

# Single cell RNA sequencing identifies unique inflammatory airspace macrophage subsets

Kara J. Mould, ... , Max A. Seibold, William J. Janssen

*JCI Insight.* 2019. <https://doi.org/10.1172/jci.insight.126556>.

**Research** In-Press Preview Immunology Pulmonology

Macrophages are well-recognized for their dual roles in orchestrating inflammatory responses and regulating tissue repair. In almost all acutely inflamed tissues, two main subclasses of macrophages co-exist. These include embryonically-derived resident tissue macrophages and bone marrow-derived recruited macrophages. While it is clear that macrophage subsets categorized in this fashion display distinct transcriptional and functional profiles, whether all cells within these categories and in the same inflammatory microenvironment share similar functions or whether further specialization exists has not been determined. To investigate inflammatory macrophage heterogeneity on a more granular level, we induced acute lung inflammation in mice and performed single cell RNA sequencing of macrophages isolated from the airspaces during health, peak inflammation, and resolution of inflammation. In doing so, we confirm that cell origin is the major determinant of AM programing and describe two previously uncharacterized, transcriptionally distinct subdivisions of AMs based on proliferative capacity and inflammatory programing.

**Find the latest version:**

<https://jci.me/126556/pdf>



# **Single cell RNA sequencing identifies unique inflammatory airspace macrophage subsets**

**Authors:** Kara J. Mould,<sup>1,2,3\*</sup> Nathan D. Jackson,<sup>3,4</sup> Peter M. Henson,<sup>2,5</sup> Max Seibold,<sup>2,4,5,6</sup> William J. Janssen<sup>1,2,6</sup>

## **Author Affiliations:**

<sup>1</sup>Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, National Jewish Health, Denver, CO, USA

<sup>2</sup>Division of Pulmonary Sciences and Critical Care Medicine, Department of Medicine, University of Colorado – Anschutz Medical Campus, Aurora, CO, USA

<sup>3</sup>These authors contributed equally

<sup>4</sup>Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA

<sup>5</sup>Program for Cell Biology, Department of Pediatrics, National Jewish Health, Denver, CO, USA

<sup>6</sup>These authors contributed equally

**\*Corresponding Author:** Kara J. Mould, National Jewish Health, 1400 Jackson Street, Denver, CO 80206. Telephone: 303-398-1787. Email: [MouldK@NJHealth.org](mailto:MouldK@NJHealth.org). ORCID 0000-0003-2967-2778.

The authors have declared that no conflict of interest exists.

## **ABSTRACT**

Macrophages are well-recognized for their dual roles in orchestrating inflammatory responses and regulating tissue repair. In almost all acutely inflamed tissues, two main subclasses of macrophages co-exist. These include embryonically-derived resident tissue macrophages and bone marrow-derived recruited macrophages. While it is clear that macrophage subsets categorized in this fashion display distinct transcriptional and functional profiles, whether all cells within these categories and in the same inflammatory microenvironment share similar functions or whether further specialization exists has not been determined. To investigate inflammatory macrophage heterogeneity on a more granular level, we induced acute lung inflammation in mice and performed single cell RNA sequencing of macrophages isolated from the airspaces during health, peak inflammation, and resolution of inflammation. In doing so, we confirm that cell origin is the major determinant of AM programing and describe two previously uncharacterized, transcriptionally distinct subdivisions of AMs based on proliferative capacity and inflammatory programing.

## INTRODUCTION

Macrophages perform multiple functions during acute inflammation, including production of inflammatory cytokines, phagocytosis of cellular debris and production of reparative molecules (1-7). As such, they are attractive candidates for therapeutic intervention in inflammatory diseases. To date, however, targeting of macrophages has been hindered by lack of precise understanding of their dynamic programs in inflamed tissues. For instance, it is unclear whether the seemingly divergent inflammatory and reparative functions are performed by macrophages within a shared microenvironment or whether discrete populations of macrophages provide ‘division of labor’ during acute inflammation.

In virtually all tissues, two main subsets of macrophages exist during acute inflammation. These include resident macrophages, which populate the tissues during homeostasis, and recruited macrophages that arise from circulating monocytes. In the lungs, resident airspace macrophages (RAMs) arise during embryogenesis and self-renew throughout life without replacement from circulating monocytes. They reside in the airspaces prior to the onset of inflammation, persist during the entirety of the inflammatory response, and remain after it resolves (8-11). In comparison, recruited airspace macrophages (RecAMs) originate from the bone marrow, traffic to the site of inflammation, and ultimately undergo apoptosis (10, 12, 13). Based on ‘bulk’ RNA sequencing (RNA-seq), we recently demonstrated that during endotoxin-induced acute inflammation, RAMs proliferate locally and have relatively stable programming (14). Conversely, the programming of RecAMs is dynamic and is characterized by production of pro-inflammatory cytokines during peak inflammation and expression of reparative factors during lung repair. However, whether these functions are shared by all cells within each group, or whether intra-group sub-specialization of function occurs remains unknown.

Macrophage programming has traditionally been described along a dichotomous spectrum of polarization. Historically, macrophage polarization denoted the opposing responses of cells to different pathogens – M1 programming described the response to bacterial or viral infections, often modeled *in vitro* with LPS or IFN $\gamma$ , while M2 programming described the response to parasitic infections, often modeled by exposure to IL-4 or IL-13 (15, 16). Later, similarities in features of M2 programming and those present during wound repair led to the evolution of ‘M2’ as a descriptor of reparative, anti-inflammatory macrophages (5, 17, 18). Innumerable studies have sought to identify specific markers of these polarization states, and have ascribed phenotypes to macrophages using selected genes or gene products. However, recent evidence suggests that macrophage programming is multi-dimensional, dynamic and exceedingly complex, and therefore unsuitable for a simplified, bipolar scheme (19, 20). To date, whether specific pro-inflammatory or anti-inflammatory programs correlate with markers of macrophage polarization has not been determined at the single cell level.

In this study, we aimed to identify heterogeneity within the inflammatory AM pool at single cell resolution. We used single cell RNA-seq to explore the transcriptional signatures of AMs within the murine lung during three discrete time points: homeostasis, peak neutrophilic inflammation, and resolution of inflammation. In doing so, we confirm that cell origin drives AM heterogeneity and describe two previously unrecognized sub-groups of RAM and RecAM cells specializing in proliferation and inflammatory signaling, respectively.

## Results

### Single cell transcriptional profiling identifies five discrete AM populations across homeostasis, acute inflammation, and resolving inflammation

We performed single cell RNA-seq of all macrophages isolated from the airspaces of C57BL/6 mice treated with intratracheal LPS. Three time points, representing homeostasis (day 0), peak of neutrophilic inflammation (day 3), and resolution of inflammation (day 6) were analyzed (12, 21). DAPI-, Ly6g-, F480+, CD64+ cells (containing both resident and recruited AM subsets) were isolated from lavage using FACS (Supplemental Figure 1). We analyzed profiles from 902 cells that passed strict quality-control thresholds and used a combination of t-distributed stochastic neighbor embedding (tSNE) and shared nearest neighbor clustering carried out with the 5,784 most variant genes across cells from all time points to identify robust cell clusters.

Five transcriptionally distinct clusters were revealed (Figure 1a). All clusters expressed macrophage-specific markers, including *CD68*, *Lgals3*, *MARCO*, and *MerTK* (Figure 1b). Overlaying time course information on the tSNE plot revealed that clusters 1 and 2 contained AMs from each of the three time points. In comparison, clusters 3 and 4 were composed almost exclusively of AMs from day 3, whereas cluster 5 contained AMs from day 6 (Figure 1c). During homeostasis more than 90% of cells belonged to cluster 1, while nearly all remaining cells were in cluster 2 (Figure 1d). In contrast, during peak inflammation (day 3), the majority of cells were members of clusters 3 or 4, with the remainder coming from clusters 1 and 2. During the resolution of inflammation (day 6), clusters 3 and 4 essentially disappeared from the sample, and cells were segregated to cluster 5 (40%) or clusters 1 and 2 (59%).

## **AM populations revealed by single cell RNA-seq reflect cell origin**

Since clusters 1 and 2 were present during homeostasis and remained throughout the inflammatory time points, we hypothesized that these populations represented the resident airspace macrophages (RAMs). Likewise, we postulated that clusters 3, 4, and 5 largely consisted of RecAMs, as they were only present during inflammation. To test these characterizations, we examined whether known RAM and RecAM markers were differentially expressed between the two hypothesized groups. We found that RAM marker genes, including *Mrc1* (CD206), *Itgax* (CD11c), *SiglecF*, and *Siglec1* (CD169) (9, 14, 22), were significantly upregulated in clusters 1 and 2 compared to clusters 3, 4, and 5. In comparison, RecAM marker genes, including *CD14*, *Ly6c1* (Ly6c), *ApoE*, *Ccr5*, *Mafb*, and *Sell* (L-selectin) (9, 14, 23), were significantly upregulated in clusters 3, 4, and 5 compared to clusters 1 and 2 (Figure 2a-b). Furthermore, mean expression across the panel of RAM markers was 1.6-fold higher in clusters 1 and 2, whereas mean expression of the panel of RecAM markers was 1.4-fold higher in clusters 3, 4, and 5 (Figure 2b-c). This confirms that cell origin is a major determinant of AM heterogeneity.

## **Macrophage polarization states are not mutually exclusive and reflect cell origin**

The presence of subgroups within the populations expressing RAM markers (clusters 1 and 2) and RecAM markers (clusters 3 and 4) during peak inflammation led us to question whether this heterogeneity could be ascribed to ‘classically activated’ (M1) or ‘alternatively activated’ (M2) polarization states (24). We first examined the expression of *Nos2* and *Arg1*, two commonly used markers to distinguish M1 and M2 activation states (25, 26). Expression of both enzymes was observed in all cell clusters but was greatest in cluster 3 (Figure 3a-b). Although M1 and M2 activation states and their corresponding transcriptional markers are often considered to be mutually exclusive (24, 26-29), we observed extensive co-expression of *Arg1* and *Nos2* within the same cells

(Figure 3c), consistent with non-exclusivity of M1 and M2 programming that has been suggested by our group and others (30-33). We next examined the mean expression of a comprehensive panel of traditional M1 and M2 markers across clusters and found that clusters 1 and 2 exhibited 1.3-fold higher expression of M2 genes compared to clusters 3 and 4, while clusters 3 and 4 expressed M1 genes at a 1.2-fold higher level compared to clusters 1 and 2 (Figure 3d). Thus, cells expressing RecAM markers at peak inflammation had the highest expression of M1 genes, whereas cells expressing RAM markers during both homeostasis and inflammatory time points expressed the highest levels of M2 genes. Notably, cluster 5 exhibited relatively low expression of both M1 and M2 genes. Taken as a whole, these data show that aggregate expression of M1 and M2 genes corresponds closely to cell origin, irrespective of inflammatory time points.

