Developmental Biology 425 (2017) 161–175

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Resource

Temporal, spatial, and phenotypical changes of PDGFRα expressing fibroblasts during late lung development☆

Mehari Endale[a], Shawn Ahlfeld[a], Erik Bao[a], Xiaoting Chen[b], Jenna Green[a], Zach Bess[a], Matthew T. Weirauch[b,c], Yan Xu[a], Anne Karina Perl[a,*]

[a] Perinatal Institute, Division of Pulmonary Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229-3039, USA
[b] Center of Autoimmune Genomics and Ethology, USA
[c] Divisions of Biomedical Informatics and Developmental Biology, USA

ARTICLE INFO

Keywords:
LungMAP
iReF
Myofibroblast
Matrixfibroblast
Lipofibroblast

ABSTRACT

Many studies have investigated the source and role of epithelial progenitors during lung development; such information is limited for fibroblast populations and their complex role in the developing lung. In this study, we characterized the spatial location, mRNA expression and Immunophenotyping of PDGFRα$^+$ fibroblasts during sacculation and alveolarization. Confocal microscopy identified spatial association of PDGFRα expressing fibroblasts with proximal epithelial cells of the branching bronchioles and the dilating acinar tubules at E16.5; with distal terminal saccules at E18.5; and with alveolar epithelial cells at PN7 and PN28. Immunohistochemistry for alpha smooth muscle actin revealed that PDGFRα$^+$ fibroblasts contribute to proximal peribronchiolar smooth muscle at E16.5 and to transient distal alveolar myofibroblasts at PN7. Time series RNA-Seq analyses of PDGFRα$^+$ fibroblasts identified differentially expressed genes that, based on gene expression similarity were clustered into 7 major gene expression profile patterns. The presence of myofibroblast and smooth muscle precursors at E16.5 and PN7 was reflected by a two-peak gene expression profile on these days and gene ontology enrichment in muscle contraction. Additional molecular and functional differences between peribronchiolar smooth muscle cells at E16.5 and transient intraseptal myofibroblasts at PN7 were suggested by a single peak in gene expression at PN7 with functional enrichment in cell projection and muscle cell differentiation. Immunophenotyping of subsets of PDGFRα$^+$ fibroblasts by flow cytometry confirmed the predicted increase in proliferation at E16.5 and PN7, and identified subsets of CD29$^+$ myofibroblasts and CD34$^+$ lipofibroblasts. These data can be further mined to develop novel hypotheses and valuable understanding of the molecular and cellular basis of alveolarization.

1. Introduction

It is well established that the interaction between the epithelium and its surrounding mesenchyme is the major driving force that induces lung branching morphogenesis and alveolar septation in the fetal lung. Due to the lack of well-established lineage markers for lung mesenchymal populations, however, the epithelial-mesenchymal cross talk has been investigated predominately from the lung epithelial side (Herriges and Morrisey, 2014; Morrisey and Hogan, 2010; Chen and Fine, 2016).

Over the last few decades, investigation of the pulmonary mesenchyme has been challenged by the remarkable complexity and diversity of fibroblast and other mesenchymal stem cell populations (Hogan et al., 2014; Bertoncello and McQualter, 2013; Volckaert et al., 2011; McQualter et al., 2009, 2013; Green et al., 2016). Mesenchymal subpopulations can be grouped by their anatomical location: tracheal, peribronchiolar, perivascular, alveolar and subpleural. Alternatively, subpopulations can be grouped by their cellular function: contractile (myofibroblasts), synthetic (matrix fibroblasts), and storage (lipofibroblasts). In contrast to a well-developed understanding of lung epithelial lineages we are just at the beginning of our understanding of lineage relationships amongst some of the fibroblast subpopulations in the lung. The most inclusive lineage marker is T-box transcription factor 4 (*Tbx4*). *Tbx4*, is expressed in a population of mesenchymal cells prior to branching morphogenesis that give rise to the entire postnatal lung mesenchyme (Zhang et al., 2013; Naiche et al., 2011; Arora et al., 2012). The transcription factor Foxd1 has been demonstrated to be a good marker specifically for pericytes (Hung et al., 2013). The lineage

☆ This work was supported by NIH grants R56 HL123969 (AKP), R01 HL104003 (AKP) and U01 HL122642 LungMAP (AKP, YX, EB), R21 HG008186 (MTW)
* Correspondence to: Children's Hospital Medical Center, Divisions of Neonatology and Pulmonary Biology, 3333 Burnet Avenue, Cincinnati, OH 45229-3039, USA.
E-mail address: Anne.Perl@cchmc.org (A.K. Perl).

http://dx.doi.org/10.1016/j.ydbio.2017.03.020
Received 17 November 2016; Received in revised form 7 March 2017; Accepted 21 March 2017
Available online 11 April 2017

of peribronchiolar and perivascular smooth muscle is defined by expression of WNT2/GLI1/ISL1 and is derived from cardiopulmonary mesoderm progenitors (CPP) (Peng and Morrisey, 2013). While CPPs do not contribute to pulmonary mesoderm populations after birth, *Gli1*-expressing mesenchymal cells are progenitors of peribronchiolar smooth muscle cells and alveolar myofibroblasts but not lipofibroblasts (Li et al., 2015). In contrast, FGF10-expressing mesenchymal cells of the developing distal lung can give rise to peribronchiolar smooth muscle and alveolar lipofibroblasts but not alveolar myofibroblasts (El Agha et al., 2014). Collectively, these data clearly show that relying on a single transcription factor or signaling pathway gene results in overlapping populations that are less exclusive than are commonly described in epithelial lineage relations. Lineage decisions of fibroblast populations are more complex and certain pathways are reactivated at different time points, due to specific functional requirements of the fibroblasts and the cell compartment they are supposed to support (i.e. epithelial versus endothelial).

Platelet-derived growth factor receptor alpha and beta (PDGFR) are believed to be common markers of lung pericytes and interstitial resident fibroblasts (Barron et al., 2016). Distal lung development requires an intact PDGFRα signaling pathway to provide normal functioning myofibroblasts that drive alveolar septation. Alveolar myofibroblasts increase in number and orient to developing septal tips during alveolarization; pharmacological or genetic inactivation of either PDGF-A or PDGFRα signaling results in loss of alveolar myofibroblasts and failure of alveolar septation (Bostrom et al., 2002, 1996; Lau et al., 2011; Lindahl et al., 1997). Defective alveolar septation is a predominate feature of bronchopulmonary dysplasia (Husain et al., 1998), and neonatal mesenchymal stromal cells obtained from human infants developing BPD show evidence that PDGFRα expression is reduced (Popova et al., 2014). These pre-clinical and clinical observations implicate a critical role for PDGF/PDGFRα signaling in resident fibroblast function and alveolar lung development. Thus, a better understanding of the contribution of various resident fibroblast cell types, their spatial-temporal localization, their gene expression and identification of additional cell markers for PDGFRα+ fibroblasts could lend insight into normal and abnormal alveolar development. Therefore, we more closely inspected the distribution and functional characteristics (i.e. myo-, lipo-, and matrix- fibroblastic) of PDGFRα-GFP$^+$ fibroblasts during normal alveolar lung development.

The definition of myofibroblasts and activated fibroblasts has been historically indistinct and includes increased proliferation, expression of alpha smooth muscle actin (αSMA) and excessive matrix production. Using a *Pdgfra*^{GFP} reporter mouse, we have previously established different functional roles, cellular markers and transcriptional profiles of PDGFRα$^+$ lung fibroblasts during adult lung regeneration following partial pneumonectomy. With PDGFRα kinase gain- or loss-of-function and in combination with manipulation of FGFR and/or PPAR-gamma signaling pathways, we identified and characterized temporally dynamic and functionally diverse subpopulations of matrix-, myo-, and lipofibroblasts that are required for injury-induced repair and regeneration (Green et al., 2016; Chen et al., 2012; Perl and Gale, 2009). Although others have reported on the temporal-spatial localization of PDGFRα$^+$ lung fibroblasts (Branchfield et al., 2016; Kimani et al., 2009; McGowan et al., 2008; Ntokou et al., 2015), a more thorough analysis that couples spatial-temporal localization with functional analysis and gene expression data may improve our understanding of the role of PDGFRα fibroblast populations during distal lung development. Moreover, assessment of proliferation and cellular differentiation in an enriched cell population will provide more insight into molecular interactions than previously published data analyzing whole lungs.

As part of the LungMAP consortium, in this study we aim to integrate anatomical location, functional phenotypes and gene expression patterns associated with PDGFRα$^+$ lung fibroblasts to dissect the differences in PDGFRα$^+$ lung fibroblasts that give rise to peribronchiolar smooth muscle and interstitial resident alveolar fibroblasts (iReFs). Additionally, we more closely inspected the distribution, association with epithelial and endothelial cells, and functional characteristics (i.e. myo-, lipo-, and matrix- fibroblastic) of PDGFRα-GFP$^+$ fibroblasts through canalicular, saccular, and alveolar lung development. Our data provide keen insights into dynamically changing spatial relationships of lung fibroblasts to epithelial and endothelial cells during canalicular, saccular and alveolar lung development. Dynamically changing gene expression profiles of PDGFRα fibroblasts reveal their changes in phenotype and function as they differentiate into proximal smooth muscle, distal alveolar myofibroblasts and finally supportive interstitial alveolar matrix and lipofibroblasts.

2. Results

2.1. PDGFRα$^+$ fibroblasts switch from a proximal bronchiolar to distal alveolar location

To assess temporal-spatial localization of PDGFRα fibroblasts and identify their potential interactions with other cell subsets, we performed immunofluorescence microscopy on lungs of *Pdgfra*^{GFP} knock in mice from E16.5 to PN28. Distribution of PDGFRα-GFP$^+$ fibroblasts along the proximal-distal axis of the lung was determined by immunofluorescence using the epithelial marker E-Cadherin, confocal microscopy with 3D reconstruction. Differentiation of PDGFRα-GFP$^+$ fibroblasts into smooth muscle cells and myofibroblasts was assessed by co-localization of the nuclear GFP with αSMA (Fig. 1).

During the canalicular stage (E16.5–17.5) a dense cell layer of PDGFRα-GFP$^+$ fibroblasts are located along bronchi and terminal bronchioles with the cells closest to the epithelium co-expressing αSMA. PDGFRα-GFP$^+$ fibroblasts line the dilating acinar tubules but are absent from the more distal acinar buds (Fig. 1A, E). During the saccular stage (E18.5 - PN4) peribronchiolar PDGFRα-GFP$^+$ fibroblasts are sparse, and predominantly negative for αSMA. The majority of PDGFRα-GFP$^+$ fibroblasts associate with the terminal saccular structures, which will give rise postnatally to the alveolar ducts and alveoli (Fig. 1B, F). During the alveolar phase (PN5-PN28), PDGFRα-GFP$^+$ fibroblasts become more restricted to the alveolar structures. At PN7, a significant number of interstitial PDGFRα-GFP$^+$ fibroblasts activate expression of αSMA and differentiate into alveolar myofibroblasts (Fig. 1C, G). By PN28, individual PDGFRα-GFP$^+$ fibroblasts are scattered throughout the alveolar interstitium; the relative number is reduced compared to other time points, and co-expression of αSMA is largely lacking (Fig. 1D, G, H). Assessment of *Pdgfra*^{GFP} lung single cell suspensions confirmed that during distal lung development, as a percentage of total lung cells, PDGFRα-GFP$^+$ fibroblasts significantly diminish (Fig. 1H). These spatiotemporal relationships are consistent with previous findings (Branchfield et al., 2016; McGowan et al., 2008; Ntokou et al., 2015). These data demonstrate that PDGFRα fibroblasts form a layer along the conducting airways directly beneath the epithelium and contribute to peribronchiolar αSMA$^+$ smooth muscle cells during terminal branching and sacculation. These peribronchiolar smooth muscle cells form a concentric pattern around mature airways, which is distinct from the perivascular mesh pattern found in mature arteries (Data in Brief Fig. 1). During saccular maturation and alveolarization, individual PDGFRα fibroblasts are localized to terminal bronchioles, terminal saccules and alveolar structures. In both compartments, and at different developmental time points, bronchiolar and alveolar PDGFRα$^+$ fibroblasts can differentiate into contractile cells expressing αSMA.

