A single-cell atlas of mouse lung development

Nicholas M. Negretti, Erin J. Plosa, John T. Benjamin, Bryce A. Schuler, A. Christian Habermann, Christopher S. Jetter, Peter Gulleman, Claire Bunn, Alice N. Hackett, Meaghan Ransom, Chase J. Taylor, David Nichols, Brittany K. Matlock, Susan H. Guttentag, Timothy S. Blackwell, Nicholas E. Banovich, Jonathan A. Kropski and Jennifer M.S. Sucre

DOI: 10.1242/dev.199512

Editor: Samantha Morris

Review timeline

Original submission:	9 February 2021
Editorial decision:	28 May 2021
First revision received:	30 September 2021
Editorial decision:	8 November 2021
Second revision received:	16 November 2021
Accepted:	19 November 2021

Original submission

First decision letter

MS ID#: DEVELOP/2021/199512

MS TITLE: A Single Cell Atlas of Lung Development

AUTHORS: Nicholas M Negretti, Erin J Plosa, John T Benjamin, Bryce A Schuler, A Christian Habermann, Christopher S Jetter, Peter Gulleman, Chase J Taylor, David Nichols, Brittany K Matlock, Susan H Guttentag, Timothy S Blackwell, Nicholas E Banovich, Jonathan A Kropski, and Jennifer MS Sucre

I have now received all the referees' reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, the referees express considerable interest in your work, but have some significant criticisms and recommend a substantial revision of your manuscript before we can consider publication. If you are able to revise the manuscript along the lines suggested, which may involve further experiments, I will be happy receive a revised version of the manuscript. In particular, the reviewers are in agreement that the novelty of your observations should be further explored, thus you shouldconsider providing some more characterization of the transitional cells or try to discuss/integrate your dataset with the recent Morrisey lab dataset.Your revised paper will be re-reviewed by one or more of the original referees, and acceptance of your manuscript will depend on your addressing satisfactorily the reviewers' major concerns. Please also note that Development will normally permit only one round of major revision.

We are aware that you may be experiencing disruption to the normal running of your lab that make experimental revisions challenging. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please attend to all of the reviewers' comments and ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion. I should be grateful if you would also provide a point-by-point response detailing how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

The manuscript by Negretti et al. presents a single cell time-series RNA-seq analysis of lung development across 8-time points from embryonic day 12 (E12) to postnatal day 14 (P14) in mice. RNA velocity analysis was used to explore dynamic transcriptomic patterns transition between the saccular and alveolar stages within epithelial, mesenchymal, and endothelial lineages.
The major concern is that the present work does not take full advantage of the well-designed single-cell developmental time-course data, fails to discover dynamics patterns associated with lung development phases and critical time points. The dynamic components in the present work purely relied on the latent time inferred from RNA velocity analysis which can be achieved using non-time-course single-cell RNA-seq data and the author's interpretation of the RNA velocity analysis is questionable. The power of time-course single-cell RNA-seq data was not demonstrated. Overall, the present study provides some single-cell supports of previous findings, providing limited new insights into lung development.

Comments for the author

Major comments:
1. Changes in the relative proportions of cell populations were used to represent developmental changes. Nevertheless, those changes were mostly identified based on single cell frequencies, which can be easily affected by sampling issues or technical factors during library preparation and sequencing. These results merit a validation using orthogonal approaches.
2. In the present study, the dynamic fate prediction was purely relied on the latent time inferred from RNA velocity analysis, authors didn't correlate or overlay the patterns with the real development time course. The power of time-course single cell RNA-seq data was not demonstrated.
3. It is very important to estimate a mapping of latent time and development time so that the latent-time dependent dynamic patterns can be understood in the context of lung developmental biology. I referred couple of examples on how to correlate computational inferred pseudotime with real developmental times (PMID: 27998929, PMID: 24658644, PMID: 30787437).
4. Figure 1 displayed UMAP embedding of 22 annotated cell-type within three major lineages. There is no brief description regarding how those cells are defined and what are the key markers were used. Some information can be found in the supplemental files. But it is difficult for readers to follow when reading the main text. A dot plot of representative markers for each cell type in Fig 1 will be helpful. Fig 1D highlighted cells from each time point in red. I noticed that epithelial cells mostly appear after birth which is inconsistent with previous studies, probably due to the sampling or technical issues.
5. The accuracy of the latent time prediction and interpretation is questionable.
- In Figure 2. The authors stated "Latent time (an estimate of a cellular maturity based on RNA velocity). From the expression patterns of Hopx and Sftpc in Fig. 2B-C, can one infer that AT1 is the most mature and AT2 is the least mature cells in the epithelial lineage? This seems to conflict with Fig. 2A which shows that there are more AT2 cells from later development time points. Would it be possible that the latent time estimation was affected by cell clustering? In Fig 2C, the expression pattern of Sftpc decreasing with the latent time which is inconsistent with the Sftpc RNA/protein expression patterns reported by other mouse lung development time course studies (PMID: 27602285; PMID: 11970905; PMID: 29516783), which showed expression of Sftpc RNAs was increased along with developmental time and reached a plateau postnatally.
- For Figure 5. Authors stated "Analysis of RNA velocity and cell trajectory mapping suggested that the prenatal proliferating myofibroblasts predominately self-renew, but also have a possibility of becoming Wnt2+ fibroblasts. Cell fate analysis did not identify a common mesenchymal progenitor population, suggesting that some degree of mesenchymal fate specification occurs earlier than E12.". Not sure I could agree with the interpretation. Fig 5B

indicated that Proliferating Wnt2+ FB is likely served as a common mesenchymal progenitor, has the potential to become Wnt2+ fibroblasts, Adventitial fibroblast, and Myofibroblast; similar to the PMP (proliferative mesenchymal progenitor) population identified and reported in previous studies (PMID: 27998929, PMID: 29590628).

6. A major finding in the present work is the role of AT1 cells in elastin assembly and extracellular matrix organization. Nevertheless, the authors only showed the expression of Fbln5 and components of the base membrane in AT1 cells. Those markers showed relative enrichment in AT1 than AT2, but those are not very selective or abundantly expressed markers to represent the mode of action of AT1 cells. No data or further analysis to support that the expression of these genes in AT1 cells contribute to the establishment of the elastin network.

7. A "primordial cell" was identified at early time points and was defined by the expression of Mdk (page 6). Is there any literature or other evidence to support this cell identity associate with this marker? My understanding of MDK is not an epithelial marker in general, it is more enriched in lung mesenchymal cells such as fibroblast and myofibroblast.

During lung development, Sox9+/Id2+ distal lung tip progenitor cells give rise to AT1 and AT2 cells. Are the "primordial cells" Sox9+/Id2+ cells?

