

A mesothelial differentiation gateway drives fibrosis

Corresponding Author: Dr Wei Han

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Review had previously been provided.

The authors have address the majority of reviewer's comments adequately although a couple of comments needs to be addressed.

1. Please add more details regarding the hydroxyproline assay on lines 687 and 688 in the methods section. It is unclear what tissue exactly was assessed. Was it from total lung tissue, specific lobes and how much tissue was assessed? How was the tissue processed?

2. Please remove the word comprehensive on line 684 when referring to the assessment of three sections per lung as 'comprehensive'.

Please check references. Reference 5, 84, 88, 90, 91, 94, 96, 97, 101 don't make sense.

(Remarks on code availability)

Not sure what this means

Reviewer #2

(Remarks to the Author)

In this revised manuscript, the authors have made some text revisions and added additional quantification of collagen content. The most substantive concern raised initially was partially addressed but at this point the provided data still do not provide conclusive evidence that this "immune modulatory" subtype is not an artifact, and some provided data are not consistent with the authors interpretation. The addition of immunofluorescence data adds strength that this cell state occurs reproducibly, but there are still several issues that warrant addressing to have confidence in the overall findings.

Major comments:

There are still several points with regard to the scRNA-seq analysis that were not directly addressed in the authors response.

1) Quality - The figure presented in the response appears to show quality metrics for all of the cells analyzed. The overall quality of the dataset (in terms of total counts and genes/cell) is on the low side (acknowledging some of this is platform related) – it appears the median number of counts is barely >500, and median number of genes is around 400. There appears to be a bimodality to the data, and a substantial number of cells appear to have only 200-300 counts. In order to have confidence these are not simply lower-quality cells, the authors should perform a sensitivity analysis excluding all cells with <500 total counts.

2) The authors need to present the percent ribosomal genes per cell-type – in general, clusters with marked enrichment of ribosomal genes are quality-related, not biological variation. If a cluster is a substantial outlier compared to the others in terms of ribosomal content, that would imply a high likelihood of a spurious/artifactual finding.

3) Ambient RNA – it does not appear there was any method used to assess or remove ambient RNA – given that C3 and

Saa3 are expressed by multiple cell-types (macrophages, inflammatory fibroblasts), ambient RNA-contamination of relatively lower quality cells could yield this result. The authors need to provide at minimum assessment of ambient RNA.

4) The velocity pseudotime analysis in Supplementary Figure 2 appears to depict the "immune modulatory" cells at pseudotime 0 (implying these are the start of a trajectory). How would this be possible if the cells are normally absent and only appear 11 days after injury? They would have to arise from some other pre-existing mesothelial cell. This seems inconsistent with the PAGA trajectory in Figure 2 and the overall concept the authors are proposing.

5) The tissue quantification (Figure S3) suggests Saa3+ cells are already present by day 5 although more abundant at day 10. It seems much more likely that if this is a real cell state, it occurs during cell-phenotype shifts (analogous to that seen in lung fibroblasts after bleomycin) which in reality occur asynchronously over a period of several days – the text and figure 2 should be modified to reflect the likely more extended time window in which this cell phenotype is observed.

6) The text revisions better reflect interpretation of trajectory analyses as "predicted." The AAV-based overexpression studies demonstrate functional consequences of these manipulations (Figure 3b, S4), although these studies were not specifically testing the proposed sequence of differentiation (if anything, these experiments suggest that any mesothelial cell can adopt one of these states without requiring a sequence of steps). To confirm the proposed trajectory, it would be necessary to specifically label cells in each state and fate-map them. This seems beyond the scope of this manuscript, but the text still needs to be softened to better reflect that these are predicted "lineage" relationships.

Minor comments:

1) The authors have clarified how the transduction was performed ameliorating concerns about off-target issues in PCLS, and provide supplemental data supporting the specificity of targeting to the mesothelium. These points are adequately addressed.

(Remarks on code availability)

The data and code are not actually available although the authors say they will make them available. It isn't possible at this point to know how reproducible these findings are.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This is a well performed study that has highlighted the importance of the mesothelial cell in disease pathogenesis, in particular lung fibrosis. The results highlight the need for further studies in mesothelial cell biology and the possibility of targeting this cell for therapeutic interventions. The authors have adequately addressed the reviewers concerns.

Reviewer #2

(Remarks to the Author)

The authors have performed additional analyses as requested. These partially address my prior concerns, though a few concerns still remain.