2.2. Spatial relation of PDGFRα-GFP$^+$ fibroblasts to distinct epithelial subpopulations

Apposing layers of epithelial and mesenchymal cells, separated by a

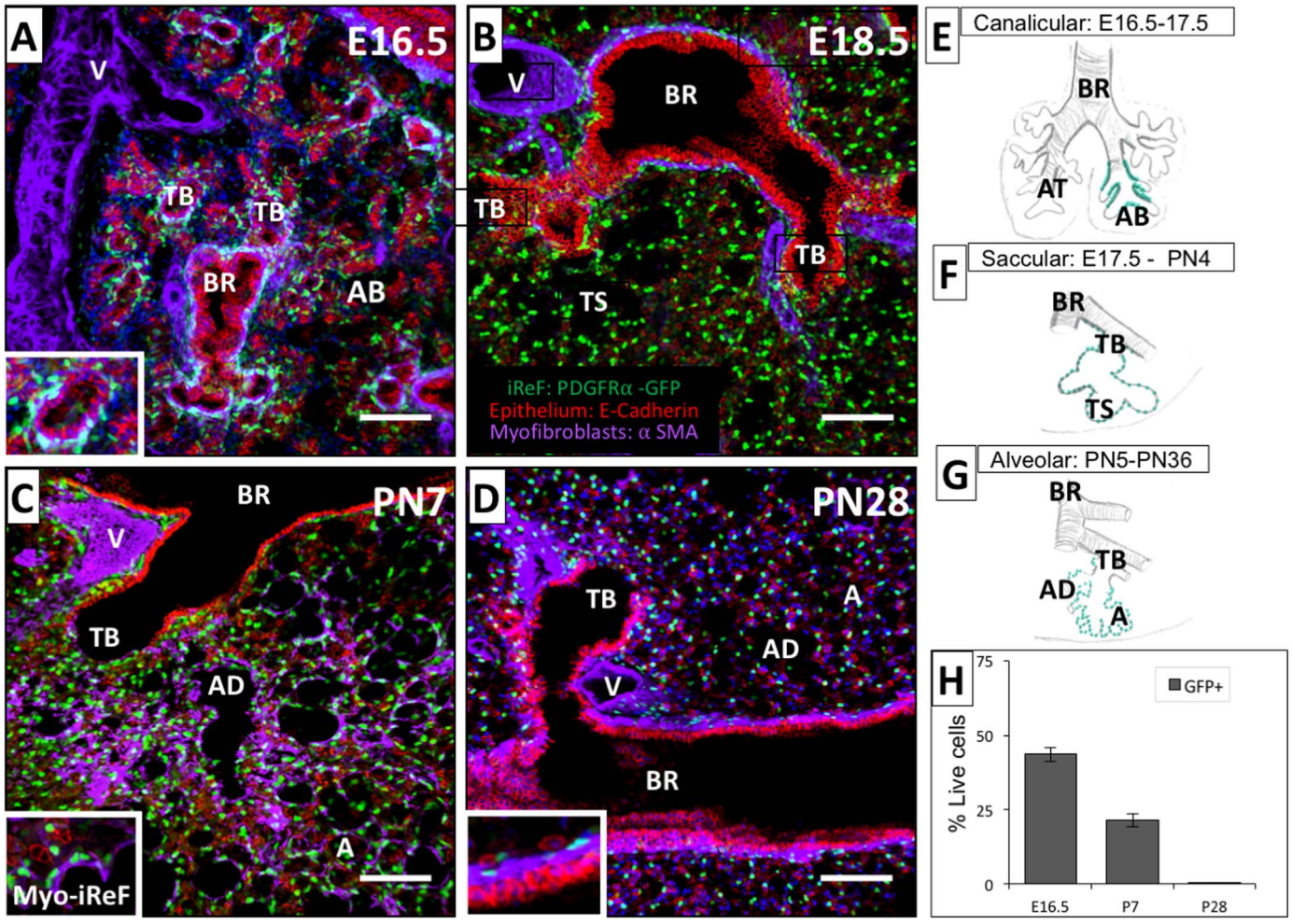

Fig. 1:. Spatial localization of PDGFRα-GFP fibroblasts to anatomical structures: Immunohistochemistry for E-cadherin (red), smooth muscle actin (αSMA, purple) on transgenic *Pdgfra*^{GFP} lungs (PDGFRα-GFP^{+}, green) and 3D confocal reconstruction of murine lung tissues at E16.5, E18.5, PN7 and PN28. At E16.5, the canalicular stage, the majority of PDGFRα-GFP^{+} fibroblasts (green) are localized along the bronchi and terminal bronchioles and some co-express αSMA (A, E). At E18.5, the saccular stage, PDGFRα-GFP^{+} fibroblasts distributed along the conducting airways thinned out and are now located as single cells in the alveolar compartment (B, F). At PN7, the alveolar stage, PDGFRα-GFP^{+} fibroblasts are distributed in the interstitium, with the majority co-expressing αSMA (myofibroblasts) (C, G). At PN28 the population density of PDGFRα-GFP^{+} fibroblasts was highly reduced and no co-localization of αSMA was found (D). (E, F, G) depicts the schematic representation of PDGFRα-GFP^{+} fibroblasts at different stages of development. (H) Flow cytometry analyses to quantify the percentage of PDGFRα-GFP^{+} fibroblasts as % of all live cells. (Red: E-cadherin; Green: PDGFRα; Pink: αSMA; Blue: DAPI). (BR: bronchioles, AT: acinar tubules, AB: acinar buds, V: vessel, TB: terminal bronchioles, TS: terminal saccules, AD: alveolar ducts, A: alveoli). All scale bars are 100 μm.

basement membrane and the capillary network, form epithelial-mesenchymal trophic units (Brewster et al., 1990; Evans et al., 1999). In the lung, two distinct epithelial-mesenchymal trophic units have been identified: 1) proximal conducting airways consisting of bronchiolar cells and 2) distal alveolar structures containing alveolar epithelial cells. The cells that comprise the mesenchymal layer of these distinct units have not been defined. To further assess temporal and spatial proximity of PDGFRα-GFP^{+} fibroblasts with distinct epithelial subpopulations, we performed immunohistochemistry on *Pdgfra*^{GFP} lungs for epithelial-specific transcription factors SOX2, HOPX and SOX9 (Que et al., 2009; Barkauskas et al., 2013; Ng et al., 1997; Perl et al., 2005; Rockich et al., 2013; Figs. 2–5). SOX2 is expressed in proximal bronchiolar cells and located in the nucleus. HOPX, is expressed in bipotential pneumocytes located in the acinar tubules during the canalicular stages and becomes a more restricted marker specific for type 1 alveolar epithelial cells (AECI) in adult lungs (Jain et al., 2015). HOPX can be found in the cytoplasm and nucleus. SOX9 is restricted to distal acinar buds and is essential for early branching morphogenesis, distal epithelial differentiation, and extracellular matrix organization (Perl et al., 2005; Rockich et al., 2013; Ustiyan et al., 2016). The SOX9 protein is localized to the nucleus and is also rarely expressed in some fibroblasts adjacent to the mature peribronchiolar smooth muscle layer.

2.3. Canalicular lung at E16.5

SOX2 expression in the respiratory epithelium gradually decreases from proximal to distal and correlates with expression of αSMA in the underlying mesenchyme. Proximal bronchiolar epithelial cells that express SOX2 protein are in close apposition to a discrete layer of peribronchiolar smooth muscle cells consisting of double-positive PDGFRα-GFP^{+}/αSMA^{+} fibroblasts (Fig. 2A). The region of the bronchiolar tree where αSMA expression in the underlying mesenchyme tapers off and gradually vanishes, identifies a bronchiolar transition zone that is continuous with the acinar tubules. In the late canalicular stage, the epithelium in this bronchiolar transition zone expresses HOPX, and the underlying fibroblasts form a PDGFRα-GFP^{+} mesenchymal layer in which αSMA is greatly diminished and absent more distally (Fig. 2B). Acinar buds are the most distal structures and that ultimately will mature into alveolar ducts and alveoli. During the canalicular stage, the epithelial cells of the acinar buds express SOX9 and the underlying mesenchyme contains almost no PDGFRα-GFP^{+}

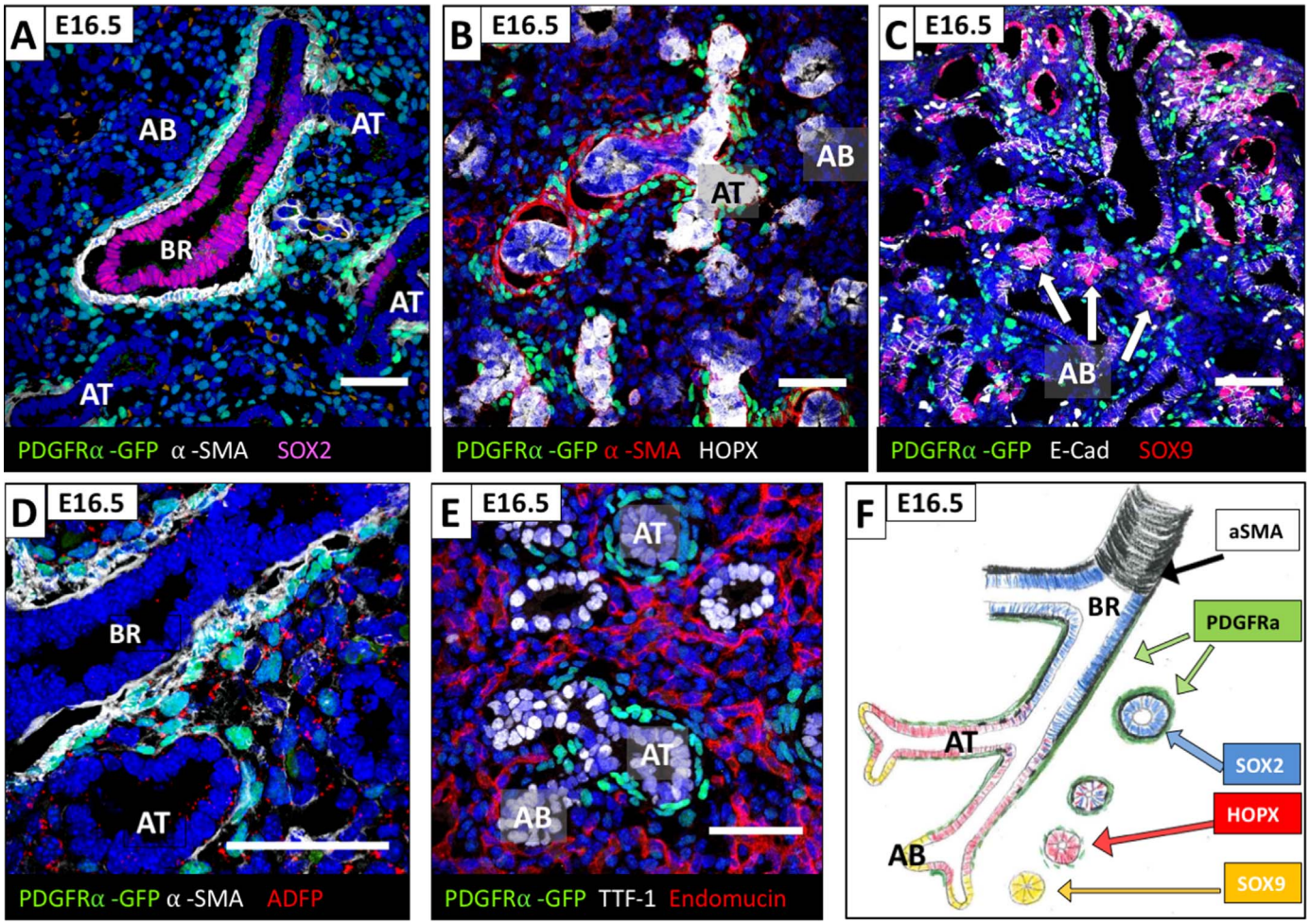

Fig. 2. Association of PDGFRα-GFP$^+$ fibroblasts with distinct epithelial populations in the canalicular lung. Assessment of proximity of PDGFRα-GFP$^+$ fibroblasts to proximal and distal lung epithelial cells using immunohistochemistry for SOX2, HOPX and SOX9 on E16.5 lungs. (A) PDGFRα-GFP$^+$/αSMA positive fibroblasts are localized adjacent to the SOX2$^+$ bronchial epithelial cells (pink) and form the peribronchiolar smooth muscle. (B) PDGFRα-GFP$^+$ and αSMA negative fibroblasts are localized adjacent to the HOPX$^+$ bipotent embryonic alveolar progenitors in the acinar tubules. (C) PDGFRα-GFP$^+$ fibroblasts are not associated with the SOX9 positive alveolar progenitor cells in the acinar buds (pink). (D) Peribronchiolar PDGFRα-GFP$^+$/αSMA positive and PDGFRα-GFP$^+$ fibroblasts associated with acinar tubules are negative for the lipofibroblast marker ADFP (adipose differentiation related protein, Perilipin2). (E) PDGFRα-GFP$^+$ fibroblasts are not associated with endomucine positive capillaries. (F) Schematic representation of the proximal to distal distribution of PDGFRα-GFP$^+$ ^{P} fibroblasts along with bronchiolar and acinar epithelial cells. (BR: bronchioles, AT: acinar tubules, AB: acinar buds, V: vessel). All scale bars are 50 μm.

fibroblasts (Fig. 2C). The lipofibroblast phenotype was assessed by immunohistochemistry for perilipin (ADFP). Perilipin staining was not co-localized with nuclear PDGFRα-GFP, suggesting the absence of lipofibroblasts in the canalicular PDGFRα$^+$ fibroblast population (Fig. 2D). Proximity of PDGFRα$^+$ fibroblasts to capillaries and vessels was assessed by immunohistochemistry for endomucin. At E16.5, PDGFRα-GFP$^+$ fibroblasts form a clear layer between bronchiolar epithelial cells and stromal cells adjacent to endomucin positive cells of the vasculature (Fig. 2E).

2.4. Saccular lung at E18.5

During the saccular stage (E18.5), SOX2-expressing bronchiolar epithelial cells are in juxtaposition to an αSMA$^+$ smooth muscle layer that contains only a few isolated PDGFRα-GFP$^+$ fibroblasts (Fig. 3A). In the terminal saccules, which will give rise to the alveolar ducts and alveoli, HOPX$^+$ epithelial cells and PDGFRα-GFP$^+$ fibroblasts are in close proximity (Fig. 3B). Moreover, PDGFRα-GFP$^+$ fibroblasts in the terminal saccules are negative for αSMA. Compared to E16.5, SOX9 expression is significantly more restricted to individual epithelial cells in the developing saccules, which are rarely adjacent to PDGFRα-GFP$^+$ fibroblasts (Fig. 3C). Perilipin and PDGFRα-GFP$^+$ did not co-localize (Fig. 3D). In contrast to E16.5, PDGFRα-GFP$^+$ fibroblasts are now adjacent to endomucin positive cells (Fig. 3E).

2.5. Alveolar lung at PN7

At PN7, SOX2-expressing epithelial cells in the bronchiolar alveolar duct junction are not associated with PDGFRα-GFP$^+$ fibroblasts (Fig. 4A). In the alveolar epithelium, HOPX is expressed in AECIs that comprise the alveolar wall (Fig. 4B), while SOX9 expression is restricted to single cells that are predominantly located in corners that are shared amongst multiple alveoli (Fig. 4C). PDGFRα-GFP$^+$ fibroblasts are located in close proximity to AECI but not SOX9$^+$ cells. The majority of PDGFRα-GFP$^+$ fibroblasts reside within the bases and tips of secondary crests, where endothelial cells (Fig. 4E), epithelial cells, PDGFRα-GFP$^+$/αSMA$^+$ myofibroblasts, and PDGFRα-GFP$^+$/ADFP$^+$ lipofibroblasts are positioned intimately (Fig. 4D–F). Closer inspection of developing secondary crests by electron microscopy revealed fibroblasts containing lipid vesicles at the base of the emerging septum (presumed lipofibroblasts), and elastin deposition in fibroblasts residing at the septal tip (presumed matrix fibroblasts). (Fig. 4F, note no myofibrils detected in this image, graphic overlay is for illustration purposes only).