I have similar concerns about how the transitional epithelial cells were defined, why choose Cdkn1a (Cyclin Dependent Kinase Inhibitor 1A) as the representative marker? Cdkn1a is not a selective marker for epithelial cells it is more enriched in immune or endothelial cells.

8. Figure 4D plotted Kdr expression increased in Car4+ population with the latent time. Kdr is a pan marker for endothelial cells, express in both Car4+ (aCap) and miEC (gCap) if I interpreted correctly. It is more relevant to plot Car4 and Car4+Pecam1+ expression here so we can related to the Fig 4F IH validation.

9. Some of the statements in the manuscript need data to support:

- Figure 7. "At E18, the lung epithelium is composed of indistinct cell types that are very rare or absent in adulthood" Maybe morphologically different than cells in adulthood, at the transcriptomic level, the cell types are quite similar at E18 and PN3.
- On page 14, "Our analysis found expression of laminin-332, …, is high at E18 in AT1 cells and continues through development." Please show the AT1 time course of the gene.
- On page 16, "The transitional epithelial cells, which we identified as Cdkn1a+, are similar to epithelial cells previously described in lung organoids as alveolar type 1 transitional cell state (PATS), Krt8+ transitional cells, and the damage associated transient progenitor (DATP), …"

No data or evidence to support these similarities. Does the Cdkn1a+ transitional epithelial cell in present work share similar marker or signature genes with PATS, Krt8+, or DATP cells? Does Cdkn1a express in PATS, Krt8+, or DATP cells?

- Fig 3 showed co-staining of Sftpc+Hopx+Cdkn1a+ transitional cells. Is it the bipotent cells identified by Desai et al. at E18.5 (PMID: 24499815) or AT1/AT2 cells identified by Guo et al at PND1 (PMID: 30604742)?
- Authors mentioned several places of lipofibroblasts population appears later in development, please provide umap of the cell cluster along the development course.

Minor:

- Fig. 2B-C, the direction of latent time needs to be specified.
- Page 16, Flbn5 should be Fbln5
- Page 17, Kdr4 should be Kdr.
- Page 20, the last sentence, this cell filtering criterion is not clear: "<0.5% 10% mitochondrial mRNA"
- The scVelo analysis needs to specify which approach was used to estimate RNA velocity.
- Page 21 : "Cell clusters with containing non-physiologic marker combinations, i.e. Epcam+/Pecam1+, were dropped at this stage". Not show how many combinations of Markers were checked for this purpose. It will be more efficient to run a doublet check using programs such as Scrublet or DoubletDecon.

Reviewer 2

Advance summary and potential significance to field

Negretti et al used longitudinal 10x based scRNASeq to generate a dynamic portrait of the cell state transitions in mouse lung development. They depleted CD45+ leukocytes and thus focused the analysis on epithelial, stromal and endothelial cell lineages. The authors use some of the newer

pseudotemporal/RNAvelocity analysis methods (scVelo/CellRank) to explore the time resolved data and point out transitional cell states. The data generated is of high quality and provides an important resource for the community. However, the study remains descriptive throughout and the recently published dataset by Zepp/Morrisey (https://pubmed.ncbi.nlm.nih.gov/33707239/) provides a very similar dataset with more in depth analysis and follow up. It would be of value to compare and possibly integrate the two datasets and identify complementary aspects.

Comments for the author

Comments:
(1) The authors point out that there is `there is tremendous coordination and interaction between epithelial, mesenchymal, and endothelial cells at every time point´. They do however not analyze this at all. Can they use cell-cell communication analysis tools such as Niche-NET to predict important mediators of the coordinated process of alveolarization?
(2) The discussion around specific ECM contribution by AT1 during alveolarization is interesting, however also descriptive and focused on cherry picked examples. Can the authors perform a more systemic analysis of the evolution of ECM during lung development and assess/visualize the contribution of individual cell types? For instance the BM between AT1 and aCAP (Car4+) is likely generated by a mixed contribution of both cell types. How does this occur in time and space during development? Some follow up both on the analysis and experimental side would be very relevant here.
(3) The authors discuss that the transitional cell state in alveolar epithelium is similar to the recently discovered intermediate state (PATS, Krt8+ADI – authors use a wrong reference, DATP). Can they show some quantitative analysis (e.g. matchScore) to which extend this is the case? How does this intermediate differ in ECM expression compared to primordial, AT2 and AT1?
(4) Nomenclature of vasculature: important to use consistent nomenclature or at least refer to other publications and identify corresponding cell labels: miEC is likely gCAP and Car4+ the aCAP in https://www.nature.com/articles/s41586-020-2822-7. Can the authors make correspondence of cell identities in other important work more clear for all cell labels throughout the study (maybe a supl table)?

Reviewer 3

Advance summary and potential significance to field

See comments below.

Comments for the author

In this manuscript, Negretti et al. combine scRNA-seq with RNA-in situ hybridization and RNA velocity to assess mouse lung development from E12 to P14, focusing on epithelial, mesenchymal, and endothelial cells. They confirmed the presence and kinetics of previously described Car4 endothelial cells, provided some clarity to Tgfbi/Wnt5a myofibroblasts and Wnt2 fibroblasts, and implicated a role of AT1 cells in extracellular matrix production. Most notably, they suggested a population of transitional epithelial cells that are predicted to come from either primordial or AT2 cells and become exclusively AT1 cells. Although their scRNA-seq data were carefully analyzed and confirmed with RNAscope, such data are readily available from other sources within the field. Considering the journal's guidelines "Techniques and Resources Articles or Reports describe a novel technique, a substantial advance of an existing technique, or a new resource that will have a significant impact on developmental biology research", we feel it is necessary to further confirm the existence of epithelial transitional cells and show their fate, which is the novel biology introduced in this manuscript.
RNAscope validation of transitional cells is not convincing. For example how was the total number of epithelial cells obtained in Fig. 3G. The automatic HALO analysis in Fig. S6D included non-epithelial cells, which would be more problematic in later lungs as cells adjacent to the alveolar lumen are often not epithelial. It is necessary to include a pan-epithelial marker such as Nkx2-1, Epcam, or Ecadherin. It would also increase confidence to show the proportions of all 4 epithelial cell types (primordial, AT1, AT2, and transitional) over time and see if they match those from scRNA-seq.

A related technical validation of the HALO analysis is to label cell types that are known to be the same (e.g. multiple pan-epithelial probes) or distinct (e.g. epithelial vs endothelial vs mesenchymal probes) to show the false positive and negative rates. In general, more experimental and analysis details are needed for RNAscope in the Methods.
The significance of the transitional cells and hence this Resources study hinges on their predicted AT1 fate and not just the result of imprecise transcriptional control of cell type markers. In the absence of Cdkn1a driver, one would like to see lineage-tracing experiments with Sftpc-CreER; Hopx-flp. The labeled cells are expected to be Cdkn1a+ve and become AT1 cells.
Additionally, the authors state that "the Wnt2+ fibroblasts becoming more randomly distributed around alveoli and the myofibroblasts becoming less aggregated over time (Figure 5D)". The text was written in a way to imply active migration, but the dispersion could be simply the result of the lung growing bigger and airspace expansion and could happen to AT2 cells as well.
Please clarify to avoid confusion.
Fig. 2H and I: why do a subset of proliferating myofibroblasts have the longest latent time?

