Decoding the IGF1 Signaling Gene Regulatory Network Behind Alveologenesis from A Mouse Model of Bronchopulmonary Dysplasia

Gao F¹, Li C¹, Smith SM¹, Peinado N¹, Kohbodi G¹, Tran E²,³, Loh E⁴, Li W⁵, Borok Z⁶ and Minoo P¹,⁷

1- Division of Neonatology, Department of Pediatrics, LAC+USC Medical Center and Children’s Hospital Los Angeles, Los Angeles, CA 90033, USA.
2- Norris Comprehensive Cancer Center, Keck School of Medicine, USC, Los Angeles, CA 90033, USA.
3- Department of Biochemistry and Molecular Medicine, Keck School of Medicine, USC, Los Angeles, CA 90033, USA.
4- Norris Medical Library, University of Southern California, Los Angeles, CA 90089, USA.
5- Department of Nephrology, Jiangsu Provincial Hospital of Traditional Chinese Medicine, Nanjing 210004, China.
6- Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA.
7- Hastings Center for Pulmonary Research, Keck School of Medicine of University of Southern California, Los Angeles, CA, 90033, USA.

*Authors for correspondence: fremontgao@gmail.com and minoo@usc.edu
Summary

Lung development is precisely controlled by underlying Gene Regulatory Networks (GRN). Disruption of genes in the network can interrupt normal development and cause diseases such as bronchopulmonary dysplasia (BPD)—a chronic lung disease in preterm infants with morbid and sometimes lethal consequences characterized by lung immaturity and reduced alveolarization.

Here, we generated a transgenic mouse exhibiting a moderate severity BPD phenotype by blocking IGF1 signaling in secondary crest myofibroblasts (SCMF) at the onset of alveologenesis. Using approaches mirroring the construction of the model GRN in sea urchin's development, we constructed the IGF1 signaling network underlying alveologenesis using this mouse model that phenocopies BPD. The constructed GRN, consisting of 43 genes, provides a bird’s-eye view of how the genes downstream of IGF1 are regulatorily connected. The GRN also reveals a mechanistic interpretation of how the effects of IGF1 signaling are transduced within SCMF from its specification genes to its effector genes and then from SCMF to its neighboring alveolar epithelial cells with WNT5A and FGF10 signaling as the bridge. Consistently, blocking WNT5A signaling in mice phenocopies BPD as inferred by the network. A comparative study on human samples suggests that a GRN of similar components and wiring underlies human BPD.

Our network view of alveologenesis is transforming our perspective to understand and treat BPD. This new perspective calls for the construction of the full signaling GRN underlying alveologenesis, upon which targeted therapies for this neonatal chronic lung disease can be viably developed.
Introduction

Development is precisely controlled by the genetic program encoded in the genome and is governed by genetic interactions (Davidson et al., 2002; Levine and Davidson, 2005). Elucidating the network of interactions among genes that govern morphogenesis through development is one of the core challenges in contemporary functional genomics research (Przybyla and Gilbert, 2021). These networks are known as developmental gene regulatory networks (GRN) and are key to understanding the developmental processes with integrative details and mechanistic perspectives. Over the past few decades, great advances have been made in decoding these networks in classical model systems (i.e. Dequeant and Pourquie, 2008; Longabaugh et al., 2017; Olson, 2006; Satou et al., 2009; Sauka-Spengler and Bronner-Fraser, 2008), as well as in forming in-depth understandings of the general design principles of these networks in development and evolution (Carre et al., 2017; Davidson, 2010; Erwin and Davidson, 2009; Gao and Davidson, 2008; Lim et al., 2013; Peter and Davidson, 2009; Royo et al., 2011). Highlighted among them is the sea urchin developmental GRN, the most comprehensive and authenticated network constructed to date (Peter and Davidson, 2017).

Disruption of any gene in the network could interrupt normal development and conceivably cause diseases. Some diseases are caused by single gene disorders while others are complex and multifactorial such as bronchopulmonary dysplasia (BPD), a common cause of morbidity and mortality in preterm infants characterized by arrested alveolar development with varied severities within the subjects’ lungs (Jain and Bancelari, 2014; Jobe, 1999; Short et al., 2007). Both prenatal insults and postnatal injury increase the risk of BPD. The multifactorial etiology of BPD has made the development of therapies a unique challenge, and currently, no effective treatment exists to prevent or cure this debilitating disease.

In recent clinical trials, therapy with recombinant human insulin-like growth factor 1 (rhIGF1) protein showed initial promise for BPD, but significant limitations did not allow further clinical trials (Ley et al., 2019; Seedorf et al., 2020). One impediment to further development of IGF1 as a therapy is the lack of comprehensive information regarding the precise role of the IGF1 signaling pathway in alveologenesis.

Alveolar development is a highly complex process that is driven by multiple signaling pathways including TGFB, SHH, WNT, PDGF, VEGF, HGF, NOTCH, BMP, and IGF1 (Juul et al., 2020; Nabhan et al., 2018; Tsao et al., 2016; Verheyden and Sun, 2020; Wu and Tang, 2021; Zepp and Morrisey, 2019). IGF1 and IGF1R expression is found throughout fetal lung development and fluctuates at different stages. Both IGF1 and IGF1R expression is significantly reduced in BPD lungs (Hellstrom et al., 2016; Lofqvist et al., 2012; Yilmaz et al., 2017), suggesting a potential role in the pathogenesis of BPD. The function of IGF1 and IGF1R has been examined in their respective constitutive and conditional knockout mice, but their specific function during alveologenesis has not been examined (Epauud et al., 2012; Lopez et al., 2016; Lopez et al., 2015). A recent study found interruption in IGF1 signaling compromised mechanosignaling and interrupted alveologenesis (He et al., 2021).

In our study here, we found Igf1 and Igf1r are primarily expressed in SCMF in postnatal mice lungs. As a result, we interrupted IGF1 signaling in SCMF by the inactivation of Igfr at the onset of alveologenesis in postnatal day 2 (PN2) mouse neonates and analyzed the resulting phenotypes. Inactivation of Igfr resulted in mutant lungs with simplified and immature alveolar structure resembling that of human BPD at a moderate severity (Short et al., 2007). We reasoned that the phenotype caused by the inactivation of Igfr in SCMF reflects interruption in the genetic program downstream of IGF1 signaling that drives the normal functions of SCMF in the process of alveologenesis. Alterations in SCMF may further impact...
specification/differentiation of other key cell types, particularly the alveolar epithelial cells. We aim to decode the genes and their interactions behind this underlying genetic program.

Taking advantage of the high throughput next generation sequencing, people have been trying to build GRNs by computational analysis using wild type gene expression data (i.e. Jia et al., 2017; Xu et al., 2012). Currently, the most reliable way is still to build GRN experimentally on data from perturbation analyses.

Genes we particularly focus on are the regulatory genes. These genes decide the outcome of a GRN as they regulate and control other genes’ expression forming a regulatory circuitry of varied hierarchy with structural and cellular genes as their terminal targets (Davidson, 2010; Erwin and Davidson, 2009). Their functions are defined by their logical control over the circuitry’s operation and their synergetic biological effects are manifested by the effector genes under the circuitry (Peter and Davidson, 2009; Peter et al., 2012).

We followed the protocol similar to that used in the construction of the sea urchin GRN (Materna and Oliveri, 2008): regulatory genes were selected at the transcriptomic scale from the LungMAP database (www.lungmap.net) where their expression during alveologenesis was reported; the cellular expression patterns of these genes were then annotated on lungMAP scRNAseq data for the screen of SCMF genes; the regulatory interactions among these genes were subsequently examined from in vivo and in vitro perturbations; and cellular communications were finally determined by Secretome-Receptome analysis. Combining this data, we constructed the IGF1 signaling gene regulatory network underlying alveologenesis from this mouse model of human BPD phenocopy.

Our GRN work on alveologenesis represents a transformative view of how to understand and perhaps even design future preventive and therapeutic strategies for BPD treatment.
Results

Postnatal Expression of Igf1 and Igf1r in the lung.

To define the expression pattern of Igf1 and Igf1r during early postnatal lung development, we performed RT-PCR using total lung RNA from embryonic day E18 to postnatal day PN30. The analysis showed Igf1 is expressed dynamically during development, peaking at the onset of alveologenesis and progressively decreasing thereafter (Fig. 1A). The expression pattern of Igf1r was biphasic, initially displaying an overlap with Igf1 in early alveologenesis with a subsequent peak occurring on PN21 (Fig. 1B). Expression profiling of Igf1 and Igf1r from the entire lung and within Myofibroblasts during lung development was calculated based on data from the LungMAP (Figure 1-Figure Supplement 1A) and from the latest scRNAseq dataset (Figure 1-Figure Supplement 1B, Negretti et al., 2021). Similar patterns of their expression were observed when compared to our data.

Alveologenesis is characterized by secondary septa/crest formation, and the Alveolar Myofibroblasts are recognized as the driving force behind it. Secondary Crest Myofibroblast was derived from this concept and is broadly adopted in independent publications (i.e. Bostrom et al., 1996; Li et al., 2015; Li et al., 2018; Sun et al., 2022; Zepp et al., 2021).

