Progressive Increases in Mesenchymal Cell Diversity Modulate Lung Development and are Attenuated by Hyperoxia

Fabio Zanini[1,+], Xibing Che[2,3], Nina Suresh[2,3], Carsten Knutsen[2,4], Paula Klavina[5], Yike Xie[1], Racquel Domingo-Gonzalez[2,4], Robert C. Jones[6], Stephen R Quake[6,7,8,*], Cristina Alvira[2,4,*], and David N. Cornfield[2,3,*,+]

*: These authors contributed equally.

+: To whom correspondence should be addressed: fabio.zanini@unsw.edu.au, davidco@stanford.edu

[1]Prince of Wales Clinical School, Lowy Cancer Research Centre, University of New South Wales, Sydney, Australia

[2]Center for Excellence in Pulmonary Biology, Stanford University School of Medicine, Stanford, CA

[3]Division of Pulmonary, Asthma and Sleep Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA

[4]Division of Critical Care Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA

[5]School of Biochemistry and Immunology, Trinity College Dublin, Dublin, Ireland

[6]Department of Bioengineering, Stanford University, Stanford, CA

[7]Chan Zuckerberg Biohub, San Francisco, CA

[8]Department of Applied Physics, Stanford University, Stanford, CA

Abstract

Early in life, the lung mesenchyme is extremely dynamic. We profiled the transcriptomes and locations of mesenchymal cells (MC) in the perinatal murine lung and observed cellular progressions of fibroblasts, myofibroblasts, airway smooth muscle (ASM), and mural cells, poised to modulate the extracellular matrix (ECM) with both subtype- and temporal-specific patterns. Within one day after birth, embryonic fibroblast precursors branched into two subtypes, alveolar and adventitial fibroblasts. ASM and myofibroblasts derived from a common precursor population that specifically expressed *Crh*, a hormone central to glucocorticoid release. Vascular smooth muscle cells and pericytes differentiated gradually. Fibroblast, myofibroblast, and pericyte proliferation peaked at 7 days after birth. Paracrine

signaling among cell types decreased after birth and was hierarchical, with pericytes at the interface between endothelial cells and other MC. Postnatal hyperoxia delayed the developmental progression in cell type abundance and gene expression, mirroring the arrested development that characterizes the neonatal lung disease, bronchopulmonary dysplasia. Hyperoxia decreased pericyte and myofibroblast abundances and proliferation, dysregulated signaling, altered ECM modulation and, in males specifically, gave rise to a novel population of contractile fibroblasts. This study identifies the mesenchymal populations orchestrating lung development at the critical transition to air-breathing life and their distinct reactions to hyperoxia.

Introduction

In utero, the lung is fluid filled and the pulmonary circulation develops in a relatively low oxygen tension environment with low blood flow and high pressure [1]. A few moments after birth, the perinatal lung undergoes a remarkable transition that enables gas exchange with establishment of a gas-liquid interface, a ten-fold increase in pulmonary blood and a marked decrease in pulmonary arterial pressure [2,3]. After birth, the distal lung undergoes structural remodeling including the formation of alveoli by secondary septation and a transition from a double to single capillary layer to increase gas exchange efficiency [4]. Alveolarization, the final stage of lung development, characterized by rapid pulmonary parenchymal and vascular growth, continues through at least the first decade of life in humans.

Lung development requires coordinated interactions between multiple cell types in the epithelial, endothelial, immune and mesenchymal populations [5,6]. Temporally and spatially regulated changes in gene expression of specific cell types mediate branching morphogenesis and the stereotypical proximal to distal outgrowth of the airways [7], vasculature and lymphatics that occurs with maturation [8]. Lung mesenchymal cells play a central role in secondary septation, thinning of the interstitium, and in transmitting mechanical forces that promote alveolarization [9]. Though the lung mesenchyme includes multiple, distinct cell types, knowledge surrounding the degree of cellular heterogeneity, dynamic changes in gene expression, and cell-cell interaction during the alveolar stage of lung development remains limited.

Insight into the lung mesenchyme during early postnatal life has significant implications for lung injury and regeneration, especially in neonates. Since the respiratory tract continues to mature in late gestation and through the first decades of life [10], premature infants are uniquely susceptible to life-threatening lung disorders, including bronchopulmonary dysplasia (BPD), a lung disease characterized by compromised alveolarization [11]. Given the importance of the mesenchyme in both physiologic and pathophysiologic lung development, we sought to interrogate the lung mesenchyme relative to cellular diversity, dynamic changes

in gene expression and cell-cell communication in the perinatal lung, during alveolarization, and in the context of hyperoxia-induced lung injury, a preclinical model of BPD [12].

In this report, we combined single cell transcriptomics (scRNA-Seq) with fluorescent multiplexed in situ hybridization (FISH) to characterize changes in composition, localization, and function of mesenchymal cells in the murine lung from just before birth through the first three weeks of postnatal life. Moreover, we applied the same strategies to cells derived from neonatal murine lung after 7 days of hyperoxia (80% oxygen). Mesenchymal cells fell into three broad categories: fibroblasts, airway smooth muscle (ASM)/myofibroblasts (MyoF), and mural cells, with 14 distinct subtypes overall. The most dynamic changes in cell phenotype occurred at birth as bipotent precursors for fibroblast and ASM/MyoF matured into distinct cell subtypes between E18.5 and P1. P7 represented the most proliferative time point across the lung mesenchyme, when a novel proliferating population of MyoF emerged. From P7 to P21, the transcriptomic profiles of MyoF and proliferative cells disappeared completely, while alveolar and adventitial fibroblasts and ASM underwent deep transcriptional changes towards cell quiescence. Hyperoxia decreased pericyte abundance and delayed maturation-related transcriptomic changes in the mesenchyme, mirroring the arrested development of the lung that characterizes BPD. Further, in hyperoxic males, but not females, a previously undescribed population of contractile fibroblasts emerged. Taken together, our data demonstrate that in the neonatal lung, the mesenchyme demonstrates remarkable cellular diversity, dynamic changes in gene expression, plays a critically important role in the communication with endothelial and immune cells, and demonstrates a marked sex-related difference to hyperoxia-induced neonatal lung injury.

Results

The perinatal lung mesenchyme is composed of three coarse cell types

To construct a single cell map of the perinatal lung mesenchyme, we generated single cell RNA-Seq (scRNA-Seq) data from eight mice at different stages of perinatal development, with two mice (one female and one male) from E18.5 (early saccular stage), P1 (late saccular), P7 (early alveolar) and P21 (late alveolar). Lung tissue was resected, perfused, and dissociated as described previously [13]. Fluorescence activated cell sorting (FACS) was used to enrich mesenchymal cells by negative selection (i.e. by depleting cells positive for CD326, CD31, and CD45 that constitute the epithelium, endothelium, and immune compartments, respectively) (**Figure 1A**). Smart-Seq2 library generation and sequencing were then constructed [13].

Mesenchymal cells are among the most diverse and least well characterized cells of the body [14]. We initially built a coarse representation of the perinatal lung mesenchyme by computing a t-distributed stochastic embedding (t-SNE) [15] and coloring each cell based on the expression of relatively established marker genes (**Figure 1B**). We thereby distinguished three coarse mesenchymal cell types: fibroblasts marked by either *Col13a1* or C*ol14a1* [16,17],

airway smooth muscle cells or myofibroblasts (ASM/MyoF) marked by *Tgfbi* [18], and mural cells marked by *Pdgfrb* [19] (**Figure 1C**). No cluster demonstrated high expression of both *Tgfbi* and *Pdgfrb*, previously described as vascular smooth muscle cells (VSM) markers [20], arguing for the importance of clarifying VSM annotations. Cell type abundance formed an interesting temporal pattern: fibroblasts were the majority cell type at all time points except P7, ASM/MyoF peaked at P7, and the relative abundance of mural cells gradually increased with maturation (**Figure 1D**). Temporal patterns were evident in the embedding with each cell colored by time point (**Figure 1E**).

Distinct mesenchymal subtypes differentially modulate the extracellular matrix composition

With this coarse cell ontology in mind, we clustered the cells [21] to identify each subtype. We identified a total of 14 distinct cell populations (**Figure 2A**). Based on our initial ontology, clusters 1-6 resembled fibroblasts, clusters 7-11 ASM/MyoF cells, and clusters 12-14 mural cells. The phenotypic richness of the pulmonary mesenchyme posed a challenge for cell type identification. We therefore harmonized our data with mesenchymal cells from Tabula Muris Senis (TMS) [22] via a batch-balanced k-nearest neighbors graph [23] and computed a separate embedding [24] colored by data source (**Figure 2B, top**) and cluster (**bottom**). Among our clusters, the ones that included only or mostly P21 cells (4, 6, 9, 12, 13, 14) merged best with TMS, which includes data from mice 1 month old or older. Clusters 13 and 14 overlapped with TMS pericytes, cluster 4 with TMS alveolar fibroblasts, cluster 6 with TMS adventitial fibroblasts, cluster 9 with TMS myofibroblasts, while cluster 12 embedded between smooth muscle-like cells and pericytes. The other clusters (1-3, 5, 7-8, 10-11) included cells predominantly from earlier developmental time points and did not overlap with adult clusters, supporting the notion of distinct mesenchymal cell phenotypes in the perinatal developing lung.

For each cluster, we identified marker genes by differential expression (**Figure 2C**). Unambiguous annotation of almost every cell subtype could be achieved by combining: (i) marker gene expression; (ii) TMS harmonization; (iii) our initial coarse ontology; and (iv) the most common time point of the cells populating the cluster. For cluster 12, the coexpression of both smooth muscle (*Tagln*, *Acta2*) and mural (*Prgfrb*) genes identified this population as vascular smooth muscle (VSM) cells, a cell type that was not annotated in TMS. The annotation of cluster 9 was the most challenging: while it segregated far from other ASM/MyoF cells in our embedding (**Figure 2A**) and consisted of cells derived from P21 mice (**Figure 1E**), it colocalized with cells annotated in TMS as myofibroblasts (MyoF) in the harmonized embedding (**Figure 2B**). Although TMS did not identify an airway smooth muscle (ASM) cell type, ASM but not MyoF are abundant in healthy, adult mice [25]. At postnatal time points, the expression of the marker genes *Hhip* and *Pdgfra* was mutually exclusive, and *Hhip* is an ASM marker [26]: we therefore labelled the *Hhip*+ *Pdgfra*- clusters 8 and 9 as early ASM and ASM, and the *Hhip*- *Pdgfra*+ cluster 10 as MyoF. At E18.5, a single MyoF/ASM cluster was found (cluster 7), which coexpressed both *Hhip* and *Pdgfra* as well as *Tgfbi*: it was therefore

annotated as MyoF/ASM precursors. Clusters 2, 11, and 14 expressed *Mki67* and were therefore labeled as proliferating fibroblasts, MyoF, and pericytes, respectively (**Figure 2C-D, bottom labels**). These cell subtypes corresponded only partially to categories identified in recent publications [27] (**Figure 2 - Supplementary 1**) and preprints [28] (**Figure 2 - Supplementary 2**). In particular, in those studies, similar to TSM, there was no distinction made between ASM and VSM.

A canonical function of mesenchymal cells is extracellular matrix (ECM) modulation. We therefore examined the expression of several ECM-related genes (**Figure 2D**). Fibroblasts shared ECM components *Mfap4*, *Fn1*, *Vcan*, and *Ogn*. In alveolar fibroblasts (AlvF), subtype-specific expression included *Spon1* which encodes a secreted adhesion protein. Adventitial fibroblasts (AdvF) specifically expressed *Podn*, encoding a molecule that constrains SMC proliferation and migration [29], *Mfap5*, which contributes to tissue elasticity [30], and *Dcn*, which confers resistance to tissue compression. ASM expressed *Aspn*, which inhibits canonical TGFβ and SMAD signaling [31], and late ASM expressed *Lum*, which is responsible for collagen binding. Interestingly, early ASM did not express *Lum*, suggesting a change in fibril diameter regulation around conducting airways at the end of alveolarization [32,33]. Conversely, *Aspn*, a modulator of ECM mineralization that competes with *Dcn* for collagen binding [34], was highly expressed in early ASM from P1 and P7 but downregulated at P21, suggesting higher activity in the temporal window surrounding late alveolarization. Pericytes expressed *Mcam*, *Mfge8*, *Postn*, and *Eng*, which modulates vascular tone, endothelial cell (EC) migration and angiogenesis [35]. VSM had lower expression of many ECM genes but the highest expression of elastin (*Eln*), which modulates tissue stiffness during alveolarization, with all cell types except pericytes expressing moderate levels of *Eln* as well [36,37].

Fibroblast and ASM/MyoF bipotent precursors split at birth

The lung mesenchyme changed dramatically between E18.5 and P1, a time window of just 48 hours. Both fibroblast precursors (Fib precursor) and ASM/MyoF precursors, the two most abundant mesenchymal types at E18.5, completely disappeared by P1 (**Figure 3A**). Fib precursors (cluster 1) are replaced by two distinct types of postnatal fibroblasts (clusters 3 and 5) (**Figure 3B**) which expressed *Col13a1* and *Col14a1* and resembled adult AlvF and AdvF fibroblasts, respectively. Both genes are coexpressed in the E18.5 precursors, indicating a common origin and a cell fate commitment during late gestation up to birth (**Figure 3C**). In terms of ECM secretion, fibroblast precursors appear to be more similar to AlvF (**Figure 2D**). ASM/MyoF precursors (cluster 7) split into ASM (cluster 8, *Hhip*+) and MyoF (cluster 10, *Pdgfra*+) (**Figure 3D**): only ASM/MyoF precursors coexpressed *Hhip* and *Pdgfra* (**Figure 3E**). Shared expression of *Tgfbi* and *Tagln* underscored a common lineage (**Figure 2D**). Unlike early AlvF and AdvF, which are already present in small numbers at E18.5, almost no ASM or MyoF cells were observed before birth.