Minor points:
Throughout the manuscript, both Sftpc (correct) and Sfptc (wrong) were used.
Figure S4E legend, "and Wnt2 (white, Car4+ cell marker)." There is a typo somewhere.
Methods: "Quality filtering was then used to remove cells with > 10% mitochondrial mRNA, < 0.5% 10% mitochondrial mRNA, and cells with < 700 detected genes." There is a typo somewhere.
In the introduction of the results, the authors describe their gating scheme for CD45, but do not address the use of Ter119.
When introducing the cell populations, the authors introduce epithelial sub-populations with 1 marker, endothelial sub-populations with 3 markers, and mesenchymal sub-populations with no markers. Perhaps the authors can address all markers used to define sub-populations in the body of the text.
When the authors list the sub-populations of the three cell populations they only use numerical lists for endothelial and mesenchymal sub-populations, not epithelial sub-populations. Perhaps they can revise and highlight the epithelial cells also using the numeric list format.

First revision

Author response to reviewers' comments

We are grateful for the editorial board's decision to receive a revision of our manuscript, "A Single-Cell Atlas of Lung Development" (199512). The thoughtful suggestions from the reviewers have improved the quality of the manuscript. Below, please find a point-by-point response to reviewers' comments and to requests by the editorial board. At the Editors' request we have 1) further explored the novelty of our findings, characterizing the transitional cells in more detail, 2) integrated our data with the recently published Morrisey lab dataset (Zepp et al, Science 2021:371) and discussed our findings in the context of this work, 3) Addressed all of the reviewers' comments in point-by-point detail below. Thank you for your consideration of this revision.

Reviewer 1:
Advance Summary and Potential Significance to Field:
The manuscript by Negretti et al. presents a single cell time-series RNA-seq analysis of lung development across 8-time points from embryonic day 12 (E12) to postnatal day 14 (P14) in mice. RNA velocity analysis was used to explore dynamic transcriptomic patterns transition between the saccular and alveolar stages within epithelial, mesenchymal, and endothelial lineages.
C1: The major concern is that the present work does not take full advantage of the well- designed single-cell developmental time-course data, fails to discover dynamics patterns associated with lung development phases and critical time points. The dynamic components in the present work purely relied on the latent time inferred from RNA velocity analysis which can be achieved using non-time course single-cell RNA-seq data and the author's interpretation of the RNA velocity analysis is questionable. The power of time-course single-cell RNA-seq data was not demonstrated. Overall, the present study provides some single-cell supports of previous findings, providing limited new insights into lung development.

R1: We appreciate the reviewer's careful reading of the manuscript and thoughtful feedback. In response, we have made several modifications to the organization of the manuscript, and we have included a substantial amount of new analysis. Critically, regarding validation, we have also done additional tissue localization with RNA in situ hybridization and immunofluorescence to further support key findings from the scRNA-seq data. We agree that the original version of this manuscript relied heavily on latent time analysis, and we now include additional analysis relying on the chronological time course, with further details below.
As this is a Techniques and Resources manuscript, we have made our dataset publicly explorable in an interactive data explorer to provide a resource to the developmental biology community (https://sucrelab.org/lungcells). Since a single manuscript cannot capture the full richness of possible findings from this data resource, we have highlighted some of the significant findings in this work, which we respectfully consider as adding depth to current knowledge (e.g., timing and characterization of transitional epithelial cells, a newly described role for type 1 pneumocytes, and the timing and spatial distribution of Wnt signaling in the mesenchyme). We expect that this public, user-friendly, searchable resource, enabled from an unusually large cell number, will freely allow other groups to probe the data to test the robustness of current hypotheses, in addition to developing new ones.

Comments for the AUTHOR
Major comments:

C2: Changes in the relative proportions of cell populations were used to represent developmental changes. Nevertheless, those changes were mostly identified based on single cell frequencies, which can be easily affected by sampling issues or technical factors during library preparation and sequencing. These results merit a validation using orthogonal approaches.

R2: We appreciate the opportunity to describe our methods in more detail, and we have modified the main text to make this aspect more clear. Our experiments were multiplexed from multiple mice using hashing antibodies (that allow tracking RNA-seq reads to each individual mouse, and very precise identification of doublets) to determine animal-to-animal variation in the cellular proportions, with n > 4 mice at each timepoint (excluding E12). We have indicated limitations to this in the discussion. As an orthogonal method of validation of these cellular proportions, we have quantified the relative number of cells in multiple key cell types using RNA in situ hybridization, including: Alveolar type I, alveolar type II, early epithelium, alveolar capillary endothelial cells, Wnt2+ fibroblasts, and myofibroblasts. While we have not quantified every cell type indicated in the single-cell experimentation, we do provide rigorous complementary validations of numerous key cellular populations, which demonstrate similar findings as estimated by sequencing and quantified by RNA ISH approaches.

C3. In the present study, the dynamic fate prediction purely relied on the latent time inferred from RNA velocity analysis, authors didn't correlate or overlay the patterns with the real development time course. The power of time-course single cell RNA-seq data was not demonstrated.

R3: We thank the reviewer for the helpful suggestions and have included additional analyses to better utilize the insights from the time-course experiment. We have included regression analysis using Monocle3 to indicate the genes that have dynamic expression profiles in particular cell-types over developmental time, and have made these analyses more substantial components of the primary figures (Figure 2D-F). We have rewritten the results section to better highlight key findings and how time affects gene expression. We have also included additional text in the discussion to indicate that these results support the notion lung development is largely asynchronous at the cellular level.

C4. It is very important to estimate a mapping of latent time and development time so that the latent-time dependent dynamic patterns can be understood in the context of lung developmental biology. I referred couple of examples on how to correlate computational inferred pseudotime with real developmental times (PMID: 27998929, PMID: 24658644, PMID: 30787437).

R4: In addressing the comment above, which we thank the reviewer for pointing out, we realized the need to reorganize the paper to first focus on development in "real" (chronological) time, with

a supporting use of latent time analysis to support specific cellular trajectories or developmental programs. Furthermore, we have added Supplementary figure 6 indicating that while latent time broadly correlates with real time, the two metrics do not always correlate. We posit that this is indicative of the fact that cellular developmental trajectories (cellular maturation) may occur asynchronously among the different cell-types and cell-fate transitions.

