



**HAL**  
open science

**How to support clinical translation in neonatal lung research? Insights from multidisciplinary approaches to enhance our knowledge on normal human lung development in utero (single-nuclei RNA sequencing), provide proof of concept for gene therapy vir**

Laurent Renesme

► **To cite this version:**

Laurent Renesme. How to support clinical translation in neonatal lung research? Insights from multidisciplinary approaches to enhance our knowledge on normal human lung development in utero (single-nuclei RNA sequencing), provide proof of concept for gene therapy vir. Human health and pathology. Université de Bordeaux, 2021. English. NNT : 2021BORD0203 . tel-03633623

**HAL Id: tel-03633623**

**<https://theses.hal.science/tel-03633623>**

Submitted on 7 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Université de Bordeaux

École doctorale Sciences de la Vie et de la Santé - ED 154

*Spécialité Biologie Cellulaire et Physiopathologie*

### **How to support clinical translation in neonatal lung research?**

**Insights from multidisciplinary approaches to enhance our knowledge on normal human lung development *in utero* (single-nuclei RNA sequencing), provide proof of concept for gene therapy viral vector transfection in human lung parenchyma (precision-cut lung slices) and develop a consensus definition and reporting guidelines for clinical trial using mesenchymal stromal cells as therapeutic agent for human lung diseases.**

Par Laurent RENESME

Thèse de doctorat de sciences

Dirigée par le Dr. Éric DUMAS DE LA ROQUE et le Pr. Bernard THEBAUD

Présentée et soutenue publiquement le 8 septembre 2021

Devant un jury composé de :

<b>Pr. Anne-Monique NUYT</b> , Professeure	Présidente du jury, Rapportrice
<b>Pr. Thierry LACAZE-MASMONTEIL</b> , Professeur	Rapporteur et examinateur
<b>Pr. Christophe DELACOURT</b> , PU-PH	Rapporteur
<b>Pr. Thierry DEBILLON</b> , PU-PH	Examineur
<b>Pr. Pierre Henri JARREAU</b> , PU-PH	Examineur
<b>Pr. Laurent STORME</b> , PU-PH	Examineur
<b>Dr. Éric DUMAS DE LA ROQUE</b> , PH	Directeur de thèse (co-direction)
<b>Pr. Bernard THEBAUD</b> , Professeur	Directeur de thèse (co-direction)



## Quelles stratégies pour soutenir la translation clinique en recherche sur le poumon néonatal?

**Apports d'approches multidisciplinaires visant à améliorer notre compréhension du développement pulmonaire humain normal (single-nuclei RNA sequencing), valider la pertinence clinique d'un vecteur viral de thérapie génique sur du tissu pulmonaire humain (precision-cut lung slices), et développer une définition consensuelle et des lignes directrices pour reporter les essais cliniques utilisant les cellules stromales mésenchymateuses.**

Dans ce travail, nous rapportons 3 projets visant à améliorer et soutenir la translation clinique en recherche sur le poumon néonatal.

Premièrement, nous décrivons l'utilisation du séquençage d'ARN sur noyaux uniques (**single-nuclei RNA sequencing** ou snRNA-seq) pour étudier le développement pulmonaire humain. Le snRNAseq sur 23,251 noyaux cellulaires isolés à partir de neuf fœtus humains (âges gestationnels de 14 à 19 semaines de gestation) a permis de décrire neuf types cellulaires, incluant des populations cellulaires rares telles que les cellules pulmonaires neuroendocrines. Pour chaque type cellulaire, des gènes marqueurs ont été identifiés et utilisés pour confirmer la répartition spatiale cellulaire au sein du tissu par hybridation in situ en fluorescence (FISH). Les analyses d'enrichissement et de trajectoires développementales ont permis de décrire les modifications des voies moléculaires et de signalisation au sein d'un même type cellulaire en fonction de l'âge gestationnel. L'analyse des interactions ligands-récepteurs a mis en évidence les voies de communications entre les différents types cellulaires. Ces résultats offrent une perspective clinique majeure pour générer des hypothèses de recherche pertinentes pour tout projet étudiant le développement pulmonaire normal ou altéré et développer et valider des modèles de substitution pour étudier le développement pulmonaire humain (organoïdes pulmonaires humains).

Le second projet décrit l'utilisation de culture organotypique dérivée de tissu pulmonaire humain pour valider l'utilisation potentielle chez l'humain d'un virus adéno-associé (AAV pour adeno-associated virus) comme vecteur de thérapie génique pour traiter le déficit en protéine B du surfactant pulmonaire. Nous avons développé un modèle de "**precision-cut lung slices**" (PCLS) obtenues à partir de poumons fœtaux humains, ces PCLS pouvaient être maintenues en culture et conserver leur activité métabolique et architecture pendant plus de 7 jours. Les PCLS ont ensuite été utilisées pour valider avec succès la transfection du vecteur AAV dans le parenchyme pulmonaire humain. Cette transfection suivait un effet dose-dépendant et ne présentait pas d'effet délétère (évalué par le niveau d'activité métabolique et l'architecture des PCLS). Ces données démontrent l'intérêt des PCLS issues de poumons humains pour évaluer des thérapeutiques, soulignant l'importance de modèles expérimentaux complexes utilisant du tissu humain en recherche translationnelle.

Le troisième projet vise à développer une définition consensuelle pour les cellules stromales mésenchymateuses (Mesenchymal Stromal Cells ou MSC) et des lignes directrices pour reporter des essais cliniques utilisant des MSC pour améliorer la rigueur, la reproductibilité, la transparence de la recherche utilisant les MSC afin de garantir une translation sûre de thérapies cellulaires efficaces en clinique. Notre protocole de recherche combine une méthodologie utilisée pour le développement de consensus appelée **méthode Delphi** en combinaison avec une **stratégie de transmission des connaissances**. Nous présentons ici notre protocole de recherche détaillé et les résultats d'une synthèse de la littérature (scoping review) menée pour décrire comment les MSC sont définies et

caractérisées en recherche préclinique et clinique. L'analyse de notre échantillon d'études originales a mis en évidence une grande variabilité et des incohérences dans la description des MSC, leurs caractéristiques et leurs conditions de culture. De plus, la majorité des essais cliniques rapportaient de façon variable et non exhaustive ces caractéristiques importantes des MSC. Ce projet souligne l'importance du choix de la méthode d'élaboration de consensus et l'aspect fondamental des stratégies de transmission de savoir en recherche translationnelle.

**Mots-clés :** Développement pulmonaire, Single-nuclei RNA sequencing, Precision-cut lung slices, Cellules stromales mésenchymateuses, Méthode Delphi, Recherche translationnelle.

## Résumé thèse - Version longue

La translation clinique, qui consiste à transférer les découvertes et connaissances issues de la recherche en laboratoire jusqu'au lit du malade ("bench to bedside") est un processus long, complexe et coûteux qui se solde souvent par un échec. Ce fossé entre laboratoire et patients ("clinical translation gap") est décrit par certains auteurs comme une "vallée de la mort" où des découvertes prometteuses échouent sans jamais atteindre leur application en clinique. La recherche translationnelle a été développée pour soutenir la translation clinique et a comme objectifs d'accélérer la transition, d'augmenter le taux de succès de thérapies prometteuses lors d'essais cliniques et d'optimiser la transmission de savoir afin de soutenir la dissémination des découvertes à la population générale, aux responsables politiques ainsi qu'aux agences de régulation et de financement. Les déterminants de la recherche translationnelle incluent une approche multidisciplinaire, l'utilisation de modèles complexes pour étudier les états physiologiques et pathologiques et le développement de stratégies soutenant la transmission de savoir.

Dans cette thèse, nous présentons trois projets s'inscrivant dans le cadre de la recherche translationnelle et visant à soutenir la translation clinique en recherche sur les pathologies respiratoires néonatales.

Dans le premier projet, nous décrivons l'apport d'une analyse du transcriptome à l'échelon de la cellule unique (single-nuclei RNA sequencing ou snRNAseq) sur poumons fœtaux humains dans notre compréhension du développement pulmonaire humain normal. Jusqu'à présent, la majeure partie des connaissances en matière de développement pulmonaire concernait le développement des voies aériennes et provenait d'études animales et plus récemment d'études sur des organoïdes pulmonaires humains. Cependant, les différences en termes d'anatomie, physiologie et mécanistique moléculaire entre humain et rongeurs, la nécessité de valider le modèle organoïde par des données provenant de tissu humain *in-vivo* et le besoin d'une approche plus holistique intégrant les différents types cellulaires et leurs interactions

soulignent l'importance des approches "omiques" sur tissu pulmonaire humain sain afin d'améliorer nos connaissances sur le développement pulmonaire humain.

Nous rapportons ici le séquençage ARN à l'échelon cellulaire sur 23 251 noyaux cellulaires (snRNAseq) isolés à partir de tissus pulmonaires humains. Ces prélèvements de tissus provenaient de 9 fœtus humains dont les âges gestationnels (de 14 à 19 semaines d'aménorrhées) couvraient le stage pseudoglandulaire et le début du stage canaliculaire. Neuf types cellulaires ont été identifiés, incluant des populations cellulaires rares telle que les cellules neuroendocrines pulmonaires. Les 9 différents types cellulaires étaient présents à tous les âges gestationnels, et nous avons décrit l'évolution au cours du temps de la composition cellulaire avec une diminution du contingent cellulaire stromal au profit de la population cellulaire des voies aériennes distales. Chaque type cellulaire présentait une signature moléculaire unique caractérisée par des gènes différentiellement exprimés. Les 2 principales populations cellulaires (stroma et épithélium des voies aériennes distales) ont été analysées à une plus haute résolution pour identifier et décrire les sous types cellulaires les composant. Nous avons ensuite exploré l'évolution de l'expression génique en fonction de l'âge gestationnel pour chaque type cellulaire. Les gènes différentiellement exprimés au cours des différents âges gestationnels ont été utilisés pour réaliser des analyses d'enrichissement. Ces analyses d'enrichissement ont permis d'identifier les changements des voies de signalisation moléculaire au sein d'un même type cellulaire en fonction de l'âge gestationnel. Afin de décrire les communications entre les différents types cellulaires au sein du poumon humain en développement, nous avons ensuite réalisé une analyse des interactions ligands-récepteurs en comparant les âges gestationnels précoces (14+1 et 14+3 semaines) et tardifs (18+2 et 19+0 semaines) et identifié des voies de signalisation moléculaires générales comme les voies NOTCH et TGFB, ainsi que des voies plus spécifiques impliquées dans le développement vasculaire, neuronal et la régulation du système immunitaire. Finalement, nous allons confirmer la répartition spatiale de certaines populations cellulaires (épithélium des voies aériennes distales, endothélium vasculaire, péricytes et cellules neuroendocrines pulmonaires) par hybridation in situ en fluorescence (FISH) sur tissu pulmonaire aux âges gestationnels précoces (14 semaines) et tardifs (19 semaines).

Ces résultats représentent un apport important pour développer des hypothèses de recherche pertinentes d'un point de vue clinique pour toute recherche étudiant le développement

pulmonaire humain normal ou anormal. Nos résultats permettront également d'aider à développer et valider des modèles complexes utilisant du tissu pulmonaire humain, tel que les organoïdes pulmonaires, visant à étudier le développement pulmonaire humain.

Le second projet décrit l'utilisation de culture organotypique à partir de tissu pulmonaire humain pour valider le potentiel de translation clinique d'un vecteur viral utilisé en thérapie génique en apportant la preuve de concept de sa transduction dans du parenchyme pulmonaire humain. Notre équipe a démontré l'intérêt de l'utilisation du virus adéno-associé (AAV pour adeno-associated virus) comme vecteur de thérapie génique dans un modèle de déficit en protéine B du surfactant pulmonaire chez la souris. Pour valider son utilisation potentielle chez l'humain nous avons utilisé un modèle de culture organotypique utilisant des sections de 300 microns d'épaisseur obtenues à partir de poumons fœtaux humains ("precision-cut lung slices" or PCLS). Nous avons démontré que ces PCLS dérivées de poumons fœtaux pouvaient être maintenues en culture et conserver une activité métabolique et une architecture pendant plus de 7 jours. Nous avons mis en évidence que l'ajout dans le milieu de culture d'AAV avec son cargo ADN (gène humain de la protéine B du surfactant) pendant 24 heures au jour 3 de culture permettait la transduction dans le parenchyme pulmonaire, évaluée par immunofluorescence et bioluminescence au jour 7 de culture. De plus, la transduction se faisait avec un effet dose dépendant. L'analyse de l'activité métabolique et de l'histologie des PCLS ne démontrait pas d'effet délétère de l'AAV sur le tissu pulmonaire humain. Nos résultats confirment l'intérêt de l'utilisation de culture organotypique dérivée de tissu humain telle que les PCLS pour évaluer la pertinence et le potentiel de translation clinique de thérapies prometteuses. L'utilisation des PCLS illustre parfaitement l'intérêt et la nécessité de développer des modèles plus complexes et holistiques en recherche translationnelle.

Le troisième projet consiste à utiliser une méthode de développement de consensus (méthode Delphi) en association avec une stratégie de transmission de savoir afin de développer une définition consensuelle pour les cellules stromales mésenchymateuses (Mesenchymal Stromal Cells ou MSC) et des lignes directrices pour reporter des essais

cliniques utilisant des MSC. Les MSC sont utilisées depuis plusieurs décennies en recherche clinique dans de nombreuses maladies avec des résultats variables, et ce malgré des résultats prometteurs dans de nombreux modèles animaux. Parmi les causes pour expliquer cet échec de translation clinique, l'absence d'une définition consensuelle des MSC partagée par la communauté scientifique et la variabilité des informations reportées dans les essais cliniques (caractéristiques des MSC, modalités de production et d'administration des MSC...etc.) sont souvent identifiées. Pour améliorer la transparence et la reproductibilité des essais cliniques évaluant des MSC nous proposons de développer, disséminer et implanter une définition pour les MSC ainsi que des lignes directrices pour le report d'essais cliniques évaluant les MSC en utilisant une méthode Delphi. La méthode Delphi permet d'éviter de nombreuses limitations liées à la méthode de consensus basée sur un groupe d'expert (pression par les pairs, nombre limite de participants, difficulté de réunir des personnes avec des opinions divergentes...etc.) et propose une méthode de développement de consensus plus rigoureuse. En associant cette méthode Delphi à une approche intégrée de transmission de savoir qui intègre dans le projet, depuis sa création et pour chaque étape, les parties prenantes et futurs utilisateurs de la définition pour s'assurer que leur préférence et avis sont pris en compte, nous espérons optimiser la dissémination et implémentation des résultats de la méthode Delphi. Nous présentons ici le protocole de recherche ainsi qu'une synthèse de la littérature (scoping review) préliminaire à notre méthode Delphi. Le protocole de recherche décrit les différentes étapes de la recherche, le recrutement de notre groupe d'experts internationaux et leurs différents domaines d'expertise ainsi que du panel de scientifiques qui participeront à l'enquête Delphi, et finalement les stratégies d'implémentation prévues pour optimiser la dissémination des résultats. La scoping review avait pour objectif de de décrire comment les MSC sont définies et caractérisées dans un échantillon d'essais pré-cliniques et cliniques. L'analyse de notre échantillon d'études originales a mis en évidence une grande variabilité et des incohérences dans la description des MSC, leur caractéristiques et leurs conditions de culture. De plus, la majorité des essais cliniques rapportaient de façon variable et non exhaustive des caractéristiques importantes des MSC. La scoping review a également permis d'identifier les items à intégrer dans le questionnaire initial de la méthode Delphi ainsi qu'un panel de scientifiques pour participer à la méthode Delphi. Les résultats de notre scoping review soutiennent la nécessité de développer et disséminer une définition consensuelle pour les MSC et des lignes directrices pour reporter les essais cliniques utilisant ces cellules. Cet

effort est nécessaire pour améliorer la rigueur, la transparence et la reproductibilité de la recherche sur les MSC et a pour objectif final d'assurer une translation clinique sûre et efficace de thérapies basées sur les MSC. Ce projet souligne l'importance de la transmission de savoir et des stratégies d'implémentation en recherche translationnelle.

En conclusion, cette thèse met en évidence plusieurs aspects complexes de la recherche translationnelle et l'importance d'une approche holistique et multidisciplinaire pour combler le fossé de la translation clinique. De plus, nous présentons ici différentes solutions et stratégies pour soutenir la collaboration multidisciplinaire, le développement de modèle préclinique complexe utilisant du tissu pulmonaire humain et la transmission de savoir. Tous ensemble, ces projets démontrent comment soutenir et améliorer la translation clinique en recherche pulmonaire néonatale.

## How to support clinical translation in neonatal lung research?

**Insights from multidisciplinary approaches to enhance our knowledge on normal human lung development *in utero* (single-nuclei RNA sequencing), provide proof of concept for gene therapy viral vector transfection in human lung parenchyma (precision-cut lung slices) and develop a consensus definition and reporting guidelines for clinical trial using mesenchymal stromal cells as therapeutic agent for human lung diseases.**

In this thesis, we present 3 different projects that aim to support clinical translation in neonatal lung research.

First, we use state-of-the-art approach to study normal human lung development at a single-cell level. We report the cellular composition, cell trajectories and cell-to-cell communication in developing human lungs with **single-nuclei RNA sequencing** (snRNA-seq) on 23,251 nuclei isolated from nine human fetuses with gestational ages between 14 to 19 weeks of gestation. We describe nine different cell types, including rare cell populations such as pulmonary neuroendocrine cells. For each cell type, marker genes are reported and use for spatial validation using fluorescent RNA in situ hybridization. Enrichment and developmental trajectory analysis provide insight into molecular mechanisms and signaling pathways changes within cell type according to gestational age. Last, ligand-receptor analysis highlights determinants of cell-to-cell communication among the different cell types. These findings will provide a clinically relevant background for research hypotheses generation in projects studying normal or impaired lung development and help to develop and validate surrogate models to study human lung development such as human lung organoids.

Second, we use a model of organotypic culture called **Precision-Cut Lung Slices (PCLS)** derived from human fetal lung tissue to validate the clinical relevance of a viral vector (Adeno-associated virus AAV) for gene therapy for surfactant B deficiency in neonates. We develop a protocol to obtain and culture PCLS from human fetal lung tissues. We reported that these PCLS can be maintained up to 7 days in culture with a preserved metabolic activity and architecture. These PCLS were then successfully used to provide proof of concept that the AAV vector can efficiently transfect human lung parenchyma. PCLS were incubated at day 3 of culture with different titers of AAV (and with different cargos: mouse SPB and human SPB) for 24 hours. On day 7 of culture, we confirmed by bioluminescence and immunofluorescence that AAV can transfect human lung parenchyma with a dose dependent effect. Furthermore, assessment of the PCLS (by metabolic activity and architecture) demonstrated the absence of adverse effects on human lung tissue transfected by AAV vector. Therefore, PCLS from human lung tissue (healthy or diseased) represent a versatile tool to assess pathophysiology or therapeutics in an organotypic culture setting. It aligns with the need for more complex and holistic models and the use of human tissue in translational research.

Last, we aim to develop a consensus definition for Mesenchymal Stromal Cell (MSC) and reporting guidelines for clinical trials of MSC therapy to enhance rigour, reproducibility, transparency in the MSC research and ultimately the safe translation of effective cell-based therapies in humans. Our research protocol uses the unique attributes of a consensus-building method called modified **Delphi method** in combination with an '**integrated knowledge translation**' approach where key stakeholders and end-users are involved from inception and participate in every step of the study to support dissemination and implementation of the findings. We present here our detailed research protocol, and the results



of a scoping review conducted to describe how MSC were defined and characterized in preclinical and clinical research. In this scoping review, we report that previous minimal criteria to define MSC provided by the International Society for Cell and Gene Therapy (ISCT) were poorly implemented with inconstant reporting among in both preclinical and clinical studies. More concerning, the clinical studies showed inconsistent completeness in reporting relevant and important information on MSC characterization and cell manufacturing processes. This project highlights important points for consensus-building strategies and the critical step of knowledge translation in translational research.

**Keywords.** Lung development, Single-nuclei RNA sequencing, Precision-cut lung slices, Mesenchymal stromal cells, Delphi method, Translational research.

Cette thèse a été préparée dans les laboratoires suivants :

**Thébaud Lab - Sinclair Center for Regenerative Medicine**

Ottawa Hospital Research Institute

501 Smyth Road, Ottawa, Ontario

Canada

**Équipe Christelle GUIBERT « Physiopathologie de la circulation pulmonaire et systémique »**

Centre de Recherche Cardio-Thoracique de Bordeaux (CRCTB) – INSERM U1045

Plateforme Technologique d'Innovation Biomédicale (PTIB)

Hôpital Xavier Arnoz, Avenue du Haut Lévêque, 33604 Pessac Cedex

France

# REMERCIEMENTS

Je tiens en premier lieu à remercier tous les membres du jury qui ont accepté de participer à l'évaluation de mon travail de thèse et qui ont accordé du temps à la lecture de ce manuscrit, en particulier à la Pre Anne-Monique Nuyt qui me fait l'honneur de présider mon jury. J'adresse mes sincères remerciements à mes rapporteurs, les Pr Christophe Delacourt et le Pr Thierry Lacaze-Masmonteil, pour l'intérêt qu'ils ont portés à ce travail et pour avoir consacré du temps à la lecture et l'évaluation de ce manuscrit. Aux Prs Thierry Debillon, Pierre Henri Jarreau et Laurent Storme, merci d'avoir accepté d'être les examinateurs de cette thèse.

A mes directeurs de thèse, le Dr Éric Dumas de la Roque et le Pr Bernard Thébaud. Éric et Bernard un immense merci pour avoir supporté ce projet (et moi par la même occasion) et d'avoir été présent tout au long de cette thèse. Bernard je me souviendrai toujours de notre première rencontre au PAS meeting en 2017 à San Francisco. Je t'ai dit que je voulais venir pour 1 an, tu m'as dit que 2 ça serait mieux. Résultat ça fait 3 ans qu'on est à Ottawa, et le retour en France ne va pas être simple tant on a apprécié notre vie ici ! Ceci étant je suis ravi de revoir Éric de l'autre côté de l'Atlantique ! Merci à tous les 2, car en plus de vos qualités professionnelles et de mentoring indéniables, ce que j'apprécie le plus chez vous sont votre ouverture d'esprit et vos qualités humaines. Alors merci encore d'avoir rendu cette incroyable expérience possible !

Merci aux personnes qui m'ont aidé dans mes projets des 2 côtés de l'Atlantique : Christelle Guibert, Paul Robillard, et Claire Marie Pilard de l'unité INSERM 1045. Arul Vadivel, Shumei Zhong, Liqun Xu, Chantal Horth from the Thébaud lab, many thanks guys for all that you have done for me! I also want to thank the Drs Kelly Cobey and Manoj Lalu from the Ottawa Hospital Research Institute for their valuable help and guidance for the Delphi project!

A huge thanks to Drs Ivana Mizikova and Martin Kang for their review of this manuscript and their insightful comments, thanks guys!

To the Thébaud lab dream team: Flore Lesage, Chanele Cyr-Depauw, Maria Hurskainen, Ivana Mizikova and Pauline Bardin. Thank you guys for being around during these years, I much appreciated your support, help, and advices and most of all, all the fun we had! Looking forward to meeting you again in Canada or somewhere in Europe!

To the Dr Stephanie Redpath and the CHEO neonatal transport team. Thanks for having giving me the opportunity to be your fellow and a part of your team. This year was amazing, and it was a pleasure to work with you all. Thanks for all the support and your understanding while I was trying to manage to finish this thesis, that was really appreciated!

A toute la gang d'Alta Vista, les Peltzers, Goubets, Simpsons, Burtons & Co, merci d'avoir été là et de nous avoir fait partager et découvrir l'Ontario et le style de vie à la canadienne, c'était génial (si, si même le canoé camping en pleine tempête !). Merci pour votre accueil et votre gentillesse, vous avez rendu ces 3 années inoubliables ! On se revoit bientôt :)

A nos familles en France, détendez-vous on rentre ;) ! On a hâte de tous vous revoir et surtout les nouveaux membres de la famille !

Enfin le 'core group' Renesme, Agnès, Sarah, Yann et Clémentine (et Myrtille pièce ajoutée certes, mais seul vrai canadien de la famille). Quelle expérience incroyable on a vécu durant ces 3 dernières années à Ottawa ! Tant de bons souvenirs et de nouveaux amis ! Et on a tenu le cap même dans les moments pas simples entre l'arrivée dans un nouveau pays, le COVID, l'école en ligne (merci Doug Ford, il ne va pas me manquer celui-là...), et la thèse. On peut aller n'importe où, je ne suis pas inquiet, tant qu'on est tous les 5 on sera toujours heureux ! Je vous aime tellement !

## Table of contents

<b>Main abbreviations .....</b>	<b>3</b>
<b>Figures and tables list .....</b>	<b>4</b>
<b>I. Figures.....</b>	<b>4</b>
<b>II. Tables.....</b>	<b>7</b>
<b>Introduction.....</b>	<b>8</b>
<b>I. General introduction. About the importance of translational research .....</b>	<b>8</b>
1. Evolution of biomedical research and the development of the clinical translation gap .....	8
2. Bridging the gap: development of translational research .....	9
A. What is translational research? .....	9
B. What are the goals of translational research?.....	10
C. The translational research model and its challenges. ....	11
D. What are the determinants of translational research? .....	13
3. Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research (INCuBAToR): optimizing translational research determinants to improve clinical translation of cell therapy for bronchopulmonary dysplasia (BPD) .....	17
<b>II. Insights into transcriptomics during human lung development at a single-cell resolution .20</b>	
1. Human lung development.....	20
A. Different stages of human lung development.....	20
B. Understanding the human lung development: insights on molecular mechanisms from animal models and human tissue-based models.....	26
a. Animal (rodents) models to study human lung development .....	26
i. Rationale for animal use to study human lung development .....	26
ii. Insight from animal (rodents) models for human lung development.....	28
iii. Limitations and pitfalls of mouse models to study human lung development .....	30
b. Human lung organoids to study human lung development. ....	33
iv. The “old” in-vitro model using human lung derived cells. ....	33
v. Human lung organoids, principles and advantages. ....	34
vi. Cell sources for human lung organoids.....	36
vii. Limitations of human lung organoids to study lung development. ....	38
2. Single cell RNA sequencing (scRNAseq).....	39
A. Principle of scRNAseq .....	39
B. Single-cell resolution transcriptomics in respiratory and lung development research .....	40
C. ScRNAseq analysis workflow .....	42
D. scRNAseq, advantages and limitations .....	47
<b>III. Human fetal lung derived Precision-cut lung slices (PCLS) to validate clinical relevance of promising pre-clinical therapies .....</b>	<b>49</b>
1. Precision-cut lung slices (PCLS): definition, principle and potential applications .....	49
2. PCLS workflow .....	51
3. Advantages and limitations of the PCLS model .....	53
<b>IV. A Delphi study to establish consensus definition and reporting guideline for Mesenchymal Stromal Cells (MSC) .....</b>	<b>56</b>
1. Mesenchymal stromal cell research .....	56
2. Delphi method and “integrated knowledge translation” approach.....	60
A. Overview of the Delphi method .....	61
B. Knowledge translation (KT) strategy .....	62
a. “integrated knowledge translation” (iKT) approach .....	63
b. End-of-project KT Plan .....	63

<b>Thesis objectives.....</b>	<b>66</b>
<b>Results.....</b>	<b>68</b>
<b>I. A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation</b>	<b>68</b>
1. Background.....	68
2. Results .....	69
3. Conclusion .....	69
4. Manuscript: “A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation” (under review in European Respiratory Journal).....	70
<b>II. A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency..</b>	<b>71</b>
1. Background.....	71
2. Results .....	72
3. Conclusion .....	73
4. Article Kang et al. <i>Nature Communication</i> 2020 .....	78
<b>III. A Delphi study to establish a consensus definition and clinical reporting guidelines for Mesenchymal Stromal Cells .....</b>	<b>79</b>
1. Background.....	79
2. Results .....	79
3. Conclusion .....	81
4. Manuscript “Establishment of a consensus definition for Mesenchymal Stromal Cells (MSC) and reporting guidelines for clinical trials of MSC therapy: a modified Delphi study protocol” (published in BMJ Open).....	82
5. Letter to the Editor “A systematic approach to enhance transparency in Mesenchymal Stromal Cell research” (published in Cytotherapy) .....	83
<b>IV. Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review.....</b>	<b>84</b>
1. Background.....	84
2. Results .....	84
3. Conclusion .....	85
4. Manuscript “Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review” (Published in Stem Cells Translational Medicine) .....	86
<b>General discussion and conclusions.....</b>	<b>87</b>
<b>References.....</b>	<b>95</b>
<b>Annex.....</b>	<b>105</b>
<b>V. Article INCuBATor Thébaud et al. <i>Stem Cells Translational Medicine</i> 2021 .....</b>	<b>105</b>
<b>VI. Article Hurskainen et al. <i>Nature Communication</i> 2021 .....</b>	<b>105</b>

# Main abbreviations

ALI	Air-liquid interface culture
AAV	Adeno-associated virus
BPD	Bronchopulmonary dysplasia
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
GA	Gestational age
HLO	Human lung organoid
hPSC	human pluripotent stem cell
iKT	Integrated knowledge translation
ISCT	International Society for Cell & Gene Therapy
MSC	Mesenchymal stromal cell
PCLS	Precision-cut lung slice
PNEC	Pulmonary neuroendocrine cell
RNA	Ribonucleic Acid
scRNAseq	single-cell RNA sequencing
snRNAseq	single-nuclei RNA sequencing
SP-B	Surfactant protein B

# Figures and tables list

## I. Figures

### Figure 1. “Translational Research” citations in PubMed

Results found on PubMed for the search “translational research”, compared to all publication for the period between 1993 and 2020. The left Y-axis represents the number of citations for “translational research”, the right Y-axis represents the number of all citations. Interrogation on May 17<sup>th</sup> 2021.

### Figure 2. Operational challenges for translational research and medicine

Translational research includes 5 stages. T0 represents the fundamental studies and applied research that define cellular mechanisms, their relationship to disease and, consequently, the identification of therapeutic targets and methods of treatment (new molecular entities); T1, first phase 1 studies in humans that aim to define proof of mechanism and proof of concept; T2, phase 2 and 3 clinical trials that are necessary for the approval of a therapeutic agent for clinical use; T3, phase 4 clinical trials that are associated with optimizing the therapeutic use of an agent in clinical practice; and T4, the outcome or comparative effectiveness research that serves to determine the ultimate utility and cost effectiveness of an agent relative to others currently in use. From (Blumberg et al. 2012).

### Figure 3. Old vs. new translational research workflows

**A.** Old one-way linear biomedical research model, where animal experiments (mouse) were a gateway between basic science and clinical research. **B.** New translational research model: more complex and holistic network, with multiple lines of evidence to refine objectives and target relevance in order to increase the chance of successful drug discovery. (Adapted from Sabroe 2007). Plain double arrows represent T1 research, with back-and-forth movements of knowledge and hypotheses between different stages. The human symbol represents both the use of human cells and tissues (healthy and diseased) and clinical trials. Dashed double arrows represent knowledge translation with T2 research. Figure created with [BioRender.com](https://BioRender.com).

### Figure 4. Overview of the knowledge translation process

Knowledge translation (KT) include four different activities to support and improve individual and public health as well as the health care system. Adapted from Kathryn Sibley, CHI KT platform (<https://medium.com/knowledgenudge/what-we-mean-when-we-say-knowledge-translation-1f81d57d5143>).

### Figure 5. Overview of INCuBAToR framework

The INCuBAToR project proposed risk mitigation strategies to improve clinical translation of cell-based therapy for BPD patients. The INCuBAToR concept relies on an evidence-based approach with 4 main pillars: **A.** preclinical and clinical systematic reviews: this knowledge synthesis will help to summarize current evidence and level of evidence to support the study rationale and identify knowledge gaps; **B.** Knowledge translation: identification of the key stakeholders (end users) as well as barriers and facilitators to further support dissemination and implementation of the future findings; **C.** Early economic evaluation to ensure the project “viability” from an economical perspective; and **D.** retrospective/prospective cohort studies to inform clinical trial design development and ensure its feasibility. From (Thébaud et al. 2021).

### Figure 6. Morphological changes in human developing lung during the pseudoglandular, canalicular and saccular stages

During the pseudoglandular stage (a), future airways are embedded in the mesenchyme, surrounded by a loose capillary network and continue to branch repeatedly to give rise to increasing generations of airways (arrow, branching morphogenesis). Distal airways are lined by simple columnar epithelium (d). The canalicular stage (b) is characterized by airspaces development and the vascularization of these airspaces: airways start to



widen, epithelium transitions from a simple cuboidal form to a flatter epithelium (initiation of mature pneumocytes differentiation), and the capillaries grow and organize in a primitive capillary network, closer to the distal airways (e). In the saccular stage (c) begins the process of the alveolar-capillary interface maturation with the development of the saccules (denoted by \*) and the formation of thick immature interspaces septa with a close proximity between capillaries and airways epithelium (f). From (Schittny 2017).

**Figure 7. Lung development stages: comparison between rodents and human**

E: embryonic day, P postnatal day. Adapted from (Salaets et al. 2017; Nardiello, Mižíková, and Morty 2017).

**Figure 8. Human lung organoids overview**

In this figure are shown the main cell sources to obtain human lung organoids (hPSC: human pluripotent stem cell, which include embryonic stem cell (ESC) and induced pluripotent stem cell (iPSC)), different culture mode (such as 3D, 3D ALI, lung-on-a-chip), and the main applications for human lung organoids. Human lung organoids can be used to study normal lung development or to serve as a platform for disease models. Adapted from (Bourguignon et al. 2020). Figure created with [BioRender.com](https://BioRender.com).

**Figure 9. Overview of directed differentiation process and HLO maturation**

The normal human embryonic development is represented at the top of the figure: ICM: inner cell mass of the blastocyst, A: anterior, D: dorsal, DP: dorsal pancreas, Int: intestine, L: lung, Li: liver, P: posterior, St: stomach, V: ventral, VP: ventral pancreas. Directed differentiation steps are represented in color boxes: gastrulation (blue) leading to definitive endoderm (SOX17<sup>+</sup>/FOXA2<sup>+</sup>), patterning (yellow) leading to anterior foregut endoderm (SOX2<sup>+</sup>/FOXA2<sup>+</sup>), and specification (green) leading to ventral lung airway progenitor (NKX2-1<sup>+</sup>/SOX2<sup>+</sup>/TUJ1<sup>-</sup>/PAX8<sup>-</sup>). The red box represents the additional step for maturation with engraftment of human lung organoids derived from hPSC or human fetal tissue in an immunocompromised mouse. Human pluripotent stem cells (hPSC) include embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC). Adapted from (Barkauskas et al. 2017; Nikolić, Sun, and Rawlins 2018).

**Figure 10. Overview of scRNAseq workflow, from tissue sampling to spatial validation**

Different steps of the scRNAseq workflow are presented here. After tissue sampling (1), the sample undergoes a dissociation protocol to obtain a single-cell suspension (2a). Single cells are isolated from the cell suspension using a droplet-based method (2b), with each droplet containing a barcoded bead (cf figure 10). Cells are then lysed, and cDNA is synthesized. Next step (3) is library construction and sequencing using Next Generation Sequencing. After pre-processing, data can be used for downstream analysis (4) which include clustering and cluster identification, trajectory analysis, differentially expressed gene analysis, enrichment analysis and ligand-receptor interaction. Last step (5) is the spatial validation of scRNAseq findings on native tissue, for example using fluorescent RNA in-situ hybridization, immunohistochemistry or spatial omics approaches. Pictures of the downstream analysis are adapted from (Hurskainen et al. 2021). Figure created with [BioRender.com](https://BioRender.com).

**Figure 11. Barcoded bead structure**

Oligonucleotides are attached to the bead. Their structure includes 1) a primer region for molecular amplification of the captured transcript (PCR Handle), 2) a unique series of nucleotides (cell barcode) that are identical for all oligonucleotides attached on the same bead to identify cell origin of all the amplified transcripts, 3) second series of nucleotide unique for each oligonucleotide (UMI) to identify amplified transcripts originating from the same mRNA molecule, and 4) poly-d(T) region to capture polyadenylated RNA.

**Figure 12. Paracrine effects of MSC**

MSC paracrine effects are mediated by (A) MSC secretome and (B) MSC-derived extracellular vesicles. These paracrine effects enhance tissue repair and regeneration by supporting immunomodulation, chemoattraction, progenitor cells proliferation and differentiation, and angiogenesis. In addition, MSC-derived exosomes decrease oxidative stress and apoptosis. From (Jimenez-Puerta et al. 2020).

**Figure 13. Precision Cut Lung Slices (PCLS) of adult mouse lung tissue (Genetic strain: FVB/N) transduced with AAV6.2FF**

**a** Study design to show transduction of adult mouse lung tissue slices by AAV6.2FF. Metabolic (resazurin) assays were carried out on days 0, 3, and 7 post-tissue harvest (marked with an \*) to ensure PCLS metabolic activity and cell viability. **b** Image of mouse PCLS. **c** Resazurin assay results as represented by mean fluorescence with SD demonstrating the metabolic activity of the mouse PCLS on days 3 and 7. **d** IVIS imaging of mouse PCLS 4 days after AAV-Luc transduction and 7 days after generating the tissue slices demonstrates luciferase expression in the adult mouse lung parenchyma following the addition of D-luciferin substrate. **e** Quantification of bioluminescence from the IVIS image in (d) as Total Flux (photons per second). All treatment groups contain n=2 animals, 1 lung each with 6 slices per lung, except AAV-Luc with n=4 animals, 1 lung each with 5-6 slices per lung. Adapted from Supplemental Figure 1 from (Kang et al. 2020).

#### **Figure 14.** AAV6.2FF transduces the human lung parenchyma

**a** Study design to demonstrate the ability of AAV6.2FF to transduce human lung tissue samples generated from precision cut lung slices (PCLS). Metabolic viability (resazurin) assays were conducted on days 2 (D2) and 7 (D7) post tissue harvest. **b** Resazurin assay results as represented by mean fluorescence with SD demonstrating the metabolic activity of the human PCLS on days 2 and 7. **c** Bioluminescence detection using IVIS in PCLS from human fetuses (n = 2; 16+0 and 16+3 weeks gestational age) either untreated or transduced with 108 vg per well of AAV-GFP, 108 vg per well of AAV-mCherry, or 105, 108, or 1010 vg per well of AAV-Luc. Images are 4 days following vector exposure, and the scale is represented as counts. **d** Quantification of AAV-Luc wells are represented by the mean total flux (in photons per sec) with SD (n =2 fetuses; with 12 replicates for the AAV-GFP, AAV-mCherry, and No Treatment groups; and 4 replicates for each of the AAV-Luc doses). **e** Images of human PCLS in a 24-well tissue culture plate 7 days after culture and 3 days after AAV-SPB transduction. An enlarged image of a PCLS in an individual well is indicated by the dashed white box. **f** Resazurin assay results as represented by mean fluorescence with SD demonstrating viability of the human PCLS 3 days after no treatment (blue), or transduction with 1010 vg per well of AAV-mSPB (red) or AAV-hSPB (pink). (n = 4 blank; and n = 6 untransduced, n =6 AAV-mSPB, and n = 6 AAV-hSPB independent tissue slices from a single set of human fetal lungs). The source data for f has been provided as a Source Data file. **g** Representative confocal images of SP-B (red) and DAPI (blue) of frozen PCLS sections 3 days after transduction with 1010 vg per well of AAV-mSPB or AAV-hSPB (Top and bottom row image scale bars, 100  $\mu$ m and 50  $\mu$ m, respectively). From Figure 10 from (Kang et al. 2020).

#### **Figure 15.** Precision Cut Lung Slices (PCLS) from human lungs of 2 human fetuses (Ages: 16<sup>0</sup> and 16<sup>+3</sup> weeks) transduced by AAV6.2FF vectors

**a** Quantification of bioluminescence from the IVIS image in Fig 10c as Total Flux (photons per second). Total Flux of AAV-Luc wells are presented as the mean with SD (n=2 fetuses; 2 lung slices per fetus per dose for AAV-Luc). **b** Representative H&E images from human fetal PCLS that were fixed in 4% PFA immediately after the IVIS imaging in Fig 14c (Scale bar, 100 $\mu$ m). The lungs demonstrate established acinar structures and are in the late pseudoglandular stage of development. **c** Representative epifluorescence images of luciferase (red), mCherry (red), and DAPI (blue) in frozen PCLS 4 days after transduction with 10<sup>8</sup> vg per well of AAV-Luc or AAV-mCherry (Scale bar, 100 $\mu$ m). **d** Representative epifluorescence image of EpCAM (red) and DAPI (blue) of frozen PCLS from human lung sections demonstrating the presence of epithelial cells in developing future alveoli. **e** Representative epifluorescence image of DAPI (blue), EpCAM (green), and luciferase (red) or mCherry (red) in frozen PCLS 4 days after transduction with 10<sup>10</sup> vg per well of AAV-Luc or 10<sup>8</sup> vg per well of AAV-mCherry (Scale bar, 100 $\mu$ m). From Supplementary Figure 11 from (Kang et al. 2020).

#### **Figure 16.** Delphi MSC project overview

#### **Figure 17.** Scoping review graphical abstract

## II. Tables

**Table 1.** Timing of human lung development, growth and ageing

From (Burri 2006).

**Table 2.** Congenital and acquired (prematurity) lung diseases and the stage of their development

Adapted from (Kotecha 2000).

**Table 3.** Examples of mutations in mouse and their reported phenotype

Gain-and-loss of function studies in genetically modified animals have helped to identify key factors in lung development. Adapted from (Cardoso and Lü 2006).

**Table 4.** Species differences between mouse and human respiratory system

From (Meyerholz et al. 2018).

**Table 5.** Comparison of Human pulmonary 3D models for translational research

ALI: Air-liquid interface, PCLS: precision-cut lung slices, hPSC: human pluripotent stem cells, ECM: extracellular matrix. From (G. Liu et al. 2019).

**Table 6.** Differences between group decision-making and Delphi Method

Adapted from (Murphy et al. 1998).

# Introduction

## I. General introduction. About the importance of translational research

### 1. Evolution of biomedical research and the development of the clinical translation gap

Before the 70's, most of the biomedical research was led and done by physician-scientists. Biomedical research in this setting was an iterative and bilateral process, with fluid movement of knowledge and hypothesis between bench and bedside. With the development of molecular biology and the incredible advances in basic science techniques and knowledge, basic and clinical research started to separate, and nowadays, the majority of biomedical research is done by basic science researchers (Butler 2008; Roberts et al. 2012). This separation has slowly led to the rise of silos between basic and clinical scientists and the development of a cultural divide (Wadman 2006; Restifo and Phelan 2011). As a result, movements of knowledge and hypotheses' back and forth between bench and bedside are hampered due to a lack of communication and division in culture and views.

In addition, since the middle of the 90's, pharmaceutical industry expenditures for Research & Development (R&D) have increased while new drug output decreased. A progressive worsening of clinical success rate has been highlighted in a study by Hay et al. The authors showed that only one-in-ten drugs entering clinical development (phase 1) is expected to advance to Food and Drug Administration (FDA) approval (Hay et al. 2014). These findings have been confirmed by a recent study showing that the probability of success for phase III

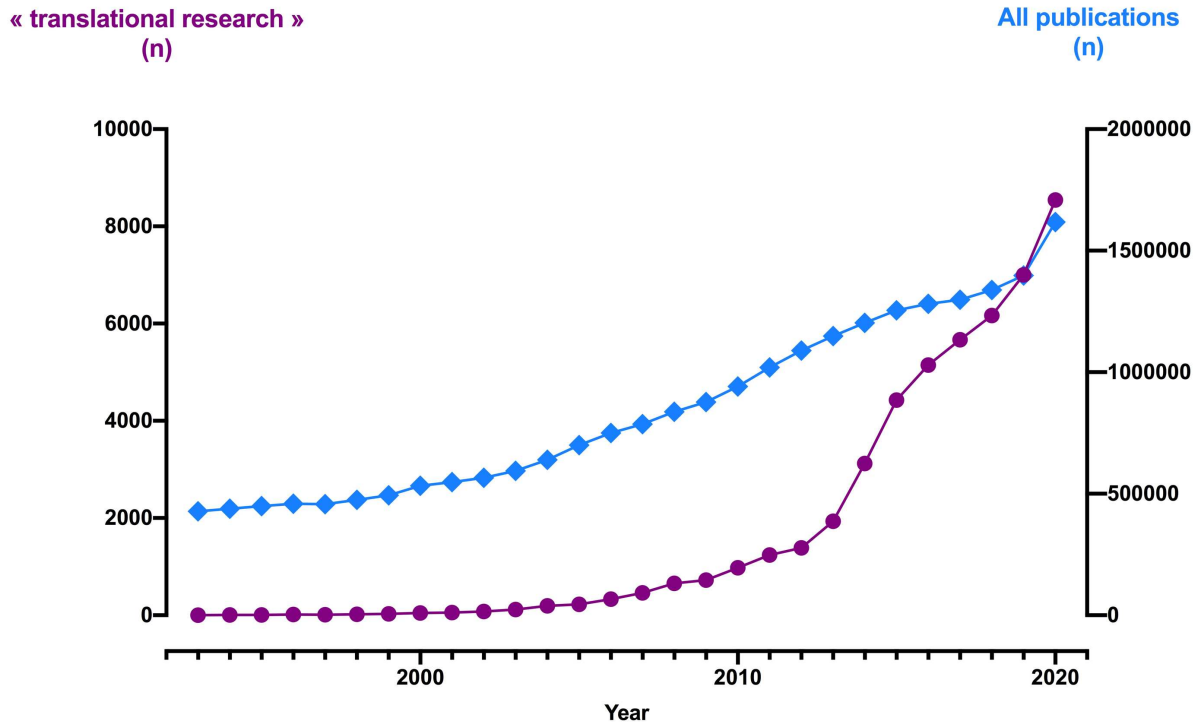
trials to be approved has increased over time (from 49% for the period between 2010-2012 to 62% for the period between 2015-2017), but the probability of success for phase I trials to be approved is still below 10% (Dowden and Munro 2019).

Therefore, the pharmaceutical industry has slowly abandoned its role in bringing academic discoveries to the patient's bedside, letting the burden of R&D, regulation and patent development fall on scientists' shoulders. Increased complexity in the ethics process, regulation, patent issues, securing funding and career advancement, lack of training in R&D specificities, and the overall overwhelming workload (scientific or clinical activities, teaching, administrative tasks...) have led many scientists to abandon this important part of biomedical research. As a result, a gap between bench and bedside research has developed, called by some authors as the 'Valley of death' where promising scientific discoveries linger and die (Butler 2008; Fernandez-Moure 2016; Roberts et al. 2012; Seyhan 2019).

## **2. Bridging the gap: development of translational research**

### **A. What is translational research?**

Translational research, also referred to as 'bench-to-bedside' research, is defined as the "process of translating discoveries in the laboratory into clinical interventions for the diagnosis, treatment, prognosis, or prevention of disease with a direct benefit of human health" (Minna and Gazdar 1996). The first entry for translational research in PubMed (<https://pubmed.ncbi.nlm.nih.gov>) was in 1993, with tremendous acceleration in publications after 2010 (Figure 1), showing a growing interest in translation research.



**Figure 1.** “Translational Research” citations in PubMed  
 Results found on PubMed for the search “translational research”, compared to all publication for the period between 1993 and 2020. The left Y-axis represents the number of citations for “translational research”, the right Y-axis represents the number of all citations. Interrogation on May 17<sup>th</sup> 2021.

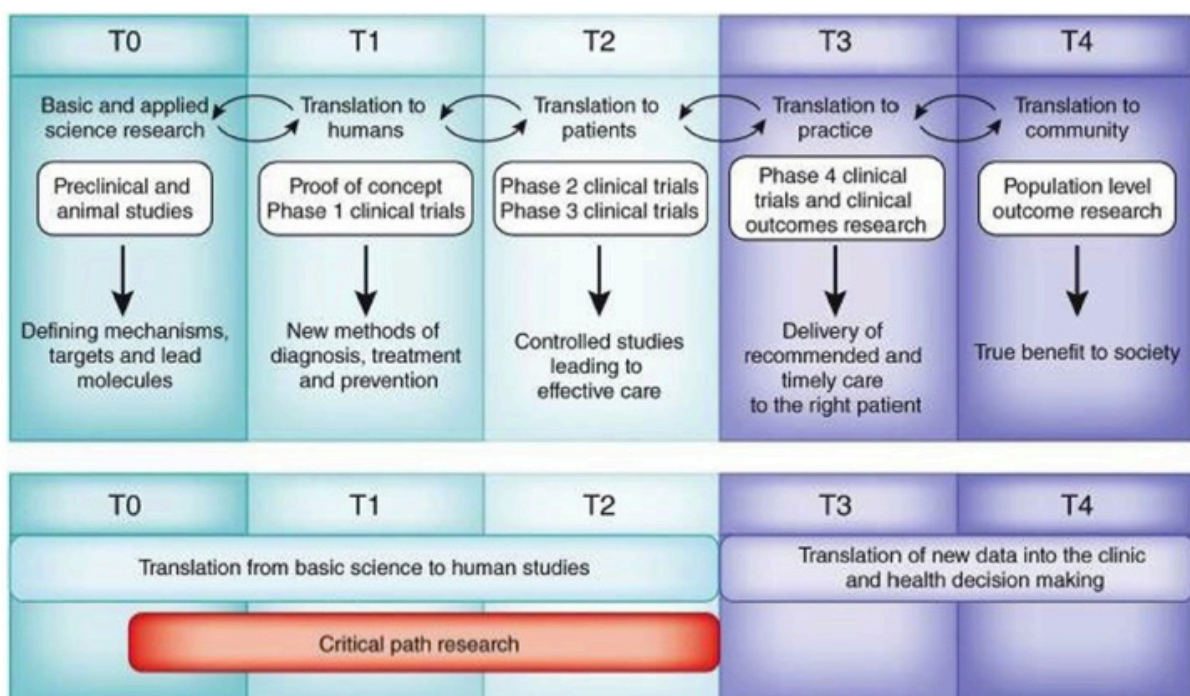
**B. What are the goals of translational research?**

Translational research was developed to bridge the clinical translation gap by tackling the root causes of the gap’s origin. Translational research aims at 1) speeding up translation of breakthrough discoveries to the patient’s bedside, 2) enhancing the success rate of promising drugs in clinical trials, and 3) ensuring proper dissemination to the general public, policy makers, regulatory and funding agencies.

### C. The translational research model and its challenges.

Translation research is often described as an organic, iterative process requiring continuous interactive feedback between different disciplines. It includes 3 main actions: discover, develop and disseminate.

The translational research model is broken down into 5 different stages (T0 to T4) presented in Figure 2.



**Figure 2.** Operational challenges for translational research and medicine

Translational research includes 5 stages. T0 represents the fundamental studies and applied research that define cellular mechanisms, their relationship to disease and, consequently, the identification of therapeutic targets and methods of treatment (new molecular entities); T1, first phase 1 studies in humans that aim to define proof of mechanism and proof of concept; T2, phase 2 and 3 clinical trials that are necessary for the approval of a therapeutic agent for clinical use; T3, phase 4 clinical trials that are associated with optimizing the therapeutic use of an agent in clinical practice; and T4, the outcome or comparative effectiveness research that serves to determine the ultimate utility and cost effectiveness of an agent relative to others currently in use. From (Blumberg et al. 2012).

Some authors used a simplified 2 translational blocks model (Woolf 2008):

- **T1 research** represents the translation from basic science to human studies (includes stages T0 to T2 in Figure 4). At this stage, the main challenges are related to the

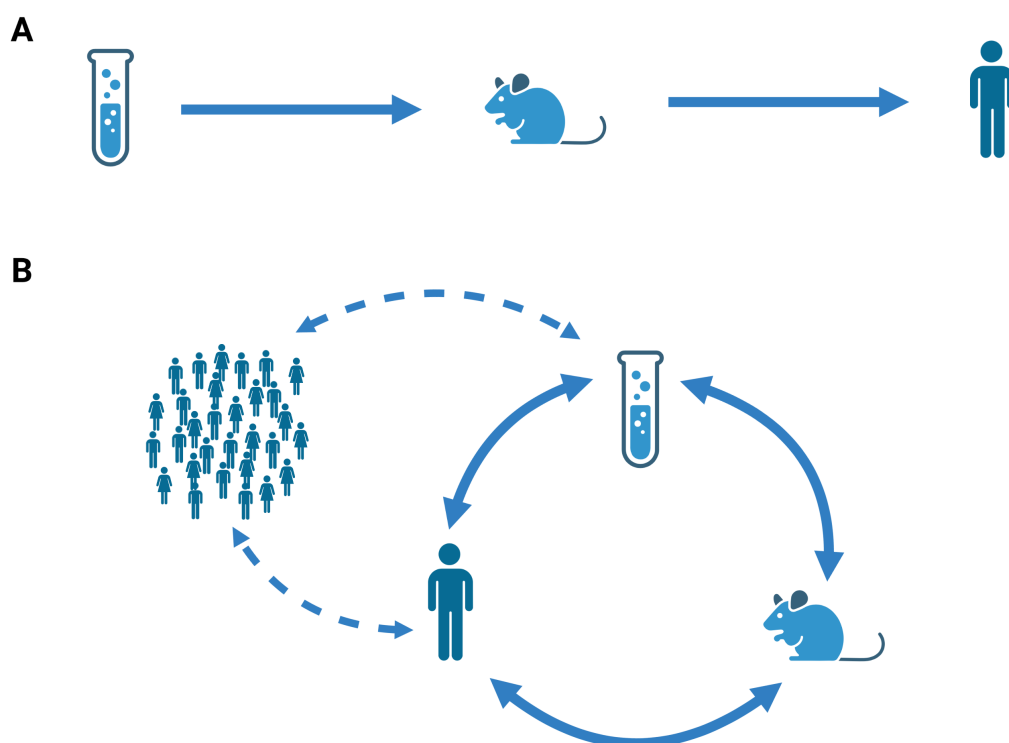
initial hypothesis' clinical relevance, the interactions between many disciplines, pre-clinical models limits (especially the poor predictive value of animal models in term of efficacy and product safety in humans) and regulatory and funding issues (Sabroe et al. 2007; Seyhan 2019; Garner et al. 2017).

- **T2 research** is described as the translation of clinical studies results into everyday clinical practice and health decision-making (includes stages T3 and T4 in figure 4). Its aim is to ensure that the discovery will be widely disseminated and ready to use in everyday healthcare decision-making. Despite the tremendous importance of T2 research, it is often overshadowed by T1 research in terms of recognition, interest and funding. Its specific challenges include human behaviors and organizational inertia, limits related to existing resources and infrastructure. However, the main challenge is the exposure of the new discovery (drug, diagnosis test, device...etc.) to the general population, a considerably less controlled environment compared to the preclinical and clinical translational block.

Both T1 and T2 research present their own challenges creating specific gaps which taken together contribute to the clinical translation gap, or the so-called “Valley of death” (Fernandez-Moure 2016).

T1 and T2 research share in common the requirement for a bilateral process with back and forth movement of knowledge between bench and bedside to generate new research hypotheses, refine experiments (both animal and human) for the next iteration, and support the process as a whole in order to deliver breakthrough discoveries to the general population (Ledford 2008; Seyhan 2019; Sabroe et al. 2007). This workflow represents an important paradigm shift from the previous one where research was a linear one-way process with the animal model being a gateway to access clinical trials (Figure 3).





**Figure 3.** Old vs. new translational research workflows

**A.** Old one-way linear biomedical research model, where animal experiments (mouse) were a gateway between basic science and clinical research. **B.** New translational research model: more complex and holistic network, with multiple lines of evidence to refine objectives and target relevance in order to increase the chance of successful drug discovery. (Adapted from Sabroe 2007). Plain double arrows represent T1 research, with back-and-forth movements of knowledge and hypotheses between different stages. The human symbol represents both the use of human cells and tissues (healthy and diseased) and clinical trials. Dashed double arrows represent knowledge translation with T2 research. Figure created with [BioRender.com](https://BioRender.com).

#### D. What are the determinants of translational research?

Translational Research's cornerstones include a multidisciplinary approach (Seyhan 2019), complex and holistic translational models (Sabroe et al. 2007; Garner et al. 2017), and knowledge translation supported by implementation science (Woolf 2008).

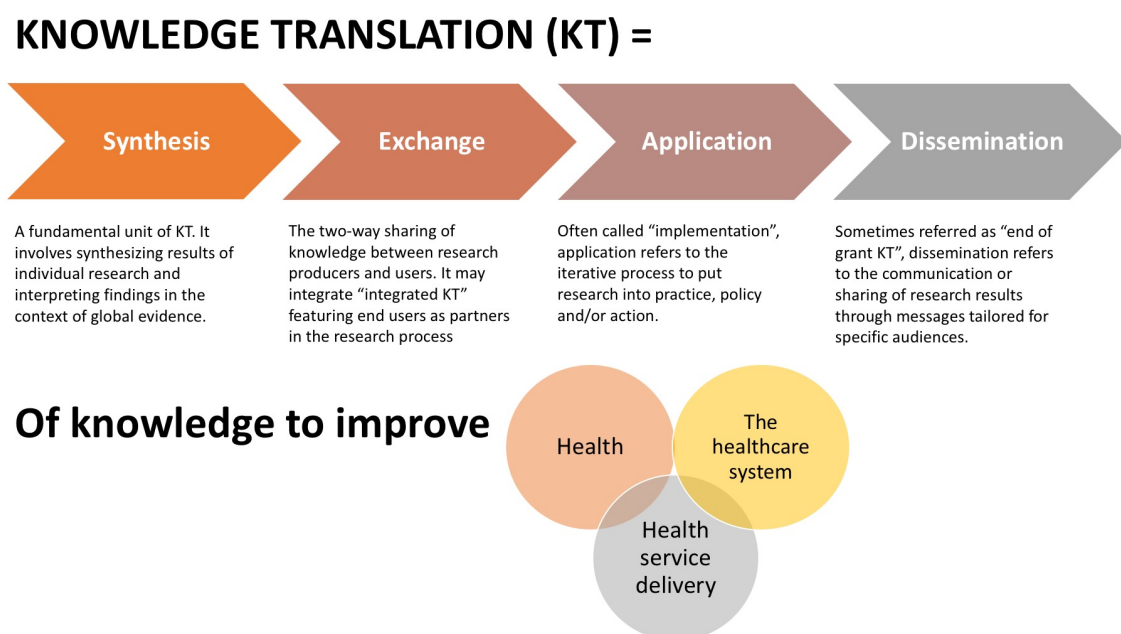
As translational research is a complex process requiring many skillsets, it requires the development of a **multidisciplinary task force** including scientists and clinicians, methodologists, computer scientists, bioengineers and industry experts, as well as experts

in knowledge translation and implementation science. In addition to offering the skillsets necessary for the study's success, multidisciplinary task force allows the lifting of silos between basic scientists and clinician scientists by restoring communication and collaboration with continuous interactive feedback between both sides (Fernandez-Moure 2016; Seyhan 2019).

To overcome its potential pitfalls and limitations, translational research scientists must abandon the traditional linear research model (figure 3) to develop more **complex and holistic models**. This effort starts with ensuring that clinical relevance is taken into account when generating a research hypothesis (Garner et al. 2017). Data or unusual findings from clinical trials (both successful and unsuccessful) represent an important source for hypothesis generation, highlighting that the two-way process (bench-to-bedside and bedside-to-bench) is a critical part of translational research (Ledford 2008). Translational scientists should not only recognize the richness of individual diversity, but also actively study it (and even introduce variability in the model in a controlled manner) to understand variation in drug response or safety profile (Garner et al. 2017; Sabroe et al. 2007). These complex models should also integrate studies (e.g., omics approaches, complex tissue models...etc.) on primary human cells and tissues from healthy and diseased subjects to further improve findings' predictive value and their relevance in human biology as well as our understanding of potential obstacles for clinical translation (Sabroe et al. 2007).

The last main determinant of translational research is **knowledge translation**, which is the core of T2 research. Failure to translate research into practice and policy is a common finding in clinical and health services research, with an estimated delay of 17 to 20 years to get clinical innovation into practice and with less than 50% of clinical innovations ever reaching general application (Bauer and Kirchner 2020). The aim of knowledge translation is

to make sure that stakeholders (knowledge users e.g., health professionals, patients and families, researchers, policy makers, industry...etc.) are aware of, and use research evidence to improve their health and healthcare decision-making (Grimshaw et al. 2012). Knowledge translation includes the following activities presented in Figure 4 (Canadian Institutes of Health Research 2012).



**Figure 4.** Overview of the knowledge translation process  
 Knowledge translation (KT) include four different activities to support and improve individual and public health as well as the health care system. Adapted from Kathryn Sibley, CHI KT platform (<https://medium.com/knowledgenudge/what-we-mean-when-we-say-knowledge-translation-1f81d57d5143>).

Implementation science is born from the need to help and support knowledge translation. Implementation science specifically focuses on how to improve knowledge translation by examining the factors, processes and strategies that influence successful integration of evidence-based intervention in practice, and is defined as “the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services” (Bauer et al. 2015).

Implementation science typically includes 4 stages (“A Practice Guide to Supporting Implementation | NIRN” n.d.):

- Exploration: Assessment for implementation needs and capacity and identification of barriers and facilitators.
- Installation: Plan and preparation for implementation (resources acquisition, staff training)
- Initial implementation: A test to monitor and evaluate the implementation and outputs used to refine the implementation strategy.
- Full implementation: Sustain and scale the implementation across different settings.

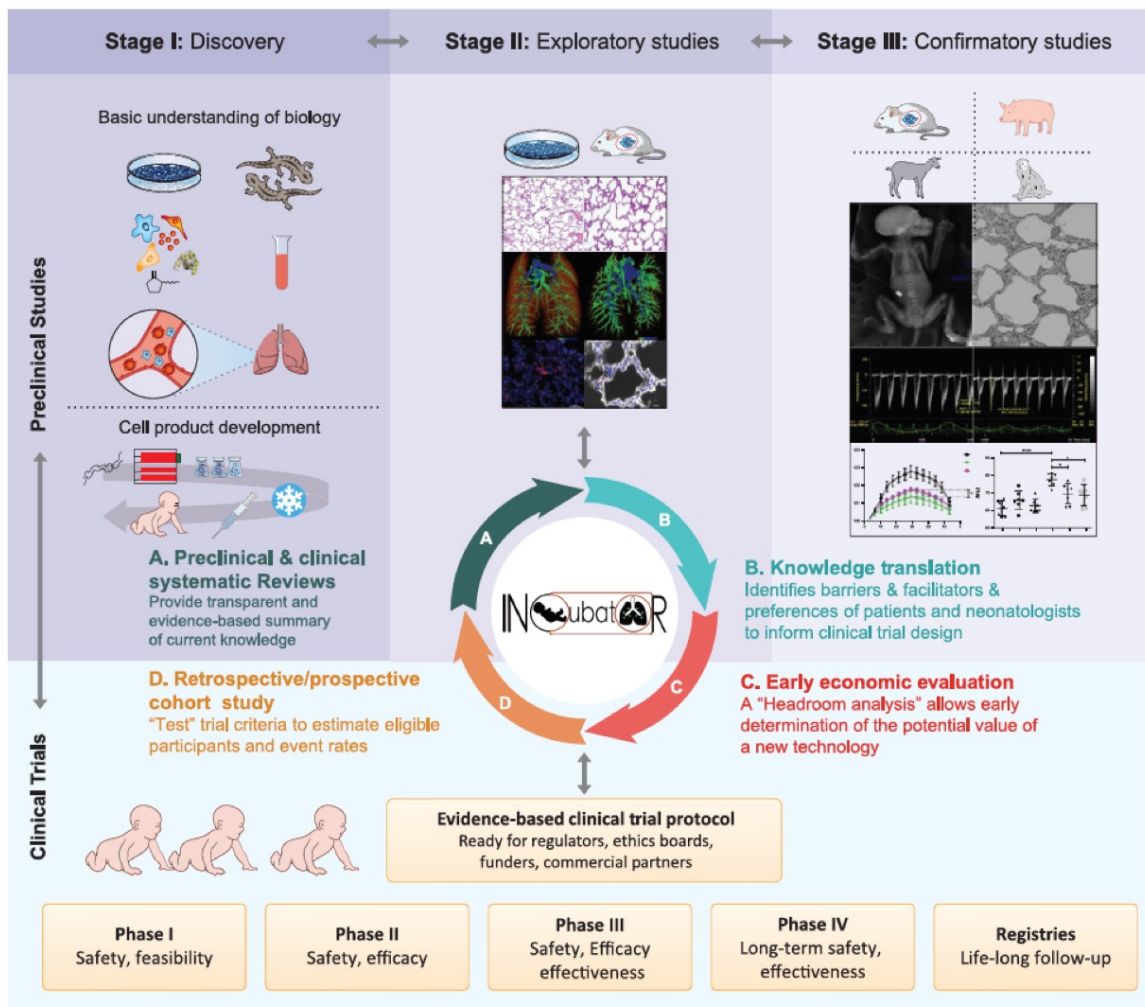
Therefore, knowledge users identification and involvement, as well as the assessment of the likely barriers and facilitators are critical to inform the choice of knowledge translation strategy and further improve knowledge transfer success (Grimshaw et al. 2012).

Finally, as translation research is time and resources consuming, these determinants must be supported by a shift in education, policy and funding. Dedicated education programs (MD-PhD programs) have been developed to prepare for careers of physician-scientists and support the creation of a new common cultural and academic identities for both clinician and basic researchers (Roberts et al. 2012). Translational research is a long (drug discovery and development can take up to 20 years), and expensive (some will argue not profitable) process with a high attrition rate in clinical trials. Therefore, measuring its outcomes is challenging and the metrics used to assess achievement should be far more complex than publication rank. Policy makers and funding agencies should develop incentives for academics and government institution to engage with industry in research collaboration and

partnership, as well as developing new metrics for funding attribution and career advancement (Fernandez-Moure 2016).

### **3. Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research (INCuBAToR): optimizing translational research determinants to improve clinical translation of cell therapy for bronchopulmonary dysplasia (BPD)**

To illustrate and give concrete examples for translational research determinants, we present here the INCuBAToR concept. The INCuBAToR was developed to enhance transparency, rigor and success in clinical translation in the field of cell therapy for BPD. We proposed a combination of an evidence-based approach and risk mitigation strategies developed by a multidisciplinary team. This concept relies on a comprehensive framework associating 1) knowledge translation strategies (integrated knowledge translation), 2) knowledge synthesis (systematic review and meta-analysis), 3) economic evaluation, 4) identification of potential reluctances and how to alleviate them, and 5) a workflow based on retrospective and prospective studies to refine study design and eligibility criteria. An overview of the INCuBAToR project framework is presented in Figure 5, and the article in the annex.



**Figure 5.** Overview of INCuBAToR framework

The INCuBAToR project proposed risk mitigation strategies to improve clinical translation of cell-based therapy for BPD patients. The INCuBAToR concept relies on an evidence-based approach with 4 main pillars: **A.** preclinical and clinical systematic reviews: this knowledge synthesis will help to summarize current evidence and level of evidence to support the study rationale and identify knowledge gaps; **B.** Knowledge translation: identification of the key stakeholders (end users) as well as barriers and facilitators to further support dissemination and implementation of the future findings; **C.** Early economic evaluation to ensure the project “viability” from an economical perspective; and **D.** retrospective/prospective cohort studies to inform clinical trial design development and ensure its feasibility. From (Thébaud et al. 2021).

This framework can be applied to any disease, as most of the biomedical research fields share the same pitfalls in term of translational potential: lack of rigorous methodology, absence of evidence synthesis to support study design development, failure to acknowledge and address ethics, regulatory and economical concerns, lack of knowledge of the translation strategies to support findings dissemination and implementation.

To further support clinical translation in neonatal lung research, we describe in this thesis three different projects that 1) provide a single-cell transcriptomic approach to enhance our understanding of normal human lung development (single-nuclei RNA sequencing), 2) describe the value of complex model derived from human lung tissue (precision-cut lung slices) to support the clinical relevance of basic research findings, and 3) combine a consensus-building method (Delphi method) with a knowledge translation strategy (integrated knowledge translation approach) to improve the reproducibility and transparency in mesenchymal stromal cell research.

## **II. Insights into transcriptomics during human lung development at a single-cell resolution**

### **1. Human lung development**

The human lung is a complex organ. While its main function is to facilitate gas exchange (oxygenation and decarboxylation of the blood), because of its exposure to external environment (air), the lung also integrates multiple other functions. These include innate immunity, mucus secretion and clearance with fluid and electrolyte transport, alveolar tension surface maintenance, environment sensing, and complex repair processes. To achieve all these tasks, up to 40 different resident cell types interact together in adult human lungs to maintain lung homeostasis (Franks et al. 2008). Moreover, to perform optimal gas exchange, the lung requires two intertwined and highly branched networks (airways and lung vasculature) and a close connection between the alveolar epithelium and the capillary endothelium to create the gas exchange surface.

Understanding and describing how this variety of cell populations contribute at a molecular and cellular level to the development of such a complex organ is critical to provide insights on lung homeostasis and regeneration, disease development, and to find new therapeutic targets (Herriges and Morrisey 2014; Ubags et al. 2020).

#### **A. Different stages of human lung development**

Human lung development is classically described in 5 stages: embryonic, pseudoglandular, canalicular, saccular and alveolar (Kotecha 2000; Morrisey and Hogan 2010; Schittny 2017). These stages were described according to histological and morphological changes during human fetal development (P H Burri 1984). Based on different observations, the transition



times between the different stages vary and overlap (Peter H. Burri 2006). This variation can be explained by the asynchronous development of the lung with the cranial regions developing faster than the caudal regions (Ernst, Ruchelli, and Huff 2011). The different stages and transition times are presented in the Table 1.

Prenatal period	
Embryonic stage	1-7 weeks
Fetal period	5 weeks to birth
Pseudoglandular stage	5-17 weeks
Canalicular stage	16-26 weeks
Saccular stage	24-38 weeks
Postnatal period	
Alveolar stage	36 weeks to 1-2 years
Microvascular maturation	Birth to 2-3 years
Late alveolarization	3-5 years to young adult age
Ageing phenomena	~ 35 years to death

**Table 1.** Timing of human lung development, growth and ageing  
From (Peter H. Burri 2006).

During the **embryonic stage** (gestational weeks 4 to 7 in human), the two primary lung buds form by outpouching of the ventral wall of the primitive foregut endoderm. At the same time, the primitive foregut tube divides into the esophagus (dorsal foregut) and the trachea (ventral foregut) which connects to the lung buds (Cardoso and Lü 2006). The primary lung buds are surrounded by mesenchyme and will undergo the process of branching morphogenesis to form the tracheobronchial tree. Branching morphogenesis uses repetitive and dichotomous branching patterns (domain branching, planar bifurcation and orthogonal bifurcation) and is strictly controlled via epithelial-mesenchymal cross-talk (Warburton 2017). In parallel, lung vasculature derives from the 6<sup>th</sup> aortic arch and accompanies the developing airways. A plexus connected to the aortic sac and the left

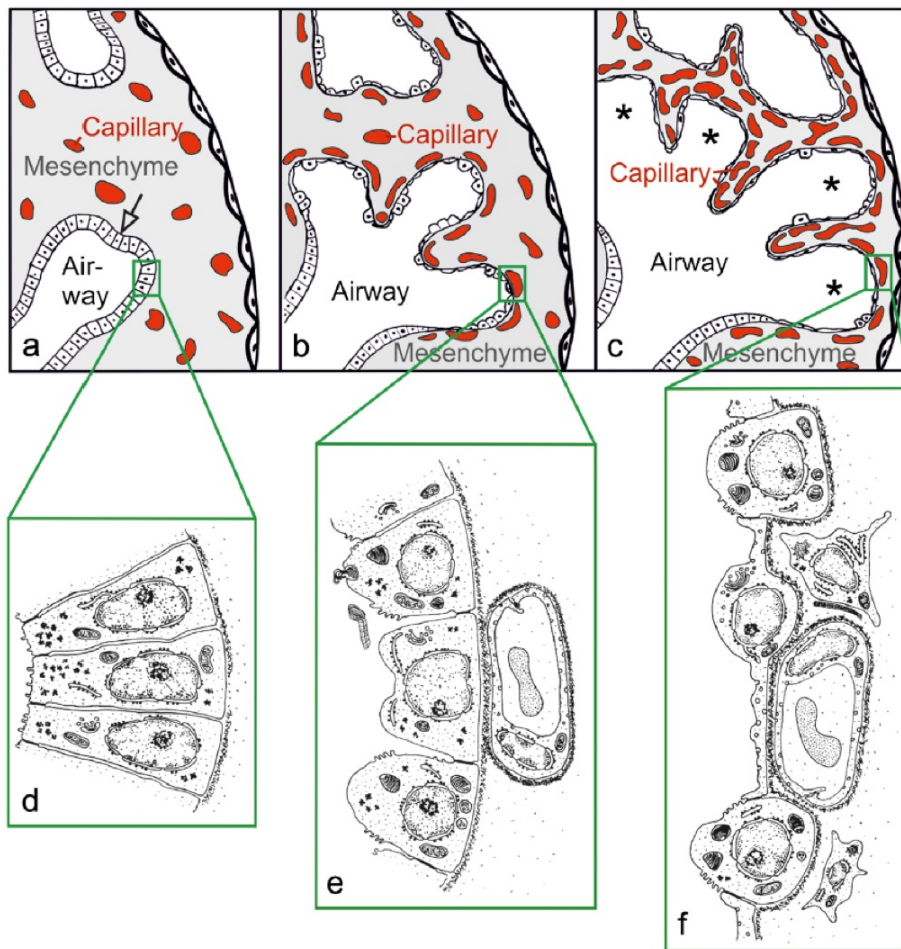
atrium is formed in the mesenchyme surrounding the lung buds by vasculogenesis (Schittny 2017).

In the **pseudoglandular stage** (gestational weeks 5 to 17 in human), the progressive branching of bronchial buds gives rise to increasing generations of airways, with the formation of pre-acinar airways (i.e., airways to the level of the terminal bronchioles) occurring at the end of this stage. The future airways are embedded in the mesenchyme and surrounded by a loose capillary network. Around 13 weeks of gestation, the epithelium starts to organize and differentiate in a proximal-to-distal pattern. Ciliated cells, goblet cells and basal cells are described in the proximal airways, while the more distal airways are lined by simple columnar epithelium. In addition, airway smooth muscle cells arise from the mesenchyme and surround the airways. These cells will start to have spontaneous contractions which generate, in combination with pulmonary fluid secretion, critical biomechanical forces for the developing lung. The lung vasculature continues to develop alongside sequential generations of branching airways by vasculogenesis. Lymphatics start to develop within the lung at around 10 weeks of gestation (Ernst, Ruchelli, and Huff 2011).

The **canalicular stage** (gestational weeks 16 to 26 in human) is characterized by the initiation of development and vascularization of the airspaces. Airways continue to develop by further peripheral branching and branch lengthening leading to the formation of canaliculi. Distal airspaces become wider at the expense of the mesenchyme, and the epithelium transitions from a simple cuboidal to a flatter epithelium, a sign of the initiation of mature pneumocytes differentiation. These newly formed acini include the respiratory bronchioles, alveolar ducts and terminal sacs (Ernst, Ruchelli, and Huff 2011). Meanwhile, capillaries grow and organize into a primitive capillary network, closer to the distal airways and protrude between airspace epithelial cells. As a result of these changes in epithelium

and capillaries during the canalicular stage, we observe the formation of the first thin air-blood barrier portions (P H Burri 1984). At this stage of the development, it is possible to distinguish conducting and respiratory airways based on their epithelial morphologies (Schittny 2017).

In the **saccular stage** (gestational weeks 24 to 38 in human), the process of alveolar-capillary interface maturation begins. Vascularized saccules develop, delimited by the formation of primary septa. These primary septa are immature inter-airspaces septa containing a double-layered capillary network. The increasing numbers of saccules leads to a marked decrease in interstitial tissue. In the primary septa interstitial tissue, smooth muscle cell precursors produce fibrous networks of elastin and collagen fibers that will be critical for the secondary septation during the alveolar stage (P H Burri 1984; Schittny 2017). Morphological changes during the pseudoglandular, canalicular and saccular stages are presented in Figure 6.



**Figure 6.** Morphological changes in human developing lung during the pseudoglandular, canalicular and sacular stages

During the pseudoglandular stage (a), future airways are embedded in the mesenchyme, surrounded by a loose capillary network and continue to branch repeatedly to give rise to increasing generations of airways (arrow, branching morphogenesis). Distal airways are lined by simple columnar epithelium (d). The canalicular stage (b) is characterized by airspaces development and the vascularization of these airspaces: airways start to widen, epithelium transitions from a simple cuboidal form to a flatter epithelium (initiation of mature pneumocytes differentiation), and the capillaries grow and organize in a primitive capillary network, closer to the distal airways (e). In the sacular stage (c) begins the process of the alveolar-capillary interface maturation with the development of the saccules (denoted by \*) and the formation of thick immature interspaces septa with a close proximity between capillaries and airways epithelium (f). From (Schittny 2017).

Finally, the **alveolar stage** starts in utero (around 36 weeks of gestation in human) and continues after birth, likely for as long as the lung continues to grow (young adulthood). During this phase, the gas exchange surface increases with secondary septation (formation of the alveoli) and matures with thinner septa and microvasculature maturation (fusion of the double layer into a sheetlike capillary layer) to improve gas exchange. During classical

(or bulk) alveolarization (from 36 weeks to approximately 3 years of age), secondary septa are formed by the upfolding of one of the capillary layers, driven by the elastin/collagen network. Additional evidence supports the existence of another alveolarization process called continued alveolarization (from approximately 2 years of age to young adulthood), during which secondary septa can rise from already mature septa and be supported by neovascularization (Schittny 2017; Herring et al. 2014; Narayanan et al. 2012).

Unravelling the complex underlying molecular mechanisms and crosstalk between different cell types that happen *in-utero* is crucial for our understanding of how perturbations at different stages of lung development can lead to congenital malformation or prematurity-associated lung diseases. Examples of congenital lung diseases in relation to lung developmental stage are presented in Table 2.

Lung development stages	Congenital malformations	Prematurity-associated lung diseases
<b>Embryonic</b>	<ul style="list-style-type: none"> <li>– Pulmonary agenesis</li> <li>– Tracheal or laryngeal agenesis or stenosis</li> <li>– Tracheo-bronchomalacia</li> <li>– Bronchial malformation</li> <li>– Ectopic lobes</li> <li>– Arterio-venous malformation</li> <li>– Congenital lobar cysts</li> </ul>	
<b>Pseudoglandular</b>	<ul style="list-style-type: none"> <li>– Cystic adenomatoid malformation</li> <li>– Pulmonary sequestration</li> <li>– Lung hypoplasia</li> <li>– Lung cysts</li> <li>– Congenital pulmonary lymphangiectasia</li> <li>– Congenital diaphragmatic hernia</li> </ul>	
<b>Canalicular</b>	<ul style="list-style-type: none"> <li>– Lung hypoplasia</li> <li>– Acinar dysplasia</li> </ul>	<ul style="list-style-type: none"> <li>– Respiratory distress syndrome</li> </ul>
<b>Saccular / Alveolar</b>	<ul style="list-style-type: none"> <li>– Pulmonary hypoplasia</li> <li>– Acinar dysplasia</li> <li>– Alveolar capillary dysplasia</li> </ul>	<ul style="list-style-type: none"> <li>– Respiratory distress syndrome</li> <li>– Chronic lung disease of the prematurity</li> </ul>

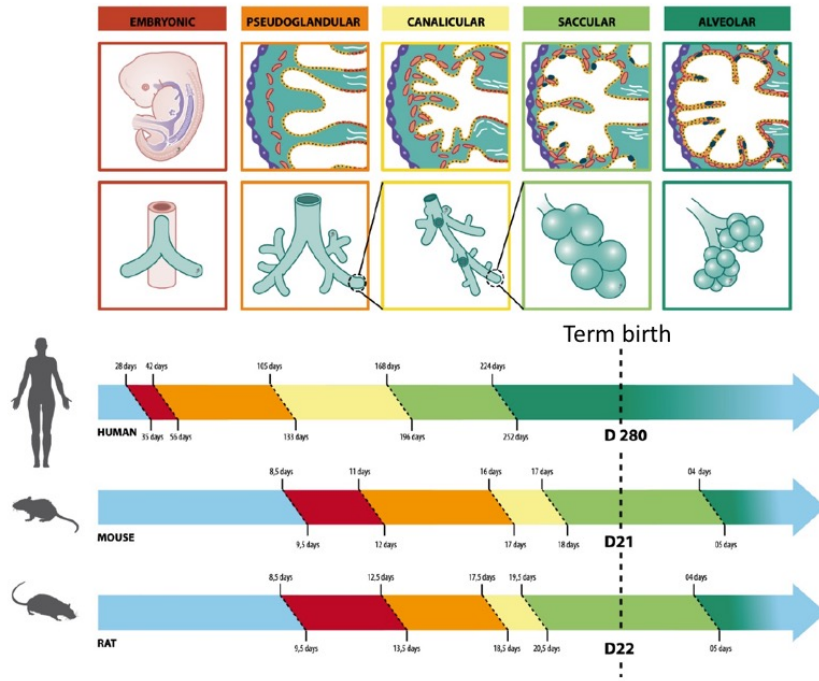
**Table 2.** Congenital and acquired (prematurity) lung diseases and the stage of their development  
Adapted from (Kotecha 2000).

## **B. Understanding the human lung development: insights on molecular mechanisms from animal models and human tissue-based models**

### a. Animal (rodents) models to study human lung development

#### *i. Rationale for animal use to study human lung development*

Rodent models have been extensively used to develop knowledge and understanding of lung development and diseases, and to find new therapeutic targets (Cardoso and Lü 2006; Morrisey and Hogan 2010). A part of the rationale to use rodent models is that mice, rats and human undergo the same stages during lung development but differ with respect to the duration of each stage and its temporal relationship to gestational age (Figure 7). Additional practical reasons to use rodents compared to larger animals include small animal size, short life cycle, ease and low-cost husbandry system, short gestation, large litters, availability of genetically modified animals, and availability of large panel of rodent antibodies.



Stage	Gestational age		
	Mouse (days)	Rat (days)	Human (weeks)
Embryonic	E9-E11.5	E8-E13	1-7
Pseudoglandular	E11.5-E16.5	E13-E18	5-17
Canalicular	E16.5-E17.5	E18-E20	16-29
Saccular	E17.5-P5	E20-P5	24-38
Alveolar	P5-P28	P5-P30	36-adolescence

**Figure 7.** Lung development stages: comparison between rodents and human  
 E: embryonic day, P postnatal day. Adapted from (Salaets et al. 2017; Nardiello, Mižíková, and Morty 2017).

Therefore rodent, and particularly mouse models have been extensively used to study and describe lung development, using different gestational ages and postnatal time points to describe morphological and molecular changes. In addition, these observational data can be confirmed by gain-and-loss of function experiments in genetically modified mice and cell lineage tracing experiments. Rodent models have also been used to study impaired lung development in both congenital lung diseases (e.g., nitrofen model of congenital diaphragmatic hernia) and acquired lung diseases due to prematurity (induction of BPD

phenotype by exposing the animals to environmental insults similar to those experienced by preterm infants such as hyperoxia, inflammation, mechanical ventilation or growth restriction) as mice are born in the saccular stage which is the same stage during which most of the preterm infants developing BPD are born in. These disease models can be used to highlight disease mechanisms and to assess potential therapeutics to prevent or treat BPD (Berger and Bhandari 2014; Giusto et al. 2021; Nardiello, Mižíková, and Morty 2017).

*ii. Insight from animal (rodents) models for human lung development*

Most of our understanding on molecular mechanism underlying lung development come from mouse studies. Branching morphogenesis and airways development have been extensively studied in mice, with a particular focus on the description of major pathways and key players such as bone morphogenetic protein (BMP), fibroblast growth factors (Fgf), Wnt, Sonic hedgehog (Shh), and retinoic acid (Cardoso and Lü 2006; Morrisey and Hogan 2010; Swarr and Morrisey 2015). Table 3 summarizes some of the data from gain-and-loss of function experiments in mice. These findings are important to describe the role of different signaling molecules in early lung development and to effectively reproduce the desired disease phenotype (Cardoso and Lü 2006). These experiments have improved our knowledge and understanding of signaling networks in the developing lung, especially those involved in the branching morphogenesis.



Gene symbol	Gene name	Expression pattern	Phenotype
<b>Signaling molecule</b>			
Fgf10	Fibroblast growth factor 10	Mesenchyme	Lung agenesis
Fgfr2b	Fibroblast growth factor receptor 2b	Epithelium	Lung agenesis
Shh	Sonic hedgehog	Epithelium	Impaired branching
Wnt7b	Wingless-related MMTV integration site 7B	Epithelium	Vascular defect, reduced mesenchyme
Wnt5a	Wingless-related MMTV integration site 5A	Epithelium and mesenchyme	Increased branching, tracheal defect
Egfr	Epidermal growth factor receptor	Epithelium and mesenchyme	Impaired branching and deficient alveolization
<b>Transcription factor</b>			
Foxa1 Foxa2	Forkhead box A1/A2	Epithelium	Impaired branching, reduced smooth muscle
Sox11	SRY-box-containing gene 11	Epithelium	Hypoplastic lung
Rarb Rara	Retinoic acid receptor alpha/beta	Epithelium and mesenchyme	Left lung agenesis and right lung hypoplasia
<b>Other</b>			
Eln	Elastin	Mesenchyme	Deficient alveolization
Lmbn1	Lamin B1	Epithelium and mesenchyme	Deficient alveolization

**Table 3.** Examples of mutations in mouse and their reported phenotype.

Gain-and-loss of function studies in genetically modified animals have helped to identify key factors in lung development. Adapted from (Cardoso and Lü 2006).

On the other hand, little is known about mesenchyme and vasculature development in the lung and the inter-cellular crosstalk between these different cell types in the lung. Mesenchymal cell diversity and functions are poorly understood and described, whereas vascular and neuronal networks begin to appear as key players in lung development. Endothelium seems to play an important role in cell-cell signaling and driving the epithelial

cells growth and differentiation during lung development (Kina et al. 2021). Moreover, evidence suggests that local parasympathetic innervation is required for airway morphogenesis, with disruption of bronchial neurons leading to abolished airway branching (Bower et al. 2014). In addition, roles of biomechanical forces and cell-extracellular matrix (ECM) interactions in developing lung need to be unraveled, with very few data from mice and humans currently available (Miller and Spence 2017).

*iii. Limitations and pitfalls of mouse models to study human lung development*

While being versatile and useful, using mouse models to understand human lung development have several limitations. Mouse and human exhibit multiple differences in anatomy, histology, physiology, molecular regulation, cells composition and organization of their respiratory system. For example, goblet cells are less abundant in mouse lung (less than 1% of the extrapulmonary epithelium) and mouse intrapulmonary airways lack submucosal glands and cartilage. Additionally, mice don't have respiratory bronchioles, and pulmonary neuroendocrine cells in human airways are mostly isolated and scattered when compared to mice (Morrisey and Hogan 2010; Swarr and Morrisey 2015). Common anatomical and histological differences between mouse and human respiratory system are presented in Table 4. When implementing an animal model to induce an insult to the lung development using environmental insults experienced by preterm infants, differences in lung physiology between mouse and human often pose a major limitation. While rodents are born at the same lung developmental stage as human preterm infants (saccular stage), these animals are born surfactant-sufficient, and therefore have a lung mature enough to

ensure proper gas exchange with normal breathing and adaptation to effort (Berger and Bhandari 2014).

Feature	Rodent	Human
<b>Gross anatomy</b>		
Lung lobes	Four right, one left	Three right, two left
Cartilaginous airways	Trachea and primary bronchi (i.e., extrapulmonary only)	Trachea through small intra-segmental bronchi in lung
<b>Histology</b>		
Respiratory bronchiole	Absent to rare in mice. In rats absent early in life (detectable by 3 weeks postnatally)	Present
Airway branching pattern	Monopodial	Dichotomous
<b>Airway Epithelial Cells (Relative Numbers)</b>		
Trachea	Non ciliated > ciliated > basal. Epithelium thinner than humans (rats>mice)	Ciliated > basal > non ciliated and mucous Epithelium thickness decreases from proximal trachea through bronchi
Primary bronchi	Non ciliated > ciliated > basal	Ciliated > non ciliated > mucous > basal
Terminal bronchioles	Non ciliated > ciliated	Ciliated > non ciliated
Alveoli	Rats: type II cells > type I Relative cell numbers not well defined in mice	Type II cells > Type I cells Type I are very thin but have a much greater cell volume than type II cells

**Table 4.** Species differences between mouse and human respiratory system  
From (Meyerholz et al. 2018).

In addition, several significant cellular and molecular differences in the underlying mechanism of lung development between mouse and human hinder prediction on how findings based on animal studies can be applied to human lung development and disease (Danopoulos, Shiosaki, and Al Alam 2019; Nikolić et al. 2017; Kho et al. 2010). In developing mouse lung, proximal epithelium is *Sox2*<sup>+</sup> and distal epithelium is *Sox9*<sup>+</sup>, whereas human distal epithelial cells express both *SOX2*<sup>+</sup> and *SOX9*<sup>+</sup> during the pseudoglandular stage and become *SOX2*<sup>-</sup>/*SOX9*<sup>+</sup> during the canalicular stage (Nikolić et al. 2017). These *SOX2*<sup>+</sup>/*SOX9*<sup>+</sup> cells are required for proper lung branching, suggesting an important role of *SOX2* in human lung branching morphogenesis (Danopoulos et al. 2018). Additional transcriptomic analysis of the *SOX2*<sup>+</sup>/*SOX9*<sup>+</sup> epithelial cells showed that, even if 96% of genes expressed by epithelial tips in human were also expressed by epithelial tips in mice, an additional 348

unique genes were found in human lung epithelial tips only (Nikolić et al. 2017). A similar discrepancy can be found in the case of the *FGF10* expression pattern. *Fgf10* and its receptor *Fgfr2* are critical for branching morphogenesis in the mouse (absence of *Fgf10* or *Fgfr2* leads to lung agenesis). *Fgf10* and *Fgfr2* play an important role in the maintenance of the proximal (*Sox2*<sup>+</sup>) - to - distal (*Sox9*<sup>+</sup>) epithelium pattern in mice (Yuan et al. 2018). However, their role in the human developing lung is less clear. Although some evidence supports abnormal *FGF10* expression in BPD patients (Benjamin et al. 2007), there is no evidence of abnormal *FGF10* expression or signaling in human lung agenesis. Furthermore, spatial distribution of *FGF10* expression in the developing lung differs between mouse (*Fgf10* expression in distal mesenchyme close to the bud tips) and human (*FGF10* expressed throughout the lung parenchyma with some expression in the airways and the vascular smooth muscle cells) (Danopoulos, Shiosaki, and Al Alam 2019). Lastly, experiments on human lung organoids often use *FGF10* in addition to the culture medium to expand the organoids. However, removing *FGF10* from the medium did not impair the organoid growth or expression of distal tip markers *SOX2* and *SOX9* in organoids constructed using either, human fetal lung tissue or human pluripotent stem cells (Nikolić et al. 2017; Miller et al. 2018).

In summary, most of our current understanding of molecular mechanisms of human lung development derives from mouse experiments, whereas little evidence is acquired from studies of the human lung itself (Cardoso and Lü 2006; Swarr and Morrisey 2015). Importantly, morphology and molecular differences between mouse and human lungs suggest that certain aspects of the human lung development can only be studied using human tissue and cells (Nikolić, Sun, and Rawlins 2018).

b. Human lung organoids to study human lung development.

As stated above, studying human specific mechanism is essential for understanding diseases pathogenesis. Based on difference between mouse and human lung, the need to access human lung tissue and develop *in-vitro* models of human lung development is critical to provide molecular level analysis of cell-cell interactions in both normal and diseased states. Furthermore, these models can be used to discover potential therapeutic targets and develop platforms for large scale screening of drugs, environmental pollutants and other external factors.

iv. *The “old” in-vitro model using human lung derived cells.*

Immortalized human cells lines have been used for a long-time in respiratory research to overcome the rarity of human lung samples, especially from patients with rare diseases. In parallel, the development of novel techniques, such as air-liquid interface (ALI) cultures, has greatly improved the clinical relevance of airway epithelium culturing. ALI culture supports the *in-vitro* reconstitution of the pseudostratified epithelium where cell polarity and functions like cilia production and motility, and mucus secretion are respected. Therefore, ALI culture constitutes a versatile platform to perform mechanistic studies on airway epithelium barrier function in studies of cystic fibrosis, pathogen host interactions, and interaction with environment pollutants (Molina et al. 2015; Krunkosky et al. 2007; Upadhyay and Palmberg 2018). The major limitation for the use of immortalized human cell lines to generate complex *in-vitro* models of the human lung, is the immortalization process these cells undergo and how it affects cell physiology when compared to normal airway cells.