Because homeostatic RAMs (day 0) should not be activated, we next examined mean expression of the gene panels across clusters and days using homeostatic RAMs (day 0) as a reference. We found that mean expression of M2 panel genes was similar across all time points for cells expressing RAM markers and was unchanged from homeostasis (Figure 3e). Conversely, mean expression of M2 panel genes was lower in cells expressing RecAM markers. Interestingly, cells expressing either RAM or RecAM markers at day 3 showed increased expression of M1 genes compared to day 0. Thus, cells that express RAM markers at day 3 exhibit dual expression of both M1 and M2 panel genes, although expression of the latter is similar to homeostatic RAMs. Notably, there was no correlation between *Arg1* expression and M2 panel genes or *Nos2* expression and M1 panel genes. Together, these data demonstrate that expression of M1 and M2 genes is not mutually exclusive and that single genes cannot be used to ascribe discrete polarization states to macrophages.

**Resident AMs contain a proliferative subpopulation present during health and inflammation.**

As a next step we sought to define the characteristics that distinguish individual clusters. We started with clusters 1 and 2 since they contained cells from all time points and had similar expression of RAM and polarization markers. In comparison to clusters 3, 4 and 5, cluster 1 contained 314 genes with significantly elevated expression and cluster 2 contained 572 such genes. Of these differentially expressed genes (DEGs), 182 were shared between clusters 1 and 2. Pathway analysis of DEGs from cluster 1, identified pathways known to be important for RAM programing, such as peroxisome proliferator-activated receptor (PPAR) signaling, lipid metabolism, and protein processing in the endoplasmic reticulum (Figure 4a, Supplemental Table 1) (34, 35). DEGs from cluster 2 were enriched for similar annotations (Figure 4A), along with multiple proliferation and cell cycle pathways (Figure 4B, Supplemental Table 2).

We next compared clusters 1 and 2 to each other. Intriguingly, cluster 1 contained no DEGs compared to cluster 2. By contrast, cluster 2 exhibited 469 upregulated DEGs, the top three of which were *Ki67*, *Pclaf*, and topoisomerase 2a (*Top2a*) (Figure 4c-e, Supplemental Table 3). Widely recognized as markers of cell proliferation, these genes were significantly upregulated in cluster 2 compared to the other clusters in the dataset. Notably, cluster 2 contained 6.5% of cells analyzed at day 0, 6.9% at day 3, and 2.1% at day 6 (Figure 1d.) These data are consistent with previously published reports of RAM proliferation based on Ki67 staining, BrdU incorporation, or [<sup>3</sup>H]thymidine labeling (10, 14, 36)

Since the vast majority of DEGs in cluster 2 were involved in regulation of the cell cycle, we sought to confirm that these genes were principally responsible for the separation of this cluster. Accordingly, the expression signature of proliferation was regressed out of the dataset, after which clustering was re-performed. As shown in Figure 4F-H, this resulted in the disappearance of the original cluster 2 and the re-assignment of most of its members to cluster 1, suggesting that proliferation genes were largely what separated cluster 2 from other cells in cluster 1. Together, these data identify for the first

time a dedicated subpopulation of proliferative AMs, with an expression profile characteristic of RAMs, present during health and disease.

Although clusters 1 and 2 contained cells isolated from three vastly different microenvironments (i.e. days 0, 3 and 6 of inflammation) relatively little partitioning of cells was apparent on the tSNE plot (Figure 1c). However, quantitative assessment of DEGs in cluster 1 revealed upregulation of 53, 414, and 39 genes at days 0, 3, and 6 (compared to the other time points), respectively. The DEGs at day 0 were enriched in pathways for lipid metabolism and transcription (Supplemental Table 4). Cells from day 3 were the most unique, with enrichment of pathways related to proteasome function, Toll-like receptor signaling, and aerobic metabolism (Supplemental Table 5). At day 6, pathways for post-translational modifications of proteins were enriched (Supplemental Table 6). Because cluster 2 contained only 43 cells, similar analyses of DEGs over time were limited by statistical power. However, a similar trend in DEGs across timepoints was observed in cluster 2. Taken as a whole, these data demonstrate that cluster 1 cells maintain their RAM identity while modifying gene function to fit the changing microenvironment posed by inflammation.

### **Novel macrophage subsets are identified during inflammation**

We next investigated the defining characteristics of clusters 3 and 4, which exhibited profiles typical of RecAMs and were only present during peak inflammation. Comparison of clusters 3 and 4 to other clusters yielded 574 upregulated genes. Top pathways enriched with these DEGs involved inflammation-related terms such as NF-kappa B signaling pathway, Toll-like receptor signaling, interferon signaling, and the ER-phagosome pathway (Figure 5a, Supplemental Table 7). Pathway analysis of 543 DEGs upregulated in cluster 3 compared to cluster 4 showed that cluster 3 is distinguished by proteolysis, antigen presentation, and oxidative metabolism (Figure 5b, Supplemental Table 8). In contrast, the 43 genes with increased expression in cluster 4 compared to

cluster 3 were enriched in pathways specific to pro-inflammatory cytokine production and cell survival (Figure 5c, Supplemental Table 9), suggesting that these cells are specialized for inflammatory signaling. Gene level analysis confirmed elevated expression of several important inflammatory cytokines in cluster 4, including *CXCL2*, *IL1b* and *TNF* (Figure 5d-f). Together, these data identify two distinct clusters of cells that express RecAM markers at the peak of inflammation: one defined by inflammatory cytokine production and one defined by protein processing and proteasome functions. Intriguingly, while RecAMs are known to undergo apoptosis (12, 13), these two clusters have differential expression of apoptotic genes, suggesting they may also have different propensities for programmed cell death. To our knowledge, such subsets have not been previously described.

In contrast to the two RecAM marker-enriched clusters present at peak inflammation, cluster 5, which contained the RecAM-like cells present at Day 6, was distinguished from all other clusters by upregulation of 450 DEGs enriched in pathways involved in production of growth factors, secretory granule formation, lysosome processing, and PIK3-AKT signaling (Figure 5g, Supplemental Table 10). In addition, pathways important for actin binding and cell migration were also identified. Further analysis of gene expression variation within cluster 5 did not reveal further subclustering of cells.

### **Unique transcriptional programs define macrophage subsets**

Having identified genes that distinguished clusters from one another, we next plotted the scaled expression of the top DEGs in each cluster across all cells to visualize the extent to which patterns of expression among these cluster-defining genes were shared or unique in respect to other clusters (Figure 6, Supplemental Table 11). Examination of the resultant heat map confirmed that genes related to cell cycle were almost exclusively expressed in cluster 2. In comparison, the DEGs that defined cluster 1 were also expressed by cells from cluster 2. Likewise, the DEGs that defined cluster 4 were also expressed by cells from cluster 3, whereas the pro-inflammatory DEGs that defined

cluster 3 were less highly expressed by cells in cluster 4 and downregulated in other clusters. Finally, genes encoding the growth factors and reparative molecules during resolution phase RecAM in cluster 5 were slightly increased in RAM (clusters 1 and 2), but were not increased in the RecAM clusters that dominated during peak inflammation (clusters 3 and 4). As a whole, these patterns illustrate that each cluster is defined by a unique transcriptional program that imparts specific functions and ‘divisions of labor’ throughout the inflammatory process.

### **Gene networks from ‘bulk’ RNA-seq are co-expressed within single cell RNA-seq-based clusters**

Because single cell sequencing is inherently prone to “dropout,” or incomplete detection of genes expressed at low levels, we sought to integrate our single cell data with our previously published ‘bulk’ RNA sequencing of RAMs and RecAMs (14, 37). In this way, we aimed to infer a more complete transcriptional profile to our newly identified single cell clusters. In addition, by comparing the two data sets, we sought to determine the extent that clustering of our single cell data explained the co-regulation of genes identified in our ‘bulk’ dataset and validate that cell origin is a major determinant of AM programming. Finally, we assessed the technical reproducibility of data from the two sequencing platforms using the same model system.

We first compared the transcriptional content of the single cell dataset to our previously published ‘bulk’ RNA sequencing data of RAMs and RecAMs sorted from BAL at equivalent time points following LPS-induced inflammation (14). Overall, 3,494 QCed genes in the bulk dataset were not present in the QCed single cell dataset, of which 1,521 had canonical gene names. Although, the majority of these missing genes were expressed at low levels in the ‘bulk’ dataset, 148 of these genes were detected at levels above the median level of expression of the ‘bulk’ data. While pathway analyses of these genes did yield enrichments, several genes encoding important macrophage

effector proteins were identified (*Myd88*, *Tlr9*), suggesting that they were prone to drop-out with single cell technology.

We next performed weighted gene co-expression network analysis (WGCNA) of the ‘bulk’ RNA sequencing data. Eleven distinct gene co-expression networks were identified based on samples collected at days 0, 3, and 6 (Supplemental Figure 2), three of which were upregulated in RAMs, five of which were upregulated in RecAMs, and three of which were upregulated in cells belonging to both groups (Supplemental Figure 3). Overall, strong correlation between ‘bulk’ and single cell heterogeneity was observed, with cells expressing a given network largely limited to one or two clusters.

Pathway analysis carried out on gene networks also revealed broad correspondence between ‘bulk’ and single cell datasets (Figure 7, Supplemental Table 12), while enhancing understanding of the processes and genes that define cell clusters. For instance, two networks were characteristic of RecAMs at day 3: the seagreen network was enriched in inflammatory signaling pathways, glucose metabolism, and apoptosis pathways, while turquoise was enriched in proteasome activity, interferon signaling, and mitochondrial respiration. Interestingly, there was no major difference in the average expression of these two networks between single cell clusters 3 and 4, even though pathway analysis based on DEGs from the clusters (Figure 4b) predicts that these networks would define them.

Two networks were most highly expressed among RecAMs at day 6 in the ‘bulk’ dataset, both of which were enriched for pathways involved in cell growth and repair (olive green and yellow). Both networks were highly and near exclusively expressed by cluster 5 of the single cell dataset.

Three networks were characteristic of RAMs. One network (mauve) had pathways enriched in lipid metabolism and PPAR signaling and was expressed most highly among day 0 cells contained within

clusters 1 and 2. Likewise, the pink network, enriched in pathways related to RNA modification, mitochondrial respiration, and proteasome activity, was most highly expressed in RAMs at day 3. This network appeared to be driven primarily by environment, rather than cell origin, in that its expression was more similar to RecAMs from the same timepoint (clusters 3 and 4) compared to RAMs at other timepoints. The third network (brown), enriched in cell proliferation genes, was strongly characteristic of cluster 2.