C5. Figure 1 displayed UMAP embedding of 22 annotated cell-type within three major lineages. There is no brief description regarding how those cells are defined and what are the key markers were used. Some information can be found in the supplemental files. But it is difficult for readers to follow when reading the main text. A dot plot of representative markers for each cell type in Fig 1 will be helpful. Fig 1D highlighted cells from each time point in red. I noticed that epithelial cells mostly appear after birth which is inconsistent with previous studies, probably due to the sampling or technical issues.

R5: We are grateful for this feedback. In response, in figure 1 we have included a dotplot indicating key marker genes for these 22 annotated cell-types. We have revised the plots indicating the number of cells at each timepoint to highlight the fact that epithelial cells are present at all sampled timepoints but tend to be more homogenous and cluster close together in the prenatal timepoints. We have attempted to provide further clarity in Supplemental Table 2 indicating relative numbers of every cell-type across development.

C6: The accuracy of the latent time prediction and interpretation is questionable.

- In Figure 2. The authors stated “Latent time (an estimate of a cellular maturity based on RNA velocity). From the expression patterns of Hopx and Sftpc in Fig. 2B-C, can one infer that AT1 is the most mature and AT2 is the least mature cells in the epithelial lineage? This seems to conflict with Fig. 2A which shows that there are more AT2 cells from later development timepoints. Would it be possible that the latent time estimation was affected by cell clustering? In Fig 2C, the expression pattern of Sftpc decreasing with the latent time which is inconsistent with the Sftpc RNA/protein expression patterns reported by other mouse lung development time course studies (PMID: 27602285; PMID: 11970905; PMID: 29516783), which showed expression of Sftpc RNAs was increased along with developmental time and reached a plateau postnatally.

R6: We have revised the figures and text to provide more clarity about our interpretation of latent time. We intended to use the latent time analysis to indicate which cell-types may give rise to other cell-types. That is to say, the data suggest that some of the AT2 cells had transcriptional profiles indicating an increase in AT1 markers and a decrease in AT2 markers. In other words, some AT2 cells may have the capability of becoming AT1 cells.

However, this does not indicate that as all AT2 cells mature, they are obligated to decrease Sftpc expression and become AT1 cells. Indeed, we have now clarified that we observe a relatively low number of AT2 cells that exist in the intermediate states, which suggests that this process is either somewhat rare or temporally quick. This analysis does not preclude prior work indicating that AT2 cells themselves may make increased Sftpc over time, and that the number of AT2 cells increases over time. We have clarified the text to more clearly indicate the findings and limitations of the latent time-based analysis.

C7: For Figure 5. Authors stated “Analysis of RNA velocity and cell trajectory mapping suggested that the prenatal proliferating myofibroblasts predominately self-renew, but also have a possibility of becoming Wnt2+ fibroblasts. Cell fate analysis did not identify a common mesenchymal progenitor population, suggesting that some degree of mesenchymal fate specification occurs earlier than E12.”. Not sure I could agree with the interpretation. Fig 5B indicated that Proliferating Wnt2+ FB is likely served as a common mesenchymal progenitor, has the potential to become Wnt2+ fibroblasts, Adventitial fibroblast, and Myofibroblast; similar to the PMP (proliferative mesenchymal progenitor) population identified and reported in previous studies (PMID: 27998929, PMID: 29590628).

R7: Given the apparent bifurcation of the mesenchymal cells into the Wnt2+ and myofibroblast populations we were hesitant to indicate a specific common progenitor population in the absence of additional evidence. We have added text to indicate that some of the cells (especially at early timepoints) have expression patterns similar to the PMP cells that have been previously identified,

and we have performed a Jaccard analysis to quantify this similarity.
Because our analysis demonstrates these cells are expressing distinguishing markers of myofibroblasts or Wnt2+ FBs, we remain hesitant to indicate that these are a generic progenitor population, but have added the possibility to our discussion. Given the suggestion from RNA velocity analysis that these cells may give rise to each other, it seems probable that a highly adaptable population of progenitor fibroblasts is represented in the single-cell dataset, and we have indicated this probability (with some qualifiers) in the revised discussion.

C7. A major finding in the present work is the role of AT1 cells in elastin assembly and extracellular matrix organization. Nevertheless, the authors only showed the expression of Fbln5 and components of the base membrane in AT1 cells. Those markers showed relative enrichment in AT1 than AT2, but those are not very selective or abundantly expressed markers to represent the mode of action of AT1 cells. No data or further analysis to support that the expression of these genes in AT1 cells contribute to the establishment of the elastin network.

R7: We have highlighted additional basement membrane components in Figure 6 and separated the genes into those involved in basement membrane and ECM and those involved in elastin assembly. Additionally, in this figure, we have compared expression of AT1 cells with other alveolar epithelium to demonstrate that expression of Col4a3, Col4a4, and Flbn5 is relatively specific to AT1 cells. We have revised the text to focus less on elastin specifically, and instead highlighted the potential role of AT1 cells in production of several components of the basement membrane. In addition to staining for Fbln5, we have found that the expression of Lama3 (Laminin alpha 3, a component of Lamanin-332) occurs in AT1 cells. Indeed, in the lung parenchyma, AT1 cells appear to be abundant expressors of both Fbln5 and Lama3. While expression of Fbln5 and Lama3 is not restricted to AT1 cells, they do exhibit increased expression. This is quantified in both the sequencing data and by RNA in situ hybridization in Figure 6.

C8: A “primordial cell” was identified at early time points and was defined by the expression of Mdk (page 6). Is there any literature or other evidence to support this cell identity associate with this marker? My understanding of MDK is not an epithelial marker in general, it is more enriched in lung mesenchymal cells such as fibroblast and myofibroblast. During lung development, Sox9+/Id2+ distal lung tip progenitor cells give rise to AT1 and AT2 cells. Are the “primordial cells” Sox9+/Id2+ cells?

R8: We chose to highlight the expression of Mdk in the early epithelium exactly because it was unexpected. Staining by RNA in situ hybridization indicates that not only is Mdk expressed in Epcam+ cells, it appears to be enriched in Epcam+ cells. The sequencing data further indicates that these early epithelial cells are Sox9 and Id2 positive. These additional markers of these early cells have been indicated in the text and supplementary figures.

C9:I have similar concerns about how the transitional epithelial cells were defined, why choose Cdkn1a (Cyclin Dependent Kinase Inhibitor 1A) as the representative marker? Cdkn1a is not a selective marker for epithelial cells, it is more enriched in immune or endothelial cells.

