Understanding the Bud Tip Progenitor Cell Niche During Human Lung Development

by

Renee Frances-Conway Hein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cell and Developmental Biology)
in the University of Michigan
2022

Doctoral Committee:

Assistant Professor Idse Heemskerk, Chair
Professor Kurt Hankenson
Professor Jason Spence
Assistant Professor Joshua Welch
Professor Rachel Zemans

Renee F.C. Hein

conwayr@umich.edu

ORCID iD: 0000-0002-1614-2420

Dedication

To my scientific mentors who cultivated my strengths, passion, and dedication into the scientist I am today. To my family and friends who have been my rock, providing encouragement and support throughout this process, and who bring out the best in me.

Acknowledgements

Thank you to my thesis mentor Dr. Jason Spence for being a scientific role model that cultivated curiosity, rigor, and innovation into my scientific mindset. Your excitement for science is contagious, and the patience you show towards your trainees is rare to find. Thank you for continually celebrating the successes of those in your lab and for showing compassion, encouragement, and perseverance during times of challenge. Thank you for providing the support I needed but also the freedom to develop my confidence as an independent scientist.

Thank you to my previous scientific mentors. To Dr. Pamela Tautz for creating the Research Methods class at University High School, which gave me the opportunity to pursue my first research project and discover my passion for science early on. Thank you to the Undergraduate Biology Research Program (UBRP) at the University of Arizona for providing resources and funding for me to pursue research throughout my undergraduate degree. The UBRP program allowed me to focus my time and efforts on research without having to take on other jobs. To my undergraduate research mentors, Drs. John Szivek, Richard Kris, Stephen Felder, and Klearchos Papas, for believing in me and teaching me research techniques, experimental design, and critical thinking from the ground-up. To my educators, in particular Barry Callesen, for showing me how mathematics can both be fun and challenging. Your class kick-started my interest in technical fields and shifted the way I think about problems.

A *special* thank you to the entire Spence Lab. To our Principal Investigator, Jason Spence, who has developed a lab space that is full of support. My colleagues are compassionate and excited to help me however they can, and I have been provided with the resources I need. I

am full of gratitude that I have had the privilege to spend my Ph.D. in this lab, and I have made some of my best friends in this space. To Sha Huang for being the glue that holds the lab together. Thank you for not only helping me to become an expert in pluripotent stem cell culture, but for making me laugh until I cry and for sharing many car rides home with me. To Dr. Yu-Hwai Tsai for being a scientific role model I can only strive to become, for always smiling, and for your love of animals. To Fei Fei Wu for keeping the lab in check (seriously) and for loving my dog Kida possibly even more than I do. To my fellow graduate students: Dr. Alyssa Miller, Dr. Emily Holloway, Dr. Meghan Capeling, Ansley Conchola, Charlie Childs, Madeline Eiken, and Mayowa Akinwale for your emotional support and scientific support. Words cannot describe how thankful I am to have had you all throughout this journey. To Joshua Wu and Zhiwei Xiao for providing exceptional computational expertise throughout my research projects. To Alexis Fine – it has been the greatest privilege and success of my academic career to mentor you. It has been so special to learn side-by-side with you and watch you develop into an incredible scientist. To the rest of the Spence Lab members for many great conversations, positive vibes, inspirational scientific journeys, and just for making up the absolute best lab.

I would like to thank the members of my dissertation committee, Drs. Kurt Hankenson, Idse Heemskerk, Joshua Welch, and Rachel Zemans, for their professional support and insightful feedback throughout my dissertation work. Thank you to the communities within Cell and Developmental Biology and the Program in Biomedical Sciences for fostering a positive and collaborative work environment. Additionally, thank you to the collaborators and funding sources that made my thesis work possible.

My closest friends deserve a huge thank-you. To Lauren Koch, who stood by my side throughout the most challenging times of graduate school. I couldn't have done this without you.

I thoroughly enjoyed all the Michigan adventures, scary movies, and reminiscing on how life brought us together in both Arizona and Michigan. To Katelyn Polemi for being there when I need a friend to talk to or eat good food with, and to Kimberly Wallace for being my longest-time best friend, for always looking out for me, and for forgiving me when I left you to fend for yourself in the nursing major (you turned out to be the greatest nurse of all time). You all make me smile. Thank you to the other friends that have been so wonderful to have throughout this time.

Thank you to my ***amazing*** family. To my mom Karen Conway, dad Kevin Conway, sister Rachel Hartshorne, brother Kyle Conway, brother-in-law Brian Hartshorne, sister-in-law Clara Williams, niece Evelyn, and nephew Owen for all of your love and support and for your continuous encouragement to strive for my dreams. Thank you all for molding me into the person I am today. Thank you to my extended family including my cousin Paige Nelson, grandma Irene Thompson, papa Tom Thompson, grandmother Nancy Conway, grandfather Eugene Conway, Aunt Lauri Martin, Aunt Donna Urling, and many others for the most incredible childhood surrounded by a loving family. Thank you to my in-laws Sally Hein, Eric Hein, Jeffery Hein, DiAnn Plunkett, Paul Plunkett, Hal Hein, and the Doucette's for always showing how proud of me you are and for showing interest in what I do. Of course, thank you to my dog Kida for giving me a big, welcoming smile and cuddle when I come home from good days or bad days. Most of all, thank you to my husband Paul Hein who has provided me with every kind of support a graduate student could hope for in a partner – thank you for walking side-by-side with me as we both make our way in this world. I love you all.

Table of Contents

List of Tables

List of Figures

Abstract

Human development is a complex but fascinating process by which a single fertilized egg grows into three tissue layers, giving rise to diverse organs and tissues that each have unique developmental requirements. Stem and progenitor cells are at the heart of development, whereby they self-renew and make cell fate decisions to create the array of cells that make up an entire organism. The mechanisms that control cell fate are critical for understanding how normal development takes place and how genetic abnormalities, premature birth, and environmental factors influence development. A subset of stem and progenitor cells persist into adulthood and often use regulatory mechanisms that are similar to those used during fetal development, so understanding how stem and progenitor cells are controlled is also important for adult tissue repair following injury and for developing cellular therapies for diseases.

The respiratory epithelium is specified from the ventral-anterior aspect of a common endoderm-derived gut tube, which gives rise to a specialized lung progenitor. These 'bud tip progenitors' are maintained through lung development and give rise to all cell types lining the lungs. Bud tip progenitors are regulated by reciprocal signaling with surrounding mesenchyme; however, mesenchymal heterogeneity and function in the developing human lung is poorly understood. The first portion of research in this thesis explores how bud tip progenitors are regulated by mesenchymal cells in their niche. To address this, we interrogated single-cell RNA sequencing data from multiple human lung specimens and identified a mesenchymal cell population present that is highly enriched for expression of the WNT agonist *RSPO2*. We found that adjacent bud tip progenitors are enriched for the RSPO2 receptor *LGR5*. Functional

experiments using organoid models, explant cultures, and FACS-isolated RSPO2^{+} mesenchyme show that RSPO2 is a critical niche cue that acts through LGR5 on bud tip progenitors to potentiate WNT signaling in bud tips, which is required for their maintenance and multipotency.

Animal models and human tissue have been instrumental in addressing questions pertaining to lung development; however, animal models limit the understanding of huma-specific biology and access to human tissue is limited and burdened by regulatory and ethical considerations. Overcoming many of these limitations are recent advances in iPSC-derived models of human development. The second portion of research in this thesis was aimed at developing an iPSC organoid model enriched with bud tip progenitor-like cells that faithfully recapitulates the bud tip progenitors found in the developing human lung. Building on prior work, we optimized a directed differentiation paradigm to generate spheroids with robust NKX2-1 expression. Spheroids were expanded into organoids that possessed bud tip progenitor-like cells, which increased in number over time. Single cell RNA-sequencing analysis revealed a high degree of transcriptional similarity between induced bud tip progenitors (iBTPs) and *in vivo* BTPs. Using FACS, iBTPs can be purified and expanded as induced bud tip progenitor organoids (iBTOs), which maintain an enriched population of bud tip progenitors. When iBTOs are directed to differentiate into airway or alveolar cell types using well-established methods, they give rise to organoids composed of organized airway or alveolar epithelium, respectively. Overall, the work in this thesis has increased our understanding in the bud tip progenitor niche influence cell fate decisions and has created an iPSC-derived bud tip progenitor organoid model that can be used to study human lung development.

Chapter 1 Introduction[1]

Human development begins with a single, fertilized egg that undergoes repeated divisions and organization into cells that will become the extraembryonic tissue and cells that will become the body, termed the inner cell mass (Niakan *et al.*, 2012; Wamaitha and Niakan, 2018; Weatherbee, Cui and Zernicka-Goetz, 2021). The cells of the inner cell mass progress into three distinct germ layers during gastrulation, including the endoderm that gives rise to the inner epithelial lining of many organs, the mesoderm that surrounds the endoderm and gives rise to supportive mesenchymal cells encapsulating organs as well as other cells like bone and muscle, and the ectoderm that gives rise to neurons and outer linings of our body (Zhai *et al.*, 2022).

Over the course of development, stem cells generally become progressively narrower in their differentiation potential, or potency, a process first described in 1957 by Conrad Waddington in "The Epigenetic Landscape". This process begins with totipotent stem cells, which give rise to all cells in a body and extraembryonic tissue (i.e. the placenta), then progresses to pluripotent stem cells that give rise to all cells in a body but not extraembryonic tissue (Zakrzewski *et al.*, 2019). Pluripotent stem cells come in the form of embryonic stem cells, which make up the inner cell mass during development, but they can also be artificially created by reprogramming somatic cells to a pluripotent state, termed induced pluripotent stem cells (Takahashi and Yamanaka, 2006; Zakrzewski *et al.*, 2019). Finally, some cells are maintained as multipotent stem and progenitor cells that have differentiation potential capped within a specific

[1] Portions of this chapter have been published: Conway, Renee F., Tristan Frum, Ansley S. Conchola, and Jason R. Spence. 2020. "Understanding Human Lung Development through In Vitro Model Systems." *BioEssays* 42 (6). https://doi.org/10.1002/bies.202000006.

tissue and/or organ (Zakrzewski *et al.*, 2019). Multipotent stem and progenitor cells exist within the developing fetus to aid in creation of cells specific to certain tissues or organs as well as within the adult body to aid in tissue maintenance and repair. Nevertheless, the more scientists have learned about stem cell fate decisions, the more we have recognized that this process is more complicated than a direct lineage tree. For example, in some contexts, mature somatic cells can change fate into a different type of mature somatic cell of the same potency (transdifferentiate) or they can backtrack and return to a stem or progenitor cell that has more potency than it had previously (de-differentiate).

The lining of many organs, such as the lung and digestive track, are derived from a common tube of endodermal progenitor cells. A specialized population of progenitor cells bud off from the ventral-anterior aspect of this tube and are responsible for generating the entire lining of the lung (Rawlins *et al.*, 2009; Miller *et al.*, 2018). The process by which these 'bud tip progenitor' cells give rise to their differentiated daughter cells is dependent on interactions with their microenvironment, or niche. A cell's niche is comprised of other cell types from all three germ layers, growth factors, extracellular matrix component, mechanical properties of the tissue, cell adhesion molecules, pH, metabolites, oxygen levels, and more (Scadden, 2006). Research about how a stem or progenitor cell's niche influences its behavior is important for understanding how tissues form during development, how aberrant development occurs, how tissues are repaired, disease manifestations within stem cell populations, and for being able to re-create or control a cell's niche for regenerative medicine or cell therapies.

The work in this thesis investigates the bud tip progenitor cell niche during human lung development, particularly how mesenchymal cells use cell signaling to control bud tip progenitor fate decisions. This thesis also aimed to utilize what we have learned about the bud tip progenitor

niche to create an iPSC-derived bud tip organoid model that can be used in future studies. The introduction chapter (Chapter 1) will delve into details about what is known about human lung development, how human lung development differs from lung development of other species, and the tools scientists employ to study questions of human-specific lung development. The end of the chapter outlines the thesis research, which is described in detail in Chapters 2 and 3. Finally, Chapter 4 discusses the future directions of the research presented in this thesis.

1.1 Introduction to Human Respiratory Development

The respiratory system is comprised of the trachea and airways of the lung, the branched network of epithelial tubes forming the bronchi and bronchioles, and the alveoli, where gas exchanges with the vascular system. Each of these structures are made of multiple specialized epithelial cell types that help carry out the lung's unique functions of air intake and gas exchange, epithelial barrier function, protection from microbes and pathogens, and the maintenance of fluid and electrolyte homeostasis (Chakraborty and Kotecha, 2013). The diverse repertoire of respiratory epithelial cells that comprise the trachea, airways, and alveoli are derived from a common population of progenitor cells, termed bud tip progenitors, that are specified in the endodermal germ layer early during development (Rawlins, 2011; Nikolić *et al.*, 2017; Miller *et al.*, 2018). In addition to the endoderm-derived epithelium, both the developing and mature respiratory systems contain cells derived from the mesoderm (e.g. smooth muscle) and ectoderm (e.g. neurons) germ layers, and the complex interactions between cells from all three germ layers are absolutely critical for proper respiratory system development and function (Alescio and Cassini, 1962; Tollet, Everett and Sparrow, 2001; Lazarus *et al.*, 2011; Zepp and Morrisey, 2019).

Development of the respiratory system is broadly divided into five stages, each representing major morphological changes that take place (Schittny, 2017) (Figure 1-1). The embryonic stage is defined by respiratory specification, the establishment of the nascent tracheal domain, and the emergence of two primary lung buds from the ventral anterior foregut endoderm (Figure 1-1A). Following these events, the lung enters the pseudoglandular stage where lung buds undergo repeated rounds of bifurcations during a process called branching morphogenesis, which establishes the arborized network of bronchi and bronchioles (Metzger *et al.*, 2008) (Figure 1-1B). The alveoli form across several stages, with alveolar cell-type specification beginning during branching morphogenesis and finalizing their differentiation in the terminal stages of lung development (Frank *et al.*, 2016), which includes the canalicular stage where alveolar ducts form at terminal bronchioles, the saccular stage where alveolar cells functionally mature and alveolar sacs form, and the alveolar stage where alveoli continue to mature and increase their surface area through septation (Figure 1-1C). The lung is one of the few organs that continues to develop in post-natal life as the alveoli continue to grow in size and complexity for seven years after birth in humans and one month after birth in mice.