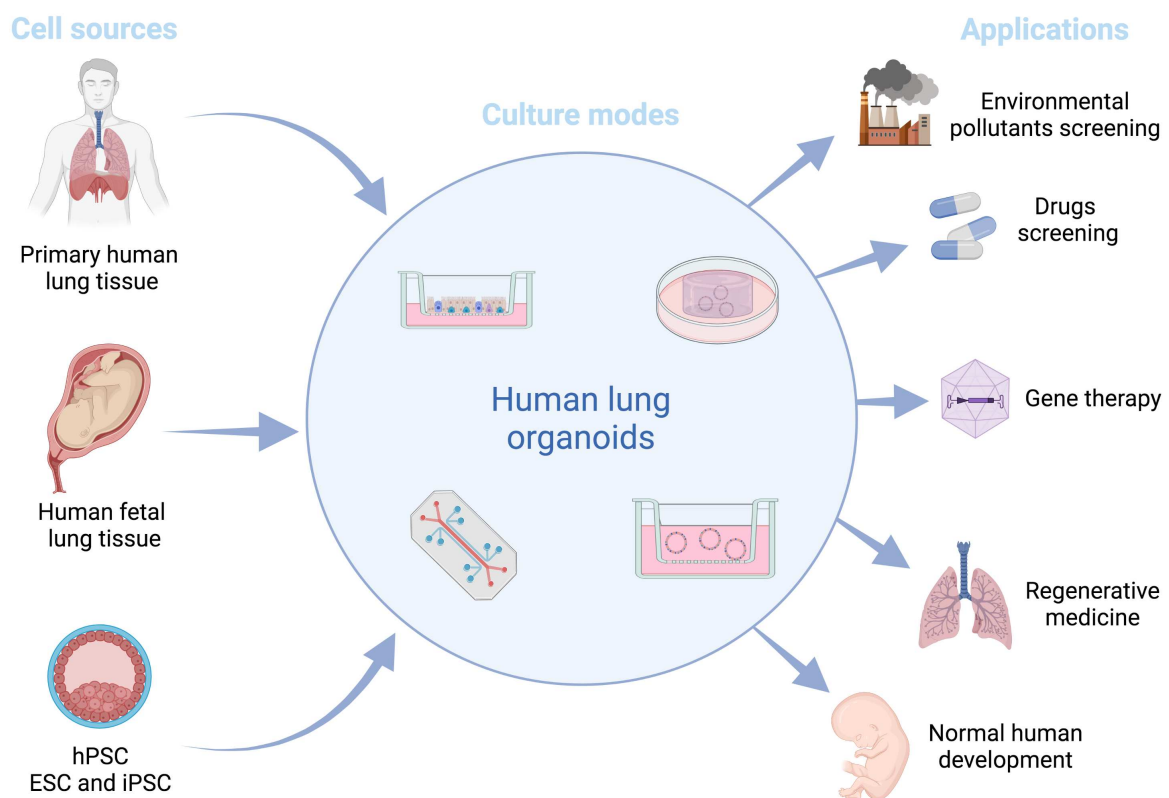
v. *Human lung organoids, principles and advantages.*

Important advances have been made with the development of complex organ-like *in-vitro* human models (organoids) which provide new opportunities to understand and study normal and diseased states. Organoids are complex three dimensional (3D) multi-cellular structures that exhibit some of the structural (including multiple cell types and cellular organization) and functional features of the native organ. Human lung organoids (HLO) are obtained by seeding spheroids obtained from cells from human primary lung tissue or human pluripotent stem cells (hPSCs) in an ECM. The ECM supports organoid growth and maturation. The most commonly used type of ECM is hydrogel, particularly Matrigel, which is popular due to its characteristics that are similar to basal membranes. In addition, seeding of other cell types, such as fibroblasts or endothelial cells in the hydrogel supports organoids structuration and maturation. HLO can be cultured for several months and have a complex and diverse epithelium including basal cells and ciliated cells. HLO can also exhibit supporting mesenchymal tissue with diverse mesenchymal cells (myofibroblasts, fibroblasts, and smooth muscle cells) which can be found in both, proximal conducting airway-like structures, as well as in the areas of alveolar cell types (Miller and Spence 2017; Aurora and Spence 2016). Different levels of the lung airways, such as tracheosphere, bronchosphere, bronchiolar organoids and alveolospheres, have already been successfully reproduced as lung organoids (Barkauskas et al. 2017; Nadkarni, Abed, and Draper 2016).

Further improvement to the HLO model has been made with the development of different culture modes including the use of biological (decellularized lungs) and bioengineered scaffolds (Miller and Spence 2017). Bioengineered scaffolds with a microfluidic system (also called "lung-on-a-chip"), which include both vascular and airway compartments, and can model mechanical forces of the lung such as stretch forces, represent a particularly

promising avenue for clinically more relevant human lung models (Huh et al. 2010). In addition, HLO are well suited to exploit CRISPR/Cas9 gene editing technology to identify important gene signatures in airway maturation or function (Barkauskas et al. 2017).

These *in-vitro* human lung models can be used to study normal lung development, crosstalk between the different components of the lung (epithelium, mesenchyme, vascular endothelium) or even be utilized to study molecular mechanisms and assess potential therapeutics for BPD (acquired impaired lung development due to preterm birth) (Aurora and Spence 2016; Nikolić, Sun, and Rawlins 2018). Figure 8 shows an overview of HLO, including cell sources, culture modes and main applications.



**Figure 8.** Human lung organoids overview

In this figure are shown the main cell sources to obtain human lung organoids (hPSC: human pluripotent stem cell, which include embryonic stem cell (ESC) and induced pluripotent stem cell (iPSC)), different culture mode (such as 3D, 3D ALI, lung-on-a-chip), and the main applications for human lung organoids. Human lung organoids can be used to study normal lung development or to serve as a platform for disease models. Adapted from (Bourguignon et al. 2020). Figure created with [BioRender.com](https://www.biorender.com).

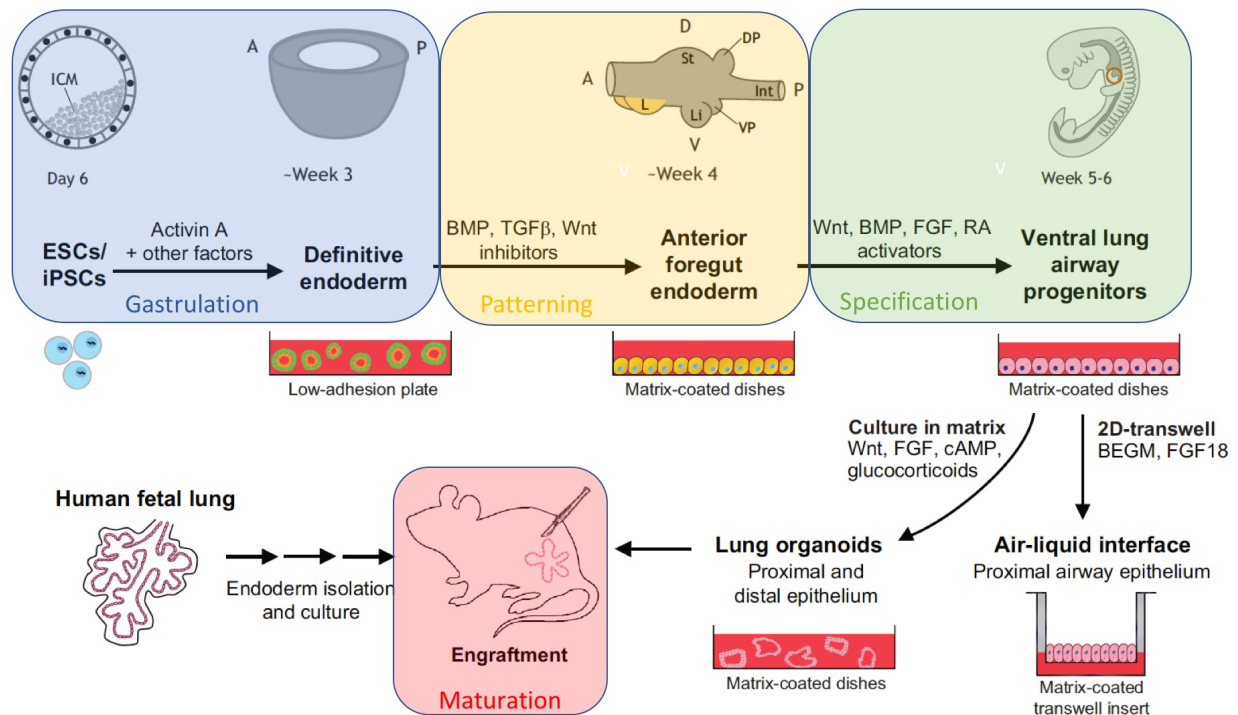
vi. *Cell sources for human lung organoids.*

HLO can be derived by differentiation of hPSCs (*de novo*) or from native tissue source (patient, human fetal lung sample...etc.) (Aurora and Spence 2016; Miller and Spence 2017; Nikolić, Sun, and Rawlins 2018).

HPSCs, which include human embryonic stem cells and induced pluripotent stem cells, need to be differentiated to obtain human lung organoids. This differentiation relies on maturing these hPSCs in a step-by-step fashion through sequential progenitor stages, mimicking as closely as possible normal embryonic development. This differentiation process is called “directed differentiation” and uses signaling cues identified through animal lung development studies to control *in-vitro* cell fate decision and mimic normal embryonic development. For lung epithelium, that includes induction of the definitive endoderm (gastrulation) followed by the induction of the anterior foregut endoderm (patterning) and induction of lung cell types (specification) (Dye, Miller, and Spence 2016). hPSCs-derived tissue has the ability to self-assemble *in-vitro* in spheroids which can be seeded on an ECM to obtain organoids. This directed differentiation process is presented in Figure 9.

A common feature of hPSCs-derived organoids is that they remain in a transcriptional and functional state similar to fetal human tissue (Dye, Miller, and Spence 2016). Therefore, strategies have been developed to address this limitation. It is now well accepted that further maturation of HLO (either derived from fetal lung tissue or hPSCs) can be achieved by a combination of *in-vitro* growth and subsequent *in vivo* engraftment into kidney capsule of immunocompromised mice. This strategy leads to functional and morphological maturation. (Barkauskas et al. 2017; Nikolić, Sun, and Rawlins 2018).





**Figure 9.** Overview of directed differentiation process and HLO maturation

The normal human embryonic development is represented at the top of the figure: ICM: inner cell mass of the blastocyst, A: anterior, D: dorsal, DP: dorsal pancreas, Int: intestine, L: lung, Li: liver, P: posterior, St: stomach, V: ventral, VP: ventral pancreas. Directed differentiation steps are represented in color boxes: gastrulation (blue) leading to definitive endoderm (SOX17<sup>+</sup>/FOXA2<sup>+</sup>), patterning (yellow) leading to anterior foregut endoderm (SOX2<sup>+</sup>/FOXA2<sup>+</sup>), and specification (green) leading to ventral lung airway progenitor (NKX2-1<sup>+</sup>/SOX2<sup>+</sup>/TUJ1<sup>+</sup>/PAX8<sup>+</sup>). The red box represents the additional step for maturation with engraftment of human lung organoids derived from hPSC or human fetal tissue in an immunocompromised mouse. Human pluripotent stem cells (hPSC) include embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC). Adapted from (Barkauskas et al. 2017; Nikolić, Sun, and Rawlins 2018).

The main limitation of the use of the native human tissue is the limited access to human fetal lung tissue, especially after 20 weeks of gestation, limiting the study of both normal *in-utero* lung development after this age, as well as prematurity-related lung diseases such as BPD. As neonatal intensive cares allow survival at 22-23 weeks of gestation (extremely low gestational age infants) and because this population of preterm infants is the most susceptible to severe BPD, studying human lung tissue at a maturation stage of 22 weeks and above is crucial to improve our knowledge and develop relevant therapeutics. One proposed approach suggests the use of human fetal lung tissue with further combined *in-*

*vitro* and *in-vivo* maturation as described above to reach the lung developmental stage of these extremely low gestational age infants (Nikolić, Sun, and Rawlins 2018).

*vii. Limitations of human lung organoids to study lung development.*

HPSCs-derived HLO allow the generation of patient specific models and overcome the barriers of human tissue availability, tissue quality and donor variability. Limitations of this model include the lack of a standardized differentiation protocol resulting in heterogenous cell populations and the immaturity of the obtained organoids. In addition, HLO mainly focus on one lung epithelium type with only few systems using organoids being able to integrate both proximal (airways) and distal (alveolar) epitheliums. Complex cell cross-talks are of critical importance in both lung development and homeostasis, and their dysfunctions can result in lung disease. The lack of interaction between organoid epithelial cells and the other lung cell types (i.e., endothelium, mesenchyme, immune, neural tissue) therefore significantly hinders the clinical relevance of this model.

Moreover, the biggest challenge for the development of these HLO is to validate their reliability against their *in-vivo* counterparts in human lung tissue. Improving our understanding of normal human lung development will help to fine-tune the human *in vitro* experimental systems and enhance their clinical relevance (Nikolić, Sun, and Rawlins 2018). Defining key factors of the molecular mechanism and cell type function during normal *in-utero* human lung development is crucial to ensure that the directed differentiation process closely mimics the human developing lung and that the maturation protocol recapitulates function and phenotypes of mature *in-vivo* lung cells (Miller and Spence 2017).

The use of a single cell “omic” approach in the study of human fetal lung tissue is a promising option to address the concerns and limitation raised by animal models of lung development and the use of human lung organoids.

## **2. Single cell RNA sequencing (scRNAseq)**

### **A. Principle of scRNAseq**

Describing how the various developmental cell populations contribute at a molecular and a cellular level to the creation of a complex organ, such as the lung, is of tremendous importance. Using an omics approaches, including transcriptomic and proteomic analysis, can help to provide a deeper insight on the complex underlying mechanisms in the developing lung. Traditional transcriptomic analysis on tissue is performed using a bulk RNA sequencing. However, a bulk RNA analysis averages the gene expression from all the cells within the sample, therefore masking any cell-to-cell differences and underlying cellular heterogeneity. The rapid recent development of transcriptomic analysis at a single-cell resolution [single cell RNA sequencing (scRNAseq) and single nuclei RNA sequencing (snRNAseq)] allows investigators to obtain information from the tissue of interest on the level of a single cell without averaging the measurements. scRNAseq enables detection of distinct gene expression profile for each cell type, revealing heterogeneity and subpopulation expression variability in thousands of individual cells. Since the first publication in 2009 (Tang et al. 2009), scRNAseq has been used in many different human tissues to identify cell heterogeneity, to obtain and validate information from lineage-tracing experiments, study developmental fate of different cell types, and to identify rare cell or pathogenic cell populations which appear only during a disease (Haque et al. 2017).

## **B. Single-cell resolution transcriptomics in respiratory and lung development research**

The Human Cell Atlas international consortium aims at creating an atlas of every human cell type using scRNAseq, with the description of distinctive gene expression pattern, physiological state and developmental trajectories for each cell type. In addition, the spatial relationship between the individual cell populations within the tissue are to be studied using spatial transcriptomic and proteomic approaches (Regev et al. 2018). The Human Cell Atlas project includes a Human Cell Atlas Lung network, which aims to provide insights on identities, functions and lineage relationships of all the cells in the normal lung. This atlas could be further used as a reference point to compare data sets from diseased lung tissues and to highlight pathological changes in cell composition, cell interactions or functions to further provide understanding on pathophysiology and relevant potential biomarkers and therapeutic targets. In addition, this consortium seeks to facilitate integration of datasets from multiple centers and different platforms, to support the development of standardized protocols for lung tissue processing, and to develop data storage solutions and shared computational approaches (Schiller et al. 2019).

scRNAseq in lung research can be used to describe lung cell complexity in both normal and diseased lung, including discovery of novel, rare cells types, such as human airway ionocytes (Plasschaert et al. 2018). scRNAseq also allows the study of underlying molecular mechanisms in normal lung processes, including development, regeneration and ageing. Its insights in lung disease include discovery of pathophysiological mechanisms, development of biomarkers to help clinical decision-making (diagnostic, prognosis, therapeutic) and identification of perturbed pathways as future potential therapeutic targets (Alexander, Budinger, and Reyfman 2020).

When it comes to the study of lung development, several studies using scRNAseq to describe lung development in animals, particularly in mice, have been published (Mižíková and Thébaud 2021). However, only very few studies were carried out using human fetal lung tissues. *Danopoulos et al.* analyzed scRNAseq dataset from 2 human fetal samples (at 11.5 and 18.5 weeks of gestation), focusing on mesenchymal cells clusters. In their study, authors described two *ACTA2*<sup>high</sup> sub-clusters corresponding to smooth muscle cells. Additionally, using spatial localization they were further able to describe unique marker genes for airway and vascular smooth muscle cells (Danopoulos et al. 2020). *Miller et al.* used human bud tip organoids to describe the bud tip-to-airway and basal cells transition via *SMAD* signaling activation. Furthermore, authors validated the clinical relevance of their model by comparing scRNAseq data from human fetal lung tissues (from 11.5 to 21 weeks of gestation) with the transcriptome of the bud tip organoids cells. They showed that transcriptome from organoids at 21 days of culture exhibited the highest correlation with the transcriptome of *EPCAM*<sup>+</sup> cells from the human fetal lung samples (Miller et al. 2020).

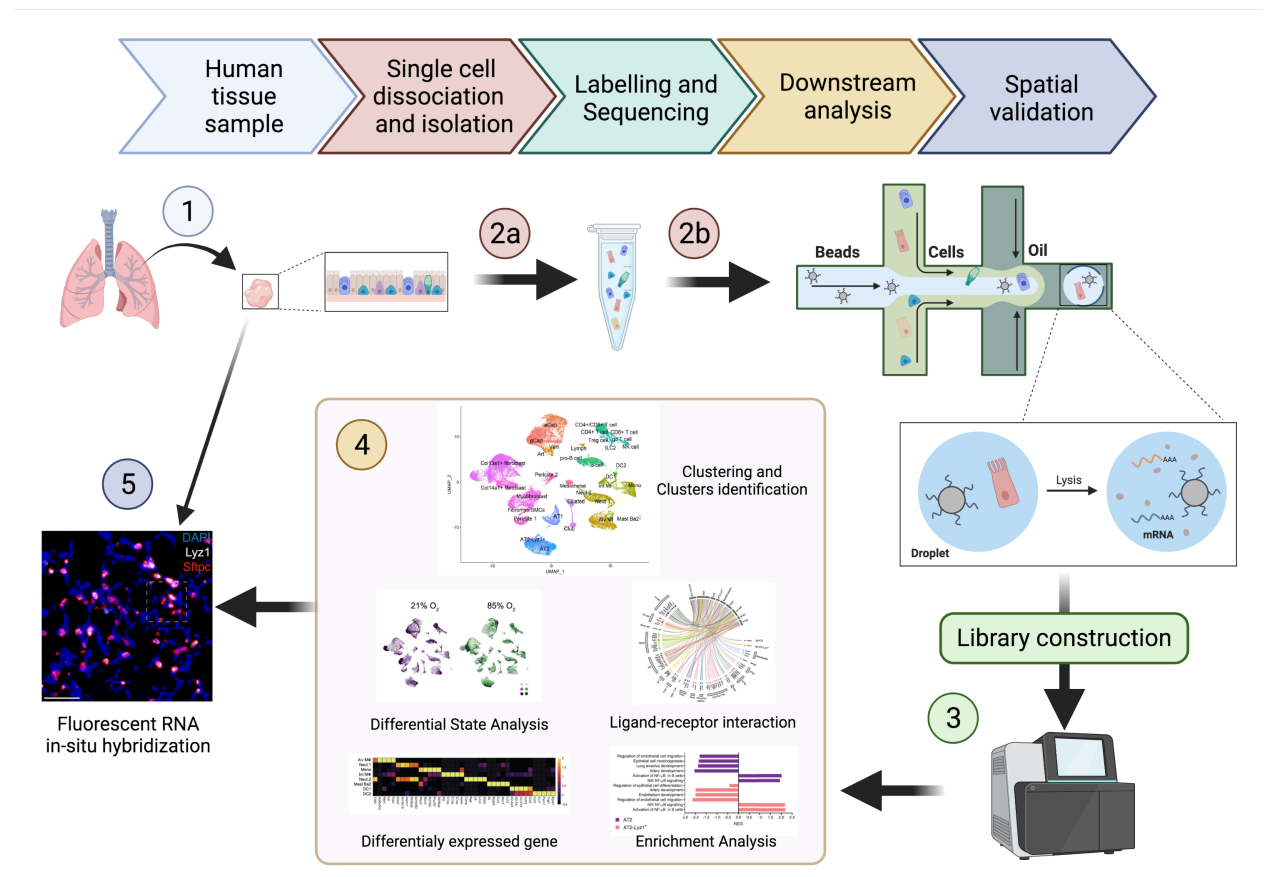
In order to support the research community in its efforts to unravel cellular composition and molecular mechanism of human lung development, the LungMap consortium (<https://lungmap.net>), funded by the National Heart, Lung and Blood Institute, has been created. LungMAP aims to provide a comprehensive developmental atlas of normal developing lung in mice and humans based on omics and imaging data. All the data are available to the research community via the BREATH (Bioinformatics REsource Atlas for the Healthy lung) database, a collaborative and open access data application. LungMap also developed analytic web tools called LungGENS (Lung Gene Expression iN Single-cell) for mapping gene expression patterns in specific pulmonary cells at a single-cell level in both, mouse and human (Du et al. 2015). An extended version of LungGENS called LGEA

webportal (Lung Gene Expression Analysis web portal) now includes data from single cell analysis (LungGENS), sorted cells (lungSortedCells) and whole lung (LungDTC) analysis, as well as integrative analytical tools such as “lung-at-a-glance”, and additional omics data from diseased human lungs (Du et al. 2017; 2021).

### **C. scRNAseq analysis workflow**

Single-cell resolution transcriptomic can be performed after a single-cell (scRNAseq) or single-nuclei (snRNAseq) isolation. scRNAseq typically requires an access to fresh or DMSO-frozen tissue whereas snRNAseq can be easily performed on frozen, or even fixed tissues. Dissociation protocol to obtain the single cells suspension from tissue sample is the most critical part of scRNAseq preparation. Most of the protocols are based on a combination of mechanical and enzymatic dissociations and are exposed to dissociation-induced bias regarding the single-cell suspension composition (e.g., depending on the level of “harshness” of the protocol, some cell type subpopulation might become overrepresented whereas other more fragile cell types can be underrepresented or completely absent). Moreover, as the dissociation process is often performed on live cells, it can induce a transcriptional stress response in the cells (artefactual genes expression). Comparison of snRNAseq vs. scRNAseq in diverse tissues (kidney, brain and lung) showed a comparable gene detection with a reduced dissociation bias (e.g., higher proportion of epithelial cells, including rare population like pulmonary neuroendocrine cells, and less immune cells isolated from mouse lung tissue) in snRNAseq. In addition, no artifactual gene expression (dissociation-induced transcriptional stress response and mitochondrial genes) was observed in the case of snRNAseq (Wu et al. 2019; Bakken et al. 2018; Lake et al. 2017; Koenitzer et al. 2020).

Both strategies share the same analysis workflow presented in Figure 10, and described below (for simplification, single cell will be used to describe both snRNAseq and scRNAseq workflow) (Haque et al. 2017; Luecken and Theis 2019).



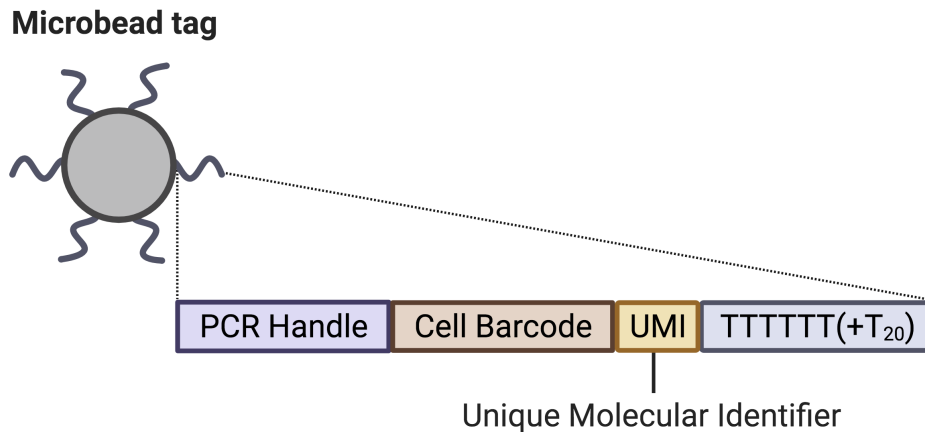
**Figure 10.** Overview of scRNAseq workflow, from tissue sampling to spatial validation  
Different steps of the scRNAseq workflow are presented here. After tissue sampling (1), the sample undergoes a dissociation protocol to obtain a single-cell suspension (2a). Single cells are isolated from the cell suspension using a droplet-based method (2b), with each droplet containing a barcoded bead (cf figure 10). Cells are then lysed, and cDNA is synthesized. Next step (3) is library construction and sequencing using Next Generation Sequencing. After pre-processing, data can be used for downstream analysis (4) which include clustering and cluster identification, trajectory analysis, differentially expressed gene analysis, enrichment analysis and ligand-receptor interaction. Last step (5) is the spatial validation of scRNAseq findings on native tissue, for example using fluorescent RNA in-situ hybridization, immunohistochemistry or spatial omics approaches. Pictures of the downstream analysis are adapted from (Hurskainen et al. 2021). Figure created with [BioRender.com](https://BioRender.com).

The first step is **tissue dissociation**, where the tissue of interest is dissociated to obtain a single-cell (or nuclei) suspension. Several dissociation protocols have been developed to target specific cell population and/or organ, all of them having their advantages and limits

(Denisenko et al. 2020; Vieira Braga and Miragaia 2019; Potter and Steven Potter 2019). As stated above, single-nuclei isolation is less prone to dissociation bias compare to single-cells isolation. Quality of the dissociation process (e.g., cell viability, cell type representation...etc.) can be assessed by fluorescence-activated single cell sorting (FACS).

The second step is **single-cell isolation**, during which each cell is isolated for further mRNA profiling at a single cell level. Droplet-based methods (Drop-seq, inDrop, 10X Genomics'Chromium) use a microfluidic approach to isolate cells by encapsulation and tag them. During this process, oil-based droplets are formed, each containing a single cell, a single bead with a unique barcode (i.e., droplet-specific cellular barcode), and molecular reagents to create barcoded cDNA within the droplet. Errors occurring during the encapsulation process (i.e., doublet or empty droplets) can be assessed and removed prior to downstream analysis (Salomon et al. 2019). The barcode beads bear oligonucleotides that have a common structure described in Figure 11. The cell barcode (common to all the oligonucleotides attached to the same bead) allows all the amplified transcripts from the captured cell to be identified as coming from the same cell. The unique molecular identifier (UMI) is specific for each oligonucleotide attached to the bead and allows to identify amplified copies from the same mRNA molecule (within the same UMI) and to differentiate them from amplified copies from other mRNA molecules present within the same droplet (different UMIs). This multi-labelling process allows to pool different samples for sequencing (MULTI-seq multiplexing technology) and therefore decrease the cost associated with scRNAseq.





**Figure 11.** Barcoded bead structure

Oligonucleotides are attached to the bead. Their structure includes 1) a primer region for molecular amplification of the captured transcript (PCR Handle), 2) a unique series of nucleotides (cell barcode) that are identical for all oligonucleotides attached on the same bead to identify cell origin of all the amplified transcripts, 3) second series of nucleotide unique for each oligonucleotide (UMI) to identify amplified transcripts originating from the same mRNA molecule, and 4) poly-d(T) region to capture polyadenylated RNA.

The third step consists of **library preparation** (i.e., molecular preparation and amplification). Library preparation is the process in which the intracellular (scRNAseq) or intranuclear (snRNAseq) mRNA is captured, reverse-transcribed to cDNA molecules, and amplified. The isolated cell is lysed within the droplet to release the mRNA which will hybridize with the bead oligonucleotide via its poly-d(T) region and be converted in cDNA. As this process happens in an isolated cell within a droplet, the obtained cellular cDNA libraries are labelled with droplet-specific cellular barcode.

The fourth step is the **sequencing** via Next Generation Sequencing (NGS). Because the cDNA libraries from the same cell share the same barcode, different samples can be pooled (multiplexing) before sequencing.

The fifth step is the **data pre-processing**, during which raw sequenced data are processed for further downstream analyses. Raw data undergo demultiplexing (sequencing read from each sample are deconvoluted based on their barcode sequence information) and are processed and aligned to obtain count matrices where each column is a cell, and each

line is a gene. The count matrices will undergo further pre-processing which can include quality control, normalization, data correction and integration (to take in account biological, technical or batch effects) and reads alignment. One important step in the data pre-processing is dimension reduction. As typical scRNAseq count matrices consist of approximately 20,000 genes in 50,000 up to a million cells, and as not all genes are equally important to describe cells' expression profile and classify cells in meaningful clusters, dimension reduction is essential for data complexity reduction (summarization) and data visualization. Most popular dimensional reduction and visualization techniques are Principal Component Analysis (PCA), t-Stochastic Neighborhood Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP).

The last step of the scRNAseq experiment consists in **downstream analysis**. Downstream analyses are used to extract biological insights and describe underlying biological and molecular processes. These include cell clustering and sub-clustering, identification of cell developmental trajectories, differential gene expression analysis, gene set enrichment analysis and ligand-receptor interaction analysis. Traditional downstream analysis starts with cell clustering, during which cells are grouped based on their level of similarity in gene expression. The same approach can be then applied on a selected cluster of interest (sub-clustering) in order to focus on more detailed substructures within the dataset. Cluster annotation is performed using marker (or signature) genes identified by differential gene expression analysis. The identified marker genes are typically compared to data from literature or a reference database (e.g., LungGENS, Human Cell Atlas) to describe cell identities. In studies using several timepoints or experimental conditions, trajectory analysis or differential state analysis are useful to describe changes in cell states and gene expression across experimental groups. In addition, enrichment analysis, where identified

differentially expressed genes are used as an input, can provide further insight into the underlying molecular functions and biological pathways, or help identify cell components involved in these changes. Last, based on differential gene expression, ligand-receptor interactions analysis can provide information on potential cell communication.

Finally, as scRNAseq is carried out using dissociated tissues, **spatial validation** of the scRNAseq findings using immunofluorescent staining or fluorescent RNA in situ hybridization (FISH), is usually performed. Such spatial validation helps to confirm that differentially expressed genes specific to a cell of interest concord with the cell localization within the tissue of origin. Additionally, in parallel to scRNAseq, spatial transcriptomic techniques have been recently developed, allowing to obtain scRNAseq data directly from the tissue section (Chen, Teichmann, and Meyer 2018).

As scRNAseq is based on computational analysis, with many different existing platforms, scRNAseq raw data, metadata and codes are usually deposited by authors on open access repositories to ensure of its access to the entire research community. This enables a collaborative approach, allowing careful reanalysis and combination with other datasets (Alexander, Budinger, and Reyfman 2020).

#### **D. scRNAseq, advantages and limitations**

scRNAseq is currently the most sensitive and unbiased way to study cell types and cell states. Limitations of scRNAseq related to tissue dissociation have been discussed above, with single nuclei approach being less prone to this dissociation bias. Technique's main limitation is the "drop out" effect when a transcript is not detected in the sequencing data because of failure to capture and amplify it. As a result, the count matrices will exhibit a zero value for this gene which can be interpreted as no gene expression (instead of technical

failure to capture the mRNA). This is especially frequent for lowly expressed genes, as a single cell might contain as little as 10 to 30 pg of RNA. Different strategies have been developed to address this issue. Lastly, due to the high cost associated with scRNAseq, the investigators might be forced to decrease the number of sequenced cells. By doing so, the low number of sequenced and analyzed cells can impede detection of very rare cell populations, which can be either under or overestimated.

In conclusion, snRNAseq represents a promising approach to study transcriptomics at a single cell resolution in the human developing lung and provide insightful findings to extend our knowledge on how the gas exchange surfaces in the lungs are built in the developing embryo and fetus. Unravelling the cells composition, developmental trajectories and communications within the human developing lung will provide critical knowledge to further support lung regeneration and repair in congenital and acquired human lung diseases.

### **III. Human fetal lung derived Precision-cut lung slices (PCLS) to validate clinical relevance of promising pre-clinical therapies**

#### **1. Precision-cut lung slices (PCLS): definition, principle and potential applications**

Precision-cut tissue slices can be generated from any species and various tissues, including lungs. First described by *Fisher et al.* in 1994 (Fisher et al. 1994), human PCLS are typically 100-500  $\mu\text{m}$  thick and represent a complex 3D *ex-vivo* lung culture model. As PCLS can be obtained from primary human lung tissue, they represent a versatile tool to support translational research (G. Liu et al. 2019; Evans and Lee 2020). In order to obtain the PCLS, the lung's airways need to be inflated with an agarose solution to prevent airway collapse during the slicing process. Once the agarose is solidified, depending on the lung size, the whole lung, or a sample selected with a tissue coring tool, can be cut into slices using a tissue slicer or a vibratome. Cultured PCLS retain many aspects of the cellular composition and cellular interactions, as well as architectural organization of the lung (organotypic culture). In addition, airways and vessels within the PCLS have preserved functions of contraction and relaxation (Sanderson 2011). Therefore, PCLS provide a valuable platform to study the function and structure of the lung in its native 3D environment, including natural interactions between cells, molecules and the ECM *ex-vivo*. In human lung derived PCLS, cellular and structural organization can be maintained in culture for up to 15 days or even up to 21 days when the PCLS were embedded in hydrogel (Neuhaus et al. 2017; Bailey et al. 2020).

Studies using PCLS traditionally include morphologic or functional endpoints, such as lung histology, or airways and pulmonary vasculature reactivity. Development of high-resolution

imaging studies has further expanded the potential use of PCLS. *Akram et al.* reported the use of live cell staining combined with long-term time-lapse imaging to study alveologenesis in PCLS derived from mice pups. With this technique, the authors unraveled the importance of epithelial cell migration during alveologenesis (*Akram et al. 2019*). Other common endpoints include PCLS viability and metabolism, apoptosis assessment, and protein and secretome analysis (*Liberati, Randle, and Toth 2010*). Only a few studies on PCLS used transcriptomic analysis as RNA extraction can be impeded due to the small amount of tissue and the agarose in the airways. However, protocols to successfully isolate RNA from PCLS have been developed (*Niehof et al. 2017*).

PCLS can be obtained from different species, but the use of human lung tissue represents the main advantage of PCLS studies. Healthy and diseased human lung tissues from surgical resection can be used to generate PCLS to facilitate studies of normal lung physiology, as well as comparative studies between healthy and diseased lung derived PCLS. In addition, *ex-vivo* disease modeling can be achieved using healthy lung PCLS (*Alsafadi et al. 2017*). The *ex-vivo* disease modeling represents an experimental model to study molecular mechanisms and cell interactions during disease, identify key actors of the diseased process and finally, provide a screening platform for potential therapeutics. PCLS obtained from diseased lungs allow baseline analysis of diseased tissue with comparison with healthy tissue and therapeutic screening (*Uhl et al. 2015*). PCLS derived from animal lung are isolated either from healthy animals in order to recreate the disease *ex-vivo*, or directly from experimental animal used for disease modelling *in-vivo*. As PCLS maintain lung architecture, retain cell types (including resident immune cells), cell-cell and cell-ECM interactions and functional airways and vessels, PCLS are used in a wide range of applications. Human and animal PCLS can be used for lung physiology studies (*Sanderson 2011*), studies on structure, function and

remodeling of pulmonary vasculature (Springer and Fischer 2003; Klouda et al. 2021) and airways (Sanderson 2011; Andrea Wohlsen et al. 2010), alveologenesi s (Pieretti et al. 2013; Akram et al. 2019), as well as lung immune response studies (Henjakovic et al. 2008). Additional application for PCLS include therapeutic screening with efficacy and safety endpoints (G. Liu et al. 2019), pharmaco-toxicological studies (Neuhaus et al. 2018; Morin et al. 2013), assessment of viral vectors transfection for gene therapy (Rosales Gerpe et al. 2018), and allergic response studies (A. Wohlsen, Uhlig, and Martin 2001).

## 2. PCLS workflow

Human PCLS workflow is usually based on the following steps: isolation of the human lung sample, lung inflation with an agarose solution, lung slicing, *ex-vivo* culture with experimental interventions, and endpoints assessment. While the workflow is rather simple, there is currently no standard for human PCLS generation and culture, and wide variations of protocols exist for each step of the process (Alsafadi et al. 2020).

The workflow is composed by the following steps:

**Lung sample isolation.** To increase the success rate in obtaining PCLS with an optimal viability, human lung tissue sample should be processed as soon as possible. After surgical resection, the lung sample can be stored in cold culture medium (+4°C) until processing. Prior to lung inflation, the lung sample can be slowly rewarmed in the culture medium using a water bath.

**Lung inflation.** Agarose solution is prepared with low-melting point agarose diluted in culture medium. Once melted, low-melting point agarose will remain fluid at 37°C and start gelling below 25°C. The agarose solution can be therefore safely instilled within live

tissue without risking a heat-induced stress response. When accessible, trachea or a main bronchus should be used to slowly and gently instill the agarose solution with a catheter and a syringe. Alternatively, for small tissue samples without access to a main airway, a 27-gauge needle can be used to slowly inject the agarose solution within the parenchyma. Lung inflation is the critical part of the PCLS generation process. Agarose solution concentration and temperature are crucial to obtain proper lung filling and stiffness required for slicing. In addition, if the study design involves assessment of the alveolarization, particular attention should be given to the instillation method. Different solutions to address this concern have been used, including constant pressure instillation and the use of an electric syringe. Once filled, the lung sample rests on ice in order to allow the agarose solution to gel.

**Lung slicing.** Once cold, the lung can be sampled for slicing with a tissue coring tool. The tissue sample can be embedded in a higher concentration agarose solution to further optimize the slicing process. Lung slicing can be performed either with a tissue slicer or a vibratome, with a slice thickness typically set between 100 to 500  $\mu\text{m}$ . PCLS are harvested in culture media and further washed before culture.

**PCLS culture.** Wide variations of culture conditions have been reported between different protocols (Table 5). Most of the studies use immersed culture while some use a dynamic system with a rolling device to allow cyclic exposure to air and medium. In addition, modification of the culture environment, such as exposure to different oxygen concentrations or addition of lipopolysaccharides (LPS) or toxic components in the culture media can be used to generate a disease model. Finally, PCLS viability assessment during culture can be performed with different assays, such as a metabolic assay (resazurin), viability staining or substance-induced bronchoconstriction.



**PCLS study endpoints.** As previously described, PCLS study endpoints include morphological and functional (broncho-reactivity and vascular reactivity) assessment, proteins and secretome analysis, as well as transcriptome analysis. Development of high-resolution imaging enables the study of morphological changes within live PCLS, allowing sequential assessments for the same PCLS at different timepoints during the culture.

### **3. Advantages and limitations of the PCLS model**

Like other complex ex-vivo systems to culture human lung tissue, the main limitation for human PCLS is the access to normal and diseased human lung tissue. However, as PCLS usually have a thickness less than 500 $\mu$ m, even a small tissue sample is sufficient to generate an adequate amount of PCLS to perform experiments. This feature also supports the use of animal lung derived PCLS as it permits the reduction in the number of animals needed for experiments (3Rs principle <https://nc3rs.org.uk>). Multiple PCLS can be generated from the same animal and randomized in different treatment groups, reducing the effect of genetic background. Another mitigation strategy to address human lung tissue paucity, is the development of protocols for human PCLS cryopreservation. Frozen-thawed PCLS were shown to maintain viability and architecture as well as cell activity (e.g., epithelium and airway smooth muscle cells) (Bai et al., n.d.; Watson et al. 2016). Additional comparisons of different human pulmonary 3D models are presented in Table 5.

	ALI	Lung organoids	PCLS
<b>Physical architecture</b>			
Structural mimics	Pseudostratified epithelium	3D multicellular tissue construct <i>in-vitro</i>	<i>Ex-vivo</i> tissue system
Cell composition	Ciliated cells, goblet cells and basal cells	Multiple differentiated cell types including basal cells, ciliated cells, goblet cells and alveolar epithelial cells	All relevant cell types including structural and immune cells
Cell origin	Primary and immortalized bronchial epithelial cells	Primary airway cells and hPSC	All resident cells
Recruitable immune cells	No (unless co-cultured)		Residents immune cells only
ECM	No	Exogenous (Matrigel)	Endogenous
<b>Physiological function</b>			
Physiological relevance	Muco-ciliary differentiation, epithelial barrier function, physiological response to insults	Morphological and functional mimicking of the airway; heterogenous cell composition and spatial organization; self-renewal and differentiation	Preservation of lung architecture (small airways, parenchyma, cell population and connective tissue); physiological responses to stimuli
Homeostasis	No	Yes	yes
Immune responses	Epithelium related responses		Collective responses from tissue components

**Table 5.** Comparison of Human pulmonary 3D models for translational research

ALI: Air-liquid interface, PCLS: precision-cut lung slices, hPSC: human pluripotent stem cells, ECM: extracellular matrix. From (G. Liu et al. 2019).

One of the limitations of the PCLS model is its inability to recruit non-resident immune cells, therefore limiting its potential to replicate *in-vivo* immune response. Furthermore, PCLS are traditionally cultured in a static system without exposure to ventilation or perfusion. The static culture limits PCLS potential to study the effect of mechanical stressors on lung epithelium and endothelium. However, *Davidovitch et al.* have developed a model of PCLS culture in which stretch forces can be applied to the PCLS to model ventilator-induced lung injury (Davidovich, Huang, and Margulies 2013).

Lastly, PCLS represent a versatile platform for drug and toxicity screenings, allowing the study of their effects on different lung cell types. In these studies, the component of interest is usually added to the culture medium. Hence, extrapolation from PCLS studies of dosing and administration route (inhaled or systemic) to clinical settings should be carefully considered and confirmed with other methodologies.

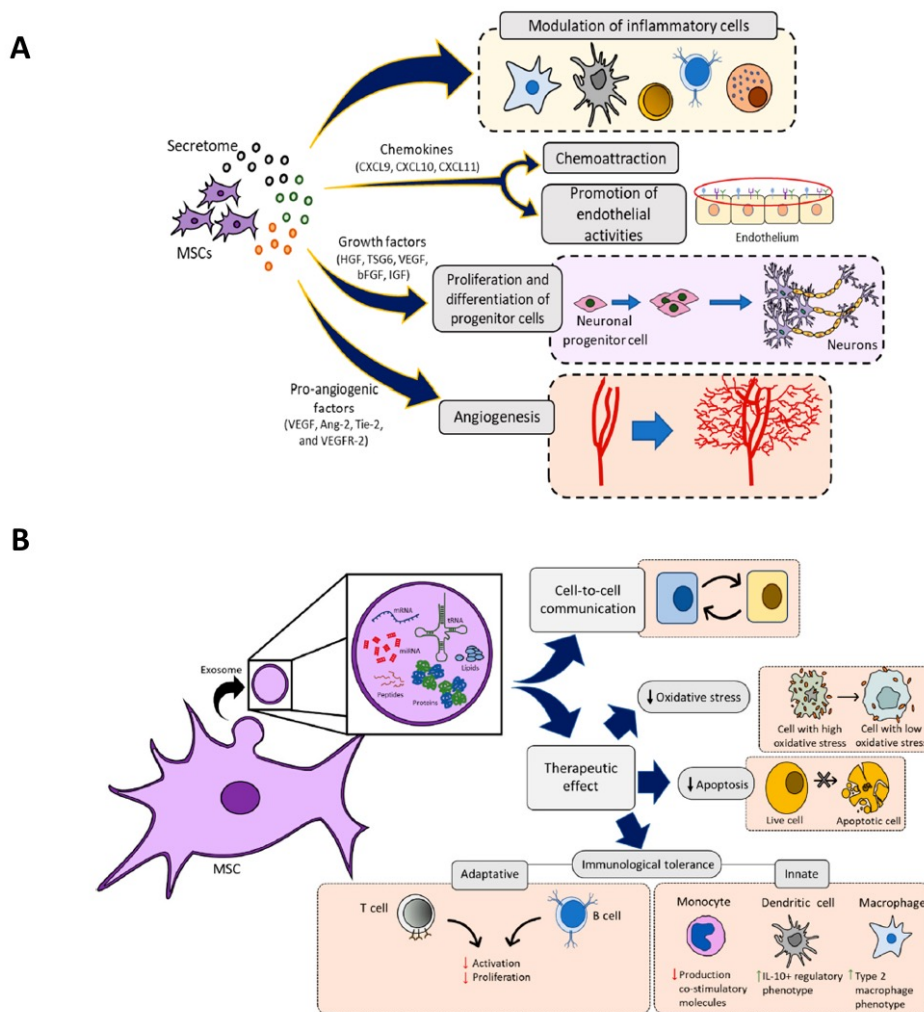
In conclusion, human PCLS (healthy or diseased) represent a versatile tool to assess pathophysiology or therapeutics in an organotypic culture setting. Additionally, use of PCLS aligns with the need for more complex and holistic models and the use of human tissue in translational research to support clinical translation.

## **IV. A Delphi study to establish consensus definition and reporting guideline for Mesenchymal Stromal Cells (MSC)**

### **1. Mesenchymal stromal cell research**

Since MSC original description in 1976 (Friedenstein, Gorskaja, and Kulagina 1976) and their first use as a therapeutic agent in humans in 1995 (Lazarus et al. 1995), over 1300 MSC clinical trials have been registered on clinicaltrials.gov (<https://clinicaltrials.gov>).

Interest in the study of MSC in regenerative medicine has originated from their availability from different tissue sources and their ability to support the repair process in injured tissues. MSC share some properties with stem-cells such as self-renewal and multipotency. However, their therapeutic effect is most likely due to their homing function (homing and transendothelial migration to injured tissue driven by chemokines) and their immunomodulatory and repair properties via paracrine effect. This complex paracrine effect is mediated by MSC secretome and MSC-derived extracellular vesicles. MSC secretome includes chemokines, growth factors and pro-angiogenic factors. MSC-derived extracellular vesicles including exosomes in association with MSC secretome play a major role in communication between MSC and surrounding cells. Paracrine effects of MSC are described in Figure 12 (Jimenez-Puerta et al. 2020).



**Figure 12.** Paracrine effects of MSC.

MSC paracrine effects are mediated by (A) MSC secretome and (B) MSC-derived extracellular vesicles. These paracrine effects enhance tissue repair and regeneration by supporting immunomodulation, chemoattraction, progenitor cells proliferation and differentiation, and angiogenesis. In addition, MSC-derived exosomes decrease oxidative stress and apoptosis. From (Jimenez-Puerta et al. 2020).

Despite promising results regarding the use of MSC in different preclinical diseases models, clinical trials using MSC in various medical conditions have provided only few encouraging results. Many potential reasons have been suggested to explain the challenges in clinical translation of MSC-based products. For instance, preclinical and clinical studies using MSC have had important disparities in MSC characteristics (e.g. characterization, immune compatibility, cell viability and dose) and study participants' characteristics, such as disease severity and associated comorbidities (Jacques Galipeau and Sensébé 2018). Many solutions

have been proposed to improve efficacy of MSC in human clinical trials. The solutions include: 1) the development of potency assays to predict *in vivo* response, 2) standardized methods for MSC manufacturing and administration (Nolta, Galipeau, and Phinney 2020), and 3) more complete and transparent reporting of both, cell characteristics and clinical trials, as well as manufacturing details. The latter solution is particularly important in order to better understand underlying factors that might contribute to efficacy (or a lack of efficacy) of MSC (Robinson et al. 2019; Rizk et al. 2016).

The International Society for Cell & Gene Therapy (ISCT) first provided minimal criteria to define MSCs in 2006 (Dominici et al. 2006). The terminology of “multipotent mesenchymal stromal cells” and the criteria described were determined through an informal consensus of a small number of experts from the ISCT. Multiple critiques of this definition have been raised since its publication, including skepticism regarding the cell surface markers phenotype, the differentiation assays which are prone to misinterpretation, limitations in functionally defining stromal cells, and that notion that the definition does not account for developmental origins of respective tissue source (Robey 2017; Sipp, Robey, and Turner 2018; Viswanathan et al. 2019). This led to an attempt to update the definition in 2019. However, the ISCT MSC committee again adopted an informal method and consensus could not be reached on many items discussed (Viswanathan et al. 2019). As a result, the uptake of the ISCT’s definition has been inconsistent in both, preclinical and clinical research, and the MSC research community still does not have an updated definition that addresses criticisms raised.

Within the clinical trials that have tested MSC therapy, there has been significant heterogeneity in reporting regarding the used products. In a report from the Food and Drug Administration, *Mendicino et al.* assessed initial filings of 66 Investigational New Drug submissions for MSC-based products. They showed important differences in cell surface marker characterization, product bioactivity assessment, as well as tissue sourcing, and product manufacturing (Mendicino et al. 2014). The largest and most cited systematic reviews of published MSC clinical trials have found extremely poor (i.e. incomplete) reporting of cell products used (Thompson et al. 2020; Lalu et al. 2012). For instance, only 13% of MSC randomized control trials described characterization of their cell product according to current criteria, and only 52% reported viability of their cells prior to administration. This lack of sufficient detail is highly problematic for several reasons. If authors do not provide sufficient details concerning the cell product they used, then readers are left with an incomplete picture of what was done. This significantly hinders the ability to judge the reliability of the results, to interpret them, or to replicate the findings.

Non-rigorous or incomplete clinical reports could lead to difficulties in study quality assessment, comparison between studies, extrapolation and even possibly influence the results. *Nowbar et al.* analyzed discrepancies in clinical reports from studies assessing the efficacy of bone-marrow derived MSC on left ventricle ejection fraction in heart diseases. They found that the rigor of the report was associated with the effect size. Studies with no discrepancies showed no effect of MSC on left ventricle function whereas studies with the highest amount of discrepancies also reported the biggest improvement in left ventricle function (Nowbar et al. 2014). This suggests that inadequate reporting is associated with biased reports.

More concerning, the uncertainty and lack of consensus on this crucial issue of how to define MSC has played a major role in enabling the marketing practices of businesses engaged in direct-to-consumer marketing of unproven and unlicensed “stem cell” products. Many of these business market “MSC stem cell therapies” claiming they can treat a wide range of diseases and injuries (Snyder, Turner, and Crooks 2018; Turner 2020; Sipp, Robey, and Turner 2018).

To better address clinical translation, reproducibility and transparency in the field of MSC research, the scientific community needs to find, adopt and enforce a consensus definition of MSC. Similarly, to improve reporting quality of clinical MSC studies, a standardized reporting guideline is needed.

## **2. Delphi method and “integrated knowledge translation” approach**

Our research proposal aims to address the above-mentioned important gaps and pitfalls of previous attempts to define MSC. We plan to use a methodologically rigorous formal consensus process (Delphi method) in combination with an “integrated knowledge transmission” approach to develop, disseminate and implement 1) an updated consensus definition of minimal criteria to define MSC and 2) Reporting guidelines for the clinical trials of MSC therapy. This reporting guideline will be a CONSORT (Consolidated Standards of Reporting Trials) extension that will focus on aspects unique to MSC trials, such as manufacturing, characterization, storage and delivery of cells.



## A. Overview of the Delphi method

Previous attempts to generate minimal criteria for MSC have failed due to divergent opinions of stakeholders from different fields (e.g., developmental biologists, translational scientists, clinicians...etc.). Our approach revolves around the **Delphi method**, which is a highly studied and well-established consensus building method, developed by the Rand Corporation in the 1950s (Dalkey 1969). The Delphi method is used to reach group consensus, even when divergent opinions are initially presented (Hsu and Sandford 2007; Okoli and Pawlowski 2004). The Delphi approach has been previously used to develop clinical reporting guidelines that are now widely adopted, and reach consensus to define highly contentious issues (e.g., biomedical editors core competencies or defining predatory journals) (Grudniewicz et al. 2019; James Galipeau et al. 2017).

The Delphi method allows for a large panel of experts and limits negative aspects of group decision-making, such as 1) logistical issues (e.g., difficulties to bring people together), 2) adequately accounting for divergent opinions, 3) limiting the number of participants, 4) meeting organisation and costs, 5) time limits for very complicated decisions, 6) peer pressure and influence in the decision-making process, and 7) group influences on individual performance. Additional limitations of group decision-making include lack of formal feedback and non-structured interactions and aggregation of opinion. Many of these limitations may explain why the most recent attempt by the ISCT, which relied on simple group decision-making, failed to reach a consensus on the definition of MSC. The Delphi method's ability to address all these pitfalls have been highly studied over a variety of fields, making it one of the most popular approaches for structured consensus building (Table 7).

Consensus development method	Formal questionnaire	Private decisions elicited	Formal feedback of group choices	Face-to-face contact	Interaction structured	Aggregation method
Group decision-making	No	No	No	Yes	No	Implicit
Modified Delphi	Yes	Yes	Yes	Yes	Yes	Explicit

**Table 6.** Differences between group decision-making and Delphi Method

Adapted from (Murphy et al. 1998).

The Delphi method is a structured communication method, which relies on a panel of experts and several rounds of questionnaires to reach a consensus (iterative process). Each expert receives a questionnaire by e-mail and privately answers each item asked. Participants are asked to rate relative importance of each item and to provide rationale for their choices as appropriate. After each round, answers are anonymously summarized and items that reach consensus as being important (kept) or un-important (discarded) are removed from subsequent rounds. For items that remain, a new questionnaire (along with summarized feedback from the previous round) is formulated for a subsequent round. After several rounds, the responses converge toward a consensus using an explicit aggregation method (Murphy et al. 1998).

## **B. Knowledge translation (KT) strategy**

Our proposal also recognizes that arriving at a consensus definition does not ensure proper knowledge translation. To further support dissemination and implementation of our findings, our approach to knowledge translation involves using: 1) an “integrated knowledge translation” (iKT) approach and 2) a tailored end of project KT plan.

a. “integrated knowledge translation” (iKT) approach

Integrated knowledge translation is defined as “a model of collaborative research, where researchers work with knowledge users who identify a problem and have the authority to implement the research recommendations. Knowledge users have unique expertise pertaining to the research topic, including knowledge of the context and the potential for implementation. Researchers bring methodological and content expertise to the collaboration. Implicit in this approach is the sharing of power between researchers and knowledge users. Sometimes referred to as the co-production of knowledge, this new way of working suggests that the synergies derived from the collaboration will result in better science, more relevant and actionable research findings, increased use of the findings in policy or practice, and mutual learning” (Kothari, McCutcheon, and Graham 2017). For our project, stakeholders include opinion leaders in the field of MSC (prominent researchers) as well as knowledge users such as patient partner, regulatory agencies, journal editors, research community and stem cell therapy companies. These stakeholders will be integrated into the project from its inception so that those with the authority to implement our recommendations have their needs and preferences considered during the consensus building process. Taking an iKT approach will help ensure that our project is relevant, timely, and ultimately useful to the knowledge users, assuring that the findings, once available, will be taken up into practice and policy more quickly.

b. End-of-project KT Plan

In addition to the iKT approach, a structured approach to end-of-project, modeled on CIHR's Knowledge Translation planning guide (Government of Canada 2012), can further help and support dissemination and implementation of our results. This essentially involves considering the specific project findings and using the following steps:

- Determining the **dissemination goal** (e.g., to increase awareness of the findings, to increase knowledge, to influence practice or policy)
- Identifying **key audiences**. Different stakeholder groups including researchers, journal editors/reviewers, funders, patients and regulatory agencies.
- Crafting **tailored messages** to specific audiences. A dissemination goal for each key message must be determined. For example, for researchers the goal will be to inform their use of MSCs. For journal editors and reviewers, it would be to influence peer-review process with structured reporting guidelines, and for funders, the dissemination goal would be to inform and influence review processes.
- Using **communication strategies** and media to reach each audience. This includes development of tools to support dissemination such as training material or website. In addition, the communication networks of our knowledge user partners and stakeholders can also be used to disseminate these findings. For example, patient partner can help to reach out to local and national social media and traditional media to inform the public and patients about our study findings.

A written KT plan will help to ensure study outputs (the consensus definition and reporting guidelines) are effectively used. The knowledge users involved in the project will play a critical role in interpretation of the Delphi results and in crafting the dissemination plan and messages for the targeted audiences.

In conclusion, this Delphi project highlights critical points for consensus-building strategies and the critical step of knowledge translation in translational research.

# Thesis objectives

In this thesis, we describe different projects that aim to improve translational research in the field of neonatal lung diseases.

In the first part, we describe new insights into normal human lung development *in-utero* from **snRNAseq on human fetal lungs samples** (9 samples with gestational ages from 14+1 to 19+0 weeks of gestation). Translational research requires a deep understanding of health and disease state in human. Therefore, these findings will help to improve our understanding of normal human lung development and provide a clinically relevant background for research hypotheses generation in projects studying normal or impaired lung development.

In the second part, we describe the generation of **PCLS from human fetal lung** tissue and their potential use in translational research. Here we present the case of a promising gene therapy for a lethal genetic lung disease, where PCLS were instrumental to assess the transduction of a viral vector (Adeno-Associated Virus as a vector for gene therapy in surfactant protein B deficiency) in human lung parenchyma. Providing the proof of concept that this viral vector and its cargo can effectively transfect human lung parenchyma highlights the clinical relevance of using such a viral vector in neonates. PCLS from human lung tissue (healthy or diseased) represent a versatile tool to assess pathophysiology or therapeutics in an organotypic culture. These findings align with the need for more complex and holistic model and the use of human tissue in translational research.

In the third and last part, we present a research project combining a **modified Delphi method** with an “**integrated knowledge translation**” approach to develop a consensus definition for MSC, as well as the reporting guidelines for clinical trials using MSC, in order to improve rigor, transparency, reproducibility, and regulation in MSC research. This project relies on an international multidisciplinary task force, the use of methodologically rigorous consensus building method (Delphi method), and inclusion of stakeholders (knowledge users) from inception. This approach ensures that stakeholders needs and preferences are considered throughout the project and therefore support dissemination and implementation of our findings (integrated knowledge translation). This project highlights several important points for knowledge translation in translational research.

# Results

## I. A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation

This manuscript is under revision in the *European Respiratory Journal* and has been submitted as a preprint on *BioRxiv* (<https://doi.org/10.1101/2021.12.23.473945>).

### 1. Background

The lung is a complex organ relying on up to 40 cell types and two complex intertwined networks (airways and vasculature) to maintain its homeostasis and carry out its vital functions. Hence, *in utero* lung development is a delicate process in which any impairment can lead to congenital lung disease. In addition, lung development can also be compromised in preterm infants, when the immature lung is exposed to *ex-utero* insults. While normal lung development has been extensively studied in animals, little is known about normal human developing lung *in utero*. Moreover, while airway branching morphogenesis is well described, but other compartments of the lung, such as stroma, vascular and neuronal networks, remain poorly explored. Unravelling cellular composition, developmental trajectories, and signaling within the normal human developing lung is critical to better understanding and addressing congenital and prematurity-related lung diseases.

Here, we describe a transcriptome analysis at a single-cell resolution on nine human fetal lungs samples and report its insights into the normal human developing lung.



## 2. Results

We performed a snRNA-seq on 23,251 nuclei isolated from nine human fetal lungs with gestational ages spanning through the pseudoglandular and canalicular stages (14 to 19 weeks of gestation). Using canonical cell lineage markers, we identified 9 different cell types present across all gestational ages. The largest cell populations included 2 stromal cells populations and a distal airway epithelium population. Other identified cell types were vascular and lymphatic endothelium, pericytes, immune cells, ciliated cells, and pulmonary neuroendocrine cells (PNEC). Further subclustering analyses were performed to explore the largest cell populations (stroma and distal airway epithelium) at a higher resolution. To identify genes differentially expressed in response to GA, we performed a Differential State Analysis (DSA). Based on results from DSA and enrichment analysis, we identified specific biological pathways at early and late gestational ages for each cell population. Next, we performed a ligand-receptor interaction analysis to identify cellular cross talk within the developing lung. We described crosstalk between diverse populations, and also crosstalk within the population itself. Enrichment analysis using the identified predicted target genes showed that the cellular cross talk was related to general biological pathways, such as *NOTCH* and *TGF $\beta$*  signaling pathways. In addition, more specific pathways involved in angiogenesis, neurogenesis and immune system regulation were also identified. Lastly, spatial validation of our findings in early and late gestational ages will be done by FISH.

## 3. Conclusion

We described here the cellular composition, fate, and cross talk within normal human developing lungs. These findings will provide a clinically relevant background for the

generation of research hypotheses in projects studying normal or impaired lung development and will help to develop and validate surrogate models to study human lung development.

**4. Manuscript: “A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation” (under review in European Respiratory Journal)**

# **A single-cell atlas of human fetal lung development between 14 and 19 weeks of gestation**

Laurent Renesme<sup>1,2,3</sup>, Flore Lesage<sup>3</sup>, David P. Cook<sup>4</sup>, Shumei Zhong<sup>3</sup>, Satu M. Hänninen<sup>5</sup>, Olli Carpén<sup>5</sup>, Ivana Mižíková\*<sup>6</sup>, Bernard Thébaud\*<sup>3, 7, 8</sup>.

<sup>1</sup> Neonatal Intensive Care Unit, University Hospital of Bordeaux, France

<sup>2</sup> INSERM, Centre de Recherche Cardio-Thoracique de Bordeaux, U1045, F-33600 Pessac, France

<sup>3</sup> Sinclair Center for Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>4</sup> Lunenfeld-Tanenbaum Research Institute, Toronto, Ontario, Canada

<sup>5</sup> Precision Cancer Pathology, department of Pathology and research Program in System Oncology, University of Helsinki and HUS Diagnostic Center, Helsinki University Hospital, Helsinki, Finland

<sup>6</sup> Experimental Pulmonology, Department of Pediatrics and Adolescent Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

<sup>7</sup> Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada

<sup>8</sup> Neonatology, Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, Ottawa, ON, Canada

\* These authors contributed equally: Ivana Mižíková and Bernard Thébaud

**Corresponding author:** Dr. Bernard Thébaud

Senior Scientist, Ottawa Hospital Research Institute;  
Professor of Pediatrics, University of Ottawa;  
Staff Neonatologist at The Ottawa Hospital and CHEO  
401 Smyth Road/501 Smyth Road  
Ottawa, Ontario  
Phone: 613-737-8899 ext: 73905  
Email: [bthebaud@toh.ca](mailto:bthebaud@toh.ca)

## **TAKE HOME MESSAGE**

Using a single-cell transcriptomic approach (single-nuclei RNA sequencing), we describe here, the cellular landscape, cell developmental trajectories, and cell-to-cell communication in the developing human lung during the pseudoglandular stage.

## ABSTRACT

**Rationale.** Human lung development has been mainly described in morphologic studies and the potential underlying molecular mechanisms were extrapolated from animal models. Therefore, there is a need to gather knowledge from native human lung tissue. In this study we describe changes at a single-cell level in human fetal lungs during the pseudoglandular stage.

**Methods.** We report the cellular composition, cell trajectories and cell-to-cell communication in developing human lungs with single-nuclei RNA sequencing (snRNA-seq) on 23,251 nuclei isolated from nine human fetuses with gestational ages between 14 to 19 weeks of gestation.

**Results.** We identified nine different cell types, including a rare pulmonary neuroendocrine cells population. For each cell type, marker genes are reported, and selected marker genes are used for spatial validation with fluorescent RNA *in situ* hybridization. Enrichment and developmental trajectory analysis provide insight into molecular mechanisms and signaling pathways within individual cell clusters according to gestational age. Lastly, ligand-receptor analysis highlights determinants of cell-to-cell communication among the different cell types through the pseudoglandular stage, including general developmental pathways (NOTCH and TGFB), as well as more specific pathways involved in vasculogenesis, neurogenesis, and immune system regulation.

**Conclusion.** These findings provide a clinically relevant background for research hypotheses generation in projects studying normal or impaired lung development and help to develop and validate surrogate models to study human lung development, such as human lung organoids.

## INTRODUCTION

Human lung development is traditionally divided into 5 stages: embryonic (1-7 weeks of gestation (WG)), pseudoglandular (5-17 WG), canalicular (16-26 WG), saccular (24 to 38 WG), and alveolar (36 WG to 3 years, with late alveolarization continuing up to early adulthood) [1]. While these stages have been defined morphologically and histologically during human fetal development, very little is known about the underlying molecular and cellular mechanisms [2, 3]. Detailed insight into the molecular pathways and cellular communication which contribute to lung development would improve our understanding of lung repair as it may recapitulate fetal lung development and would enable us to develop new therapeutic targets [4, 5]. Due to developmental similarities in human and rodent lungs, mouse and rats have been used extensively to study lung development [6]. Airway branching morphogenesis in rodents is well described, and gain- and loss-of-function studies using genetically modified animals have identified the critical factors of this developmental process [3, 7]. However, some compartments of the lung, such as stroma or vascular and neuronal networks, remain poorly explored during the fetal developmental period despite their potential key roles in lung development [8, 9]. While versatile in experimental settings, rodent models of lung development have several limitations, including species-specific maturation and function at birth, and molecular regulation [10, 11]. Therefore, there is a need to unravel cellular composition, developmental trajectories, and signaling within the developing human lung.

Single-cell and single-nuclei RNA sequencing (scRNA-seq and snRNA-seq, respectively) can reveal complex and rare populations, uncover regulatory relationships, and track distinct cell lineages in development. In addition, computational methods have been developed to analyze cellular communications and to determine which extracellular signals influence the cell population of interest.

In the present study, we report the cellular composition, developmental pathways and cell-to-cell communication in fetal human developing lungs using snRNA-seq approach on

23,251 nuclei isolated from nine lungs with gestational ages (GAs, described in WG) spanning through the pseudoglandular and canalicular stages (14 to 19 WG). We describe nine different cells populations, including rare cell population of pulmonary neuroendocrine cells (PNEC). For each identified population we report marker genes and enrichment and developmental state analysis. Additionally, we provide a spatial validation for selected populations using fluorescent RNA *in situ* hybridization (FISH). Lastly, ligand-receptor analysis highlights determinants of cell-to-cell communication among the different cell types.

## **METHODS**

A full description of the experimental procedures can be found in the supplementary methods.

### ***Human fetal tissue collection***

De-identified human fetal lung samples were obtained under REB approval by The Ottawa Hospital Review Ethical Board (20170603-01H). Lung tissue was snap frozen and stored at -80°C.

### ***Single-nuclei suspension preparation***

Single nuclei isolation was performed according to Martelotto *et al.* [12] with minor adjustments. Sample multiplexing was performed using the MULTI-seq protocol as reported before [13, 14]. Nuclei integrity was confirmed by fluorescent microscopy (Axio Imager M2, Carl Zeiss, Toronto, ON, Canada) and single diploid nuclei were sorted using a flow cytometer (BD LSR Fortessa, Beckton Dickinson Biosciences, Franklin Lakes, NJ, USA) (Supplemental figure 1a-b)

### ***Single-nuclei sequencing***

Cell capture and library production was performed with the Chromium system (10X Genomics, Pleasanton, CA, USA). Sequencing was performed with NextSeq500 (Illumina, San Diego, CA, USA). Raw sequencing reads were processed using CellRanger v3.1.0. All analyses were performed with Seurat v4.0.0[15]. SCTransform() was used to normalize samples, select highly variable genes, and to regress out cell cycle and cell stress effects. To eliminate batch effects or biological variability effects on clustering, the data integration method implemented by Seurat for SCTransform-normalized data was performed, using the SelectIntegrationFeatures(), PrepSCTIntegration(), FindIntegrationAnchors(), and

IntegrateData() functions. Differential State Analysis (DSA) was performed using muscat R package[16]. Gene set enrichment analysis (GSEA) was performed with Metascape[17]. Cell communications were inferred using NicheNet (v1.0.0) R package [18].

### ***In situ hybridization***

FISH was performed on formalin-fixed, paraffin embedded human fetal lungs. 3µm thick sections were analysed using the RNAscope technology following the manufacturer's protocol, and as previously described [14].



## RESULTS

### ***Cellular landscape of the human developing lung***

To characterize the cellular composition in developing human lungs, we generated snRNA-seq profiles of 23,251 nuclei isolated from nine fetal lung samples with GAs between 14+1 and 19+0 weeks (Figure 1a, and Supplemental figure 1c). We identified nine distinct cell populations, present across all GAs (Figure 1b-d), and described the lung cellular composition variation over time. While proportions of both stromal populations decreased over time (83.7 and 6.1% at 14+1 weeks vs. 47.6 and 4% at 19+0 weeks), the size of distal airway epithelium cluster increased (from 3.6% to 39.5%) (Figure 1c and Supplemental table 1). Each cell cluster displayed a unique molecular signature characterized by multiple DEGs (Figure 1e and Supplemental table 2), and cell populations were annotated based on the expression of canonical cell lineage markers (Figure 1f and 2a). Additionally, among the top ten DEGs in some clusters, we identified less commonly used transcriptional patterns (Figure 1e). For example, pericytes expressed genes involved in the cGMP-PKG pathway such as *MEF2C*, *PDE3A*, *TRPC6*, and *PLCB*. Vascular and lymphatic endothelium expressed genes involved in angiogenesis, respectively *EPAS1*, *PTPRB*, *CALCRL* and *NRP2*, *VAV3*, *STAB2*. Finally, the immune cells cluster expressed *HDAC9*, *SAMSN1* and *DOCK8* that are involved in lymphocyte activation.

To further confirm clusters identities, and to gain further insights into the underlying molecular functions and biological processes, we performed an GSEA. Cell-type specific terms were associated with all the identified clusters. As expected, stromal populations 1 and 2 were associated with general terms such as extracellular matrix and mesenchyme development. Distal airway epithelium was associated with respiratory gaseous exchange, as well as apical and basolateral plasma membrane terms, while ciliated cells were characterized by cilium-specific pathways. Blood vessel development was associated with both, pericytes and vascular endothelium, while lymph vessel development pathways were specific to lymphatic endothelium cluster. Immune cells cluster was characterized by

interleukin signalling and myeloid cell differentiation pathways. Finally, transsynaptic signaling was associated specifically with pulmonary neuroendocrine cells (Figure 2b, Supplemental table 3).

### ***Developmental changes in gene expression in developing lung stromal cells***

Within the stromal population (stroma 1 and 2), we identified three subclusters with distinct gene expression patterns represented across different GAs (Figure 3a-e, Supplemental table 4). Stromal subclusters 0 and 1 expressed traditional matrix fibroblast markers (*FN1*, *MEOX2*) (Figure 3f). Pulmonary fibroblasts can be typically segregated into two subtypes based on the distinct expression of *COL13A1* and *COL14A1* [19]. In our dataset the expression of *COL13A1* was most prominent in stromal subcluster 0, which was also characterized by expression of additional genes associated with *COL13A1*<sup>+</sup> fibroblasts (*ITGA8*, *LIMCH1*, *MYLK*, *PLXDC2*, *MACF1*) (Figure 3f). A sizeable fraction of subcluster 1 expressed *COL14A1*, as well as other matrix fibroblast markers (*FBLN1*, *COL1A2*, *AKAP12*, *COL1A1*) (Figure 3f). Stromal subcluster 2 showed similarities to mesenchymal progenitors as described by *Xie et al.* [19] and the LungGENS database [20], characterized by the expression of *SDC2* and *SMARCC1* (Figure 3f). However, this sub-population did not express any proliferative genes such as *TOP2A* or *MKI67* which are typically expressed in self-renewing progenitor populations. GSEA of stromal subclusters revealed supportive roles for the stromal cells in lung and mesenchyme development, with subcluster 0 and 1 being associated with regulation of organ architecture (connective tissue development, cell junction assembly, cytoskeleton organization, negative regulation of locomotion/cell motility) and subcluster 2 being associated with proliferative/stemness functions (DNA binding transcription factor activity, negative regulation of cell differentiation) (Supplemental figure 2a and c, Supplemental table 5).

Next, we performed DSA-based GSEA (Supplemental tables 6, 7 and 8). In the two main stromal clusters, Stroma 1 cluster was associated with response to growth factor and

MTORC1 signaling for early GA (14+1 to 14+3 WG), and regulation of cell communication by electrical coupling, regulation of vascular associated smooth muscle contraction for late GA (18+2 to 19+0 WG) (Figure 3g, Supplemental figure 3). For stroma 2, prominent regulated enriched pathways included TNF $\alpha$  signaling via NF $\kappa$ B for early GA, as well as muscle structure development and Kit receptor signaling pathway for late GA (Figure 3h, Supplemental figure 3).

### ***Developmental changes in gene expression in developing airway epithelium***

Upon subclustering of the distal airway epithelium, we identified three subclusters with distinct gene expression represented across the different GAs (Figure 4a-e, Supplemental table 4). Epithelial subclusters shared common marker genes with bud tip progenitors (subcluster 0: *SFTPC*, *ETV5*), bud tip adjacent (subcluster 1: *DMD*, *MECOM*), and secretory progenitor (subcluster 2: *SCGB3A2*, *CFTR*) epithelial cells as identified by *Miller et al.* [21] in human fetal lung tissue of similar GA (Figure 4f). In addition, a GSEA was performed with DEG from these three subclusters (Supplemental figure 2b and d, Supplemental table 5). Based on DEG and DSA results, we selected gene markers for spatial validation on the fetal lung tissue by FISH (Supplemental tables 2 and 6). Due to their cluster specificity, we selected *SEMA3C* and *SFTPB* as potential early and late distal airway epithelium markers in fetal lung (Supplemental figure 4a). As seen by FISH (Figure 4g), distal airway epithelium (marked by *FGFR2*<sup>+</sup>) in earlier GAs co-expresses *SEMA3C*, with only few cells also co-expressing *SFTPB*. However, the majority of *FGFR2*<sup>+</sup> epithelial cells surrounding the airway lumen at later GAs were *SEMA3C*<sup>-</sup> and *SFTPB*<sup>+</sup>. Notably, an adjacent layer of *FGF2*/*SEMA3C*<sup>+</sup>/*SFTPB*<sup>-</sup> cells can also be seen (Figure 4g). Finally, based on the DSA, we performed a GSEA. Among the most regulated pathways in distal airway epithelium were sensory organ development and blood vessel development (Figure 4h, Supplemental figure 3, Supplemental tables 7 and 8).

In the case of ciliated cells and PNECs, DSA identified only few DEGs, confirming previous reports on the early specification of these cell populations (Supplementary figures 5 and 6, respectively) [22, 23]. Interestingly, only a very small fraction of the PNEC population showed high expression of typical PNEC markers such as *GRP1*, *CALCA*, *ASCL1*, and *SYN* [21], whereas markers associated with the synapse, signal release and transsynaptic signaling (*NRXN1*, *PTPRN2*, *PCLO*, *CADPS*) were expressed in the majority of the PNECs in our dataset. FISH was performed using *NRXN1* and *GRP* as PNEC markers (Supplemental figure 4b). The *NRXN1*<sup>+</sup>/*GRP*<sup>+</sup> PNECs were present at all GAs and were either adjacent to or localized within the airway epithelium. Both, isolated cells and small cells clusters (neuroendocrine bodies) were noted. Interestingly, while all *GRP*<sup>+</sup> cells were *NRXN1*<sup>+</sup>, *GRP*/*NRXN1*<sup>+</sup> cells were also noted, suggesting a possible existence of two subpopulations of PNECs in the developing lung (Figure 5).

### ***Developmental changes in gene expression in developing endothelial, pericyte and immune cells populations***

Vascular endothelium at early GA (14+3 WG) was characterized by cell type specific pathways such as VEGFA-VEGFR2 signaling pathway, endothelium development and vasculature development (Figure 6a, Supplemental figure 3). In the lymphatic endothelium, genes expressed at early GAs were involved in ubiquitin protein ligase binding and IL-6 signaling. Interestingly, ubiquitination plays a central role in NOTCH signaling [24] and IL-6 is involved in lymph angiogenesis via VEGF-C [25], whereas later GAs were characterized by expression of genes involved in “cholesterol biosynthesis” (Supplemental figure 7). Next, to visualize the vascular endothelial cells in the fetal lung we performed a FISH (Figure 6b). Based on the DSA (Supplemental table 6), we selected *KITLG* as an early, and *CAV1* as a late marker for vascular endothelium during the pseudoglandular phase (Supplemental figure 4c). However, both *KITLG*<sup>+</sup>/*PECAM*<sup>+</sup> and *CAV1*<sup>+</sup>/*PECAM*<sup>+</sup> double positive cells could be found at all investigated GAs (Figure 6b).

The pericytes cluster showed gene expression associated with regulation of vascular wound healing and cellular response to hormone stimulus at early GAs. In addition, cellular components involved in pericytes functions, such as filopodium and Z disc were described at later GAs (16+0 and 16+3 WG, respectively) [26, 27] (Figure 6c). *PDGFRB*<sup>+</sup> pericytes (Supplemental figure 4d) could be found in immediate proximity to vascular endothelium (defined by the expression of *PECAM1*<sup>+</sup>) in the lungs throughout the pseudoglandular development (Figure 6d).

Finally, in the case of immune cells population, the DEGs at early GAs were associated with “lymphocyte activation” and “TNF $\alpha$  signaling via NF $\kappa$ B” which is known to be crucial in lymphocyte activation, as well the specification of innate and adaptive immunity [28] (Supplemental figure 8).

### ***Cellular crosstalk in the human developing lung***

In order to understand cell communication networks within the developing human fetal lung samples, we performed a ligand-receptor interaction analysis using NicheNet [18]. The analysis focused on time-dependent changes in cell communication rather than stable cell crosstalk, and the set of gene of interest used for the analysis was defined by differential expression between early GAs (14+1 and 14+3 WG, reference condition) and late GAs (18+0 and 19+0 WG, condition of interest). Cellular crosstalks identified at later GAs are presented in Figure 7, with the identified receptors and predicted target genes involved (Figure 8). GSEA on predicted target genes are presented in Supplemental table 9. Inferred cellular communications at later GAs were associated with general pathways such as NOTCH and TGFB pathways, as well as in pathways involved in immune system regulation, and neurogenic tissue and vasculature development.

We identified key factors in the NOTCH signaling pathway such as *DLL1* (expressed by the ciliated cells) and its receptors *NOTCH2* (expressed in stroma 1 and distal airway epithelium) and *NOTCH3* (distal airway epithelium). Predicted target genes were associated

with T cell differentiation and TNF signaling in stroma 1 cluster, and with cellular senescence in airway epithelium. Moreover, *DLL4* expressed by the vascular endothelium targeted changes in gene expression associated with INF $\gamma$ -mediated signaling and leukocytes cell-cell adhesion in lymphatic endothelium via its receptor *NOTCH1*. *ADAM17*, an important NOTCH signaling activator, was expressed by majority of cell populations, while its receptor *ITGB1* was expressed by distal airway epithelium. Potential *ADAM17* target genes were primarily associated with inflammatory response.

*TGFB1* was expressed by lymphatic and vascular endothelium and the immune cells. Its receptors, *TGFBR1* and *TGFBR2*, were expressed in distal airway epithelium and the predicted target genes were associated with cellular senescence and gastrin signaling. *BMP5*, another member of the TFGF family, was expressed in both stromal clusters, while its receptors (*BMPR2*, *BMPR1A*, *ACVR2A* and *ACVR1*) were specifically expressed by stroma 2. The targeted genes were associated with insulin resistance, T cell differentiation, IL-18 signaling, and with regulation of cell adhesion and motility.

*CD44* was expressed in stroma 1 and immune cells and its potential ligands HAS2 and FN1 were expressed by stroma 1 and 2 clusters. Predicted target genes for CD44 activation in stroma 1 were associated with T cell differentiation and regulation of adaptive immune response. Target genes in immune cells cluster were associated with IL-18 and NF-kappa B signaling pathways.

Further regulated were also genes involved in neurogenic tissue development, including stroma-expressed *NTNG1*, and its receptor *LRRC4C* (stroma1, pericytes and lymphatic endothelium), *NEGR1* (expressed by ciliated cells and PNECs) and its receptor *NEGR1* (stroma 1), as well as *NCAM1* (stroma, lymphatic, and PNEC clusters) and its receptor *ROBO1* (distal airway epithelium and vascular endothelium). TNC, an extracellular matrix protein implicated in the guidance of migrating neurons, was expressed in stroma 2 and distal airway epithelium clusters, with its receptors expressed in distal airway epithelium and vascular endothelium.

Lastly, in regard to vasculature development, VEGF ligands were expressed by distal airway epithelium (*VEGFA*), vascular endothelium (*VEGFC*), and PNEC (*VEGFB*). *VEGFA* receptors were expressed in pericytes (*NRP1*, *NRP2*, *FLT1*) with predicted target genes associated with TGF-beta receptor signaling and positive regulation of cell migration, and in lymphatic endothelium (*NRP1*, *NRP2*, *KDR*) with predicted target genes associated with INF $\gamma$ -mediated signaling and regulation of leukocytes cell-cell adhesion. *VEGFB* receptor *NRP1* was expressed in stroma 1, while *NRP2* was expressed in immune cells, and the predicted target genes were associated with NF-kappa B signaling. In addition, we reported the expression of *SEMA3E* and its receptor *PLXND1*, which are required for normal vascular patterning during embryogenesis. *SEMA3E* was expressed by the airway epithelium (ciliated cells, PNEC and distal airway epithelium) and *PLXND1*, was expressed by the vascular endothelium.

## DISCUSSION

This is the first report of an unbiased transcriptomic analysis of human fetal lung samples, providing a detailed cell atlas of the human developing lung during pseudoglandular and early canalicular stages at a single-cell resolution. We identified 9 cell types, including rare cell type PNEC, and described developmental changes in gene expression and cell-cell communication across GAs. In addition to canonical marker genes, we identified an array of marker genes for each cell population and validated the spatial localization of specific cell populations such as PNEC and pericytes in the human developing lung. Lastly, the cell communication analysis provided unique insight into detailed cellular crosstalk and main pathways involved in the pseudoglandular stage of lung development.

Only few studies to date have reported the use of single-cell transcriptomic on human developing lung, typically including a small number of samples or focusing on specific cell types [21, 29, 30]. Our study reported transcriptomics data from 9 fetuses, representing each developmental week from 14 and 19 WG, allowing detailed analyses of gene expression changes overtime. Additionally, choosing single-nuclei isolation from frozen material, rather than single-cell isolation from fresh tissue, enabled us to isolate a relatively large number of distal airway epithelial cells (up to 39.5% of all cells at 19+0 WG). This allowed us to show a steady increase in distal airway epithelial cells across GAs in relation to airway development, whereas previous studies using single-cell isolation approach seemed more prone to dissociation bias (with epithelial cells representing less than 4% of the total cells) [29, 30]. These findings suggest that transcriptomic analysis of tissue with high stromal content, such as human fetal lungs, might be more efficient when using a single-nuclei, rather than a single-cell isolation preparation.

The main strength of this study lies within the detailed description of the cellular crosstalk changes between early (14+1 and 14+3 WG) and late (18+2 and 19+0 WG) GAs. We identified several critical pathways associated with normal lung development and lung disease. For instance, VEGF plays a central role in lung development and maintenance, and



is involved in the pathogenesis of many lung diseases, including emphysema, chronic obstructive pulmonary disease, bronchopulmonary dysplasia, pulmonary hypertension, acute lung injury, and asthma [31]. ADAM17, a NOTCH activator, plays a key role in lung inflammation regulation by increasing epithelium and smooth muscle cells permeability, secretion of inflammatory mediators, and trans-endothelial leukocyte migration, and its deficiency leads to decreased airway branching [32]. In addition, the role of the neurogenic tissue and vasculature in the developing lung is rapidly gaining interest [8, 9]. Here we have identified key factors for neurogenic tissue development also known to be involved in abnormal lung development, such as *ROBO1* (expressed in distal airway epithelium and vascular endothelium) and TNC (expressed in stroma 2 and distal airway epithelium). *Robo1* knock-out mice exhibit delayed lung maturation with increased mesenchymal cellularity and reduced terminal air spaces leading to respiratory distress and death at birth [33]. TNC inactivation in mice induces abnormal lung development and persistent abnormal lung function after birth [34]. Lastly, we described crosstalk between immune cells and stroma populations via CD44-HAS2. CD44 plays an important role in the regulation of lung fibroblasts senescence and apoptosis and the interaction HAS2-CD44 can induce pulmonary fibrosis via its receptor [35]. Immune cells population in our dataset were not well differentiated but additional data suggest that lung immune cells complexity and heterogeneity increase later during the gestation and after birth and play an important role in stroma remodelling and angiogenesis in the developing lung [36].

Distal airway epithelial cells were among the most active signal-receiving cell types in our dataset, displaying active communication with all other cell populations including themselves. Our dataset further identifies a cluster of rare PNEC, and its interaction with both stroma 1 and 2, pericytes and vascular endothelium populations. The exact role of PNEC in lung development is yet to be described [37]. However, GRP expression had previously been detected in human fetal lung at 8 WG with peak of expression at mid-gestation. Moreover, elevated urine GRP was previously associated with the development of

BPD suggesting an important role for PNECs in normal and impaired lung development [38]. In addition, FISH identified 2 populations of PNEC within the fetal lung tissue (*GRP<sup>+</sup>/NRXN1<sup>+</sup>* and *GRP<sup>-</sup>/NRXN1<sup>+</sup>* cells). These findings suggest that *NRXN1* might be a more specific marker for PNEC than *GRP*.

Our study has several limitations. Firstly, the study doesn't consider the potential effects of fetus sex on the lung development whereas sex-related differences in human fetal lung transcriptome during the pseudoglandular stage had been previously reported [39]. However, the same study also found GA to have a more dominant effect on transcriptome than sex. Secondly, we assumed the fetuses were healthy as they were issued from selective abortions with no fetal indications. Nevertheless, all the fetuses had a GA below 20 weeks of gestation, prior to typical ultrasound morphological assessment for congenital malformation. Additional limitations relate to the downstream analysis. During development, important changes in cell differentiation and signalling can be induced by upregulation or downregulation in gene expression. While the presented DSA approach provides information on up-regulated genes across the different GA, it does not allow to study down-regulation in gene expression.

We report here, for the first time, an unbiased transcriptomic analysis of human fetal lungs during pseudoglandular stage at a single-cell resolution. This transcriptomic approach at a single cell level, combined with other novel approaches such as proteomics and metabolomics analysis [40, 41], are critical to unravel molecular pathways and cell communication in the human developing lung. Altogether, they will provide a clinically relevant background for the generation of novel research hypotheses in studies of normal or impaired lung development, help to develop and validate surrogate models to study human lung development and ultimately identify new therapeutic targets.

## **ACKNOWLEDGEMENTS**

The authors thank the EORLA staff for the human fetal lung sample isolation, especially Janet Stinson and Korrine Hutt-Acres.

## **FUNDINGS**

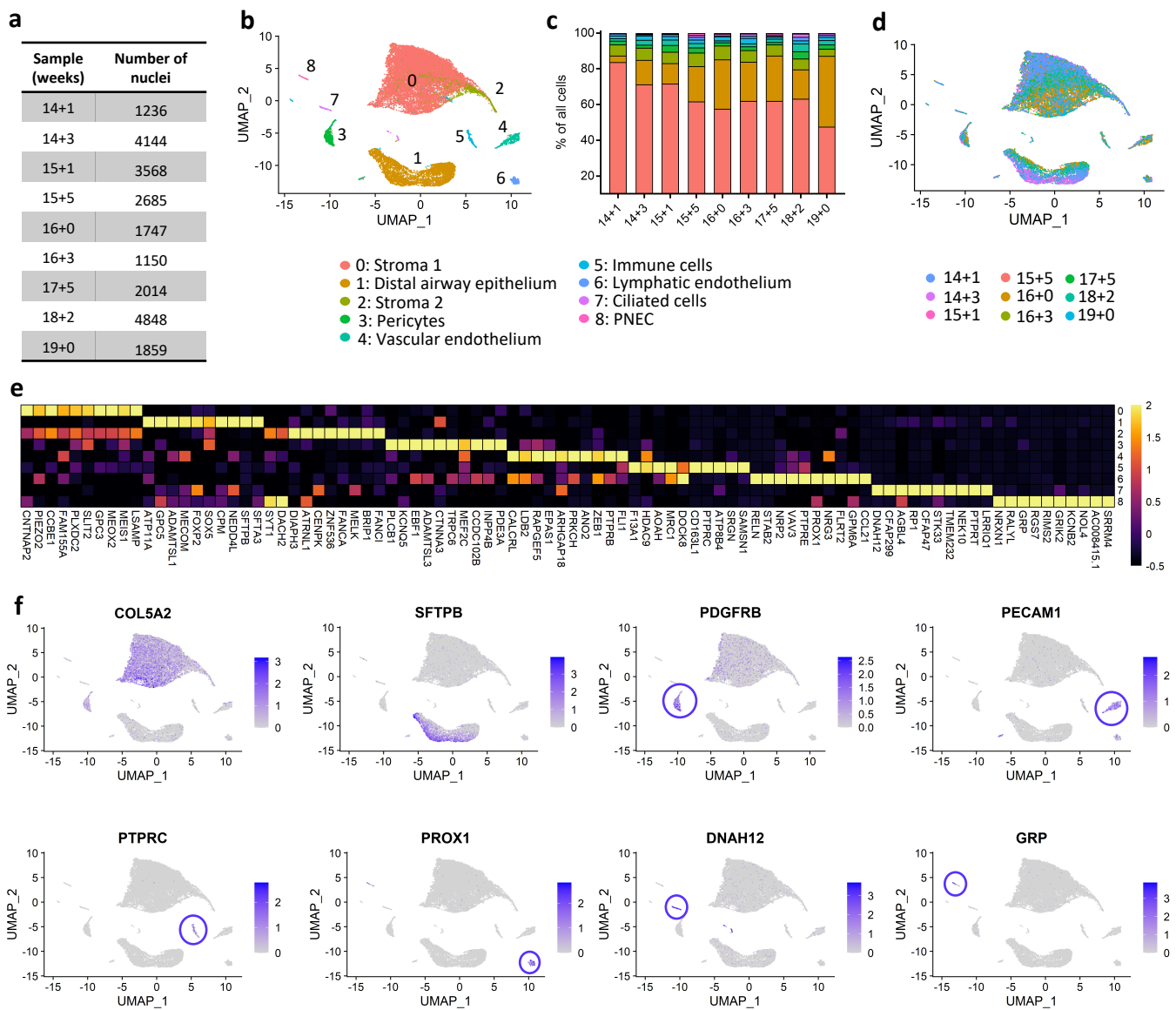
This work was supported in part by funding from the European Respiratory Society (Long Term Research fellowship) to LR, the CHIR Banting Fellowship to DPC, the German Research Foundation (Deutsche Forschungsgemeinschaft) to IM

## REFERENCES

1. Burri PH. Fetal and Postnatal Development of the Lung. *Annu. Rev. Physiol.* 1984; 46: 617–628.
2. Schittny JC. Development of the lung. *Cell Tissue Res* 2017; 367: 427–444.
3. Morrissey EE, Hogan BLM. Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. *Dev Cell* 2010; 18: 8–23.
4. Herriges M, Morrissey EE. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* 2014; 141: 502–513.
5. Ubags NDJ, Alejandre Alcazar MA, Kallapur SG *et al.* Early origins of lung disease: towards an interdisciplinary approach. *Eur Respir Rev* 2020; 29.
6. Nardiello C, Mižiková I, Morty RE. Looking ahead: where to next for animal models of bronchopulmonary dysplasia? *Cell Tissue Res* 2017; 367: 457–468.
7. Cardoso WV, Lü J. Regulation of early lung morphogenesis: questions, facts and controversies. *Development* 2006; 133: 1611–1624.
8. Kina YP, Khadim A, Seeger W *et al.* The Lung Vasculature: A Driver or Passenger in Lung Branching Morphogenesis? *Front Cell Dev Biol* 2021; 8: 623868.
9. Bower DV, Lee H-K, Lansford R *et al.* Airway branching has conserved needs for local parasymphetic innervation but not neurotransmission. *BMC Biol* 2014; 12: 92.
10. Danopoulos S, Shiosaki J, Al Alam D. FGF Signaling in Lung Development and Disease: Human Versus Mouse. *Front Genet* 2019; 10: 170.
11. Nikolić MZ, Caritg O, Jeng Q *et al.* Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *eLife* 2017; 6: e26575.
12. Martelotto L. 'Frankenstein' protocol for nuclei isolation from fresh and frozen tissue for snRNAseq V.3. Available from: <https://www.protocols.io/view/frankenstein-protocol-for-nuclei-isolation-from-f-bqxymxpw>. Date last accessed: January 11, 2022.
13. McGinnis CS, Patterson DM, Winkler J *et al.* MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 2019; 16: 619–626.
14. Hurskainen M, Mižiková I, Cook DP *et al.* Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat Commun* 2021; 12: 1565.
15. Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021; 184: 3573-3587.
16. Crowell HL, Soneson C, Germain P-L *et al.* muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat Commun* 2020; 11: 6077.

