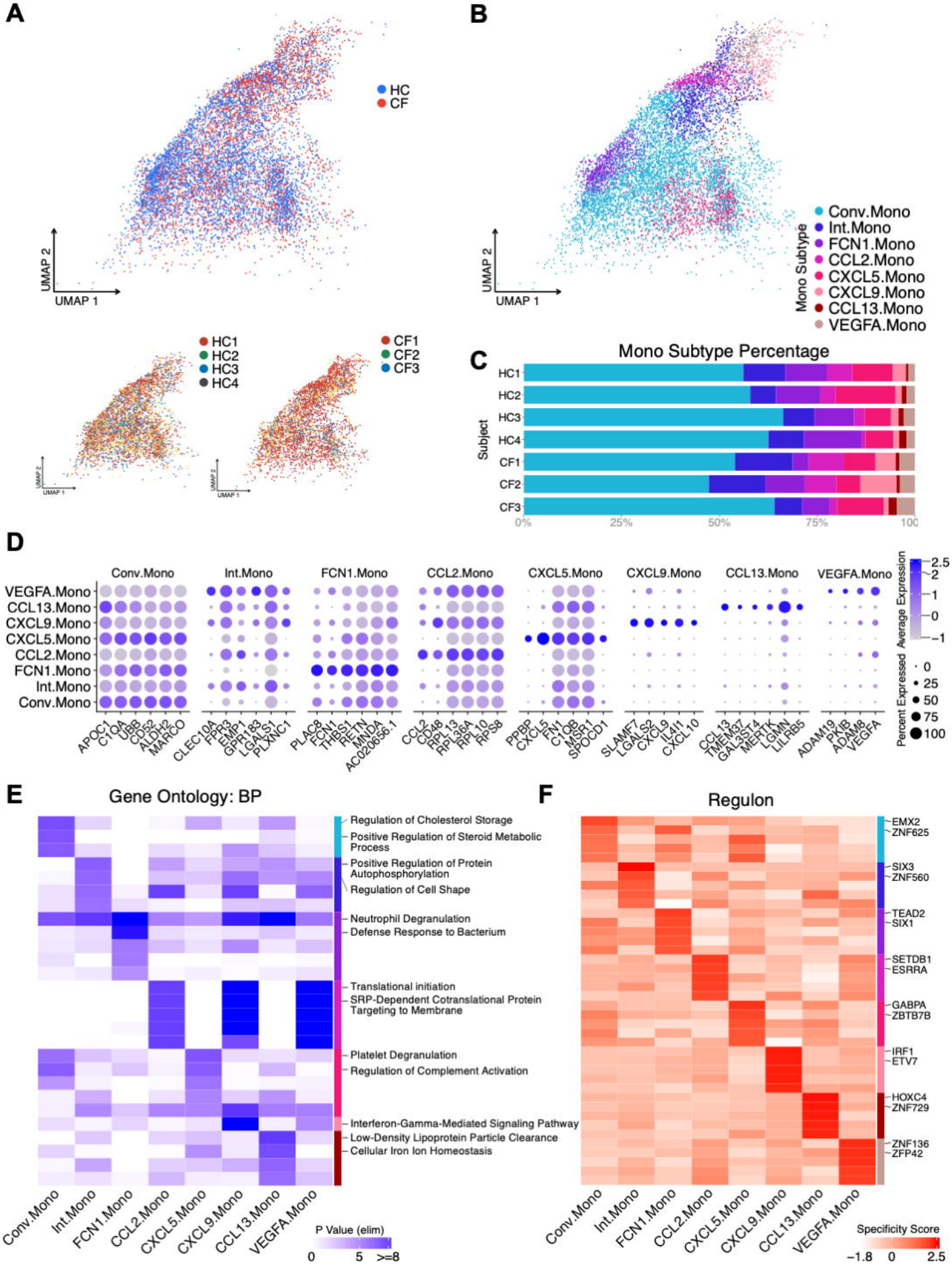

A
UMAP 2
UMAP 1
B
AM Subtype Percentage
Subject
HC1
HC2
HC3
HC4
CF1
CF2
CF3
0%
25%
50%
75%
100
AM Subtype
CK.AMs.c1
CK.AMs.c2
CK.AMs.c3
CK.AMs.c4
C
CXCR4
CCL4
CXCL10
CXCL3
Expression
High
Low
D
Differentially Expressed Genes
Curated Genes
CK.AMs.c1
CK.AMs.c2
CK.AMs.c3
CK.AMs.c4
PDK4
CXCR4
TXNIP
ZFP36L2
APOE
CD84
CCL4
CCL20
CCL4L2
IL1B
CCL3
CXCL8
CXCL10
CXCL11
GBP1
GBP5
WARS
GBP4
THBD
MCEMP1
TGM2
RGCC
THBS1
GPD1
CCL18
CCL23
CXCR4
CCL4
CCL20
CCL4L2
CCL3
CXCL8
CCL3L1
CXCL2
CXCL5
CCRL2
CCR1
ACKR3
CXCL10
CXCL11
CXCL9
CXCL16
CXCL3
Average Expression
2.5
2
1
0
−1
Percent Expressed
0
25
50
75
100
E
Gene Ontology: BP
Genitalia Morphogenesis
Organic Acid Transport
Neutrophil Degranulation
Cellular Response to Lipopolysaccharide
Type I Interferon Signaling Pathway
Positive Regulation of Apoptotic Process
Detoxification of Copper Ion