The definition, specification, and function of SCMFs have not been systematically characterized. Presently, a consensus marker for these cells hasn’t been established although a list of markers, such as Acta2 (Kugler et al., 2017), Stc1 (Zepp et al., 2021), Tagln (Li et al., 2018), Fgf18 (Hagan et al., 2020) and Pdgfra (Li et al., 2015) have been suggested from previous work. To help resolve this issue, we compiled a comprehensive list of these markers suggested from literature and examined their expression across different mesenchymal cell types as clustered on the two latest scRNAseq datasets (Figure 1-Figure Supplement 1C&F, Negretti et al., 2021, Zepp et al., 2021). Pdgfra was revealed as a good SCMF marker agreed by both datasets (Figure 1-Figure Supplement 1C&F). RNAscope showed highly overlapping localization of Igf1/Igf1r and Pdgfra (Fig.1C&D), indicating SCMF as a principal cellular site of Igf1 and Igf1r expression in lungs undergoing alveologenesis.

Mesodermal-specific Inactivation of Igf1 and Igf1r.

As SCMF are derived from the lung mesenchyme early in lung development and Dermo1 is activated at the onset of mesodermal lineage specification (Li et al., 2008), we used Dermo1-cre (Dcre) to inactivate Igf1 and Igf1r separately, specifically within mesodermal lineages, and examined the mutant lung phenotype during embryonic and postnatal development (Figure 2-Figure Supplement 1A).

Loss of Igf1 or Igf1r in mesodermal progenitors using floxed alleles of these genes decreased their mRNA by approximately two folds (Figure 2-Figure Supplement 1B). Consistent with previous reports (Epaud et al., 2012; Lopez et al., 2015), reduction in body weight was observed in our Igf1r mutant pups (Figure 2-Figure Supplement 1C&D).

The Dcre;Igf1r-f/f mutant displayed clearly visible lung defects: thickened saccular walls at E18; dilated sacculi and severe reductions in the number of secondary crests at PN14 with the latter extending into PN30 (Figure 2-Figure Supplement 1E). Similar defects of less severity were observed in the Dcre;Igf1-f/f mutant (Figure 2-Figure Supplement 1F). However, due to the timing of Dermo1-cre activation, it is difficult to determine whether the impact of Igf1r on alveologenesis occurs postnatally or is a carryover impact of events occurring prior to onset of alveologenesis.
Postnatal inactivation of Igf1r in SCMF profoundly arrests alveologenesis.

The SHH targeted (i.e. Gli1+) fibroblasts have been rigorously examined through lineage tracing using Gli1creERT2;mTmG in our lab (i.e. Li et al., 2015, Li et al., 2019). Cell lineage analysis shows the SHH signaling targets different mesenchymal cell lineages through lung development.

There was a window of time in the early postnatal stage during which the derived GFP+ cells were observed primarily localized to the secondary septa and secondarily to parabronchial & perivascular smooth muscle fibers (Li et al., 2015). When Tamoxifen was titrated down to a certain dosage, the smooth muscle fibers were no longer labeled by GFP (Figure 2-Figure Supplement 2C). This very specific regimen was employed in our current paper. Consistent with our observation, it was found that Gli1 is predominantly expressed by proliferative SCMF/Myofibroblast cells as calculated from the two largest scRNAseq datasets recently published (Figure 1-Figure Supplement 1D&G; Negretti et al., 2021; Zepp et al., 2021).

Using Gli1-creERT2 (Gcre), we inactivated floxed alleles of Igf1r on PN2 at the onset of alveologenesis (Fig.2A). Inactivation of Igf1r was validated by genotyping (Figure 2-Figure Supplement 2A) and its downregulation verified by RT-PCR with RNA from both entire lung and FACS-isolated SCMF (Figure 2-Figure Supplement 2B). A successful recombination was assessed by Cre-induced green fluorescent protein (GFP) (Figure 2-Figure Supplement 2C). The mutant mice were slightly runted compared to the controls (Figure 2-Figure Supplement 2D).

Histology of multiple Gcre;Igf1r-f/f lungs at the timepoints PN7, PN14, and PN30, revealed a phenotype of profoundly arrested alveolar formation as measured by the mean linear intercept (MLI) (Fig.2B). In addition, ImageJ analysis showed decreased perimeter to area ratio of airspace (Peri/Area) as well as the number of airspaces per unit of area (# of Airspace/Area) – all consistent with the alveolar hypoplasia phenotype in the mutant lungs (Fig.2B, Figure 2-Figure Supplement 2E,F). The largest deviation between the controls and mutants in these measurements occurred at PN14, which marks the midpoint in the alveologenesis phase. Still, the severity of hypoplasia here is eclipsed by that seen in Gcre;Tgfbr-f/f induced BPD phenocopies (Gao et al., 2022).

Immunohistochemical analysis revealed no significant gross changes between the control and mutant lungs in either proliferation or apoptosis (Figure 2-Figure Supplement 2G). Similarly, examination of specific markers for various lung alveolar cell lineages, including the mesenchyme (ELN, TPM1), myofibroblasts (ACTA2, PDGFRA), endothelial cells (EMCN), lipofibroblasts (ADRP), epithelial cells (NKX2.1, SFTPC, HOPX), and pericytes (NG2) did not reveal significant differences (Fig.2C, Figure 2-Figure Supplement 2H). Nonetheless, it was observed in the mutant lungs that the deposition of elastin seemed aggregated at the secondary crest tips and the number of AT1 and AT2 cells were reduced (Fig.2C). The altered elastin deposition has been reported in previous studies (He et al., 2021; Li et al., 2019), but whether this alteration is the cause or the consequence of the impaired alveolar formation remains unknown.

To examine the genetic changes in Gcre;Igf1r-f/f mutant lungs, we tested two selected groups of genes: lung signature (Figure 2-Figure Supplement 2I) and angiogenesis (Figure 2-Figure Supplement 2J) genes. The analysis showed significant alterations in a few genes including Fgf10, Igf1r, and Pdgfra in the mutant lungs. It is noteworthy that the RNA used in the test was from whole lung tissue, while inactivation of Igf1r by Gcre was only targeted to SCMF. To determine the cell specific impacts of Igf1r inactivation, we characterized gene expression in FACS sorted cells in the following studies.
Identification of SCMF genes altered in mouse lungs of BPD phenotype.

Both GFP+ (SCMF) and Tomato+ (non-SCMF) cells were isolated by FACS from lungs dissected from Gcre;mTmG (control) and Gcre;mTmG;Igf1r-f/f (mutant) mice. Using the sorted cells, we compared the gene expression between them to identify genes enriched within SCMFs, those altered within SCMFs, and those altered within non-SCMFs in the mutant lungs (Fig.3A, Figure 3-Figure Supplement 1A).

Our analysis didn’t focus on the components of the pathway itself which are self-conserved and usually tissue-independent (Pires-daSilva and Sommer, 2003), but was instead directed at identifying genes downstream of the pathway, especially the regulatory genes through which the developmental GRN is specified (Erwin and Davidson, 2009). 47 genes known to encode signaling molecules and transcription factors were screened from the LungMAP transcriptomic database (www.lungmap.net) where they were indicated to be expressed from SCMF between PN3 to PN14 in the mouse developmental window (Figure 3-Source Data 1). Expression of the selected genes, together with 15 known SCMF cell markers and 3 non-SCMF genes (negative control), were examined and compared by quantitative RT-PCR using RNA from the sorted cells.

Three criteria were applied to identify IGF1 signaling targets within SCMF: (1) Functionally active in SCMF with its deltaCT <= 9 (relative to Gapdh in Ctrl_GFP, Figure 3-Source Data 1), equivalent to >= 10 copies of transcripts per cell (copies of Gapdh transcripts per cell based on Barber et al., 2005); (2) highly enriched in SCMF with FC >= 10, p<=0.05 (Ctrl_GFP vs Ctrl_Tomato, Fig.3B left column, Figure 3-Source Data 1); (3) significantly altered with FC >=2, p<=0.05 (Mut_GFP vs Ctrl_GFP, Fig.3B right column, Figure 3-Source Data 1).

Nineteen genes were identified that met all 3 criteria (genes highlighted red in the first column of Figure 3-Source Data 1). Spatial localization of a select number of these genes – including Foxd1, Sox8, Tbx2, Bmper, Actc1, Wnt5a, and Fgf10 – in SCMF was validated by IHC (Fig.3C, Figure 3-Figure Supplement 1B) or RNAscope (Fig.3D). Consistently, Wnt5a was recently reported as a signature gene of SCMF (Negretti et al., 2021), and data analysis on the two latest scRNAseq datasets revealed Wnt5a is dominantly expressed from SCMF/Myofibroblast along with smooth muscle cells (Figure 1-Figure Supplement 1E&H).

Regulation of the latter 19 genes by Igf1 signaling was illustrated on Biotapestry (Longabaugh, 2012) as displayed in Fig.3E. Notedly, all connections were drawn directly from the signaling, whereas the regulation may happen indirectly through the cross regulation among these genes (to be examined below).

The constructed link map points to IGF1 signaling as a positive regulator for 17 out of the 19 putative downstream genes (Fig.3E). Based on their molecular and cellular functions, the 19 genes can be stratified into two separate groups. One group is comprised of regulatory genes encoding transcription factors (Gli1, Foxd1, Tbx2, Sox8, Dermoi, Rxra) and signaling molecules (Fgf10, Wnt5a, Bmper, Cyr61, Pdgfra). The second group is comprised of Acta2, Actc1, BGN, Des, Eln, Cnn1, FN1, and TNC, representing genes that encode structural & cellular molecules. Based on the GRN’s hierarchical design (Erwin and Davidson, 2009), it is most likely that IGF1 signaling first targets the regulatory genes, referred to as SCMF specifiers, whose expression subsequently targets the downstream structural and cellular genes in SCMF (SCMF effectors).