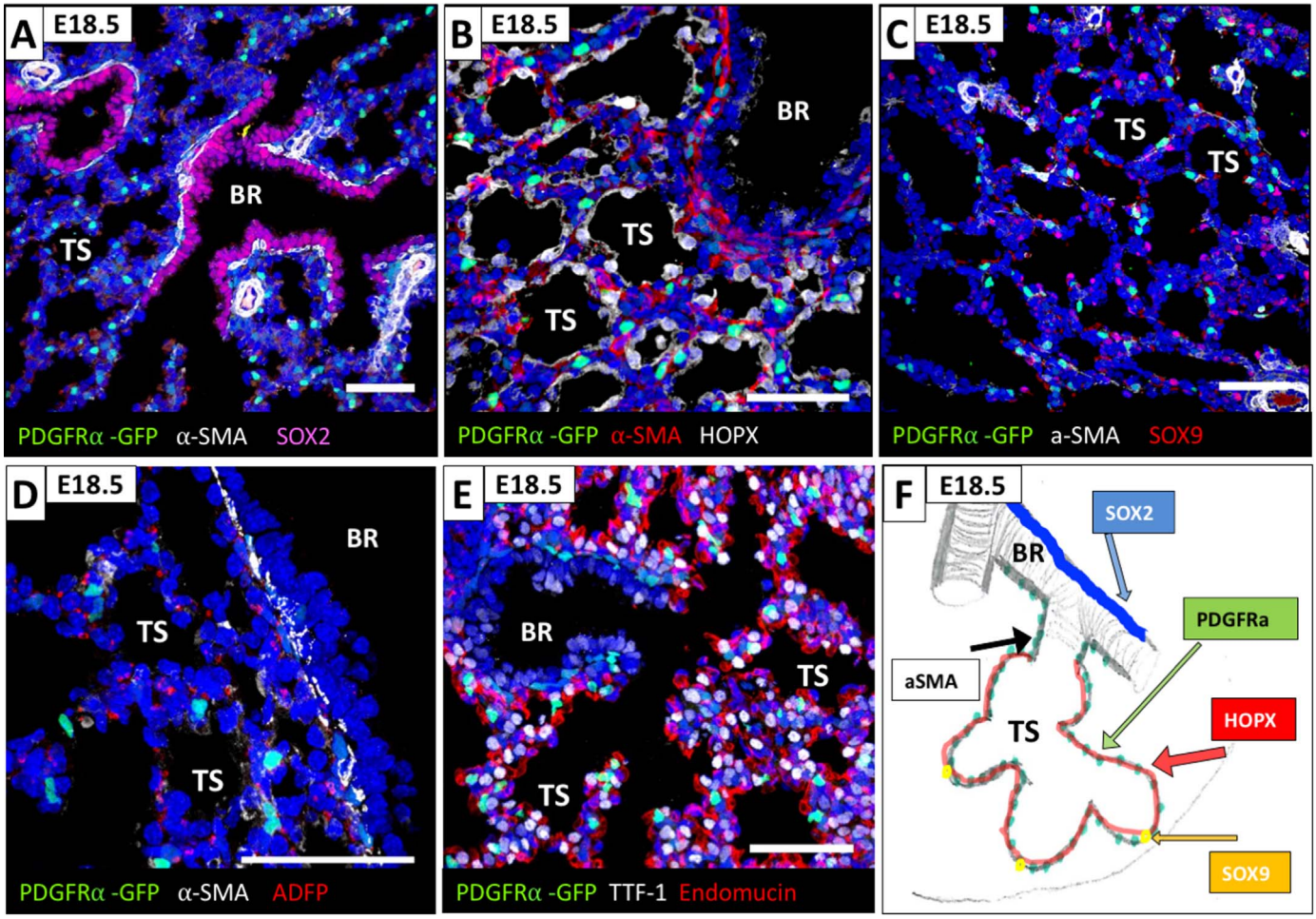

Fig. 3. Association of PDGFRα-GFP$^+$ fibroblasts with distinct epithelial populations in saccular lung. Assessment of proximity of PDGFRα-GFP$^+$ fibroblasts to proximal and distal lung epithelial cells using immunohistochemistry for SOX2, HOPX and SOX9 on E18.5 lungs. In contrast to canalicular lungs PDGFRα-GFP$^+$ fibroblasts are not forming fibroblast layers around conducting airways but are found as individual cells along tubular saccules and are αSMA and lipid negative. (A) Peribronchiolar αSMA positive fibroblasts adjacent to the SOX2$^+$ bronchial epithelial cells (pink) are negative for PDGFRα-GFP$^+$. (B) PDGFRα-GFP$^+$ are negative for αSMA and localized adjacent to the HOPX$^+$ epithelial cells (white) in the terminal saccules. (C) PDGFRα-GFP$^+$ fibroblasts are randomly associated with the individual SOX9 positive alveolar epithelial progenitor cells (pink). (D) PDGFRα-GFP$^+$ fibroblasts associated with terminal saccules are negative for the lipofibroblast marker ADFP (adipose differentiation related protein, Perilipin2). (E) PDGFRα-GFP$^+$ fibroblasts and endomucin positive capillaries are in close proximity. (F) Schematic representation of the proximal to distal distribution of PDGFRα-GFP$^+$ fibroblasts along with bronchioles and terminal saccules. (BR: bronchioles, V: vessel, TS: terminal saccules). All scale bars are 50 μm.

2.6. End of alveolarization PN28

Toward the end of the alveolar stage at PN28, SOX2 is expressed in the bronchiolar epithelium, including the epithelium of the bronchiolar-alveolar duct junction (Fig. 5A). Some PDGFRα-GFP$^+$ fibroblasts are found adjacent to but predominately distinct from the αSMA$^+$ smooth muscle layer (Fig. 1D, insert). This spatial proximity of PDGFRα-GFP$^+$ fibroblasts to HOPX$^+$ and SOX9$^+$ alveolar epithelial cells is comparable to PN7 (Fig. 5B, C). Contrary to the earlier period of alveolar development, where there was robust expression of αSMA, now there are only sparse interstitial PDGFRα-GFP$^+$/αSMA$^+$ myofibroblasts but there are clearly identifiable interstitial PDGFRα-GFP$^+$/ADFP$^+$ lipofibroblasts (Fig. 5B–D).

2.7. Dynamic changes in gene expression profiles in PDGFRα positive fibroblasts from the canalicular to the alveolar stages

PDGFRα-fibroblasts compartmentalize with proximal bronchiolar epithelial cells in the canalicular stage and later with distal alveolar epithelial cells during saccular and alveolar development and exhibit different functions co-expressing αSMA and perilipin. We therefore compared their gene expression profiles throughout the canalicular, saccular and alveolar stages. PDGFRα$^+$ fibroblasts were isolated from mouse lungs at E16.5, E18.5, PN7 and PN28, and RNA obtained from the resulting enriched PDGFRα population was subjected to RNA-Seq analysis. Bayesian Analysis of Time Series (BATS) was performed and 1908 genes were identified that were differentially expressed across these developmental stages (Bayesian Factor < 0.01) (Angelini et al., 2008; Fig. 6A). Hierarchical clustering and principle component analyses demonstrated clear separation of gene expression profiles in PDGFRα$^+$ fibroblasts isolated at different stages of distal lung development (Fig. 6B, C). Differentially expressed genes were subjected to pattern recognition using Short Time series Expression Miner, STEM (Ernst and Bar-Joseph, 2006). STEM clustered 1607 genes into 7 major gene expression profiles (Fig. 6D–J). Unique functional enrichments of genes within each profile were identified by comparing all profiles to each other using Toppcluster (See Tables 1 and 2) (http://toppcluster.cchmc.org/) (Kaimal et al., 2010).

2.8. Gene expression and functional enrichment at E16.5

Profile #10 consists of genes that are gradually downregulated throughout distal lung development and it inversely correlates with profiles #39 and #23 which consist of genes that are gradually upregulated (Fig. 6F, J). To account for molecular agonists and antagonists of biological processes we looked for shared functional

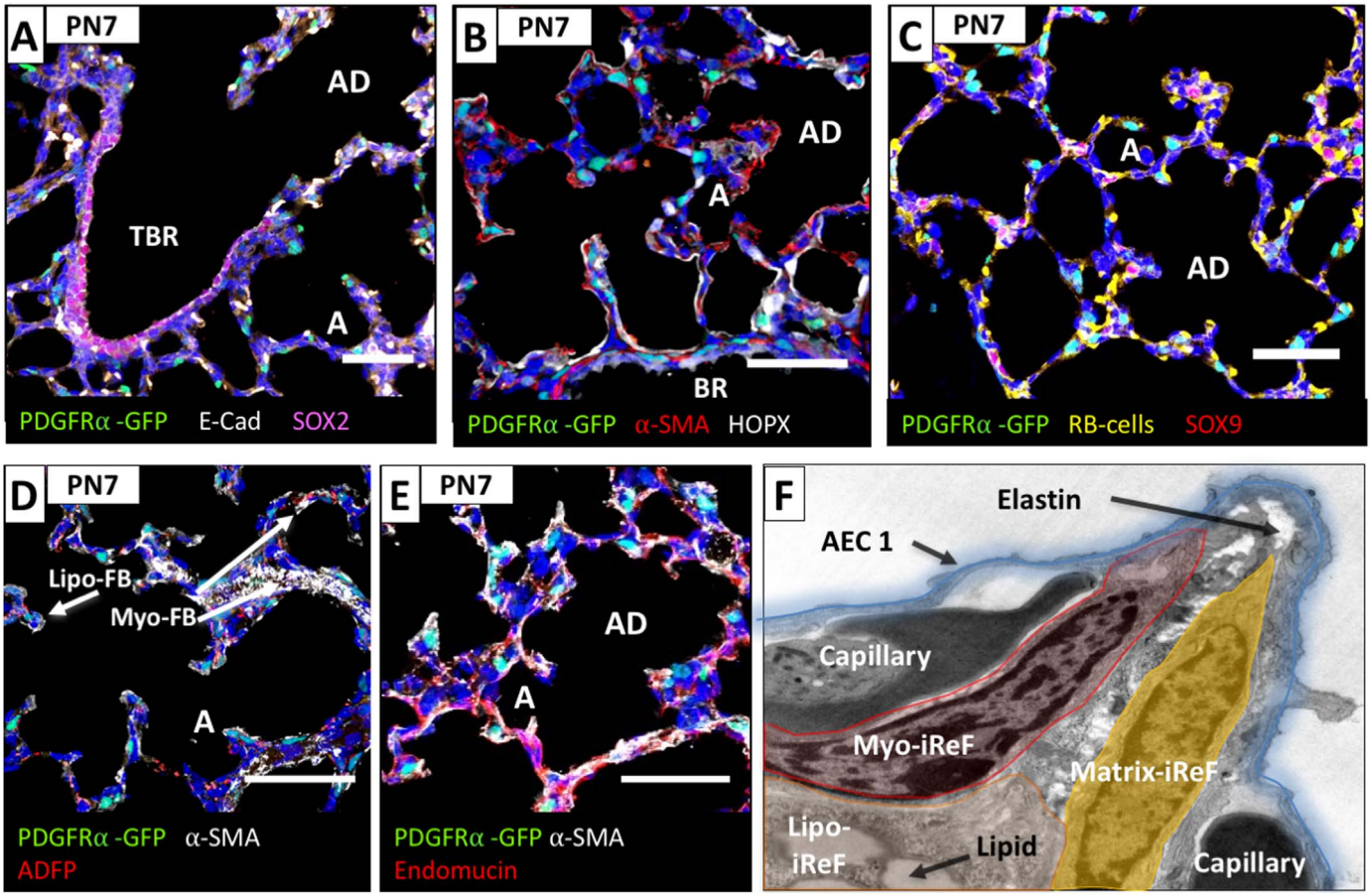

Fig. 4. Distribution and myofibroblast differentiation of PDGFRα-GFP$^+$ fibroblasts during secondary septation in the alveolar lung. Assessment of proximity of PDGFRα-GFP$^+$ fibroblasts to proximal and distal lung epithelial cells using immunohistochemistry for SOX2, HOPX and SOX9 on Pn7 lungs. PDGFRα-GFP$^+$ fibroblasts were found as individual cells throughout the alveolar ducts and alveoli and are αSMA or lipid positive. (A) Peribronchiolar αSMA positive fibroblasts adjacent to the SOX2$^+$ bronchial epithelial cells (pink) are negative for PDGFRα-GFP$^+$. (B) PDGFRα-GFP$^+$ fibroblasts in the emerging septae are αSMA positive and in close proximity to HOPX$^+$ alveolar epithelial type I cells (white). (C) PDGFRα-GFP$^+$ fibroblasts are not associated with the individual SOX9 positive alveolar epithelial progenitor cells (pink). (D) PDGFRα-GFP$^+$ fibroblasts located at the base of alveolar septae are positive (arrow) for the lipofibroblast marker ADFP (adipose differentiation related protein, Perilipin2). PDGFRα-GFP$^+$/αSMA, myofibroblasts, are found in newly forming septae that project into the alveoli (arrows). (E) PDGFRα-GFP$^+$ fibroblasts and endomucin positive capillaries are in close proximity. (F) Electron microscope photograph was overlaid with a schematic drawing of a PN7 emerging septum. Fibroblasts with readily detectable lipid vesicles can be found at the base of the growing septum (lipo-iReF). Elastin deposition (arrow) originating from the myofibril containing myofibroblast (myo-iReF) and matrix fibroblast (matrix-iReF) can be found in the center of the septum and at the septal tip. Alveolar epithelial cells form a thin layer over the capillaries and fibroblast core of the septum. (BR: bronchioles, TBR: terminal bronchioles, AD: alveolar ducts, A: alveoli, AEC1: alveolar epithelial type I cell, iReF: interstitial resident fibroblast). All scale bars are 50 μm.

enrichment categories of either up or downregulated genes. With this approach, we identified gene enrichment for negative regulation of cell migration, regulation of response to wounding, and negative regulation of cell death (See Table 2). These profiles suggest that cell migration decreases over time; that developmental processes are also used in responses to wounding; and that cell death is negatively regulated upon completion of distal lung development.

2.9. Gene expression and functional enrichment at E18.5

During the saccular stage at E18.5, the majority of profiles indicate that gene expression is relatively downregulated (Profile #20 and #13). Only 66 of the 1607 genes are significantly upregulated at E18.5 when compared to all other time points (Data in Brief Table 3). Unique gene enrichment categories for these genes include “protein kinase inhibitor activity” and “delayed wound healing.” Co-expression analysis revealed high correlation with “genes upregulated in cultured stromal stem cells from adipose tissue.” These gene expression data suggest that during saccular lung development, cells are less mitotic and are shifting from the smooth muscle phenotype of the canalicular bronchioles to the interstitial matrix- and lipofibroblasts.

