

Genes, other than Muc5b, are associated with bleomycin-induced lung injury

Corresponding Author: Dr David Schwartz

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

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This new version of the manuscript is better aligned with the scope of the study, I believe this data will be a great resource to mine for the field. While some comments were not fully addressed they were more related to scientific curiosity and maybe not aligned with the current scope of this study. The key comments supporting this version were addressed by the authors and the present work should be published.

Reviewer #2

(Remarks to the Author)

The authors perform a detailed QTL analysis of radiomic features of lung fibrosis and injury after bleomycin in DO mice who all carry the MUC5B orthologous rs35705950 locus. The authors find and focus on a chromosome 7 locus associated with radiomic features after bleomycin injury that is near, but independent of the MUC5B locus and prioritize QTLs in the region that also associate with gene expression. The manuscript represents a significant body of work and the authors should be commended for the scope of the analyses. The authors have revised and addressed most major comments in my previous review of the paper. Some comments remain.

1. The results related to OAT expression in human lungs remains unsupported by the objective data presented in Figure 6 as there is no image analysis/quantification shown. As such, the following statements should either come with corresponding data analysis, or changed to emphasize its descriptive, rather than analytic, results: "Our results demonstrate that in controls OAT protein was detected in airway epithelia, with slightly higher expression in Keratin 5+ basal cells and MUC5B- cells in a rs35705950 genotype-dependent manner (Fig. 6d, Supplementary Fig. 14). Importantly, we observed an increase in OAT expression in airway epithelia in tissues from IPF donors. In addition, we observed substantial increase of OAT expression in Keratin 5+ basal cells and MUC5B- cells in IPF airways, honeycomb cysts, and cells overlaying fibroblastic foci in a genotype dependent manner (Fig. 6d, Supplementary Fig. 14)."

2. The results section and Figure 6e describe modifiers of MUC5B. The language also describes "penetrance" which is not supported by the manuscript. It is fine to suggest a framework based on preliminary data, but neither the mouse nor human data shown confirm existence of MUC5B modifiers. If the human lung histopathology data shows supportive genotype-dependent data, then would be fine to include. Alternatively, the authors could temper the results and remove mention of "penetrance" from Figure 6E and leave the human lung histology results as more descriptive.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have fully addressed my remaining comments. The manuscript is well-done and will be a valuable resource for

the community.

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/>

Dear Reviewers,

We sincerely thank you for taking the time to carefully read our manuscript and provide such valuable comments. We have addressed each point raised and made corresponding revisions. Below, we provide a point-by-point response to your comments. All changes in the manuscript text file have been highlighted in yellow to facilitate your review.

Reviewer #1 (Remarks to the Author):

This new version of the manuscript is better aligned with the scope of the study, I believe this data will be a great resource to mine for the field. While some comments were not fully addressed they were more related to scientific curiosity and maybe not aligned with the current scope of this study. The key comments supporting this version were addressed by the authors and the present work should be published.

We thank the reviewer for their positive feedback and constructive comments, which helped improve the manuscript.

Reviewer #2 (Remarks to the Author):

The authors perform a detailed QTL analysis of radiomic features of lung fibrosis and injury after bleomycin in DO mice who all carry the MUC5B orthologous rs35705950 locus. The authors find and focus on a chromosome 7 locus associated with radiomic features after bleomycin injury that is near, but independent of the MUC5B locus and prioritize QTLs in the region that also associate with gene expression. The manuscript represents a significant body of work and the authors should be commended for the scope of the analyses. The authors have revised and addressed most major comments in my previous review of the paper. Some comments remain.

1. The results related to OAT expression in human lungs remains unsupported by the objective data presented in Figure 6 as there is no image analysis/quantification shown. As such, the following statements should either come with corresponding data analysis, or changed to emphasize its descriptive, rather than analytic, results: "Our results demonstrate that in controls OAT protein was detected in airway epithelia, with slightly higher expression in Keratin 5+ basal cells and MUC5B- cells in a rs35705950 genotype-dependent manner (Fig. 6d, Supplementary Fig. 14). Importantly, we observed an increase in OAT expression in airway epithelia in tissues from IPF donors. In addition, we observed substantial increase of OAT expression in Keratin 5+ basal cells and MUC5B- cells in IPF airways, honeycomb cysts, and cells overlaying fibroblastic foci in a genotype dependent manner (Fig. 6d, Supplementary Fig. 14)."