Cross regulation of altered SCMF genes in the mouse BPD phenocopy model.

Within the inventory of altered SCMF genes, as defined above, are transcription factor Foxd1, one of the most reduced regulatory genes from our RT-PCR measurements, and two growth...
factor signaling molecules, *Fgf10* and *Wnt5a*, both known to function in postnatal lung development (Chao et al., 2016; Li et al., 2020; Zhang et al., 2020). Secretome-Receptome computation using the Fantom algorithm (Ramilowski et al., 2015) revealed the cognate receptors, FGFR1/3/4 and ROR1/2, involve in FGF10 and WNT5a signaling transduction within SCMF (Fig.4B). The results were validated on bulk RNAseq (Figure 4-Figure Supplement 1A) and the latest scRNAseq datasets (Figure 4-Figure Supplement 1B,C; Figure 4-Source Data 1).

**WNT5A** is a highly evolutionary conserved non-canonical Wnt ligand. In spite of being identified as a non-canonical ligand, it can, under certain circumstances, signal through canonical Wnt signaling pathways directly or indirectly (i.e. Mikels and Nusse, 2006). Current analysis is focused on the signaling’s non-canonical aspect and does not exclude other possibilities.

To investigate the possibility of cross regulation of these genes on the other altered SCMF genes, GFP+ SCMFs were isolated by FACS from postnatal *Gcre;mTmG* lungs and cultured in vitro (Fig.4 A). Cultures were treated with inhibitors which target the gene or pathway of interest. The inhibitor’s dose was determined from the literature and a series of testing (Figure 4-Source Data 2), and inhibition was validated by examining its target genes (Figure 4-Figure Supplement 2A), Blocking IGF1 signaling using the IGF1R inhibitor (PPP at 4uM as in Chen et al., 2017; Jin et al., 2018; Wang et al., 2019), led to altered expression of all genes – except *Gli1* – in the same direction as observed in vivo (Figure 4-Figure Supplement 2B), indicating that isolated cells in vitro recapitulate the observed in vivo findings. Expression level of Gli1 was too low to be reliably detected in culture, likely due to absence of HH signaling which in vivo is provided exclusively by the lung epithelium.

The signaling of FGF10 and the expression of *Wnt5a* and *Foxd1* were blocked in culture with the FGFR inhibitor (Infigratinib at 1uM as in Manchado et al., 2016; Nakamura et al., 2015; Wong et al., 2018), si*Wnt5a* (at 20nM as in Nemoto et al., 2012; Sakisaka et al., 2015; Zhao et al., 2017) and si*Foxd1* (at 20nM as in Li et al., 2021; Nakayama et al., 2015; Wu et al., 2018) respectively, and their effects on other SCMF genes were quantified by RT-PCR (Fig.4C). Identified genes with significant change are designated as targets of the gene/pathway perturbed, and their connections are plotted on the network (Fig.4D).

The network constructed on the new perturbation data has built connections between *Wnt5a* and certain fibroblast effector genes (i.e. *Acta2, Eln*) confirming its regulatory control on fibroblast growth and development. This finding in regard to this aspect of Wnt5a’s function resonates well with the clinical data from the study of some IPF patients (Martin-Medina et al., 2018). The architecture of the new construct reveals the identity of 3 transcription factors FOXD1, TBX2, and SOX8 forming a presumed connection hub. They are targeted by all 3 signaling pathways within the network and with each other as demonstrated by FOXD1’s regulation of *Tbx2* and *Sox8*, likely representing a core network subcircuit (Peter and Davidson, 2009) tasked to lockdown a regulatory state so that the signaling effect from the upstream can be stabilized.

**Alveolar epithelial genes are affected through WNT5a and FGF10 signaling.**

As mentioned above, both AT1 and AT2 cells were reduced in *Gcre:Igf1r-f/f* mutant lungs (Fig.2C). To identify any genetic alterations within these cells, we analyzed our FACS-sorted Tomato+ cells (Fig.3A), in which the lung epithelial cells reside, for the expression of a broad list of known AT1/AT2 signature genes, including their canonical markers and the ones identified by
scRNAseq (i.e. Treutlein et al., 2014; Nabhan et al., 2018; LungMAP). The comparison between the control and the mutant data identified the genes as significantly altered (Fig.5A).

Since our Gcre;Igf1r-f/f mutant was a SCMF targeted deletion, these changes must be a secondary consequence of intercellular cross communication. Within the altered SCMF gene list there are a total of 3 ligands: Wnt5a, Fgf10, and Cyr61 (Fig.4D). CYR61 is more commonly recognized as a matricellular protein (i.e. Lau, 2011) rather than a classical growth factor such as FGF10 and WNT5A. Secretome-Receptome analysis indicates that FGF10 and WNT5a from SCMF communicate with alveolar epithelial cells through their cognate receptors, FGFR2 and ROR1 (Fig.4B, Figure 4-Figure Supplement 1).

The intercellular connections discovered above, plus the affected epithelial genes, were also added to the network as shown in Fig.5B. Multiple evidence has shown that IGF1 signaling can promote lung epithelial growth and development (Ghosh et al., 2013; Narasaraju et al., 2006; Wang et al., 2018b). Our work here reveals a nearly comprehensive look at the genetic pathway behind that.

**WNT5a is required for alveolar formation as inferred by the IGF1 signaling GRN.**

Our effort so far has led to construction of the IGF1 signaling GRN during alveologenesis. The network offers firstly a bird’s-eye view of how the genes involved in IGF1 signaling are connected in the form of a molecular circuitry for alveologenesis, and then a mechanistic perception of how this circuitry is initially turned on in SCMF for the function of these cells and then advances to the neighboring alveolar epithelial cells to influence their cellular activities.

Under this mechanistic view, the position where a gene is located on the architecture of the network and the connections it has with other genes manifests its role and impact on the network’s operation and outcome, hence its specific biological function in the process of alveolar formation. The network makes it possible to assess and predict a gene’s function before it is tested in vivo.

Mainly derived from our in vitro data, it was discovered WNT5a signaling on the current network is located immediately downstream of IGF1 and the signaling has connections with many IGF1 downstream genes (Fig.5B). This indicates WNT5a’s feed-forward role on the transduction of IGF1 signaling and its biological effects. Indeed, when Wnt5a was inactivated postnatally in mice, the mutant lungs exhibited a BPD-like phenotype with arrested alveologenesis, similar to what was seen in the Igf1r mutant (Figure 5-Figure Supplement 1; Li et al., 2020). Inversely, the in vivo perturbation data collected from this Wnt5a mutant mouse model can be used to further define the regulatory connections where WNT5A is involved on the current IGF1 GRN.

**A GRN of similar components and wiring underlies human BPD.**

The expression of the mouse IGF1 signaling GRN regulatory genes in SCMF was also examined in lung samples from postmortem human BPD samples (Figure 6-Source Data 1). The regulatory genes, on the upper hierarchy of the network, determine the network’s outcome. In comparison to non-BPD lungs, 9 of the 12 genes examined were altered and 8 were altered in the same direction as they were defined in the mouse GRN (Fig.6). These findings indicate a genetic program of similar components and wiring underlying in human BPD.
**Discussion**

With massive gene expression data available (e.g., the LungMAP database) in this postgenomic era, one pressing task is to identify the function of genes, in particular, how they interact with one another (Przybyla and Gilbert, 2021). With gene regulatory network analysis as our targeted approach, we have constructed the IGF1 signaling GRN underlying alveologenesis using a mouse model of BPD.

The application of the GRN approach in the lung field is novel as it hasn’t been previously reported. Although not all the genes on the network have been examined by perturbation and the links haven’t been determined as direct or indirect, the regulatory connections running along the constructed GRN are already able to provide potential mechanistic explanation as to how the effect of IGF1 signaling is transduced from one gene to a constellation of its downstream genes, and from one cell type to another. Indeed, blocking WNT5A signaling is confirmed to produce a mouse BPD-like phenotype that mimics Igf1r-/- lungs as inferred by the network (Figure 5-Figure Supplement 1; Li et al., 2020). Signaling by IGF1, FGF10 and WNT5A have long been recognized in having roles in alveologenesis (i.e. Chao et al., 2016; He et al., 2021; Li et al., 2020; Zhang et al., 2020). The present GRN reveals a full genetic program and the cross-talk among them. The fact that all these signaling pathways have connections on the specification and effector genes of both alveolar mesenchymal and epithelial cells provides a direct causal link between the signaling and alveolar development.

Behind the overall process of alveologenesis is a much larger signaling gene regulatory network where several different cell types are involved. Within the mesenchymal cell type only, it is known that blocking IGF1, WNT5a, PDGFa, and TGFb signaling leads to impairment in development of alveoli, though with varying severities (Gao et al., 2022; He et al., 2021; Li et al., 2020; Li et al., 2019; Zhang et al., 2020). The sum of these leads to the construction of the hierarchical regulatory connections of these pathways within SCMF (Fig.7). Once signaling pathways from alveolar epithelium, endothelium, and immune cells are included, a much larger and more comprehensive signaling GRN behind alveologenesis is expected to emerge.