By comparing co-regulated gene networks in both ‘bulk’ and single cell datasets, the relative contributions of cell origin and environment to the overall heterogeneity and discrete functions that define each AM cluster can begin to be ascertained. Further, given the high overlap of ‘bulk’ sequencing-defined modules, which contained many genes not detected by the single cell sequencing platform, the integration of these technologies provides significantly more genic definition to the single cell dataset and validates this approach to more fully inform the biology of single cell populations.

## Discussion

In this study, we applied single cell RNA-seq technology to re-examine and refine our previous characterization of AM heterogeneity in a self-resolving model of lung inflammation. In doing so, we have identified five discrete populations of airspace macrophages that exist in the lungs during acute inflammation, each with unique transcriptional signatures and putative functions. Three populations contain cells that express RecAM markers and are only present during inflammation, whereas two populations contain cells that express RAM markers are present during both inflammation and homeostasis. While this study is not sufficient to prove cell origin, together with our previous finding that embryonic versus post-natal origin dictates the majority of AM programming, it provides compelling supporting evidence. Further, we have identified additional determinants of cell heterogeneity. This includes a proliferative subset of RAMs that is present during all time points, and novel subsets of pro-inflammatory and proteolytic RecAMs present only during peak inflammation. A uniform group of reparative RecAMs present during resolution of inflammation is also identified.

The existence of a proliferative pool of RAMs is consistent with their known self-renewing properties (10, 36, 38-42). RAMs populate the airspaces during embryogenesis and maintain a steady population throughout life with little contribution from bone marrow-derived cells (11). During acute inflammation, they continue to proliferate and remain after inflammation has resolved. To our knowledge, our data provide the first transcriptional characterization of the proliferative subset of RAMs during both health and disease. As has been previously shown, we found overall low expression of transcriptional repressors of proliferation, *Mafb* and *Maf* (cMaf), in RAMs (43). However, in contrast to macrophages in other organs, levels of *Mafb* and *Maf* were not further decreased in actively proliferating (cluster 2) versus quiescent (cluster 1) RAM subsets (43). Rather, our data demonstrate that *Mafb* and *Maf* gene expression levels are largely unchanged throughout

inflammation. Notably, our data do not distinguish between a specialized stem-like population of cells, or as suggested by previous studies, a regular proportion of cells stochastically entering and leaving the cell cycle (10, 36) Because the proliferative RAMs we identified show only minor changes in expression of gene transcripts that encode cell surface proteins, studies using cell labeling techniques or proliferation reporters will likely be required to further investigate the specific stimuli that signal AMs to maintain the niche during homeostasis and inflammation.

Strikingly, despite existing in drastically different microenvironments during homeostasis, peak inflammation, and lung repair, RAMs displayed remarkably stable programing. The largest differences among RAMs were observed at peak inflammation, where proteasome and immune response genes were upregulated, echoing the concomitant, but more dramatic patterns observed in RecAMs. While it is possible that direct exposure to LPS had significant effects on the transcriptional signature of RAMs at earlier time points that were not included in our dataset, our study shows that after three days, any major transcriptional consequences of LPS treatment have resolved, underscoring the role of resident macrophages in maintenance of tissue homeostasis (7, 8, 44). Besides highly proliferative RAMs, no other obvious subsets of RAMs were detected. Specifically, we did not identify the subset of sessile RAMs suggested by Westphalen et. al, which may be important in immunosuppression (2). The absence of a sessile RAM subpopulation (characterized by high expression of connexins and calcium signaling molecules) may reflect the relative ineffectiveness of lavage for retrieving cells from the alveolus (45). Future comparison of cell populations isolated from lavage versus cells isolated by lung digestion will be important to expand the repertoire of AM functions during inflammation.

The finding of two subsets of RecAMs during peak neutrophilic inflammation has major implications for macrophage programing during inflammatory lung disease. RecAMs are thought to derive from circulating monocytes that are recruited to inflamed tissues in response to specific chemoattractant molecules. RecAMs persist during inflammation and repair, then undergo apoptosis, leaving very few

bone marrow-derived AMs once homeostasis has been restored (11-13, 46-48). Our previous ‘bulk’ sequencing data identified RecAMs as pro-inflammatory and suggested that they are responsible for the majority of pro-inflammatory cytokines produced by macrophages during peak inflammation. However, our single cell data indicate that a distinct subset of RecAMs, cluster 4, is largely responsible for cytokine production. In comparison, cluster 3 is characterized by expression of enzymes involved in protein processing and antigen presentation, raising the possibility that cells in this cluster are programmed for functions distinct from those of cluster 4 - despite existing in a shared microenvironment. Indeed, altered ratios of constitutive versus immunoproteasomal subunit transcripts in cluster 4 compared to cluster 3 are in keeping with the emerging role of proteasomal protease activities as master regulators of the inflammatory response (49-53). Given that several of the DEGs that distinguish cluster 4 from cluster 3 are also expressed in neutrophils at steady-state (*S100a9*, *Lcn2*, *Il2r*) (38) we considered whether this population may represent macrophage/neutrophil complexes (i.e. “doublets” containing two cells bound together). However, we found no differences in total expressed genes between clusters 3 and 4, as would be expected in the case of cell doublets. We also considered the possibility that the cells in cluster 4 were macrophages that phagocytosed neutrophils. However, many other genes that are highly expressed in neutrophils such as *Ly6G* were absent. Moreover, there was no enrichment in phagocytosis pathways in cluster 4 versus cluster 3. Rather, the high expression of protein processing and proteolysis genes in cluster 3 suggest this population may be poised for active phagocytosis. We suspect that the genes in question are truly expressed by this pro-inflammatory macrophage subset, and that this reflects their monocytic origin. Interrogation of the Immgen database (54) lends support for this concept, and demonstrates that the genes are highly expressed by Ly6C<sup>hi</sup> bone marrow monocytes.

In addition to differences in inflammatory programming, RecAM subclusters at day 3 may have differential cell fate. Although it is commonly held that monocytes arrive in bulk during the initial phase of inflammation, it is currently unknown whether the RecAMs that exist at day 6 derive from those

present at day 3, or whether ongoing turnover and replacement of RecAMs occurs. One possibility is that clusters 3 and 4 represent different phases of monocyte influx into the lungs. In this context, cluster 3 may contain mature RecAMs that are destined for apoptosis prior to day 6, while cluster 4 may represent more recently recruited AMs that mature to reparative RecAMs at day 6. In support of this hypothesis, cluster 3 demonstrates increased expression of pro-apoptotic caspase 8, while cluster 4 is notable for increased expression of anti-apoptotic molecules, *Bcl2L1*, *Ier1*, mitogen-activated protein kinase 2 (MAPK2), and *Mapk14* (p38) (55). Likewise, cluster 4 cells express high levels of *Nrf2* and its downstream products, *Slc7a11* and *Sipi*, genes implicated in cytoprotective and antioxidant functions (56).

Alternatively, the basis for heterogeneity within RecAMs may be explained by heterogeneity of circulating monocytes. Recent studies investigating monocyte progenitors have identified two distinct sources for circulating monocytes: granulocyte-monocyte progenitors (GMPs) and monocyte-dendritic cell progenitors (MDPs). GMPs produce neutrophil-like, inflammatory monocytes, while MDPs produce a subset of monocyte-derived dendritic cells (57). While the markers used to identify these populations in the bone marrow are not clearly confined to either clusters 3 or 4, the division of pro-inflammatory cytokine production and antigen-presenting functions between clusters aligns with the observed neutrophil-like GMP and dendritic cell-like MDP-derived monocyte phenotypes. Identification of markers to distinguish cells within clusters 3 and 4 will allow for further study of the source and fate of these two subsets.

Single cell analysis allows for direct correlation of traditional markers of macrophage programming with cell function. In our LPS model, we see the majority of M1 marker expression present during peak inflammation with a return to baseline expression of M1 markers during resolution of inflammation (Figure 3F). This trend is most pronounced in RecAMs, but also present in the RAM populations. Notably, we detected the expression of several M2 activation markers in homeostatic RAMs. In fact,

mean expression of our panel of M2 markers was significantly higher in naïve (Day 0) RAMs compared to cells captured at peak inflammation or resolution time points (one-sided t-test p-values = 4.54e-21 and 2.28e-04, respectively). This invalidates these transcripts as markers of cell activation, and rather suggests they represent markers of the normal process of differentiation into mature macrophages. Importantly, M2 marker expression did not correlate with expression of growth factors and molecules thought to be important for tissue repair in cluster 5. Together, these analyses complement an emerging literature that suggests that the traditional “M1 versus M2” dichotomy derived from in vitro studies is inadequate in ascribing a molecular basis for the multi-dimensional, dynamic programs required for specific functions of AMs *in vivo* (19, 20). Specifically, while *Arg1* and *Nos2* have been used for decades to assign phenotypes to macrophages, we show they are co-expressed within the same clusters of cells, and even in the same individual cell, as has been suggested previously (30-33). In our study, these markers alone, or as part of a panel of markers, are neither sufficiently sensitive nor specific to functionally characterize the cells that express them.

It is important to note that our study did not include all myeloid populations present in the lung lavage. For instance, while widely accepted in the literature as sensitive and specific markers of AMs, the use of F480 and CD64 for positive selection may have excluded AMs with lower expression of these molecules (58). However, our previous work with reporters specific for RAM or bone marrow-derived myeloid cells (12, 47) suggest that F480<sup>low</sup> and CD64<sup>low</sup> cells represent an insignificant portion of AMs. Likewise, the goal of this study was to focus on *bona fide* AMs. As such, dendritic cells and monocytes were also excluded during FACS and were not analyzed. Our study included data from a single animal at each time point and was not powered to detect macrophage heterogeneity between animals or varying levels of inflammation.

By using the same murine model of LPS-induced lung inflammation as our prior work (14), the current study allowed for direct comparison and examination of reproducibility of data between single cell and

'bulk' RNA sequencing platforms. Using WGCNA, we have illustrated a high level of technical reproducibility between the two techniques and have highlighted the added power of single cell analysis. While 'bulk' sequencing identified RAMs and RecAMs as transcriptionally distinct cell types with different inflammatory and proliferative programs, the addition of single cell RNA-seq allowed for identification of additional subgroups of cells and assignment of putative functions to these groups. Further, sequencing of single cells allows for direct correlation of cell markers with cell function, enabling re-examination of the adequacy of traditional markers of cellular programming. In this study, application of single cell technologies to a murine model of lung inflammation has yielded the discovery of pro-inflammatory and proteolytic subsets of RecAM present only during peak inflammation and identification of a subset of proliferative RAMs during both health and disease. Our results confirm that AMs exist in heterogeneous populations during lung inflammation and suggest that targeting the 'division of labor' between AM subgroups may provide therapeutic targets for inflammatory lung diseases.