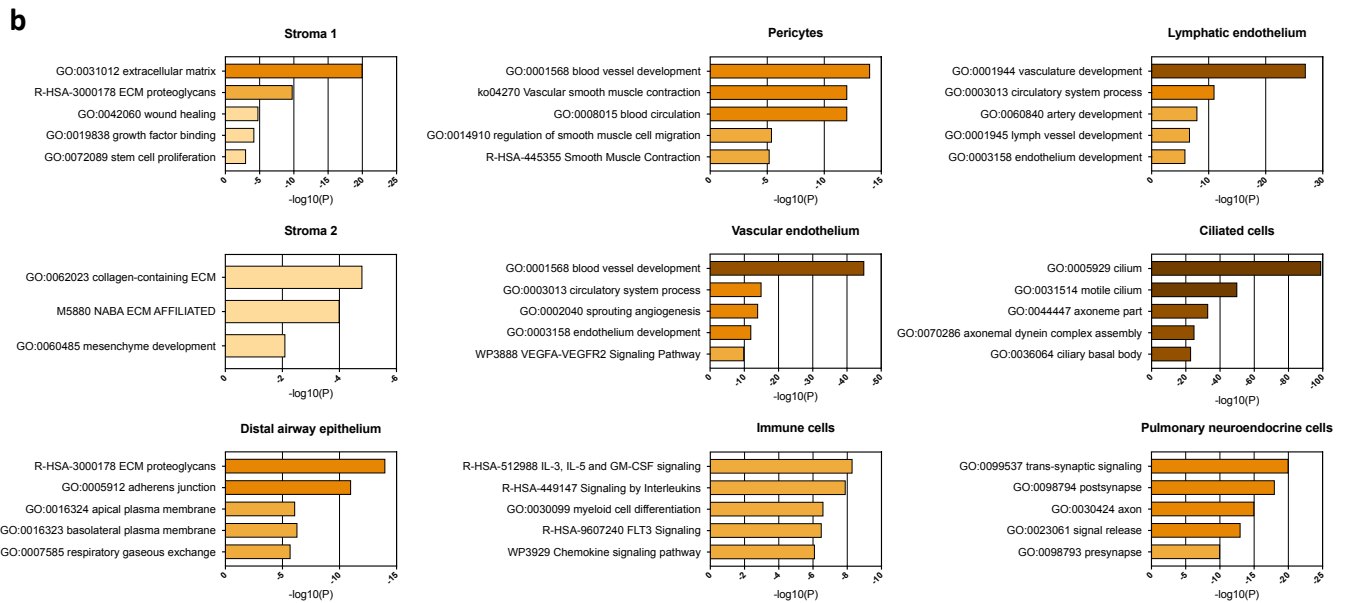
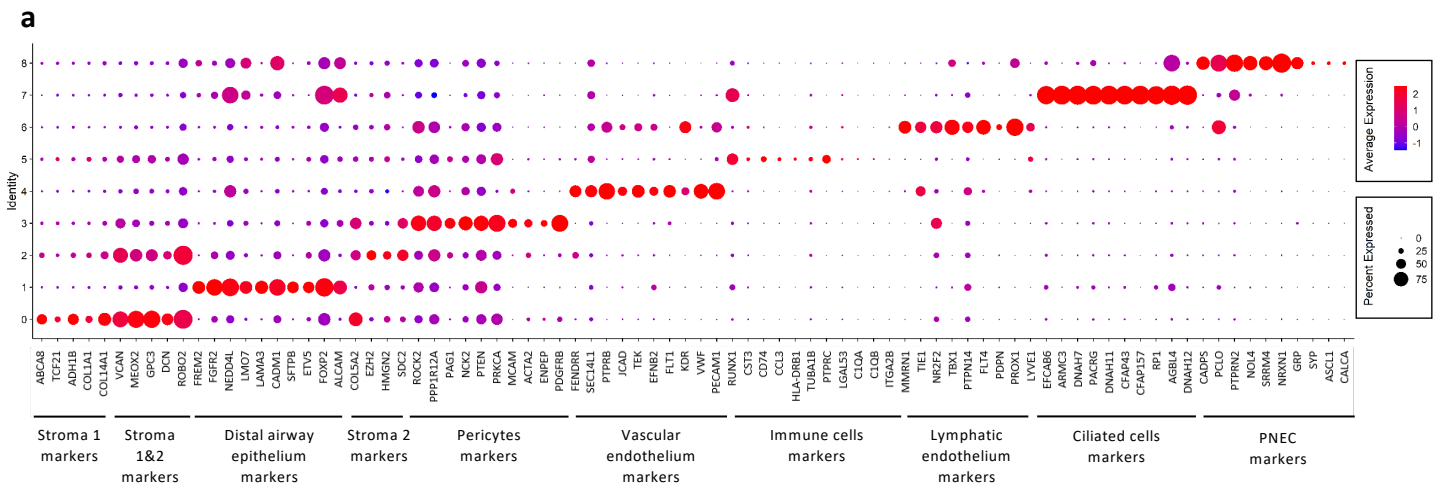
17. Zhou Y, Zhou B, Pache L *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; 10: 1523.
18. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* 2020; 17: 159–162.
19. Xie T, Wang Y, Deng N *et al.* Single-Cell Deconvolution of Fibroblast Heterogeneity in Mouse Pulmonary Fibrosis. *Cell Rep* 2018; 22: 3625–3640.
20. Du Y, Guo M, Whitsett JA *et al.* “LungGENS”: a web-based tool for mapping single-cell gene expression in the developing lung. *Thorax* 2015; 70: 1092–1094.
21. Miller AJ, Yu Q, Czerwinski M *et al.* In Vitro and In Vivo Development of the Human Airway at Single-Cell Resolution. *Dev Cell* 2020; 53: 117-128.e6.
22. Cutz E, Gillan JE, Bryan AC. Neuroendocrine cells in the developing human lung: morphologic and functional considerations. *Pediatr Pulmonol* 1985; 1: S21-29.
23. Portia A. Kreiger. Respiratory Tract – Lung. In: Ernst LM, Ruchelli ED, Huff DS. *Color Atlas of Fetal and Neonatal Histology*. Springer, New York, NY; 2011; pp 21-35. Available from: <http://link.springer.com/10.1007/978-1-4614-0019-6>. Date last accessed: January 11, 2022.
24. Moretti J, Brou C. Ubiquitinations in the Notch Signaling Pathway. *IJMS* 2013; 14: 6359–6381.
25. Huang Y-H, Yang H-Y, Huang S-W *et al.* Interleukin-6 Induces Vascular Endothelial Growth Factor-C Expression via Src-FAK-STAT3 Signaling in Lymphatic Endothelial Cells. *PLoS One* 2016; 11: e0158839.
26. Hung CF, Wilson CL, Schnapp LM. Pericytes in the Lung. *Adv Exp Med Biol* 2019; 1122: 41-58.
27. Yamazaki T, Mukoyama Y. Tissue Specific Origin, Development, and Pathological Perspectives of Pericytes. *Front Cardiovasc Med* 2018; 5: 78.
28. Bonizzi G, Karin M. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol* 2004; 25: 280–288.
29. Danopoulos S, Bhattacharya S, Mariani TJ *et al.* Transcriptional characterisation of human lung cells identifies novel mesenchymal lineage markers. *Eur Respir J* 2020; 55: 1900746.
30. Bhattacharya S, Myers JL, Baker C *et al.* Single cell transcriptomic profiling identifies molecular phenotypes of newborn human lung cells. Available from: <http://biorxiv.org/lookup/doi/10.1101/2020.06.16.156042>. Date last accessed: January 11 2022.
31. Voelkel NF, Vandivier RW, Tuder RM. Vascular endothelial growth factor in the lung. *Am J Physiol Lung Cell Mol Physiol* 2006; 290: L209–L221.
32. Drey Mueller D, Uhlig S, Ludwig A. ADAM-family metalloproteinases in lung inflammation: potential therapeutic targets. *Am J Physiol Lung Cell Mol Physiol* 2015; 308: L325-343.

33. Xian J, Clark KJ, Fordham R *et al.* Inadequate lung development and bronchial hyperplasia in mice with a targeted deletion in the *Dutt1/Robo1* gene. *Proceedings of the National Academy of Sciences* 2001; 98: 15062–15066.
34. Gremlich S, Roth-Kleiner M, Equey L *et al.* Tenascin-C inactivation impacts lung structure and function beyond lung development. *Sci Rep* 2020; 10: 5118.
35. Li Y, Liang J, Yang T *et al.* Hyaluronan synthase 2 regulates fibroblast senescence in pulmonary fibrosis. *Matrix Biology* 2016; 55: 35–48.
36. Domingo-Gonzalez R, Zanini F, Che X *et al.* Diverse homeostatic and immunomodulatory roles of immune cells in the developing mouse lung at single cell resolution. *Elife* 2020; 9: e56890.
37. Xu J, Yu H, Sun X. Less Is More: Rare Pulmonary Neuroendocrine Cells Function as Critical Sensors in Lung. *Dev Cell* 2020; 55: 123–132.
38. Degan S, Lopez GY, Kevill K *et al.* Gastrin-Releasing Peptide, Immune Responses, and Lung Disease. *Annals of the New York Academy of Sciences* 2008; 1144: 136–147.
39. Kho AT, Chhabra D, Sharma S *et al.* Age, Sexual Dimorphism, and Disease Associations in the Developing Human Fetal Lung Transcriptome. *Am J Respir Cell Mol Biol* 2016; 54: 814–821.
40. Clair G, Bramer LM, Misra R *et al.* Proteomic Analysis of Human Lung Development. *Am J Respir Crit Care Med* 2021; .
41. Du Y, Clair GC, Al Alam D *et al.* Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am J Physiol Lung Cell Mol Physiol* 2019; 317: L347–L360.



**Figure 1: Cellular composition of human fetal lung tissue between 14 and 19 weeks of gestation**

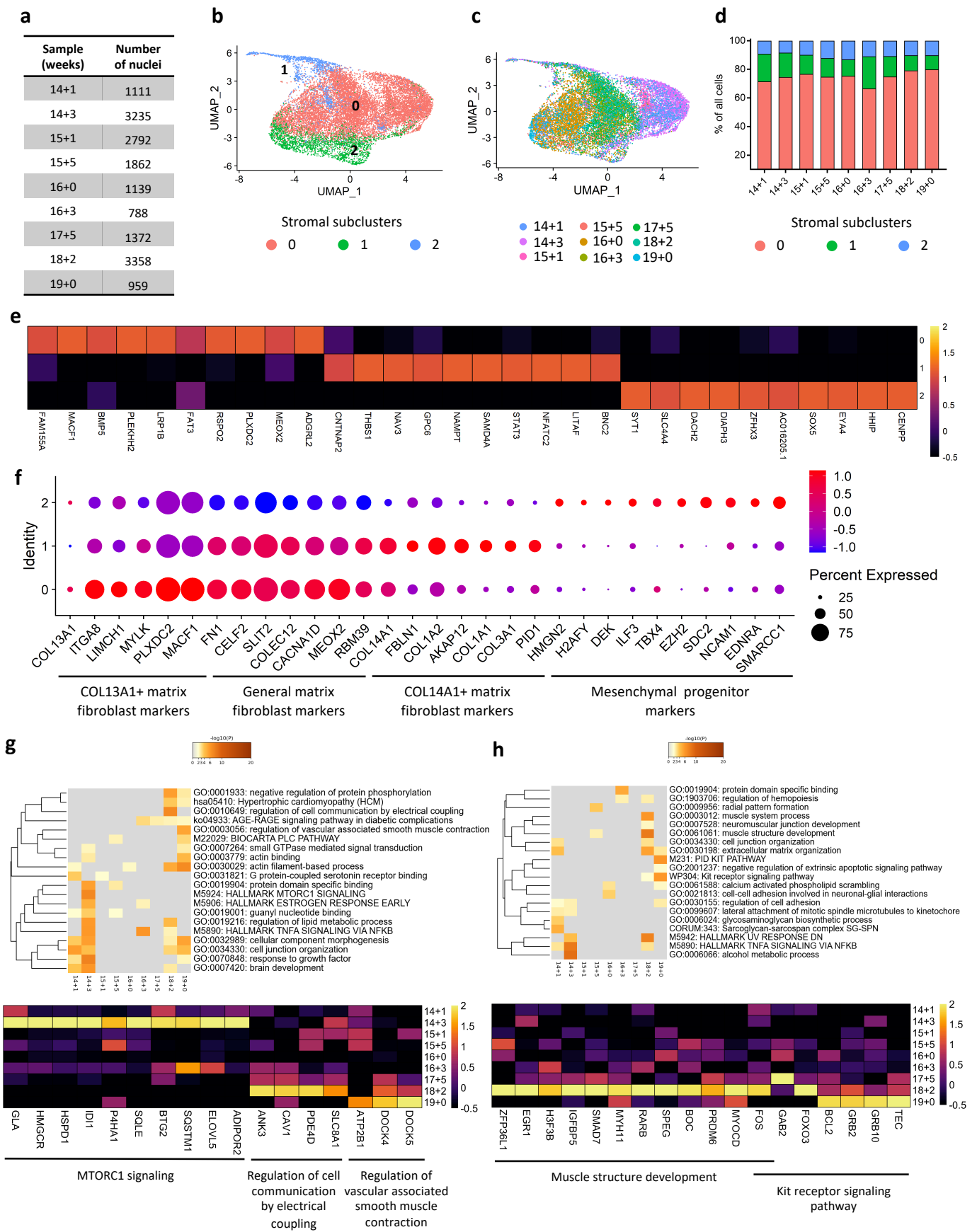
a) Table depicting the number of nuclei sequenced in each sample. b) UMAP plot depicting 9 cell clusters identified by snRNA-seq analysis of fetal lung tissue collected between 14 and 19 weeks of gestation. c) Bar graph depicting the percentual contribution of individual cell populations at different developmental stages. d) UMAP plot depicting the distribution of fetal lung cells based on developmental age. e) Heatmap of top ten most differentially expressed genes across fetal lung cell clusters depicted in panel b. f) UMAP plots showing expression levels for canonical markers of lung cell populations, including stroma (COL5A2), distal airway epithelium (SFTPB), pericytes (PDGFRB), vascular endothelium (PECAM1), immune cells (PTPRC), lymphatic endothelium (PROX1), ciliated cells (DNAH12) and pulmonary neuroendocrine cells (GRP). The intensity of expression is indicated by purple coloring. Expression values in heatmaps and UMAP plots represent Z-score-transformed  $\log(TP10k+1)$  values.  $\log(TP10k+1)$  corresponds to log-transformed UMIs per 10k.



**Figure 2: Identification of cell clusters comprising human fetal lung between 14 and 19 weeks of gestation**

a) Dotplot depicting the expression of most commonly used markers for each cell cluster as described in literature. Expression levels in the dotplot are presented as  $\log(\text{TP10k}+1)$  values.  $\log(\text{TP10k}+1)$  corresponds to log-transformed UMIs per 10k. b) Selected regulated pathways for each cluster identified by gene set enrichment analysis. All terms are significantly enriched (adjusted p-value < 0.05). Plotted values represent  $-\log_{10}(\text{p-value})$ .

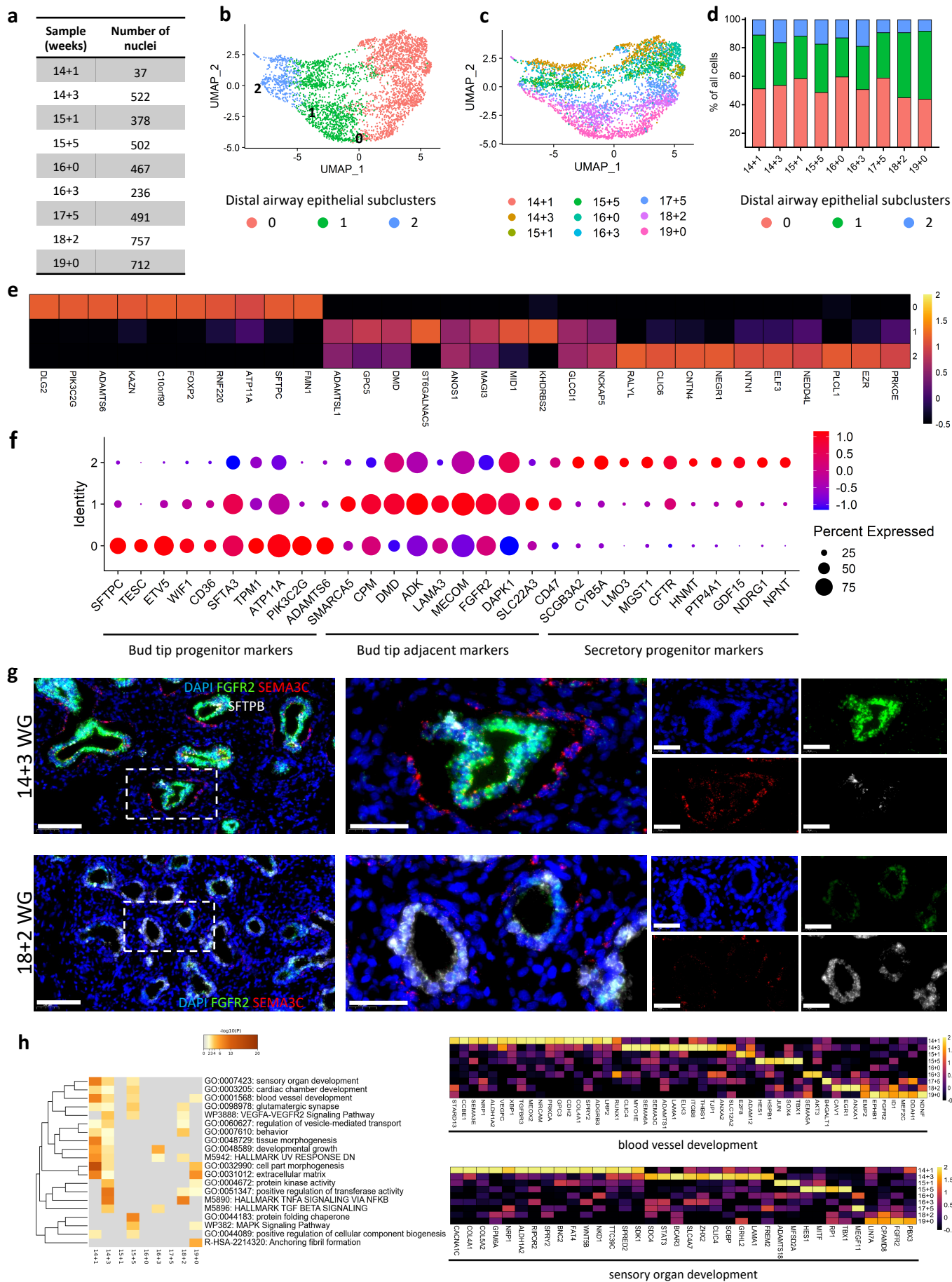




**Figure 3. Cellular composition and developmental changes in gene expression of stromal clusters in human fetal lung between 14 and 19 weeks of gestation**

**Figure 3. Cellular composition and developmental changes in gene expression of stromal clusters in human fetal lung between 14 and 19 weeks of gestation**

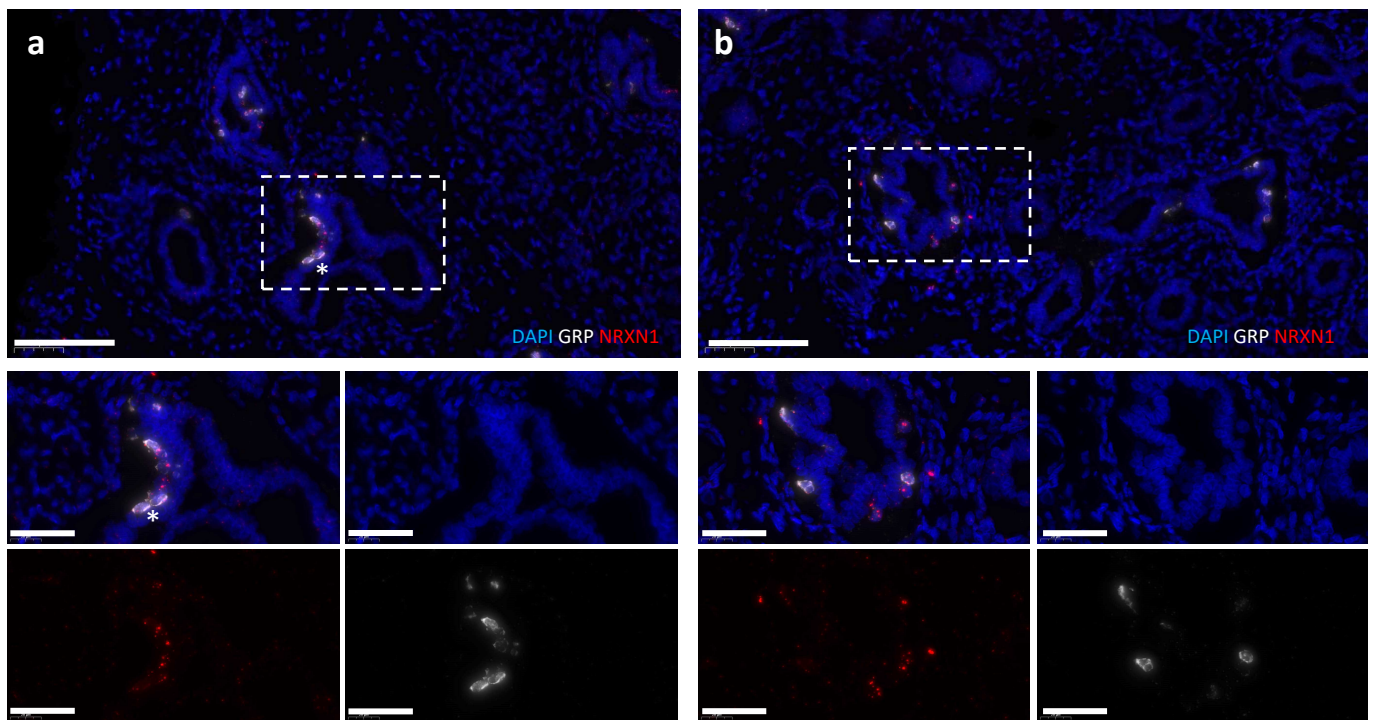
a) Table depicting the number of stromal cells nuclei sequenced in each sample. b) UMAP plot depicting the 3 stromal subclusters identified within the fetal lung stroma. c) UMAP plot depicting the distribution of fetal lung stromal cells based on the developmental age. d) Percentual contribution of the different lung stromal subclusters in each individual sample. e) Heatmap depicting the top ten most differentially expressed genes across fetal lung stromal subclusters shown in panel b). f) Dotplot depicting the expression of stromal markers known to be associated with *COL13A1*<sup>+</sup> matrix fibroblasts (subcluster 0), *COL14A1*<sup>+</sup> matrix fibroblasts (subcluster 1), general matrix fibroblasts (subcluster 0 and 1), and mesenchymal progenitors (subcluster 2) in stromal subclusters. g) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in the main cluster stroma 1 (top), and heatmap depicting the expression level of genes associated with selected enriched terms (bottom). h) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in the main cluster stroma 2 (top), and heatmap depicting the expression level of genes associated with selected enriched terms (bottom). Expression levels in the heatmaps and dotplots are presented as  $\log(\text{TP}10\text{k}+1)$  values.  $\log(\text{TP}10\text{k}+1)$  corresponds to log-transformed UMIs per 10k.



**Figure 4. Cellular composition, developmental changes in gene expression, and spatial localization of distal airway epithelial cells in human fetal lung between 14 and 19 weeks of gestation**

**Figure 4. Cellular composition, developmental changes in gene expression, and spatial localization of distal airway epithelial cells in human fetal lung between 14 and 19 weeks of gestation**

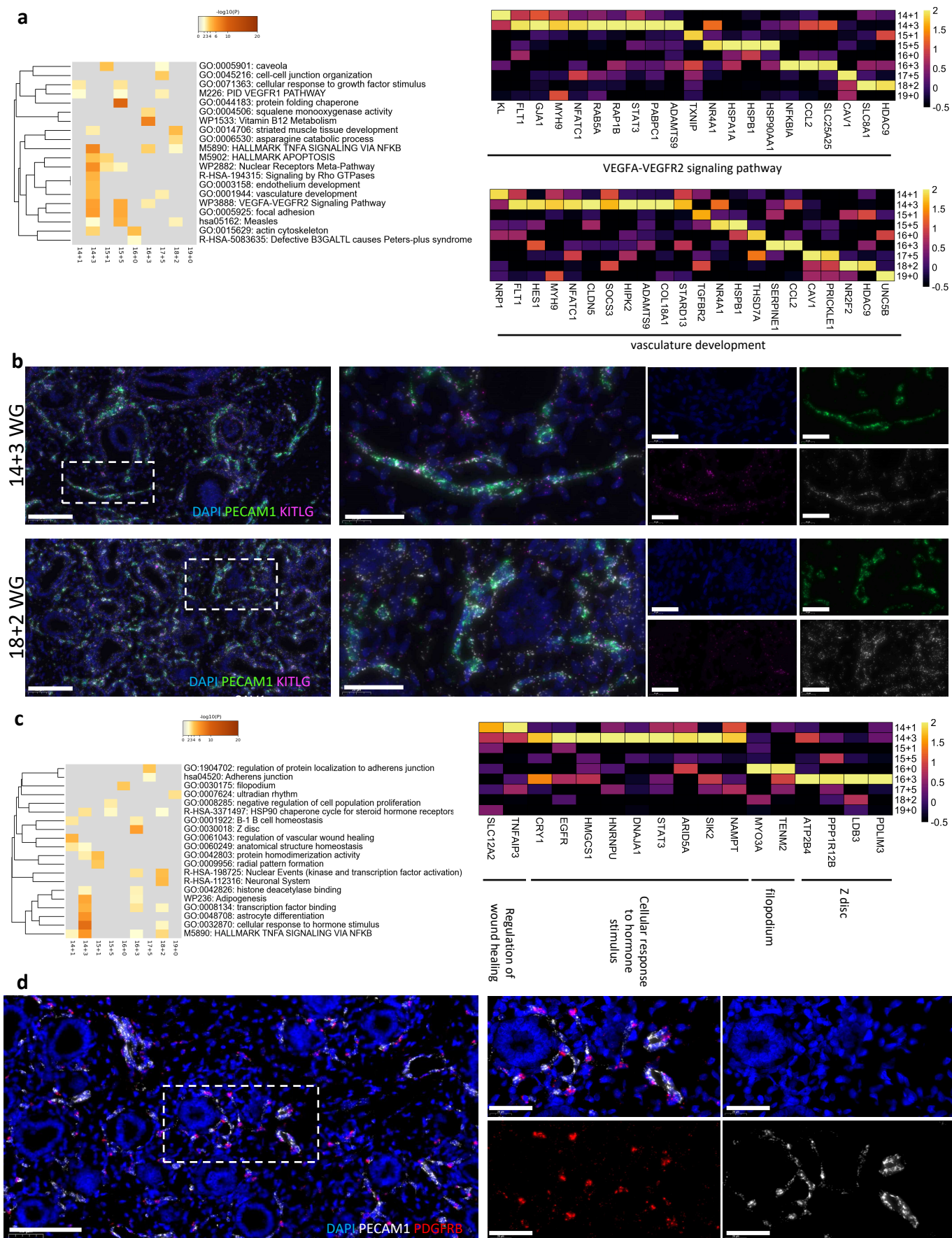
a) Table depicting the number of distal airway epithelial cells nuclei sequenced per sample. b) UMAP plot depicting the 3 epithelial subclusters identified within the distal airway epithelium cluster; c) UMAP plot depicting the distribution of fetal lung distal airway epithelium cells based on developmental age. d) Percentage contribution to the different lung epithelial subclusters per sample. e) Heatmap depicting the top ten most differentially expressed genes across fetal lung epithelial subclusters shown in panel b). f) Dotplot depicting the expression of epithelial markers known to be associated with bud tip progenitor cells (subcluster 0), bud tip adjacent cells (subcluster 1) and secretory progenitors (subcluster 2). g) Representative fluorescent RNA *in situ* hybridization pictures showing co-expression of *FGFR2* (green), *SEMA3C* (red) and *SFTPB* (white) in the airway epithelium in human fetal lungs at 14+3 (top) and 18+2 (bottom) weeks of gestation. Magnification at x20 (scale bar 100  $\mu\text{m}$ ) and x63 (scale bar 40  $\mu\text{m}$ ). Expression levels in the heatmaps and dotplots are presented as  $\log(\text{TP10k}+1)$  values.  $\log(\text{TP10k}+1)$  corresponds to log-transformed UMIs per 10k. h) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in the main cluster distal airway epithelium (left), and heatmap depicting the expression level of genes associated with selected enriched terms (right).



**Figure 5. Fluorescent RNA *in situ* hybridization of pulmonary neuroendocrine cells in the human fetal lung**

Representative fluorescent RNA *in situ* hybridization pictures showing co-expression of *NRXN1* (red) and *GRP* (white) in PNEC in human fetal lungs at a) 14+3 and b) 15+5 weeks of gestation. Both, isolated *NRXN1*<sup>+</sup>/*GRP*<sup>+</sup> PNECs and small PNECs clusters (neuroendocrine bodies, white \*) were noted. Magnification at x20 (top, scale bar 100 μm) and x63 (bottom, scale bar 40 μm).

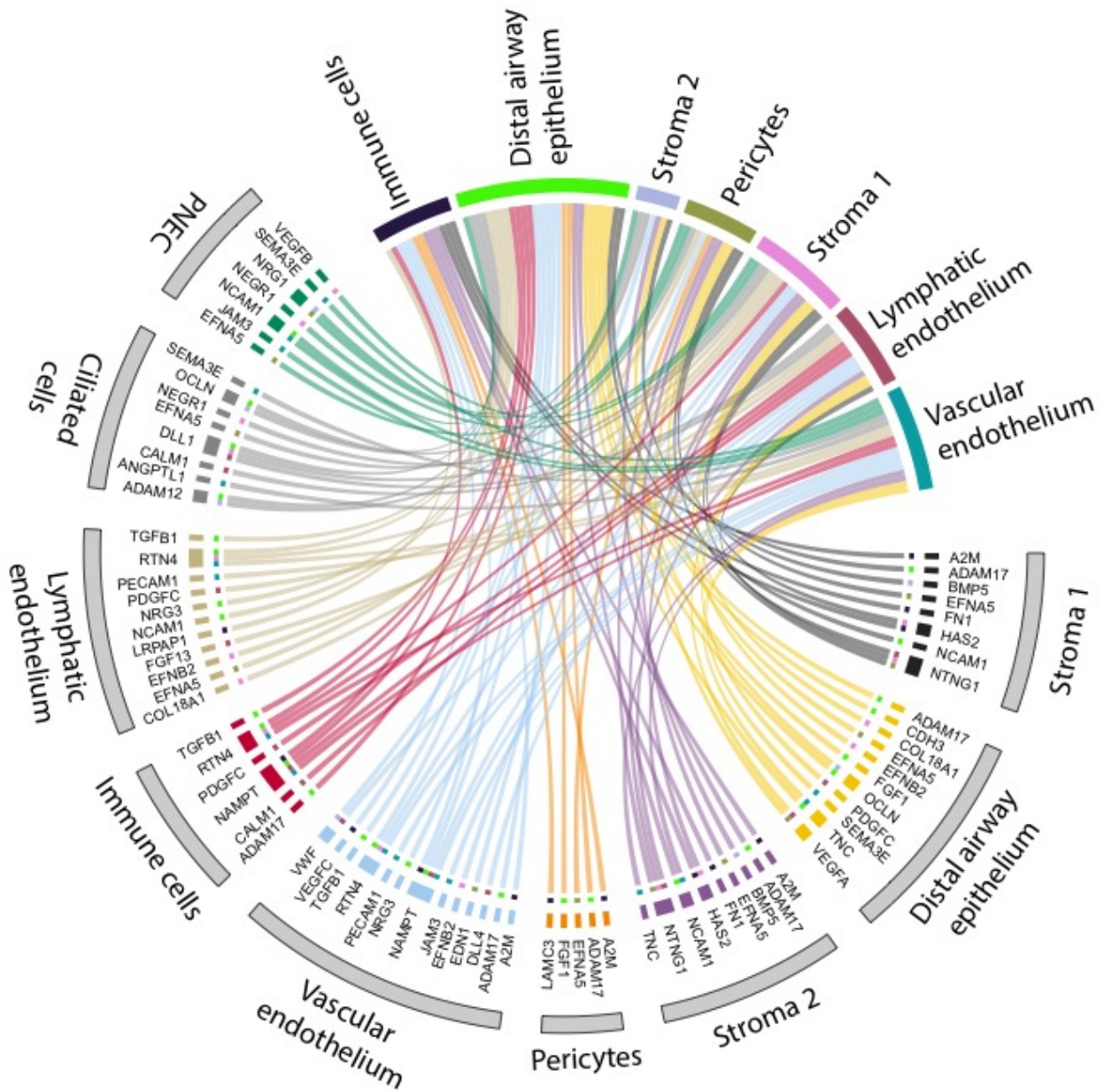




**Figure 6. Developmental changes in gene expression and spatial localization of vascular endothelium and pericytes**

**Figure 6. Developmental changes in gene expression and spatial localization of vascular endothelium and pericytes**

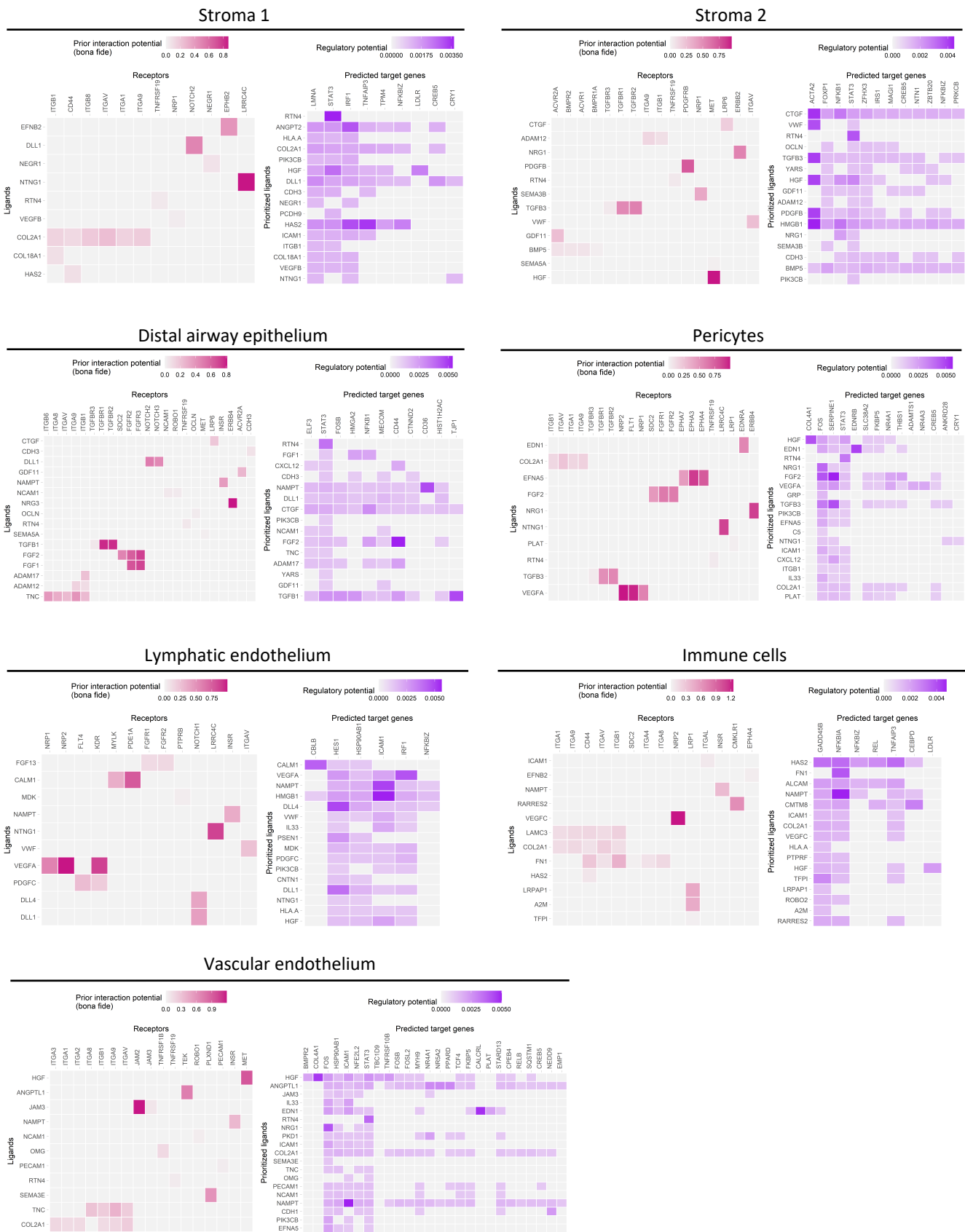
a) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in the main cluster vascular endothelium (left), and heatmap depicting the expression level of genes associated with selected enriched terms (right). b) Representative fluorescent RNA *in situ* hybridization pictures showing co-expression of *PECAM1* (green), *KITLG* (magenta) and *CAV1* (white) in the vascular endothelium in human fetal lungs at 14+3 (top) and 18+2 (bottom) weeks of gestation. Magnification at x20 (scale bar 100  $\mu$ m) and x63 (scale bar 40  $\mu$ m). c) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in the main cluster pericytes (left), and heatmap depicting the expression level of genes associated with selected enriched terms (right). Expression levels in the heatmaps are presented as  $\log(\text{TP10k}+1)$  values.  $\log(\text{TP10k}+1)$  corresponds to log-transformed UMIs per 10k. d) Representative fluorescent RNA *in situ* hybridization pictures showing co-localization of *PECAM1+* endothelial cells (white) and *PDGFRB+* pericytes (red) in human fetal lungs at 18+2 weeks of gestation. Magnification at x20 (scale bar 100  $\mu$ m) and x63 (scale bar 40  $\mu$ m).



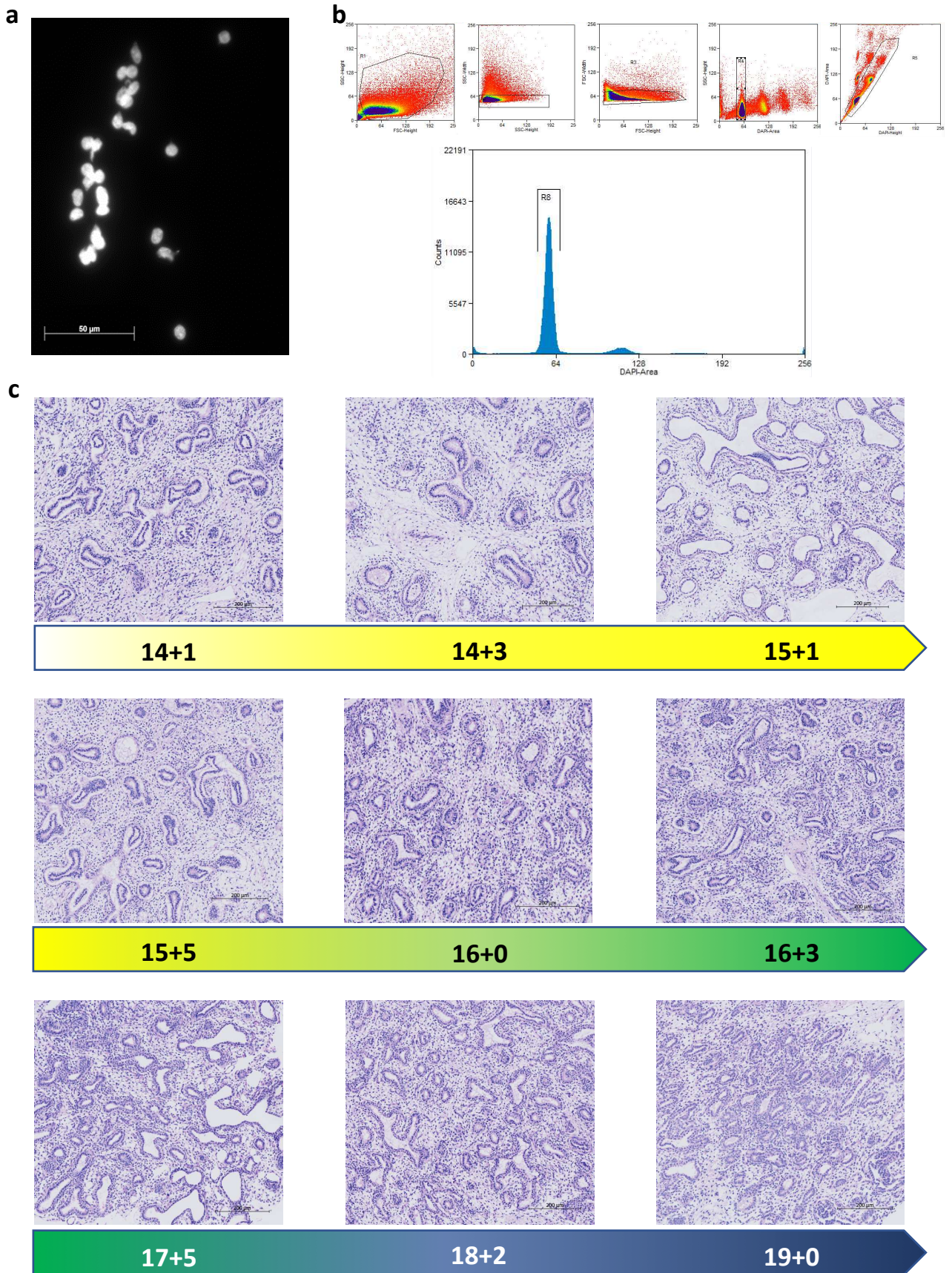
**Figure 7: Cell communication pathways induced over time**

Circos plot showing inferred cell communications identified with NicheNet by comparing the two latest GA (18+2 and 19+0 weeks) with the two earliest GA (14+1 and 14+3 weeks). Colored populations in the upper part of the plot represent the receiving cell populations, while grey-colored populations in the lower part of the plot represent the signal senders and their ligands. Each ligand is connected to the respective receiving population.





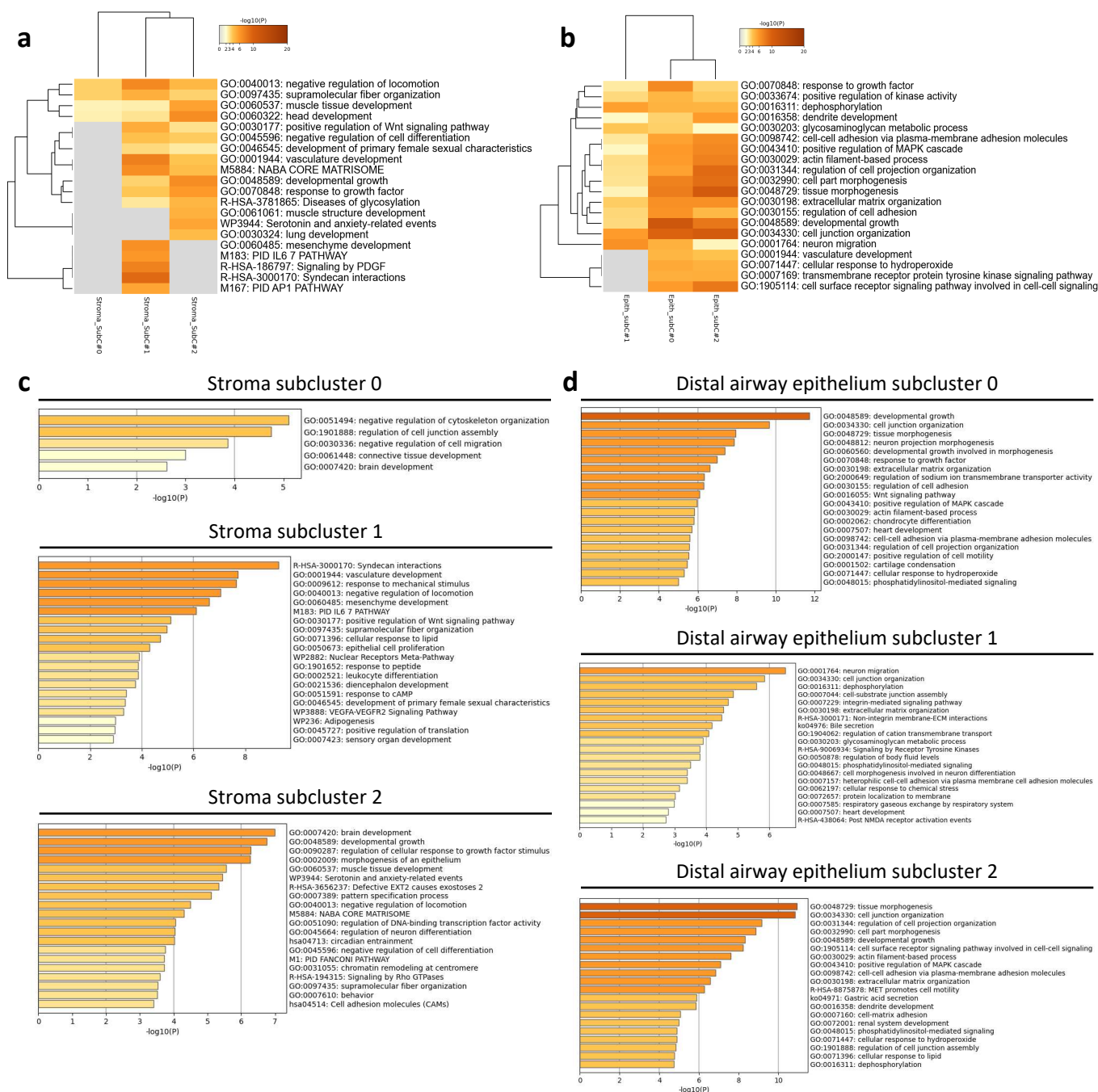
**Figure 8. Identified receptors and predicted target genes involved in cell communication**  
 Heatmaps showing bona-fide receptors and predicted target genes for each receiving cell population identified with NicheNet by comparing the two latest GA (18+2 and 19+0 weeks) with the two earliest GA (14+1 and 14+3 weeks). Receiving cell populations include stroma 1 and 2, pericytes, Lymphatic endothelium, distal airway epithelium, immune cells and vascular endothelium. For each cell population, pink heatmap reports bona fide receptors and their ligands and purple heatmap shows predicted target genes according to prioritized ligands.



**Supplemental figure 1. Quality control of the nuclei used for snRNA-seq and morphology of the human fetal lung samples**

a) Representative image showing isolated single nuclei stained with DAPI and visualized by fluorescent microscopy. Magnification x40. b) Gating strategy to obtain single diploid nuclei and histogram showing the purity and quality of the single nuclei suspension as evaluated by flow cytometry. c) Representative images showing the lung architecture in all fetal samples stained with hematoxylin and eosin. Magnification x20, scale = 200  $\mu\text{m}$ .

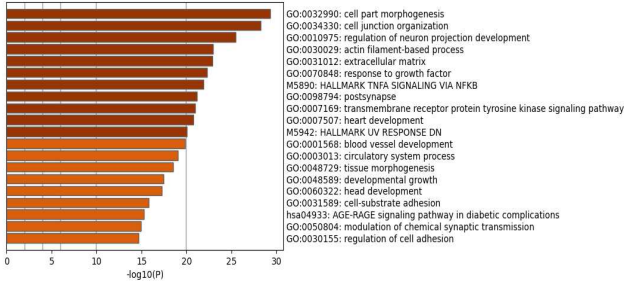




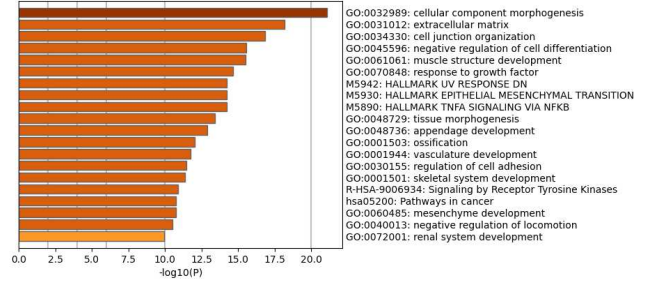
## Supplemental figure 2. Gene set enrichment analysis for stromal and distal airway epithelial subclusters

a) Heatmaps depicting the enriched terms associated with different stromal subclusters as identified by multi-list gene set enrichment analysis. b) Heatmaps depicting the enriched terms associated with different distal airway epithelium subclusters as identified by multi-list enrichment analysis. c) Enriched terms associated with individual stromal subclusters as identified by enrichment analysis. d) Enriched terms associated with individual distal airway epithelium subclusters as identified by enrichment analysis.

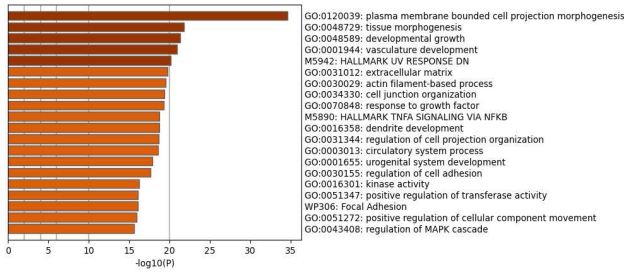
## Stroma 1



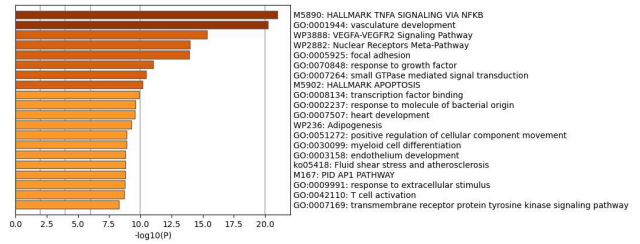
## Stroma 2



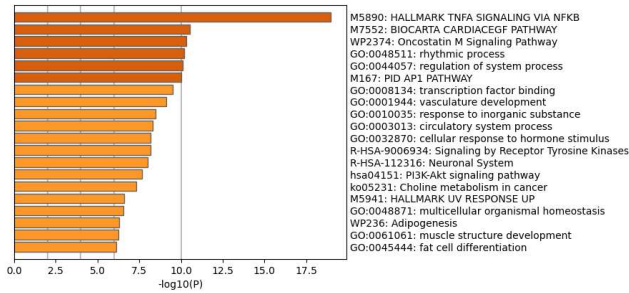
## Distal airway epithelium



## Vascular endothelium

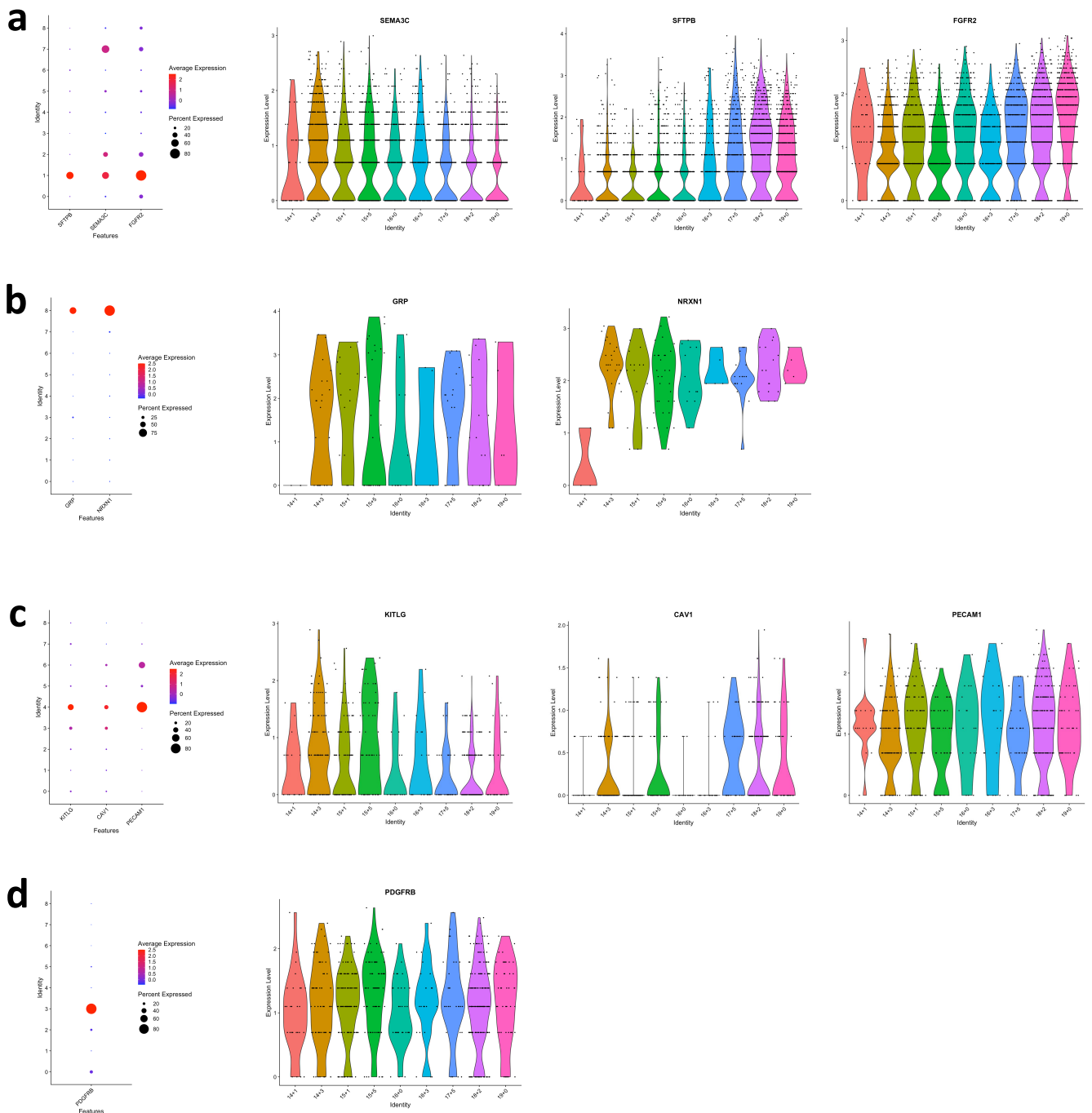


## Pericytes



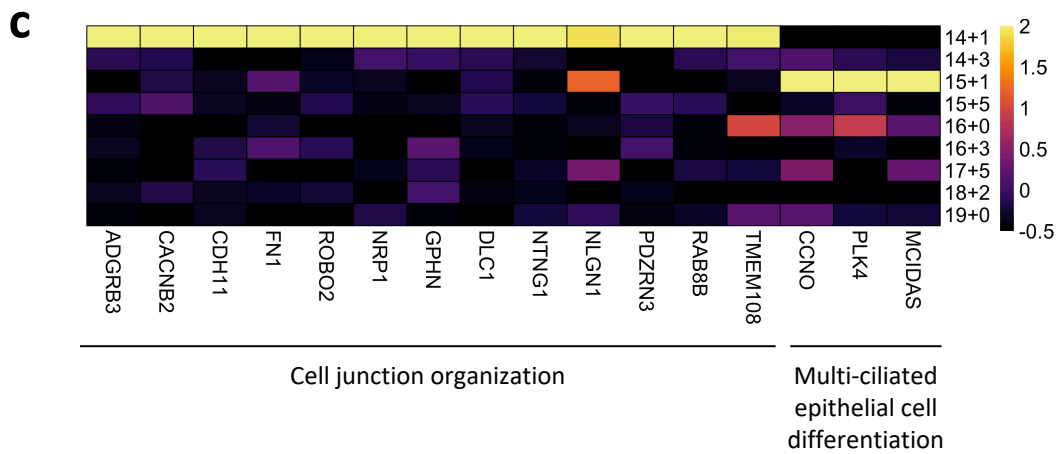
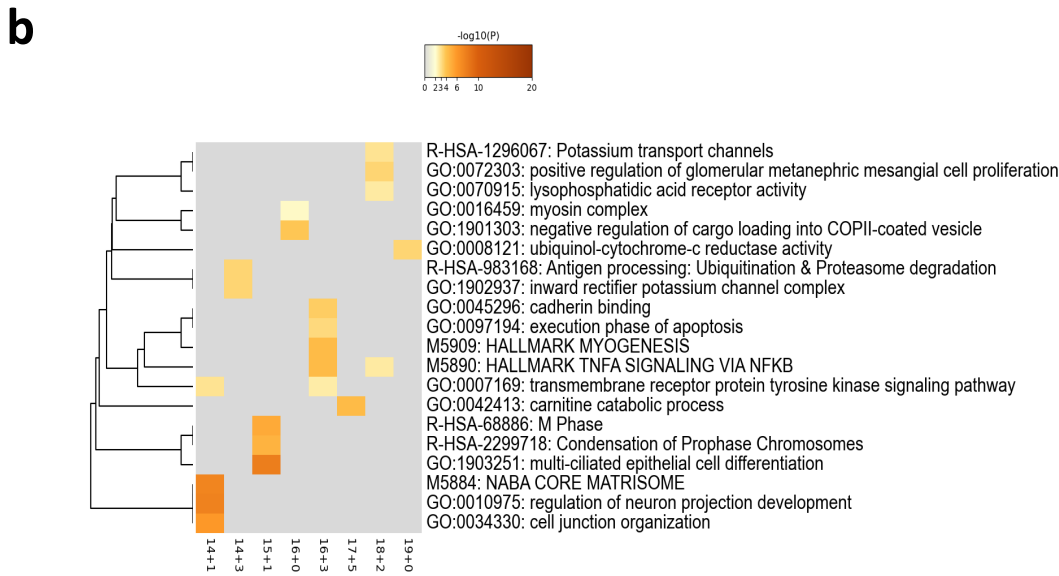
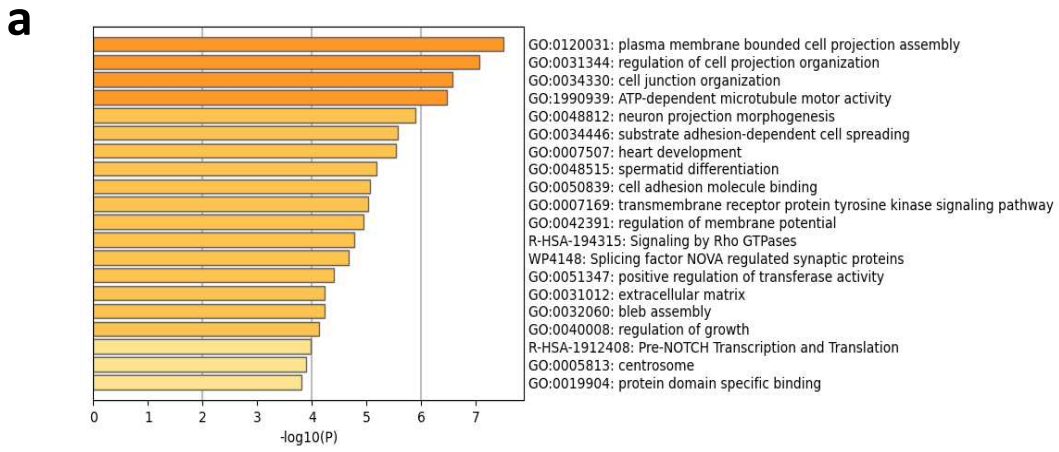
### Supplemental figure 3. Pathways induced by increasing age between 14 and 19 weeks of gestation in the fetal lung stromal, distal airway epithelium, vascular endothelium and pericytes populations

Pathways induced during fetal lung development (14+1 until 19+0 weeks) as identified by gene set enrichment analysis based on the DSA analysis in stroma 1, stroma 2, distal airway epithelium, vascular endothelium and lung pericytes.

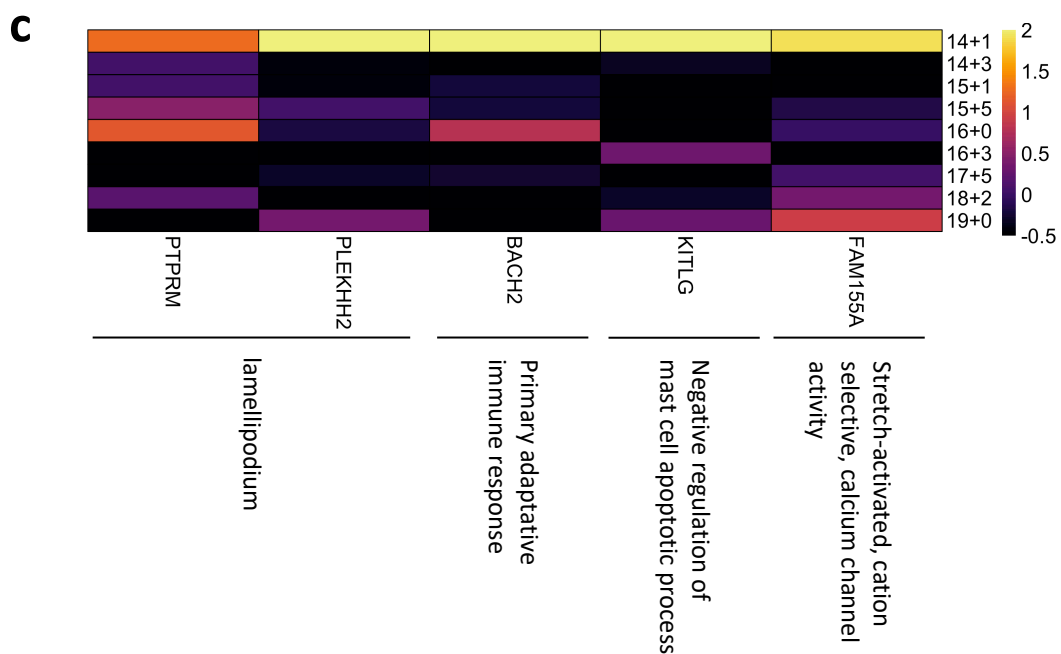
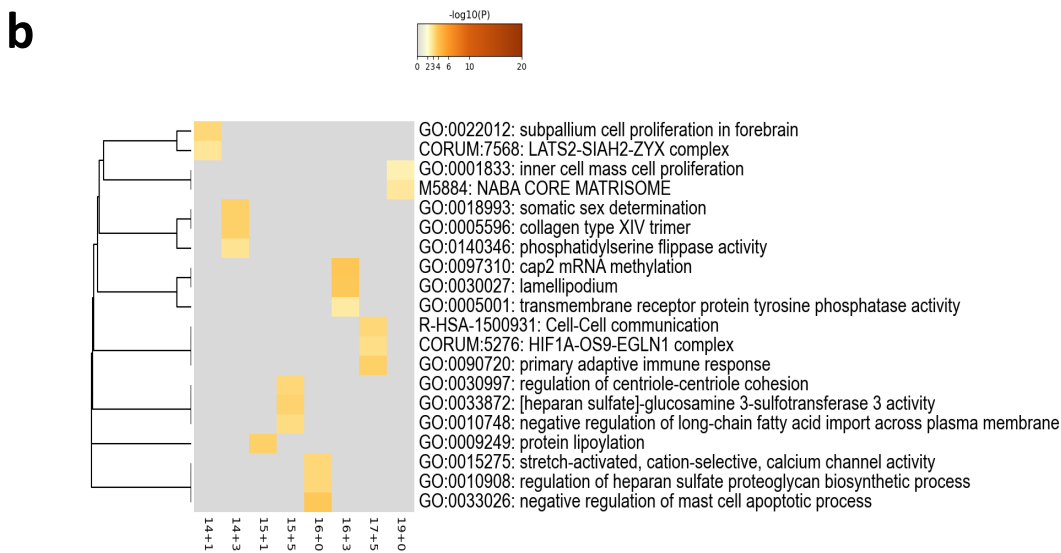
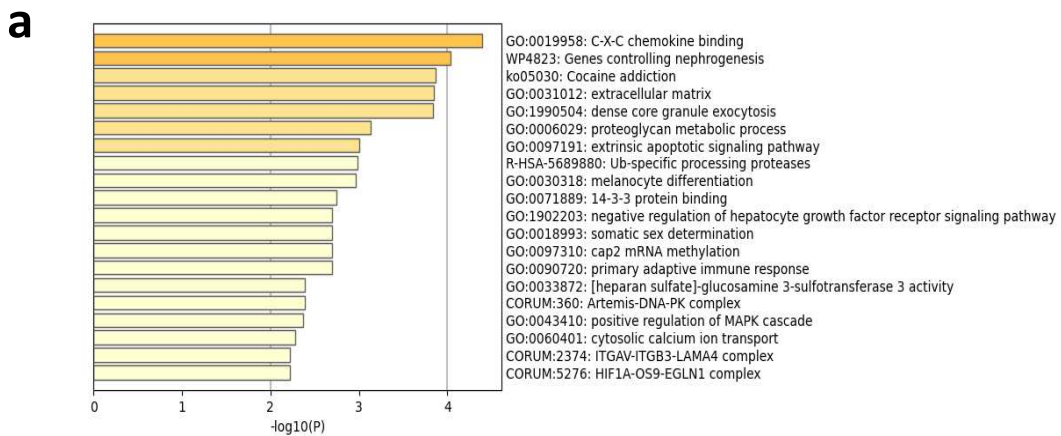


### Supplemental figure 4. Gene selection for spatial validation by Fluorescent *In Situ* Hybridization (FISH)

a) Dotplot depicting distal airway epithelium cluster specificity for *FGFR2*, *SEMA3C* and *SFTPB* (left). Violin plots (right) depicting gene expression level across different GAs (14+1 to 19+0 weeks of gestation, x-axis) and identifying *SEMA3C* as an early marker, *SFTPB* as a late marker and *FGFR2* as a general marker of the distal airway epithelium cluster. b) Dotplot depicting pulmonary neuroendocrine cell cluster specificity for *GRP* and *NRXN1* (left). Violin plot (right) depicting *GRP* and *NRXN1* expression level across different GAs (14+1 to 19+0 weeks of gestation, x-axis) and identifying *GRP* and *NRXN1* as a general marker of the pulmonary neuroendocrine cell cluster. Expression values in violin plots represent Z-score-transformed  $\log(\text{TP10k}+1)$  values. c) Dotplot depicting vascular endothelium cluster specificity for *KITLG*, *CAV1* and *PECAM1* (left). Violin plots (right) depicting gene expression level across different GAs (14+1 to 19+0 weeks of gestation, x-axis) and identifying *KITLG* as an early marker, *CAV1* as a late marker and *PECAM1* as a general marker of the vascular endothelium cluster. d) Dotplot depicting pericyte cluster specificity for *PDGFRB* (left). Violin plot (right) depicting *PDGFRB* expression level across different GAs (14+1 to 19+0 weeks of gestation, x-axis) and identifying *PDGFRB* as a general marker of the pericyte cluster.



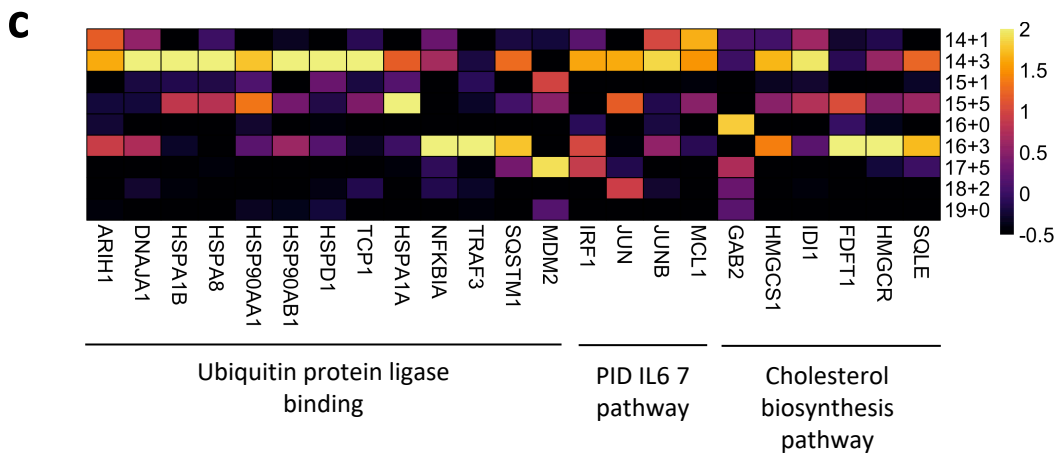
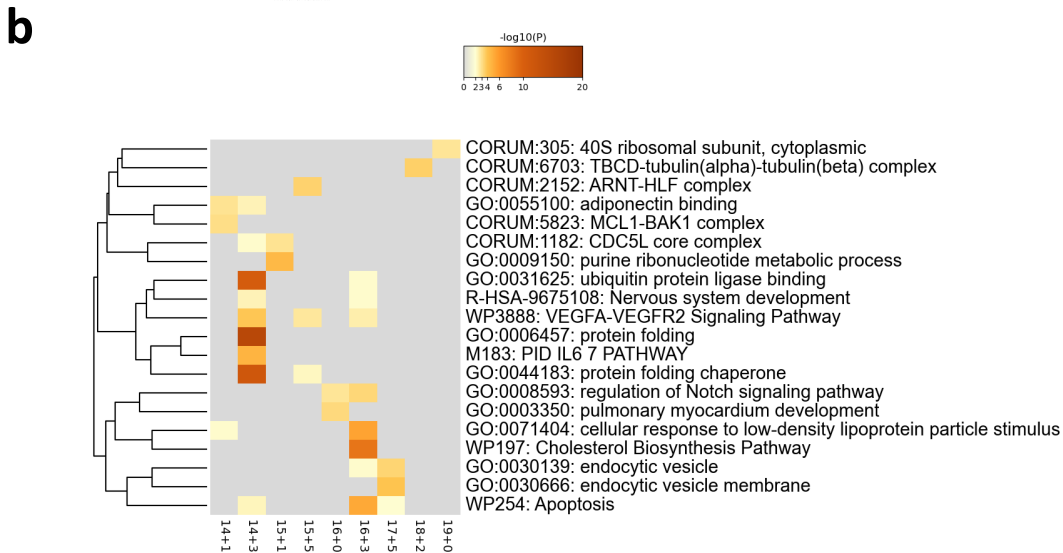
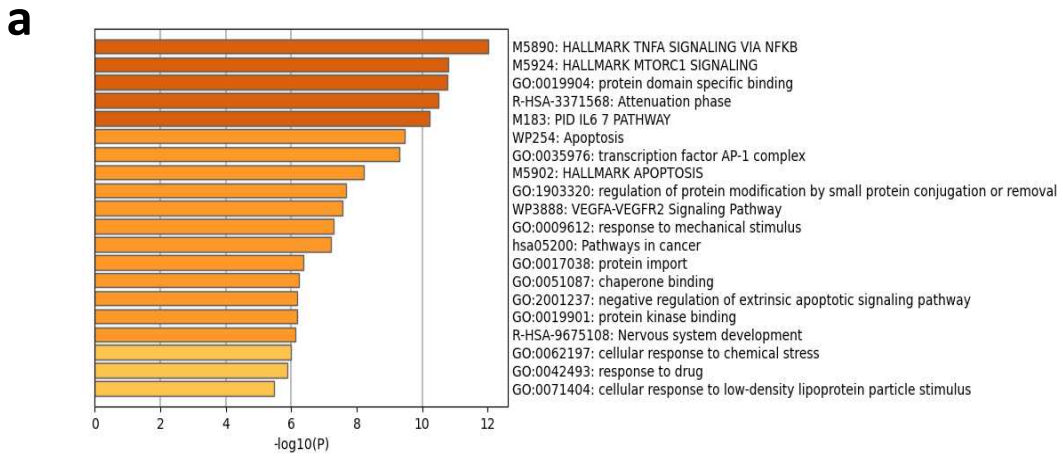
**Supplemental figure 5. Developmental changes in gene expression in the fetal lung ciliated cells**  
 a) Pathways induced during fetal lung development (14+1 until 19+0 weeks) in the ciliated cells as identified by gene set enrichment analysis based on the DSA. b) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in immune cells. c) Heatmap depicting the expression level of genes associated with selected enriched terms as identified in panel b). Expression levels in the heatmaps and dotplots are presented as  $\log(TP10k+1)$  values.  $\log(TP10k+1)$  corresponds to log-transformed UMIs per 10k.



**Supplemental figure 6. Developmental changes in gene expression in the fetal lung pulmonary neuroendocrine cells**

a) Pathways induced during fetal lung development (14+1 until 19+0 weeks) in the pulmonary neuroendocrine cells (PNEC) as identified by gene set enrichment analysis based on the DSA. b) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in immune cells. c) Heatmap depicting the expression level of genes associated with selected enriched terms as identified in panel b). Expression levels in the heatmaps and dotplots are presented as  $\log(TP10k+1)$  values.  $\log(TP10k+1)$  corresponds to log-transformed UMIs per 10k.

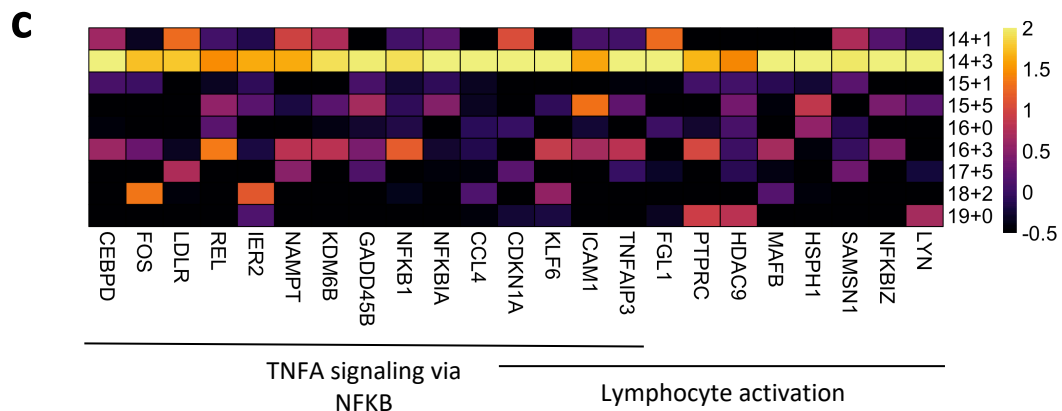
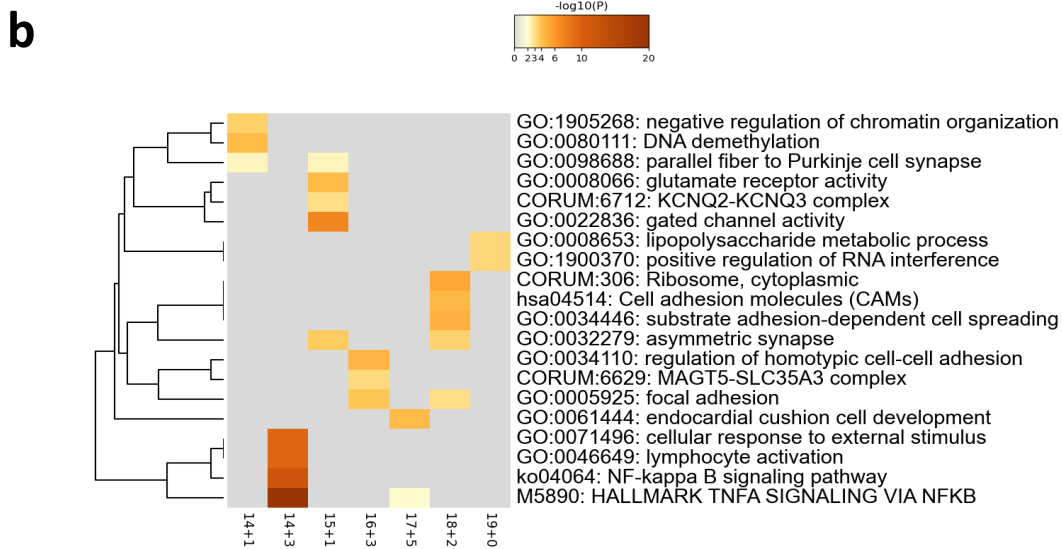
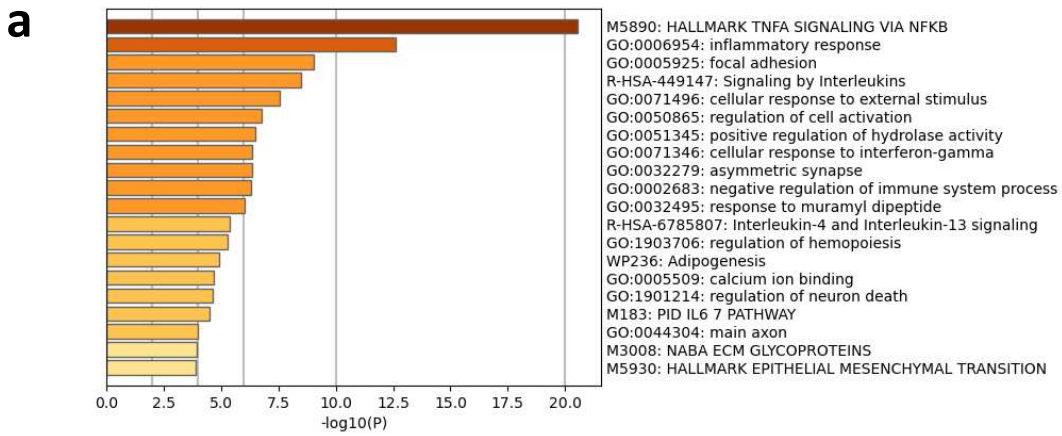




**Supplemental figure 7. Developmental changes in gene expression in the fetal lung lymphatic endothelial cells**

a) pathways induced during fetal lung development (14+1 until 19+0 weeks) in the lymphatic endothelial cells as identified by gene set enrichment analysis based on the DSA. b) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in immune cells. c) Heatmap depicting the expression level of genes associated with selected enriched terms as identified in panel b). Expression levels in the heatmaps and dotplots are presented as  $\log(TP10k+1)$  values.  $\log(TP10k+1)$  corresponds to log-transformed UMIs per 10k.





**Supplemental figure 8. Developmental changes in gene expression in the fetal lung immune cells**

a) pathways induced during fetal lung development (14+1 until 19+0 weeks) in the immune cells as identified by gene set enrichment analysis based on the DSA. b) Heatmaps depicting the enriched terms associated with individual GAs as identified by multi-list enrichment analysis based on DSA in immune cells. c) Heatmap depicting the expression level of genes associated with selected enriched terms as identified in panel b). Expression levels in the heatmaps and dotplots are presented as  $\log(TP10k+1)$  values.  $\log(TP10k+1)$  corresponds to log-transformed UMIs per 10k.

## Supplemental Methods

### 1. Human fetal lung collection.

De-identified human fetal lung samples were obtained from women undergoing elective termination of pregnancy with accurate dating (either last menstrual period or an ultrasound) and live, singleton pregnancy. Pregnancies with fetal morphological or chromosomal anomalies were excluded. Tissue collection was approved by The Ottawa Hospital Review Ethical Board (approval number 20170603-01H).

Lung samples were collected for precision-cut lung slices [1]. Briefly fetal lungs were sampled by the pathologist. Lungs were subsequently inflated with a low gelling temperature agarose solution (1.5% solution, #A9414 Sigma-Aldrich, Oakville, ON, Canada), and sampled with an 8mm biopsy punch (Robbins True-Cut Disposable Biopsy Punch 8mm, #RBP-80, Robbins instruments, Chatham, NJ, USA) prior to slicing with a tissue slicer. Spare lung tissue was snap frozen in liquid nitrogen and stored at -80°C for single nuclei isolation.

### 2. Single nuclei isolation.

Single nuclei isolation was performed according to Martelotto *et al.* [2] with minor adjustments.

#### 2.1 Tissue homogenisation.

From each fetal lung, 4 small pieces (each the size of a grain of rice) were randomly sampled. These 4 pieces were divided in two 1.5 ml tubes. The two pieces per 1.5 ml tube were mechanically dissociated on ice using a single-use plastic pestle (Bel-Art® Disposable Polypropylene Pestles and 1.5ml tubes, #66001-104, VWR, Mont-Royal, QC, Canada). Homogenates were resuspended in 500 µl of chilled Nuclei EZ Lysis buffer (Nuclei Isolation Kit: Nuclei EZ Prep, #NUC101-1KT, Sigma-Aldrich, Oakville, ON, Canada). Samples corresponding to same fetal lung sample were pooled together and 500 µl of the homogenate was transferred to a new 2 ml tube.

## 2.2 Nuclei isolation, multiplexing and staining.

One milliliter of Nuclei EZ Lysis buffer (Nuclei Isolation Kit: Nuclei EZ Prep, #NUC101-1KT, Sigma-Aldrich, Oakville, ON, Canada) was added to each sample, mixed gently, and incubated on ice for 5 minutes. Homogenates were filtered through a 50 $\mu$ l nylon mesh (ThermoFisher Scientific, Burlington, ON, Canada). Flow-through fractions were centrifuged at 500x g for 5 minutes.

Multiplexing was performed according to the MULTI-seq protocol [3]. The resulting pellets were re-suspended in 200  $\mu$ l of 200 nM anchor/200 nM barcode solution, each sample receiving a different sample barcode. Samples were incubated on ice for 10 minutes, after which 16  $\mu$ l of common lipid-modified co-anchor mix (2  $\mu$ M) was added. Samples were incubated on ice for 5 minutes and diluted in 500  $\mu$ l of Nuclei wash buffer (1x DPBS, Lonza, Basel, Switzerland) with 1% (v/w) Bovine Serum Albumin (Sigma-Aldrich, Oakville, ON, Canada) and 0.2 IU/ $\mu$ l RNase inhibitor (NProtector RNase inhibitor, #3335402001, Sigma-Aldrich, Oakville, ON, Canada). Single nuclei were counted with an automated cell counter (Eve Automatic cell Counter, Nano Entek, Waltham, MA, USA) and pelleted at 500 $\times$  g for 5 minutes. Pellets were re-suspended with Nuclei Wash Buffer in order to achieve equal concentrations. Nuclei samples were pooled at equal ratio into a 1.5ml tube and pelleted at 500x g for 5 min. The pooled nuclei were washed with 500  $\mu$ l of Nuclei wash buffer and pelleted at 500 x g for 5 minutes.

The pooled nuclei were re-suspended in 500  $\mu$ l Nuclei Wash Buffer with DAPI (10  $\mu$ g/ml). Nuclei integrity was confirmed using a fluorescent microscope (Axio Imager M2, Carl Zeiss, Toronto, ON, Canada).

## 2.3 Nuclei sorting.

Nuclei were sorted using a flow cytometer (BD LSR Fortessa, Beckton Dickinson Biosciences, Franklin Lakes, NJ, USA) at the Ottawa Hospital Research Institute (OHRI) core facility. Sample compensation was performed using Summit v.5.4 software and data analysis were performed with MoFlo XDP software (XDP, Beckman Coulter, Fullerton, CA, USA). The

flow cytometer was gated on the single diploid DAPI peak, and 50,000 - 60,000 nuclei were sorted in a small volume of Nuclei Wash Buffer (~50  $\mu$ l). Sorted nuclei were immediately processed by 10x Chromium at the OHRI StemCore facility.

### **3. Single nuclei RNA sequencing and pre-processing.**

#### 3.1 Library preparation and sequencing.

The 10x Chromium gene expression libraries were prepared according to manufacturer's protocol (10X Genomics, Pleasanton, CA, USA), and the sequencing was performed at the OHRI StemCore facility using NextSeq500 (Illumina, San Diego, CA, USA).

#### 3.2 Pre-processing.

Raw sequencing reads were processed using CellRanger v3.1.0, aligning reads to the mm10 build of the human genome. MULTI-seq barcode libraries were trimmed prior to demultiplexing using Trimmomatic (v0.36). Demultiplexing was performed using the deMULTIplex R package (v1.0.2) [3, 4]. Only cells positive for a single barcode were kept for further analysis. Sample annotations were added to all cells in the data set. Pre-processing steps and downstream analysis were performed with Seurat (Seurat v4.0.0) [5].

#### 3.3 Quality control, normalization, and integration

Expression matrices were loaded as Seurat objects into R [6]. Only cells with > 200 genes detected and < 20% of UMIs mapped to mitochondrial genes were used. Each unique sample was split based on MULTI-seq sample barcodes into a separate Seurat object. SCTransform() was used to normalize samples, select highly variable genes, and to regress out cell cycle and cell stress effects. To eliminate batch effects or biological variability effects on clustering, the data integration method implemented by Seurat for SCTransform-normalized data was performed, using the SelectIntegrationFeatures(), PrepSCTIntegration(), FindIntegrationAnchors(), and IntegrateData() functions.

## 4. Downstream analysis.

### 4.1 Clustering and clusters annotation.

Dimensionality reduction was performed using principal component analysis (PCA) with RunPCA() function on the top 3000 variable genes and the data was clustered at a low resolution (dims: 20, resolution: 0.05, re-embedding dims: 20) with the Louvain algorithm implemented in the FindClusters() function in Seurat. Cell populations were identified with a simple Wilcoxon rank sum test with the FindAllMarkers() function in Seurat. For cluster annotation, we used a differential expression testing approach using FindAllMarkers() function to identify clusters marker genes (e.g., genes up-regulated in the cluster of interest). These marker genes were compared to known marker genes sets from LunGENS (Lung Gene Expression iN Single-cell) datasets on the LGEA (Lung Gene Expression Analysis) web portal [7, 8], and literature to define clusters identities. We next isolated and reprocessed the stromal and epithelial subsets using the same normalization and integration approach. In this case data were re-clustered at a higher resolution (Stroma dims: 20, resolution: 0.075, re-embedding dims:20; and for Distal airway epithelium dims: 20, resolution: 0.10, re-embedding dims:20). Upon subclustering the distal airway epithelium, we identified a small fourth subclustering showing high expression of stromal canonical markers. This population of cells was manually removed from the dataset. Gene set enrichment analysis (GSEA) was performed with Metascape [9].

### 4.2 Differential State Analysis.

To identify differentially expressed genes in response to gestational age (GA), we performed a Differential State Analysis (DSA) using muscat R package [10]. Genes with an adjusted p-value <0.01 were considered significant and used to generate a heat map. From this heat map, genes were classified according to their up-regulated expression among different GA (individual samples).

To identify gene sets associated with differentially expressed genes we performed a GSEA using Metascape [9]. For each cluster, enrichment analysis was done using a single list (DSA results gene adj-p<0.01). In addition, we performed a meta-analysis using multiple gene list where each column represented a different GA and contained the up-regulated genes for this specific GA.

Differential gene expression across GA within a subpopulation was also used to identify potential early, late, and general (expressed across all the GA) marker genes for this subpopulation and further assess for spatial localization by FISH.

#### 4.3 Cell communication

To describe and understand cell communication networks within the developing human fetal lung samples, we used NicheNet (v1.0.0) R package [11]. To prioritize results, analysis was limited to larger cell populations with higher number of differentially expressed genes across GA. Clusters 0, 1, 2, 3, 4, 5, and 6 were considered as signal receivers, while all the clusters were considered as potential signal senders. All the cell types were considered as sending cells, whereas ciliated cells and PNEC were not considered as receiving cell types as they represented small clusters with lack of changes in gene expression across the GA. In addition to crosstalk between diverse populations, crosstalk within the population itself was also observed in all the cell types, except PNEC and ciliated cells. To define a gene set of interest, the two latest GA timepoints (18+2w and 19+0w) were used as a condition of interest and two earliest GA timepoints (14+1w and 14+3w) as the reference condition. Background expression of genes was specified using all genes with adjusted p-value < 0.05, log2FC value > 0.05, and >10% detection in a given cluster. For each “receiver” cell population, top 20 ligands predicted to drive developmental age were selected based on the Pearson correlation coefficient. Quantile cut-off on the ligand-target scores of the input weighted ligand-target network was set to 0.33. To further validate the clinical relevance of our cell communication results, ligands expression in sending cell types and receptors expression in receiving cell

types were compared to the Human Protein Atlas (<https://www.proteinatlas.org>) to confirm the expression of the protein (ligand or receptor) within the cell type of interest in human lung. Cell communication results are presented as a circos plot in Figure 6, and bona-fide receptors and predicted target genes for the receiving cell types are presented in Figure 7. When the predicted target genes list contained 3 or more genes, a GSEA was performed, and the results are presented in supplemental table 9.

## **5. Data and code availability.**

All RNA sequencing data including raw fastq sequencing files, gene expression matrices, and cell metadata reported in this article is NCBI Gene Expression Omnibus (GEO) database and can be made available by request to the corresponding author. Code used for the analysis of the snRNA-seq data will be made available at the public GitHub repository and is available by request to the corresponding author.

## **6. Fluorescent in situ hybridization.**

RNA in situ hybridization was performed on fresh 4  $\mu$ m formalin-fixed paraffin embedded (FFPE) tissue sections using RNAscope Multiplex Fluorescent Reagent Kit v2 (#323100, Advanced Cell Diagnostics, CA, USA) according to the manufacturers's instructions. Briefly, tissue sections were baked for 1 h at 60°C, deparaffinized and treated with hydrogen peroxide for 10 min at room temperature (RT). Target retrieval was performed for 15 min at 98°C, followed by protease plus treatment for 15 min at 40°C. The sections were then hybridized with probes for 2 h at 40°C followed by signal amplification and developing of HRP channels. The RNAscope probes used in this study were: Hs CAV1 (#452071), Hs FGFR2 (#311171-C2), Hs GRP (#465261) with 1:5 dilution, Hs KITLG (#407671-C3), Hs NRXN1 (527151-C3), Hs PDGFRB (#548991), Hs PECAM1-O1 (#455931-C2), Hs SFTBP-O1 (#1087181-C3), Hs SEMA3C (#549241), 3-plex positive control probe Hs (#320861), and 3-plex negative control probe (#320871). The signals were detected with TSA Plus fluorophores

fluorescein (1:750 dilution), Cyanine 3 (1:1500 dilution), and Cyanine 5 (1:3000 dilution) (NEL741001KT, NEL744001KT, and NEL745001KT, respectively, Akoya Biosciences, MA, USA). The sections were counterstained with DAPI and mounted with ProLong Gold Antifade Mountant (P36930, Life Technologies Limited, Thermo Fisher Scientific, UK). Tissue sections were scanned using 3DHISTECH Pannoramic 250 FLASH II digital slide scanner at 40x magnification with extended focus and 7 focus levels at Genome Biology Unit supported by HiLIFE and the Faculty of Medicine, University of Helsinki, and Biocenter Finland.



## References.

1. Kang MH, van Lieshout LP, Xu L *et al.* A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency. *Nat Commun* 2020; 11: 3929.
2. Martelotto L. 'Frankenstein' protocol for nuclei isolation from fresh and frozen tissue for snRNAseq V.3.  
Available from: <https://www.protocols.io/view/frankenstein-protocol-for-nuclei-isolation-from-f-bqxymxpw>. Date last accessed: January 11, 2022.
3. McGinnis CS, Patterson DM, Winkler J *et al.* MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 2019; 16: 619–626.
4. Hurskainen M, Mižíková I, Cook DP *et al.* Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat Commun* 2021; 12: 1565.
5. Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021; 184: 3573-3587.
6. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: <https://www.r-project.org/>. Date last accessed: January 11, 2022.
7. Du Y, Guo M, Whitsett JA *et al.* "LungGENS": a web-based tool for mapping single-cell gene expression in the developing lung. *Thorax* 2015; 70: 1092–1094.
8. Du Y, Kitzmiller JA, Sridharan A *et al.* Lung Gene Expression Analysis (LGEA): an integrative web portal for comprehensive gene expression data analysis in lung development. *Thorax* 2017; 72: 481–484.
9. Zhou Y, Zhou B, Pache L *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; 10: 1523.
10. Crowell HL, Sonesson C, Germain P-L *et al.* muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat Commun* 2020; 11: 6077.
11. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* 2020; 17: 159–162.

## II. A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency

This study is published in *Nature Communication* (Kang et al. 2020). The figures presented after this section are the original figures from the manuscript, except for Figure 13 which the modified Supplemental Figure 1 in which panels b (PCLS image) and c (resazurin assay) have been added.

### 1. Background

Surfactant protein B (SP-B) deficiency is a rare, but lethal neonatal condition. SP-B plays a central role in surfactant homeostasis and prevents alveoli collapse by reducing the alveoli surface tension. SP-B deficiency is an autosomal recessive condition in which patient's alveolar type 2 (AT2) cells do not produce SP-B. As a result, these patients present with a severe respiratory distress syndrome within hours after birth. Ultimately, this progressive respiratory distress leads to lethal respiratory failure before 6 months of life. Currently, lung transplantation is the only curative treatment option for these patients.

Here, we showed in a model of SP-B deficient mice that an engineered adeno-associated virus (AAV) called AAV6.2FF can be used as vector for gene therapy to deliver proSFTPB cDNA to AT2 cells. Treated mice exhibited a dramatic increase in survival and a restored surfactant homeostasis with improved lung physiology. As some evidence suggested that AAV vector transduction in airway epithelium may differ between mice and human (X. Liu et al. 2006; Meyer-Berg et al. 2020), we wanted to further explore this AAV vector's potential for clinical translation. We generated a model of PCLS using human fetal lungs to assess AAV6.2FF transduction potential in human lung parenchyma.

## 2. Results

To generate and culture the PCLS, we used a protocol from *Rosales Gerpe et al.* (Rosales Gerpe et al. 2018). Our study protocol was based on a 7 days culture, with AAV6.2FF added to the culture medium for 24 hours on day 3 (Figure 13a and 14a). AAV delivered the transgenes with GFP (AAV-GFP), mCherry (AAV-mCherry) or luciferase (AAV-Luc). Mouse (mSPB) or human (hSPB) SPB genes were also used as cargo transgenes. PCLS viability was assessed by resazurin assays performed at days 0, 3 (pre-AAV transduction) and 7 of culture. AAV transduction was evaluated at day 7 by bioluminescence (In Vivo Imaging System, IVIS) and immunofluorescence.

