


## ARTICLE



# A single-cell lung atlas of complement genes identifies the mesothelium and epithelium as prominent sources of extrahepatic complement proteins

Neha Chaudhary<sup>1,4</sup>, Archana Jayaraman<sup>1,4</sup>, Christoph Reinhardt<sup>2</sup>, Joshua D. Campbell<sup>3</sup> and Markus Bosmann<sup>1,2</sup> 

© The Author(s), under exclusive licence to Society for Mucosal Immunology 2022

To understand functional duality of the complement system in host defense and lung injury, a more comprehensive view of its localized production in the lung, and the impact of age on complement production are essential. Here, we explored the expression of complement genes through computational analysis of preexisting single cell RNA sequencing data from lung transcriptomes of healthy young (3 months) and old C57BL/6 mice (24 months), and humans. We characterized the distribution of 48 complement genes. Across 28 distinct immune and non-immune cell types in mice, mesothelial cells expressed the greatest number of complement genes (e.g., *C1ra*, *C2*, *C3*), and regulators (e.g., *Serpining1*, *Cfh*). *C5* was abundant in type II alveolar epithelial cells and *C1q* in interstitial lung macrophages. There were only moderate differences in gene expression between young and old mice. Among 57 human lung cell types, mesothelial cells showed abundant complement expression. A few differences in gene expression (e.g., *FCN1*, *CFI*, *C6*, *C7*) were also evident between mice and human lung cells. Our findings present a novel perspective on the expression patterns of complement genes in normal lungs. These findings highlight the potential functions of complement in tissue-specific homeostasis and immunity and may foster a mechanistic understanding of its role in lung health and disease.

*Mucosal Immunology* (2022) 15:927–939; <https://doi.org/10.1038/s41385-022-00534-7>

## INTRODUCTION

Lungs are large barrier organs with an extensive mucosal surface and are characterized by remarkable and complex cellular heterogeneity. Innovations in omics-based technologies have facilitated the identification of new cell types<sup>1,2</sup>. A recent study reported as many as 57 distinct cell phenotypes in the human lung<sup>3</sup>. These cell types include various non-immune cell populations such as epithelial, endothelial, and mesenchymal/stromal cells. The prominent leukocyte populations are mononuclear phagocytes, lymphocytes, and rare megakaryocytes<sup>3</sup>. Each of these cell populations exhibit functional plasticity and are involved in multiple regulatory pathways.

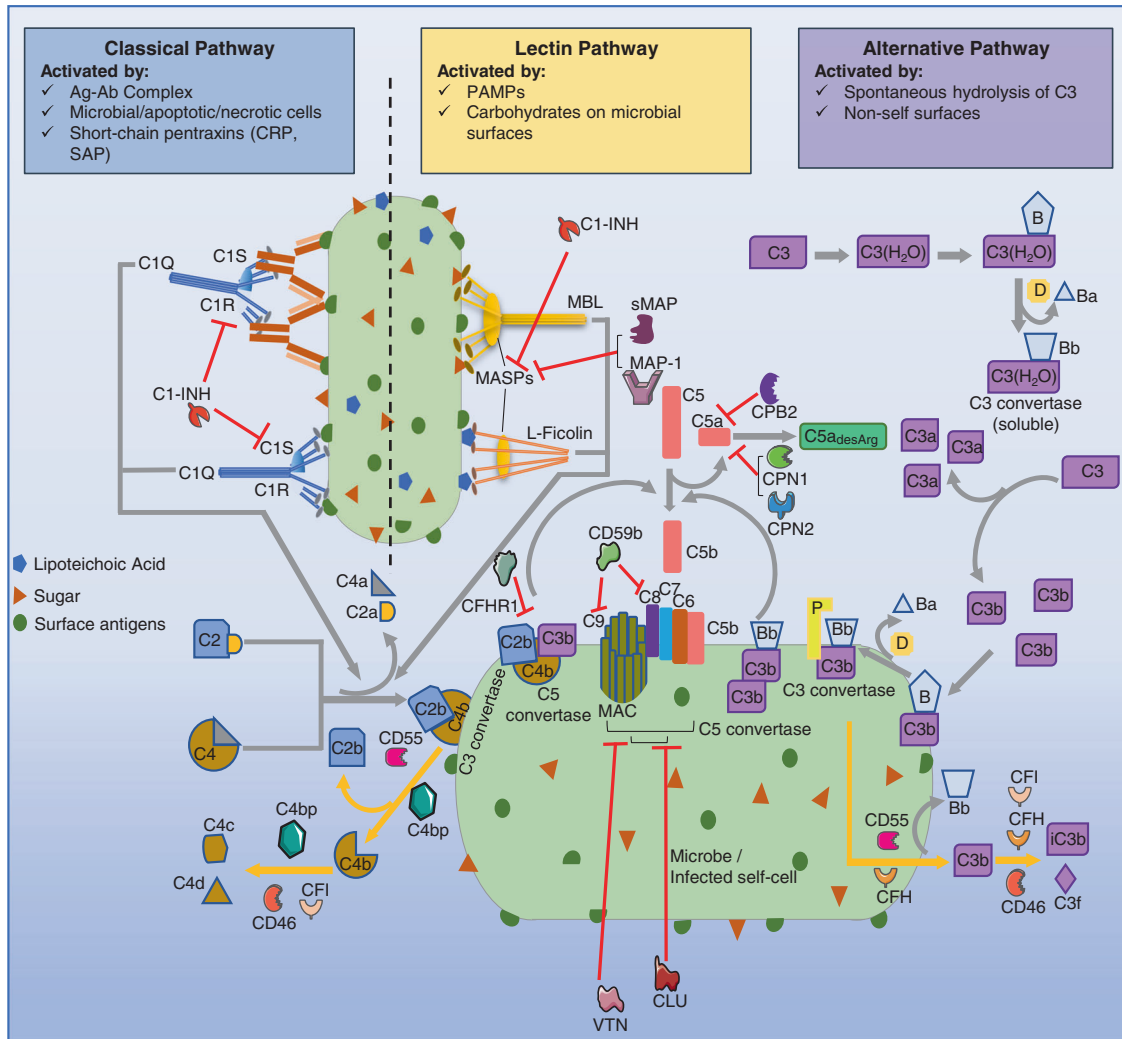
Airway epithelial cells such as goblet and ciliated cells are required for mucociliary clearance of inhaled particles. Alveolar epithelial type 1 (AT1) cells provide a short-distance barrier for diffusion gas exchange<sup>4</sup>. AT2 cells maintain homeostasis, secrete surfactant, mediate defense against pathogens (including MHCII-dependent antigen presentation<sup>5</sup>), and serve as AT1 progenitors<sup>4</sup>. Pulmonary capillary endothelial cells facilitate gas exchange and leukocyte migration<sup>6</sup>. Other subsets of lung endothelial cells from the pulmonary lymphatics, pulmonary blood vessels, and systemic blood vessels in the lung await functional annotations<sup>7</sup>. Mesenchymal cells, including smooth muscle cells, fibroblasts and mesothelial cells, are involved in homeostasis, repair, and wound healing<sup>8</sup>. Mesothelial cells cover the pleural spaces, secrete

glycosaminoglycans and surfactant, and initiate inflammatory responses to pathogens via pro-inflammatory mediators such as interleukins (e.g., IL-6, IL-8) and interferons<sup>9,10</sup>. Amongst myeloid cells, alveolar macrophages are potent airway scavengers and modulate the magnitude of immune responses and tissue homeostasis<sup>11</sup>. Interstitial macrophages, localized in the parenchyma, function in inflammation, wound healing, and tissue repair<sup>11</sup>. Monocytes patrol the blood and play a role in surveillance of endothelial integrity and are recruited to the inflamed lungs<sup>12</sup>. Dendritic cells activate naïve T-cells, initiating adaptive immunity<sup>11</sup>. Granulocytes such as neutrophils and eosinophils modulate immune response and are involved in pathogen killing and clearance<sup>11</sup>. Resident memory CD4<sup>+</sup> and CD8<sup>+</sup> T-cells function in adaptive immune responses through the secretion of cytokines, activation of macrophages and lung stromal cells, and/or elimination of infected cells<sup>13</sup>. Tissue resident B-cells and plasma cells mediate humoral immunity and pathogen clearance through antibody production<sup>14,15</sup>.

The proteolytic complement system is an essential part of the innate immune response to pathogens that infect the lung<sup>16</sup>. Activation of the complement system is mediated by one of three pathways: the classical pathway (CP) is potentially initiated via the C1 protein complex through IgM/IgG antibodies bound to antigens, especially when clustered as pentameric or hexameric structural motifs<sup>17</sup>. The lectin pathway (LP) is

<sup>1</sup>Pulmonary Center, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. <sup>2</sup>Center for Thrombosis and Hemostasis, University Medical Center Mainz, Mainz, Germany. <sup>3</sup>Division of Computational Biomedicine, Boston University School of Medicine, Boston, MA, USA. <sup>4</sup>These authors contributed equally: Neha Chaudhary, Archana Jayaraman. ✉email: mbosmann@bu.edu

Received: 28 February 2022 Revised: 5 May 2022 Accepted: 22 May 2022  
Published online: 7 June 2022