The most discriminating genes for the both "Fibrogenic" and "Immune Modulatory" clusters are highly enriched for ribosomal genes, and the total ribosomal expression is higher in these clusters than others - It is still difficult to have total confidence that this is not a data-quality artifact. I do appreciate the considerable efforts the authors have made to address this issue, and this is really a reflection of a dataset that is just overall not of the highest quality. At this point, I think text needs to be added to the discussion to acknowledge the quality-related limitations of the dataset.

Regarding prior comment 5 - the schematic summarizing the cell phenotypes in Figure 2 still needs revising - in particular, the multiple peaks for "fibrogenic" and multiple nadirs for the "metabolically primed" states just don't make physiologic sense - this seems likely an artifact of small numbers of cells - it would seem this should be smoothed.

Regarding the trajectory analysis - the velocity stream plots still seem to just show noise - would recommend removing these as they are at best confusing.

Version 3:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have made revisions that address my previous comments.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Review had previously been provided.

The authors have address the majority of reviewer's comments adequately although a couple of comments needs to be addressed.

1. Please add more details regarding the hydroxyproline assay on lines 687 and 688 in the methods section. It is unclear what tissue exactly was assessed. Was it from total lung tissue, specific lobes and how much tissue was assessed? How was the tissue processed?

Response: thank you very much for your suggestions. We have added more detailed information in the Method part: "To quantify collagen content, a hydroxyproline assay was performed using a commercially available kit (Hydroxyprolin Assay-Kit, ab222941, abcam). Hydroxyproline content was assessed from whole lung tissue. Every 10 mg of lung tissue was homogenized in 100 μ L distilled water using an ultrasonic probe homogenizer. A 100 μ L aliquot of the homogenate was transferred to a pressure-tight polypropylene vial, followed by the addition of 100 μ L of 10 N NaOH. Samples were hydrolyzed at 120°C for 1 hour, cooled on ice, and neutralized with 100 μ L of 10 N HCl. After centrifugation at 10,000 $\times$ g for 5 minutes, the supernatant was collected and used for hydroxyproline quantification. The hydrolyzed samples were evaporated at 65°C and subjected to oxidation and colorimetric detection according to the manufacturer's protocol. Absorbance was measured at 560 nm using a microplate reader, and hydroxyproline concentration was calculated based on a standard curve."

2. Please remove the word comprehensive on line 684 when referring to the assessment of three sections per lung as 'comprehensive'.

Response: thank you very much for your input. We have deleted "comprehensive" as you suggested: "The scores from the selected regions were averaged for each lobe to obtain a fibrosis assessment."

Please check references. Reference 5, 84, 88, 90, 91, 94, 96, 97, 101 don't make sense.

Response: thank you very much for pointing it out. We have thoroughly checked all the references to make sure all the references are up-to-date.

However, we would like to make some clarifications:

5. Se, M., Cm, P., S, P. & Se, H. Mesothelial cells and peritoneal homeostasis. Fertil. Steril. 106, (2016).

In this paper, it was written: "They provide a nonadhesive protective surface through the production of phospholipids and a well developed surface glycocalyx (fuzzy coat on the external surface of their plasma membrane consisting of glycolipids and glycoproteins) to allow movement of organs within the serosal cavities." This is consistent with the sentences in our paper: "Mesothelial cells secrete a carbohydrate-rich glycocalyx across the serosal membranes³ that coats organ surfaces with a

protective barrier^{5,6}." That's why we cited it. We have changed the format to: **Mutsaers, S. E., Prêle, C. M.-A., Pengelly, S. & Herrick, S. E. Mesothelial cells and peritoneal homeostasis. Fertil. Steril. 106, 1018–1024 (2016).**

84. C, M. et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. Eur. Respir. J. 54, (2019).

This paper is cited as we includes the IPF lung datasets into our analysis. We have changed the format to: **Morse, C. et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. Eur. Respir. J. 54, 1802441 (2019).**

88. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure - PubMed. <https://pubmed.ncbi.nlm.nih.gov/35959412/>.

While the study primarily focuses on heart failure, we specifically utilized the datasets of non-disease samples in our analysis (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183852>). We have changed the format to: **Koenig, A. L. et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. Nat. Cardiovasc. Res. 1, 263–280 (2022).**

90. The C1q-ApoE complex: A new hallmark pathology of viral hepatitis and nonalcoholic fatty liver disease - PubMed.

<https://pubmed.ncbi.nlm.nih.gov/36304458/>.