We then used pathway enrichment analysis to compare prenatal precursors and postnatal subtypes [38]. Both fibroblast and ASM/MyoF bipotent precursors expressed higher levels of genes related to vascular development and muscle cell proliferation. In contrast, fibroblasts precursors specialized in mesenchyme development compared to postnatal fibroblasts, while ASM/MyoF precursors specialized in response to hormones compared to postnatal ASM and MyoF (**Figure 3F, upper, red bars**). Pathways down-regulated in precursors compared to distinct postnatal counterparts also differed between the two cell lineages (**Figure 3F, lower, blue bars**). MyoF/ASM precursors were the only cell type expressing corticotropin releasing hormone (*Crh*, **Figure 3G**), an early factor within the hypothalamic–pituitary–adrenal (HPA) axis that stimulates glucocorticoid release. No genes encoding *Crh* receptors were expressed by any cell type in the perinatal lung, including in immune [13], endothelial [20] and epithelial [39] cells (**Figure 3 - Supplementary 1**). Further, no cells in the adult lung expressed *Crh* or its receptors [22], raising the possibility that MyoF/ASM precursor-secreted *Crh* might be signaling through the HPA axis. Interestingly, *Crh* plays an essential role in lung maturation [40] without significantly impacting surfactant production by epithelial cells [41,42]. Embeddings of ASM/MyoF precursors at E18.5 colored by expression of *Tgfbi* and *Crh* indicate higher heterogeneity for the latter gene (**Figure 3H**). *Crh*-expressing MyoF/ASM precursors were localized by in-situ imaging via RNA-Scope within the lung parenchyma (**Figure 3I**). Fibroblast precursors expressed high levels of *Hsd11b1* (**Figure 3G**), the main enzyme responsible for converting cortisone into its bioactive form cortisol, indicating a cell-type-specific amplification of glucocorticoid action [43]. Several genes related to response to hypoxia were lost in both lineages, including *Hif3a*, a negative regulator of *Hif1a*, suggesting tight regulation of this pathway at the critical transition between fetal environment and air-breathing life. *Lox*, a key modulator of elastin fiber assembly, was also down-regulated after birth except in AdvF, indicating a cell function specialization in that direction (**Figure 3G**).

Mesenchymal cell proliferation increases transiently during postnatal development

Through the first seven days of life, the relative abundance of fibroblasts decreased, while ASM and particularly MyoF increased, with MyoF also branching into a distinct proliferating population representing the only novel cell type at P7 (cluster 11, see **Figure 4A, left**). The subsequent temporal transition from P7 to P21 indicated a widespread maturation of early postnatal fibroblast and ASM cell types into the adult types (**Figure 4A, right**) that harmonized with TMS (**Figure 2B**). While MyoF were the most abundant cell type at P7, no cells with this transcriptomic profile were observed at P21, while *Hhip*+ cells remained. This is consistent with MyoF playing a specific function during alveolarization [44].

To understand the gene expression changes underlying the P7 to P21 transition, we performed pathway analysis of differentially expressed genes in ASM, AlvF and AdvF (**Figure 4C**). In ASM cells, this analysis indicated a loss of cell proliferation and reduced expression in pathways related to alveolar septation (trabecula formation, molecules associated with elastic fibers). In both AlvF and AdvF, up-regulated pathways included immune functions,

suggesting an increased role in immunity for these cells in adults compared to neonates. Down-regulated pathways indicated a stabilization of the ECM.

In mice, during the first week of life the lung has transitioned from the saccular to the alveolar stage by P7, corresponding to an explosive growth in lung surface area which entails substantial cell proliferation. We therefore computed the fraction of proliferating cells within fibroblasts, MyoF, and pericytes at all three postnatal time points and observed a peak at P7 in all populations (**Figure 4C**), consistent with our previous report in endothelial cells [20]. At P21 there were no MyoF, proliferating pericytes, or proliferating fibroblasts.

We then used fluorescence in-situ RNA hybridization via RNA-Scope to define the location of ASM, MyoF, and proliferating MyoF within the P7 lung (**Figure 4D**). First, we detected *Mki67*+ *Pdgfra*+ *Tgfbi*+ proliferating MyoF in the distal lung (**left**, blue is nuclear DAPI stain). We then aimed to distinguish ASM from MyoF by probing *Tgfbi*, *Pdgfra*, and *Hhip* and detected both cell types in the distal lung parenchyma (**middle**): *Tgfbi*+ *Hhip*+ Pdgfra- cells (**dashed arrow**) and *Tgfbi*+ *Pdgfra*+ *Hhip*- cells (**solid arrow**). We also found *Tgfbi*+ *Hhip*+ cells lining conducting airways (**right**). This suggests that all three cell populations are present in the murine lung at P7: (i) *Hhip*- MyoF populating the distal lung, (ii) *Hhip+* ASM cells around proximal airways, and (iii) *Hhip+* cells that share a transcriptomic profile with ASM but, like MyoF, are found in the distal parenchyma. To better delineate the gene expression boundary between MyoF and ASM, we plotted the expression of several differentially expressed genes between these two populations on top of the embedding (**Figure 4E**) and found that the mutually exclusive expression of *Hhip* and *Pdgfra* is accompanied by additional gradients in gene expression, with one subset of ASM expressing *Actc1* (top left corner) and one MyoF subset expressing *Cd34* (bottom right corner), which has been implicated in MyoF activation [45,46].

Mural cells change gradually during perinatal development

The transcriptomes of mural cells changed more slowly and gradually between E18.5 and P21 compared to other mesenchymal cells. Relative cell type abundance of mural cells as a whole (**Figure 1D**) and of both pericytes and VSM (**Figure 5A**) increased over time. Like MyoF, pericytes were accompanied by a twin, proliferating population at P7 (**Figure 4C**). Unlike other mesenchymal cells, pericytes and VSM did not split into clear subclusters based on developmental stage. Nonetheless, the embedding of pericytes (**Figure 5B**) and VSM (**Figure 5C**) colored by time point indicates a transcriptomic shift over time (arrows are guides to the eye). To better understand this shift, we visualized the expression of ECM-related genes stratified by time points in pericytes (**Figures 5D)** and VSM (**Figure 5E)**. In pericytes, *Pdgfrb* was stable over time while *Postn* showed widespread expression with a steady increase over time (green to yellow dots). Genes involved in cell contractility (*Acta2*, *Myl9*) were

up-regulated at both E18.5 and P7. A number of genes were expressed only or mostly at P21 (*Enpp2*, *Gap43*, *Npnt*, *Prelp*, *Cxcl14*). In VSM cells, expression of *Sparc* was constant throughout, while elastin (*Eln*) and contractility genes peaked at P7. *Sparcl1*, *Mgp* and *Timp3* instead dipped at P7. Like in pericytes, several genes show similar expression at E18.5 and P7 (*Fbn1*, *Mfap4*, *Fn1*). Other ECM modulators (*Aspn*, *Mfge8* and *Postn*) peaked earlier in development. Overall, this evidence supports the notion of mural cells undergoing a process of slow maturation to gradually modulate the ECM.

Cell-cell communications decrease after birth and offer insight into tissue organization

Mesenchymal cells often provide support to neighboring cell types. To quantify the magnitude of this process, we counted the number of ligand-receptor pairs in the CellPhoneDB database [47] expressed in at least 10% of a mesenchymal cell type and a partner cell type among immune and endothelial cells (EC) and found a hierarchical structure, especially at P7 (**Figure 6A** and **Figure 6 - Supplementary 1**). The largest number of interactions were found between AdvF and macrovascular EC, especially lymphatics. Pericytes play a natural mediating role between EC and parenchyma: we quantified the number of interactions with a more stringent cutoff (20% of cells expressing each gene) and confirmed more interactions with macro- than microvascular cell subtypes at all time points (**Figure 6B**). The number of interactions with most cell types decreased from E18.5 to P7, and this conclusion was further supported when we included other pairs of cell types beyond pericytes (**Figure 6C**), indicating less intense cell-cell communications at P7. To further understand at the single gene level what kind of communications pericytes are entertaining, we computed a dot plot of several highly expressed genes involved in paracrine communications (**Figure 6D**). Expression of partner genes across EC types was heterogeneous, with the highest expression of *Pdgfb* in *Car4*+ microvascular EC, while interactions with both MyoF and ASM included the angiotensin pathway (*Agtr1a-Agt*) which regulates and responds to blood pressure, the Wnt (*Wnt5a-Fzd1*) pathways which is a central developmental signaling hub, and *Axl-Pros1*, which might be important during lung injury [48].

Paracrine signaling via secreted ligands is not the only way cells communicate with one another. We investigated the developmental expression patterns of connexins, which enable direct exchange of biochemicals across the cytoplasmic membranes of two adjacent cells (**Figure 6E**). Four connexins were expressed in the neonatal pulmonary mesenchyme. *Gja5* (aka CX40) marked arterial endothelial cells (ECs) while *Gja4* (CX37) was expressed by both arterial ECs and pericytes, which are in physical contact, as well as a few additional cell types. *Gjc1* (CX45) was expressed by multiple cell types including pericytes and MyoF. *Gja1* (CX43) was not detected in ASM/MyoF precursors at E18.5 but was expressed in ASM at P1 and later time points, as well as in ECs of lymphatic vessels, and in epithelial alveolar cells types 1 and 2. *Gja1* expression is consistent with homotypic junctions (i) within the contractile ASM rings surrounding airways; (ii) in the alveolar epithelium, mirroring previous studies in the eye lens epithelium [49]; and (iii) in lymphatic valves preventing lymph backflow [50]. Taken together,

these data show that specific connexins are expressed by distinct groups of cell subtypes within the lung and at specific time points, indicating that the direct exchange of cytoplasmic biomolecules between adjacent cells through these channels might contribute to the orchestration of tissue development.

Hyperoxia delays development, dysregulates communications, and gives rise to a new male-specific population of fibroblasts

Having characterized the distinct mesenchymal populations during perinatal development, we used a murine model of bronchopulmonary dysplasia (BPD) in which newborn mice were exposed to 80% oxygen (hyperoxia) from birth through P7, whereupon mice were sacrificed and tissue obtained for scRNA-Seq and in-situ analysis. We quantified relative abundances of mesenchymal populations and observed a marked expansion of fibroblasts including proliferating ones at the expense of MyoF and pericytes (**Figure 7A**). The distribution of cell type abundances resembled that observed in healthy P1 lungs (**Figure 7A, right**), suggesting a global developmental arrest. To test this hypothesis, we computed 1,000 bootstraps over cells of cell type abundance distributions at P1, P7, and P7 hyperoxia and asked whether the hyperoxia composition was more similar to a healthy P1 lung (right, blue) or to a healthy P7 lung (left, red). In 999 out of 1,000 randomizations, hyperoxic mice were closer to P1 than healthy P7 (**Figure 7B**), which is consistent with previous histological analyses [51,52]

We next asked how gene expression within distinct kinds of mesenchymal cells is affected by hyperoxia. First, we found that cells from hyperoxia-exposed mice segregated in the embedding (**Figure 7C**). Embedding shifts represent changes in hundreds of genes at the same time. To exemplify these changes at the level of single genes, we plotted the distributions of expression of *Btg2* and *Spry2* in AlvF (**Figure 7D**, all cell types in **Figure 7 - Supplementary 1**). *Btg2*, an inhibitor of proliferation, is highly expressed at P1 (yellow line) and down-regulated in healthy P7 mice (red), but much less affected upon hyperoxia exposure (purple dashed line, **Figure 7D, top**), indicating residual inhibition leading to reduced cell proliferation, in agreement with the observed loss of several proliferative cell subtypes such as pericytes and MyoF (**Figure 7A**). *Spry2*, a modulator of fibroblast growth factor (FGF) pathway-related mesenchymal cell differentiation [53], is even higher in hyperoxia P7 mice than in P1 and P7 healthy mice (**Figure 7D, bottom**), indicating that hyperoxia can affect gene expression beyond a simple developmental delay.

Second, hyperoxia caused changes in the expression of genes involved in tissue stiffness and ECM secretion (**Figure 7E**). We measured an increase in *Myh9*, *Tpm3*, *Itga5* across multiple cell types, and of *Itga6* specifically in the remaining pericytes, suggesting the mesenchyme becomes more reactive to mechanical forces. Genes that contribute to or modulate the ECM were also mostly upregulated, often across multiple mesenchymal subtypes such as *Mgp*, *Mmp2*, *Eng*, and *Mfge8*. Expression of some genes was increased in a cell specific manner.

For example, *Cd34* increased mostly in the remaining MyoF and to a lesser degree in both AlvF and AdvF, but did not change in other cells.

A direct consequence of the observed severe depletion in MyoF and pericytes and the changes in gene expression is the risk of an imbalanced paracrine signaling network across the tissue. Indeed, when we plotted the heatmap of interaction numbers in hyperoxic mice (**Figure 7F**) and compared it with healthy P7 mice (**Figure 6A**), we observed a more active but less organized network, suggesting widespread dysregulation of paracrine signaling within the mesenchyme and with other cell types.

Although most cells from hyperoxic mice fit into previously characterized cell types (**Figure 2A**), we identified a population of 55 cells specific to our male mouse that formed a novel cluster (**Figure 7G, red circles**, also visible in **Figure 7C**). Most of these cells clustered together next to AlvF. No such cells were found in the female hyperoxic mouse or in the normoxic mice. Therefore, we called these cells male hyperoxic fibroblasts (MHF). Like AlvF, MHF expressed *Col13a1* and *Wnt2* but not *Col14a1*, *Tgfbi*, or *Pdgfrb* (**Figure 7H**). They also expressed a number of marker genes, including contractility genes *Acta1* and *Actc1* and microtubule-related genes *Tubb3* and *Tuba1c*, surface markers *Cd9* and *Aqp3*, the enzyme *Aldoart2* and the lectin *Lgals3*, which plays a role in fibrosis [54]. In-situ imaging of lungs from healthy and hyperoxia-exposed mice at P7 confirmed the presence of cells double positive for *Tubb3* and *Wnt2*, which are markers of MHF, only in male hyperoxia-exposed mice (**Figure 7I, top left**). No double positive cells were observed in female (**top right**) or healthy mice of the same age (**bottom**). Overall, these data indicate that a new type of contractile cells related to AlvF emerge in the lung parenchyma of males only after hyperoxia exposure.

