nature communications

Article

https://doi.org/10.1038/s41467-025-60371-7

Bronchopulmonary dysplasia with pulmonary hypertension associates with semaphorin signaling loss and functionally decreased FOXF1 expression

Received: 2 September 2024

Accepted: 20 May 2025

Shawyon P. Shirazi[1,2], **Nicholas M. Negretti**[1,2], **Christopher S. Jetter**[1], **Alexandria L. Sharkey**[1], **Shriya Garg**[1], **Meghan E. Kapp**[3], **Devan Wilkins**[1], **Gabrielle Fortier**[2], **Saahithi Mallapragada**[4], **Nicholas E. Banovich**[4,5], **Laurie C. Eldredge**[6], **Gail H. Deutsch**[6], **Christopher V. E. Wright**[2,5,7], **David B. Frank**[8], **Jonathan A. Kropski**[2,5,7,9,10,11] & **Jennifer M. S. Sucre**[1,2,5,7,11]

Lung injury in preterm infants leads to structural and functional respiratory deficits, with a risk for bronchopulmonary dysplasia (BPD) that in its most severe form is accompanied by pulmonary hypertension (PH). To identify potential cellular and molecular drivers of BPD in humans, we performed single-cell RNA sequencing of preterm infant lungs with evolving BPD and BPD + PH compared to term infants. Examination of endothelial cells reveals a unique, aberrant capillary cell-state in BPD + PH defined by *ANKRD1* expression. Within the alveolar parenchyma in infants with BPD/BPD + PH, predictive signaling analysis identifies surprising deficits in the semaphorin guidance-cue pathway, with decreased expression of pro-angiogenic transcription factor *FOXF1*. Loss of semaphorin signaling is replicated in a murine BPD model and in humans with causal *FOXF1* mutations for alveolar capillary dysplasia (ACDMPV), suggesting a mechanistic link between developmental programs underlying BPD and ACDMPV and uncovering a critical role for semaphorin signaling in normal lung development.

Proper consideration of the later stages of lung organogenesis requires an appreciation for the vulnerability of the developing lung to permanent respiratory deficits after neonatal injury. Indeed, in infants born preterm, abnormal lung development leading to chronic lung disease is a leading complication, resulting in bronchopulmonary dysplasia (BPD)[1]. In its most severe form, BPD is associated with life-threatening pulmonary hypertension (BPD + PH), which is thought to be caused by abnormal development of the pulmonary vasculature,

---

[1]Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN, USA. [2]Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA. [3]Department of Pathology, Case Western Reserve University Hospitals, Cleveland, OH, USA. [4]Division of Bioinnovation and Genome Science, Translational Genomics Research Institute, Phoenix, AZ, USA. [5]Biodevelopmental Origins of Lung Disease (BOLD) Center, Vanderbilt University Medical Center, Nashville, TN, USA. [6]Center for Respiratory Biology and Therapeutics, Seattle Children's Research Institute, Seattle Children's Hospital, Seattle, WA, USA. [7]Program in Developmental Biology, Vanderbilt University, Nashville, TN, USA. [8]Department of Pediatrics, Division of Cardiology, Children's Hospital of Philadelphia, Penn Cardiovascular Institute, Penn-CHOP Lung Biology Institute, Philadelphia, PA, USA. [9]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [10]Department of Veterans Affairs Medical Center, Nashville, TN, USA. [11]These authors contributed equally: Jonathan A. Kropski, Jennifer M. S. Sucre. e-mail: jon.kropski@vumc.org; jennifer.sucre@vumc.org

with decreased numbers of alveolar-capillary units formed for gas exchange[2,3]. Despite increasing awareness of the environmental and perinatal risk factors that increase the risk of BPD, there are very few disease-modifying therapies, and our understanding of the molecular mechanisms and potential therapeutic targets remains limited.

The advent of single-cell multiomics technology as applied to human lung samples has enabled an increasingly nuanced understanding of the cell populations and cell-states under homeostatic and diseased conditions in adults[4–6]. However, because neonatal biopsies are infrequently performed and infant autopsies are rare in general, lung tissue from preterm infants and healthy neonates is largely underrepresented in most of the currently available large sequencing series. Here we report single-cell transcriptomic findings from former preterm infants with evolving BPD, and infants with both BPD and PH, compared to term infants at approximately the same corrected gestational age as the preterm infants.

We anticipate that this transcriptional dataset will be a resource for future work in neonatal lung injury and BPD in general (and have made these data publicly available as a searchable data explorer), and we were particularly interested in the endothelial cell populations in infants with BPD + PH and how they might differ from corrected gestational age-matched controls and patients with BPD at a similar age. In these sequencing data, we identified a capillary endothelial cell-state highly specific to patients with BPD + PH and spatially localized these capillary cells to the distal alveolar parenchyma exclusively; these cells were absent from proximal lung tissue. Further, we employed predictive signaling analysis to generate hypotheses about the role of semaphorin guidance-cue signaling during alveolar formation, a pathway we identified as transcriptomically deficient in both BPD and BPD + PH patients, and then spatially localized to the developing alveolar regions. Our parallel identification of semaphorin signaling deficits in alveolar epithelial and endothelial cells in an established mouse model of BPD suggests that this pathway is central to normal alveologenesis. Moreover, we find evidence of dysregulated semaphorin signaling in patients with alveolar capillary dysplasia with misalignment of pulmonary veins (ACDMPV), implicating semaphorin signaling as a potential mechanism driving abnormal lung development.

Results

To examine the cellular dynamics and candidate mediators of neonatal lung injury and the subsequent development of severe BPD, we used the 10x Genomics Chromium Platform to perform single-cell RNA sequencing (scRNAseq) of human neonatal lungs at varying stages of lung injury and chronic lung disease: one acute preterm infant with lung injury, two BPD samples at term corrected gestational age (CGA), two BPD + PH at term CGA, and two term infants in the first three weeks of life for age-matched comparison (Fig. 1a, and Supplementary Table 1). For histologic analysis, an expanded set of FFPE samples (independent of those used for sequencing) was used that included term infants, infants with BPD, infants with BPD + PH and infants born 24–26 weeks gestation who survived less than 6 h after birth representing preterm infant uninjured lungs (Fig. 1b, and Supplementary Table 2). After filtering and quality control to remove doublets and ambient RNA (see Methods), we recovered 43,607 single-cell transcriptomes with a range of 2199 to 8339 cells per subject, a median of 4128 unique molecular identifiers (UMI) (interquartile range 1188 – 4670) per cell, and a median of 1,795 genes (interquartile range 728 – 1,944) per cell (Fig. 1c–g). Unbiased clustering with cell-type annotation informed by prior early human, neonatal, and adult studies identified seven epithelial, seven endothelial, seven mesenchymal, and twelve immune-cell clusters (Figs. 1c–g, 2a)[4–12]. These sequencing data are available through an interactive data explorer at www.sucrelab.org/lungcells and the LungMAP website, www.lungmap.net.

Transcriptomic landscapes of term infants, infants with BPD and infants with BPD and PH revealed differences within cell lineages and cell types

Comparing the epithelial, mesenchymal, endothelial, and immune cell lineages across conditions, we identified differentially expressed genes between each lineage (Fig. 2b) and within each parenchymal cell type (Fig. 2c, and S1). While BPD and BPD + PH patients are often grouped together as a single clinical entity, global examination of transcriptional trends suggests significant differences between these two groups of patients (Fig. 2b, c). These data identify distinguishing differences in BPD + PH at the cellular and molecular level, with the greatest number of expression changes identified within alveolar type 1 (AT1) cells. Comparing transcription profiles from AT1s in BPD + PH with BPD and term infants, we found decreased expression of components of the basement membrane (e.g., *LAMA3*, *LAMC2*) and tight junctions (e.g., *CLDN4*) when compared to BPD and to term infants. Differential gene expression between term infant, BPD, and BPD + PH was demonstrated further in all parenchymal cell lineages in Fig. S1 and Supplementary Data 1–3.

Bronchopulmonary dysplasia with pulmonary hypertension is associated with a specific population of capillary endothelial cells

To focus on potential vascular cell differences in BPD patients with and without PH, the endothelial cluster was extracted, re-embedded, and separately examined to identify cell-type-specific differences across homeostasis and disease. Analysis of the endothelial cells identified a population primarily obtained from BPD + PH patients with transcriptional hallmarks common to both alveolar capillaries (aCaps, CAP2) and general capillaries (gCaps, CAP1) (Fig. 3a, b, and S2a). This group of 647 cells is marked by high expression of *ANKRD1*, a gene that is nearly absent among all other endothelial cell clusters, with limited detection in AT1 cells (Fig. 3c). This aberrant capillary cell-state (abCap) expresses some aCap-associated genes, e.g. *APLN*, *EMCN*, *S100A4*, but has markedly reduced expression of aCap hallmark genes *EDNRB* and *SOSTDC1* when compared to other aCaps. Similarly, abCaps express some genes associated with gCaps, e.g., *PTPRB* and *EDN1*, but lack expression of canonical gCap genes *FCN3*, *IL7R*, and *BTNL9* (Fig. 3d). The abCaps also express certain pan-capillary markers *PDGFB*, *RGCC*, and *ADGRL2*, but not *CA4*, *CD36*, and *AFF3*. In addition to the YAP/TAZ signaling target *ANKRD1*, this population expresses relatively high levels of other YAP targets *CCN1* (*CYR61*), *CCN2* (*CTGF*), *SERPINE1*, and *CRIM1*, compared to other endothelial cells (Fig. 3d). Notably, the abCaps present a gene-expression profile similar to previously described endothelial tip cells[13,14], a population present in early lung development that directionally guides the growth and migration of capillary endothelium. Relative to other endothelial cells, abCaps have increased expression of tip cell hallmarks *ESM1*, *ANGPT2*, *APLN*, *MCAM*, *COTL1*, but lower expression of tip cell–associated receptors such as *KDR*, *FLT4*, *UNC5B*, and *NRP1*, relative to other endothelial cells (Fig. 3e).