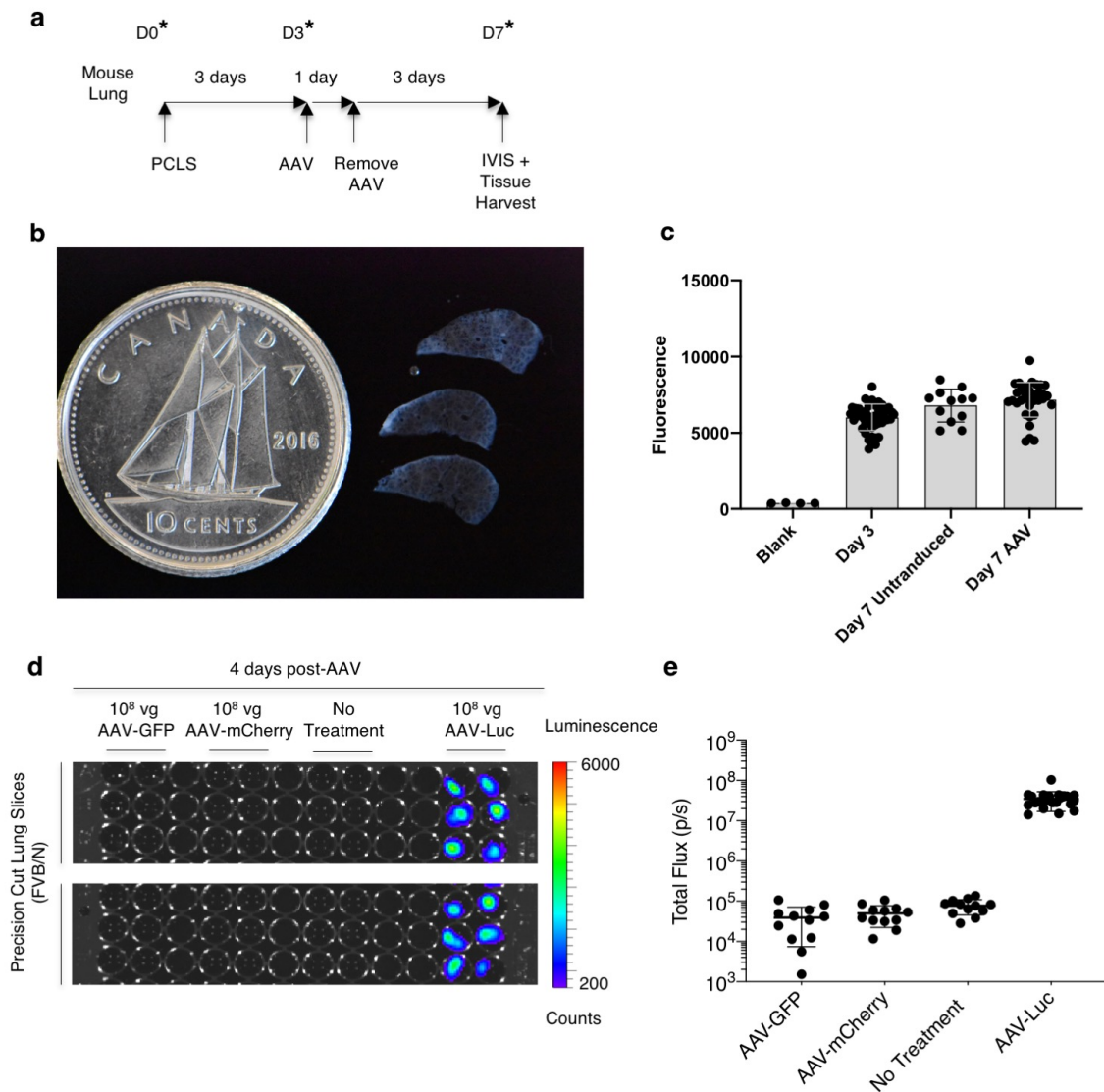
A first set of experiments was performed on adult mice PCLS as a proof-of-concept study. IVIS imaging four days after AAV transduction confirmed luciferase expression in AAV-Luc transduced PCLS. The resazurin assay showed similar level of metabolic activity before and after AAV transduction, as well as in AAV transduced PCLS compared to untransduced control PCLS (Figure 13).

Next, PCLS were generated from human fetal lungs obtained from abortion (tissue collection approved by The Ottawa Hospital Review Ethical Board, approval number 20170603-01H). Two different human fetal lungs were used to generate PCLS (gestational age 16+0 and 16+3 weeks). IVIS imaging confirmed AAV transduction in human lung parenchyma 4 days after AAV treatment with a dose-dependent effect (Figure 14c and d, Figure 15a). An immunofluorescent study confirmed the co-localization of AAV-mCherry or AAV-Luc with EPCAM<sup>+</sup> cells. PCLS from a third fetal lung (16+3 weeks) were transduced with AAV-mSPB and AAV-hSPB (Figure 15c, d and e). An immunofluorescent study showed that both vectors were associated with an increase in SPB expression compared to untransduced PCLS (Figure

14g). Resazurin assay showed similar level of metabolic activity before and after AAV transduction, as well as in AAV transduced PCLS compared to untransduced control PCLS on day 7 (Figure 14b and f). In addition, Hematoxylin and eosin staining of PCLS showed comparable lung architecture and morphology at day 7 of culture (Figure 15b). The morphological analysis and the resazurin assay results suggested that AAV transduction was not associated with any adverse effects on the human tissues.

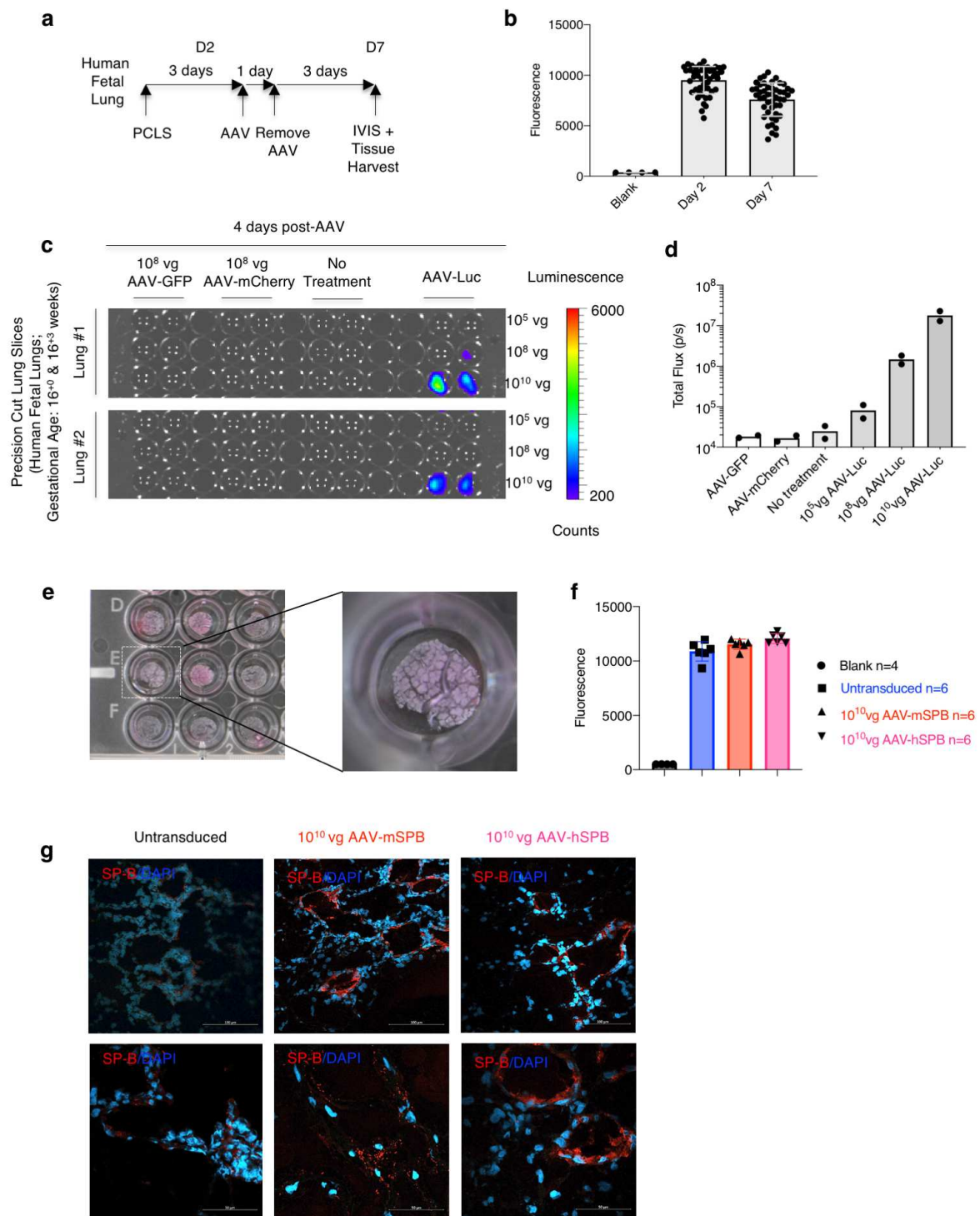
### **3. Conclusion**

Human fetal lung derived PCLS were used as a proof-of-concept for AAV6.2FF transduction in the human lung parenchyma. Our findings demonstrated that AAV6.2FF can effectively transduce the lung parenchyma with a dose response effect. AAV6.2FF was colocalized in EPCAM<sup>+</sup> cells and associated with increased SPB expression. Finally, no major adverse effects on lung parenchyma structure and viability were reported. All together these results suggest a high potential for clinical translation of this AAV vector for SP-B gene therapy in humans.



**Figure 13.** Precision Cut Lung Slices (PCLS) of adult mouse lung tissue (Genetic strain: FVB/N) transduced with AAV6.2FF.

**a** Study design to show transduction of adult mouse lung tissue slices by AAV6.2FF. Metabolic (resazurin) assays were carried out on days 0, 3, and 7 post-tissue harvest (marked with an \*) to ensure PCLS metabolic activity and cell viability. **b** Image of mouse PCLS. **c** Resazurin assay results as represented by mean fluorescence with SD demonstrating the metabolic activity of the mouse PCLS on days 3 and 7. **d** IVIS imaging of mouse PCLS 4 days after AAV-Luc transduction and 7 days after generating the tissue slices demonstrates luciferase expression in the adult mouse lung parenchyma following the addition of D-luciferin substrate. **e** Quantification of bioluminescence from the IVIS image in (d) as Total Flux (photons per second). All treatment groups contain n=2 animals, 1 lung each with 6 slices per lung, except AAV-Luc with n=4 animals, 1 lung each with 5-6 slices per lung. Adapted from Supplemental Figure 1 from (Kang et al. 2020).

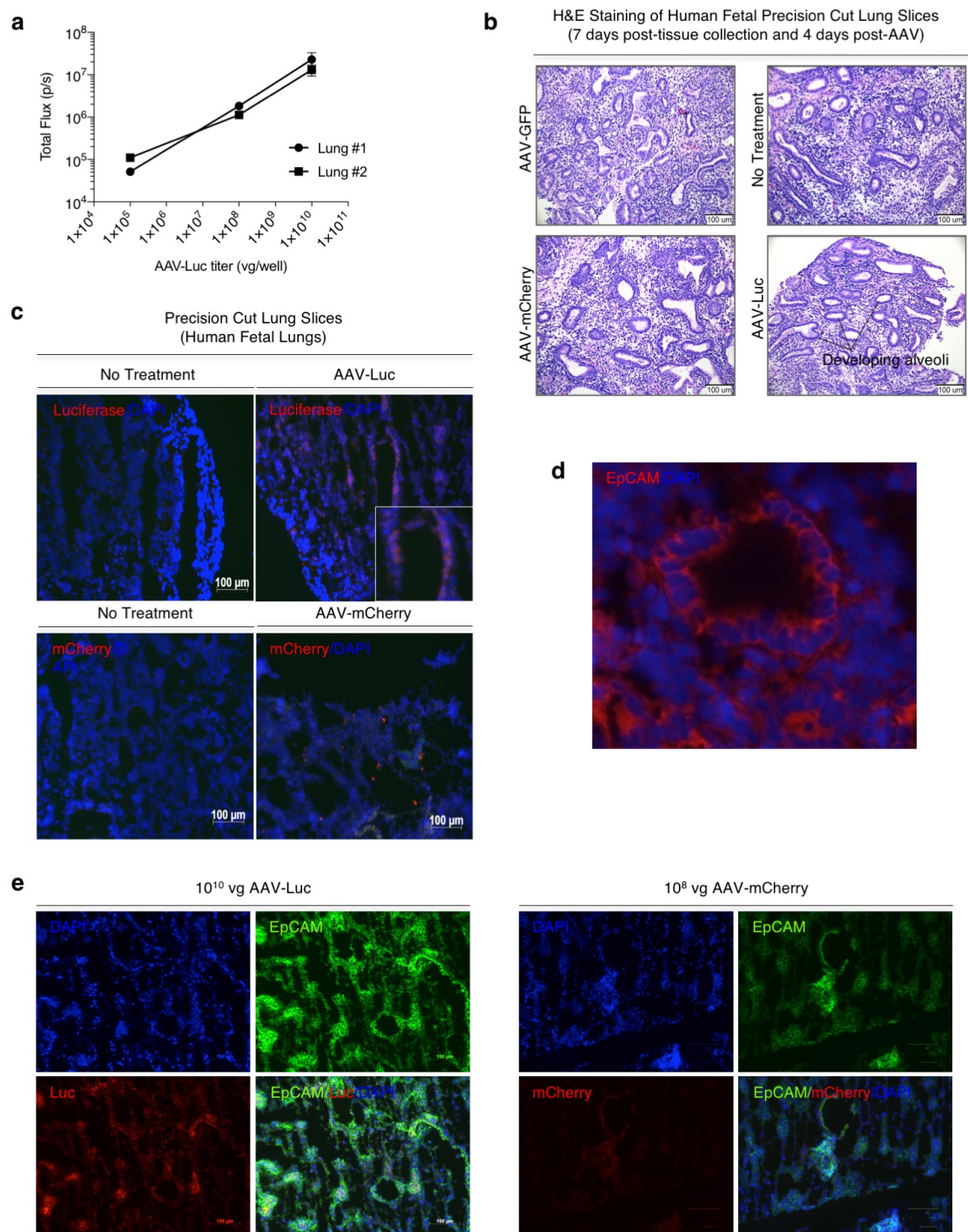


**Figure 14.** AAV6.2FF transduces the human lung parenchyma.

**a** Study design to demonstrate the ability of AAV6.2FF to transduce human lung tissue samples generated from precision cut lung slices (PCLS). Metabolic viability (resazurin) assays were conducted on days 2 (D2) and 7 (D7) post tissue harvest. **b** Resazurin assay results as represented by mean fluorescence with SD demonstrating the metabolic activity of the human PCLS on days 2 and 7. **c** Bioluminescence detection using IVIS in PCLS from

human fetuses (n = 2; 16+0 and 16+3 weeks gestational age) either untreated or transduced with 108 vg per well of AAV-GFP, 108 vg per well of AAV-mCherry, or 105, 108, or 1010 vg per well of AAV-Luc. Images are 4 days following vector exposure, and the scale is represented as counts. **d** Quantification of AAV-Luc wells are represented by the mean total flux (in photons per sec) with SD (n =2 fetuses; with 12 replicates for the AAV-GFP, AAV-mCherry, and No Treatment groups; and 4 replicates for each of the AAV-Luc doses). **e** Images of human PCLS in a 24-well tissue culture plate 7 days after culture and 3 days after AAV-SPB transduction. An enlarged image of a PCLS in an individual well is indicated by the dashed white box. **f** Resazurin assay results as represented by mean fluorescence with SD demonstrating viability of the human PCLS 3 days after no treatment (blue), or transduction with 1010 vg per well of AAV-mSPB (red) or AAV-hSPB (pink). (n = 4 blank; and n = 6 untransduced, n =6 AAV-mSPB, and n = 6 AAV-hSPB independent tissue slices from a single set of human fetal lungs). The source data for f has been provided as a Source Data file. **g** Representative confocal images of SP-B (red) and DAPI (blue) of frozen PCLS sections 3 days after transduction with 1010 vg per well of AAV-mSPB or AAV-hSPB (Top and bottom row image scale bars, 100  $\mu$ m and 50  $\mu$ m, respectively). From Figure 10 from (Kang et al. 2020).





**Figure 15.** Precision Cut Lung Slices (PCLS) from human lungs of 2 human fetuses (Ages:  $16^{+0}$  and  $16^{+3}$  weeks) transduced by AAV6.2FF vectors.







**a** Quantification of bioluminescence from the IVIS image in Fig 10c as Total Flux (photons per second). Total Flux of AAV-Luc wells are presented as the mean with SD ( $n=2$  fetuses; 2 lung slices per fetus per dose for AAV-Luc). **b** Representative H&E images from human fetal PCLS that were fixed in 4% PFA immediately after the IVIS imaging in Fig 14c (Scale bar,  $100\mu\text{m}$ ). The lungs demonstrate established acinar structures and are in the



late pseudoglandular stage of development. **c** Representative epifluorescence images of luciferase (red), mCherry (red), and DAPI (blue) in frozen PCLS 4 days after transduction with  $10^8$  vg per well of AAV-Luc or AAV-mCherry (Scale bar,  $100\mu\text{m}$ ). **d** Representative epifluorescence image of EpCAM (red) and DAPI (blue) of frozen PCLS from human lung sections demonstrating the presence of epithelial cells in developing future alveoli. **e** Representative epifluorescence image of DAPI (blue), EpCAM (green), and luciferase (red) or mCherry (red) in frozen PCLS 4 days after transduction with  $10^{10}$  vg per well of AAV-Luc or  $10^8$  vg per well of AAV-mCherry (Scale bar,  $100\mu\text{m}$ ). From Supplementary Figure 11 from (Kang et al. 2020).

#### **4. Article Kang et al. *Nature Communication* 2020**

# A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency

Martin H. Kang <sup>1,2</sup>, Laura P. van Lieshout<sup>3,8</sup>, Liqun Xu<sup>1,2,8</sup>, Jakob M. Domm <sup>3</sup>, Arul Vadivel<sup>1,2</sup>, Laurent Renesme<sup>1,2</sup>, Christian Mühlfeld<sup>4</sup>, Maria Hurskainen<sup>1,2</sup>, Ivana Mižiková<sup>1,2</sup>, Yanlong Pei<sup>3</sup>, Jacob P. van Vloten<sup>3</sup>, Sylvia P. Thomas<sup>3</sup>, Claudia Milazzo<sup>1,2</sup>, Chanèle Cyr-Depauw <sup>1,2</sup>, Jeffrey A. Whitsett <sup>5</sup>, Lawrence M. Noguee <sup>6</sup>, Sarah K. Wootton <sup>3,9</sup>✉ & Bernard Thébaud<sup>1,2,7,9</sup>✉

Surfactant protein B (SP-B) deficiency is an autosomal recessive disorder that impairs surfactant homeostasis and manifests as lethal respiratory distress. A compelling argument exists for gene therapy to treat this disease, as *de novo* protein synthesis of SP-B in alveolar type 2 epithelial cells is required for proper surfactant production. Here we report a rationally designed adeno-associated virus (AAV) 6 capsid that demonstrates efficiency in lung epithelial cell transduction based on imaging and flow cytometry analysis. Intratracheal administration of this vector delivering murine or human *proSFTPB* cDNA into SP-B deficient mice restores surfactant homeostasis, prevents lung injury, and improves lung physiology. Untreated SP-B deficient mice develop fatal respiratory distress within two days. Gene therapy results in an improvement in median survival to greater than 200 days. This vector also transduces human lung tissue, demonstrating its potential for clinical translation against this lethal disease.

<sup>1</sup>Sinclair Center for Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, ON K1Y 4E9, Canada. <sup>2</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON K1N 6N5, Canada. <sup>3</sup>Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada. <sup>4</sup>Institute of Functional and Applied Anatomy, Hannover Medical School, 30625 Hannover, Germany. <sup>5</sup>Divisions of Neonatology and Pulmonary Biology, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA. <sup>6</sup>Division of Neonatology, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>7</sup>Neonatology, Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, Ottawa, ON K1H 8L1, Canada. <sup>8</sup>These authors contributed equally: Laura P. van Lieshout, Liqun Xu. <sup>9</sup>These authors jointly supervised this work: Sarah K. Wootton, Bernard Thébaud. ✉email: [kwootton@uoguelph.ca](mailto:kwootton@uoguelph.ca); [bthebaud@ohri.ca](mailto:bthebaud@ohri.ca)

Surfactant Protein B (SP-B) is crucial in proper pulmonary surfactant assembly by influencing lipid packing, stabilizing lipid layers, and reducing surface tension<sup>1</sup>. Infants born with SP-B deficiency (OMIM#265120) present with rapidly progressive respiratory failure<sup>2</sup>, and without lung transplantation this disease is lethal within 3 to 6 months of birth<sup>3</sup>. Lung transplantation remains a rare procedure in neonates due to preferential allocation of donor lungs for adults or older pediatric patients<sup>4</sup>, and unique anatomical and physiological requirements of neonatal patients such as size matching<sup>5</sup>. Genetic causes of pulmonary surfactant protein deficiencies also contribute to an estimated 25% of all severe refractory diffuse lung diseases<sup>6</sup>. This highlights the need to identify alternative therapeutic strategies.

Treatment with exogenous pulmonary surfactant even when enriched for SP-B protein is ineffective<sup>7</sup>, and only selective expression of SP-B from alveolar type 2 (AT2) cells maintains proper surfactant homeostasis<sup>8</sup>. Because the lung is a barrier organ in contact with the external environment, targeted delivery of gene therapy to AT2 cells on the alveolar surface has been considered a promising therapeutic strategy. However, earlier attempts at using adenoviral vectors were unsuccessful in expressing SP-B protein in vivo<sup>9,10</sup>. The challenges in delivering transgenes to AT2 cells are formidable due to obstacles such as the immune response, the requirement of cell-surface receptors for entry, and the barrier properties of respiratory mucus and alveolar fluid<sup>11,12</sup>. Alternative therapeutic strategies including *SFTPB* mRNA therapy<sup>13</sup>, gene editing<sup>14</sup>, and electroporation of *SFTPB* cDNA<sup>15</sup> have been unable to establish a median survival longer than 30 days in SP-B deficient mice<sup>16</sup>.

Recently, we engineered an AAV capsid containing an amino acid substitution (F129L) that facilitates heparin binding (AAV6.2) at the cell surface<sup>17</sup>, and 2 mutations (Y445F, Y731F) that abrogate ubiquitin-mediated degradation (AAV6.2FF). Although this vector displayed the ability to transduce the lung parenchyma<sup>18</sup>, its ability in targeting lung epithelial cells and in vivo therapeutic potential were unexplored.

In this study, we demonstrate that AAV6.2FF transduces airway and alveolar epithelial cells. This rationally designed AAV6 based vector primarily targets cells that demonstrate high expression levels of the cell surface epithelial cell adhesion molecule (EpCAM) marker in the lung parenchyma, which includes AT2 cells. Intratracheal administration of AAV6.2FF delivering either murine or human *SFTPB* cDNA transgene into a SP-B deficient mouse model restores SP-B expression, maintains lamellar body (LB) structure, and improves lung function resulting in extended survival. The clinical relevancy of this gene therapy is demonstrated by the rapid expression of SP-B within days of administration, long-term expression of therapeutic SP-B protein levels, efficacy and safety in neonatal mice, the absence of adverse effects as observed by increases in body weight and the lack of a pro-inflammatory cytokine profile, and the ability to transduce human lung tissue. Due to its propensity to transduce both airway and alveolar epithelial cells, this vector may have the potential to target a number of other monogenetic respiratory diseases.

## Results

**AAV6.2FF transduces airway and alveolar epithelial cells.** To establish the suitability of AAV6.2FF as a vector for lung epithelial cells, the extent and duration of transgene expression in the lungs following intratracheal (IT) administration of 10<sup>11</sup> vector genomes (vg) per mouse of AAV6.2FF-Luciferase (AAV-Luc) was determined. All plasmids contained the inverted terminal repeats from serotype 2 (rAAV2), the CAG or composite CASI promoter validated for lung expression<sup>19</sup>,

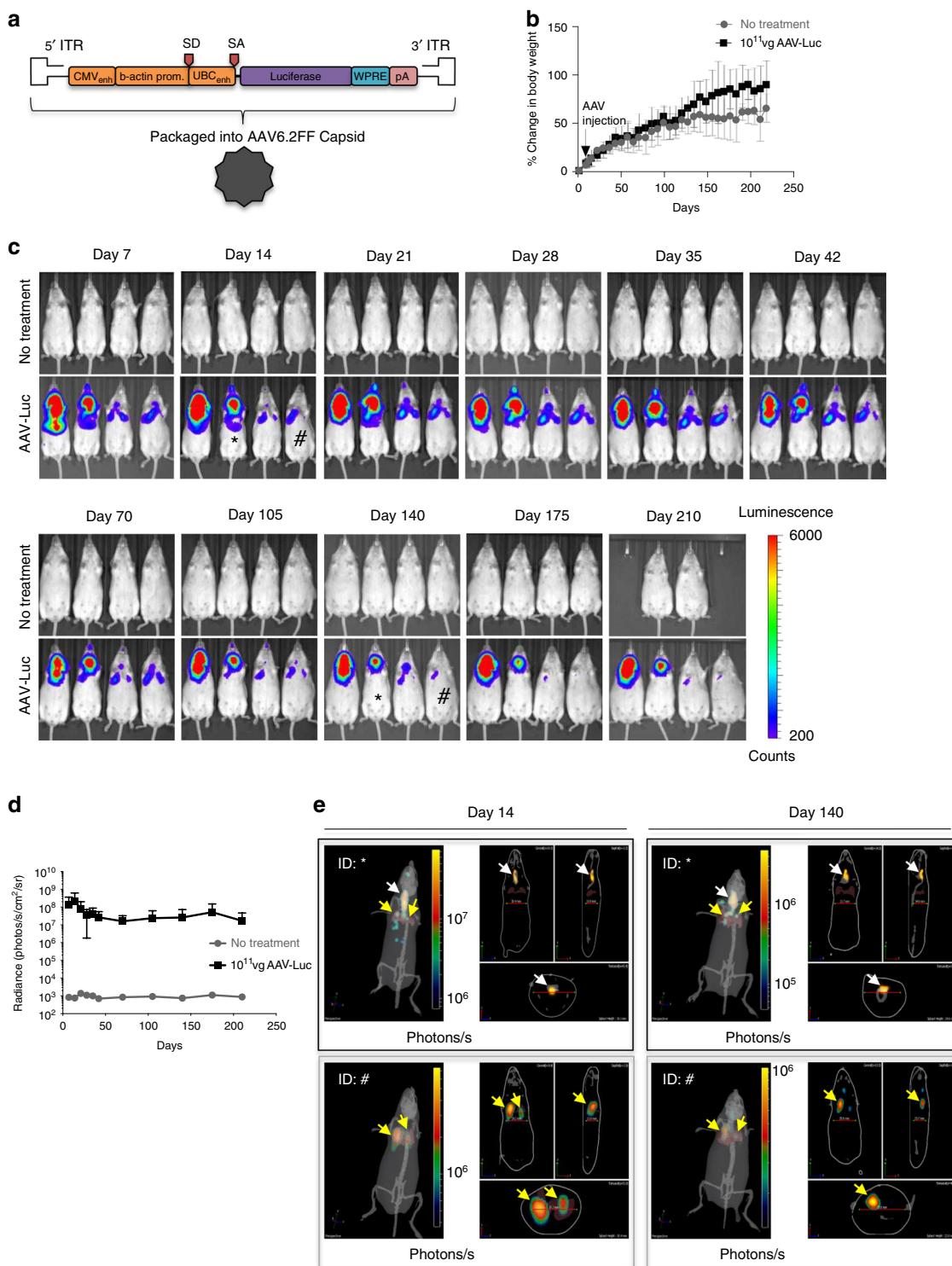
the woodchuck posttranscriptional regulatory element, and a SV40 polyadenylation tail<sup>20</sup> (Fig. 1a). Comparable increases in body weight to untreated mice indicated that exposure to AAV6.2FF vector did not affect general health (Fig. 1b). Signal as measured by radiance from the thorax (lung) region was observed 7 days post-injection, with a peak at 14 days, followed by a slightly lower but sustained expression over the ensuing 200 days (Fig. 1c, d). We confirmed that luciferase expression originated from the lungs (yellow arrows) through tomographic reconstructions of the In Vivo Imaging System (IVIS) figures (Fig. 1e), and IVIS imaging of precision cut lung slices (PCLS)<sup>21</sup> following AAV-Luc transduction (Supplementary Fig. 1a–c).

To demonstrate cell specificity of AAV6.2FF, 10<sup>11</sup> vg of AAV6.2FF-emGFP-nlsCre vector was administered into Rosa 26 floxed-LacZ reporter mice (Rosa26-Flox/LacZ; JAX Stock No 003474) by intranasal (IN) delivery (Supplementary Fig. 2a, b). Any cell from this transgenic reporter mouse line transduced with this vector results in Cre recombinase-mediated removal of its loxP sites. Subsequent  $\beta$ -galactosidase ( $\beta$ -gal) expression can be observed by X-gal staining. Three weeks after vector delivery, cells lining the nasal cavity, airways, and distal lung were X-gal positive including AT2 and club (Clara) cells (Fig. 2a, b, Supplementary Fig. 2c). However, IN administration did not result in AAV6.2FF vector transduction of non-respiratory organs and tissues (Supplementary Fig. 2d).

Flow cytometry studies further confirmed that AAV6.2FF targets AT2 or ciliated lung epithelial cells. SP-B deficient mice were administered 10<sup>11</sup> vg per mouse of either AAV-Luc (control) or AAV6.2FF-GFP (AAV-GFP) (Supplementary Fig. 3a, b). AT2 cells were stratified from whole dissociated lung tissues based on the lack of CD45 or CD31 staining combined with high (++) epithelial cellular adhesion molecule (EpCAM) expression<sup>22</sup> (Supplementary Fig. 3c–e). This cell population demonstrated high intracellular granularity ostensibly due to the presence of lamellar bodies (LB) (Supplementary Fig. 3f), reinforcing their identity as AT2 cells. In mice administered AAV-GFP, 21.6% of AT2 or ciliated cells expressed GFP (Fig. 2c), while more than 78% of all the cells in the lungs that stained for GFP were AT2 or ciliated cells (Fig. 2d, Supplementary Data File 1). Transduction of AT2 cells was visually supported by immunofluorescence (IF) staining of alveoli from 10<sup>11</sup> vg of AAV-Luc treated mice, which demonstrated co-localization of Luc with proSP-C expressing AT2 cells (Supplementary Fig. 3g, i). Although AAV6.2FF vector transduces multiple cell types following intramuscular (18), or intravenous administration (Supplementary Fig. 4), exclusive targeting of respiratory cells can be ensured by restricting vector delivery through the nasal passage and trachea.

## Murine *proSftpb* cDNA (AAV-mSPB) attenuates lung injury.

To assess the feasibility of AAV6.2FF delivery of *proSftpb* cDNA as a gene therapy for SP-B deficiency, we utilized an inducible transgenic mouse that expresses SP-B in the presence of doxycycline (dox)<sup>16</sup>, as the loss of function of both *Sftpb* alleles results in lethal respiratory distress within 20 min of birth<sup>23</sup>. Following dox removal, SP-B expression ceases and the majority of death from respiratory distress occurred within 2 to 7 days, which is similar to the previously reported estimate of 7.5 days  $\pm$  3.5 days (4 to 11 days)<sup>16</sup>. Codon optimized murine *proSftpb* cDNA with a C-terminal myc tag was cloned into the rAAV2 expression plasmid (AAV-mSPB; Fig. 3a). AAV-mSPB transduction of HEK293 cells (Multiplicity of Infection = 20,000 vg per cell)<sup>24</sup>, or the transient transfection of the pAAV-mSPB expression plasmid resulted in proSP-B expression at the correct



molecular weight (42 kDa)<sup>25</sup> (Fig. 3b). We resolved whether AAV-mSPB prevents the lung phenotype of SP-B deficient mice by comparing mice on dox, to mice administered 10<sup>11</sup> vg per mouse of AAV-Luc or AAV-mSPB off dox (Fig. 4a, Supplementary Fig. 5a). Following IT delivery, we waited 4 weeks before removing the dox, as clinical trials show that transgenes from AAV vectors can take 1 month before reaching peak expression levels<sup>26</sup>. We assessed the lung phenotype 3 days following dox removal, which was before the earliest reported incidence of mortality (4 days)<sup>16</sup>.

Body weight revealed itself as a marker of impending respiratory distress. While there were no body weight differences on dox following AAV administration (Fig. 4b, Supplementary Fig. 5b, c), AAV-mSPB mitigated the weight loss observed in AAV-Luc mice following dox removal (Fig. 4c). Three AAV-Luc mice died before the 3-day experimental endpoint and were not included in the assessment of lung function, while all AAV-mSPB mice survived until the harvest on day 3.

Macroscopic tissue injury as indicated by lung lesions (yellow arrows) were observed in 6 out of 7 AAV-Luc mice (Fig. 4d).

**Fig. 1 AAV6.2FF mediates long-term transgene expression in the lungs.** **a** Schematic of the recombinant AAV2 (rAAV2) vector genome containing the firefly luciferase reporter gene pseudotyped with the AAV6.2FF capsid (AAV-Luc). ITR, inverted terminal repeat; CMV<sub>enh</sub>, human cytomegalovirus immediate early gene enhancer region; b-actin prom, chicken beta actin promoter; SD, splice donor; SA, splice acceptor; UBC<sub>enh</sub>, human ubiquitin C promoter; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element; pA, simian virus 40 polyadenylation signal. **b** Mean percent change in body weight of untreated mice (gray circles) versus AAV-Luc (black squares) treated mice with standard deviation (SD).  $n = 4$  biologically independent animals per group. **c** Bioluminescence detection using the In Vivo Imaging System (IVIS) in transgenic SP-B mice either untreated or intratracheally (IT) administered with  $10^{11}$  vector genomes (vg) per mouse of AAV-Luc. Two untreated mice died before imaging on Day 210. All IVIS images were normalized for a signal count ranging from 200 to 6000. The counts scale represents an uncalibrated measurement of photon incidents in a pixel and allows for the normalization between images acquired on different days. **d** Quantification of the IVIS images from the thorax (lung) region in untreated mice (gray circles) versus AAV-Luc (black squares) treated mice. Data are presented as the mean radiance with SD. Radiance is a calibrated measurement of photon emissions and is expressed as the number of photons per second that leave a square centimeter of tissue and radiate into a solid angle of 1 steradian (photons per sec per cm<sup>2</sup> per sr; Caliper Life Sciences). ( $n = 4$  biologically independent animals per group). **e** Confirmation of lung (yellow arrows) and tracheal (white arrows) region expression by 3D diffuse light imaging tomography (DLIT) reconstruction of the IVIS images on days 14 and 140 in the mice identified by the [\*] and [#] symbols. The color scale for the DLIT images are presented as the total flux, which is expressed as the number of photons per second.

Mice on dox showed no visible signs of lung injury, and 3 out of 10 AAV-mSPB mice demonstrated some gross lung damage (yellow arrows). Hematoxylin and Eosin (H&E) and Wright-Giemsa Jenner (WGI) staining revealed hallmarks of diffuse alveolar damage<sup>27</sup>, including infiltration of alveolar spaces by red-blood cells in some AAV-Luc mice (Fig. 4e).

*Sftpb* expression from the vector was confirmed by real-time qPCR (SYBR Green) measurements. We designed the forward primer to bind the murine *proSftpb* sequence, and the reverse primer to the myc sequence in order to differentiate between exogenous transgene and endogenous SP-B expression. An increase greater than 400,000-fold of murine *proSftpb*-Myc was observed in AAV-mSPB treated mice ( $n = 10$ ) compared to mice on dox ( $n = 10$ ;  $P = 0.0185$ , one-way ANOVA, Dunnett's post hoc test) in their right lungs (Fig. 4f).

Functionally, AAV-Luc mice appeared to display decreased lung distensibility (Fig. 4g), as well as a decreased %V<sub>10</sub> indicating high alveolar surface tension (Fig. 4h). The %V<sub>10</sub> quantifies the stability of the lung during deflation and decreased values are associated with surfactant inhibition<sup>28–30</sup>. The percentage of the lung volume at 10 cmH<sub>2</sub>O divided by the lung volume at 30 cmH<sub>2</sub>O yields the %V<sub>10</sub> as demonstrated in Eq. (1).

$$\%V_{10} = V_{10} \div V_{30} \quad (1)$$

In addition, total lung capacity (TLC), residual volume (RV), and compliance<sup>31</sup> were all significantly reduced in AAV-Luc mice compared to mice on dox or in mice treated with AAV-mSPB (Fig. 4i–k). All lung function data was normalized to the body weight of each individual mouse.

IF staining confirmed the presence of proSP-C in all groups; however, SP-B expression was consistent only in mice on dox or following AAV-mSPB treatment. In the AAV-Luc mice, SP-B staining was periodically observed (Supplementary Fig. 5d), suggesting that some endogenous SP-B remained present within 3 days of dox removal. To demonstrate that SP-B expression was solely from the vector, SP-B was stained in the lungs of AAV-mSPB treated mice off dox for 60 days. At the time of harvest these mice showed no signs of respiratory distress, with one mouse displaying only a small region of lung damage (yellow arrow) (Fig. 5a). SP-B staining was comparable to mice on dox, while no SP-B staining was detected in untreated mice off dox after 9 days (Fig. 5b, Supplementary Fig. 5e). All IF-imaging controls are shown in Supplementary Fig. 5f, g.

Pulmonary surfactant is stored and secreted from LB<sup>32,33</sup>, the lysosomal-like structures which present with abnormal assembly in SP-B deficient patients<sup>1,2</sup>. Ultrastructural images from AAV-Luc mice revealed small LB often containing central accumulations of amorphous material (white arrows) with few surrounding

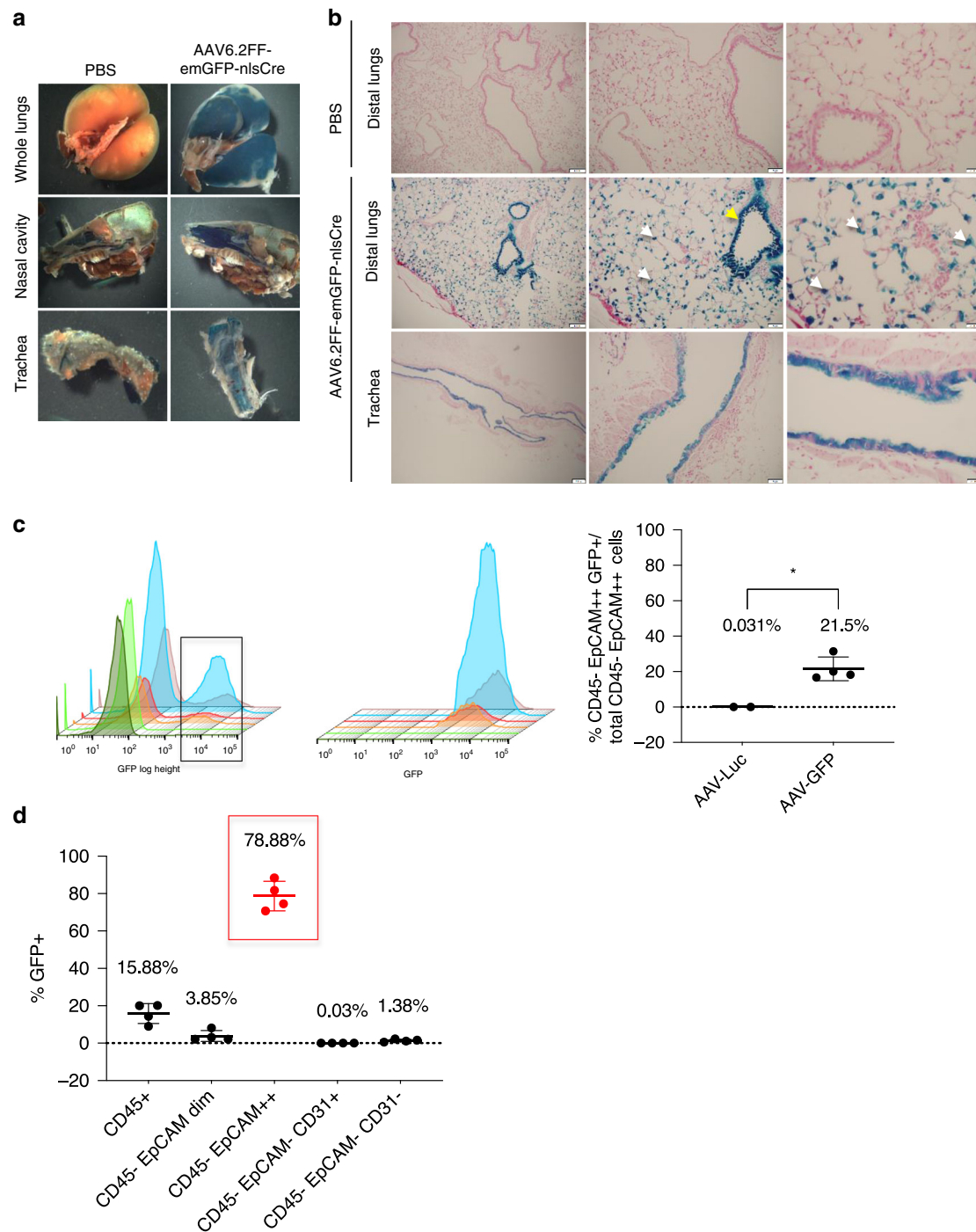
lamellae (Fig. 5c, Supplementary Fig. 5h, i). The LB from mice on dox and following AAV-mSPB treatment exhibited the characteristic multi-layered appearance of densely stacked phospholipid membranes<sup>32,33</sup> typical of AT2 cells with proper surfactant homeostasis.

**AAV-mSPB improves survival in a dose-dependent manner.** To determine whether AAV-mSPB improves viability, survival studies were carried out where we monitored for signs of respiratory distress indicating euthanasia (Supplementary Data File 2), or recorded when mice were found expired using the outlined study design (Fig. 6a). All mice were IT administered with a single injection of AAV-mSPB or a negative control (1 × PBS, AAV-Luc, or untreated). Equivalent mortality outcomes using different negative control groups demonstrated the consistency in lethal respiratory distress development in this animal model. A low dose ( $10^{10}$  vg per mouse) of AAV-mSPB improved median survival to 5.05 days compared to 3.89 days in 1 × PBS mice ( $P = 0.0067$ , Log-rank, Mantel-Cox), with the longest surviving mouse living 40 days off dox. At an intermediate dose ( $10^{11}$  vg per mouse) of AAV-mSPB, median survival increased to 20 days compared to 3 days in AAV-Luc treated mice ( $P < 0.0001$ , Log-rank, Mantel-Cox), with 1 mouse surviving for 71 days off dox. In the highest dose ( $5 \times 10^{11}$  vg per mouse), we observed a median survival of 128 days ( $P < 0.0001$ , Log-rank, Mantel-Cox), with 1 mouse surviving for over 6 months (200 days) without dox (Fig. 6b,c). Comparison of the survival curves indicates that AAV-mSPB operates in a dose-dependent manner. The majority (57 out of 102) of negative control animals in our study suffered fatal respiratory distress before the earliest expected endpoint of 4 days. However, some animals (3 out of 102) demonstrated a longer than expected (greater than 11 days) survival, but either showed a considerable decrease in body weight before death ( $n = 2$ ), (Supplementary Fig. 6a), or exhibited substantial alveolar injury by H&E staining (Supplementary Fig. 6b–f).

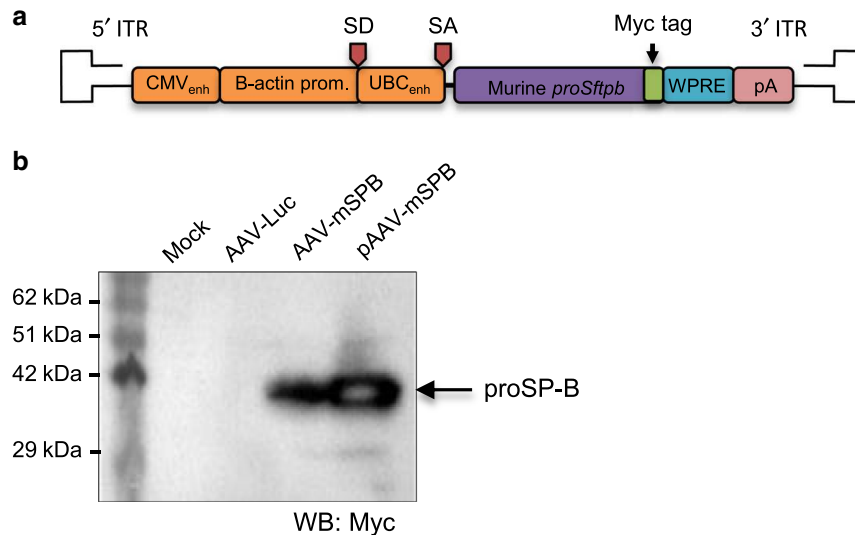
#### Administration methods for AAV delivery to the distal lung.

An issue for many pre-clinical gene therapy studies in small animal models has been the poor translation of promising results into larger animal models or humans due to a lack of clinical relevancy<sup>34,35</sup>, so we explored multiple strategies to address this issue in our study. First, we examined clinically relevant methods for vector distribution in the distal lung. In the clinic, SP-B patients are intubated and placed on a mechanical ventilator as their lungs exhibit a high surface tension often in conjunction with cellular and fluid infiltration<sup>36</sup>. Intubation is one of the most efficient administration routes for gene therapy delivery to the





**Fig. 2 AAV6.2FF transduces both airway and alveolar epithelial cells.** **a** Whole-mount X-gal staining of the lungs, nasal cavity and trachea 3 weeks post-vector administration from either 1 × PBS or AAV6.2FF-emGFP-nlsCre treated Rosa26-Flox/LacZ mice. **b** Representative nuclear fast red counterstained paraffin sections (4 μm) from X-gal stained lungs. LacZ-positive cells were found in the distal airway epithelium and alveoli following delivery of AAV6.2FF-CMV-emGFP-nlsCre. Morphologic criteria demonstrate that both AT2 (white arrows) and club cells (yellow arrow) are X-gal positive (Distal lung scale bars represent 100 μm, 50 μm, and 20 μm from left to right; Trachea scale bars represent 200 μm, 50 μm, and 20 μm from left to right). **c** 3D histograms from flow cytometry analysis of dissociated whole lungs harvested from transgenic SP-B mice 8 days after IT administration of 10<sup>11</sup> vg per mouse of AAV-Luc control (n = 2), or 10<sup>11</sup> vg per mouse of AAV6.2FF-GFP (AAV-GFP; n = 4). The first set of histograms represent the total CD45 negative, EpCAM++ (high expressing) cell population (first peak), and the CD45 negative, EpCAM++ cells that are GFP positive (second peak) in each mouse. The second set of histograms are the rescaled CD45 negative, EpCAM++, GFP positive cell populations as indicated by the black box in the first histogram. The column graph represents the mean percentage of total CD45 negative, EpCAM++ cells that are GFP positive with SD (P value = two-tailed Student’s t test, \*P = 0.0127). **d** This column graph represents the mean percentage of different lung cell populations with SD that are GFP positive. CD45 positive staining represents hematopoietic cells, CD45 negative, EpCAM dim staining represents alveolar type 1 (AT1) epithelial cells, CD45 negative, EpCAM++ staining represents AT2 or ciliated cells (red circles and red box), CD45 negative, CD31 positive staining represents endothelial cells, and CD45 negative, EpCAM negative, CD31 negative represents all other cell types found in dissociated lung tissue (n = 4 biologically independent animals per group). The source data for **c** and **d** have been provided as a Source Data file.



**Fig. 3** Generation of AAV6.2FF murine *proSftpb* cDNA (AAV-mSPB). **a** Schematic of codon optimized murine *proSftpb* cDNA transgene in the rAAV2 vector genome. A myc tag is attached to the C-terminal end of the *proSftpb* sequence. **b** Western blot of cell lysates from HEK293 cells transduced with AAV6.2FF-murine *proSftpb* cDNA (AAV-mSPB) at a Multiplicity of Infection (MOI) of 20,000 or transiently transfected with a *proSftpb* expression plasmid (pAAV-mSPB), probed with an anti-myc-tag antibody. The expression of proSP-B protein is indicated by an arrow. The source data for **b** has been provided as a Source Data file.

distal lung<sup>37</sup>, and pulmonary surfactant has been effective in promoting adenovirus-mediated gene transfer by facilitating vector dispersion in the distal lung by mucus clearance, and alveolar recruitment through reopening of collapsed or fluid-filled alveoli<sup>38,39</sup>. Both the addition of pulmonary surfactant (Bovine Lipid Extract Surfactant; BLES) as a vehicle and administration through an endotracheal cannula reduced the variability (standard deviation, SD) in the signal of lung reporter expression (Supplementary Fig. 7a–i, Supplementary Data File 3). BLES addition to the IT administration of an intermediate dose ( $10^{11}$  vg per mouse) of AAV-mSPB resulted in a median survival of 37 days (Supplementary Fig. 8a, b), and BLES addition to the intubation delivery of an intermediate dose of AAV-mSPB had a median survival of 96 days (Supplementary Fig. 8c, d).

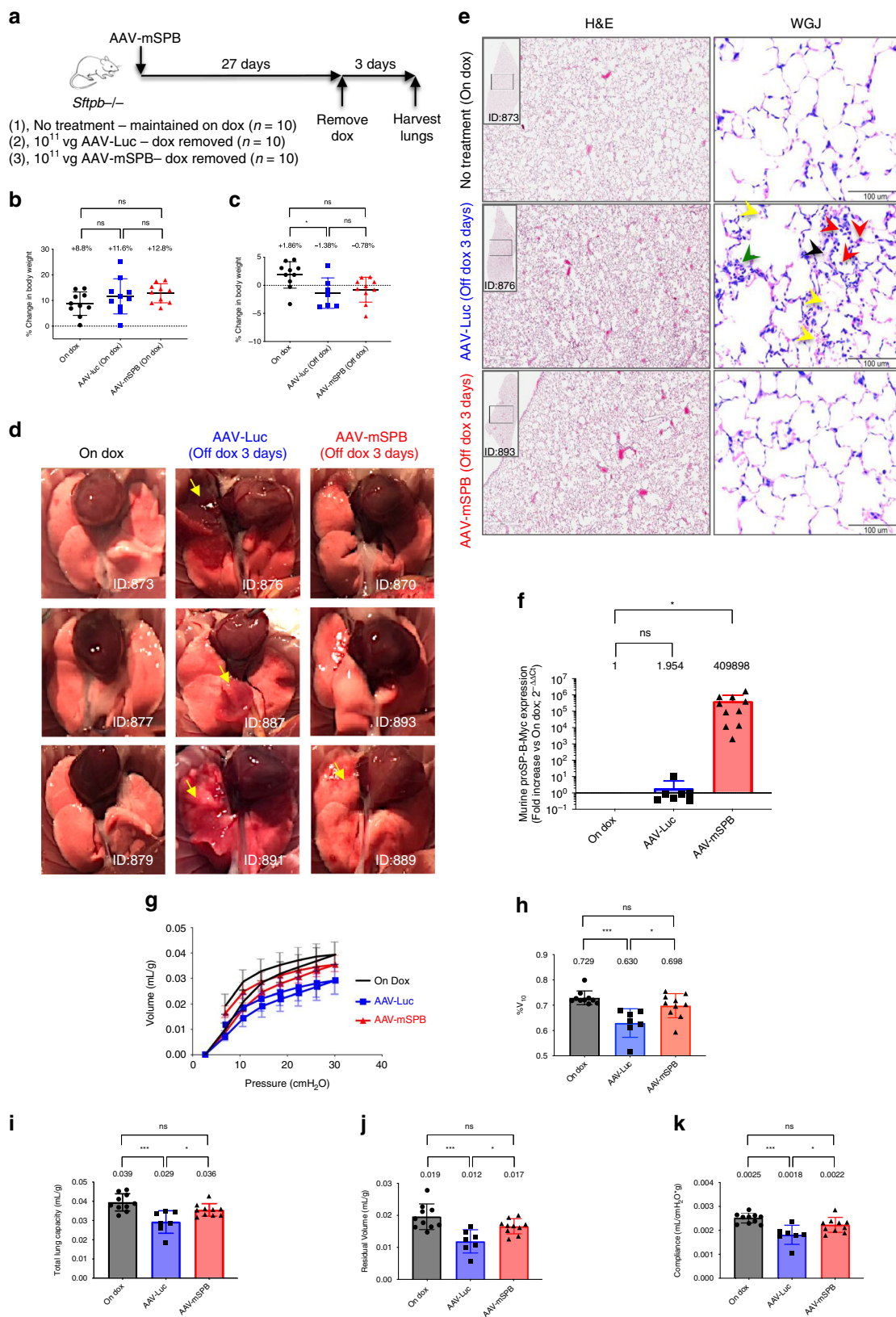
**SP-B rapidly expresses to therapeutic levels.** Second, we determined how soon therapeutic levels of SP-B protein are expressed after AAV administration as rapid SP-B expression from vectors may minimize the alveolar damage caused by surfactant deficiency<sup>36</sup>. SP-B must be expressed above 25% of normal levels in order for proper surfactant homeostasis and respiratory function<sup>16</sup>. We removed dox at 3 (D3) and 7 (D7) days following IT administration of an intermediate dose ( $10^{11}$  vg per mouse) of AAV-mSPB. At both time points, enough SP-B protein was present to eliminate the fatalities associated with SP-B deficiency (Fig. 6d, Supplementary Fig. 8e). The median survival for the D3 group was 137 days and for the D7 group 194 days, with one mouse living for 287 days off dox in this latter group.

**AAV-mSPB does not cause an inflammatory cytokine profile.** Third, the safety of AAV-mSPB was assessed by determining whether it caused systemic inflammation in an experiment comparing 3 groups: IT delivery of  $1 \times$  PBS maintained on dox ( $n = 7$ ), IT delivery of an intermediate dose of AAV-mSPB with dox removed 7 days after administration ( $n = 8$ ); and intraperitoneal injection of 3 mg per kg of lipopolysaccharides (LPS) from *E. coli* O111:B4 maintained on dox ( $n = 8$ ) (Fig. 6e). LPS is a highly endotoxic component of Gram-negative bacteria and its administration into animals elicits a wide-ranging physiological

response including systemic inflammation. LPS-induced systemic inflammation results in peak inflammatory cytokine levels 2 to 8 h after injection, with a return to baseline within 12 to 24 h<sup>40</sup>. Serum was collected 7 days before (–7D) and 1 (+1D), 7 (+7D), and 29 days (+29D) after AAV or LPS administration. A 3 mg per kg dosage of LPS is considered sublethal in adult mice, however 5 of 8 mice died post-injection, while 1 AAV-mSPB treated mouse died within 5 days following dox removal (Fig. 6f, g, Supplementary Fig. 8f, g). Thirteen humoral factors were analyzed by multiplex technology using the LEGENDplex (BioLegend) mouse inflammation panel. From –7D to +1D, AAV-mSPB administration did not result in a significant increase in any of the 13 cytokines measured (Fig. 6h). Although the effects of LPS on humoral factors was expected to diminish within 24 h, LPS still displayed a significant increase in the inflammatory cytokines IL-23 ( $P = 0.020$ , Mixed Effects Model, Tukey’s post hoc test), TNF $\alpha$  ( $P = 0.036$ ), and GM-CSF ( $P = 0.045$ ) from –7D to +1D.

**Human *proSFTPB* cDNA (AAV-hSPB) also prevents lung damage.** Another approach in improving clinical relevancy was to test whether human *proSFTPB* cDNA (AAV-hSPB) (Fig. 7a) prevents lung injury and improves survival comparable to murine *proSftpb* cDNA (AAV-mSPB) in the SP-B mouse model. A sequence comparison using BLAST demonstrates that there is 69% identity between the murine and human proSP-B amino acid sequence<sup>41</sup>. We compared mice on dox to mice IT administered  $1 \times$  PBS + BLES (negative control), or  $5 \times 10^{10}$  vg AAV-hSPB + BLES. We waited 4 weeks before removing dox, with lung structure and function assessed 4 days following dox removal (Fig. 7b, Supplementary Fig. 9a, Supplementary Data File 4).

Following AAV-hSPB delivery, these mice displayed similar increases in body weight as the other groups while on dox (Fig. 7c; Supplementary Fig. 9b, c), and significantly mitigated the loss in body weight (–1.7%) following dox removal compared to negative control mice (–11.1%;  $P = 0.0221$ , one-way ANOVA, Tukey’s post hoc test) (Fig. 7d). Seven  $1 \times$  PBS + BLES mice suffered lethal respiratory distress before the 4-day experimental endpoint and presented with pronounced and extensive lung



lesions. Macroscopic lung injury (yellow arrows) was also observed in the remaining 1 × PBS + BLES mice that survived, while only 4 out of 11 AAV-hSPB mice demonstrated signs of lung damage (yellow arrows) (Fig. 7e). Representative H&E and WGJ staining revealed large levels of cellular infiltration in the alveolar spaces of 1 × PBS treated negative control mice, but no

signs of alveolar damage in AAV-hSPB treated mice (Fig. 7f, Supplementary Fig. 9d). To demonstrate that lung injury was not caused by post-mortem decomposition, healthy mice on dox euthanized 24 h before lung collection did not exhibit comparable levels of damage as the 1 × PBS treated negative control mice (Supplementary Fig. 9e, f). IF images demonstrated proSP-C



**Fig. 4 AAV-mSPB prevents the lung phenotype of SP-B deficient mice.** **a** Study design to determine whether  $10^{11}$  vg per mouse of AAV-mSPB improves lung structure and function in SP-B deficient mice. **b** Percentage change in body weight following AAV administration while still on doxycycline (dox) (27-day duration). **c** Percentage change in body weight after dox removal (3-day duration). All body weight data in **c** and **d** are presented as the mean with SD ( $n = 10$ , except AAV-Luc post-dox where  $n = 7$ ;  $*P = 0.0294$ ). **d** Representative gross lung images 3 days following dox removal. Yellow arrows indicate regions of lung injury. Mouse IDs are provided to match with histological or epifluorescence lung images when available. **e** Representative Hematoxylin and Eosin (H&E) and Wright-Giemsa Jenner (WGJ) staining of paraffin embedded whole left lungs 3 days after dox removal. In the WGJ staining of the AAV-Luc lung, the red arrows indicate the ring-shaped and/or pretzeloid-shaped nuclei of neutrophils, the black arrow indicates hyaline membrane deposition, the green arrow indicates alveolar septal wall thickening, and the yellow arrows indicate red blood cells (H&E scale bar, 500  $\mu\text{m}$ ; WGJ scale bar, 100  $\mu\text{m}$ ). **f** The expression levels of murine proSP-B-Myc from right lung tissues 3 days after dox removal. All comparative qPCR data are presented as the mean with SD, and the On Dox group acts as a reference point with a value of 1 ( $n = 10$ , except AAV-Luc where  $n = 7$ ;  $*P = 0.0185$ ). **g** Pressure-volume curve 3 days following dox removal corrected for body weight (in mL per g) for mice on dox (black circles), mice treated with AAV-Luc off dox (blue squares), and mice treated with AAV-mSPB off dox (red triangles). The source data for **g** has been provided as a Source Data file. **h**  $\%V_{10}$  corrected for body weight ( $*P = 0.0113$ ,  $***P = 0.0003$ ). **i** Total Lung Capacity (TLC) corrected for body weight (in mL per g;  $*P = 0.0243$ ,  $***P = 0.0003$ ). **j** Residual Volume (RV) corrected for body weight (in mL per g;  $*P = 0.0239$ ,  $***P = 0.0003$ ). **k** Compliance corrected for body weight (in mL per  $\text{cmH}_2\text{O}$  per g;  $*P = 0.0282$ ,  $***P = 0.0002$ ). All lung function data are presented as the mean with SD ( $n = 10$  except AAV-Luc post-dox where  $n = 7$ ). (All  $P$  values in Fig. 4 = 2-tailed ordinary one-way ANOVA with Tukey's or Dunnett's (**f** only) multiple comparisons post hoc test, ns = not significant).

staining in all groups; however, SP-B expression was consistent only in the mice on dox or following AAV-hSPB treatment (Fig. 7g, Supplementary Fig. 9g, h).

Functionally,  $1 \times$  PBS + BLES mice displayed reduced lung distension (Fig. 7h), as well as surfactant inhibition through a lower  $\%V_{10}$  calculated using Eq. (1) compared to AAV-hSPB treated mice (0.571 versus 0.694, respectively;  $P = 0.0169$ , one-way ANOVA, Tukey's post hoc test) (Fig. 7i). Although TLC, RV, and compliance were all reduced in the  $1 \times$  PBS + BLES mice (Supplementary Fig. 9i–k), the differences compared to AAV-hSPB treated mice were not statistically significant due to the reduced sample size of the negative control group ( $n = 4$ ).

#### AAV-hSPB improves survival in a dose-dependent manner.

Survival studies following IT administration of  $5 \times 10^{10}$  vg per mouse of AAV-hSPB + BLES resulted in a significant improvement in median survival from 3 days ( $1 \times$  PBS + BLES) to 11 days ( $P = 0.0019$ , Log-rank, Mantel-Cox) (Fig. 8a, Supplementary Fig. 9l, Supplementary Movies 1, 2), resembling our findings with the low dose of AAV-mSPB (Fig. 6b). We observed a dose-dependent increase in survival following IT administration of an intermediate dose ( $10^{11}$  vg per mouse) of AAV-hSPB + BLES, in which dox was removed 7 days following AAV delivery. In this group, median survival increased to 141 days compared to 6.25 days in  $1 \times$  PBS treated mice ( $P < 0.0001$ , Log-rank, Mantel-Cox), with 1 mouse surviving for 206 days off dox. (Fig. 8a, b, Supplementary Movie 3, Supplementary Data Files 5, 6).

As SP-B deficiency is a neonatal genetic disease, a survival study using neonatal mice was performed to further enhance the clinical relevancy of our gene therapy. Neonatal transgenic SP-B deficient mice (postnatal (P)8 to P10) were maintained on dox, treated with  $1 \times$  PBS + BLES, or treated with  $10^{11}$  vg per mouse AAV-hSPB + BLES (Fig. 8c, Supplementary Fig. 9m–o). At P21, the mice were weaned from their mothers and off dox onto a regular chow diet. The  $1 \times$  PBS + BLES mice had a median survival of 5 days, while the AAV-hSPB + BLES neonatal mice had a median survival of 205 days with 3 of 7 mice still alive at 230 days at the time of manuscript submission (Fig. 8d, e).

#### AAV-hSPB improves survival in models of inherited SP-B deficiency.

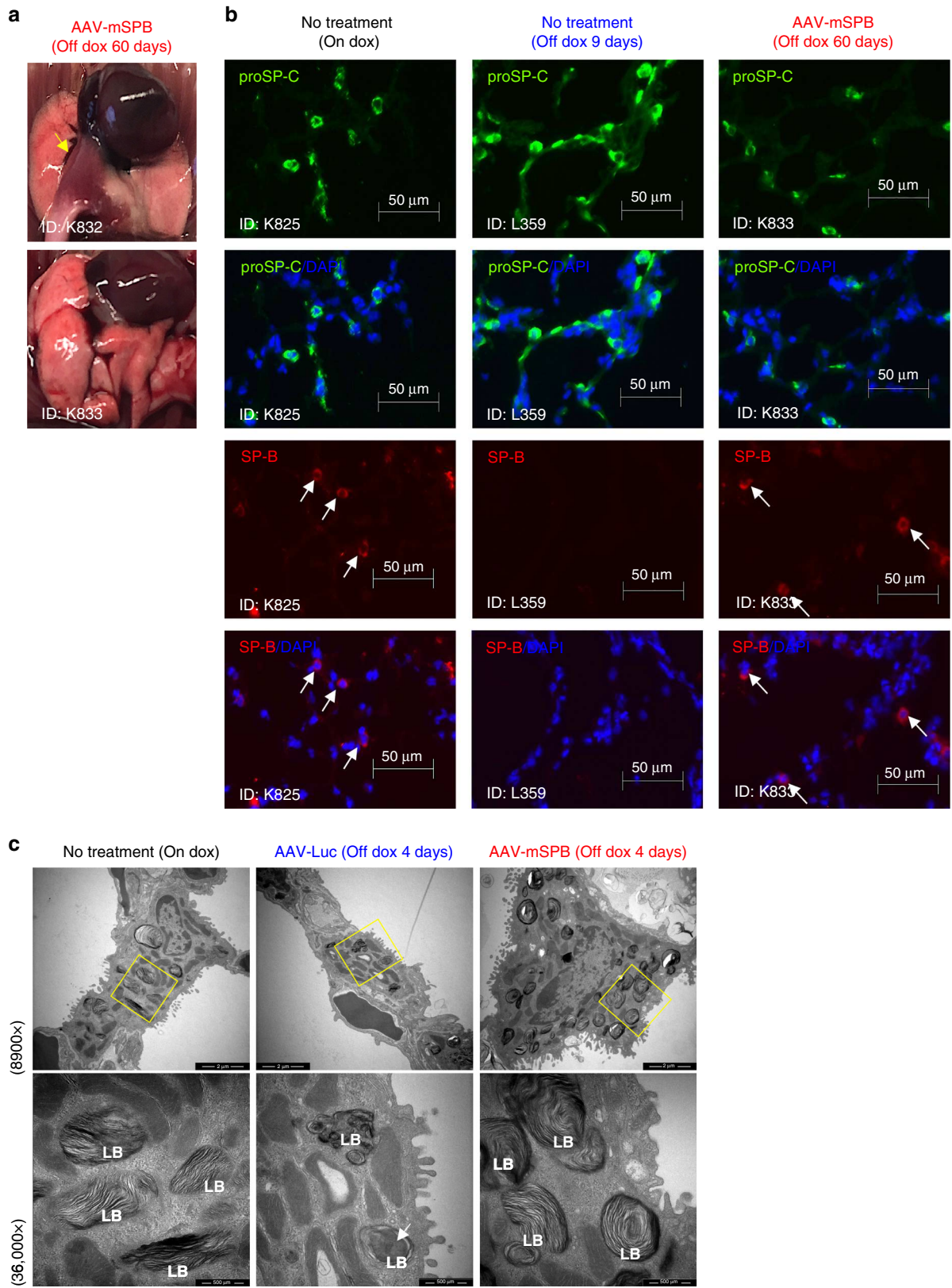
In all the preceding experimental designs, AAV-SPB was administered before dox removal. To assess the therapeutic efficacy of AAV-hSPB to treat inherited SP-B deficiency, dox was removed 2 days (–2D) and 1 day (–1D) before IT administration of an intermediate dose ( $10^{11}$  vg per mouse) of AAV-hSPB + BLES (Fig. 9a). A control group with dox removed 1 day before 1

$\times$  PBS + BLES administration (–1D) had a median survival of 4 days compared to a median survival of 140 days in the –1D group, and 155 days in the –2D group following AAV-hSPB + BLES treatment (Fig. 9b). AAV-hSPB restored initial weight loss caused by dox removal, and both treatment groups demonstrated increasing body weight within days of vector administration (Fig. 9c).

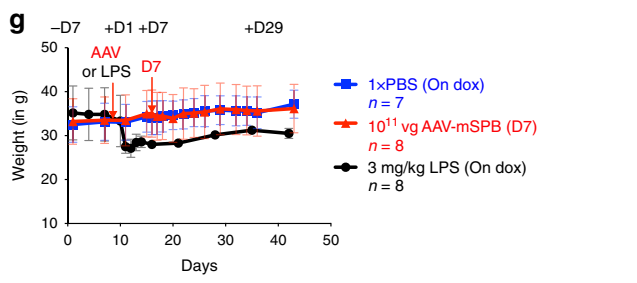
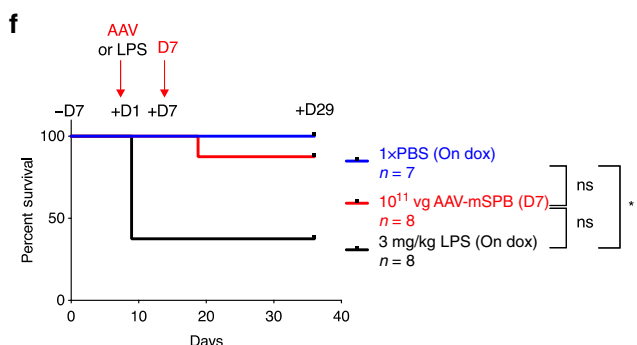
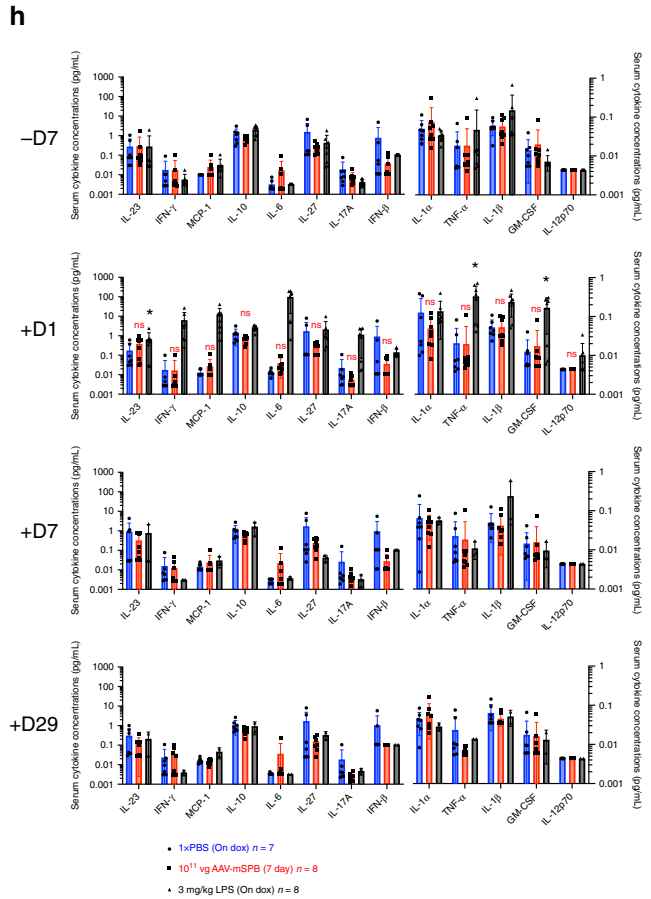
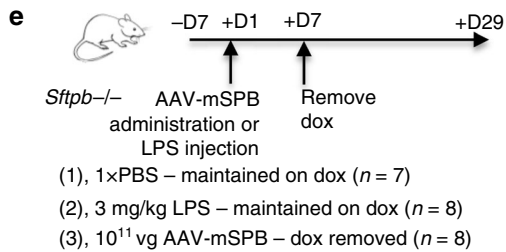
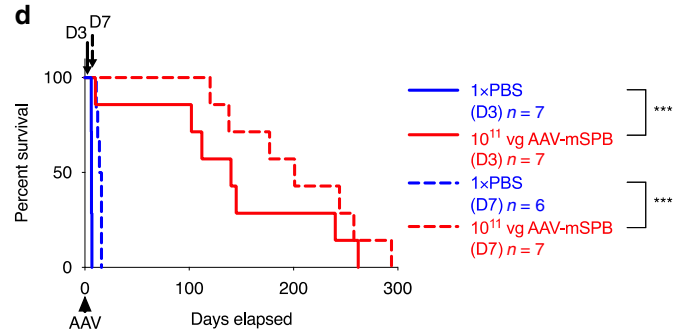
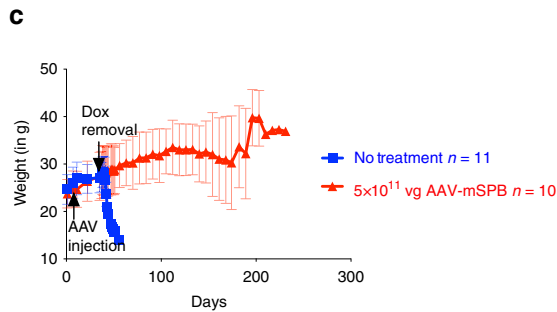
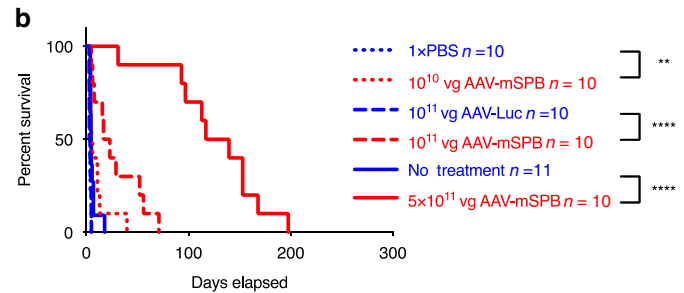
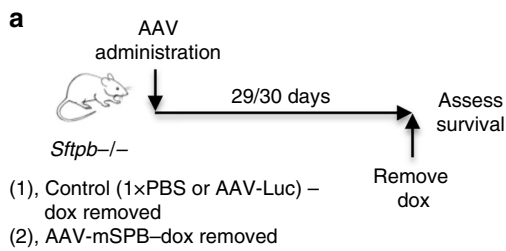
To determine whether our vector could treat inherited surfactant deficiency in a neonatal mouse model, P13 to P15 transgenic SP-B pups were removed from dox immediately following IT administration of  $10^{11}$  vg per mouse AAV-hSPB + BLES. This was accomplished by placing them with nursing FVB/N (the genetic background of the inducible SP-B model is FVB/N)<sup>16</sup> foster mothers that had recently given birth to similar aged litters (Fig. 9d). Median survival in AAV-hSPB + BLES treated mice was 123 days with 1 mouse still alive at 207 days after dox removal at the time of manuscript submission (Fig. 9e, f, Supplementary Movies 4, 5, Supplementary Data Files 5, 6). The examination of lung histology 3 days after vector administration and dox removal revealed a similar alveolar structure to neonatal mice on dox (H&E staining), with no detectable presence of inflammatory cells such as neutrophils in the alveoli (WGJ staining) (Fig. 9g, h).

#### Human lung tissue is transduced by AAV6.2FF.

Finally, we reveal that AAV6.2FF can transduce human fetal lung tissue. PCLS from 2 human fetal lungs (gestational ages  $16^{+0}$  and  $16^{+3}$  weeks; Fig. 10a, b) were transduced with increasing dosages of AAV-Luc at  $10^5$ ,  $10^8$ , and  $10^{10}$  vg per well, and demonstrated a dose-dependent increase in luciferase expression (Fig. 10c, d, Supplementary Fig. 10a). Transduction of human epithelial lung tissue in developing alveoli was confirmed by co-staining for Luciferase or mCherry with EpCAM in AAV-Luc or AAV-mCherry transduced PCLS, respectively (Supplementary Fig. 10b–e). PCLS from a third pair of human fetal lungs (gestational age  $16^{+3}$  weeks; Fig. 10e) were also transduced with  $10^{10}$  vg per well of AAV-mSPB or AAV-hSPB. The lung tissues maintained metabolic activity at an equivalent level to untransduced PCLS 3 days after treatment (Fig. 10f), demonstrating the absence of adverse effects on human tissue transduced with AAV6.2FF vector. The transduced human lung tissue slices exhibited an increase in SP-B expression from both vectors by IF staining (Fig. 10g). Although the ages of the fetuses indicate that lung development is in the late pseudoglandular stage<sup>42</sup> and therefore are too early for the development of mature AT2 cells (Supplementary Fig. 10b, d), this finding demonstrates the capability of AAV6.2FF to safely target the human lung parenchyma.



**Fig. 5 AAV-mSPB restores SP-B expression and maintains lamellar body structure.** **a** Gross lung images from 2 AAV-mSPB treated mice 60 days following dox removal. The yellow arrow indicates a region of lung damage. **b** Representative epifluorescence images of proSP-C (green), SP-B (red), and DAPI (blue) from frozen lung sections of a mouse on dox; an untreated mouse off dox for 9 days; and a  $10^{11}$  vg AAV-mSPB treated mouse off dox for 60 days. The white arrows indicate SP-B staining (Scale bar, 50  $\mu$ m). **c** Representative transmission electron microscope images of AT2 cells 4 days following dox removal. Magnified images (36,000 $\times$ ) of lamellar bodies (LB) are indicated by the yellow boxes on the 8900 $\times$  images. The white arrow indicates the accumulation of amorphous material within the lamellar body (LB) of an AT2 cell from an AAV-Luc lung (Low magnification image scale bar, 2  $\mu$ m; High magnification image scale bar, 500 nm).



**Discussion**

The onset of respiratory impairment in SP-B deficient patients is rapid and lethal<sup>2</sup>, and expeditious, long-term SP-B expression is required to feasibly treat this disease. This study demonstrates a safe, rapidly acting, long-term viral-mediated gene therapy

with improved survival in a pre-clinical small animal model of SP-B deficiency. We achieved this by overcoming a number of obstacles that confounded earlier attempts at identifying a successful viral-mediated gene therapy strategy for this genetic disorder<sup>9,10</sup>.



**Fig. 6 AAV-mSPB improves survival in a dose dependent manner.** **a** Study design to assess survival following AAV-mSPB gene therapy in SP-B deficient mice. Control mice were either untreated or underwent sham surgeries with IT injections of 1× PBS or AAV-Luc. Mice were maintained on dox following AAV administration for 29 to 30 days. **b** Kaplan-Meier survival curves from IT injections of a low (dotted red line;  $10^{10}$  vg per mouse;  $**P = 0.0067$ ), intermediate (dashed red line;  $10^{11}$  vg per mouse;  $****P < 0.0001$ ), and high (solid red line;  $5 \times 10^{11}$  vg per mouse;  $****P < 0.0001$ ) dose of AAV-mSPB. **c** Mean weight measurements with SD of untreated versus mice treated with  $5 \times 10^{11}$  vg of AAV-mSPB from **b**. **d** Kaplan-Meier survival curves following dox removal 3 (D3; solid red line;  $***P = 0.0002$ ) and 7 (D7; dashed red line;  $***P = 0.0005$ ) days after IT administration of  $10^{11}$  vg of AAV-mSPB. **e** Study design to assess inflammation by measuring cytokines from serum in SP-B deficient mice with dox removed 7 days after IT injection of  $10^{11}$  vg of AAV-mSPB (red line). The negative control group was administered 1× PBS and maintained on dox (blue line), while a positive control group was administered 3 mg per kg of *E. coli* O111:B4 lipopolysaccharides (LPS) by intraperitoneal (IP) injection (black line). Serum was collected from mice 7 days before (−7D), and 1 (+1D), 7 (+7D), and 29 days (+29D) after AAV or LPS administration. **f** Kaplan-Meier survival curves following dox removal 7 (D7) days after AAV-mSPB or LPS administration ( $*P = 0.0133$ ). The curve is represented until the final day of serum collection (+29D). The complete survival curve is displayed in Supplementary Fig. 8f. **g** Mean weight measurements with SD of all surviving mice from **f**. **h** Measurements of 13 mouse inflammatory cytokines 7 (−7D) days before and 1 (+1D), 7 (+7D), or 29 (+29D) days after AAV or LPS administration. All cytokine measurements are presented as the mean concentration in pg per mL with SD. The source data for **h** has been provided as a Source Data file. (*P* values in **h** = 2-tailed Mixed Effects Model with the Geisser-Greenhouse correction followed by Tukey's multiple comparisons post hoc test, ns = not significant; IL-23  $*P = 0.0203$ ; TNF- $\alpha$   $*P = 0.0362$ ; GM-CSF  $*P = 0.0456$ ) (All survival curve *P* values in Fig. 6 = 2-tailed Log-rank, Mantel-Cox test, except **f** = 2-tailed Log-rank, Mantel-Cox test with Bonferroni correction, significance is set at  $P = 0.0167$ , ns = not significant).

A major challenge was to identify a vector capable of efficiently transducing AT2 cells in the distal lung with minimal transduction of other cell types<sup>34</sup>. Using imaging and flow cytometry techniques, AAV6.2FF does transduce AT2 cells, but also other cell types and tissues. Despite this, all the evidence from our pre-clinical experiments using the transgenic SP-B small animal model or human PCLS suggests that AAV6.2FF transduction of multiple cell types is non-toxic. This is based on multiple lines of evidence. The first was body weight measurements. Body weight is a primary indicator of the overall well-being and health of the animal<sup>43</sup>, and has been used in previous pre-clinical studies in small animals to demonstrate the safety and effectiveness of gene therapies such as Zolgensma (onasemnogene abeparvovec-xioi)<sup>44</sup>. In our study, AAV6.2FF administration resulted in consistent increases in body weight, although at slightly lower levels than age-matched mice on dox. The second indication for vector safety was the lack of an inflammatory cytokine profile at 1, 7, and 29 days following IT administration of AAV-mSPB. An additional demonstration of vector reliability were the survival studies in which no adverse effects, abnormal behaviors or unexplained injuries were observed in the animals during the course of these experiments. AAV-mSPB mediated long-term survival in these mice with the highest median survival reaching more than 200 days in multiple experiments (Figs. 6d, 8d, Supplementary Fig. 8f), and the longest living subject surviving for 287 days off dox (Fig. 6d). The final piece of evidence supporting vector safety were the resazurin viability assays used to measure metabolic activity in the precision cut lung slices (PCLS) from human embryos. Three days after transduction with AAV-mSPB or AAV-hSPB, tissue slices demonstrated similar levels of viability as untransduced PCLS. From in vivo Cre recombinase studies, we also demonstrate that transduction is localized to the airways or distal lung, with no discernable transduction of non-respiratory tissues, when administration occurs through the trachea or nasal passages.

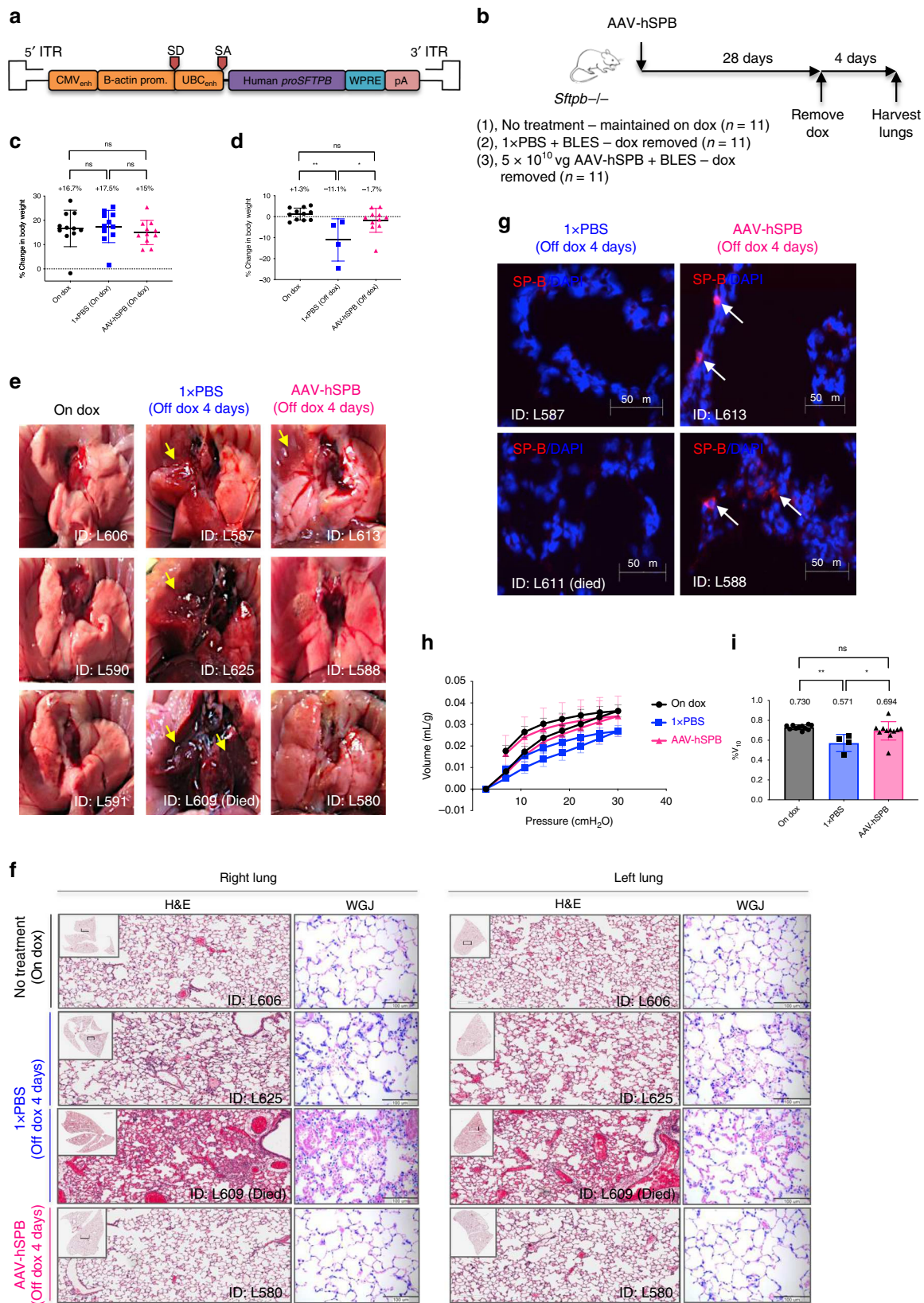
A potential obstacle in the efficacy of this AAV-mediated gene therapy was our use of single-stranded AAV and the associated delays in transgene expression caused by second strand synthesis<sup>45</sup>. However, survival studies demonstrated that therapeutic levels of SP-B protein were generated within 3 days of AAV-SPB administration, and it is possible that therapeutic levels are generated even more rapidly. In the inherited model of SP-B deficiency experiment (Fig. 9a–c), dox was removed 2 days before AAV administration. With a median survival of 3.65 days in the negative control group, SP-B protein would have been required to express at therapeutic levels within 1 to 2 days of administration

in order to rescue the lethal respiratory distress phenotype. Although our gene therapy shows promising results, the study designs and mouse model still do not faithfully recapitulate the disease progression of human SP-B deficient patients that are born without SP-B protein expression. Even in this experiment where dox was removed before AAV administration (Fig. 9a–c), there was likely endogenous SP-B present when AAV was administered as suggested by SP-B staining in IF images of lungs from AAV-Luc negative control mice 3 days off dox (Supplementary Fig. 5d).

AAV presents a low risk of insertional mutagenesis due to its persistence as episomes, however, long-lasting transgene expression occurs only when non-dividing or self-renewing cell populations are transduced<sup>34</sup>. A subset of AT2 cells are considered to be progenitors, and although the exact rate of AT2 cell turnover is unknown it is estimated to be low<sup>46</sup>. Using a luciferase reporter, we demonstrated stable transgene expression for more than 200 days, while survival studies had a number of mice ( $n = 15$  at the time of submission) surviving longer than 200 days (Figs. 6d, 8a, d, Supplementary Fig. 8f).

AAV-SPB acts in a dose-dependent manner with higher dosages demonstrating the most promising therapeutic effects on survival. However, the longest median survivals were observed using an intermediate dose combined with early (day 7 to 13) dox removal (Figs. 6d, 8d, Supplementary Fig. 8f). Part of the improved metrics in survival can be accounted for by the recording of days in which the animals were not on dox while therapeutic expression levels of SP-B were present. This also suggests that the survival studies in which dox was removed 4 weeks after AAV administration markedly underestimates median survival. Eventually the effects of our gene therapy did wane, and we hypothesize that the loss in therapeutic effectiveness is caused by AT2 cell turnover. However, even transient gene therapy can be beneficial in affected individuals. Between 1993 and 2005, 29% of SP-B deficient patients died while waiting for a lung transplant<sup>47</sup>, and AAV-SPB therapy represents a practical bridge treatment until surgery can be performed. Re-administration studies potentially combined with steroids or other immunosuppressants are currently being carried out to assess the feasibility of AAV-SPB as a long-term therapy for SP-B deficiency.

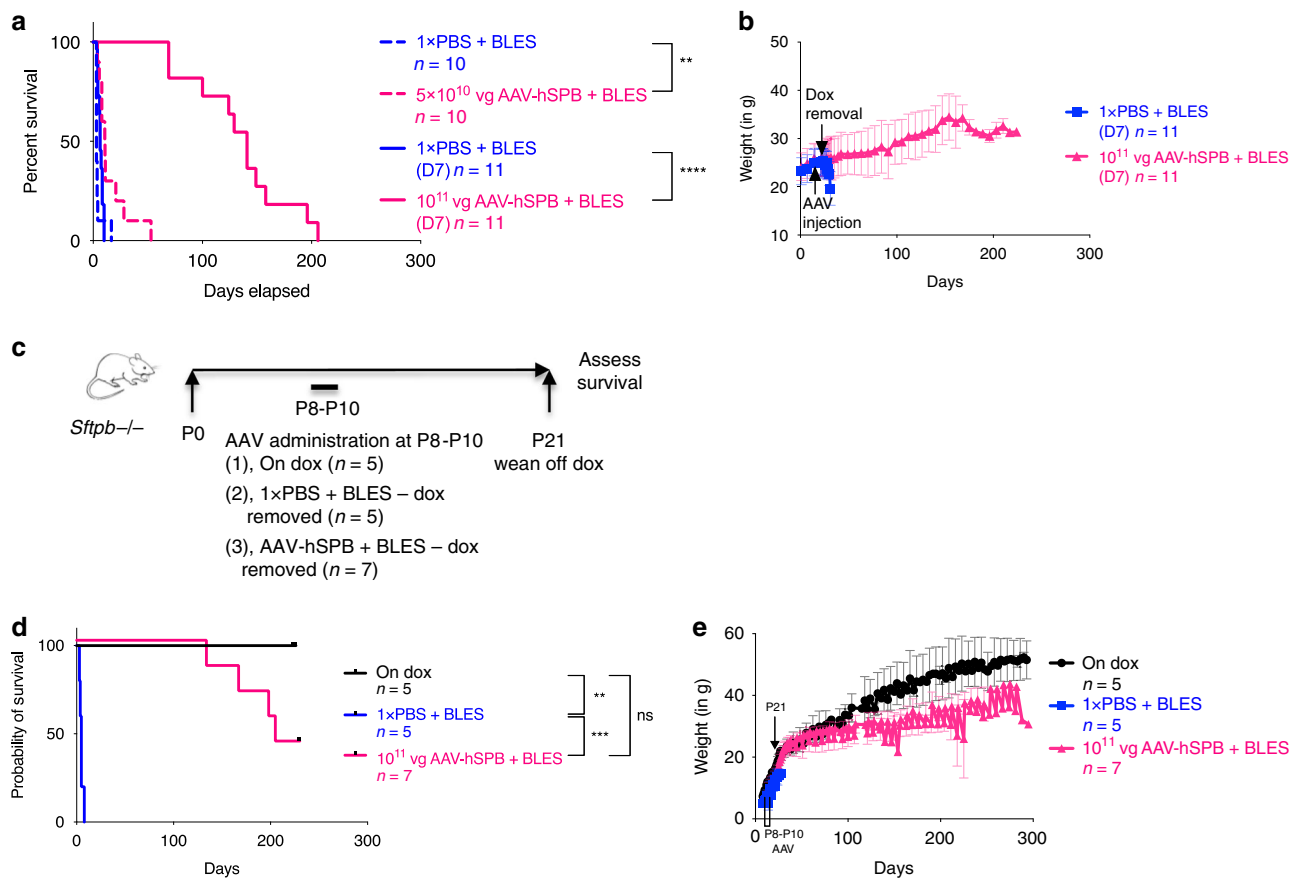
AAV6.2FF has applications beyond gene therapy for SP-B deficiency. This vector can provide a therapeutic platform for other genetic disorders of surfactant deficiency affecting AT2 cells including ABCA3 deficiency (OMIM#610921), or SP-C dominant negative mutations (OMIM#610913)<sup>1</sup>. Patients heterozygous for *SFTPB* mutations may also benefit from gene therapy as lung



disease can develop when other factors reduce their remaining SP-B levels<sup>36</sup>. An alternative therapeutic strategy for SP-B deficiency is genome editing and AAV6.2FF represents a potential vector for delivery of gene editing constructs into AT2 cells or their progenitors.

In summary, this study demonstrates an effective, long-term viral-mediated gene therapy for genetic SP-B deficiency. Due to the lethal nature of this disease and the lack of therapeutic options, we anticipate rapid clinical translation of this vector.

**Fig. 7 AAV6.2FF human *proSFTPB* cDNA (AAV-hSPB) prevents lung damage.** **a** Schematic of the human *proSFTPB* cDNA transgene in the rAAV2 vector genome. **b** Study design to determine whether  $5 \times 10^{10}$  vg per mouse of AAV-hSPB improves lung structure and function in SP-B deficient mice. **c** Percentage change in body weight following AAV injection while still on dox (duration of 28 days). ( $n = 11$  biologically independent animals per group). **d** Percentage change in body weight after dox removal (duration of 4 days). All body weight measurements are presented as the mean with SD ( $n = 11$ , except  $1 \times$  PBS plus BLES post-dox where  $n = 4$ ;  $*P = 0.0221$ ). **e** Representative gross lung images at harvest 4 days following dox removal. Yellow arrows indicate regions of lung injury. The lung image from the mouse found dead was taken within 24 h of death. **f** Representative H&E and WGJ staining of paraffin embedded whole right and left lungs 4 days after dox removal (H&E scale bar, 200  $\mu\text{m}$ ; WGJ scale bar, 100  $\mu\text{m}$ ). **g** Representative epifluorescence images of SP-B (red) and DAPI (blue) of frozen lung sections from  $1 \times$  PBS plus BLES treated mice off dox ( $n = 2$ ), and  $5 \times 10^{10}$  vg AAV-hSPB plus BLES treated mice off dox ( $n = 2$ ) for 4 days (Scale bar, 50  $\mu\text{m}$ ). **h** Pressure-volume curve 4 days following dox removal corrected for body weight (in mL per g) for mice On Dox (black circles), mice treated with  $1 \times$  PBS plus BLES off dox (blue squares), and mice treated with AAV-hSPB plus BLES off dox (pink triangles). The source data for **h** has been provided as a Source Data file. **i** The  $\%V_{10}$  corrected for body weight. All lung function data are presented as the mean with SD (For **h** and **i**  $n = 11$  except  $1 \times$  PBS  $n = 4$ ;  $*P = 0.0169$ ,  $**P = 0.0021$ ). (All  $P$  values in Fig. 7 = 2-tailed ordinary one-way ANOVA with Tukey's multiple comparisons post hoc test, ns = not significant) (BLES = Bovine Lipid Extract Surfactant).