We utilized healthy liver tissue samples from this database (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE192742>). We have changed the format to: **Habenicht, L. K. L. et al. The C1q-ApoE complex: A new hallmark pathology of viral hepatitis and nonalcoholic fatty liver disease. Front. Immunol. 13, 970938 (2022).**

91. Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution - PubMed.

<https://pubmed.ncbi.nlm.nih.gov/33691112/>.

We replaced this reference with the original article of the database they used (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116222>) : **Parikh, K. et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature 567, 49–55 (2019).**

94. Inhibition of hyperglycolysis in mesothelial cells prevents peritoneal fibrosis - PubMed. <https://pubmed.ncbi.nlm.nih.gov/31167927/>.

We used the normal peritoneum datasets from this paper in our study. (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130888>). We have now changed the format to: **Si, M. et al. Inhibition of hyperglycolysis in mesothelial cells prevents peritoneal fibrosis. Sci. Transl. Med. 11, eaav5341 (2019).**

96. N, K. et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat. Commun. 11, (2020).

We used lung adenocarcinoma datasets from this reference. (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM3755693>). We changed the format to: **Kim, N. et al. Single-cell RNA sequencing demonstrates the**

molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat. Commun. 11, 2285 (2020).

97. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis - PubMed.
<https://pubmed.ncbi.nlm.nih.gov/30554520/>.

We used pulmonary fibrosis datasets from this study, and we have changed the format as follows: Reyfman, P. A. et al. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am. J. Respir. Crit. Care Med. 199, 1517–1536 (2019).

101. Lotfollahi, M. et al. Mapping single-cell data to reference atlases by transfer learning. Nat. Biotechnol. 40, 121–130 (2022).

This paper introduced a deep learning strategy for mapping query datasets on top of a reference called single-cell architectural surgery (scArches). scArches uses transfer learning and parameter optimization to enable efficient, decentralized, iterative reference building and contextualization of new datasets with existing references without sharing raw data. We have utilized scArches in our paper therefore we cited it.

Reviewer #1 (Remarks on code availability):

Not sure what this means

Response: We have provided all the data and code in this link:
<https://zenodo.org/records/15386604>
Doi: 10.5281/zenodo.15386603

Reviewer #2 (Remarks to the Author):

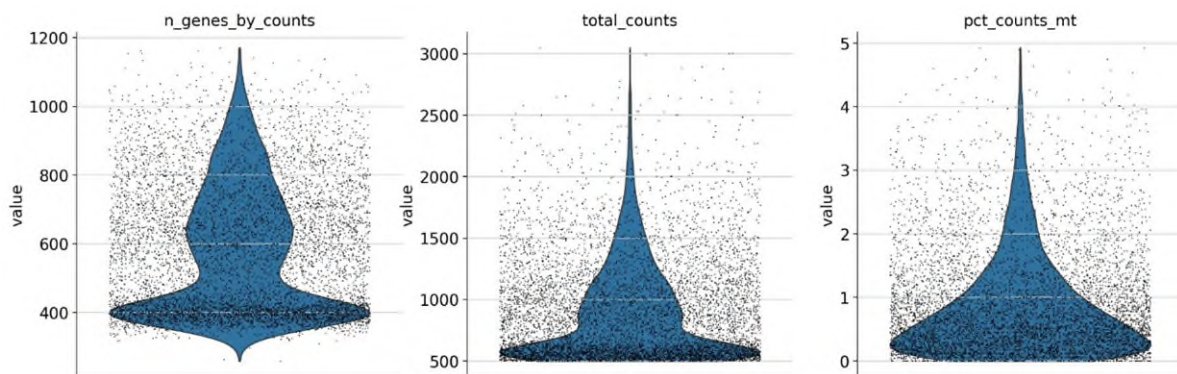
In this revised manuscript, the authors have made some text revisions and added additional quantification of collagen content. The most substantive concern raised initially was partially addressed but at this point the provided data still do not provide conclusive evidence that this “immune modulatory” subtype is not an artifact, and some provided data are not consistent with the authors interpretation. The addition of immunofluorescence data adds strength that this cell state occurs reproducibly, but there are still several issues that warrant addressing to have confidence in the overall findings.