To spatially validate and localize the abCap cell-state, we performed RNA in situ hybridization (RNA ISH) using formalin-fixed paraffin-embedded (FFPE) lung tissue blocks from preterm infants with BPD with and without PH, as well as from term infant control lungs (Fig. 3f). These FFPE tissues are from an expanded set of samples, independent of those interrogated by scRNAseq (Supplementary Table 2). In lung tissue from patients with BPD + PH, *ANKRD1*+/*PECAM1*+ endothelial cells were localized to the alveolar parenchyma, near *SFTPC*+ alveolar type 2 (AT2) cells, consistent with their expression of capillary hallmark genes. Quantification demonstrated a marked enrichment of this population in patients with BPD + PH when compared with age-matched BPD patients and normal term infant controls of approximately the same corrected gestational age as well as preterm controls between 23–26 weeks (Fig. 3g).

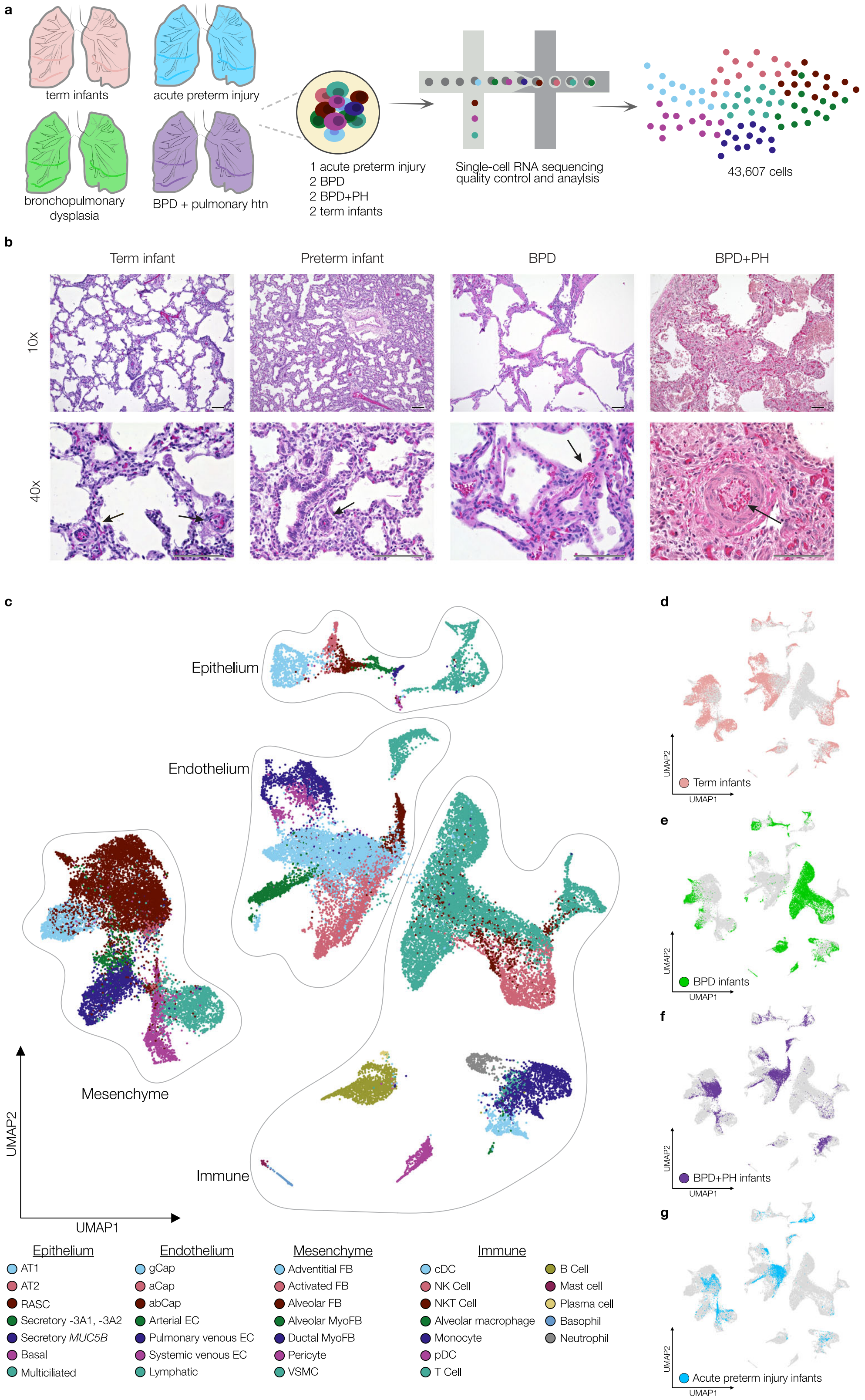


Fig. 1 | Analysis of neonatal human lung single-cell transcriptomics identifies distinct populations of epithelial, endothelial, mesenchymal, and immune cells. **a** Experimental design overview for single-cell sequencing of two term control infants, one acute neonatal lung injury, two infants with severe bronchopulmonary dysplasia (BPD), and two patients with severe bronchopulmonary dysplasia and pulmonary hypertension (BPD + PH). **b** Hematoxylin/eosin (H&E) stained lung samples from uninjured term ($n = 3$) and preterm control infants ($n = 5$), patients with BPD ($n = 2$), and with BPD + PH ($n = 4$) at 10x and 40x magnification highlighting differences in parenchymal vasculature (arrows). At least 5 images were taken at 10x and at least 10 images taken at 40x per sample. Scale bar = 100 μm. **c** Uniform Manifold Approximation and Projection (UMAP) embedding of the neonatal human lung dataset ($n = 43{,}607$ cells). Cells labeled by lineage and by cell type. **d**–**g** Breakdown of the dataset's UMAP embedding by disease-state condition annotation.

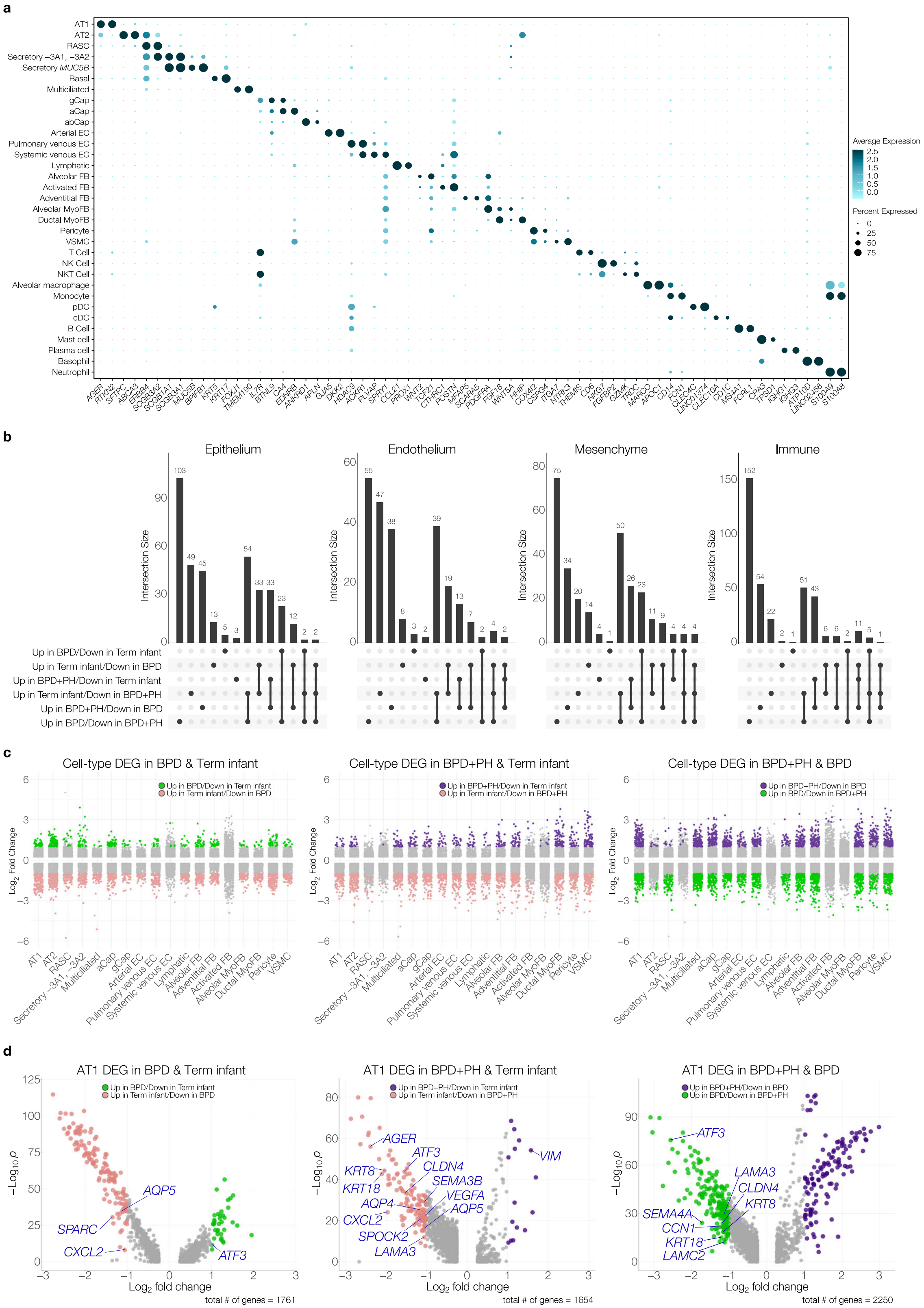
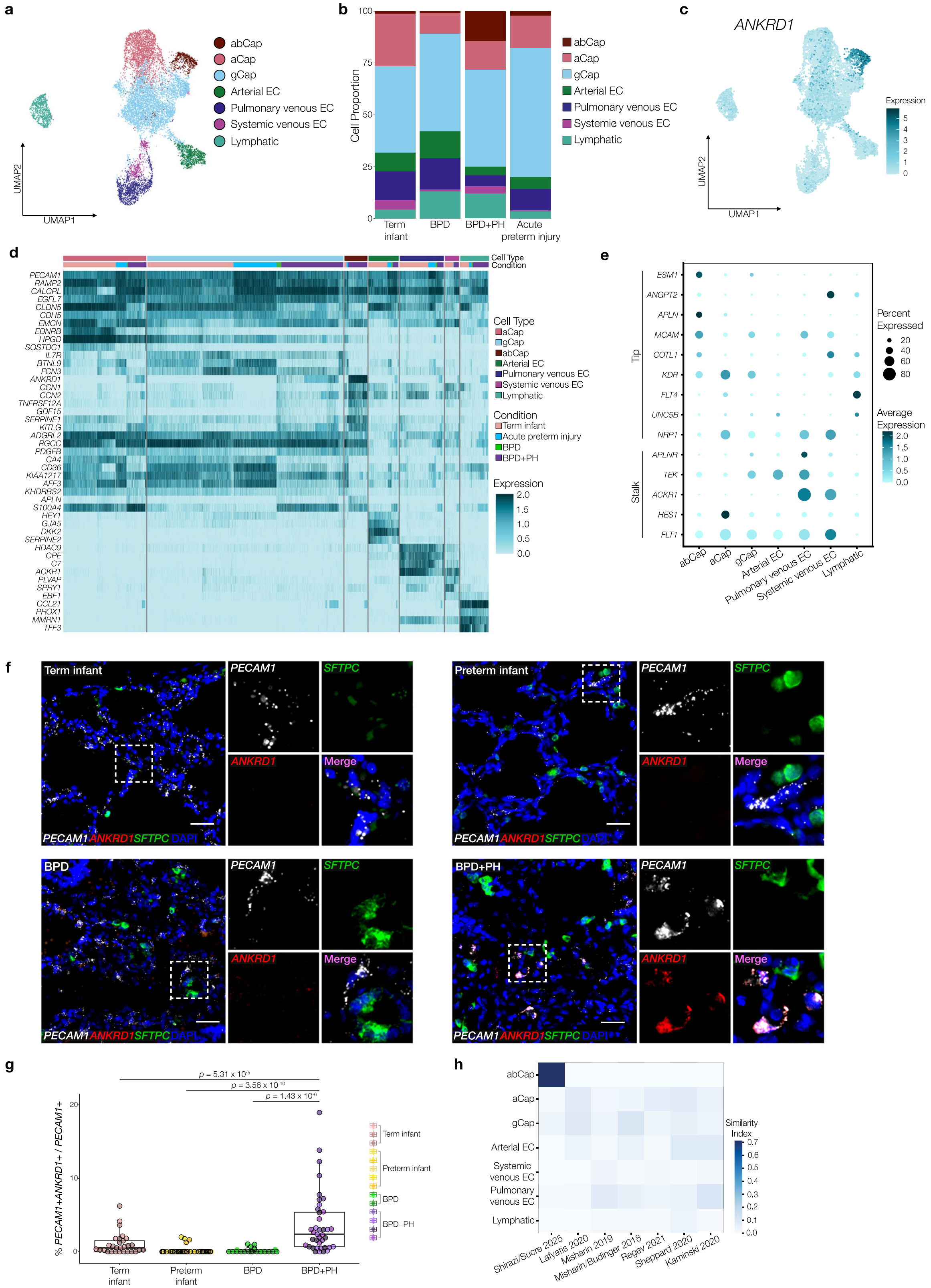