Discussion

In mammalian life, organismal survival depends upon sufficient lung development to permit establishment of an air-liquid interface, a ten-fold increase in pulmonary blood flow, and marked decrease in pulmonary arterial blood pressure as gas exchange moves suddenly from the placenta to the lungs [3]. Lung development continues after birth with a rapid increase in the number of alveoli [10]. While insight into the contributions of the cellular constituents of the lung across development is increasing [13,20], definitive information surrounding the composition, interactions and functions of lung mesenchymal cells remains elusive. In particular, there is still great ambiguity in the definition and naming of cell subtypes across the lung mesenchyme, especially during development [27,28,39]. The transcriptomic and imaging data presented in this study provide sufficient molecular texture to increase the consistency of mesenchymal cell type identification, motivating unambiguous annotations across development and in a model of neonatal lung injury.

VSM, though identified as a distinct cluster by Zepp et al. [55], have not been annotated specifically in most previous reports, including TMS [22]. Nonetheless, some TMS cells harmonize on top of our VSM (**Figure 2B**). Cohen et al. [39] and Liu et al. [28] report no VSM either, although in Liu et al.'s embedding in Figure 1, *Ebf1*+ cells are compatible with a distinct subpopulation that may have been overlooked owing to clustering parameters [21]. Guo et al. report a cell type called Pericyte-1, which had low expression of *Pdgfrb* but high expression of *Acta2*, a marker of smooth muscle cells [27]. As shown in **Figure 2 - Supplementary 1**, that population aligns with our VSM, so these authors likely isolated the cluster successfully, but did not annotate it as VSM.

In contrast to VSM, there is broad acceptance of the existence of a heterogeneous population of lung fibroblasts. However, consensus surrounding the nomenclature has not been reached. *Col13a1*+ fibroblasts have been recently labeled matrix fibroblasts by Cohen et al. [39], MatrixFB-1 by Guo et al. [27], and lipofibroblasts by Liu et al. [28]. We named them alveolar fibroblasts or AlvF, in agreement with TMS [22] and others [16] and to highlight their anatomic location. Further, there is no direct evidence that AlvF always store large lipid reserves. *Col14a1*+ fibroblasts have been labeled MatrixFB-2 [27], intermediate fibroblasts [28], or aggregated into fibroblast progenitors [39], while we call them adventitial fibroblasts to emphasize their spatial niche within the lung parenchyma, consistent with TMS. The presence of a distinct pericyte fraction is also generally acknowledged, however Liu et al. opted to label them *Ebf1*+ fibroblasts at all developmental stages, as opposed to explicit recognition of pericytes. The embedding in Figure 1 of Cohen et al. appears to imply that pericytes are transcriptionally contiguous with fibroblast precursors, which is inconsistent with our data from E18.5 and later times. Guo et al. report two pericyte subtypes, with only Pericyte-2 mapping to our pericyte transcriptomic profile. None of those studies report proliferating pericytes, although they constitute >10% of pericytes in our P7 data (**Figure 4C**).

Myofibroblasts and ASM have also been annotated inconsistently. In healthy adult mice, TMS annotated a relatively abundant *Tgfbi*+ *Hhip*+ *Pdgfra*- *Actc1*- population labeled as MyoF and a rare *Tgfbi*- *Hhip*- *Pdgfra*- *Actc1*+ population labeled as ASM [22]. However, MyoF are absent or rare in healthy adults and emerge only with injury or disease [25]. Moreover, *Hhip* is considered an ASM marker [26] and *Pdgfra* appears necessary for MyoF function [56]. Independently, Liu et al. found specific expression of *Hhip* in lipofibroblasts at P7 (their Figure 4B) [28] and Guo et al. describe smooth muscle cells that matches our early ASM (**Figure 2 - Supplementary 1**) and two clusters of MyoF, one of which (MyoFB-2) does not express specific markers (their Figure 1b) [27]. Taken together, these observations indicate that previous annotations of MyoF may merit reconsideration on the basis of our present results. First, before birth, *Pdgfra+ Hhip+* precursors are transcriptionally relatively homogeneous. Second, there is a large transcriptional shift in ASM/MyoF between P7 and P21, with the latter

time point being essentially adult-like and much less heterogeneous, as supported by our harmonization with TMS (**Figure 2B**). Third, around birth a very diverse group of ASM/MyoF cells emerges (our clusters 8 and 10): while it becomes more proliferative at P7 (**Figure 4C**), its transcriptional profile is lost by P21. These cells are characterized by a subspecialization into *Hhip+ Pdgfra-* cells, which are found both in the distal lung and circling airways, and *Hhip- Pdgfra+* cells, found exclusively in the distal lung (**Figure 4D**). A small fraction of the *Hhip+ Pdgfra-* cells also express *Actc1*, while a larger fraction of the *Hhip- Pdgfra+* cells express *Cd34*; further transcriptional gradients are observed (**Figure 4E**), perhaps indicating additional layers of specialization.

The dramatic temporal shifts in the transcriptome of fibroblasts and ASM/MyoF resemble those seen in Mac I-II-III macrophages [13] and *Car4-* capillary endothelial cells [20]. In contrast, phenotypic shifts in mural cells were gradual but progressive, without marked changes between time points, reminiscent of alterations in multiple immune subtypes including monocytes (Mac V) [13] and in *Car4+* capillaries [20,57], macrovascular cells, and lymphatic cells within the endothelium [20]. These findings indicate that select cell types perform a similar biological function throughout this developmental window, while others need to adapt transcriptionally to the distinct needs of each stage of tissue growth.

The present data identify a previously unknown role for a specific subpopulation of mesenchymal cells, as precursors of airway smooth muscle cells and myofibroblasts express corticotropin releasing hormone (*Crh*). *Crh* expression is absent in all other lung cells but high in ASM/MyoF precursors. The expression in these cells, which are positioned adjacent to the air spaces, is limited in time, present at E18.5 immediately prior to the onset of air-breathing life, and virtually lost by P1. For mammalian organisms, maturation of the surfactant system is biologically imperative [58]. *Crh* expression in precursors of airway smooth muscle cells and myofibroblasts supports the notion that these specific mesenchymal cells may sense and respond to select physiologic signals to release *Crh* and promote systemic cortisol release and maturation of alveolar type 2 cells [59] to increase surfactant production and release [41,59–61]. Moreover, we identified *Hsdb11b1*, a gene that encodes a protein that metabolizes cortisone into active cortisol [43,62], in fibroblasts, which may serve to amplify these effects locally. Overall, these observations suggest that specialized cells in the lung might be signaling to the brain, through corticotropin releasing hormone, to promote surfactant production and release at a time point when the molecule is most essential.

Correct tissue development requires orchestration by dozens of cell types via paracrine cell-cell interactions, which are influenced by ECM composition. Our finding that the number of detected interactions between mesenchymal and other cells show hierarchical structure

(**Figure 6A**) mirrors our previous findings in the endothelium [20] and suggest that AdvF and lymphatic EC might play pivotal roles as "signaling hubs" of early alveolarization at P7. The observation that the number of interactions at P7 is consistently lower than E18.5 (**Figure 6C**) is remarkable considering that P7 is a time of rapid growth and cell proliferation (**Figure 4C**). Additional analyses to determine what specific interactions are lost will be necessary to fully understand the physiological implications of this result. Algorithmic improvements beyond current practices [47] are also needed to efficiently prioritize biologically relevant interactions from the huge tables of expressed ligand-receptor pairs found in scRNA-Seq data.

The expression patterns of connexins in our data support a coordinated action between pericytes and endothelium via *Gja4* (encoding CX37). Differential expression of specific connexins enabled the distinction between VSM, which only express *Gjc1* (aka CX45) and ASM, which express both *Gjc1* and *Gja1* (aka CX43), a gene directly involved in heart [63], bladder [64], and uterine contraction [65]. CX43 plays an important role in multiple lung cell types, including endothelial cells, alveolar type 1, alveolar type 2 cells, and fibroblasts. Global deletion of *Gja1* in murine models leads to hypoplastic lungs, marked increases in the lung interstitium and neonatal death [66]. Taken together, these data indicate an untapped potential to utilize single cell transcriptomic data to investigate the direct cell-cell communication channels established by homo- and heteromultimers of connexins across tissues and development.

We employed a well-established, preclinical model of chronic lung disease of infancy to investigate whether neonatal hyperoxia possesses a specific molecular signature [12]. Approximately 75% of infants born at less than 750 grams possess chronic lung disease of infancy, initially described as bronchopulmonary dysplasia (BPD) in 1967 [11]. The transcriptomic data generated in this study was entirely consistent with this construct and indicates that throughout all mesenchymal cell types, the lung transcriptome of the P7 hyperoxic lung more closely resembled the normoxic P1 than the normoxic P7 lung (**Figure 7B**). Moreover, the observed increase in the expression of genes that help the cell adhere to an unstable ECM substrate provides indirect evidence of stiffness changes across the pulmonary ECM (**Figure 7E**), as previously observed in pulmonary fibrosis [67]. The cellular constituents of the lung demonstrated differential susceptibility to hyperoxia-induced injury. Pericytes, and not endothelial cells, were the most markedly affected by hyperoxia, with an almost complete loss of pericytes at 7 days of hyperoxia (**Figure 7A**). Pericytes seem to play a pivotal role in receiving signal from the pulmonary artery endothelial cells and delivering signals to the ASM/MyoF, cells that are critically important for postnatal secondary alveolar septation (**Figure 6D**). From a translational standpoint, the observation of differential vulnerability among the lung mesenchymal or vascular constituents has implications for emerging cell based, or exosome-driven therapeutic interventions [68].

Interestingly, in human infants, male infants generally are more frequently and severely affected than female infants [69]. The underlying reasons for the sexually dimorphic outcomes remain unknown. In these experiments, we identified a previously undescribed cluster of fibroblasts (MHF), present only in male hyperoxia exposed mice (**Figures 7G-I**). No novel cell types, including sex-specific ones, were observed in the immune [13] or endothelial [20] compartments in our previous studies. The spatial distribution of MHF, superimposed upon expression of cell contractility genes, support the proposition that the cells possess meaningful pathophysiologic significance. Positioned in the distal lung, these cells might negatively affect both lung compliance and gas exchange, both clinical hallmarks of BPD. The marked increase in *Col13a1*, *Actc1* and *Acta1*, all of which likely contribute to lung stiffness, in the male hyperoxic fibroblasts compared to the early alveolar fibroblasts further highlights the potentially pathogenic role of these cells.

Overall, our data portray the perinatal lung mesenchyme as both extremely heterogeneous at each developmental stage and temporally dynamic, with entire transcriptomic profiles appearing or disappearing within 48 hours as well as upon hyperoxia exposure. Although much remains to be discovered, the combination of single cell resolution, careful cell type annotation, in-situ experiments, and comparison with publicly available data employed here provides a blueprint to tissue and disease profiling that could be used to systematically improve our understanding of mesenchymal cells across organs, conditions, and organisms.

Methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (M. musculus)	C57BL/6J	The Jackson Laboratory	Stock No:000664	
Commercial assay or kit	Liberase TM TL Research Grade	Sigma Aldrich	Cat No. 5401020001	
Antibody cocktail	Purified Anti-Mouse CD16 / CD32 (Fc Shield) (Rat monoclonal)	Tonbo Biosciences	clone: 2.4G2 Cat No. 70-0161-U100	Pre-FACS (1:100)

Antibody	Anti-CD31 (Rat monoclonal)	BioLegend	clone: MEC13.3 Cat No. 102501	FACS (1:100)
Antibody	Anti-CD326 (EpCAM) Monoclonal Antibody, eFluor 450	eBioscience/ThermoFisher	clone: G8.8 Cat# 48-5791-82	FACS (1:100)
Antibody	Anti-CD45 (Rat monoclonal)	eBioscience/ThermoFisher	clone: 30-F11 Cat#14-0451-82	FACS (1:100)
Commercial assay or kit	RNAscope® Multiplex Fluorescent Detection Kit v2	Advanced Cell Diagnostics (ADC)	Cat No. 323110	
Commercial assay or kit	Probe-Mm-Tgfbi	ADC	Cat No. 494551	RNAscope: 50 µl per slide
Commercial assay or kit	Probe-Mm-Pdgfra-C2	ADC	Cat No. 480661-C2	RNAscope: (1:50)
Commercial assay or kit	Probe-Mm-Hhip-C3	ADC	Cat No. 448441-C3	RNAscope: (1:50)
Commercial assay or kit	Probe-Mm-Crh-C2	ADC	Cat No. 418011-C2	RNAscope: (1:50)
Commercial assay or kit	Probe-Mm-Mki67-C3	ADC	Cat No. 416771-C3	RNAscope: (1:50)
Commercial assay or kit	Probe-Mm-Wnt2	ADC	Cat No. 313601	RNAscope: 50 µl per slide
Commercial assay or kit	Probe-Mm-Tubb3-C2	ADC	Cat No. 423391-C2	RNAscope: (1:50)
Commercial assay or kit	Probe-Mm-Cspg4-C3	ADC	Cat No. 404131-C3	RNAscope: (1:50)

Mouse lung cell isolation

C57BL/6 mice were obtained from Charles River Laboratories. For studies using E18.5, P1, and P7 murine lungs, pregnant dams were purchased, and pups aged prior to lung isolation. At E18.5, the dam was asphyxiated with CO2 and pups extracted. At P1, P7, and P21 pups were euthanized with euthanasia solution (Vedco Inc). Genetic sex of mice at developmental stages E18.5 and P1 was determined by performing PCR amplification of the Y chromosome gene Sry. P7 and P21 mice were sexed through identification of a pigment spot on the scrotum of male mice (Wolterink-Donselaar et al., 2009). For all timepoints, female and male

mice were randomly selected for the studies. For all timepoints, except E18.5, the pulmonary circulation was perfused with ice cold heparin in 1x PBS until the circulation was cleared of blood. Lungs were minced and digested with Liberase (Sigma Aldrich) in RPMI for 15 (E18.5, P1, and P7) or 30 (P21) minutes at 37C, 200 rpm. Lungs were manually triturated and 5% fetal bovine serum (FBS) in 1x PBS was used to quench liberase solution. Red blood cells were lysed with 1x RBC lysis buffer (Invitrogen) as indicated by the manufacturer and total lung cells counted on Biorad cell counter (BioRad). Protocols for the murine studies adhered to American Physiological Society/US National Institutes of Health guidelines for humane use of animals for research and were prospectively approved by the Institutional Animal Care and Use Committee at Stanford (APLAC #19087).