The structure and function of the adult mouse and adult human respiratory systems have multiple differences, including anatomical differences such as the number of airway branches, the identity and localization of adult stem cells, and the morphology of alveoli (Rock and Hogan, 2011) (Figure 1-1D). These physiological differences likely contribute to the failure of most human clinical trials using lung therapeutics developed in mouse models (Perrin, 2014). Until recently, it has been difficult to assess the mechanistic differences that emerge during respiratory development that led to differences in the mature lungs of mice and humans. However, contemporary research has addressed this issue by using primary human tissue and by

developing *in vitro* model systems that mimic human respiratory development. Coupled with technological advances such as single cell RNA sequencing (scRNA-seq) (Treutlein *et al.*, 2014; Brazovskaja, Treutlein and Camp, 2019; Guo *et al.*, 2019; Travaglini *et al.*, 2020; Sun *et al.*, 2021), these studies have shed light on many of the similarities and differences between mouse and human respiratory development. This introduction discusses our understanding of human lung development during each stage of respiratory development, focusing on signaling and transcriptional networks that regulate the developing human respiratory system. It also discusses the current state of human model systems to accurately model human respiratory development and disease and highlight the challenges that remain.

1.2 The Embryonic Stage of Respiratory Development

1.2.1 Contribution of Cellular Signaling Pathways to Respiratory Endoderm Specification

The respiratory system is specified at E9.5 in mice and at 4 weeks of gestation in humans as the trachea and primary lung buds separate ventrally from the esophagus in the anterior foregut endoderm (Zaw-Tun, 1982; Perl *et al.*, 2002; Que *et al.*, 2007). The respiratory system is first marked by the transcription factor NKX2-1 (Lazzaro *et al.*, 1991; Hawkins *et al.*, 2017), which is also necessary for lung specification (Minoo *et al.*, 1995). Respiratory specification in mouse models has been reviewed extensively (Warburton *et al.*, 2005; Cardoso *et al.*, 2006; Morrisey and Hogan, 2010; Rankin and Zorn, 2014; Swarr and Morrisey, 2015). These studies have identified many of the signaling pathways that are essential during respiratory specification and have been used as a framework to differentiate human pluripotent stem cells (hPSCs) into respiratory lineages *in vitro*. This strategy, known as ‘directed differentiation’, is an attempt to recapitulate a series of developmental events in a stepwise manner by modifying the growth factor signaling environment in the tissue culture dish. This approach has allowed us to gain an

appreciation of the signaling and transcriptional regulators that are necessary for respiratory specification in a human-specific context. The major developmental milestones for lung specification using directed differentiation include definitive endoderm differentiation (D'Amour *et al.*, 2006; Loh *et al.*, 2014), followed by anterior-posterior patterning into anterior foregut endoderm (Green *et al.*, 2011), at which point NKX2-1^{+} respiratory progenitor cells can be specified (Wong *et al.*, 2012; Gotoh *et al.*, 2014; Huang *et al.*, 2014, 2015; Dye *et al.*, 2015; Rankin *et al.*, 2016; Hawkins *et al.*, 2017; Serra *et al.*, 2017; Ng *et al.*, 2022).

Studies using directed differentiation from hPSCs as well as studies in animal models have stressed the importance of WNT signaling for initiating the expression of NKX2-1 from anterior foregut endoderm (Goss *et al.*, 2009; Harris-Johnson *et al.*, 2009; Rankin *et al.*, 2016; Szenker-Ravi *et al.*, 2018) (Figure 1-2A,B). However, activation of WNT signaling that induces NKX2-1 expression requires cooperation from multiple other signaling pathways (Figure 1-2A,B). The complex signaling network that induces the respiratory fate is dependent on retinoic acid (RA) signaling, which is required prior to respiratory specification and renders the ventral foregut endoderm competent to respond to cues that induce the respiratory lineage. The mechanisms through which RA signaling acts are just beginning to be unveiled, and current research suggests that RA signaling directly activates Sonic Hedgehog (SHH) in the endoderm (Huang *et al.*, 2014, 2015; Rankin *et al.*, 2016, 2018, 2021). In mice, SHH ligands emanating from the endoderm induce the expression of WNT ligands in the mesoderm, which signal back to the endoderm to activate NKX2-1 expression (Litingtung *et al.*, 1998; Motoyama *et al.*, 1998; Rankin *et al.*, 2016). NKX2-1^{+} cells have been induced from hPSCs without the addition of SHH signaling components to the media; however, since SHH signaling acts upstream of WNT in mice, it is possible that directed differentiation strategies using hPSCs bypass the need for SHH

components through the addition of exogenous WNT ligands. In mice and humans, BMP signaling represses SOX2 in the endoderm, which is required for the endoderm to properly respond to WNT ligands and express NKX2-1 (Domyan *et al.*, 2011; Rankin *et al.*, 2016). Genetic gain- and loss-of-function studies in mice have also established a role for FGF signaling during respiratory specification (Bellusci *et al.*, 1997; Min *et al.*, 1998; Sekine *et al.*, 1999; Serls *et al.*, 2005); however, like SHH, FGF has not played a prominent role in differentiation of hPSC into NKX2-1$^+$ respiratory progenitor cells, and its role in human respiratory specification remains unknown.

1.2.2 Signaling Involved in Self-organization of 3D Lung Models

Many directed differentiation protocols that induce anterior foregut endoderm lineages from hPSCs use 2D cultures; however, it is also possible to generate 3D anterior foregut endoderm structures, called spheroids, using directed differentiation techniques (McCracken, Catá, Crawford, Sinagoga, Schumacher, Rockich, Y.-H. Tsai, *et al.*, 2014; Dye *et al.*, 2015). Spheroids are immature multicellular tissue structures that arise during directed differentiation through unknown mechanisms and which mimic a primitive gut tube-like structure. Spheroids provide an opportunity to direct the differentiation of hPSCs into lung cells with the correct cellular organization. The cues that are needed to pattern hPSCs into 3D lung spheroids seem to require a different set of signals compared to cells grown in 2D. For example, Dye *et al.* has shown it is possible to derive 3D ventral anterior foregut structures that can give rise to mature lung lineages by simultaneously inhibiting SMAD, which is required for anterior foregut patterning, and by activating FGF4, WNT, and SHH, which are required for both inducing 3D spheroid formation and robust NKX2-1 expression (Dye *et al.*, 2015, 2016; Miller *et al.*, 2019). The necessity of FGF4, WNT, and SHH for the formation of 3D structures suggest that these

signaling pathways may be responsible for cell migration and patterning during respiratory fate specification in humans.

1.2.3 Different Signaling Pathways Contribute to Mouse and Human Respiratory Mesoderm Specification

In the mouse, respiratory mesoderm is *Nkx2-1*$^{-}$ but is marked by *Tbx4* and *Tbx5*, both of which are necessary for respiratory mesoderm development and specification of the lung and trachea (Arora, Metzger and Papaioannou, 2012). A study from Kishimoto *et al*. showed that WNT signaling originating from the mouse endoderm induces *Tbx4* expression in the primitive lung mesoderm independent of *Nkx2-1* expression (Figure 1-2A). Using mouse pluripotent stem cells and hPSCs, they showed that tracheal mesoderm (chondrocytes and proximal smooth muscle cells) could be specified from lateral plate mesoderm by BMP4 and WNT signaling in the mouse and SHH, BMP4, and WNT signaling in the human (Kishimoto *et al.*, 2020). This demonstrates that the primary molecular mechanisms responsible for tracheal mesoderm specification are different between the mouse and human as mouse tracheal mesoderm specification does not require SHH.

The same group was more recently able to derive multiple organ-specific mesenchyme lineages from hPSCs *in vitro* (Kishimoto *et al.*, 2022). After specification of mid primitive streak, specification of lateral plate mesoderm required treatment with TGFβ and WNT inhibitors as well as BMP4 while the addition of RA with these components was required specifically for foregut lateral plate mesoderm specification, and anteriorization of the foregut lateral plate mesoderm required a SHH agonist. Differentiation into anterior foregut splanchnic mesoderm then required continued treatment with TGFβ and WNT inhibitors, BMP4, RA, and a SHH agonist while also activating FGF signaling. Separation between esophageal and respiratory

mesoderm lineages from anterior foregut splanchnic mesoderm required initial treatment with RA and a SHH agonist and then the addition of Noggin (BMP inhibition) for esophageal mesoderm. For respiratory mesoderm specification, the addition of BMP4 was required for both steps while the second step required activation of WNT instead of BMP inhibition via Noggin. This data suggests that specification into respiratory mesoderm requires a complicated and coordinated effort between the WNT, SMAD, RA, FGF, and SHH signaling pathways. The mechanisms regulating specification into specific respiratory mesenchymal cell types are currently unknown.

1.3 The Pseudoglandular Stage of Respiratory Development

The pseudoglandular stage occurs between E10.5 to E16.5 in mice and 5 to 17 weeks of gestation in humans. This stage is defined by branching morphogenesis, where progenitor-rich lung bud tips begin to undergo repeated bifurcations to create the complex arborized network of the airways (Hogan and Yinalina, 1998; Cardoso *et al.*, 2006; Metzger *et al.*, 2008; Rawlins, 2011; Rock and Hogan, 2011; Dye, Miller and Spence, 2016; Miller and Spence, 2017; Zepp and Morrisey, 2019) (Figure 1-1B). Humans undergo extended rounds of branching relative to mice (17 – 21 in humans, 7 – 17 in mice) (Irvin and Bates, 2003; Metzger *et al.*, 2008), raising the possibility of regulatory divergence in human branching morphogenesis. Complex reciprocal signaling between the epithelium and mesenchyme during this stage creates a unique hurdle in characterizing the signaling pathways important for branching. Other changes in the lung during pseudoglandular development include the emergence of smooth muscle and vasculature, which both contribute to the environment that influences branching morphogenesis. This section discusses the emergence of lung cell types during branching morphogenesis, their role in establishing the lung microenvironment, and how these environments dictate local signaling.

1.3.1 Cellular Differentiation During Branching Morphogenesis

A significant event during the pseudoglandular stage is the specification of airway cell types in the lung epithelium. As branching tips of the epithelium continue to grow and bifurcate, bud tip progenitors leave progeny behind, which differentiate into airway cell types including basal, multiciliated, secretory, and neuroendocrine cells (Figure 1-3B). Lineage tracing in mice suggest there is a specific developmental window where bud tip progenitors preferentially give rise to airway cell types (Rawlins *et al.*, 2009; Yang *et al.*, 2018). Until recently, there was limited knowledge about how these processes differ in humans. Several groups performed scRNA-seq on human fetal lung samples (Nikolić *et al.*, 2017; Miller *et al.*, 2020), and these studies established important *in vivo* benchmarks of cellular transcriptional states that can be directly compared with *in vitro*-derived cells, providing a roadmap for developing directed differentiation approaches to generate specific airway cell types. For example, methods to direct the differentiation of bud tip progenitors to airway progenitors including $TP63^+$ basal cells and recently described fetal airway secretory (FAS) cells have been developed by manipulating SMAD signaling (Mou *et al.*, 2016; Miller *et al.*, 2020; Hawkins *et al.*, 2021) (Figure 1-3B). A recent study utilizing a scRNA-seq-based lineage tracing method in cultured human bud tip progenitors showed that FAS cells give rise to neuroendocrine cells and a subset of multiciliated cells while basal cells give rise to club cells and a separate subset of multiciliated cells (Figure 1-3B) (Conchola *et al.*, 2022). It is unclear if a FAS-like cell exists in murine lung as it has classically been thought that airway epithelium is derived from basal cells in mice.

The molecular mechanisms governing lineage decisions in the airway is still under investigation; however, inhibition of Notch signaling directs hPSC-derived lung epithelium to differentiate into multiciliated and neuroendocrine cells (Konishi *et al.*, 2016; Hor *et al.*, 2020)

(Figure 1-3B). There remains debate over the role of WNT on bud tip progenitor fate (Chen *et al.*, 2017; Katherine B. McCauley *et al.*, 2017a). Some groups conclude that high WNT signaling supports a proximal airway cell fate and other groups conclude it supports alveolar cell types. As organoid models continue to improve, coupled with single cell studies, it is likely that more questions can be answered about cell lineage specification in the human airway.

1.3.2 Regulation of Bud Tips During Branching Morphogenesis

Branching morphogenesis is a complex morphological process that relies on highly proliferative progenitor-rich distal bud tips of the developing lung. RNA-seq data on bud tips found differences in both gene and protein expression between human and mouse (Rawlins *et al.*, 2009; Dye *et al.*, 2016; Nikolić *et al.*, 2017; Miller *et al.*, 2019). In mice, bud tip progenitors express *N-myc*, *Id2*, and *Sox9* but are *Sox2* negative (Okubo, 2005; Rawlins *et al.*, 2009; Rockich *et al.*, 2013; Danopoulos *et al.*, 2016). This contrasts with humans where bud tip progenitors express *SOX9* in addition to *SOX2* (Nikolić *et al.*, 2017; Danopoulos, Alonso, *et al.*, 2018; Miller *et al.*, 2018) (Figures 2C, 3D). Loss of *SOX2* does not occur in human bud tips until the canalicular stage. In cultured human lung explants where RAC1 inhibition causes decreased $SOX9^{+}$/ $SOX2^{+}$ bud tip progenitors, there is also decreased epithelial proliferation and impaired branching (Danopoulos, Alonso, *et al.*, 2018). As *Sox2* has been shown to be essential for airway cell fates in mice (Gontan *et al.*, 2008), longer perdurance of *SOX2* expression in human bud tip progenitors may suggest that human bud tips retain the potential to differentiate into airway cell fates much later into development than in mice.

Studies of branching morphogenesis in the mouse have elucidated important mechanisms that regulate this process, which are reviewed extensively elsewhere (Morrisey and Hogan, 2010; Hines and Sun, 2014; McCulley, Wienhold and Sun, 2015; Prince, 2018). This section focusses

on comparing the signaling regulation of murine and human bud tips during branching. A thoroughly investigated signaling pathway in branching is FGF10, which is expressed in the mesenchyme near the most distal bud tips and is critical for branching and proximal-distal patterning in mice (Danopoulos, Thornton, *et al.*, 2018; Danopoulos, Shiosaki and Al Alam, 2019) (Figure 1-3A). *Fgf10*$^{-/-}$ mice do not undergo branching and conditional knock-outs of *Fgf10* or *Fgfr2* also disrupt lobe growth and have fewer branches (Min *et al.*, 1998; Sekine *et al.*, 1999; Abler, Mansour and Sun, 2008). In the developing human lung, *FGF10* is expressed from 10 – 21 weeks (Al Alam *et al.*, 2015; Danopoulos, Thornton, *et al.*, 2018; Danopoulos, Shiosaki and Al Alam, 2019) diffusely throughout the lung parenchyma (Danopoulos, Thornton, *et al.*, 2018; Jones *et al.*, 2019). Murine lung explants cultured with FGF10 show increased branching, while in contrast, human lung explants cultured with FGF10 exhibit enlarged buds and fewer branches (Danopoulos, Thornton, *et al.*, 2018; Danopoulos, Shiosaki and Al Alam, 2019). Human lung organoid models suggest that FGF10 is not required for bud tip progenitor maintenance, rather FGF7 appears to be the primary FGF ligand responsible for bud tip maintenance (Chen *et al.*, 2017; Nikolić *et al.*, 2017; Miller *et al.*, 2018; Rabata *et al.*, 2020). Nevertheless, FGF signaling during human lung development is complex and multi-faceted as other FGF ligands appear to play some role (Danopoulos, Thornton, *et al.*, 2018; Danopoulos, Shiosaki and Al Alam, 2019).