Negative Regulation of Viral Genome Replication
Negative Regulation of Transcription by RNA Polymerase II
Blood Vessel Remodeling
Regulation of Cholesterol Storage
Antigen Processing and Presentation of Exogenous Peptide Antigen via MHC Class II
Positive Regulation of Phagocytosis
Positive Regulation of Transcription by RNA Polymerase II
Regulation of Alternative mRNA Splicing, via Spliceosome
AMs.s1
AMs.s2
AMs.s3
CK.AMs.c1
CK.AMs.c2
CK.AMs.c3
CK.AMs.c4
MT.AMs
IFN.AMs
Cyc.AMs
IGF1.AMs
CCL18.AMs
CXCL5.AMs
Cholesterol.AMs
GF.AMs
SNHG.AMs
P Value (elim)
0
5
>=10
F
Regulon
HMGN3
TBX21
DLX3
OVOL1
EOMES
BACH2
HNF1A
E2F7
ETS1
IRF7
MAZ
HOXB7
HES2
ZFY
SREBF2
ZSCAN31
EGR1
AMs.s1
AMs.s2
AMs.s3
AMs.s4
CK.AMs.c1
CK.AMs.c2
CK.AMs.c3
CK.AMs.c4
MT.AMs
IFN.AMs
Cyc.AMs
IGF1.AMs
CCL18.AMs
CXCL5.AMs
Cholesterol.AMs
SNHG.AMs
GF.AMs
Specificity Score
−2.5
0
>=2.5