2.10. Gene expression and functional enrichment at PN7

Profiles #1, #18 and #39 comprise genes that are upregulated specifically during active alveolar septation. Additionally, genes in #39 increase until PN28. Analysis revealed that expression is enriched for signal transduction, actin cytoskeleton, ECM organization, cell migration, angiogenesis and epithelial development (Table 1). These enrichment categories identify myo-, matrix-, and lipo-fibroblast functions within the PDGFRα$^+$ population, and imply that PDGFRα$^+$ interstitial fibroblasts play an important role in regulating both angiogenesis and epithelial differentiation during alveolar septation.

2.11. Gene expression and functional enrichment at PN28

Profiles #39, #13, and #23 are comprised of genes that have high expression in the late phases of distal lung development. Expression profile #39 contains genes that steadily increase throughout lung development. It consists of 271 genes, 18 of which encode transcription factors. The major unique enriched gene categories include “cell motility”, “cell matrix adhesion”, “tube development” and “renal development”. Profiles #13 and #23 include genes that are increasing from E18.5 onwards. These gene ontology categories contain functional

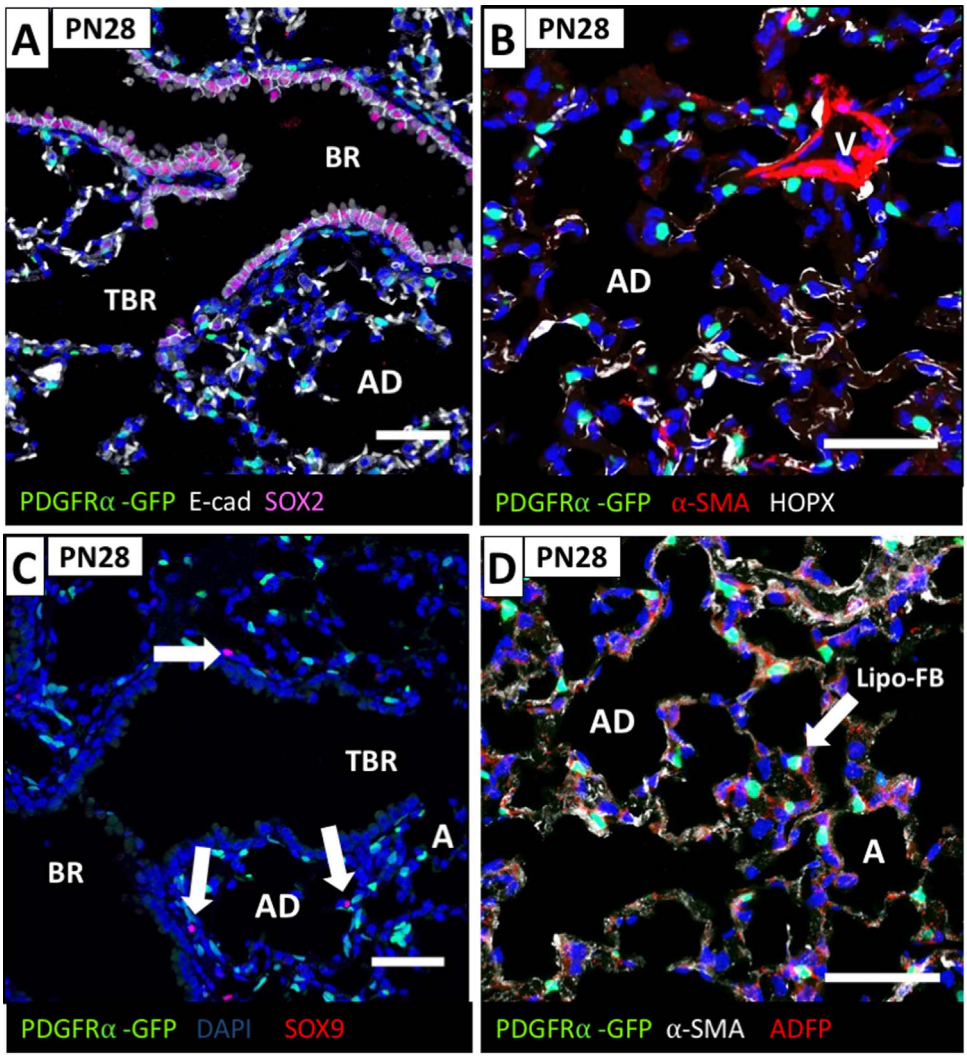

Fig. 5. Spatial localization of PDGFRα-GFP fibroblasts to the bronchiolar tree at PN28. A marked reduction in the density of PDGFRα-GFP$^+$ cells in the interstitial space were observed at this stage and a representative image in (A) depicts a layer of SOX2$^+$ epithelial cells lining the conducting air way until the level of the bronchoalveolar junction where it terminates. Assessment of proximity of PDGFRα-GFP$^+$ fibroblasts to proximal and distal lung epithelial cells using immunohistochemistry for SOX2, HOPX and SOX9 on PN7 lungs. A marked reduction in the density of PDGFRα-GFP$^+$ cells in the interstitial space was observed, cells are αSMA negative but some are lipid positive. (A) Peribronchiolar αSMA positive fibroblasts adjacent to the SOX2$^+$ bronchial epithelial cells (pink) are negative for PDGFRα-GFP. (B) PDGFRα-GFP$^+$ fibroblasts are in close proximity to HOPX$^+$ alveolar epithelial type I cells (white) and negative for αSMA. (C) PDGFRα-GFP$^+$ fibroblasts are not associated with the sparse individual SOX9 positive alveolar epithelial progenitor cells (pink, arrows). (D) PDGFRα-GFP$^+$ fibroblasts (arrow) located at the base of alveolar septae are positive for the lipofibroblast marker ADFP (adipose differentiation related protein, Perilipin2). (BR: bronchioles, TBR: terminal bronchioles, AD: alveolar ducts, A: alveoli, AEC1: alveolar epithelial type I cell, iReF: interstitial resident fibroblast). All scale bars are 50 μm.

terms that suggest differentiation of lipid and matrix fibroblasts ("regulation of lipid metabolic process", "vesicle mediated transport", and "Golgi membrane") (Eyden, 2008; Tsupykov et al., 2016). The predominance of PDGFRα-GFP$^+$/Perilipin lipofibroblasts that we detected at PN28 corroborates these observations (Fig. 5D).

2.12. Smooth muscle versus myofibroblasts

Despite the use of single or combinations of expression markers, including PDGFRα and αSMA, the ability to identify the distinct developmental functions of the heterogeneous population of lung fibroblasts remains challenging. Previously, single cell RNA-Seq analysis performed by the LungMAP consortium of E16.5 and E18.5 lungs classified individual fibroblast populations based on unique gene expression signatures and named them based on functional gene enrichment categories. Searching this single cell RNA- Seq dataset for *Pdgfra* and *Acta2* expressing cells identified these smooth muscle and myofibroblast progenitors in 4 different fibroblast populations with distinct functional phenotypes, Myofibroblast/SmoothMuscle, Proliferative Mesenchymal Progenitor cells, Intermediate Fibroblasts and MatrixFBs (Du et al., 2015; Guo et al., 2016). Our present work expands upon this data by examining gene expression patterns of PDGFRα expressing fibroblasts throughout the spectrum of distal lung development and comparing/contrasting patterns of gene expression at each distinct stage. Additionally, we employed immunohistochemistry and confocal microscopy to map the unique locations of these cells in the developing lung to better understand their individual microenvironments and infer how cell-cell interactions might specify each population's distinct function. Our data demonstrate that at E16.5, PDGFRα-GFP$^+$/αSMA$^+$ smooth muscle cells are located proximally and adjacent to the SOX2$^+$ bronchiolar epithelium (insert Fig. 1 A and Fig. 2). Starting at E18.5, expression of PDGFRα-GFP can be detected in the developing alveolar compartment. Previously, it was postulated that interstitial septal myofibroblasts were derived from the migration of peribronchiolar PDGFRα-GFP$^+$/αSMA$^+$ smooth muscle cells to the distally developing alveolar compartment (Lindahl et al., 1997). While migration of proximal PDGFRα$^+$ cells cannot be excluded, recent linage tracing using a constitutive Pdgfra^{Cre} demonstrated that proximal

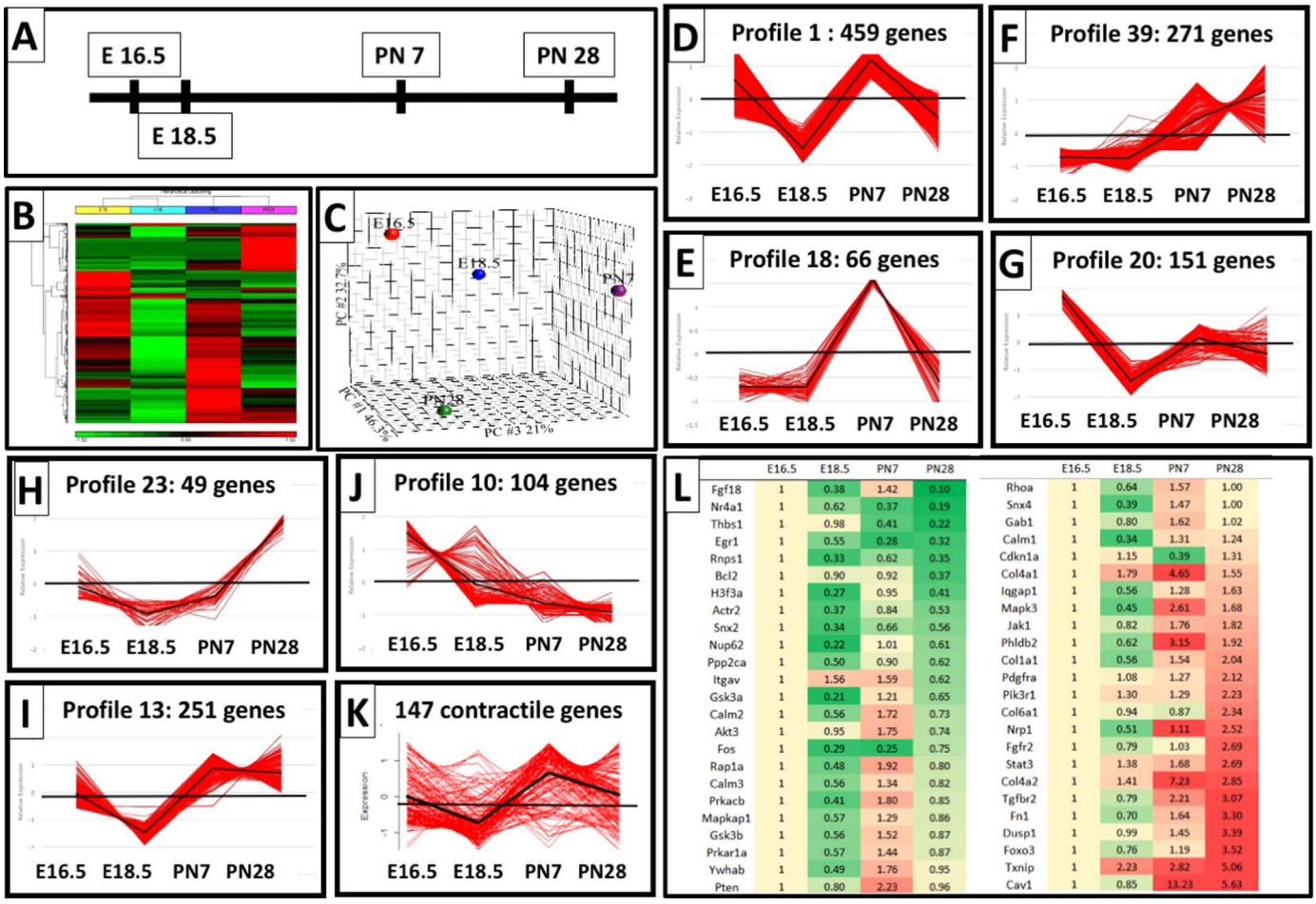

Fig. 6. Gene expression profiling of PDGFRα expressing fibroblasts from the canalicular stage throughout alveolarization: (A) Schematic representation of the time points selected for RNA-Seq analysis of CD140a positive lung fibroblasts. (B) Two-dimensional hierarchical clustering of 1908 genes that were significantly differentially expressed at one of the time points. Red blocks represent upregulated and green blocks represent downregulated expression. (B) Principal Component Analysis (PCA) using all probe sets clearly separate the subpopulations between time points (D–J): 7 gene profiles identified by STEM method that represent genes with similar changes in expression at each developmental time point. Note: profiles 1 and 18 (D, E) with peak expression at E16.5 and PN7, when myofibroblasts significantly contribute to the PDGFRα fibroblasts. (K) Expression of 147 genes that belong to the GO category"contractile function". (L) heatmap of 48 differentially expressed genes that are known PDGF downstream genes. Gene lists are provided in the Data in Brief.

smooth muscles cells do not contribute to PDGFRα-GFP$^+$/αSMA$^+$ alveolar interstitial myofibroblasts (Ntokou et al., 2015). Together with our RNA-Seq profiling data, the notion that E16.5 PDGFRα-GFP$^+$/αSMA$^+$ fibroblasts do not migrate from their proximal position to the alveolar compartment and rather fully differentiate into proximal PDGFRα^{neg} smooth muscle cells is probable. While further linage tracing experiments and or co-expression of markers will be necessary to fully answer the question, we speculate that distal interstitial PDGFRα-GFP$^+$/αSMA$^+$ myofibroblasts are derived from the intermediate fibroblast population identified by LungMAP single cell RNA-Seq and that these cells undergo de novo activation of PDGFRα expression after E16.5 (Lindahl et al., 1997; Ntokou et al., 2015; Guo et al., 2016, 2015).