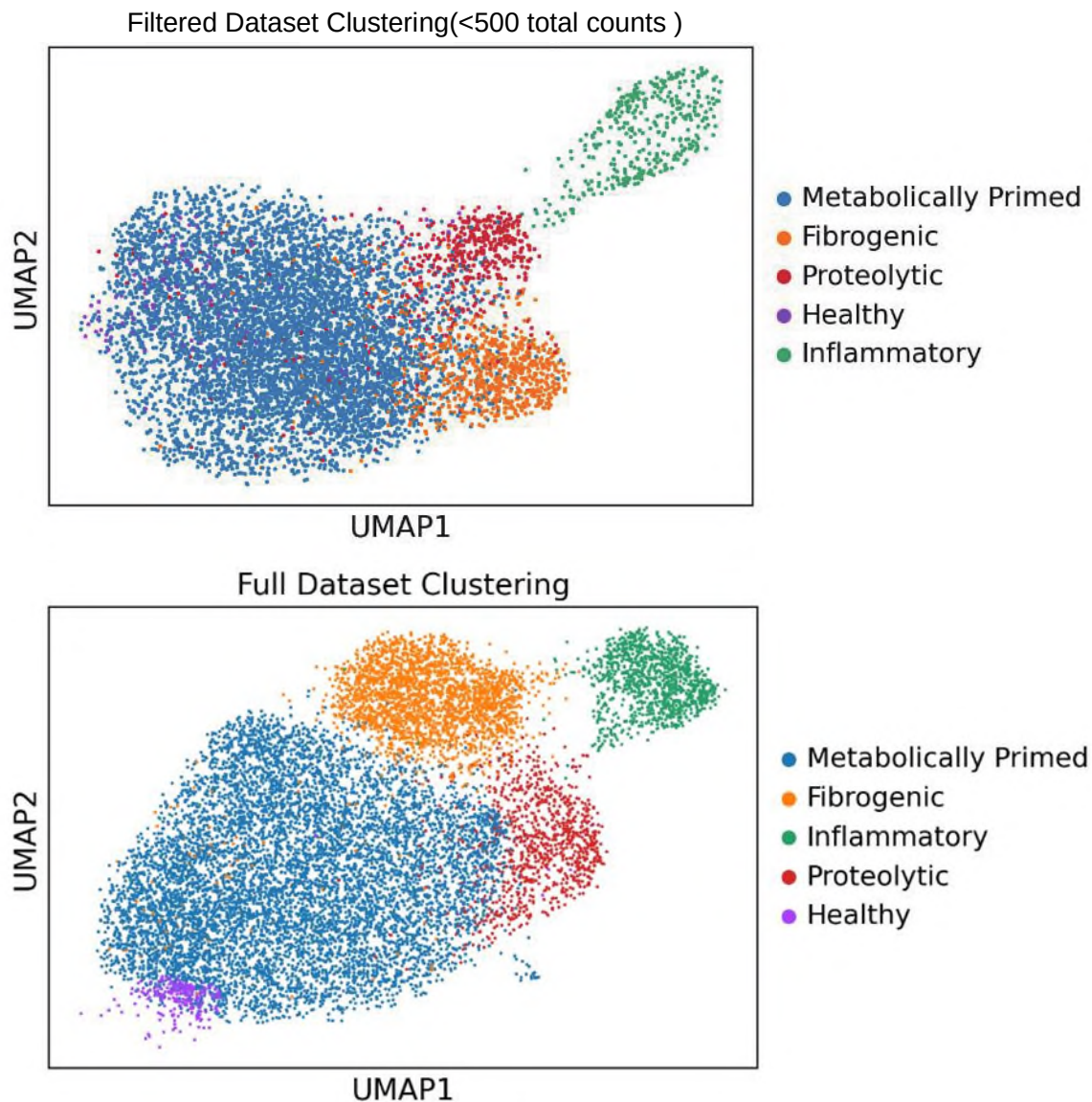
Major comments:

There are still several points with regard to the scRNA-seq analysis that were not directly addressed in the authors response.

1) Quality - The figure presented in the response appears to show quality metrics for all of the cells analyzed. The overall quality of the dataset (in terms of total counts and genes/cell) is on the low side (acknowledging some of this is platform related) – it appears the median number of counts is barely >500, and median number of genes is around 400. There appears to be a bimodality to the data, and a substantially number of cells appear to have only 200-300 counts. In order to have confidence these are not simply lower-quality cells, the authors should perform a sensitivity analysis excluding all cells with <500 total counts.