**Fig. 8 AAV-hSPB improves survival in neonatal and adult mice.** **a** Kaplan-Meier survival curves following IT delivery of  $5 \times 10^{10}$  vg per mouse (dashed pink line;  $**P = 0.0019$ ) of AAV-hSPB plus BLES with dox removal 28 days after AAV administration, and IT delivery of  $10^{11}$  vg per mouse (solid pink line;  $****P < 0.0001$ ) of AAV-hSPB plus BLES with dox removal 7 days (D7) after AAV administration. **b** Mean weight measurements with SD of all surviving mice either treated with  $1 \times$  PBS (blue squares) or  $10^{11}$  vg of AAV-hSPB plus BLES (pink triangles) from **a**. **c** Study design to assess survival following IT delivery of  $10^{11}$  vg per mouse AAV-hSPB plus BLES gene therapy in postnatal (P8) to P10 SP-B deficient mice. Mice were removed from dox upon weaning at P21. **d** Kaplan-Meier survival curves following IT delivery of  $10^{11}$  vg per mouse of AAV-hSPB plus BLES (pink line;  $***P = 0.0004$ ) at P8 to P10 with dox removal at P21. Control mice were untreated and maintained on dox (black line;  $**P = 0.0021$ ) or underwent sham surgeries with IT injections of  $1 \times$  PBS plus BLES (blue line). **e** Mean weight measurements with SD of all surviving mice untreated and maintained on dox (black circles), treated with  $1 \times$  PBS plus BLES with dox removed (blue squares), or treated with  $10^{11}$  vg of AAV-hSPB plus BLES with dox removed (pink triangles) from **d**. (Survival curve  $P$  values of **a** = 2-tailed Log-rank, Mantel-Cox test; Survival curve  $P$  values of **d** = 2-tailed Log-rank, Mantel-Cox test with Bonferroni correction, significance is set at  $P = 0.0167$ , ns = not significant) (BLES = Bovine Lipid Extract Surfactant).

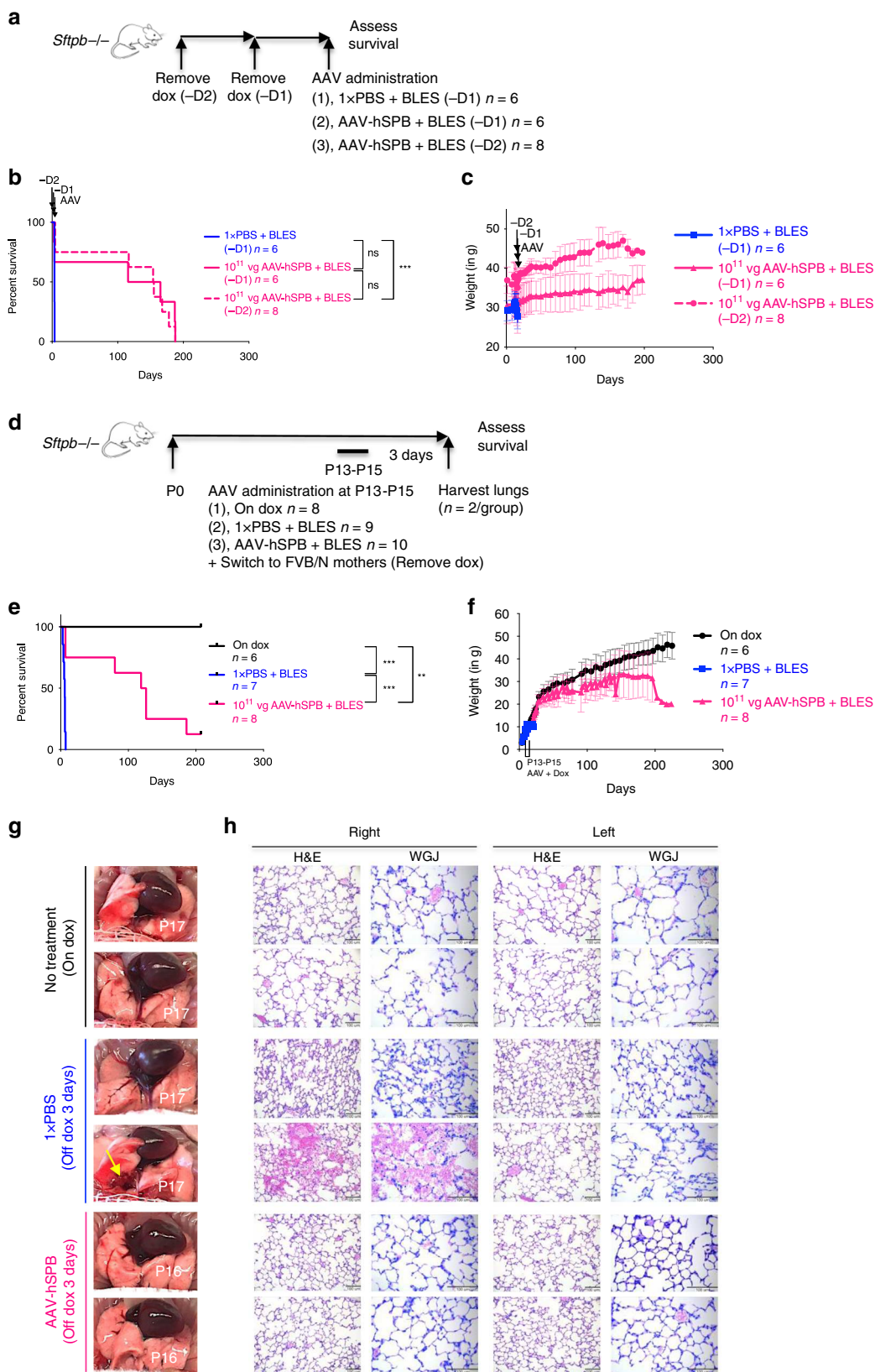
## Methods

**Institutional approval for animal and human studies.** All mouse experiments were performed in compliance with the guidelines set forth by the Canadian Council on Animal Care. All protocols and procedures involving animals were approved by the Animal Care and Veterinary Service Committee (ACVS) at the University of Ottawa, and the Animal Care Committee at the University of

Guelph. All human fetal tissues in this study were obtained after receiving institutional review board approval from The Ottawa Hospital (Research Protocol 20170603-01H), and after informed parental consent from donors.

**Recombinant AAV plasmid generation.** AAV genome plasmids contain the composite CASI promoter consisting of the human cytomegalovirus immediate





early gene enhancer region, the chicken beta actin promoter, and the human ubiquitin C promoter, as well as the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) and simian virus 40 polyadenylation sequence downstream of the transgene with flanking adeno-associated virus 2 (AAV2) inverted terminal repeats (ITRs)<sup>48</sup>. The murine *proSftpb* sequence (NM\_147779.2) was synthesized (GenScript) after codon optimization for murine expression using the GenScript codon optimization tool, and contains a myc epitope tag at the

C-terminus. Human *proSFTPB* cDNA (NM\_000542) was synthesized without codon optimization (GenScript) or an epitope tag.

**AAV6.2FF vector production.** All AAV6.2FF vectors were generated by plasmid transfection of adherent HEK293 cells using polyethylenimine at the University of Guelph<sup>49</sup>, except for the AAV6.2FF-murine SP-B cDNA (AAV-mSPB) vector used

**Fig. 9 AAV-hSPB rescues respiratory distress in an inherited model of SP-B deficiency.** **a** Study design to assess survival following AAV-hSPB plus BLES gene therapy in inherited SP-B deficient adult mice. Control mice underwent sham surgeries with IT injections of 1× PBS plus BLES 1 day (-D1) after dox removal (blue line). Treatment mice had dox removed 2 days (-D2; dashed pink line) and 1 day (-D1; solid pink line) before IT administration of  $10^{11}$  vg per mouse of AAV-hSPB plus BLES. **b** Kaplan-Meier survival curves from IT injections of a  $10^{11}$  vg dose of AAV-hSPB plus BLES 1 (-D1) and 2 (-D2; \*\*\* $P$  = 0.0001) days after dox removal. **c** Mean weight measurements with SD of all surviving mice from **b**. **d** Study design to assess survival following IT delivery of  $10^{11}$  vg per mouse AAV-hSPB plus BLES into P13 to P15 neonatal SP-B deficient mice. The pups were removed from dox immediately following AAV administration by placing them with wild type FVB/N foster mothers that had recently given birth to similar aged pups. Two mice were euthanized from each group 3 days after surgery for histological lung examination. **e** Kaplan-Meier survival curves following IT delivery of  $10^{11}$  vg per mouse of AAV-hSPB plus BLES into P13 to P15 neonatal mice with dox removal immediately following surgery (\*\*\* $P$  = 0.0004 compared to 1× PBS). Neonatal pups were untreated and maintained on dox (black line), underwent sham surgeries with IT injections of 1× PBS (blue line; \*\*\* $P$  = 0.0004 compared to on dox), or were IT administered  $10^{11}$  vg of AAV-hSPB plus BLES (pink line; \*\* $P$  = 0.0021 compared to on dox). The survival curve data do not include the mice removed for lung histology. **f** Mean weight measurements with SD of all surviving mice from **e**. **g** Gross lung images from all 3 groups ( $n$  = 2 per group) at harvest 3 days following AAV administration plus dox removal. Yellow arrows indicate regions of lung injury. **h** Representative H&E and WGJ staining of paraffin embedded whole right and left lungs 3 days after AAV administration and dox removal from all 3 groups in **d** (H&E and WGJ scale bar, 100  $\mu$ m). (All survival curve  $P$  values in Fig. 9 = 2-tailed Log-rank, Mantel-Cox test with Bonferroni correction, significance is set at  $P$  = 0.0167, ns = not significant).

in the intubation (Supplementary Fig. 8c) and high dose (Fig. 6b) survival experiments which was manufactured at the Senator Paul D Wellstone Muscular Dystrophy Cooperative Research Center vector core (Seattle, WA), and the AAV8-luciferase vector which was generated by the University of Pennsylvania vector core (Supplementary Fig. 7) (Philadelphia, PA). Vector genomes were quantified by TaqMan qPCR assay by targeting the inverted terminal repeat sequence of the AAV2 genome using the primers Forward: 5'-GGA ACC CCT AGT GAT GGA GTT-3'; Reverse: 5' CGG CCT CAG TGA GCG A-3'; and Probe: 5'-FAM-CAC TCC CTC TCT GCG CGC TCG-BHQ - 3'<sup>50</sup>.

**In vitro cell culture and transfections.** Human embryonic kidney 293 cells (HEK293; ATCC® CRL-1573) were maintained in high glucose DMEM (HyClone; SH30022.01) with 10% cosmic calf serum (HyClone; SH30087), 2 mM L-glutamine (HyClone; SH3003401), and 1% pen-strep (HyClone; SV30010). Murine lung epithelial 12 cells (MLE12; ATCC® CRL-2110) were maintained in DMEM/F12 1:1 media including HEPES (HyClone; SH30261) with the addition of 2% fetal bovine serum (FBS; HyClone; SH30088), insulin (0.005 mg per mL)-transferrin (0.01 mg per mL)-selenium 30 nM (Gibco; 41400045), 10 nM hydrocortisone (Sigma; H6909), 10 nM Beta-estradiol (Sigma; E2758), 2 mM L-glutamine (HyClone; SH3003401), and 1% pen-strep (HyClone; SV30010). At 75% confluency, HEK293 cells were transfected with linear polyethylenimine MW 25000, while lipofectamine was used to transfect MLE12 cells. Cell lysates were generated using a RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 10 mM EDTA, 1% sodium deoxycholate) containing  $\text{Na}_3\text{VO}_4$  (1 mmol per L), NaF (50 mM) and protease inhibitors (Sigma) 48 h post transfection, and prepared for immunoblotting using a 4 × SDS-PAGE reducing buffer. AAV-mSPB was incubated with HEK293 cells at a multiplicity of infection (MOI) of 20,000<sup>24</sup> and cells were harvested 72 h later and subsequently prepared for immunoblotting as described above.

**Immunoblotting.** Cell lysates were separated on a 10% tris-glycine polyacrylamide gel and transferred to a PDVF membrane at 100 V for 1 h. The membrane was blocked in 5% bovine serum albumin with PBS plus 1% Tween-20 (PBS-T) for 1 h while rotating, and washed 3 times with PBS-T. The membranes were incubated overnight at 4 °C with a c-Myc primary antibody at a 1 in 2000 dilution (Cell Signaling; D84C12), and washed 3 times with PBS-T before the addition of an anti-rabbit HRP polyclonal secondary antibody at a 1 in 4000 dilution (Invitrogen IgG (H + L); LS656120). Membranes were washed twice with PBS-T, once with PBS and imaged using HRP substrate (Luminata; WBLUF0100) in a BioRad Chemidoc. An uncropped version of the Western blot in Fig. 3b is supplied in the Source Data files.

**Animal model for AAV6.2FF targeting analysis.** The following procedure was carried out at the University of Guelph. Five 8-week old male Rosa26-Flox/LacZ (B6.129S4-Gt(ROSA)26Sor<sup>tm1Sor</sup>); strain code 003474) mice purchased from Jackson Laboratory (Bar Harbor, ME) were housed at the University of Guelph in a specific pathogen-free isolation facility. The animals were maintained on a 12 h on and 12 h off light cycle, room temperature (RT) was 23 to 24 °C, and humidity ranged from 30% to 60%. Mice were housed in groups of 4 and food (Teklad Global 14% Protein Rodent Maintenance Diet) and water (tap) were provided ad libitum. Mice were acclimated to the environment for 7 days prior to study initiation.

Using a modified intranasal delivery method<sup>37</sup> or via tail vein injection, groups of 3 mice were administered  $10^{11}$  vg of an AAV6.2FF vector expressing Cre recombinase containing a nuclear localization signal (nls) and fused to GFP (AAV6.2FF-CMV-emGFP-nlsCre) under the control of the CMV promoter. Control mice received 1 × PBS. At 3 weeks post-AAV administration, mice were humanely euthanized by isoflurane and all major organs harvested. Briefly, mice

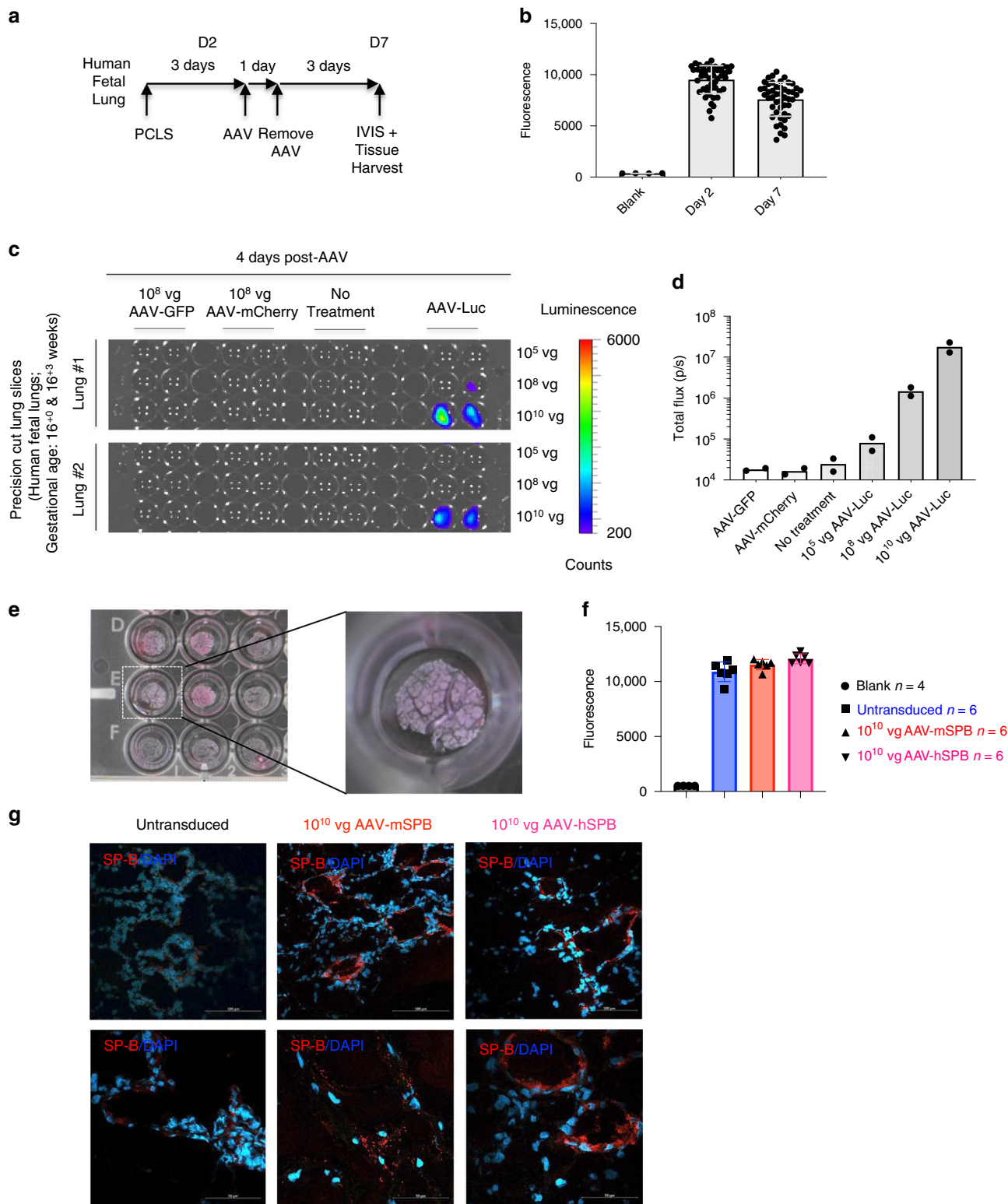
were exsanguinated via heart puncture and perfused with 1 × PBS via the right ventricle of the heart to remove excess blood. Next, a 27-gauge needle was inserted into the cartilage rings of the trachea and 5 mL of fixative (0.5% glutaraldehyde, 2 mM  $\text{MgCl}_2$ , 0.02% NP40, 0.01% deoxycholate in 1 × PBS) was administered to fully inflate the lungs while still inside the rib cage to avoid over-inflation. The trachea and lungs were removed en bloc and fixed for 2 h at RT. The remaining organs including the heart, spleen, liver, kidneys, pancreas, and brain were fixed for 16 h at RT while shaking. After fixation, tissues were washed 6 times for 5 min in wash solution (2 mM  $\text{MgCl}_2$  in 1 × PBS), before being stained in the dark with X-gal staining solution (5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM  $\text{MgCl}_2$ , 0.01% sodium deoxycholate, 0.02% Nonidet P-40 in 1 × PBS, pH 7.0 supplemented with 1 mg per mL X-gal [5-Bromo-4-chloro-3-indolyl-beta-D-galactopyranoside; 40 mg/mL stock in dimethylformamide] added immediately before use) for 2 h at 37 °C in the case of the lungs, and 8 h at 37 °C for all other tissues. Tissues were rinsed 6 times in 2 mM  $\text{MgCl}_2$  in 1 × PBS until the solution no longer turned yellow, placed in a cryomold, embedded in optimal cutting temperature compound (OCT) and stored at -80 °C. Lung and tracheal tissue from mice receiving the vector intranasally was paraffin embedded, sectioned (4  $\mu$ m) and counterstained with nuclear fast red. All other tissues were cryosectioned and re-stained with X-gal staining solution for 4 h at 37 °C in the dark and counterstained with nuclear fast red. Tissue sections were imaged using an Olympus BX45 light microscope.

**Animal model for lung analysis and survival studies.** The following procedures were carried out at The Ottawa Hospital Research Institute and The University of Ottawa. We utilized an inducible model of SP-B deficiency under the control of a doxycycline (dox) dependent promoter<sup>16</sup>. Mice were maintained on dox-supplemented feed (0.625 g per kg doxycycline hyclate; Teklad) unless otherwise noted. The animals were maintained on a 12 h on and 12 h off light cycle, RT was 21 °C, and humidity was 45%.

Mice were genotyped and maintained according to previously published protocols<sup>5</sup>. Briefly, 4 regions from the transgenic mouse model were genotyped. DNA extraction for PCR was obtained from clipped ear tissue using the AccuStart II Mouse Genotyping Kit (Quanta Biosciences) according to the manufacturer's instructions. For the rat CCSP promoter-rtTA genotyping-Forward: 5'-ACT GCC CAT TGC CCA AAC AC-3'; Reverse: 5'-AAA ATC TTG CCA GCT TTC CCC-3'; PCR Cycling Protocol: (1) 94 °C 5 min (2) 94 °C 30 s (3) 64 °C 30 s (4) 72 °C 40 s (5) Repeat 2-4 39 times (6) 72 °C 5 min (7) 4 °C. For the tetO7 minimal CMV promoter (murine *Sftpb* cDNA) genotyping-Forward: 5'-TGC TGC CAG GAG CCC TCT TG-3'; Reverse: 5'-AAG GCA CGG GGG AGG GGC AAA-3'; PCR Cycling Protocol: Same as the rat CCSP promoter-rtTA genotyping. For the phosphoglycerate kinase (Neomycin cassette) genotyping-Forward: 5'-TGA CCG CTT CCT CGT GCT TTA C-3'; Reverse: 5'-CCC CCC AGA ATA GAA TGA CAC CTA C-3'; PCR Cycling Protocol: (1) 94 °C 5 min (2) 94 °C 30 s (3) 64 °C 30 s (4) 72 °C 15 s (5) Repeat 2-4 39 times (6) 72 °C 5 min (7) 4 °C. For the murine SP-B exon 4 genotyping-Forward: 5'-CCA GGC TAA TCC TCC CTT CT-3'; Reverse: 5'-CCC ACT TAG GCA CAT GCA C-3'; PCR Cycling Protocol: (1) 94 °C 5 min (2) 94 °C 30 s (3) 62 °C 30 s (4) 72 °C 15 s (5) Repeat 2-4 39 times (6) 72 °C 5 min (7) 4 °C.

All mice were caged in either open or individually ventilated cage systems. Up to 4 animals were housed per cage and grouped with littermates and by sex. Mice were assigned a numerical ID. The treatment of mice (either control or AAV-mSPB and -hSPB) was assigned based on alternating numerical IDs. For example, in a cage of mice with the IDs L777, L778, L779, and L780: L777 and L779 would be administered AAV-mSPB while L778 and L780 would receive 1 × PBS (Fig. 6d). The order of surgeries were performed based on their numerical order: 1st: L777 (AAV-mSPB); 2nd: L778 (1 × PBS); 3rd: L779 (AAV-mSPB); 4th: L780 (1 × PBS).





To study the effects of AAV-SPB in the animal model, mice were typically maintained on dox for 27 to 30 days (4 weeks) following AAV injection to allow sufficient time for transgene expression<sup>26</sup>. We also assessed the effects following dox removal 3 and 7 days after AAV administration, upon weaning (6 to 13 days after dox), and dox removal 1 or 2 days before AAV administration. Mice were regularly weighed before and after AAV delivery and dox removal. Treatment or negative control groups were taken off dox and placed on a regular chow diet (Teklad 2018, 18% protein rodent diet; Envigo) for 3 to 4 days before euthanizing and harvesting the lungs for structure and function analysis (Figs. 4a, 7b).

For neonatal survival studies, transgenic SP-B deficient mice at postnatal (P) ages P8 to P10 were either left on dox, treated with 1 × PBS + BLES, or treated with 10<sup>11</sup> vg per mouse AAV-hSPB + BLES by IT administration. At P21, the mice were weaned off dox (obtained through the milk of nursing mothers) onto a regular chow diet (except for the mice maintained on dox), and body weight and survival were recorded. The volume of fluid instilled into the lungs of P8 to P10 pups resulted in approximately 50% mortality due to the inability of their lungs to accommodate the injection volume (20 μL). Administration at P8 resulted in 7 out of 10 (70%) deaths, at P9 4 out of 11 (36%) deaths, and at P10 2 out of 5 (40%) deaths within 1 to 5 min of administration. Death occurred regardless of whether

**Fig. 10 AAV6.2FF transduces the human lung parenchyma.** **a** Study design to demonstrate the ability of AAV6.2FF to transduce human lung tissue samples generated from precision cut lung slices (PCLS). Metabolic viability (resazurin) assays were conducted on days 2 (D2) and 7 (D7) post tissue harvest. **b** Resazurin assay results as represented by mean fluorescence with SD demonstrating the metabolic activity of the human PCLS on days 2 and 7. **c** Bioluminescence detection using IVIS in PCLS from human fetuses ( $n = 2$ ;  $16^{+0}$  and  $16^{+3}$  weeks gestational age) either untreated or transduced with  $10^8$  vg per well of AAV-GFP,  $10^8$  vg per well of AAV-mCherry, or  $10^5$ ,  $10^8$ , or  $10^{10}$  vg per well of AAV-Luc. Images are 4 days following vector exposure, and the scale is represented as counts. **d** Quantification of AAV-Luc wells are represented by the mean total flux (in photons per sec) with SD ( $n = 2$  fetuses; with 12 replicates for the AAV-GFP, AAV-mCherry, and No Treatment groups; and 4 replicates for each of the AAV-Luc doses). **e** Images of human PCLS in a 24-well tissue culture plate 7 days after culture and 3 days after AAV-SPB transduction. An enlarged image of a PCLS in an individual well is indicated by the dashed white box. **f** Resazurin assay results as represented by mean fluorescence with SD demonstrating viability of the human PCLS 3 days after no treatment (blue), or transduction with  $10^{10}$  vg per well of AAV-mSPB (red) or AAV-hSPB (pink). ( $n = 4$  blank; and  $n = 6$  untransduced,  $n = 6$  AAV-mSPB, and  $n = 6$  AAV-hSPB independent tissue slices from a single set of human fetal lungs). The source data for **f** has been provided as a Source Data file. **g** Representative confocal images of SP-B (red) and DAPI (blue) of frozen PCLS sections 3 days after transduction with  $10^{10}$  vg per well of AAV-mSPB or AAV-hSPB (Top and bottom row image scale bars, 100  $\mu$ m and 50  $\mu$ m, respectively).

$1 \times$  PBS or AAV-hSPB was instilled. To reduce post-surgery death, we repeated this experiment using older neonatal pups (P13 to P15) which significantly reduced the mortality (1 out of 20 deaths or 5%). In this latter experiment, pups were removed off dox immediately following AAV administration by placing them with nursing FVB/N foster mothers (on a regular chow diet) who had recently given birth to similar age-matched litters. This required the setting up of breeding pairs to deliver litters at the same time from both transgenic SP-B deficient mice on dox (6 pairs), and FVB/N mice on a regular chow diet (6 pairs). The inducible transgenic SP-B mouse model was generated by injecting murine *Sftpb* cDNA under the control of a (teto)<sub>7</sub> promoter into fertilized FVB/N oocytes<sup>16</sup>.

For survival studies mice were monitored at least 2 times per day following dox removal. When mice began to display initial signs of respiratory distress and weight loss, monitoring took place every 2 to 3 h. A scoring and endpoint chart for signs of respiratory distress requiring euthanasia was provided by the ACVS at The University of Ottawa (Supplementary Data File 2). Once respiratory distress was identified, animals were euthanized within 1 h of observing signs and symptoms of M2, and were immediately euthanized upon displaying signs and symptoms of M3. The decision for identifying the respiratory signs and symptoms for euthanasia were made by the staff at ACVS or by members of our group blinded to the identity and treatment groups of the animals. Mice were euthanized by 100% carbon dioxide (CO<sub>2</sub>) inhalation.

Although it was reported that the inducible SP-B deficient mice suffer lethal respiratory distress at 7.5 days  $\pm$  3.5 days (4 to 11 days) off dox<sup>16</sup>, the majority of untreated mice suffered lethal respiratory distress before 4 days. We also observed a few cases ( $n = 3$ ) where untreated mice survived beyond their expected endpoints (Supplementary Fig. 6). To ensure that AAV-SPB improvements in survival or lung structure and function were not due to stochastic incidences of extended survival, negative control groups were almost always used in our studies.

**Intratracheal (IT) injections.** One hour before IT surgery, animals were injected subcutaneously (SC) with 0.1 mg per kg buprenorphine (Champion Alstoe). Just before surgeries, animals were individually anesthetized with isoflurane (Fresenius Kabi) and injected SC with 1 mL of sterile 0.9% sodium chloride saline solution (Baxter Corp). The fur covering the tracheal region was then shaved to the skin. A tracheotomy was performed under constant isoflurane anesthesia, with mice receiving a single IT injection of AAV vector. Different titers of AAV-mSPB ( $10^{10}$  vector genomes (vg);  $10^{11}$  vg; or  $5 \times 10^{11}$  vg) or AAV-hSPB ( $5 \times 10^{10}$  vg;  $10^{11}$  vg) were diluted to a total volume of 42 to 83  $\mu$ L with  $1 \times$  PBS or pulmonary surfactant (Bovine Lipid Extract Surfactant; BLES Biochemicals Inc). Injections were carried out using a 3/10 mL insulin syringe 29 gauge  $\times$  1/2" (Covidien). Openings were sutured and topical 2% transdermal bupivacaine HCl as monohydrate (Chiron) was applied on the surgical site post-surgery, and 4 to 6 h after surgery. Mice were allowed to recover in a 37 °C incubator for 1 h post-surgery.

**Intubation (Int) injections.** Intubations were performed based on a combination of previously described methodologies<sup>37,51</sup>. Briefly, mice were injected SC in the posterior dorsal region with 1 mL of saline to prevent dehydration. Mice were then anesthetized intraperitoneally (IP) with 100 mg Ketalean (Ketamine; Bimeda-MTC Animal Health Inc) per 10 mg Rompun (Xylazine; Bayer Inc) per kg of body weight. Mice older than 6 weeks of age and weighing 20 to 35 g were intubated with a 1 inch long, 22 gauge IV catheter (BD Insyte). To view the tracheal opening, approximately 60 cm of 0.5 mm optical cable (Edmund Optics) was used with one end inserted through the 22-gauge IV catheter and the other end inserted through a rubber stopper connected to a light source. The anesthetized animal was placed on a vertical support (Harvard Apparatus) and suspended by its upper incisors using surgical thread. The ventral side of the mouse was positioned facing away from the operator and the tongue was gently pulled out using forceps. The fiber optic cable plus IV catheter was inserted with the light source used to visualize the vocal cords. The cable plus catheter was advanced past the vocal cords into the trachea, the fiber optic cable was withdrawn, and a 1 mL syringe containing 50 to 60  $\mu$ L of AAV

vector was administered through the catheter. This was followed by three 100  $\mu$ L injections of air with the 1 mL syringe, in an attempt to improve AAV distribution throughout the distal lung. The mouse was left in this standing or upright position for 1 to 2 min to allow time for vector dispersion throughout the lungs. Eye lube (CLC Medica) was continuously applied to protect against corneal drying and injury. Mice were allowed to recover for 1 to 2 h in a 37 °C incubator post-surgery.

**In vivo imaging system and diffuse light imaging tomography.** Mice were administered with a single injection of  $10^{11}$  vg of AAV6.2FF-Luciferase (AAV-Luc) or AAV8-Luc. Beginning 1-week post injection, IVIS or DLIT images were acquired. D-luciferin, Sodium Salt (BioVision) was sterilely prepared 24 h before injection at a concentration of 15 mg per mL in  $1 \times$  PBS. All mice were SC injected with 150 mg per kg of D-Luciferin 15 to 20 min before IVIS imaging. Up to 4 mice were imaged by IVIS simultaneously, while DLIT was performed on a single mouse. For DLIT imaging, surface topography was generated, followed by 3D reconstruction of bioluminescence as per the manufacturer's instructions (PerkinElmer). All DLIT images were performed within 30 to 45 min of D-Luciferin injection.

All quantifications of regions of interest (ROI) were generated using the Living Image 3.2 software according to the manufacturer's instructions (Caliper LifeSciences). Briefly, each mouse was identified through the generation of subject ROIs automatically by the Living Image software. Subsequently, bioluminescence signal intensity was manually identified in each mouse by an operator. ROIs larger than the visual boundaries of the bioluminescence signal were defined to ensure that the entire area of signal diffusion was included in each measurement. For each animal, the total flux (photons per second; p per s), average radiance (photons per second per square centimeter per solid angle of 1 steradian; p per s per cm<sup>2</sup> per sr), the standard deviation (SD) of radiance, as well as the minimum and maximum radiance were obtained from the nasal and tracheal, thorax (lung), and abdominal regions. All images are presented as count measurements which are uncalibrated measurements of photons in a pixel. Counts allow normalization between images acquired at different times. The total flux (p per s) or radiance (p per s per cm<sup>2</sup> per sr) was used to quantitate bioluminescence signal from IVIS figures in this study, which are calibrated measurements of photon emissions.

**Generation of precision cut lung slices (PCLS).** PCLS were processed based on previous descriptions<sup>21</sup>. Briefly, mice were euthanized with an IP injection of Euthanyl (Pentobarbital Sodium Injection; Bimeda-MTC Animal Health Inc), and the lungs were instilled with a low gelling point agarose solution (0.7% weight per volume [w per v] in  $1 \times$  PBS) under a constant pressure of 20 cmH<sub>2</sub>O. After cooling the mice in an ice bath, the left lung was isolated and embedded in a 5% low gelling point agarose solution.

Human fetal samples were obtained after elective abortions. The lungs were isolated by a pathologist and transported to the laboratory in ice cold DMEM GlutaMAX (Gibco; 10569-D10). Samples were rewarmed in 37 °C DMEM GlutaMAX for 30 min before low gelling point agarose solution (0.7% to 1.5% weight per volume) was slowly instilled into the lungs via the trachea or the main bronchus. Instillation was terminated when the lungs were inflated and the agarose solution overflowed from the airway. After cooling in an ice bath, the lungs were sampled and embedded in a 5% low gelling point agarose solution.

The tissue slicer (Kruldieck Tissue Slicer, TSE Systems) was set to a slicing thickness of 300  $\mu$ m at a speed of 30 slices per minute, and PCLS were collected in lung slice wash medium (LSWM: DMEM GlutaMAX (Gibco; 10569-D10), Anti-Anti (Gibco; 15240-062), gentamycin (ThermoFisher; 15710-064), 10  $\mu$ M 8-bromo-cAMP (Sigma; B5386), 100  $\mu$ M IBMX (Sigma; D4902), 100 nM Dexamethasone (Sigma; D4902)). The PCLS were maintained in a 48-well tissue culture plate (1 slice per well) in lung slice maintenance media (LSMM: LSWM with 10% FBS (Sigma; F1051), 2 mM L-glutamine (Gibco; 25030-081), 10 ng/mL recombinant KGF (FGF-7) (ThermoFisher; PHG0094)). The PCLS were acclimated for 3 days with fresh LSMM added daily. The viability of the PCLS were

assessed by performing resazurin (Sigma; 199303) fluorescence assays which was used to measure metabolic activity on the day of tissue collection (day 0), before or on the day of AAV transduction (day 2 or 3), and the day of IVIS imaging (day 7). On day 3, the adult mice or human fetal PCLS were transduced with  $10^8$  vg per well of AAV6.2FF-GFP, -mCherry, -Luc, -mSPB, or -hSPB. In the human fetal PCLS, 3 different AAV-Luc titers were used ( $10^5$ ,  $10^8$ , and  $10^{10}$  vg per well), and  $10^{10}$  vg per well of either AAV-mSPB or AAV-hSPB was used. Untransduced PCLS were used as a negative control. The viral vectors were applied to each well in a total volume of 150 to 300  $\mu$ L of neat DMEM GlutaMAX (depending on their size) for 24 h with constant rocking. Seven days after initial collection of PCLS and 4 days after vector exposure, the PCLS underwent IVIS imaging and were collected in 4% PFA, or embedded in VWR Premium Frozen Section Compound (VWR; 95057-838) and frozen on dry ice.

**Flow cytometry to determine cell-types transduced by AAV.** Eight days after IT administration of  $10^{11}$  vg of either AAV-Luc or AAV-GFP, mice were IP injected with 1 mUnits per g of heparin, euthanized by Euthanyl (Pentobarbital Sodium Injection; Bimeda-MTC Animal Health Inc), and had the thorax region surgically opened. The lung vasculature was perfused with 5 to 10 mL of cold  $1 \times$  PBS (no  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ ) plus undiluted heparin (25  $\mu$ L per 1 mL) through the right atrium to remove red blood cells (RBC). Following perfusion, an enzyme mix containing 30 U neutral protease (Worthington; LS02104), 2500 U Collagenase I (Worthington; LS004196), 10  $\mu$ g DNase I (Sigma; D5025-150KU) in DPBS with  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  was instilled into the lungs. The lungs were removed and placed in 5 mL of the enzyme mix for 1 h at 37 °C. Following removal of all traces of the trachea, bronchi, and heart tissue, the lung lobes were dissociated into small pieces and passed through a 40  $\mu$ m cell strainer. The filtered lung suspensions were collected and washed in FACS buffer ( $1 \times$  PBS, 5% FBS (Sigma; F1051), 1 mM EDTA). All centrifugation steps were carried out at 500 rcf for 5 min. Cold RBC lysis buffer was added to the cells and washed 3 times before the total cell number was counted for each sample. A total of  $5 \times 10^5$  cells in 200  $\mu$ L were placed into individual wells of a 96-well plate for staining. A pooled lung sample from the remaining AAV-Luc and AAV-GFP suspensions was used as a control at a concentration of  $10^5$  cells in 200  $\mu$ L. All samples were blocked with Heavy chain (Fc) block at RT in the dark. The samples were resuspended in 100  $\mu$ L of FACS buffer with the following antibody dilutions: 1 in 100 CD31-BV421 (Biolegend; Cat#102424); 1 in 100 CD45-AF647 (Biolegend; Cat#103124); and 1 in 100 EpCam-PE-CY7 (Biolegend; Cat#118216), and incubated for 30 min. The cells were washed 3 times with FACS buffer, and then resuspended in FACS buffer with 1  $\mu$ L of 7-AAD (Biolegend; Cat#420403) for 5 to 10 min before flow cytometry. No stain, single stain, and fluorescence minus one (FMO) control stains were all carried out. The samples were analyzed on the BeckmanCoulter MoFlo XDP flow cytometry machine (BeckmanCoulter), sample compensation was performed with the BD FACSDiva software version 8.0, and data was analyzed using FlowJo version 10 (Becton, Dickinson and Company).

**Lung function analysis.** Lung function including a measure of pulmonary surfactant function was calculated following the generation of individual pressure-volume (PV) curves for each mouse<sup>52,53</sup>. Mice were euthanized with an IP injection of Euthanyl (Pentobarbital Sodium Injection; Bimeda-MTC Animal Health Inc). Within 10 to 15 min of Euthanyl injection, pressure-volume curves were obtained using a small animal mechanical ventilator (flexiVent, Scireq). Briefly, an 18-gauge cannula attached to the flexiVent was inflated to the trachea of euthanized animals in a supine position. The lungs were secured with regular increasing intervals of pressure to a maximum of 30 cmH<sub>2</sub>O. Lungs were subsequently deflated with regular decreasing intervals of pressure to obtain pressure-volume curves. All data were obtained using the flexiWare version 7.0 (Scireq) software. Pressure-volume curves were normalized to the body weight of each animal.

$\%V_{10}$ , total lung capacity (TLC), residual volume (RV), and lung compliance were extracted from the pressure-volume curves<sup>31</sup>. Briefly,  $\%V_{10}$  quantifies the shape of the deflation limb and decreases with surfactant inhibition and is calculated using Eq. (1). The TLC is the volume at the defined maximal pressure of 30 cmH<sub>2</sub>O. The RV is the residual air trapped in alveoli at maximal expiration at the end of the deflation curve. The compliance is the slope at any linear region of the deflation limb and was calculated by the ratio in the change in volume between 10 cmH<sub>2</sub>O and 7 cmH<sub>2</sub>O ( $\Delta V$ ) over the change in pressure ( $\Delta P$ ).

**Quantitative PCR expression analysis of lung tissue.** RNA was isolated from snap frozen right lung tissue using QIAamp® Viral RNA Mini Kit (Qiagen). A total of 1  $\mu$ g of DNase treated RNA was reverse transcribed into complementary DNA (cDNA) using reverse transcriptase (Invitrogen). Real time qPCR (SYBR Green) was performed using the following primers for proSP-B-Myc (Murine proSP-B Forward: 5'-CCA AGA TCT CAG GAC GCC-3'; Myc Reverse: 5'-TCA CAG ATC CTC TTC TGA GAT G-3') and for Gapdh (Gapdh Forward: 5'-GTT GTC TCC TGC GAC TTC A-3'; Gapdh Reverse: 5'-GGT GGT CCA GGG TTT CTT A-3'). All qPCR reactions were carried out on a LightCycler 480 II (Roche Diagnostics) with Ct values automatically generated using the StepOnePlus software 2.3.

**Lung histology.** Macroscopic lung images were obtained with an iPhone 6S camera (Apple), or a Canon PowerShot SX620 HS point-and-shoot camera. For

microscopic lung images, lungs were perfused with 4% paraformaldehyde (4% PFA; Sigma-Aldrich) and fixed for 2 days. On day 3, the 4% PFA was removed and replaced with 70% ethanol. The left lungs were embedded in paraffin and cut coronally to obtain 4  $\mu$ m longitudinal sections of the lung. Serial sections were stained with hematoxylin and eosin (H&E) or Wright-Giemsa Jenner (WGJ) by the University of Ottawa histology core (Ottawa, ON). Scanned images were obtained at  $\times 20$  with the Aperio CS2 digital brightfield scanner (Leica) using Aperio's ImageScope version 12.3 software, and at  $\times 20$  and  $\times 40$  with the Leica DM4000 upright brightfield microscope.

**Lung immunofluorescence (IF).** All IF images were obtained from lung sections perfused with a 1 to 1 ratio of  $1 \times$  PBS and VWR Premium Frozen Section Compound (VWR; 95057-838), frozen on dry ice and stored at  $-80$  °C for at least 24 h. All sections were cut into 6 to 8  $\mu$ m sections using the Leica CM1860 cryostat. Sections were air-dried for 3 h at RT and stored at  $-20$  °C for at least 24 h. Prior to fixing and staining, frozen sections were thawed at 37 °C for 2 to 3 h. Briefly, sections were fixed in  $-20$  °C acetone for 15 min and washed with 0.1% Tween-20 in  $1 \times$  PBS 2 times (5 min per wash) and  $1 \times$  PBS once (5 min). Antigen retrieval using 10 mM sodium citrate pH 6.0 solution with 0.05% Tween-20 heated to boiling in a microwave was carried out 3 times (10 min per incubation). The sections were permeabilized in 0.1% Triton X-100 in  $1 \times$  PBS for 10 min at RT and blocked in 10% FBS in  $1 \times$  PBS for 1 h at RT. Sections were stained with a dilution of 1 in 250 proSP-C (rabbit anti-proSP-C; Millipore Sigma; AB3786) or 1 in 500 SP-B (rabbit anti-SP-B; Seven Hills Bioreagents; WRAB-48604) primary antibodies for 2 to 3 days at 4 °C in a light-resistant slide box. Human fetal PCLS were stained with a dilution of 1 in 50 firefly luciferase (goat anti-luciferase; Novus; NB100-1677), 1 in 200 mCherry (mouse anti-mCherry; Novus; NBP-1-96752), 1 in 50 EpCAM (rabbit anti-EpCAM; abcam; ab71916), 1 in 250 proSP-C, or 1 in 500 SP-B primary antibodies. Secondary antibody staining was with a dilution of 1 in 500 of goat  $\alpha$ -rabbit AF-568 (Life technologies; A11011), 1 in 500 of donkey  $\alpha$ -rabbit AF-488 (Invitrogen; A21206), 1 in 500 of goat  $\alpha$ -mouse AF-568 (Invitrogen; A21124), or 1 in 500 of donkey  $\alpha$ -goat AF-568 (Invitrogen; A11057) for 1 h at RT. Final washing steps were with  $1 \times$  PBS 3 times (5 min per wash) and coverslips were mounted on sections with Fluoroshield with DAPI (Sigma; F6057).  $\times 20$  images were obtained with an epi-fluorescence microscope (Zeiss Axio Imager.M2) using AxioVision 40  $\times$  64 version 4.9.1.0. Confocal images of human PCLS transduced with  $10^{10}$  vg per well of AAV-mSPB or AAV-hSPB were obtained by the Zeiss LSM800 AvioObserver Z1 microscope using Zen 2 (Blue edition).

**Transmission electron microscopy (TEM) of AT2 cells.** Lung tissue for imaging by TEM was processed based on previous descriptions<sup>54</sup>. Briefly, the right bronchus was tied off and a fixative solution containing 1.5% glutaraldehyde (Sigma-Aldrich) and 1.5% PFA in 0.15 M HEPES buffer (Sigma-Aldrich; 83264) was instilled into the left lung. The trachea was tightly tied to ensure the fixative remained in the lungs, before the left lung was excised and placed into 15 mL of the fixative solution and stored at 4 °C. Lung samples were subsequently processed as rapidly as possible. The lungs were incubated with 1% osmium tetroxide, stained en bloc with half-aqueous uranyl acetate, dehydrated in an ascending acetone series, before embedding in epoxy resin. From the embedded samples, ultrathin sections (60 to 80 nm thickness) were cut, mounted on formvar-coated copper support grids, and post-stained with lead citrate and uranyl acetate. The sections were analyzed using a Morgani TEM (Field Electron and Ion Company, FEI).

**Pulmonary surfactant as a vehicle for AAV vector delivery.** Clinical grade exogenous pulmonary surfactant Bovine Lipid Extract Surfactant (BLES; BLES Biochemicals Inc) enriched for hydrophobic phospholipids (27 mg phospholipids/mL of suspension) and the SP-B and SP-C surfactant proteins was used as a vehicle for AAV vectors. Clinically, BLES is used to treat neonatal respiratory distress syndrome at a recommended concentration of 135 mg phospholipids per kg (BLES Biochemicals). BLES was administered by IT or intubation injections at various concentrations including: 1.34 mg of total phospholipids per mouse (Supplementary Fig. 7), 0.54 mg of total phospholipids per mouse (Supplementary Fig. 8a), 1.21 mg phospholipids per mouse (Supplementary Fig. 8c), 0.82 mg of total phospholipids per mouse (Fig. 8a), 0.92 mg of total phospholipids per mouse (Figs. 8a, 9b), or 0.11 mg of total phospholipids per mouse (Figs. 8d, 9e). Differences in BLES concentration between the experiments were due to differences in viral titers from individual AAV preparations, combined with attempts to maintain instillation volumes at approximately 50  $\mu$ L per mouse. In the intubation plus BLES group (Supplementary Fig. 7), 1 mouse died within 10 min of administration of the vector plus BLES, presumably due to difficulties in respiration following fluid instillation into the lungs.

**Inflammatory cytokine measurements from mouse serum.** Serum was collected from the lateral saphenous vein of SP-B deficient mice in a survival study 7 days before ( $-7D$ ), and 1 ( $+1D$ ), 7 ( $+7D$ ), and 29 days ( $+29D$ ) after AAV or LPS administration in each of the following three groups:  $1 \times$  PBS ( $n = 7$ ) by IT administration maintained on dox;  $10^{11}$  vg AAV-mSPB ( $n = 8$ ) by IT administration with dox removal 7 days after AAV treatment; and 3 mg per kg of lipopolysaccharides (LPS) from *Escherichia coli* O111:B4 (Sigma; L2630) ( $n = 8$ ) by a



single bolus IP injection maintained on dox. The 3 mg per kg dosage of LPS is considered sublethal in adult mice<sup>40</sup>, however 5 out of 8 mice died post-LPS administration. Thirteen murine factors were analyzed by multiplex technology using the LEGENDplex (BioLegend) mouse inflammation panel according to the manufacturer's instructions. This assay is a bead-based immunoassay operating in a similar principle to the sandwich immunoassay. Briefly, the samples were analyzed in a V-bottom plate on a BD FACSCanto II flow cytometer using a high throughput autosampler, and data was collected using BDFACSDiva version 8.0. All statistics were performed using the Mixed Effects Model with the Geisser-Greenhouse correction followed by a post hoc Tukey's multiple comparison's test. Note that data was analyzed by fitting the Mixed Model rather than by repeated measures ANOVA as there were missing values due to the death of animals (5 LPS and 1 AAV-mSPB mice) during the course of the study.

**Statistical analyses and reproducibility.** Graphpad Prism 8 software was used to perform all statistical analyses. Graphically, means are always presented with standard deviation (SD) error bars. For 2 groups, statistical analysis was by the 2-tailed Student's *t*-test. For 3 groups, statistical analysis was by ordinary one-way ANOVA with Tukey's or Dunnett's multiple comparisons post hoc test. Kaplan Meier survival curves were analyzed with the Log-rank, Mantel-Cox test. When more than 2 variables were compared in survival studies, Bonferroni correction was applied. Cytokine measurement data were analyzed by fitting the Mixed Effects Model with Tukey's multiple comparison post hoc test. A *P* value less than 0.05 was considered significant, unless Bonferroni correction was applied. All instances of Bonferroni correction and significant *P* values are indicated in the manuscript text or figure legends.

For representative experiments such as micrographs, the following list provides the number of times an experiment was independently repeated for a given figure: Fig. 2b (*n* = 2); Fig. 3b (*n* = 3); Fig. 4e (*n* = 10 for all groups except AAV-Luc, *n* = 7); Fig. 5b (*n* = 2 per group); Fig. 5c (*n* = 2 per group); Fig. 7f (*n* = 2 per group); Fig. 7g (*n* = 2 per group); Fig. 9h (*n* = 2 per group); and Fig. 10g (*n* = 6 PCLS slices per group). For the supplementary figures: Supplementary Fig. 1c (*n* = 3); Supplementary Fig. 2c (different field of view from image in Fig. 2b); Supplementary Fig. 2d (*n* = 1); Supplementary Fig. 3i (*n* = 2); Supplementary Fig. 4 (*n* = 1); Supplementary Fig. 5d (*n* = 10 for all groups except AAV-Luc, *n* = 7); Supplementary Fig. 5f-h (*n* = 3 for all groups); Supplementary Fig. 5j (from the same experiment as Fig. 5c, *n* = 2 per group); and Supplementary Fig. 6e, f (*n* = 1).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The authors declare that the data supporting the findings of this study are available within the paper, the associated Supplementary Figures and Tables, and in the Source Data files. Source data are provided with this paper. All other information is freely available from the corresponding authors upon request.

Received: 31 July 2019; Accepted: 4 July 2020;

Published online: 06 August 2020

## References

- Whitsett, J. A., Wert, S. E. & Weaver, T. E. Alveolar surfactant homeostasis and the pathogenesis of pulmonary disease. *Annu. Rev. Med.* **61**, 105–119 (2010).
- Whitsett, J. A. & Weaver, T. E. Hydrophobic surfactant proteins in lung function and disease. *N. Engl. J. Med.* **347**, 2141–2148 (2002).
- Whitsett, J. A., Noguee, L. M., Weaver, T. E. & Horowitz, A. D. Human surfactant protein B: structure, function, regulation, and genetic disease. *Physiol. Rev.* **75**, 749–757 (1995).
- Sweet, S. C. Update on pediatric lung allocation in the United States. *Pediatr. Transplantation* **13**, 808–813 (2009).
- Kirkby, S. & Hayes, D. Jr. Pediatric lung transplantation: indications and outcomes. *J. Thorac. Dis.* **6**, 1024–1031 (2014).
- Kurland, G. et al. and the American Thoracic Society Committee on Childhood Interstitial Lung Disease (chILD) and the chILD Research Network. An official American Thoracic Society clinical practice guideline: classification, evaluation, and management of childhood interstitial lung disease in infancy. *Am. J. Respiratory Crit. Care Med.* **188**, 376–394 (2013).
- Thompson, M. W. Surfactant protein B deficiency: insights into surfactant function through clinical surfactant protein deficiency. *Am. J. Med. Sci.* **321**, 26–32 (2001).
- Lin, S. et al. Surfactant protein B (SP-B) *-/-* mice are rescued by restoration of SP-B expression in alveolar type II cells but not Clara cells. *J. Biol. Chem.* **274**, 19168–19174 (1999).
- Yei, S. et al. Adenoviral-mediated gene transfer of human surfactant protein B to respiratory epithelial cells. *Am. J. Respiratory Cell Mol. Biol.* **11**, 329–336 (1994).
- Korst, R. J., Bewig, B. & Crystal, R. G. In vitro and in vivo transfer and expression of human surfactant SP-A- and SP-B-associated protein cDNAs mediated by replication-deficient, recombinant adenoviral vectors. *Hum. Gene Ther.* **6**, 277–287 (1995).
- Sanders, N., Rudolph, C., Braeckmans, K., De Smedt, S. C. & Demeester, J. Extracellular barriers in respiratory gene therapy. *Adv. Drug Deliv. Rev.* **61**, 115–217 (2009).
- Geiger, J., Aneja, M. K. & Rudolph, C. Vectors for pulmonary gene therapy. *Int. J. Pharmaceutics* **390**, 84–88 (2010).
- Kormann, M. S. et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.* **29**, 154–157 (2011).
- Mahiny, A. J. et al. In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency. *Nat. Biotechnol.* **33**, 584–586 (2015).
- Barnett, R. C. et al. Featured Article: electroporation-mediated gene delivery of surfactant protein B (SP-B) restores expression and improves survival in mouse model of SP-B deficiency. *Exp. Biol. Med.* **242**, 1345–1354 (2017).
- Melton, K. R. et al. SP-B deficiency causes respiratory failure in adult mice. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **285**, L543–L549 (2003).
- Limberis, M. P., Vandenberghe, L. H., Zhang, L., Pickles, R. J. & Wilson, J. M. Transduction efficiencies of novel AAV vectors in mouse airway epithelium in vivo and human ciliated airway epithelium in vitro. *Mol. Ther.* **17**, 294–301 (2009).
- van Lieshout, L. P. et al. A novel triple-mutant AAV6 capsid induces rapid and potent transgene expression in the muscle and respiratory tract of mice. *Mol. Ther. Methods Clin. Dev.* **9**, 323–329 (2018).
- Payne, J. G. et al. Multilineage transduction of resident lung cells in vivo by AAV2/8 for  $\alpha$ 1-antitrypsin gene therapy. *Mol. Ther. Methods Clin. Dev.* **3**, 16042 (2016).
- Loeb, J. E., Cordier, W. S., Harris, M. E., Weitzman, M. D. & Hope, T. J. Enhanced expression of transgenes from adeno-associated virus vectors with the woodchuck hepatitis virus posttranscriptional regulatory element: implications for gene therapy. *Hum. Gene Ther.* **10**, 2295–2305 (1999).
- Rosales Gerpe, M. C. et al. Use of precision-cut lung slices as an ex vivo tool for evaluating viruses and viral vectors for gene and oncolytic therapy. *Mol. Ther. Methods Clin. Dev.* **10**, 245–256 (2018).
- Hasegawa, K. et al. Fraction of MHCII and EpCAM expression characterizes distal lung epithelial cells for alveolar type 2 cell isolation. *Respiratory Res.* **18**, 150 (2017).
- Clark, J. C. et al. Targeted disruption of the surfactant protein B gene disrupts surfactant homeostasis, causing respiratory failure in newborn mice. *Proc. Natl. Acad. Sci. USA* **92**, 7794–7798 (1995).
- Pillay, S. et al. An essential receptor for adeno-associated virus infection. *Nature* **530**, 108–112 (2016).
- Banfi, C. & Agostoni, P. Surfactant protein B: from biochemistry to its potential role as diagnostic and prognostic marker in heart failure. *Int. J. Cardiol.* **221**, 456–462 (2016).
- Flotte, T. R. et al. Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing  $\alpha$ 1-antitrypsin: interim results. *Hum. Gene Ther.* **22**, 1239–1247 (2011).
- Matute-Bello, G. et al. & Acute Lung Injury in Animals Study Group An official American Thoracic Society workshop report: features and measurements of experimental acute lung injury in animals. *Am. J. Respiratory Cell Mol. Biol.* **44**, 725–738 (2011).
- Johnson, J. W., Permutt, S., Sipple, J. H. & Salem, E. S. Effect of intra-alveolar fluid on pulmonary surface tension properties. *J. Appl. Physiol.* **19**, 769–777 (1964).
- Faridy, E. E. Effect of distension on release of surfactant in excised dogs' lungs. *Respiratory Physiol.* **27**, 99–114 (1976).
- Faridy, E. E., Permutt, S. & Riley, R. L. Effect of ventilation on surface forces in excised dogs' lungs. *J. Appl. Physiol.* **21**, 1453–1462 (1966).
- Limjunyawong, N., Fallica, J., Horton, M. R. & Mitzner, W. Measurement of the pressure-volume curve in mouse lungs. *J. Visualized Exp.* **95**, 52376 (2015).
- Askin, F. B. & Kuhn, C. The cellular origin of pulmonary surfactant. *Lab. Invest.* **25**, 260–268 (1971).
- Fehrenbach, H. Alveolar epithelial type II cell: defender of the alveolus revisited. *Respiratory Res.* **2**, 33–46 (2001).
- Aneja, M. K., Geiger, J. P., Himmel, A. & Rudolph, C. Targeted gene delivery to the lung. *Expert Opin. Drug Deliv.* **6**, 567–583 (2009).
- Colella, P., Ronzitti, G. & Mingozzi, F. Emerging issues in AAV-mediated in vivo gene therapy. *Mol. Ther. Methods Clin. Dev.* **8**, 87–104 (2017).
- Nogee, L. M. Genetic mechanisms of surfactant deficiency. *Biol. Neonate* **85**, 314–318 (2004).
- Santry, L. A. et al. AAV vector distribution in the mouse respiratory tract following four different methods of administration. *BMC Biotechnol.* **17**, 43 (2017).

38. Jobe, A. H., Ueda, T., Whitsett, J. A., Trapnell, B. C. & Ikegami, M. Surfactant enhances adenovirus-mediated gene expression in rabbit lungs. *Gene Ther.* **3**, 775–779 (1996).
39. Weiss, D. J. et al. Comparison of surfactant and perfluorochemical liquid enhanced adenovirus-mediated gene transfer in normal rat lung. *Mol. Ther.* **6**, 43–49 (2002).
40. Seemann, S., Zohles, F. & Lupp, A. Comprehensive comparison of three different animal models for systemic inflammation. *J. Biomed. Sci.* **24**, 60 (2017).
41. Altshul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
42. Burri, P. H. Fetal and postnatal development of the lung. *Annu. Rev. Physiol.* **46**, 617–628 (1984).
43. Burkholder, T., Foltz, C., Karlsson, E., Linton, C. G. & Smith, J. M. Health evaluation of experimental laboratory mice. *Curr. Protoc. Mouse Biol.* **2**, 145–165 (2012).
44. Foust, K. D. et al. Rescue of the spinal muscular atrophy phenotype in a mouse model by early postnatal delivery of SMN. *Nat. Biotechnol.* **28**, 271–274 (2010).
45. Samulski, R. J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).
46. Barkauskas, C. E. et al. Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
47. Palomar, L. M. et al. Long-term outcomes after infant lung transplantation for surfactant protein B deficiency related to other causes of respiratory failure. *J. Pediatrics* **149**, 548–553 (2006).
48. Balazs, A. B. et al. Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature* **481**, 81–84 (2011).
49. van Lieshout, L. P., Domm, J. M. & Wootton, S. K. AAV-mediated gene delivery to the lung. *Methods Mol. Biol.* **1950**, 361–372 (2019).
50. Aurnhammer, C. et al. Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum. Gene Ther. Methods* **23**, 18–28 (2012).
51. Das, S., MacDonald, K., Chang, H. Y. & Mitzner, W. A simple method of mouse lung intubation. *J. Visualized Exp.* **73**, e50318 (2013).
52. Avery, M. E. & Mead, J. Surface properties in relation to atelectasis and hyaline membrane disease. *Am. J. Dis. Child.* **97**, 517–523 (1959).
53. Mitzner, W., Johnson, J. W., Scott, R., London, W. T. & Palmer, A. E. Effect of betamethasone on pressure-volume relationship of fetal rhesus monkey lung. *J. Appl. Physiol. Respiratory Environ. Exerc. Physiol.* **47**, 377–382 (1979).
54. Kling, K. M., Lopez-Rodriguez, E., Pfarrer, C., Mühlfeld, C. & Brandenberger, C. Aging exacerbates acute lung injury-induced changes of the air-blood barrier, lung function, and inflammation in the mouse. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **312**, L1–L12 (2017).

## Acknowledgements

This work was supported in part by funding from a CIHR Foundation Scheme Grant to B.T. (FDN-143316), a CHIR Project grant to S.K.W. and B.T. (420518), grants from The Lung Association–Ontario (34998) and The Mason Research Fund (44610) to S.K.W., NIH grant support to J.A.W. (U01HL134745 and U01HL148856), a European Respiratory Society Fellowship to L.R. (LTRF 2018), and Deutsche Forschungsgemeinschaft funding to I.M. The authors thank all the ACVS veterinary and technical staff at the University of Ottawa, and in particular Dr Mark Liepmann for the development of the respiratory distress scoring chart; Dr Greg Cron, for training of IVIS imaging; all the staff at the University of Ottawa Histology Core, especially Zaida Ticas and Sharlene Faulkes; Jeff McClintock at the CHEO Electron Microscope lab; and

EORLA staff for the human fetal lung sample isolation, especially Janet Stinson and Korrine Hutt-Acres.

## Author contributions

M.H.K., S.K.W., and B.T. designed all the experiments and drafted the manuscript. M.H.K., L.P.v.L., L.X., J.M.D., A.V., L.R., C.M., M.H., I.M., Y.P., J.P.v.V., S.P.T., C.M., and C.C.-D. performed all the experiments. L.P.v.L., J.M.D., Y.P., and S.K.W. designed, generated, and purified the AAV vectors. M.H.K., L.X., A.V., L.R., M.H., I.M., C.M., and C.C.-D. performed all the in vivo experiments. L.X. maintained and genotyped all transgenic mice used in this study. A.V. performed all the IT surgeries. L.X. performed all the intubation surgeries. M.H.K. and L.X. designed and performed all the IVIS and DLIT studies. L.R. designed and generated the PCLS studies and results. C.M. processed and generated all the TEM images. M.H., and I.M. designed and generated the flow cytometry analyses. Y.P. performed the qPCR analysis. J.P.v.V. performed the cytokine measurements. S.P.T. performed the Cre recombinase studies. M.H.K., L.X., and C.C.-D. performed the flexiVent analyses and harvested the lungs for histology. M.H.K. and C.M. generated all the immunofluorescence images. M.H.K. and L.X. monitored the mice for respiratory distress in the survival studies, and regularly recorded their body weights. L.P.v.L., L.X., J.M.D., A.V., L.R., C.M., M.H., I.M., J.A.W., L.M.N., S.K.W., and B.T. provided critical comments and suggestions during the drafting of the manuscript. J.M.D. and A.V. contributed equally to this work.

## Competing interests

S.K.W., L.P.v.L., B.T., and M.H.K. have a patent pending entitled: ‘Adeno-Associated Virus Particle With Mutated Capsid And Methods Of Use Thereof.’ The remaining authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-17577-8>.

**Correspondence** and requests for materials should be addressed to S.K.W. or B.Téb.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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



**Open Access** This article 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. The images or other third party material in this article 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 <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

### **III. A Delphi study to establish a consensus definition and clinical reporting guidelines for Mesenchymal Stromal Cells**

This protocol is published in *BMJ Open* and submitted as a preprint on *MedRxiv* (<https://doi.org/10.1101/2021.06.18.21259151>). A letter to the Editor for this project is published in *Cytotherapy*.

#### **1. Background**

Despite more than two decades of research, MSC-based treatments are still struggling to cross the clinical translation gap, as many MSC-based clinical therapies have yielded suboptimal outcomes or failed to meet their primary endpoints. Two key issues that likely contribute to these failures are 1) the lack of clear definition for MSC (despite previous attempts to define MSC) and 2) poor quality of reporting in MSC clinical studies. To address these issues, we propose here a protocol of a modified Delphi study to establish a consensus definition for MSCs and clinical reporting guidelines for MSC.

#### **2. Results**

We have designed a detailed protocol to conduct a three-round international modified Delphi Survey combined to an integrated knowledge translation (iKT) approach. The project overview is presented in Figure 16.

The first step of this project consists in the recruitment of a core group, including representative stakeholders from different research fields including developmental biology, translational science, research methods, regulatory practices and scholarly journal editing, industry, as well as a patient partner. To follow the iKT approach, this core group will be involved in the process since inception and will be part of each step of the project. This

includes the development of a tailored end-of-project knowledge translation plan to support and ensure dissemination and implementation of the Delphi results.

For the second step, we will conduct a scoping review to describe how MSCs are defined and reported in preclinical and clinical studies (cf. Part IV Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review). This scoping review will help to identify and inform relevant items for the initial Delphi survey. In addition, corresponding authors from the selected original studies will be contacted to be a part of the researcher panel.

The modified Delphi survey will have two survey rounds online, and the final round will take place in person. Participants will include the core group (participating in all rounds), as well as researchers whose work was captured in our scoping review (will participate only in the first round). Each participant will be asked to rate their agreement on potential MSC definition characteristics and reporting items using a Likert scale. For each item, participants will be asked to provide a rationale for their choice. In addition, participants will be allowed to provide any additional relevant items. After each round, we will analyse data to determine which items have reached consensus for inclusion/exclusion. Data will be then summarized and a revised questionnaire for any new items, or items that did not reach consensus, will be developed and sent out for the next round. The first online round of the Delphi is scheduled to start in September 2021.

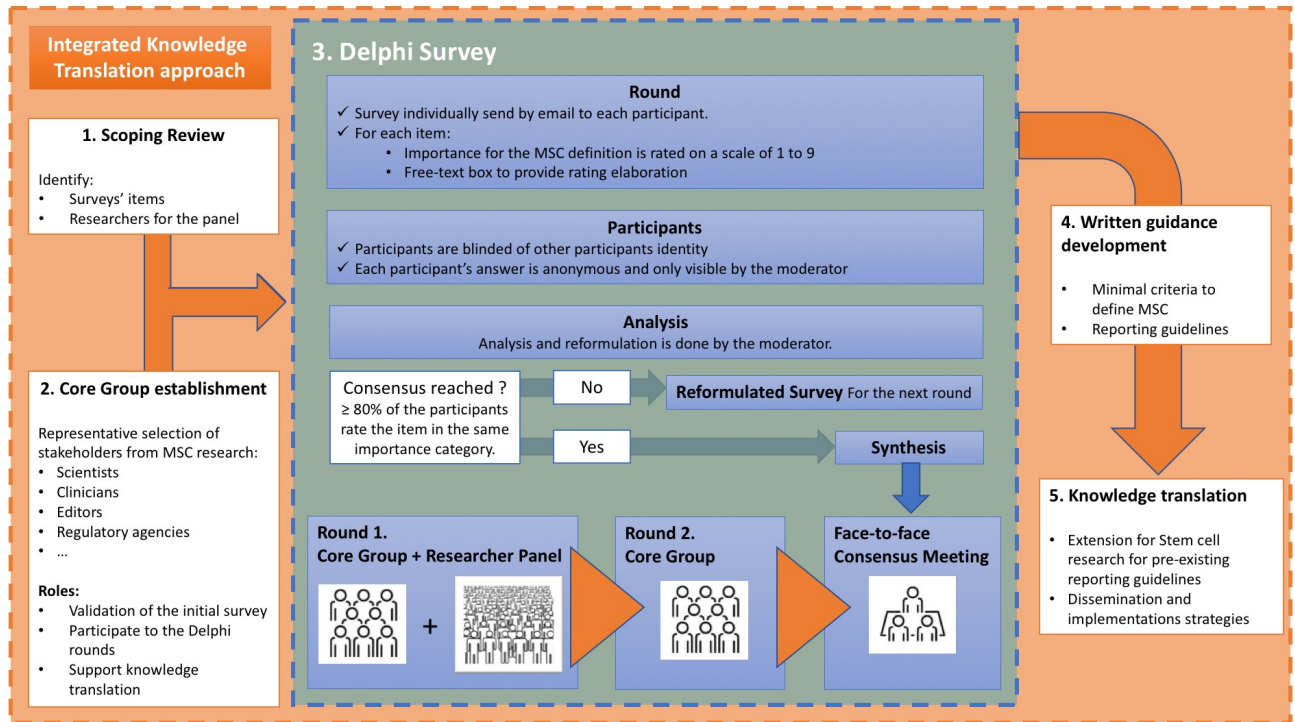


Figure 16. Delphi MSC project overview.

### 3. Conclusion

To address the pitfalls of the previous attempts to define MSC and further support dissemination and implementation, we propose here an original research protocol using the unique attributes of a consensus-building method called modified Delphi method in combination with an ‘integrated knowledge translation’ approach. By developing a consensus definition for MSC and reporting guidelines for clinical trials using MSC, we hope to better address clinical translation, reproducibility, and transparency in the field of MSC research.



4. Manuscript “Establishment of a consensus definition for Mesenchymal Stromal Cells (MSC) and reporting guidelines for clinical trials of MSC therapy: a modified Delphi study protocol” (published in BMJ Open)

# BMJ Open Establishment of a consensus definition for mesenchymal stromal cells (MSC) and reporting guidelines for clinical trials of MSC therapy: a modified Delphi study protocol

Laurent Renesme <sup>1</sup>, Kelly D Cobey <sup>2,3</sup>, Maxime Le,<sup>4</sup> Manoj M Lalu <sup>3,5</sup>, Bernard Thebaud <sup>1,6,7</sup>

**To cite:** Renesme L, Cobey KD, Le M, *et al.* Establishment of a consensus definition for mesenchymal stromal cells (MSC) and reporting guidelines for clinical trials of MSC therapy: a modified Delphi study protocol. *BMJ Open* 2021;**11**:e054740. doi:10.1136/bmjopen-2021-054740

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-054740>).

Received 21 June 2021

Accepted 16 September 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Bernard Thebaud;  
[bthebaud@toh.ca](mailto:bthebaud@toh.ca)

## ABSTRACT

**Introduction** Despite being more than two decades of research, mesenchymal stromal cell (MSC) treatments are still struggling to cross the translational gap. Two key issues that likely contribute to these failures are (1) the lack of clear definition for MSC and (2) poor quality of reporting in MSC clinical studies. To address these issues, we propose a modified Delphi study to establish a consensus definition for MSC and reporting guidelines for clinical trials of MSC therapy.

**Methods and analysis** We will conduct a three-round international modified Delphi survey. Findings from a recent scoping review examining how MSCs are defined and reported in preclinical and clinical studies were used to draft the initial survey for round 1 of our Delphi. Participants will include a ‘core group’ of individuals as well as researchers whose work was captured in our scoping review. The core group will include stakeholders from different research fields including developmental biology, translational science, research methods, regulatory practices, scholarly journal editing and industry. The first two survey rounds will be online, and the final round will take place in person. Each participant will be asked to rate their agreement on potential MSC definition characteristics and reporting items using a Likert scale. After each round, we will analyse the data to determine which items have reached consensus for inclusion/exclusion, and then develop a revised questionnaire for any new items, or items that did not reach consensus.

**Ethics and dissemination** This study received ethical approval from the Ottawa Health Research Network Research Ethics Board. To support the dissemination of our findings, we will use an evidence-based ‘integrated knowledge translation’ approach to engage knowledge users from the inception of the research. This will allow us to develop a tailored end-of-project knowledge translation plan to support and ensure dissemination and implementation of the Delphi results.

## INTRODUCTION

Since their original description<sup>1</sup> and their first use as a therapeutic agent in humans,<sup>2</sup> interest for mesenchymal stromal cells (MSC)

## Strengths and limitations of this study

- We proposed to address the current limitations in mesenchymal stromal cell experimental and clinical research with a rigorous and methodological consensus building method (Delphi method) that will allow for structured communication on controversial issues.
- To support dissemination and implementation of our results, we will engage stakeholders and end users from the inception of the project—such as patient partners—and will develop a tailored end-of-project knowledge translation plan (integrated knowledge translation approach) in order to overcome historical issues related to community uptake.
- To address the main limitations of a Delphi method (eg, lack of participation, no in-person interaction or information exchange), we use a modified Delphi survey with a core group of stakeholders and a face-to-face meeting.

from the scientific and patient community keeps growing exponentially: a PubMed search with the query MSCs between 1995 and 2021 found 73 876 results; more than half of the results were from the last 5 years. From a clinical research perspective, over 1300 MSC clinical trials have been registered on ClinicalTrials.gov.<sup>3</sup> Despite promising results of MSC in different preclinical diseases models, clinical trials using MSC in various medical conditions failed to deliver encouraging results.<sup>4 5</sup> Many potential explanations for this disparity between preclinical and clinical results exist, including MSC characteristics, cell manufacturing processes, administration protocols or study participants’ characteristics such as disease severity and associated comorbidities.<sup>6</sup> In addition, MSC characteristics (eg, definition, characterisation, immune

compatibility, cell viability and dose) raise two main issues of (1) lack of consensus definition for MSC and (2) broad variability in MSC characteristics reported. For example, a report from the Food and Drug Administration showed a significant heterogeneity in MSC products used in the clinical trials, with important differences in cell surface marker characterisation, product bioactivity assessment, as well as tissue sourcing and product manufacturing.<sup>7</sup> Regarding reporting of MSC characteristics, members of our group have performed systematic reviews of published MSC clinical trials and have found extremely poor (ie, incomplete) reporting of cell products used.<sup>8–10</sup> These observations have pushed some scientists to call to ‘clear up this stem cell mess’.<sup>4</sup>

The first important initiative to develop a consensus definition for MSC was provided by the International Society for Cell & Gene Therapy (ISCT), which in 2006 established the minimal criteria for human bone marrow-derived MSC and these were updated in 2019.<sup>11 12</sup> These criteria were determined through a group decision-making method (eg, informal consensus of a small number of experts), which is a consensus development method with several limitations.<sup>13</sup> However, implementation of this MSC definition has been inconsistent and several researchers have expressed scepticism and highlighted limitations (eg, cell surface markers phenotype, misinterpretation of differentiation assays, limitations in functionally defining stromal cells, etc).<sup>4 11 14</sup> For example, a scoping review conducted by our group to describe how MSCs are defined and characterised in both preclinical and clinical research showed that only 18% of the articles from our sample explicitly referred to the ISCT criteria, and only 7% of clinical studies selection reported the three minimal criteria from the 2006 definition. The uncertainty and lack of consensus on this crucial issue of how to define MSC has also allowed questionable for-profit business to sell unproven and unlicensed ‘stem cell’ therapies to patients and public.<sup>15 16</sup>

Therefore, there is an urgent need to develop a consensus definition of MSC and reporting guidelines to create standards for complete and transparent reporting of both cell characteristics and clinical trial/manufacturing details. A clear definition and reporting guidelines will be the cornerstones of a more robust, reproducible and transparent research, and are mandatory to better understand underlying factors that may contribute to efficacy (or a lack of efficacy) of MSC. Moreover, for patients and members of the public, having a clear definition of MSC will allow them to make informed decisions into which related treatments, clinical trials or products they elect to partake in.

To address limitations of previous attempts to define MSC and support the dissemination and implementation of MSC definition and reporting guidelines, our research protocol is based on a modified Delphi method combined with an integrated knowledge translation (iKT) approach.

The objectives of the current study are to: (1) develop, disseminate and implement an updated consensus

definition of minimal criteria to define MSC; and (2) develop reporting guidelines for the clinical trials of MSC therapy. The work is not hypothesis testing, and therefore we have no a priori predictions related to study outcomes.

## METHODS AND ANALYSIS

### Project overview

We will conduct a three-round Delphi survey. The Delphi method is a process that uses several rounds of surveying in order to reach consensus on a topic.<sup>13 17</sup> Between rounds, responses to survey items are aggregated—items that reach a priori threshold for inclusion/exclusion are omitted, and any remaining or new items are shared with participants in the subsequent round to reconsider and vote on again. The Delphi method addresses group decision-making method’s limitations such as (1) group influences on individual performance, (2) adequately accounting for divergent opinions, (3) limiting the number of participants, (4) meeting organisation and costs, (5) time limits for very complicated decisions, (6) peer pressure and influence in the decision-making process, and (7) logistical issues.<sup>13</sup> The Delphi method has been successful in the past in solving contentious issues.<sup>18–20</sup>

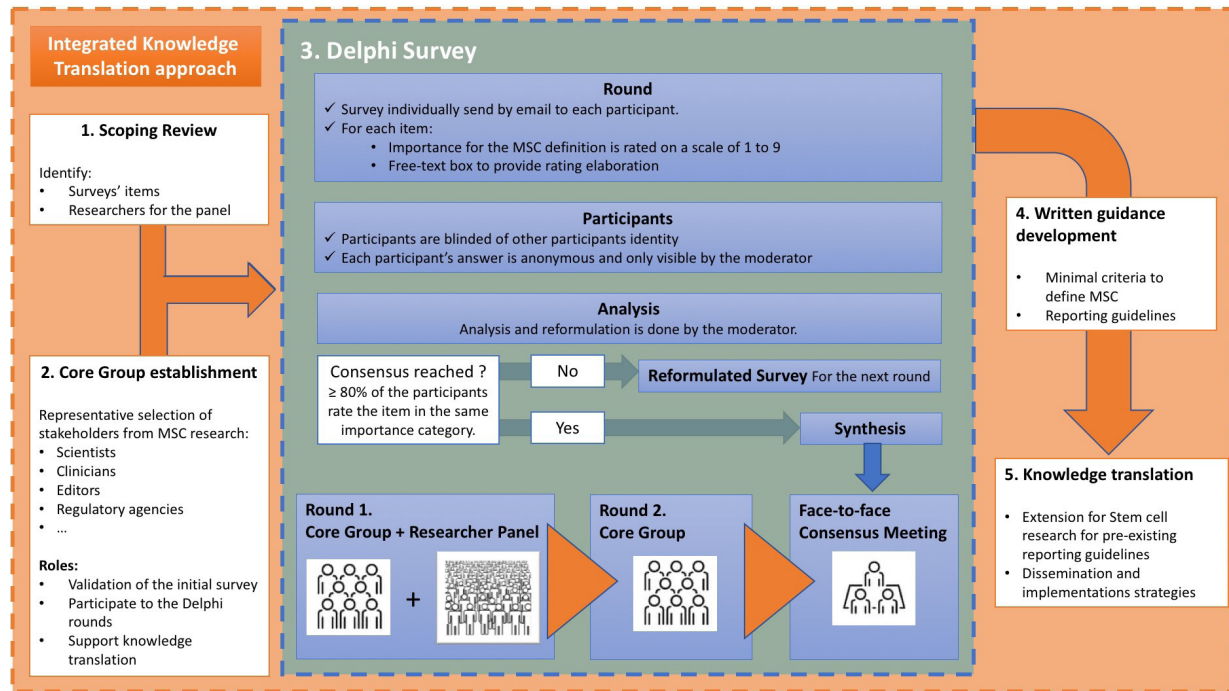
### Participants

Participants for this Delphi study will be identified in two ways:

- A. First, we will invite members of a ‘core group’ of experts that have been identified for the purpose of planning and implementing this research programme. The core group members are diverse in their make-up, including in vitro, preclinical and clinical researchers, and include member with backgrounds in developmental biology, translational science, research methods, regulatory practices, scholarly journal editing and industry.
- B. Second, we will invite a ‘researcher panel’ to participate in the Delphi. This panel will comprise 311 researchers that we identified in our previous scoping review of MSC research conducted between March and May 2020 (submitted). Our scoping review (<https://osf.io/3dsqx/>) was developed to describe how MSCs are currently defined in preclinical and clinical research, which will provide valuable themes to consider as items for the current Delphi survey. We extracted the corresponding author’s name, email and country of primary affiliation from all original articles included in this scoping review. These authors will be contacted by a member of the research team and invited to participate in the Delphi survey.

### Recruitment

Participants on the core group and on the researcher panel will be approached for participation in the study using a standardised email recruitment script (see online supplemental appendix 1). This script will provide participants with more information about the study and provide



**Figure 1** Study design. MSC, mesenchymal stromal cell.

a link to access the informed consent form and the round 1 online Delphi survey. The core group will take part in all three rounds of the Delphi. The researcher panel will be invited to the initial round only. This approach was taken to maximise the initial reach of our survey while maintaining a manageable number of individuals to facilitate our round 3 in-person consensus meeting.

Participants will be given 3 weeks to complete each round of the Delphi survey. Reminder emails will be sent 7 and 14 days following the dissemination of each online questionnaire.

### Study design

Our modified Delphi study will involve three rounds of surveying; the first two will take place via an online survey, while the third round will take place via an in-person consensus meeting. The study design is presented in [figure 1](#). The online Delphi surveys will be administered using Surveylet (<https://calibrium.com>), a cloud-based platform specifically designed for Delphi surveys.

- ▶ Round 1 (online): completed by the researcher panel and the core group.
- ▶ Round 2 (online): completed by the core group only.
- ▶ Round 3 (virtual consensus meeting): all items will be discussed by the core group members during a virtual meeting in 2022.

### Delphi survey

We have generated an initial survey (<https://osf.io/u579g/>) which has three sections: (1) participant demographics, (2) items for MSC definition, and (3) items for reporting guidelines. Based on participants' self-identified field of expertise, they will access different parts of the survey: MSC definition, reporting guidelines

or both. This survey was piloted initially by six members of our research group. Then, the questionnaire was sent to eight leaders of opinion in MSC research (core group members and leadership within international stem cell scientific societies) to ensure relevance and completeness of our items for definition and reporting guidelines.

After completing the demographics section, participants will be asked to rate each survey item's importance for MSC definition and/or as a minimal reporting criterion (eg, 'A description of MSC capacity to adhere to a plastic surface when maintained in standard culture condition, is essential to define them') on a scale of 1 (strongly disagree) to 9 (strongly agree). A free text response option will accompany each listed item to allow participants to elaborate on their rating. At the end of the survey, participants will also have the opportunity to suggest any new relevant items to define MSC or report clinical trials of MSC therapy that were not already listed on the survey. All the additional input will be analysed and added to the new questionnaire for the next round. This process allows any additional items (eg, haemocompatibility of MSC), to be discussed and rated by all the participants. After each round, data from the survey will be aggregated. A revised survey will be developed for the subsequent rounds that incorporates any newly suggested items, and items that did not reach consensus. Items that did not reach consensus will be presented with verbatim feedback participants made in the previous rounds.

Study conceptualisation and survey development began in the second and third quarters of 2021. We are currently refining and piloting the survey, and plan to administer the three rounds in the fourth quarter of 2021. The study should be completed by the first quarter of 2022.



## Data analysis

### Data analysis

Data will be analysed using Excel. Data will be collected anonymously. The total number of participants, item rating scores and demographic data will be summarised using frequencies and percentages. For each Likert scale item, we will report the mean and range. Mean rating scores for each item will be categorised into three groups (unessential, potentially essential, essential). A median rating of 1–3 means an item is deemed unessential, 4–6 implies that an item is potentially essential and a median rating of 7–9 is deemed essential to report.

Consensus will be achieved when at least 80% of Delphi participants rate the item in the same category of importance: unessential or essential. This threshold was described as the most commonly used in a systematic review of consensus in Delphis.<sup>21</sup> Items that achieve consensus, as being ‘essential or unessential’, will be labelled as such, and in subsequent rounds participants will have the opportunity to comment on these items. Items that did not achieve consensus, achieved consensus as ‘potentially essential’, had major wording changes or were added by Delphi participants will automatically move on to the subsequent rounds. For each round, response rate will be recorded and reported.

### First round analysis

Any items from round 1 with disagreement between the core group and researcher panel will be provided to the core group in round 2 to revote on. For transparency, all the results will be reported separately in a final manuscript so that any discrepancies between the core group and the research group will be clearly stated.

### Second round analysis

After the second round with the core group, items that reach consensus (>80% agreement) in the essential category will be considered for inclusion in the MSC definition and the reporting guidelines. Items that reach consensus in the unessential category will be considered for exclusion. All remaining items that have not yet reached consensus will be presented and discussed at the final face-to-face virtual meeting where they will then be electronically voted on anonymously using Surveylet software (<https://calibrium.com>).

### Potential limitations and mitigation strategies

One feature of the traditional Delphi method is that the participants are isolated from each other, with no in-person interaction or information exchange to limit group influences and peer pressure. Face-to-face interaction can be useful to identify reasons for any disagreements. To address this limitation, we plan to organise the third round as face-to-face virtual meeting, after which participants again vote anonymously. A lack of participation overall is another concern for any survey-based study. We have mitigated this risk by involving thought leaders from MSC science and other stem cell areas who have

committed to contributing. Furthermore, our previously conducted scoping review captured 311 MSC researchers, and as such if not all researchers respond we still anticipate sufficient recruitment numbers. Another potential limitation of the Delphi method is respondent attrition between rounds. We have addressed this by engaging our core group from project inception. During the rounds, reminder emails will be sent to solicit greater responses, and to limit participant attrition between rounds, we plan to analyse the results of each round and provide feedback and a new questionnaire rapidly to maintain the interest and engagement of participants. Respondent attrition per round will be recorded and reported.

## ETHICS AND DISSEMINATION

### Ethics

Ethics approval was obtained from the Ottawa Health Research Network Research Ethics Board (REB protocol ID 20210187-01K).

### iKT approach

The iKT will ensure that the MSC research community is aware of the recommendations described in our consensus definition and reporting guidelines in order to maximise the impact of the work. This entails the early inclusion of key stakeholders who have the ability to implement our recommendations, including patient partners. This participation begins from project inception and continues through to dissemination and implementation of study findings. The goal of the iKT approach is to ensure that the needs and preferences of stakeholders are considered throughout the project, with the idea being that successful partnership upstream will facilitate improved implementation and uptake downstream.<sup>22</sup>

Finally, to further support the dissemination of our findings (definition and reporting guidelines), we will use a structured approach to end of project, modelled on Canadian Institutes of Health Research’s Knowledge Translation planning guide (<https://cihr-irsc.gc.ca/e/45321.html>). This essentially will involve considering the specific project findings and determining the dissemination goal (eg, to increase awareness of the findings, to increase knowledge, to influence practice or policy), identifying key audiences, crafting messages tailored to specific audiences and using strategies and media to reach each audience. A written knowledge translation plan will help ensure study outputs are effectively used. Our knowledge users will help interpret the Delphi results and craft dissemination messages. We plan to create an extension to existing clinical study reporting guidelines (eg, Consolidated Standards of Reporting Trials) that will focus on aspects unique to MSC therapy trials (eg, manufacturing, characterisation, storage and delivery of cells). To support its implementation, the guideline draft will be published as a preprint to obtain feedback from the scientific community and associated explanatory documents

such as a guidance development statement and an Explanation and Elaboration document will be provided.

### Registration and data availability

This study protocol was registered using the Open Science Framework (OSF) (<https://osf.io/3dsqx/>). Data and study materials will be made publicly available at the time of publication using OSF.

### Patient and public involvement

A patient partner is involved in our project since inception. The patient partner will provide feedback on the consensus definition of MSC and reporting guidelines development and on knowledge translation strategies. They will help codevelop a communication plan for the general public by supporting the development of non-technical summaries for the scoping review and the Delphi findings as well as their dissemination to general public via patients' associations and social media.

### Author affiliations

<sup>1</sup>Sinclair Center for Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

<sup>2</sup>Clinical Epidemiology Program, Centre for Journalology, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

<sup>3</sup>Faculty of Medicine, School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada

<sup>4</sup>Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

<sup>5</sup>Department of Anesthesiology and Pain Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

<sup>6</sup>Neonatology, Department of Pediatrics, CHEO Research Institute, Ottawa, Ontario, Canada

<sup>7</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

**Twitter** Maxime Le @MaxLeOTT and Manoj M Lalu @manojlalu

**Acknowledgements** The authors want to thank the following members from the Thebaud lab for their valuable feedbacks on the initial Delphi questionnaire: Pauline Bardin, Chanele Cyr-Depauw, Flore Lesage, Marissa Lithopoulos, Ivana Mizikova and Lihuan Xu.

**Contributors** LR: conceptualisation, methodology and writing (original draft preparation). MML, KDC, BT: conceptualisation, funding acquisition, methodology, supervision and writing (review and editing). ML (patient partner): manuscript review and editing.

**Funding** This work is supported by the Canadian Stem Cell Network, 'Translation and Society Team Awards' 2020 grant. MML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and holds a University of Ottawa Junior Research Chair in Innovative Translational Research.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially,

and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

### ORCID iDs

Laurent Renesme <http://orcid.org/0000-0001-6848-3617>

Kelly D Cobey <http://orcid.org/0000-0003-2797-1686>

Manoj M Lalu <http://orcid.org/0000-0002-0322-382X>

Bernard Thebaud <http://orcid.org/0000-0003-1844-7145>

### REFERENCES

- Friedenstein AJ, Gorskaja JF, Kulagina NN. Fibroblast precursors in normal and irradiated mouse hematopoietic organs. *Exp Hematol* 1976;4:267–74.
- Lazarus HM, Haynesworth SE, Gerson SL, et al. Ex vivo expansion and subsequent infusion of human bone marrow-derived stromal progenitor cells (mesenchymal progenitor cells): implications for therapeutic use. *Bone Marrow Transplant* 1995;16:557–64.
- Clinicaltrials.Gov. Available: <https://clinicaltrials.gov/ct2/results?recrs=&cond=&term=mesechymal+AND+%28Stromal+OR+stem%29+AND+cell&cntry=&state=&city=&dist=> [Accessed 10 Apr 2021].
- Sipp D, Robey PG, Turner L. Clear up this stem-cell mess. *Nature* 2018;561:455–7.
- Galipeau J, Sensébé L. Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. *Cell Stem Cell* 2018;22:824–33.
- Levy O, Kuai R, Siren EMJ, et al. Shattering barriers toward clinically meaningful MSC therapies. *Sci Adv* 2020;6:eaba6884.
- Mendicino M, Bailey AM, Wonnacott K, et al. MSC-based product characterization for clinical trials: an FDA perspective. *Cell Stem Cell* 2014;14:141–5.
- Pierro M, Thébaud B, Soll R. Mesenchymal stem cells for the prevention and treatment of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev* 2017;11:CD011932.
- Lalu MM, McIntyre L, Pugliese C, et al. Safety of cell therapy with mesenchymal stromal cells (safecell): a systematic review and meta-analysis of clinical trials. *PLoS One* 2012;7:e47559.
- Thompson M, Mei SHJ, Wolfe D, et al. Cell therapy with intravascular administration of mesenchymal stromal cells continues to appear safe: an updated systematic review and meta-analysis. *EClinicalMedicine* 2020;19:100249.
- Viswanathan S, Shi Y, Galipeau J, et al. Mesenchymal stem versus stromal cells: international society for cell & gene therapy (ISCT®) mesenchymal stromal cell committee position statement on nomenclature. *Cytotherapy* 2019;21:1019–24.
- Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The international society for cellular therapy position statement. *Cytotherapy* 2006;8:315–7.
- Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:1–88.
- Robey P. "Mesenchymal stem cells": fact or fiction, and implications in their therapeutic use. *F1000Res* 2017;6. doi:10.12688/f1000research.10955.1. [Epub ahead of print: 20 04 2017].
- Snyder J, Turner L, Crooks VA. Crowdfunding for unproven stem interventions. *JAMA* 2018;319:1935.
- Turner L. Preying on public fears and anxieties in a pandemic: businesses selling unproven and unlicensed "stem cell treatments" for COVID-19. *Cell Stem Cell* 2020;26:806–10.
- Dalkey NC. The delphi method: an experimental study of group opinion, 1969. Rand Corporation. Available: [https://www.rand.org/pubs/research\\_memoranda/RM5888.html](https://www.rand.org/pubs/research_memoranda/RM5888.html)
- Grudniewicz A, Moher D, Cobey KD, et al. Predatory journals: no definition, no defence. *Nature* 2019;576:210–2.
- Rivara FP, Tennyson R, Mills B, et al. Consensus statement on sports-related concussions in youth sports using a modified delphi approach. *JAMA Pediatr* 2020;174:79.
- McGinn R, Fergusson DA, Stewart DJ, et al. Surrogate humane endpoints in small animal models of acute lung injury: a modified delphi consensus study of researchers and laboratory animal veterinarians. *Crit Care Med* 2021;49:311–23.
- Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of delphi studies. *J Clin Epidemiol* 2014;67:401–9.
- Kothari A, McCutcheon C, Graham ID. Defining integrated knowledge translation and moving forward: a response to recent commentaries. *Int J Health Policy Manag* 2017;6:299–300.

**5. Letter to the Editor “A systematic approach to enhance transparency in Mesenchymal Stromal Cell research” (published in Cytotherapy)**



ELSEVIER

Contents lists available at ScienceDirect

# CYTOTHERAPY

journal homepage: [www.isct-cytotherapy.org](http://www.isct-cytotherapy.org)

International Society  
**ISCT**  
 Cell & Gene Therapy®

Letter to the Editor

## A systematic approach to enhance transparency in mesenchymal stromal cell research

For any topic to be studied systematically, there is a need for clarity in definition. The definition of mesenchymal stromal cells (MSCs) is presently ambiguous in the literature, and the misleading interchangeable terminology of MSCs as mesenchymal stromal or stem cells is common [1,2]. This causes confusion among scientists and clinicians, which can hinder their ability to interpret and judge the reliability of studies or replicate the findings. This lack of clarity may contribute to patients' and the general public's perception of a "one size fits all" MSC and contribute to the rise of problematic stem cell medical clinics worldwide hawking unproven therapies. These observations have led to calls to "clear up this stem cell mess" [3] and a vision to create a consensus definition of MSCs as well as standards for complete and transparent reporting of both cell characteristics and clinical trial/manufacturing details.