Another important signaling pathway during branching is WNT (Figure 1-3A), which has been shown to act upstream of Fgf10 in murine mesenchyme (Goss *et al.*, 2011; Yuan *et al.*, 2018). In mice, loss of both *Wnt2* and *Wnt2b* results in complete lung agenesis (Goss *et al.*, 2009), and conditional epithelial knock-out of *β-catenin* results in malformed distal airways with aberrant proximal airways (Mucenski *et al.*, 2003). Similarly, RAC1-mediated WNT inhibition

in human lung explants decreases branching and results in loss of bud tip progenitors, although the molecular mechanisms remain to be investigated (Danopoulos *et al.*, 2016; Danopoulos, Alonso, *et al.*, 2018). It was recently discovered that humans with mutations in the WNT activator *R-spondin 2* (*RSPO2*) exhibit lung agenesis (Szenker-Ravi *et al.*, 2018), which is a surprising contrast to murine lung, where *Rspo2* mutants have more mild branching defects (Bell *et al.*, 2008). The continued use and advancement of human *in vitro* models are required to fully appreciate the molecular mechanisms of FGF and WNT signaling in human lung branching morphogenesis.

1.3.3 Molecular and Mechanical Cues from the Mesenchyme During Branching Morphogenesis

The mesenchyme undergoes significant morphological changes as the branching epithelium continues to bifurcate and alter the landscape of the lung. We are only beginning to understand the diversity of mesenchymal cell types and changes they undergo during human lung development (Danopoulos *et al.*, 2019). Therefore, i*n vitro* human models of lung mesenchyme are less developed compared to epithelial models. Engineering approaches using microfluidic chambers with mouse lung explants as well as *in silico* modeling have begun to examine the changes that occur during branching morphogenesis and show promising innovation for human models (Varner *et al.*, 2015; Nelson *et al.*, 2017). In mice, it has been shown that both *Fgf10*$^{+}$ mesenchymal cells and *Pdgfra*$^{+}$ mesenchymal cells give rise to airway smooth muscle (Mailleux *et al.*, 2005), the latter through WNT2 and WNT7b signaling (Miller *et al.*, 2012). Blocking smooth muscle differentiation also prevents epithelial buds from bifurcating (Kim *et al.*, 2015), and it was recently shown that smooth muscle differentiation defines specific domains along the airways that propagate branches in mice (Goodwin *et al.*, 2022). In humans, α-SMA^{+}

smooth muscle cells support the proximal fate of the human airway and branching (Danopoulos, Alonso, *et al.*, 2018), although the signaling mechanisms involved are unknown (Figure 1-3A). More *in vitro* models using organoid, co-culture, and explant-like cultures will be required for understanding the signaling changes in the mesenchyme during human lung development. Single cell analysis will continue to help shed light on the complexity of the mesenchyme and identify key signaling factors involved in the morphing landscape of the lung.

1.4 The Canalicular Stage of Respiratory Development

During the canalicular stage of respiratory development, the lung transitions from generating airway (bronchi, bronchioles) to generating the gas-exchange units of the lung, the alveoli. This is characterized by the continued differentiation of bud tip progenitors towards alveolar fates (Shannon, Gebb and Nielsen, 1999; Rawlins *et al.*, 2009; Chang *et al.*, 2013; Desai, Brownfield and Krasnow, 2014; Laresgoiti *et al.*, 2016; Nikolić *et al.*, 2017; Miller *et al.*, 2018; Frank *et al.*, 2019) and by the formation of the bronchoalveolar duct junction (BADJ) in mice (Bal and Ghoshal, 1988; Alanis *et al.*, 2014), which demarcates airway-fated epithelial cells from alveolar-fated epithelial cells and can be identified in mice by the terminal border of *Sox2* expression (Cole *et al.*, 2004; Manwani *et al.*, 2010; Treutlein *et al.*, 2014). The existence of a BADJ in humans has not been demonstrated, but both human and mouse bud tips lack expression of the airway cell fate marker SOX2 prior to generating alveolar cells (Nikolić *et al.*, 2017; Danopoulos, Alonso, *et al.*, 2018; Miller *et al.*, 2018) (Figure 1-4A).

1.4.1 Alveolar Cell Specification During the Canalicular Stage

Classic models of alveolar development proposed that alveolar cell types are specified in a sequential manner, with bud tip progenitors giving rise to alveolar progenitors, which give rise

to alveolar type II (ATII) cells, which give rise to alveolar type 1 (ATI) cells (Adamson and Bowden, 1975). More recent studies of alveolar cell specification at the single cell level have proposed an alternative model that ATI and ATII cells are specified from a bipotent alveolar progenitor (Desai, Brownfield and Krasnow, 2014; Treutlein *et al.*, 2014), whose existence in mice was recently proved with lineage tracing strategies (Frank *et al.*, 2019). However, this latter study suggested that bipotent progenitors are rare and most likely remnant undifferentiated cells that remain at the end of branching morphogenesis (Frank *et al.*, 2019). It was further demonstrated that commitment to alveolar fates occurs much earlier than previously appreciated, taking place concurrently with branching morphogenesis, rather than afterwards. The majority of mature alveolar cells are the progeny of unipotent alveolar progenitor cells fated towards either an ATI or ATII cell early in development, with ATII cells being specified first at the most distal tip of the lung and ATI cells being specified just after ATII cells slightly more proximal (Frank *et al.*, 2019). Interestingly, in humans, markers of ATI and ATII cell fate aren't detected before 16 weeks of gestation (canalicular stage) and are not robust even at 20 weeks of gestation. It is likely that epithelial bud tip progenitors choose their eventual alveolar cell fate at the molecular level days before becoming morphologically and functionally distinct in mice, but alveolar specification may occur much later in humans. This data also poses a question about whether or not an alveolar progenitor cell state even exists; it is possible that ATI and ATII cells are directly specified from multipotent bud tip progenitors, obfuscating the timing of when it is appropriate to term a bud tip progenitor an alveolar progenitor.

In mice and humans, both paracrine signals from the mesenchyme and endocrine signals involving glucocorticoids appear to direct bud tip progenitors to give rise to alveolar cells (Deterding *et al.*, 1994; Shannon, Gebb and Nielsen, 1999; Jacob *et al.*, 2017; Yamamoto *et al.*,

2017; Sucre *et al.*, 2018; de Carvalho *et al.*, 2019; Bridges *et al.*, 2020). Premature human infants are frequently given glucocorticoids in order to speed the maturation of ATII cells such that they begin producing surfactant to have functional lungs (Liggins and Howie, 1972; Dluholucký, Babic and Taufer, 1976; Morrison *et al.*, 1978; Baud *et al.*, 2016; Roberts *et al.*, 2017). Although glucocorticoids are used to mature already specified alveolar cells in the human, studies using mice suggest that endocrine glucocorticoid signaling drives the formation of the BADJ. Interestingly, manipulation of glucocorticoid signal timing or strength alters the size of the future alveolar compartment of the lung without disrupting the appearance of mature alveolar cell types (Alanis *et al.*, 2014; Laresgoiti *et al.*, 2016), suggesting that glucocorticoid signaling acts to restrict the developmental potential of bud tip progenitors away from airway fates without being required for alveolar differentiation. Laresgoiti *et al.* showed that glucocorticoid signaling may interact with inflammatory pathways via STAT3 to initiate the switch from bud tip progenitors giving rise to airway cell types to alveolar cell types in the mouse (Laresgoiti *et al.*, 2016). Beyond this data, the signaling mechanisms that glucocorticoid signaling works through to propel alveolar formation and maturation is unknown, and given the clinical applications of glucocorticoid signaling in the developing human lung, understanding the precise role of glucocorticoid signaling in alveolar specification could have drastic impacts on preventing chronic respiratory disease in premature infants.

Signals originating locally from the mesenchyme are also involved in regulating the differentiation of bud tip progenitors into alveolar cells. Mesenchyme surrounding the bud tips in both humans and mice secrete FGF ligands (Bellusci *et al.*, 1997; Hirashima, Iwasa and Morishita, 2009; Danopoulos, Thornton, *et al.*, 2018). In mice, it has been shown that mesenchyme-derived FGF acts on the epithelium through KRAS to maintain the progenitor state

of bud tips (Chang *et al.*, 2013; Alanis *et al.*, 2014; Volckaert and De Langhe, 2015). Secretion of FGF from the mesenchyme is promoted by WNT ligands, which are thought to partially originate from the epithelium, creating a positive feedback loop (Goss *et al.*, 2009; Chen *et al.*, 2010; Goss *et al.*, 2011; Volckaert *et al.*, 2013, 2017). HIPPO signaling terminates branching morphogenesis and promotes alveolar differentiation through degradation of *β*-catenin in the epithelium, disrupting the WNT-FGF feedback loop and directing bud tip progenitors to differentiate (Mahoney *et al.*, 2014; Volckaert *et al.*, 2017) (Figure 1-4A). Active FGF and WNT signaling are known to be important for maintaining bud tip progenitor identity in the human as well (Nikolić *et al.*, 2017; Danopoulos, Alonso, *et al.*, 2018; Danopoulos, Thornton, *et al.*, 2018; Miller *et al.*, 2018); therefore, these pathways may perform analogous roles in maintaining the progenitor state of human canalicular stage bud tips. It is important to note that the specific FGF ligand(s) involved in human bud tip progenitor maintenance are likely different than those in mice (Danopoulos, Thornton, *et al.*, 2018). A role for HIPPO signaling in the human lung has not been examined.

1.4.2 Mesenchyme Development During the Canalicular Stage

While the diversity of mesenchymal cell types in the developing lung is still being uncovered (Guo *et al.*, 2019), two distinct mesenchymal populations in mice have been defined to undergo significant changes during the canalicular stage: *Fgf10*$^+$ mesenchymal cells and *Pdgfra*$^+$ mesenchymal cells. Importantly, the appearance of lipofibroblasts in the human lung has not been confirmed (Ahlbrecht and McGowan, 2014). However, in mice *Fgf10*$^+$ mesenchymal cells give rise to lipofibroblasts (LIFs), which are lipid droplet-containing mesenchymal cells that have a regulatory role during alveolar development (Al Alam *et al.*, 2015). Unlike during the pseudglandular stage when *Fgf10*$^+$ mesenchymal cells give rise to

myofibroblasts (MYFs), smooth muscle cells, and LIFs in the distal lung, the large majority of *Fgf10*$^{+}$ cells give rise to LIFs (and other unknown mesenchymal cell types) but not MYFs during the canalicular and later stages of development (El Agha *et al.*, 2014). Likewise, it was found that the majority of *Pdgfra*$^{+}$ mesenchymal cells give rise to MYFs, which lay down much of the ECM important for alveolar formation and function (Lindahl *et al.*, 1997; Green *et al.*, 2016; Li *et al.*, 2018b), during the canalicular and later stages of development (Li *et al.*, 2018b). TGFβ signaling negatively regulates FGF10 signaling in the mesenchyme to control the differentiation of mesenchymal progenitor cells to MYFs versus LIFs such that higher FGF10 signaling favors LIF identify and lower FGF10 signaling favors MYF identity (McQualter *et al.*, 2013; Al Alam *et al.*, 2015; A. Li *et al.*, 2016).

1.5 Saccular and Alveolar Stages of Respiratory Development

Sacculation and alveologenesis are the terminal stages of lung development, beginning late during development and completing sometime during the first decade of life. The saccular stage culminates with the formation of primitive alveoli called saccules. Saccules are further divided during alveologenesis through a process called septation that maximizes the area available for gas-exchange. Similar to previous stages of lung development, formation of saccules (sacculation) and their maturation into alveoli is driven by changes occurring in both the epithelium and mesenchyme. In the epithelium, ATI cells transition from a cuboidal to a squamous morphology and then stretch to 10x their original size to form the majority of the surface area within the alveolar epithelium (Yang *et al.*, 2016) (Figure 1-4B). ATII cells become highly proliferative and build specialized organelles dedicated to surfactant production called lamellar bodies (Figure 1-4B). Meanwhile, new cell types appear in the alveolar mesenchyme that secrete extracellular matrix (ECM) and further remodel it, thus contributing to development

of saccules and setting the stage for further septation during alveologenesis. Importantly, defects in these late stages of lung development in model organisms mimic features of bronchopulmonary dysplasia (Husain, Siddiqui and Stocker, 1998; Jobe, 1999; Makita *et al.*, 2008; Mitani *et al.*, 2009; Branchfield *et al.*, 2016), a disease prevalent in premature births that leads to chronic respiratory difficulty throughout life. Thus, the mechanisms that ensure proper sacculation and alveologenesis are of great importance for developing interventions that will rescue lung function in the neonatal ward.

1.5.1 Sacculation and Alveologenesis – Signaling Active in the Epithelium

Sacculation occurs relatively late in human gestation (3rd trimester), making access to human lung tissue at this stage rare. To circumvent this limitation, several groups have developed methods to differentiate hPSCs into progenitors of the lung epithelium (Gotoh *et al.*, 2014; Huang *et al.*, 2014; Dye *et al.*, 2015; Chen *et al.*, 2017; Miller *et al.*, 2018, 2019), which can then give rise to alveolar cell types, partially recapitulating the development of the lung epithelium during sacculation and alveologenesis. Methods also exist to generate more purified populations of alveolar cells called alveolospheres (Gotoh *et al.*, 2014; Hawkins *et al.*, 2017; Jacob *et al.*, 2017; Yamamoto *et al.*, 2017; Tamò *et al.*, 2018), which have already proven useful for modeling congenital disease of the alveoli (Korogi *et al.*, 2019; Leibel *et al.*, 2019) and alveolar injury (Heo *et al.*, 2019). These methods, although state-of-the-art, provide an incomplete picture of sacculation and alveolar development in humans because they either lack mesenchyme (Chen *et al.*, 2017; Jacob *et al.*, 2017; Miller *et al.*, 2018), require exogenous mesenchyme for alveolar differentiation (Gotoh *et al.*, 2014; Yamamoto *et al.*, 2017), generate immature alveolar cells stochastically (Huang *et al.*, 2014; Dye *et al.*, 2015; Chen *et al.*, 2017), or give rise to ATII cells only (Jacob *et al.*, 2017; Yamamoto *et al.*, 2017). Furthermore,

although the methods mentioned above generate alveolar cells, it is unknown whether these cells pass through intermediate states that represent the true signaling, timing, and cell fate trajectories that occur *in vivo*. Never-the-less, establishment of these *in vitro* models has provided insights into cues necessary for alveolar cell specification and maturation in humans.