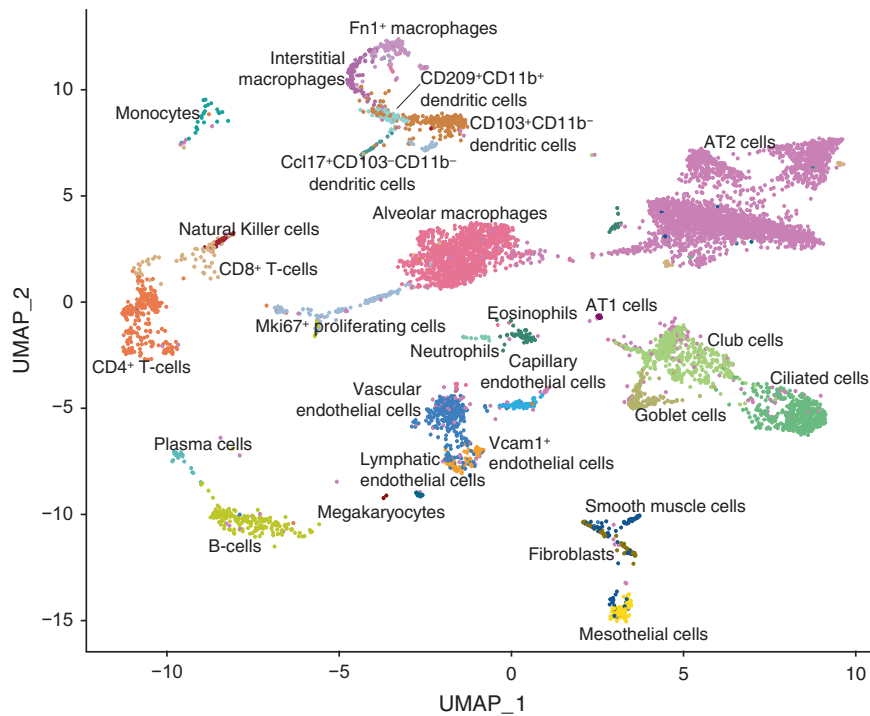


**Fig. 1 Overview of the complement pathways.** The complement system is activated when its proteins encounter pathogens or damaged/infected self-cells and recognize a pattern on the surface of these cells. These events trigger the activation of complement through one of three pathways: the classical pathway (CP), the lectin pathway (LP) and the alternative pathway (AP). CP can be triggered by antigen-antibody (Ag-Ab) complexes, microbes, apoptotic cells, or short-chain pentraxins (e.g., C-reactive protein [CRP], serum amyloid P [SAP]) binding to the C1Q-C1R-C1S complex. This binding cleaves complement C4 and C2 to their respective peptide fragments and leads to the formation of the CP C3 convertase, C4b2b, and the C5 convertase (C4b2bC3b). LP is activated by attachment of mannose-binding lectin to cell surface carbohydrates on the pathogen and pathogen associated molecular patterns (PAMPs). This results in the MASP1 and MASP2 serine proteases, associated with the lectin pattern recognition molecules (MBL, ficolin), facilitating C3 convertase (C4b2b) and C5 convertase (C4b2bC3b) formation by cleaving C4 and C2 to peptide fragments. AP is activated in response to cellular damage and mediated by spontaneous hydrolysis of C3 or stimulation of C3 by properdin (P/CFP). The hydrolyzed C3 then forms a complex with factor B (B/CFB) and factor D (D/CFD) to form the AP C3 convertase (C3bBb). The CP, LP and AP C3 convertases can cleave C3 into C3a (anaphylatoxin) and C3b (opsonin) peptide fragments. Two C3b molecules can further associate with the factor B peptide Bb to form the AP C5 convertase C3bBb3b. The C5 convertases cleave C5 into C5a (anaphylatoxin) and C5b peptide fragments. The C5b fragment forms a membrane attack complex (MAC) together with C6–C9 complement component proteins that insert into the pathogen cell membrane, forming a pore and leading to cell lysis. C3a and C5a mediate phagocytosis, chemotaxis of immune cells and immunomodulation by binding to their respective G-protein coupled transmembrane receptors (C3AR1, C5AR1, C5AR2). Complement regulators such as CD46, CD55, CD59b (all membrane-anchored), C1-INH, sMAP, MAP-1, factor H (CFH), factor I (CFI) and carboxypeptidases (CPN1, CPN2, CPB2) closely monitor and modulate the complement activation and propagation<sup>16,19,24</sup>.

triggered by attachment of mannose-binding lectin and other pattern recognition molecules (e.g., ficolins) to cell surface carbohydrates on the pathogen. The alternative pathway (AP) is activated in response to cellular damage on non-self surfaces (CD46<sup>−</sup>CD55<sup>−</sup>CD59<sup>−</sup>) either by spontaneous hydrolysis of effector C3 or activated as an amplification loop for the other pathways<sup>18,19</sup> (Fig. 1; Supplementary Table 1). All of these events result in the formation of C3 convertases that cleave C3 into C3a (anaphylatoxin) and C3b (opsonin) fragments, and C5 convertases that cleave C5 into C5a and C5b with subsequent formation of the membrane attack complex (C5b-C9). These

pathways involve regulated interactions amongst a network of nearly 50 complement proteins which can be classified as pattern recognition molecules, proteases, complement components, receptors, and regulators<sup>18–20</sup>.

Although the traditional view favors hepatic synthesis of inactive complement precursors, newer concepts appreciate localized production of complement genes in peripheral tissues and at mucosal surfaces such as the intestines (e.g., C3)<sup>21</sup> and lung (e.g., C2-C5, factor B)<sup>16,22</sup>. However, a comprehensive analysis of the expression of complement genes in the context of the evolving cellular landscape of the lung is not available.



**Fig. 2 Visualization of different cell types identified in the mouse lung transcriptome.** Uniform Manifold Approximation and Projection (UMAP) plot of lung single cell transcriptomes from young C57BL/6 mice ( $n = 7473$  cells from 8 mice, 3 months old).

In the current study, we have profiled the expression of 48 complement genes in different lung cells through bioinformatics analysis of single cell RNA sequence data (scRNA-Seq) from healthy young and old mice (3 versus 24 months of age) generated by Angelidis et al.<sup>23</sup>, and from healthy tissue from human lungs by Travaglini et al.<sup>3</sup>.

Aberrant activation of the complement has been linked to the pathogenesis of age-associated diseases such as age-related macular degeneration<sup>23</sup>. However, the impact of age on complement production in different lung cell types has not been elucidated. Hence, we have also examined age-associated differences in the complement transcriptome profiles in immune and non-immune cell types in mice.

Overall, we observed that the classical and alternative pathways genes were expressed more compared with lectin pathway in both mice and humans, suggesting potential localized enforcement of complement activation through these pathways in the lung. Aging did not have a large impact on the expression of complement genes in the mouse dataset. These findings may provide a useful perspective in understanding cell type specific dysregulation of the complement system in the pathogenesis of lung diseases.

## RESULTS

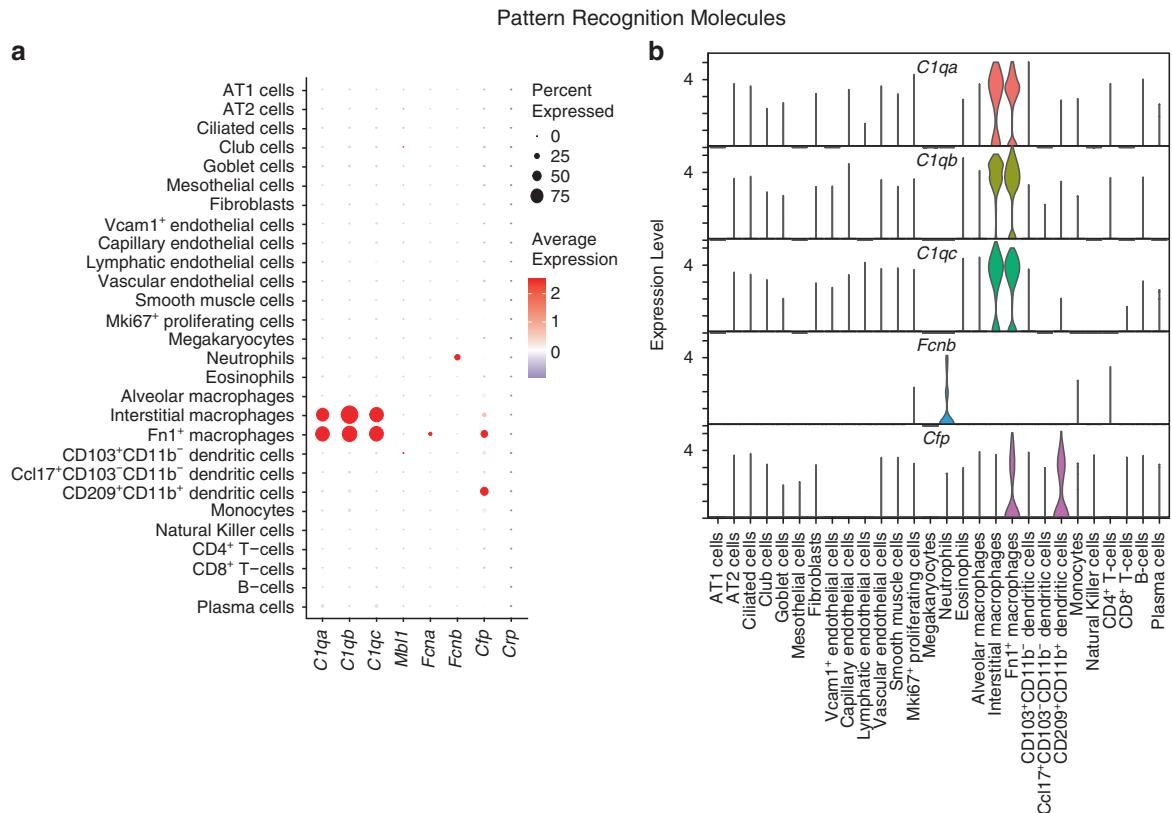
To elucidate the cell type-specific expression of complement genes in lung, while complying with IACUC guidelines of reducing the numbers of mice needed for research, we performed a search in NCBI's Gene Expression Omnibus database for a dataset of untreated, healthy C57BL/6 mice. The scRNA-Seq study dataset by Angelidis et al.<sup>23</sup> that included lung tissues from eight young mice (3 months) and seven old mice (24 months), was used in our analysis (Supplementary Fig. 1a). Quality control parameters such as unique molecular identifiers (UMI) per cell and percentage of mitochondrial genes were comparable across age groups (Supplementary Fig. 1b–i). After pre-processing and unsupervised clustering (see Methods for details), we

categorized the dataset into 28 distinct cell types based on expression of markers from the original publication and literature<sup>23</sup>. The dataset comprised a wide variety of immune cells (e.g., alveolar macrophages, dendritic, T- and B-cells) and non-immune cells (e.g., AT1, AT2, ciliated, mesothelial, and endothelial cells) (Fig. 2; Supplementary Figs. 2 and 3). In digested lungs of both young and old mice, AT2 cells were most abundant, followed by alveolar macrophages, however their absolute numbers were higher in young mice (Supplementary Fig. 1j). A few cell types such as CD4<sup>+</sup> and CD8<sup>+</sup> T-cells, B-cells and plasma cells, ciliated cells, and monocytes were more abundant in old mice (Supplementary Fig. 1j).

We analyzed the cell type specific expression of 48 complement genes, which were classified in 5 groups as pattern recognition molecules, proteases, complement component genes, receptors and regulators<sup>19</sup>.