R9: In response to this feedback and the feedback from other reviewers, we have updated the text to indicate that these transitional epithelial cells are highly similar to epithelial cell-types seen during repair in the bleomycin model of induced injury in adult mice, and we performed a Jaccard index analysis to quantify that similarity in Figure 3. In the bleomycin injury model, these cells have been termed alveolar differentiation intermediate (ADI) or Krt8+ cells (Strunz et al, Nature Communications 2020). We have updated the discussion to explain that in our data, Cdkn1a appeared to be a specific marker by the scRNAseq data, and so we used it as a marker gene. Cdkn1a is a gene that is also enriched in the ADI, pre-alveolar type 1 transitional state (PATS), damage-association transition progenitors (DATP) and Krt8+ cells that have been previously observed (Kobayashi et al Nature Cell Bio 2020, Choi et al Cell Stem Cell, 2020).
Because these cells have no apparent unique positive identifier (exclusive to these cells), we have done additional staining using PanCK antibodies to further support the observed expression pattern in epithelial cells. In addition, to characterize the transitional cells more fully, we have performed analysis of the spatial distribution of the transitional cells indicates that they are often found within 10 μm of AT2 cells, and they are somewhat cuboidal in shape.

C10: Figure 4D plotted Kdr expression increased in Car4+ population with the latent time. Kdr is a pan marker for endothelial cells, express in both Car4+ (aCap) and miEC (gCap) if I interpreted correctly. It is more relevant to plot Car4 and Car4+Pecam1+ expression here so we can related to the Fig 4F IH validation.

R10: While most endothelial cells express a low amount of Kdr, Car4+ aCap endothelial cells express increased levels of Kdr as evaluated by RNA in situ hybridization and quantified in Figure 4 I-J.

Some of the statements in the manuscript need data to support:
C11: Figure 7. "At E18, the lung epithelium is composed of indistinct cell types that are very rare or absent in adulthood" Maybe morphologically different than cells in adulthood, at the transcriptomic level, the cell types are quite similar at E18 and PN3.

R11: This is a typographical error in the figure, which should read "E15" not "E18", We have fixed the typo in the figure as we intended to indicate E15. At E15, the epithelium does not strongly express markers of mature AT1 or AT2 cells.

C12: On page 14, "Our analysis found expression of laminin-332, …, is high at E18 in AT1 cells and continues through development." Please show the AT1 time course of the gene.

R12: We have reformatted figure 6 to more clearly indicate the findings by separating out basement membrane components from elastin components and to show the time-course of the three genes that comprise laminin-332 (Lama3, Lamb3, Lamc2). We have also performed RNA- ISH to show a time-course of *Lama3* expression in Hopx+ AT1 cells.

C13: On page 16, "The transitional epithelial cells, which we identified as Cdkn1a+, are similar to epithelial cells previously described in lung organoids as alveolar type 1 transitional cell state (PATS), Krt8+ transitional cells, and the damage associated transient progenitor (DATP), …" No data or evidence to support these similarities. Does the Cdkn1a+ transitional epithelial cell in present work share similar marker or signature genes with PATS, Krt8+, or DATP cells? Does Cdkn1a express in PATS, Krt8+, or DATP cells?

R13: We appreciate this comment, as it inspired a reanalysis of our data to place it in the context of other studies of transitional epithelial cells during injury. By Jaccard index analysis, our transitional epithelial cells share many common features of previously described DATP / ADI/PATS / Krt8+ cells (including *Cdkn1a* expression), one of the first times these kind of cells have been described during normal development. We have added additional comments of the similarities in the discussion.

C14: Fig 3 showed co-staining of Sftpc+Hopx+Cdkn1a+ transitional cells. Is it the bipotent cells identified by Desai et al. at E18.5 (PMID: 24499815) or AT1/AT2 cells identified by Guo et al at PND1 (PMID: 30604742)?

R14: Based on the Jaccard analysis indicated above, we conclude that the 'transitional' cells are similar to both the AT1/AT2 cells identified by Guo and the bipotent cells identified by Desai. Interestingly, these cells also show some similarity to DATP, PATS, and ADI cells, as noted above. This has been determined by calculating Jaccard indices based on the marker genes described in this study and others and is now shown in Figure 3D.

C15: Authors mentioned several places of lipofibroblasts population appears later in development, please provide umap of the cell cluster along the development course.

R15: Unsupervised clustering of the developmental data did not identify a separate lipofibroblast cluster per se, but rather that cells expressing Tcf21 and Plin2 appear at ~P0 within the Wnt2+ fibroblast population. This finding has been indicated in the text and supplementary figures. We attribute the lack of a distinct lipofibroblast cluster to the fact that the cells are all clustered together in the context of mature and immature cells over developmental time – we believe that this approach is a strength in that it helped highlight groups of cells that share the same features over time.

C16: Minor:
R16: All minor issues have been addressed as described below.

Fig. 2B-C, the direction of latent time needs to be specified.
An arrow and relative scale has been added.
Page 16, Flbn5 should be Fbln5
This error has been corrected.
Page 17, Kdr4 should be Kdr.
This error has been corrected.
Page 20, the last sentence, this cell filtering criterion is not clear: "<0.5% 10% mitochondrial mRNA"
The second '10%' is repeated and has been removed.
The scVelo analysis needs to specify which approach was used to estimate RNA velocity. We have indicated that dynamical modeling was used.
Page 21 : "Cell clusters with containing non-physiologic marker combinations, i.e. Epcam+/Pecam1+, were dropped at this stage". Not sure how many combinations of Markers were checked for this purpose. It will be more efficient to run a doublet check using programs such as Scrublet or DoubletDecon.

Preliminary testing indicated that Scrublet and DubletDecon performed poorly on this dataset. Therefore, we have applied rigorous doublet removal using the following strategy: We utilized the fact that our sequencing experiments included four mice that were individually identified with oligo-tagged antibodies (cell hashing). Using this approach, we removed clusters composed of cells containing multiple hashtags (doublets from different mice). While this approach may miss heterotypic doublets from the same mouse, it is expected to remove most of the doublets in the dataset. After this step, we removed any residual Epcam+/Pecam+ cells, however this was the only noted non-physiologic marker combination noted. This detail has been added to the methods section.

Reviewer 2:
C17: Negretti et al used longitudinal 10x based scRNASeq to generate a dynamic portrait of the cell state transitions in mouse lung development. They depleted CD45+ leukocytes and thus focused the analysis on epithelial, stromal and endothelial cell lineages. The authors use some of the newer pseudotemporal/RNAvelocity analysis methods (scVelo/CellRank) to explore the time resolved data and point out transitional cell states. The data generated is of high quality and provides an important resource for the community. However, the study remains descriptive throughout and the recently published dataset by Zepp/Morrisey (https://pubmed.ncbi.nlm.nih.gov/33707239/) provides a very similar dataset with more in depth analysis and follow up. It would be of value to compare and **possibly integrate the two datasets** and identify complementary aspects.