We fully agree that the IF results provide descriptive observations intended to support and spatially contextualize the RNA-seq findings, rather than serve as quantitative measurements. In response, we have revised the Results section to clearly emphasize that these findings are qualitative rather than quantitative. The updated description now reads as follows: "Our qualitative observations demonstrate that in controls OAT protein was detected in airway epithelia, with what appears to be higher expression in Keratin 5+ basal cells and MUC5B- cells, which appears to be dependent on the MUC5B promoter risk variant (rs35705950) (Fig. 6d, Supplementary Fig. 14). Importantly, we observed a qualitative increase in OAT expression in airway epithelia in tissues from IPF donors. In addition, we observed higher expression of OAT in Keratin 5+ basal cells and MUC5B- cells in IPF airways, honeycomb cysts, and cells overlaying fibroblastic foci in a MUC5B genotype dependent manner (Fig. 6d, Supplementary Fig. 14)."

2. The results section and Figure 6e describe modifiers of MUC5B. The language also describes "penetrance" which is not supported by the manuscript. It is fine to suggest a framework based on preliminary data, but neither the mouse nor human data shown confirm existence of MUC5B modifiers. If the human lung histopathology data shows supportive genotype-dependent data, then would be fine to include. Alternatively, the authors could temper the results and remove mention of "penetrance" from Figure 6E and leave the human lung histology results as more descriptive.

We appreciate the reviewer's thoughtful feedback. We agree that our current data do not provide sufficient evidence to conclude that the identified genes modify MUC5B penetrance. In response, we have carefully revised the manuscript to remove all wording implying "MUC5B modifiers". We have also updated Figure 6e and its legend, replacing "disease penetrance and severity" with the more appropriate term "**disease outcome**".

3 Follow-up on multiple testing burden

We thank the reviewer for raising the issue of "multiple testing burden" in the previous round. In this revised manuscript, we have further addressed this concern. In the QTL summary table (Table 1), we have added the following clarifications: **p-values** represent genome-wide significance after family-wise error rate correction across all tested loci, and **q-values** represent significance controlled both across the genome and across the multiple outcome radiomics features, using family-wise error rate and false discovery rate for correction. Compared with the LOD score threshold used in the previous version, these values are more intuitive and easier to interpret. We have also expanded the Methods section "QTL significance assessment" to provide detailed descriptions of how both p-values and q-values were calculated.

Required revisions for consideration at Communications Biology:

*Reframe as a resource paper emphasising the DO mouse dataset, radiomics-QTL integration, and cross-species analyses.

We have reframed the manuscript to include both the DO mouse dataset and downstream analyses that include radiomics–QTL integration and cross-species comparisons.

*Temper claims about MUC5B; clarify that findings relate to radiomics-associated loci rather than direct fibrosis modifiers.

This was a very helpful suggestion, and we have substantially modified the text to address this concern. Specifically, we have revised the title of the paper and tempered the discussion regarding the relationship between *MUC5B* and candidate genes. Additionally, we have revised the terminology from “fibrosis-QTL” to “radiomics-QTL” throughout the manuscript to clarify that our findings pertain to radiomics-associated loci rather than direct fibrosis modifiers.

*Improve narrative flow and figure legends; add histology for strain differences and clarify QTL significance thresholds.

We have improved the narrative flow and revised figure legends to emphasize the conclusions rather than technical details. Table 1 now includes the significant LOD thresholds for each radiomics feature, as well as the thresholds after adjusting for multiple testing across all 90 radiomics features. For histology of strain differences, these data (H&E staining and SHG imaging after 3-week bleomycin treatment) were previously published (DOI:10.1152/ajplung.00615.2020); to avoid redundancy, we have not repeated the figures here, but we have cited the publication and highlighted the relevant key information in the text.

*Provide stronger rationale for candidate gene prioritisation and include key statistical details.