In 2006, the International Society for Cell & Gene Therapy (ISCT) provided the first consensus definition for the minimal criteria to define MSCs [4], which was recently updated [5]. The terminology of "multipotent mesenchymal stromal cells" and the criteria described were determined through an informal consensus of a small number of experts from the ISCT. Numerous limitations of the resulting definition have been raised since its publication, including skepticism regarding the cell surface markers phenotype; differentiation assays, which are prone to misinterpretation; limitations in functionally defining stromal cells; and inadequate consideration of tissue source [6]. As a result, uptake of the ISCT's definition has been inconsistent in both preclinical and clinical MSC research. In our recent scoping review of recently published MSC studies, we found that just 18% of the references explicitly referred to the ISCT's definition (29% of clinical studies, 13% of animal studies and 19% of *in vitro* studies) [7].

To address issues pertaining to the development and implementation of a consensus definition, our group aims to re-approach this challenge and create an international consensus definition of MSCs, as well as minimal reporting standards for MSC studies. This Stem Cell Network-funded project will build on previous approaches for consensus in several ways. To ensure optimal development and implementation of a consensus definition of MSCs, two considerations are critical. First, diverse stakeholders in the MSC research community need to be consulted and directly involved in the definition's development. This necessitates an integrated knowledge translation approach, which is a model of collaborative research in which knowledge users are involved in the development of the definition from the inception of the work [8]. This integration ensures that their needs and preferences are considered during the development of the definition, which helps to ensure the definition developed will have maximum uptake within the community. To this end, we have recruited a core group of internationally renowned developmental

biologists, translational scientists, clinical researchers, methodologists, regulatory agency representatives, journal editors and private stem cell company representatives. We also include implementation experts to ensure the output of the work is effectively integrated into the community. In addition, we have participation of several key organizations, including ISCT, TERMIS, EuroStemCells and Canadian Stem Cell Network. Further, we conducted a scoping review to identify how MSCs are currently described in the literature, as a baseline to inform discussion among these experts [7]. We extracted corresponding author contact details from identified documents in this scoping review, which will enable us to reach out to the broader researcher community beyond our core group and help to ensure diverse, representative participation in this process.

A second key consideration is that methods used to generate consensus need to be transparent, systematic, and structured to reduce bias. We are adopting the Delphi method, a rigorous social science method that provides structured, facilitated communication, designed to allow expert stakeholders from diverse locations and areas of expertise to reach consensus on a controversial issue [9,10]. This method is routinely used with great success to develop clinical and reporting guidelines. Our planned Delphi involves three rounds of surveying of MSC community stakeholders. During each round, participants will vote using Likert scales to indicate whether potential characteristics are essential to the definition of MSC. Between rounds, participants are provided with the group's mean ratings for each item, their previous answers and written participant feedback for each potential item, and new items suggested by participants may also be introduced for voting. Thus, the Delphi method provides an opportunity for participants to communicate their opinions and knowledge on this controversial issue in an anonymous and equitable manner. This will allow them to assess how their evaluation of issues aligns with others, and it will offer opportunity to change their opinion, if desired, after reviewing and considering the collective findings of the group's ideas.

It may be critical to return to the drawing board and reconsider the fundamental, original description of stem cells provided by McCulloch and Till [11]. This may involve avoiding the term "stem" for the more descriptive phenomenon they observed of a "colony-forming cell" [11]. Similarly, Friedenstein *et al.*'s original description [12] of MSCs as a bone marrow-derived "colony-forming unit-fibroblast" may be most appropriate until more knowledge is available to satisfactorily describe these cells [12]. The issue is timely, as highlighted by a recent review concluding that "more than a decade after the first attempts, clarifying this nomenclature by agreeing on a name and a clear set of cellular properties for MSCs should be a priority for the field" [13]. We will be able to clarify these questions



through our Delphi that will actively seek the input of the large and diverse MSC research community. Ultimately, its unbiased approach provides us with tools to enhance transparency, reportability and potentially translation of research and to clear up the stem cell mess. The planned integrated knowledge translation approach [14] (<https://osf.io/3dsqx/>) will ensure that the definition we arrive at is acceptable to the community and integrated into community norms.

## References

- [1] Caplan AI. What's in a name? *Tissue Eng Part A* 2010;16:2415–7.
- [2] Galipeau J, Sensébé L. Mesenchymal Stromal Cells: Clinical Challenges and Therapeutic Opportunities. *Cell Stem Cell* 2018;22:824–33.
- [3] Sipp D, Robey PG, Turner L. Clear up this stem-cell mess. *Nature* 2018;561:455–7.
- [4] Dominici M, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* 2006;8:315–7.
- [5] Viswanathan S, et al. Mesenchymal stem versus stromal cells: International Society for Cell & Gene Therapy (ISCT®) Mesenchymal Stromal Cell committee position statement on nomenclature. *Cytotherapy* 2019;21:1019–24.
- [6] Robey P. Mesenchymal stem cells': fact or fiction, and implications in their therapeutic use. *F1000Res* 2017;6.
- [7] Renesme, L. et al. Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review. *STEM CELLS Transl Med* Accepted for publication.
- [8] Kothari A, McCutcheon C, Graham ID. Defining Integrated Knowledge Translation and Moving Forward: A Response to Recent Commentaries. *Int J Health Policy Manag* 2017;6:299–300.
- [9] Murphy MK, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:1–88. i–iv.
- [10] Taylor E. We Agree, Don't We? The Delphi Method for Health Environments Research. *HERD* 2020;13:11–23.
- [11] McCulloch EA, Till JE. The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. *Radiat Res* 1960;13:115–25.
- [12] Friedenstein AJ, Gorskaja JF, Kulagina NN. Fibroblast precursors in normal and irradiated mouse hematopoietic organs. *Exp Hematol* 1976;4:267–74.
- [13] Soliman H, et al. Multipotent stromal cells: One name, multiple identities. *Cell Stem Cell* 2021;28:1690–707.
- [14] Renesme L, Cobey KD, Le M, Lalu MM, Thebaud B. Establishment of a consensus definition for mesenchymal stromal cells (MSC) and reporting guidelines for clinical trials of MSC therapy: a modified Delphi study protocol. *BMJ Open* 2021;11:e054740.

Laurent Renesme

Kelly D Cobey

Manoj M Lalu

Bernard Thébaud\*

*Regenerative Medicine Program, Ottawa Hospital Research Institute,  
Ottawa, ON, Canada*

*Centre for Journalology, Clinical Epidemiology Program, The Ottawa  
Hospital Research Institute, Ottawa, Canada*

*School of Epidemiology and Public Health, Faculty of Medicine,  
University of Ottawa, Ottawa, Canada*

*Department of Anesthesiology and Pain Medicine, Faculty of Medicine,  
University of Ottawa, Ottawa, Canada*

*Department of Cellular and Molecular Medicine, University of Ottawa,  
Ottawa, ON, Canada*

*Neonatology, Department of Pediatrics, Children's Hospital of Eastern  
Ontario (CHEO) and CHEO Research Institute, Ottawa, ON, Canada*

\*Corresponding Author: Bernard Thébaud, Ottawa Hospital Research  
Institute, CCW Room W6120, 501 Smyth Road, Ottawa K1H 8L6, ON,  
Canada

E-mail address: [bthebaud@ohri.ca](mailto:bthebaud@ohri.ca) (B. Thébaud).

Received 13 December 2021

Accepted 17 December 2021

## **IV. Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review**

This study is published in *Stem Cells Translational Medicine*. The graphical abstract is presented in Figure 17.

### **1. Background**

MSC are widely used in preclinical and clinical research. Despite minimal criteria to define MSC provided by the International Society for Cell and Gene Therapy (ISCT), concerns have been raised about inconsistent descriptions of cell products used. To address the question ‘How are MSC currently defined and characterized?’ we conducted a scoping review on original MSC preclinical and clinical studies. In this scoping review, we used a systematic search to map the current literature, and to identify key concepts and knowledge gaps. In addition, the scoping review was used to identify and inform items for the initial Delphi survey, as well as to identify the researchers among corresponding authors from selected original studies to participate to the Delphi survey.

### **2. Results**

Original preclinical and clinical studies using MSC were identified from a systematic search of MEDLINE and Embase over a 3-month period (March 1<sup>st</sup> to May 31<sup>st</sup>, 2020). Studies screening was performed by 2 independent reviewers and categorized as: clinical, animal, biology, or biomaterial studies. Data were extracted from a randomly selected sample of studies and included information on study demographics, study design, ISCT criteria, and MSC characterization and culture condition.

A total of 1053 articles were included and among them 318 articles were analyzed (42 clinical studies, 77 animal studies, 160 biology studies, and 39 MSC and biomaterial studies). The top 3 countries of affiliation were China (37%), United States (13%), and Korea (6%). Overall, only 18% of the articles explicitly referred to the ISCT minimal criteria for MSC. Minimal criteria from the ISCT princeps publication in 2006, i.e., plastic adherence, cell surface markers and differentiation assay, were poorly reported, with 36%, 53% and 40% of our sample of studies, respectively. MSC characteristics and culture conditions were inconsistently reported (e.g., viability assay reported in only 18 % of the articles), which is of concern, as culture condition can greatly impact MSC phenotype. Only 20% of the studies reported at least one functional assay. Clinical studies showed inconstant completeness in reporting relevant information on the MSC characterization and cell manufacturing processes. Clinical studies also reported a broad variety of MSC administration protocols in regards of administration route, MSC dose (range  $10^5$  to  $10^7$  cells/kg) and number of doses administered (range 1 to 25 doses).

### **3. Conclusion**

Our scoping review highlighted a wide variability in reporting quality, consistence, and completeness in both MSC definition and characterization, and culture condition. These results suggest that further efforts are required to ensure the adoption of a consensus definition of MSCs and reporting guidelines in order to enhance rigour, reproducibility, transparency in the MSC literature, and ultimately the safe translation of effective cell-based therapies.

## 'How are Mesenchymal Stromal Cells (MSCs) defined and characterized?'

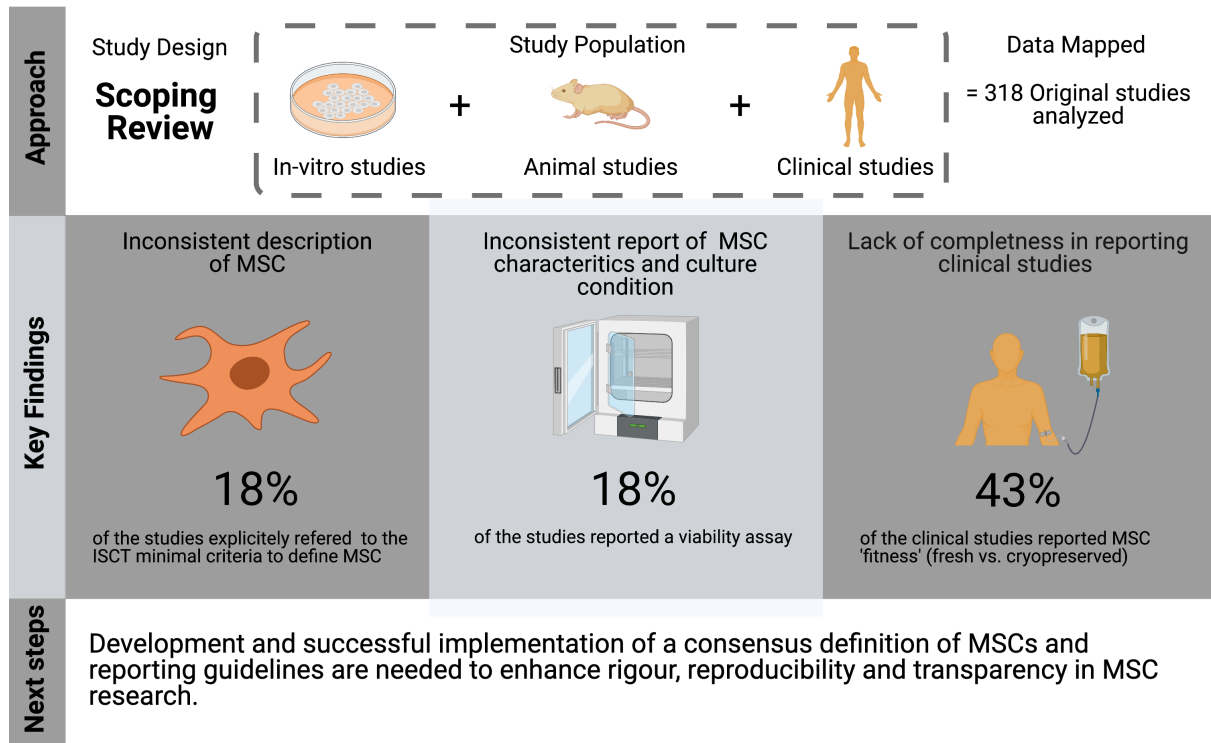


Figure 17. Scoping review graphical abstract.

### 4. Manuscript "Definition and characteristics of mesenchymal stromal cells (MSC) in preclinical and clinical studies: a scoping review" (Published in Stem Cells Translational Medicine)

# Definition and Characteristics of Mesenchymal Stromal Cells in Preclinical and Clinical Studies: A Scoping Review

Laurent Renesme<sup>1</sup>, Maria Pierro<sup>2</sup>, Kelly D. Cobey<sup>3,4</sup>, Rhea Mital<sup>1,5</sup>, Kennedy Nangle<sup>1</sup>, Risa Shorr<sup>3</sup>, Manoj M. Lalu<sup>1,3,6,7</sup>, Bernard Thébaud<sup>1,5,7,\*</sup>

<sup>1</sup>Regenerative Medicine Program, The Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>2</sup>Neonatal and Paediatric Intensive Care Unit, M. Bufalini Hospital, AUSL Romagna, Cesena, Italy

<sup>3</sup>Centre for Journalology, Clinical Epidemiology Program, The Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>4</sup>School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

<sup>5</sup>Neonatology, Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, Ottawa, ON, Canada

<sup>6</sup>Department of Anesthesiology and Pain Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

<sup>7</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada

\*Corresponding author: Bernard Thébaud, MD, PhD, Regenerative Medicine Program, The Ottawa Hospital Research Institute, CCW Room W6120, 501 Smyth Road, Ottawa, ON K1H 8L6, Canada. Email: [bthebaud@ohri.ca](mailto:bthebaud@ohri.ca)

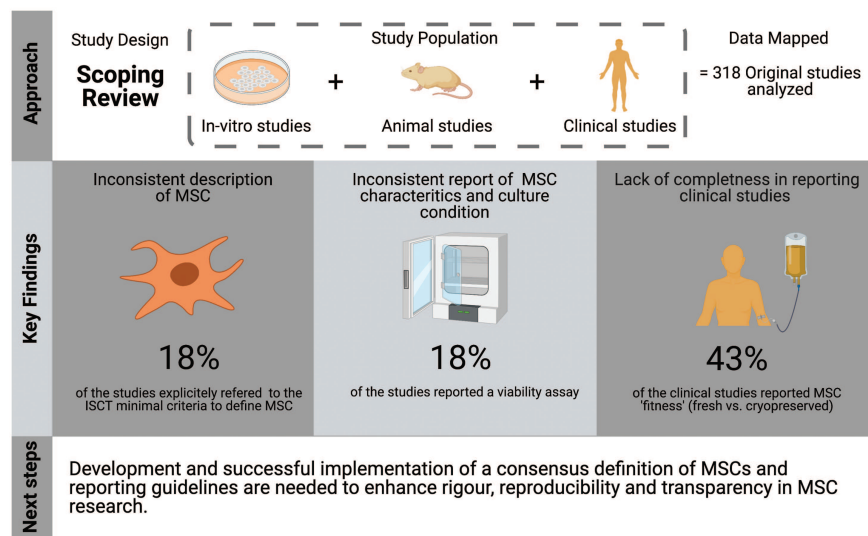
## Abstract

Mesenchymal stromal cells (MSCs) are widely used in preclinical and clinical research. Despite minimal criteria to define MSCs provided by the International Society for Cell and Gene Therapy (ISCT), concerns have been raised about inconsistent descriptions of cell products used. To address the question “How are MSCs currently defined and characterized?” we conducted a scoping review on original MSC preclinical and clinical studies published over a 3-month period. Selected studies identified from a systematic search of MEDLINE and Embase were categorized as follows: Clinical, Animal, Biology, or Biomaterial studies. Data were extracted from a randomly selected subsample of studies. We extracted information, including epidemiological characteristics of studies, study design, ISCT criteria, and MSC characterization and culture condition. A total of 1053 articles were included and among them, 318 articles were analyzed. Overall, 18% of the articles explicitly referred to the ISCT minimal criteria for MSC. MSC characteristics and culture conditions were inconsistently reported (eg, viability assay reported in only 18% of the articles). Only 20% of documents reported at least 1 functional assay. Clinical studies showed inconsistent completeness in reporting relevant information on the MSC characterization and cell manufacturing processes. These results suggest that further development and implementation of a consensus definition of MSCs and reporting guidelines are needed to enhance rigor, reproducibility, and transparency in MSC research.

**Key words:** mesenchymal stromal cell; MSC definition; MSC characteristics; scoping review; preclinical; clinical studies.

## Graphical Abstract

‘How are Mesenchymal Stromal Cells (MSCs) defined and characterized?’



Received: 27 April 2021; Accepted: 5 October 2021.

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Significance Statement

Our scoping review highlights several findings that require the attention of the stem cell community. The minimal criteria to define MSC proposed by the International Society for Cell and Gene Therapy (ISCT) were poorly implemented with inconsistent reporting. More concerning, the clinical studies showed inconsistent completeness in reporting relevant and important information on MSC characterization and cell manufacturing processes. Further efforts are required to ensure the adoption of a consensus definition of MSCs and reporting guidelines in order to enhance rigor, reproducibility, transparency in the MSC literature, and ultimately the safe translation of effective cell-based therapies.

## Introduction

Since mesenchymal stromal cells (MSC) were first tested as a therapeutic agent in 1995,<sup>1</sup> more than 1300 MSC clinical trials have been registered on [clinicaltrials.gov](https://clinicaltrials.gov).<sup>2</sup> Despite highly promising results of MSCs in different preclinical disease models, results of clinical trials using MSCs in various medical conditions have been less encouraging, with currently only 2 clinical approvals for graft versus host disease and Crohn's associated perianal fistula.<sup>3,4</sup> Although many issues have contributed to failures in translation (eg, patient characteristics and comorbidities), disparities in MSC characteristics (eg, definition, characterization, immune compatibility, cell viability, and dose) appear to be critical factors.<sup>4</sup> Within the clinical trials that have tested MSC therapy there has been significant heterogeneity in reporting of products used, despite attempts from the International Society for Cell and Gene Therapy (ISCT) to provide minimal criteria to define MSC.<sup>5,6</sup> For example, a report from the Food and Drug Administration showed important differences in cell surface marker characterization, product bioactivity assessment, as well as tissue sourcing and product manufacturing.<sup>7</sup> Compounding the issue of heterogeneous cell products is incomplete reporting.<sup>8-12</sup>

To better address clinical translation, reproducibility, and transparency in the field of MSC research, the scientific community needs to agree a consensus definition of MSCs and supports its dissemination and implementation. The absence of a consensus definition will lead to ongoing difficulties in study quality assessment, comparison between studies, extrapolation from study findings, and even possibly influence the results of preclinical and clinical reports.

Here we seek to address the question "How are MSCs currently defined and characterized?" The objective of the current study is to describe how MSCs are defined and characterized in preclinical and clinical research assessing MSCs' therapeutic potential. In this scoping review, we used a systematic search to map the current literature and identified key concepts and knowledge gaps.<sup>13,14</sup> This scoping review is the first step in a larger research program that seeks to establish a new consensus definition of MSCs.<sup>15</sup> Our results will inform a subsequent Delphi study<sup>16</sup> to establish and implement an international definition of MSCs.

## Materials and Methods

Our research protocol was drafted according to the methodological framework for scoping reviews proposed by Arksey et al and adapted by Levac et al<sup>13,14</sup> and further updated by the Joanna Briggs Institute.<sup>17</sup> This protocol was registered prospectively using the Open Science Framework<sup>18</sup> (see <https://osf.io/3dsqx/>) The data charting form was continuously updated from the protocol version as part of an iterative process

as data was charted. The methods and findings of this study are reported in accordance with the PRISMA Extension for Scoping Review.<sup>19</sup> Study materials and data can be found here: <https://osf.io/3dsqx/>.

## Research Question

Our research question was: "How are MSCs currently defined and characterized in published preclinical and clinical studies?" The purpose of this scoping review was to describe how MSCs are described and defined in a sample of preclinical and clinical research literature.

## Search Strategy: Identifying Relevant Studies

We identified relevant original preclinical and clinical MSC studies published over a 3-month period (March 1-May 31, 2020). A 3-month period was chosen a priori due to the large number of publications related to MSCs.

The search strategy (see [Supplemental material](#)) was developed by an experienced information specialist (R.S.) and further refined through team discussion. This search strategy was modified from previous systematic reviews of MSCs by our group,<sup>10,20</sup> and underwent Peer Review of Electronic Search Strategy (PRESS) to ensure adequate sensitivity and specificity.<sup>21</sup> To identify potentially relevant studies, the following MEDLINE and Embase were searched. We ran simultaneous searches of Ovid Medline All and Embase using a broad search strategy. MeSH terms for MSCs were searched along with synonymous text words in the titles or abstracts such as multipotent stromal cells.

Results of this search were limited to English language articles, de-duplicated, and then uploaded into DistillerSR<sup>22</sup> (Evidence Partners, Ottawa), a cloud-based, audit-ready software that facilitates screening and selection of articles and allows transparent and reproducible work.

## Study Selection

### The Study Population

The study population included original preclinical and clinical MSC studies. To be included, articles needed to report original research using MSC as a main intervention/focus and/or assess its therapeutic potential (for Animal and Clinical studies). We excluded studies if MSC were not the main intervention/focus, if the study did not investigate mesenchymal stromal/stem cells and if it was not an original study (eg, editorial, review). Systematic reviews and meta-analyses were also excluded.

### Screening

Two independent reviewers (L.R., M.P.) performed the study selection using DistillerSR. For each screening step (title



and abstract and full text), calibration exercises were done on 10 random articles to ensure inter-reviewer reliability. Conflicts were resolved by consensus among screeners, or if needed by a third independent reviewer (B.T.). First, the 2 independent reviewers screened article titles and abstracts in duplicate using an initial screening questionnaire. Subsequently, full-text screening for all the articles retained was conducted against our eligibility criteria. Selected studies were stratified according to 3 categories: In Vitro, Animal, and Clinical studies. In vitro studies were stratified using the following categories: MSC Biology, MSC, and Biomaterial. Finally, if the number of articles included a category was >100, we used a random sample of 25% for data charting. The random sample was selected using a function in DistillerSR.

## Charting the Data

### Data-Extraction Forms

Three data-extraction forms were developed *a priori* for each category (Clinical, preclinical Animal, and preclinical In Vitro) and pilot-tested by our team. These forms were designed to capture epidemiological characteristics of the original study and detailed information on MSC descriptions and their use. The data-extraction forms collected information on:

- *Epidemiological characteristics.* Information on publication year, corresponding author (name, email, and country of affiliation), and funding (reported funding and funding source) were captured.
- *Study design (Clinical and Animal categories only).* Information on the disease studied, intervention group (MSC dose, administration route, concentration, etc.), and control group were collected.
- *MSC description and reference to ISCT minimal criteria.* This section of the form was designed to evaluate how authors referred to the ISCT criteria (eg, plastic adherence, cell markers, tissue source, differentiation assays) and provide information on which criteria were used and how detailed were these criteria.
- *MSC characteristics and culture condition.* This section was designed to inform the MSC characteristics (species sources, compatibility for Clinical and Animal studies, “fitness,” and viability assessment) and what were the culture condition, including a number of passages prior to MSC administration, cell confluence during culture, oxygen condition, medium, and serum used.

The full data-extraction forms are accessible on the Open Science Framework (<https://osf.io/3dsqx/>).

### Data Charting Process

Two independent reviewers (L.R., M.P.) performed the data extraction using a single charting and audit approach using the quality control function in DistillerSR. Each reviewer charted half of the articles and audited the other half. In case of disagreement between the reviewers, a third independent reviewer (B.T.) was consulted. The extraction forms were piloted on 5 random studies of each sample (in vitro, animal, and human) to ensure the approach to data charting was consistent and in line with the research question and purpose. Then a calibration exercise was done on the next 10 articles for each category. The team discussed results, and the data

charting form was continuously updated in an iterative process in order to be inclusive of other aspects of the cell characterization, manufacturing, delivery, etc. not listed a priori.

Extracted information will be used to generate our initial set of items for the future Delphi survey. In addition, the corresponding author’s names/email and affiliation country were extracted from all original articles selected for charting. These authors will be contacted by a member of the research team to participate in the Delphi survey.

## Collating, Summarizing, and Reporting Results

We conducted our data analysis distinctly for each study design group (Biology, Biomaterial, Animal, and Clinical studies) and involved both quantitative (ie, frequencies) and qualitative (ie, thematic analysis) methods. We have reported frequencies and percentages for original studies’ epidemiological characteristics (country listed in the first stated affiliation of the first listed corresponding author on each article for each study category, funding sources presented, and disease model) and study design (intervention group with MSC route of administration and dose, control group, cell nomenclature). Study participants are described as species for preclinical Animal studies (frequency and percentages) or participants’ age category for Clinical studies (pediatric, adult, or both). MSC description, frequencies, and percentages were reported for: ISCT minimal criteria for MSC (and functional matrix assays for Clinical studies), and MSC characteristics (tissue source, “fitness,” and culture condition).

At the request of reviewers, we also conducted an unplanned post hoc analysis to provide a description of how completeness of reporting related to journal impact factor. To do so, we obtained journal impact factors for included journals from Clarivate’s Journal Citation Reports for the year 2020 (<https://jcr.clarivate.com>). An additional post hoc analysis was done to compare the completeness of reporting between randomized controlled trials (RCT) and other clinical studies.

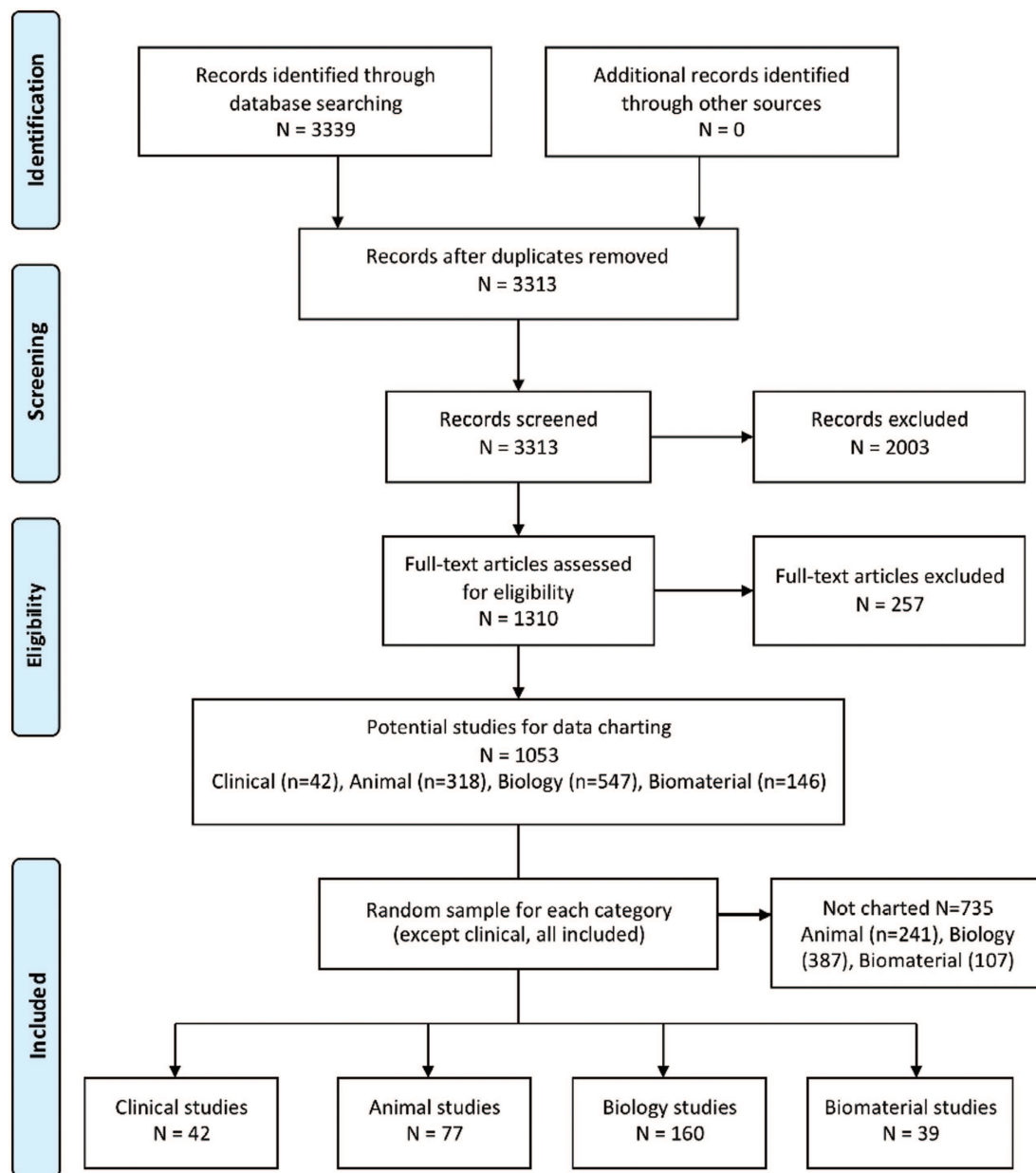
## Results

### Study Selection

In total, 3339 potential articles were identified and following screening 1053 articles met our inclusion criteria (Fig. 1). Among these articles, we selected all the included Clinical studies, and random samples of the Animal, Biology, and Biomaterial studies to conduct data extraction. A total of 318 articles were included for charting: 42 Clinical studies, 77 Animal studies, 160 Biology studies, and 39 MSC and Biomaterial studies. For data charting, the weighted overall Kappa was 0.93 for Clinical studies, 0.98 for Animal studies, 0.95 for Biology studies, and 0.89 for Biomaterial studies.

### Epidemiological Characteristics

Among all the selected articles ( $n = 318$ ), the top 3 countries were China (37%), US (13%), and Korea (6%). Funding was reported in 92% of the articles, 39% of the articles reported multiple funding sources, and Government was the most frequent source of funding (reported in 64% of the articles). A summary of article epidemiological characteristics is presented in Table 1 and more details are provided in Supplemental Table S1.



**Figure 1.** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

## General Study Characteristics

### Clinical Studies

Forty-two clinical studies were analyzed. Thirty-two (76%) were observational studies (27 descriptive studies and 5 analytical studies—1 case-control and 4 cohort studies). There were 10 (24%) experimental studies, of which 9 were RCT.

Of the 42 studies, 35 (83%) studies used an adult population, 2 (5%) a pediatric population, and 5 (12%) both adult and pediatric populations. Regarding the disease model, the 3 most frequently investigated systems were musculoskeletal system (29%), respiratory system (17%), and nervous system and genitourinary systems, both reported in 14% of the studies (Supplemental Table S2). These 42 Clinical articles described an important variety of administration protocols, with wide ranges for MSC doses and the number of doses administered (Table 2). MSC dose was reported in cells/kg (39%) with a median dose of  $10^6$  cells/kg (range  $10^5$  to  $10^7$

cells/kg) or in total cells per dose (58%) with a median dose of  $2.5 \times 10^7$  cells per dose (range  $5 \times 10^5$  to  $325 \times 10^6$  cells per dose). Multiple MSC doses were used in 29% of the Clinical articles with a median of 3.5 doses (range 2–25 doses). A control group was reported in 15 studies (36%). Control groups included a vehicle placebo (5 studies, 33%), standard care with no placebo (8 studies, 53%), other cell types (1 study, 7%), or unreported (1 study, 7%).

### Animal Studies

Seventy-seven studies were randomly selected and analyzed. All the articles reported the animal species used for the model; most of them were rodents: rat,  $n = 45$  (58%); mouse,  $n = 24$  (31%); rabbit,  $n = 5$  (6%); other (swine, sheep, and dog),  $n = 3$  (4%). Regarding the disease model, the 3 most reported systems were nervous system (25%), musculoskeletal system (16%), and cardiovascular system (12%). The



**Table 1.** Article epidemiological characteristics.

	Clinical (n = 42)	Animal (n = 77)	Biology n = 160)	Biomaterial (n = 39)	All articles (n = 318)
Top 3 journals (n)	- Stem Cell Research & Therapy (3)- Stem Cells Translational Medicine (2)- Knee Surgery, Sports Traumatology, Arthroscopy (2) - China (7)- US (6)- Spain (4)	- Stem Cell Research & Therapy (3)- International Journal of Stem Cells (3)- Stem Cells and Development (3)- Stem Cells International (3) - China (39)- US (6)- India (4)	- Stem Cell Research & Therapy (6)- Cells (5)- Molecular Medicine Reports (4) - China (58)- US (19)- Germany (13)- Korea (13) 145 (91) 60 (41)	- Journal of Materials Chemistry B (11)- ACS Biomaterials Science & Engineering (3)- ACS Applied Materials & Interfaces (3) - China (13)- US (6)- Iran (3) 37 (95) 20 (54)	- Stem Cell Research & Therapy (13)- Journal of Materials Chemistry B (12)- Stem Cells International (8) - China (117)- US (42)- Korea (20) 291 (92) 123 (42)
Top 3 countries (n)					
Funding reported	32 (76)	77 (100)	145 (91)	37 (95)	291 (92)
Multiple sources	11 (34)	32 (42)	60 (41)	20 (54)	123 (42)
Funding sources					
Government	15 (36)	53 (69)	107 (67)	28 (72)	203 (64)
Academic	7 (17)	26 (34)	50 (31)	21 (54)	104 (33)
Foundation	9 (21)	9 (12)	31 (19)	6 (15)	55 (17)
Industry	3 (7)	3 (4)	5 (3)	2 (5)	13 (4)
Hospital	9 (21)	9 (12)	8 (5)	1 (3)	27 (8)
Reported as not funded	6 (14)	2 (3)	5 (3)	0 (0)	13 (4)

Funding data presented as n (%).

**Table 2.** Intervention group in clinical studies.

	N (%)
Administration route reported	41 (98)
Multiple administration routes	4 (10)
Intravenous	13 (32)
Infusion rate reported in 6 of 13 studies (46%) using IV route	
Intratracheal	1 (2)
Intramuscular	4 (10)
Intrathecal	3 (7)
Intra-articular	3 (7)
Other routes	21 (51)
Cardiovascular system (intra-myocardium)	2 (10)
Digestive system	3 (14)
Ear-nose-throat	1 (5)
Eye	1 (5)
Genitourinary system	2 (10)
Musculoskeletal tissue (tendon)	8 (38)
Nervous system	2 (10)
Skin and subcutaneous tissue	2 (10)
MSC dose reported	38 (90)
Dose in cells/kg	15 (39)
Range $10^5$ to $10^7$ cells/kg	
Dose in a total amount of cell administered	22 (58)
Range $5 \times 10^5$ to $325 \times 10^6$ cells per dose	
Other dose reported	1 (3)
Dose reported in $\mu\text{L}/\text{cm}^2$ of the defect area	
Different dose groups used in the study	15 (39)
Single administration	28 (74)
Multiple administration	11 (29)
Range 2-25 doses	
MSC concentration reported	13 (31)
Use of dimethyl sulfoxide (DMSO) reported	9 (21)

Abbreviations: IV, intra-venous; MSC, mesenchymal stromal cells.

different disease models used in the animal studies are presented in [Supplemental Table S2](#). For the intervention group, the administration route was reported in all studies and 91% (70/77) of the animal studies reported the MSC dose. The most common administration route was intravenous, reported in 43% of the articles. MSC dose was reported in cells/kg in 24% of the articles with a median dose of  $2.25 \times 10^6$  cells/kg (range  $1.25 \times 10^5$  to  $10^7$  cells/kg) or in total cells per dose in 71% of the articles, with a median dose of  $10^6$  cells per dose (range  $3 \times 10^4$  to  $10^7$  cells per dose). Sixteen percent of the articles reported using multiple MSC doses, with a median of 3 doses (range 2-6 doses). The reported intervention group for animal studies is detailed in [Supplemental Table S3](#). A control group was reported in 76 studies (99%) with vehicle injection (67%) or no injection (28%).

## MSC Description and Reference to the ISCT Minimal Criteria

### Cell Nomenclature

Mesenchymal stem cell was used more frequently than MSC (Clinical 69 vs. 24%; Animal 82 vs. 17%, Biology 73 vs. 24%;

Biomaterial 82 vs. 15%, all articles 76 vs. 21%). A small proportion of the studies used both terms (2% of all the articles).

### Reference to ISCT Criteria and Recommendations

Overall, only 18% of the articles explicitly referred to the ISCT minimal criteria for MSC, with the highest percentage of articles in the Clinical category (29%) and the lowest for the Animal category (12%). The reported ISCT criteria according to each study category are presented in Fig. 2 and detailed in Table 3. To further explore the completeness in reporting the initial minimal criteria to define MSC from the ISCT (plastic adherence, cell markers, and in vitro differentiation assay), we described its association with the journal impact factor in Supplemental Table S4, and comparison between RCT and other clinical studies is presented in Supplemental Table S5. Among all the articles, 55 different positive or negative markers were used to define MSC. The top reported positive and negative cell markers are presented in Supplemental Fig. S1, and Supplemental Table S5 compared the top reported cell markers between RCT and other clinical studies. In addition to the minimal criteria to define MSC, to ensure product consistency, the ISCT also recommends to report tissue source as well as functional assays including in vitro MSC licensing with pro-inflammatory cytokines in order to mimic the in vivo environment in patients with systemic inflammation or abnormal immune response.<sup>23</sup> Tissue source was reported in 94% of all the articles with a wide variety of tissue used as MSC source. MSC licensing was used in 7% of all the articles, the molecule or substance used was always reported and a

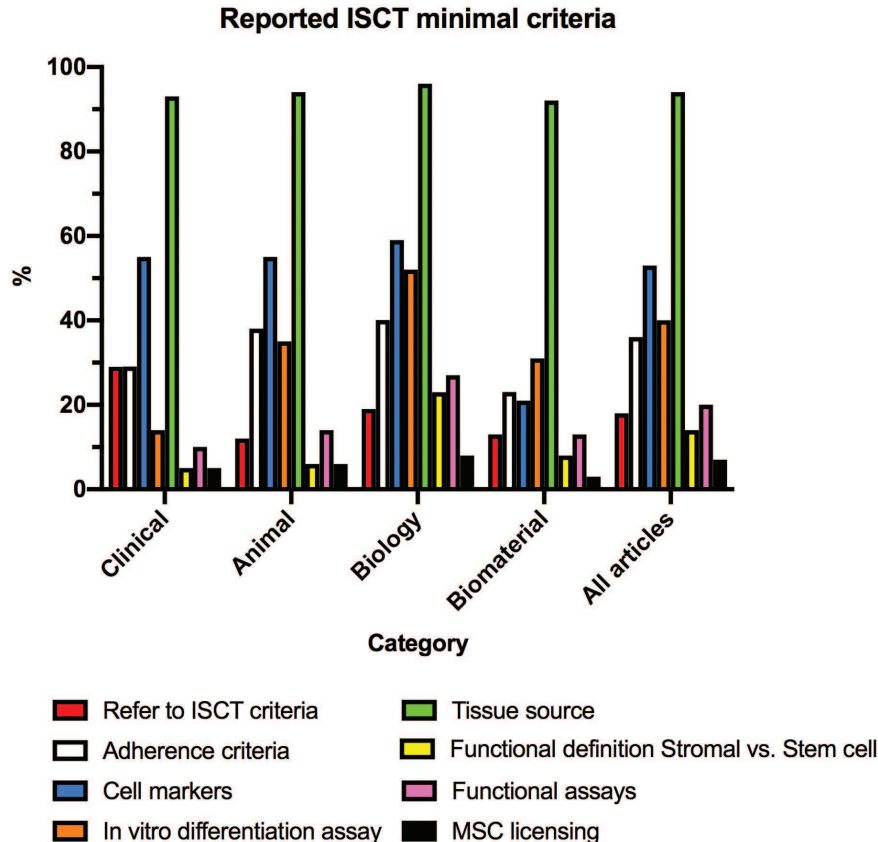
resting MSC was used as control in 24% of the articles reporting MSC licensing.

### MSC Characteristics and Culture Conditions

The reported MSC characteristics and culture conditions are presented in Fig. 3 and detailed in Table 4. For MSC characteristics, MSC source (ie, human, animal or commercial) was almost always reported (in 98% of all articles), but MSC “fitness” (ie, fresh or cryopreserved) as well as MSC viability assessment prior to administration were reported in only 16% and 18% of all articles, respectively. Reporting items describing the culture condition were inconsistently reported, culture medium type being the most reported (in 85% of all articles), and oxygen culture condition being the least reported (8% of all articles).

### Discussion

The aim of this scoping review was to describe how MSC are currently defined in preclinical (in vitro and animal studies) and clinical studies and to describe the characteristics of this published literature. In our selected articles, we found that only 29% of the clinical studies and 18% of all articles explicitly refer to the ISCT minimal criteria for MSC, and found important variations in criteria used to define MSC and the tissue sources, cell characteristics, and culture conditions. Both clinical and animal studies showed important variation in MSC dose. For example, in clinical studies, we found a two-log difference between the lowest and the highest dose



**Figure 2.** Reported International Society for Cell and Gene Therapy (ISCT) minimal criteria for mesenchymal stromal cells (MSC). This graph shows, for each study category, the percentage of studies reporting the criterion. The criterion “refer to ISCT criteria” refers to the number of studies where the authors stated that they were using the ISCT criteria.

**Table 3.** Reported International Society Cell and Gene Therapy (ISCT) minimal criteria to define mesenchymal stromal cells (MSC).

	Clinical ( <i>n</i> = 42)	Animal ( <i>n</i> = 77)	Biology ( <i>n</i> = 160)	Biomaterial ( <i>n</i> = 39)	All articles ( <i>n</i> = 318)
Refer to ISCT criteria	12 (29)	9 (12)	31 (19)	5 (13)	57 (18)
Studies without any criteria	2 (5)	5 (7)	6 (4)	3 (8)	16 (5)
Adherence criteria reported	12 (29)	29 (38)	64 (40)	9 (23)	114 (36)
Plastic	6 (50)	7 (24)	14 (22)	2 (22)	29 (25)
Surface not reported	6 (50)	22 (76)	50 (78)	7 (78)	85 (75)
Cell markers reported	23 (55)	42 (55)	95 (59)	8 (21)	168 (53)
In vitro differentiation assay reported	6 (14)	27 (35)	83 (52)	12 (31)	128 (40)
Adipocyte	6 (100)	24 (89)	60 (72)	8 (67)	98 (77)
Chondrocyte	4 (67)	14 (52)	39 (47)	5 (42)	62 (48)
Osteoblast	6 (100)	24 (89)	77 (93)	10 (83)	117 (91)
Tri-lineage	4 (67)	13 (48)	37 (45)	4 (33)	58 (45)
Tissue source reported	39 (93)	72 (94)	153 (96)	36 (92)	300 (94)
Bone marrow	17 (44)	40 (56)	95 (62)	26 (72)	178 (59)
Adipose tissue	9 (23)	9 (13)	29 (19)	5 (14)	52 (17)
Umbilical cord	10 (26)	18 (25)	17 (11)	2 (6)	47 (16)
Placenta	0 (0)	3 (4)	0 (0)	1 (3)	4 (1)
Amnion	0 (0)	1 (1)	1 (1)	1 (3)	3 (1)
Synovial	0 (0)	2 (3)	4 (3)	0 (0)	6 (2)
Peripheral blood	0 (0)	0 (0)	3 (2)	0 (0)	3 (1)
Other	4 (10)	1 (1)	26 (17)	1 (3)	32 (11)
Functional definition MSC stromal versus stem reported	2 (5)	5 (7)	36 (23)	3 (8)	46 (15)
Self-renewal assay	2 (100)	5 (100)	25 (69)	3 (100)	35 (76)
Multilineage differentiation	0 (0)	0 (0)	20 (56)	0 (0)	20 (44)
Functional assays reported	4 (10)	11 (14)	43 (27)	5 (13)	63 (20)
Quantitative RNA analysis	0 (0)	4 (36)	28 (65)	4 (80)	36 (57)
Mixed lymphocyte reaction	2 (50)	1 (9)	3 (7)	0 (0)	6 (10)
MSC secretome analysis	0 (0)	2 (18)	14 (33)	0 (0)	16 (25)
Migration assay	0 (0)	8 (73)	7 (16)	4 (80)	19 (30)
Other	2 (50)	2 (18)	2 (5)	0 (0)	6 (10)

Data presented as *n* (%).

For in vitro differentiation assay, studies can report none, 1, 2, or tri-lineage. The data are reported in this table as the count number for each lineage differentiation.

Tissue source: some studies reported the use of different tissue sources, the description of the tissue source is shown as count number and therefore can be superior to the number of studies reporting tissue source.

“Other” tissue source details:

Clinical: menstrual blood (2), endometrial tissue (1), gingival connective tissue (1).

Animal: cardiac tissue (1).

Biology: abdominal aortic aneurysm wall (1), amniotic fluid (2), bone fracture site (1), dental apical papilla tissue (1), dental follicle (1), dental pulp (1), dermis (3), endometrial tissue (1), gingival tissue (2), hair follicle (1), peri-cardiac fat (1), liver (1), peri-tumor normal tissue (1), olfactory mucosa (1), Teeth (1), tongue epithelium (1), tonsil (3), femoral marrow fat (1), coronary corium (1), alveolar bone (1).

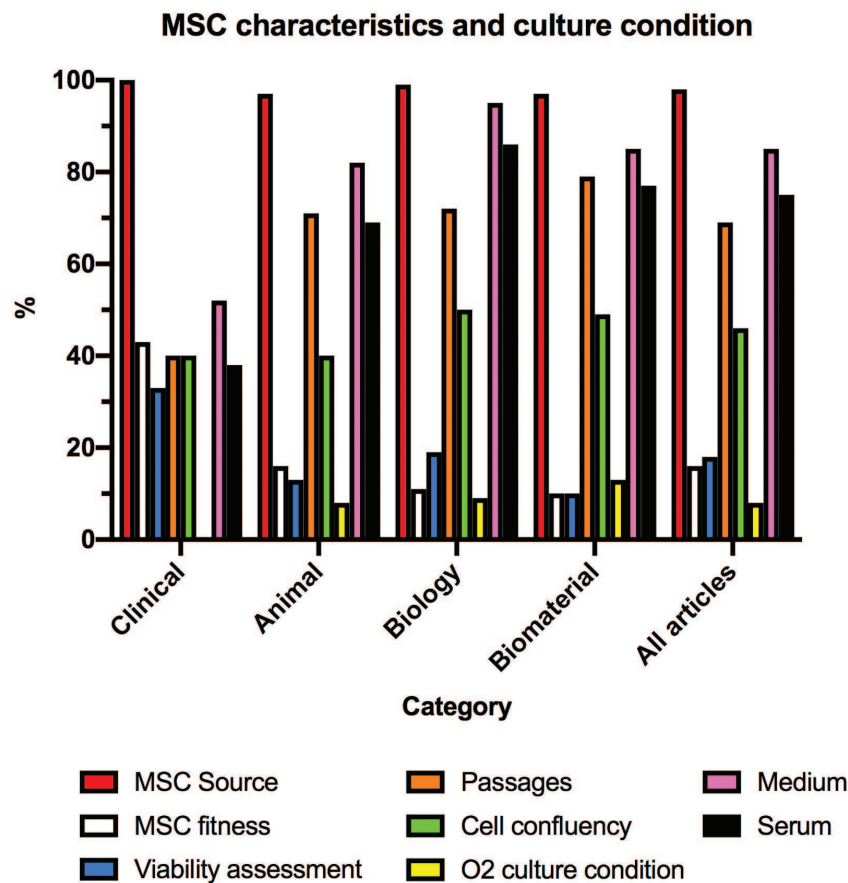
Biomaterial: dental pulp (1).

reported and a >600-fold difference in the studies reporting total cells infused. In addition to these variations, we found that the quality and the rigor in reporting were inconsistent.

Driven by concerns about the inconsistent characterization of MSC as well as different cell manufacturing protocols, the ISCT issued a statement article for minimal criteria for defining MSC in 2006.<sup>5</sup> In this statement, the ISCT committee supported the use of the recommended designation “Multipotent mesenchymal stromal cells” for MSC<sup>24</sup> and defined 3 minimal criteria to describe MSC: “(i) adherence

to plastic; (ii) specific surface antigen expression; and (iii) multipotent differentiation potential.” This statement was updated in 2019, where the committee stressed again the importance of the cell nomenclature and the need for additional criteria to report such as tissue of origin and functional assays to better characterize these cells.<sup>6</sup>

Our results show that the uptake for the ISCT definition is inconsistent among a contemporary sample of articles. Most authors still use the term of “Mesenchymal Stem Cell” to describe MSC, even without providing any evidence



**Figure 3.** Reported mesenchymal stromal cells (MSC) characteristics and culture condition. For MSC characteristics, the following items were assessed: MSC source (eg, patient, donor, commercial), MSC fitness (fresh or cryopreserved MSC), viability assessment prior to MSC use. For culture condition, the following items were assessed: number of cell passages prior to cell use/administration, cell confluence before cell harvest, oxygen ( $O_2$ ) condition for culture (5% vs. 21% of oxygen), type of medium, and serum use.

of the stemness of their cells. Initial minimal criteria from the 2006 statement article, plastic adherence, cell markers, and differentiation assay were inconsistently reported 36%, 53%, and 40%, respectively, of our sample of studies. Most of the articles reporting the use of cell markers to describe their cell population did not provide any information on the flow cytometry cutoff used to define positive and negative markers. In a post hoc test, we examined the association between journal impact factor and level of completeness in reporting the minimal criteria to define MSC from the ISCT. We found no correlation between the level of completeness and the journal impact factor which is consistent with the systematic review from Saginur et al, where they described little to no association between journal impact factor and study methodological quality.<sup>25</sup> In addition, RCT did not show a better completeness in reporting minimal criteria to define MSC compared to other clinical studies. Tissue source was the most reported ISCT criteria. The broad variety of tissue sources reported echoes with some authors' concern that, considering cell markers are nonspecific and artifacts and misinterpretation are frequent in differentiation assays, so-called MSCs can be isolated from any kind of tissue.<sup>26</sup> Another concern is that, in our study, only 20% of all articles (10% of the clinical studies) reported using a functional assay to describe MSCs' potency and properties. These functional potency assays seem to be critical to better characterize MSC and provide a prediction of these MSCs' effectiveness in clinical settings

(depending on the disease, patient demographics),<sup>4</sup> and are required for FDA biologics license application.<sup>7</sup>

We also found that critical information for MSC descriptions (eg, viability assessment and MSC fitness) and culture conditions (eg, oxygen level) were often missing. Among reported items, we found a wide variety of reported culture parameters such as number of passages and cell confluence in the different study categories. The importance of culture conditions in the field of MSC is well recognized as these conditions can dramatically change MSCs' phenotype.<sup>27</sup> MSCs may exhibit different functional properties depending on how they are produced, handled, and administered. In order to enhance reproducibility and transparency in MSC research, it is critical to report these important culture parameters as well as MSC viability and fitness.

In addition to these findings, we also showed in our sample of clinical studies using MSC an important variation in the MSC dose regimen (dose range was  $10^5$  to  $10^7$  cells/kg) and a number of MSC doses (range from 1 to 24 doses). Although MSC administration protocols (administration route, dose, number of doses) were overall well reported, other important information about the cell product used was absent. For instance, only 33% of the clinical articles reported a MSC viability assay at some point in their protocol and most of them did not provide the results of this assay. Administration of senescent or apoptotic cells can mitigate the results in terms of efficacy but also raise the concern about safety as these cells secrete or release paracrine

**Table 4.** Reported mesenchymal stromal cells (MSC) characteristics and culture condition.

	Clinical ( <i>n</i> = 42)	Animal ( <i>n</i> = 77)	Biology ( <i>n</i> = 160)	Biomaterial ( <i>n</i> = 39)	All articles ( <i>n</i> = 318)
MSC source reported	42 (100)	75 (97)	158 (99)	38 (97)	313 (98)
Patient/donor	41 (98)	34 (45)	113 (72)	24 (63)	212 (68)
Animal	N/A	37 (49)	51 (32)	13 (34)	101 (32)
Commercial	3 (7)	10 (13)	37 (23)	10 (26)	60 (19)
Compatibility reported	38 (90)	53 (69)	N/A	N/A	91 (76)
Autologous	27 (71)	0 (0)			27 (30)
Matched allogenic	3 (8)	1 (2)			4 (4)
Unmatched allogenic	8 (21)	19 (36)			27 (30)
Xenogenic	N/A	33 (62)			33 (36)
MSC fitness reported	18 (43)	12 (16)	17 (11)	4 (10)	51 (16)
Fresh	4 (22)	0 (0)	3 (18)	0 (0)	7 (14)
Cryopreserved	14 (78)	12 (100)	14 (82)	4 (100)	44 (86)
Viability assessment reported	14 (33)	10 (13)	30 (19)	4 (10)	58 (18)
Number of passages reported	17 (40)	55 (71)	115 (72)	31 (79)	218 (69)
Range ( <i>n</i> passages)	1-7	1-20	1-38	2-14	1-38
Cell confluence reported	17 (40)	31 (40)	80 (50)	19 (49)	147 (46)
Range (%)	70-90	30-100	50-100	60-90	30-100
O <sub>2</sub> culture condition reported	0 (0)	6 (8)	14 (9)	5 (13)	25 (8)
O <sub>2</sub> 5%	0 (0)	2 (33)	4 (29)	0 (0)	6 (24)
O <sub>2</sub> 21%	0 (0)	4 (67)	10 (71)	5 (100)	19 (76)
Culture medium reported	22 (52)	63 (82)	152 (95)	33 (85)	270 (85)
α-MEM	8 (36)	8 (13)	32 (21)	7 (21)	55 (20)
DMEM	1 (5)	23 (37)	48 (32)	13 (39)	85 (31)
DMEM/F12	7 (32)	9 (14)	25 (16)	2 (6)	43 (16)
LG-DMEM	3 (14)	12 (19)	20 (13)	6 (18)	41 (15)
Serum use reported	16 (38)	53 (69)	138 (86)	30 (77)	237 (75)
No serum	1 (6)	1 (1)	0 (0)	0 (0)	2 (1)
FBS	11 (69)	52 (68)	130 (94)	30 (100)	223 (94)

Data presented as *n* (%).

MSC fitness and clinical studies. For cryopreserved MSC, 50% reported frozen/thawed/MSC administration and 50% reported frozen/thawed/cultured/MSC administration.

MSC sources reported.

Some studies reported the use of different sources.

Animal sources—Animal studies: rat (19), mouse (7), rabbit (4), dog (1), pig (1), sheep (1), unknown (4); Biology studies: rat (18), mouse (20), bovine (1), dog (3), goat (1), horse (1), ovine fetus (1), rabbit (1), shrew (1); Biomaterial studies: rat (4), rabbit (3), mouse (3), sheep (1).

Commercial, details on product and company names for clinical and animal studies—Clinical studies: Bionet Corp. (1), Cartistem Medipost (1), Basic Medical Sciences (1); Animal studies: Cyagen Biosciences (3), Lonza (1), hUC-MSC CHA Biotech (1), Shanghai Yiyuan Biotechnology (1), Shanghai Saibaikang Biotechnology (1), Noor Genetics Laboratory of Ahva (1), Kalang Technology (1), RoosterBio (1).

Serum in clinical studies. Four studies reported other serum than FBS: autologous serum, human serum B, HyClone, newborn calf serum.

Abbreviations: DMEM, Dulbecco's modified Eagle's medium; LG-DMEM, low-glucose DMEM; FBS, fetal bovine serum; MEM, minimum essential medium; O<sub>2</sub>, oxygen.



factors which can negatively regulate the host cells. Some authors have reported that the acute inflammation triggered by the dead stem cells could be at the origin of tissue regeneration, more than the cell product itself.<sup>28</sup> Similarly, functional assays were reported in only 10% of the clinical studies. As stated above these functional assays provide insight into the MSC potency and their potential effectiveness in human diseases.

Overall, these findings from clinical studies are in line with a recent meta-analysis reported that MSC administration seems to be safe in humans but also stressed that the study design quality and reporting transparency of the included studies was sub-optimal.<sup>10</sup> It is critical to have an extensive characterization of the MSC product used as well as a detailed cell manufacturing process to ensure reproducibility, comparison, and transparency between clinical trials using MSC. In addition, given indications that some trial participants have been “non-responders” to MSC therapy, unraveling this heterogeneous response to therapy and the development of predictive biomarkers will ultimately rely on the use of well-characterized and “standardized” MSC products.<sup>4</sup>

A key challenge of this scoping review was managing the large amount of preclinical and clinical MSC research being produced to create our evidence map. To address this challenge, we limited the literature search to 3 months, categorized the selected studies according to the research field (in vitro, animal, human), and for data charting, we randomly selected samples for each category with more than 100 included articles. A potential limitation of this scoping review was our choice to exclude non-English articles, meaning we may have missed relevant information published in another language.

## Conclusion

Our study highlighted a broad variability in reporting quality and completeness in both MSC definition and product characterization. This finding is of concern as many authors consider that for MSCs the “process is the product,” stressing the importance of limiting the sources of variability in the resulting cell product by clearly defining the cell (source, functional assays) and culture condition.<sup>27</sup> In the light of the new evidence provided by our study, it is not surprising that there have been calls to “clearing-up the stem cells mess.”<sup>3</sup> Therefore, we propose to develop a research protocol combining a rigorous consensus development method (modified Delphi method)<sup>16</sup> to address the lack of consensus definition for MSC and to provide reporting guidelines for clinical studies using MSC.<sup>15</sup> Equally important are strategies to support its dissemination and implementation. A science-based approach such as “integrated knowledge translation”<sup>29</sup> may help by engaging knowledge users in the Delphi process since inception and developing a tailored end of project knowledge translation plan to support dissemination and implementation of the Delphi results. Results of this scoping review will be used to generate our initial Delphi survey and to identify potential participants among corresponding authors of the selected articles to contribute to the development of an internationally accepted consensus definition of MSCs.

## Acknowledgments

The authors thank Alexandra Davis for the Peer Review of Electronic Search Strategy (PRESS).

## Funding

This study was funded by a Canadian Stem Cell Network Translation and Society Team Grant 2020. M.M.L. is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and holds a University of Ottawa Junior Research Chair in Innovative Translational Research. B.T. holds a University of Ottawa Partnership Research Chair in Regenerative Medicine.

## Conflict of Interest

The authors declared no potential conflicts of interest.

## Author Contributions

L.R.: conceptualization, methodology, investigation, formal analysis, and writing—original draft preparation; M.P.: methodology, investigation, and formal analysis; R.M., K.N.: methodology and investigation; R.S.: methodology; M.M.L., K.D.C., B.T.: conceptualization, funding acquisition, methodology, supervision, and writing—review and editing.

## Data Availability

No new data were generated or analyzed in support of this research.

## Supplementary Material

Supplementary material is available at *Stem Cells Translational Medicine* online.

## References

- Lazarus HM, Haynesworth SE, Gerson SL, Rosenthal NS, Caplan AI. Ex vivo expansion and subsequent infusion of human bone marrow-derived stromal progenitor cells (mesenchymal progenitor cells): implications for therapeutic use. *Bone Marrow Transplant.* 1995;16(4):557-564.
- ClinicalTrials.gov. Accessed April 10, 2021. <https://clinicaltrials.gov/ct2/results?recrs=&cond=&term=mesechymal+AND+%28Stromal+OR+stem%29+AND+cell&cntry=&state=&city=&dist=>
- Sipp D, Robey PG, Turner L. Clear up this stem-cell mess. *Nature.* 2018;561(7724):455-457.
- Galipeau J, Sensébé L. Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. *Cell Stem Cell.* 2018;22(6):824-833.
- Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy.* 2006;8(4):315-317.
- Viswanathan S, Shi Y, Galipeau J, et al. Mesenchymal stem versus stromal cells: International Society for Cell & Gene Therapy (ISCT®) Mesenchymal Stromal Cell committee position statement on nomenclature. *Cytotherapy.* 2019;21(10):1019-1024.
- Mendicino M, Bailey AM, Wonnacott K, Puri RK, Bauer SR. MSC-based product characterization for clinical trials: an FDA perspective. *Cell Stem Cell.* 2014;14(2):141-145.
- Pierro M, Thébaud B, Soll R. Mesenchymal stem cells for the prevention and treatment of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev.* 2017;11:CD011932.
- Lalu MM, McIntyre L, Pugliese C, et al.; Canadian Critical Care Trials Group. Safety of cell therapy with mesenchymal stromal cells (SafeCell): a systematic review and meta-analysis of clinical trials. *PLoS One.* 2012;7(10):e47559.

10. Thompson M, Mei SHJ, Wolfe D, et al. Cell therapy with intravascular administration of mesenchymal stromal cells continues to appear safe: an updated systematic review and meta-analysis. *EClinicalMedicine*. 2020;19:100249.
11. Augustine S, Avey MT, Harrison B, et al. Mesenchymal stromal cell therapy in bronchopulmonary dysplasia: systematic review and meta-analysis of preclinical studies. *Stem Cells Transl Med*. 2017;6(12):2079-2093.
12. Avey MT, Moher D, Sullivan KJ, et al.; Canadian Critical Care Translational Biology Group. The devil is in the details: incomplete reporting in preclinical animal research. *PLoS One*. 2016;11(11):e0166733.
13. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8:19-32.
14. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010;5:69.
15. Renesme L, Pierro M, Shorr R, Cobey K, Lalu M, Bernard T. A rational approach to clearing up the stem-cell mess : a modified Delphi protocol. 2020. Accessed April 2, 2021. <https://osf.io/3dsqx/>.
16. Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess*. 1998;2(3):i-iv, 1.
17. Peters M, Godfrey C, McInerney P, Munn Z, Trico A, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. JBI; 2020. <https://doi.org/10.46658/JBIMES-20-12>
18. Open Science Framework OSF. Accessed April 10, 2021. <https://osf.io/>
19. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473.
20. Lalu MM, Moher D, Marshall J, et al.; Canadian Critical Care Translational Biology Group. Efficacy and safety of mesenchymal stromal cells in preclinical models of acute lung injury: a systematic review protocol. *Syst Rev*. 2014;3:48.
21. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol*. 2016;75:40-46.
22. DistillerSR. Systematic review and literature review software by evidence partners. Accessed April 2, 2021. <https://www.evidencepartners.com/products/distillers-systematic-review-software/>.
23. Galipeau J, Krampere M, Barrett J, et al. International Society for Cellular Therapy perspective on immune functional assays for mesenchymal stromal cells as potency release criterion for advanced phase clinical trials. *Cytotherapy*. 2016;18(2):151-159.
24. Horwitz EM, Le Blanc K, Dominici M, et al.; International Society for Cellular Therapy. Clarification of the nomenclature for MSC: the International Society for Cellular Therapy position statement. *Cytotherapy*. 2005;7(5):393-395.
25. Saginur M, Fergusson D, Zhang T, et al. Journal impact factor, trial effect size, and methodological quality appear scantily related: a systematic review and meta-analysis. *Syst Rev*. 2020;9(1):53.
26. Robey P. "Mesenchymal stem cells": fact or fiction, and implications in their therapeutic use. *F1000Res*. 2017;6. <https://doi.org/10.12688/f1000research.10955.1>
27. Pittenger MF, Discher DE, Péault BM, Phinney DG, Hare JM, Caplan AI. Mesenchymal stem cell perspective: cell biology to clinical progress. *NPJ Regen Med*. 2019;4:22.
28. Vagnozzi RJ, Maillat M, Sargent MA, et al. An acute immune response underlies the benefit of cardiac stem cell therapy. *Nature*. 2020;577(7790):405-409.
29. Kothari A, McCutcheon C, Graham ID. Defining integrated knowledge translation and moving forward: a response to recent commentaries. *Int J Health Policy Manag*. 2017;6(5):299-300.

# General discussion and conclusions

The clinical translation gaps, bench-to-bedside, and between clinical studies to everyday clinical practice, represent the main obstacles in translating the promising preclinical discoveries to the general population. These clinical gaps concern all fields of biomedical research and medicine and are particularly critical in case of the rare diseases, such as congenital or prematurity-related lung diseases. Translational research was created to address these gaps by developing solutions to speed up translation of breakthrough discoveries to the patient's bedside, to enhance the success rate of promising drugs in clinical trials, and ensure proper dissemination to the general public, policy makers, and regulatory and funding agencies.

In this thesis, we presented three different projects related to neonatal lung research which support the main focus of translational research: 1) a multidisciplinary approach, 2) the use of complex and holistic models including human tissue derived models, and 3) knowledge translation.

Nowadays, the increased complexity of biomedical research protocols renders multidisciplinary collaboration mandatory, with inclusion of individuals with different but complementary skillsets and perspectives. A multidisciplinary research team ensures that the critical determinants for successful biomedical research are addressed: scientific rigour of study design and data analysis, clinical relevance of initial hypothesis, study endpoints



and results interpretation, and strategies to disseminate and implement study findings. All three projects presented here rely on a **multidisciplinary collaboration** to enhance the project's overall quality and findings. The snRNAseq and PCLS projects used human fetal lungs, therefore the quality and viability of these samples were of the highest importance. To ensure optimal tissue quality, our protocol required a close collaboration and teamwork between the obstetrical team (patient recruitment and coordination for abortion process), pathologists (careful dissection and harvest of the lungs), and lab team (coordination, transport and process of the samples within few hours after delivery). In addition, the snRNAseq project further required collaboration with bioinformaticians for the data pre-processing and downstream analyses.

The Delphi project is another good example of how a multidisciplinary approach is critical for project's success. Previous attempts to provide minimal criteria to define MSC by the ISCT failed to ensure proper intake of this definition. Despite the fact that the ISCT project was led by world-renowned experts in the field of MSC research, major pitfalls in the consensus method and in the dissemination and implementation processes were not anticipated. The definition and characterization of MSC are a sensitive and controversial issue, involving many different fields of research, from developmental biology to regulatory agencies and policy makers. The ISCT definition relied on group decision-making process which usually fails to bring to the table people with divergent opinions. Group decision-making presents several limitations when it comes to generate consensus, especially on controversial topics, whereas the Delphi method is a more rigorous and inclusive method, which has been previously successfully used to reach consensus on highly contentious issues. In addition, the ISCT proposal lacked a clear knowledge translation plan. Therefore, we identified the need for our Delphi project to bring onboard both MSC content experts

and experts in methodology and implementation science. We developed our Delphi project in close collaboration with members from the Ottawa Method Centre with specific expertise in consensus-building methods, guidelines development, knowledge translation, and implementation science. They provided guidance for each step of the project, from generating current knowledge synthesis (scoping review) and conducting the Delphi survey, to the development of knowledge translation strategies.

Translational research supports the use of **complex and holistic models** including **human tissue-derived models** to enhance clinical relevance. The use of human tissue, both healthy and diseased, is extremely valuable to bridge preclinical and clinical research. In this thesis, we presented two studies using human fetal lungs obtained from abortions.

Using the human fetal PCLS, we provided a proof of concept of human lung parenchyma transduction of a promising AAV vector for SP-B deficiency gene therapy. The dose-dependent effect, colocalization of the AAV with EPCAM<sup>+</sup> cells and the lack of severe adverse effects strongly suggest a major potential for this AAV vector in term of clinical translation. Preclinical proof of concept studies on primary human tissue are of tremendous importance to assess the clinical potential of a therapy, especially in rare diseases such as SP-B deficiency. SP-B deficiency has an estimated incidence of 1 in 1 million live births. Conducting clinical trials for diseases with such a rare incidence is difficult and complex, therefore optimizing preclinical data for potential therapies with complex human tissue-derived models is critical. PCLS provide a good model to study viral vector transduction as they provide an organotypic culture of a human lung, where complex cell composition and

relation, as well as important features of airway epithelium such as cell polarity are preserved.

In the human developing lung project, we performed an snRNAseq on nine samples from nine human fetal lungs. To our knowledge, this is the first reported unbiased transcriptome analysis at a single-cell resolution of human fetal lungs with gestational ages spanning through the pseudoglandular and the canalicular stages (14 to 19 weeks of gestation). Our findings provide valuable insights on normal *in utero* human lung development during this period. Description of cell composition, gene expression patterns across gestational ages for each cluster, as well as signaling pathways involved in cellular crosstalk, will help to extend our knowledge on human developing lung. Identification of critical biological pathways and underlying molecular mechanisms can be used to generate clinically relevant research hypothesis, develop new animal models of impaired lung development, validate surrogate *ex-vivo* human developing lung model (HLO), and develop new therapies targeting these pathways. In addition, single-cell transcriptomics provides a more holistic and comprehensive picture of the cellular composition and interactions within primary human tissue, whereas other *ex-vivo* models to study human lung development usually allow the study of interactions limited to several cell types.

The last important determinant of translation research is **knowledge translation**. In our Delphi project we described a knowledge translation approach called integrated knowledge translation (iKT). This approach stresses the importance of identifying knowledge users, as they have a unique expertise pertaining to the research topic, including knowledge of the context and the potential for implementation. Once identified, these knowledge users

should be included in the project since its inception and participate in every important step to ensure their needs and preferences are considered. This two-way sharing of knowledge between research producers and users is critical to ensure proper implementation. Too many discoveries failed to cross the knowledge translation gap because they were not sufficiently relevant or useful to the knowledge users. Lastly, these knowledge users will support the dissemination and implementation of the findings by participating in the development of tailored messages to specific audience as well as using their communication network to disseminate the findings. Knowledge translation represents a critical, and often underrated, part of translational research. Every translational research protocol should include a dedicated knowledge translation plan. This plan should at least involve knowledge synthesis, identification of the knowledge users, identification of facilitators and potential obstacles for knowledge translation, as well as an end-of-project knowledge translation plan.

As insightful as these 3 projects can be, they also have their limitations. We have shown that human PCLS is a valuable model to study transduction of viral vectors for gene therapy. One limitation of gene therapy using viral vectors is related to the cellular turn-over, requiring further administrations of the viral vector with its cargo. Little is known on viral vectors immunogenicity and how it could affect the efficacy of viral vector readministration. As PCLS only contain resident immune cells, they don't allow the study of systemic immune response and immune cells recruitment. Another limitation is the short duration for PCLS culture, which hinders its capacity to study more chronic changes. Therefore, there is a need

to develop a more complex model to study the effect of AAV vector readministration on human lung parenchyma, such as co-culturing with human PCLS and human leukocytes.

The snRNAseq and PCLS projects shared the main limitation of all experimental studies using primary human tissue: the access to human tissue. This is particularly true in neonatal lung research where we aim at studying lung development at later stages of the pregnancy. Since the World Health Organization has set the limit of viability for human fetuses at 22 weeks of gestation, access to normal human fetal tissue at 22 or more weeks of gestation is extremely rare. One possible solution would be to develop a maturation protocol for human fetal PCLS. This option however has several limitations, such as the short duration for PCLS culture and the need to validate maturation protocol (like for the HLO) with the *in-vivo* counterpart.

The Delphi project has 2 critical elements that can hinder findings uptakes: the lack of participation in the Delphi survey and the dissemination and implementation of the Delphi results. Both of these limitations can be addressed by the iKT approach, as described above. Our core group is composed of representatives from the different fields of MSC research and we have involved this core group in all the project's steps to ensure their participation in the Delphi survey. Another limit of the Delphi method is the absence of in-person interaction or information exchanges, as face-to-face interaction can be useful to identify reasons for any disagreement. To mitigate this issue, we proposed a modified Delphi, where the last round will be conducted in a form of an in-person meeting with the Core group.

Future directions will include the use of snRNAseq to compare normal vs. impaired human lung development and highlight important modifications in cellular composition, crosstalk,

or molecular mechanisms. This will improve our understanding of congenital or acquired lung disease and facilitate the discovery of new potential therapeutic targets. To achieve this goal, we might consider 2 options.

The first option would be to use human fetal PCLS to generate prematurity-related impaired lung development. The PCLS from the same fetus will be randomized in “normal” culture condition, or “disease” culture condition. In the “disease” culture condition, PCLS will be exposed to *ex-utero* insults experienced by preterm infants, such as hyperoxia, inflammation, or combination of both. snRNAseq will be performed on PCLS at the day of slicing and at different timepoints during the culture period. Comparison of transcriptomic data between “healthy” and “diseased” PCLS could provide insights on immature lung reaction against *ex-utero* insults and highlight important changes in cell fate and communication, as well as in signaling pathways to improve our understanding of BPD and discover potential therapeutic targets. The principal advantage of this option relies in the possibility to assess the human lung tissue responses to different environments within PCLS sharing the same genetic and epigenetic background.

The second option would be to create and develop a biobank of human fetal lung tissues, both from healthy fetuses and fetuses with congenital lung disease. These fetal lung tissues could be used for transcriptomic analysis to compare healthy vs. diseased tissues (with additional adjustment on gestational age and gender). The principal advantage of this option is the use of primary human tissue with no further processing or culture. Therefore, the differences observed in the transcriptomic analysis are more likely due to the condition (healthy vs. diseased) rather than an artifact generated by the culture condition.

In conclusion, this thesis highlighted several complex aspects of translational research and the importance of a holistic and multidisciplinary approach to bridge the clinical translation gap. In addition, I presented here different solutions and mitigation strategies to support multidisciplinary collaboration, development of complex preclinical model using primary human lung tissue, and knowledge translation. All together, these projects demonstrated how to support and improve the clinical translation in neonatal lung research.

## References

- “A Practice Guide to Supporting Implementation | NIRN.” n.d. Accessed May 22, 2021. <https://nirn.fpg.unc.edu/resources/practice-guide-supporting-implementation>.
- Akram, Khondoker M., Laura L. Yates, Róisín Mongey, Stephen Rothery, David C. A. Gaboriau, Jeremy Sanderson, Matthew Hind, Mark Griffiths, and Charlotte H. Dean. 2019. “Live Imaging of Alveologenesis in Precision-Cut Lung Slices Reveals Dynamic Epithelial Cell Behaviour.” *Nature Communications* 10 (1): 1178. <https://doi.org/10.1038/s41467-019-09067-3>.
- Alexander, Michael J., G.R. Scott Budinger, and Paul A. Reyfman. 2020. “Breathing Fresh Air into Respiratory Research with Single-Cell RNA Sequencing.” *European Respiratory Review* 29 (156): 200060. <https://doi.org/10.1183/16000617.0060-2020>.
- Alsafadi, Hani N., Claudia A. Staab-Weijnitz, Mareike Lehmann, Michael Lindner, Britta Peschel, Melanie Königshoff, and Darcy E. Wagner. 2017. “An Ex Vivo Model to Induce Early Fibrosis-like Changes in Human Precision-Cut Lung Slices.” *American Journal of Physiology. Lung Cellular and Molecular Physiology* 312 (6): L896–902. <https://doi.org/10.1152/ajplung.00084.2017>.
- Alsafadi, Hani N., Franziska E. Uhl, Ricardo H. Pineda, Kolene E. Bailey, Mauricio Rojas, Darcy E. Wagner, and Melanie Königshoff. 2020. “Applications and Approaches for Three-Dimensional Precision-Cut Lung Slices. Disease Modeling and Drug Discovery.” *American Journal of Respiratory Cell and Molecular Biology* 62 (6): 681–91. <https://doi.org/10.1165/rcmb.2019-0276TR>.
- Aurora, Megan, and Jason R. Spence. 2016. “HPSC-Derived Lung and Intestinal Organoids as Models of Human Fetal Tissue.” *Developmental Biology* 420 (2): 230–38. <https://doi.org/10.1016/j.ydbio.2016.06.006>.
- Bai, H., Y. Zhao, C. Wang, Z. Wang, J. Wang, H. Liu, Y. Feng, Q. Lin, Z. Li, and H. Liu. n.d. “Enhanced Osseointegration of Three-Dimensional Supramolecular Bioactive Interface through Osteoporotic Microenvironment Regulation.” *Theranostics*, no. 11: 4779–94.
- Bailey, Kolene E., Christopher Pino, Mallory L. Lennon, Anne Lyons, Jeffrey G. Jacot, Steven R. Lammers, Melanie Königshoff, and Chelsea M. Magin. 2020. “Embedding of Precision-Cut Lung Slices in Engineered Hydrogel Biomaterials Supports Extended *Ex Vivo* Culture.” *American Journal of Respiratory Cell and Molecular Biology* 62 (1): 14–22. <https://doi.org/10.1165/rcmb.2019-0232MA>.
- Bakken, Trygve E., Rebecca D. Hodge, Jeremy A. Miller, Zizhen Yao, Thuc Nghi Nguyen, Brian Aevermann, Eliza Barkan, et al. 2018. “Single-Nucleus and Single-Cell Transcriptomes Compared in Matched Cortical Cell Types.” *PloS One* 13 (12): e0209648. <https://doi.org/10.1371/journal.pone.0209648>.
- Barkauskas, Christina E., Mei-I Chung, Bryan Fioret, Xia Gao, Hiroaki Katsura, and Brigid L. M. Hogan. 2017. “Lung Organoids: Current Uses and Future Promise.” *Development* 144 (6): 986–97. <https://doi.org/10.1242/dev.140103>.
- Bauer, Mark S., Laura Damschroder, Hildi Hagedorn, Jeffrey Smith, and Amy M. Kilbourne. 2015. “An Introduction to Implementation Science for the Non-Specialist.” *BMC Psychology* 3 (1): 32. <https://doi.org/10.1186/s40359-015-0089-9>.
- Bauer, Mark S., and JoAnn Kirchner. 2020. “Implementation Science: What Is It and Why Should I Care?” *Psychiatry Research* 283 (January): 112376. <https://doi.org/10.1016/j.psychres.2019.04.025>.



Benjamin, John T., Rebekah J. Smith, Brian A. Halloran, Timothy J. Day, David R. Kelly, and Lawrence S. Prince. 2007. "FGF-10 Is Decreased in Bronchopulmonary Dysplasia and Suppressed by Toll-like Receptor Activation." *American Journal of Physiology-Lung Cellular and Molecular Physiology* 292 (2): L550–58. <https://doi.org/10.1152/ajplung.00329.2006>.

Berger, Jessica, and Vineet Bhandari. 2014. "Animal Models of Bronchopulmonary Dysplasia. The Term Mouse Models." *American Journal of Physiology. Lung Cellular and Molecular Physiology* 307 (12): L936–947. <https://doi.org/10.1152/ajplung.00159.2014>.

Blumberg, Richard S., Bonnie Dittel, David Hafler, Matthias von Herrath, and Frank O. Nestle. 2012. "Unraveling the Autoimmune Translational Research Process Layer by Layer." *Nature Medicine* 18 (1): 35–41. <https://doi.org/10.1038/nm.2632>.

Bourguignon, Chloé, Charlotte Vernisse, Joffrey Mianné, Mathieu Fieldès, Engi Ahmed, Aurélie Petit, Isabelle Vachier, et al. 2020. "Les Organoïdes Pulmonaires." *Médecine/Sciences* 36 (4): 382–88. <https://doi.org/10.1051/medsci/2020056>.

Bower, Danielle V., Hyung-Kook Lee, Rusty Lansford, Kai Zinn, David Warburton, Scott E. Fraser, and Edwin C. Jesudason. 2014. "Airway Branching Has Conserved Needs for Local Parasympathetic Innervation but Not Neurotransmission." *BMC Biology* 12 (November): 92. <https://doi.org/10.1186/s12915-014-0092-2>.

Burri, P H. 1984. "Fetal and Postnatal Development of the Lung." *Annual Review of Physiology* 46 (1): 617–28. <https://doi.org/10.1146/annurev.ph.46.030184.003153>.

Burri, Peter H. 2006. "Structural Aspects of Postnatal Lung Development – Alveolar Formation and Growth." *Neonatology* 89 (4): 313–22. <https://doi.org/10.1159/000092868>.

Butler, Declan. 2008. "Translational Research: Crossing the Valley of Death." *Nature* 453 (7197): 840–42. <https://doi.org/10.1038/453840a>.

Canadian Institutes of Health Research. 2012. "Guide to Knowledge Translation Planning at CIHR: Integrated and End-of-Grant Approaches - CIHR." December 6, 2012. <https://cihr-irsc.gc.ca/e/45321.html>.

Cardoso, Wellington V., and Jining Lü. 2006. "Regulation of Early Lung Morphogenesis: Questions, Facts and Controversies." *Development (Cambridge, England)* 133 (9): 1611–24. <https://doi.org/10.1242/dev.02310>.

Chen, Xi, Sarah A. Teichmann, and Kerstin B. Meyer. 2018. "From Tissues to Cell Types and Back: Single-Cell Gene Expression Analysis of Tissue Architecture." *Annual Review of Biomedical Data Science* 1 (1): 29–51. <https://doi.org/10.1146/annurev-biodatasci-080917-013452>.

Dalkey, Norman Crolee. 1969. *The Delphi Method: An Experimental Study of Group Opinion*. RAND Corporation. [https://www.rand.org/pubs/research\\_memoranda/RM5888.html](https://www.rand.org/pubs/research_memoranda/RM5888.html).

Danopoulos, Soula, Irving Alonso, Matthew E. Thornton, Brendan H. Grubbs, Saverio Bellusci, David Warburton, and Denise Al Alam. 2018. "Human Lung Branching Morphogenesis Is Orchestrated by the Spatiotemporal Distribution of ACTA2, SOX2, and SOX9." *American Journal of Physiology-Lung Cellular and Molecular Physiology* 314 (1): L144–49. <https://doi.org/10.1152/ajplung.00379.2017>.

Danopoulos, Soula, Soumyaroop Bhattacharya, Thomas J. Mariani, and Denise Al Alam. 2020. "Transcriptional Characterisation of Human Lung Cells Identifies Novel Mesenchymal Lineage Markers." *European Respiratory Journal* 55 (1): 1900746. <https://doi.org/10.1183/13993003.00746-2019>.

Danopoulos, Soula, Jessica Shiosaki, and Denise Al Alam. 2019. "FGF Signaling in Lung Development and Disease: Human Versus Mouse." *Frontiers in Genetics* 10 (March): 170. <https://doi.org/10.3389/fgene.2019.00170>.

Davidovich, N., J. Huang, and S. S. Margulies. 2013. "Reproducible Uniform Equibiaxial

Stretch of Precision-Cut Lung Slices.” *American Journal of Physiology-Lung Cellular and Molecular Physiology* 304 (4): L210–20. <https://doi.org/10.1152/ajplung.00224.2012>.

Denisenko, Elena, Belinda B. Guo, Matthew Jones, Rui Hou, Leanne de Kock, Timo Lassmann, Daniel Poppe, et al. 2020. “Systematic Assessment of Tissue Dissociation and Storage Biases in Single-Cell and Single-Nucleus RNA-Seq Workflows.” *Genome Biology* 21 (1): 130. <https://doi.org/10.1186/s13059-020-02048-6>.

Dominici, M., K. Le Blanc, I. Mueller, I. Slaper-Cortenbach, Fc Marini, Ds Krause, Rj Deans, A. Keating, Dj Prockop, and Em Horwitz. 2006. “Minimal Criteria for Defining Multipotent Mesenchymal Stromal Cells. The International Society for Cellular Therapy Position Statement.” *Cytotherapy* 8 (4): 315–17. <https://doi.org/10.1080/14653240600855905>.

Dowden, Helen, and Jamie Munro. 2019. “Trends in Clinical Success Rates and Therapeutic Focus.” *Nature Reviews Drug Discovery* 18 (7): 495–96. <https://doi.org/10.1038/d41573-019-00074-z>.

Du, Yina, Minzhe Guo, Jeffrey A. Whitsett, and Yan Xu. 2015. “‘LungGENS’: A Web-Based Tool for Mapping Single-Cell Gene Expression in the Developing Lung.” *Thorax* 70 (11): 1092–94. <https://doi.org/10.1136/thoraxjnl-2015-207035>.

Du, Yina, Joseph A. Kitzmiller, Anusha Sridharan, Anne K. Perl, James P. Bridges, Ravi S. Misra, Gloria S. Pryhuber, et al. 2017. “Lung Gene Expression Analysis (LGEA): An Integrative Web Portal for Comprehensive Gene Expression Data Analysis in Lung Development.” *Thorax* 72 (5): 481–84. <https://doi.org/10.1136/thoraxjnl-2016-209598>.

Du, Yina, Weichen Ouyang, Joseph A. Kitzmiller, Minzhe Guo, Shuyang Zhao, Jeffrey A. Whitsett, and Yan Xu. 2021. “Lung Gene Expression Analysis Web Portal Version 3: Lung-at-a-Glance.” *American Journal of Respiratory Cell and Molecular Biology* 64 (1): 146–49. <https://doi.org/10.1165/rcmb.2020-0308LE>.

Dye, Briana R., Alyssa J. Miller, and Jason R. Spence. 2016. “How to Grow a Lung: Applying Principles of Developmental Biology to Generate Lung Lineages from Human Pluripotent Stem Cells.” *Current Pathobiology Reports* 4 (2): 47–57. <https://doi.org/10.1007/s40139-016-0102-x>.

Ernst, Linda M., Eduardo D. Ruchelli, and Dale S. Huff, eds. 2011. *Color Atlas of Fetal and Neonatal Histology*. New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4614-0019-6>.

Evans, Kelly V., and Joo-Hyeon Lee. 2020. “Alveolar Wars: The Rise of in Vitro Models to Understand Human Lung Alveolar Maintenance, Regeneration, and Disease.” *Stem Cells Translational Medicine* 9 (8): 867–81. <https://doi.org/10.1002/sctm.19-0433>.

Fernandez-Moure, Joseph S. 2016. “Lost in Translation: The Gap in Scientific Advancements and Clinical Application.” *Frontiers in Bioengineering and Biotechnology* 4 (June). <https://doi.org/10.3389/fbioe.2016.00043>.

Fisher, Robyn L., Mary S. Smith, Steven J. Hasal, Katherine S. Hasal, A. Jay Gandolfi, and Klaus Brendel. 1994. “The Use of Human Lung Slices in Toxicology.” *Human & Experimental Toxicology* 13 (7): 466–71. <https://doi.org/10.1177/096032719401300703>.

Franks, Teri J., Thomas V. Colby, William D. Travis, Rubin M. Tuder, Herbert Y. Reynolds, Arnold R. Brody, Wellington V. Cardoso, et al. 2008. “Resident Cellular Components of the Human Lung: Current Knowledge and Goals for Research on Cell Phenotyping and Function.” *Proceedings of the American Thoracic Society* 5 (7): 763–66. <https://doi.org/10.1513/pats.200803-025HR>.

Friedenstein, A. J., J. F. Gorskaja, and N. N. Kulagina. 1976. “Fibroblast Precursors in Normal and Irradiated Mouse Hematopoietic Organs.” *Experimental Hematology* 4 (5): 267–74.

Galipeau, Jacques, and Luc Sensébé. 2018. “Mesenchymal Stromal Cells: Clinical

Challenges and Therapeutic Opportunities.” *Cell Stem Cell* 22 (6): 824–33.  
<https://doi.org/10.1016/j.stem.2018.05.004>.

Galipeau, James, Kelly D. Cobey, Virginia Barbour, Patricia Baskin, Sally Bell-Syer, Jonathan Deeks, Paul Garner, et al. 2017. “An International Survey and Modified Delphi Process Revealed Editors’ Perceptions, Training Needs, and Ratings of Competency-Related Statements for the Development of Core Competencies for Scientific Editors of Biomedical Journals.” *F1000Research* 6 (September). <https://doi.org/10.12688/f1000research.12400.1>.

Garner, Joseph P, Brianna N Gaskill, Elin M Weber, Jamie Ahloy-Dallaire, and Kathleen R Pritchett-Corning. 2017. “Introducing Therioepistemology: The Study of How Knowledge Is Gained from Animal Research.” *Lab Animal* 46 (4): 103–13.  
<https://doi.org/10.1038/labanimal.1224>.

Giusto, Kiersten, Heather Wanczyk, Todd Jensen, and Christine Finck. 2021. “Hyperoxia-Induced Bronchopulmonary Dysplasia: Better Models for Better Therapies.” *Disease Models & Mechanisms* 14 (2). <https://doi.org/10.1242/dmm.047753>.

Government of Canada, Canadian Institutes of Health Research. 2012. “Guide to Knowledge Translation Planning at CIHR: Integrated and End-of-Grant Approaches - CIHR.” December 6, 2012. <https://cihr-irsc.gc.ca/e/45321.html>.

Grimshaw, Jeremy M, Martin P Eccles, John N Lavis, Sophie J Hill, and Janet E Squires. 2012. “Knowledge Translation of Research Findings.” *Implementation Science* 7 (1): 50.  
<https://doi.org/10.1186/1748-5908-7-50>.

Grudniewicz, Agnes, David Moher, Kelly D. Cobey, Gregory L. Bryson, Samantha Cukier, Kristiann Allen, Clare Arden, et al. 2019. “Predatory Journals: No Definition, No Defence.” *Nature* 576 (7786): 210–12. <https://doi.org/10.1038/d41586-019-03759-y>.

Haque, Ashraful, Jessica Engel, Sarah A. Teichmann, and Tapio Lönnberg. 2017. “A Practical Guide to Single-Cell RNA-Sequencing for Biomedical Research and Clinical Applications.” *Genome Medicine* 9 (1): 75. <https://doi.org/10.1186/s13073-017-0467-4>.

Hay, Michael, David W. Thomas, John L. Craighead, Celia Economides, and Jesse Rosenthal. 2014. “Clinical Development Success Rates for Investigational Drugs.” *Nature Biotechnology* 32 (1): 40–51. <https://doi.org/10.1038/nbt.2786>.

Henjakovic, M., K. Sewald, S. Switalla, D. Kaiser, M. Müller, T.Z. Veres, C. Martin, S. Uhlig, N. Krug, and A. Braun. 2008. “Ex Vivo Testing of Immune Responses in Precision-Cut Lung Slices.” *Toxicology and Applied Pharmacology* 231 (1): 68–76.  
<https://doi.org/10.1016/j.taap.2008.04.003>.

Herriges, Michael, and Edward E. Morrisey. 2014. “Lung Development: Orchestrating the Generation and Regeneration of a Complex Organ.” *Development (Cambridge, England)* 141 (3): 502–13. <https://doi.org/10.1242/dev.098186>.

Herring, Matt J., Lei F. Putney, Gregory Wyatt, Walter E. Finkbeiner, and Dallas M. Hyde. 2014. “Growth of Alveoli during Postnatal Development in Humans Based on Stereological Estimation.” *American Journal of Physiology-Lung Cellular and Molecular Physiology* 307 (4): L338–44. <https://doi.org/10.1152/ajplung.00094.2014>.

Hsu, C-C, and BA Sandford. 2007. “The Delphi Technique: Making Sense of Consensus.” *Practical Assessment, Research, and Evaluation*: 12.

Huh, Dongeun, Benjamin D. Matthews, Akiko Mammoto, Martín Montoya-Zavala, Hong Yuan Hsin, and Donald E. Ingber. 2010. “Reconstituting Organ-Level Lung Functions on a Chip.” *Science (New York, N.Y.)* 328 (5986): 1662–68.  
<https://doi.org/10.1126/science.1188302>.

Hurskainen, Maria, Ivana Mižíková, David P. Cook, Noora Andersson, Chanèle Cyr-Depauw, Flore Lesage, Emmi Helle, et al. 2021. “Single Cell Transcriptomic Analysis of Murine Lung Development on Hyperoxia-Induced Damage.” *Nature Communications* 12 (1): 1565. <https://doi.org/10.1038/s41467-021-21865-2>.

Jimenez-Puerta, Gonzalo José, Juan Antonio Marchal, Elena López-Ruiz, and Patricia Gálvez-Martín. 2020. “Role of Mesenchymal Stromal Cells as Therapeutic Agents: Potential Mechanisms of Action and Implications in Their Clinical Use.” *Journal of Clinical Medicine* 9 (2): E445. <https://doi.org/10.3390/jcm9020445>.

Kang, Martin H., Laura P. van Lieshout, Liqun Xu, Jakob M. Domm, Arul Vadivel, Laurent Renesme, Christian Mühlfeld, et al. 2020. “A Lung Tropic AAV Vector Improves Survival in a Mouse Model of Surfactant B Deficiency.” *Nature Communications* 11 (1): 3929. <https://doi.org/10.1038/s41467-020-17577-8>.

Kho, Alvin T., Soumyaroop Bhattacharya, Kelan G. Tantisira, Vincent J. Carey, Roger Gaedigk, J. Steven Leeder, Isaac S. Kohane, Scott T. Weiss, and Thomas J. Mariani. 2010. “Transcriptomic Analysis of Human Lung Development.” *American Journal of Respiratory and Critical Care Medicine* 181 (1): 54–63. <https://doi.org/10.1164/rccm.200907-1063OC>.

Kina, Yelda Pakize, Ali Khadim, Werner Seeger, and Elie El Agha. 2021. “The Lung Vasculature: A Driver or Passenger in Lung Branching Morphogenesis?” *Frontiers in Cell and Developmental Biology* 8 (January): 623868. <https://doi.org/10.3389/fcell.2020.623868>.

Klouda, Timothy, Hyunbum Kim, Jiwon Kim, Gary Visner, and Ke Yuan. 2021. “Precision Cut Lung Slices as an Efficient Tool for Ex Vivo Pulmonary Vessel Structure and Contractility Studies.” *Journal of Visualized Experiments: JoVE*, no. 171 (May). <https://doi.org/10.3791/62392>.