Methods to generate alveolospheres generally follow directed differentiation paradigms in order to induce lung progenitors from anterior foregut endoderm progenitors, followed by purification of putative alveolar progenitors. Alveolar progenitors are placed into various media types, but common to many protocols is the stimulation of cyclic AMP as well as the WNT, FGF, and glucocorticoid signaling pathways (Gotoh *et al.*, 2014; Jacob *et al.*, 2017; Yamamoto *et al.*, 2017). These studies suggest that WNT, FGF, glucocorticoid, and cAMP-signaling pathways act to specify and/or mature alveolar cells in humans (Figure 1-4C). These findings are corroborated by more recent studies generating alveolar-like cells from human fetal bud tip progenitor organoids by manipulation of WNT and FGF signaling (Lim *et al.*, 2021).

A role for WNT in ATII cell specification and/or maturation is consistent with insights from animal models. During late sacculation and early alveologenesis, ATII cells exhibit an increase in WNT signaling activity that correlates with an expansion in ATII cell number (Frank *et al.*, 2016), and constitutive WNT increases ATII cell number while loss of *β-catenin* during sacculation reduces ATII cell number and leads to an increase in ATI cell number (Frank *et al.*, 2016). Together, this suggests that WNT signaling promotes the proliferation of ATII cells and may regulate the identity of alveolar progenitor progeny. The role of FGF and cAMP-signaling in alveolar cell fate specification in animal models has yet to be elucidated.

Studies from animal models also suggest a key role for the HIPPO-signaling pathway in promoting ATI cell fate. HIPPO-signaling pathway mouse mutants exhibit defects in saccular

architecture that phenocopy aspects of human emphysema and bronchopulmonary dysplasia (Makita *et al.*, 2008; Mitani *et al.*, 2009; Isago *et al.*, 2019). Of note, mutations in the HIPPO-signaling pathway leading to overactive TAZ activity generate lung epithelium with precocious and ectopic expression of markers of ATI cell identity (Nantie *et al.*, 2018), suggesting that defects in the saccular architecture reflect a specific role for TAZ in promoting ATI cell fate. Intriguingly, physical association between the lung epithelial transcription factor NKX2-1 and TAZ has been demonstrated (Park *et al.*, 2004), and more recently, NKX2-1 was demonstrated to perform a role in ATI cell specification distinct from its earlier role in specification of lung epithelium (Little, Gerner-Mauro, Flodby, Crandall, Borok, Akiyama, *et al.*, 2019), suggesting that NKX2-1 and TAZ may partner to drive development of ATI cells. Despite the important role of HIPPO signaling in lung epithelial progenitor specification (see canalicular section) and the development of ATI cells in mice, the dynamics of HIPPO signaling in human models of lung development are not yet known.

1.5.2 Sacculation and Alveologenesis – Contributions from the Mesenchyme

Three major populations of alveolar fibroblasts have been defined that guide the development of alveolar epithelium through sacculation during the formation of mature alveoli. Myofibroblasts (MYFs), expressing α-SMA, localize to developing septal tips where they remodel existing networks of elastin, which is necessary for proper formation of alveoli and provides elasticity for the lung during respiration (Boström *et al.*, 1996; Lindahl *et al.*, 1997; Hrycaj *et al.*, 2015; Branchfield *et al.*, 2016; Luo *et al.*, 2018). Cues for remodeling the lung ECM may be primarily physical, as stretching induces the activity of elastase (Kim *et al.*, 2015). MYFs are also thought to play a key role in driving secondary septation during alveologenesis (Boström *et al.*, 1996; Lindahl *et al.*, 1997; McGowan *et al.*, 2008; Li *et al.*, 2020). Similar to

MYFs, matrix fibroblasts are intimately associated with the saccule during its development and are distinguished from MYFs by high levels of PDGF-signaling activity and high levels of WNT5a production (Green *et al.*, 2016; Endale *et al.*, 2017). Matrix fibroblasts secret collagen and other ECM components (Endale *et al.*, 2017) that are essential for sacculation and alveolar maturation (Kida and Thurlbeck, 1980; Willem *et al.*, 2002; Bader *et al.*, 2005; Loscertales *et al.*, 2016; Fumoto *et al.*, 2019). In contrast to MYFs and matrix fibroblasts, which are thought to play more structural roles in sacculation and alveolar maturation, lipofibroblasts (LIFs) are thought to guide development and maturation of ATII cells through trafficking of lipids to ATII cells to assist in production of surfactant (McGowan and Torday, 1997). Interestingly, development of LIFs is dependent on signaling from ATII cells, which secrete PTHRP to antagonize Hedgehog and WNT signaling in LIF progenitors, which in turn leads to PPARy mediated transcription of Leptin and ADRP (Rubin *et al.*, 1997; Torday *et al.*, 2002; Torday and Rehan, 2006), molecules that induce surfactant production in ATII cells (Torday *et al.*, 2002; Rubin *et al.*, 2004), thus linking the co-maturation of ATII cells and LIFs.

hPSC-derived models of human lung development highlight the important contribution of mesenchyme to human sacculation and alveologenesis. For instance, hPSC-derived alveolospheres normally contain only ATII cells (Jacob *et al.*, 2017) but will give rise to cells with features of ATI cells when co-cultured with fetal lung fibroblasts (Figure 1-4C) (Gotoh *et al.*, 2014; Yamamoto *et al.*, 2017). Likewise, fetal-derived lung bud tip progenitors cultured *in vitro* readily differentiate into airway cell types but require co-culture with fetal lung mesenchyme for alveolar cell fate specification to occur (Nikolić *et al.*, 2017). Together, these studies suggest that human fetal lung mesenchyme provides cues that induce alveolar cell fates in human lung epithelium.

Mesenchyme-derived cues for alveolar cell fate specification are likely partially ECM-derived, as decellularized lung ECM supports the development of multiple alveolar cell types in hPSC-derived lung epithelium (Ghaedi *et al.*, 2013; Gilpin *et al.*, 2014). Notably, many hPSC-derived alveolospheres are grown in hydrogels that don't necessarily recapitulate the properties of the lung ECM during alveolar development. How ECM instructs alveolar differentiation is not known; however, given the mechanosensitivity of the HIPPO signaling pathway, the evidence for a central role of HIPPO signaling in ATI cell specification and maturation (Mitani *et al.*, 2009; Mahoney *et al.*, 2014; Nantie *et al.*, 2018; Isago *et al.*, 2019; Volckaert *et al.*, 2019), and a recent study showing that IGF1R controls mechanosignaling in myofibroblasts that is required for alveologenesis (He *et al.*, 2021), it is tempting to speculate that an ECM to HIPPO signaling axis guides the development of ATI cells *in vivo*. Interestingly, a recent epigenetic study in mice shows that loss of YAP/TAZ redirects NKX2-1 from ATI-specific to ATII-specific binding sites (Little *et al.*, 2021). A greater understanding of the roles of mesenchyme during sacculation and alveologenesis will be essential to recapitulate cues that instruct hPSC-derived lung epithelium to specify alveolar cells.

1.6 The Need for Models of Human Lung Development

We can never fully understand the unique aspects of human respiratory development without the use of *in vitro* model systems. In order to continue answering unknown questions in human respiratory development and properly model disease and genetic defects, several challenges must be overcome. For example, most *in vitro* human lung model systems are still overly simplistic, where the epithelium is cultured alone and relies on the addition of signaling components to media, or where epithelium is co-cultured with poorly characterized mesenchymal cells that organize in an unclear way. Neither of these approaches meticulously

recapitulate an *in vivo* environment, and it would be invaluable to develop model systems where the mesenchyme and epithelium are cultured together in the correct organization. It is also important to note that these systems often lack a functional vasculature and a nervous system, although efforts to improve complexity have been reported recently (Tan *et al.*, 2017; Holloway, Capeling and Spence, 2019). Access to developing human tissues as well as the advancement of technologies such as scRNA-seq coupled with *in situ* hybridization and immunofluorescence have begun to provide temporal and spatial gene expression patterns and have laid a strong foundation for the description of cell types, cell type-associated gene expression signatures, transcription factors, and signaling pathway components (Treutlein *et al.*, 2014; Brazovskaja, Treutlein and Camp, 2019; Miller *et al.*, 2020; Travaglini *et al.*, 2020). Translation of genetic manipulation techniques such as CRISPR to *in vitro* human model systems is evolving and will be instrumental to the functional understanding of signaling pathways during human lung development (Howden *et al.*, 2019; Sun *et al.*, 2021).

In addition to molecular mechanisms and cellular functions guiding lung development, it is also appreciated that mechanical cues play important roles in lung development and function; thus, establishing complex *in vitro* human model systems that incorporate and/or mimic aspects cellular, signaling, and biomechanical cues important for human lung development remains a critical obstacle. Current challenges include incorporating mechanical forces that occur during lung development, such as local forces involved in branching morphogenesis, peristaltic contractions observed in the developing lung, blood sheer stress, transmural pressure, and surface tension (Kim *et al.*, 2015; Varner *et al.*, 2015; Nelson *et al.*, 2017; Goodwin *et al.*, 2022). It is likely that lung-on-chip (Stucki *et al.*, 2018) and microfluidic technologies (Nelson *et al.*, 2017) will serve as useful tools to understand the influence of mechanical forces on human lung

development. As all of these technologies continue to be integrated into *in vitro* human model systems of respiratory development, we will better appreciate the mechanisms conserved among species as well as the uniqueness of human biology.

In vitro models of human lung development will likely also play an important role in personalized medicine. With the ability to use cultured primary patient tissue or generate patient-derived induced pluripotent stem cells to generate human *in vitro* models, we have the capability to model human lung disease and perform large-scale screens for patient-specific reactions to toxins, new drugs, and therapies. This could serve as a powerful tool for diseases such as cystic fibrosis and COPD, where current therapies are often ineffective or can be extremely costly; personalized screens could save months of trial-and-error with various medications to determine the optimal drug regime for a patient (Ferkol and Schraufnagel, 2014; Pittman and Ferkol, 2015). As chronic lung disease is a major cause of death worldwide (Ferkol and Schraufnagel, 2014), the need for new therapies and better treatments is critical, and *in vitro* model systems of the human lung will provide a high-throughput opportunity to develop personalized treatments for lung diseases.

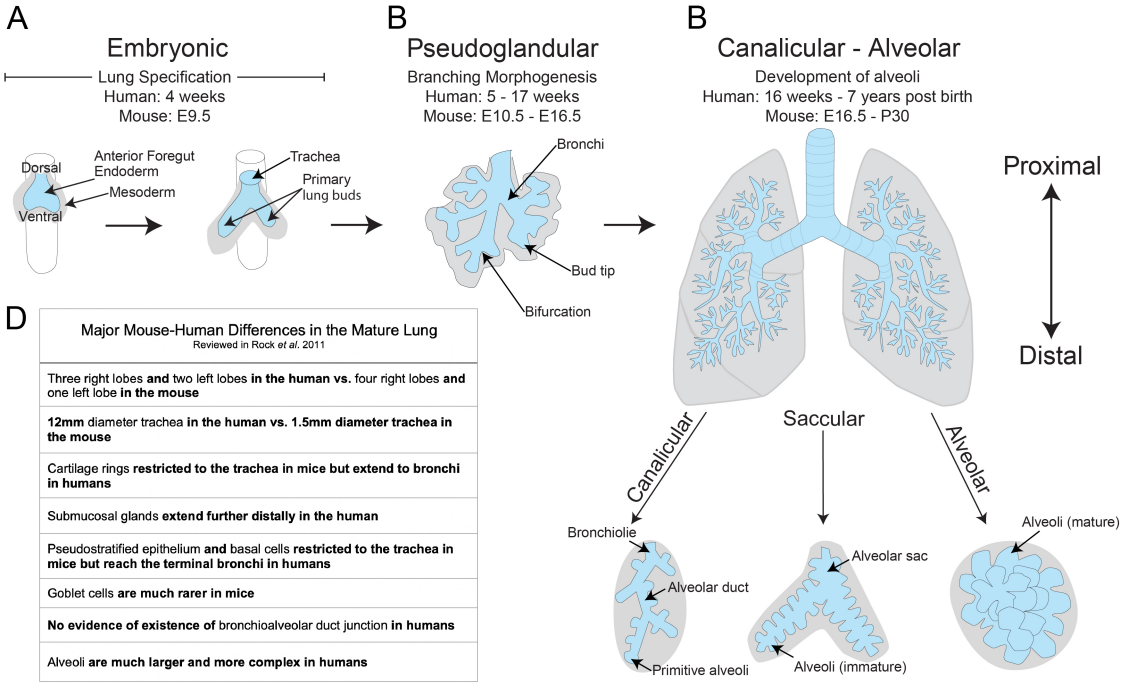
1.7 Overview of Thesis Research

Years of literature from murine and other animal models have laid a foundational body of work that has uncovered many of the intricacies of respiratory development, including morphological changes, cell lineage trees, and the molecular and mechanical mechanisms governing these processes. The more recent advent of human-specific model systems has begun to unveil the uniqueness of human biology, and although a great deal of progress has been made with respect to human respiratory development, there is much to learn. A significant hole in this field has been the lack of description of mesenchymal cell types present in the bud tip progenitor

domain and how each of these mesenchymal cell types uniquely influence bud tip progenitor behavior. Therefore, the first body of my research has explored the heterogeneity of cell types in the bud tip domain of the developing human lung and has investigated the cell signaling mechanisms behind mesenchymal cell influence on bud tip progenitor maintenance.