Contrary to the proximal association with bronchiolar epithelium at E16.5, PDGFRα-GFP$^+$/αSMA$^+$ fibroblasts associate with alveolar epithelial cells distally at PN7 (insert Fig. 1C, Fig. 4). These fibroblasts transiently differentiate into contractile αSMA positive myofibroblasts during active septation, but once septation is complete, αSMA expression does not persist. To identify differences between proximal E16.5 smooth muscle fibroblasts (E16.5) and distal PDGFRα-GFP$^+$/αSMA$^+$ myofibroblasts (PN7), we interrogated gene profiles consisting of high gene expression at the time of rapid alveolarization (PN7) and contrasted them to high expression at the time of bronchiolar smooth muscle development (E16.5). Profiles #1 and #18 identify genes with high expression at E16.5 and PN7 and predict a contractile cellular function. A list of 676 actin-related genes was obtained by Ingenuity Pathway Analysis (IPA). Venn diagram comparison of these actin-related genes with the 1908 differentially expressed genes identified 147 genes associated with muscle contractile function that are differentially expressed at PN7 compared to E16.5 and, thus, distinguish the intermediate myofibroblast program of the interstitial PDGFRα$^+$ fibroblast at PN7 from the E16.5 peribronchiolar smooth muscle program. (Fig. 6 K, Data in Brief Table 4). Persistence and overabundance of myofibroblasts in the alveolar interstitium, because of premature birth or perinatal hyperoxia exposure, is found in association with arrested alveolar septation in bronchopulmonary dysplasia (BPD) (Ahlfeld et al., 2013; Bozyk et al., 2012; Toti et al., 1997). Further analysis of Profile #18, which contains 66 genes that are uniquely upregulated on PN7, in combination with gene expression studies from BPD samples and other septation defects will provide valuable insight into the regulatory networks that result in fibroblast differentiation to smooth muscle cells or transient myofibroblasts.

Profile #1 contains 459 genes, including 75 transcription factors, which are up-regulated at E16.5 and PN7 (Fig. 6D). The up-regulated genes comprise unique enriched functions in cell cycle, cell projection and muscle cell differentiation (Table 1). One hundred fourteen out of the 1908 differentially expressed genes were associated with cell cycle and their expression was up at E16.5 and PN7 (Data in Brief Table 2). Profile #18 contains 66 genes, including 6 transcription factors, which are uniquely up-regulated at PN7. The unique gene enrichment

Table 1
Gene Enrichment results of 7 major co-regulated gene profiles.

Profile	Genes	TF	Pattern	Unique Representative term
#1	459	75	Peak on E16. and PN7	• Cytoskeletal part (76 genes) • Negative regulation of signaling (62 genes) • Cell projection morphogenesis (57 genes) • Cell division (45 genes) • Wnt signaling pathway (37 genes) • Regulation of cell projection organization (34 genes) • Negative regulation of cell cycle (34 genes) • Apoptotic signaling pathways (28 genes) • Regulation of mitotic cell cycle phase transition (27 genes) • Cellular response to FGF stimulus (26 genes) • Muscle cell differentiation (26 genes) • EGFR signaling pathway (25 genes)
#39	271	18	Increase E16.5 to PN28	• Regulation of cell motility (42 genes) • Regulation of anatomical structure morphogenesis (35 genes) • Calcium ion binding (33 genes) • Tube development (29 genes) • Peptidase regulator activity (18 genes) • Cell matrix adhesion (18 genes) • Renal development (18 genes)
#13	251	24	Dip at E18.5	• Vesicle mediated transport (56 genes) • Membrane organization (46 genes) • Golgi membrane (26 genes) • Establishment of organelle localization (25 genes) • Response to ER stress (23 genes) • Glycosylation (21 genes) • Endosomal part (21 genes) • Actin filament bundle assembly (10 genes) • Stress fibers (8 genes)
#20	151	14	High at E16.5	• Mitochondrial part (35 genes) • Oxidation reduction process (34 genes) • Nucleolus (27 genes) • ATP metabolism (22 genes) • Vegf signaling (13 genes) • NIK/NfkappaB signaling (7 genes)
#10	104	10	Decrease E16.5 to PN28	• Immune system development (23 genes) • Positive regulation of Cell death (19 genes) • Chemotaxis (18 genes) • TNF production (8 genes)
#18	66	6	Spike at PN7	• Positive regulation of Cell Diff. (14 genes) • Muscle contraction (14 genes) • Proteinaceous ECM (10 genes) • Tube morphogenesis (9 genes) • Cellular response to TGFβ (8 genes) • Focal adhesion (7 genes) • ECM disassembly (7 genes) • Basement membrane (7 genes)
#23	49	4	Increase PN7 to PN28	• Positive regulation of cell proliferation (11 genes) • Regulation of lipid metabolic process (7 genes) • Abnormal pulmonary circulation (5 genes) • Negative regulation of ER-stress induced apoptotic signaling (3 genes)

categories of profile #18 are cell differentiation, tube morphogenesis, response to TGFβ signaling, and muscle contraction (Fig. 6E).

We compiled a list of 288 PDGFR signaling pathway genes by combining GO categories for “Signaling by PDGF”, “PDGFR interactions”, “PDGFR signaling path” and “PDGFR binding”. Out of these 288 genes 48 were significantly differentially expressed at one of the time points (Fig. 6L).

2.13. Transcription factor promoter enrichment analysis

To predict potential transcriptional regulators of the gene profiles that we identified, we used two complementary computational approaches. For each profile, we identified all genes encoding transcription factors (TFs), and assessed the enrichment of their (1) predicted binding sites or (2) ChIP-Seq peaks, in the promoter regions of all genes contained within that profile (see Methods). The first approach identified 9 enriched TF binding site motifs in profiles #1 and #13, 2 enriched motifs in profile #10 and 1 motif in #18, #23 and #39 (Data in Brief Table 5). Several of these TFs are involved in adipogenesis, fibrosis, tumor progression, smooth muscle differentiation, or cell-matrix adhesion. However, a role in lung development has only been demonstrated for NFIA and NFIX (Hsu et al., 2011).

The second approach identified three TFs whose ChIP-Seq peaks significantly overlap the promoters of the genes contained in a particular profile: RE1-silencing transcription factor (REST/NRSF) in profile 1, CCCTC-binding factor (CTCF) in profile 1, and myc-associated factor X (MAX) in profile 13 (Data in Brief Table 6 and 7). Of note, CTCF and MAX were also implicated in the first computational approach.

The NRSF ChIP-Seq dataset that overlaps 207 of the 465 gene promoters in profile 1 has been performed in murine C2C12 myoblasts, that were differentiated into myocytes in vitro (GEO:GSM915175). Profile #1 has the two peak expression levels at E16.5 and PN7, which are the time points of smooth muscle and interstitial myofibroblast differentiation. While PDGFRα$^+$ fibroblasts are not C2C12 myoblasts, our analysis infers that both cell types use a similar gene expression program for muscle and myofibroblast differentiation.

The CTCF ChIP-Seq dataset that overlaps with 286 of the 465 gene promoters in profile 1 has been performed in murine erythroleukemia cells (MEL) (GEO:GSM918744). MEL have been extensively studied for their ability to specifically bind fibronectin, and lose adhesion during cellular differentiation (Fraser and Gordon, 1994; Patel and Lodish, 1986). PDGFRα$^+$ fibroblasts produce and bind fibronectin and change fibroblast phenotypes dependent on matrix adherence. An overlap with 286 genes of the CTCF ChIP-Seq dataset in MEL cells with the expression of genes in PDGFRα$^+$ fibroblasts imply a cellular function for these genes in fibronectin interaction and cell adhesion.

Among the 207 NRSF and 286 CTCF genes, 144 genes were common (Data in Brief Table 7). Gene ontology analysis using Toppcluster identified “cell cycle” (32 genes), “non-canonical Wnt signaling” (10 genes), and “cell junctions” (29 genes) as significantly enriched biological processes and cellular compartments among these shared genes. In the 142 CTCF-specific genes, “mitotic cell cycle” (25 genes) were identified, and in the 63 NRSF-specific genes “regulation of cell morphogenesis” (11 genes) and “regulation of anatomical structure morphogenesis” (14 genes) were the most strongly enriched. “Interaction with fibronectin” was another strongly enriched term in both the CTCF-bound gene list (21 genes) and NRSF-bound gene list (21 genes).

The MAX ChIP-seq experiment has been performed in murine B-cell lymphoma cells (GEO:GSM912908). 192 of the 257 profile #13 gene promoters overlap with this dataset (Data in Brief Table 7). Gene ontology analysis using ToppFun identified the ER as the major cellular component (59 genes); “ER to Golgi vesicle-mediated transport” as one of the most significant biological process (18 out of 178 genes in the annotation), and “protein processing in ER” (18/167),

Table 2
Gene enrichment results of shared gene ontology terms.

Profile	Trend	Representative Terms
#10, #39	inverse correlation	• Regulation of response to wounding (#10: 19 genes; #39: 22 genes) • Cell migration (#10: 22 genes; #39: 37 genes) • Negative regulation of cell migration (#39: 16 genes)
#10, #23	inverse correlation	• Negative regulation of cell death (#23:15 genes, #10: 20 genes)
#1, #10, #20, #23	Down between E16.5 - E18.5	• Negative regulation of G1/S transition (#1 :15 genes; #20: 9 genes) • Canonical Wnt signaling pathway (#1: 29 genes; #20: 12 genes) • Negative regulation of cell death (#1: 53 genes; #23: 15 genes; #10: 19 genes; #20: 24 genes)
#1, #39, #18	Genes up at PN7	• Response to growth factor (#1: 50 genes; #39: 36 genes) • Positive regulation of signal transduction (#1: 75 genes; #39: 46 genes) • Actin cytoskeleton (#1: 28 genes; #39: 21 genes) • Cell migration (#18: 11 genes: #39: 41 genes) • Angiogenesis (#18: 12 genes: #39: 26 genes) • ECM organization (#18: 12 genes: #39: 30 genes) • Tissue morphogenesis (#18: 13 genes: #39: 29 genes) • Epithelium development (#18: 15 genes: #39: 41 genes)
#1, #39, #13, #18	Up between E18.5 - PN7	• Cell junction: #1: 54 genes; #39: 51 genes;#13: 32 genes, #18: 17 genes
#23, #39	Increase PN7-PN28	• Coexpression Atlas GudMAP: renal capsule (#23: 13 genes; #39: 29 genes)

"fatty acid elongation" (6/23) and "lysosome" (10/122) as major KEGG pathways.

2.14. Temporal changes in the composition of PDGFRα fibroblast population

To complement conclusions from gene expression data about functional differences and proliferation in PDGFRα$^+$ cell populations at different stages throughout lung development, we performed flow cytometry of lung single-cell suspensions at E16.5, E18.5, PN7 and PN28 (Figs. 7 and 8). Contrary to previous approaches that have utilized differential adherence to enrich for lung fibroblasts prior to flow cytometry analysis (Chen et al., 2012; Kimani et al., 2009; McGowan and McCoy, 2011), we utilized a whole-lung single-cell suspension and gated for CD45$^-$, CD326$^-$, CD31$^-$, PDGFRα$^+$(CD140a$^+$) fibroblasts (Fig. 7A). Compared to a whole-lung approach, isolating fibroblasts by differential adherence results in a 10% loss of all PDGFRα-GFP$^+$ fibroblasts and reduces by 13% the relative contribution of PDGFRα-GFP with high GFP expression (Data in Brief Fig. 2).

Temporal changes in the overall composition of the single cell digest were assessed in PDGFRα$^+$(CD140a$^+$) and CD140a^{neg} stromal cells, leukocytes (CD45$^+$), epithelial cells (CD326$^+$), and endothelial cells (CD31$^+$) (Fig. 7B). Consistent with our PDGFRα-GFP immunofluorescence data, the contribution of PDGFRα$^+$(CD140a$^+$) positive cells to total lung cells was more than 50% at E16.5, declined substantially to less than 10% by E18.5 and PN7, and contributed ~5% of total lung cells at PN21 and PN28 (Fig. 7B, E–H). Flow cytometry analysis of analysis of single cell suspensions from *Pdgfra^{GFP}* lungs corroborated these findings (Fig. 1H). CD140^{neg} stromal cells peaked at E18.5, at which time they represented ~30% of total lung cells, but were around 10% at all other time points. Leukocytes and epithelial cells comprised 30–50% of all total lung cells from E18.5 to PN28. Endothelial cells comprised ~10% of the total lung cell population throughout lung development (Fig. 7A and B).

The PDGFRα$^+$ fibroblast population was further analyzed for composition of myo- and lipofibroblasts by using flow cytometry for CD140a$^+$, αSMA and lipidTOX (Fig. 7C, D). At PN7 over 75% of PDGFRα$^+$ fibroblasts were myofibroblasts (CD140a$^+$/αSMA$^+$), which is in accordance with immunofluorescence staining and the enrichment for expression of genes with contractile function (Figs. 1C, 4, 6K). Although not discernible at other time points, at PN7 PDGFRα-GFP$^+$ fibroblasts could be differentiated in PDGFRα-GFP^{dim} and PDGFRα-GFP^{bright} subpopulations. Notably, the majority of lipidTOX$^+$ lipofibroblasts were PDGFRα-GFP^{dim} and αSMA$^+$ myofibroblasts were PDGFRα-GFP^{bright} (Data in Brief Fig. 3).

PDGFRα$^+$(CD140a$^+$/lipidTOX$^+$) lipofibroblasts were most abundant at PN7 but were only a minority fibroblast population when compared to the size of the myofibroblast population, which might explain why a lipofibroblast signature at PN7 was not readily detectable (Fig. 7D, G). To identify lipofibroblast distribution, we performed cytohistochemistry for Oil-red-O to assess lipid droplets in PDGFRα-GFP$^+$ cells and immunohistochemistry for ADFP staining (Figs. 4D, 5D, 7E–H). Based on flow cytometry not all myo- and lipofibroblasts were PDGFRα$^+$(CD140a$^+$), as αSMA and lipidTOX were also found in CD140^{neg} stromal cells (Data in Brief Fig. 4E). These findings are in accordance with the gene expression data, which indicated postnatal lipofibroblast differentiation.

To validate the prediction of increased cell-cycle activity at E16.5 and PN7, we assessed cell proliferation by flow cytometry using Ki67 expression. From E16.5 until PN7, PDGFRα$^+$ fibroblasts demonstrated high rates of proliferative activity (> 50%), which is subsequently reduced dramatically by PN21 and PN28 (Fig. 7D, Data in Brief Fig. 4D). These findings are in accordance with previously published data on fibroblasts isolated by differential adherence (Kimani et al., 2009).