Clinically, BPD is caused predominantly by extrinsic factors which in the first place would interfere with normal cell’s extracellular activities including cell signaling and communication. From the GRN perspective, BPD is a developmental disease when the signaling GRN for alveologenesis is derailed and disrupted by these extrinsic factors (Fig.7).

The type and effect of these extrinsic factors must be taken into consideration when studying this disease. There was a reported increase in Wnt5A expression by mesenchymal cells from hyperoxia BPD models (Sucre et al., 2020)—a finding seemingly contradictory to our data on Wnt5a. The etiology of BPD is multifactorial in that it involves a plethora of factors such as lung immaturity, injury, inflammation, and genetic defects. In the injury model, the two opposite conditions, hyperoxia and hypoxia, can both cause lung injury and lead to BPD-like phenotypes. In our model, the genetic defect from the targeted genetic manipulation is the culprit behind the lungs of the BPD phenotype. It is expected that genes on the network may respond differently to these different conditions and challenges. The derailed gene expression on the network, either increased or decreased, can both disrupt the signaling GRN for alveologenesis and cause BPD.

The aforementioned network disruption may occur within different cell types, onto different signaling pathways, and upon different hierarchy of the network (Fig.7). From the network’s structural view, the disruption at the higher hierarchy has greater manifestation, thus leading to a higher severity of disease. TGFB signaling, atop the whole hierarchy of the signaling GRN on
Fig. 7 exhibits in its mutant mouse model the most severe BPD-like phenotype when compared to mutants from the other signaling pathways under it (Chao et al., 2017; Gao et al., 2022; He et al., 2021; Li et al., 2020; Li et al., 2019). Also, different signaling sitting at various hierarchical levels has different coverage (sum of genes under the signaling’s control) on the network. If disruption occurred beyond a signaling’s coverage, the treatment by targeting this signaling would be off target. IGF1 signaling is found to be in the middle hierarchical level of the constructed network (Fig. 7). By targeting simply one such signaling and hoping to treat BPD altogether will surely not always work.

Our GRN view on alveologenesis presents a transformative viewing of BPD and could potentially help in designing novel strategies to prevent and treat it. As for prevention, the GRN suggests a study centered on the network’s periphery with its focus on cutting off the extrinsic factors and blocking their connections to the intrinsic alveologenesis GRN. For treatment, it is a study on the network itself focusing on recovering and rescuing the signaling pathways that have been disrupted. However, all this relies on the foundational construction of the whole alveologenesis GRN—which includes all major signaling pathways involved within and between each cell type and the network assembly of the hierarchical regulatory connections among them. Though the network construction is currently not experimentally accessible in human models, our work shows the alveologenesis GRN is well conserved between mouse and human lungs. With the mouse model, the whole signaling GRN behind alveologenesis can be decoded as we did in this paper. The impact of extrinsic factors/signaling pathways and their manipulation on the network can be modeled and tested as well. The insights collected can then be used to guide health delivery for BPD clinically. The general process would then be as follows: clinical patient exams and tests are used to determine the extrinsic factors; cell type gene expression data from bronchoalveolar lavage and/or biopsy are used to reveal genes altered within the alveolar compartments in patient lungs; implied extrinsic factors and altered genes are mapped onto the alveologenesis GRN where points of network disruptions can thus be defined; therapies proven successful from mouse modeling data can then be clinically pursued to recover/rescue the disruption points in the patient.

The GRN is traditionally a network built solely on chemical interactions (RNA, DNA, protein and other chemicals). Having said that however, it has been recognized that these interactions can also happen at mechanical and electrical level (i.e. Azeloglu and Iyengar, 2015). Mechanosignaling is of particular interest in the lung as the organ itself is born to function through its non-stopping mechanical movement. New technologies, including whole genome CRISPR perturbation and screening, epigenomic profiling, high resolution ChIP-seq, spatial genomics and cis-regulatory analysis, are sprouting with their potential use for a high throughput GRN construction (i.e. Cusanovich et al., 2018; Eng et al., 2019; Sanson et al., 2018; Skene and Henikoff, 2015). Our journey to decode the alveologenesis GRN is also an adventure to construct the new generation GRNs with rising technologies.

A web resource of the network and the data presented in this paper is made available to the public: https://sites.google.com/view/the-alveologenesis-grn/home, which we hope can be used as an online repository to promote any further studies and collaborations in this direction.
MATERIALS AND METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>Reagent or Resource</th>
<th>Source</th>
<th>Identifier</th>
<th>Working Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTA2</td>
<td>Abcam</td>
<td>Cat#: AB5694</td>
<td>1:300</td>
</tr>
<tr>
<td>ACTC1</td>
<td>Santa Cruz</td>
<td>Cat#: SC-58670</td>
<td>1:100</td>
</tr>
<tr>
<td>ADRP</td>
<td>Abcam</td>
<td>Cat#: AB108323</td>
<td>1:200</td>
</tr>
<tr>
<td>AQP5</td>
<td>Alomone</td>
<td>Cat#: AQP-005</td>
<td>1:100</td>
</tr>
<tr>
<td>BMPER</td>
<td>Santa Cruz</td>
<td>Cat#: SC-377502</td>
<td>1:200</td>
</tr>
<tr>
<td>CAS3</td>
<td>Cell Signaling</td>
<td>Cat#: 9664</td>
<td>1:100</td>
</tr>
<tr>
<td>ELN</td>
<td>Abcam</td>
<td>Cat#: AB21600</td>
<td>1:200</td>
</tr>
<tr>
<td>EMCN</td>
<td>R&amp;D Systems</td>
<td>Cat#: AF4666</td>
<td>1:100</td>
</tr>
<tr>
<td>GFP</td>
<td>Santa Cruz</td>
<td>Cat#: SC-9996</td>
<td>1:100</td>
</tr>
<tr>
<td>HOPX</td>
<td>Santa Cruz</td>
<td>Cat#: SC-30216</td>
<td>1:50</td>
</tr>
<tr>
<td>KI67</td>
<td>R&amp;D Systems</td>
<td>Cat#: AF7649</td>
<td>1:50</td>
</tr>
<tr>
<td>NKX2.1</td>
<td>Seven Hills</td>
<td>Cat#: 8G7G3-1</td>
<td>1:50</td>
</tr>
<tr>
<td>NG2</td>
<td>Abcam</td>
<td>Cat#: AB5320</td>
<td>1:100</td>
</tr>
<tr>
<td>PDGFRa</td>
<td>Cell Signaling</td>
<td>Cat#: 3174</td>
<td>1:50</td>
</tr>
<tr>
<td>PDPN</td>
<td>Thermo Fisher</td>
<td>Cat#: 14-5381-82</td>
<td>1:300</td>
</tr>
<tr>
<td>RFP</td>
<td>Rockland</td>
<td>Cat#: 600-401-379S</td>
<td>1:300</td>
</tr>
<tr>
<td>SFTPC</td>
<td>Abcam</td>
<td>Cat#: AB3786</td>
<td>1:200</td>
</tr>
<tr>
<td>SOX8</td>
<td>Santa Cruz</td>
<td>Cat#: SC-374446</td>
<td>1:50</td>
</tr>
<tr>
<td>TBX2</td>
<td>Santa Cruz</td>
<td>Cat#: SC-514291</td>
<td>1:50</td>
</tr>
<tr>
<td>TPM1</td>
<td>Sigma Aldrich</td>
<td>Cat#: T2780</td>
<td>1:500</td>
</tr>
</tbody>
</table>

**Chemicals and Inhibitors**

<table>
<thead>
<tr>
<th>Chemical / Inhibitor</th>
<th>Source</th>
<th>Identifier</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamoxifen</td>
<td>Sigma</td>
<td>Cat#: T5648</td>
<td>8mg/ml</td>
</tr>
<tr>
<td>Infigratinib</td>
<td>Selleckchem</td>
<td>Cat#: S2788</td>
<td>1uM</td>
</tr>
<tr>
<td>PPP</td>
<td>Selleckchem</td>
<td>Cat#: S7668</td>
<td>4uM</td>
</tr>
<tr>
<td>Product</td>
<td>Manufacturer</td>
<td>Catalog#</td>
<td>Concentration</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>siFoxd1</td>
<td>Dharmaco</td>
<td>L-046204-00-0005</td>
<td>20nM</td>
</tr>
<tr>
<td>siWnt5a</td>
<td>Dharmaco</td>
<td>L-065584-01-0005</td>
<td>20nM</td>
</tr>
<tr>
<td>siRNA non-targeting</td>
<td>Dharmaco</td>
<td>D-001810-01-05</td>
<td>20nM</td>
</tr>
</tbody>
</table>

**Sequencing**

<table>
<thead>
<tr>
<th>Library Prep Kit</th>
<th>Manufacturer</th>
<th>Catalog#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next Ultra DNA</td>
<td>New England Biolabs</td>
<td>E7370</td>
</tr>
</tbody>
</table>

**Bulk RNAseq**

<table>
<thead>
<tr>
<th>Library</th>
<th>Manufacturer</th>
<th>Catalog#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk RNAseq</td>
<td>Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech</td>
<td>GEO: GSE182886</td>
</tr>
</tbody>
</table>