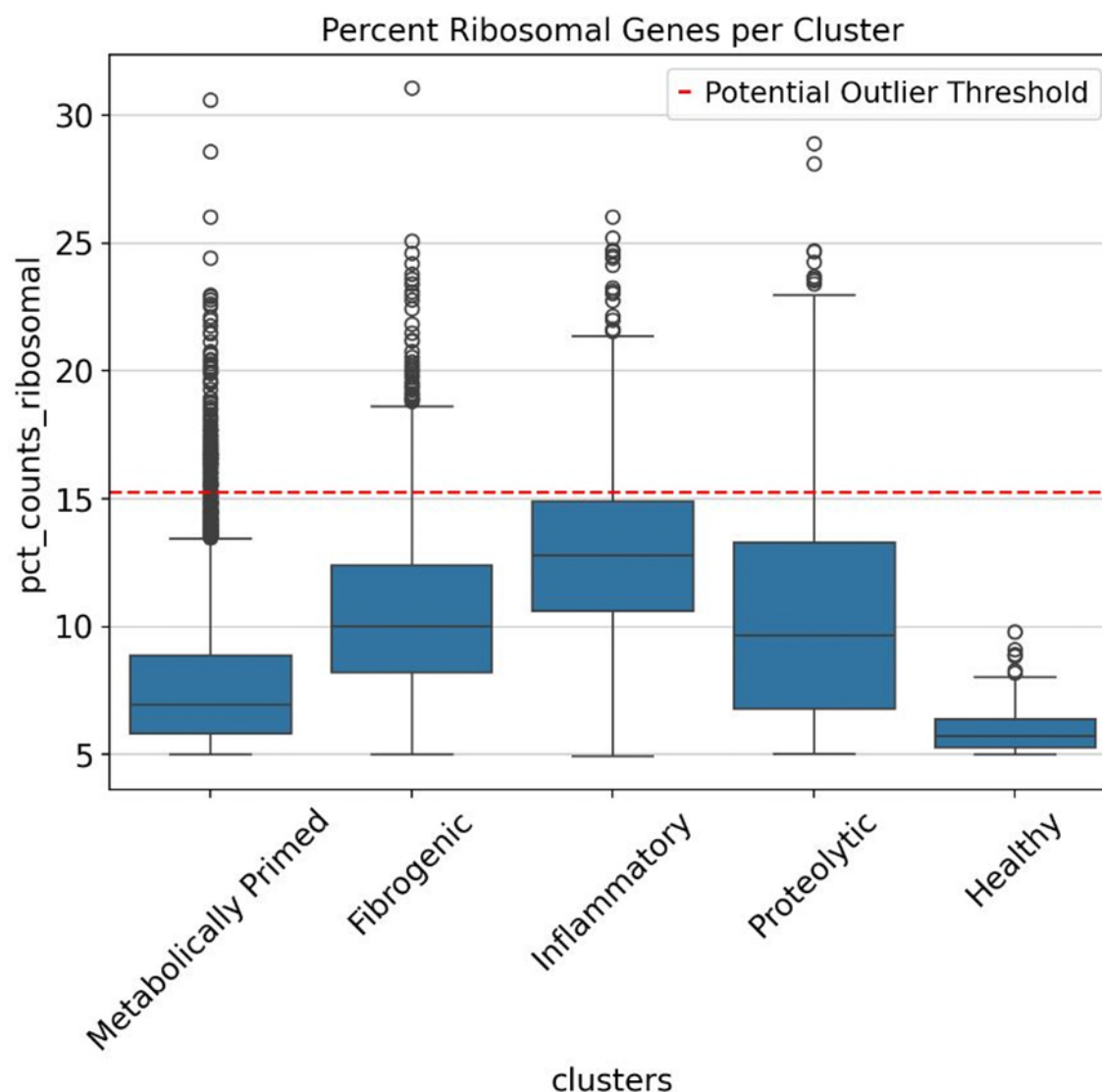
Response: We appreciate the reviewer’s concern regarding the overall quality of the dataset. To ensure data quality, we performed rigorous quality control and confirmed that all included cells had total counts >500. In the figure provided below, the full dataset clustering includes 11,952 cells, and after filtering, 7,615 cells remain. This demonstrates that our QC criteria effectively excluded low-quality cells, addressing the reviewer’s concern.