Koenitzer, Jeffrey R., Haojia Wu, Jeffrey J. Atkinson, Steven L. Brody, and Benjamin D. Humphreys. 2020. “Single-Nucleus RNA-Sequencing Profiling of Mouse Lung. Reduced Dissociation Bias and Improved Rare Cell-Type Detection Compared with Single-Cell RNA Sequencing.” *American Journal of Respiratory Cell and Molecular Biology* 63 (6): 739–47. <https://doi.org/10.1165/rcmb.2020-0095MA>.

Kotecha, S. 2000. “Lung Growth for Beginners.” *Paediatric Respiratory Reviews* 1 (4): 308–13. <https://doi.org/10.1053/prrv.2000.0069>.

Kothari, Anita, Chris McCutcheon, and Ian D. Graham. 2017. “Defining Integrated Knowledge Translation and Moving Forward: A Response to Recent Commentaries.” *International Journal of Health Policy and Management* 6 (5): 299–300. <https://doi.org/10.15171/ijhpm.2017.15>.

Krunkosky, Thomas M., Jarrat L. Jordan, Emily Chambers, and Duncan C. Krause. 2007. “Mycoplasma Pneumoniae Host-Pathogen Studies in an Air-Liquid Culture of Differentiated Human Airway Epithelial Cells.” *Microbial Pathogenesis* 42 (2–3): 98–103. <https://doi.org/10.1016/j.micpath.2006.11.003>.

Lake, Blue B., Simone Codeluppi, Yun C. Yung, Derek Gao, Jerold Chun, Peter V. Kharchenko, Sten Linnarsson, and Kun Zhang. 2017. “A Comparative Strategy for Single-Nucleus and Single-Cell Transcriptomes Confirms Accuracy in Predicted Cell-Type Expression from Nuclear RNA.” *Scientific Reports* 7 (1): 6031. <https://doi.org/10.1038/s41598-017-04426-w>.

Lalu, Manoj M., Lauralyn McIntyre, Christina Pugliese, Dean Fergusson, Brent W. Winston, John C. Marshall, John Granton, Duncan J. Stewart, and Canadian Critical Care Trials Group. 2012. “Safety of Cell Therapy with Mesenchymal Stromal Cells (SafeCell): A Systematic Review and Meta-Analysis of Clinical Trials.” *PloS One* 7 (10): e47559. <https://doi.org/10.1371/journal.pone.0047559>.

Lazarus, H. M., S. E. Haynesworth, S. L. Gerson, N. S. Rosenthal, and A. I. Caplan. 1995. “Ex Vivo Expansion and Subsequent Infusion of Human Bone Marrow-Derived Stromal Progenitor Cells (Mesenchymal Progenitor Cells): Implications for Therapeutic Use.” *Bone Marrow Transplantation* 16 (4): 557–64.

Ledford, Heidi. 2008. “Translational Research: The Full Cycle.” *Nature* 453 (7197): 843–45. <https://doi.org/10.1038/453843a>.

Liberati, Teresa A, Michelle R Randle, and Linda A Toth. 2010. “In Vitro Lung Slices: A Powerful Approach for Assessment of Lung Pathophysiology.” *Expert Review of Molecular Diagnostics* 10 (4): 501–8. <https://doi.org/10.1586/erm.10.21>.

Liu, Guanghui, Catherine Betts, Danen M. Cunoosamy, Per M. Åberg, Jorrit J. Hornberg, Kinga Balogh Sivars, and Taylor S. Cohen. 2019. “Use of Precision Cut Lung Slices as a Translational Model for the Study of Lung Biology.” *Respiratory Research* 20 (1): 162. <https://doi.org/10.1186/s12931-019-1131-x>.

Liu, Xiaoming, Ziyang Yan, Meihui Luo, and John F. Engelhardt. 2006. “Species-Specific Differences in Mouse and Human Airway Epithelial Biology of Recombinant Adeno-Associated Virus Transduction.” *American Journal of Respiratory Cell and Molecular Biology* 34 (1): 56–64. <https://doi.org/10.1165/rcmb.2005-0189OC>.

Luecken, Malte D, and Fabian J Theis. 2019. “Current Best Practices in Single-cell RNA-seq Analysis: A Tutorial.” *Molecular Systems Biology* 15 (6). <https://doi.org/10.15252/msb.20188746>.

Mendicino, Michael, Alexander M. Bailey, Keith Wonnacott, Raj K. Puri, and Steven R. Bauer. 2014. “MSC-Based Product Characterization for Clinical Trials: An FDA Perspective.” *Cell Stem Cell* 14 (2): 141–45. <https://doi.org/10.1016/j.stem.2014.01.013>.

Meyer-Berg, Helena, Lucia Zhou Yang, María Pilar de Lucas, Alberto Zambrano, Stephen C. Hyde, and Deborah R. Gill. 2020. “Identification of AAV Serotypes for Lung Gene Therapy in Human Embryonic Stem Cell-Derived Lung Organoids.” *Stem Cell Research & Therapy* 11 (1): 448. <https://doi.org/10.1186/s13287-020-01950-x>.

Meyerholz, David K., Carlos J. Suarez, Suzanne M. Dintzis, and Charles W. Frevert. 2018. “Respiratory System.” In *Comparative Anatomy and Histology*, 147–62. Elsevier. <https://doi.org/10.1016/B978-0-12-802900-8.00009-9>.

Miller, Alyssa J., David R. Hill, Melinda S. Nagy, Yoshiro Aoki, Briana R. Dye, Alana M. Chin, Sha Huang, et al. 2018. “In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells.” *Stem Cell Reports* 10 (1): 101–19. <https://doi.org/10.1016/j.stemcr.2017.11.012>.

Miller, Alyssa J., and Jason R. Spence. 2017. “In Vitro Models to Study Human Lung Development, Disease and Homeostasis.” *Physiology* 32 (3): 246–60. <https://doi.org/10.1152/physiol.00041.2016>.

Miller, Alyssa J., Qianhui Yu, Michael Czerwinski, Yu-Hwai Tsai, Renee F. Conway, Angeline Wu, Emily M. Holloway, et al. 2020. “In Vitro and In Vivo Development of the Human Airway at Single-Cell Resolution.” *Developmental Cell* 53 (1): 117–128.e6. <https://doi.org/10.1016/j.devcel.2020.01.033>.

Minna, J. D., and A. F. Gazdar. 1996. “Translational Research Comes of Age.” *Nature Medicine* 2 (9): 974–75. <https://doi.org/10.1038/nm0996-974>.

Mižíková, I., and B. Thébaud. 2021. “Looking at the Developing Lung in Single-Cell Resolution.” *American Journal of Physiology-Lung Cellular and Molecular Physiology* 320 (5): L680–87. <https://doi.org/10.1152/ajplung.00385.2020>.

Molina, Samuel A., Brandon Stauffer, Hannah K. Moriarty, Agnes H. Kim, Nael A. McCarty, and Michael Koval. 2015. “Junctional Abnormalities in Human Airway Epithelial Cells Expressing F508del CFTR.” *American Journal of Physiology. Lung Cellular and Molecular Physiology* 309 (5): L475–487. <https://doi.org/10.1152/ajplung.00060.2015>.

Morin, Jean-Paul, Jean-Marc Baste, Arnaud Gay, Clément Crochemore, Cécile Corbière, and Christelle Monteil. 2013. “Precision Cut Lung Slices as an Efficient Tool for in Vitro Lung Physio-Pharmacotoxicology Studies.” *Xenobiotica* 43 (1): 63–72. <https://doi.org/10.3109/00498254.2012.727043>.

Morrissey, Edward E., and Brigid L.M. Hogan. 2010. “Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development.” *Developmental Cell* 18 (1): 8–23.

<https://doi.org/10.1016/j.devcel.2009.12.010>.

Murphy, M. K., N. A. Black, D. L. Lamping, C. M. McKee, C. F. Sanderson, J. Askham, and T. Marteau. 1998. "Consensus Development Methods, and Their Use in Clinical Guideline Development." *Health Technology Assessment (Winchester, England)* 2 (3): i–iv, 1–88.

Nadkarni, Rohan R., Soumeya Abed, and Jonathan S. Draper. 2016. "Organoids as a Model System for Studying Human Lung Development and Disease." *Biochemical and Biophysical Research Communications* 473 (3): 675–82. <https://doi.org/10.1016/j.bbrc.2015.12.091>.

Narayanan, Manjith, John Owers-Bradley, Caroline S. Beardsmore, Marius Mada, Iain Ball, Ruslan Garipov, Kuldeep S. Panesar, et al. 2012. "Alveolarization Continues during Childhood and Adolescence: New Evidence from Helium-3 Magnetic Resonance." *American Journal of Respiratory and Critical Care Medicine* 185 (2): 186–91.

<https://doi.org/10.1164/rccm.201107-1348OC>.

Nardiello, Claudio, Ivana Mižíková, and Rory E. Morty. 2017. "Looking Ahead: Where to next for Animal Models of Bronchopulmonary Dysplasia?" *Cell and Tissue Research* 367 (3): 457–68. <https://doi.org/10.1007/s00441-016-2534-3>.

Neuhaus, Vanessa, Olga Danov, Sebastian Konzok, Helena Obernolte, Susann Dehmel, Peter Braubach, Danny Jonigk, et al. 2018. "Assessment of the Cytotoxic and Immunomodulatory Effects of Substances in Human Precision-Cut Lung Slices." *Journal of Visualized Experiments*, no. 135 (May): 57042. <https://doi.org/10.3791/57042>.

Neuhaus, Vanessa, Dirk Schaudien, Tatiana Golovina, Ulla-Angela Temann, Carolann Thompson, Torsten Lippmann, Claus Bersch, et al. 2017. "Assessment of Long-Term Cultivated Human Precision-Cut Lung Slices as an Ex Vivo System for Evaluation of Chronic Cytotoxicity and Functionality." *Journal of Occupational Medicine and Toxicology* 12 (1): 13. <https://doi.org/10.1186/s12995-017-0158-5>.

Niehof, Monika, Tobias Hildebrandt, Olga Danov, Kirsten Arndt, Jeannette Koschmann, Franziska Dahlmann, Tanja Hansen, and Katherina Sewald. 2017. "RNA Isolation from Precision-Cut Lung Slices (PCLS) from Different Species." *BMC Research Notes* 10 (1): 121. <https://doi.org/10.1186/s13104-017-2447-6>.

Nikolić, Marko Z, Oriol Caritg, Quitz Jeng, Jo-Anne Johnson, Dawei Sun, Kate J Howell, Jane L Brady, et al. 2017. "Human Embryonic Lung Epithelial Tips Are Multipotent Progenitors That Can Be Expanded in Vitro as Long-Term Self-Renewing Organoids." *ELife* 6 (June): e26575. <https://doi.org/10.7554/eLife.26575>.

Nikolić, Marko Z., Dawei Sun, and Emma L. Rawlins. 2018. "Human Lung Development: Recent Progress and New Challenges." *Development* 145 (16): dev163485.

<https://doi.org/10.1242/dev.163485>.

Nolta, Jan A., Jacques Galipeau, and Donald G. Phinney. 2020. "Improving Mesenchymal Stem/Stromal Cell Potency and Survival: Proceedings from the International Society of Cell Therapy (ISCT) MSC Preconference Held in May 2018, Palais Des Congrès de Montréal, Organized by the ISCT MSC Scientific Committee." *Cytotherapy* 22 (3): 123–26.

<https://doi.org/10.1016/j.jcyt.2020.01.004>.

Nowbar, Alexandra N., Michael Mielewicz, Maria Karavassilis, Hakim-Moulay Dehbi, Matthew J. Shun-Shin, Siana Jones, James P. Howard, Graham D. Cole, Darrel P. Francis, and DAMASCENE writing group. 2014. "Discrepancies in Autologous Bone Marrow Stem Cell Trials and Enhancement of Ejection Fraction (DAMASCENE): Weighted Regression and Meta-Analysis." *BMJ (Clinical Research Ed.)* 348 (April): g2688.

<https://doi.org/10.1136/bmj.g2688>.

Okoli, C, and SD Pawlowski. 2004. "The Delphi Method as a Research Tool: An Example, Design Considerations and Applications." *Information & Management* 42 (1): 15–29.

Pieretti, Alberto C, Alwiya M Ahmed, Jesse D Roberts, Jr., and Cassandra M Kelleher. 2013. "A Novel *in Vitro* Model to Study Alveologenesis." *American Journal of Respiratory Cell*

*and Molecular Biology*, September, 130925125950002. <https://doi.org/10.1165/rcmb.2013-0056OC>.

Plasschaert, Lindsey W., Rapolas Žilionis, Rayman Choo-Wing, Virginia Savova, Judith Knehr, Guglielmo Roma, Allon M. Klein, and Aron B. Jaffe. 2018. “A Single-Cell Atlas of the Airway Epithelium Reveals the CFTR-Rich Pulmonary Ionocyte.” *Nature* 560 (7718): 377–81. <https://doi.org/10.1038/s41586-018-0394-6>.

Potter, Andrew S., and S. Steven Potter. 2019. “Dissociation of Tissues for Single-Cell Analysis.” *Methods in Molecular Biology (Clifton, N.J.)* 1926: 55–62. [https://doi.org/10.1007/978-1-4939-9021-4\\_5](https://doi.org/10.1007/978-1-4939-9021-4_5).

Regev, Aviv, Sarah Teichmann, Orit Rozenblatt-Rosen, Michael Stubbington, Kristin Ardlie, Ido Amit, Paola Arlotta, et al. 2018. “The Human Cell Atlas White Paper.” *ArXiv:1810.05192 [q-Bio]*, October. <http://arxiv.org/abs/1810.05192>.

Restifo, Linda L., and Gerald R. Phelan. 2011. “The Cultural Divide: Exploring Communication Barriers between Scientists and Clinicians.” *Disease Models & Mechanisms* 4 (4): 423–26. <https://doi.org/10.1242/dmm.008177>.

Rizk, Mina, Madeline Monaghan, Risa Shorr, Natasha Kekre, Christopher N. Bredeson, and David S. Allan. 2016. “Heterogeneity in Studies of Mesenchymal Stromal Cells to Treat or Prevent Graft-versus-Host Disease: A Scoping Review of the Evidence.” *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation* 22 (8): 1416–23. <https://doi.org/10.1016/j.bbmt.2016.04.010>.

Roberts, Scott F., Martin A. Fischhoff, Stacey A. Sakowski, and Eva L. Feldman. 2012. “Perspective: Transforming Science into Medicine: How Clinician-Scientists Can Build Bridges across Research’s ‘Valley of Death.’” *Academic Medicine: Journal of the Association of American Medical Colleges* 87 (3): 266–70. <https://doi.org/10.1097/ACM.0b013e3182446fa3>.

Robey, Pamela. 2017. “‘Mesenchymal Stem Cells’: Fact or Fiction, and Implications in Their Therapeutic Use.” *F1000Research* 6. <https://doi.org/10.12688/f1000research.10955.1>.

Robinson, Patrick G., Iain R. Murray, Christopher C. West, Ewan B. Goudie, Li Y. Yong, Timothy O. White, and Robert F. LaPrade. 2019. “Reporting of Mesenchymal Stem Cell Preparation Protocols and Composition: A Systematic Review of the Clinical Orthopaedic Literature.” *The American Journal of Sports Medicine* 47 (4): 991–1000. <https://doi.org/10.1177/0363546518758667>.

Rosales Gerpe, María C., Jacob P. van Vloten, Lisa A. Santry, Jondavid de Jong, Robert C. Mould, Adrian Pelin, John C. Bell, Byram W. Bridle, and Sarah K. Wootton. 2018. “Use of Precision-Cut Lung Slices as an Ex Vivo Tool for Evaluating Viruses and Viral Vectors for Gene and Oncolytic Therapy.” *Molecular Therapy - Methods & Clinical Development* 10 (September): 245–56. <https://doi.org/10.1016/j.omtm.2018.07.010>.

Sabroe, Ian, David H. Dockrell, Stefanie N. Vogel, Stephen A. Renshaw, Moira K. B. Whyte, and Steven K. Dower. 2007. “Identifying and Hurdling Obstacles to Translational Research.” *Nature Reviews. Immunology* 7 (1): 77–82. <https://doi.org/10.1038/nri1999>.

Salaets, Thomas, Andre Gie, Bieke Tack, Jan Deprest, and Jaan Toelen. 2017. “Modelling Bronchopulmonary Dysplasia in Animals: Arguments for the Preterm Rabbit Model.” *Current Pharmaceutical Design* 23 (38): 5887–5901. <https://doi.org/10.2174/1381612823666170926123550>.

Salomon, Robert, Dominik Kaczorowski, Fatima Valdes-Mora, Robert E. Nordon, Adrian Neild, Nona Farbehi, Nenad Bartonicek, and David Gallego-Ortega. 2019. “Droplet-Based Single Cell RNAseq Tools: A Practical Guide.” *Lab on a Chip* 19 (10): 1706–27. <https://doi.org/10.1039/C8LC01239C>.

Sanderson, Michael J. 2011. “Exploring Lung Physiology in Health and Disease with Lung Slices.” *Pulmonary Pharmacology & Therapeutics* 24 (5): 452–65.

<https://doi.org/10.1016/j.pupt.2011.05.001>.

Schiller, Herbert B., Daniel T. Montoro, Lukas M. Simon, Emma L. Rawlins, Kerstin B. Meyer, Maximilian Strunz, Felipe A. Vieira Braga, et al. 2019. “The Human Lung Cell Atlas: A High-Resolution Reference Map of the Human Lung in Health and Disease.” *American Journal of Respiratory Cell and Molecular Biology* 61 (1): 31–41. <https://doi.org/10.1165/rcmb.2018-0416TR>.

Schittny, Johannes C. 2017. “Development of the Lung.” *Cell and Tissue Research* 367 (3): 427–44. <https://doi.org/10.1007/s00441-016-2545-0>.

Seyhan, Attila A. 2019. “Lost in Translation: The Valley of Death across Preclinical and Clinical Divide – Identification of Problems and Overcoming Obstacles.” *Translational Medicine Communications* 4 (1): 18. <https://doi.org/10.1186/s41231-019-0050-7>.

Sipp, Douglas, Pamela G. Robey, and Leigh Turner. 2018. “Clear up This Stem-Cell Mess.” *Nature* 561 (7724): 455–57. <https://doi.org/10.1038/d41586-018-06756-9>.

Snyder, Jeremy, Leigh Turner, and Valorie A. Crooks. 2018. “Crowdfunding for Unproven Stem Cell–Based Interventions.” *JAMA* 319 (18): 1935. <https://doi.org/10.1001/jama.2018.3057>.

Springer, J., and A. Fischer. 2003. “Substance P-Induced Pulmonary Vascular Remodelling in Precision Cut Lung Slices.” *The European Respiratory Journal* 22 (4): 596–601. <https://doi.org/10.1183/09031936.03.00027903>.

Swarr, Daniel T., and Edward E. Morrisey. 2015. “Lung Endoderm Morphogenesis: Gasping for Form and Function.” *Annual Review of Cell and Developmental Biology* 31 (1): 553–73. <https://doi.org/10.1146/annurev-cellbio-100814-125249>.

Tang, Fuchou, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, et al. 2009. “MRNA-Seq Whole-Transcriptome Analysis of a Single Cell.” *Nature Methods* 6 (5): 377–82. <https://doi.org/10.1038/nmeth.1315>.

Thébaud, Bernard, Manoj Lalu, Laurent Renesme, Sasha Katwyk, Justin Presseau, Kednapa Thavorn, Kelly D. Cobey, et al. 2021. “Benefits and Obstacles to Cell Therapy in Neonates: The INCuBATOR (Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research).” *STEM CELLS Translational Medicine*, February, sctm.20-0508. <https://doi.org/10.1002/sctm.20-0508>.

Thompson, Mary, Shirley H. J. Mei, Dianna Wolfe, Josée Champagne, Dean Fergusson, Duncan J. Stewart, Katrina J. Sullivan, et al. 2020. “Cell Therapy with Intravascular Administration of Mesenchymal Stromal Cells Continues to Appear Safe: An Updated Systematic Review and Meta-Analysis.” *EClinicalMedicine* 19 (February): 100249. <https://doi.org/10.1016/j.eclinm.2019.100249>.

Turner, Leigh. 2020. “Preying on Public Fears and Anxieties in a Pandemic: Businesses Selling Unproven and Unlicensed ‘Stem Cell Treatments’ for COVID-19.” *Cell Stem Cell* 26 (6): 806–10. <https://doi.org/10.1016/j.stem.2020.05.003>.

Ubags, Niki D. J., Miguel A. Alejandro Alcazar, Suhas G. Kallapur, Sylvia Knapp, Sophie Lanone, Clare M. Lloyd, Rory E. Morty, et al. 2020. “Early Origins of Lung Disease: Towards an Interdisciplinary Approach.” *European Respiratory Review: An Official Journal of the European Respiratory Society* 29 (157). <https://doi.org/10.1183/16000617.0191-2020>.

Uhl, Franziska E., Sarah Vierkotten, Darcy E. Wagner, Gerald Burgstaller, Rita Costa, Ina Koch, Michael Lindner, Silke Meiners, Oliver Eickelberg, and Melanie Königshoff. 2015. “Preclinical Validation and Imaging of Wnt-Induced Repair in Human 3D Lung Tissue Cultures.” *The European Respiratory Journal* 46 (4): 1150–66. <https://doi.org/10.1183/09031936.00183214>.

Upadhyay, Swapna, and Lena Palmberg. 2018. “Air-Liquid Interface: Relevant In Vitro Models for Investigating Air Pollutant-Induced Pulmonary Toxicity.” *Toxicological Sciences: An Official Journal of the Society of Toxicology* 164 (1): 21–30.



<https://doi.org/10.1093/toxsci/kfy053>.

Vieira Braga, Felipe A., and Ricardo J. Miragaia. 2019. "Tissue Handling and Dissociation for Single-Cell RNA-Seq." *Methods in Molecular Biology (Clifton, N.J.)* 1979: 9–21. [https://doi.org/10.1007/978-1-4939-9240-9\\_2](https://doi.org/10.1007/978-1-4939-9240-9_2).

Viswanathan, S., Y. Shi, J. Galipeau, M. Krampera, K. Leblanc, I. Martin, J. Nolte, D. G. Phinney, and L. Sensebe. 2019. "Mesenchymal Stem versus Stromal Cells: International Society for Cell & Gene Therapy (ISCT®) Mesenchymal Stromal Cell Committee Position Statement on Nomenclature." *Cytotherapy* 21 (10): 1019–24. <https://doi.org/10.1016/j.jcyt.2019.08.002>.

Wadman, Meredith. 2006. "Medical Research: Them and Us No Longer." *Nature* 439 (7078): 779–80. <https://doi.org/10.1038/439779a>.

Warburton, David. 2017. "Overview of Lung Development in the Newborn Human." *Neonatology* 111 (4): 398–401. <https://doi.org/10.1159/000458465>.

Watson, Christa Y., Flavia Damiani, Sumati Ram-Mohan, Sylvia Rodrigues, Priscila de Moura Queiroz, Thomas C. Donaghey, Jamie H. Rosenblum Lichtenstein, Joseph D. Brain, Ramaswamy Krishnan, and Ramon M. Molina. 2016. "Screening for Chemical Toxicity Using Cryopreserved Precision Cut Lung Slices." *Toxicological Sciences* 150 (1): 225–33. <https://doi.org/10.1093/toxsci/kfv320>.

Wohlsen, A., S. Uhlig, and C. Martin. 2001. "Immediate Allergic Response in Small Airways." *American Journal of Respiratory and Critical Care Medicine* 163 (6): 1462–69. <https://doi.org/10.1164/ajrccm.163.6.2007138>.

Wohlsen, Andrea, Andrea Hirrle, Hermann Tenor, Degenhard Marx, and Rolf Beume. 2010. "Effect of Cyclic AMP-Elevating Agents on Airway Ciliary Beat Frequency in Central and Lateral Airways in Rat Precision-Cut Lung Slices." *European Journal of Pharmacology* 635 (1–3): 177–83. <https://doi.org/10.1016/j.ejphar.2010.03.005>.

Woolf, Steven H. 2008. "The Meaning of Translational Research and Why It Matters." *JAMA* 299 (2): 211–13. <https://doi.org/10.1001/jama.2007.26>.

Wu, Haojia, Yuhei Kirita, Erinn L. Donnelly, and Benjamin D. Humphreys. 2019. "Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis." *Journal of the American Society of Nephrology: JASN* 30 (1): 23–32. <https://doi.org/10.1681/ASN.2018090912>.

Yuan, Tingting, Thomas Volckaert, Diptiman Chanda, Victor J. Thannickal, and Stijn P. De Langhe. 2018. "Fgf10 Signaling in Lung Development, Homeostasis, Disease, and Repair After Injury." *Frontiers in Genetics* 9 (September): 418. <https://doi.org/10.3389/fgene.2018.00418>.



# Annex

V. Article INCuBATor Thébaud et al. *Stem Cells Translational Medicine* 2021

VI. Article Hurskainen et al. *Nature Communication* 2021

**HUMAN CLINICAL ARTICLE**

# Benefits and obstacles to cell therapy in neonates: The INCuBAToR (Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research)

Bernard Thébaud<sup>1,2,3</sup>  | Manoj Lalu<sup>1,3,4,5</sup> | Laurent Renesme<sup>2</sup> |  
 Sasha van Katwyk<sup>5</sup>  | Justin Presseau<sup>5,6</sup> | Kednapa Thavorn<sup>5,6</sup> |  
 Kelly D. Cobey<sup>5,7</sup> | Brian Hutton<sup>5</sup> | David Moher<sup>5</sup> | Roger F. Soll<sup>8</sup> |  
 Dean Fergusson<sup>5</sup>

<sup>1</sup>Regenerative Medicine Program, The Ottawa Hospital Research Institute (OHRI), Ottawa, Ontario, Canada

<sup>2</sup>Neonatology, Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, Ottawa, Ontario, Canada

<sup>3</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

<sup>4</sup>Department of Anesthesiology and Pain Medicine, University of Ottawa, Ottawa, Ontario, Canada

<sup>5</sup>Clinical Epidemiology Program, The Ottawa Hospital Research Institute (OHRI), Ottawa, Ontario, Canada

<sup>6</sup>School of Public Health and Preventive Medicine, University of Ottawa, Ottawa, Ontario, Canada

<sup>7</sup>Centre for Journalism, The Ottawa Hospital Research Institute (OHRI), Ottawa, Ontario, Canada

<sup>8</sup>Department of Pediatrics, Larner College of Medicine, University of Vermont, Burlington, Vermont

**Correspondence**

Bernard Thébaud, MD, PhD, Ottawa Hospital Research Institute, CCW Room W6120, 501 Smyth Road, Ottawa K1H 8L6, ON, Canada.  
 Email: bthebaud@ohri.ca

**Funding information**

Stem Cell Network; Ontario Institute for

**Abstract**

Cell-based therapies hold promise to substantially curb complications from extreme preterm birth, the main cause of death in children below the age of 5 years. Exciting preclinical studies in experimental neonatal lung injury have provided the impetus for the initiation of early phase clinical trials in extreme preterm infants at risk of developing bronchopulmonary dysplasia. Clinical translation of promising therapies, however, is slow and often fails. In the adult population, results of clinical trials so far have not matched the enticing preclinical data. The neonatal field has experienced many hard-earned lessons with the implementation of oxygen therapy or postnatal steroids. Here we briefly summarize the preclinical data that have permitted the initiation of early phase clinical trials of cell-based therapies in extreme preterm infants and describe the INCuBAToR concept (Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research), an evidence-based approach to mitigate the risk of translating advanced therapies into this vulnerable patient population. The INCuBAToR addresses several of the shortcomings at the preclinical and the clinical stage that usually contribute to the failure of clinical translation through (a) systematic reviews of preclinical and clinical studies, (b) integrated knowledge transfer through engaging important stakeholders early on, (c) early economic evaluation to determine if a novel therapy is viable, and (d) retrospective and prospective studies to define and test ideal eligibility criteria to optimize clinical trial design. The INCuBAToR concept can be applied to any novel therapy in order to enhance the likelihood of success of clinical translation in a timely, transparent, rigorous, and evidence-based fashion.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. STEM CELLS TRANSLATIONAL MEDICINE published by Wiley Periodicals LLC on behalf of AlphaMed Press

**KEYWORDS**

clinical translation, lung injury, mesenchymal stromal cells, preterm birth, regenerative medicine

**1 | MESENCHYMAL STROMAL CELLS FOR COMPLICATIONS OF PRETERM BIRTH**

Bronchopulmonary dysplasia (BPD), defined as need for oxygen and/or respiratory support at 36 weeks' corrected age, is the most frequent sequela of prematurity.<sup>1,2</sup> BPD contributes to lifelong respiratory and neurological impairment resulting in increased health care costs and parental burden.<sup>3</sup> BPD occurs in neonates born during the late canalicular stage of lung development when critical components of lung vascularization and gas exchange are just being established.<sup>2</sup> BPD is a multifactorial disease in which inflammation, oxidative stress, and mechanical stretch disrupt the normal sequence of lung growth.<sup>2</sup> As a consequence, the prevention of lung injury has become increasingly more challenging with no progress over the past decade.<sup>4</sup>

Preclinical proof-of-concept<sup>5,6</sup> and exploratory studies<sup>7</sup> demonstrate that mesenchymal stromal cells (MSCs) from various sources prevent oxygen-induced lung injury in a widely used neonatal rodent model mimicking some aspects of BPD. MSCs have been ascribed pleiotropic effects (eg, anti-inflammatory, proangiogenic, antifibrotic, and antioxidative), making MSC therapy very appealing for a multifactorial disease such as BPD.<sup>8</sup> Furthermore, many of the healing molecules released by MSCs, such as keratinocyte growth factor, insulin growth factor-I, and angiogenic growth factors, are known to promote lung growth, and this is of specific interest for the preterm lung. These data provide strong biological plausibility for the use of MSCs in BPD. Fueled by these promising preclinical findings, early phase clinical trials testing the safety and feasibility of MSC and other cell therapies for BPD have already been completed.<sup>9-11</sup> Many more trials are ongoing or in the planning phase, indicating the heightened recognition of the potential benefits of neonatal cell therapy for BPD and other complications of prematurity.

This current enthusiasm is reminiscent of the “Heroic” years (1950-1970) of neonatology when a “great spirit of innovation, somewhat lacking in discipline” was accompanied by the most striking care changes and errors in neonatology.<sup>12</sup> Avoiding errors of the past<sup>13,14</sup> and overcoming obstacles to progress for cell therapy in this vulnerable patient population is the focus of the INCuBAToR (Innovative Neonatal Cellular Therapy for BPD: Accelerating Translation of Research), a multidisciplinary and evidence-based engine to ensure safe and timely translation of promising advanced therapies from the bench to the bedside. The importance of evidence-based medicine has again come to the forefront in the current pandemic and the race to find efficient treatments for COVID-19-related complications.<sup>15</sup>

**Lessons learned**

The INCuBAToR concept is an evidence-based approach to mitigate the risk of translating advanced therapies into vulnerable extreme preterm infants but can be applied to any population and any novel therapy. The INCuBAToR presents a unique translational research platform that offers a comprehensive infrastructure to design translational research including expertise in trial design, knowledge synthesis, economic evaluation, and knowledge translation. It aims to facilitate and enable timely and robust evaluation of promising therapeutics to help bring effective therapeutics to patients sooner.

**Significance statement**

Cell-based therapies offer exciting opportunities to curb complications of extreme preterm birth, the main cause of death in children aged 5 years or younger. Although early phase clinical trials have begun, the translation of promising therapies often fails. Here the INCuBAToR concept, an evidence-based approach to mitigate the risk of translating cell-based therapies into a vulnerable patient population, is introduced. The INCuBAToR concept can be applied to any novel therapy to enhance the likelihood of success of clinical translation in a timely, transparent, rigorous, and evidence-based fashion.

**2 | THE INCuBAToR TO ENHANCE THE SUCCESS OF CLINICAL TRANSLATION OF MSCS FOR BPD**

Translation of potentially life-saving therapies is unacceptably slow and often fails. In neonatology, it took decades to bring breakthrough therapies from discovery to approval by the U.S. Food and Drug Administration: over 30 years for surfactant<sup>16</sup> and 20 years for inhaled nitric oxide.<sup>17</sup> More worrisome is that less than 5% of high impact preclinical reports are being clinically translated and only 11% of those tested clinically ultimately receive licensing.<sup>18</sup> The underlying reasons for failure are multifactorial; however, issues highlighted by others and our group include (a) lack of rigorous preclinical and clinical research methodology; (b) failure to properly synthesize current evidence to help justify and inform trial design<sup>19</sup>; (c) logistical, institutional, and regulatory obstacles at the clinical stage that add years to scientific development; and (d) failure to address concerns of ethics

boards and regulatory agencies as the result of failing to engage these stakeholders early in the translation process.<sup>20</sup> There is also failure to appreciate concerns of critical stakeholders, such as patients and their surrogates where relevant (in our case, parents) and clinicians, when establishing eligibility criteria, specifics of the intervention, and outcomes to be assessed in a clinical trial. This ultimately contributes to poor patient recruitment to trials. Moreover, eligibility criteria for trials of novel therapies are often too restrictive, resulting in failure to accrue sufficient patients within the target time. Finally, even in the case of successful therapies, the early economic consideration is often overlooked by investigators, leading to delayed commercialization of the therapy and reimbursement, which are necessary steps for clinical adoption. The INCuBAToR (Figure 1) is a novel, translational engine that can enhance speed, efficiency, rigor, and thus success for the clinical development of MSC therapy in neonatology. As outlined below, this approach was successfully applied to launch HULC-I, a phase I trial for MSC in BPD (ClinicalTrials.gov NCT04255147). A similar evidence-based approach—EPICOT (previous evidence, population to include, intervention to evaluate, comparison groups to identify, outcomes to define, and time over which the outcomes will be assessed)—has been successfully used to establish a consensus on designing efficient and consistent clinical trials for the intravenous use of MSCs for inflammatory bowel disease.<sup>21</sup>

## 2.1 | Systematic reviews of preclinical and clinical studies

Systematic reviews and meta-analyses (SRMAs) provide the most robust and evidence-based overview of the evidence on a topic and can point out knowledge gaps and thus guide future studies. Yet, their application to first-in-human trials has been limited and preclinical SRMAs for neonatal interventions were until recently nonexistent. We undertook two systematic reviews (SRs) to justify and inform our clinical trial: (a) a preclinical SR of studies testing MSCs in experimental models of neonatal lung injury and (b) a clinical SR to examine evidence for MSCs in BPD.

### 2.1.1 | Preclinical systematic review of evidence

Using SR methodologies described by the Cochrane Collaboration<sup>22</sup> and modified for preclinical SRs,<sup>23</sup> we performed an SRMA on preclinical studies testing MSCs in experimental models of neonatal lung injury. We developed systematic strategies to search MEDLINE, Embase, BIOSIS, and Cochrane databases in collaboration with an information specialist. Validated filters were applied to improve search efficiency and strategies underwent peer review of electronic search strategy.<sup>24</sup> Two reviewers independently screened studies. Relevant data were extracted and summarized and meta-analysis performed where appropriate. The study protocol was registered through CAMARADES (www.CAMARADES.info). This SRMA included 25 studies with over 450 animals used in 33 individual experiments and

suggested the beneficial effects of MSCs on lung structure, inflammation, and other parameters.<sup>7</sup> Of note, all studies used exclusively newborn rodents exposed to hyperoxia, highlighting the importance of experiments in large animal models to address important endpoints not obtainable in small rodents and additional safety data critical for regulatory agencies. Importantly, unclear risk of bias and incomplete reporting in the primary studies revealed nonadherence to reporting standards, emphasizing potential obstacles to successful clinical translation based on flawed preclinical data and the need to reinforce implementation of reporting standards such as the ARRIVE guidelines<sup>25</sup> ([www.nc3rs.org.uk/arrive-guidelines](http://www.nc3rs.org.uk/arrive-guidelines)). Because of the burst of preclinical MSC studies in the most recent literature and the emergence of novel cellular therapies, we updated our SR and included a network meta-analysis to compare these various cell products. Fifty-three studies assessed 15 different cell-based therapies, and 35 of those studied the effects of MSCs almost exclusively in hyperoxic rodent models of BPD.<sup>26</sup> The exploratory meta-network analysis suggested that MSCs are the most effective therapy with few head-to-head comparisons, highlighting again the relative youth of cell therapy in the neonatal arena. The unclear risk of bias still existed in most studies; however, many preclinical journals are now enforcing data transparency and reporting guidelines—at least for confirmatory research<sup>27</sup>—pushing a methodology resembling norms for randomized clinical trials.<sup>28</sup> These measures should decrease waste, improve quality and rigor in the reporting of preclinical studies, and thus generate a stronger foundation on which to make sound decisions to initiate clinical trials.<sup>29</sup>

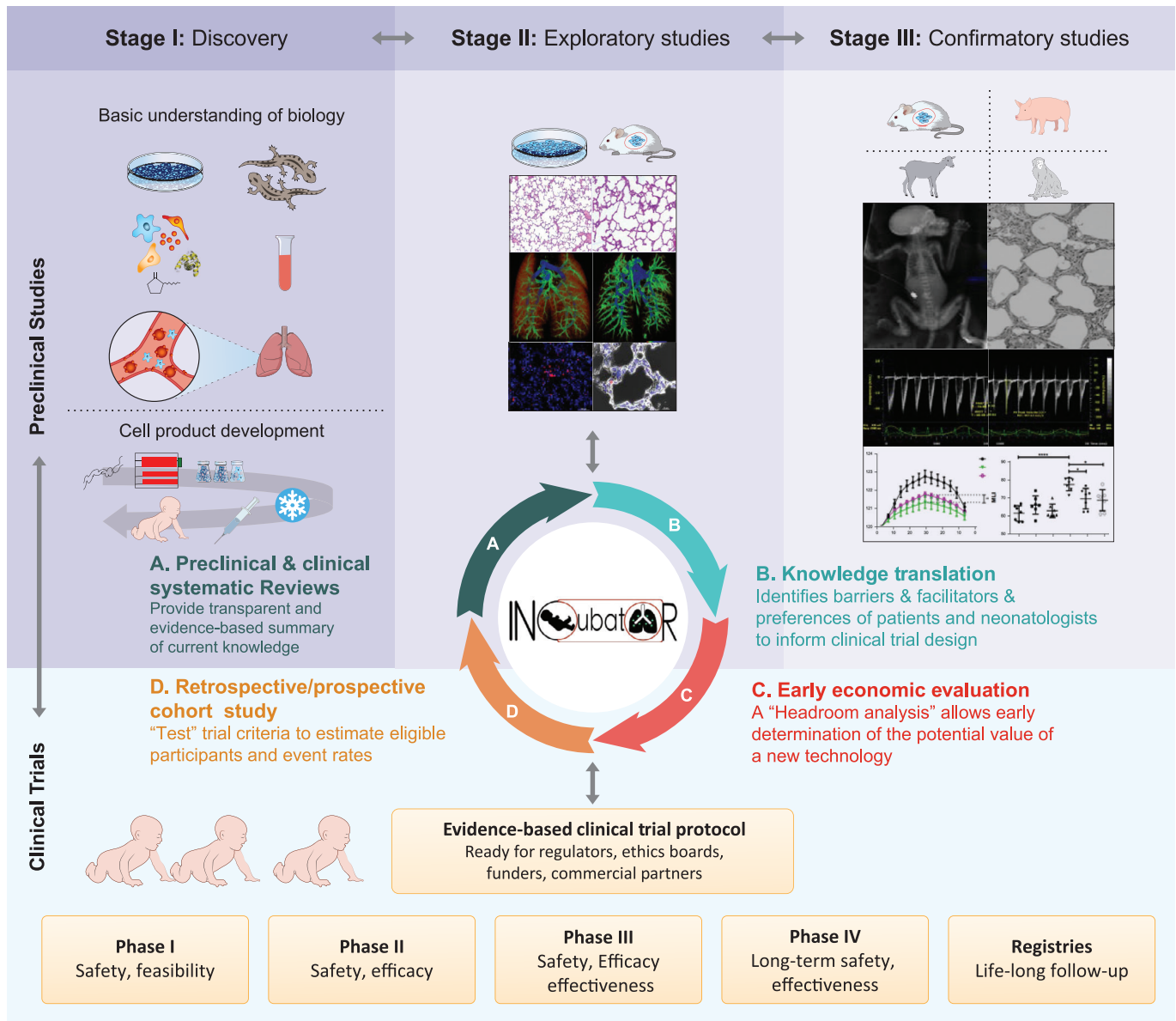
### 2.1.2 | Clinical systematic review of evidence

A systematic review to establish the current clinical evidence for MSCs in BPD further revealed the early stage of MSC therapy in neonatology.<sup>30</sup> We used the standard search strategy of the Cochrane Neonatal review group (<https://neonatal.cochrane.org/resources-review-authors>). We also searched clinical trials databases, conference proceedings, and the reference lists of retrieved articles for randomized controlled trials and quasirandomized trials. As of 2017, there was one published phase I trial using highly expanded MSCs from *cord blood* (Pneumostem; MEDIPOST Co., Ltd., Seongnam, South-Korea) from a commercial entity.<sup>9</sup> This first trial was a phase I dose-escalation trial in preterm infants with evolving BPD (5–14 days of life requiring continuous ventilatory support) to assess the safety and feasibility of cord blood MSCs. A single intratracheal injection of this allogeneic MSC product, starting with a dose of  $1.0 \times 10^7$  cells/kg for the first three patients and progressing to a dose of  $2.0 \times 10^7$  cells/kg for the next six patients, caused no serious adverse events or dose-limiting toxicity. Levels of proinflammatory cytokines in tracheal aspirates were reduced after MSC transplantation. When compared with historical controls, BPD severity was lower in MSC recipients, and rates of other adverse outcomes did not differ between the comparison group and MSC recipients.

Since publication of this clinical SR, the 2-year follow-up study of these nine preterm infants reported no adverse effects on growth,

respiratory, and neurodevelopmental outcomes.<sup>31</sup> A second phase I dose-escalation trial using the same *cord blood*-derived MSC product (Pneumostem) and a similar trial design was published in 2019, confirming feasibility and absence of serious adverse events in 12 preterm infants.<sup>11</sup> These studies indicate feasibility and absence of short-term

toxicity in a small number of patients. As of 18 November 2020, there were 18 registered trials in ClinicalTrials.gov under the terms “bronchopulmonary dysplasia” and “mesenchymal stem cells” at various stages of development with one completed phase II trial (NCT01828957).



**FIGURE 1** The INCuBAToR (Innovative Neonatal Cellular Therapy for BPD: Accelerating Translation of Research) for the successful clinical translation of cell-based therapies in neonates. The classical pathway to clinical translation includes the preclinical stage of discovery, exploratory, and confirmatory studies that provide the biological plausibility for the use of a novel therapy for a given disease and the rationale for initiating clinical trials. Numerous clinical trials fail because of shortcomings in the preclinical stages and/or lack of integration of critical information into clinical trial design. The INCuBAToR is designed to mitigate the risk of translating cell therapies into the clinic by providing an evidence-based approach through the following. A, Preclinical and clinical systematic reviews and meta-analyses to evaluate, synthesize, and quantitatively assess the best available evidence and identify knowledge gaps. B, Integrated knowledge translation to engage pertinent stakeholders, including patients, parents, physicians, nurses, and regulatory agencies, to identify opportunities and barriers in clinical trial implementation. C, Early economic evaluation to establish a “headroom analysis” for early determination of the potential value of the cell therapy. D, Retrospective cohort studies to estimate sample size, adverse events, and understanding of current patterns of care; this information can then be further refined by prospective observational cohort studies to provide “real world, real time” evidence to ensure feasibility and ultimately success of the clinical translation

Both the preclinical and clinical SRs provide some safety profile of MSCs and help inform trial design and potential trial participants, regulators, and research ethics boards of risks associated with MSC therapy in newborns. SRs from the adult literature further support the safety profile of MSCs from various sources.<sup>32-34</sup>

## 2.2 | Identifying barriers and enablers to conducting and participating in clinical trials of MSC therapy in preterm infants

Failure to enroll patients is a major concern to trial feasibility.<sup>18</sup> Understanding underlying beliefs, concerns, and perspectives are crucial to directly improve processes surrounding consent, refine eligibility criteria and trial outcomes to optimize the experience of deciding to enroll their child in a trial (parents) and optimize recruitment, retention, and trial delivery practices (clinicians).<sup>35</sup> In a novel application of the theoretical domain framework<sup>36</sup> and first systematic evaluation of stakeholder beliefs prior to embarking on any cellular therapy study, semistructured interviews were used with directed content analysis to identify barriers and enablers that may influence parents' and neonatologists' participation in clinical trials of MSC for BPD.<sup>37</sup> One-on-one interviews with parents of extremely preterm infants (n = 18) and neonatologists (n = 16) revealed key barriers for parents, including lack of knowledge about clinical trials and stem cells, concerns about their risks and side effects, and preferences for who should help them make the decision. Physicians reported competing priorities, time commitment, costs, and lack of institutional support as significant barriers to their ability to recruit patients. Using this methodological approach—which is typically used to inform later phase trials—helped to identify barriers that were unanticipated and may not have otherwise been flagged. The approach allowed us to systematically identify how to better support parents and clinicians by considering their concerns in the development of this early phase clinical trial of MSC therapy. As a result, we were able to directly address issues that could have compromised recruitment. These findings led to the development of an animated information video to enhance parent education with the goal of increasing trial enrollment. This video will first be evaluated in the observational arm of our phase I clinical trial. Parent foundations such as the Canadian Premature Babies Foundation or the European Foundation for the Care of Newborn Infants provide invaluable insight and are authentic drivers in the quest for successful clinical translation of promising therapies.<sup>38</sup>

Early engagement of parents and clinicians directly informs trial design, improves informed consent documents and process, and identifies feasibility issues associated with clinical adoption of MSC therapy.

## 2.3 | Assessing the potential value of MSC therapy for BPD

Early economic evaluation is now recognized as a tool to support product investment decision making.<sup>39</sup> Such evaluation is novel in neonatology, yet crucial to ensure that MSC therapy for BPD will be

economically viable. Given the lack of available data in the literature and the difficulty of obtaining reliable information from routinely collected databases, we developed a new, flexible tool to reliably forecast short- and long-term costs and health outcomes of BPD. The tool used an individual sample Markov model with seven health states in preterm infants born at 23-28 weeks. According to this tool, we have shown that BPD patients will incur over CAD\$700 000 in lifetime health systems costs associated with BPD and resulting complications.<sup>40</sup> This new model will now enable the “headroom” and the “value of information (VOI)” analyses. The headroom analysis presents the “cost-effectiveness gap” or maximum cost for which the MSC therapy can be brought to market and still be considered cost effective from the perspective of health care payers.<sup>41</sup> The VOI analysis identifies parameters that have large impact on the cost-effectiveness profile of MSC therapy for BPD and estimates optimal sample size and follow-up period in future randomized controlled trials. The VOI analysis will also identify research areas that will have the highest impact on reimbursement decisions.<sup>42</sup> As we are in the early phase of the development of MSCs for extremely preterm infants, the actual cost of MSCs is currently unknown. Unlike conventional therapeutics, there has been little published data on the cost of cell-based therapies, including MSCs. In general, biological products have higher prices in comparison with other drugs. For example, Zolgensma (onasemnogene abeparvovec-xioi; Novartis, Basel, Switzerland)—a gene therapy product for young children with spinal muscular atrophy—is priced at US \$2.1 million. Chimeric antigen receptor T-cell-based products, such as Kymriah (tisagenlecleucel; Novartis, Basel, Switzerland) and Yescarta (axicabtagene ciloleucel; Gilead Sciences, Inc, Foster City, California), cost US \$475 000 and US \$373 000, respectively, for a one-time dose. As part of INCuBAToR, we will use an early economic evaluation to estimate the potential price of MSCs at which the therapy is still considered cost effective for BPD. This estimated price will be based on early evidence on the potential impact of MSCs on health outcomes of infants with BPD.

## 2.4 | Retrospective and prospective observational cohort studies to ensure recruitment targets are met

Defining and “testing” optimal eligibility criteria are of particular importance in first-in-human trials as they enhance safety by excluding patients with an unacceptably high risk of treatment-related toxicity (relative to benefit) and/or insufficient expectation of efficacy. Too restrictive eligibility criteria can significantly reduce trial feasibility, as they limit patient accrual, and results may not be generalizable. Trials can also experience significant delays related to recruitment. As many as 86% of clinical trials do not reach recruitment targets within their specified time periods.<sup>43</sup> This is of particular concern in cell therapy trials, as failure to enroll patients within anticipated periods has been a major threat to trial feasibility.<sup>44</sup> Such an approach, paired with the abovementioned early economic evaluation, may have averted the shelving of a promising ventilation strategy for preterm infants.<sup>45</sup>



## 2.4.1 | Retrospective cohort study

To provide estimates of the number of eligible patients expected during the study period, as well as the expected event rates for these patients, retrospective cohort studies are useful and increasingly facilitated by national or international repositories (eg, Canadian Neonatal Network, Vermont Oxford Network, German Neonatal Network, etc.) that gather data on antenatal characteristics, risk stratification, resource utilization, and outcomes from neonatal intensive care unit patients.

## 2.4.2 | Prospective cohort study

Using eligibility criteria and outcomes refined and justified by the retrospective cohort study, a prospective observational cohort study provides “real world, real time” evidence to further refine proposed criteria to ensure feasibility while balancing concerns of safety. This is a novel approach to highly refine and evidence-inform a trial protocol prior to conducting a high-stakes, resource-intensive, interventional study. Given the acuity and expected high incidence of adverse events in extreme premature infants, data from a prospective cohort of patients are needed for appropriate comparative assessment of safety in phase I/II trials. This is of particular use for investigators, data safety monitoring boards, research ethics boards, and regulators. Thus, an observational study serves a number of important purposes for phase I/II interventional trials: (a) characterize the type and incidence of serious adverse events in a population meeting eligibility criteria, (b) serve as a practical “lead-in” phase for the investigator team to gain experience enrolling patients just prior to the initiation of a phase I/II interventional study by assessing the feasibility of potential patient recruitment by gaining insights into the parents' hypothetical willingness to participate in such a trial, (c) provide prospective measures of trial conduct feasibility such as consent and data collection procedures, (d) further define current patterns of care of BPD patients against which MSCs may be tested, and (e) refine and justify sample size calculation for a definitive interventional study. This observational cohort strategy is timely as the current pandemic seems to indicate a reduction in extreme preterm birth<sup>46</sup> and may lead to reevaluated timelines.

In summary, the INCuBAToR provides a rational, evidence-based approach to ensure safe and successful translation of MSC therapy in a vulnerable patient population. As with any disruptive innovation, the INCuBAToR will go through several iterations to improve over time. Indeed, although it enabled the rapid launch of a phase I trial, it also uncovered—but did not address—some major obstacles to progress of MSC therapy.

## 3 | MAJOR OBSTACLES TO PROGRESS OF MSC THERAPY—“CLEANING UP THE MESS”

The SRMAs revealed important disparities in MSC characterization, indication, and administration strategies.<sup>32-34</sup> This has also been the

experience of the Food and Drug Administration, which reported important differences in cell characterization, product bioactivity assessment, and tissue sourcing and product manufacturing in initial filings of 66 investigational new drug submissions for MSC-based therapies.<sup>47</sup> These inconsistencies and the stark contrast between promising results in the lab vs mitigated success in the clinic have led to major criticism regarding MSC therapy.<sup>48</sup> Current shortcomings, including an incomplete definition<sup>49</sup> and the lack of (a) potency assays to predict in vivo response, (b) standardized methods for manufacturing and use at bedside,<sup>50,51</sup> and (c) complete and transparent reporting of both cell characteristics and clinical trial details,<sup>52,53</sup> contribute to the controversy.<sup>48</sup> This lack of sufficient details concerning the cell product is highly problematic, as it significantly hinders the ability to judge the reliability of the results, interpret them, and replicate the findings. These consequences have been illustrated recently in the pandemic literature.<sup>54,55</sup> Likewise, an analysis of discrepancies (defined as at least two reported facts that cannot both be true because they are logically or mathematically incompatible) in clinical studies assessing the efficacy of bone marrow-derived cells on left ventricle ejection fraction in heart disease<sup>56</sup> revealed that the rigor of the report was associated with the effect size: studies with no discrepancies showed no effect of MSC on left ventricle function, whereas studies with the highest number of discrepancies also reported the biggest improvement in left ventricle function. This suggests that inadequate reporting is associated with biased reports.

To better address clinical translation, reproducibility, and transparency in the field of MSC research, the scientific community needs a consensus definition of MSCs. Similarly, to improve reporting quality of clinical MSC studies, a standardized reporting guideline is needed. A query on the EQUATOR network (Enhancing the Quality and Transparency of Health Research; <https://www.equator-network.org/>) found only one published reporting guideline for studies evaluating biologics in orthopedics (platelet-rich plasma and MSCs).<sup>57</sup> We propose a method to establish a consensus definition of MSC and to establish relevant reporting guidelines. Our approach will directly address the pitfalls and criticisms of previous attempts to generate consensus in the MSC field by using the Delphi method, a highly studied and well-established social science approach to reach group consensus on highly contentious issues (eg, biomedical editors core competencies, defining predatory journals).<sup>58,59</sup> The Delphi method allows for broader input beyond a small panel of experts, encourages independent reflection, and limits negative aspects of group decision making such as peer pressure, limited time to express point of view and reach agreement, lack of formal feedback, and nonstructured interactions and aggregation of opinion. Importantly, we will take an integrated knowledge translation approach where diverse stakeholders are part of the program from its inception,<sup>60</sup> which will help ensure that the definition and related reporting guidelines created are relevant to the community and ultimately are effectively adopted to improve quality, transparency, and reproducibility in basic and translational MSC research.

Nuances in the manufacturing processes can significantly influence bioactivity and functional outcomes of the MSC preparations.



The challenge here may lie in the fact that specific manufacturing processes may be proprietary and thus not disclosed particularly by commercial entities. This caveat will need to be addressed as the field matures further.

## 4 | CONCLUSION

MSC therapy has created much hope in neonatology with the promise to curb complications of extreme prematurity and to substantially improve the outcome of extreme preterm infants. The multidisciplinary INCuBAToR engine provides a rigorous and evidence-based approach to address the multiple obstacles to successful clinical translation of MSC and other cell therapies. The INCuBAToR has emanated from the Exceerator framework (<http://www.ohri.ca/blueprint/>) and as such can be applied to any promising novel therapy. The next iteration of this approach will include an attempt to tackle one of the remaining obstacles to progress of MSC therapy. Without stifling innovation in this still burgeoning field, the clinical translation of MSC requires an unbiased robust definition of MSCs and clinical reporting criteria. The Delphi method has previously provided solutions to contentious issues and may enable attaining this ambitious goal in order to further improve the rigor in translating promising MSC therapy into patient care.

## ACKNOWLEDGMENTS

B.T. is supported by the Canadian Institutes of Health Research (CIHR), the Ontario Institute for Regenerative Medicine (OIRM), and the Stem Cell Network.

## CONFLICT OF INTEREST

B.H. has a consultant/advisory role with Eversana Inc. The other authors declared no potential conflicts of interest.

## AUTHOR CONTRIBUTIONS

B.T.: conception/design, financial support, manuscript writing; M.L., L.R., S.v.K., J.P., K.T., K.D.C., B.H., D.M., and R.F.S.: manuscript writing; D.F.: conception/design, manuscript writing.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

Bernard Thébaud  <https://orcid.org/0000-0003-1844-7145>

Sasha van Katwyk  <https://orcid.org/0000-0003-3026-2063>

## REFERENCES

- Baraldi E, Filippone M. Chronic lung disease after premature birth. *N Engl J Med*. 2007;357:1946-1955.
- Thebaud B, Goss KN, Laughon M, et al. Bronchopulmonary dysplasia. *Nat Rev Dis Primers*. 2019;5:78.
- Lawn JE, Blencowe H, Oza S, et al.; Lancet Every Newborn Study Group. Every newborn: progress, priorities, and potential beyond survival. *Lancet*. 2014;384:189-205.
- Jobe A. The search for treatment of bronchopulmonary dysplasia. *JAMA Pediatr*. 2016;170:322-324.
- Aslam M, Baveja R, Liang OD, et al. Bone marrow stromal cells attenuate lung injury in a murine model of neonatal chronic lung disease. *Am J Respir Crit Care Med*. 2009;180:1122-1130.
- van Haften T, Byrne R, Bonnet S, et al. Airway delivery of mesenchymal stem cells prevents arrested alveolar growth in neonatal lung injury in rats. *Am J Respir Crit Care Med*. 2009;180:1131-1142.
- Augustine S, Avey MT, Harrison B, et al. Mesenchymal Stromal Cell Therapy in Bronchopulmonary dysplasia: systematic review and meta-analysis of preclinical studies. *stem cells translational med*. 2017;6:2079-2093.
- Lesage F, Thebaud B. Nanotherapies for micropreemies: stem cells and the secretome in bronchopulmonary dysplasia. *Semin Perinatol*. 2018;42:453-458.
- Chang YS, Ahn SY, Yoo HS, et al. Mesenchymal stem cells for bronchopulmonary dysplasia: phase 1 dose-escalation clinical trial. *J Pediatr*. 2014;164:966-972.e6.
- Lim R, Malhotra A, Tan J, et al. First-in-human administration of allogeneic amnion cells in premature infants with bronchopulmonary dysplasia: a safety study. *stem cells translational med*. 2018;7:628-635.
- Powell SB, Silvestri JM. Safety of intratracheal administration of human umbilical cord blood derived mesenchymal stromal cells in extremely low birth weight preterm infants. *J Pediatr*. 2019;210:209-213.e2.
- Robertson AF. Reflections on errors in neonatology: II. The "Heroic" years, 1950 to 1970. *J Perinatol*. 2003;23:154-161.
- Bizzarro MJ. Optimizing oxygen saturation targets in extremely preterm infants. *JAMA*. 2018;319:2173-2174.
- Marlow N. Reevaluating postnatal steroids for extremely preterm infants to prevent lung disease. *JAMA*. 2017;317:1317-1318.
- Rice TW, Janz DR. In defense of evidence-based medicine for the treatment of COVID-19 acute respiratory distress syndrome. *Ann Am Thorac Soc*. 2020;17:787-789.
- Avery ME. Surfactant deficiency in hyaline membrane disease: the story of discovery. *Am J Respir Crit Care Med*. 2000;161:1074-1075.
- Abman SH. Inhaled nitric oxide for the treatment of pulmonary arterial hypertension. *Handb Exp Pharmacol*. 2013;218:257-276.
- Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. *Am J Med*. 2003;114:477-484.
- Schmidt-Pogoda A, Bonberg N, Koecke MHM, et al. Why most acute stroke studies are positive in animals but not in patients: a systematic comparison of preclinical, early phase, and phase 3 clinical trials of neuroprotective agents. *Ann Neurol*. 2020;87:40-51.
- Davis JM, Pursley DM, Pediatric Policy Council. Cell-based therapies in neonates: the emerging role of regulatory science. *Pediatr Res*. 2019;86:145-146.
- Ciccocioppo R, Baumgart DC, Dos Santos CC, et al. Perspectives of the International Society for Cell & Gene Therapy Gastrointestinal Scientific Committee on the intravenous use of mesenchymal stromal cells in inflammatory bowel disease (PeMeGi). *Cytotherapy*. 2019;21:824-839.
- Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. London, England: Cochrane Collaboration; 2011.
- Sena ES, Currie GL, McCann SK, et al. Systematic reviews and meta-analysis of preclinical studies: why perform them and how to appraise them critically. *J Cereb Blood Flow Metab*. 2014;34:737-742.
- Leenaars M, Hooijmans CR, van Veggel N, et al. A step-by-step guide to systematically identify all relevant animal studies. *Lab Anim*. 2012;46:24-31.

25. Percie du Sert N, Hurst V, Ahluwalia A, et al. The ARRIVE guidelines 2.0: updated guidelines for reporting animal research. *PLoS Biol.* 2020; 18:e3000410.
26. Augustine S, Cheng W, Avey MT, et al. Are all stem cells equal? Systematic review, evidence map, and meta-analyses of preclinical stem cell-based therapies for bronchopulmonary dysplasia. *stem cells translational med.* 2020;9:158-168.
27. Han S, Olonisakin TF, Pribis JP, et al. A checklist is associated with increased quality of reporting preclinical biomedical research: a systematic review. *PLoS One.* 2017;12:e0183591.
28. Muhlhauser BS, Bloomfield FH, Gillman MW. Whole animal experiments should be more like human randomized controlled trials. *PLoS Biol.* 2013;11:e1001481.
29. Moher D, Shamseer L, Cobey KD, et al. Stop this waste of people, animals and money. *Nature.* 2017;549:23-25.
30. Pierro M, Thebaud B, Soll R. Mesenchymal stem cells for the prevention and treatment of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev.* 2017;11:CD011932.
31. Ahn SY, Chang YS, Kim JH, Sung SI, Park WS. Two-year follow-up outcomes of premature infants enrolled in the phase I trial of mesenchymal stem cells transplantation for bronchopulmonary dysplasia. *J Pediatr.* 2017;185:49-54.e2.
32. Can A, Celikkan FT, Cinar O. Umbilical cord mesenchymal stromal cell transplantations: a systemic analysis of clinical trials. *Cytotherapy.* 2017;19:1351-1382.
33. Lalu MM, McIntyre L, Pugliese C, et al.; Canadian Critical Care Trials Group. Safety of cell therapy with mesenchymal stromal cells (SafeCell): a systematic review and meta-analysis of clinical trials. *PLoS One.* 2012;7:e47559.
34. Thompson M, Mei SHJ, Wolfe D, et al. Cell therapy with intravascular administration of mesenchymal stromal cells continues to appear safe: an updated systematic review and meta-analysis. *EClinicalMedicine.* 2020;19:100249.
35. Francis JJ, Johnston M, Burr J, et al. Importance of behaviour in interventions. *BMJ.* 2008;337:a2472.
36. Duncan EM, Francis JJ, Johnston M, et al. Learning curves, taking instructions, and patient safety: using a theoretical domains framework in an interview study to investigate prescribing errors among trainee doctors. *Implement Sci.* 2012;7:86.
37. Guillot M, Asad S, Lalu MM, et al. So you want to give stem cells to babies? Neonatologists and parents' views to optimize clinical trials. *J Pediatr.* 2019;210:41-47.e1.
38. Zimmermann LJI, Kostenzer J, Mader S. Tackling bronchopulmonary dysplasia to improve preterm health: a call for family-centered care at world prematurity day 2020. *Am J Physiol Lung Cell Mol Physiol.* 2020; 319:L867-L870.
39. Ijzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. *Appl Health Econ Health Policy.* 2011;9: 331-347.
40. van Katwyk S, Augustine S, Thebaud B, et al. Lifetime patient outcomes and healthcare utilization for BPD and extreme preterm infants: a microsimulation study. *J Pediatr.* 2020;25:136.
41. Girling A, Lilford R, Cole A, Young T. Headroom approach to device development: current and future directions. *Int J Technol Assess Health Care.* 2015;31:331-338.
42. Claxton K, Thompson KM. A dynamic programming approach to the efficient design of clinical trials. *J Health Econ.* 2001;20: 797-822.
43. Kadam RA, Borde SU, Madas SA, Salvi SS, Limaye SS. Challenges in recruitment and retention of clinical trial subjects. *Perspect Clin Res.* 2016;7:137-143.
44. Bonfiglio GA. Cell therapy clinical trials: why they fail. Paper presented at: IBC Cell Therapy Clinical Development Conference; September 10-11, 2012; Arlington, VA.
45. Leach CL, Greenspan JS, Rubenstein SD, et al. Partial liquid ventilation with Perflubron in premature infants with severe respiratory distress syndrome. *N Engl J Med.* 1996;335:761-767.
46. Hedermann G, Hedley PL, Baekvad-Hansen M, et al. Danish premature birth rates during the COVID-19 lockdown. *Arch Dis Child Fetal Neonatal Ed.* 2021;106:93-95.
47. Mendicino M, Bailey AM, Wonnacott K, Puri RK, Bauer SR. MSC-based product characterization for clinical trials: an FDA perspective. *Cell Stem Cell.* 2014;14:141-145.
48. Sipp D, Robey PG, Turner L. Clear up this stem-cell mess. *Nature.* 2018;561:455-457.
49. Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy.* 2006;8:315-317.
50. Galipeau J, Sensebe L. Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. *Cell Stem Cell.* 2018;22:824-833.
51. Nolta JA, Galipeau J, Phinney DG. Improving mesenchymal stem/stromal cell potency and survival: proceedings from the International Society of Cell Therapy (ISCT) MSC preconference held in May 2018, Palais des Congres de Montreal, organized by the ISCT MSC Scientific Committee. *Cytotherapy.* 2020;22:123-126.
52. Rizk M, Monaghan M, Shorr R, Kekre N, Bredeson CN, Allan DS. Heterogeneity in studies of mesenchymal stromal cells to treat or prevent graft-versus-host disease: a scoping review of the evidence. *Biol Blood Marrow Transplant.* 2016;22:1416-1423.
53. Robinson PG, Murray IR, West CC, et al. Reporting of mesenchymal stem cell preparation protocols and composition: a systematic review of the clinical orthopaedic literature. *Am J Sports Med.* 2019;47:991-1000.
54. Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Retraction: cardiovascular disease, drug therapy, and mortality in COVID-19. *N Engl J Med.* 2020;382:2582.
55. Mehra MR, Ruschitzka F, Patel AN. Retraction-Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet.* 2020;395:1820.
56. Nowbar AN, Mielewicz M, Karavassilis M, et al. DAMASCENE writing group. Discrepancies in autologous bone marrow stem cell trials and enhancement of ejection fraction (DAMASCENE): weighted regression and meta-analysis. *BMJ.* 2014;348:g2688.
57. Murray IR, Geeslin AG, Goudie EB, Petrigliano FA, LaPrade RF. Minimum information for studies evaluating biologics in orthopaedics (MIBO): platelet-rich plasma and mesenchymal stem cells. *J Bone Joint Surg Am.* 2017;99:809-819.
58. Galipeau J, Cobey KD, Barbour V, et al. An international survey and modified Delphi process revealed editors' perceptions, training needs, and ratings of competency-related statements for the development of core competencies for scientific editors of biomedical journals. *F1000Res.* 2017;6:1634.
59. Grudniewicz A, Moher D, Cobey KD, et al. Predatory journals: no definition, no defence. *Nature.* 2019;576:210-212.
60. Boland L, Kothari A, McCutcheon C. Integrated knowledge translation research network. Building an integrated knowledge translation (IKT) evidence base: colloquium proceedings and research direction. *Health Res Policy Syst.* 2020;18:8.

**How to cite this article:** Thébaud B, Lalu M, Renesme L, et al. Benefits and obstacles to cell therapy in neonates: The INCuBATORe (Innovative Neonatal Cellular Therapy for Bronchopulmonary Dysplasia: Accelerating Translation of Research). *STEM CELLS Transl Med.* 2021;10:968-975. <https://doi.org/10.1002/sctm.20-0508>

# Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage

Maria Hurskainen<sup>1,2,3,4,11</sup>, Ivana Mižíková <sup>1,4,11</sup>, David P. Cook<sup>4,5</sup>, Noora Andersson<sup>3,6</sup>, Chanèle Cyr-Depauw <sup>1,4</sup>, Flore Lesage<sup>1,4</sup>, Emmi Helle <sup>2,3,7</sup>, Laurent Renesme <sup>1,4</sup>, Robert P. Jankov<sup>4,8,9</sup>, Markku Heikinheimo <sup>3</sup>, Barbara C. Vanderhyden <sup>4,5,10</sup> & Bernard Thébaud <sup>1,4,8</sup>✉

During late lung development, alveolar and microvascular development is finalized to enable sufficient gas exchange. Impaired late lung development manifests as bronchopulmonary dysplasia (BPD) in preterm infants. Single-cell RNA sequencing (scRNA-seq) allows for assessment of complex cellular dynamics during biological processes, such as development. Here, we use MULTI-seq to generate scRNA-seq profiles of over 66,000 cells from 36 mice during normal or impaired lung development secondary to hyperoxia with validation of some of the findings in lungs from BPD patients. We observe dynamic populations of cells, including several rare cell types and putative progenitors. Hyperoxia exposure, which mimics the BPD phenotype, alters the composition of all cellular compartments, particularly alveolar epithelium, stromal fibroblasts, capillary endothelium and macrophage populations. Pathway analysis and predicted dynamic cellular crosstalk suggest inflammatory signaling as the main driver of hyperoxia-induced changes. Our data provides a single-cell view of cellular changes associated with late lung development in health and disease.

<sup>1</sup> Sinclair Centre for Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, ON, Canada. <sup>2</sup> Division of Pediatric Cardiology, New Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland. <sup>3</sup> Pediatric Research Center, New Children's Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. <sup>4</sup> Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada. <sup>5</sup> Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada. <sup>6</sup> Research Programs unit, Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland. <sup>7</sup> Research Programs Unit, Stem Cells and Metabolism, University of Helsinki, Helsinki, Finland. <sup>8</sup> Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, University of Ottawa, Ottawa, ON, Canada. <sup>9</sup> Molecular Biomedicine Program, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada. <sup>10</sup> Department of Obstetrics and Gynecology, University of Ottawa/The Ottawa Hospital, Ottawa, ON, Canada. <sup>11</sup> These authors contributed equally: Maria Hurskainen, Ivana Mižíková. ✉ email: [bthebaud@toh.ca](mailto:bthebaud@toh.ca)

Late lung development is responsible for the formation of intricate structures enabling the exchange of inspired oxygen from the atmosphere and carbon dioxide from the blood, which is the primary function of the mammalian lung. This complex task of gas exchange is achieved in the smallest, most distal respiratory units of the lung (the alveoli) and occurs across a thin structure (0.2–2  $\mu\text{m}$ ) of the alveolo-capillary barrier covering a vast surface area of the lung (~75  $\text{m}^2$ ). The formation of this complex structure is achieved via interconnected events of secondary septa formation and microvascular maturation during the period of late lung development. These processes are facilitated by temporarily and spatially coordinated crosstalks between multiple cell types in the lung microenvironment. In addition to gas exchange, the lung acts as an important immune barrier, requiring resident alveolar macrophages to transition toward a mature anti-inflammatory phenotype. However, the signals driving these processes and the landscape of resident cells during late lung development remain largely uncharacterized<sup>1–3</sup>.

In humans, impaired late lung development presents as bronchopulmonary dysplasia (BPD), the most common chronic lung disease in children. BPD is a multifactorial disease with some hereditary component<sup>4,5</sup>, occurring as a consequence of premature birth and the result of an aberrant reparative response to both antenatal and repetitive postnatal injury to the developing lungs<sup>6</sup>. In addition to impaired alveolar and microvascular formation, immune development of the lung is interrupted, leading to recurrent bacterial and viral respiratory infections. To mimic these injuries, rodent models of BPD utilize various levels of hyperoxia and/or other pro-inflammatory stimuli. Sustained exposure of neonatal mice to hyperoxia leads to a BPD-like lung phenotype, making it an ideal model to identify and study pivotal developmental steps during late lung development<sup>7</sup>.

The role of various cell types during late lung development has been extensively studied, establishing important functions for myofibroblasts in secondary septation, endothelial cell (EC) signaling in the processes of microvascular maturation and coordination of inflammatory cell signaling<sup>7</sup>. The abundance and identity of individual cell types are dynamic throughout lung development. Identification and classification of lung cells become even more complex under pathological conditions, in particular when the nature of the disease is heterogeneous. Traditional methods to assess molecular characteristics of pathologies have depended on bulk measurements of protein or RNA, but given the heterogeneity of lung tissue and its dynamics during late development, these measurements are confounded by changes in cellular composition. As a result, changes in individual cell types cannot be identified. This is more problematic when responses are limited to rare populations, as these changes will be masked by the signal from more-abundant cell types. To circumvent these obstacles, we herein employed multiplexed single-cell RNA sequencing (scRNA-seq) to resolve changes in cellular composition and state during both normal and impaired late lung development.

Here, we report an extensive profiling of the cellular composition in the developing mouse lung by generating scRNA-seq profiles of 66,200 cells from 36 normally and aberrantly ( $\text{O}_2$ -exposed) developing mouse lungs at three time points (P3, P7, and P14). We also validate some of our findings by fluorescent RNA *in situ* hybridization in lung samples from BPD patients. In this study, we observe greatly diverse and dynamic populations of cells. Hyperoxia exposure alters the phenotype of all major cell types, particularly capillary endothelium, stromal, and macrophage populations. Our data suggest that inflammatory activation is the major driver of the observed transcriptional changes in hyperoxia. In conclusion, we identify multiple cell-specific gene signatures, which provide a detailed

cell and molecular atlas of normal and impaired postnatal lung development.

## Results

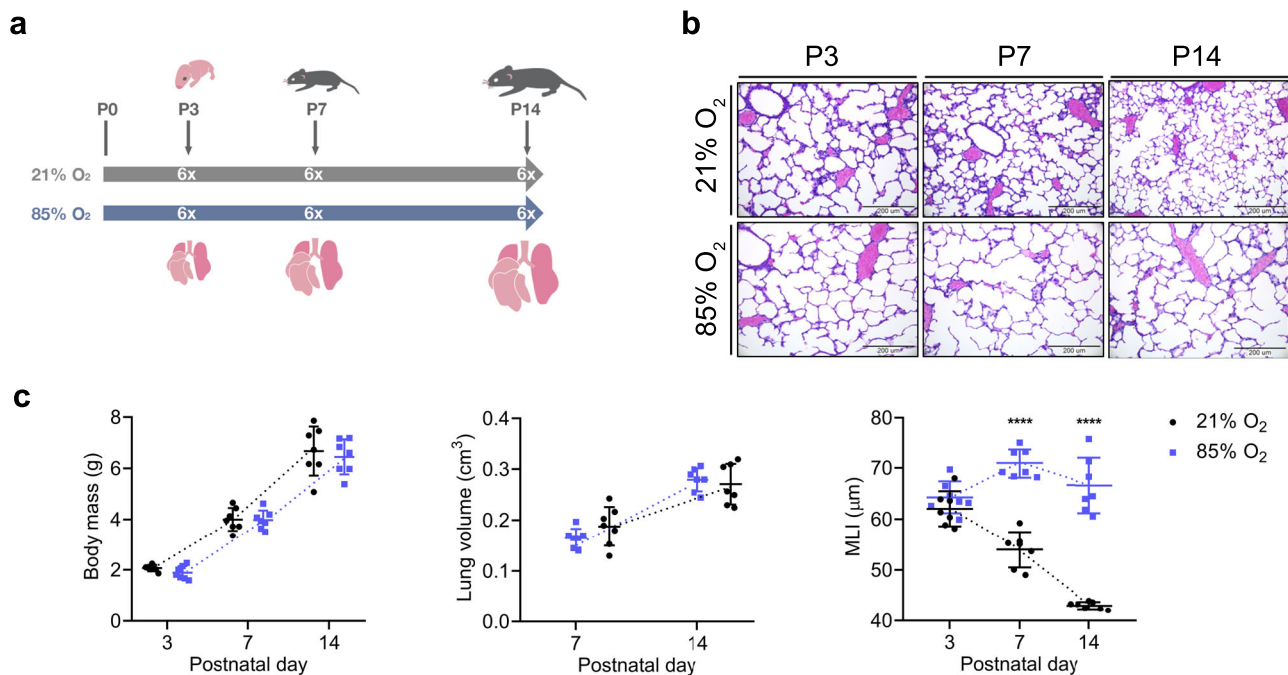
**Detailed map of cellular composition during normal and impaired late murine lung development.** In order to create a comprehensive cellular map of the normal and impaired developing lung, we generated scRNA-seq profiles of 36 mice on postnatal days (P)3, 7, and 14 (Fig. 1a). Impaired lung development was induced by normobaric hyperoxia (85%  $\text{O}_2$ ) from day of birth (P0) to P14 (Fig. 1a–c) and was independent of body weight or lung volume (Fig. 1a–c). In order to capture diverse cell populations present in the lung, we optimized the single-cell preparation protocol by testing several digestion conditions as assessed by FACS and scRNA-seq analysis (Supplementary Fig. 1a, b). To evaluate the actual cell contribution *in vivo* prior to tissue digestion, we performed a stereological assessment of alveolar epithelial type 2 (AT2) cells. The number of AT2 cells *in vivo* was not impacted by hyperoxia as assessed by stereology, supporting the validity of our observations (Supplementary Fig. 1c; Supplementary Table 1). In total, we generated scRNA-seq expression profiles for 66,200 cells (~11,033 cells/group) (Supplementary Fig. 1d). Single-cell suspensions from individual mice were multiplexed using MULTI-seq<sup>8</sup> (Supplementary Fig. 1e, f). No major sex-dependent bias in cell distribution could be observed (Supplementary Fig. 1g, h).

Cells were clustered based on their expression profile, and cell types were annotated based on established cell markers available on LungMap<sup>9</sup> (<https://lungmap.net/>), CellMarker<sup>10</sup>, The Human Protein Atlas<sup>11</sup> (<http://www.proteinatlas.org>), and in the published literature (Fig. 2a; Supplementary Data 1). A total of 34 clusters were identified, corresponding to six major cell groups: epithelial, stromal, endothelial, myeloid, lymphoid, and mesothelial cells (Fig. 2b; Table 1; Supplementary Fig. 2a, b, Supplementary Movie 1). We observed dynamic changes in the cellular composition of normally developing lungs. Most changes occurred between P7 and P14. Impaired alveolar development induced by hyperoxia changed the cellular distribution of the lung at all time points (Fig. 2c, d; Supplementary Fig. 2a, b; Supplementary Data 2). The expression of BPD-associated genes in our data showed gene-specific distinct cellular expression patterns with the highest expression in capillary endothelial cells, macrophages and neutrophils, NK, mast and basophil and T cells, AT1 and AT2 cells, fibroblasts and mesothelial cells, many of which were indicated as most reactive to hyperoxia in our other data analysis (Supplementary Fig. 3). We then used the NicheNet tool<sup>12</sup> to infer cellular communications specific to hyperoxia-associated gene expression patterns based on expression of ligands, receptors, associated pathway components, and genomic targets of these pathways (Fig. 2e). According to differentially expressed genes in hyperoxia samples, our cell communication inference indicated that the effects of hyperoxia were mediated mostly by inflammatory signals (Fig. 2e). Our signaling pathway and cell communication analysis suggests that hyperoxia initiates inflammation with activation of particular endothelial, epithelial, stromal and resident lung immune cell subpopulations promoting innate and adaptive immune responses, fibrosis and several pathways disturbing endothelial development and homeostasis.

**Hyperoxia alters AT2 cell populations during late murine lung development.** We identified five clusters of epithelial cells with distinct expression profiles (Fig. 3a–c; Supplementary Fig. 4a; Supplementary Data 3).

Two bronchial epithelial clusters—club cells and ciliated cells—were identified. The frequency of club cells in normally





**Fig. 1** Exposure to hyperoxia induced an arrest in alveolarization in the developing mouse lung. **a** Mouse pups were exposed from day of birth to room air (21% O<sub>2</sub>, gray) or hyperoxia (85% O<sub>2</sub>, blue). A total of 36 lungs were harvested on postnatal days (P)3, 7, and 14. *n* = 6/group. **b** Representative histological sections from lungs developing in 21% O<sub>2</sub> (black circles) or 85% O<sub>2</sub> (purple squares) at P3, P7, and P14. Seven animals/group were evaluated. Scale bar = 200 µm. **c** Body weight was assessed at P3, P7, and P14. Lung volume was assessed by Archimedes principle at P7 and P14. Precise measurements at P3 were not possible due to the small size of the organ. Lung morphometry was quantified by the mean linear intercept (MLI) measurement. Data are presented as means ± SD. Statistical analyses were performed with GraphPad Prism 8.0. The presence of potential statistical outliers was determined by Grubbs' test. Significance was evaluated by multiple unpaired Student's *t*-test with Holm-Sidak correction. *P* values \*\*\*\* = <0.0001. *n* = 7 animals/group.

developing lungs remained constant between P3 and P7, but decreased slightly by P14. This dynamic was absent in club cells population in hyperoxic lungs (Supplementary Fig. 4b, Supplementary Data 2).

Within the alveolar epithelium, we identified one alveolar epithelial type 1 cell (AT1) cluster and two distinct AT2 clusters (Fig. 3a). The AT1 cluster was associated with expression of *Hopx*, *Akap5*, and *Vegfa* (Fig. 3b, c). The proportion of AT1 cells (cluster 1) decreased during normal development. This decline was not observed between P7 and P14 in hyperoxia-exposed lungs (Supplementary Figs. 2a and 4b).

The two AT2 cell clusters had largely similar transcriptional profiles in developing lungs, with *Lyz1* serving as the single most distinctive identifier of the secondary AT2 population (AT2-*Lyz1*<sup>+</sup>; Fig. 3b-d; Supplementary Fig. 4a and c; Supplementary Data 3).

While decreased lysozyme content in airway secretions was associated with BPD<sup>13</sup>, unlike in humans, lysozyme orthologs in mice are encoded by two genes (*Lyz1* and *Lyz2*). The AT2-*Lyz1*<sup>+</sup> population may be mouse-specific. Both clusters showed shifts in gene expression in hyperoxia-exposed lungs at all investigated time points (Fig. 3d and f; Supplementary Fig. 4a; Supplementary Data 4). Among the hyperoxia-specific genes were multiple factors associated with BPD and epithelial damage, including *Slpi*—a protease inhibitor protecting from epithelial damage—and the innate immune response regulator *Mif*<sup>14</sup> (Fig. 3d). MIF was previously shown to promote production of IL6 and IL1β, and has been implicated in the arrest of lung development and angiogenesis<sup>15,16</sup>. *Lcn-2* expression—a gene associated with BPD<sup>17</sup>—was also higher in the aberrant AT2 cells. Among the top downregulated genes by hyperoxia were *Meg3* and *Abca3* (Fig. 3d). *Meg3* is a known regulator of epithelial cell

differentiation, while *Abca3* is a crucial factor in surfactant and lamellar bodies' metabolism<sup>18,19</sup>. Mutations of *Abca3* have also been associated with interstitial lung disease and respiratory distress<sup>20</sup>. Furthermore, gene set enrichment analysis (GSEA) of AT2 populations suggested that pathways associated with epithelial, endothelial, and lung alveolus developments were downregulated following hyperoxia exposure (Fig. 3g; Supplementary Data 5).

Finally, we sought to identify a subpopulation of *Axin2*<sup>+</sup> AT2 progenitors, postulated to play a role in alveolar regeneration<sup>21</sup>. We identified a small number of epithelial cells expressing *Axin2*, mostly located to AT1 and AT2 clusters. *Axin2*<sup>+</sup> cells were not confined to a separate cluster, but were primarily present within AT1 and AT2 clusters (Supplementary Fig. 4d).

Cell communication inference suggested that AT2 and AT2-*Lyz1*<sup>+</sup> are very interactive with other cell types during the response to hyperoxia. This analysis highlighted multiple hyperoxia-related effects in both AT2 clusters, possibly mediated by signals from stromal, endothelial and immune cells, including *Vegfa* and *Tnf* signaling (Fig. 2e; Supplementary Fig. 4e, f). Hyperoxia upregulated *Tnf* expression in interstitial macrophages, which was associated with expression of *Tnf* receptors *Nrp1* and *Dag1* in AT2 cells (Fig. 2e; Supplementary Fig. 4e, f). *Tnf* is a potent pro-inflammatory cytokine increased in both BPD patients and in animal models of BPD<sup>22,23</sup>. *Nrp1* is critical for normal branching morphogenesis<sup>24</sup>, while *Dag1* plays an important role in airway epithelial wound repair<sup>25</sup>. In addition, the analysis revealed a potential regulation of AT2 cells by *Angpt2*, which was upregulated by hyperoxia in two capillary clusters (Cap, Cap-a) (Fig. 2e; Supplementary Fig. 4e). *Angpt2* is a ligand of the cell surface receptor ITGB1, critical in epithelia stratification during lung branching morphogenesis<sup>26</sup>.



**Fig. 2 Map of cellular composition in normal and hyperoxia-impaired late murine lung development.** **a** UMAP plot of all scRNA-seq data, showing a total of 34 distinct cell types that were identified. **b** UMAP plots showing expression levels for canonical markers of epithelial, mesenchymal, endothelial, immune, and mesothelial populations. The intensity of expression is indicated by purple coloring. **c** UMAP plots of normally (21% O<sub>2</sub>-exposed; left) and aberrantly (85% O<sub>2</sub>-exposed; right) developing lungs. Each cell is colored by mouse age as indicated by the legend. **d** Cluster distribution in lungs of normally and aberrantly developing mice at P3, P7, and P14.  $n = 6$  animals/group. **e** Circos plot showing inferred cell communications. Cell types in the top right correspond to those with the largest changes in response to hyperoxia. These cell types are connected to the cell types expressing ligands predicted to promote this response. Ligands expressed by the same cell population are colored the same. Expression levels in UMAP plots are presented as log (TP10k + 1) values. Log(TP10k + 1) corresponds to log-transformed UMIs per 10k. Cell populations in **a** and **d** are colored as indicated by the legend in **d**.

**Table 1 Identified cell populations.**

Abbreviation	Cell type	Abbreviation	Cell type
gCap	General capillary endothelial cells	Treg cells	Regulatory T cells
aCap	Capillary endothelial cells - aerocytes	ILC2	Innate lymphoid cells 2
Art	Arterial endothelial cells	$\gamma\delta$ T cells	$\gamma\delta$ T cells
Vein	Venous endothelial cells	pro-B cells	B-cell progenitors
Lymph	Lymphatic endothelial cells	CD4 <sup>+</sup> /CD8 <sup>+</sup> T cells	CD4 <sup>+</sup> /CD8 <sup>+</sup> T cells
Alv M $\Phi$	Alveolar macrophages	AT2	Alveolar type 2 cells ( <i>Lyz1</i> )
Neut 1	Neutrophils 1	AT1	Alveolar type 1 cells
Mono	Monocytes	AT2- <i>Lyz1</i> <sup>+</sup>	Alveolar type 2 cells ( <i>Lyz1</i> <sup>+</sup> )
Int M $\Phi$	Interstitial macrophages	Ciliated	Ciliated cells
Neut 2	Neutrophils 2	Club	Club cells
Mast Ba2	Mast basophils 2	<i>Col13a1</i> <sup>+</sup> fib	<i>Col13a1</i> <sup>+</sup> fibroblasts
DC1	Dendritic cells 1	Myofib.	Myofibroblasts
DC2	Dendritic cells 2	<i>Col14a1</i> <sup>+</sup> fib	<i>Col14a1</i> <sup>+</sup> fibroblasts
B cells	B cells	Pericyte 1	Pericytes 1
CD4 <sup>+</sup> T cells	CD4 <sup>+</sup> T cells	Pericyte 2	Pericytes 2
CD8 <sup>+</sup> T cells	CD8 <sup>+</sup> T cells	Fibromyo./SMCs	Fibromyocytes/Smooth muscle cells
NK cells	NK cells	Mesothelial	Mesothelial cells

representing ~12.5% of stromal cells by P14. While Pericytes 1 were not impacted by hyperoxia, the size of Pericytes 2 cluster was significantly decreased by P14 (Supplementary Figs. 2a and 5a, b).

Two distinct fibroblast clusters were identified (Fig. 4a–c). As proposed previously<sup>29</sup>, we categorized fibroblasts based on the expression of *Col13a1* and *Col14a1* (Fig. 2a, b).

The population of *Col13a1*<sup>+</sup> fibroblasts expressed additional *Col13a1*<sup>+</sup> fibroblast markers<sup>29</sup> and multiple pan-fibroblast markers, as well as the lipofibroblast marker *Plin2* (Supplementary Fig. 5c). The *Col13a1*<sup>+</sup> fibroblasts cluster represented ~60% of all stromal cells at P3 in both healthy and diseased lungs (Supplementary Fig. 2a; Supplementary Fig. 5a; Supplementary Data 2). While the size of the population rapidly decreased in healthy lungs, this decrease was largely absent in hyperoxia-exposed lungs. By P14, hyperoxia exposure induced large shifts in the expression profiles of *Col13a1*<sup>+</sup> fibroblasts (Fig. 4d; Supplementary Fig. 5b; Supplementary Data 7). The top marker for normoxic and hyperoxic *Col13a1*<sup>+</sup> fibroblasts were *Inmt* and *Saa3*, respectively (Fig. 4e; Supplementary Fig. 5d). To confirm this finding, we further analyzed the expression pattern of *Inmt*/*INMT* and *Saa3*/*SAA3* RNA in *Col13a1*<sup>+</sup>/*COL13A1*<sup>+</sup> fibroblasts in mouse and human BPD lung tissues, respectively (Fig. 4f and Supplementary Fig. 5e, respectively). *Saa3* was increased in a lamb preterm lung injury model<sup>30</sup> and regulated *Pdgfra*<sup>31</sup>, which is critical for normal alveolar development, and have been reported to be decreased by hyperoxia-exposure<sup>32,33</sup>. In addition, decreased PDGFRA expression was associated with increased risk for male patients to develop BPD<sup>34</sup>. The expression of *Pdgfra* was specifically decreased by hyperoxia in both fibroblast clusters (Fig. 4e; Supplementary Fig. 5f; Supplementary Data 7). This was however not associated with sex genotype within the hyperoxic portion of the *Col14a1*<sup>+</sup> fibroblasts (Supplementary Fig. 5f). While further studies of SAA3-PDGFRA interactions are needed, these data support the

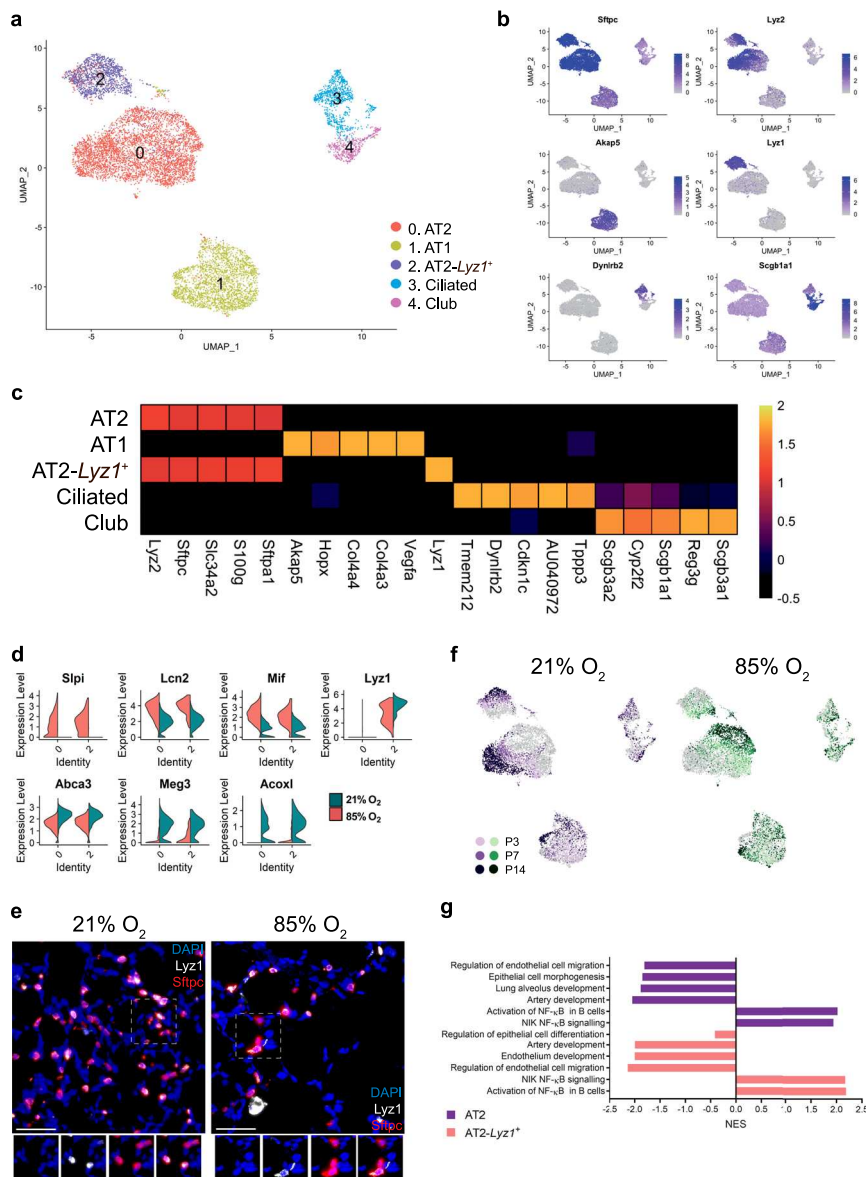
hypothesis that PDGFRA (possibly via regulation by SAA3) may play a role in hyperoxia-induced changes in *Col13a1*<sup>+</sup> fibroblasts.

The *Col14a1*<sup>+</sup> cluster expressed known markers of *Col14a1*<sup>+</sup> fibroblasts, including *Meg3*, *Dcn*, and *Fbln1*<sup>29</sup>. In addition, this population expressed *Coll1a1* and *Coll1a2*, consistent with the expression signature of interstitial matrix fibroblasts<sup>1,35</sup> (Supplementary Fig. 5c). Within the cluster, we detected a subpopulation of *Dcn*<sup>+</sup> cells co-expressing a progenitor marker *Ly6a* (*Sca1*), identifying a potential population of lung resident mesenchymal stromal cells (MSCs; Supplementary Fig. 5g)<sup>36</sup>. The *Ly6a*<sup>+</sup> population increased significantly between P3 and P14 in both healthy and BPD lungs. Correspondingly, gene expression levels of multiple MSC markers (*Cd44*, *Eng*, and *Lepr*) were increased in *Col14a1*<sup>+</sup> cluster by hyperoxia at P14 (Supplementary Fig. 5g).