Fig. 2 | Transcriptomic analysis of differentially expressed genes in cell types reveals cellular and molecular differences between term infant, BPD, BPD + PH. **a** Dot plot of the hallmark-identifying genes for each cell-type cluster. Dot size indicates the percentage of cells within a cluster expressing a certain gene. Color intensity indicates the relative average gene expression level. **b** UpSet plots for each lung cell lineage displaying the intersecting sets of differentially expressed genes (DEGs) when comparing term infant to BPD, term infant to BPD + PH, and BPD to BPD + PH. **c** Strip plots of the DEGs from parenchymal cell types with comparisons across term infant, BPD, BPD + PH conditions. **d** Volcano plots of the DEGs from AT1 cells with comparisons across term infant, BPD, and BPD + PH conditions. DEGs calculated in (**b**, **c** and **d**) are determined by average $\log_2$ fold-change of >|±1| and *p*-value < 1×10^{-5} from the MAST test in Seurat's FindMarkers function.



The abCap endothelial cell state is transcriptomically distinct from endothelial cells previously described in various human lung diseases

To understand abCaps in the context of previously reported atlases of human pulmonary endothelium under normal and disease conditions, we performed analysis with SCimilarity[15], a computational tool that compares a given cell population with cell populations from nearly 400 previously published human datasets, including 41 from the lung. Comparing abCaps across the 1.12-million endothelial cells in the 22.7-million cell corpus of SCimilarity, we identified the 100 human samples (contained within 34 source studies) with the highest SCimilarity score to our BPD + PH capillary cells. Of these 34 source studies, six studies containing 178 patient samples were performed on human lung tissue with adequate sample size ($n > 2$) and accessible RNA data[4,12,16–19].

Fig. 3 | Analysis of the endothelium reveals a cluster of aberrant capillary cells (abCaps) enriched in patients with BPD + PH. a UMAP embedding of the neonatal human lung endothelial cells ($n = 11{,}816$) labeled by cell type. **b** Stacked area plot of cellular proportions across the four disease-state conditions. **c** UMAP of the expression level of *ANKRD1*, the highest expressing gene in the abCaps. **d** Heatmap of hallmark gene expression of the seven identified endothelial cell-type clusters, demonstrating a distinct profile for abCaps as compared with aCaps and gCaps. The cell types are subdivided by disease condition. Each column represents a single cell. Maximum expression values are clipped at 2. **e** Dot plot of the hallmark genes for endothelial tip and stalk cells. Dot size indicates the percentage of cells within a cell type cluster expressing a certain gene. Color intensity indicates relative average expression level. **f** RNA ISH for *ANKRD1* (red), *PECAM1* (white), and *SFTPC* (green) demonstrating co-localization of *ANKRD1* within *PECAM1*+ cells in the alveolar parenchyma of BPD + PH ($n = 4$) patients in contrast to term infant ($n = 3$), preterm infant ($n = 5$), and BPD ($n = 2$) patients. 8–12 images were taken per sample. Scale bar = 30 μm. **g** RNA ISH quantification using HALO of the percentage of *PECAM1*+ cells that are also *ANKRD1*+ over the total amount of *PECAM1*+ cells (*PECAM1*+ *ANKRD1*+ / all *PECAM1*+) across disease condition. Box plots represent the median (center line), 25th and 75th percentiles (box bounds), and whiskers extend to 1.5 times the interquartile range. *p*-values (indicated in the figure) calculated by post-hoc pairwise *t*-tests with Bonferroni correction after conducting Kruskal-Wallis, one-way ANOVA ($p\text{-value} = 1.75 \times 10^{-10}$). Each dot represents an image. Each color represents an individual. Source data are provided in a Source Data file. **h** Heatmap showing relative homology from top six datasets identified by SCimiliarity, with intensity of color indicating greater similarity, abCap cell included in upper left as positive control.

(Fig. S2b). With the abCap gene-expression module score (see Methods), we compared the abCap state to the endothelial cells of healthy controls and five disease states – COVID, lymphangioleiomyomatosis (LAM), idiopathic pulmonary fibrosis (IPF), systemic sclerosis-associated interstitial lung disease (SSc-ILD), and pulmonary arterial hypertension (PAH) in these six source studies. Comparison of the abCap state revealed little homology with other previously described endothelial cells, with fewer than 12% of cells in any previously reported endothelial sub-population having transcriptional overlap with abCaps (Fig. 3h). As a positive control that our selected module score can identify abCap features, directly applying this metric to our entire dataset confirmed that the cells containing the genes with the highest gene-module score were the BPD + PH-associated abCap cluster.

Analyzing putative ligand-receptor interactions identifies deficient semaphorin signaling between abCaps and the alveolar niche in BPD and BPD + PH

To characterize how abCaps might interact with other cells in the alveolar niche, we interrogated our scRNAseq data with CellChat, which predicts signaling interactions by expression of ligand-receptor pairs. In intracellular communication, cells can be classified as having more outgoing or incoming signaling information[20]. Broadly, CellChat identified endothelial capillaries as acting predominantly as “receivers” of cell-signaling inputs, with AT1 cells and mesenchymal cells having a higher “sender” score (Fig. 4a). CellChat identified semaphorin signaling as a leading differentially expressed candidate pathway that differentiated term controls versus infants with BPD and BPD + PH, with significantly decreased interactions in both epithelial-capillary and mesenchymal-capillary interactions (Fig. 4b). Interrogation of the transcriptomic dataset for the specific ligands and receptors of this pathway revealed that BPD + PH patients had markedly decreased expression of *SEMA3B* by AT1 cells, *SEMA3C* by alveolar and ductal myofibroblasts, and *SEMA6A* by endothelial capillary cells, with some decreased expression of these ligands in patients with severe BPD without PH, compared to controls (Fig. 4c). Expression of semaphorin receptors *NRP1* and *PLXNA2* was also decreased (Fig. 4c) in BPD and BPD + PH patients. Some of these predicted expression changes were spatially validated by RNA ISH, showing decreased expression of *SEMA3B* in *AGER*-expressing AT1 cells and decreased *SEMA6A* in *PECAM1*+ alveolar endothelial cells (Fig. 4d, e).