R17: We appreciate the reviewer's careful reading of the manuscript and thoughtful feedback. We have made several modifications to the organization of the manuscript, included a substantial amount of new analysis, adding an E16 timepoint to the scRNA-seq study, and additional tissue staining to further support key findings. In addition, we have made our dataset publicly explorable in an interactive viewer to better provide a resource to the lung community (https://sucrelab.org/lungcells).

We agree that the Zepp/Morrisey dataset (published after the initial submission of this manuscript) provides an opportunity for integration with our own dataset. We have downloaded the Zepp/Morrisey raw data and integrated it with our own time series. We have added this integration and joint analysis to the discussion and included cell-type identification comparisons in a supplementary figure. One of the central findings in our work is that AT1 cells express genes critical to establishing the basement membrane and elastin assembly, and this finding is independently validated in analysis of the Zepp/Morrisey dataset.

Reviewer 2 Comments for the Author: Comments:
C18: The authors point out that there is `there is tremendous coordination and interaction between epithelial, mesenchymal, and endothelial cells at every time point´. They do however not

analyze this at all. Can they use cell-cell communication analysis tools such as Niche-NET to predict important mediators of the coordinated process of alveolarization?

R18: We have taken your suggestion and applied Niche-NET to the data to explore the potential ligands and downstream signaling pathways that are involved in development. To further utilize the time-course nature of the experiment, we have opted to specifically explore ligands that have expression patterns that change with developmental time. The results section and supplemental figures have been updated with this information.

C19: The discussion around specific ECM contribution by AT1 during alveolarization is interesting, however also descriptive and focused on cherry picked examples. Can the authors perform a more systemic analysis of the evolution of ECM during lung development and assess/visualize the contribution of individual cell types? For instance the BM between AT1 and aCAP (Car4+) is likely generated by a mixed contribution of both cell types. How does this occur in time and space during development? Some follow up both on the analysis and experimental side would be very relevant here.

R19: To cast a wider net for ECM interactions, we have added additional genes to this dataset, including a large collection of ECM genes that are not expressed in AT1 cells (shown in the supplement). To further address the specific timing of these cells in the context of saccular stage development, we have added an additional E16 timepoint to the sequencing experiment and we have shown the time course of AT1 expression of these genes in Figure 6 and Supplemental Figure 8.

To further clarify the cells that express basement membrane components at the AT1 / aCap barrier, we have focused on expression of Collagen 4. RNA In situ hybridization indicates that AT1 cells make Col4a3 while aCap cells make Col4a1. Additional information about ECM expression patterns in aCap cells is included in the supplement. We agree that the larger question about the specific functional roles and source of the shared basement membrane between AT1 and aCap cells is important, however, further mechanistic studies into the mechanisms of this interaction are beyond the scope of this paper.

C20: The authors discuss that the transitional cell state in alveolar epithelium is similar to the recently discovered intermediate state (PATS, Krt8+ADI – authors use a wrong reference, DATP). Can they show some quantitative analysis (e.g. matchScore) to which extend this is the case? How does this intermediate differ in ECM expression compared to primordial, AT2 and AT1?

R20: We agree with the reviewer that the ‘transitional cells’ that we identified during development are similar to the DATP / PATS / ADI cells identified following injury in an adult mouse. To address this comment specifically, we applied a Jaccard index (the algorithm used by matchScore) to quantify the similarity between the transitional cells in our dataset and the transitional cells identified in these other studies, indicating a high degree of similarity (Figure 3). We have also used this analysis to quantify the relative similarity to the AT1/AT2 cells identified by Guo (PMID: 30604742) or the and bipotent progenitors identified by Desai (PMID: 24499815).

In terms of ECM expression, we have added data about the expression of ECM associated genes in all of the alveolar epithelial cells to Figure 6. To highlight additional expression patterns in the transitional cells, we have added dotplots to the supplementary information.

C21: Nomenclature of vasculature: important to use consistent nomenclature or at least refer to other publications and identify corresponding cell labels: miEC is likely gCAP and Car4+ the aCAP in https://www.nature.com/articles/s41586-020-2822-7. Can the authors make correspondence of cell identities in other important work more clear for all cell labels throughout the study (maybe a supl table)?

R21: We have adjusted the nomenclature to indicate aCap and gCap. Further, we have done extensive comparisons to the cell types called in other studies as indicated above using Jaccard index analysis.

Reviewer 3:
Reviewer 3 Comments for the Author:
C22: In this manuscript, Negretti et al. combine scRNA-seq with RNA-in situ hybridization and RNA velocity to assess mouse lung development from E12 to P14, focusing on epithelial, mesenchymal, and endothelial cells. They confirmed the presence and kinetics of previously described Car4 endothelial cells, provided some clarity to Tgfbi/Wnt5a myofibroblasts and Wnt2 fibroblasts, and implicated a role of AT1 cells in extracellular matrix production. Most notably, they suggested a population of transitional epithelial cells that are predicted to come from either primordial or AT2 cells and become exclusively AT1 cells. Although their scRNA-seq data were carefully analyzed and confirmed with RNAscope, such data are readily available from other sources within the field. Considering the journal's guidelines "Techniques and Resources Articles or Reports describe a novel technique, a substantial advance of an existing technique, or a new resource that will have a significant impact on developmental biology research", **we feel it is necessary to further confirm the existence of epithelial transitional cells and show their fate, which is the novel biology introduced in this manuscript**.

R22: We appreciate the reviewer's careful reading of the manuscript and thoughtful feedback. We have made several extensive modifications to the organization of the manuscript, included a substantial amount of new analysis, adding an E16 timepoint to the scRNA-seq study, and additional tissue staining to further support key findings. In addition, we have made our dataset publicly explorable in an interactive viewer to better provide a resource to the lung community (https://sucrelab.org/lungcells), in keeping with the spirit of a Techniques and Resources Article. We have characterized the transitional epithelial population in greater detail by spatial analysis, and we have used Jaccard index analysis to demonstrate the similarity and differences with this population in our manuscript and transitional populations of epithelial cells identified in other studies, placing this biological discovery in context with other work in lung injury and development, with additional details below.

C23: RNAscope validation of transitional cells is not convincing. For example, how was the total number of epithelial cells obtained in Fig. 3G. The automatic HALO analysis in Fig. S6D included non-epithelial cells, which would be more problematic in later lungs as cells adjacent to the alveolar lumen are often not epithelial. It is necessary to include a pan-epithelial marker such as Nkx2-1, Epcam, or Ecadherin. It would also increase confidence to show the proportions of all 4 epithelial cell types (primordial, AT1, AT2, and transitional) over time and see if they match those from scRNA-seq.