### Pattern recognition molecules

Specific patterns on pathogens or damaged/infected self-cells including carbohydrate chains (non-self), pathogen associated molecular patterns (PAMPs), damage associated molecular patterns (DAMPs), clustered antibodies and short-chain pentraxins such as C-reactive protein (CRP) or serum amyloid P (SAP) can attach to the surface of suspicious cells and trigger complement activation<sup>24</sup>. C1q is a member of the pattern recognition molecule family and responsible for activation of CP. A single C1q protein is composed of 3 subunits (*C1qa*, *C1qb* and *C1qc*) which are combined as a hexamer<sup>17</sup>. In the lung cells of young C57BL/6 mice, all three subunits were highly expressed in >75% of Fn1<sup>+</sup> macrophages and interstitial macrophages (Fig. 3a, b; Supplementary file 1). Fn1<sup>+</sup> macrophages were recently described as a macrophage subset that arises through early differentiation of monocytes<sup>25</sup>. Ficolin-1 (*Fcna*) and Ficolin-2 (*Fcnb*), which recognize carbohydrates and activate LP, were expressed at high levels albeit in a very small percentage of Fn1<sup>+</sup> macrophages and neutrophils, respectively (Fig. 3a). Properdin (*Cfp*) binds to PAMPs and DAMPs, initiates AP and enhances AP activation by stabilizing



**Fig. 3** Expression profile of complement pattern recognition molecules. **a** Dot plot depicting expression of genes involved in pattern recognition in the complement system in different cell types of mouse lungs. Size of the dots represents the percentage of cells expressing the gene and color intensity represents the average expression level. **b** Violin plot of highly expressed pattern recognition genes across all cell types. Vertical lines with no violin indicate that a particular gene is expressed in only a few cells within the cell type while most cells have no expression. Sections with no violin or line indicate that the gene is not expressed in the particular cell type. Data from young C57BL/6 mice (3 months old).

the AP C3/C5 convertases<sup>26</sup>. *Cfp* was highly and selectively expressed in <40% of  $\text{Fn1}^+$  macrophages and  $\text{CD209}^+\text{CD11b}^+$  dendritic cells in the lungs of young mice (Fig. 3a, b). *Crp*, which activates the complement system via CP, showed no detectable expression in the mouse dataset.

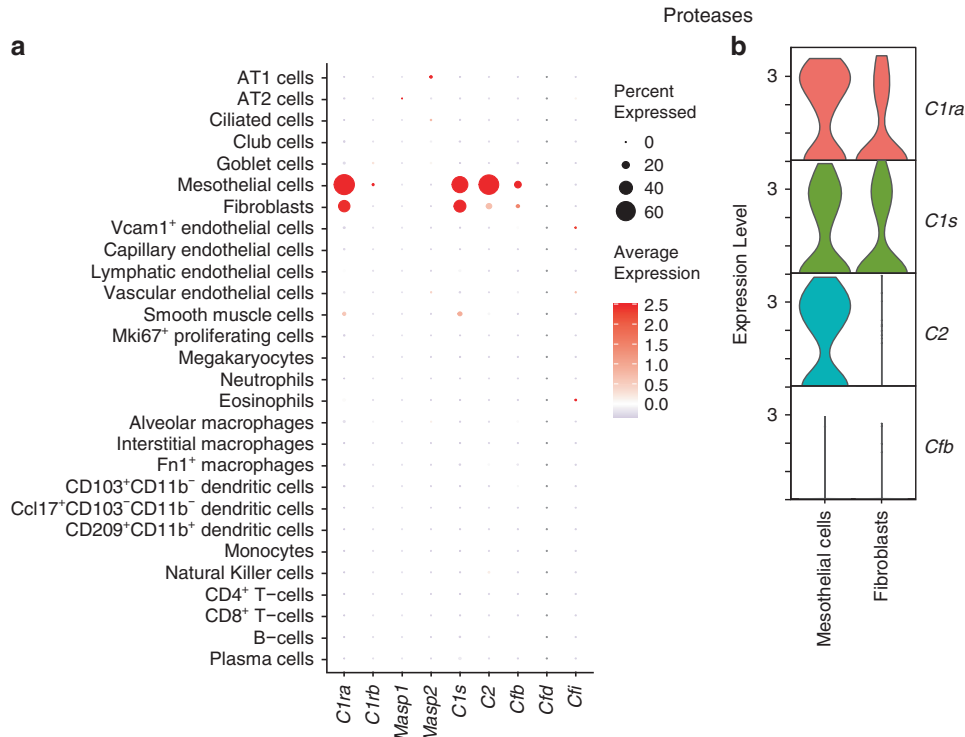
### Proteases

Serine proteases help initiate, propagate, and regulate the complement pathway, and their activation occurs sequentially with each protease activating another. Association of pattern recognition molecules (CP, LP) or complement component C3 (AP) with zymogen proteases initiates the complement pathway<sup>19</sup>. Of the nine complement serine proteases examined, we found high expression of Complement component 1, r subcomponent A (*C1ra*), Complement component 1, s subcomponent (*C1s*), and Complement component 2 (*C2*) in the lung cells of young mice (Fig. 4; Supplementary file 1). Complement factor B (*Cfb*) was expressed at comparatively lower levels. These proteases function in the CP (*C1ra*, *C1s*), AP (*Cfb*) and CP/LP (*C2*)<sup>19</sup>. In addition, expression of these genes was observed only in non-immune cells, specifically fibroblasts (*C1ra*, *C1s*), and, interestingly, mesothelial cells (*C1ra*, *C1s*, *C2*; Fig. 4; Supplementary Figs. 4a and 5a). In fact, all three proteases showed the highest expression in mesothelial cells. Other serine proteases that function in CP (Complement component 1, r subcomponent B [*C1rb*]), LP (Mannan binding lectin serine peptidase 1 and 2 [*Masp1*, *Masp2*]), and AP (Complement factors d and i [*Cfd*, *Cfi*]) showed no detectable expression in any cell types in the mouse lung dataset.

### Complement components

In addition to *C1* and *C2*, the soluble complement components *C3*, *C4*, *C5*, *C6*, *C7*, *C8* and *C9* are critical for complement activation and terminal effector functions. *C3*, *C4* and *C5* are cleaved into two fragments (a and b). The smaller fragments (*C3a* and *C5a*) are powerful anaphylatoxins<sup>19</sup>. *C3b* can act as an opsonin and serve as the *C3* convertase (AP) or *C5* convertase (CP, LP and AP). *C4b* is part of both *C3* (*C4b2b*) and *C5* (*C4b2b3b*) convertases in CP and LP. In humans, mice, and rats, the *C4* gene exists in isotopic forms, chiefly *C4a* (acidic) and *C4b* (basic) with specific functions such as immunoclearance and hemolysis<sup>27</sup>. In the following text, *C4a* and *C4b* refers to the isotypes of the *C4* gene, rather than the cleavage product of *C4* protein. *C5b* triggers the terminal complement pathway wherein binding of *C5b* to *C6* initiates formation of membrane attack complex (MAC) by components *C5b*, *C6*, *C7*, *C8* and *C9*, leading to lysis of the target cell. *C8* encodes a trimeric complex comprising *C8a*, *C8b* and *C8g* subunits (as described in humans<sup>28</sup>).

Both *C3* and *C4b* were highly expressed in a large percentage of mesothelial cells (Fig. 5a, b; Supplementary file 1), with lower expression in fibroblasts (<50% of cells) of young mice. *C5*, on the other hand, was highly expressed in ~75% of AT2 cells, and at lower levels in many other cell types, predominantly eosinophils, smooth muscle cells and  $\text{CD8}^+$  T-cells (Fig. 5a, b). Of note, *C3* and *C5* were largely not co-expressed in the same cell types (Fig. 5c), and only detected together in  $n = 395$  (~5%) cells of the entire dataset (Fig. 5d). The expression of the other complement components, namely *C4a*, *C6*, *C7*, *C8* (*C8a* and *C8g*) and *C9*, was low in the mouse dataset (Fig. 5a). *C8b* expression was not detectable at all in these mice (data not shown).



**Fig. 4 Expression profile of complement proteases.** **a** Dot plot depicting expression of genes encoding proteases in the complement system. **b** Violin plot showing differential expression of the proteases in selected cell types that showed highest expression levels in **a**. Data from young C57BL/6 mice (3 months old).

### Receptors

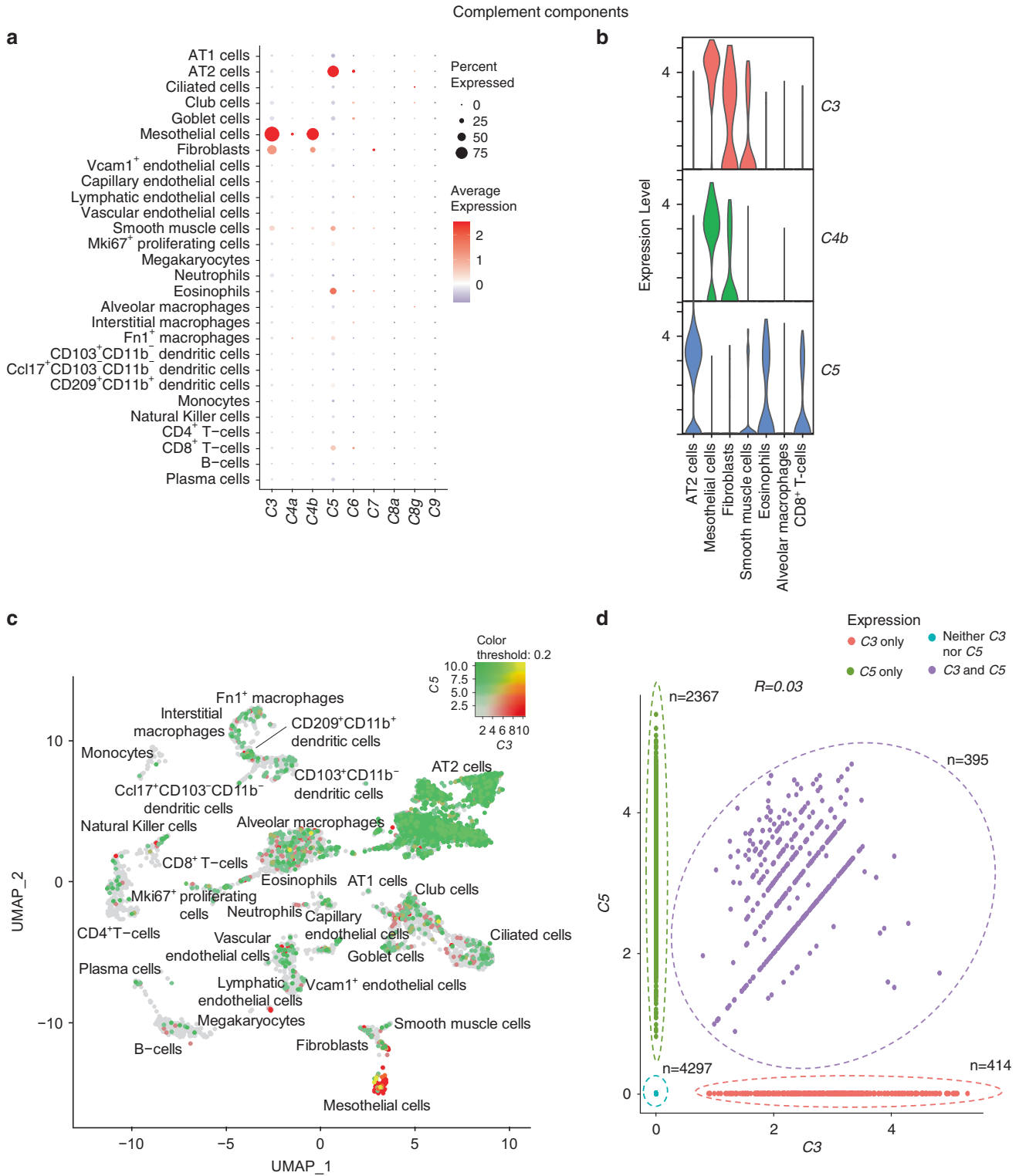
Complement receptors are a diverse group of proteins that play key roles in inflammatory responses through leukocyte recruitment and formation of phagosomes<sup>19</sup>. The complement receptors we studied included complement component (3b/4b) receptor 1-like (*Cr1l* in mice, *CR1* in humans), Complement receptor 2 (*Cr2*), *CR3* (complex of integrins *Itgax* [CD11b] and *Itgb2* [CD18]), *CR4* (integrins *Itgax* [CD11c] and *Itgb2* [CD18]), Complement component 3a receptor 1 (*C3ar1*), Complement component 5a receptors 1 and 2 (*C5ar1*, *C5ar2*), V-set and immunoglobulin domain containing 4 (*Vsig4*/CRlg), Calreticulin (*Calr*/cC1qR), Complement C1q binding protein (*C1qbp*/gC1qR), and the Complement component 1q receptor 1 (*Cd93*/C1qRP)<sup>19</sup>. *Itgb2* was highly expressed in ~40% of the monocytes, and at lower levels in neutrophils and alveolar macrophages of young mice (Fig. 6a, b; Supplementary file 1). *Itgam* was expressed in the rare population of neutrophils at low levels and in <20% of monocytes, CD209<sup>+</sup>CD11b<sup>+</sup> dendritic cells, interstitial macrophages and Fn1<sup>+</sup> macrophages, while *Cd93* was expressed in vascular endothelial cells and at lower levels in Vcam1<sup>+</sup> endothelial cells (Fig. 6a, b). *Itgax* was expressed in ~20% of alveolar macrophages (Fig. 6a). Calreticulin (*Calr*) was more ubiquitous, with high expression observed in ~60–80% of mesothelial cells, plasma cells and nearly 50% of fibroblasts (Fig. 6a, b). Comparatively fewer percentages of other cell types such as AT2, alveolar macrophages, megakaryocytes and Vcam1<sup>+</sup> endothelial cells also weakly expressed *Calr* (Fig. 6). Of the anaphylatoxin receptors, *C3ar1* expression was high in <20% of interstitial and Fn1<sup>+</sup> macrophages, *C5ar1* in <20% of alveolar, interstitial and Fn1<sup>+</sup> macrophages while *C5ar2* was hardly detectable in any of the studied cell types. *Cr1l* was expressed in <20% of monocytes (Fig. 6a). Overall, expression of receptors trended more towards immune cells when compared with non-immune cells (Supplementary Figs. 4a and 5a), although this was only statistically significant for *Itgb2* and *Itgax* (adjusted  $p < 0.05$ ; Supplementary file 2).