Decreased semaphorin ligand expression replicated in a murine BPD model

To determine whether the decreased semaphorin signaling seen in human lung could be modeled in vivo, we examined a murine model of saccular-stage injury for semaphorin ligand expression. Injury by exposure to hyperoxia (70% O_2) and inflammation (by intranasal LPS) in the saccular stage at postnatal day (P) 1-P5 mimics many of the features of BPD, including impaired alveologenesis (as measured by increased mean linear intercept, a measurement of terminal airspace subdivision) (Fig. 5a–c)[21,22]. RNA ISH on samples from this model demonstrated decreased expression of semaphorin ligands *Sema3b* and *Sema6a* in a cell-type specific pattern (Fig. 5d, e) similar to that observed in human patients (Fig. 4d, e). Notably, by RNA ISH, there was abundant expression of *Sema3b* and *Sema6a* in uninjured AT1 cells and endothelial cells, respectively, suggesting a conserved role of semaphorin signaling in early alveologenesis under normal developmental conditions.

Human infants with ACDMPV demonstrate decreased semaphorin ligand and receptor expression

Previous work has linked semaphorin signaling to regulation by endothelial-cell transcription factor FOXF1[23], a known upstream mediator of angiogenesis. Mutations in *FOXF1* result in alveolar capillary dysplasia with misalignment of pulmonary veins (ACDMPV), a rare and lethal form of developmental lung disease with severe pulmonary hypertension. To further explore the relationship between FOXF1 and semaphorin signaling, we interrogated a recently published single-nucleus RNA sequencing dataset from infants and children with ACDMPV and age-matched controls[11] (Fig. 6a, and S3) for expression of semaphorin ligands and receptors. While the expression deficits in semaphorin signaling are not identical in ACDMPV and BPD + PH (Fig. 6b), we note that patients with ACDMPV have decreased *SEMA3B* expression in AT1 cells and decreased *SEMA6A* signaling in alveolar capillaries (aCaps + gCaps), reminiscent of the deficits in BPD + PH.

BPD and BPD + PH lung tissues demonstrate decreased FOXF1 expression

Interrogation of our neonatal transcriptomic dataset demonstrated significantly decreased *FOXF1* in the capillary endothelium in BPD and BPD + PH conditions (Fig. 7a). Immunofluorescence for FOXF1 and the capillary hallmark protein ERG on BPD, BPD + PH, term control, and preterm control lung tissue substantiates the decreased nuclear localization of FOXF1 within endothelial cells of both disease states (Fig. 7b), suggesting a functional deficit of this critical transcription factor in BPD and BPD + PH.

Discussion

In this single-cell transcriptomic analysis of human neonatal lungs across stages of injury, we identified a previously undefined endothelial cell state, abCap, that is enriched in lungs from infants with bronchopulmonary dysplasia and pulmonary hypertension (BPD + PH). Using predictive analysis tools and tissue validation, we discovered deficits in semaphorin signaling and an apparent functional decrease in FOXF1 in infant tissue representing BPD and BPD + PH. The finding of decreased semaphorin signaling also in tissues from the rare disease ACDMPV suggests an underlying cell-signaling commonality with the severe forms of abnormal lung development that occur with preterm birth. The concordance of these findings in a murine BPD model provides additional foundational data to support a pivotal role for semaphorin signaling as a mediator of normal lung development

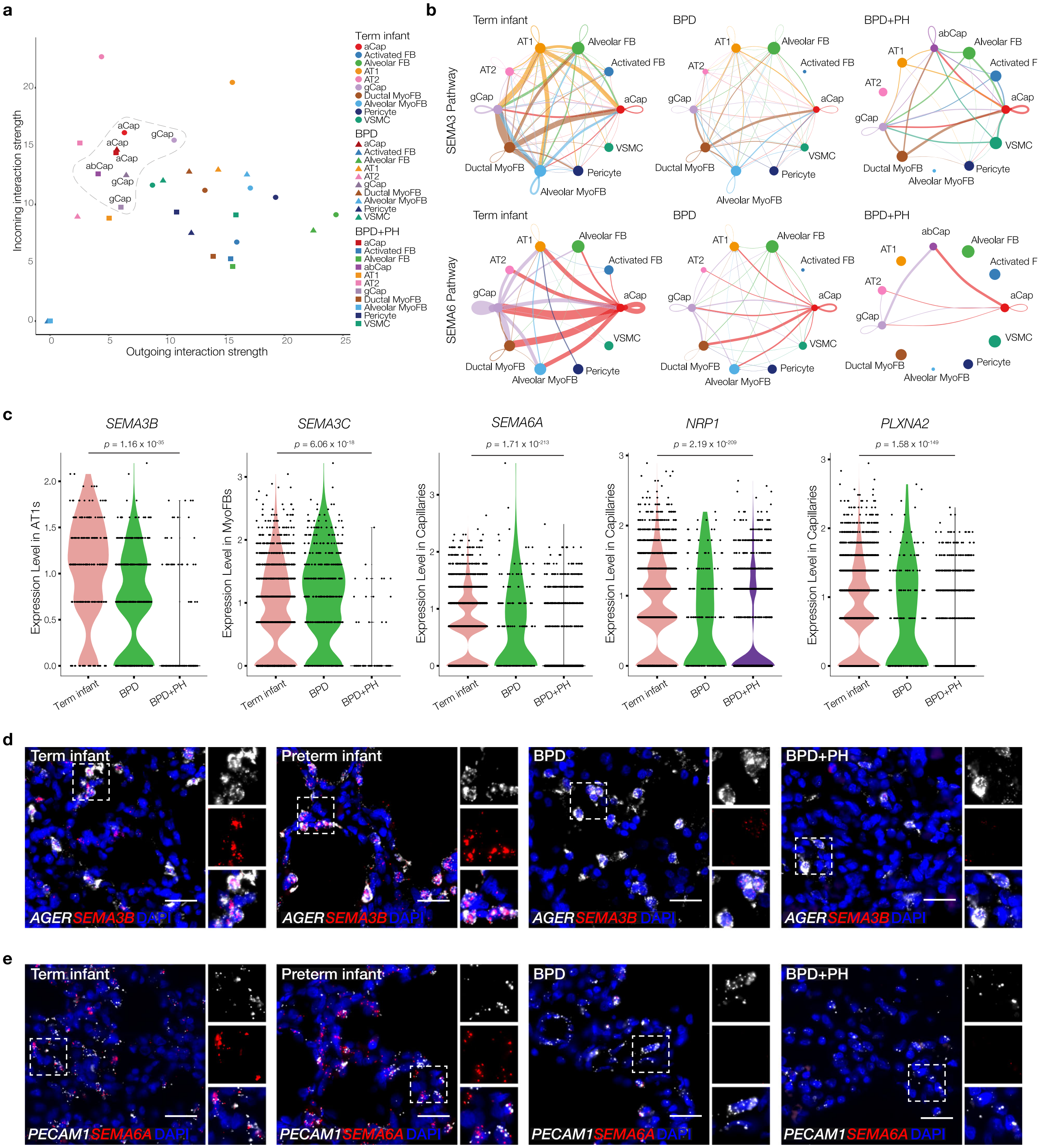


Fig. 4 | Predicted ligand-receptor interactions by single-cell sequencing analysis identifies decreased semaphorin signaling between endothelial cells and the alveolar niche in BPD and BPD + PH. a Cell types plotted by predicted outgoing versus incoming information averaged for all detected pathway interactions from the CellChat ligand and receptor database. **b** Circle plots for the SEMA3 and SEMA6 pathways displayed by disease state. Line weights indicate signal interaction strength from RNA levels of ligands and receptors for each respective pathway. Dot sizes indicate interaction strength per cell type. **c** Violin plots for expression levels of semaphorin pathway genes *SEMA3B*, *SEMA3C*, *SEMA6A*, *NRP1*, and *PLXNA2* by disease state, with cell-type specificity indicated on y-axis. *p*-values (indicated in the figure) were calculated by one-way ANOVA. **d**) RNA ISH for *SEMA3B* (red) and *AGER* (white), demonstrating co-localization of *SEMA3B* within *AGER* + AT1 cells in term ($n = 3$) and preterm ($n = 5$) uninjured infants, with decreased expression in patients with BPD ($n = 2$) and BPD + PH ($n = 4$). 10–12 images were taken per sample. Scale bar = 25 μm. **e** RNA ISH for *SEMA6A* (red) and *PECAM1* (white), demonstrating co-localization of *SEMA6A* within *PECAM1*+ endothelial cells in term ($n = 3$) and preterm ($n = 5$) uninjured controls, with decreased expression in BPD ($n = 2$) and BPD + PH ($n = 4$). 10–12 images were taken per sample. Scale bar = 25 μm.

whose disruption may lead to the abnormal alveologenesis and vascular development observed after preterm birth[24].

There are potential limitations in this study. While transcriptomic data from neonates at varying stages of development and evolving injury are informative in generating new hypotheses about molecular mechanisms that drive abnormal lung development after preterm birth, these data come from a limited number of subjects, and it is important to recognize that autopsy samples in general have an inherent bias toward the sickest individuals who have died from BPD and BPD + PH. Nonetheless, these data fill a critical gap regarding the