We have added a detailed rationale for candidate gene prioritization in the manuscript, along with the relevant statistical details.

Not required:

*Mechanistic validation (e.g., OAT targeting) and extensive single-cell RNA-seq integration are optional.

Reviewer #1 (Remarks to the Author):

Overall comments for authors:

The authors utilize DO mice to understand genetic modifiers of MUC5B protein levels and fibrosis following bleomycin injury. All DO founder mice have the risk allele orthologous to the human MUC5B risk SNP. The approach is highly innovative and represents an admirable amount of work. However, the results should be interpreted with caution. There is no QTL for MUC5B protein levels that reach genome-wide significance threshold. There is no QTL for post-injury fibrosis measured by hydroxyproline or SHG. Radiomic features are not shown to be equivalent to fibrosis. The genome-wide significant QTLs found only for radiomic features calls into question if the proposed modifier is a true fibrosis QTL. Additionally, since all DO founder mice carry the MUC5B risk allele, it is not appropriate to comment on effect modification as no analysis of non-MUC5B mice were performed. While the group has previously shown that transgenic overexpression of MUC5B in epithelial cells leads to fibrosis in mice, it has not been shown that mice carrying the syntenic MUC5B risk SNP have an increased susceptibility to fibrosis. As it stands, the findings seem to be better characterized as the discovery of novel

post-injury radiomics-associated genetic loci that may be otherwise unrelated to MUC5B, contrary to the title and discussion. These post-injury radiomic-associated genetic loci are highly relevant to pulmonary and radiologic imaging, and if robust to orthogonal analyses, are likely relevant to pathways underlying mouse lung injury and repair.

We thank the reviewer for these insightful comments and have reoriented the paper around the discovery of post-injury radiomic-associated genetic loci that are unrelated to MUC5B. We have modified the title and text accordingly.

Specific Comments:

1. There are no genome-wide significant MUC5B pQTLs nor fibrosis hydroxyproline/SHG QTLs. The only significant QTLs are for radiomic features. The authors do not show a direct link between radiomic features and pulmonary fibrosis, or a link between these QTLs and MUC5B. As the evidence currently stands, there is insufficient proof linking the novel genetic loci to fibrosis or MUC5B.

Thank you for pointing this out. We agree that, in this study, we did not identify genome-wide significant QTLs for MUC5B protein levels or for fibrosis end points measured by hydroxyproline content or SHG. The significant loci detected were associated with radiomic features derived from lung imaging. Radiomic features capture spatial and textural characteristics of lung injury and remodeling that may reflect aspects of disease heterogeneity not fully captured by hydroxyproline or SHG alone.

In response to this comment, we have reframed the manuscript to more clearly emphasize that this work explores genetic modifiers of lung imaging traits in the DO bleomycin model, unrelated to MUC5B. We have revised the title and text to avoid over-interpretation and to clarify that the identified loci represent modifiers of radiographic phenotypes rather than definitive genetic determinants of fibrosis severity or MUC5B regulation.

2. The DO founder mice all carry the orthologous MUC5B risk allele. Modification of MUC5B cannot be identified in a cohort where every mouse carries the MUC5B risk allele. As such, it is not appropriate to say that a modifier or modulator of MUC5B and MUC5B-driven events has been discovered. Similar language should be tempered, unless analysis of non-MUC5B risk allele carrying mice is done showing true effect modification.

We agree with the reviewer. Because all DO founder mice carry the orthologous *Muc5b* risk allele, it is not possible in this cohort to identify genetic modification of *Muc5b* itself or to demonstrate effect modification of *Muc5b*-driven events.

In response to this comment, we have modified the title and revised the manuscript to temper this language throughout. Specifically, we now avoid referring to the identified loci as modifiers or modulators of MUC5B or MUC5B-driven fibrosis. Instead, we frame our findings as the identification of genetic loci associated with lung imaging traits in the DO bleomycin model, within a uniform *Muc5b* background.

3. Variants within human syntenic regions corresponding to the QTLs were not found to be associated with IPF. These novel radiomics-associated genetic loci may be distinct from human genetic risks underlying IPF. As such, human IPF transcriptomic analyses may be less relevant than murine post-bleomycin transcriptomic analyses.