### Regulators

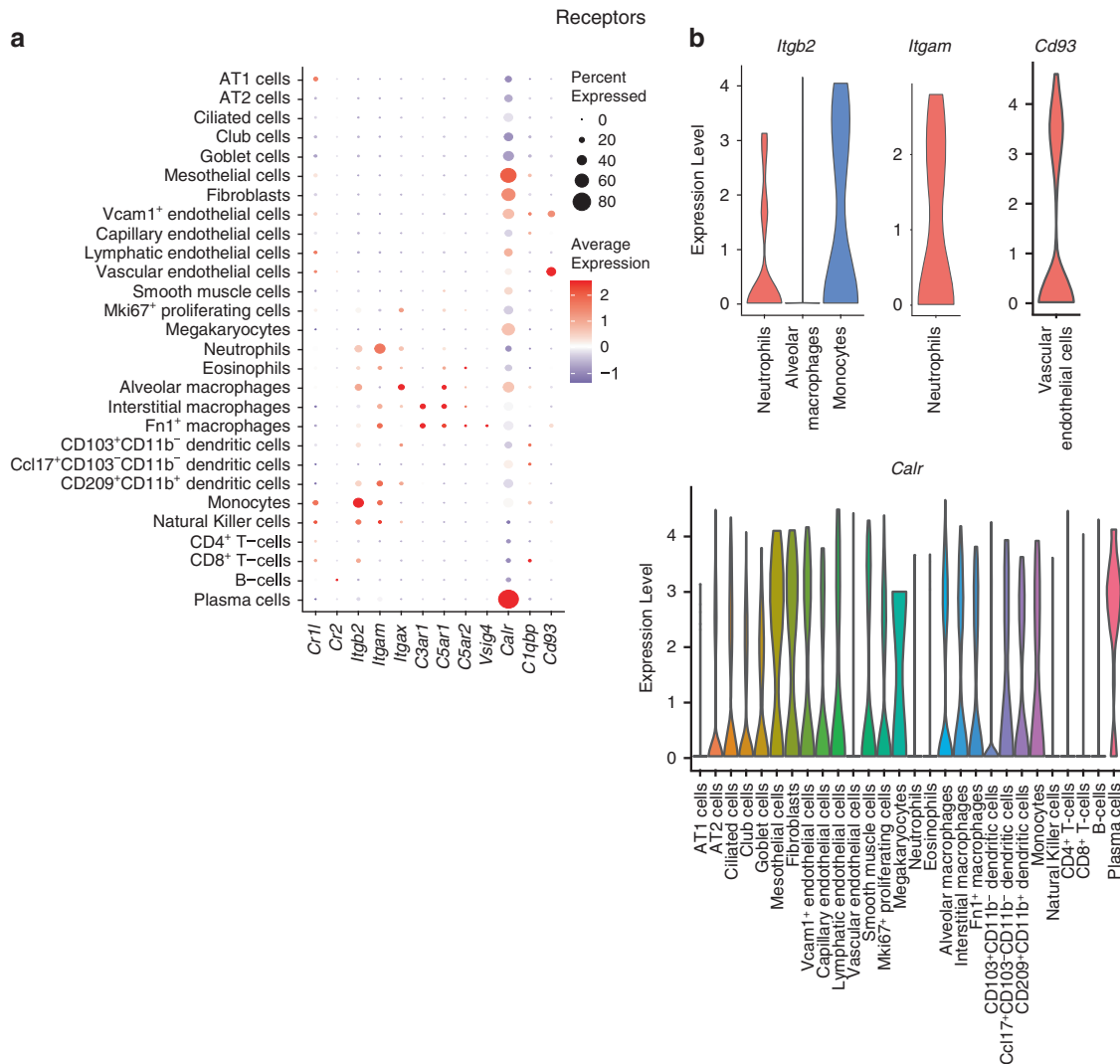
Complement regulators modulate complement activation, persistence, and function either through inhibition of complement proteases, or competitive binding to pattern recognition molecules and complement convertases<sup>19</sup>. Among the different regulators of the complement system studied, C1-inhibitor (C1-INH/*Serp1*), Factor H (*Cfh*), and Clusterin (*Clu*) were highly expressed in the lungs of young mice (Fig. 7a, b; Supplementary file 1). Expression of *Serp1*, *Cfh* and *Clu* was observed in at least 70% of the mesothelial cells; *Serp1* and *Cfh* were also expressed in fibroblasts. *Clu* was also expressed in goblet cells, capillary endothelial cells, and megakaryocytes. Decay-Accelerating Factor (*Cd55*/DAF) was detected in a few ciliated cells and club cells (Fig. 7a, b). Expression of regulator genes was not detectable in most of the immune cell types annotated in young mice (Supplementary Fig. 4a). Other regulators such as Membrane Cofactor Protein (*Cd46*/MCP), Complement Component 4 Binding Protein (*C4bp*), Complement Factor H Related 1 (*Cfhr1*), Membrane Attack Complex Inhibition Factor (*Cd59b*), Vitronectin (*Vtn*), Carboxypeptidase-N (*Cpn1*, *Cpn2*) and Carboxypeptidase-B2 (*Cpb2*; data not shown) did not show any relevant mRNA expression in the lung cells identified in the murine dataset.

### Complement expression in old mice

A comparative analysis of expression trends of the pattern recognition molecules between old and young mice showed much more overlap than differences, with some notable exceptions (Fig. 8; Supplementary Figs. 4 and 5). Complement component *C5* appeared even more strictly confined to AT2 cells of old mice (Fig. 8; Supplementary Fig. 5). Interestingly, *C5* expression was not detected in eosinophils and CD8<sup>+</sup> T-cells in old mice in contrast to young mice (adjusted  $p < 0.05$ ; Supplementary file 3). There were also minor differences in expression profiles of complement receptors *Itgb2* and *Itgam* between old and young mice, though these differences were not statistically



**Fig. 5 Expression profile of complement component genes.** **a** Dot plot depicting expression of genes encoding complement components C3–C9. **b** Violin plot showing selected cell types with differential expression of complement component genes. **c** Uniform Manifold Approximation and Projection (UMAP) plot illustrating differential expression of C3 and C5 in lung cells. Clusters expressing C3 are colored red, C5 colored green, clusters expressing neither are colored grey and clusters co-expressing both are colored yellow. Intensities of the colors for each gene vary from light to dark depending on the expression level (low to high). **d** Scatter plot showing expression of C3 and C5 across the cell types. Pearson’s correlation coefficient ( $R$ ) at the top of the plot indicates the extent of correlation between the genes. Number of cells expressing the gene(s) or not expressing either gene are indicated. Data from young C57BL/6 mice (3 months old).