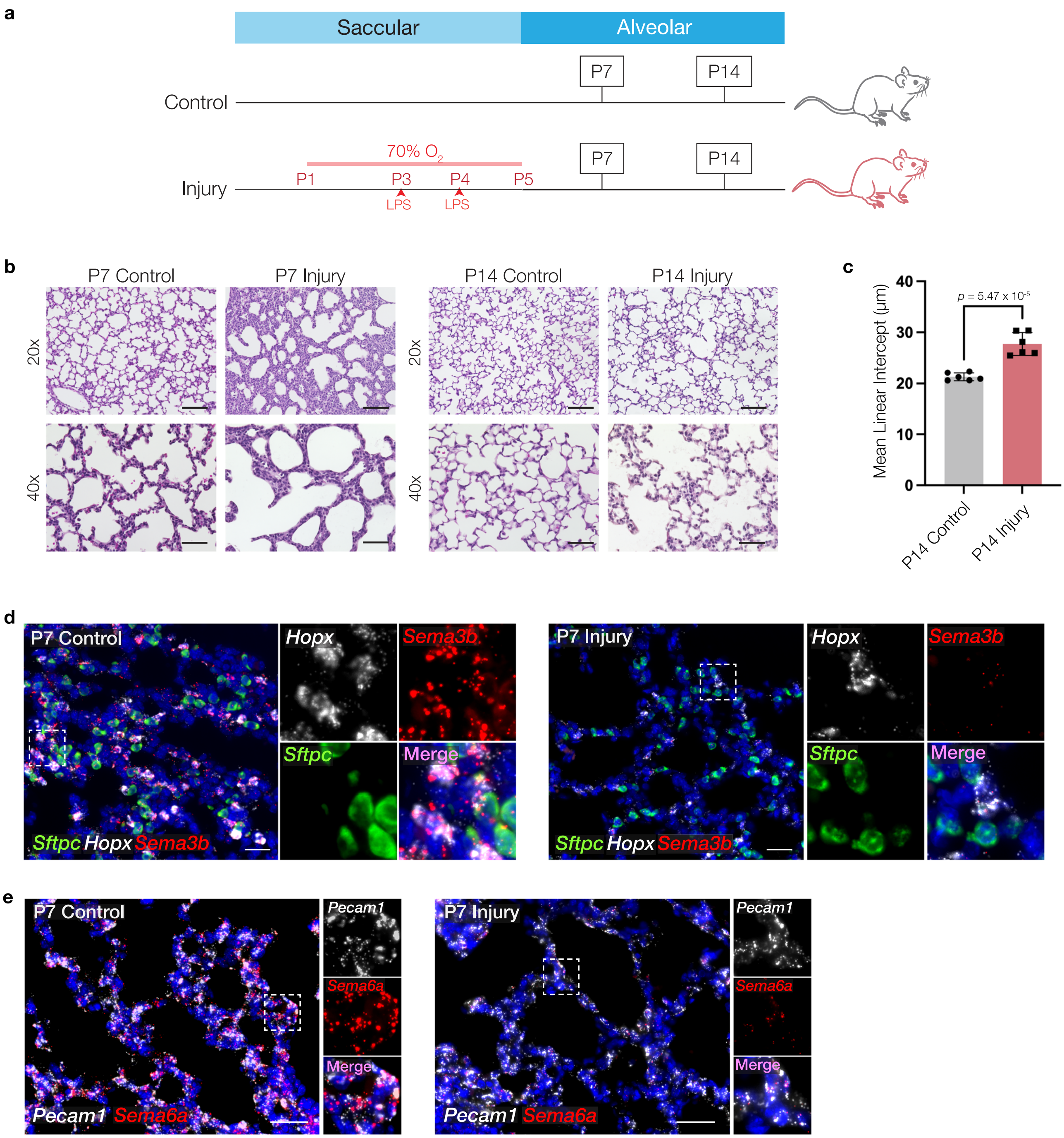


Fig. 5 | Murine injury models of BPD and BPD + PH display decreased semaphorin signaling. a Schematic of murine BPD model with hyperoxia exposure from postnatal day (P) 1-P5 and intranasal LPS on P3 and P4. **b** H&E of the BPD murine injury model and control mice at P7 (control $n = 6$, injury n = 6) and P14 (control $n = 6$, injury $n = 6$) at 20x and 40x magnification. Scale bar = 110 μm (20x), 55 μm (40x). **c** Mean linear intercept on the P14 control ($n = 6$) and injured ($n = 6$) mouse H&E images at 40x magnification. At least 10 images were taken at 40x per sample. p-value (indicated in the figure) was calculated by two-sided Welch's t-test. Data are presented as mean values ± standard deviation. Each dot represents an individual mouse. Source data are provided in a Source Data file. **d** RNA ISH of *Sftpc* (green), *Hopx* (white), and *Sema3b* (red) on the P7 control ($n = 5$) and injured ($n = 5$) mouse lungs exhibiting co-localization of *Sema3b*+ within *Hopx*+ epithelial cells in the alveolar parenchyma. 10 images were taken per sample. Scale bar = 20 μm. **e** RNA ISH of *Pecam1* (white) and *Sema6a* (red) on the P7 control ($n = 5$) and injured ($n = 5$) mouse lungs exhibiting co-localization of *Sema6a*+ within *Pecam1*+ endothelial cells in the alveolar parenchyma. 10 images were taken per sample. Scale bar = 35 μm.

deep transcriptomic analysis of normal and injured lungs across the lifespan, because there are so very few published reports on neonatal premature injured and uninjured lung samples.

Moreover, these data suggest a previously underappreciated cellular and molecular distinction between infants with BPD and BPD + PH. By definition, BPD is a clinical diagnosis determined by an infant's oxygen needs at term corrected gestational age[25]. Clinically, we appreciate different endotypes of BPD, with some patients having proximal airway malacia, some having lower airway disease, some having deficits in distal alveologenesis with resultant hypoxia, and some having pulmonary hypertension with impaired vasculogenesis[26]. As we compile and curate repositories of neonatal human lung tissue and other related samples, it will be essential to gather the kind of clinical metadata that will discern between these BPD endotypes, with

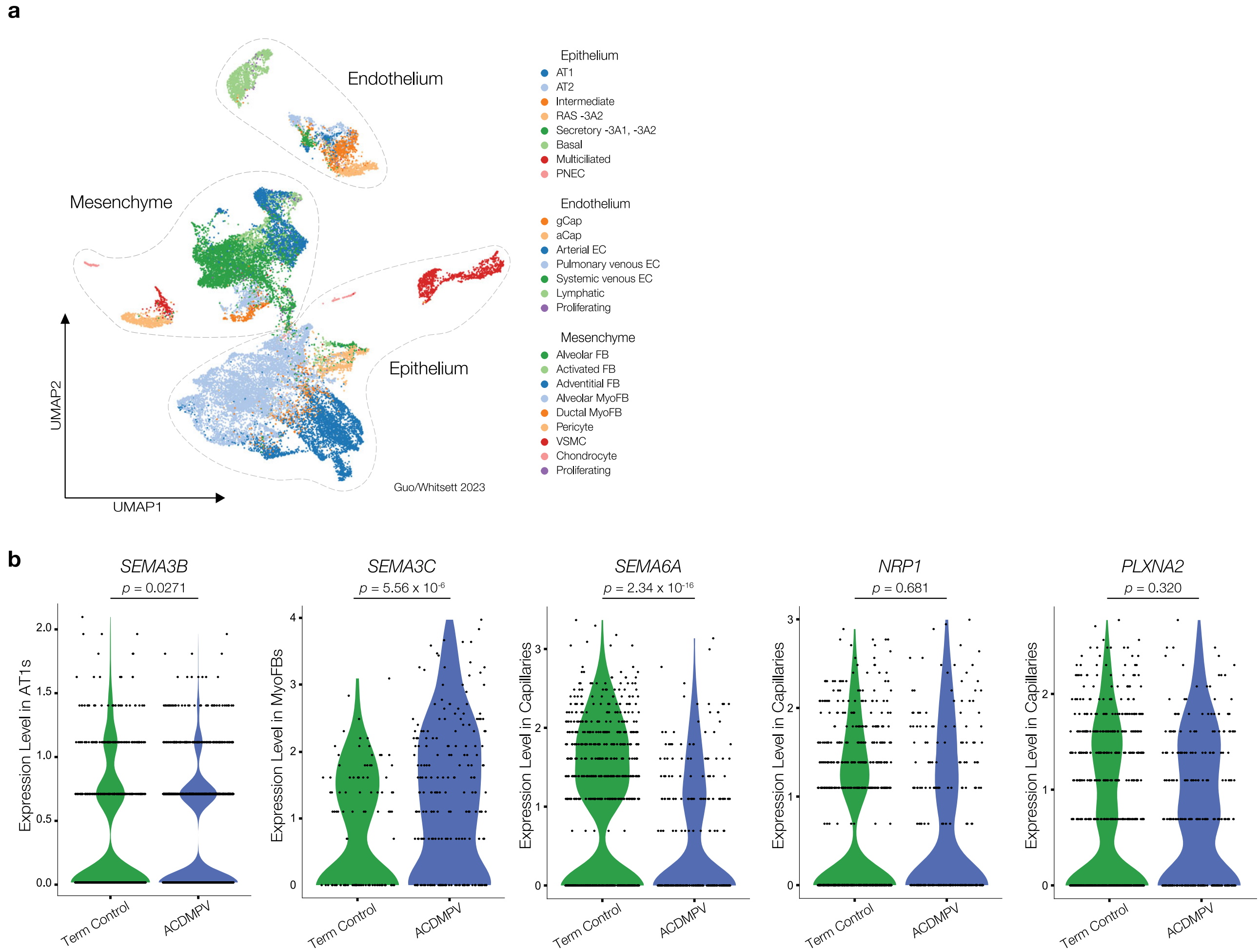


Fig. 6 | Human infant ACDMPV single-nucleus RNA sequencing identifies decreased semaphorin signaling. a UMAP embedding of single-nucleus RNA sequencing of five ACDMPV neonatal lung samples[11], three preterm, and three term control lung samples (*n* = 29,387 cells) with an age range of 29 weeks gestational age to 3.5 years old annotated by lineage and cell type. **b** Violin plots for expression levels of semaphorin pathway genes *SEMA3B*, *SEMA3C*, *SEMA6A*, *NRP1*, and *PLXNA2* by disease state, with cell-type specificity indicated on y-axis. *p*-values (indicated in the figure) were calculated by two-sided Welch's *t*-test.

an overall goal of moving towards hypotheses about the molecular mechanisms that drive these divergent clinical manifestations.