Immunostaining and fluorescence-activated cell sorting (FACS) of single cells

Lungs were plated at 1 × 10^6 cells per well and stained with Fc block (CD16/32, 1:100, Tonbo Biosciences) for 30 min on ice. Cells were surface stained with the endothelial marker CD31 (1:100, eBioscience/ThermoFisher), epithelial marker Epcam (1:100, eBioscience/ThermoFisher), and immune marker CD45 (1:100, eBioscience/ThermoFisher) for 30 min on ice. The live/dead dye, Sytox Blue (Invitrogen), was added to cells and incubated for 3 min prior to sorting into 384-well plates (Bio-Rad Laboratories, Inc) prefilled with lysis buffer using the Sony LE-SH800 cell sorter (Sony Biotechnology Inc), a 100 µm sorting chip (Catalog number: LE-C3110) and ultra-purity mode. Single color controls were used to perform fluorescence compensation and generate sorting gates. 384-well plates containing single cells were spun down, immediately placed on dry ice and stored at −80C.

cDNA library generation using Smart-Seq2

RNA from sorted cells was reverse transcribed and amplified using the Smart-Seq2 protocol on 384-well plates as previously described [13]. Concentration of cDNA was quantified using Quant-it Picogreen (Life Technologies/Thermofisher) to ensure adequate cDNA amplification and cDNA was normalized to 0.4 ng/uL. Tagmentation and barcoding of cDNA was prepared using in-house Tn5 transposase and custom, double barcoded indices (Tabula Muris Consortium et al., 2018). Library fragment concentration and purity were quantified by Agilent bioanalyzer. Libraries were pooled and sequenced on Illumina NovaSeq 6000 with 2 × 100 base kits and at a depth of around 1 million read pairs per cell.

Data analysis and availability

Sequencing reads were mapped against the mouse genome (GRCm38) using STAR aligner (Dobin et al., 2013) and gene expression was quantified using HTSeq (Anders et al., 2015). FZ has been the main developer of HTSeq since 2016. To coordinate mapping and counting on Stanford's high-performance computing cluster, snakemake was used (Koster and Rahmann, 2012). Gene expression count tables were converted into loom objects (https://linnarssonlab.org/loompy/) and cells with less than 50,000 uniquely mapped reads or less than 400 genes per cell were discarded. Doublets were discarded by excluding small clusters and single cells that coexpress markers for incompatible cell tyes. Counts for the

remaining NNN cells were normalized to counts per million reads. For t-distributed stochastic embedding (t-SNE) (van der Maaten and Hinton, 2008), 500 features were selected that had a high Fano factor in most mice, and the restricted count matrix was log-transformed with a pseudocount of 0.1 and projected onto the top 25 principal components using scikit-learn (Pedregosa et al., 2011). Unsupervised clustering was performed using Leiden (C++/Python implementation) (Traag et al., 2019). Singlet (https://github.com/iosonofabio/singlet) and custom Python3 scripts were used: the latter are available at https://github.com/iosonofabio/lung_neonatal_mesenchymal/. Pathway analysis on the differentially expressed genes was performed via Metascape (Zhou et al., 2019) on the top 100 most differentially expressed genes for each comparison: precursors versus sample-balanced joint progenies (Fig 3), subsequent time points within each cluster (Fig 4 and 5). The most enriched pathways against a permutation test are shown ordered by significance from top to bottom (negative log of the P-value). Batch-corrected KNN [23] was used to compare our P21 data with Tabula Muris Senis [22]. Raw fastq files, count tables, and metadata are available on NCBI's Gene Expression Omnibus (GEO) website: GSE172251.

In-situ validation using RNAscope and immunofluorescence (IF)

Embryonic and post-natal mice were euthanized as described above. Female and male mice were randomly selected from the litter, and at least two litters were used to source the lung tissue for all validation studies. E18.5 lungs were immediately placed in 10% neutral buffered formalin following dissection. P1, P7, and P21 murine lungs were perfused as described above, and P7 and P21 lungs inflated with 2% low melting agarose (LMT) in 1xPBS and placed in 10% neutral buffered formalin. Following 20 hr incubation at 4C, fixed lungs were washed twice in 1xPBS and placed in 70% ethanol for paraffin-embedding. In situ validation of genes identified by single cell RNA-seq was performed using the RNAscope Multiplex Fluorescent v2 Assay kit (Advanced Cell Diagnostics) and according to the manufacturer's protocol. Formalin-fixed paraffin-embedded (FFPE) lung sections (5 μm) were used within a day of sectioning for optimal results. Nuclei were counterstained with DAPI (Life Technology Corp.) and extracellular matrix proteins stained with hydrazide [70]. Opal dyes (Akoya Biosciences) were used for signal amplification as directed by the manufacturer. Images were captured with Zeiss LSM 780 and Zeiss LSM 880 confocal microscopes, using 405 nm, 488 nm, 560 nm and 633 nm excitation lasers. For scanning tissue, each image frame was set as 1024 × 1024 and pinhole 1AiryUnit (AU). For providing Z-stack confocal images, the Z-stack panel was used to set z-boundary and optimal intervals, and images with maximum intensity were processed by merging Z-stacks images. For all both merged signals and split channels were collected.

Statistical analyses

To identify differentially expressed genes within cell populations, Kolmogorov Smirnov tests on the distributions of gene expression were performed on all genes, and either the genes with the largest absolute value of the test statistic or the genes above a certain KS statistic (e.g. 0.3) with the largest fold change were chosen.

Acknowledgements

We thank Sai Saroja Kolluru (Stanford University) for assistance with library submission to the Chan Zuckerberg Biohub, Yuan Xue (Stanford University) for assistance with the initial single cell RNA-seq data acquisition, Astrid Gillich for technical support with the RNAscope experiments, and Maya Kumar for providing hydrazide. We also thank the Stanford Shared FACS Facility, Lisa Nichols, Meredith Weglarz, and Tim Knaak for assistance with the flow cytometry instrumentation and antibody panel design. Flow cytometry data was collected on an instrument in the Stanford Shared FACS Facility obtained using NIH S10 Shared Instrument Grant (S10RR027431-01). This work was supported by National Institutes of Health grants HL122918 (CMA), HD092316 (CMA, DNC), the Stanford Maternal Child Health Institute Tashia and John Morgridge Faculty Scholar Award (CMA), the Stanford Center of Excellence in Pulmonary Biology (DNC), Bill and Melinda Gates Foundation (SRQ), and the Chan Zuckerberg Biohub (DNC. and SRQ). MAS is supported by the NSF-GRFP.

References

1. Rudolph A M. Distribution and regulation of blood flow in the fetal and neonatal lamb. Circ Res. 1985;57: 811–821.

2. Cassin S, Dawes GS, Mott JC, Ross BB, Strang LB. THE VASCULAR RESISTANCE OF THE FOETAL AND NEWLY VENTILATED LUNG OF THE LAMB. J Physiol. 1964;171: 61–79.

3. Dawes GS, Mott JC, Widdicombe JG, Wyatt DG. Changes in the lungs of the new-born lamb. J Physiol. 1953;121: 141–162.

4. Schittny JC. Development of the lung. Cell Tissue Res. 2017;367: 427–444.

5. Branchfield K, Li R, Lungova V, Verheyden JM, McCulley D, Sun X. A three-dimensional study of alveologenesis in mouse lung. Dev Biol. 2016;409: 429–441.

6. Peng T, Tian Y, Boogerd CJ, Lu MM, Kadzik RS, Stewart KM, et al. Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature. 2013;500: 589–592.

7. Metzger RJ, Klein OD, Martin GR, Krasnow MA. The branching programme of mouse lung development. Nature. 2008;453: 745–750.

8. Potente M, Mäkinen T. Vascular heterogeneity and specialization in development and disease. Nat Rev Mol Cell Biol. 2017;18: 477–494.

9. Li R, Li X, Hagood J, Zhu M-S, Sun X. Myofibroblast contraction is essential for generating and regenerating the gas-exchange surface. J Clin Invest. 2020;130: 2859–2871.

10. Herring MJ, Putney LF, Wyatt G, Finkbeiner WE, Hyde DM. Growth of alveoli during postnatal development in humans based on stereological estimation. Am J Physiol Lung Cell Mol Physiol. 2014;307: L338–44.

11. Thébaud B, Goss KN, Laughon M, Whitsett JA, Abman SH, Steinhorn RH, et al. Bronchopulmonary

dysplasia. Nat Rev Dis Primers. 2019;5: 78.

12. Hilgendorff A, Reiss I, Ehrhardt H, Eickelberg O, Alvira CM. Chronic Lung Disease in the Preterm Infant. Lessons Learned from Animal Models. Am J Respir Cell Mol Biol. 2014;50: 233–245.

13. Domingo-Gonzalez R, Zanini F, Che X, Liu M, Jones RC, Swift MA, et al. Diverse homeostatic and immunomodulatory roles of immune cells in the developing mouse lung at single cell resolution. Elife. 2020;9. doi:10.7554/eLife.56890

14. Hoogduijn MJ, Lombardo E. Mesenchymal Stromal Cells Anno 2019: Dawn of the Therapeutic Era? Concise Review. Stem Cells Transl Med. 2019;8: 1126–1134.

15. Maaten L van der, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res. 2008;9: 2579–2605.

16. Tsukui T, Sun K-H, Wetter JB, Wilson-Kanamori JR, Hazelwood LA, Henderson NC, et al. Collagen-producing lung cell atlas identifies multiple subsets with distinct localization and relevance to fibrosis. Nat Commun. 2020;11: 1920.

17. Xie T, Wang Y, Deng N, Huang G, Taghavifar F, Geng Y, et al. Single-Cell Deconvolution of Fibroblast Heterogeneity in Mouse Pulmonary Fibrosis. Cell Rep. 2018;22: 3625–3640.

18. Desmoulière A, Geinoz A, Gabbiani F, Gabbiani G. Transforming growth factor-beta 1 induces alpha-smooth muscle actin expression in granulation tissue myofibroblasts and in quiescent and growing cultured fibroblasts. J Cell Biol. 1993;122: 103–111.

19. Winkler EA, Bell RD, Zlokovic BV. Pericyte-specific expression of PDGF beta receptor in mouse models with normal and deficient PDGF beta receptor signaling. Mol Neurodegener. 2010;5: 32.

20. Zanini F, Che X, Knutsen C, Liu M, Suresh N, Domingo-Gonzalez R, et al. Phenotypic diversity and sensitivity to injury of the pulmonary endothelium during a period of rapid postnatal growth. bioRxiv. 2021. p. 2021.04.27.441649. doi:10.1101/2021.04.27.441649

21. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9: 5233.

22. Tabula Muris Consortium. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature. 2020;583: 590–595.

23. Polański K, Young MD, Miao Z, Meyer KB, Teichmann SA, Park J-E. BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics. 2020;36: 964–965.

24. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv [stat.ML]. 2018. Available: http://arxiv.org/abs/1802.03426

25. Hinz B, Phan SH, Thannickal VJ, Galli A, Bochaton-Piallat M-L, Gabbiani G. The myofibroblast: one function, multiple origins. Am J Pathol. 2007;170: 1807–1816.

26. Danopoulos S, Bhattacharya S, Mariani TJ, Al Alam D. Transcriptional characterisation of human lung cells identifies novel mesenchymal lineage markers. Eur Respir J. 2020;55. doi:10.1183/13993003.00746-2019

27. Guo M, Du Y, Gokey JJ, Ray S, Bell SM, Adam M, et al. Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. Nat Commun. 2019;10: 37.

28. Liu X, Rowan SC, Liang J, Yao C, Huang G, Deng N, et al. Definition and Signatures of Lung Fibroblast Populations in Development and Fibrosis in Mice and Men. bioRxiv. 2020. p. 2020.07.15.203141. doi:10.1101/2020.07.15.203141

29. Ross MD, Bruggeman LA, Hanss B, Sunamoto M, Marras D, Klotman ME, et al. Podocan, a novel small leucine-rich repeat protein expressed in the sclerotic glomerular lesion of experimental HIV-associated nephropathy. J Biol Chem. 2003;278: 33248–33255.

30. Combs MD, Knutsen RH, Broekelmann TJ, Toennies HM, Brett TJ, Miller CA, et al. Microfibril-associated glycoprotein 2 (MAGP2) loss of function has pleiotropic effects in vivo. J Biol Chem. 2013;288: 28869–28880.

31. Tomoeda M, Yamada S, Shirai H, Ozawa Y, Yanagita M, Murakami S. PLAP-1/asporin inhibits activation of BMP receptor via its leucine-rich repeat motif. Biochem Biophys Res Commun. 2008;371: 191–196.

32. Rada JA, Cornuet PK, Hassell JR. Regulation of corneal collagen fibrillogenesis in vitro by corneal proteoglycan (lumican and decorin) core proteins. Exp Eye Res. 1993;56: 635–648.

33. Lucariello A, Trabucco E, Boccia O, Perna A, Sellitto C, Castaldi MA, et al. Small leucine rich proteoglycans are differently distributed in normal and pathological endometrium. In Vivo. 2015;29: 217–222.

34. Kalamajski S, Aspberg A, Lindblom K, Heinegård D, Oldberg A. Asporin competes with decorin for collagen binding, binds calcium and promotes osteoblast collagen mineralization. Biochem J. 2009;423: 53–59.

35. Sugden WW, Meissner R, Aegerter-Wilmsen T, Tsaryk R, Leonard EV, Bussmann J, et al. Endoglin controls blood vessel diameter through endothelial cell shape changes in response to haemodynamic cues. Nat Cell Biol. 2017;19: 653–665.

36. Luo Y, Li N, Chen H, Fernandez GE, Warburton D, Moats R, et al. Spatial and temporal changes in extracellular elastin and laminin distribution during lung alveolar development. Sci Rep. 2018;8: 8334.

37. Wendel DP, Taylor DG, Albertine KH, Keating MT, Li DY. Impaired distal airway development in mice lacking elastin. Am J Respir Cell Mol Biol. 2000;23: 320–326.

38. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10: 1523.

39. Cohen M, Giladi A, Gorki A-D, Solodkin DG, Zada M, Hladik A, et al. Lung Single-Cell Signaling Interaction Map Reveals Basophil Role in Macrophage Imprinting. Cell. 2018;175: 1031–1044.e18.

40. Muglia L, Jacobson L, Dikkes P, Majzoub JA. Corticotropin-releasing hormone deficiency reveals major fetal but not adult glucocorticoid need. Nature. 1995;373: 427–432.

41. Emanuel RL, Torday JS, Asokananthan N, Sunday ME. Direct effects of corticotropin-releasing hormone and thyrotropin-releasing hormone on fetal lung explants. Peptides. 2000;21: 1819–1829.

42. Jobe AH, Ikegami M. Surfactant homeostasis in corticotropin-releasing hormone deficiency in

adult mice. Am J Respir Crit Care Med. 1998;158: 995–997.

43. Seckl JR, Walker BR. Minireview: 11beta-hydroxysteroid dehydrogenase type 1- a tissue-specific amplifier of glucocorticoid action. Endocrinology. 2001;142: 1371–1376.

44. Zepp JA, Zacharias WJ, Frank DB, Cavanaugh CA, Zhou S, Morley MP, et al. Distinct Mesenchymal Lineages and Niches Promote Epithelial Self-Renewal and Myofibrogenesis in the Lung. Cell. 2017;170: 1134–1148.e10.

45. Chauhan H, Abraham A, Phillips JRA, Pringle JH, Walker RA, Jones JL. There is more than one kind of myofibroblast: analysis of CD34 expression in benign, in situ, and invasive breast lesions. J Clin Pathol. 2003;56: 271–276.

46. Matsuo T, Tashiro H, Sumiyoshi R, Saito S, Shirasaki R, Shirafuji N. Functional expression cloning of molecules inducing CD34 expression in bone marrow-derived stromal myofibroblasts. Biochem Biophys Res Commun. 2020;533: 1283–1289.

47. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020;15: 1484–1506.

48. Ng C-T, Dheen ST, Yip W-CG, Ong C-N, Bay B-H, Lanry Yung L-Y. The induction of epigenetic regulation of PROS1 gene in lung fibroblasts by gold nanoparticles and implications for potential lung injury. Biomaterials. 2011;32: 7609–7615.

49. Gao Y, Spray DC. Structural changes in lenses of mice lacking the gap junction protein connexin43. Invest Ophthalmol Vis Sci. 1998;39: 1198–1209.

50. Munger SJ, Davis MJ, Simon AM. Defective lymphatic valve development and chylothorax in mice with a lymphatic-specific deletion of Connexin43. Dev Biol. 2017;421: 204–218.

51. Mižíková I, Ruiz-Camp J, Steenbock H, Madurga A, Vadász I, Herold S, et al. Collagen and elastin cross-linking is altered during aberrant late lung development associated with hyperoxia. Am J Physiol Lung Cell Mol Physiol. 2015;308: L1145–58.

52. Dieperink HI, Blackwell TS, Prince LS. Hyperoxia and apoptosis in developing mouse lung mesenchyme. Pediatr Res. 2006;59: 185–190.

53. Kuracha MR, Siefker E, Licht JD, Govindarajan V. Spry1 and Spry2 are necessary for eyelid closure. Dev Biol. 2013;383: 227–238.

54. Li L-C, Li J, Gao J. Functions of galectin-3 and its role in fibrotic diseases. J Pharmacol Exp Ther. 2014;351: 336–343.

55. Zepp JA, Morley MP, Loebel C, Kremp MM, Chaudhry FN, Basil MC, et al. Genomic, epigenomic, and biophysical cues controlling the emergence of the lung alveolus. Science. 2021;371. doi:10.1126/science.abc3172

56. Gouveia L, Betsholtz C, Andrae J. PDGF-A signaling is required for secondary alveolar septation and controls epithelial proliferation in the developing lung. Development. 2018;145. doi:10.1242/dev.161976

57. Gillich A, Zhang F, Farmer CG, Travaglini KJ, Tan SY, Gu M, et al. Capillary cell-type specialization in

the alveolus. Nature. 2020;586: 785–789.

58. Avery ME, Mead J. Surface properties in relation to atelectasis and hyaline membrane disease. AMA J Dis Child. 1959;97: 517–523.

59. Muglia LJ, Bae DS, Brown TT, Vogt SK, Alvarez JG, Sunday ME, et al. Proliferation and differentiation defects during lung development in corticotropin-releasing hormone-deficient mice. Am J Respir Cell Mol Biol. 1999;20: 181–188.

60. DeLemos RA, Shermeta DW, Knelson JH, Kotas R, Avery ME. Acceleration of appearance of pulmonary surfactant in the fetal lamb by administration of corticosteroids. Am Rev Respir Dis. 1970;102: 459–461.

61. Platzker AC, Kitterman JA, Mescher EJ, Clements JA, Tooley WH. Surfactant in the lung and tracheal fluid of the fetal lamb and acceleration of its appearance by dexamethasone. Pediatrics. 1975;56: 554–561.

62. Odermatt A, Arnold P, Stauffer A, Frey BM, Frey FJ. The N-terminal anchor sequences of 11beta-hydroxysteroid dehydrogenases determine their orientation in the endoplasmic reticulum membrane. J Biol Chem. 1999;274: 28762–28770.

63. Ya J, Erdtsieck-Ernste EB, de Boer PA, van Kempen MJ, Jongsma H, Gros D, et al. Heart defects in connexin43-deficient mice. Circ Res. 1998;82: 360–366.

64. Neuhaus J, Weimann A, Stolzenburg J-U, Wolburg H, Horn L-C, Dorschner W. Smooth muscle cells from human urinary bladder express connexin 43 in vivo and in vitro. World J Urol. 2002;20: 250–254.

65. Geimonen E, Jiang W, Ali M, Fishman GI, Garfield RE, Andersen J. Activation of protein kinase C in human uterine smooth muscle induces connexin-43 gene transcription through an AP-1 site in the promoter sequence. J Biol Chem. 1996;271: 23667–23674.

66. Nagata K, Masumoto K, Esumi G, Teshiba R, Yoshizaki K, Fukumoto S, et al. Connexin43 plays an important role in lung development. J Pediatr Surg. 2009;44: 2296–2301.

67. Jaffar J, Yang S-H, Kim SY, Kim H-W, Faiz A, Chrzanowski W, et al. Greater cellular stiffness in fibroblasts from patients with idiopathic pulmonary fibrosis. Am J Physiol Lung Cell Mol Physiol. 2018;315: L59–L65.

68. Willis GR, Fernandez-Gonzalez A, Anastas J, Vitali SH, Liu X, Ericsson M, et al. Mesenchymal Stromal Cell Exosomes Ameliorate Experimental Bronchopulmonary Dysplasia and Restore Lung Function through Macrophage Immunomodulation. Am J Respir Crit Care Med. 2018;197: 104–116.

69. Keller RL, Feng R, DeMauro SB, Ferkol T, Hardie W, Rogers EE, et al. Bronchopulmonary Dysplasia and Perinatal Characteristics Predict 1-Year Respiratory Outcomes in Newborns Born at Extremely Low Gestational Age: A Prospective Cohort Study. J Pediatr. 2017;187: 89–97.e3.

70. Clifford Philip S., Ella Srikanth R., Stupica Aaron J., Nourian Zahra, Li Min, Martinez-Lemus Luis A., et al. Spatial Distribution and Mechanical Function of Elastin in Resistance Arteries. Arterioscler Thromb Vasc Biol. 2011;31: 2889–2896.

Figures

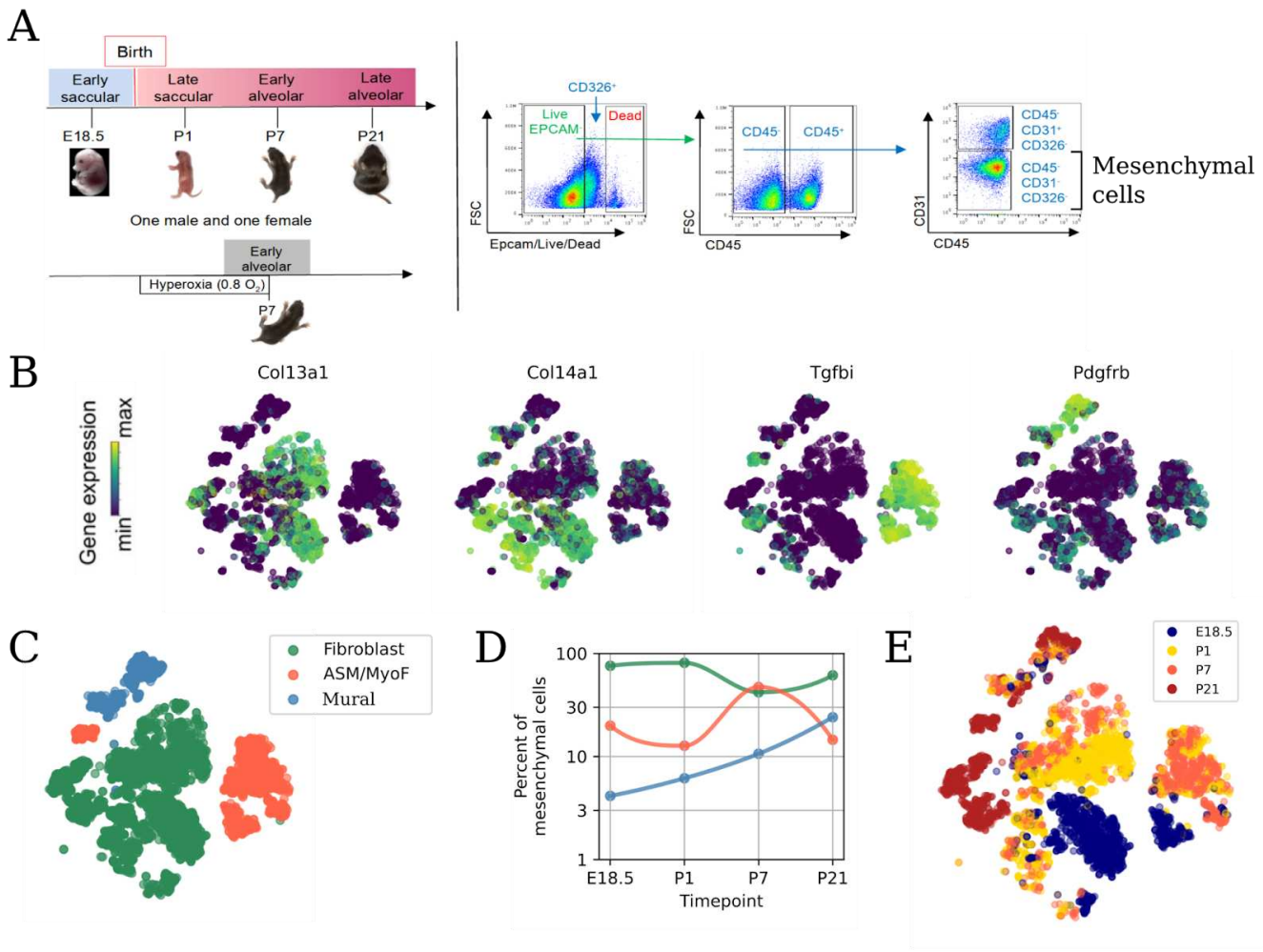

Figure 1. Three coarse cell types encompass the perinatal lung mesenchyme.

(A) Experimental design including time points sampled, tissue dissociation, and cell sorting to isolate mesenchymal cells. (B) Embedding of lung mesenchymal cells, colored by relative expression of four distinguishing marker genes. (C) As in B, but colored by coarse cell type. (D) Relative abundance for each coarse cell type during normal development, with ASM/MyoFs peaking at P7 and mural cells gradually increasing over time. (E) As in B, but colored by time point of each cell. ASM: airway smooth muscle. MyoF: myofibroblast.

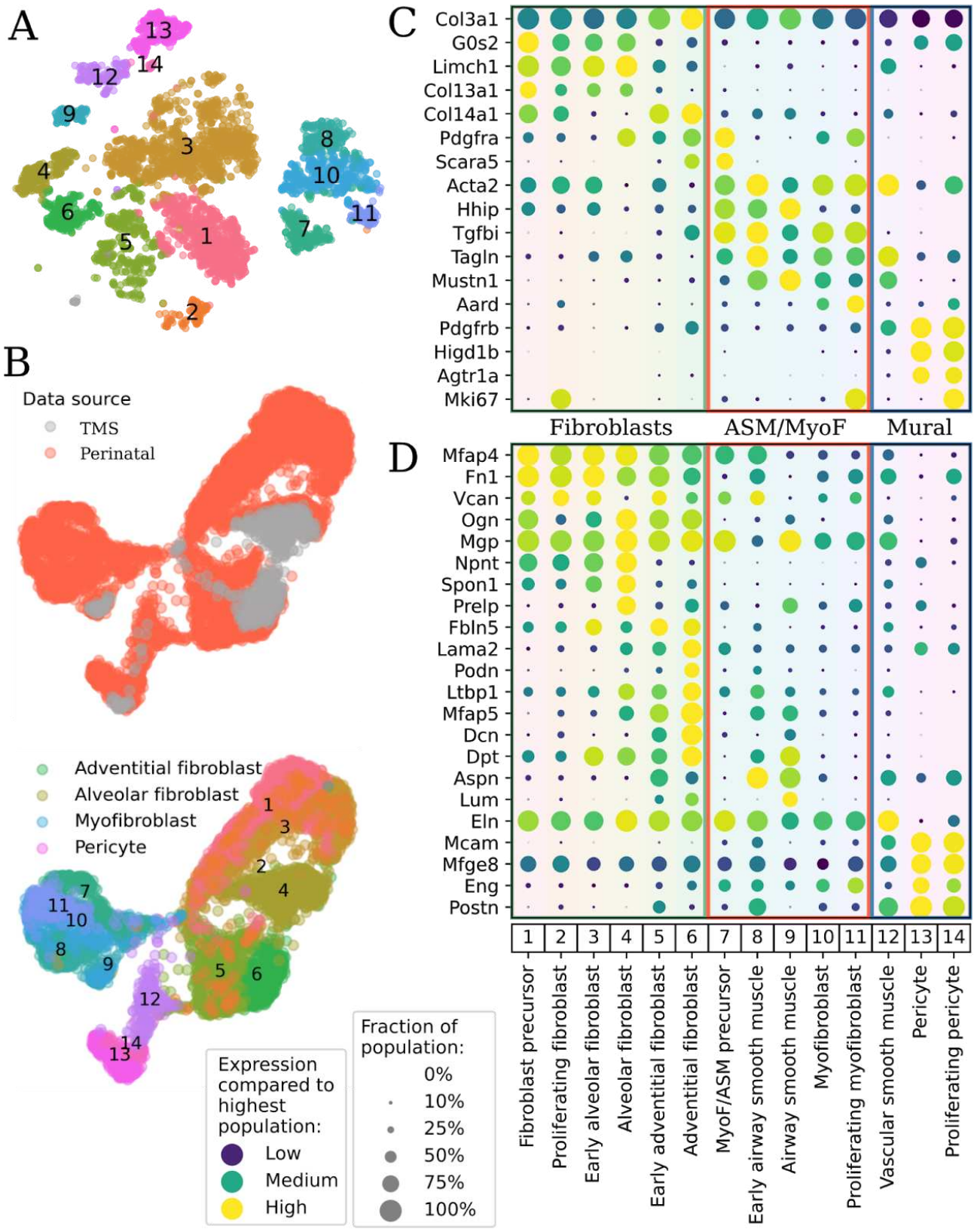

Figure 2. Fine composition of the perinatal mesenchyme and the extracellular matrix.