## METHODS

### Animal Model

*Animals.* All mice used for RNA sequencing were C57BL/6J males, age 10-12 weeks of age, from the Jackson Laboratory (Bar Harbor, ME). This study was approved by and performed in accordance with the ethical standards of IACUC at National Jewish Health, Denver, Colorado.

*Acute lung inflammation model.* 20 micrograms of LPS (*Escherichia coli* O55:B5 from List Biological (59)Laboratories, Campbell, CA) in 50 microliters of PBS were instilled directly into the tracheas of mice sedated with isofluorane (Baxter, Deerfield, IL) using a modified feeding needle. Mice were sacrificed using peritoneal injection of pentobarbital sodium solution.

*Macrophage isolation.* Bronchoalveolar lavage (BAL) was performed with 10 serial lavages of 1ml PBS containing 5 mM EDTA. One animal was used for each time point. BAL cells were washed twice in PBS. Fc $\gamma$ R was blocked using anti-CD16/CD32 (EBioscience, clone 93) for 20 minutes. For day 3 sample, neutrophils were depleted over a column by binding Ly6G prior to surface staining. Cells were incubated with primary antibody on ice for 30 minutes, washed twice, and then taken to FACS. Dead cells were excluded on the basis of scatter and DAPI exclusion. Monocytes were excluded based on size and light scattering characteristics (14). Antibodies and cell stains used are as follows: Anti-Ly6G (Biolegend, clone 1A8), anti-CD3 (ThermoFischer Scientific, clone 17A2), anti-CD19 (ThermoFischer Scientific, clone eBio1D3), anti-NK1.1 (ThermoFischer Scientific, clone PK136), DAPI (Biolegend, #422801), anti-CD64 (BD Biosciences, clone X54-5/7.1), anti-CD11c (ThermoFischer Scientific, clone N418), anti-CD11b (ThermoFischer Scientific, clone M1/70), and anti-F4/80 (ThermoFischer Scientific, clone BM8). FACS was performed using a BD FACSAria Cell

Sorting System with BD FACSDiva Software (BD Biosciences, Franklin Lake, NJ) on unfixed specimens.

*Single cell RNA isolation and sequencing.* Sorted cells were dispensed into a 5,184-nanolwell version 1 microchip using the ICELL8 platform (WaferGen Biosystems, Fremont, CA) from which single cells were selected for library preparation and sequencing. Library preparation was carried out using ICell8 single-cell poly-(A)+ transcriptome amplification hands-on workflow (with Triton X-100 for cell lysis), protocol D07-000040-003 rev4. Libraries were generated for a total of 1,134 selected cells, which were then sequenced as 100 base pair single-end reads on an Illumina HiSeq 2500 (Illumina, Inc., San Diego, CA) in three batches (two flow cell lanes per run).

## Statistics

*Pre-processing of single cell RNA-seq data.* Raw cDNA reads in FASTQ files were trimmed using Cutadapt (60), where 5' and 3' quality trimming was carried out ( $q < 20$ ) and both poly A tails and reads shorter than 25 base pairs were removed. Trimmed reads were aligned to the GRCm38/mm10 genome with GSNAp, setting “max-mismatches=0.05” and accounting for both known Ensembl splice sites and SNPs. We quantified gene expression using HTSeq with “stranded=yes”, “mode=intersection-nonempty”, and “t=gene.” We summed the number of unique molecular identifiers (UMIs) for each gene across runs to obtain a UMI count matrix that we used for all downstream analysis.

We quantified the number of genes with expression greater than zero in each cell and then removed 222 cells expressing fewer than 2,000 genes. Although samples comprising multiple cells were removed during the cell selection stage, to safeguard against undetected doublets, we removed seven outlier cells expressing  $> 11,800$  genes or  $> 90,000$  UMIs. We removed three additional cells

for which < 40% of original reads mapped to a unique gene-annotated region. All remaining cells exhibited unique mapping rates > 50% and expressed genes of which < 30% were aligned to the mitochondrial genome. Prior to downstream analysis, we also removed select mitochondrial and ribosomal genes (those beginning with Mtat, Mt-, Mrpl, Mrps, Rpl, Rps) and very lowly expressed genes (those genes not expressed in at least 0.5% of cells).

The final QCed dataset contained 902 cells with a mean of 20,199 UMI's and 4,832 genes per cell. To account for differences in coverage across cells, we normalized UMI counts by dividing each by the sum of its cell's UMIs, multiplying by 10,000, and then taking the log. For input into dimensionality reduction and clustering analyses and for plotting relative expression in heat maps and bubble plots, we also fitted normalized expression of each gene to the sum of the UMI counts per cell, and then mean centered (subtracted from each gene count the average expression of that gene) and scaled (divided each centered gene count by that gene's standard deviation) the residuals. We used the Seurat R package (61) to carry out data normalization and scaling as well as downstream dimensionality reduction, clustering, tSNE plot overlaying, and differential expression.

*Dimensionality reduction and clustering.* We focused on the most informative genes for use in dimensionality reduction. To obtain these, we modeled the relationship between mean expression and the log of the ratio of expression variance to its mean (dispersion) and then selected the 5,784 genes for which fitted dispersion was > 0.2 and mean expression was > 0.15. We subjected these variable genes to principle components (PC) analysis and then, based on visualizing a relative leveling off of the standard deviation across PCs, we selected the top six PCs for subsequent clustering and visualization. To reduce dimensionality further, we carried out the Barnes-Hut implementation of t-distributed neighborhood embedding (tSNE) based on the top six PCs and plotted this embedding in two-dimensional space (perplexity = 50, number of iterations = 1,000).

We clustered single cells in an unsupervised manner based on their expression by first constructing a shared nearest neighbor graph based on k-nearest neighbors calculated from the first six PCs of the scaled data ( $k = 30$ ). The number of clusters was then determined using a modularity function optimizer based on the Louvain algorithm (resolution = 0.465).

To remove the cell cycle signature from the dataset, we used Seurat to calculate a cell cycle score for each cell based on mean expression across a list of S and G2/M markers compiled by Tirosh et al. (62). We then performed dimensionality reduction and clustering using scaled residual expression obtained from modeling the relationship between gene expression and cell cycle score.

*Plotting expression across cells.* For overlaying expression onto the tSNE plot for single genes or for the average across a single panel of genes, we plotted normalized expression along a continuous color scale, with the extreme color values being set to the 5<sup>th</sup> and 95<sup>th</sup> quantile expression values. For module expression overlays in Figure 7, we plotted eigengene expression (eigenvectors from the first principle component of a given network's gene expression) rather than mean expression. For plotting the expression of two genes on a single plot, "high" versus "low" expression for one or both genes was relative to the 85<sup>th</sup> quantile of normalized expression. The bubble plot and heat map show scaled normalized expression along a continuous color scale. We produced the heat map in Figure 6 using Heatmap3 (63).

*Differential expression analysis* Differential expression for each gene between various groups specified in the text was tested using a non-parametric Wilcoxon rank sum test. We limited each comparison to genes exhibiting both an estimated log fold change  $> +0.25$  and detectable expression in  $> 10\%$  of cells in one of the two clusters being compared. We corrected for multiple hypothesis testing by calculating FDR-adjusted p-values. Genes were considered to be differentially expressed when FDR  $< 0.05$ . To create the heat map in Figure 6 , we carried out differential expression analysis

between each of the five focal clusters and other individual clusters of comparative interest, based on our previous characterization of clusters using biomarkers of cell origin and state. Specifically, we compared cluster 1 with cluster 3, cluster 4, and cluster 5; cluster 2 with cluster 1; cluster 3 with cluster 1, cluster 4, and cluster 5; cluster 4 with cluster 1 and cluster 3; and cluster 5 with cluster 1, cluster 3, and cluster 4.

*Functional enrichment analysis.* We tested for gene overrepresentation of all target lists (e.g., from DEGs and WGCNA networks) within a panel of annotated gene databases (Gene Ontology [GO] Biological Process [BP], GO Molecular Function [MF], GO Cellular Component [CC], Kyoto Encyclopedia of Genes and Genomes [KEGG], Reactome) using hypergeometric tests implemented with Enrichr (64), as automated using the python script, EnrichrAPI (<https://github.com/russell-stewart/enrichrAPI>). We report only terms and pathways that were enriched with FDR < 0.05. Box plots comparing expression of enriched pathways across clusters show mean expression for genes in the target list that are also annotated for the pathway being plotted.

*Weighted gene co-expression network analysis (WGCNA).* We carried out WGCNA using samples from a previous bulk RNA-seq experiment on flow sorted resident and recruited AMs that was similar in design to the current study (14). We first downloaded the raw count table from GEO and then filtered samples such that only RAMs from days 0, 3, and 6 and RecAMs from days 3 and 6 were retained (3 replicates per treatment = 15 samples total). Genes without counts > 1 in at least 1 sample were removed, leaving 19,857 genes to analyze. We normalized the counts using DESeq2's implementation of variance stabilizing transformation (VST), which accounts for heterogeneity in sequencing coverage across cells while also minimizing mean-dependent variance (65).

We then analyzed the VST-normalized counts using the WGCNA R package (66) in order to infer distinct networks (or modules) of co-expressed genes that may share common biological functions or

be regulated by the same upstream mechanisms (67). Briefly, the WGCNA method measures pairwise Pearson correlations among all expressed genes in the dataset, constructs a gene network based on these correlations, calculates topological overlap among genes in the network, and then, by carrying out hierarchical clustering of a distance metric based on the topological overlap matrix, identifies discrete clusters of genes. We set the soft-thresholding power to three, which best maximized the number of connections in the gene network, while still meeting the scale-free network assumptions of the method, and we retained the direction of correlations among indirect connections in the network (i.e., analysis was ‘signed’). In addition, minimum network size was set to 30, maxCoreScatter=0.84, minGap=0.15, and cutHeight=0.78 (the 90% quantile of the tree height). We then merged any networks whose eigengenes were highly correlated ( $> 0.85$ ).

## **Statistics overview**

When determining whether gene expression was significantly different in one group compared to another, we used the Wilcoxon rank sum test. Genes were considered differentially expressed when p-values (controlled for multiple comparisons using the Benjamini–Hochberg false discovery rate (FDR) procedure) were  $< 0.05$ . Similarly, for determining significance in pathway analysis, pathways were only deemed significant if the hypergeometric test FDR was  $< 0.05$  and there were at least three overlapping genes. For comparisons of mean expression between groups, we used 1-sided t-tests, with p-values  $< 0.05$  considered significant.

## **Data and software availability**

All datasets have been deposited in the National Center for Biotechnology Information/Gene Expression Omnibus under accession number GSE120000.

## **Study approval**

Animal studies were approved by the Institutional Animal Care and Use Committee at National Jewish Health.