With the advent of single-cell transcriptomics, one can envision that granular clinical phenotyping based on molecular features could help stratify patients for potential disease-specific targeted therapies. To do so, one must identify the correct kinds of samples to be comparators for diseased lung – a critical and complex undertaking since BPD is a disease in which developmental trajectories are altered and all preterm infants experience injurious exposures of oxygen and mechanical ventilation. In this study, we employed two different comparisons for tissue samples as controls without lung injury for our BPD and BPD + PH lungs: 1) healthy term newborn lungs from infants who did not have respiratory disease, 2) preterm infants from 23-26 weeks gestation who survived < 6 h after birth. Within the recent and historical literature about BPD, there are inconsistent approaches for which kinds of human/infant lung tissue to use for comparison with BPD, with tissue samples labeled as "controls" having significant variability in chronological age. We recognize that the interplay of gestational age at birth, environmental exposure, and chronological age at sample collection contribute to both BPD pathology and to the variability of gene expression and histology observed in non-BPD subjects, and that the selection of tissue for comparison is limited by the rare number of pediatric samples. Ergo, reporting gestational and chronologic age for comparator tissue in current and future biorepositories is vital to our ability to interpret the data in our BPD patients and to uncovering aberrant cellular and molecular processes in developmental lung disease.

A central strength of our study is the transcriptomic integration of these rare BPD and BPD + PH samples with samples from ACDMPV patients. Although we do not find the abCap state in ACDMPV, the observed deficits in semaphorin signaling common to BPD + PH and ACDMPV suggest that the functional deficits in *FOXF1* associated with the developmental phenotypes of BPD and BPD + PH may be mechanistically related to the deficits in ACDMPV. Emerging technologies in spatial transcriptomic analysis offer increasing potential for interrogation of archival FFPE tissue blocks, which we anticipate will facilitate rapid growth in studies of normal development and disease across the neonatal and pediatric populations. Making these emergent data accessible and searchable could promote discovery of plausible mechanistic targets for promoting lung growth and regeneration and allow for a larger representation of samples across institutions and platforms.

While further work is required to define when abCaps emerge in disease and the role they may play in the pathogenesis of BPD, the discovery and validation of the abCap cell-state broadens our understanding of the endothelium as a receiver and coordinator of cell signaling and cell movements during this specific and crucial developmental stage. Analysis across hundreds of published datasets reveals that abCaps are distinct from the diversity of endothelial cell types and states observed in other diseases (e.g. adult PH, COVID,

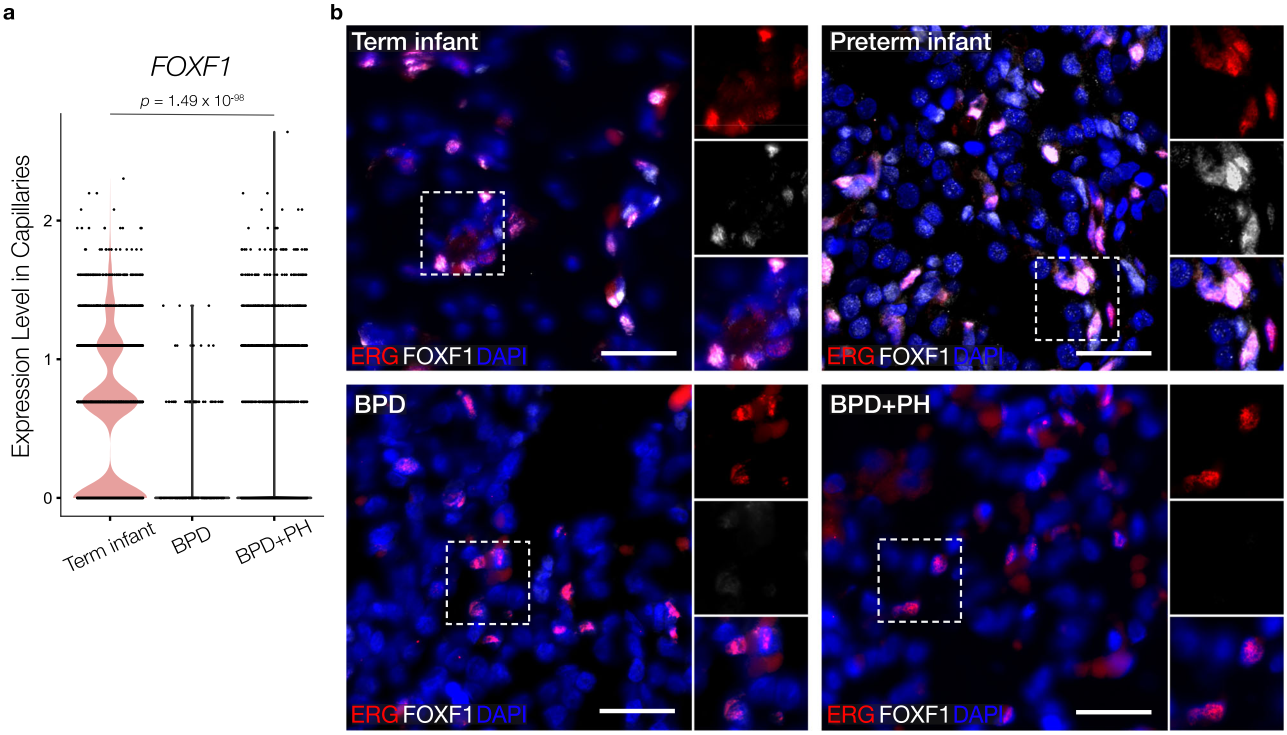


Fig. 7 | FOXF1 has decreased expression in human neonatal lung injury. **a** Violin plots for expression levels of endothelial transcription factor *FOXF1* in the human neonatal BPD and BPD + PH single-cell RNA sequencing dataset comparing disease state in capillaries, *p*-value (indicated in the figure) was calculated by one-way ANOVA. **b** Immunofluorescence for ERG (red) and FOXF1 (white) demonstrating co-localization of FOXF1 within ERG+ endothelial cells in the distal lung parenchyma in term ($n = 3$) and preterm ($n = 5$) uninjured controls, with decreased FOXF1 in patients with BPD ($n = 2$) and BPD + PH ($n = 4$). At least 10 images were taken per sample. Scale bar = 25 μm.

ARDS), and suggests that abCaps may be a cell state specific to BPD + PH (although single-cell analyzes of pediatric PH remain underrepresented in published datasets). Deeper characterization of these cells demonstrates increased expression of YAP signaling targets[27–29] (e.g., *CCN1*, *CCN2*, *SERPINE1*, and *CRIM1*, along with abCap hallmark and YAP target gene *ANKRD1*). While the VEGF-induced YAP/TAZ signaling responsiveness in endothelial cells is required for vasculogenesis[27–30], elevated YAP/TAZ target expression in abCaps in BPD + PH patients could indicate an attempted cell-autonomous compensatory response within abCaps themselves or arising as a response to abnormal cell-signaling by injury-affected adjacent cells. Indeed, in murine models of hyperoxia-induced injury in neonates and hypoxia-induced PH in adults, an increase in endothelial *CCN* and *Ankrd1* expression was reported[31–33]. Understanding how failed injury repair after preterm birth leads to severe BPD + PH in some patients, and how abCaps affect and respond to the complex signaling milieu of the post-injury alveolar niche is critical to untangling the mechanisms of abnormal repair in BPD and BPD + PH[34].

CellChat analysis predicted alveolar capillaries as receivers of signaling information from adjacent alveolar epithelial and mesenchymal cells. In BPD and BPD + PH, and murine hyperoxia BPD models, we noted dramatic reductions in expression of factors involved in semaphorin signaling. The semaphorin pathway is well-defined in neuronal axon directional movement and synapse coordination, especially in the cellular targeting of growth cones toward target cells[35]. While the role of this pathway in lung development has generally received little attention, a global Sema3a inactivation resulted in high perinatal lethality with marked impairment in alveologenesis in rare surviving KO pups[36], global knockdown of *Sema3C* impaired alveologenesis[37], and ectopic *Sema3a* overexpression attenuated apoptosis in neonatal rats exposed to hyperoxia[38]. Further, genetic inactivation of semaphorin receptor Nrp1 (neuropilin-1) in the adult mouse impaired airspace remodeling after smoke exposure[39], suggesting a role for semaphorins in repair after injury. Additionally, genomic analysis of patients with anomalous pulmonary venous connections and associated PH identified an association with mutations in *SEMA3D*[40]. Our current study builds from those foundational papers to dissect the cell-type specific expression of distinct semaphorins in the setting of alveologenesis and the functional loss of semaphorin signaling during injury and repair. Our findings on semaphorins add to previous work identifying axonal-guidance signaling among several pathways with differential DNA methylation in BPD[41] as well as those made on another axon-guidance pathway (Netrin/DCC/Unc) in fine-tuning the size and shape of emerging epithelial buds during early lung development[42], although the latter study is of a much earlier developmental stage than the alveolar stage reported here.

A central role of semaphorins is in regulating movement of cells toward or away from each other. Recent 4D-live imaging experiments suggest that alveologenesis results from precisely coordinated migration and shape change of epithelial and endothelial cells as they move away from the alveolar duct to form a common functional unit of gas exchange[43], with AT1 and aCap cells tightly aligned on a shared basement membrane. Our data implicate the semaphorin pathway as regulating these complex cellular movements and shape changes. In BPD + PH, whether abCaps still positively contribute to alveolar function or are the major contributor to deficit is unclear. Although abCaps present with characteristics notably distinct from prior published cell populations, they do share some transcriptional overlap with endothelial tip cells, a specialized epithelial cell that provides guidance for the growth and migration of capillaries during angiogenesis. Expressing some but not all features of endothelial tip cells in evolving BPD + PH may be indicative of compensatory response during repair or representative of a misaligned developmental program in abCaps because of decreased cues from semaphorins and other intercellular signals. Together, the decreased expression of *SEMA3B* by AT1 cells, *SEMA6A* by aCaps and gCaps, *SEMA3C* in the mesenchyme and decreased *NRP1* expression in endothelial cells under arrested lung development brings semaphorin signaling to the foreground in the injured and normal contexts. A deeper characterization of the timing and cell specificity of semaphorin signaling during lung development may reveal critical windows in which this pathway could be targeted to promote regrowth and structural repair after injury.

Methods

Ethical statement

The Human Infant Lung Repository at Vanderbilt University Medical Center has been reviewed and approved by the Vanderbilt University Institutional Review Board and the Biorepository at Seattle Children's Hospital Institutional Review Board, with a non-human subjects determination (no additional consent required). The murine neonatal lung injury model protocol was approved by the Institutional Animal Care and Use Committee of Vanderbilt University and was in compliance with the Public Health Services policy on humane care and use of laboratory animals.