We thank the reviewer for this important point. We agree that the radiomic traits analyzed in our mouse model capture lung features that extend beyond established human lung fibrosis and likely reflect a broader spectrum of injury, inflammation, and tissue remodeling following

bleomycin exposure. This may, at least in part, explain why variants within the human syntenic regions corresponding to these QTLs are not associated with IPF risk.

We also agree that murine post-bleomycin transcriptomic analyses may be more directly informative for interpreting these loci in the context of the phenotypes studied. In the present manuscript, we leveraged eQTL data for a subset of genes to support candidate QTL driver gene prioritization. However, comprehensive transcriptomic analyses of post-bleomycin mouse lungs are ongoing, and we anticipate that these data will be presented and discussed in detail in a separate manuscript.

The inclusion of human IPF transcriptomic data in the current study was intended to explore the potential translational relevance of the loci identified in mice, rather than to imply direct genetic concordance with human IPF risk. This point has been clarified in the text.

4. There are no analyses of murine single cell RNA datasets. Existing mouse single cell RNA datasets of bleomycin-related lung injury and fibrosis could be analyzed to test the relevance of these novel radiomics-associated QTLs. In humans, MUC5B is highly expressed in terminal bronchioles, but mice lack terminal bronchioles. The lack of this relevant cell types in mice may temper conclusions reached in projecting genetic discoveries in mice to transcriptomic studies in humans, and vice versa.

We agree that analysis of murine single-cell RNA-seq datasets from bleomycin-induced lung injury and fibrosis would be highly informative. We plan to address this in a separate manuscript focused specifically on mouse RNA-seq analyses that includes eQTLs. While we acknowledge that translation between mouse and human systems is not straightforward, particularly given anatomical differences between mouse and human lungs, we believe it remains important to explore cross-species trends to assess potential biological relevance to human IPF.

5. Please clarify if Figure 4E represents the result of 90 genome-wide QTL analyses. If so, did genome-wide significance threshold for the RF QTLs account for multiple comparison testing of 90 tests and ~ 1M genetic loci ($\sim 2.8 \times 10^{-9}$)?

Figure 4E summarizes the results of genome-wide QTL analyses performed for 90 radiomic features. QTL mapping was conducted using qtl2, in which genome-wide significance thresholds were determined by 1,000 permutations for each radiomic feature, with the LOD threshold corresponding to $P = 0.05$ derived empirically. Because permutations were performed independently for each feature, the resulting genome-wide LOD thresholds differ across radiomic traits.

We recognize the reviewer's concern regarding multiple testing across 90 radiomic features. To address this, we have added two columns to Table 1: one reporting the genome-wide LOD threshold corresponding to $P = 0.05$ for each individual feature, and a second reporting a more stringent threshold adjusted for multiple testing across all 90 radiomic features.

6. Figures 6b and 6c show that expression of NKX1-2 is more differentially regulated in human IPF lung samples vs control, and more highly correlated with MUC5B, than OAT. These figures also show FGFR2 as another differentially expressed gene. Both NKX1-2 and FGFR2 reside in the chr 7 QTL. It is unclear why some genes were excluded from further analysis.

As detailed in the manuscript, OAT was selected based on two predefined, sequential criteria. First, six candidate genes were prioritized based for their potential to drive the identified QTLs. More specifically, genes exhibiting both eQTL and radiomics-QTL signals at the same locus are more likely to be causal. We first evaluated whether each local gene harbored an eQTL that overlapped the radiomic QTL and whether the haplotype exerting the strongest effects on gene expression and radiomics traits originated from the same founder strain. We then tested

whether including gene expression as a covariate in the QTL mapping attenuated or eliminated the radiomic QTL signal, indicating a potential mediating role. Second, expression of orthologs for all QTL-associated genes was assessed for differential expression and correlation with MUC5B expression in human IPF lung datasets. By integrating these two analyses, OAT emerged as the only gene that met the criteria for a QTL-driving candidate and also showed differential expression and correlation with MUC5B in human IPF samples. Although NKX1-2 and FGFR2 exhibit stronger differential expression or correlation signals in the human data, they did not meet the criteria for QTL-driving candidate genes and were therefore not prioritized for further validation. We have revised the manuscript text to clarify this selection rationale.