Fig 4


A
HC
CF
UMAP 2
UMAP 1
HC1
HC2
HC3
HC4
CF1
CF2
CF3
B
Conv.Mono
Int.Mono
FCN1.Mono
CCL2.Mono
CXCL5.Mono
CXCL9.Mono
CCL13.Mono
VEGFA.Mono
Mono Subtype
C
Mono Subtype Percentage
Subject
0%
25%
50%
75%
100
D
Conv.Mono
Int.Mono
FCN1.Mono
CCL2.Mono
CXCL5.Mono
CXCL9.Mono
CCL13.Mono
VEGFA.Mono
APOC1
C1QA
UBB
CD52
ALDH2
MARCO
CLEC10A
FPR3
EMP1
GPR183
LGALS1
PLXNC1
PLAC8
FCN1
THBS1
RETN
MNDA
AC020656.1
CCL2
CD48
RPL13
RPL36A
RPL10
RPS8
PPBP
CXCL5
FN1
C1QB
MSR1
SPOCD1
SLAMF7
LGALS2
CXCL9
IL4I1
CXCL10
CCL13
TMEM37
GAL3ST4
MERTK
LGMN
LILRB5
ADAM19
PKIB
ADAM8
VEGFA
Average Expression
Percent Expressed
E
Gene Ontology: BP
Regulation of Cholesterol Storage
Positive Regulation of Steroid Metabolic Process
Positive Regulation of Protein Autophosphorylation
Regulation of Cell Shape
Neutrophil Degranulation
Defense Response to Bacterium
Translational initiation
SRP-Dependent Cotranslational Protein Targeting to Membrane
Platelet Degranulation
Regulation of Complement Activation
Interferon-Gamma-Mediated Signaling Pathway
Low-Density Lipoprotein Particle Clearance
Cellular Iron Ion Homeostasis
P Value (elim)
F
Regulon
EMX2
ZNF625
SIX3
ZNF560
TEAD2
SIX1
SETDB1
ESRRA
GABPA
ZBTB7B
IRF1
ETV7
HOXC4
ZNF729
ZNF136
ZFP42
Specificity Score