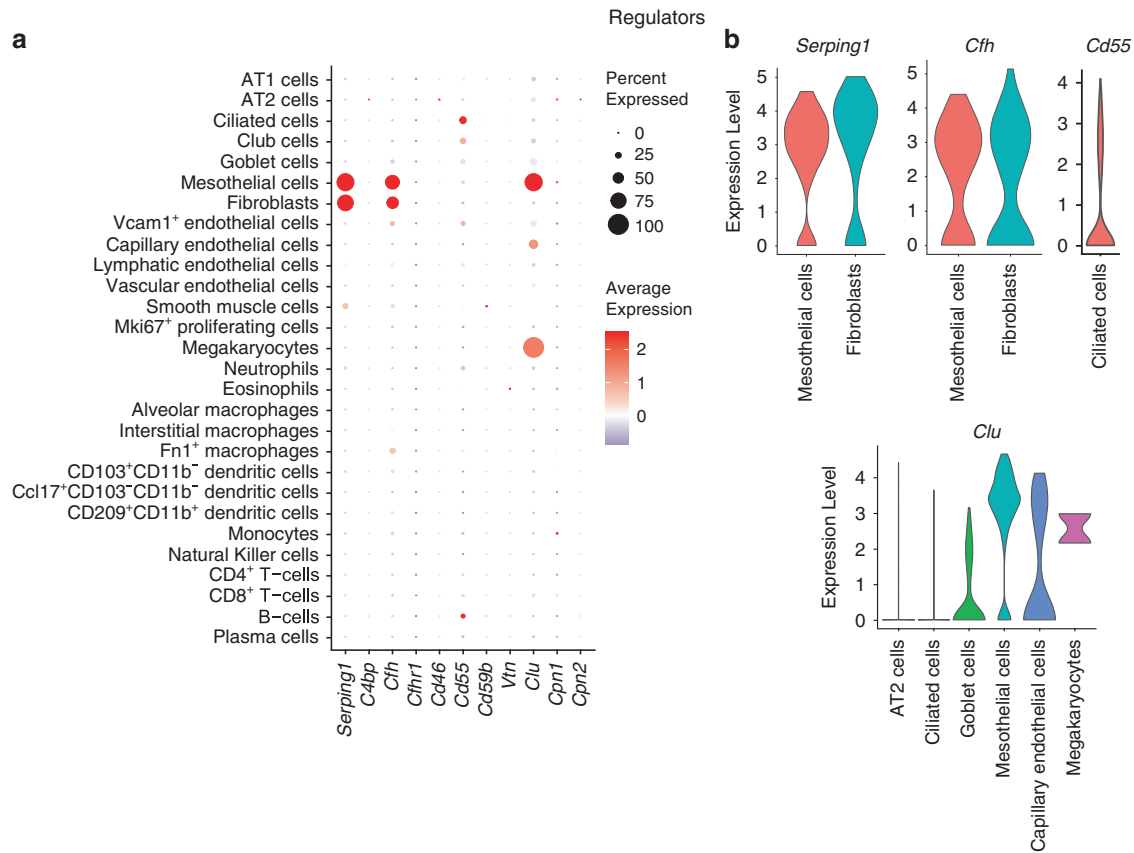
**Fig. 6** Expression profile of complement receptor genes. **a** Dot plot showing expression of receptor genes in different cell types. **b** Violin plot highlighting distribution of highly expressed receptor genes in various cell types. Data from young C57BL/6 mice (3 months old).

significant. *Itgb2* expression was comparatively higher in alveolar macrophages, Fn1<sup>+</sup> macrophages, interstitial macrophages, CD209<sup>+</sup>CD11b<sup>+</sup> dendritic cells, eosinophils, and CD4<sup>+</sup> T-cells in old mice versus young mice. *Itgam* was expressed in a greater number of neutrophils, Fn1<sup>+</sup> macrophages, monocytes, and Natural Killer cells from old mice versus young mice (Fig. 8; Supplementary Fig. 4). Expression of complement regulator *Cfh* was relatively higher in Vcam1<sup>+</sup> endothelial cells in old mice compared with young mice (Fig. 8, adjusted  $p > 0.05$ ; Supplementary file 3). The interpretation of age-dependent complement gene expression in lungs may need to take into consideration that differences in numbers of some cell populations existed. For example, the samples from old mice contained fewer AT2 cells and alveolar macrophages, but higher numbers of T-cells and B-cells (Supplementary Fig. 1).

### Complement expression in human lungs

To understand the similarities and differences in complement expression between mice and humans<sup>29</sup>, we also analyzed a human lung dataset generated by Travaglini et al.<sup>3</sup> This dataset comprised of 57 distinct cell types as reported earlier (Supplementary Fig. 6).<sup>3</sup> There were similarities and a few differences in the expression patterns of complement genes between human and mouse lungs. The most striking similarity was the high

expression of several complement genes in the mesothelial cells (Fig. 9; Supplementary file 4), as observed in mouse lungs (Fig. 8). In human lungs, *C1QA–C* were expressed additionally in dendritic cells and platelets/megakaryocytes and *FCP* in monocytes. Interestingly, we observed expression of *FCN1* and *FCN3* in human monocytes and endothelial cells (including capillary intermediate cells), respectively (Fig. 9), while they were hardly detectable in the mouse lungs. High expression of *C1R*, *C1S* and *C2* was observed in human mesothelial cells and certain fibroblast sub-types (Fig. 9), quite similar to mice (Fig. 8). Moreover, human *C2* expression was also identified in macrophages, TREM2<sup>+</sup> dendritic cells and platelets/megakaryocytes (Fig. 9). In contrast to mice, *CFI* was expressed highly in human mesothelial cells and lipofibroblasts, and at lower levels in a few epithelial and endothelial cell sub-types. *C3* was expressed at high levels in human mesothelial cells and lipofibroblasts, with lower expression in adventitial fibroblasts, goblet cells, mucous cells and AT2 cells. Expression patterns of complement components *C4–C8* were different between human and mouse lungs (Figs. 8 and 9). *C4B* was not detectable in human lungs (data not shown). *C5* was expressed in ~25% of AT2 cells, which was a lower percentage than for mouse AT2 cells. *C3* and *C5* were only detected together in  $n = 741$  (~1%) cells of the entire dataset (Supplementary Fig. 7). *C6* was detectable in human ciliated cells and *C7* in human fibroblast clusters, vein endothelial



**Fig. 7 Expression profile of complement regulator genes.** **a** Dot plot showing expression pattern of regulator genes. **b** Violin plot illustrating the cell-specific distribution of highly expressed regulator genes. Data from young C57BL/6 mice (3 months old).

cells and bronchial vessels 1. *C8B* was expressed in 25–50% of cells in the macrophage clusters (Fig. 9; Supplementary file 4). The complement receptors *ITGAM*, *ITGAX*, *ITGB2*, *C3AR1*, *C5AR1*, *VSIG4* were chiefly expressed in the myeloid cell clusters. *ITGB2* was also detectable in lymphocyte clusters, including subtypes of Natural Killer cells,  $CD4^+$  T-cells and  $CD8^+$  T-cells. *CALR* expression was ubiquitous with highest levels observed in endothelial, goblet and plasma cells, ionocytes, and some of the mesenchymal clusters. *C1QB* was expressed mostly in goblet cells, ionocytes, basal cells, endothelial cells, lipofibroblasts, fibromyocytes, and plasmacytoid dendritic cells. *CD93* was expressed in endothelial cell clusters, similar to mouse lungs (<50% of cells; Fig. 8; Supplementary file 1), dendritic cells and monocytes; however, *CD93* was more abundant in human lungs (>50% of cells; Fig. 9; Supplementary file 4). The complement regulator *SERPING1* was expressed abundantly in mesenchymal cell clusters including mesothelial cells and fibroblasts, and at lower levels in the macrophage clusters. *C4BPA* showed high expression in ~60% of AT1, >70% of AT2 and signaling AT2 cell clusters, and ~50% of club cells (Fig. 9; Supplementary file 4). Expression of *CFH* was prominent in the human mesothelial and fibroblast cell clusters (Fig. 9), similar to mouse lungs (Fig. 8). *CD46*, *CD55* and *CD59* appeared more ubiquitously expressed in human lungs, with high levels observed in epithelial and endothelial cell clusters. Approximately 50% of human neuroendocrine cells expressed *VTN*. *CLU* was highly expressed in human myofibroblasts, fibromyocytes, neuroendocrine cells and vein endothelial cells (Fig. 9). Interestingly, *CPB2* was expressed in 30–50% of human AT2 cell clusters, whereas it was not detectable in mouse lungs (data not shown). Notably, ionocytes, which were not present in the mouse dataset, showed high expression of *CALR*, *C1QB*, *CD46* and *CD59* (Fig. 9). Capillary aerocytes, a recently described specialized endothelial cell sub-

type involved in diffusion at the gas exchange surface and leukocyte migration in the lung<sup>6</sup>, displayed low expression of *CD59*, *CD46*, *CALR* and *SERPING1*.

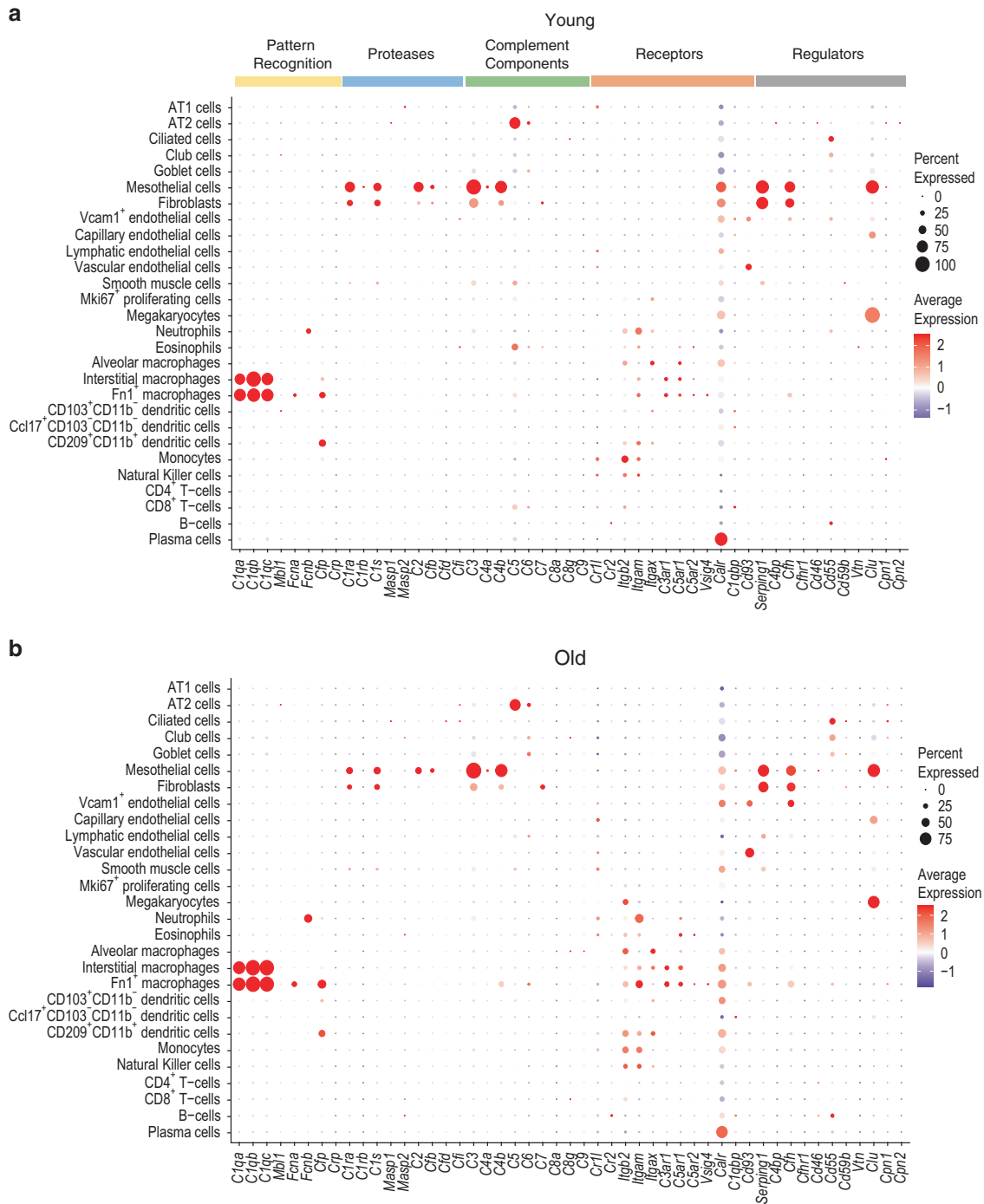
In conclusion, the dataset from three human patients recapitulated major findings in mice in terms of expression of several complement proteins in mesothelial cells and *C5* in AT2 cells, although differences for some cell types and complement genes were also noted.