Additionally, progress towards understanding human respiratory development will require advances in the complexity of human-specific model systems as well as improvements in their ability to accurately model *in vivo* development. The use of primary tissue has significantly advanced our understanding of the developing human lung (Miller *et al.*, 2018, 2019, 2020; Little, Gerner-Mauro, Flodby, Crandall, Borok, Spence, *et al.*, 2019); however, there are significant limitations and concerns when it comes to using human fetal tissue for research. These include ethical considerations, regulatory roadblocks, limited access to the number of tissue specimens available and which researchers are able to access the specimens, and inaccessibility of later-stage tissue. These limitations and concerns have prompted scientists to develop models using hPSCs, which come in the form of embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs), to generate tissues of interest using directed differentiation. A major concern of hPSC-derived models is the ability of these models to accurately represent *in vivo* tissue or primary tissue-derived models. However, technologies such as scRNA-seq have given us the opportunity to benchmark hPSC-derived cultures to *in vivo* tissue. Given the limitations of primary tissue and the shortcomings of current hPSC-derived models, the second body of my research has aimed to leverage what is known about human respiratory development and in specific bud tip progenitor specification and maintenance, to improve hPSC-derived models of bud tip progenitors.

1.8 Chapter 1 Figures



D

Major Mouse-Human Differences in the Mature Lung Reviewed in Rock *et al.* 2011
Three right lobes **and** two left lobes **in the human vs.** four right lobes **and** one left lobe **in the mouse**
12mm diameter trachea **in the human vs. 1.5mm diameter trachea in the mouse**
Cartilage rings **restricted to the trachea in mice but extend to bronchi in humans**
Submucosal glands **extend further distally in the human**
Pseudostratified epithelium **and** basal cells **restricted to the trachea in mice but reach the terminal bronchi in humans**
Goblet cells **are much rarer in mice**
No evidence of existence of bronchioalveolar duct junction **in humans**
Alveoli **are much larger and more complex in humans**

Figure 1-1: Overview of the Five Stages of Human Respiratory Development. (A) During the embryonic stage, the lung arises ventrally from the anterior foregut endoderm, giving rise to two primary lung buds that branch off from the trachea into the surrounding mesoderm. **(B)** The pseudoglandular stage is characterized by the processes of branching morphogenesis, whereby distal bud tips undergo repeated rounds of bifurcations to create the arborized network of airways. **(C)** The alveoli, the air sacs that allow for gas exchange, are formed during the canalicular, saccular, and alveolar stages. This occurs as alveolar ducts form at the most distal airways, the bronchioles, which then form terminal sacs that will give rise to functional alveoli. **(D)** The adult mouse and adult human lungs contain many morphological differences.

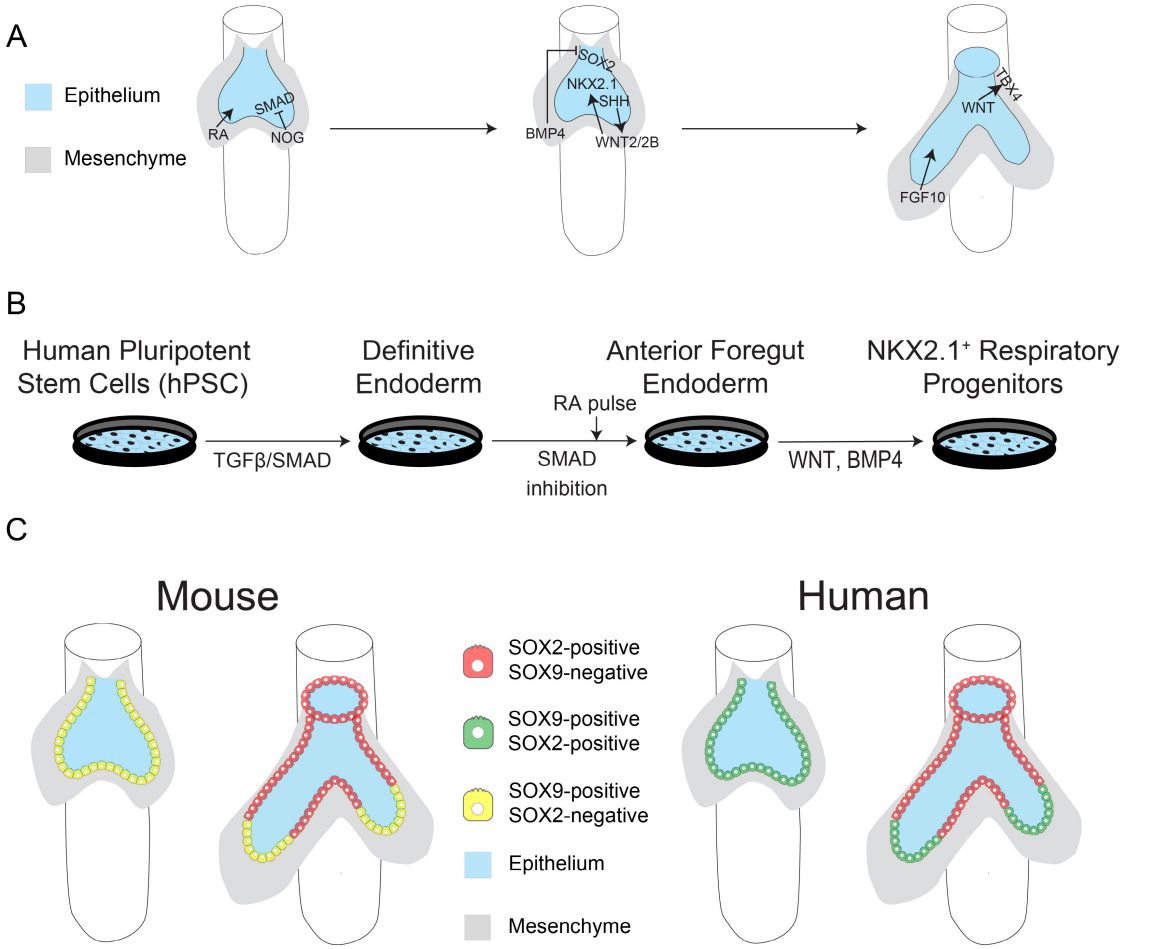


Figure 1-2: The Embryonic Stage of Respiratory Development. Signaling mechanisms required for respiratory specification *in vivo* **(A)** and *in vitro* **(B)**. TGFβ/SMAD signaling drives definitive endoderm specification, and SMAD inhibition through Noggin drives anteriorization of definitive endoderm. BMP4 from the mesoderm inhibits SOX2 expression in the mesoderm while SHH from the endoderm activates WNT ligands in the mesoderm that turn on NKX2-1 expression. RA is required for this process. WNT, BMP4, and SHH (*humans only*) from the endoderm specify the tracheal mesoderm, which is marked by TBX4. FGF10 is required for lung bud outgrowth. **(C)** The mouse respiratory epithelium (left) is initially made of $SOX9^+$ bud tip progenitors, which become restricted to the budded tips of the lung as the primary lung buds grow out from the trachea. The bud tip progenitors that are left behind proximally become $SOX2^+$. The human respiratory epithelium (right) is initially made of $SOX2^+/SOX9^+$ bud tip progenitors, which become restricted to the budded tips of the lung as the primary lung buds grow out from the trachea. The bud tip progenitors that are left behind proximally lose SOX9 expression but remain $SOX2^+$.

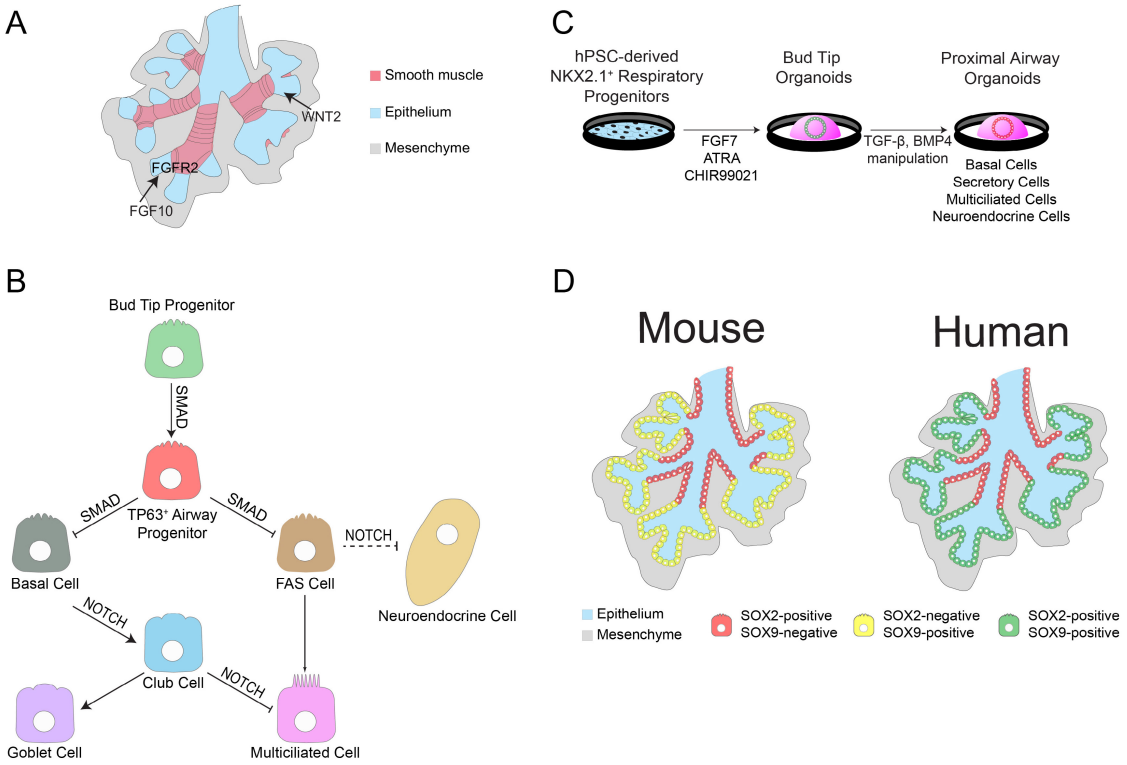


Figure 1-3: The Pseudoglandular Stage of Respiratory Development. (A) FGF10 signaling from the mesenchyme interacts with FGFR2 receptors on epithelium. WNT signaling from the mesenchyme also supports branching of the epithelium. An important physical cue for branching is the smooth muscle (pink). **(B)** The signaling pathways important for airway cell differentiation include SMAD signaling from bud tip progenitors to $TP63^+$ basal cells. Terminal differentiation into club/secretory cells is facilitated by active Notch signaling, and inhibition of Notch gives rise to multiciliated cells. Neuroendocrine cells also form from an epithelial progenitor through Notch inhibition, although it is less clear if they are specified directly from bud tip progenitors. **(C)** In vitro directed differentiation approaches have enabled expansion of bud tip progenitors as well as their differentiation into airway cell types using mechanisms that mimic in vivo signaling. **(D)** Organization of $SOX2^+$ cells and $SOX9^+$ cells vary between mice (left) and humans (right) where $SOX2^+$ cells are limited to proximal airway cells, but bud tip progenitors are $SOX9^+$ in mice and $SOX2^+/SOX9^+$ in humans.

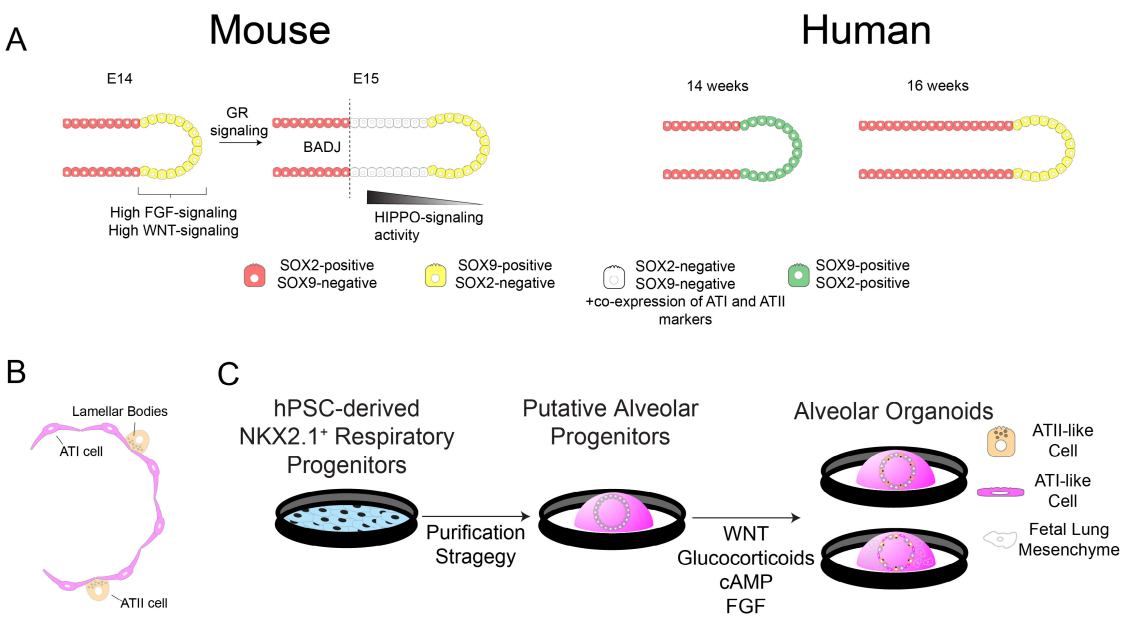


Figure 1-4: Canalicular, Saccular, and Alveolar Stages of Respiratory Development. (A) In mice, glucocorticoid signaling establishes the bronchoalveolar duct junction (BADJ), which demarcates the airway from the future site of alveoli formation. Bud tip progenitor identity is maintained by high levels of WNT and FGF signaling. High levels of HIPPO signaling in bud tip daughter cells born after BADJ formation leads to differentiation into alveolar progenitors, which co-express markers of ATI and ATII cells. In humans, whether BADJ formation occurs is unknown. Human bud tip progenitors downregulate SOX2 by week 16 of development, suggesting a change in the developmental potential of human bud tips occurs at 16 weeks. **(B)** Morphology of alveoli. ATI cells are thin to facilitate gas-exchange. ATII cells contain lamellar bodies, a surfactant producing organelle. **(C)** *In vitro* models of alveolar cell fate specification. Putative alveolar progenitors are purified from hPSC-derived NXK2.1^{+} lung epithelium. Conditions of high WNT, glucocorticoid, cAMP, and FGF signaling lead to the formation of ATII-like cells containing lamellar bodies. Incorporation of human fetal lung mesenchyme leads to the generation of ATI-like cells.