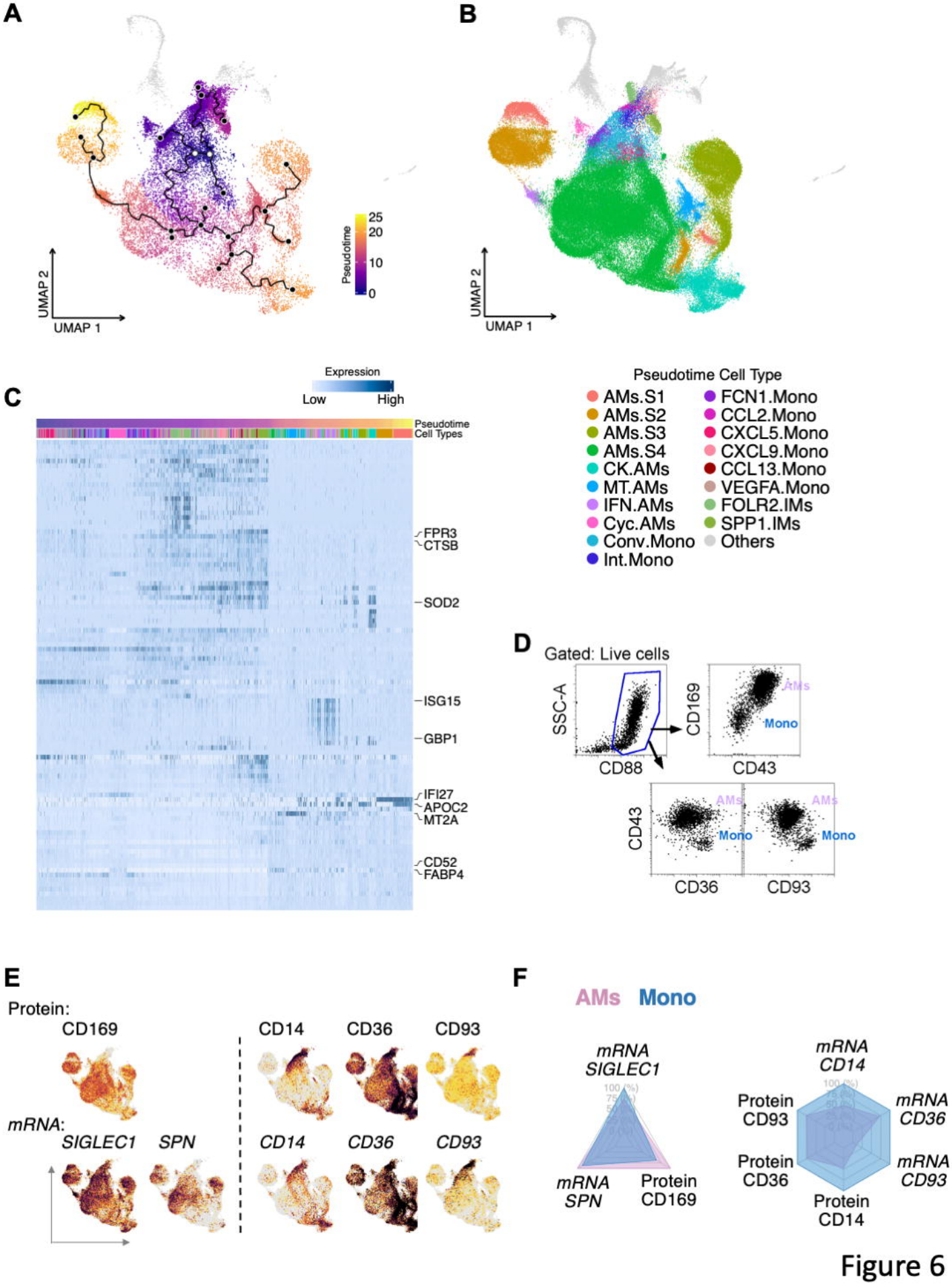

A
B
UMAP 2
UMAP 1
Pseudotime
25
20
10
0
C
Expression
Low
High
Pseudotime
Cell Types
FPR3
CTSB
SOD2
ISG15
GBP1
IFI27
APOC2
MT2A
CD52
FABP4
Pseudotime Cell Type
AMs.S1
AMs.S2
AMs.S3
AMs.S4
CK.AMs
MT.AMs
IFN.AMs
Cyc.AMs
Conv.Mono
Int.Mono
FCN1.Mono
CCL2.Mono
CXCL5.Mono
CXCL9.Mono
CCL13.Mono
VEGFA.Mono
FOLR2.IMs
SPP1.IMs
Others
D
Gated: Live cells
SSC-A
CD88
CD169
CD43
AMs
Mono
CD36
CD93
E
Protein:
CD169
CD14
CD36
CD93
mRNA:
SIGLEC1
SPN
CD14
CD36
CD93
F
AMs
Mono
mRNA SIGLEC1
mRNA SPN
Protein CD169
mRNA CD14
mRNA CD36
mRNA CD93
Protein CD14
Protein CD36
Protein CD93