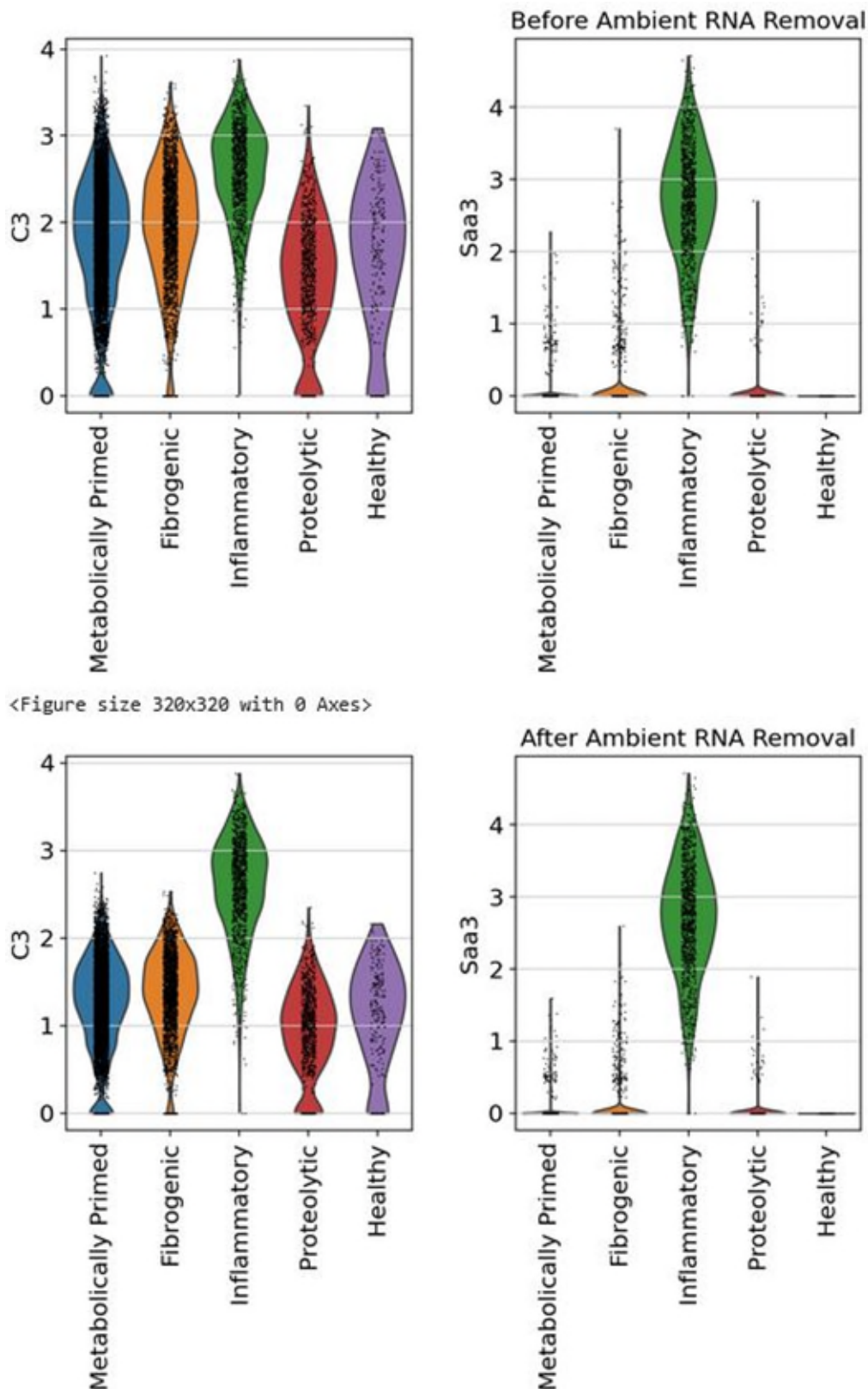
2) The authors need to present the percent ribosomal genes per cell-type – in general, clusters with marked enrichment of ribosomal genes are quality-related, not biological variation. If is cluster is a substantial outlier compared to the others in terms of ribosomal content, that would imply a high likelihood of a spurious/artifactual finding.

Response: We appreciate the reviewer's suggestion. We have performed the requested analysis, and the results are shown in the figure below. As illustrated, the percentage of ribosomal genes per cell type has been assessed, and no cluster exhibits a substantial outlier in terms of ribosomal content. This indicates that our clustering results are not driven by technical artifacts but rather reflect biological variation.