Chapter 2 R-Spondin^{+} Mesenchymal Cells Form the Bud Tip Progenitor Niche During Human Lung Development[2]

2.1 Introduction

Human lung development begins approximately four weeks post conception as the lung buds develop from the ventral anterior foregut endoderm. Soon after, the buds begin branching morphogenesis, leading to a network of epithelial tubes that include the trachea and series of bronchi that become progressively smaller, terminating at the gas-exchanging alveoli (Miller and Spence, 2017; Conway *et al.*, 2020). During branching morphogenesis, the tip of each budding branch possesses a population of transient epithelial progenitor cells called 'bud tip progenitors.' *In vivo* lineage tracing in animal models has shown that bud tip progenitors give rise to all cell types in the lung epithelium, including those that line the airways and alveoli (Rawlins *et al.*, 2009; Yang *et al.*, 2018). More recently, a population of bud tip progenitors in the human lung has been identified, and functional experiments have shown that they have the ability to generate a broad spectrum of lung epithelial cell types (Nikolić *et al.*, 2017; Miller *et al.*, 2018, 2020).

In mice, a specialized niche that supports bud tip progenitors is made up of surrounding mesenchyme that provides physical and biochemical support and determines whether bud tip progenitors will self-renew or differentiate into different epithelial cell types (Shu *et al.*, 2002; Weaver, Batts and Hogan, 2003; Alejandre-Alcázar *et al.*, 2007; Li *et al.*, 2008; Rajagopal *et al.*, 2008; Tsao *et al.*, 2008; Goss *et al.*, 2009; Morrisey and Hogan, 2010; McCulley, Wienhold and

[2] This chapter has been published: Hein, Renee F.C., Joshua H. Wu, Emily M. Holloway, Tristan Frum, Ansley S. Conchola, Yu-Hwai Tsai, Angeline Wu, et al. 2022. "R-SPONDIN2+ Mesenchymal Cells Form the Bud Tip Progenitor Niche during Human Lung Development." *Developmental Cell*, June. https://doi.org/10.1016/J.DEVCEL.2022.05.010.

Sun, 2015; Zepp and Morrisey, 2019; Riccetti *et al.*, 2020). Genetic gain- and loss-of-function studies have identified many of the signaling pathways important for bud tip progenitor maintenance and for determining cell-fate choices during differentiation (Morrisey and Hogan, 2010). Drawing from these studies, a minimal set of essential signaling cues required to maintain isolated human bud tip progenitors in long-term *in vitro* culture has been described (Nikolić *et al.*, 2017; Miller *et al.*, 2018); however, the specific mesenchymal cells and signaling components that make up the *in vivo* bud tip progenitor niche are unclear. Moreover, we have only begun to understand similarities and differences between animal models and humans (Danopoulos, Shiosaki and Al Alam, 2019; Conway *et al.*, 2020; Miller *et al.*, 2020).

Here, we investigate mesenchymal cell populations in the developing human distal lung from 8.5 – 19 weeks post-conception, a location where and time when the lung possesses an actively branching bud tip progenitor population. Leveraging single-cell RNA sequencing (scRNA-seq) data and unsupervised clustering analysis, we identified transcriptionally distinct mesenchymal cell populations in the distal lung domain during this time frame, including smooth muscle cells and four, non-smooth muscle mesenchymal cell clusters. Two clusters are highly enriched for expression of the secreted ligand *R-SPONDIN 2* (*RSPO2*). RSPO2 is an agonist of the canonical WNT signaling pathway (de Lau *et al.*, 2014). When RSPOs are absent in a cell's environment, ZNRF3 ubiquitinates WNTs co-receptor FZD, targeting it for proteasome degradation. However, when RSPO binds its receptor LGR (LGR4, 5, and 6 with high affinity), this forms a complex with ZNRF3, which frees FZD allowing WNT to bind and activate transcription of downstream target genes through activation (non-phosphorylation) of the transcriptional co-factor β-Catenin, which binds TCF/LEF transcription factors in the nucleus.

It is notable that *Rspo2*$^{-/-}$ mice have mild lung defects compared to lung aplasia that is seen in humans with *RSPO2* mutations (Bell *et al.*, 2008; Szenker-Ravi *et al.*, 2018). Indeed, mutations in human *RSPO2* are lethal (Szenker-Ravi *et al.*, 2018); however, the specific role of RSPO2 in the developing human lung has not been interrogated.

Based on the severe phenotype linked with *RSPO2* mutations in humans as well as the large population of *RPSO2*$^{+}$ mesenchymal cells identified in scRNA-seq data, we interrogated the spatial localization of *RPSO2*$^{+}$ cells along with the functional role of RSPO2. By fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF), we show that *RSPO2* is expressed in mesenchymal cells physically located adjacent to bud tip progenitors. On the other hand, SM22^{+} airway smooth muscle cells lack *RSPO2* expression and line the newly differentiating proximal (airway) epithelium in domains adjacent to bud tip progenitors. In addition, we found that the RSPO2 receptor *LGR5*, but not other *LGR* family members, is uniquely expressed in bud tip progenitors, as are other canonical WNT target genes such as *AXIN2*. Using *in vitro* lung explant cultures with functional inhibition experiments, we show that blocking endogenous RSPO signaling leads to a loss of the high WNT signaling environment in the bud tip domain and stochastic differentiation of bud tip progenitors into multiple airway, but not alveolar, epithelial cell types. LGR5 knock-down experiments in bud tip progenitor organoids suggest that RSPO2 acts through LGR5 in bud tips to maintain their progenitor fate. Lastly, we identified LIFR as a cell surface marker for RSPO2^{+} cells and show that FACS-isolated LIFR^{HI} cells support the ability of bud tip organoids to give rise to both airway and alveolar epithelium while LIFR^{-} cells only support airway differentiation in co-cultures. Collectively, this work identifies an RSPO2-producing niche cell in the human fetal distal lung mesenchyme and reveals a critical

RSPO2/LGR5-mediated WNT signaling axis that supports bud tip progenitor maintenance and multipotency throughout early lung development.

2.2 Results

2.2.1 Single Cell RNA Sequencing Identifies Mesenchymal Cell Populations in the Fetal Distal Lung

To interrogate the mesenchymal cells present in the developing human lung in the bud tip progenitor domain, we re-analyzed scRNA-seq data from the physically-isolated distal portion of 8.5 – 19 week post-conception lungs (n = 5 lungs) (Figure 2-1A) (Miller *et al.*, 2020). We used principal component analysis for dimensionality reduction, Louvain clustering, and UMAP for visualization (Wolf, Angerer and Theis, 2018; Becht *et al.*, 2019). Expression analysis of canonical marker genes identified major cell classes within the data, including epithelial, mesenchymal, immune, and endothelial cell types as well as a cluster of proliferating cells (Figures 2-2A, 2-2B). To specifically interrogate mesenchymal cell types that may comprise the bud tip progenitor niche, we extracted and re-clustered the non-vascular smooth muscle mesenchymal cell clusters 0 and 1.

Re-clustering identified 5 transcriptionally distinct mesenchymal cell clusters (Figure 2-1B). Differential expression analysis showed that cluster 4 contains cells expressing canonical markers of smooth muscle cells (e.g., *SM22*, *ACTA2*) and is enriched for expression of *HHIP*, *FOXF1*, and *WNT5A* (Figure 2-1D). Cells in clusters 0 and 1 have the most distinct transcriptome compared to smooth muscle cells and share very similar gene expression profiles, expressing enriched levels of *RSPO2*, *WNT2*, *CDO1*, *BMP5*, *LIFR*, and *FGFR4*, but differ in expression of *RPS4Y1*, *EGR1*, and *CA3* (Figure 2-1D). Cluster 2 is uniquely enriched for genes

including *SOX11*, *PDGFRα*, *STC1*, and *KCKN17* and cluster 3 is uniquely enriched for genes including *ELN*, *MGP*, *SERPINF1*, *AGTR2*, *FGF7*, and *BMP4*.

Of note, transcriptional differences between clusters 0 and 1 may largely be based on gestational age since we observed a nonequivalent distribution over time (Figure 2-1C). Consistent with this, we found that expression of *EGR1*, distinguishing cells in clusters 0 and 1, changes over gestational age (Figure 2-2C). Importantly, clusters 0 and 1 are enriched for the WNT signaling molecules *WNT2* and *RSPO2* (Figure 2-1D). Given that the growth of human bud tip progenitors in culture requires exogenous WNT stimulation (Nikolić and Rawlins, 2017; Miller *et al.*, 2018), we hypothesized these cells might make up an important part of the bud tip progenitor niche.

2.2.2 RSPO2$^+$ Mesenchymal Cells are Localized Adjacent to Bud Tip Progenitors

To spatially profile *RSPO2*$^+$ mesenchymal cells, we used FISH combined with IF on human fetal lung tissue sections spanning 8.5 – 19 weeks post-conception. *RSPO2*$^+$ mesenchymal cells were co-visualized with airway smooth muscle cells (SM22$^+$) because the scRNA-seq data showed that these populations are molecularly distinct. Co-FISH/IF for *RSPO2* and SM22 confirmed that these markers are expressed in different mesenchymal cell populations where *RSPO2*$^+$ mesenchymal cells are located physically adjacent to bud tip progenitors and SM22$^+$ airway smooth muscle cells line the more proximal, bud tip-adjacent epithelium (Figure 2-1E).

Additional FISH data showed that *RSPO2*$^+$ cells express *WNT2* and *FGFR4*, confirming scRNA-seq data and supporting previous reports revealing expression of *WNT2* and *FGFR4* in the distal lung near bud tip progenitors (Figure 2-2C) (Goss *et al.*, 2009; Miller *et al.*, 2012; Danopoulos, Alonso, *et al.*, 2018; Danopoulos, Shiosaki and Al Alam, 2019; Yu *et al.*, 2021).

FGFR4 expression was also found in bud tip progenitors (Figure 2-2C). Also complimenting scRNA-seq data, *PDGFRα* expression was found in both *RSPO2*$^+$ cells and in SM22$^+$ cells and appears particularly enriched in SM22$^+$ cells directly adjacent to bud tip regions (Figure 2-2D, arrowhead). Enriched co-expression of *AGTR2* with *BMP4* or *SERPINF1* (cluster 3 in Figure 2-1B) is evident in mesenchymal cells adjacent to SM22$^+$ smooth muscle cells, especially surrounding larger airways (Figure 2-2E). Based on scRNA-seq and FISH, the other *R-SPONDIN* transcripts were detected at much lower levels compared to *RSPO2* and not specifically localized near bud tip or differentiating epithelium (Figures 2-2F, 2-2J). In contrast to the distal lung domain, *RSPO2* expression is nearly absent from lung mesenchyme surrounding the bronchi and trachea and from all epithelia (Figures 2-1E, 2-2G). The proximity and specificity of *RSPO2*$^+$ mesenchyme to bud tip progenitors further suggested that RSPO2 comprises an important component of the bud tip progenitor niche.

2.2.3 LGR5 is Expressed in Bud Tip Progenitors

One mechanism by which R-SPONDIN proteins are known to amplify WNT signaling is by binding to LGR receptors and sequestering ubiquitin ligases RNF43 and ZNRF3, subsequently freeing the WNT receptor Frizzled from protein degradation (de Lau, Snel and Clevers, 2012; Niehrs, 2012; Chen *et al.*, 2013; de Lau *et al.*, 2014; Park *et al.*, 2018; Raslan and Yoon, 2019). To determine which cells RSPO2$^+$ mesenchymal cells may signal to, we used FISH to characterize the localization of *LGR* receptors within the fetal lung. *LGR5* (but not *LGR4* or *LGR6*) is highly specific to bud tip progenitors in the distal lung (Figures 2-1F, 2-2H). *LGR5* expression is largely excluded from mesenchymal cells and from differentiated epithelial cell types, except for a subset of basal cells in the more proximal airways (Figure 2-2H). In contrast, *LGR4* is expressed broadly throughout the mesenchyme in the distal and proximal lung and is

excluded from the distal epithelium while *LGR6* is expressed specifically in airway smooth muscle cells in the distal and proximal lungs (Figure 2-2H). In addition, the RSPO co-receptors *RNF43* and *ZNRF3* are bud tip-enriched (Figure 2-2I). The specific expression of *LGR5* in bud tip progenitors suggests that LGR5 may be a cognate receptor for RSPO2 present in the bud tip progenitor niche, and the enriched expression of *RNF43* and *ZNRF3* suggest these may also play a role in signaling.

2.2.4 WNT Target Gene Expression is Enriched in Bud Tip Progenitors

Given the expression pattern of *RSPO2* and *LGR5*, we predicted bud tip progenitors would display higher levels of WNT-mediated target gene expression compared to other cell types in the distal lung. Using *AXIN2* expression as a read-out for WNT signaling and SOX9 as a marker for bud tips (Rawlins, 2008), quantification of FISH data revealed that *AXIN2* is significantly enriched in bud tip progenitors compared to other cell types in the distal lung (Figure 2-1G). Based on our collective data, this further supports our hypothesis that RSPO2 from the mesenchyme may act on bud tip progenitors via LGR5 to support a high WNT signaling domain to maintain bud tip progenitors.

2.2.5 RSPO2-Mediated WNT Signaling in Bud Tips is Required for Proximal-distal Patterning

To test the necessity of RSPO2-mediated WNT signaling for bud tip progenitor maintenance, we used an adenovirus (ad) expressing the soluble ectodomain of LGR5 (hereafter termed LGR5 ECD ad), which was previously shown to bind and inhibit endogenous RSPO2, leading to reduced levels of WNT signaling in infected tissues (Yan *et al.*, 2017). We infected human fetal lung explants placed in an air-liquid-interface culture system with the LGR5 ECD ad

or a control ad encoding murine immunoglobulin IgG2a (Yan *et al.*, 2017) twice over a 4-day culture period. Both the control and the LGR5 ECD ad-infected explants grew over the 4-day culture period, and in both conditions, mesenchymal cells were retained, and the epithelium expanded (Figure 2-3A). We confirmed successful infection of the virus via antibody staining against murine IgG2a_Fc for the control and against FLAG for the LGR5 ECD ad (Figure 2-4A). Both viruses appeared to target the mesenchyme more than the epithelium and infect the edge of the explants more strongly than the center (Figure 2-4A), so we focused our analysis to the periphery of explants when possible.