Human infant lung repository

At the time of autopsy/biopsy, fresh lung tissue was isolated and processed for single-cell sequencing (scRNAseq) as described previously[5] with additional details below. In addition, we studied formalin-fixed paraffin embedded tissue blocks from human infant lung samples from autopsies and biopsies in varying stages of lung injury and without evidence of lung injury (as determined by a clinical pathologist) have been collected. All samples are de-identified with the exception of sex, gestational age at birth, age at time of death, cause of death, and whether or not the patient had evidence of BPD (defined as a need for supplemental oxygen at 36 weeks correct gestational age) or pulmonary hypertension (as determined by clinical echocardiography). Given the small number of human samples, this study was not powered to detect differences by sex. A table with the clinical characteristics of the infant lungs used for FFPE tissue and scRNAseq is found in the supplement (Supplementary Tables 1, 2). The preterm uninjured archived FFPE lung sections were collected and processed at Seattle Children's Hospital, and all other lung tissue was collected at Vanderbilt.

Single-cell sequencing

Human lung tissue from infants who died with BPD, BPD + PH, acute preterm lung injury, and term infant controls were processed into single-cell suspension as described previously[5]. scRNAseq libraries were generated using the 10x Chromium platform 3' or 5' library preparation kits (10x Genomics) following the manufacturer's recommendations and targeting 10,000–20,000 cells per sample. Next-generation sequencing was performed using an Illumina Novaseq 6000. Reads with read quality less than 30 were filtered out and CellRanger Count v7.1.0 (10x Genomics) was used to align reads onto the GRCh38 reference genome. Full code used in the analysis of these data is available at https://github.com/SucreLab/HumanNeonatal. A table with the clinical characteristics of infants whose lungs were sequenced is in the supplement (Supplementary Table 1).

Single-cell RNA sequencing analysis

The ambient RNA was removed from the CellRanger count matrix h5 file outputs using CellBender v0.2.1[44]. The scRNAseq data was analyzed using Seurat (v4.4.0)[45] in R (v4.3.2). The quality control of the ambient RNA corrected cells filtered out cells with > 10% mitochondrial mRNA mapped UMIs, < 500 identified genes, and > 7000 identified genes. The data was normalized using SCTransform (v0.4.1)[46] with the glmGamPoi (v1.14.0)[47] method by stabilizing the variance of the total molecular counts. The default parameters were used with SCTranform except for: batchvar set to correct batch effects from the tissue samples processed, vars.to.regress set to regress out the effects on differential gene expression analysis due to the presence of increased mitochondrial mRNAs and ribosomal rRNAs, and ncells set to NULL. The reduction of dimensionality was calculated via PCA. The PCA embeddings were further integrated with Harmony (v1.2.0)[48] to account for any remaining batch effects from the samples. The Harmony reduction was then applied in obtaining the UMAP embeddings and the shared nearest neighbor graph which subsequently allowed for the Louvain algorithm to find the cell clusters.

The clusters were split into the four main cell subtypes of epithelial, endothelial, mesenchymal, and immune based on respective hallmark genes (Supplementary Data 4). Each subtype went through an iterative process of SCTransform, PCA, Harmony, UMAP, shared nearest neighbors, and clustering to isolate and remove any remaining doublet or nonbiological cells. For each iteration, the cluster resolution would start in the range of $1 < n \leq 2$, and with each iteration, n would be brought down slightly until the last iteration which would have $n = 1$. Cell type labels were assigned after quality control and cleaning of the data from their respective hallmark genes (Supplementary Data 4, 5). After annotation, the four subtype-group Seurat objects were merged to produce the final Seurat object.

Differential expressed gene (DEG) analysis comparing across conditions was conducted with the FindMarkers function in the Seurat package using the MAST method[49]. Genes with an average $\log_2$ fold change of either greater than 1 or less than −1 and with an adjusted p-value of less than 1×10^{-5} were marked as significant up- or down-regulated in the condition comparison. The basal, secretory *MUC5B*, and abCap cell-types were excluded from the DEG analysis due to having 3 or less cells within at least one of the conditions being compared. Mitochondrial genes and ribosomal genes (rRNA) were excluded from the list of significant DEG, as were genes *FP236383.3* and *FP671120.4* (due to > 95% homology with rRNA[50]). The full list of DEGs of the parenchymal cell-types comparing term infant to BPD, term infant to BPD + PH, and BPD to BPD + PH is in Supplementary Data 1–3, respectively.

CellChat (v1.5.0)[20] was used to predict possible ligand-receptor pair interactions between cell types of the same condition to generate hypotheses about possible signaling mechanisms associated with BPD and BPD + PH. We employed SCimilarity (0.2.0)[15] to query cell populations or states that could have similar gene expression as the BPD + PH endothelial cell state (abCaps). To do this, we used a gene signature list curated from the intersection of the top Wilcoxon rank sum method ranked genes, top $\log_2$ fold change expression genes, and genes with the highest proportion expressed within abCaps versus all other cell types: *ANKRD1*, *CCN2*, *KITLG*, *SERPINE1*, *TNFRSF12A*, *CRIM1*, *GDF15*, *CCND1*, *EDN1*, *RBFOX2*, *CD151*, *PLS3*, *CD59*, *GBE1*, *PODXL*. We then applied this gene module score to all cell populations or clusters in each study independently, with a cutoff of a cell score greater than two standard deviations from the mean cell score in the dataset. From the top 100 studies by SCimilarity score, we excluded 7 studies of iPSCs, 14 studies that did not contain lung tissue, two with samples from patients in the 7th–8th decade of life, two with a sample size < 2, and two with inaccessible scRNAseq data files. This subtraction left six studies with 178 samples for comparison. In summary, we looked for cells with features of the expression pattern identified in the abCap gene module. We then plotted the percentage of cells within each cluster that passed this cutoff.

Reanalysis of the recently published single-nucleus RNA sequencing dataset of infants who died of ACDMPV and term control infants was performed using the same strategy used above for scRNAseq.

Murine model of neonatal lung injury

C57BL/6 mice (from Jackson Laboratories) were exposed to 70% O_2 (hyperoxia) or 21% O_2 normoxia conditions from P1–P5, with hyperoxia animals receiving intranasal LPS at 1 μg/g on P3 and P4. On P7 and P14, lungs were inflation-fixed by gravity filling with 10% buffered formalin and paraffin embedded as described previously[51]. The mice were housed under 12:12 h dark/light cycle with free access to food and water, ambient temperature of 72° ± 4°, and relative humidity of 50% ± 15%. Equal numbers of male and female mice were used, with no differences noted in phenotype when analyzed separately.

RNA in situ hybridization

RNAScope technology (ACDBio) was used to perform all RNA ISH experiments according to the manufacturer's instructions as described previously. RNAScope probes to the following human genes were used (*SEMA3B*, *AGER*, *SFTPC*, *SEMA6A*, *PECAM1*). RNAScope probes to the following mouse genes were used*:* *Sftpc*, *Hopx*, *Pecam1*, *Sema3b*, *Sema6a*. Positive control probe (PPIB) and negative control probe (DapB) were purchased from ACDBio and performed with each experimental replicate.

Immunofluorescence

Immunofluorescence was performed as described previously[43,52,53], with the following primary antibodies used: anti-FOXF1 (1:100, R&D systems AF4798), anti-ERG1 (1:100, Abcam ab214341) and the following secondary antibodies used: 1:200 Donkey anti-Goat (Invitrogen A21447), 1:200 Donkey anti-Mouse (Invitrogen A31570). Prior to adding primary antibodies, slides were blocked in 5% normal donkey serum in PBST at room temperature for 2-3 hrs.

Image acquisition and analysis

We acquired fluorescent and brightfield images using a Keyence BZ-X810 with BZ-X Viewer software and with 10x, 20x, and 40x objectives. Images were collected with the following filter sets: 405, 488, 561 and 647 nm. Automated image analysis was performed with HALO software (Indica Labs) to quantify the number of *ANKRD1* + *PECAM1*+ endothelial cells within all *PECAM1*+ endothelial cells in the alveolar parenchyma. After HALO analysis, the rate of false co-localization of non-physiological expression combinations (*SFTPC* + *PECAM1*+) was low, with a mean of < 1%.

Lung morphometry

Mean linear intercept (MLI) was performed on a minimum of ten 40x images per mouse ($n = 6$ mice per group) by measuring the distance between tissue intersections of a minimum of 3 straight lines, excluding vessels and large airways[51,52], with the use of ImageJ and AlveolEye for automated calculation of MLI measurements (https://github.com/SucreLab/AlveolEye).

Statistics

Statistical analysis was performed in R and with Prism Graphpad version 10.0.03, with specific details on statistical tests provided in Fig. legends.

Reporting summary

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

Data availability

The data supporting the findings from this study are available within the manuscript and its supplementary information. The raw and processed single-cell sequencing data generated in this study have been deposited in the GEO database under accession code GSE275938. The processed sequencing data are available as a freely searchable tool at www.sucrelab.org/lungcells and in the LungMAP website https://www.lungmap.net/dataset/?experiment_id=LMEX0000004401. Any additional raw data will be available from the corresponding authors upon request. Source data are provided with this paper. The single-cell RNA sequencing data used for the similarity analysis using SCimilarity in Fig. 3h are available in the GEO database under the accession codes GSE128169, GSE158127, GSE122960, GSE171668, GSE132771, and GSE136831. The ACDMPV single-nucleus sequencing data used in this study are from a previously published manuscript[11] and are available at the LungMAP website: https://data-browser.lungmap.net/projects/fdadee7e-2097-45d5-bf81-cc280bd8348e. Source data are provided with this paper.

Code availability

All code used in analysis and generation of sequencing data display can be found at https://github.com/SucreLab/HumanNeonatal.