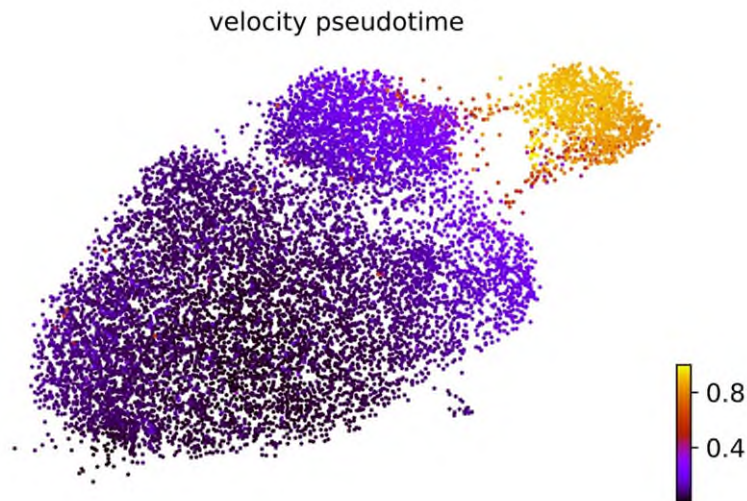
3) Ambient RNA – it does not appear there was any method used to assess or remove ambient RNA – given that C3 and Saa3 are expressed by multiple cell-types (macrophages, inflammatory fibroblasts), ambient RNA-contamination of relatively lower quality cells could yield this result. The authors need to provide at minimum assessment of ambient RNA.

Response: We appreciate the reviewer's comment regarding ambient RNA contamination. To address this concern, we performed ambient RNA analysis using SoupX, and this method has been included in the Methods section. The result of this analysis is provided in the figure below.



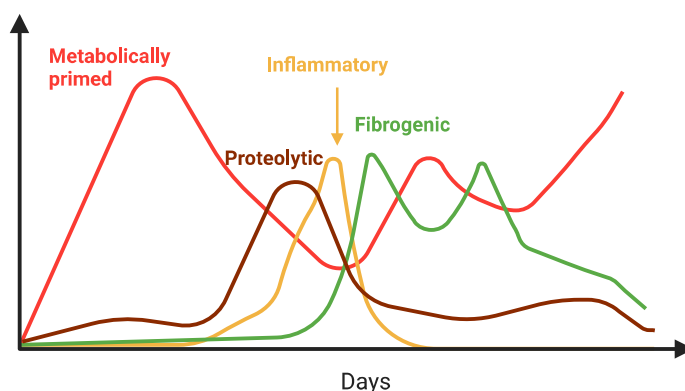
4) The velocity pseudotime analysis in Supplementary Figure 2 appears to depict the “immune modulatory” cells at pseudotime 0 (implying these are the start of a trajectory). How would this be possible if the cells are normally absent and only appear 11 days after injury? They would have to arise from some other pre-existing mesothelial cell. This seems inconsistent with the PAGA trajectory in Figure 2 and the overall concept the authors are proposing.

Response: Thank you very much for pointing it out. We acknowledge that there was an error in the initial analysis, where the pseudotime trajectory was inadvertently inverted. We have now corrected this, and the revised figure is provided below, and Supplementary Figure 2 has been updated.



5) The tissue quantification (Figure S3) suggests Saa3+ cells are already present by day 5 although more abundant at day 10. It seems much more likely that if this is a real cell state, it occurs during cell-phenotype shifts (analogous to that seen in lung fibroblasts after bleomycin) which in reality occur asynchronously over a period of several days – the text and figure 2 should be modified to reflect the likely more extended time window in which this cell phenotype is observed.

Response: We appreciate the reviewer for pointing out this issue. We have updated Figure 2e (provided below) to depict a more gradual increase in Immune Modulator cells starting from day 5, rather than a sharp peak. Correspondingly, we have also revised the relevant text in the manuscript: This immune phenotype emerges around day 5 and gradually increases, reaching its peak at day 10. Then, these immune modulatory cells transition progressively into fibrogenic cells, coinciding with the onset of lung fibrosis at day 11.



6) The text revisions better reflect interpretation of trajectory analyses as “predicted.”

The AAV-based overexpression studies demonstrate functional consequences of these manipulations (Figure 3b, S4), although these studies were not specifically testing the proposed sequence of differentiation (if anything, these experiments suggest that any mesothelial cell can adopt one of these states without requiring a sequence of steps). To confirm the proposed trajectory, it would be necessary to specifically label cells in each state and fate-map them. This seems beyond the scope of this manuscript, but the text still needs to be softened to better reflect that these are predicted “lineage” relationships.