Following 4 days of culture, the explants infected with the LGR5 ECD ad exhibited reduced staining for the distal bud tip marker SOX9 in the epithelium, with $SOX9^{+}$ cells making up 36.0% of total cells in the control but only 12.8% of cells in the LGR5 ECD ad-infected explants (Figure 2-3B). In comparison to *in vivo*, uncultured lung tissue of a similar gestational age and previous reports (Abler *et al.*, 2017; Miller *et al.*, 2018), the control ad-infected explants maintained proper SOX2 and SOX9 epithelial patterning, with $SOX9^{HI}/SOX2^{LOW}$ bud tips and $SOX9^{-}/SOX2^{HI}$ proximal (airway) epithelium (Figure 2-3B). Much of the epithelium in the LGR5 ECD ad-infected explants became $SOX9^{-}/SOX2^{HI}$ (Figure 2-3B), indicative of a loss of bud tip identity and differentiation into airway epithelium. WNT target gene expression, measured by *AXIN2* FISH, was also reduced in LGR5 ECD ad-infected explants compared to the control (Figure 2-3C), and non-phosphorylated (active) β-Catenin staining was more robust in control versus LGR5 ECD ad-infected explants (Figure 2-4B). *LGR5,* also a WNT target gene, was reduced via FISH (Figure 2-3D). In control ad-infected explants, *LGR5* was expressed and restricted to $SOX9^{+}$ bud tips (Figure 2-3D). Although *LGR5* was still detected in the few remaining bud tip progenitors in the LGR5 ECD ad-infected explants, its endogenous expression

appeared reduced relative to bud tip progenitors in control ad-infected explants, consistent with loss of bud tip progenitor identity (Figure 2-3D). Note that many ectopic *LGR5*$^{+}$ cells could be detected, revealing abundant expression in strongly-infected cells, which made it difficult to quantify changes in endogenous *LGR5* expression (Figure 2-3D).

RSPO2 expression was not affected by control or LGR5 ECD ad infection as expression was maintained at similar levels in SM22^{-} mesenchyme in both conditions (Figure 2-3E). There were no significant differences in KI67 expression between the control and LGR5 ECD ad-infected explants (Figure 2-4C). In some cases, independent of viral infection, explants grew abnormally large, leading to cell death or necrosis towards the center of the explant; however, Cleaved Caspase 3 (CCASP3) staining was low or absent in most explants (Figure 2-4D).

2.2.6 The Bud Tip Transcriptional Profile is Dependent on RSPO2-mediated Signaling in Bud Tips

Human fetal lung explants infected with control or LGR5 ECD ad for 4 days were dissociated, and single cells were sequenced via scRNA-seq. Louvain clustering and UMAP visualization revealed clusters of epithelial, mesenchymal, vascular smooth muscle, endothelial, neuroendocrine, and proliferating cells, which were identified by examining expression of canonical marker genes for these cell types (Figures 2-6A, 2-6C). There is also a cell cluster composed of only LGR5 ECD ad-infected cells, which appears to be clustered based on high expression of *LGR5* (Figures 2-6A – C). As a control, we sequenced non-infected explants from the same experiment, which contained the same general cell populations (Figures 2-6D – F). We noted that most clusters (with the exception of the LGR5^{+} cluster) possessed cells from each sample that were evenly distributed across clusters; however, part of the epithelial cell cluster showed separation between non-infected and control ad-infected cells compared to LGR5 ECD

ad-infected cells (Figures 2-6D, 2-6E). This separation was also noticed in the UMAP embedding excluding non-infected cells (Figures 2-6A, 2-6B).

To gain better resolution of changes occurring in the epithelium, the epithelial cell cluster (cluster 1) from the UMAP embedding that includes cells from control and LGR5 ECD ad-infected explants (Figure 2-6A) was extracted and re-clustered (Figure 2-5A). We identified cluster 0 as bud tip progenitors based on enriched expression of canonical bud tip progenitor marker genes (Figure 2-5C). Based on visual inspection of this cluster and expression of bud tip markers, we noted that cells from each sample as well as gene expression were not evenly distributed across the cluster (Figures 2-5B, 2-5C). To gain insights into possible differences between the control and LGR5 ECD ad-infected bud tip progenitors, cluster 0 was again extracted and re-clustered (Figure 2-5D). When re-clustered, there were 4 predicted clusters, with sub-cluster 2 primarily consisting of cells from the control ad-infected explants and sub-cluster 0 primarily consisting of cells from the LGR5 ad-infected explants (sub-clusters 1 and 3 contained both conditions) (Figures 2-5D, 2-5E). Bud tip progenitor genes (*SOX9*, *TESC*, *ETV5*, *CA2*) were more highly expressed in control ad-infected cells (Figure 2-5F). Individual cells from each sample were evaluated against a panel of the top 22 most differentially expressed genes from published *in vivo* bud tip progenitors (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022), thus assigning a "bud tip progenitor score" to every cell (Holloway, Wu, *et al.*, 2020). Consistent with the reduced expression of individual bud tip progenitor genes in LGR5 ECD ad-infected explants, cells from LGR5 ECD ad-infected explants scored lower for having a bud tip progenitor identity (Figure 2-5G). Together, this data suggests that loss of endogenous RSPO2 activity during human lung development causes a reduction of bud tip progenitor gene expression.

2.2.7 RSPO2-potentiated Signaling in Bud Tips Prevents Differentiation into Airway Cell Types

Based on the reduced bud tip progenitor transcriptional profile in LGR5 ECD ad-infected explants compared to control ad-infected explants (see Figure 2-5), and because epithelial SOX2 expression was higher in the LGR5 ECD ad-infected explants compared to the control via IF (Figure 2-3), we hypothesized that the bud tips from these explants might be differentiating into airway lung cell types. Using scRNA-seq data from control and LGR5 ECD ad-infected explants and complementary IF and FISH stains on tissue sections from these explants, we evaluated expression of proximal (airway) and distal (alveolar) differentiated cell type markers. When possible, we calculated cell type scores (Holloway, Wu, *et al.*, 2020) for specific cell types by evaluating the average expression of the published 50 most differentially expressed genes from *in vivo* cells of the listed cell type (Miller *et al.*, 2020). We limited the scRNA-seq analysis to cells within the bud tip cluster (Figure 2-5) to specifically determine how the bud tips in the LGR5 ECD ad-infected cells are changing relative to controls.

The biggest difference in cell type marker expression from the control and LGR5 ECD ad-infected explants was expression of secretory cell genes. By FISH, the secretory cell marker *SCGB3A2* was properly localized to airway structures and differentiating, bud tip-adjacent cells in control ad-infected explants, as seen in *in vivo* tissue (Miller *et al.*, 2020) (Figure 2-7Ai.). In contrast, remaining $SOX9^{+}$ bud tip regions in LGR5 ECD ad-infected explants expressed *SCGB3A2* mRNA, suggesting airway differentiation (Figure 2-7Ai.). In agreement with this, scRNA-seq data revealed that LGR5 ECD ad-infected cells from the bud tip cluster had higher cell type scores for secretory progenitor and club cells, and to a lower extent, for goblet cells compared to the control (Figures 2-7Aii. – v.). Additionally, there was an increase in $TP63^{+}$ cell

numbers in the LGR5 ECD ad-infected explants compared to the controls (Figures 2-7Bi. – ii.), although basal cell scores for LGR5 ECD and control ad-infected cells were similar (Figure 2-7Biii.). Although the cell type scores for neuroendocrine and multiciliated cells were moderately increased in cells from LGR5 ECD ad-infected explants, $CHGA^+$ cells and $FOXJ1^+$ cells were only sparsely detected by IF in both conditions (Figure 2-8A). Overall, LGR5 ECD ad-infected explants had higher expression of airway lineage markers compared to control ad-infected explants, particularly with respect to secretory cell lineages.

To determine if bud tip progenitors in LGR5 ECD ad-infected explants were undergoing general differentiation or differentiating specifically to airway cell types, we examined LGR5 ECD ad-infected bud tips for the presence of differentiated alveolar cell types. By scRNA-seq, *SFTPB* was increased in LGR5 ECD ad-infected explants (Figure 2-8Bi.). It has recently been shown that *SFTPB* is expressed in some airway secretory cell types in addition to alveolar cells (Miller *et al.* 2020), which could account for higher expression in the LGR5 ECD ad-infected explants. Remaining alveolar cell type markers were similarly expressed between control and LGR5 ECD ad-infected explants (Figures 2-8Bii. – vii.). Overall, there does not appear to be a differentiation bias of bud tips into alveolar cell types between the two conditions, rather, we see an increase in differentiation towards airway cell types upon LGR5 ECD ad-infection.

2.2.8 LGR5 can Respond to RSPO2 to Maintain Bud Tip Progenitor Cell Fate

Based on the specificity of *LGR5* expression in bud tips, we hypothesized that exogenous RSPO2 would be able to maintain the bud tip progenitor fate in human fetal-derived bud tip progenitor organoids (BTOs) (Miller *et al.*, 2018). Therefore, we cultured BTOs in the presence of recombinant human RSPO2, WNT3a-afamin conditioned media (Mihara *et al.*, 2016; Nanki *et al.*, 2018), or both, and included a positive control (+CHIR99021, shown to maintain BTOs)

(Miller *et al.*, 2018) and negative control (-CHIR99021; CHIR). After 4 days of culture, both conditions containing RSPO2 maintained both cystic structures (similar to the positive control) and denser structures (similar to the negative control) while bud tips cultured in WNT3a alone contained mostly dense structures (Figure 2-9A). qRT-PCR analysis confirmed that the addition of RSPO2 maintained expression of bud tip genes and *AXIN2* at levels close to the positive control and at higher levels than the negative control or WNT3a-only condition; however, airway genes *SCGB3A2* and *TP63* were increased in all conditions compared to the positive control (Figure 2-9B). This suggests RSPO2 can act to maintain bud tip progenitors in BTOs but is less robust at doing so than the positive control containing CHIR.

Next, we confirmed that *LGR5* is expressed in BTOs when cultured in published media containing FGF7, ATRA, and CHIR (Miller *et al.*, 2018) (Figure 2-9C). To more directly test the requirement of LGR5 for maintaining a high WNT environment in bud tip progenitors, we used a published shRNA against LGR5 (Jha *et al.*, 2017) to knock-down *LGR5* in BTOs (Figure 2-9D). Cohorts of scrambled controls and LGR5 shRNA BTOs were switched into media containing FGF7 and ATRA either with or without CHIR (positive/negative control) or with RSPO2 in place of CHIR. After 4 days of culture, qRT-PCR analysis of bud tip markers and airway markers revealed that removing CHIR in scrambled controls led to a loss of bud tip markers (*NPC2*, *TESC*, *SFTPC*) that could be rescued by the addition of exogenous RSPO2. In contrast, RSPO2 was unable to rescue loss of bud tip markers in LGR5 shRNA BTOs. No differences were observed for expression of airway markers (*SCGB3A2*, *TP63*) (Figure 2-9D). This data suggests that bud tip progenitors with normal levels of LGR5 can respond to RSPO2 to maintain the bud tip progenitor fate, whereas LGR5 knockdown renders bud tips unable to respond to exogenous RSPO2.

2.2.9 Isolated RSPO2^{+} Mesenchymal Cells Support Bud Tip Multipotency in Organoid Co-cultures

To determine how the RSPO2 and SM22 mesenchymal cell populations differentially regulate bud tip progenitor behavior in culture, we performed 3D co-cultures with BTOs (Miller *et al.*, 2018) and isolated RSPO2^{+} or SM22^{+} mesenchyme in Matrigel. In order to use fluorescence activated cell sorting (FACS) to isolate these populations, we used published approaches to find putative cell surface markers that are co-expressed in RSPO2^{+} cells [SurfaceGenie – (Waas *et al.*, 2020)], which identified LIFR. We confirmed that *LIFR* is strongly co-expressed with *RSPO2* in mesenchymal cells by scRNA-seq (Figure 2-10A). To specifically enrich for LIFR^{HI} (RSPO2^{+}) and LIFR^{-} (SM22^{+}) mesenchyme, we used a combinatorial staining approach that allowed for the isolation of non-epithelial (EPCAM^{-}), non-endothelial (CD31^{-}), LIFR^{HI} or LIFR^{-} cells using FACS (Figure 2-10B). qRT-PCR analysis on the isolated populations confirmed that genes enriched in *RSPO2*$^{+}$ mesenchyme (*RSPO2*, *FGFR4*) were enriched in the isolated LIFR^{HI} population and genes enriched in airway smooth muscle cells (*SM22*, *FOXF1*) were enriched in the isolated LIFR^{-} population (Figure 2-10C). We also confirmed that epithelial cells and endothelial cells were successfully depleted from the LIFR^{HI} and LIFR^{-} populations (Figure 2-10C).

BTOs were cultured with LIFR^{HI} or LIFR^{-} mesenchymal cells directly after FACS-isolation in a media including FGF7 and ATRA (Miller *et al.*, 2018), but excluding any WNT ligands or small molecule activators. A positive control without mesenchyme but with CHIR as well as a negative control without mesenchyme or CHIR were included in each experiment. After 10 – 11 days of bud tip/mesenchyme co-culture, the mesenchyme caused condensation of the Matrigel and bud tips into a more tightly compacted structure compared to bud tips cultured

in Matrigel without mesenchyme; however, this was much more pronounced in the $LIFR^-$ co-cultures (Figure 2-11A). Positive control (+CHIR) bud tips retained a normal, cystic phenotype while negative control (-CHIR) bud tips became dense, as expected (Figure 2-11A) (Miller *et al.*, 2018).

By FISH, $RSPO2^+$ cells were found in the $LIFR^{HI}$ co-culture after 11 days of culture (Figure 2-11Bi., top panel). Surprisingly, $SM22^+$ cells were also detected in the $LIFR^{HI}$ co-culture at this time point, but at lower numbers and expression levels compared to the $LIFR^-$ co-culture (Figure 2-11Bi., top panel). Within $LIFR^{HI}$ co-cultures, *RSPO2* expression was largely absent from $SM22^+$ cells while *RSPO2* expression was almost entirely absent from the $LIFR^-$ co-culture, as expected (Figure 2-11Bi., top panel). *PDGFR*α expression was found throughout the mesenchyme in both co-cultures, similar to the *in vivo* lung (Figure 2-11Bi., 2-2D). Interestingly, when isolated $LIFR^{HI}$ cells are cultured in 2D, they largely maintain their expression profile in short-term cultures (6 hours) but gain expression of airway smooth muscle cell markers by 7 days in culture, whereas isolated $LIFR^-$ cells maintain smooth muscle cell markers in short-term and long-term cultures (Figures 2-10D, 2-10E). This suggests that $LIFR^{HI}/RSPO2^+$ cells can give rise to $LIFR^-/SM22^+$ cells in 2D culture conditions. To corroborate this, we also performed pseudotime analysis on scRNA-seq data of 8.5 – 19-week human fetal distal lung mesenchyme and found predicted lineage trajectories from $RSPO2^+$ cells to smooth muscle cells (Figure 2-10F) (Haghverdi *et al.*, 2016).