7. Statistical analysis of OAT protein expression across photomicrographs or between subjects stratified by MUC5B genotype (Fig 6) are not provided.

Thank you for your comment. We performed quantitative analysis for gene expression on published scRNAseq data. The IF images were performed for visualization purposes. We agree that protein quantification would be valuable information, but we don't have enough subjects with all different structures (honeycomb cysts and fibroblastic foci) present on the slides for statistical comparisons. However, to address the reviewer's suggestion, we have added a supplement figure (Supplementary Figure 14) with images from small airways from each of stained samples.

8. Supplementary Tables. Please improve readability of the pdf files. Multiple abbreviations are used for which there are no explanations.

We have added explanations for abbreviations used.

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3 (Remarks to the Author):

In this manuscript, Wang et al. present a systems genetics analysis of lung fibrosis in Diversity Outbred (DO) mice among the 8 founder strains, focusing on MUC5B expression and fibrotic responses. They identify several genetic loci and candidate genes that modulate fibrosis severity, with variable effects among mouse strains in bleomycin-induced fibrosis. The study highlights the complexity of MUC5B-dependent and independent mechanisms in fibrosis.

While the manuscript is well-written and the figures are clear, it reads more as a resource of modulated genes in IPF using human samples and extensive mouse data, rather than a mechanistic or hypothesis-driven story. The story builds on foundational data previously published by the group. The purpose and integration of MUC5B in the current findings are unclear, with both MUC5B-dependent and independent effects presented somewhat haphazardly. The most interesting phenotype which is the difference in fibrosis level between some strains (B6 vs A/J for instance) in their ability to get fibrosis is not explored. Understanding which mutations are associated with increased susceptibility to fibrosis in the pre-clinical model to identify novel targets rather than focus on MUC5B which is not showing major differences is

limiting the impact of this manuscript. OAT is highlighted as an interesting target, but its therapeutic relevance and rationale in IPF are not explored. Major revision is required, or the manuscript may be better suited for a journal focused on genetic resources or datasets.

Major comments:

- The manuscript should be revised for narrative structure to present a cohesive story rather than a data resource, if that is the intention. This is particularly noticeable in the figure legends, where the titles do not provide conclusions, but rather, technical descriptions of the data shown (e.g., “Study overview”; “Phenotypic distribution...following bleomycin challenge”; “Allele frequency....and heritability of all measured traits”; “Genetic mapping of...”). Many figures, albeit in supplemental, appear as repeated graphical representations for different genes. Do these provide mechanistic insights or a means of cataloging data?

We have improved the narrative flow and revised figure legends to emphasize the conclusions rather than technical details.

- Histology from the difference strains should be added to figure 1 on top of hydroxyproline quantification. H&E, Picro-sirius red, etc.

H&E staining and SHG imaging for each strain after 3-week bleomycin treatment were previously published (DOI:10.1152/ajplung.00615.2020); to avoid redundancy, we have not repeated the figures here, but we have cited the publication and highlighted the relevant key information in the text.

- What should be the “take-home-message” from the mouse strains in IPF models and MUC5B expression? While data from the 8 founder strains and DO mice are shown, with expected variability and complexity, what are the actionable conclusions and mouse model preferences for future studies?

We thank the reviewer for this comment. In our previous publication (DOI: 10.1152/ajplung.00615.2020), we performed detailed statistical tests across strains following bleomycin exposure. We found that the percent change following treatment showed significant strain-dependent effects, with WSB/EiJ exhibiting significantly elevated hydroxyproline and A/J, NZO/H1LtJ, and WSB/EiJ showing significant increases in fibrillar collagen relative to their respective saline controls. We have added a description of this to the revised manuscript.

- The role of MUC5B should be more clearly articulated, as the manuscript mixes MUC5B-dependent and independent effects. For example, Chr7-fibrosis-QTL is stated to contribute to lung fibrosis, independent of Muc5b regulation, in the same figure in which Oat is presented. It is therefore unclear how such genes fit within the manuscript title “Gene modulators of MUC5B-driven lung fibrosis”.