**Mouse**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Manufacturer</th>
<th>Catalog#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gli1-creERT2</td>
<td>The Jackson Laboratory</td>
<td>007913</td>
</tr>
<tr>
<td>Rosa26mTmG</td>
<td>The Jackson Laboratory</td>
<td>007676</td>
</tr>
<tr>
<td>CAGTomato</td>
<td>The Jackson Laboratory</td>
<td>007914</td>
</tr>
<tr>
<td>Dermo1-cre</td>
<td>The Jackson Laboratory</td>
<td>028701</td>
</tr>
<tr>
<td>Igfl(flox/flox)</td>
<td>The Jackson Laboratory</td>
<td>016831</td>
</tr>
<tr>
<td>Igflr(flox/flox)</td>
<td>The Jackson Laboratory</td>
<td>012251</td>
</tr>
<tr>
<td>CAG-creER</td>
<td>The Jackson Laboratory</td>
<td>004453</td>
</tr>
<tr>
<td>Wnt5a(flox/flox)</td>
<td>Kuruvilla Laboratory</td>
<td>N/A</td>
</tr>
<tr>
<td>Foxd1GCE</td>
<td>McMahon Laboratory</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**RNAscope**

<table>
<thead>
<tr>
<th>Reagent Kit</th>
<th>Manufacturer</th>
<th>Catalog#</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNAscope® 3-Fluorescent Reagent Kit V2</td>
<td>Advanced Cell Diagnostics</td>
<td>323100</td>
</tr>
<tr>
<td>Igfl</td>
<td>Advanced Cell Diagnostics</td>
<td>443901-C1</td>
</tr>
<tr>
<td>Igflr</td>
<td>Advanced Cell Diagnostics</td>
<td>417561-C3</td>
</tr>
<tr>
<td>Pdgfra</td>
<td>Advanced Cell Diagnostics</td>
<td>480661-C2</td>
</tr>
<tr>
<td>Fgf10</td>
<td>Advanced Cell Diagnostics</td>
<td>446371-C1</td>
</tr>
<tr>
<td>Wnt5a</td>
<td>Advanced Cell Diagnostics</td>
<td>316791-C3</td>
</tr>
<tr>
<td>RNAscope® 3-Fluorescent Reagent Kit V2</td>
<td>Advanced Cell Diagnostics</td>
<td>320881</td>
</tr>
</tbody>
</table>
plex Positive Control Probe_Mm

RNAscope® 3-plex Negative Control Probe

Software and Algorithms

<table>
<thead>
<tr>
<th>Software and Algorithms</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image J</td>
<td>NIH</td>
</tr>
<tr>
<td>STAR 2.5</td>
<td>PMCID: PMC3530905</td>
</tr>
<tr>
<td>Fantom5 Cell Connectome</td>
<td>FANTOM5 project</td>
</tr>
<tr>
<td>Imaris</td>
<td>BitPlane</td>
</tr>
<tr>
<td>R 3.2</td>
<td>R Project</td>
</tr>
<tr>
<td>LAS X</td>
<td>Leica</td>
</tr>
<tr>
<td>BioTapestry</td>
<td>Institute for Systems Biology</td>
</tr>
</tbody>
</table>

1. Mouse Breeding and Genotyping

All animal studies were conducted strictly according to protocols approved by the USC Institutional Animal Care and Use Committee (IACUC) (Los Angeles, CA, USA). The mice were housed and maintained in pathogen-free conditions at constant room temperature (20–22 °C), with a 12 h light/dark cycle, and free access to water and food. *Dermo1-cre, Gli1-cre*<sup>ERT2</sup>, *Rosa26<sup>mtmG</sup>, CAG<sup>tdTomato</sup>, Igf1<sup>(flox/flox)</sup>, Igf1r<sup>(flox/flox)</sup>*, and *CAG-cre<sup>ER</sup>* mice were purchased from the Jackson Laboratory. *Foxd1GCE* mice were generated by McMahon lab at USC. *SPC<sup>GFP</sup>* mice were generated by Wright lab at Duke University. *Wnt5a<sup>(flox/flox)</sup>* mice were provided by Kuruvilla lab at Johns Hopkins University.

*Dermo1-cre;Igf1<sup>(flox/flox)</sup>* (*Dcre;Igf1-f/f*) and *Dermo1-cre;Igf1r<sup>(flox/flox)</sup>* (*Dcre;Igf1r-f/f*) mice were generated by breeding *Dermo1-cre* with *Igf1<sup>(flox/flox)</sup>* and *Igf1r<sup>(flox/flox)</sup>* respectively.

*Gli1-cre*<sup>ERT2</sup>; *Rosa26<sup>mtmG</sup>* mice were generated by breeding *Gli1-cre*<sup>ERT2</sup> and *Rosa26<sup>mtmG</sup>* mice.

*Gli1-cre*<sup>ERT2</sup>; *Igf1r<sup>(flox/flox)</sup>* (*Gcre;Igf1r-f/f*) mice were generated by breeding *Gli1-cre*<sup>ERT2</sup> mice with the *Igf1r<sup>(flox/flox)</sup>* mice.
Rosa26^mTmG;Igf1r(fox/flox) (mTmG;Igf1r-f/f) mice were generated by breeding Rosa26^mTmG mice with the Igf1r(fox/flox) mice.

Gli1-cre^ERT2;Rosa26^mTmG;Igf1r(fox/flox) (Gcre;mTmG;Igf1r-f/f) mice were generated by breeding Gli1-cre^ERT2;Igf1r(fox/flox) mice with the Rosa26^mTmG;Igf1r(fox/flox) mice.

Foxd1GCE;CAG^tdTomato (Fcre;Tomato) mice were generated by breeding Foxd1GCE mice with the CAG^tdTomato mice.

CAG-cre^ER,Wnt5a(fox/flox) (Ccre;Wnt5a-f/f) mice were generated by breeding CAG-cre^ER mice with the Wnt5a(fox/flox) mice.

Genotyping of the transgenic mice was performed by PCR with genomic DNA isolated from mouse tails. The forward (F) and reverse primers (R) for transgenic mouse genotyping are listed in Figure 3-Source Data 2.

2. Tamoxifen Administration
A single dose of Tamoxifen (8mg/mL in peanut oil) was administered by oral gavage to neonates at postnatal day 2 (PN2, 100μg per pup) with a plastic feeding needle (Instech Laboratories, PA). Neonatal lungs were collected between PN7 and PN30 for morphological, immunohistochemical, cellular, and molecular biological analyses.

3. Mouse lung tissue
Mice were euthanized by CO2 inhalation at the time of tissue harvest. Chest cavity was exposed and lungs cleared of blood by perfusion with cold PBS via the right ventricle. Lungs were inflated with 4% formaldehyde under constant 30cm H2O pressure and allowed to fix overnight at 4°C. Tissue was dehydrated through a series of ethanol washes after which they were embedded in paraffin and sectioned.

4. Immunohistochemistry
H&E staining was performed as usual, and morphometric measurements were made using ImageJ. Immunofluorescent staining was performed as previously described using paraffin-embedded lung sections (Li et al., 2019). In brief, five micrometer (μm) tissue sections were deparaffinized, rehydrated, and subjected to antigen retrieval. After blocking with normal serum, the sections were probed with primary antibodies at 4°C overnight. Combinations of Alexa Fluor Plus secondary antibodies (Thermo Fisher Scientific) were applied for fluorescent detection of above specific primary antibodies. Nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI). Primary antibodies used and their sources are listed in the Key Resources Table below. Images were made with Leica DMi8 fluorescence microscope and processed with Leica LAS X and ImageJ.

5. RNAscope
Samples were fixed in 10% neutral buffered formalin, dehydrated with ethanol, and embedded in paraffin wax. 5μm Sections from paraffin blocks were processed using standard pretreatment conditions per the RNAscope multiplex fluorescent reagent kit version 2 (Advanced Cell Diagnostics) assay protocol. TSA-plus fluorescein, Cy3 and Cy5 fluorophores were used at different dilutions optimized for each probe. RNAscope probes used are listed in the Key Resources Table below. Images were made with Leica DMi8 fluorescence microscope and processed with Leica LAS X and ImageJ.
6. Mouse lung single-cell dissociation

Single-cell suspension was prepared as described in Adam et al. (2017) with all the procedures performed on ice or in cold room. Mice were euthanized and the lungs were perfused with PBS as described above. The lungs were inflated with cold active protease solution (5mM CaCl2, 10 mg/ml Bacillus Licheniformis protease), dissected, and transferred to a petri dish where the heart, thymus, and trachea were removed. The lobes were minced using a razor blade. The minced tissue was then immersed in extra cold active protease solution for 10 min and trituated using a 1ml pipette. This Homogenate was transferred to a Miltenyi C-tube with 5ml HBSS/DNase (Hank's Balanced Salt Buffer) and the Miltenyi gentleMACS lung program was run twice on GentleMACs Dissociator. Subsequently, this suspension was passed through a 100um strainer, pelleted at 300g for 6 minutes, suspended in 2ml RBC (Red Blood Cell) Lysis Buffer (BioLegend), and incubated for 2 min. At this point, 8ml HBSS was added and centrifuged again. The pellet was suspended in HBSS, then filtered through a 30um strainer. The suspension was pelleted again and finally suspended in MACS separation buffer (Miltenyi Biotec) with 2% FBS for Fluorescence-activated cell sorting (FACS). Cell separation and viability were examined under the microscope and through Vi-CELL Cell Counter after staining with Trypan blue.