Proliferation was also assessed in non-stromal cell populations. Peak proliferation of CD45$^+$ cells occurred between E16.5 and E18.5, while peak proliferation in CD326$^+$ epithelial and CD31$^+$ endothelial cells occurred at E16.5 (Data in Brief Fig. 4A–C). These data demonstrate that different cell compartments proliferate at different developmental stages, which adds an additional layer of complexity to gene expression analysis from whole lungs versus preparations enriched for specific cell populations.

Expression of CD90 (Thy-1), CD73 (Nt5e), CD105 (Endoglin), CD44, and CD146 (Mcam), and have previously been described in the PDGFRa-GFP expressing fibroblast lineage (Ntokou et al., 2015). Relative mRNA expression of these potential progenitor markers was extracted from our RNA-Seq expression analyses and demonstrate a dynamically changing expression pattern as previously described. Relative mRNA expression of CD34 and CD29 (Itgb1) was two to four-fold higher than expression of these previously-described fibroblast markers and was significantly different at the peak of interstitial myofibroblast differentiation at PN7 (Fig. 7I, J).

2.15. Immunophenotyping PDGFRα$^+$ fibroblasts using CD34 and CD29

Although PDGFRα$^+$ myo- and lipo- fibroblasts can be identified by

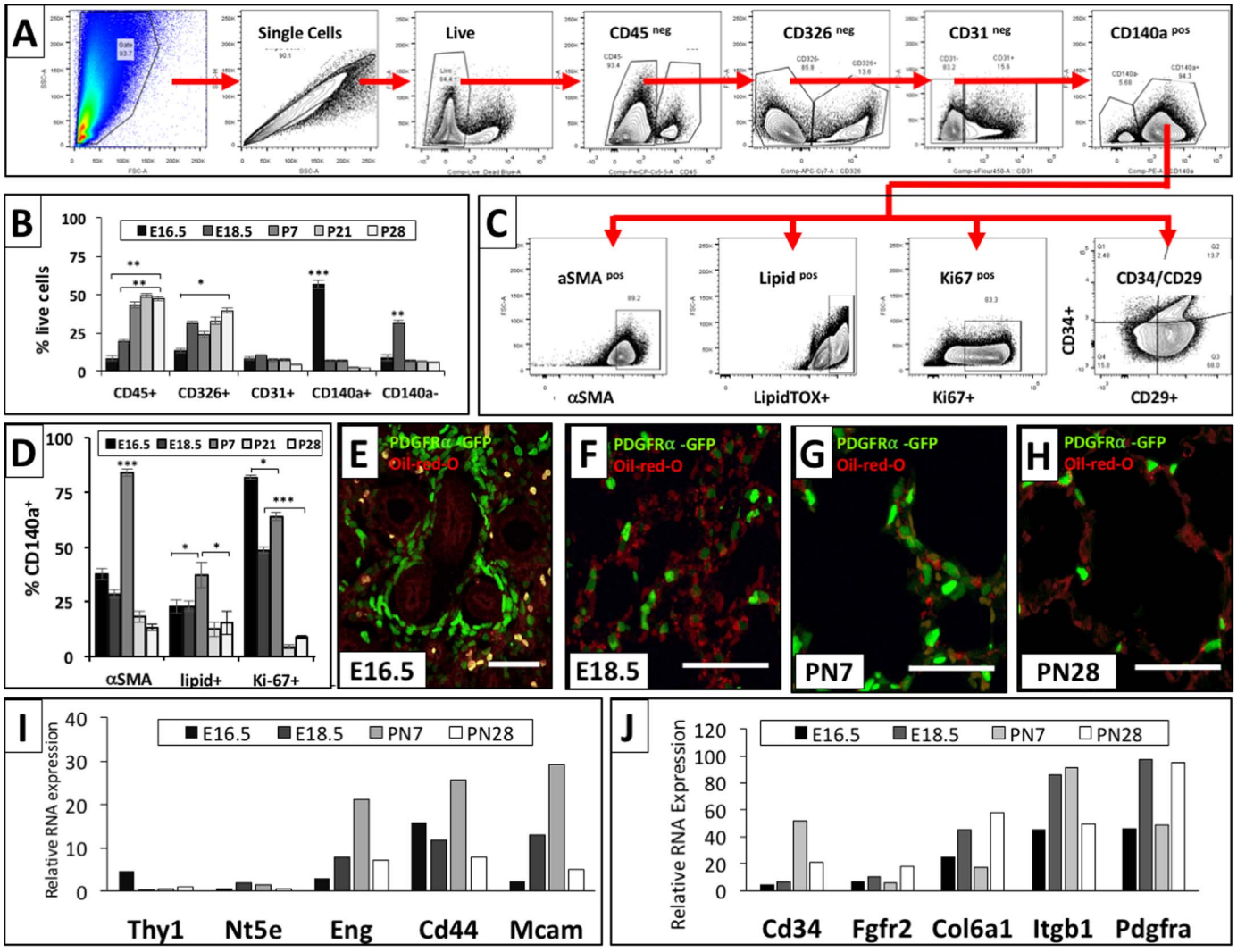

Fig. 7. Immunophenotyping of PDGFRα$^+$ fibroblasts across the developmental stages by flow cytometry. Single cell suspension of lungs at E16.5, E18.5, PN7, PN21 and PN28 were stained for viability and the different cell lineage markers (CD45, CD31, CD326, CD140a) and further characterized for αSMA expression, lipid content, proliferation, and expression of CD34 and CD29. (A, C) Panels show the representative gating strategy for different lung cell subpopulations at E16.5. (B) Bar graph of the proportion of different cell lineages across the developmental stages indicated. (D) Bar graph of αSMA expression, lipid content, and proliferation of CD140a$^+$ cells at different developmental stages. (E–H). The spatial - temporal distribution of PDGFRα-GFP$^+$ fibroblasts (green) with lipids (red: oil-red O stain) was analyzed at E16.5 (E), E18.5 (F), PN7 (G) and PN28 (H). (I, J) Relative RNA expression of mesenchymal progenitor and cell specific markers, extracted from RNA-Seq data. Bar graph data was expressed as mean ± SEM and comparisons among groups were made by ANOVA, with $p < 0.05$ considered significant. All scale bars are 50 μm.

staining with αSMA and lipidTOX, respectively, there is a population of interstitial PDGFRα$^+$ fibroblasts that are neither myo- nor lipofibroblasts. In a re-alveolarization model following partial pneumonectomy, we previously identified αSMA/lipidTOX -negative matrix fibroblasts that express CD34. In contrast, expression of CD29 (β1 integrin) identified proliferative myofibroblasts. Our previously published realveolarization data suggested that PDGFRα$^+$/CD29$^+$/CD34$^+$ fibroblasts contribute to the transient re-emergence of αSMA$^+$ myofibroblasts as they lose CD34 expression (Green et al., 2016). We performed flow cytometry to determine whether similar functional differences within PDGFRα$^+$(CD140a$^+$) fibroblasts are associated with CD29 and CD34 during lung development. Our comparison of CD140a$^+$ fibroblasts isolated by lung digest versus differential plate-down revealed a loss of CD29+ cells due to trypsinization after in vitro culture (Brown et al., 2007) (Data in Brief Fig. 4).

At E16.5 and E18.5, the majority of PDGFRα$^+$(CD140a$^+$) fibroblasts are CD29$^+$(Fig. 8A). Starting at PN7 there is a shift toward CD29$^+$/34$^+$ co-expression followed by a loss of both CD29 and CD34 by PN28 (Fig. 8A). To determine whether CD29 or CD34 are associated with proliferation, αSMA or lipid we analyzed CD140a$^+$/Ki67$^+$, CD140a$^+$/αSMA$^+$ and CD140a$^+$/lipidTOX$^+$ cells for the presence of CD29 and CD34 (Fig. 8B–D). Prenatally, the majority of CD140a$^+$ fibroblasts express CD29. At PN7 and PN21 the percentage of cells that express both CD29 and CD34 increases to ~20%. The percentage of cells that express neither CD29 nor CD34 also increases from PN7 to PN28. At both E16.5 and PN7, the majority of proliferative CD140a$^+$ cells express CD29, but neither CD29 nor CD34 is expressed by PN21 and PN28 (Fig. 8B). The majority of CD140a$^+$/αSMA$^+$ cells express CD29 at all time points (Fig. 8C). CD140a$^+$/lipidTOX$^+$ cells were only found at the postnatal time points (Fig. 7G, H). At PN7, CD140a$^+$/lipidTOX$^+$ cells express CD29, at PN21 they co-express both CD29 and CD34, and at PN28 they express neither (Fig. 8D). These data suggest that both peribronchiolar and alveolar myofibroblasts are proliferative and express CD29. Although there was no evidence that myo- and lipofibroblasts express CD34 alone, at PN21 there is a significant peak of CD29/CD34 co-expression in lipofibroblasts.

3. Discussion

PDGFRα$^+$ lung fibroblasts are essential for alveolarization and

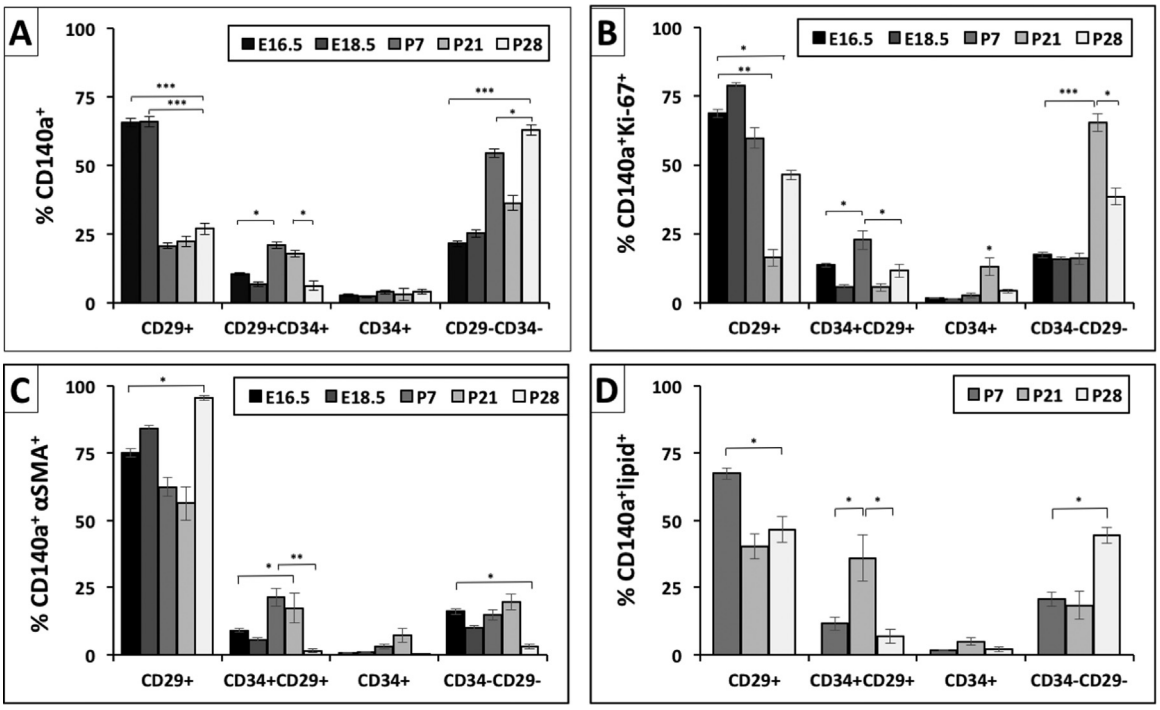

Fig. 8. Characterization of subpopulations of CD140a fibroblasts throughout lung development. (A) Subpopulation of CD140a$^+$ fibroblasts were grouped into 4 subgroups based on surface expression of CD29 and/or CD34. (B) Proliferative, (C) αSMA positive, (D) lipid containing CD140a fibroblasts were analyzed for the expression of CD29 and CD34. Bar graph data was expressed as mean ± SEM and comparisons among groups were made by ANOVA, with $p < 0.05$ considered significant * < 0.05, ** < 0.01, *** < 0.001.

realveolarization (Green et al., 2016; Bostrom et al., 2002, 1996; Lindahl et al., 1997; Chen et al., 2012; Perl and Gale, 2009). As early as E12.5, a population of PDGFRα-expressing, SHH-responsive *Gli1*+ mesenchymal cells can be traced throughout postnatal lung development and contribute to nearly all of the alveolar myofibroblasts required for septation (Li et al., 2015; Liu et al., 2013). During the canalicular phase, PDGFRα$^+$ fibroblasts are associated with proximal airway structures, whereas during saccular and alveolar lung development, PDGFRα$^+$ fibroblasts are found distally within the walls of the developing distal airspaces where they contribute to both αSMA$^+$ myofibroblasts and αSMA$^-$ lipofibroblasts (Kimani et al., 2009; McGowan et al., 2008; Ntokou et al., 2015; McGowan and McCoy, 2011, 2013, 2014). Finally, although PDGFRα$^+$ fibroblasts are still present in significant numbers toward the end of alveolarization, at P28, few express αSMA or contain lipid. Thus, in a relatively short period of time, PDGFRα$^+$ interstitial fibroblasts undergo dynamic changes in location and function. *Pdgfra*^{GFP} and *Pdgfra*^{Cre} transgenic animals have been used to follow the cell fate of PDGFRα$^+$ fibroblasts during alveolarization and alveolar reseptation. However, little is known about the diversity of fibroblast phenotypes within the PDGFRα$^+$ population at specific time points throughout distal lung development (Chen et al., 2012; Perl and Gale, 2009; Kimani et al., 2009; McGowan et al., 2008; McGowan and McCoy, 2011). This study identifies temporal, spatial, and functional differences in the PDGFRα$^+$ fibroblast populations and integrates these with temporal changes in the transcriptome of isolated PDGFRα$^+$ fibroblasts from E16.5 to PN28. Our data demonstrate that PDGFRα$^+$ fibroblasts comprise a heterogeneous mesenchymal population, consisting of dynamically changing proportions of airway smooth muscle cells, interstitial myo- and lipofibroblasts and an underappreciated population of interstitial matrix fibroblasts.