References

1. Baraldi, E. & Filippone, M. Chronic lung disease after premature birth. *N. Engl. J. Med* **357**, 1946–1955 (2007).
2. Berkelhamer, S. K., Mestan, K. K. & Steinhorn, R. H. Pulmonary hypertension in bronchopulmonary dysplasia. *Semin Perinatol.* **37**, 124–131 (2013).
3. Frank, D. B. et al. Primary pulmonary vein stenosis during infancy: state of the art review. *J. Perinatol.* **41**, 1528–1539 (2021).
4. Adams, T. S. et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **6**, eaba1983 (2020).
5. Habermann, A. C. et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci. Adv.* **6**, eaba1972 (2020).
6. Schupp, J. C. et al. Integrated single-cell atlas of endothelial cells of the human lung. *Circulation* **144**, 286–302 (2021).
7. He, P. et al. A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates. *Cell* **185**, 4841–4860.e25 (2022).
8. Guo, M. et al. Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. *Nat. Commun.* **10**, 37 (2019).
9. Sikkema, L. et al. An integrated cell atlas of the lung in health and disease. *Nat. Med* **29**, 1563–1577 (2023).
10. Duong, T. E. et al. A single-cell regulatory map of postnatal lung alveologenesis in humans and mice. *Cell Genom.* **2**, 100108 (2022).
11. Guo, M. et al. Single cell multiomics identifies cells and genetic networks underlying alveolar capillary dysplasia. *Am. J. Respir. Crit. Care Med* **208**, 709–725 (2023).
12. Tsukui, T. et al. Collagen-producing lung cell atlas identifies multiple subsets with distinct localization and relevance to fibrosis. *Nat. Commun.* **11**, 1920 (2020).
13. Chen, W. et al. The endothelial tip-stalk cell selection and shuffling during angiogenesis. *J. Cell Commun. Signal* **13**, 291–301 (2019).
14. del Toro, R. et al. Identification and functional analysis of endothelial tip cell–enriched genes. *Blood* **116**, 4025–4033 (2010).
15. Heimberg, G. et al. A cell atlas foundation model for scalable search of similar human cells. *Nature* **638**, 1085–1094 (2025).
16. Delorey, T. M. et al. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* **595**, 107–113 (2021).
17. 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).
18. Valenzi, E. et al. Single-cell analysis reveals fibroblast heterogeneity and myofibroblasts in systemic sclerosis-associated interstitial lung disease. *Ann. Rheum. Dis.* **78**, 1379–1387 (2019).
19. Bharat, A. et al. Lung transplantation for patients with severe COVID-19. *Sci. Transl. Med.* **12**, eabe4282 (2020).
20. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

21. Hilgendorff, A., Reiss, I., Ehrhardt, H., Eickelberg, O. & Alvira, C. M. Chronic lung disease in the preterm infant. Lessons learned from animal models. *Am. J. Respir. Cell Mol. Biol.* **50**, 233–245 (2014).
22. Shrestha, A. K. et al. Interactive and independent effects of early lipopolysaccharide and hyperoxia exposure on developing murine lungs. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **319**, L981–L996 (2020).
23. Karolak, J. A. et al. Perturbation of semaphorin and VEGF signaling in ACDMPV lungs due to FOXF1 deficiency. *Respiratory Res.* **22**, 212 (2021).
24. Abman, S. H. & Lakshminrusimha, S. Pulmonary hypertension in established bronchopulmonary dysplasia: physiologic approaches to clinical care. *Clin. Perinatol.* **51**, 195–216 (2024).
25. Jensen, E. A. et al. The diagnosis of bronchopulmonary dysplasia in very preterm infants. an evidence-based approach. *Am. J. Respir. Crit. Care Med* **200**, 751–759 (2019).
26. Pierro, M., Van Mechelen, K., van Westering-Kroon, E., Villamor-Martínez, E. & Villamor, E. Endotypes of prematurity and phenotypes of bronchopulmonary dysplasia: toward personalized neonatology. *J. Pers. Med* **12**, 687 (2022).
27. Boopathy, G. T. K. & Hong, W. Role of hippo pathway-YAP/TAZ signaling in angiogenesis. *Front. Cell Dev. Biol.* **7**, 49 (2019).
28. Isago, H. et al. The epithelial expressions of YAP and TAZ are Sequentially Required in Lung Development. *Am J Respir Cell Mol Biol* https://doi.org/10.1165/rcmb.2019-0218OC. (2019).
29. Hooglugt, A., van der Stoel, M. M., Boon, R. A. & Huveneers, S. Endothelial YAP/TAZ signaling in angiogenesis and tumor vasculature. *Front Oncol.* **10**, 612802 (2021).
30. Wang, X. et al. YAP/TAZ Orchestrate VEGF signaling during developmental angiogenesis. *Developmental Cell* **42**, 462–478.e7 (2017).
31. Lingappan, K. et al. Analysis of the transcriptome in hyperoxic lung injury and sex-specific alterations in gene expression. *PLOS ONE* **9**, e101581 (2014).
32. Perkowski, S. et al. Gene expression profiling of the early pulmonary response to hyperoxia in mice. *Am. J. Respir. Cell Mol. Biol.* **28**, 682–696 (2003).
33. Lee, S. et al. CCN1 suppresses pulmonary vascular smooth muscle contraction in response to hypoxia. *Pulm. Circ.* **5**, 716–722 (2015).
34. Vila Ellis, L. & Chen, J. A cell-centric view of lung alveologenesis. *Developmental Dyn.* **250**, 482–496 (2021).
35. Jackson, R. E. & Eickholt, B. J. Semaphorin signalling. *Curr. Biol.* **19**, R504–R507 (2009).
36. Becker, P. M. et al. Semaphorin 3A contributes to distal pulmonary epithelial cell differentiation and lung morphogenesis. *PLoS One* **6**, e27449 (2011).
37. Vadivel, A. et al. The axonal guidance cue semaphorin 3C contributes to alveolar growth and repair. *PLoS One* **8**, e67225 (2013).
38. Liang, Z., Zhang, X., Liu, Y., Wu, Q. & You, C. SEMA3A protects against hyperoxia-induced lung injury in a bronchopulmonary dysplasia model of newborn rat by inhibiting ERK pathway. *Allergol. Immunopathol. (Madr.)* **49**, 8–15 (2021).
39. Le, A. et al. Pulmonary epithelial neuropilin-1 deletion enhances development of cigarette smoke–induced emphysema. *Am. J. Respir. Crit. Care Med* **180**, 396–406 (2009).
40. Degenhardt, K. et al. Semaphorin 3d signaling defects are associated with anomalous pulmonary venous connections. *Nat. Med* **19**, 760–765 (2013).
41. Cuna, A. et al. Alterations in gene expression and DNA methylation during murine and human lung alveolar septation. *Am. J. Respir. Cell Mol. Biol.* **53**, 60–73 (2015).
42. Liu, Y. et al. Novel role for Netrins in regulating epithelial behavior during lung branching morphogenesis. *Curr. Biol.* **14**, 897–905 (2004).
43. Negretti, N. M. et al. Epithelial outgrowth through mesenchymal rings drives alveologenesis. *JCI Insight* e187876 https://doi.org/10.1172/jci.insight.187876 (2025).
44. Fleming, S. J. et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using cellbender. *Nat. Methods* **20**, 1323–1335 (2023).
45. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e29 (2021).
46. Choudhary, S. & Satija, R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **23**, 27 (2022).
47. Ahlmann-Eltze, C. & Huber, W. glmGamPoi: fitting gamma-poisson generalized linear models on single cell count data. *Bioinformatics* **36**, 5701–5702 (2021).
48. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
49. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **10**, 278 (2015).
50. Loi, D. S. C., Yu, L. & Wu, A. R. Effective ribosomal RNA depletion for single-cell total RNA-seq by scDASH. *PeerJ* **9**, e10717 (2021).
51. Sucre, J. M. S. et al. Hyperoxia injury in the developing lung is mediated by mesenchymal expression of Wnt5A. *Am. J. Respir. Crit. Care Med* **201**, 1249–1262 (2020).
52. Schuler, B. A. et al. Age-determined expression of priming protease TMPRSS2 and localization of SARS-CoV-2 in lung epithelium. *J. Clin. Invest.* https://doi.org/10.1172/JCI140766 (2020).
53. Negretti, N. M. et al. A single-cell atlas of mouse lung development. *Development* **148**, dev199512 (2021).

Acknowledgements

This work was supported by the Francis Family Foundation (JAK, DBF, JMSS, NMN, LCE), R01HL168556 (JMSS), K08HL143051 (JMSS) R01HL145372(JAK/NEB), R56HL167937 (DBF), U01HL175444 (JAK, NEB, JMSS), K08HL140104 (LCE). We are grateful to Anne Hilgendorff and Vivian Siegel for thoughtful discussions and insights.

Author contributions

Conceptualization: J.A.K., J.M.S.S., D.B.F., C.V.E.W., Methodology: S.P.S., J.A.K., N.M.N., C.S.J., A.L.S., Validation: S.G., G.F., D.W., Resources: M.E.K., D.B.F., L.C.E., G.H.D., Investigation and Interpretation: C.S.J., J.A.K., J.M.S.S., S.P.S., N.M.N., B.D.F., C.V.E.W., S.M., N.E.B. Manuscript preparation (original): S.P.S., J.M.S.S., Manuscript preparation (revision): S.P.S., N.M.N., C.S.J., L.C.E., G.H.D.,A.L.S., S.G., M.E.K., G.F., S.M., N.E.B., C.V.E.W., D.B.F., J.A.K., J.M.S.S.

Competing interests

J.A.K. reports grant funding from Boehringer Ingelheim and Bristol-Myers Squibb outside the scope of this work, and serves on the Scientific Advisory Board for APIE and ARDA, outside of the scope of this work. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-60371-7.

Correspondence and requests for materials should be addressed to Jonathan A. Kropski or Jennifer M. S. Sucre.

Peer review information *Nature Communications* thanks Steven Abman and Yuxuan Zheng for their contribution to the peer review of this work. A peer review file is available.