## DISCUSSION

In this study, we present a comprehensive profile of the localized expression of complement genes using bioinformatics analysis of open access mouse and human scRNA-Seq atlases. The datasets included 28 different cell types from mouse and 57 cell types from human lungs. Our results highlight expression of complement pattern recognition molecules, proteases, components, receptors, and regulators by different cell types in the lung, encompassing both immune and non-immune cells. These findings support the presence of localized, tissue-specific complement proteins in the lungs during inflammation and indicate a potential role for their non-canonical functions through molecular crosstalk with other pathways, and cell-specific homeostasis.

An interesting observation from our analysis is the expression of several proteases, complement components, receptors, and regulators at mid-to-high levels in the mesothelial cell cluster in both mice and humans. Mesothelial cells form a border surface and play an important role in both innate and adaptive immunity of the pleural compartment<sup>9</sup>. A study in patients with tuberculosis pleurisy reported that complement activation in the mesothelial cells might lead to increased stimulation of chemokines and monocyte chemotaxis into the pleural space contributing to



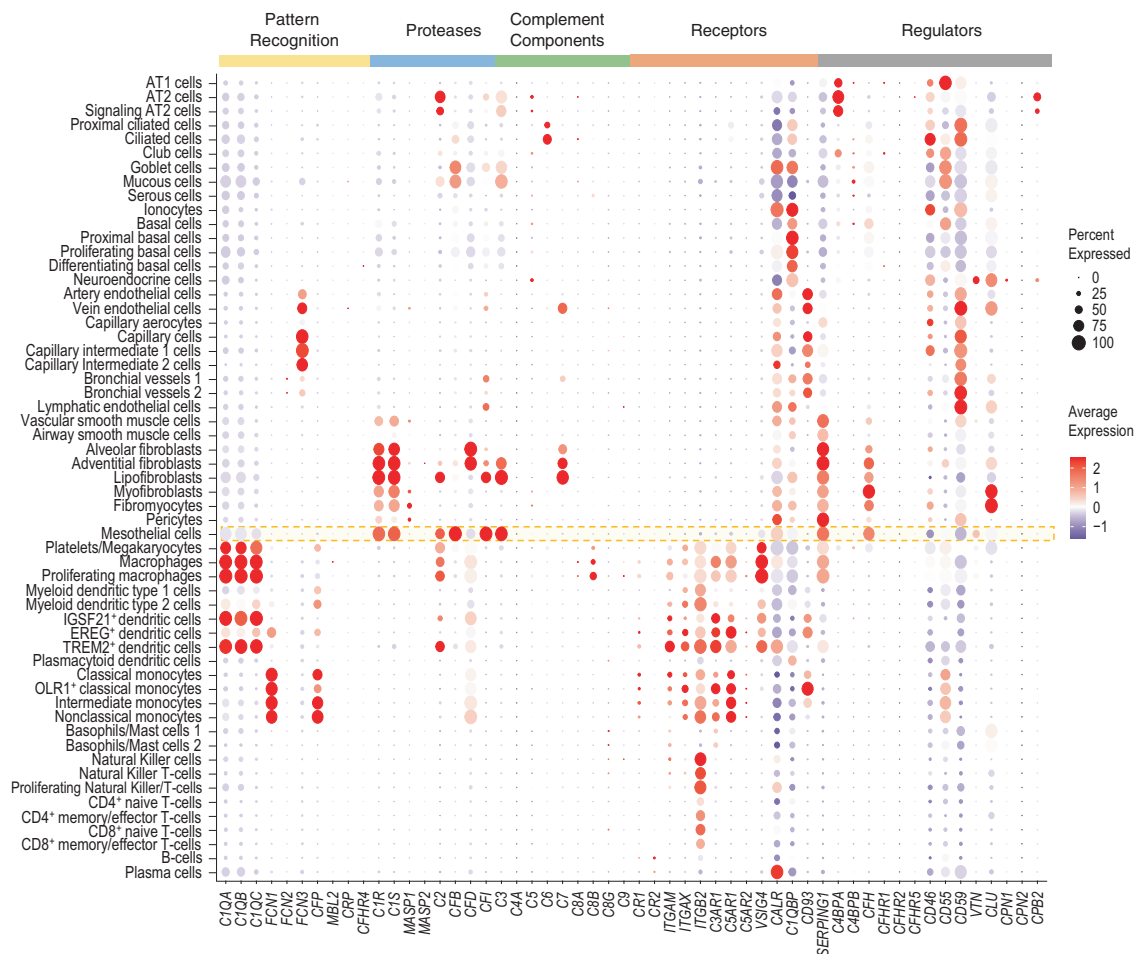


**Fig. 8 Comparison of complement genes in the lungs of young and old C57BL/6 mice.** Dot plot illustrating the expression pattern of all complement genes in various cell types identified in the single cell transcriptome data for **(a)** young mice (aged 3 months,  $n = 8$ ). **(b)** Old mice (aged 24 months,  $n = 7$ ). Size of the dots represents the percentage of cells expressing a gene and color intensity represents the average expression level. Panel ‘a’ was prepared from the same raw data analyzed for Figs. 2–7.

disease pathophysiology<sup>30</sup>. Expression of mouse and human complement proteases (*C1R*, *C1S*, *C2*) and regulators (*SERPING1*, *CFH*, *CLU*) in steady-state mesothelial cells supports their dual role to fuel local complement activation against invading pathogens but limit unwanted complement-mediated damage to host cells.

Another key finding from our analysis is that *C3* and *C5* are highly expressed by distinct cell types in the homeostatic state. *C3* was detected primarily in mesothelial cells and fibroblasts in murine and human lungs, as well as goblet, mucous, club and AT2

cells in humans. *C5* was mainly detected in AT2 cells. Lung fibroblasts play an important role in host immune responses through production of cytokines and lipid mediators. Intracellular expression and non-canonical cleavage of *C3* in fibroblasts might be a mechanism for homeostasis and rapid cell-specific responses to inflammation. Studies of steady-state conditions suggest that localized pools of *C3* have a role in cellular processes such as cell cycle and gene regulation<sup>31</sup>, and serve as reservoirs that are activated by intracellular proteases (e.g., cathepsins)<sup>32</sup>. *C3* expression has also been observed in human alveolar macrophages,



**Fig. 9 Complement gene expression in healthy human lung tissues.** Dot plot illustrating the expression pattern of all complement genes in various cell types annotated in the single cell transcriptome data of healthy lung tissue of human patients ( $n = 3$ ). Size of the dots represents the percentage of cells expressing a gene and color intensity represents the average expression level. The data is based on scRNA-sequencing from healthy, uninvolved lung regions of patients undergoing lobectomy for pulmonary tumors. Complement gene expression in mesothelial cells has been highlighted in orange.

fibroblasts, B-cells, monocytes, neutrophils, epithelial, endothelial cells, and mouse AT2 cells<sup>33,34</sup>, highlighting similarities and differences in expression patterns compared to our analysis. Constitutive C5 expression might facilitate the immunoregulatory and repair functions of AT2 cells<sup>35</sup>, with levels regulated during infection and lung injury. Intracellular expression of C5 may play a role in inflammasome assembly, wherein activation of cells results in increased production of anaphylatoxin C5a<sup>36</sup>. This in turn stimulates C5aR1 and subsequently induces reactive oxygen species (ROS), TNF $\alpha$  and IL-1 $\beta$  production, associated with NLRP3 inflammasome assembly<sup>32,36</sup>. A similar role has been ascribed to C3 in NLRP3 inflammasome activation<sup>32,36</sup>.

A previous study reported human Properdin (*CFP*) expression in epithelial cells<sup>16</sup> and alveolar macrophages<sup>37</sup>. However, we observed *CFP* expression only in murine Fn1<sup>+</sup> macrophages, CD209<sup>+</sup>CD11b<sup>+</sup> dendritic cells as well as human monocytes, dendritic cells, and platelets/megakaryocytes. In healthy mice, Ficolin expression was detected in neutrophils, macrophages, and lymphocytes, through western blotting<sup>38</sup>. Although RNA expression of *Fcna* and *Fcnb* in Fn1<sup>+</sup> macrophages and neutrophils, respectively, supports the previous findings from mice, their expression was limited to a few cells in our analysis. These differences might be due to variations in transcriptional and translational rates<sup>39</sup> and, possibly, sparsity in single cell data<sup>40</sup>. Human *FCN3* does not have an ortholog in mouse and was mainly expressed in the endothelial cell clusters. The high expression of

*C1qa*, *C1qb* and *C1qc* in mouse interstitial macrophages and Fn1<sup>+</sup> macrophages corroborate other murine lung studies<sup>41</sup>. Expression of *C1QA–C* in human lung cells has not been well characterized, though expression has been reported in macrophages and dendritic cells similar to our analysis<sup>42</sup>.

While high expression of the C1 inhibitor, *Serp1*, has been reported earlier in the normal mouse lung<sup>43</sup>, its cellular sources have not been characterized. We observed *Serp1/SERP1* expression in mouse and human mesothelial cells and fibroblasts, where it may limit unwanted complement activation in the lung interstitium and pleural space. The highest expression of Cd55/*CD55*, *CD46* and *CD59* was detected in airway epithelial cells and/or endothelial cells, comparable to earlier reports<sup>16,44</sup>. *Cfh* was mainly detected in mouse and human mesenchymal cells and in lower quantities in human epithelial cells in our analysis. *Cfh* expression has not been characterized previously in lung cell types. Carboxypeptidase *N* (*Cpn1*, *Cpn2*) and *B* (*Cpb2*) were not expressed in any of the mouse cell types in our study, while human *CPB2* was detected in AT2 cell clusters. Human *CPN* and *CPB2* have been largely characterized in the plasma and pancreas, respectively, and although mouse *Cpn1* has been detected in lungs, its cellular origin was not defined<sup>45</sup>.