## AUTHOR CONTRIBUTIONS

K.J.M designed research, performed experiments, analyzed data, and wrote the manuscript. N.D.J. performed bioinformatics, analyzed data, and wrote the manuscript. P.M.H contributed to the manuscript. M.S. and W.J.J. designed research, analyzed data, and wrote the manuscript.

## **ACKNOWLEDGEMENTS**

This work was supported by the National Institutes of Health grants R35 HL140039, R01 HL130938, R01 HL114381, and K08 HL141648. The authors declare no conflicts of interest.

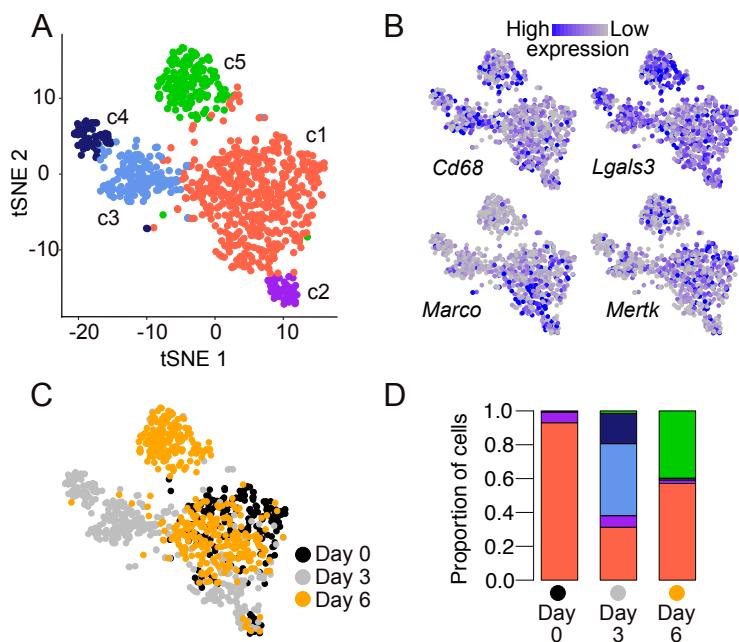
## REFERENCES

1. Herold S, Tabar TS, Janssen H, Hoegner K, Cabanski M, Lewe-Schlosser P, et al. Exudate macrophages attenuate lung injury by the release of IL-1 receptor antagonist in gram-negative pneumonia. *American journal of respiratory and critical care medicine*. 2011;183(10):1380-90.
2. Westphalen K, Gussarova GA, Islam MN, Subramanian M, Cohen TS, Prince AS, et al. Sessile alveolar macrophages communicate with alveolar epithelium to modulate immunity. *Nature*. 2014;506(7489):503-6.
3. Chen Q, Luo AA, Qiu H, Han B, Ko BH, Slutsky AS, et al. Monocyte interaction accelerates HCl-induced lung epithelial remodeling. *BMC Pulm Med*. 2014;14:135.
4. Cakarova L, Marsh LM, Wilhelm J, Mayer K, Grimminger F, Seeger W, et al. Macrophage tumor necrosis factor-alpha induces epithelial expression of granulocyte-macrophage colony-stimulating factor: impact on alveolar epithelial repair. *American journal of respiratory and critical care medicine*. 2009;180(6):521-32.
5. Golpon HA, Fadok VA, Taraseviciene-Stewart L, Scerbavicius R, Sauer C, Welte T, et al. Life after corpse engulfment: phagocytosis of apoptotic cells leads to VEGF secretion and cell growth. *FASEB J*. 2004;18(14):1716-8.
6. Liu J, Copland DA, Horie S, Wu WK, Chen M, Xu Y, et al. Myeloid cells expressing VEGF and arginase-1 following uptake of damaged retinal pigment epithelium suggests potential mechanism that drives the onset of choroidal angiogenesis in mice. *PLoS One*. 2013;8(8):e72935.
7. Zigmond E, Samia-Grinberg S, Pasmanik-Chor M, Brazowski E, Shibolet O, Halpern Z, et al. Infiltrating monocyte-derived macrophages and resident kupffer cells display different ontogeny and functions in acute liver injury. *J Immunol*. 2014;193(1):344-53.
8. Davies LC, Rosas M, Smith PJ, Fraser DJ, Jones SA, and Taylor PR. A quantifiable proliferative burst of tissue macrophages restores homeostatic macrophage populations after acute inflammation. *Eur J Immunol*. 2011;41(8):2155-64.
9. Gundra UM, Girgis NM, Ruckerl D, Jenkins S, Ward LN, Kurtz ZD, et al. Alternatively activated macrophages derived from monocytes and tissue macrophages are phenotypically and functionally distinct. *Blood*. 2014;123(20):e110-22.
10. Hashimoto D, Chow A, Noizat C, Teo P, Beasley MB, Leboeuf M, et al. Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity*. 2013;38(4):792-804.
11. Yona S, Kim KW, Wolf Y, Mildner A, Varol D, Breker M, et al. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity*. 2013;38(1):79-91.
12. Janssen WJ, Barthel L, Muldrow A, Oberley-Deegan RE, Kearns MT, Jakubzick C, et al. Fas determines differential fates of resident and recruited macrophages during resolution of acute lung injury. *American journal of respiratory and critical care medicine*. 2011;184(5):547-60.
13. Kearns MT, Barthel L, Bednarek JM, Yunt ZX, Henson PM, and Janssen WJ. Fas ligand-expressing lymphocytes enhance alveolar macrophage apoptosis in the resolution of acute pulmonary inflammation. *American journal of physiology Lung cellular and molecular physiology*. 2014;307(1):L62-70.
14. Mould KJ, Barthel L, Mohning MP, Thomas SM, McCubbrey AL, Danhorn T, et al. Cell Origin Dictates Programming of Resident Versus Recruited Macrophages During Acute Lung Injury. *Am J Respir Cell Mol Biol*. 2017.
15. Mills CD, Kincaid K, Alt JM, Heilman MJ, and Hill AM. M-1/M-2 macrophages and the Th1/Th2 paradigm. *Journal of immunology*. 2000;164(12):6166-73.

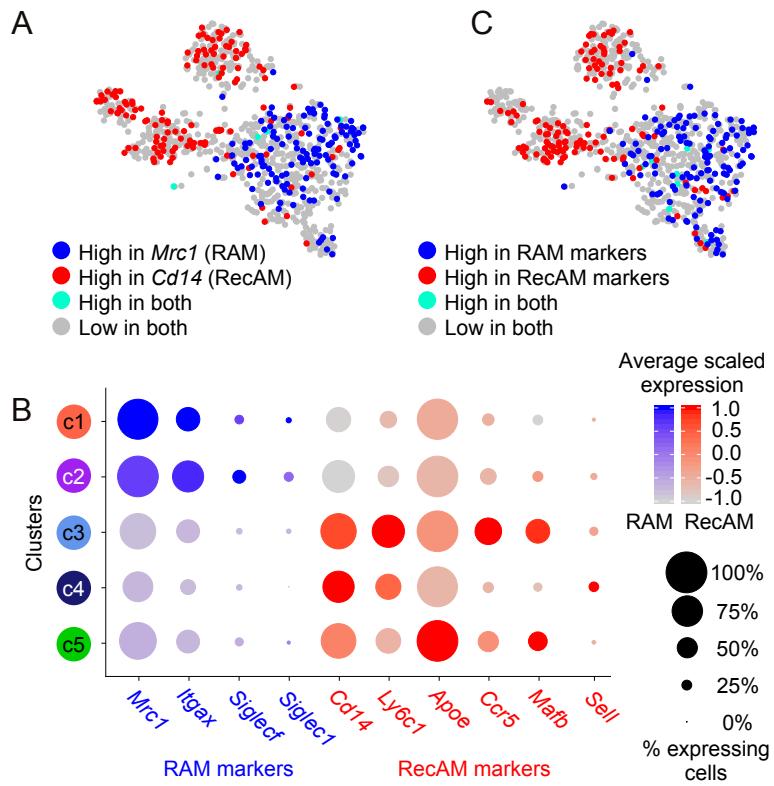
16. Doyle AG, Herbein G, Montaner LJ, Minty AJ, Caput D, Ferrara P, et al. Interleukin-13 alters the activation state of murine macrophages in vitro: comparison with interleukin-4 and interferon-gamma. *Eur J Immunol*. 1994;24(6):1441-5.
17. Troidl C, Mollmann H, Nef H, Masseli F, Voss S, Szardien S, et al. Classically and alternatively activated macrophages contribute to tissue remodelling after myocardial infarction. *J Cell Mol Med*. 2009;13(9B):3485-96.
18. Mosser DM, and Edwards JP. Exploring the full spectrum of macrophage activation. *Nature reviews Immunology*. 2008;8(12):958-69.
19. Martinez FO, and Gordon S. The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000Prime Rep*. 2014;6:13.
20. Xue J, Schmidt SV, Sander J, Draffehn A, Krebs W, Quester I, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity*. 2014;40(2):274-88.
21. O'Brien J, Lyons T, Monks J, Lucia MS, Wilson RS, Hines L, et al. Alternatively activated macrophages and collagen remodeling characterize the postpartum involuting mammary gland across species. *Am J Pathol*. 2010;176(3):1241-55.
22. Misharin AV, Morales-Nebreda L, Mutlu GM, Budinger GR, and Perlman H. Flow cytometric analysis of macrophages and dendritic cell subsets in the mouse lung. *Am J Respir Cell Mol Biol*. 2013;49(4):503-10.
23. Gibbings SL, Goyal R, Desch AN, Leach SM, Prabagar M, Atif SM, et al. Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. *Blood*. 2015;126(11):1357-66.
24. Liu YC, Zou XB, Chai YF, and Yao YM. Macrophage polarization in inflammatory diseases. *Int J Biol Sci*. 2014;10(5):520-9.
25. Rodriguez PC, Ochoa AC, and Al-Khami AA. Arginine Metabolism in Myeloid Cells Shapes Innate and Adaptive Immunity. *Front Immunol*. 2017;8:93.
26. Rath M, Muller I, Kropf P, Closs EI, and Munder M. Metabolism via Arginase or Nitric Oxide Synthase: Two Competing Arginine Pathways in Macrophages. *Front Immunol*. 2014;5:532.
27. Barron L, Smith AM, El Kasmi KC, Qualls JE, Huang X, Cheever A, et al. Role of arginase 1 from myeloid cells in th2-dominated lung inflammation. *PLoS One*. 2013;8(4):e61961.
28. Das A, Sinha M, Datta S, Abas M, Chaffee S, Sen CK, et al. Monocyte and macrophage plasticity in tissue repair and regeneration. *Am J Pathol*. 2015;185(10):2596-606.
29. Piccolo V, Curina A, Genua M, Ghisletti S, Simonatto M, Sabo A, et al. Opposing macrophage polarization programs show extensive epigenomic and transcriptional cross-talk. *Nat Immunol*. 2017;18(5):530-40.
30. Redente EF, Higgins DM, Dwyer-Nield LD, Orme IM, Gonzalez-Juarrero M, and Malkinson AM. Differential polarization of alveolar macrophages and bone marrow-derived monocytes following chemically and pathogen-induced chronic lung inflammation. *J Leukoc Biol*. 2010;88(1):159-68.
31. Schleicher U, Paduch K, Debus A, Obermeyer S, Konig T, Kling JC, et al. TNF-Mediated Restriction of Arginase 1 Expression in Myeloid Cells Triggers Type 2 NO Synthase Activity at the Site of Infection. *Cell Rep*. 2016;15(5):1062-75.
32. Maloney J, Keselman A, Li E, and Singer SM. Macrophages expressing arginase 1 and nitric oxide synthase 2 accumulate in the small intestine during Giardia lamblia infection. *Microbes Infect*. 2015;17(6):462-7.
33. El Kasmi KC, Qualls JE, Pesce JT, Smith AM, Thompson RW, Henao-Tamayo M, et al. Toll-like receptor-induced arginase 1 in macrophages thwarts effective immunity against intracellular pathogens. *Nat Immunol*. 2008;9(12):1399-406.
34. Reddy RC, Keshamouni VG, Jaigirdar SH, Zeng X, Leff T, Thannickal VJ, et al. Deactivation of murine alveolar macrophages by peroxisome proliferator-activated receptor-gamma ligands. *American journal of physiology Lung cellular and molecular physiology*. 2004;286(3):L613-9.