A major finding in this spatial-temporal analysis was that PDGFRα$^+$ fibroblasts switch from early association with the proximal airways, to association with the terminal saccules and acini in the saccular and alveolar phase. In the canalicular phase, the PDGFRα$^+$ fibroblast population consists of αSMA positive and negative populations. The more proximal αSMA positive cells are associated with SOX2 positive bronchiolar epithelial cells, whereas αSMA negative cells are more distal and associated with the HOPX positive bipotent embryonic alveolar progenitors in the acinar tubules. Transcriptome analysis and computational promoter analysis of the E16.5 population identified a predominant gene signature that regulates smooth muscle differentiation as well as cellular interaction with fibronectin that regulates cell migration. Immunophenotyping and functional analysis by flow cytometry identified the E16.5 PDGFRα$^+$ fibroblast population as highly proliferative. Additionally, 75% of all PDGFRα$^+$(CD140a$^+$) proliferative cells expressed both CD29$^+$ and αSMA, while the remaining 25% were CD29, αSMA and lipid negative. These findings are in accordance with previously published lineage tracing studies confirming that the peribronchiolar population will give rise to postnatal peribronchiolar smooth muscle (Ntokou et al., 2015). However, the lineage fate of αSMA negative PDGFRα$^+$ fibroblasts and their role in the epithelial crosstalk with the HOPX positive bipotent embryonic alveolar progenitors in the transition zone (between the bronchioles and acinar tube region) remains unclear.

Alveolarization, whether during initial development or in response to pneumonectomy, occurs in two phases. During the initial phase of septal eruption (~P4-7) there is an abrupt increase in αSMA expression organized along the developing septal crest (Green et al., 2016; Chen et al., 2012; Branchfield et al., 2016; McGowan et al., 2008; Ntokou et al., 2015; Mariani et al., 2002). Following eruption, there is a phase of septal elongation and maturation in which αSMA expression is dramatically reduced, elastin expression and organization proceeds, and there is thinning of the alveolar interstitium (Branchfield et al., 2016; McGowan and McCoy, 2011; Mariani et al., 2002). This transient differentiation of PDGFRα$^+$ fibroblasts into contractile alveolar myofibroblasts from PN4-7 is in sharp contrast to the smooth muscle differentiation in peribronchiolar PDGFRα fibroblasts at E16.5. To address the molecular differences between permanent smooth muscle differentiation and transient myofibroblast differentiation, we performed Gene Ontology analyses and identified “muscle cell differentiation” at E16.5 and “muscle contraction” at PN7 as unique ontology

terms with 26 and 14 differentially expressed genes. The importance of a WNT-Tenascin-PDGFRα signaling cascade has been previously described for pulmonary smooth muscle precursors (Cohen et al., 2009). While the Wnt signaling, pathway was significantly represented at both time points, expression of Wnt2 peaked at E16.5 and Wnt5a at PN7. It has been previously published that peribronchiolar fibroblasts and interstitial fibroblasts are derived from mesenchymal lineages that express Fgf10, Tbx4, and Gli1, while Wnt2 only gives rise to peribronchiolar fibroblasts (Li et al., 2015; Mailleux et al., 2005; Peng et al., 2013; Xie et al., 2016). We have recently demonstrated that increased PDGFRα kinase activity increases Wnt5a expression and promotes matrix fibroblast differentiation, while Wnt2 expression is associated with alveolar myofibroblast differentiation (Green et al., 2016). Taken together these data suggest that Wnt2 is important for permanent muscle differentiation while Wnt5a promotes a switch from myo- to matrix fibroblasts. The transition from myo- to matrix fibroblasts occurs during the phase of septal elongation and maturation, during which we demonstrated, by both immunofluorescence and flow cytometry, a dramatic decrease in αSMA expression while elastin expression and ECM organization proceeds (Branchfield et al., 2016; McGowan and McCoy, 2011; Mariani et al., 2002). We have previously demonstrated that inhibition of FGFR2 signaling between E16.5 and E18.5 did not disturb peribronchiolar smooth muscle differentiation but blocked myofibroblast differentiation and postnatal alveolarization (Hokuto et al., 2003). In contrast FGFR3/4 double knock out mice have increased myofibroblasts and elastin deposition (Srisuma et al., 2010; Weinstein et al., 1998). Collectively these data suggest that peribronchiolar smooth muscle and interstitial myo- and matrix fibroblast differentiation are both regulated by complex, synergistic signaling of Wnt, Fgf and Pdgf, through the use of distinct ligands and receptors. Further understanding of the signaling partners, on a cellular level (e.g., alveolar versus bronchiolar) and on a molecular level (e. g., Wnt2 versus Wnt5a, FGFR2 versus FGFR3/4) could provide new insights into the pathogenesis of abnormal alveolarization and aberrant alveolar repair.

Computational analysis of our co-expressed gene profiles identified 3 transcription factors, NRSF, CTCF and MAX, which have been demonstrated by ChIP-seq to bind a significant number of promoter sequences of genes within these profiles. NRSF/REST is a component of a transcriptional repressor complex and is involved in negative regulation of mesenchymal stem cell differentiation (GO: 0017053, GO: 2000740). CTCF has epigenetic regulatory functions and maintains DNA methylation (GO: 0040030, GO: 0010216); in addition, SMAD proteins localize to CTCF binding sites in a CTCF dependent manner, which connects TGF-β/BMP/SMAD with CTCF and myofibroblast differentiation (Van Bortle et al., 2015). One of the genes CTCF binds is REST. REST and CTCF bind 349 promoter sequences out of 459 of the genes in the co-expression profile #1, that peaks at E16.5 and again at PN7. Gene-ontology analysis of these REST and CTCF target genes identified “focal adhesions” (32 of 393 genes) as a cellular component and coexpression with the “16.5 embryonic lung matrix fibroblast intermediate fibroblast cluster” (10 of 32 genes). The third transcription factor that has been demonstrated by ChIP-Seq to bind a significant number of genes within our profiles is MAX. MAX is induced upon response to starvation and heterodimerizes with MYC. The MYC/MAX complex has the ability to stimulate the transcription of genes that are epigenetically poised for transcription and to amplify the transcription of actively transcribed genes (Link and Hurlin, 2015). Gene-ontology analysis of the MAX target genes identified 56 genes associated with the ER as a cellular component and coexpression with the “16.5 embryonic lung proliferative and migrating myofibroblast cluster c3 and c1″ (10/95 and 9/71). These computational analyses suggest that REST and CTCF regulate cellular adhesion and matrix fibroblast differentiation, while MAX promotes myofibroblast differentiation and migration. Immunohistochemistry for cell-type specific expression of REST, CTCF and MAX were inconclusive, since these transcription factors are expressed in many other non-stromal cell types (data not shown). Further analysis of these transcription factors and their implicated target gene expression in specific fibroblasts, will provide new insights into the differentiation of myo- and matrix-fibroblasts, which can be used to understand excessive myofibroblast differentiation and extracellular matrix deposition in lung fibrosis.

Fibroblasts are an exceptionally heterogeneous and plastic cell population, and the lack of specific markers for isolation and characterization of fibroblasts has made it difficult to study their role in lung development and disease. GFP knock in mice and a few transgenic mice expressing Cre have been used to shed light on the lineage relationships within fibroblast populations. The *Pdgfra*^{GFP} knock in mice have been used to study fibroblasts during alveolarization, with the assumption that nuclear GFP expression correlates with *Pdgfra* gene expression (Hamilton et al., 2003; McGowan et al., 2008; Perl and Gale, 2009). However, due to the stability of the GFP protein beyond the period of *Pdgfra* gene transcription and/or receptor expression, PDGFRα receptor expression and nuclear GFP do not overlap, thus complicating interpretation of some results (Green et al., 2016). Moreover, PDGFRα-GFP labeled cells have been distinguished into GFP bright and dim populations using flow cytometry but these results depend heavily on the cell isolation protocol used. A standard procedure to isolate fibroblasts is by differential plating of a single cell digest. In this study, we show that a proportion of the PDGFRα$^+$ fibroblasts and PDGFRα-GFP$^+$ fibroblasts do not adhere within the standard one hour of plating and will thus be lost in subsequent flow cytometry and transcriptome analysis. This confounding artifact needs to be considered when comparing studies with variations in the fibroblast isolation methods. To further characterize and distinguish interstitial fibroblasts, we have previously screened several mesenchymal stromal cell markers and identified integrin β1 (CD29) and CD34 expression as a way to differentiate between the myofibroblasts and matrix/lipofibroblasts, respectively (Green et al., 2016). While CD29 is being expressed on 70% of the Tbx4-lineage fibroblasts, CD34 is only expressed on very few Tbx4-lineage fibroblasts (Xie et al., 2016). In a sham operated mature adult lung, lipid-containing PDGFRα$^+$ fibroblasts are CD29$^+$/CD34$^+$ or CD34$^+$, and few PDGFRα$^+$/CD29$^+$ fibroblasts express αSMA (Green et al., 2016). During the eruption phase of re-alveolarization following partial pneumonectomy, we previously demonstrated that double positive CD29$^+$/34$^+$ and single positive CD29$^+$ PDGFRα$^+$ fibroblasts transiently proliferate and shift from dual C–D29$^+$/34$^+$ to single CD29$^+$ (Green et al., 2016). CD34 is a cell surface marker that is expressed by a broad range of cells including hematopoietic, stromal, epithelial, and endothelial cells. In all tissues the CD34$^+$ population has a higher differentiation capacity and loss of CD34$^+$ results in differentiation, suggesting a lineage commitment of the progenitor cell. We concluded that co-expression of PDGFRα$^+$, CD29$^+$, and CD34$^+$ defined a reservoir of matrix/lipofibroblasts that, following injury, contribute to the transient re-emergence of αSMA$^+$ myofibroblasts required for septal eruption (Green et al., 2016). During septal elongation, when elastin production and organization peaks (~PN7-28), expression of both CD29 and CD34 in PDGFRα$^+$ fibroblasts was gradually lost, suggesting that αSMA$^-$/lipid$^-$/CD29$^-$/CD34$^-$/PDGFRα$^+$ fibroblasts are losing their progenitor status as they commit to the elastin-producing matrix fibroblasts (Branchfield et al., 2016; Mariani et al., 2002).

While CD29 and CD34 may be markers that are useful in identifying progenitor populations and myofibroblast populations, they are not sufficient for the characterization of a cell type or a lineage relationship. The RNA-Seq data from this study will serve as a great resource to identify new markers and to further interrogate the differences of these fibroblast sub-populations at different stages of distal lung development. These datasets are publically available through the user-friendly and interactive the LungMAP website and Lung Gene Expression Analysis Web Portal (LGEA) (Du et al., 2017). URL: https://research.cchmc.org/pbge/lunggens/mainportal.html. Identifying

additional surface markers to further characterize and understand this heterogeneous and extremely plastic PDGFRα$^+$ fibroblast population will provide new insights into how a loss of this diversity and plasticity results in various pathogenesis of alveolar malformations and the repair and recovery of alveolarization following perinatal lung injury.

4. Methods

4.1. Materials and methods

Additional detail on the methods is provided in the Data in Brief.

4.2. Confocal microscopy

Lung tissues were harvested, fixed with 4% PFA in PBS and frozen, or paraffin embedded, sectioned and further processed. Tissue preparations were blocked with donkey or goat serum and incubated with anti-αSMA (Sigma-Aldrich, St. Louis, MO), SOX2, SOX9, TTF1, Endomucin, E-cadherin, HOPX (Santa Cruz, CA, USA), Pro-SPC and ADFP (Abcam, Cambridge, MA) followed by incubation with secondary antibodies conjugated to Alexa Fluor-488, 595 or 647 (Invitrogen, Carlsbad, CA) and counterstained with DAPI-containing mounting media. GFP signals were detected using a chicken polyclonal anti-GFP antibody (Abcam, Cambridge, MA). Data was analyzed by Imaris software, version 7.6.

4.3. Transgenic mice

B6.129S4-*PDGFRα$^{tm11(EGFP)Sor}$*/J herein designated *Pdgfra^{GFP}* (Hamilton et al., 2003): The endogenous PDGFRα promoter drives expression of the H2B-eGFP fusion gene.

4.4. PDGFRα$^+$ fibroblast sorting, RNA extraction and RNA-Seq

Lung tissue was harvested and digested into single cell suspension. PDGFRα$^+$ fibroblasts were isolated, using microbeads (Miltenyi Biotec MACS technology, Gladbach, Germany), by negative selection for CD45, CD326, and CD31 (to deplete leukocytes, epithelia, and endothelial cells, respectively) followed by positive selection for CD140a. Cells of 6–10 pubs per time point were pooled. RNA was extracted from MACS-sorted PDGFRα$^+$(CD140α) cells using RNeasy Mini Kit (No. 74104) (QIAGEN, Germantown, MD). RNA was quality checked, transcribed into complementary DNA using the Verso complementary DNA synthesis kit (AB-1453; LifeTechnologies, Carlsbad, CA), sheared, amplified, adaptor ligated and sequenced for single-end and pair-end RNA-Seq reads (Illumina Inc. San Diego, CA, USA).

4.5. Bioinformatics data analysis

RNA-Seq data was quantitated using TopHat and Cufflinks (Trapnell et al., 2012), normalized using trimmed means and genes were pre-filtered out if the expression level (FPKM) was less than 1 in all samples. Normalized data were subject to Bayesian Analysis of Time Series (BATS) using a normal error distribution model with a cutoff of Bayesian Factor < 0.01 for differentially expressed genes (Angelini et al., 2008). Differentially expressed genes were subject to pattern recognition using STEM (Ernst and Bar-Joseph, 2006). Gene set enrichment analysis was carried out using Toppgene (http://toppgene.cchmc.org) (Chen et al., 2009). Unique functional enrichments within each profile were identified and all profiles were compared to each other using Toppcluster (http://toppcluster.cchmc.org/) (Kaimal et al., 2010). Bayesian-based analytic tools that measure time-dependent differential gene expression require at least 5 time points - therefore three additional time points (E17, PN0 and PN18) were interpolated using the R script smooth.spline function (Bioconductor 3.3).

4.6. Transcription factor promoter enrichment analysis

We used two complementary computational methods to identify transcription factors (TFs) that might regulate the genes contained in each gene expression profile. The first method performs TF binding site motif enrichment analysis on the promoters of each profile. Promoters were defined as −200 to +100 relative to the transcription start site. To this end, we applied the HOMER motif enrichment algorithm (Heinz et al., 2010) and a large library of mouse position weight matrix binding models obtained from the CisBP database (Weirauch et al., 2014).