Based on multiple reviewer comments, we have implemented the following revisions: we reduced discussion directly linking candidate genes discovered to MUC5B, and clarified that different genes exert effects at different time points following bleomycin exposure: MUC5B-related QTL effects were observed at 3 weeks, whereas radiomics-associated QTLs were identified at 10 weeks post-treatment. To better reflect the revised scope and focus of the study, we have also modified the manuscript title to “Genes, other than Muc5b, associated with Bleomycin-Induced Lung Injury”

- Line 233: authors state MUC5B is involved in early response, but it remains upregulated 10 weeks post-bleomycin (Fig.2), well after peak fibrosis. Wouldn't this suggest MUC5B is a stable factor rather than an early responder?

In the Diversity Outbred cohort, MUC5B expression rises at 3 weeks post-bleomycin and declines by 10 weeks. Correspondingly, we observed a MUC5B-related QTL signal at 3 weeks but not at 10 weeks. Based on these temporal patterns, we conclude that the concentration of MUC5B is relevant in the early response phase following bleomycin exposure.

- Figure 2: authors should show the weight loss and histological data referenced on lines 112-113. These are critical for validating the biological relevance of the outlier data points mentioned.

We thank the reviewer for this suggestion. We have included the weight loss and histological check results in the Phenotype Data uploaded to ZENODO.

<https://zenodo.org/records/17316693?preview=1&token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6ImQxZjQxZWZiLTZlbnUtNGUxNC04Y2NjLWQzODM3YzlhYmZmYSIsImRhdGEiOiOnt9LCJyYW5kb2OiOiJiOGRhbjA0YjdjMzA0NWYyMjg3YzUyNjFiNDBiZjczNiJ9.ThSOiKQGyICNCABCTysHCwOh5kVSAusM3vi4i7NYjZ29H9tvlQsTi0Jji7-nx1FnnQM4jxWfRaZHV2tAFyDIDQ>

- Figure 3, line 130: what is the “strong biological evidence” claimed as supportive?

We appreciate the reviewer's request for clarification. Although the LOD score for this QTL does not reach the genome-wide significance threshold corresponding to $P = 0.05$, the haplotype effects observed across the eight founder strains align with the MUC5B phenotype in these strains. This concordance provides the "biological evidence" supporting the QTL, and we have revised the manuscript text to make this rationale clearer.

- Figure 3: can the authors correlate the heritability findings to fibrosis severity and/or MUC5B expression?

We thank the reviewer for the suggestion. However, we do not believe this analysis is feasible. Heritability was estimated at the cohort level rather than for individual animals. Because heritability represents a population-wide parameter, it cannot be directly correlated with individual-level traits such as fibrosis severity or MUC5B expression.

- Figure 4: can the authors elaborate on why only lung lavage provided significant QTLs? Why would hydroxyproline content or fibrillar collagen not lead to QTL detection? Was this due to the large number of outliers in the DO groups, as shown in Figures 2b-c and can this be corrected?

We thank the reviewer for this question. We initially expected to detect strong QTL signals for fibrosis-related readouts at 3 weeks post-bleomycin, and were likewise surprised that genome-wide significant QTLs were not observed for hydroxyproline content or fibrillar collagen. Instead, significant QTLs were identified for radiomic features at 10 weeks, which may reflect genetic influences on later-stage processes such as injury resolution and tissue remodeling rather than peak fibrosis burden. Unfortunately, given the experimental design, we weren't able to obtain hydroxyproline or fibrillar collagen measures at 10 weeks post bleomycin exposure.

We do not believe that the lack of QTL detection for hydroxyproline or fibrillar collagen is due to the presence of outliers. An inherent feature of the Diversity Outbred (DO) population is its extensive genetic and phenotypic diversity, which results in broad trait distributions across all measured phenotypes. This variability is observed consistently across hydroxyproline, fibrillar collagen, lung lavage measures, and radiomic features. Values that may appear as statistical outliers represent biologically valid measurements rather than technical artifacts, and therefore were not corrected or excluded.