Mouse and human *Cr1/CR1*, *CR3* and *CR4* expression were largely aligned with previous reports in immune cells<sup>46</sup>. *C3aR1* expression has been previously reported in human and murine bronchial epithelial and smooth muscle cells<sup>47</sup>, *C5AR1* in human

alveolar epithelial cells, endothelial cells, and smooth muscle cells<sup>48</sup>, and murine *C5ar1* and *C5ar2* in neutrophils, eosinophils, alveolar macrophages, CD11b<sup>+</sup> dendritic cells, and monocyte-derived dendritic cells<sup>49</sup>. We detected low mouse *C3ar1* and *C5ar1* expression, and almost negligible *C5ar2* expression, restricted to macrophages. Human *C3AR1* and *C5AR1* showed high expression in myeloid cells in our analysis. CR1g (*Vsig4*), reported earlier in murine alveolar macrophages through western blotting<sup>50</sup>, showed almost no expression in any mouse cell type, whereas human *VSIG4* was strongly expressed in macrophages, dendritic cells, and platelets/megakaryocytes. While not specific to the lung, the two C1q receptors, cC1QR (*CALR*) and gC1QR (*C1QBP*) were reported in human macrophages, dendritic, B-cells, T-cells, and endothelial cells<sup>51</sup>. C1qRP (*CD93*) has been reported largely in human non-lung endothelial cells, monocytes, and neutrophils<sup>52</sup>, and in murine endothelial (non-lung) and alveolar epithelial cells<sup>53</sup>. The variation in the expression profiles between published reports and our analysis could be due to the differences in experimental approaches, or possibly localized profiles in the lung compared to other tissues, although further studies are needed.

Given the important role that complement genes play in modulating immune responses and homeostasis, any imbalance in the complement pathway can result in deleterious effects, as reported in various lung diseases. These include the C5a-mediated cellular damage observed in acute lung injury and the highly fatal acute respiratory distress syndrome<sup>54,55</sup>. In both bacterial and viral pneumonia, the pathogens escape complement-mediated eradication through either inhibition of CP or AP genes<sup>56</sup>. C3, C5 and their receptors have also been implicated in pathogenesis of asthma, cystic fibrosis, and idiopathic pulmonary fibrosis<sup>16,55</sup>. In COVID-19, excessive activation of C3 and C5a have been linked to increased viral loads, coagulation, and lung injury<sup>20,57</sup>.

Although our study provides a comprehensive analysis of complement expression in different cell types, there exist a few limitations. Firstly, several newly identified cell types such as ionocytes were absent in the mouse dataset, which may be due to the enzymatic isolation protocol. Nevertheless, the 28 cell types that were identified in mice represent a broad range that have versatile functions. Secondly, we may underestimate the expression of some of the complement genes due to low sensitivity (low sequencing depth of the mouse dataset), sequencing bias, and lower numbers for some cell types. This is an inherent limitation in single cell sequencing technology where many factors determine the balance between sequencing depth and the number of cells being sequenced. Thus, scRNA-Seq might be less sensitive in detecting expression when compared to technologies such as RT-qPCR. RNA-based data is also observed to have low correlation with protein abundance, with studies reporting <40% correlation, ascribed to differences in regulation of transcription and translation. Hence, the expression patterns in this analysis may not necessarily reflect the functional activity of the encoded proteins in the cell types analyzed. Another limitation is that our study is based on steady state. Complement expression patterns in pathophysiological conditions may be substantially different compared to healthy lungs. Finally, we based the cell type identities in the mouse dataset largely on the Angelidis et al. publication and extant markers<sup>23</sup>. Due to the evolving landscape of cell type mapping in the lungs, absence of consensus on nomenclature of certain cell subtypes and relatively low numbers of some cells in the mouse and human datasets, we suggest caution while considering findings from heterogeneous populations such as fibroblasts. Moreover, the mesothelial cells are either rare or not present in typical samples of lung tissue for single cell sequencing. Despite these limitations, our findings provide a comprehensive profile of cell type specific expression of complement genes that may serve as a basis for more in-depth and targeted investigations.

Recent advances in single cell technologies, including multiomics approaches that can profile the transcriptome, proteome, and epigenome together, may help in developing a more comprehensive map in future studies. Since C3 and C5 are considered the central hubs of complement activation, and dysfunction in disease, therapeutic targeting has been mainly focused on blockade of these proteins, with several drugs such as eculizumab, ravulizumab (C5), avacopan (C5aR1) approved for use, and several more drug candidates in development. Given the complex functioning and interactivity of complement genes within multiple pathways, an inclusive design strategy driven by cell-specific cues might result in novel, targeted therapies with reduced side effects. Our findings from mouse and human lungs might serve as an initial road map that can help guide the design of such therapies in the long-term future.

## MATERIALS AND METHODS

### Data retrieval

Raw counts and metadata for the scRNA-Seq mouse lung atlas was obtained from the NCBI Gene Expression Omnibus database (accession number GSE124872)<sup>23</sup>. The dataset comprised 14813 cells: 7672 cells from 8 mice aged 3 months ('young') and 7141 cells from 7 mice aged 24 months ('old'). Angelidis et al.<sup>23</sup> isolated single cells from the lung tissue of male and female mice through enzymatic digestion, performed red blood cell lysis, and assessed the viability of cells by counting in a Neubauer chamber. The cells (100 cells/μL) were further processed using a Dropseq protocol and were sequenced on the Illumina HiSeq4000 platform<sup>23</sup>. The Dropseq computational workflow was used to demultiplex the data, align the sequences to mm10 reference genome and quantify the reads to generate the raw counts matrix comprising the number of UMIs per gene per cell<sup>23</sup>. This raw count matrix was the input for our analysis.

For human lung atlas, data generated by Travaglini et al. was analyzed<sup>3</sup>. Travaglini et al. isolated and dissociated lung tissues from uninvolved regions of three patients (aged 46 years [male], 51 years [female], 75 years [male]) undergoing lobectomy for lung tumors, together with peripheral blood<sup>3</sup>. The cells were sequenced using 10X 3' v2 Chromium and Smart-seq2 protocols using the Illumina NovaSeq 6000 platform. The sequences were demultiplexed, aligned to human reference genome (GRCh38.p12) and quantified using 10X Genomics Cell Ranger v2. The data was filtered to only include cells with greater than 500 genes and 50,000 mapped reads. The raw counts matrix was processed using Seurat and annotated as reported by Travaglini et al. using canonical markers, excluding doublets based on expression of multiple cell type markers<sup>3</sup>. The processed data from lungs with cell type annotations by Travaglini et al.<sup>3</sup> was retrieved from cellxgene<sup>58</sup> in h5ad format for our analysis. The Smart-seq2 dataset was not used for our analysis as it did not include mesothelial cells.

### Data analysis

The mouse and human datasets were analyzed using the R framework (v 4.0.2)<sup>59</sup> in RStudio Server (v 1.3.1073)<sup>60</sup>.

**Mouse dataset.** Based on quality control analysis, the dataset was filtered to exclude low-quality cells and only cells expressing greater than 200 genes, greater than 200 UMIs/cell, and less than 15% mitochondrial genes were used for downstream analysis. To exclude any variation in the data due to cell cycle, we normalized the data using Seurat's (v. 4.0.1)<sup>61</sup> `NormalizeData` function and, calculated and examined cell cycle phase scores based on expression of canonical markers. The dataset was split into young and old based on age of the mice, and the data subsets were normalized and scaled separately using `NormalizeData` and `SCTransform` functions, regressing out variation due to high mitochondrial gene expression<sup>61</sup>. By default, `SCTransform` retains the top 3000 most variable genes, which were used as anchors to integrate the two datasets (young and old) for dimensional reduction and clustering. Seurat integrates datasets based on shared cell populations, correcting for any technical differences (e.g., sequencing depth and batch effects), while preserving biological heterogeneity. Linear dimensional reduction was performed through principal component analysis and the top 42 principal components (PCs) that explained the highest variance in the dataset, were used for clustering. Nearest neighbors were identified based on the K-nearest neighbor (KNN) graph using `FindNeighbors` function and the cells were

grouped based on the Louvain algorithm using the FindClusters function, resulting in 44 clusters. Differential expression analysis was performed using the FindAllMarkers function. Cluster annotations were based on canonical markers from literature and included subsetting and re-clustering a few groups to enhance identification of cell types. In total, the dataset was classified into 28 cell types. Clusters that showed substantial expression of multiple cell type markers in a single population were identified as doublets and excluded from further analysis. Analysis using Single Cell Toolkit (singleCellTK)<sup>52</sup> package estimated a very low proportion of doublets (0.6–4%) and low ambient RNA (median contamination score: <0.1) in the processed mouse dataset. Non-linear dimensional reduction was performed using Uniform Manifold Approximation and Projection (UMAP) and was used to visualize the clusters in two-dimensional space. All dot plots, violin plots and heatmaps, and differential expression analysis were based on the RNA assay.

**Human dataset.** The processed data from cellxgene was converted to Seurat file format using SeuratDisk (v0.0.0.9019)<sup>63</sup>. The dataset was subset to exclude cells from the peripheral blood samples. No filtering was applied for the quality metrics for this processed dataset. The data was re-normalized using NormalizeData function in Seurat and integrated across patients using Seurat wrapper for Harmony<sup>64</sup>. Non-linear dimensional reduction was performed using UMAP based on the 30 principal components encoding the highest variance in the data and was used to visualize the clusters in two-dimensional space. Cell type (free) annotations by Travaglini et al.<sup>3</sup> were retained and used to analyze complement gene expression. Neutrophils, reported in the original dataset, were not present in the annotated cellxgene dataset that we analyzed.

### Statistical analysis

The Wilcoxon Rank Sum Test (two-sided) through Seurat's FindAllMarkers function was used to investigate differential gene expression between clusters/cell types, with the threshold parameters set to log<sub>2</sub> fold-change (logFC) >0.20, min.pct of 0.1 to only test genes that are expressed in at least 10% of the cells, and multiple testing using Bonferroni correction. In addition, to test statistical significance between groups of mice (old versus young; immune versus non-immune), the two-sided Wilcoxon Rank Sum Test using Seurat's FindMarkers function was run with the default Seurat parameters of logFC threshold of 0.25 (old versus young; immune versus non-immune cells) and min.pct of 0.1. All differential analysis were based on expression levels. The threshold for significance was set at 0.05 and adjusted *p*-values < 0.05 were considered statistically significant.

### DATA AVAILABILITY

The mouse dataset for Angelidis et al. can be accessed from Gene Expression Omnibus with accession number GSE124872<sup>23</sup>, and the human dataset from Travaglini et al.<sup>3</sup> can be accessed via cellxgene<sup>58</sup> or Synapse (<https://www.synapse.org/#!Synapse:syn21041850>).