7. Flow Cytometry and Cell Sorting

FACS was performed on a BD FACS Aria II at stem cell Flow Cytometry Core at the Keck School of Medicine of USC. The sorting was gated on viability, singlets, GFP and/or Tomato. GFP+ and/or Tomato+ cells were collected as needed. For cell culture, cells were sorted in DMEM containing 10% FBS. For RNA, cells of interest were collected in Trizol-LS reagent (ThermoFisher).

8. Bulk RNA-seq and scRNAseq data analyses

Lungs dissected from Sftpc-GFP mouse at PN14 were dissociated into single cell suspension and GFP+ cells were sorted and collected as described above. RNA was extracted using Qiagen RNeasy Micro kit and then submitted to the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech for sequencing, which was run at 50bp, single end, and 30 million reading depth. The unaligned raw reads from aforementioned sequencing were processed on the PartekFlow platform. In brief, read alignment, and gene annotation and quantification, were based on mouse genome (mm10) and transcriptome (GENECODE genes-release 7). Tophat2 and Upper Quartile algorithms were used for mapping and normalization. The RNA-seq data have been deposited with GEO under the accession number GSE182886.

The raw scRNAseq datasets (GSE160876 and GSE165063 from Negretti et al., 2021, GSE149563 from Zepp et al., 2021) were downloaded from the SRA server. The 10x Genomics’ Cell Ranger pipeline was used to demultiplex raw base call (BCL) files into FASTQ files, perform alignment, filtering, barcode counting, and UMI counting, combine and normalize counts from multiple samples, generate feature-barcode matrices, run the dimensionality reduction, clustering, and gene expression analysis using parameters once they were provided in the original papers. Quality control was done additionally in Partek Flow to filter out cells with excess mitochondrial reads and possible doublets and remove batch effects. Loupe Browser was used for data visualization and analysis including cell clustering, cell counting, cell type classification, gene expression and comparative analysis.

9. Neonatal lung myofibroblast culture and treatment
FACS sorted GFP+ cells from PN5 neonatal lungs of Gli1-creER<T>2;Rosa26<sup>mTmG</sup> mice were suspended in DMEM containing 10% FBS, plated in 24-well culture plates at 50,000 cells/well, and incubated at 37°C with 5% CO2 on day 1 after sorting. On day 2 after sorting, the attached myofibroblasts were washed with PBS and cultured in fresh medium with inhibitors or siRNAs as indicated in each experiment. The ON-TARGETplus SMARTpool siRNAs from Dharmacon were used, and the transfection was done using DharmaFECT™ Transfection Reagents. On day 4, the cells were collected for RNA analyses. The cells were authenticated for absence of contaminations.

10. Real-time Quantitative Polymerase Chain Reaction (real-time RT-PCR)

Neonatal mouse lung cells were collected from FACS or cell culture as described above. The RNA was isolated with Direct-zol™ RNA MiniPrep kit according to the manufacturer’s protocol (ZYMO Research). Following RNA purification, cDNA was generated using the SuperScript IV First-Strand Synthesis System (ThermoFisher). Expression of selected genes was quantified by Quantitative Real Time RT-PCR performed on a Light Cycler (Roche) or 7900HT fast real-time PCR system (Applied Biosystems) using SYBR green reagents (ThermoFisher). The deltaCT method was used to calculate relative ratios of a target gene mRNA in mutant lungs compared to littermate control lungs. Gapdh was used as the reference gene. Primers for each gene were designed on IDT website and the specificity of their amplification was verified by their melting curve. Sequences of the primers are listed in Figure 3-Source Data 2.

11. Human neonatal lung samples

BPD and non-BPD postnatal human lung tissues were provided by the International Institute for the Advancement of Medicine and the National Disease Research Interchange, and were classified exempt from human subject regulations per the University of Rochester Research Subjects Review Board protocol (RSRB00056775).

12. Secretome-Receptome analyses

Cell to cell communications were predicted using a published Fantom5 Cell Connectome dataset linking ligands to their receptors (STAR Methods) (Ramilowski et al., 2015). The ligands and receptors were identified from the following bulk RNAseq datasets: GSE126457 for SCMF (Li et al., 2019), GSE182886 for AT2 (submitted with this paper), GSE106960 for AT1 (Wang et al., 2018a), and the following scRNAseq datasets: GSE160876 (Schuler et al., 2021), GSE165063 (Negretti et al., 2021), and GSE149563 (Zepp et al., 2021). The IGF1 signaling GRN was drawn in BioTapestry software developed by Longabaugh et al., 2012.

13. Quantification and Statistical Analysis

In gene expression quantification using RT-PCR, at least three biological replicates (in different cases including lungs/FACS sorted cells/cultured cells) for each experimental group (Ctrl vs Mut, FACS sorted cell lineage #1 vs #2, treated vs untreated, BPD vs nonBPD) were used. Measurement for each biological replicate was repeated three times. The Ct (cycle threshold) was normalized to Gapdh, and the final result was presented as deltaCt or fold change. In morphometric quantification and cell counting, four lungs for each experimental group (Ctrl vs Mut) were used. Left lobe and right inferior lobe from each lung were targeted. Five images from each lobe after staining were analyzed for morphometric quantification (at 10x magnification) and cell counting (at 40x magnification). A two-tailed Student's t-test was used for the comparison between two experimental groups and a one-way ANOVA was used for multiple
comparisons. Quantitative data are presented as mean values +/-SD. Data were considered significant if p < 0.05.
Acknowledgements: We thank Dr. Andrew P. McMahon (Keck School of Medicine of USC) for providing the *Foxd1GCE;CAGtdTomato* mice. We thank Dr. Gloria S. Pryhuber (University of Rochester Medical Center) for providing the human BPD samples. We thank Arnold Sipos (Keck School of Medicine of USC) for help with imaging and Sean Gao (Arcadia High School/Duke University) for data analysis, editing, and the construction of the Alveologenesis GRN website.

Competing Interests: No.

Fundings: This work was supported by the National Institutes of Health [HL144932, HL122764 (C.L. & P.M.), HL143059 (P.M.), R35 HL135747 (Z.B. & P.M.)], the Hastings Foundation (P.M., Z.B.).

Data Availability Statement: Additional data that support the findings of this study are available on request from the corresponding author.
REFERENCES


FIGURES AND SUPPLEMENTS

Figure 1: Temporal and spatial expression of Igf1 and Igf1r during neonatal lung development in mice.

A, B: Temporal expression of Igf1 (A) and Igf1r (B) from E18 to PN30, quantified by RT-PCR and normalized to Gapdh. RNA used was collected from the whole lung and each stage was represented by at least five lungs. C: Spatial localization of mRNA for Igf1 (C) and Igf1r (D) in PN7 lungs as detected by RNAscope and their overlapped expression with Pdgfra. Outlined area on the left is magnified on the right. Scale bars: 20um under the whole view and 10um under the magnified view.

Figure 1-Figure Supplement 1: Igf1, Igf1r and SCMF/MyoFB marker gene’s expression in mouse lung from published public data sources.

A, B: Expression profiling of Igf1 and Igf1r from whole lung and in SCMF/MyoFB cells during lung development as calculated on data from LungMap (A) and from Negretti et al., 2021 (B). C-H: Heatmap and violin plots of collected SCMF/MyoFB marker gene’s expression within mesenchymal sub cell clusters as calculated on data from Zepp et al., 2021 (C-E) and Negretti et al., 2021 (F-H).

Figure 2: Postnatal inactivation of Igf1r from lung secondary crest myofibroblast cells.

A: Schematic of the experimental protocol. B: H&E staining of lung sections from control (a-c) and Gli1-creERT2 mutant (d-f) mice and their morphometric measurements by MLI (g-i) and Peri/Area ratio (j-l) at PN7, PN14 and PN30. See Figure 2-Figure Supplement 1 for the definition and calculation of these indices. C: Immunostaining of lung sections from control (a-d) and mutant (e-h) mice for Elastin (a,e), ACTA2 (b,f), AQP5/HOPX (c,g) and SFTPC/PDPN (d,h), and the comparison of the number of AT1 (i) and AT2 (j) cells between the control and mutant. Scale bar: 100um in B and 25um in C. p-value: * stands for 0.05-0.01, ** for 0.01-0.001, *** for <0.001, n.s. for not significant. The same designation is used throughout the paper.

Figure 2-Figure Supplement 1: Mesoderm constitutive deletion of Igf1 and Igf1r from mouse lung.

A: Schematic of the experimental protocol. B: Impact of inactivation of Igf1 (left) and Igf1r (right) by Dermo1-cre on their respective transcripts in PN14 lungs. Data is represented as mean +/- SD of three lungs from each experimental group. C: Body weight variation of pups at PN14 of one single litter from the breeding of Dermo1-cre:Igf1r-f/- and Igf1r-f/f (left graph). Body weight comparison of pups between the control and the homozygous mutant (right graph). Data is represented as mean +/- SD of 4 mice from each experimental group. D: Same experiment as done in C for pups at PN30. E: H&E staining of lung sections from Igf1r control (a-c) and Dermo1-cre mutant (d-f) mice and their morphometric measurements (MLI&%Airspace – see Figure 2-Figure Supplement 2E,F for more information) (g-i) at E18, PN14, and PN30. F: H&E staining of lung sections from Igf1 control (a-c) and Dermo1-cre mutant (d-f) mice and their morphometric measurements (g-i) at E18, PN14, and PN30. Scale bar: 100um for all images. p-value: * stands for 0.05-0.001, ** for 0.01-0.001, *** for <0.001, n.s. for not significant. The same designation is used throughout the paper.