- Figure 5: the authors previously demonstrated that MUC5B and Muc5b expression is associated with pathologic ER stress. What are the effects on the genes Xbp1, Hspa5, Atf4,

Atf6, Ddit3?

We thank the reviewer for this insightful suggestion. While examining the effects on genes such as Xbp1, Hspa5, Atf4, Atf6, and Ddit3 would be interesting, this analysis is beyond the scope of the current manuscript.

- Figure 5/6: the authors should consider analyzing genes involved downstream or upstream of MUC5B, including non-coding regulatory elements, as these may reveal additional or unconventional mechanisms of MUC5B regulation.

We thank the reviewer for this valuable suggestion. We are currently investigating genes upstream and downstream of MUC5B. However, these analyses are beyond the scope of the current manuscript and, therefore, are not included in the present discussion.

- Figure 6: while IF is shown for human tissue with GG and GT genotypes, the authors should show the TT genotype, homozygous for the risk allele. These genotypes should be better discussed. Additional images should also be presented, as MUC5B expression in the honeycomb cyst seems unspecific. May IF quantification be provided?

Thank you for your comment. Yes, we agree that lungs from TT genotype would provide additional information, however, we don't have tissues collect from the patients with TT genotypes yet. The MUC5B staining in the lumens of honeycomb cysts represent mucus accumulations with trapped cells. We added that in the figure legend description. We agree that protein quantification would be valuable information, unfortunately, we don't have enough subjects with all different structures present on the slides for statistical comparisons. In Supplementary Figure 14 of the revised manuscript, we present immunofluorescent staining for MUC5B (magenta), OAT (yellow), and Keratin 5 (white) in small-airway lung tissues from control (n=4) and IPF (n=6) donors, stratified by MUC5B genotype.

- Figure 6: The prioritization of the 6 candidate genes requires further justification. These genes do not all correlate with MUC5B, let alone IPF vs healthy control samples. In each panel, OAT shows borderline significance.

Our candidate gene prioritization strategy involved two sequential steps. First, we identified potential driver genes underlying the radiomics-QTL, resulting in six candidate genes. Second, we evaluated the orthologs of these QTL genes in human datasets to assess differential expression and correlation with MUC5B. Among the six genes, only OAT met both criteria, acting as a putative driver of the radiomics-QTL and showing differential expression as well as correlation with MUC5B in human data. For this reason, OAT emerges as the strongest candidate, despite showing borderline significance in some panels.

- If OAT is a key driving factor, can the authors show the therapeutic potential of targeting this factor in mechanistic studies with compounds and/or gene downregulation?

We thank the reviewer for this suggestion. Evidence for the therapeutic potential of targeting OAT has been reported in a recent study (DOI: 10.1038/s12276-024-01170-w), showing that the OAT inhibitor L-canaline attenuates bleomycin-induced lung inflammation and fibrosis in a mouse model. We have included this information in the revised manuscript.

Minor comments:

- Ensure all figures include appropriate statistical tests and p-values. Can any statistical significance be reported for Figure 2a-c plots for MUC5B protein levels and hydroxyproline content? The data analysis methods paragraph is particularly brief and should be revised.

We thank the reviewer for this comment. In our previous publication (DOI: 10.1152/ajplung.00615.2020), we performed detailed statistical tests across strains following

bleomycin exposure. We found that at the same time point, differences between strains were not statistically significant, whereas the percent change following treatment showed significant strain-dependent effects. In the current manuscript, we present the distribution of each strain at each time point; since differences between strains at individual time points are not significant, statistical significance is not reported for these plots. We have, however, added a description of this to the revised manuscript.

- Sample sizes should be reported and verified. Show all data points in volcano/box-and-whisker plots were possible (e.g., Figure 2). While Figure 1 presents n=545 DO mice used at 10 weeks post-bleomycin exposure, n=544 is used within the text.

We thank the reviewer for raising this point. The minor discrepancy between sample sizes arises because different phenotypes were measured at the same time point, and each phenotype includes slightly different numbers of samples. We have now clarified the exact sample sizes for each phenotype in the dataset description to ensure transparency and consistency.