### REFERENCES

- Franks, T. J. et al. Resident cellular components of the human lung: current knowledge and goals for research on cell phenotyping and function. *Proc. Am. Thorac. Soc.* **5**, 763–766 (2008).
- Plasschaert, L. W. et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).
- Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
- Weibel, E. R. On the tricks alveolar epithelial cells play to make a good lung. *Am. J. Respir. Crit. Care Med.* **191**, 504–513 (2015).
- Shenoy, A. T. et al. Antigen presentation by lung epithelial cells directs CD4(+) T (RM) cell function and regulates barrier immunity. *Nat. Commun.* **12**, 5834 (2021).
- Gillich, A. et al. Capillary cell-type specialization in the alveolus. *Nature* **586**, 785–789 (2020).
- Schupp, J. C. et al. Integrated single-cell atlas of endothelial cells of the human lung. *Circulation* **144**, 286–302 (2021).
- Sveiven, S. N. & Nordgren, T. M. Lung-resident mesenchymal stromal cells are tissue-specific regulators of lung homeostasis. *Am. J. Physiol. Lung Cell Mol. Physiol.* **319**, L197–L210 (2020).
- Batra, H. & Antony, V. B. Pleural mesothelial cells in pleural and lung diseases. *J. Thorac. Dis.* **7**, 964–980 (2015).
- Mutsaers, S. E. Mesothelial cells: Their structure, function and role in serosal repair. *Respirology* **7**, 171–191 (2002).
- Ardain, A., Marakalala, M. J. & Leslie, A. Tissue-resident innate immunity in the lung. *Immunology* **159**, 245–256 (2020).
- Ginhoux, F. & Jung, S. Monocytes and macrophages: Developmental pathways and tissue homeostasis. *Nat. Rev. Immunol.* **14**, 392–404 (2014).
- Chen, K. & Kolls, J. K. T cell-mediated host immune defenses in the lung. *Annu. Rev. Immunol.* **31**, 605–633 (2013).
- Allie, S. R. et al. The establishment of resident memory B cells in the lung requires local antigen encounter. *Nat. Immunol.* **20**, 97–108 (2019).
- Barker, K. A., et al. Lung-resident memory B cells protect against bacterial pneumonia. *J. Clin. Invest.* **131**, e141810 (2021).
- Kulkarni, H. S., Liszewski, M. K., Brody, S. L. & Atkinson, J. P. The complement system in the airway epithelium: An overlooked host defense mechanism and therapeutic target? *J. Allergy Clin. Immunol.* **141**, 1582–1586.e1581 (2018).
- Sharp, T. H. et al. Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc. Natl. Acad. Sci. USA.* **116**, 11900–11905 (2019).
- Mollnes, T. E. & Huber-Lang, M. Complement in sepsis—when science meets clinics. *FEBS Lett.* **594**, 2621–2632 (2020).
- Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. Complement: A key system for immune surveillance and homeostasis. *Nat. Immunol.* **11**, 785–797 (2010).
- Bosmann, M. Complement control for COVID-19. *Sci. Immunol.* **6**, eabj1014 (2021).
- Jain, U., Otle, A. R., Van Limbergen, J. & Stadnyk, A. W. The complement system in inflammatory bowel disease. *Inflamm. Bowel Dis.* **20**, 1628–1637 (2014).
- Ackerman, S. K., Friend, P. S., Hoidal, J. R. & Douglas, S. D. Production of C2 by human alveolar macrophages. *Immunology* **35**, 369–372 (1978).
- Angelidis, I. et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).
- Degn, S. E. & Thiel, S. Humoral pattern recognition and the complement system. *Scand. J. Immunol.* **78**, 181–193 (2013).
- Rizzo G., et al. Single-cell transcriptomic profiling maps monocyte/macrophage transitions after myocardial infarction in mice. *bioRxiv* 2020:2020.2004.2014.040451.
- Kemper, C., Atkinson, J. P. & Hourcade, D. E. Properdin: Emerging roles of a pattern-recognition molecule. *Annu. Rev. Immunol.* **28**, 131–155 (2010).
- Roos, C., Dressel, R., Schmidt, B., Günther, E. & Walter, L. The rat expresses two complement factor C4 proteins, but only one isotype is expressed in the liver. *J. Immunol.* **174**, 970–975 (2005).
- Bubeck, D. et al. Structure of human complement C8, a precursor to membrane attack. *J. Mol. Biol.* **405**, 325–330 (2011).
- Jackson, H. M. et al. A novel mouse model expressing human forms for complement receptors CR1 and CR2. *BMC Genet* **21**, 101 (2020).
- Luo, L. et al. Anaphylatoxins enhance recruitment of nonclassical monocytes via chemokines produced by pleural mesothelial cells in tuberculous pleural effusion. *Am. J. Respir. Cell Mol. Biol.* **60**, 454–464 (2019).
- Liszewski, M. K. et al. Intracellular complement activation sustains T cell homeostasis and mediates effector differentiation. *Immunity* **39**, 1143–1157 (2013).
- Reichhardt, M. P. & Meri, S. Intracellular complement activation—An alarm raising mechanism? *Semin. Immunol.* **38**, 54–62 (2018).
- Cole, F. S. et al. Complement biosynthesis by human bronchoalveolar macrophages. *Clin. Immunol. Immunopathol.* **27**, 153–159 (1983).
- Rothman, B. L., Merrow, M., Bamba, M., Kennedy, T. & Kreutzer, D. L. Biosynthesis of the third and fifth complement components by isolated human lung cells. *Am. Rev. Respir. Dis.* **139**, 212–220 (1989).
- Fehrenbach, H. Alveolar epithelial type II cell: Defender of the alveolus revisited. *Respir. Res.* **2**, 33–46 (2001).
- Arbore, G. & Kemper, C. A novel “complement-metabolism-inflamasome axis” as a key regulator of immune cell effector function. *Eur. J. Immunol.* **46**, 1563–1573 (2016).
- Mangogna, A. et al. Prognostic value of complement properdin in cancer. *Front Immunol.* **11**, 614980 (2020).
- Wu, X. et al. Ficolin A derived from local macrophages and neutrophils protects against lipopolysaccharide-induced acute lung injury by activating complement. *Immunol. Cell Biol.* **98**, 595–606 (2020).
- Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232 (2012).
- Lähnemann, D. et al. Eleven grand challenges in single-cell data science. *Genome Biol.* **21**, 31 (2020).
- Gibbins, S. L. et al. Three unique interstitial macrophages in the murine lung at steady state. *Am. J. Respir. Cell Mol. Biol.* **57**, 66–76 (2017).
- Fraser, D. A., Laust, A. K., Nelson, E. L. & Tenner, A. J. C1q differentially modulates phagocytosis and cytokine responses during ingestion of apoptotic cells by human monocytes, macrophages, and dendritic cells. *J. Immunol.* **183**, 6175–6185 (2009).
- Lener, M., Vinci, G., Duponchel, C., Meo, T. & Tosi, M. Molecular cloning, gene structure and expression profile of mouse C1 inhibitor. *Eur. J. Biochem.* **254**, 117–122 (1998).

44. McQuaid, S. & Cosby, S. L. An immunohistochemical study of the distribution of the measles virus receptors, CD46 and SLAM, in normal human tissues and subacute sclerosing panencephalitis. *Lab Invest.* **82**, 403–409 (2002).
45. Sato, T. et al. Pro-carboxypeptidase R is an acute phase protein in the mouse, whereas carboxypeptidase N is not. *J. Immunol.* **165**, 1053–1058 (2000).
46. Vandendriessche, S., Cambier, S., Proost, P. & Marques, P. E. Complement receptors and their role in leukocyte recruitment and phagocytosis. *Front. Cell Dev. Biol.* **9**, 624025 (2021).
47. Drouin, S. M. et al. Expression of the complement anaphylatoxin C3a and C5a receptors on bronchial epithelial and smooth muscle cells in models of sepsis and asthma. *J. Immunol.* **166**, 2025–2032 (2001).
48. Haviland, D. L. et al. Cellular expression of the C5a anaphylatoxin receptor (C5aR): demonstration of C5aR on nonmyeloid cells of the liver and lung. *J. Immunol.* **154**, 1861–1869 (1995).
49. Karsten, C. M. et al. Monitoring C5aR2 expression using a floxed tdTomato-C5aR2 knock-in mouse. *J. Immunol.* **199**, 3234–3248 (2017).
50. Nagre, N. et al. Inhibition of macrophage complement receptor CRlg by TRIM72 polarizes innate immunity of the lung. *Am. J. Respir. Cell Mol. Biol.* **58**, 756–766 (2018).
51. Ghebrehiwet, B. & Peerschke, E. I. cC1q-R (calreticulin) and gC1q-R/p33: Ubiquitously expressed multi-ligand binding cellular proteins involved in inflammation and infection. *Mol. Immunol.* **41**, 173–183 (2004).
52. Nepomuceno, R. R. & Tenner, A. J. C1qRP, the C1q receptor that enhances phagocytosis, is detected specifically in human cells of myeloid lineage, endothelial cells, and platelets. *J. Immunol.* **160**, 1929–1935 (1998).
53. Dean, Y. D., McGreal, E. P. & Gasque, P. Endothelial cells, megakaryoblasts, platelets and alveolar epithelial cells express abundant levels of the mouse AA4 antigen, a C-type lectin-like receptor involved in homing activities and innate immune host defense. *Eur. J. Immunol.* **31**, 1370–1381 (2001).
54. Bosmann, M. & Ward, P. A. Role of C3, C5 and anaphylatoxin receptors in acute lung injury and in sepsis. *Adv. Exp. Med Biol.* **946**, 147–159 (2012).
55. Pandya, P. H. & Wilkes, D. S. Complement system in lung disease. *Am. J. Respir. Cell Mol. Biol.* **51**, 467–473 (2014).
56. Andre, G. O. et al. Role of Streptococcus pneumoniae proteins in evasion of complement-mediated immunity. *Front Microbiol* **8**, 224 (2017).
57. Afzali, B., Noris, M., Lambrecht, B. N. & Kemper, C. The state of complement in COVID-19. *Nat. Rev. Immunol.* **22**, 77–84 (2022).
58. Travaglini K. J. et al. Processed data from cellxgene for: A molecular cell atlas of the human lung from single-cell RNA sequencing. Available from: <https://cellxgene.cziscience.com/collections/5d445965-6f1a-4b68-ba3a-b8f765155d3a>. (2020). Accessed on 23 March 2022.
59. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
60. RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.
61. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
62. Hong, R. et al. Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *Nat. Commun.* **13**, 1688 (2022).
63. Hoffman P. SeuratDisk: Interfaces for HDF5-Based Single Cell File Formats (2021).
64. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

## ACKNOWLEDGEMENTS

Animated figures were created using some elements from servier medical art by servier under a creative commons attribution 3.0 unported license. We thank Dr. Joseph P. Mizgerd, Director, Pulmonary Center, Boston University School of Medicine, for reading and critical review of the manuscript. Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Numbers T32HL007035 and R01HL139641. C.R. was awarded a Fellowship of the Gutenberg Research College at the Johannes Gutenberg-University Mainz. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. All requests for materials and correspondence should be addressed to M.B.

## AUTHOR CONTRIBUTIONS

N.C. and A.J. analyzed data and wrote the manuscript. C.R. and J.D.C. contributed to the data analysis, provided comments and guidance. M.B. conceived and supervised the study, interpreted data, provided funding, and edited the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41385-022-00534-7>.

**Correspondence** and requests for materials should be addressed to Markus Bosmann.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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