(A) Embedding as in Figure 1B, but colored by fine cell type and annotated with cluster numbers 1-14. This color code is used in subsequent figures. (B) Harmonized embedding of perinatal mesenchymal cells with mesenchymal cells from TMS colored by source (top panel, grey TMS cells drawn on top) and cell type (bottom, merging concordant cell types between our data and TMS). (C-D) Dotplots of all mesenchymal cell types in this study for (C) marker genes and (D) components of the extracellular matrix. TMS: Tabula Muris Senis.

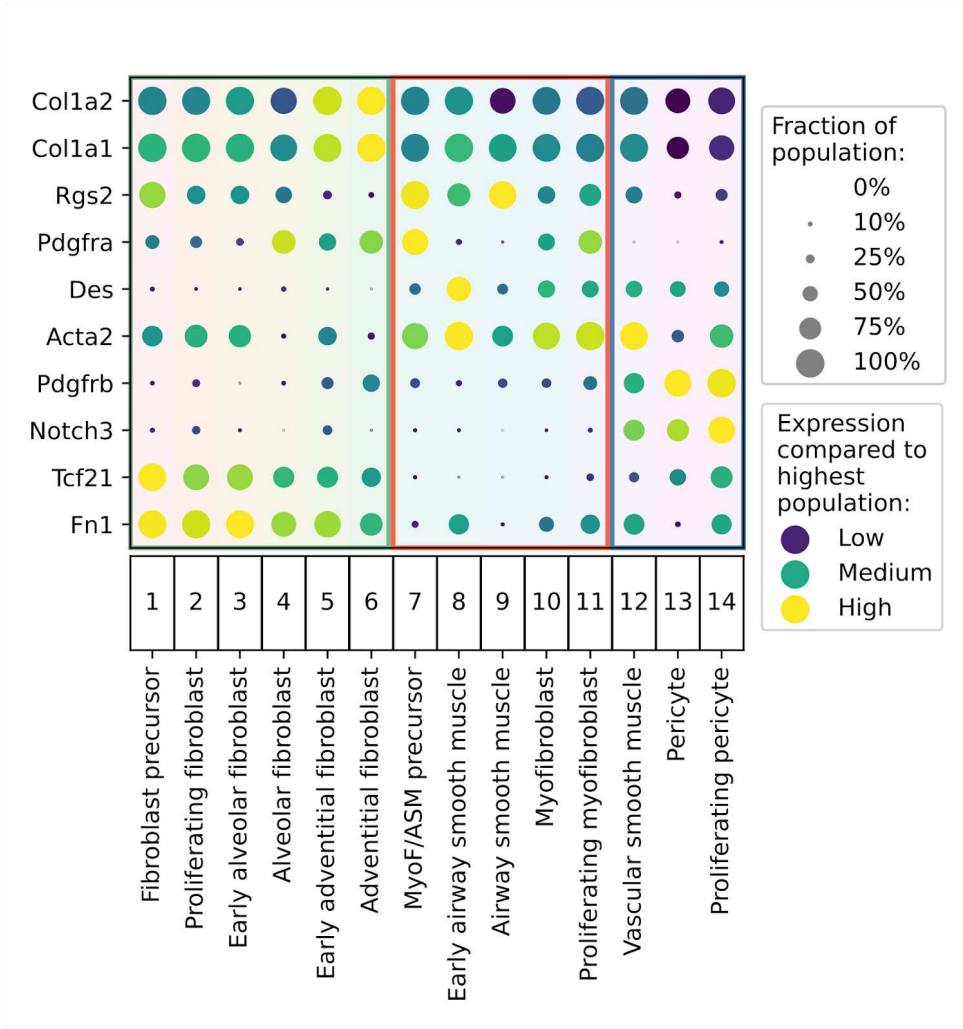

Figure 2 - Supplementary 1. Dot plot of the marker genes by Guo et al. across our subtypes. These genes are the same as in ref [27], Figure 1b, limited to the markers for their mesenchymal populations.

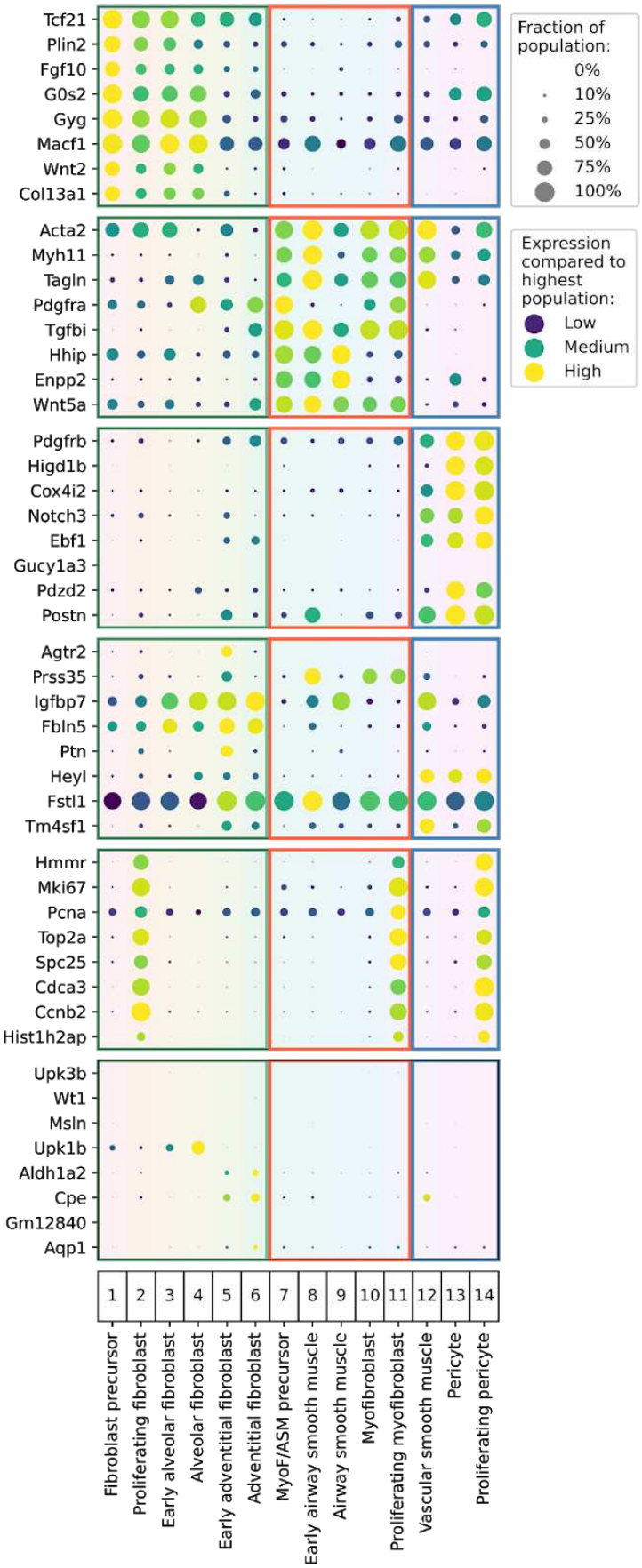

Figure 2 - Supplementary 2. Dot plot of the marker genes by Liu et al. across our subtypes. These genes are the same as in ref [28], Figure 1E.

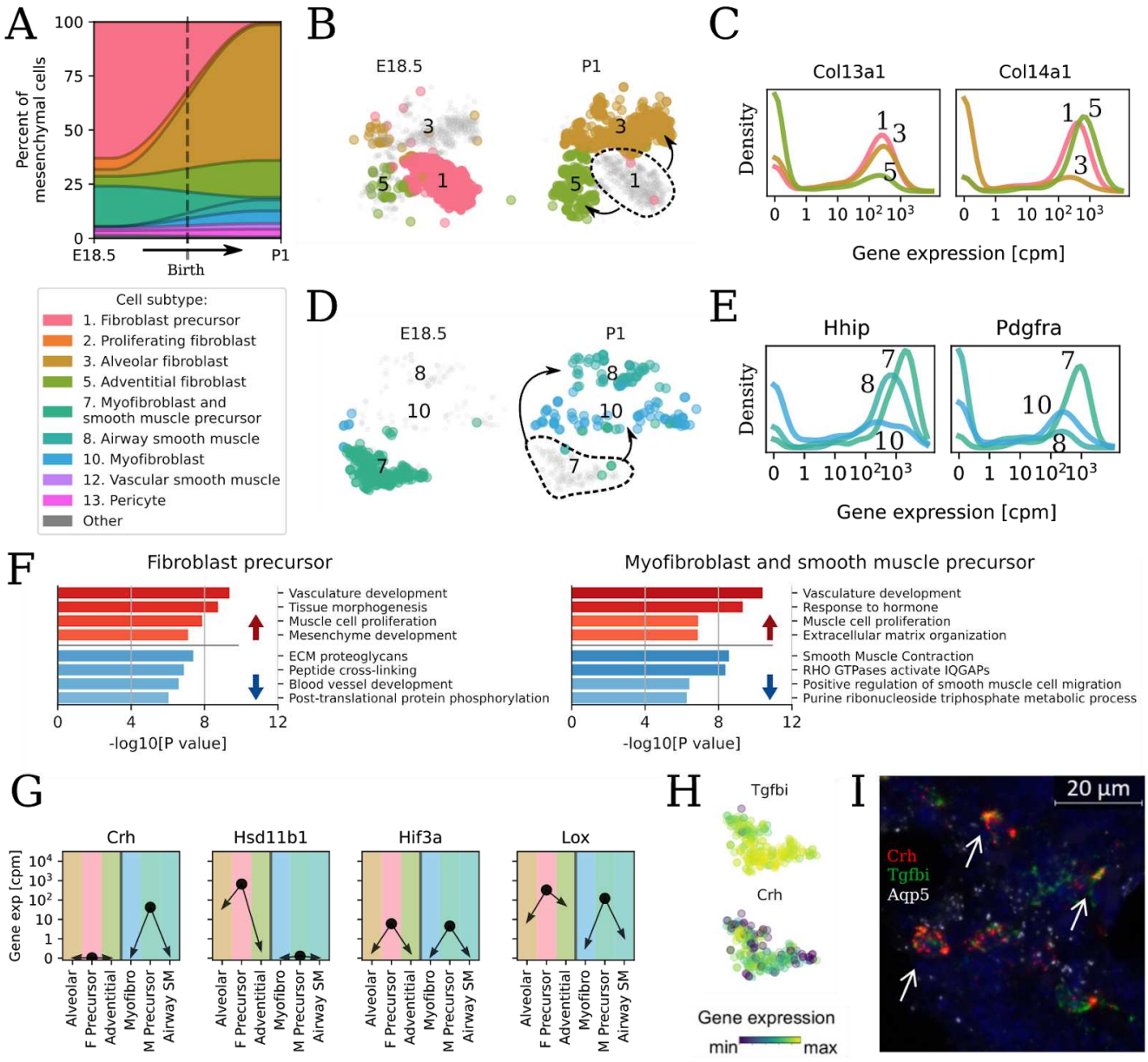

Figure 3. Bipotent precursors and distinct postnatal subtypes in fibroblasts and ASM/MyoF.

(A) Changes in cell type composition in the lung mesenchyme between E18.5 and P1. Colors and cluster numbering as in **Figure 2A**. (B) Zoom-in of the embedding on fibroblast precursors (cluster 1), early alveolar (3), and early adventitial fibroblasts (5) at E18.5 (left) and P1 (right). (C) Gaussian kernel density estimate of expression for genes shared between fibroblast precursors and either one of the progenies. (D) As in B, but focused on ASM/MyoF precursors (7), early airway smooth muscle (8), and myofibroblasts (10). (E) As in C but for ASM/MyoF precursors versus their progenies. (F) Negative log-P values of gene set enrichment tests for

each precursor against a balanced sample of their two postnatal subtypes [38], limited to the top 4 up-regulated (red, top) and 4 down-regulated (blue, bottom) pathways each. (G) Geometric averaged expression of biologically relevant genes in the bipotent precursors versus their progenies. (H) Zoom-in of the embedding on ASM/MyoF precursors showing the expression of *Tgfbi* and *Crh*. (I) In-situ imaging of *Crh*+ *Tgfbi*+ cells (arrows) in E18.5 lungs.