35. Schneider C, Nobs SP, Kurrer M, Rehrauer H, Thiele C, and Kopf M. Induction of the nuclear receptor PPAR-gamma by the cytokine GM-CSF is critical for the differentiation of fetal monocytes into alveolar macrophages. *Nat Immunol.* 2014;15(11):1026-37.
36. Coggle JE, and Tarling JD. The proliferation kinetics of pulmonary alveolar macrophages. *J Leukoc Biol.* 1984;35(3):317-27.
37. Gong W, Kwak IY, Pota P, Koyano-Nakagawa N, and Garry DJ. DrImpute: imputing dropout events in single cell RNA sequencing data. *BMC Bioinformatics.* 2018;19(1):220.
38. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, et al. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell.* 2018;172(5):1091-107 e17.
39. Tarling JD, Lin HS, and Hsu S. Self-renewal of pulmonary alveolar macrophages: evidence from radiation chimera studies. *J Leukoc Biol.* 1987;42(5):443-6.
40. Jenkins SJ, Ruckerl D, Cook PC, Jones LH, Finkelman FD, van Rooijen N, et al. Local macrophage proliferation, rather than recruitment from the blood, is a signature of TH2 inflammation. *Science.* 2011;332(6035):1284-8.
41. Lavin Y, Winter D, Blecher-Gonen R, David E, Keren-Shaul H, Merad M, et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell.* 2014;159(6):1312-26.
42. Rosas M, Davies LC, Giles PJ, Liao CT, Kharfan B, Stone TC, et al. The transcription factor Gata6 links tissue macrophage phenotype and proliferative renewal. *Science.* 2014;344(6184):645-8.
43. Soucie EL, Weng Z, Geirsdottir L, Molawi K, Maurizio J, Fenouil R, et al. Lineage-specific enhancers activate self-renewal genes in macrophages and embryonic stem cells. *Science.* 2016;351(6274):aad5510.
44. Zaslona Z, Przybranowski S, Wilke C, van Rooijen N, Teitz-Tennenbaum S, Osterholzer JJ, et al. Resident alveolar macrophages suppress, whereas recruited monocytes promote, allergic lung inflammation in murine models of asthma. *J Immunol.* 2014;193(8):4245-53.
45. Downey GP, Worthen GS, Henson PM, and Hyde DM. Neutrophil sequestration and migration in localized pulmonary inflammation. Capillary localization and migration across the interalveolar septum. *Am Rev Respir Dis.* 1993;147(1):168-76.
46. Maus UA, Waelsch K, Kuziel WA, Delbeck T, Mack M, Blackwell TS, et al. Monocytes are potent facilitators of alveolar neutrophil emigration during lung inflammation: role of the CCL2-CCR2 axis. *J Immunol.* 2003;170(6):3273-8.
47. McCubbrey AL, Barthel L, Mould KJ, Mohning MP, Redente EF, and Janssen WJ. Selective and inducible targeting of CD11b+ mononuclear phagocytes in the murine lung with hCD68-rtTA transgenic systems. *American journal of physiology Lung cellular and molecular physiology.* 2016;311(1):L87-L100.
48. Janssen WJ, Muldrow A, Kearns MT, Barthel L, and Henson PM. Development and characterization of a lung-protective method of bone marrow transplantation in the mouse. *J Immunol Methods.* 2010;357(1-2):1-9.
49. Chen S, Kammerl IE, Vosyka O, Baumann T, Yu Y, Wu Y, et al. Immunoproteasome dysfunction augments alternative polarization of alveolar macrophages. *Cell Death Differ.* 2016;23(6):1026-37.
50. Silswal N, Reis J, Qureshi AA, Papasian C, and Qureshi N. Of Mice and Men: Proteasome's Role in LPS-Induced Inflammation and Tolerance. *Shock.* 2017;47(4):445-54.
51. Keller IE, Vosyka O, Takenaka S, Kloss A, Dahlmann B, Willems LI, et al. Regulation of immunoproteasome function in the lung. *Sci Rep.* 2015;5:10230.
52. Reis J, Guan XQ, Kissilev AF, Papasian CJ, Qureshi AA, Morrison DC, et al. LPS-induced formation of immunoproteasomes: TNF-alpha and nitric oxide production are regulated by altered composition of proteasome-active sites. *Cell Biochem Biophys.* 2011;60(1-2):77-88.
53. Reis J, Hassan F, Guan XQ, Shen J, Monaco JJ, Papasian CJ, et al. The immunoproteasomes regulate LPS-induced TRIF/TRAM signaling pathway in murine macrophages. *Cell Biochem Biophys.* 2011;60(1-2):119-26.

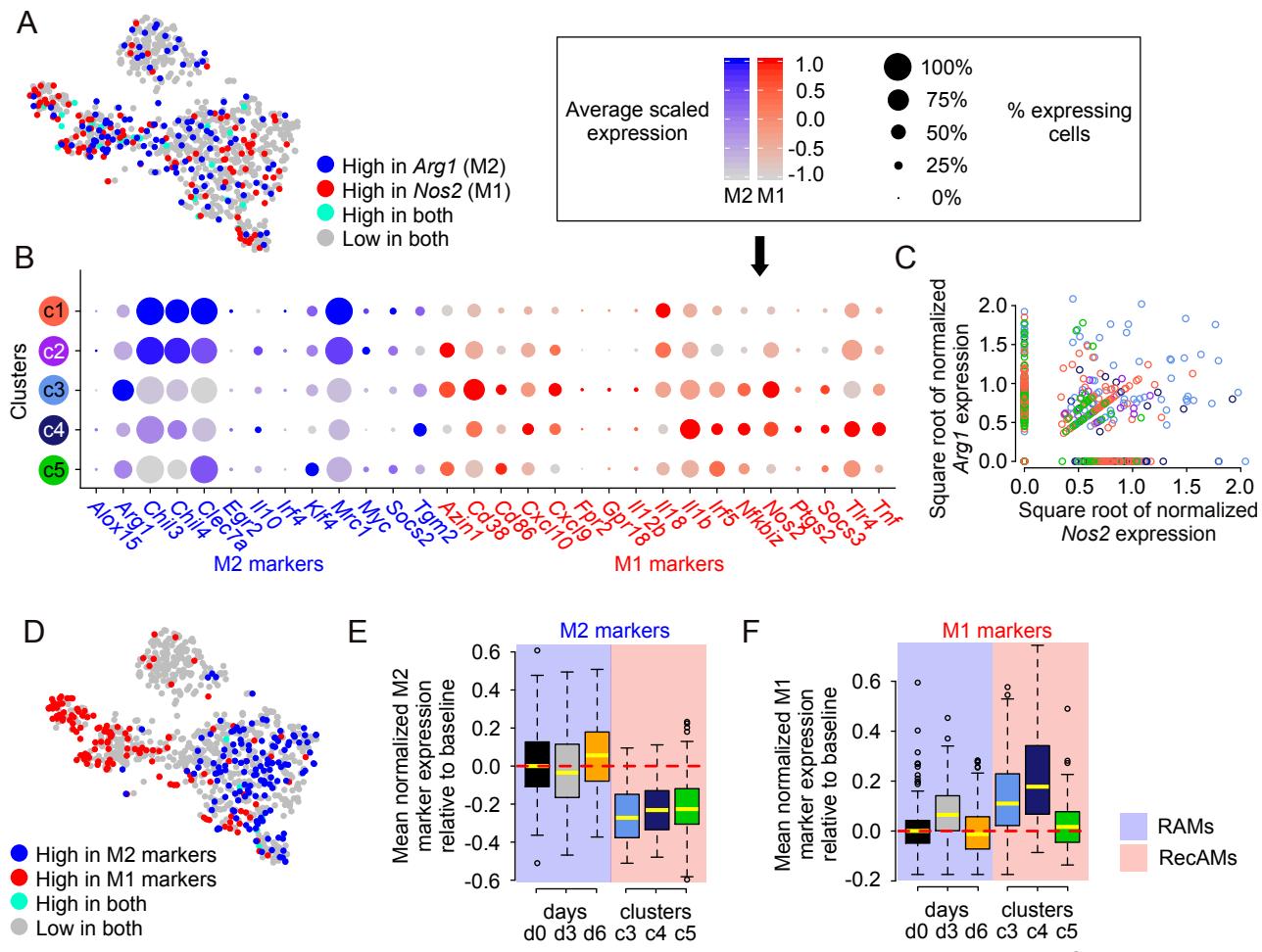
54. Heng TS, Painter MW, and Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-4.
55. Rijal D, Ariana A, Wight A, Kim K, Alturki NA, Aamir Z, et al. Differentiated Macrophages Acquire a Pro-Inflammatory and Cell Death-Resistant Phenotype Due to Increasing XIAP and P38-mediated inhibition of RipK1. *J Biol Chem*. 2018.
56. Kikuchi K, Iida M, Ikeda N, Moriyama S, Hamada M, Takahashi S, et al. Macrophages Switch Their Phenotype by Regulating Maf Expression during Different Phases of Inflammation. *J Immunol*. 2018.
57. Yanez A, Coetzee SG, Olsson A, Muench DE, Berman BP, Hazelett DJ, et al. Granulocyte-Monocyte Progenitors and Monocyte-Dendritic Cell Progenitors Independently Produce Functionally Distinct Monocytes. *Immunity*. 2017;47(5):890-902 e4.
58. Gautier EL, Shay T, Miller J, Greter M, Jakubzick C, Ivanov S, et al. Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat Immunol*. 2012;13(11):1118-28.
59. Agbayani G, Gurnani K, Zafer A, Sad S, and Krishnan L. Lack of functional selectin-ligand interactions enhances innate immune resistance to systemic Listeria monocytogenes infection. *J Leukoc Biol*. 2018;103(2):355-68.
60. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*. 2011;17(1):10.
61. Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*. 2018;36:411-20.
62. Tirosh I, Venteicher AS, Hebert C, Escalante LE, Patel AP, Yizhak K, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature*. 2016;539(7628):309-13.
63. Zhao S, Guo Y, Sheng Q, and Shyr Y. Advanced heat map and clustering analysis using heatmap3. *Biomed Res Int*. 2014;2014:986048.
64. Chen EY, Tan CM, Lou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;128(14).
65. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12).
66. Langfelder P, and Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
67. Zhang B, and Horvath S. *A general framework for weighted gene co-expression network analysis*. The Berkeley Electronic Press; 2005.