The second method compares the promoter regions of a profile with a large library of experimentally determined TF-genome interactions. We created the library by compiling 498 mouse ChIP-seq datasets from a variety of sources, including ENCODE (Consortium, 2012), Cistrome (Liu et al., 2011), and PAZAR (Portales-Casamar et al., 2009). As input, our method takes a set of genomic regions of interest (in this case, the promoter regions of the genes in a given profile) and systematically overlaps them with each ChIP-Seq dataset. The *observed overlap* of the input set with each ChIP-Seq dataset is calculated by counting the number of input regions it overlaps by at least one base. Next, a P-value describing the significance of this overlap is estimated using a simulation-based procedure. A distribution of *expected overlap* values is created from 1000 iterations of randomly choosing RefSeq gene promoters with the same length as the input set (e.g., if 50 promoters of length 300 bp are used as input, then 50 randomly chosen promoters of length 300 bp will be used in each simulation). The distribution of the expected overlap values from the randomized data resembles a normal distribution, and is used to generate a Z-score and P-value estimating the significance of the observed number of input regions that overlap each ChIP-Seq dataset. The resulting data thus provide ranked lists of datasets, based on enrichment for experimentally determined ChIP "peaks" located in the promoters of each gene set. From the results of our analysis all of these ChIP experiments were performed in mouse cell lines.

Conflict of interest statement

None of the authors have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Acknowledgement

We thank Susan Wert for electron microscopy and Sheila Bell for critical input on the manuscript. We want to thank the FACS Core and the "Nikon Core for Excellence" at CCHMC for technical assistance. This work was supported by NIH grants R01 HL 104003 01 (AKP), R56 HL 123969 (AKP), U01 HL122642 (LungMAP, AKP, YX), and R21 HG008186 (MTW).

References

Ahlfeld, S.K., et al., 2013. Periostin downregulation is an early marker of inhibited neonatal murine lung alveolar septation. Birth Defects Res. A Clin. Mol. Teratol. 97 (6), 373–385.
Angelini, C., et al., 2008. BATS: a Bayesian user-friendly software for analyzing time series microarray experiments. BMC Bioinforma. 9, 415.
Arora, R., Metzger, R.J., Papaioannou, V.E., 2012. Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system. PLoS Genet. 8 (8), (e1002866).
Barkauskas, C.E., et al., 2013. Type 2 alveolar cells are stem cells in adult lung. J. Clin. Investig. 123 (7), 3025–3036.
Barron, L., Gharib, S.A., Duffield, J.S., 2016. Lung pericytes and resident fibroblasts: busy multitaskers. Am. J. Pathol. 186 (10), 2519–2531.
Bertoncello, I., McQualter, J.L., 2013. Lung stem cells: do they exist? Respirology 18 (4), 587–595.
Bostrom, H., et al., 1996. PDGF-A signaling is a critical event in lung alveolar myofibroblast development and alveogenesis. Cell 85 (6), 863–873.
Bostrom, H., Gritli-Linde, A., Betsholtz, C., 2002. PDGF-A/PDGF alpha-receptor signaling is required for lung growth and the formation of alveoli but not for early

lung branching morphogenesis. Dev. Dyn. 223 (1), 155–162.
Bozyk, P.D., et al., 2012. Neonatal periostin knockout mice are protected from hyperoxia-induced alveolar simplication. PLoS One 7 (2), (e31336).
Branchfield, K., et al., 2016. A three-dimensional study of alveologenesis in mouse lung. Dev. Biol. 409 (2), 429–441.
Brewster, C.E., et al., 1990. Myofibroblasts and subepithelial fibrosis in bronchial asthma. Am. J. Respir. Cell Mol. Biol. 3 (5), 507–511.
Chen, F., Fine, A., 2016. Stem cells in lung injury and repair. Am. J. Pathol. 186 (10), 2544–2550.
Chen, J., et al., 2009. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 37, W305–W311.
Chen, L., et al., 2012. Dynamic regulation of platelet-derived growth factor receptor alpha expression in alveolar fibroblasts during realveolarization. Am. J. Respir. Cell Mol. Biol. 47 (4), 517–527.
Cohen, E.D., et al., 2009. Wnt signaling regulates smooth muscle precursor development in the mouse lung via a tenascin C/PDGFR pathway. J. Clin. Investig. 119 (9), 2538–2549.
Consortium, E.P., 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489 (7414), 57–74.
Du, Y., et al., 2015. 'LungGENS': a web-based tool for mapping single-cell gene expression in the developing lung. Thorax 70 (11), 1092–1094.
Du, Y., et al., 2017. Lung Gene Expression Analysis (LGEA): an integrative web portal for comprehensive gene expression data analysis in lung development. Thorax.
El Agha, E., et al., 2014. Fgf10-positive cells represent a progenitor cell population during lung development and postnatally. Development 141 (2), 296–306.
Ernst, J., Bar-Joseph, Z., 2006. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinforma. 7, 191.
Evans, M.J., et al., 1999. The attenuated fibroblast sheath of the respiratory tract epithelial-mesenchymal trophic unit. Am. J. Respir. Cell Mol. Biol. 21 (6), 655–657.
Eyden, B., 2008. The myofibroblast: phenotypic characterization as a prerequisite to understanding its functions in translational medicine. J. Cell Mol. Med. 12 (1), 22–37.
Fraser, I.P., Gordon, S., 1994. Murine erythroleukemia (MEL) cells bear ligands for the sialoadhesin and erythroblast receptor macrophage hemagglutinins. Eur. J. Cell Biol. 64 (2), 217–221.
Green, J., et al., 2016. Diversity of interstitial lung fibroblasts Is regulated by platelet-derived growth factor receptor alpha kinase activity. Am. J. Respir. Cell Mol. Biol. 54 (4), 532–545.
Guo, M., et al., 2015. SINCERA: a Pipeline for Single-Cell RNA-Seq profiling analysis. PLoS Comput. Biol. 11 (11), (e1004575).
Guo, M., et al., 2016. SLICE: determining cell differentiation and lineage based on single cell entropy. Nucleic Acids Res..
Hamilton, T.G., et al., 2003. Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. Mol. Cell Biol. 23 (11), 4013–4025.
Heinz, S., et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38 (4), 576–589.
Herriges, M., Morrisey, E.E., 2014. Lung development: orchestrating the generation and regeneration of a complex organ. Development 141 (3), 502–513.
Hogan, B.L., et al., 2014. Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. Cell Stem Cell 15 (2), 123–138.
Hokuto, I., Perl, A.K., Whitsett, J.A., 2003. Prenatal, but not postnatal, inhibition of fibroblast growth factor receptor signaling causes emphysema. J. Biol. Chem. 278 (1), 415–421.
Hsu, Y.C., et al., 2011. Mesenchymal Nuclear factor I B regulates cell proliferation and epithelial differentiation during lung maturation. Dev. Biol. 354 (2), 242–252.
Hung, C., et al., 2013. Role of lung pericytes and resident fibroblasts in the pathogenesis of pulmonary fibrosis. Am. J. Respir. Crit. Care Med. 188 (7), 820–830.
Husain, A.N., Siddiqui, N.H., Stocker, J.T., 1998. Pathology of arrested acinar development in postsurfactant bronchopulmonary dysplasia. Hum. Pathol. 29 (7), 710–717.
Jain, R., et al., 2015. Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. Nat. Commun. 6, 6727.
Kaimal, V., et al., 2010. ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic Acids Res. 38, W96–W102.
Kimani, P.W., et al., 2009. PDGF-Ralpha gene expression predicts proliferation, but PDGF-A suppresses transdifferentiation of neonatal mouse lung myofibroblasts. Respir. Res. 10, 119.
Lau, M., et al., 2011. Long-term failure of alveologenesis after an early short-term exposure to a PDGF-receptor antagonist. Am. J. Physiol. Lung Cell Mol. Physiol. 300 (4), L534–L547.
Li, C., et al., 2015. Progenitors of secondary crest myofibroblasts are developmentally committed in early lung mesoderm. Stem Cells 33 (3), 999–1012.
Lindahl, P., et al., 1997. Alveogenesis failure in PDGF-A-deficient mice is coupled to lack of distal spreading of alveolar smooth muscle cell progenitors during lung development. Development 124 (20), 3943–3953.
Link, J.M., Hurlin, P.J., 2015. The activities of MYC, MNT and the MAX-interactome in lymphocyte proliferation and oncogenesis. Biochim. Biophys. Acta 1849 (5), 554–562.
Liu, L., et al., 2013. Hedgehog signaling in neonatal and adult lung. Am. J. Respir. Cell Mol. Biol. 48 (6), 703–710.
Liu, T., et al., 2011. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 12 (8), R83.
Mailleux, A.A., et al., 2005. Fgf10 expression identifies parabronchial smooth muscle cell progenitors and is required for their entry into the smooth muscle cell lineage. Development 132 (9), 2157–2166.
Mariani, T.J., Reed, J.J., Shapiro, S.D., 2002. Expression profiling of the developing mouse lung: insights into the establishment of the extracellular matrix. Am. J. Respir. Cell Mol. Biol. 26 (5), 541–548.
McGowan, S.E., et al., 2008. Platelet-derived growth factor receptor-alpha-expressing cells localize to the alveolar entry ring and have characteristics of myofibroblasts during pulmonary alveolar septal formation. Anat. Rec. 291 (12), 1649–1661.
McGowan, S.E., McCoy, D.M., 2011. Fibroblasts expressing PDGF-receptor-alpha diminish during alveolar septal thinning in mice. Pediatr. Res. 70 (1), 44–49.
McGowan, S.E., McCoy, D.M., 2013. Platelet-derived growth factor-A and sonic hedgehog signaling direct lung fibroblast precursors during alveolar septal formation. Am. J. Physiol. Lung Cell Mol. Physiol. 305 (3), L229–L239.
McGowan, S.E., McCoy, D.M., 2014. Regulation of fibroblast lipid storage and myofibroblast phenotypes during alveolar septation in mice. Am. J. Physiol. Lung Cell Mol. Physiol. 307 (8), L618–L631.
McQualter, J.L., et al., 2009. Endogenous fibroblastic progenitor cells in the adult mouse lung are highly enriched in the sca-1 positive cell fraction. Stem Cells 27 (3), 623–633.
McQualter, J.L., et al., 2013. TGF-beta signaling in stromal cells acts upstream of FGF-10 to regulate epithelial stem cell growth in the adult lung. Stem Cell Res. 11 (3), 1222–1233.
Morrisey, E.E., Hogan, B.L., 2010. Preparing for the first breath: genetic and cellular mechanisms in lung development. Dev. Cell 18 (1), 8–23.
Naiche, L.A., et al., 2011. Identity and fate of Tbx4-expressing cells reveal developmental cell fate decisions in the allantois, limb, and external genitalia. Dev. Dyn. 240 (10), 2290–2300.
Ng, L.J., et al., 1997. SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. Dev. Biol. 183 (1), 108–121.
Ntokou, A., et al., 2015. Characterization of the platelet-derived growth factor receptor-alpha-positive cell lineage during murine late lung development. Am. J. Physiol. Lung Cell Mol. Physiol. 309 (9), L942–L958.
Patel, V.P., Lodish, H.F., 1986. The fibronectin receptor on mammalian erythroid precursor cells: characterization and developmental regulation. J. Cell Biol. 102 (2), 449–456.
Peng, T., et al., 2013. Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature 500 (7464), 589–592.
Peng, T., Morrisey, E.E., 2013. Development of the pulmonary vasculature: current understanding and concepts for the future. Pulm. Circ. 3 (1), 176–178.
Perl, A.K., et al., 2005. Normal lung development and function after Sox9 inactivation in the respiratory epithelium. Genesis 41 (1), 23–32.
Perl, A.K., Gale, E., 2009. FGF signaling is required for myofibroblast differentiation during alveolar regeneration. Am. J. Physiol. Lung Cell Mol. Physiol. 297 (2), L299–L308.
Popova, A.P., et al., 2014. Reduced platelet-derived growth factor receptor expression is a primary feature of human bronchopulmonary dysplasia. Am. J. Physiol. Lung Cell Mol. Physiol. 307 (3), L231–L239.
Portales-Casamar, E., et al., 2009. The PAZAR database of gene regulatory information coupled to the ORCA toolkit for the study of regulatory sequences. Nucleic Acids Res. 37, D54–D60.
Que, J., et al., 2009. Multiple roles for Sox2 in the developing and adult mouse trachea. Development 136 (11), 1899–1907.
Rockich, B.E., et al., 2013. Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proc. Natl. Acad. Sci. USA 110 (47), E4456–E4464.
Srisuma, S., et al., 2010. Fibroblast growth factor receptors control epithelial-mesenchymal interactions necessary for alveolar elastogenesis. Am. J. Respir. Crit. Care Med. 181 (8), 838–850.
Toti, P., et al., 1997. Bronchopulmonary dysplasia of the premature baby: an immunohistochemical study. Pediatr. Pulmonol. 24 (1), 22–28.
Trapnell, C., et al., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7 (3), 562–578.
Tsupykov, O., et al., 2016. Ultrastructural study of mouse adipose-derived stromal cells induced towards osteogenic direction. Microsc. Res. Tech. 79 (6), 557–564.
Ustiyan, V., et al., 2016. beta-catenin and Kras/Foxm1 signaling pathway are critical to restrict Sox9 in basal cells during pulmonary branching morphogenesis. Dev. Dyn. 245 (5), 590–604.
Van Bortle, K., et al., 2015. CTCF-dependent co-localization of canonical Smad signaling factors at architectural protein binding sites in D. melanogaster. Cell Cycle 14 (16), 2677–2687.
Volckaert, T., et al., 2011. Parabronchial smooth muscle constitutes an airway epithelial stem cell niche in the mouse lung after injury. J. Clin. Investig. 121 (11), 4409–4419.
Weinstein, M., et al., 1998. FGFR-3 and FGFR-4 function cooperatively to direct alveogenesis in the murine lung. Development 125 (18), 3615–3623.
Weirauch, M.T., et al., 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158 (6), 1431–1443.
Xie, T., et al., 2016. Transcription factor TBX4 regulates myofibroblast accumulation and lung fibrosis. J. Clin. Investig. 126 (9), 3626.
Zhang, W., et al., 2013. Spatial-temporal targeting of lung-specific mesenchyme by a Tbx4 enhancer. BMC Biol. 11, 111.