Response: We appreciate the reviewer’s feedback. We have revised the text to clarify that the trajectory analyses and proposed lineage relationships are predicted rather than definitively proven: “We next employed RNA velocity analysis, based on unspliced and mature mRNAs, to **predict** the trajectory and differentiation process between the four cell states in the Bleomycin model (Supplementary fig 2g,h). The earliest disease-specific mesothelial state **was predicted to be** metabolically active cells expressing Ifi2712a, and Crip1. Metabolically active cells **were predicted to** differentiate into proteolytic-active cells overexpressing metalloprotease and endopeptidases, including Matrix Metalloproteinase 8 and 14 (Mmp8, Mmp14) Cathepsin B (Ctsb), Dcn and Plac8. Proteolytic cells transiently **were predicted to** differentiate into immune-active cells expressing chemoattractant Saa3. Finally, immune modulatory cells **were predicted to** differentiate into fibrogenically active cells, overexpressing Mgp, Sparc and collagens 1 and 3 (fig 2d and Supplementary fig2i). This unprecedented differentiation pathway is likely key to the fibrotic process.”The **predictive** differentiation steps were mirrored by sequential enrichment of the mesothelial cell states over the time-course.”

Minor comments:

1) The authors have clarified how the transduction was performed ameliorating concerns about off-target issues in PCLS, and provide supplemental data supporting the specificity of targeting to the mesothelium. These points are adequately addressed.

Reviewer #2 (Remarks on code availability):

The data and code are not actually available although the authors say they will make them available. It isn't possible at this point to know how reproducible these findings are.

Response: We appreciate your feedback! Here is the link which included all the data and code: <https://zenodo.org/records/15386604>
Doi: [10.5281/zenodo.15386603](https://doi.org/10.5281/zenodo.15386603)

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This is a well performed study that has highlighted the importance of the mesothelial cell in disease pathogenesis, in particular lung fibrosis. The results highlight the need for further studies in mesothelial cell biology and the possibility of targeting this cell for therapeutic interventions. The authors have adequately addressed the reviewers concerns.

Response: Thank you very much for acknowledging our work. We sincerely appreciate your professional feedback and the time you have dedicated to reviewing our manuscript.

Reviewer #2 (Remarks to the Author):

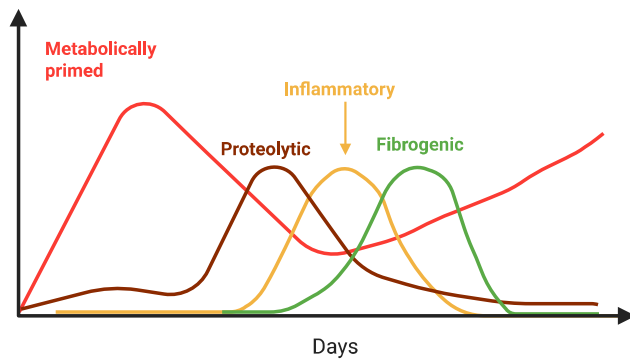
The authors have performed additional analyses as requested. These partially address my prior concerns, though a few concerns still remain.

1. The most discriminating genes for the both "Fibrogenic" and "Immune Modulatory" clusters are highly enriched for ribosomal genes, and the total ribosomal expression is higher in these clusters than others - It is still difficult to have total confidence that this is not a data-quality artifact. I do appreciate the considerable efforts the authors have made to address this issue, and this is really a reflection of a dataset that is just overall not of the highest quality. At this point, I think text needs to be added to the discussion to acknowledge the quality-related limitations of the dataset.

Response: Thank you very much for acknowledging our revision. We have added the limitations of the dataset in our discussion part: While our single-cell analysis supports the sequential relationships of these mesothelial states, we acknowledge that some of the observed transcriptional patterns may be influenced by technical limitations associated with the method used. Despite applying rigorous quality control, we cannot fully exclude the possibility of residual technical artifacts. Future studies will be important to further validate these findings.

2. Regarding prior comment 5 - the schematic summarizing the cell phenotypes in Figure 2 still needs revising - in particular, the multiple peaks for "fibrogenic" and multiple nadirs for the "metabolically primed" states just don't make physiologic sense - this seems likely an artifact of small numbers of cells - it would seem this should be smoothed.

Response: Thank you for your feedback. We have adjusted the Figure 2 as suggested (shown below).



3. Regarding the trajectory analysis - the velocity stream plots still seem to just show noise - would recommend removing these as they are at best confusing.
Response: Thank you for the suggestion. We have removed the velocity plots as requested in supplementary figure 2 from the revised version.