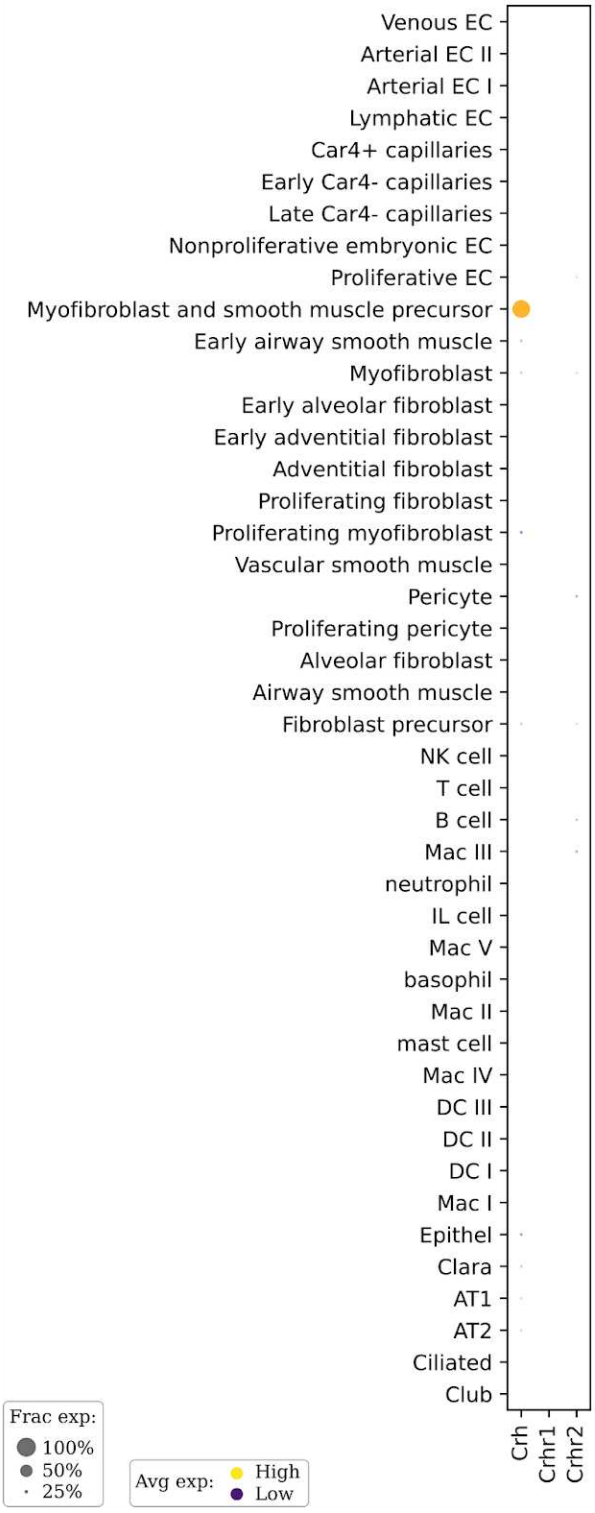

Figure 3 - Supplementary 1. Dot plot of Crh and its receptors in developing murine lungs. Mesenchymal cell data from this article, endothelial cells from [20], immune cells from [13], epithelial from [39]. Neither of the receptors is expressed in the developing lung in any detected cell type.

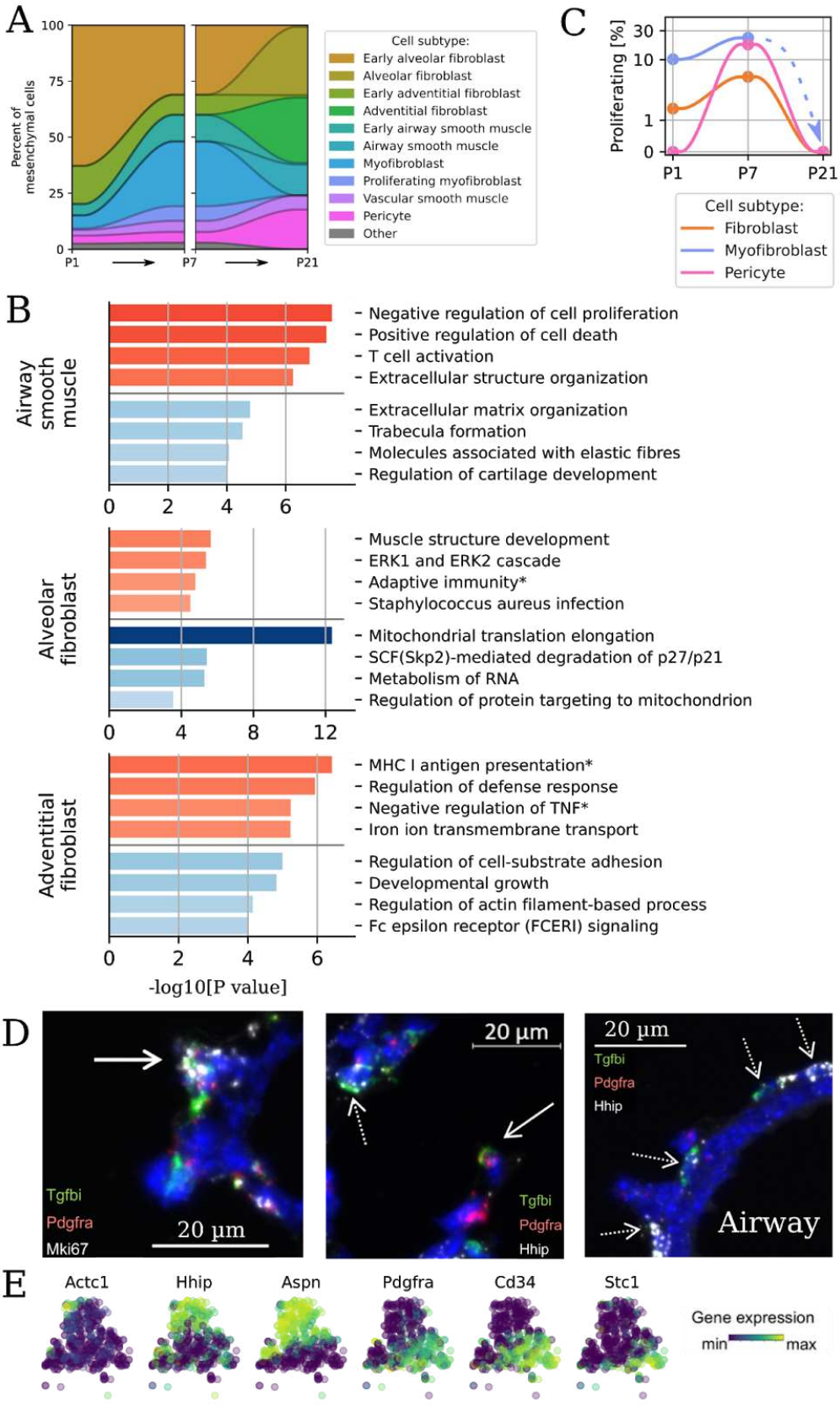

Figure 4. Postnatal changes during alveolarization in fibroblasts and MyoF/ASM.

(A) Cell type composition changes between P1, P7, and P21. Colors as in Figure 2A. (B) Negative log-P values of gene set enrichment analysis for pathways up-regulated (red, top) and down-regulated (blue, bottom) in adult versus early postnatal Alv F, Adv F, and airway SM. (C) Fraction of proliferative fibroblasts, myofibroblasts, and pericytes over total number of cells of that type over time. MyoF (proliferating or not) were not detected at P21. (D) In-situ imaging of proliferating myofibroblasts at P7 (left), of distal *Tgfbi+ Hhip+* (dashed arrow) and *Tgfbi+ Pdgfra+* cells (solid arrow), indicating two distinct ASM/MyoF subpopulations in agreement with the scRNA-Seq (center), and of *Tgfbi+ Hhip+ Pdgfra-* cells around conducting airways, consistent with ASM function (right). (E) Zoom-in of the embedding on early postnatal ASM and MyoF cells, colored by gene expression.

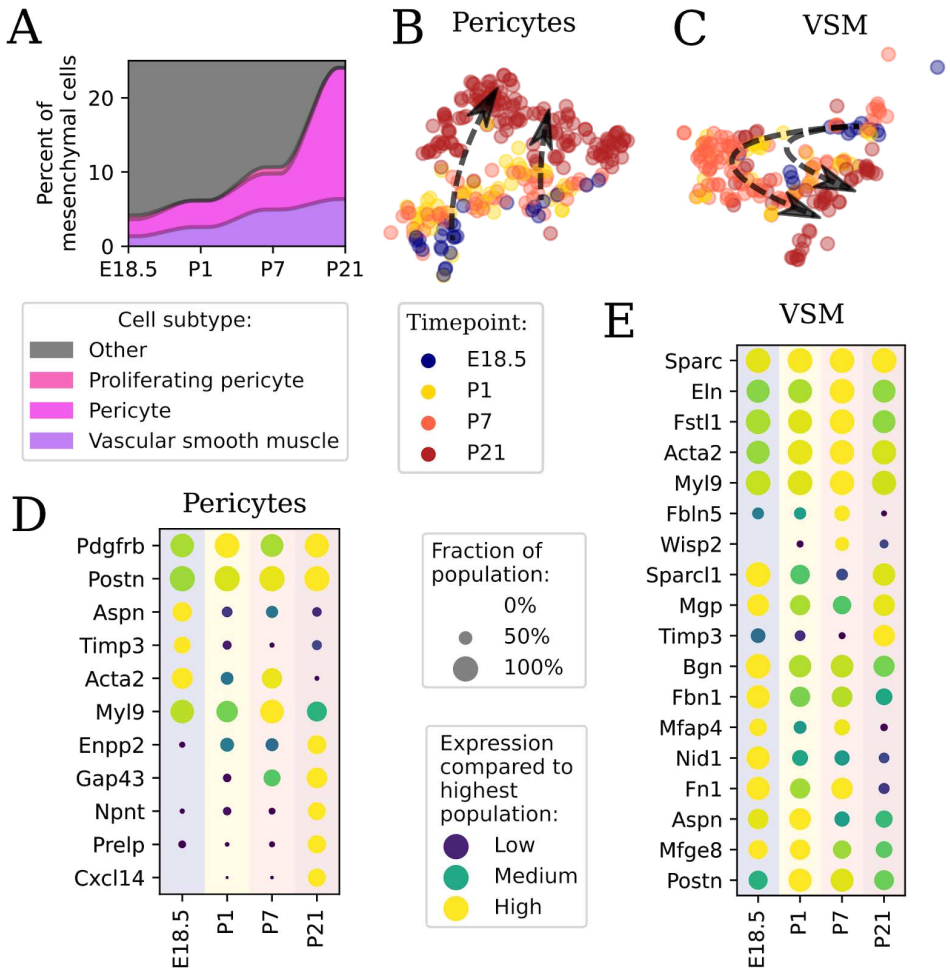

Figure 5. Mural cells slowly and gradually shift phenotypes over time.

(A) Mural cell type composition changes between E18.5 and P21. Colors as in Figure 2A. (B-C) Zoom-in of the embedding on pericytes (B) and vascular SM (C), colored by time point. Arrows are guides to the eye, indicating temporal shifts. (D-E) Dot plots divided by time point for pericytes (D) and vascular smooth muscle (E) VSM: vascular smooth muscle.

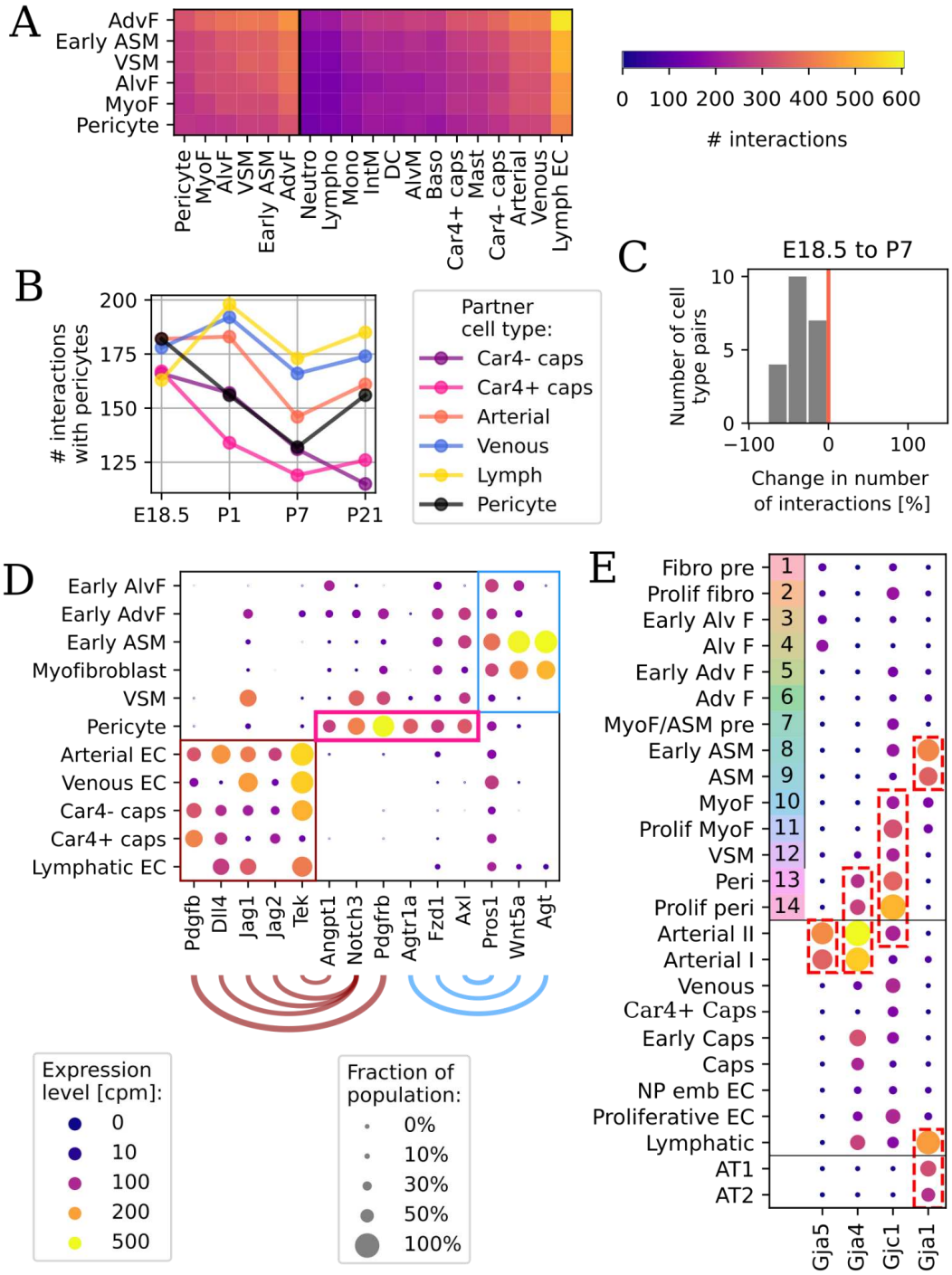

Figure 6. Cellular communications within and outside the perinatal lung mesenchyme.