Figure 6

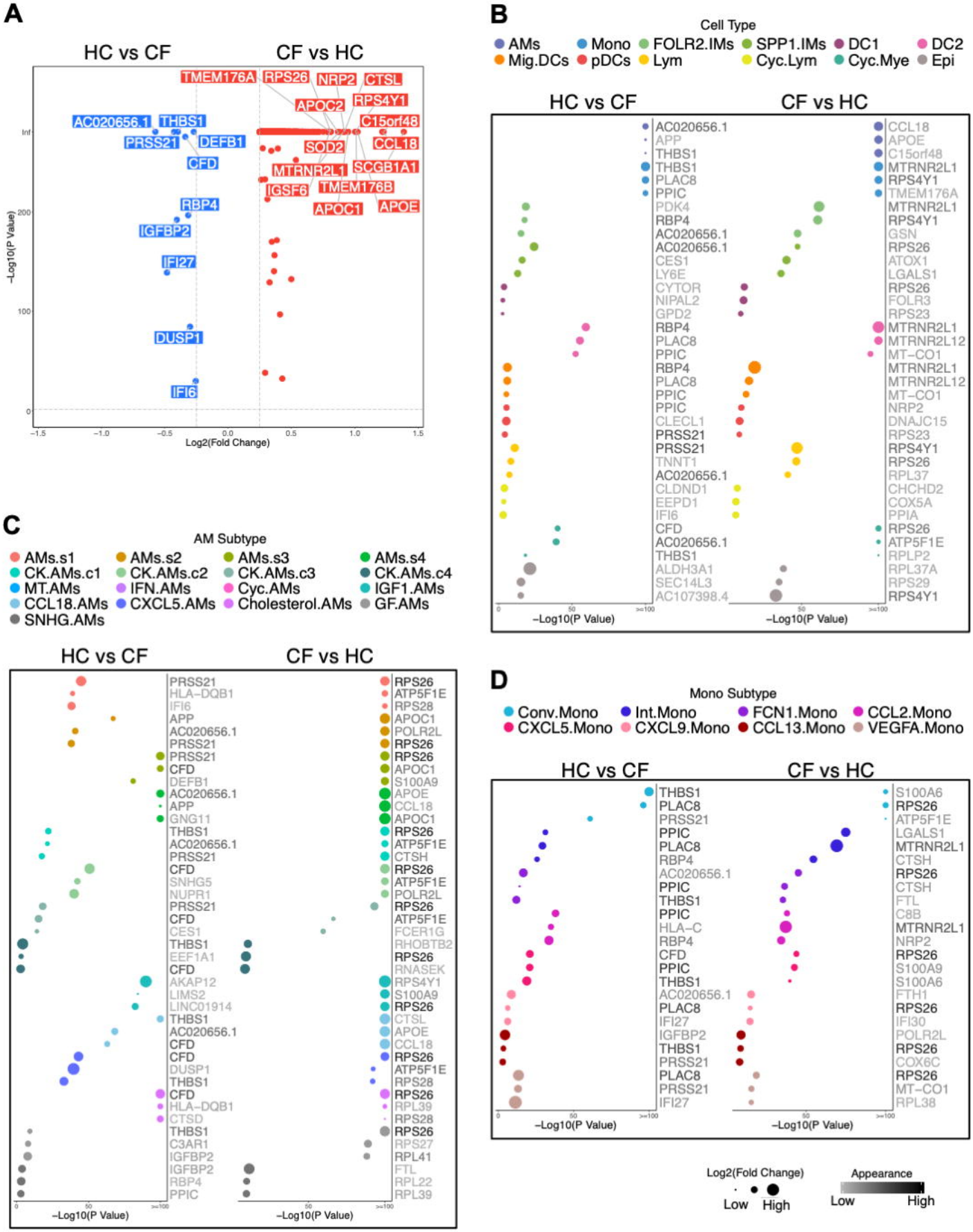

A
HC vs CF
CF vs HC
TMEM176A
RPS26
NRP2
CTSL
APOC2
RPS4Y1
AC020656.1
THBS1
C15orf48
PRSS21
DEFB1
SOD2
CCL18
CFD
MTRNR2L1
SCGB1A1
IGSF6
TMEM176B
RBP4
APOC1
APOE
IGFBP2
IFI27
DUSP1
IFI6
−Log10(P Value)
Log2(Fold Change)
B
Cell Type
AMs
Mono
FOLR2.IMs
SPP1.IMs
DC1
DC2
Mig.DCs
pDCs
Lym
Cyc.Lym
Cyc.Mye
Epi
HC vs CF
CF vs HC
−Log10(P Value)
C
AM Subtype
AMs.s1
AMs.s2
AMs.s3
AMs.s4
CK.AMs.c1
CK.AMs.c2
CK.AMs.c3
CK.AMs.c4
MT.AMs
IFN.AMs
Cyc.AMs
IGF1.AMs
CCL18.AMs
CXCL5.AMs
Cholesterol.AMs
GF.AMs
SNHG.AMs
HC vs CF
CF vs HC
−Log10(P Value)
D
Mono Subtype
Conv.Mono
Int.Mono
FCN1.Mono
CCL2.Mono
CXCL5.Mono
CXCL9.Mono
CCL13.Mono
VEGFA.Mono
HC vs CF
CF vs HC
−Log10(P Value)
Log2(Fold Change)
Low
High
Appearance
Low
High

Figure 7