Figure 2-Figure Supplement 2: Postnatal inactivation of Igf1r in secondary crest myofibroblasts.
A: genotyping of pups from the breeding between Gli1-cre\textsuperscript{ERT2};Igf1r-f/f and Igf1r-f/f;mTmG mice. B: Igf1r expression in control and mutant lungs at PN14 using whole lung (left) and FACS sorted GFP+ cells (right). Data is represented as mean +/- SD of three lungs from each experimental group. C: Immunostaining for Eln, Acta2, and GFP on PN14 lung sections from Gli1-cre\textsuperscript{ERT2};mTmG mice which were tamoxifen-treated at PN2 with 400ug (a-c) and 100ug (d-f) per pup. Smooth muscle fibers (outlined in c&f) were labeled by GFP at higher but not at lower TAM dosage. Scale bar: 25um for all images. D: Weight comparison between Ctrl and Mut at PN14. Data is represented as mean +/- SD of three mice from each experimental group. E: A sample of how to use ImageJ on cell counting (a,b) and morphometric measurements (c,d) where number of separated airspaces (N), perimeter (Peri) and area (Area) of each airspace can be calculated. Scale bar: 25um in a and 100um in c. F: Morphometric comparison between Ctrl and Mut at PN14 using different indexes. MLI is defined as in Crowley et al., 2019 and the others are quantified as below: (Peri/Area) = \( \sum_{i}^{N} \) (Peri)/\( \sum_{i}^{N} \) (Area), (%Airspace=(Area of Airspaces/Area Total) = \( \sum_{i}^{N} \) (Area)/(Area Total). G: Immunostaining for Ki67 (a,b) and cleaved Caspase-3 (d,e) on control and mutant lungs at PN14. The percentage of cells identified from above staining was compared between the control and mutant (c,f). Data is represented as mean +/- SD of three lungs from each experimental group. Scale bar: 25um for all images. H: Immunostaining for mesenchymal marker ELN (a,d), TPM1 (a,d), ACTA2 (b,e), PDGFRA (h,k), epithelial marker NKX2.1 (h,k), CLDN18 (i,l), endothelial marker EMCN (b,c,e,f,g,j), lipofibroblast marker ADRP (g,j), and pericyte marker NG2 (c,f) using control and mutant lungs at PN14. Scale bar: 50um for i, l and 25um for all the other images. I: Comparative expression of selected lung marker genes between control and mutant lungs at PN14. Data is represented as mean +/- SD of three lungs from each experimental group. J: Comparative expression of angiogenesis related genes between control and mutant lungs at PN14. Data is represented as mean +/- SD of three lungs from each experimental group.

**Figure 3: Identification of SCMF genes altered in Gli1-cre\textsuperscript{ERT2};Igf1r-f/f mutant lungs.**

A: Schematic of the experimental protocol. B: RT-PCR data from the selected genes displaying their enrichment in SCMF and alterations in the mutant. Genes marked with the orange line: regulatory genes selected from LungMap database; Genes with the green line: common SCMF markers; Genes with the blue line: non-SCMF genes. Red circles: data points meeting the cutoff criteria described in the text; Empty circles: data points failing the cutoff criteria. This designation is used throughout the manuscript. C: Spatial expression of Foxd1-Tomato/ELN (a-c), SOX8/ELN (d-f), and TBX2/ELN (g-i) in the alveolar compartment of PN14 lungs as detected by immunostaining. Specific antibodies were used for SOX8, TBX2, and ELN. RFP antibody was used for Foxd1-Tomato on lungs dissected from Foxd1GCE;CAG\textsuperscript{Tomato} mice. Scale bar: 25um for all images. D: Spatial expression of Wnt5a/Pdgfra (a) and Fgf10/Pdgfra (b) in the alveolar compartment of PN14 lungs as detected by RNAscope. Outlined area on the left is magnified on the right. Scale bars: 20um under the whole view and 10um under the magnified view. E: Biotapestry network illustration of altered SCMF genes and their connections to IGF1 signaling within SCMF. The source of IGF1 can be both autocrine and paracrine. Genes (nodes) are shown in the territories (colored boxes) in which they are expressed. Edges show regulation by the originating upstream factor and are either positive (arrow) or repressive (bar). Signalings across cell membranes are indicated as double arrow heads.

**Figure 3-Figure Supplement 1: Identification of SCMF genes altered in Gli1-cre\textsuperscript{ERT2};Igf1r-f/f mutant lung.**
A: GFP+ and Tomato+ cells are separated by FACS. B: Spatial expression of SOX8/ELN (a-f), TBX2/ELN (g-l), BMPER/ELN (m-r), and ACTC1/ELN (s-x) in the alveolar compartment in control and mutant lungs at PN14 as detected by immunostaining. Scale bar: 25um for all images.

**Figure 3-Source Data 1:** Table of genes with their RT-PCR data showing their level of expression in SCMF, enrichment in SCMF, and differential expression between control and mutant lungs, and annotation of their cellular expression in the lung based on LungMAP scRNAseq data. Cells in red: data points which meet the cutoff criteria designated in the paper or the cell type where the gene is annotated as expressing.

**Figure 3-Source Data 2:** List of the primers used in the paper. Primers are for *Mus musculus* by default and for *Homo sapiens* when denoted with Hs.

**Figure 4:** Cross regulation of the altered SCMF genes.

A: Schematic of the experimental protocol. B: Flow chart of the secretome-receptome analysis among SCMF, AT1 and AT2 cells. The ligands and receptors were identified from the following RNAseq datasets: GSE126457 for SCMF (Li et al., 2019), GSE182886 for AT2, GSE106960 for AT1 (Wang et al., 2018a). C: RT-PCR data from the altered SCMF genes demonstrating their response to the treatments by Infigratinib, siWnt5a, and siFoxd1. D: Biotapestry network illustration of the cross regulation among the altered SCMF genes.

**Figure 4-Figure Supplement 1:** Secretome-receptome analysis of cellular communications between *IGF1*, *WNT5A* and *FGF10* ligands from SCMF and receptors from SCMF, AT2 and AT1 as calculated on different data sources: bulk RNAseq data (A), scRNAseq data from Negretti et al., 2021 (B) and scRNAseq data from Zepp et al., 2021 (C).

Receptors are rendered as magenta for those from SCMF, yellow for AT2, and cyan for AT1; are arranged based on their abundance within each cell type with higher quantities on top/left and lower quantities on bottom/right. Ligands are connected to the receptors predicted to be restrictive and decisive for the function of the respective signaling (labeled in black). Other receptors (labeled in gray) are ignored at this time due to the following: 1. their lower abundance—only the top 150 highly expressed receptors from each cell type are considered; 2. less binding affinity—i.e. IGF1R vs INSR; 3. broad expression—i.e. LRP5 and FZD proteins are ubiquitously expressed across all cell types studied here; 4. known gene spatial expression—i.e. Fgfr2 is specifically expressed in alveolar epithelium. See Figure 4-Source Data 1 for the full list of these signaling genes selected for each cell type and from each dataset.

**Figure 4-Figure Supplement 2:** SCMF cell culture treatments.

A: RT-PCR data showing the effect of various treatments (top) on their respective targeted genes (bottom). B: RT-PCR data from the altered SCMF genes demonstrating their response to the treatment from PPP.

**Figure 4-Source Data 1:** List of the ligands and receptors used for Secretome-Receptome analysis.

**Figure 4-Source Data 2:** List of the inhibitors and their concentrations used in cell culture treatment.
Figure 5: Epithelial genes affected in Gli1-cre$^{ERT2}$;Igf1r-f/f mutant lung and their connections to IGF1 signaling from SCMF.

A: RT-PCR data from selected AT1 and AT2 genes revealing their alteration in the mutant lung. B: The IGF1 signaling GRN during alveologenesis consisting of genes downstream of IGF1 signaling from three cell types (SCMF, AT2 and AT1) and the intracellular and intercellular regulatory connections among them.

Figure 5-Figure Supplement 1: Postnatal inactivation of Wnt5a.

A: Schematic of the experimental protocol. B: H&E staining of lung sections obtained from control (a) and Wnt5a-/- (b) mice as generated and described in Li et al., 2020 and analyzed for this study. C: Morphometric measurements by MLI (c) and Peri/Area ratio (d). Scale bar: 100um for all images.

Figure 6: Regulatory genes from IGF1 signaling GRN and their expression in human bronchopulmonary dysplasia (BPD) lungs.

Figure 6-Source Data 1: List of clinical data for human neonatal lung samples.