R23: We appreciate the feedback and have done additional experiments to characterize the transitional epithelial population at the tissue level. To address the concern about specificity of determining epithelial cells, we have stained the tissues with a pan-cytokeratin antibody in addition to RNA in situ hybridization with *Cdkn1a* and Sftpc. Analysis of the spatial distribution of the transitional cells indicates that they are often found within 10 μm of AT2 cells, and they are often somewhat cuboidal in shape. Quantification of cellular numbers by RNA in situ hybridization of epithelial cells has supported the proportions of epithelial cells as determined by the scRNA sequencing experiment. Furthermore, we have taken advantage of the fact that our experiments were multiplexed from multiple mice using hashing antibodies to determine animal- to-animal variation in the cellular proportions in n > 4 mice at each timepoint (excluding E12).

C24: A related technical validation of the HALO analysis is to label cell types that are known to be the same (e.g. multiple pan-epithelial probes) or distinct (e.g. epithelial vs endothelial vs mesenchymal probes) to show the false positive and negative rates. In general, more experimental and analysis details are needed for RNAscope in the Methods.

R24: To address the level of specificity in the HALO analysis, we have calculated the fraction of cells that are determined to be both Wnt5a+ (myoFB) and Sftpc+ (AT2) in addition to both Tgfbi+ and Wnt2+. These are gene combinations that are not expected biologically, and aren't present in the sequencing data. We found that the false positive rate was less than 1% on average. These false-positive data are now included in the supplement.

C25: The significance of the transitional cells and hence this Resources study hinges on their predicted AT1 fate and not just the result of imprecise transcriptional control of cell type markers.

In the absence of Cdkn1a driver, one would like to see lineage-tracing experiments with Sftpc-CreER; Hopx-flp. The labeled cells are expected to be Cdkn1a+ve and become AT1 cells.

R25: We have updated the text to indicate that these transitional epithelial cells are highly similar to epithelial cell-types seen during bleomycin induced injury of adult mice. These cells have been termed DATP, PATS, and ADI + cells. There have been extensive lineage labeling experiments described for these cells after injury and we've leveraged that to include additional discussion about these cells and their likelihood of developing into AT1 cells. While our Jaccard index analysis shows that these cells have been described before in adult injury, this study is of the first to identify them in a developmental context. In addition, the augmentation of our time series with an E16 timepoint further helps characterize the evolution of these cells during developmental time.

We are extremely interested in these cells as a matter of understanding the correlates between development and injury in the lung epithelium and have considered/evaluated a number of strategies to lineage-trace these cells. We suspect that the proposed dual-recombinase-based lineage tracing, while elegant, will unfortunately not answer this question as our transcriptomic data do not suggest coincident high-level gene expression of both *Sftpc* and Hopx in these transitional cells – we doubt this strategy would efficiently label transitional cells. Further, the Hopx proposed driver is likely to be problematic given that Hopx is less specific at earlier timepoints, both in our data and in the work of others (Zepp et al, 2021). Additional experimental models, including developmental injury, are likely needed to understand the role of these cells in a satisfactory manner, and we believe this work is critical but beyond the scope of this manuscript.

C26: Additionally, the authors state that "the Wnt2+ fibroblasts becoming more randomly distributed around alveoli and the myofibroblasts becoming less aggregated over time (Figure 5D)". The text was written in a way to imply active migration, but the dispersion could be simply the result of the lung growing bigger and airspace expansion and could happen to AT2 cells as well. Please clarify to avoid confusion.

R26: The text has been modified to indicate that we are only attempting to describe patterns of localization and we have been more precise with our word choice here. Specifically, we have highlighted the observation that the Tgfbi+ Wnt5a+ myofibroblasts are found clustered together during early lung development, but are often found in isolation later in development.

C27: Fig. 2H and I: why do a subset of proliferating myofibroblasts have the longest latent time?

We have rewritten the text and analysis to clarify the role of latent time as a marker of cellular maturity (or how long a cell has been present in the tissue). In this context, it is simply an indication that those cells are mature and don't become another cell type.

C28: Minor points:
R 28: All of these points have been addressed below:

Throughout the manuscript, both Sftpc (correct) and Sfptc (wrong) were used. Figure S4E legend, "and Wnt2 (white, Car4+ cell marker)." There is a typo somewhere.
These typos have been corrected.

Methods: "Quality filtering was then used to remove cells with > 10% mitochondrial mRNA, < 0.5% 10% mitochondrial mRNA, and cells with < 700 detected genes." There is a typo somewhere.
The second '10%' was included in error and has been removed.

In the introduction of the results, the authors describe their gating scheme for CD45, but do not address the use of Ter119.
The use of Ter119 to remove red blood cells has been addressed.

When introducing the cell populations, the authors introduce epithelial sub-populations with 1 marker, endothelial sub-populations with 3 markers, and mesenchymal sub-populations with no markers. Perhaps the authors can address all markers used to define sub-populations in the body of the text.

All Marker genes have been addressed more specifically in Figure 1 and in the text.

When the authors list the sub-populations of the three cell populations, they only use numerical lists for endothelial and mesenchymal sub-populations, not epithelial sub-populations. Perhaps they can revise and highlight the epithelial cells also using the numeric list format.

We have updated the text to have a consistent style.

Second decision letter

MS ID#: DEVELOP/2021/199512

MS TITLE: A Single-Cell Atlas of Lung Development

AUTHORS: Nicholas M Negretti, Erin J Plosa, John T Benjamin, Bryce A Schuler, A Christian Habermann, Christopher S Jetter, Peter Gulleman, Claire Bunn, Meaghan Ransom, Alice N Hackett, Chase J Taylor, David Nichols, Brittany K Matlock, Susan H Guttentag, Timothy S Blackwell, Nicholas E Banovich, Jonathan A Kropski, and Jennifer MS Sucre

I have now received all the referees reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.
The overall evaluation is positive and we would like to publish a revised manuscript in Development, provided that the referees' comments can be satisfactorily addressed. Please attend to all of the reviewers' comments in your revised manuscript and detail them in your point-by-point response. If you do not agree with any of their criticisms or suggestions explain clearly why this is so.

We are aware that you may currently be unable to access the lab to undertake experimental revisions. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Reviewer 1

Advance summary and potential significance to field

Many improvements have been made in the revision, including additional data at E16.5, additional tissue localization with RNA in situ hybridization, and immunofluorescence to further support key findings from the scRNA-seq data.
Together with the publicly available dataset and web portal, will be a useful resource to the lung developmental biology community.

Comments for the author

The revision has addressed most of my concerns, the manuscript quality is largely improved.

Reviewer 2

Advance summary and potential significance to field

the authors have adressed and satisfied all the points I have raised.

Comments for the author

the authors have adressed and satisfied all the points I have raised.

Reviewer 3

Advance summary and potential significance to field

See below.