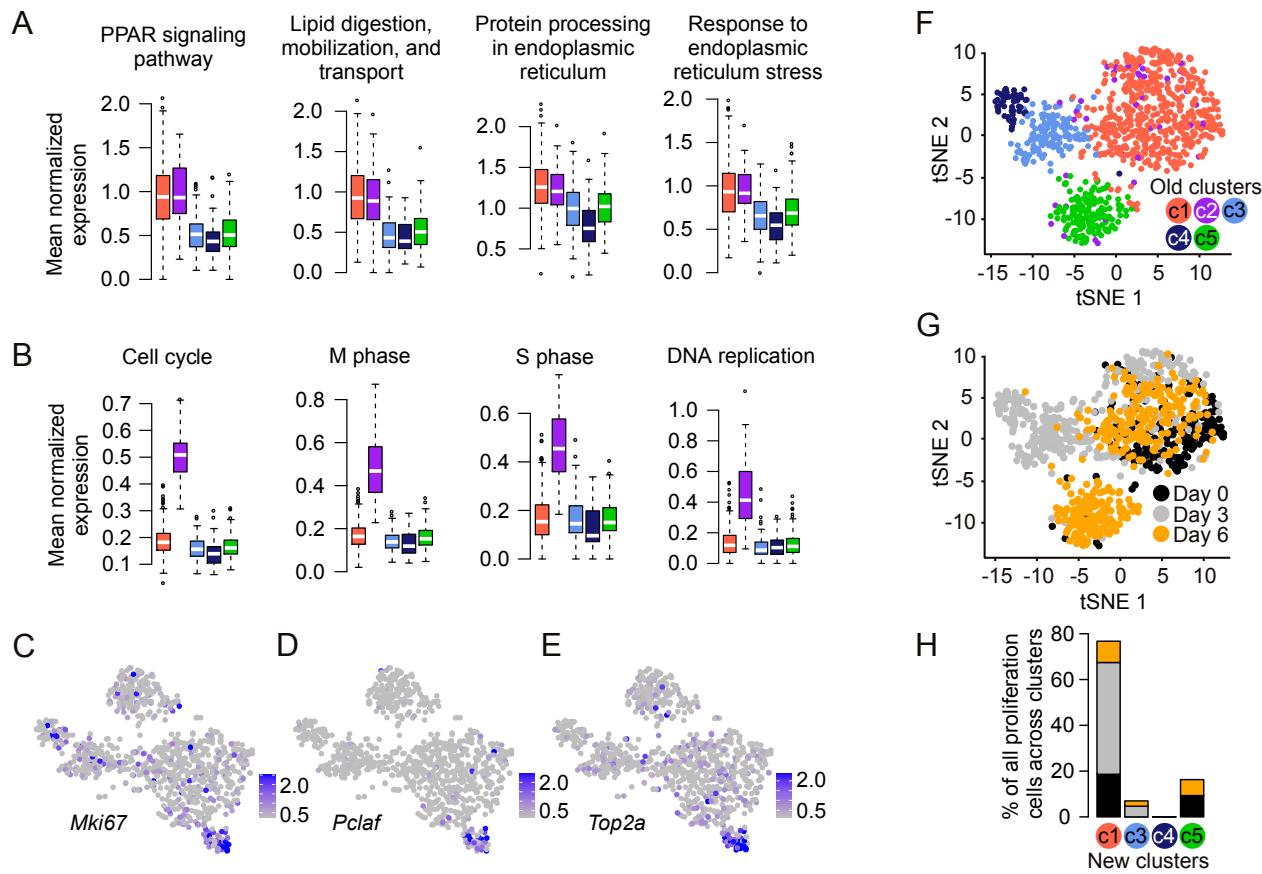
**Figure 1. Single cell transcriptional profiling identifies five discrete AM populations across homeostasis, acute inflammation, and resolving inflammation.** Mice were treated with intratracheal LPS and macrophages were isolated from lavage at days 0 (homeostasis), 3 (peak neutrophil inflammation), and 6 (resolution of lung inflammation). **A**) t-distributed stochastic neighbor embedding (tSNE) plot shows clustering of 902 cells based on gene expression. Point coordinates are based on tSNE dimensionality reduction of the top six principle components calculated from the 5,784 most informative genes. Cell color specifies assignment of cells to one of five clusters (c1-5) inferred using shared nearest neighbor clustering. **B**) Normalized expression of macrophage markers overlaid on tSNE plot. **C**) Time course information overlaid on tSNE plot. **D**) Relative proportion of cells in each cluster versus time.



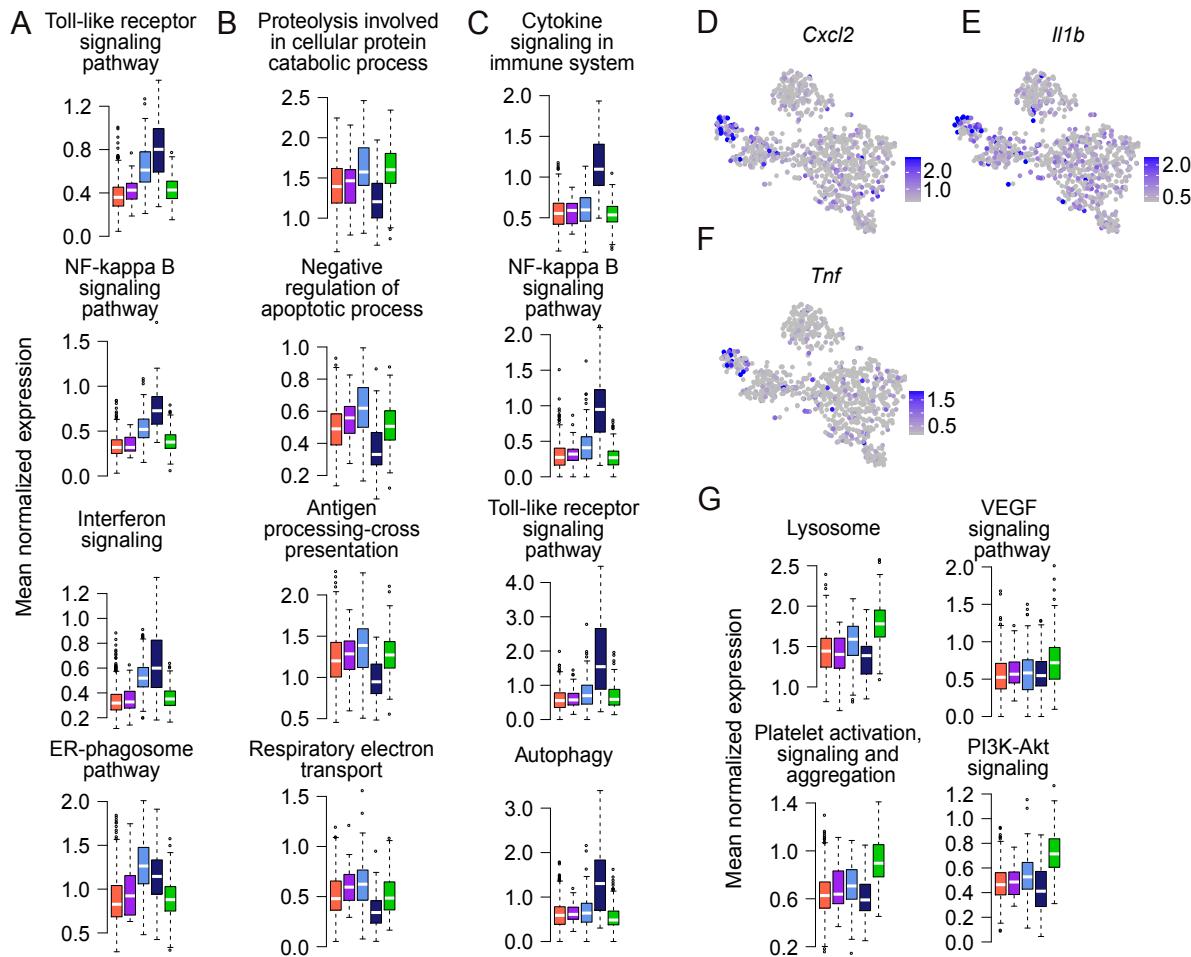
**Figure 2. AM populations revealed by single cell RNA-seq reflect cell origin.** **A)** Relative expression of *Mrc1* and *CD14* overlaid on tSNE plot. Cells that express both markers are turquoise. “High” versus “low” expression is defined relative to the 85<sup>th</sup> percentile. **B)** Bubble plot comparing expression of resident (blue) and recruited (red) biomarkers across the five macrophage clusters. Bubble size is proportional to percentage of cells in a cluster expressing a gene and color intensity is proportional to average scaled gene expression within a cluster. **C)** Summary expression of four resident biomarkers (*Mrc1*, *Itgax*, *Siglecf*, and *Siglec1*) and six recruited biomarkers (*Cd14*, *Ly6c1*, *Apoe*, *Ccr5*, *Mafb*, and *Sell*) overlaid on tSNE plot.