(A) Number of potential ligand-receptor interactions between mesenchymal and other cell types at P7 with a threshold of 10% expressing cells. (B) Number of interactions between pericytes and endothelial cells at each time point, with a threshold of 20% expressing cells. (C) Histogram of percentage changes in number of interactions across many pairs of lung cell types from E18.5 to P7. (D) Dot plot of specific genes involved in pericyte interactions with the

endothelium (red arcs connecting ligand-receptor pairs) and the rest of the mesenchyme (blue arcs). (E) Dot plot of Gap junction genes across many cell types in the neonatal lung. Immune cells do not express high levels of Gap junction genes and are excluded. Mesenchymal cell types are numbered as in **Figure 2A**. Other cell type short names: EC = endothelial cell, NP emb EC = nonproliferative embryonic EC, AT1/2 = epithelial alveolar type 1/2. Data sources: endothelial cells [20], epithelial cells [39], and immune cells [13].

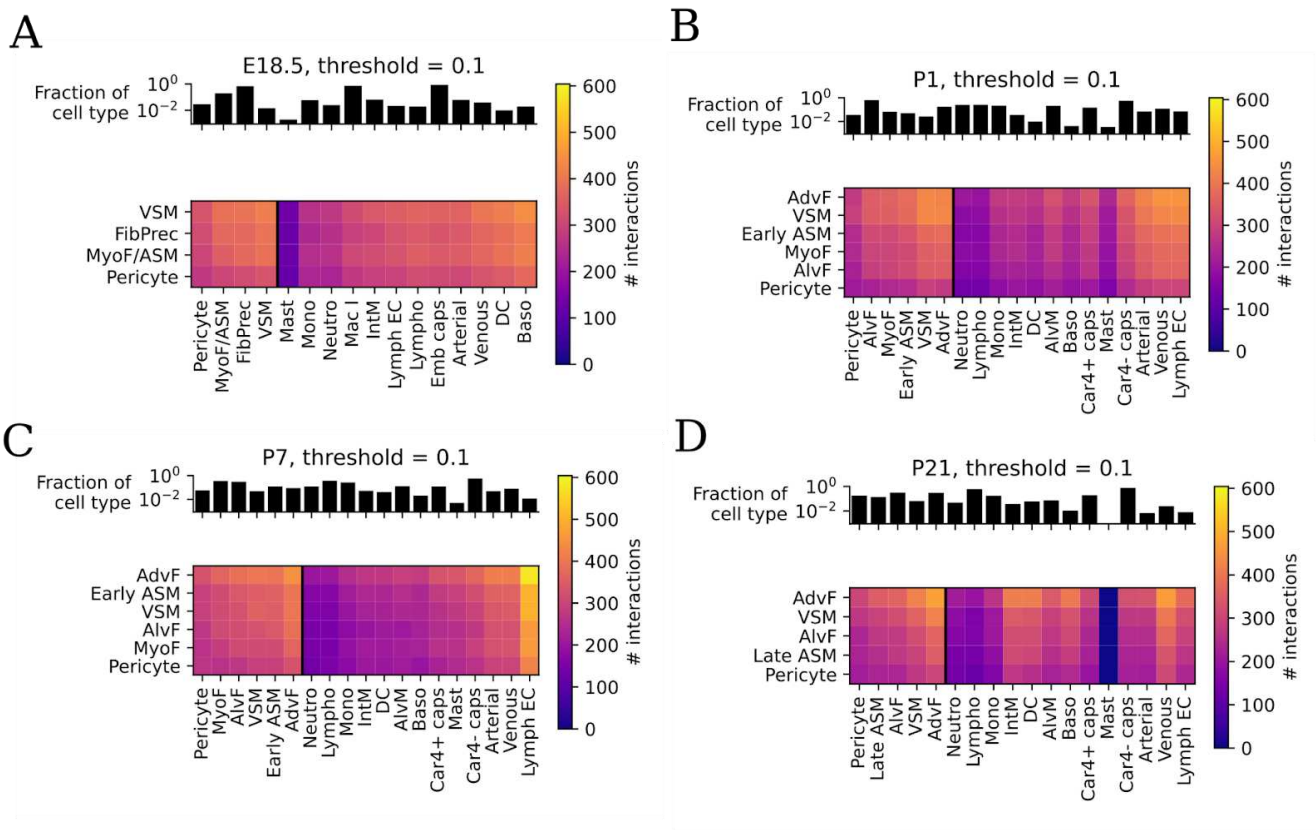

Figure 6 - Supplementary 1. Heatmaps with the number of interactions with the mesenchyme at all four time points. For context, on top of each heatmap is plotted the relative abundance of each subtype (e.g. AlvF) over the total of the cell type (e.g. mesenchymal), with a log-scaled y axis. The cell types are: endothelial, immune, and mesenchymal. Endothelial data from [20], immune data from [13].

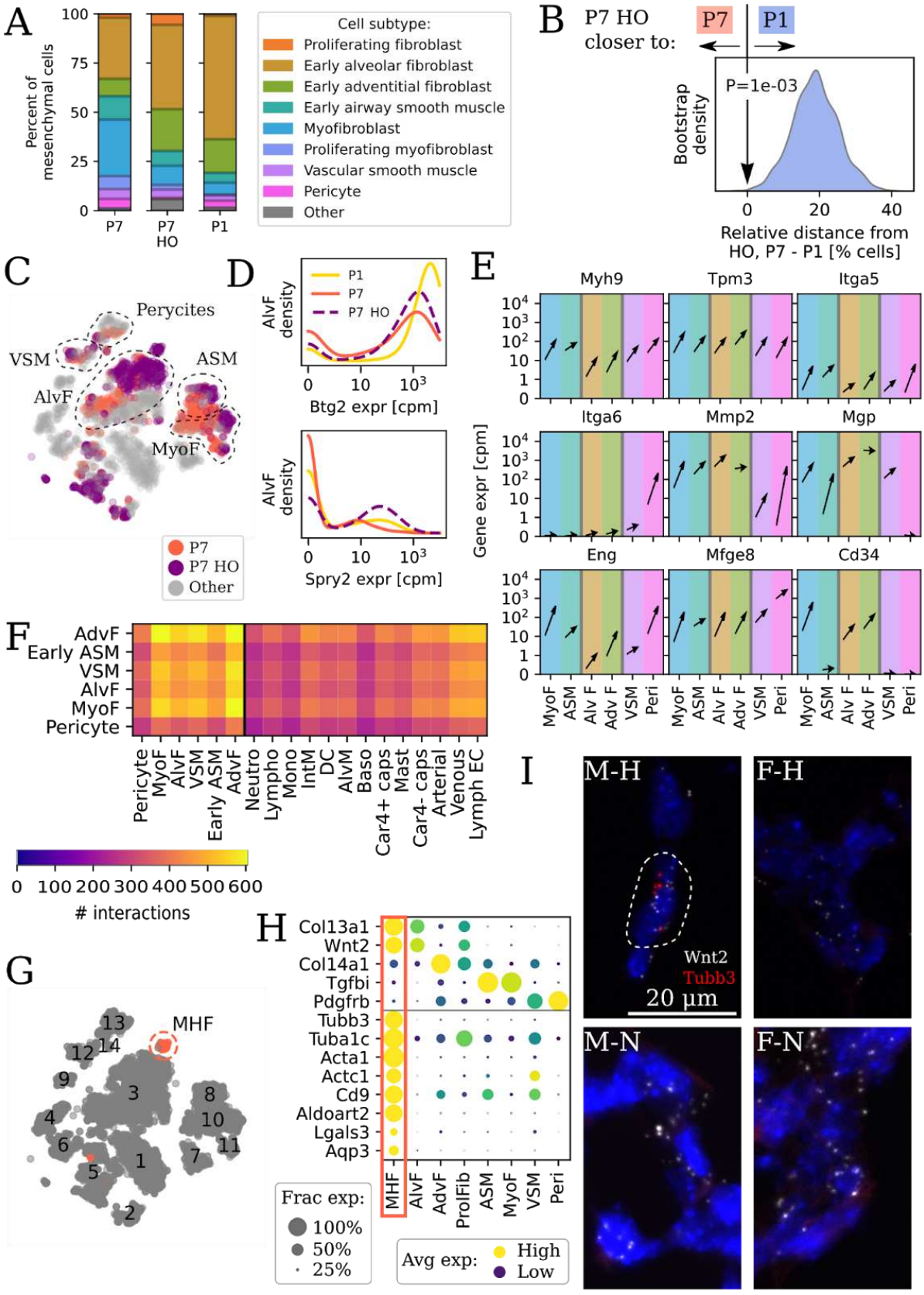

Figure 7. Hyperoxia profoundly affects mesenchymal composition and gene expression.

(A) Cell type composition in between normoxia and hyperoxia at P7 and in normoxia at P1. Colors as in **Figure 2A**. (B) Kernel density estimate of 1,000 bootstraps over cells (balanced at each time point) for the difference in distance on cell composition between hyperoxic mice and P7, minus hyperoxic mice and P1. A positive number indicated that the hyperoxic mesenchyme composition is closer to a healthy P1 mouse than a P7 one. Only 1 out of 1000 simulations showed hyperoxia closer to P7 than P1. (C) Embedding as in **Figure 2A** but coloring only P7 cells from healthy (red) and hyperoxia-exposed mice (purple), indicating a distinct separation between the two in many cell types. Several subtypes are highlighted by dashed lines. Cells from other time points are colored in grey. (D) Distribution of expression of *Btg2* and *Spry2* in alveolar fibroblasts at P1 (yellow solid), P7 healthy (red solid) and P7 hyperoxia-exposed mice (purple dashed). (E) Expression changes of genes involved in cytoskeleton, ECM adhesion, contractility, and ECM modulation from normoxia to hyperoxia across cell subtypes. (F) Heatmap of cell-cell interactions as in **Figure 6A** but for P7 mice exposed to hyperoxia. (G) Embedding as in **Figure 2A** but with the male hyperoxic fibroblast (MHF) population added in red. Other cells are colored in grey. (H) Dot plot of genes used to identify (top) and mark (bottom) MHF. (I) In-situ imaging of P7 lungs showing detection of *Tubb3*+ *Wnt2*+ MHF only in male, hyperoxia-exposed mice (M-H, top left) but not in females (F-H, top right) or healthy mice (bottom, M-N male left and F-N female right).

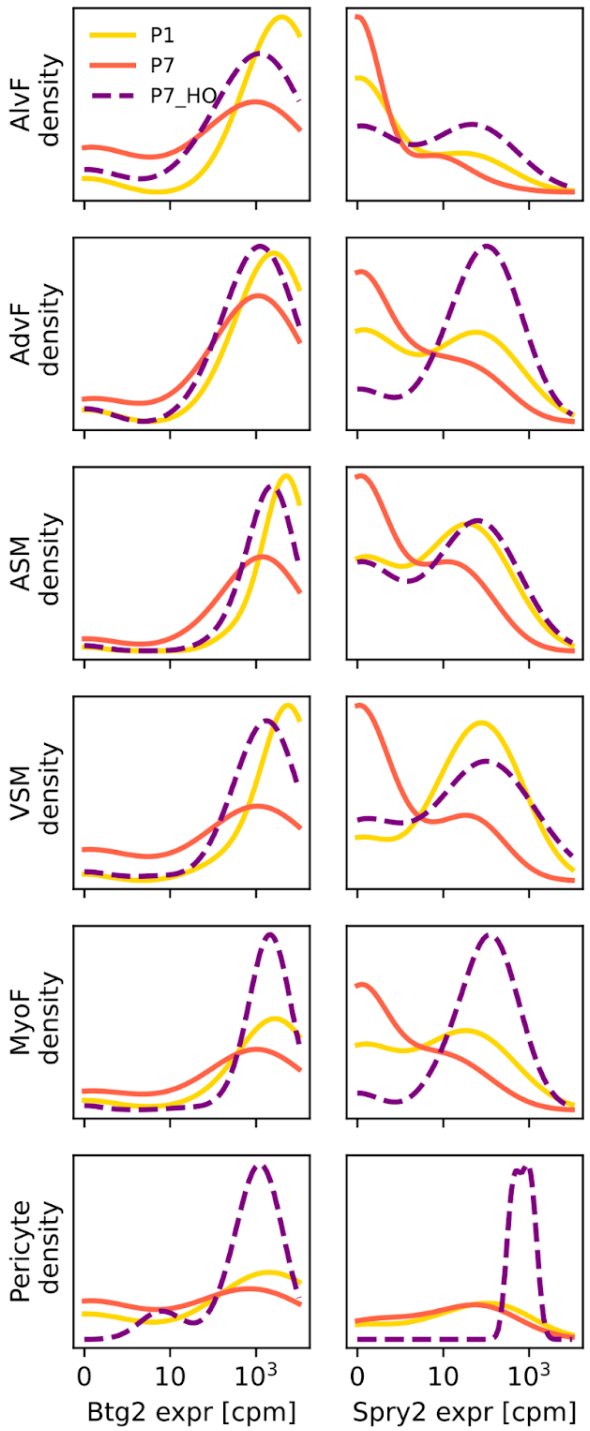

Figure 7 - Supplementary 1. Distributions of expression of Btg and Spry2 in all mesenchymal cell types at P1, P7, and at P7 after hyperoxia exposure. While hyperoxia delayed down-regulation of expression in some cell types (e.g. ASM), it was associated with distinct up-regulation in MyoF and pericytes.