Comments for the author

The revised manuscript has substantially improved, but the evidence for the transitional cells remains weak. Without robust methods to identify them, it is difficult for follow-up studies by others in the field. For example, the UMAP in Fig. 3A does not show transitional cells as a discrete cluster; they are instead embedded within the AT1 cell cluster. Also, their RNAscope validation defines transitional cells as Cdkn1a/Sftpc/Hopx triple positive, but the images (Fig. 3I, S2C) are not convincing, mainly because Cdkn1a dots are sporadic and over many cells. Can they show each channel separately and label negative, single double, and triple cells so that others repeating such an experiment would know how to interpret the data? If stronger evidence is not available, the conclusions about transitional cells need to be significantly softened.
A methodological detail needs clarification. Some single-cell samples were enriched via Epcam sorting. It's necessary to state which ones and whether such enrichment affects conclusions about cell proportions.
Fig. 2A and 2B are the same.
The title should include mouse or murine.

Second revision

Author response to reviewers' comments

We are grateful for the editorial board's decision to receive a revision of our manuscript, "A Single-Cell Atlas of Lung Development" (199512). We were especially encouraged by the responses of Reviewers 1 and 2, who were satisfied with our responses to their critiques. At the Editors' request we have 1) made revisions to our manuscript in response to the concerns of Reviewer 3 about our characterization and conclusions about the transitional cell population, 2) generated new UMAPs for pre- and postnatal epithelial cells, which allows for better visualization of the transitional cell cluster, 3) provided single-channel data for RNA *in situ* hybridization and clarification of our analysis of triple positive cells, 4) taken additional high-resolution confocal microscopy images to further spatially localize these transitional cells within the developing lung, 5) made changes to the text and title as requested, 6) responded to all of Reviewer 3's comments in point-by-point detail below. Thank you for your consideration of this revision.

Reviewer 1

Advance Summary and Potential Significance to Field:
Many improvements have been made in the revision, including additional data at E16.5, additional tissue localization with RNA in situ hybridization, and immunofluorescence to further support key findings from the scRNA-seq data. Together with the publicly available dataset and web portal, will be a useful resource to the lung developmental biology community.

Reviewer 1 Comments for the Author:
C1: The revision has addressed most of my concerns, the manuscript quality is largely improved.

R1: We are grateful for the helpful feedback and are encouraged by the comments of this reviewer.

Reviewer 2 Advance Summary and Potential Significance to Field: The authors have addressed and satisfied all the points I have raised.

Reviewer 2 Comments for the Author:
C2: The authors have addressed and satisfied all the points I have raised.

R2: We thank the reviewer for their thoughtful comments which have helped us improve this manuscript.

Reviewer 3 Advance Summary and Potential Significance to Field: See below.

Reviewer 3 Comments for the Author:
C3: The revised manuscript has substantially improved, but the evidence for the transitional cells remains weak. Without robust methods to identify them, it is difficult for follow-up studies by others in the field. For example, the UMAP in Fig. 3A does not show transitional cells as a discrete cluster; they are instead embedded within the AT1 cell cluster. Also, their RNAscope validation defines transitional cells as Cdkn1a/Sftpc/Hopx triple positive, but the images (Fig. 3I, S2C) are not convincing, mainly because Cdkn1a dots are sporadic and over many cells. Can they show each channel separately and label negative, single, double, and triple cells so that others repeating such an experiment would know how to interpret the data? If stronger evidence is not available, the conclusions about transitional cells need to be significantly softened.

R3: We appreciate the encouragement and this reviewer's observation that the manuscript has substantially improved.

This reviewer's critiques focus primarily on the identification and conclusions drawn about the transitional cell population. We appreciate the reviewer highlighting this opportunity to improve the clarity and overall conclusions drawn in our manuscript. In response to these comments, we have made the following revisions to the manuscript and figures:

1. Because developmental time is a strong driver of clustering in our dataset, we have generated additional UMAPs for the epithelial cells, separating prenatal and postnatal timepoints (Fig 3D, E). In these UMAPs, the transitional cell cluster is more clearly observed as distinct from the AT1 cell cluster.
2. We have added dot plots for marker genes of epithelial cells from the pre- and post-natal lung (Fig 3D, E) in order to better demonstrate the distinct expression patterns for the transitional cell population. In our dataset, this population is notable for high expression of *Cdkn1a* and *Krt8*, as well as low-level expression of both AT1 and AT2 marker genes, similar to other similar populations previously described in development and adult injury.
3. We have added to the supplement RNAscope images of Cdkn1a/Sftpc/Hopx as both merged and single-channel images, highlighting the triple-positive cells at E15, E18, P0, P3, and P5. Our sequencing analysis predicts that these transitional cells become exceedingly rare at P7 and P14, a finding confirmed by RNA *in situ* hybridization, and we have added this to the results section for emphasis.
4. To add to our spatial localization of these cells relative to the AT2 cells, we have obtained additional high-resolution confocal images and added these images to the main panel (Fig 3M).
5. We have revised the language used in the description of and conclusions about these transitional cells in the results and discussion sections. Specifically, we have distinguished between the prenatal and postnatal source/fate predictions of this population (Figure 3G and 3H). By using Jaccard comparative analysis, we have placed our findings in the context of prior work in lung development and lung injury.

With these revisions, our findings can be used in the field as a basis for follow-up mechanistic studies about this population to delineate the specific mediators for the expansion and differentiation of these cells during development and regeneration after injury.

C4: A methodological detail needs clarification. Some single-cell samples were enriched via Epcam sorting. It's necessary to state which ones and whether such enrichment affects conclusions about cell proportions.

R4: We have clarified that the P7 samples were enriched for epithelial cells with Epcam+ sorting in the methods section. Because our analysis of relative proportions compared subpopulations with groups (epithelial, mesenchymal, endothelial) after *in silico* sorting and not between these groups, this enrichment does not affect our conclusions about relative proportions. We have added this clarification to our methods section. Across all time points, the relative cell proportions of subpopulations of cells aligned with tissue validation by RNA ISH.

C5: Fig. 2A and 2B are the same.

R5: This error has been corrected.

C6: The title should include mouse or murine.

R6: We have added the word "mouse" to the title.

Third decision letter

MS ID#: DEVELOP/2021/199512

MS TITLE: A Single-Cell Atlas of Mouse Lung Development

AUTHORS: Nicholas M Negretti, Erin J Plosa, John T Benjamin, Bryce A Schuler, A Christian Habermann, Christopher S Jetter, Peter Gulleman, Claire Bunn, Meaghan Ransom, Alice N Hackett, Chase J Taylor, David Nichols, Brittany K Matlock, Susan H Guttentag, Timothy S Blackwell, Nicholas E Banovich, Jonathan A Kropski, and Jennifer MS Sucre
ARTICLE TYPE: Techniques and Resources Article

I am happy to tell you that your manuscript has been accepted for publication in Development, pending our standard ethics checks.