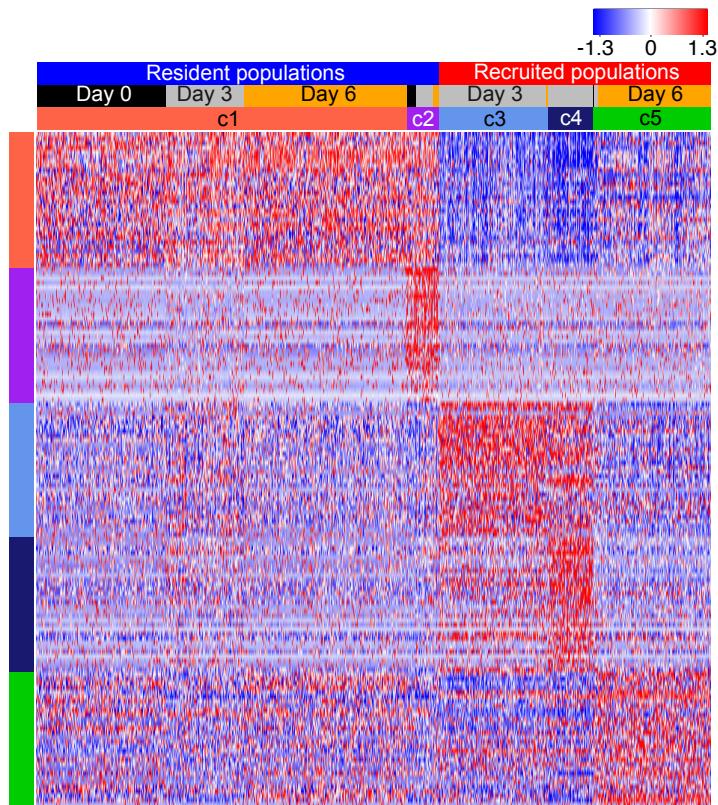
**Figure 3. Macrophage polarization states are not mutually exclusive and reflect cell origin.** **A)** Cells with high expression of *Arg1* and/or *Nos2* overlaid on tSNE plot (“high” versus “low” expression defined relative to the 85<sup>th</sup> percentile). **B)** Bubble plot shows relative expression of M2 (blue) and M1 (red) markers across clusters. Bubble size is proportional to percentage of cells expressing a gene and color intensity is proportional to average scaled gene expression within a cluster. **C)** Co-expression of *Arg1* and *Nos2* shown as square root-transformed expression of one gene against the other. Sample color denotes cluster from Figure 1A. **D)** Summary expression of genes from B overlaid on tSNE plot. **E,F)** Mean normalized expression of M2 (**E**) and M1 (**F**) markers relative to homeostatic RAMs. Data are shown across days for RAMs and clusters for RecAMs. Relative-to-baseline expression obtained by subtracting median value observed in clusters 1 and 2 at day 0.



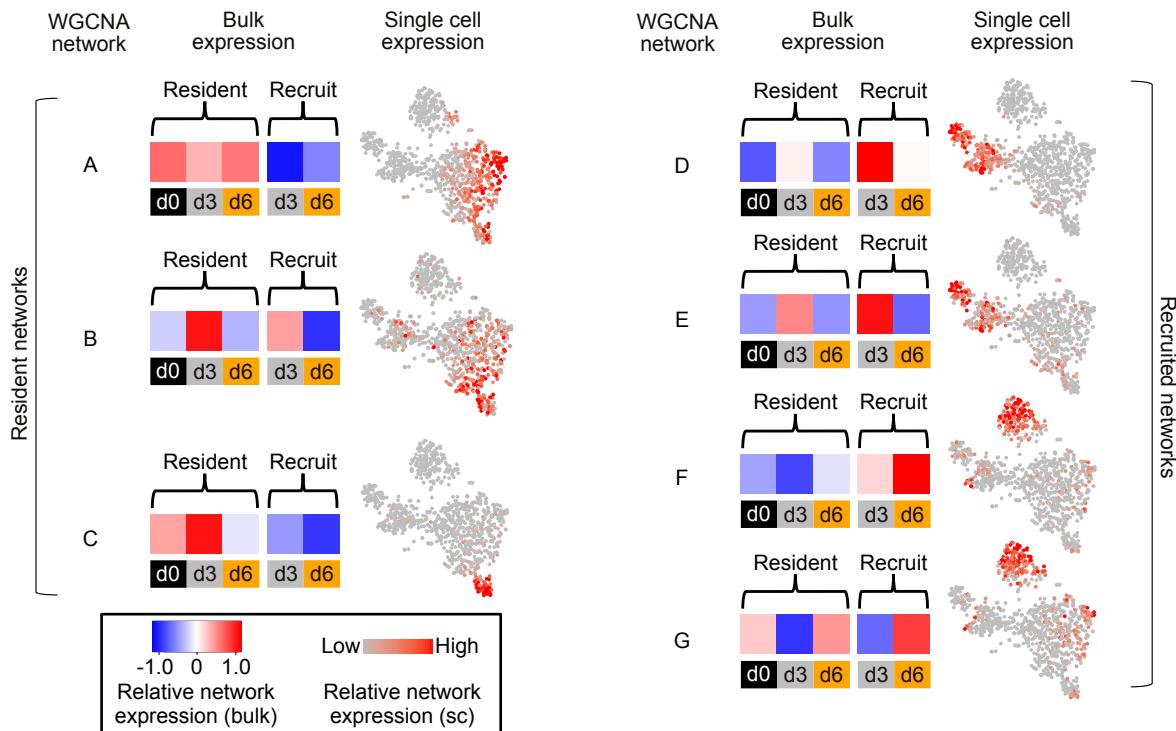
**Figure 4. Resident AMs contain a proliferative subpopulation present during health and inflammation.** **A,B)** Mean normalized expression of genes annotated for enriched pathways indicated that were also upregulated in cluster 1 (**A**) or cluster 2 (**B**) when compared to RecAM clusters. **C-E)** Normalized expression of three classic markers of proliferation overlaid onto tSNE plot. **F-H)** Characterization of proliferating RAMs in cluster 2 after statistically removing expression heterogeneity related to cell cycle from dataset. **F)** t-SNE plot based scaled residual expression obtained from modeling relationship between gene expression and estimated cell cycle score. Cells colored based on original clusters in Figure 1A. **G)** Time course information overlaid on new t-SNE plot. **H)** Distribution of cells from original proliferation cluster across new clusters inferred using cell cycle-corrected dataset. New clusters were broadly similar to original clusters and were thus named and colored to match cluster analogs in Figure 1A.



**Figure 5. Novel macrophage subsets are identified during inflammation.** **A-C**) Mean normalized expression of genes annotated for enriched pathways upregulated in clusters 3 and 4 together when compared to all other clusters (**A**), or cluster 3 (**B**) or cluster 4 (**C**) when compared to each other. **C-E**) Normalized expression of three inflammatory cytokines overlaid on tSNE plot. **F**) Bar plots comparing mean normalized expression of genes annotated for enriched pathways upregulated in cluster 5 when compared to all other clusters.



**Figure 6. Expression of identified gene sets distinguishes clusters.** Scaled expression of top 30 differentially expressed genes from each cluster. Individual cells are represented on the horizontal axis and grouped by cluster and day of inflammation.



**Figure 7. Gene networks from ‘bulk’ RNA-seq are co-expressed within single cell RNA-seq-based clusters.** Seven co-expressed gene networks derived from bulk RNA-seq data exhibit peak expression within single cell-derived resident (networks A-C) or recruited (networks D-G) macrophage clusters (See also Supplemental Figure 3). For each network, mean scaled eigengene expression is shown across ‘bulk’ RNA-seq resident and recruited AMs from each day (heat maps). For comparison with the single cell dataset, relative eigengene expression of each network is overlaid onto the single cell tSNE plot. Select results from pathway analysis of each network are shown in Table 1.

**Table 1.** Functional descriptions of WGCNA networks

Network	Functional summaries	Enriched terms	Gene set	Overlap	FDR
A	Metabolism of lipids & other macromolecules	Metabolism of lipids and lipoproteins	Reactome	125/659	9.05E-11
		Peroxisome	GO (CC)	36/94	3.76E-11
B	RNA modification	Fatty acid metabolism	KEGG	21/48	9.88E-08
		Peroxisomal lipid metabolism	Reactome	12/27	3.06E-04
B	Metabolism	mRNA splicing - major pathway	Reactome	26/134	5.77E-11
		Spliceosome	KEGG	23/134	9.94E-09
B	Protein modification & transport	Mitochondrial inner membrane	GO (CC)	40/330	1.48E-10
		The citric acid (TCA) cycle and RET	Reactome	22/153	3.62E-07
B	ER stress/UPR	Respiratory electron transport,...	Reactome	15/109	9.68E-05
		Protein processing in endoplasmic reticulum	KEGG	23/169	5.59E-07
B	Proteasome	Asparagine N-linked glycosylation	Reactome	26/259	2.10E-05
		ER to Golgi Anterograde Transport	Reactome	15/131	4.76E-04
B	ER stress/UPR	Vesicle-mediated transport	Reactome	32/492	1.38E-03
		IRE1-mediated unfolded protein response	GO (BP)	9/57	4.04E-03
C	Cell proliferation	Unfolded Protein Response (UPR)	Reactome	10/86	3.70E-03
		XBP1(S) activates chaperone genes	Reactome	7/53	1.06E-02
D	Immune signaling	Antigen processing-cross presentation	Reactome	13/82	9.68E-05
		Proteasome	KEGG	8/44	1.69E-03
D	Glycolysis	Cell cycle	Reactome	164/566	6.72E-88
		DNA replication	GO (BP)	53/106	1.21E-40
D	Apoptosis	Centrosome	GO (CC)	61/401	1.52E-15
		Cytokine signaling in immune system	Reactome	107/620	1.60E-20
D	Glycolysis	TNF signaling pathway	KEGG	36/110	2.92E-15
		Toll-like receptor signaling pathway	KEGG	30/106	2.31E-11
D	Apoptosis	NF-kappa B signaling pathway	KEGG	27/93	1.46E-10
		HIF-1 signaling pathway	KEGG	24/103	1.34E-07
D	Apoptosis	Glycolysis	Reactome	12/32	1.43E-05
		Glucose metabolism	Reactome	16/79	6.17E-04
E	Proteasome	Apoptosis	KEGG	26/140	2.95E-06
		Extrinsic apoptotic signaling pathway	GO (BP)	9/29	2.12E-03
E	Interferon signaling	Programmed cell death	Reactome	20/166	2.01E-02
		Proteasome	KEGG	25/44	1.51E-25
E	Metabolism	Antigen processing-cross presentation	Reactome	27/82	6.99E-21
		Cytokine signaling in immune system	Reactome	57/620	2.16E-14
F	Cell growth & repair	Interferon signaling	Reactome	20/196	3.48E-06
		Oxidative phosphorylation	KEGG	22/133	1.29E-09
F	Cell growth & repair	Respiratory electron transport	Reactome	17/88	1.94E-09
		Lysosome	KEGG	16/123	4.62E-03
F	Cell growth & repair	Signaling by EGFR	Reactome	29/355	1.73E-02
		Signaling by PDGF/FGFR1	Reactome	28/364	1.99E-02
F	Cell growth & repair	Signaling by VEGF	Reactome	24/328	5.45E-02
		PI3K/AKT activation	Reactome	19/125	1.54E-02
F	Cell growth & repair	mTOR signaling pathway	KEGG	12/60	2.27E-02