Further, we identified one myofibroblast (Myofib.) cluster expressing *Tgfbi* and *Acta2* (Fig. 4b, c). While the size of the myofibroblast population gradually increased during development in the healthy lungs, this increase was less obvious in the diseased lungs (Supplementary Fig. 5a; Supplementary Data 2). This was accompanied by a major transcriptional shift in the myofibroblast population from diseased lungs at P14 (Fig. 4g). GSEA of hyperoxia-induced gene expression further suggested a hyperoxia-induced myofibroblast-mediated activation of extracellular matrix-related and immune pathways (Fig. 4h; Supplementary Data 8).

Finally, we identified an additional small population of fibromyocytes/smooth muscle cells (Fibromyo/SMCs). This cluster expressed a mixture of fibroblast, myofibroblast, and SMC markers, partially corresponding to the fibromyocyte population described by Travaglini et al.<sup>37</sup> (Fig. 4a–c). The Fibromyo/SMCs cluster displayed limited dynamic during the development and relatively few hyperoxia-induced changes in gene expression (Figs. 2a, 4d; Supplementary Data 7).





**Fig. 3 Cellular composition of epithelial cells during normal and hyperoxia-impaired late murine lung development.** **a** A total of five clusters of epithelial cells were identified in developing lungs. Cell populations are colored as indicated by the legend. **b** UMAP plots showing expression of principal identifiers of different epithelial cell types. The intensity of expression is indicated by purple coloring. **c** Heatmap of top five most differentially expressed genes across epithelial clusters. The intensity of expression is indicated as specified by the color legend. **d** Violin plots depicting changes in gene expression of some normoxia (21% O<sub>2</sub>, green) and hyperoxia-specific (85% O<sub>2</sub>, red) genes in the two AT2 clusters at P14. **e** Representative images from fluorescent RNA in situ hybridization for AT2-*Lyz1*<sup>+</sup> marker *Lyz1* (white) and pan-AT2 marker *Sftpc* (red) in developing mouse lungs. Magnification: 40×. Scale bar = 40 μm. Three 14-days old animals/group were analyzed. **f** UMAP plots depicting cell identity in regard to developmental time points in normally (purple) and aberrantly (green) developing lung epithelium. The intensity of expression is indicated as specified by the color legend. **g** Selected hyperoxia-impacted signaling pathways in AT2 (purple) and AT2-*Lyz1*<sup>+</sup> (pink) clusters, as identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted *p* value < 0.05) and normalized enrichment scores (NES) are shown. NES values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap and violin plots represent Z-score-transformed log(TP10k + 1) values. Expression levels in UMAP plots are presented as log(TP10k + 1) values. Log(TP10k + 1) corresponds to log-transformed UMIs per 10k.

Cell communication inference revealed *Col13a1*<sup>+</sup> fibroblasts, myofibroblasts and Pericytes 2 as potent signal senders and receivers among the stromal clusters in hyperoxia, essentially affecting all other cellular compartments (Fig. 2e). Hyperoxia-induced signals received by stromal cells largely originated from immune cells, particularly alveolar and interstitial macrophages, and from within the stromal compartment itself (Supplementary Fig. 6a). Similar to epithelial cells, this analysis revealed induction of the pro-inflammatory TNF signaling pathway in stromal cells,

particularly *Col13a1*<sup>+</sup> fibroblasts. Additionally, we observed an induction of *Anxa1*, reported to act as a regulator of TNF-induced proliferation and inflammatory responses in lung fibroblasts<sup>38</sup>. This is in agreement with the hyperoxia-mediated induction of multiple immune pathways in *Col13a1*<sup>+</sup> fibroblasts as suggested by GSEA (Fig. 4h; Supplementary Data 8). This increase in *Anxa1* expression was further associated with increased expression of its receptors *Cd44* and *Vcam1* in *Col13a1*<sup>+</sup> fibroblasts (Supplementary Fig. 6b; Supplementary Data 7). Increased expression of





**Fig. 4 Cellular composition of stromal cells during normal and hyperoxia-impaired late lung development.** **a** A total of six clusters of stromal cells were identified in developing lungs. Cell populations are colored as indicated by the legend. **b** UMAP plots showing expression of principal identifiers of different stromal cell types. The intensity of expression is indicated by purple coloring. **c** Heatmap of top five most differentially expressed genes across stromal clusters. The intensity of expression is indicated as specified by the color legend. **d** UMAP plots depicting cell identity in regard to developmental time points in normally (purple) and aberrantly (green) developing lung endothelium. The intensity of expression is indicated as specified by the color legend. **e** Dotplot depicting expression of oxygen-specific markers in *Col13a1*<sup>+</sup> fibroblasts at P14. The intensity of expression is indicated by the color legend. Size of the cell population expressing the gene of interest is indicated by the size of the circle as specified by the legend. **f** Fluorescent RNA in situ hybridization showing co-expression of *Inmt* (pink) and *Saa3* (pink) with *Col13a1* (green) in normal and aberrant mouse lungs. Magnification: 40×. Scale bar = 40 μm. Two 14-days old animals/group were analyzed. **g** Dotplot depicting expression of oxygen-specific markers in Myofib. cluster at P14. The intensity of expression is indicated by the color legend. Size of the cell population expressing the gene of interest is indicated by the size of the circle as specified by the legend. **h** Selected hyperoxia-impacted signaling pathways in *Col13a1*<sup>+</sup> fib (pink), Myofib. (yellow), and Pericytes 2 (blue) clusters as identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted *p* value < 0.05) and normalized enrichment scores (NES) are shown. NES values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap and violin plots represent Z-score-transformed log(TP10k + 1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k + 1) values. Log(TP10k + 1) corresponds to log-transformed UMIs per 10k.

**Hyperoxia induces inflammatory and anti-angiogenic gene expression in murine *Car4*<sup>+</sup> capillary endothelial cells.** We identified five distinct EC clusters based on their expression profiles. These included arterial, venous, capillary and lymphatic endothelial cells.

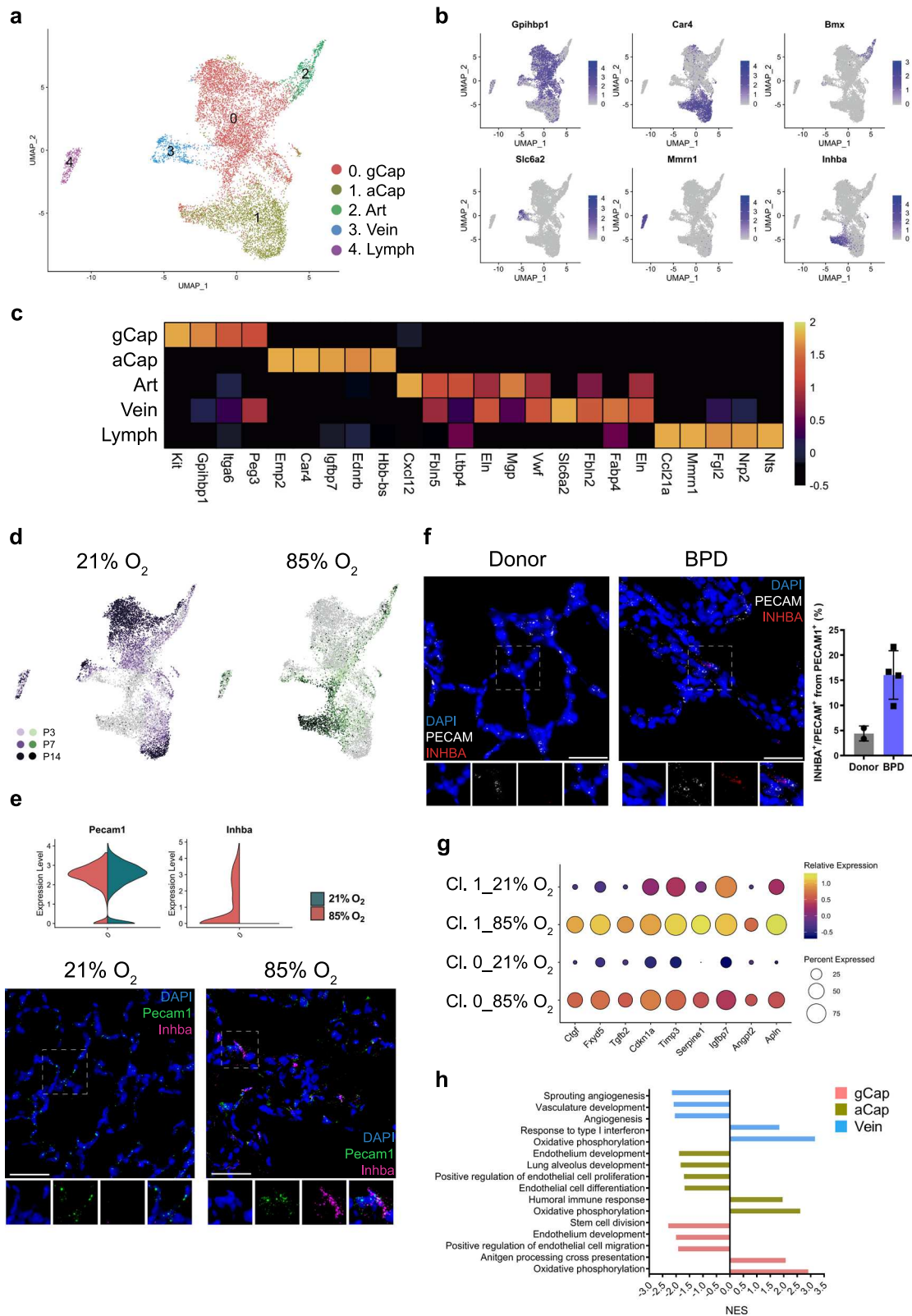
Capillary cells formed two distinct clusters gCap and aCap<sup>37,44,45</sup> (Fig. 5a–c; Supplementary Data 9), corresponding to capillary cell phenotypes recently characterized by Gillich et al.<sup>46</sup>. Throughout normal development, the proportion of gCap cells increased significantly (Supplementary Fig. 7a, b; Supplementary Data 2). In contrast to the gCap cluster, the number of aCap cells seemed to decrease with time, indicating the importance of this population in early postnatal development (Supplementary Fig. 7a, b; Supplementary Data 2). Hyperoxia significantly reduced the number of gCap cells (Supplementary Fig. 2a), but increased the number of aCap cells, distinguished by the expression of *Car4* (Fig. 5b–d; Supplementary Fig. 7a, b; Supplementary Data 2). *Car4*<sup>+</sup> endothelial cells contribute to lung septation<sup>47</sup> and alveolar revascularization after injury<sup>48</sup>. Our data suggest that hyperoxia particularly affected the gene expression of both gCap and *Car4*<sup>+</sup> aCap cells as assessed by the number of differentially regulated genes (Fig. 5d; Supplementary Fig. 7b; Supplementary Data 10). The top differentially regulated gene in the hyperoxic aCap cells was *Inhba* (Fig. 5e; Supplementary Data 10), a member of the TGFβ superfamily suggested to contribute to the pathology of BPD<sup>49</sup>. To validate this finding, we further showed the expression of *Inhba/INHBA* in *Pecam/PECAM* positive endothelial cells in both hyperoxic mouse and human BPD lung tissues and in healthy controls (Fig. 5e, f). Several genes known to be induced by cellular stress or inflammation<sup>50–52</sup> were upregulated in the hyperoxic aCap cells. We observed increased expression of *Ctgf* and *Fxyd5*, known to contribute to inflammatory lung injury<sup>53,54</sup> and of *Tgfb2*, shown to be associated with profibrotic responses in the lung<sup>55</sup> (Fig. 5g). In addition, an anti-angiogenic gene expression profile was observed in hyperoxia, characterized by an increase in expression of *Cdkn1a*, *Timp3*, *Serpine1*, and *Igfbp7* (Fig. 5g), in accordance with the crucial role of angiogenic growth factors in normal lung development<sup>56</sup>. Consistent with previous reports<sup>50</sup>, the cell cycle inhibitor *Cdkn1a*, which may protect the lung from oxidative stress<sup>57,58</sup>, was overexpressed in hyperoxia (Fig. 5g). Another potentially protective gene that was overexpressed in hyperoxia was *Apln*, which was shown to reduce potential pulmonary inflammation, fibrin deposition, and partially restore alveolarization in rat pups with neonatal hyperoxic lung injury<sup>59</sup> (Fig. 5g). GSEA of hyperoxia-induced effects in gCap and aCap cells suggested downregulation of several angiogenic, developmental,

and stem cell-related pathways in hyperoxia (Fig. 5h; Supplementary Data 11).

Cell communication inference suggested aCap to be an interactive cell type in both sending and receiving signals to and from other cell types in the lung in hyperoxia (Fig. 2e; Supplementary Fig. 7c, d). Several hyperoxia-associated gene expression changes in the aCap cells were predicted to be driven by ligands largely expressed by immune and stromal cells. We noticed *Bmpr2* upregulation in aCap cells along with significant upregulation of its ligand *Bmp5* in *Col13*<sup>+</sup> fibroblasts. *Bmpr2* protects endothelial cells from dysfunction and loss of *Bmpr2* leads to pulmonary arterial hypertension<sup>60</sup>. In addition, our analysis predicted activation of inflammatory immune response in aCap by activation of several receptors involved in immune cell adhesion (*Itgb1*, *Itga2*, *Icam1* and *2*, *Esam*) with simultaneous upregulation of their ligands in stromal cells such as *Col13*<sup>+</sup> fibroblasts, myofibroblasts, Pericyte 1 (*Tnc*) and immune cells such as Alveolar Macrophage, Neutrophil, Interleukin Macrophage, γδ T cells, and NK cells (*Itgam* and *Selp1g*) (Supplementary Fig. 7c, d).

In contrast to capillary endothelial cells, hyperoxia caused less changes in the venous and only minor changes in the arterial and lymphatic ECs (Fig. 5a–d; Supplementary Data 10). As described in the literature, the expression of arterial-specific genes formed a continuum, with the arterial gene expression present in the Cap cluster, whereas in the Vein cluster, the expression of arterial-specific genes was absent or considerably downregulated (Supplementary Fig. 7e). Venous cell GSEA suggested hyperoxia-induced upregulation of oxidative pathways and IFNγ related pathways, and downregulation of some angiogenic pathways (Fig. 5h; Supplementary Data 11). In hyperoxia, venous cells expressed several putative receptors involved in inflammatory immune response that were predicted to be activated by ligands produced by many immune and stromal cells (*Itgam*, *Selp1g*, *Apoe*, *Itga4*, *Anxa1*, *Vcam1*) (Supplementary Fig. 7c, d). In contrast, no significant hyperoxia-related interactions were predicted for the lymphatic and arterial cells.

The EC progenitors in the lung are not well understood<sup>61</sup>. gCap cells expressed *Kit*, a previously reported endothelial progenitor marker<sup>62</sup> (Supplementary Fig. 7f). Hyperoxia significantly reduced expression of *Kit* after P7, a crucial stage of capillary development in the lung (Supplementary Fig. 7f; Supplementary Data 10). Recently, gCap cells were shown to have progenitor properties and contribute to maintenance and repair of capillary endothelium<sup>46</sup>. In contrast to *Kit*, the expression of *Bst1*<sup>63</sup> and *Procr*<sup>64</sup>, which are suggested to be specifically expressed by EC progenitors, were upregulated in the lungs by hyperoxia, but the cells did not form a unique progenitor



cluster (Supplementary Fig. 7g). This is in concordance with recent findings in adult lungs showing that the regenerative potential of endothelial cells was not restricted to a particular subset of cells<sup>48</sup>.

**Hyperoxia exposure activates murine myeloid macrophages and neutrophils.** Immune cell clusters were identified<sup>1,2,65,66</sup>

(Fig. 2a) and grouped as belonging to either myeloid or lymphoid lineage. During normal lung development, the relative proportion of lymphoid cells from total lung cells increased with time (Fig. 6a). In hyperoxia, the myeloid cells remained the most abundant immune cell type in the lung during development.

We identified eight distinct clusters of myeloid cells, including macrophages, monocytes, neutrophils, dendritic cells (DCs), and



**Fig. 5 Cellular composition of lung endothelium during normal and hyperoxia-impaired late lung development.** **a** A total of five clusters of endothelial cells were identified in developing lungs. Cell populations are colored as indicated by the legend. **b** UMAP plots of principal identifiers of different types of endothelial cells. The intensity of expression is indicated by purple coloring. **c** Heatmap of top 5 most differentially expressed genes across endothelial clusters. The intensity of expression is indicated as specified by the color legend. **d** UMAP plots depicting cell identity in regards to developmental time points in normally (purple) and aberrantly (green) developing lung endothelium. The intensity of expression is indicated as specified by the color legend. Violin plots depicting changes in gene expression of *Pecam* and *Inhba* in normoxia (21% O<sub>2</sub>, green) and hyperoxia-specific (85% O<sub>2</sub>, red) at P14. **e** Representative images from fluorescent RNA in situ hybridization for hyperoxia-specific marker *Inhba* (pink) and pan-endothelial marker *Pecam* (green) in developing mouse lungs. Magnification: 40×. Scale bar = 40 μm. Two 14-days old animals/group were analyzed. **f** Representative images and quantitative analysis of fluorescent RNA in situ hybridization for hyperoxia-specific marker *INHBA* (red) and pan-endothelial marker *PECAM* (white) in BPD patients' and donors' lungs. Samples from five BPD patients and two donor lungs were analyzed. For quantitative analysis, cells in 15 randomly chosen fields of view/sample were analyzed. Magnification: 40×. Scale bar = 40 μm. **g** Dot plot depicting hyperoxia-induced changes in the expression of inflammatory, anti-angiogenic and protective genes in capillary cells at P14. The intensity of expression is indicated by the color legend. Size of the cell population expressing the gene of interest is indicated by the size of the circle as specified by the legend. **h** Hyperoxia-impacted signaling pathways in gCap (pink), aCap (yellow), and Vein (blue) clusters, as identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted *p* value < 0.05) and normalized enrichment scores (NES) are shown. NES values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap represent Z-score-transformed log(TP10k + 1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k + 1) values. Log(TP10k + 1) corresponds to log-transformed UMIs per 10k.

basophils (Fig. 6b–d, Supplementary Data 12). Hyperoxia dramatically changed the cell distribution, particularly in the macrophage populations (Fig. 6e; Supplementary Fig. 8a, b).

We identified two distinct macrophage populations, representing alveolar and interstitial macrophages based on the expression of *Clqb*, *Lgals3*, and *Adgre1* genes<sup>37,67</sup> (Fig. 6b; Supplementary Fig. 8c). In hyperoxia, the number of Alv Mφ cells decreased, whereas the number of Int Mφ cells increased (Supplementary Fig. 8b; Supplementary Data 2). Hyperoxia induced the expression of *Inhba* and *Marco* in Alv Mφ population, whereas the expression of *Ear* was decreased (Supplementary Fig. 8d; Supplementary Data 13). To further validate this finding, we showed that in lung tissue sections from hyperoxic mice and BPD patients, the expression of *Inhba/INHBA* and *Marco/MARCO* was more strongly expressed in *Ptpnc/PTPNC* + leukocytes morphologically resembling alveolar macrophages, compared to the healthy control lungs (Fig. 6f and Supplementary Fig. 8e, respectively). GSEA of hyperoxia-induced expression changes in Alv Mφ cells suggested activation of pathways involved in inflammatory response, leukocyte migration, response to wounding and oxidative phosphorylation (Fig. 6g; Supplementary Data 14). Cell communication inference suggested Alv Mφ as interactive cell types in sending and receiving signals from other cell types in the lung in hyperoxia (Fig. 2e; Supplementary Fig. 9a, b).

The Int Mφ population expressed high levels of inflammatory mediators, consistent with data by Gibbings et al.<sup>68</sup> (Supplementary Data 12). The hyperoxic macrophages clustered together in a particular region of Int Mφ cluster (Fig. 6e). GSEA of hyperoxia effects in Int Mφ cells suggested upregulation of pathways involved in oxidative phosphorylation, innate immune response, chemotaxis, and antigen presentation via MHC class I (Fig. 6g; Supplementary Data 14). Of all the lung cell types, cell communication inference suggested Int Mφ, along with *Coll3a*<sup>+</sup> fibroblasts to be strong mediators of paracrine signaling, releasing factors putatively driving gene expression changes across many cell types (Fig. 2e; Supplementary Fig. 9a, b).

Interestingly, we observed upregulation of *Ccr2* and *Csf1r* in Int Mφ population after hyperoxia exposure. Furthermore, the expression of their ligands, *Ccl2* and *Csf1*, was upregulated in the Neut and Alv Mφ, respectively (Supplementary Fig. 9c). These observations are consistent with the study by Kalymbetova et al., where *Ccr2*<sup>-/-</sup> and CSF1R-depleted mice developed milder structural changes in hyperoxia-exposed lungs<sup>69</sup>.

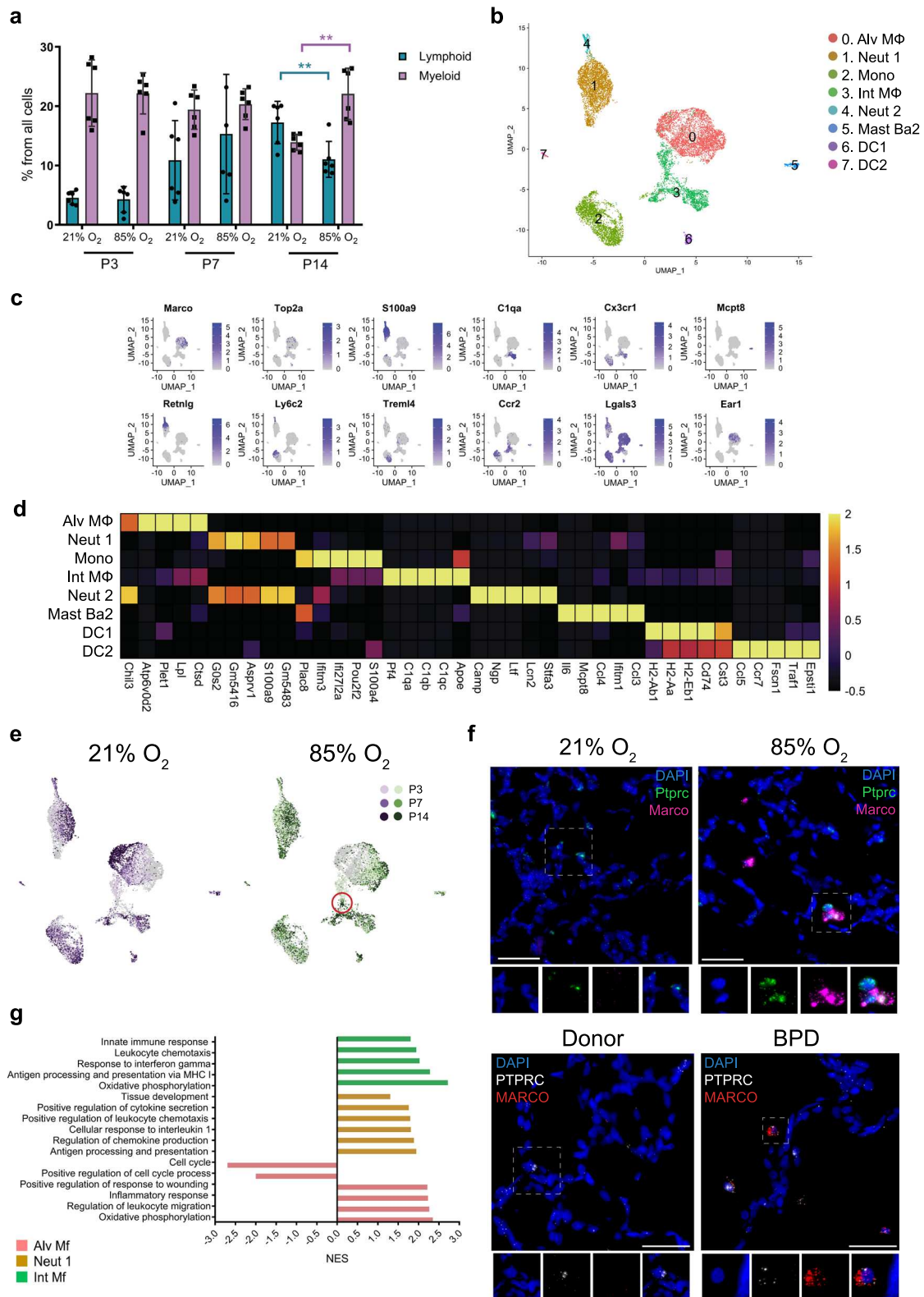
The concept of M1/M2 polarization of macrophages has been linked to normal development, as well as to several lung pathologies, including BPD<sup>2,70</sup>. Consistent with previous studies,

our results suggested that hyperoxia slightly enhanced the M1 signature of both Int Mφ and Alv Mφ populations and decreased the M2 signature of Alv Mφ population (Supplementary Fig. 9d). In hyperoxia, several receptors for both M1 (*Ifngr1*, *Ifngr2*, *Adrb2*, *Tnfrsf1b*) and M2 (*Il4ra*, *Il13ra*), as well as receptors associated with both M1- and M2- pathways (*Plaur*, *Cd44*)<sup>71</sup>, were predicted to activate Alv Mφ cells by ligands expressed from many types of immune, stromal cells, and Cap-a cells according to cell communication predictions (Fig. 2e; Supplementary Fig. 9a, b).

Monocytes are commonly identified as functionally distinct groups of classical or non-classical monocytes. In our study, we found a single monocyte cluster, part of which was positive for classical, and part for non-classical monocyte markers (Fig. 6b; Supplementary Fig. 9a). The number of classical monocytes defined by the expression of *Ly6c2* and *Ccr2* increased after hyperoxia exposure, whereas the number of non-classical monocytes defined by the expression of *Fcgr4* was reduced (Supplementary Fig. 9e). These results suggest that, in addition to the hyperoxia-induced upregulation of the *Ccl2-Ccr2* axis in macrophages, the number of *Ccr2* expressing classical inflammatory monocytes also increased.

We identified two neutrophil clusters: Neut 1 and 2 (Fig. 6b). Exposure to hyperoxia caused a considerable increase in the size of the Neut 1 cluster and considerable hyperoxia-induced changes in gene expression (Supplementary Fig. 8b; Supplementary Data 2 and 13). GSEA results suggest hyperoxia-induced upregulation of several pathways involved in immune response in these cells (Fig. 6g; Supplementary Data 14). The size and transcriptional profile of the Neut 2 cluster showed fewer changes in hyperoxia compared with Neut 1 (Supplementary Data 13 and 14). Cell communication inference suggested Neut 1 cells as one of the most active cell types in sending signals to other cell types in the hyperoxic lung (Fig. 2e; Supplementary Fig. 9a, b).

**Hyperoxia activates murine lymphoid NK, CD8<sup>+</sup>, and CD4<sup>+</sup> T cells.** Based on known cell markers, we identified nine lymphoid cell clusters including B cells, T cells, NK cells and ILC2 (Fig. 7a–c; Supplementary Data 15). Hyperoxia exposure considerably decreased the proportion of B-cell and CD4<sup>+</sup> T-cell populations in the lung at P14, whereas the number of NK cells increased significantly (Fig. 7d, e). Differential gene expression and GSEA suggested that the largest changes in expression occurred in the B cell and CD8<sup>+</sup> T-cell populations and the altered pathways were involved in inflammatory response, response to INFγ, cell cycle, and oxidative phosphorylation



(Fig. 7e, f; Supplementary Data 16 and 17). NK cells, CD8<sup>+</sup> T cells, and Treg cells were predicted to be interactive lymphoid cell types, sending signals to other cell types in the lung driving changes associated with hyperoxia (Fig. 2e). Our results suggest that hyperoxia altered the normal development of adaptive lung immunity due to the hyperoxia-induced activation of the innate immune system, particularly causing a reduction in the number of B-cells and CD4<sup>+</sup> T cells, resulting in the myeloid

cells remaining the major immune cell population of the lung at P14.

**Hyperoxia alters the expression profile of murine mesothelial cells.** A single mesothelial cluster (Fig. 2a, b) was identified based on the expression of *Msln* and other known markers<sup>35</sup> (Fig. 8a, Supplementary Data 18). Although mesothelial cells represented

**Fig. 6 Cellular composition of lung myeloid populations during normal and hyperoxia-impaired late lung development.** **a** The relative proportion of myeloid (purple) and lymphoid (teal) cells in developing lungs was significantly impacted by hyperoxia exposure.  $n = 6$  animals/group. Data are presented as means  $\pm$  SD. Statistical analyses were performed with GraphPad Prism 8.0. Significance for each population at each time point was evaluated by unpaired, two-tailed Student's *t*-test.  $P$  value = 0.0013 for Myeloid population, and 0.0086 for Lymphoid population. **b** A total of eight clusters of myeloid cells were identified in developing lungs. Cell populations are colored as indicated by the legend. **c** UMAP plots of principal identifiers of different types of myeloid cells. The intensity of expression is indicated by purple coloring. **d** Heatmap of top five most differentially expressed genes across myeloid clusters. The intensity of expression is indicated as specified by the color legend. **e** UMAP plots depicting cell identity of myeloid cells in regard to developmental time points in normally (21% O<sub>2</sub>-exposed, purple) and aberrantly (85% O<sub>2</sub>-exposed, green) developing lung. Each cell is colored by mouse age as indicated by the legend. **f** Fluorescent RNA in situ hybridization showing co-expression of *Marco/MARCO* (pink/red) with *Ptpcr/PTPRC* (green/white) positive leukocytes morphologically resembling alveolar macrophages in normal and mouse hyperoxic/human BPD lungs, respectively. Magnification: 40 $\times$ . Scale bar = 40  $\mu$ m. Two 14-days old animals/group were analyzed and samples from five BPD patients and two donor lungs were analyzed. **g** Hyperoxia-impacted signaling pathways in Alv Mf (pink), Neut 1 (yellow) and Int Mf (green) clusters as identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted  $p$  value < 0.05) and normalized enrichment scores (NES) are shown. NES values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap and violin plots represent Z-score-transformed  $\log(\text{TP10k} + 1)$  values. Expression levels in UMAP plots are presented as  $\log(\text{TP10k} + 1)$  values.  $\log(\text{TP10k} + 1)$  corresponds to  $\log$ -transformed UMIs per 10k.

only 1.20% of all analyzed cells, a clear pattern of gradual decrease was observed in healthy lungs between P3 and P14, which was absent in hyperoxic lungs (Fig. 8b). The developmental arrest induced by hyperoxia was associated with multiple changes in gene expression and signaling pathways (Fig. 8c, d; Supplementary Data 19 and 20). We observed an increase in the pro-angiogenic factor *Angptl2*, accompanied with a decrease in anti-angiogenic *Igfbp6* (Fig. 8d). Among the most de-regulated expression profile were also several components of the extracellular matrix: *Timp1*, *Timp3*, and *Fbln5* (Fig. 8d). FBLN5 is critical for normal alveolar development since *Fbln5*<sup>-/-</sup> mice exhibited arrest in alveolarization<sup>41</sup>. Aberrant *Timp1* expression was reported in ventilated preterm human lungs and murine BPD models<sup>41</sup>, while increased *Timp3* expression was associated with BPD severity<sup>72</sup>.

## Discussion

The heterogeneity and origin of lung cell lineages in the context of lung development have been the focus of significant research efforts during the last decades. Here we provide an extensive profiling of cellular composition in normal and impaired late lung development using multiplexed scRNA-seq to assess the expression profile of 66,200 cells. Our study provides insight into the pathogenesis of impaired alveolarization by characterizing several pathological cell populations with distinct molecular expression profiles.

We followed the developing lung through three crucial time points of late lung development, across which we identified 5 epithelial, 6 stromal, 5 endothelial, 8 myeloid, 9 lymphoid, and 1 mesothelial cell clusters. Cluster annotations were largely consistent with previously published data<sup>1,2,73</sup>. By assessing tissues at multiple time points, we captured developmental trajectories of these populations during normal and impaired lung development. Hyperoxia exposure altered all cellular compartments, particularly alveolar epithelium, stromal fibroblasts, capillary endothelium, and macrophage populations. Analysis of BPD-associated genes pointed to marked changes in expression in many of the cell populations sensitive to hyperoxia in our data, supporting the importance of these cell populations in the disease process. Pathway analysis and predicted dynamic cellular crosstalk suggested inflammatory signaling as the main driver of hyperoxia-induced changes (Fig. 9).

Importantly, exposure to hyperoxia impaired the composition and expression patterns of all cellular compartments necessary for normal alveolarization. The changes were gradual, appearing from P7 on. Multiple affected transcriptional programs were related to activation of the inflammatory response, suggesting inflammation as one of the main drivers of hyperoxia-induced changes. Previously, fibroblasts (*Col13*<sup>+</sup> and *Col14*<sup>+</sup>), SMCs,

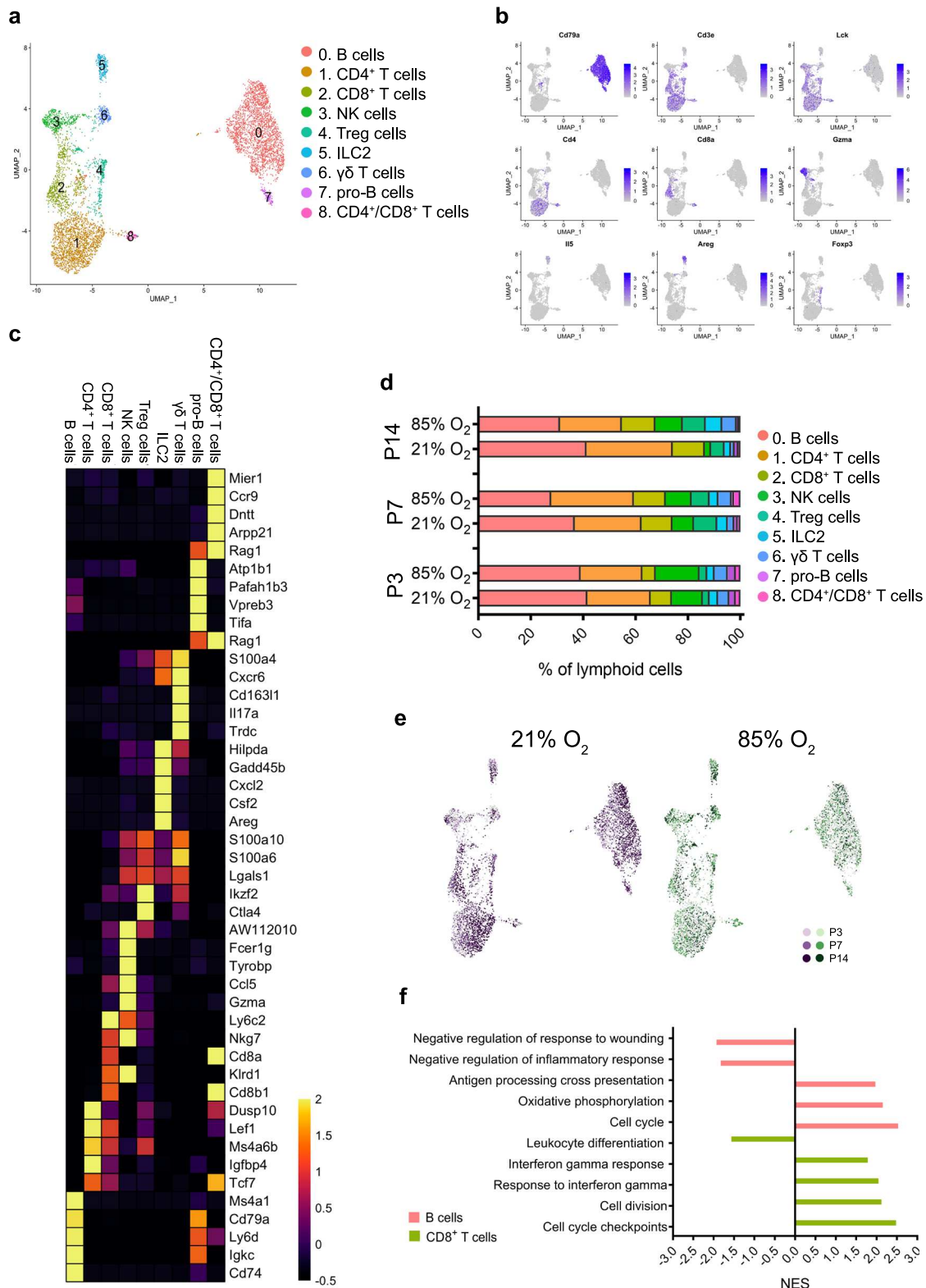
macrophages, AT1, and EC lymph cells were predicted to have the strongest network of interactions in lung homeostasis<sup>74</sup>. Our analysis identified hyperoxia-induced interactions between cell compartments in the lung, pointing out lung cell types most actively involved in cellular crosstalk in hyperoxia and specifying the activated receptor pathways. We included selected analyses of interactions induced by hyperoxia, but our dataset also allows further systems level analyses in homeostasis during normal development, not addressed here.

Within epithelial clusters, AT2 and AT2-*Lyz1*<sup>+</sup> cells were interactive in both sending and receiving signals in the diseased lungs. Signals induced in AT2 clusters originated from stromal, endothelial, and immune compartments.

Within the stroma, Pericytes 2, *Col13a1*<sup>+</sup> fibroblasts, and myofibroblasts found in diseased lungs showed prominent hyperoxia-induced gene expression pattern changes. *Col13a1*<sup>+</sup> fibroblasts and Pericytes 2 were further predicted among the most active signal receivers and senders within the developing lung stroma in hyperoxia. Additional signals were received by these cells predominantly from macrophages and epithelial populations. Most dominant pathways included pro-inflammatory and extracellular matrix signaling. Our analysis suggests an important role for *Col13a*<sup>+</sup> fibroblasts, pericytes and myofibroblasts in hyperoxia exposure.

Particularly striking was a loss of healthy capillary endothelial cells in diseased lungs. Our data support previous studies showing that capillaries are most vulnerable to hyperoxic conditions<sup>61,75,76</sup>. Importantly, hyperoxia seemed to reduce the number of gCap cells, which are putative distal lung vascular progenitor cells<sup>46</sup>. The depletion of these cells may contribute to the lack of repair capability of the injured preterm lung and lead to the occurrence of pulmonary vascular disease<sup>56</sup> and early onset emphysema in BPD patients<sup>77</sup>, lending some rationale for endothelial-cell-derived therapies<sup>61</sup>. Simultaneously, *Car4*<sup>+</sup> capillary endothelial cells (aCap)<sup>46</sup>, which are important for normal alveolar development and in alveolar revascularization post-injury<sup>47,48</sup>, increased in number. aCap cells displayed a pathological gene expression profile in hyperoxia characterized by upregulation of *Inhba* and many other pro-inflammatory as well as anti-angiogenic markers. aCap were predicted as one of the most active endothelial cell types in hyperoxia-induced cellular cross-talk. The predicted interactions took place with many immune cell types, smooth muscle cells, and epithelial cells. The study by Niethamer et al. described the predicted adult mouse endothelial cell interactions between endothelium and epithelium, using primary epithelial lineages as their epithelial dataset<sup>48</sup>. The authors' analysis predicted strong interactions between AT1 and *Car4*<sup>+</sup> capillary endothelial cells after influenza exposure<sup>48</sup>,



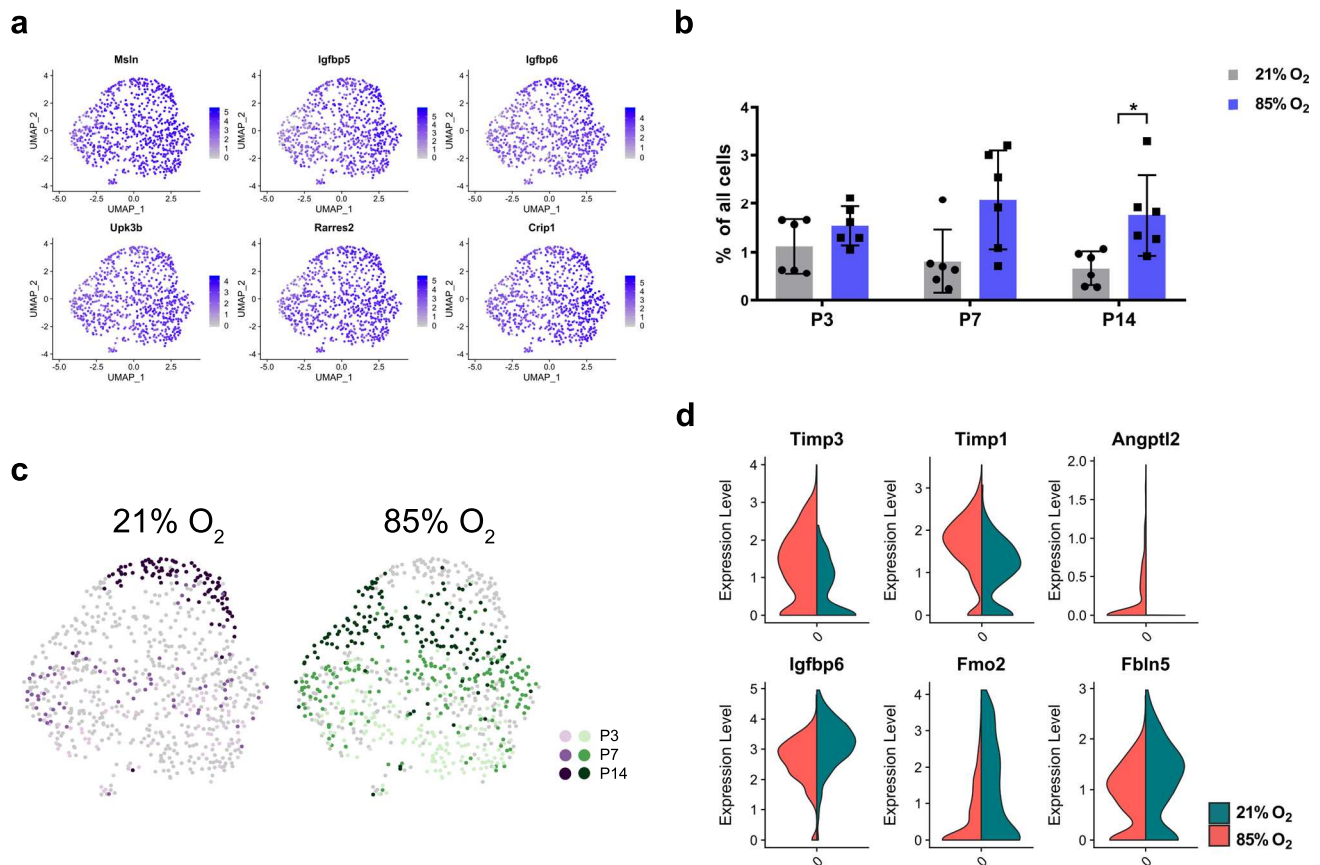


whereas our analysis did not predict any significant hyperoxia-induced interactions between AT1 and aCap cells. However, our data predicted *Angpt2*-mediated signaling between *Car4*<sup>+</sup> aCap cells and AT2 cells.

Our study highlights the importance of activation of the inflammatory response in impaired lung development caused by hyperoxia exposure, which has been previously indicated<sup>78-80</sup>. The majority of hyperoxia-induced signals in the lung were

involved in the inflammatory response. In addition to activation of many immune cell types, hyperoxia-induced activation of subpopulations of endothelial, stromal, and epithelial cells, all of which participated in inflammatory signaling. In particular, our data identified innate immune cells such as macrophages and neutrophils as important immune cells in the dynamic cellular cross-talk during hyperoxia. Hyperoxia induced both pro- and anti-inflammatory activation of macrophages, resulting in active

**Fig. 7 Cellular composition of lung lymphoid populations during normal and hyperoxia-impaired late lung development.** **a** A total of nine clusters of lymphoid cells were identified in developing lungs. Cell populations are colored as indicated by the legend. **b** UMAP plots of principal identifiers of identified types of lymphoid cells. The intensity of expression is indicated by purple coloring. **c** Heatmap of top five most differentially expressed genes across the lymphoid clusters. The intensity of expression is indicated as specified by the color legend. **d** Relative contribution of individual lymphoid clusters changed significantly during development and exposure to hyperoxia.  $n = 6$  animals/group. Cell populations are colored as indicated by the legend. **e** UMAP plots depicting cell identity in regard to developmental time points in (21% O<sub>2</sub>-exposed, purple) normally and aberrantly (85% O<sub>2</sub>-exposed, green) developing lymphoid populations. Each cell is colored by mouse age as indicated by the legend. **f** Hyperoxia-impacted signaling pathways in B cells (pink) and CD8<sup>+</sup> T cells (bright green) clusters as identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted  $p$  value < 0.05) and normalized enrichment scores (NES) are shown. NES values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap represent Z-score-transformed  $\log(\text{TP10k} + 1)$  values. Expression levels in UMAP plots are presented as  $\log(\text{TP10k} + 1)$  values.  $\log(\text{TP10k} + 1)$  corresponds to log-transformed UMIs per 10k.



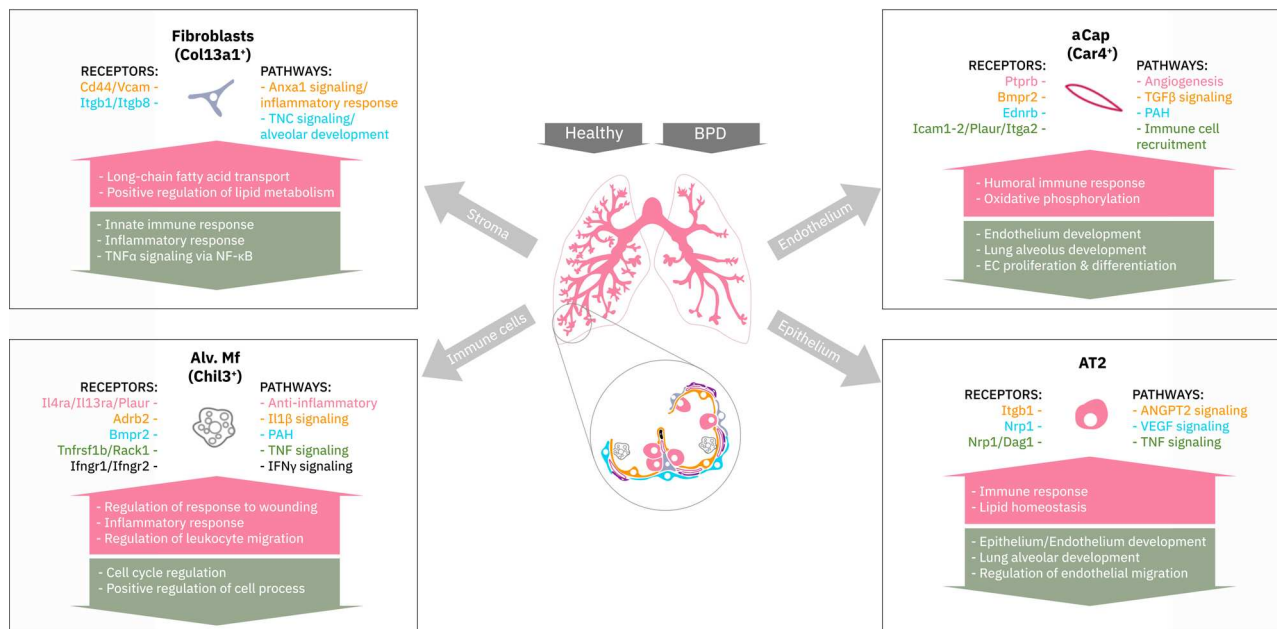
**Fig. 8 Cellular composition of lung mesothelium during normal and hyperoxia-impaired late lung development.** **a** UMAP plots of principal identifiers of mesothelial cells in developing lung. Intensity of expression is indicated by purple coloring. **b** Percentual proportion of mesothelial cells in normally (21% O<sub>2</sub>-exposed, gray) and aberrantly (85% O<sub>2</sub>-exposed, purple) developing lungs at P3, P7 and P14.  $n = 6$  animals/group. Significance was evaluated by multiple unpaired multiple Student's  $t$ -test with Holm-Sidak correction. Data are presented as means  $\pm$  SD.  $P$  value = 0.0426. **c** UMAP plots depicting cell identity in regard to developmental time points in (21% O<sub>2</sub>-exposed, purple) normally and aberrantly (85% O<sub>2</sub>-exposed, green) developing lung mesothelium. Each cell is colored by mouse age as indicated by the legend. **d** Exposure to 85% O<sub>2</sub> altered gene expression in developing lung mesothelium. 21% O<sub>2</sub>-exposed: green, 85% O<sub>2</sub>-exposed: red. Expression values in violin plots represent Z-score-transformed  $\log(\text{TP10k} + 1)$  values. Expression levels in UMAP plots are presented as  $\log(\text{TP10k} + 1)$  values.  $\log(\text{TP10k} + 1)$  corresponds to log-transformed UMIs per 10k.

signaling from macrophages to other lung cell types. Many of these signaling pathways can be targeted by mesenchymal stromal cell-based therapies, specifically the modulation of macrophages toward an anti-inflammatory phenotype<sup>70</sup>. Also, activation of neutrophils resulted in major inflammatory signaling. In addition to the innate immune cells, the recruitment of lymphoid cells such as NK cells as well as CD8<sup>+</sup> and CD4<sup>+</sup> T cells induced inflammatory pathways in many target cell types in hyperoxia. The role of immune cell signaling, such as IL-1 $\beta$ <sup>81</sup>, IFN $\gamma$ <sup>82</sup>, and TGF $\beta$ <sup>83</sup> has been shown in impaired lung development by several studies. Our data support the importance of these pathways in

hyperoxia, as we showed and addressed their activation in specific cell types.

Here we provide data-driven transcriptomic evidence of the most important cell subpopulations and molecular pathways responsible for impaired late lung development caused by hyperoxia exposure. We demonstrate extensive changes in cellular composition caused by hyperoxia, and identify crosstalk between immune, stromal, endothelial, and epithelial cells resulting in impaired lung development. Furthermore, to reflect the relevance of our murine data on the most common chronic lung disease of infancy, BPD, we have validated some of our





**Fig. 9 Summary of the cellular crosstalk and pathway analysis in hyperoxia-impaired late murine lung development.** Mouse lung cell subtypes with the most significant gene expression changes in hyperoxia mimicking bronchopulmonary dysplasia (BPD). Activated receptors and their involved biological pathways are indicated by the cell communication inference analysis. Biological processes in the arrows are indicated by the gene set enrichment analysis.

findings in lung tissues from BPD patients. Limitations of the study include bias from tissue dissociation and bias caused by the low number of cells used for quantification in scRNAseq compared to traditional single-cell methods, such as FACS. These may limit the use of scRNA-seq data for cell quantification. We assessed the magnitude of this bias by stereological cell counting (a gold standard for lung cell quantification) of AT2 cells as an example. However, our analysis is limited to one cell type only. In the future, improved knowledge of markers for cell subpopulations and morphological quantification of these cell populations may provide further insight into the lung dissociation bias. Another limitation was the unequal sex distribution of animals at P3 and P7. However, the largest transcriptional changes were observed at P14, where the sex distribution was equal among study samples. Next, other systems biology approaches such as spatial transcriptomics will allow the construction of a more complete map of events leading to impaired alveolarization. Finally, analyzing the recovery phase after hyperoxia exposure will identify the most severely impaired cell populations unable to recover, which may be responsible for the permanent changes in the lung architecture. In this study, we have described multiple aspects of hyperoxia exposure and identified pathological pathways as putative drug targets of BPD.

## Methods

**Experimental animals.** Pregnant C57BL/6 at embryonic day (E)14 or E17 were purchased from Charles Rivers Laboratories, Saint Constant, QC, Canada. Mice were housed by the Animal Care and Veterinary Service of the University of Ottawa in accordance with institutional guidelines. Newborn mouse pups from dams that delivered on the same day, were randomized at day of birth [postnatal day (P) 0] and divided to equal-sized litters of 6 to 8. Following randomization, mice cages were either maintained in room air (normoxia, 21% O<sub>2</sub>) or in normobaric hyperoxia (85% O<sub>2</sub>) from P0 until day of harvest. The hyperoxic environment was maintained in sealed plexiglass chambers with continuous oxygen monitoring (BioSpherix, Redfield, NY). In order to avoid oxygen toxicity and associated confounding factors, nursing dams were rotated between normoxic and hyperoxic group every 48 h. All mice were maintained in 12/12 h light/dark cycle and received food *ad libitum*. All developing mice and their nursing dams were euthanized either at P3, P7 or P14 by 10 µl/g intraperitoneal (i.p.) injection of Pentobarbital Sodium (CDMV, Saint-Hyacinthe, QC, Canada). Animals designated for scRNA-seq and FACS analyses received an additional i.p. injection of 10 mU/g

heparin sodium (LEO Pharma Inc., Thornhill, ON, Canada). All animal procedures were approved by the Animal Care Committee of the University of Ottawa under animal ethics protocol OHRI-1696.

**Sex genotyping of mice.** Determination of sex in mouse pups was performed as described previously<sup>84</sup>. Briefly, genomic DNA was isolated from mice tail cuts and regions of interest were amplified by polymerase chain reaction (PCR) in order to determine expression of male-specific Sry gene, as well as *Ii3* gene present in mice of both sexes. Forward and reverse primer sequences are listed in Supplementary Table 2. Amplified sequences were visualized by ethidium bromide on 1.5% agarose gel.

**Lung isolation.** Pups were euthanized at P3, P7, or P14 by 10 µl/g i.p. injection of pentobarbital sodium (CDMV, Saint-Hyacinthe, QC, Canada). Following euthanasia, mice were tracheotomized and lungs were installation-fixed for 5 min at 20 cm H<sub>2</sub>O hydrostatic pressure with 1.5% (w/v) paraformaldehyde (PFA) (Sigma-Aldrich, Oakville, ON, Canada) and 1.5% (w/v) glutaraldehyde (Sigma-Aldrich, Oakville, ON, Canada) in 150 mM HEPES (Sigma-Aldrich, Oakville, ON, Canada) fixation solution with pH 7.4. After isolation, lungs were kept in the fixation solution for 48 h at 4 °C and collected for embedding in paraffin. Paraffin-embedded tissue blocks were sectioned at 4 µm and stained with hematoxylin and eosin (H&E) stain. Tissue dehydration, paraffin embedding, sectioning, and staining were performed by the University of Ottawa Louis Pelletier Histology Core Facility.

**Lung volume measurement.** Inflated and fixated lungs were carefully dried with soft tissue in order to remove any droplets. Lung volume was assessed by Archimedes principle (water displacement). The mean value of three measurements/lung was considered.

**Mean linear intercept (MLI) measurement.** The MLI was estimated in a blinded fashion using the Quorum Analysis (Quorum Technologies Inc., Guelph, ON, Canada) software. Briefly, MLI quantification was performed in a semi-automated fashion using a 155.34 µm line grid moving through sections of interest in defined intervals, where number of intersections between grid line located within alveolar parenchyma and alveolar walls was noted. The average MLI was computed using the formula:  $MLI = (FOVs \times \frac{155.34}{I})$ , where FOVs = fields of view within which intersections were counted, I = number of intersections and 155.34 = the length of the grid line. A total of at least 200 FOVs were assessed in each lung, corresponding to 10 sections analyzed from lungs at P3, 6 sections from lungs at P7 and 4-5 sections from lungs at P14.

**Stereological estimation of number of alveolar type II cells.** The number of AT2 cells was determined in the lungs by stereological principles as described previously<sup>85</sup>. Developing pups were euthanized at P14 by 10 µl/g i.p. injection of

pentobarbital sodium (CDMV, Saint-Hyacinthe, QC, Canada), tracheotomized and their lungs were installation-fixed for 5 min at 20 cm H<sub>2</sub>O hydrostatic pressure with 4% (w/v) paraformaldehyde (PFA) (Sigma-Aldrich, Oakville, ON, Canada), pH 7.4. Following fixation, lungs were stored in PFA solution for 48 h at 4 °C, embedded *in toto* in 2% agar (Diamed, Mississauga, ON, Canada) and sliced into 2 mm slices. Lung volume was then estimated by the Cavalieri principle exactly as described before<sup>84</sup>. Sliced pieces of each lung were collected for embedding in paraffin. Each paraffin block, containing one pair of lungs from one mouse, was sectioned in agreement with the rules of serial uniform random sampling, where pairs of consecutive, 3 μm thick sections, were collected every 200 sections throughout the block. All sections were stained for the AT2 marker Prosurfactant protein C (ProSPC) and quantified using Stereo Investigator<sup>®</sup> software (MBF Bioscience, Williston, VT, USA). Essentially, AT2 cells were counted in all consecutive sections at 40× magnification by dissector counting. For every pair of lungs, 0.5% of total surface area was analyzed. The number of cells was calculated using the following formula (1):  $N_{(AT2)} = \frac{1}{ssf} \times \frac{1}{asf} \times \frac{1}{2} \times \sum Q_{(AT2)}$ , where *ssf* = slide sampling fraction ( $\frac{1}{200}$ ), *asf* = area sampling fraction ( $\frac{1}{200}$ ) and  $\sum Q_{(AT2)}$  = number of AT2 as counted in both direction in the dissector setting. The number of AT2 cells was normalized to total surface area of the lung. Surface area was calculated as (2):  $S_v \times N_{par} \times V_{lung}$ , where *S<sub>v</sub>* = surface density in mm<sup>-1</sup>, *N<sub>par</sub>* = parenchymal fraction and *V<sub>lung</sub>* = lung volume in mm<sup>3</sup> as estimated by Cavalieri principle. Parenchymal fraction was assessed as described previously. Surface density was assessed stereologically as described before and calculated using following formula (3):  $S_v = \frac{2 \times I}{l_p \times P}$ , where *S<sub>v</sub>* = surface density in mm<sup>-1</sup>, *I* = number of intersections of probe with alveolar surface, *l<sub>p</sub>* = length of probe/point and *P* = number of points of the probe falling within parenchymal region of the lung<sup>84</sup>. Tissue dehydration, paraffin embedding, sectioning, and immunohistochemistry were performed by the University of Ottawa Louis Pelletier Histology Core Facility.

**Immunohistochemistry.** Briefly, 3 μm thick paraffin sections were deparaffinized in xylene and rehydrated in decreasing ethanol series. Retrieval was accomplished using an ethylenediaminetetraacetic acid (EDTA) buffer (Bond epitope retrieval solution 2, Leica, Concord, ON, Canada), pH9 for 20 min. Sections were stained for proSP-C using an anti-proSP-C antibody (Millipore/Sigma, Entobicoke, ON, Canada) at 1:1500 dilution for 30 min followed by detection with horseradish peroxidase (HRP)-conjugated polymer system (Bond Polymer Refine Detection Kit, Leica, Concord, ON, Canada). All sections were then stained using 3,3'-Diaminobenzidine (DAB) as chromogen, counterstained with Hematoxylin, and mounted.

**Human tissues.** Paraffin-embedded lung sections from BPD patients and age-matched donors were kindly provided by the LungMAP Human Tissue Core, Biorepository for Investigation of Neonatal Diseases of Lung-Normal (BRINDL-NL, Dr. Gloria Pryhuber). Written informed consent was obtained from the parents/guardians of the minor participant. Ethical approval was permitted by the University of Rochester, Rochester, NY. Detailed information about all samples is provided in Supplementary Data 21.

**Fluorescent RNA in situ hybridization.** RNA in situ hybridization was performed for target detection on fresh 4% PFA-fixed paraffin embedded 3 μm tissue sections using RNAscope Multiplex Fluorescent Reagent Kit Version 2 (Advanced Cell Diagnostics, Newark, CA, USA) according to the manual. Firstly, tissue sections were baked for 1 h at 60 °C, then deparaffinized and treated with hydrogen peroxide for 10 min at room temperature (RT). Target retrieval was performed for 15 min at 98 °C, followed by protease plus treatment for 15 min at 40 °C. All probes were hybridized for 2 h at 40 °C followed by signal amplification and developing of HRP channels was done according to manual. The following RNAscope probes were used in the study: 3-Plex negative control probe dapB (#320871), 3-Plex positive control probe\_Mm (#320881), Mm-Inhba-C2 (#455871-C2), Mm-Inmt (#486371), Mm-Marco (#510631), Mm-Pecam1 (#316721), Mm-Saa3 (#446841), Mm-Ptprc-C3 (#318651-C3), Mm-Col13a1-C2 (#837001-C2), Mm-Sftpc-C1 (#314101-C1) and Mm-Lyz1-C2 (#415131-C2). For human samples the following probes were used: Hs-INHBA-O1 (#569271), Hs-COL13A1-C2 (#857781-C2), Hs-INMT (#459961), Hs-MARCO (#512231), Hs-PECAM1-No-XMm-O1-C2 (#455931-C2), Hs-PTPRC-C2 (#601991-C2), and Hs-SAA3P (#857771). TSA Plus fluorophores fluorescein Cyanine 3 (1:1500 dilution) and Cyanine 5 (1:3000 dilution) (Perkin Elmer, Waltham, MA, USA) were used for signal detection. Sections were counterstained with DAPI and mounted with ProLong Gold Antifade Mountant (Invitrogen, Carlsbad, CA, USA). Tissue sections were scanned using 3DHISTECH Panoramic 250 FLASH II digital slide scanner at Genome Biology Unit (Research Programs Unit, Faculty of Medicine, University of Helsinki, Biocenter Finland) using 1 × 40 magnification with extended focus and 7 focus levels. Images were generated using 3DHISTECH Panoramic 250 FLASH II digital slide scanner at Genome Biology Unit supported by HILIFE and the Faculty of Medicine, University of Helsinki, and Biocenter Finland.

**Lung isolation and tissue dissociation.** Developing pups designated for single-cell RNA sequencing (scRNA-seq) and Fluorescence-activated cell sorting (FACS)

analyses were euthanized at P3, P7, or P14 by 10 μl/g i.p. injection of pentobarbital sodium (CDMV, Saint-Hyacinthe, QC, Canada) and received an additional i.p. injection of 10 mU/g (in 10 μl/g volume) heparin sodium (LEO Pharma Inc., Thornhill, ON, Canada). Following euthanasia, the chest was opened and the abdominal aorta and vena cava were cut above the liver. The left atrium was perforated and lungs were perfused through the right ventricle with 5 ml of 25 U/ml Heparin Sodium until completely white. Lungs were removed, dissected into individual lobes, and shortly rinsed with Dulbecco's PBS (DPBS, Lonza, Basel, Switzerland). Dissected lungs were then digested in 5 ml of enzyme mixture at 37 °C by gentleMACS<sup>™</sup> Octo Dissociator (Miltenyi Biotech, Bergisch Gladbach, Germany). Following costumed dissociation program was used for the digestion: loop 6 × (spin 300 rpm, 10''; spin -300rpm, 10''); loop 2 × (spin 150 rpm, 5''; spin -150rpm, 5''); loop 2 × (spin 20 rpm, 5' 0''; spin -20rpm, 5' 0''); loop 6 × (ramp 360 rpm, 15''; ramp -360rpm, 15'').

The following customized enzymatic mixture was used for lung digestion: (i) 2500U Collagenase I (Worthington Biochem, Lakewood, NJ, USA), 30U Neutral Protease (Worthington Biochem, Lakewood, NJ, USA), 500U Deoxyribonuclease (DNase) I (Sigma-Aldrich, Oakville, ON, Canada); (ii) Elastase (Worthington Biochem, Lakewood, NJ, USA), 500U DNase I; (iii) 2500U Collagenase I, 500U DNase I; (iv) Collagenase/Dispase (Sigma-Aldrich, Oakville, ON, Canada), 500U DNase I. All enzyme mixtures were diluted in 5 ml of DPBS supplemented with Mg<sup>2+</sup>/Ca<sup>2+</sup> (ThermoFisher Scientific, Burlington, ON, Canada). Combination of Collagenase I, Neutral Protease, and DNase (i) resulted in the most equal distribution and was used in all following experiments.

The resulting suspension was filtered through a 100 μm nylon mesh (ThermoFisher Scientific, Burlington, ON, Canada) and the enzymatic reaction was terminated by 0.9 mM EDTA. The cell suspension was then centrifuged and the resulting pellet was resuspended in 5 ml DPBS (Lonza, Basel, Switzerland), thoroughly filtered through 40 μm filter (Corning Life Sciences, Tewksbury, MA, USA) and centrifuged again. According to its size, the resulting pellet was resuspended in 500–1000 μl of cold RBC lysis buffer (ThermoFischer Scientific, Burlington, ON, Canada) for 3–5 min until white appearance of the suspension was achieved. The cell suspension was then diluted by DPBS (Lonza, Basel, Switzerland) to a total volume of 5 ml, centrifuged and washed twice. Cells were counted using both, the Scepter<sup>™</sup> automated cell counter (Millipore-Sigma, Burlington, MA, USA) and a manual hemacytometer (Bright-Line<sup>™</sup> Hemacytometer; Sigma-Aldrich, Oakville, ON, Canada).

**Fluorescent activated cell sorting.** The number of cells in single-cell suspension was estimated using a Scepter<sup>™</sup> automated cell counter (Millipore-Sigma, Burlington, MA, USA) and a total of 0.5 × 10<sup>6</sup> cells/sample were resuspended in 200 μl of PBS in 96-well plate. Cells were incubated in the dark with 2 μl/1 × 10<sup>6</sup> cells of CD16/32 antibody (Fc block; BD Biosciences, Mississauga, ON, Canada) for 15 min at RT. Following blocking, cells were centrifuged and resulting pellets were resuspended in 1:100 mixture of panel of antibodies: FITC-conjugated CD31 (BD Biosciences, Mississauga, ON, Canada), AF647-conjugated CD45 (Southern Biotech, Birmingham, AL, USA) and Pe/Cy7-conjugated CD326 (EpcAM; ThermoFischer Scientific, Burlington, ON, Canada). Cells were incubated with antibodies at RT for 30 min in dark. Following staining, cells were pelleted by centrifugation and washed 3 × with FACS buffer (5% (v/v) FBS and 1 mM EDTA in 1 × DPBS). All samples were fixed by 4% (w/v) PFA prior to analysis. Flow cytometry was performed using a BD LSR Fortessa (Beckton Dickinson Biosciences, Franklin Lakes, NJ, USA) at the Ottawa Hospital Research Institute (OHRI) core facility. Sample compensation was performed using BD FACSDIVA software and data analysis were performed with FlowJo v10 software (FlowJo LLC, Ashland, OR, USA) (Supplementary Fig. 10).

**Multiplexing individual samples for scRNA-seq.** Multiplexing was performed according to the MULTI-seq protocol<sup>8</sup>. Following the preparation of single-cell suspension, cells were counted and a total of 0.5 × 10<sup>6</sup> cells/sample were resuspended and pelleted in a 96-well plate at 400 × g for 5 min. The resulting pellet was resuspended in 150 μl of 200 nM anchor/200 nM barcode solution (kindly provided by Prof. Zev Gartner from University of California, San Francisco). The lipid-modified DNA oligonucleotide (LMO) anchor and a unique "sample barcode" oligonucleotides were added to each sample in order to be multiplexed, with each sample receiving a different sample barcode (Supplementary Fig. 1e). Samples were then incubated for 10 min at room temperature (RT). After 10 min, samples were supplemented with 200 nM common lipid-modified co-anchor to stabilize the membrane residence of barcodes. Samples were incubated on ice for additional 5 min and pelleted at 400 × g for 5 min. Barcode-containing media was then removed, and the resulting cell pellet was washed twice with 1% FBS (Sigma-Aldrich, Oakville, ON, Canada) in 1 × DPBS (Lonza, Basel, Switzerland). After the final wash, cells were resuspended in 1 × DPBS + 1% FBS, counted and samples were pooled together at 1:1 ratio while maintaining the final concentration of 500–1000 cells/μl. Viability and cell counts were assessed using a manual hemacytometer (Bright-Line<sup>™</sup> Hemacytometer; Sigma-Aldrich, Oakville, ON, Canada), and only samples with viability ≥ 80% were further processed by 10 × Chromium.

**scRNA-seq library preparation and sequencing.** Single-cell suspensions were processed using the 10x Genomics Single Cell 3' v3 RNA-seq kit. Gene expression libraries were prepared according to the manufacturer's protocol. MULTI-seq barcode libraries were retrieved from the samples and libraries were prepared independently, as described previously<sup>8</sup>. Final libraries were sequenced on the NextSeq500 platform (Illumina) to reach an approximate depth of 20,000–25,000 reads/cell.

**Processing of raw sequencing reads.** Raw sequencing reads from the gene expression libraries were processed using CellRanger v3.0.2, aligning reads to the mm10 build of the mouse genome. Except for explicitly setting `expect-cells=25000`, default parameters were used for all samples. MULTI-seq barcode libraries were simply trimmed to 28 bp using Trimmomatic (v0.36) prior to demultiplexing.

**Demultiplexing expression data with MULTI-seq barcode libraries.** Demultiplexing was performed using the deMULTiplex R package (v1.0.2) (<https://github.com/chris-mcginnis-ucsf/MULTI-seq>). The key concepts for demultiplexing are described in McGinnis et al.<sup>8</sup>. Briefly, the tool considers the barcode sequencing reads and counts the frequency with which each of the sample barcodes appears in each cell. Then, for each barcode, the distribution of counts in cells is assessed and an optimal quantile threshold to deem a cell positive for a given barcode is determined. Cells positive for more than one barcode are classified as doublets and removed. Only cells positive for a single barcode are retained for further analysis (Supplementary Fig. 1b). As each barcode corresponds to a specific sample in the experiment, the sample annotations can then be added to all cells in the data set.

**Data quality control, integration, and clustering.** All main processing steps were performed with Seurat v3.1.5<sup>86</sup>. Quality control was first performed independently on each library to find appropriate filtering thresholds for each. Expression matrices for each sample were loaded into R as Seurat objects, retaining only cells in which more than 200 genes were detected. Poor quality cells with a high percentage (>20%) of UMIs mapped to mitochondrial genes were removed.

We then split each unique sample (based on MULTI-seq sample barcodes) into a separate Seurat object. SCTransform<sup>87</sup> was used to normalize each sample and select highly variable genes. This was also used to regress out effects of cell cycle and cell stress (percentage of mitochondrial reads) to ensure that downstream clustering would effectively capture cell type patterns rather than cell cycle stage, for example. To further ensure that clustering would not be impacted by batch effects or biological variability associated with mouse age or oxygen conditions, we performed the data integration method implemented by Seurat v3 for SCTransform-normalized data, using the `SelectIntegrationFeatures()`, `PrepSCTIntegration()`, `FindIntegrationAnchors()`, and `IntegrateData()` functions with default options. PCA was then run on the top 3000 variable genes and the data was then clustered at a low resolution (`dims=1:30`, `Resolution=0.2`) with the Louvain algorithm implemented in the `FindClusters()` function in Seurat.

To identify major cell type subsets (epithelial, endothelial, immune, stromal, mesothelial), we assessed expression of the canonical markers *Epcam*, *Pecam1*, *Ptprc*, *Colla1*, and *Msln*, respectively. We also assessed additional genes discriminating these clusters by performing a simple Wilcoxon rank-sum test with the `FindAllMarkers()` function in Seurat. We next isolated each major subset, and reprocessed the subsets using the same normalization and integration approach. We first clustered the data at a higher resolution (subset-specific) and removed rare clusters comprising dozens to a couple hundred of cells expressing canonical markers of contaminant cell types in the subset (e.g., a *Ptprc*<sup>+</sup> cluster in the epithelial subset), which likely represent rare doublets that were not identified during demultiplexing. With contaminants removed, we re-clustered the integrated data to identify the individual cell types comprising the subset based on previously reported expression signatures.

Cell type labels for each subset were then added to the Seurat object containing all data. While the integration approach prevented clustering and cell-type annotation from being driven by biological and technical effects, these signals are important for data exploration. While retaining integrated cell type labels, we reprocessed data without integration. SCTransform was used to normalize the data, identify the top 3000 variable features, and regress out cell cycle and cell stress. PCA was then performed and UMAP embeddings were generated on the top 40 PCs. This ultimately resulted in UMAP embeddings that show variation associated with experimental conditions with labels resulting underlying cell types.

**Differential expression analysis and gene set enrichment analysis (GSEA).** To identify genes differentially expressed in response to hyperoxia or as a result of mouse age, we used the R package *muscat* (v1.1.6), which was designed specifically for multi-sample, multi-condition scRNA-seq experiments. We used a standard workflow with the tool, generating pseudobulk expression profiles for each sample in each cluster and then testing for differential expression between groups associated with the queried experimental conditions. Significant genes were those with an adjusted *p* value <0.05 and a detection rate of at least 10% in at least one of the conditions tested.

To identify gene sets associated with differentially expressed genes, we used R package *fgsea* (v1.12.0) on the fold-change-ranked list of genes for a given condition. We queried and aggregated a list of gene sets comprising all GO terms,

KEGG pathways, Reactome pathways, and the MSigDB Hallmark gene sets. These gene sets were acquired from the Molecular Signatures Database (v6)<sup>88,89</sup>. Significantly enriched gene sets were those with an adjusted *p* value of <0.05 and the normalized enrichment score (NES) was used to assess whether these gene sets were associated with upregulated or downregulated genes in a given condition.

**Cell communication inference.** As we were specifically interested in understanding cell communication networks that give rise to hyperoxia-specific effects, we used the R package NicheNet<sup>12</sup>. NicheNet uses information about expression of cognate ligands, receptors, signaling pathways, and genomic targets to infer cell communication patterns that contribute to a specific queried gene set. Differential gene expression in hyperoxia vs normoxia was used for the queried gene set of interest in the NicheNet analysis. To prioritize results, we only performed this analysis to identify signaling that could contribute to the effects in cell types with fairly large responses to hyperoxia (over 200 differentially expressed genes), but included all cell types as potential sources of ligand expression. Background expression of genes was specified with default approach used in NicheNet's pipeline, i.e. all genes with >10% detection in a given cluster. Cells from both normoxia and hyperoxia conditions were used in the analysis, but in our synthesis of most relevant changes, we prioritized hyperoxia-induced ligands or those from cell types that increase in proportion in hyperoxia samples. For each "receiver" cell type, we selected the top 10 ligands predicted to drive hyperoxia responses based on the Pearson correlation coefficient between the ligand-target regulatory potential score of each ligand and the target indicator vector, as implemented in NicheNet<sup>12</sup>. In each case, we also assessed whether the specific ligands and receptors were upregulated in hyperoxia samples, or whether the cell types expressing the ligands increase in proportion in hyperoxia samples. This information is all included in heatmaps provided in Supplementary Figures. The circo plot in Fig. 2c represents a summary of conditions where the ligands are upregulated in hyperoxia, or the cell type expressing the ligand increases in proportion.

**Statistical analysis.** Data are presented as means ± SD. All statistical analyses were performed with GraphPad Prism 8.0. The presence of potential statistical outliers was determined by Grubbs' test. Differences in case of two-member groups were evaluated either by unpaired Student's *t*-test, or multiple Student's *t*-test with correction for multiple comparisons using Holm-Sidak method. *P* values < 0.05 were considered as significant and depicted as following: *P* values < 0.05: \*; *P* values < 0.01: \*\*; *P* values < 0.001: \*\*\*; *P* values < 0.0001: \*\*\*\*.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All scRNA sequencing data, including raw fastq sequencing files, gene expression matrices, and associated cell metadata generated in this study have been deposited in the NCBI's Gene Expression Omnibus (GEO) database under accession code G4.

## Code availability

Code used for the analysis of scRNA-seq data is available at the public GitHub repository at [https://github.com/dpcook/thebaud\\_lung\\_BPD](https://github.com/dpcook/thebaud_lung_BPD).

Received: 14 December 2019; Accepted: 9 February 2021;

Published online: 10 March 2021

## References

- Guo, M. et al. Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. *Nat. Commun.* **10**, 37 (2019).
- Cohen, M. et al. Lung single-cell signaling interaction map reveals basophil role in macrophage imprinting. *Cell* **175**, 1031–1044 (2018). e18.
- Warburton, D. Overview of lung development in the newborn human. *Neonatology* **111**, 398–401 (2017).
- Leong, M. Genetic approaches to bronchopulmonary dysplasia. *NeoRev.* **20**, e272–e279 (2019).
- Lal, C. V. & Ambalavanan, N. Genetic predisposition to bronchopulmonary dysplasia. *Semin. Perinatol.* **39**, 584–591 (2015).
- Thebaud, B. et al. Bronchopulmonary dysplasia. *Nat. Rev. Dis. Prim.* **5**, 78 (2019).
- Surate Solaligue, D. E., Rodriguez-Castillo, J. A., Ahlbrecht, K. & Morty, R. E. Recent advances in our understanding of the mechanisms of late lung development and bronchopulmonary dysplasia. *Am. J. Physiol. -Lung Cell. Mol. Physiol.* **313**, L1101–L1153 (2017).
- McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).



9. Ardini-Poleske, M. E. et al. LungMAP: the molecular atlas of lung development program. *Am. J. Physiol. -Lung Cell. Mol. Physiol.* **313**, L733–L740 (2017).
10. Zhang, X. et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* **47**, D721–D728 (2019).
11. Thul, P. J. et al. A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
12. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
13. Revenis, M. E. & Kaliner, M. A. Lactoferrin and lysozyme deficiency in airway secretions: Association with the development of bronchopulmonary dysplasia. *J. Pediatr.* **121**, 262–270 (1992).
14. Marino, R. et al. Secretory leukocyte protease inhibitor plays an important role in the regulation of allergic asthma in mice. *J. Immunol.* **186**, 4433–4442 (2011).
15. Kevill, K. A. et al. A role for macrophage migration inhibitory factor in the neonatal respiratory distress syndrome. *J. Immunol.* **180**, 601–608 (2008).
16. Perveen, S. et al. MIF inhibition enhances pulmonary angiogenesis and lung development in congenital diaphragmatic hernia. *Pediatr. Res.* **85**, 711–718 (2019).
17. Inoue, H. et al. Serum neutrophil gelatinase-associated lipocalin as a predictor of the development of bronchopulmonary dysplasia in preterm infants. *Early Hum. Dev.* **89**, 425–429 (2013).
18. Gokey, J. J. et al. MEG3 is increased in idiopathic pulmonary fibrosis and regulates epithelial cell differentiation. *JCI Insight* **3**, e122490 (2018).
19. Beers, M. F. & Mulugeta, S. The biology of the ABCA3 lipid transporter in lung health and disease. *Cell Tissue Res.* **367**, 481–493 (2017).
20. Cheong, N. et al. Functional and trafficking defects in ATP binding cassette A3 mutants associated with respiratory distress syndrome. *J. Biol. Chem.* **281**, 9791–9800 (2006).
21. Zepp, J. A. et al. Distinct mesenchymal lineages and niches promote epithelial self-renewal and myofibrogenesis in the lung. *Cell* **170**, 1134–1148 (2017). e10.
22. Jónsson, B., Tullus, K., Brauner, A., Lu, Y. & Noack, G. Early increase of TNF $\alpha$  and IL-6 in tracheobronchial aspirate fluid indicator of subsequent chronic lung disease in preterm infants. *Arch. Dis. Child. - Fetal Neonatal Ed.* **77**, F198–F201 (1997).
23. Zhong, Y. et al. Maternal omega-3 PUFA supplementation prevents hyperoxia-induced pulmonary hypertension in the offspring. *Am. J. Physiol. -Lung Cell. Mol. Physiol.* **315**, L116–L132 (2018).
24. Joza, S., Wang, J., Tseu, I., Ackerley, C. & Post, M. Fetal, but not postnatal, deletion of semaphorin-neuropilin-1 signaling affects murine alveolar development. *Am. J. Respir. Cell Mol. Biol.* **49**, 627–636 (2013).
25. White, S. R., Wojcik, K. R., Gruenert, D., Sun, S. & Dorscheid, D. R. Airway epithelial cell wound repair mediated by  $\alpha$ -dystroglycan. *Am. J. Respir. Cell Mol. Biol.* **24**, 179–186 (2001).
26. Chen, J. & Krasnow, M. A. Integrin beta 1 suppresses multilayering of a simple epithelium. *PLoS ONE* **7**, e52886 (2012).
27. Lambrechts, D. et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289 (2018).
28. Barron, L., Gharib, S. A. & Duffield, J. S. Lung pericytes and resident fibroblasts. *Am. J. Pathol.* **186**, 2519–2531 (2016).
29. Xie, T. et al. Single-cell deconvolution of fibroblast heterogeneity in mouse pulmonary fibrosis. *Cell Rep.* **22**, 3625–3640 (2018).
30. Wilson, T. C., Bachurski, C. J., Ikegami, M., Jobe, A. H. & Kallapur, S. G. Pulmonary and systemic induction of SAA3 after ventilation and endotoxin in preterm lambs. *Pediatr. Res.* **58**, 1204–1209 (2005).
31. Djurec, M. et al. Saa3 is a key mediator of the protumorigenic properties of cancer-associated fibroblasts in pancreatic tumors. *Proc. Natl Acad. Sci.* **115**, E1147–E1156 (2018).
32. Ruiz-Camp, J. et al. Targeting miR-34a/ Pdgfra interactions partially corrects alveologenesis in experimental bronchopulmonary dysplasia. *EMBO Mol. Med.* **11**, e9448 (2019).
33. Li, C. et al. Secondary crest myofibroblast PDGFR $\alpha$  controls the elastogenesis pathway via a secondary tier of signaling networks during alveologenesis. *Development* **146**, dev176354 (2019).
34. Fulton, C. T., Cui, T. X., Goldsmith, A. M., Bermick, J. & Popova, A. P. Gene expression signatures point to a male sex-specific lung mesenchymal cell PDGF receptor signaling defect in infants developing bronchopulmonary dysplasia. *Sci. Rep.* **8**, 17070 (2018).
35. Angelidis, I. et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).
36. McQualter, J. L. Endogenous lung stem cells for lung regeneration. *Expert Opin. Biol. Ther.* **19**, 539–546 (2019).
37. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
38. Jia, Y. et al. Regulation of lung fibroblast activation by annexin A1. *J. Cell. Physiol.* **228**, 476–484 (2013).
39. Li, M. et al. Inhibition of calcineurin/NFATc4 signaling attenuates ventilator-induced lung injury. *Mol. Med. Rep.* **21**, 607–614 (2019).
40. Windt, G. J. W., van der Schouten, M., Zeerleder, S., Florquin, S. & van der Poll, T. CD44 is protective during hyperoxia-induced lung injury. *Am. J. Respir. Cell Mol. Biol.* **44**, 377–383 (2011).
41. Mižíková, I. & Morty, R. E. The extracellular matrix in bronchopulmonary dysplasia: target and source. *Front. Med.* **2**, 91 (2015).
42. Gremlich, S. et al. Tenascin-C inactivation impacts lung structure and function beyond lung development. *Sci. Rep.* **10**, 5118 (2020).
43. Olave, N. et al. Regulation of alveolar septation by microRNA-489. *Am. J. Physiol. -Lung Cell. Mol. Physiol.* **310**, L476–L487 (2016).
44. Fleming, R. E., Crouch, E. C., Ruzicka, C. A. & Sly, W. S. Pulmonary carbonic anhydrase IV: developmental regulation and cell-specific expression in the capillary endothelium. *Am. J. Physiol. -Lung Cell. Mol. Physiol.* **265**, L627–L635 (1993).
45. Wang, M. M. et al. Expression of periaxin (PRX) specifically in the human cerebrovascular system: PDZ domain-mediated strengthening of endothelial barrier function. *Sci. Rep.* **8**, 10042 (2018).
46. Gillich, A. et al. Capillary cell-type specialization in the alveolus. *Nature* **586**, 785–789 (2020).
47. Vila Ellis, L. et al. Epithelial vegfa specifies a distinct endothelial population in the mouse lung. *Dev. Cell* **52**, 617–630 (2020). e6.
48. Niethamer, T. K. et al. Defining the role of pulmonary endothelial cell heterogeneity in the response to acute lung injury. *eLife* **9**, e53072 (2020).
49. Lim, R. et al. Activin A contributes to the development of hyperoxia-induced lung injury in neonatal mice. *Pediatr. Res.* **77**, 749–756 (2015).
50. Harper, W. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**, 805–816 (1993).
51. Bartram, U. & Speer, C. P. The role of transforming growth factor  $\beta$  in lung development and disease. *Chest* **125**, 754–765 (2004).
52. Wujak, Ł. A. et al. FXYD1 negatively regulates Na<sup>+</sup>/K<sup>+</sup>-ATPase activity in lung alveolar epithelial cells. *Respir. Physiol. Neurobiol.* **220**, 54–61 (2016).
53. Brazeo, P. L. et al. FXYD5 is an essential mediator of the inflammatory response during lung injury. *Front. Immunol.* **8**, 623 (2017).
54. Pi, L. et al. Vascular endothelial cell-specific connective tissue growth factor (CTGF) is necessary for development of chronic hypoxia-induced pulmonary hypertension. *Front. Physiol.* **9**, 138 (2018).
55. Bottoms, S. E., Howell, J. E., Reinhardt, A. K., Evans, I. C. & McAnulty, R. J. TGF- $\beta$  isoform specific regulation of airway inflammation and remodelling in a murine model of asthma. *PLoS ONE* **5**, e9674 (2010).
56. Thébaud, B. & Abman, S. H. Bronchopulmonary dysplasia: where have all the vessels gone? Roles of angiogenic growth factors in chronic lung disease. *Am. J. Respir. Crit. Care Med.* **175**, 978–985 (2007).
57. O'Reilly, M. A. et al. The cyclin-dependent kinase inhibitor p21 protects the lung from oxidative stress. *Am. J. Respir. Cell Mol. Biol.* **24**, 703–710 (2001).
58. McGrath-Morrow, S. A., Cho, C., Soutiere, S., Mitzner, W. & Tuder, R. The effect of neonatal hyperoxia on the lung of p21 Waf1/Cip1/Sdi1-deficient mice. *Am. J. Respir. Cell Mol. Biol.* **30**, 635–640 (2004).
59. Visser, Y. P., de, Walther, F. J., Laghmani, E. H., van der Laarse, A. & Wagenaar, G. T. M. Apelin attenuates hyperoxic lung and heart injury in neonatal rats. *Am. J. Respir. Crit. Care Med.* **182**, 1239–1250 (2010).
60. Hiepen, C. et al. BMPR2 acts as a gatekeeper to protect endothelial cells from increased TGF $\beta$  responses and altered cell mechanics. *PLoS Biol.* **17**, e3000557 (2019).
61. Alphonse, R. S. et al. Existence, functional impairment, and lung repair potential of endothelial colony-forming cells in oxygen-induced arrested alveolar growth. *Circulation* **129**, 2144–2157 (2014).
62. Ren, X. et al. Postnatal alveologenesis depends on FOXF1 signaling in c-KIT<sup>+</sup> endothelial progenitor cells. *Am. J. Respir. Crit. Care Med.* **200**, 1164–1176 (2019).
63. Wakabayashi, T. et al. CD157 marks tissue-resident endothelial stem cells with homeostatic and regenerative properties. *Cell Stem Cell* **22**, 384–397 (2018). e6.
64. Yu, Q. C., Song, W., Wang, D. & Zeng, Y. A. Identification of blood vascular endothelial stem cells by the expression of protein C receptor. *Cell Res.* **26**, 1079–1098 (2016).
65. Mould, K. J., Jackson, N. D., Henson, P. M., Seibold, M. & Janssen, W. J. Single cell RNA sequencing identifies unique inflammatory airspace macrophage subsets. *JCI Insight* **4**, e126556 (2019).
66. Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334 (2019). e10.
67. Tan, S. Y. S. & Krasnow, M. A. Developmental origin of lung macrophage diversity. *Development* **143**, 1318–1327 (2016).
68. Gibbings, S. L. et al. Three unique interstitial macrophages in the murine lung at steady state. *Am. J. Respir. Cell Mol. Biol.* **57**, 66–76 (2017).

69. Kalymbetova, T. V. et al. Resident alveolar macrophages are master regulators of arrested alveolarization in experimental bronchopulmonary dysplasia: Macrophages mediate aberrant lung alveolarization. *J. Pathol.* **245**, 153–159 (2018).
70. Willis, G. R. et al. Mesenchymal stromal cell exosomes ameliorate experimental bronchopulmonary dysplasia and restore lung function through macrophage immunomodulation. *Am. J. Respir. Crit. Care Med.* **197**, 104–116 (2018).
71. Orecchioni, M., Ghosheh, Y., Pramod, A. B. & Ley, K. Macrophage polarization: different gene signatures in M1(LPS+) vs. classically and M2 (LPS-) vs. alternatively activated macrophages. *Front. Immunol.* **10**, 1084 (2019).
72. Yang, M. et al. Angiogenesis-related genes may be a more important factor than matrix metalloproteinases in bronchopulmonary dysplasia development. *Oncotarget* **8**, 18670–18679 (2017).
73. Reyfman, P. A. et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **199**, 1517–1536 (2019).
74. Raredon, M. S. B. et al. Single-cell connectomic analysis of adult mammalian lungs. *Sci. Adv.* **5**, eaaw3851 (2019).
75. Thébaud, B. et al. Vascular endothelial growth factor gene therapy increases survival, promotes lung angiogenesis, and prevents alveolar damage in hyperoxia-induced lung injury: evidence that angiogenesis participates in alveolarization. *Circulation* **112**, 2477–2486 (2005).
76. Greco, F. et al. Hyperoxia-induced lung structure–function relation, vessel rarefaction, and cardiac hypertrophy in an infant rat model. *J. Transl. Med.* **17**, 91 (2019).
77. Wong, P. M. et al. Emphysema in young adult survivors of moderate-to-severe bronchopulmonary dysplasia. *Eur. Respir. J.* **32**, 321–328 (2008).
78. Liao, J. et al. The NLRP3 inflammasome is critically involved in the development of bronchopulmonary dysplasia. *Nat. Commun.* **6**, 8977 (2015).
79. Pryhuber, G. S. Postnatal infections and immunology affecting chronic lung disease of prematurity. *Clin. Perinatol.* **42**, 697–718 (2015).
80. D'Angio, C. T., Johnston, C. J., Wright, T. W., Reed, C. K. & Finkelstein, J. N. Chemokine mRNA alterations in newborn and adult mouse lung during acute hyperoxia. *Exp. Lung Res.* **24**, 685–702 (1998).
81. Stouch, A. N. et al. IL-1 $\beta$  and inflammasome activity link inflammation to abnormal fetal airway development. *J. Immunol.* **196**, 3411–3420 (2016).
82. Harijith, A. et al. A role for matrix metalloproteinase 9 in IFN $\gamma$ -mediated injury in developing lungs: relevance to bronchopulmonary dysplasia. *Am. J. Respir. Cell Mol. Biol.* **44**, 621–630 (2011).
83. Sureshbabu, A. et al. Conditional overexpression of TGF $\beta$ 1 promotes pulmonary inflammation, apoptosis and mortality via TGF $\beta$ 2 in the developing mouse lung. *Respir. Res.* **16**, 4 (2015).
84. Nardiello, C. et al. Standardisation of oxygen exposure in the development of mouse models for bronchopulmonary dysplasia. *Dis. Model. Mech.* **10**, 185–196 (2017).
85. Dzhuraev, G. et al. Estimation of absolute number of alveolar epithelial type 2 cells in mouse lungs: a comparison between stereology and flow cytometry. *J. Microsc.* **275**, 36–50 (2019).
86. Butler, A., Hoffman, P., Smibert, P., Papalexis, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
87. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
88. Liberzon, A. et al. The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
89. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci.* **102**, 15545–15550 (2005).

## Acknowledgements

The authors acknowledge the assistance of Martin Kang, PhD. (Ottawa Hospital Research Institute) with interpretation of scRNA-seq analyses of epithelial cells, Shumei Zhong (Ottawa Hospital Research Institute) and Sharlene Faulkes (University of Ottawa Louis Pelletier Histology Core Facility) for expert technical assistance. We acknowledge Prof. Zev Gartner (University of California, San Francisco) for kindly providing barcodes for multiplex labeling and Prof. Dr. Gloria Pryhuber from LungMAP Human Tissue Core, Biorepository for Investigation of Neonatal Diseases of Lung-Normal (BRINDL-NL) for kindly providing the human tissue material. Authors also acknowledge Michaela Kmaková for assistance with graphical design. This study was supported by the Canadian Institutes of Health Research (CIHR), the Finnish Sigrid Juselius Foundation, the Finnish Foundation for Pediatric Research, the German Research Foundation (Deutsche Forschungsgemeinschaft), the Ontario Institute for Regenerative Medicine (OIRM), the Stem Cell Network, Heart and Stroke Foundation Canada, the Ontario Graduate Scholarship, the Lung Association Breathing as One, the Molly Towel Perinatal Research Foundation, the European Respiratory Society, Société Française de Néonatalogie and the Canada Foundation for Innovation John R. Evans Leaders Fund.

## Author contributions

M.H. and I.M. conceived and directed the study, performed in vivo experiments, cell isolations and scRNA-seq analyses and wrote the manuscript. D.C. performed scRNA-seq analyses and edited the manuscript. C.D.C. performed FACS and stereological analyses. F.L. performed stereological analyses. N.A. performed fluorescent RNA in situ hybridization. E.H. analyzed scRNA-seq data. L.R. analyzed data from in vivo studies. R. P.J., M.H., B.C.V., and B.T. supervised experiments and provided essential equipment and infrastructure. B.T. edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21865-2>.

**Correspondence** and requests for materials should be addressed to B.Téb.

**Peer review information** *Nature Communications* thanks Anne Hilgendorff, Kerstin Meyer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

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



**Open Access** This article 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. The images or other third party material in this article 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 <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021