Figure 7: The hierarchical connections among the signaling pathways within SCMF during alveologenesis.
Figure 1: Temporal and spatial expression of Igf1 and Igf1r during neonatal lung development in mice. A, B: Temporal expression of Igf1 (A) and Igf1r (B) from E18 to PN30, quantified by RT-PCR and normalized to Gapdh. RNA used was collected from the whole lung, and each stage was represented by at least five lungs. C: Spatial localization of mRNA for Igf1 (C) and Igf1r (D) in PN7 lungs as detected by RNAscope and their overlapped expression with Pdgfra. Outlined area on the left is magnified on the right. Scale bars: 20µm under the whole view and 10µm under the magnified view.
Supplemental Figure 1: Igf1, Igf1r and SCMF/MyoFB marker gene’s expression in mouse lung from published public data sources. A, B: Expression profiling of Igf1 and Igf1r from whole lung and in SCMF/MyoFB cells during lung development as calculated on data from LungMap (A) and from Negretti et al., 2021 (B). C-H: Heatmap and violin plots of collected SCMF/MyoFB marker gene’s expression within mesenchymal sub cell clusters as calculated on data from Zepp et al., 2021 (C-E) and Negretti et al., 2021 (F-H).
Figure 2: Postnatal inactivation of Igf1r from lung secondary crest myofibroblast cells. A: Schematic of the experimental protocol. B: H&E staining of lung sections from control (a-c) and Gli1-creERT2 mutant (d-f) mice and their morphometric measurements by MLI (g-i) and Peri/Area ratio (j-l) at PN7, PN14 and PN30. See Supplemental Figure 2 for the definition and calculation of these indices. C: Immunostaining of lung sections from control (a-d) and mutant (e-h) mice for Elastin (a,e), ACTA2 (b,f), AQP5/HOPX (c,g) and SFTPC/PDPN (d,h), and the comparison of number of AT1 (i) and AT2 (j) cells between the control and mutant. Scale bar: 100um in B and 25um in C. p-value: * stands for 0.05-0.01, ** for 0.01-0.001, *** for <0.001, n.s. for not significant, and the same designation is used through the paper.
Supplemental Figure 2: Mesoderm constitutive deletion of Igf1 and Igf1r from mouse lung. A: Schematic of the experimental protocol. B: Impact of inactivation of Igf1 (left) and Igf1r (right) by Dermo1-cre on their respective transcripts in PN14 lungs. Data are represented as mean +/- SD from three lungs from each experimental group. C: Body weight variation of pups at PN14 within a litter from the breeding of Dermo1-cre;Igf1r-f/f and Igf1r-f/f (left graph). Body weight comparison of pups between the control and the homozygous mutant (right graph). Data are represented as mean +/- SD from 4 mice from each experimental group. D: Same experiment as in E for pups at PN30. E: H&E staining of lung sections from Igf1r control (a-c) and Dermo1-cre mutant (d-f) mice and their morphometric measurements (g-i) at E18, PN14 and PN30. F: H&E staining of lung sections from Igf1 control (a-c) and Dermo1-cre mutant (d-f) mice and their morphometric measurements (g-i) at E18, PN14 and PN30. Scale bar: 100um for all images. p-value: * stands for 0.05-0.001, ** for 0.01-0.001, *** for <0.001, n.s. for not significant, and the same designation is used throughout the paper.
**Figure A**: Representative images of Igf1r+/+ (top row) and Igf1r-/- (bottom row) lungs stained with DAPI and anti-ELN antibodies.

**Figure B**: Quantification of ELN expression in the whole lung and isolated cells of Igf1r+/+ and Igf1r-/- mice.

**Figure C**: Immunofluorescence images of TAM @ 400ug and 100ug showing the expression of ELN and GFP in Ctrl and Mut groups.

**Figure D**: Graphs showing the weight measurements for Ctrl and Mut groups.

**Figure E**: Images showing DAPI staining, ImageJ tracing, H&E staining, and ImageJ tracing.

**Figure F**: Bar graphs showing the MLI, Peri/Area, # of Airspaces/Area, and Area of Airspace/Area comparisons between Ctrl and Mut groups.

**Figure G**: Immunofluorescence images of Igf1r+/+ and Gcre;Igf1r+/+ lungs showing K87 and Cas3 expression.

**Figure H**: Images showing TPMI/ELN, ACTA2/ELN, ECMN/NG2, and ECMN/ADRP expression in Igf1r+/+, Gcre;Igf1r+/+, Igf1r+/+, and Gcre;Igf1r+/+ groups.

**Figure I**: Bar graph showing the lung marker gene expression in Gcre;Igf1r+/+.

**Figure J**: Bar graph showing the angiogenesis gene expression in Gcre;Igf1r+/+.

**Table**: Summary of gene expression changes in different groups.
Figure 3: Identification of SCMF genes altered in Gli1-creERT2;Igf1r-f/f mutant lungs. A: Schematic of the experimental protocol. B: RT-PCR data from the selected genes displaying their enrichment in SCMF and alterations in the mutant. Genes marked with orange line: regulatory genes selected from LungMap database; Genes with green line: common SCMF markers; Genes with blue line: non-SCMF genes. Red circles: data points meeting the cutoff criteria described in the text, Empty circles: data points failing the cutoff criteria. This designation is used throughout the manuscript. C: Spatial expression of Foxd1-Tomato/ELN (a-c), SOX8/ELN (d-f), and TBX2/ELN (g-i) in the alveolar compartment of PN14 lungs as detected by immunostaining. Specific antibodies were used for SOX8, TBX2 and ELN, and GFP antibody was used for Foxd1-Tomato on lungs dissected from Foxd1GCE;CAGTomato mice. Scale bar: 25um for all images. D: Spatial expression of Wnt5a/Pdgfra (a) and Fgf10/Pdgfra (b) in the alveolar compartment of PN14 lungs as detected by RNAscope. Outlined area on the left is magnified on the right. Scale bars: 20um under the whole view and 10um under the magnified view. E: Biotapestry network illustration of altered SCMF genes and their connections to IGF1 signaling within SCMF. The source of IGF1 can be both autocrine and paracrine. Genes (nodes) are shown in the territories (colored boxes) in which they are expressed. Edges show regulation by the originating upstream factor and are either positive (arrow) or repressive (bar). Signaling across the cell membrane is indicated as double arrow heads.
Supplemental Figure 4: Identification of SCMF genes altered in Gli1-creERT2;Igf1r-f/f mutant lung. A: GFP+ and Tamato+ cells are separated by FACS. B: Spatial expression of SOX8/ELN (a-f), TBX2/ELN (g-l), BMPER/ELN (m-r), and ACTC1/ELN (s-x) in the alveolar compartment in control and mutant lungs at PN14 as detected by immunostaining. Scale bar: 25um for all images.
Figure 4: Cross regulation of the altered SCMF genes. A: Schematic of the experimental protocol. B: Flow chart of the secretome-receptome analysis among SCMF, AT1 and AT2 cells. The ligands and receptors were identified from the following RNAseq datasets: GSE126457 for SCMF (Li et al., 2019a), GSE182886 for AT2, GSE106960 for AT1 (Wang et al., 2018a). C: RT-PCR data from the altered SCMF genes demonstrating their response to the treatments by Infigratinib, siWnt5a and siFodxl. D: Biotapestry network illustration of the cross regulation among the altered SCMF genes.
**Supplemental Figure 5:** Secretome-receptome analysis of cellular communications between *IGF1*, *WNT5A* and *FGF10* ligands from SCMF and receptors from SCMF, AT2 and AT1 as calculated on different data sources: bulk RNAseq data (A), scRNAseq data from Negretti et al., 2021 (B) and scRNAseq data from Zepp et al., 2021 (C). Receptors are rendered as magenta for those from SCMF, yellow for AT2 and cyan for AT1, and are arranged on their abundance within each cell type with high on top/left and low on bottom/right. Ligands are connected to the receptors (labelled in black) predicted to be restrictive and decisive for the function of the respective signaling while other receptors (labelled in gray) are ignored at this time because of: 1. their lower abundance – only the top 150 highly expressed receptors from each cell type are considered; 2. less binding affinity – i.e. IGF1R vs INSR; 3. broad expression – i.e. LRP5 and FZD proteins are ubiquitously expressed across all cell types studied here; 4. known gene spatial expression – i.e. Fgfr2 is specifically expressed in alveolar epithelium. See Supplemental Table for the full list of these signaling genes selected for each cell type and from each dataset.
Supplemental Figure 6: SCMF cell culture treatments. A: RT-PCR data showing the effect of various treatments (top) on their respective targeted genes (bottom). B: RT-PCR data from the altered SCMF genes demonstrating their response to the treatment from PPP.
Figure 5: Epithelial genes affected in Glil-creERT2;Igf1r-f/f mutant lung and their connections to IGF1 signaling from SCMF. A: RT-PCR data from selected AT1 and AT2 genes revealing their alteration in the mutant lung. B: The IGF1 signaling GRN during alveologenesis with the IGF1 signaling-targeted epithelial genes and their intercellular connections to the signaling using WNT5A and FGF10 signaling as the bridge.
Figure 6: Postnatal inactivation of Wnt5a. A: Schematic of the experimental protocol. B: H&E staining of lung sections from control (a) and Wnt5a-/- (b) mice and their morphometric measurements by MLI (c) and Peri/Area ratio (d). Scale bar: 100um for all images.
Figure 7: Regulatory genes from IGF1 signaling GRN and their expression in human bronchopulmonary dysplasia (BPD) lungs.
Figure 8: The hierarchical connections among the signaling pathways within SCMF during alveologenesis.