By IF, we observed that $aSMA^+$ cells (a marker for smooth muscle) in the $LIFR^{HI}$ co-culture surrounded epithelium that had low levels of SOX9 and had high expression of the airway marker SOX2, while epithelial cells surrounded by $RSPO2^+$ cells retained SOX9 expression (Figures 2-11Bi. – iii.). Epithelium in the $LIFR^-$ co-cultures was $SOX2^{HI}$ and $SOX9^{-\text{ or}}$

^{LOW} (Figure 2-11B), indicative of bud tips undergoing airway differentiation. Bud tip genes *SOX9*, *ETV5*, *NPC2*, and *LGR5* were increased in the LIFR^{HI} co-cultures compared to the LIFR^{-} co-cultures by qRT-PCR (Figure 2-12Bi.).

Next, we wanted to investigate if differentiated proximal (airway) or distal (alveolar) cell types were present in each co-culture. Airway cell types, including secretory, basal, multiciliated, and goblet cell markers were detected by IF and qRT-PCR at levels above the positive control BTOs in both co-culture conditions but were more significantly upregulated in LIFR^{-} co-cultures (Figures 2-11A, 2-12Bii.). Interestingly, the alveolar cell type markers *SFTPC*, *ABCA3*, and *RAGE* were upregulated in the LIFR^{HI} co-culture compared to all other conditions by qRT-PCR (Figure 2-12Biii.). At the protein level, SFTPC was detected at high levels in the LIFR^{HI} co-culture, and some SFTPC^{+} cells co-expressed other alveolar type II proteins, including SFTPB and HTII-280 (Figures 2-12Ai., 2-12Aii., 2-12Avi.). The alveolar type I marker RAGE was also sporadically detected in the LIFR^{HI} co-culture (Figure 2-12Av.). This data suggests some SOX9^{+} bud tips progenitors in the LIFR^{HI} co-culture have begun alveolar differentiation. Abundant SFTPB^{+} cells were detected in the LIFR^{-} co-culture and in the negative control, but they co-expressed *SCGB3A2* (Figure 2-12Aii.), indicative of an airway secretory cell (Miller *et al.*, 2020). Together, this data shows that smooth muscle cells (LIFR^{-}/SM22^{+}) provide cues for differentiation into airway cell types while LIFR^{HI}/RSPO2^{+} mesenchymal cells support bud tip differentiation into alveolar cell types and airway cell types, potentially through RSPO2^{+} differentiation into smooth muscle cells for the latter.

In addition to differentiation, we wanted to determine if there were changes in cell proliferation between the two co-culture conditions. We found that KI67 expression was significantly upregulated in the LIFR^{HI} co-culture (Figures 2-11Ci., 2-11Ciii.). This correlated

with higher *AXIN2* expression in the $LIFR^{HI}$ co-culture compared to the $LIFR^{-}$ co-culture and negative control (Figures 2-11Cii., 2-11Civ.). This data suggests that $RSPO2^{+}$ mesenchymal cells support a high WNT signaling niche conducive for self-renewal (proliferation) and differentiation of bud tip progenitors into both airway and alveolar epithelium.

2.3 Discussion

The lung development field has well-established literature interrogating the diversity of cell types and functions in the epithelium; however, less is known with respect to the developing lung mesenchyme, and particularly in the context of the human lung. Most of what is known about the role of the lung mesenchyme during development has come from animal models and has primarily focused on understanding airway smooth muscle cells or mesenchymal heterogeneity and function during alveolar and later stages of development (Torday, Torres and Rehan, 2003; Chen *et al.*, 2012; McQualter *et al.*, 2013; El Agha and Bellusci, 2014; Li *et al.*, 2015, 2018a, 2020; Green *et al.*, 2016; Zepp *et al.*, 2017; Endale *et al.*, 2017; Kishimoto *et al.*, 2018; Wu *et al.*, 2018; Noe *et al.*, 2019; Guo *et al.*, 2019; Bridges *et al.*, 2020; Riccetti *et al.*, 2020; Gouveia *et al.*, 2020; Han *et al.*, 2020; Kalucka *et al.*, 2020; Negretti *et al.*, 2021; Goodwin *et al.*, 2022). There are also far fewer similar studies in humans (Rehan *et al.*, 2006; Danopoulos, Shiosaki and Al Alam, 2019; Du *et al.*, 2019; Leeman *et al.*, 2019; Shiraishi, Nakajima, *et al.*, 2019; Shiraishi, Shichino, *et al.*, 2019; Danopoulos *et al.*, 2020; Güney *et al.*, 2021; Goodwin *et al.*, 2022). Nevertheless, recent advances in single-cell analytical tools and *in vitro* human-specific model systems have allowed us to begin addressing unknowns in human lung development (Treutlein *et al.*, 2014; Brazovskaja, Treutlein and Camp, 2019; Du *et al.*, 2019; Danopoulos *et al.*, 2020; Kishimoto *et al.*, 2020; Miller *et al.*, 2020; Travaglini *et al.*,

2020; Yu *et al.*, 2021). Here, we aimed to understand how mesenchymal cells are involved in creating an epithelial bud tip progenitor cell niche.

Single-cell RNA sequencing analysis predicted airway smooth muscle cells and four additional non-smooth muscle populations that have similar but not identical gene expression profiles in the developing human lung. We observed that two mesenchymal cell clusters identified by scRNA-seq are enriched for expression of the WNT-agonist *RSPO2* throughout the developmental time frame analyzed, and spatial localization showed that *RSPO2* is expressed adjacent to the bud tip domain. It is known that high WNT signaling conditions are necessary for the maintenance of the bud tip progenitor state *in vitro* (Nikolić *et al.*, 2017; Miller *et al.*, 2018; Rabata *et al.*, 2020), which made this cell population a strong bud tip-associated mesenchymal cell candidate.

By using human tissue specimens and human *in vitro* model systems to explore this cell population further, we provide evidence that $RSPO2^{+}$ mesenchymal cells signal to bud tip progenitors through LGR5 to maintain a high WNT signaling zone in the bud tips. We show that this RSPO2/LGR5-mediated WNT signaling niche provides support for the bud tips to maintain their progenitor state and give rise to both alveolar and airway cell types. The process of RSPO ligands signaling through LGR receptors to maintain high WNT signaling in progenitor and stem cells has been described in other organs and tissues, such as in the intestinal crypt, skin, hair follicle, and mammary tissue (Barker *et al.*, 2007; Jaks *et al.*, 2008; Trejo *et al.*, 2017; Yan *et al.*, 2017; Dame *et al.*, 2018; Baulies, Angelis and Li, 2020; Holloway, Czerwinski, *et al.*, 2020). However, an alternative mechanism where RSPO2 acts independently of LGRs, and instead through RNF43/ZNRF3, has been described in the context of limb development (Szenker-Ravi *et al.*, 2018). Work in cell lines has shown that RSPOs can act through heparan sulfate

proteoglycans, rather than LGRs, to promote WNT signaling (Dubey *et al.*, 2020). It is possible that other receptors may act independent of or synergistically with LGR5 to promote WNT signaling in the bud tips, and further genetic gain- and loss-of-function experiments would be required to test this.

Our data suggests a potential differentiation bias towards airway secretory cells when the RSPO2/LGR5-mediated WNT signaling axis is disrupted in bud tips. However, previous data using cultured bud tip progenitors shows that removal of the WNT component of the media causes an upregulation of secretory cell markers, suggesting secretory cells may just be a default differentiation state (Miller *et al.*, 2018). Moreover, in mice, multiciliated, goblet, and neuroendocrine cells appear later in development compared to secretory and basal cell types (Rock *et al.*, 2009; Treutlein *et al.*, 2014; Ardini-Poleske *et al.*, 2017; Montoro *et al.*, 2018; Miller *et al.*, 2020). Our short-term (4-day) cultures may not have provided enough time for non-secretory cell types to emerge.

Co-culture experiments with bud tip organoids and isolated $RSPO2^+$ mesenchymal cells or $SM22^+$ smooth muscle cells support the idea that $RSPO2^+$ mesenchymal cells provide cues for bud tip progenitor maintenance and differentiation towards alveolar cells while smooth muscle cells provide cues for bud tip differentiation into airway cells. Additional experiments culturing these mesenchymal populations separately in 2D short and long-term suggest that $RSPO2^+$ mesenchymal cells differentiate to $SM22^+$ airway smooth muscle cells, and pseudotime analysis on scRNA-seq data propose this could also be true *in vivo*. Together, this suggests a potential model by which *RSPO2*$^+$ cells support distal (bud tip and alveolar) cell fates and differentiate to smooth muscle to promote airway cell fates. Nevertheless, our experiments do not exclude the possibility that $SM22^+$ mesenchymal cells captured in the $LIFR^{HI}$ FACS-isolated population due

to sorting error were able to expand over the culture period. We believe this is unlikely because RSPO2$^+$ mesenchymal cells are more proliferative and were not detected in the FACS-isolated LIFR$^-$ population.

Of particular interest for the current study, *RSPO2* mutations in humans are lethal at birth, causing nearly complete lung aplasia with lung development ceasing just after the primary lung buds emerge from the trachea (Szenker-Ravi *et al.*, 2018), supporting a critical role for RSPO2$^+$ mesenchymal cells during human lung development. Although non-functional *Rspo2* in mice is also lethal at birth, with mutant lungs exhibiting reduced branching, laryngeal-tracheal defects, and reduced Wnt signaling in the bud tips, mutant lungs undergo some branching and do not show nearly as severe lung aplasia seen in humans (Bell *et al.*, 2008). There are multiple possible explanations for why humans and mice have different severities of developmental defects caused by *RSPO2* mutations. First, it is possible that *Rspo1*, *Rspo3*, or *Rspo4* can compensate for the loss of *Rspo2* in mice, which may not exist in humans. Our scRNA-seq and FISH data for *RSPO1*, *RSPO3*, and *RSPO4* in the human distal lung indicate that they are expressed in the same broad cell population as *RSPO2*, but at much lower levels. It would be valuable to determine the expression patterns of the other *Rspo* transcripts and proteins in the murine lung and determine if other RSPOs can replace the role of RSPO2 in the murine and human lungs. The advent of *in vitro* model systems of the developing human lung provides an excellent opportunity to explore these questions further (Conway *et al.*, 2020).

Another possibility for the mouse/human phenotype difference is if RSPO2 is necessary for initiating branching morphogenesis during the earliest stages of human, but not mouse, lung development. In mice, deletion of *Wnt2* and *Wnt2b* together inhibit lung progenitors from ever being specified, and deletion of *Wnt2* and *Wnt7b* together result in defective branching and non-

localized SOX9 expression (Goss *et al.*, 2009; Miller *et al.*, 2012). The possible necessity of RSPO2 to promote WNT signaling that may be necessary for maintaining lung progenitors, and preventing precocious airway differentiation, in the primary lung buds and/or to initiate branching morphogenesis in humans could explain why lungs in humans lacking functional *RSPO2* fail to develop past the primary lung bud stage. Through infection of human fetal lung explants with an LGR5 ECD adenovirus that reduces RSPO2-mediated WNT signaling, we show that premature differentiation of bud tip progenitors into airway lineages occurs. Although we show this phenomenon after branching morphogenesis has already begun, because of *RSPO2's* persistence throughout all the time points sequenced in this study, this could be the case beginning at the primary lung buds. Therefore, if premature differentiation of lung progenitors occurs at the primary lung bud stage, the lung could fail to develop further.

The findings in this study have also prompted additional questions. It is known that WNT signaling is an important regulator of human bud tip progenitor maintenance (Miller *et al.*, 2018); however, for the first time, we can appreciate the much larger signaling network involved in maintaining WNT signaling in the bud tip niche. We interrogated scRNA-seq data for many WNT signaling components and found an enrichment of many components throughout multiple cell types, suggesting a complex WNT signaling environment in the developing human lung. How other cell types and signaling pathways may be integrated to control bud tip progenitor behavior is a fascinating avenue of future exploration. For example, BMP4 is known to be critical for the bud tip niche and branching morphogenesis (Morrisey and Hogan, 2010; Hines and Sun, 2014; Miller *et al.*, 2018), but it was recently shown that RSPOs can also antagonize BMP signaling (Lee *et al.*, 2020). How BMP is mechanistically involved in bud tip progenitor behavior will be important to investigate. Additionally, although *RSPO2*$^{+}$ mesenchymal cells are

localized adjacent to bud tip progenitors, expression of *RSPO2* extends far beyond the cells that sit near bud tip progenitors. Combined with the unique expression pattern of *LGR4* throughout the mesenchyme and *LGR6* in airway smooth muscle cells, it would be interesting to understand the role that RSPO2^{+} cells have a role in regulating the behavior of other mesenchymal cells. The role of RSPOs may also extend beyond the distal lung and into the proximal lung.

2.4 Materials and Methods

2.4.1 Data and Code Availability

Sequencing data used in this study is deposited at EMBL-EBI ArrayExpress. Single-cell RNA sequencing of human fetal lung and human fetal lung explants: human fetal lung (ArrayExpress: E-MTAB-8221) (Miller *et al.*, 2020), human fetal lung explants (ArrayExpress: E-MTAB-10662) (this study). Code used to process data can be found at: https://github.com/jason-spence-lab/Hein_2021. CellProfiler pipelines used in this study will be made freely available to readers upon request. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

2.4.2 Experimental Models and Subject Details

Human Lung Tissue

Research involving human lung tissue (8.5 – 19 weeks post conception) was approved by the University of Michigan Institutional Review Board. All human lung tissue used in these experiments was normal, de-identified tissue obtained from the University of Washington Laboratory of Developmental Biology. Specimens from both male and female sexes were used. The tissue was shipped overnight in Belzer-UW Cold Storage Solution (Thermo Fisher, Cat#NC0952695) on ice, and all experiments were performed within 24 hours in Belzer solution.

Lung Explants: Culture Establishment

Three unique human tissue samples spanning 11 – 13 weeks post-conception were used. For each unique tissue, one to three explants were included for each type of analysis.

For air-liquid-interface culture, Nucleopore Track-Etched Membranes (13 mm, 8 µm pore, polycarbonate) (Sigma, Cat#WHA110414) placed in 24-well tissue culture plates (Thermo Fisher, Cat#12565163) were pre-coated with 20 µg/cm^2 Collagen Type I (Thermo Fisher, Cat#A10483-01) diluted in 0.01 N ice-cold acetic acid for 30 minutes on ice followed by 2 hours at 37ºC. The membranes were then washed with 1X PBS directly before use. To prepare the explants, the lung was placed in a petri dish in ice-cold 1X PBS and approximately 1 mm^2 pieces of tissue were cut from the most distal edge of the lung under a stereomicroscope using forceps and a scalpel. 500 µL culture media containing Advanced DMEM/F-12 (Thermo Fisher, Cat#12634010), 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122), 2 mM L-Glutamine (Thermo Fisher, Cat#25030081), 10 mM HEPES (Corning, Cat#25060CI), 1 bottle B-27 Supplement (Thermo Fisher, Cat#17504044), 1 bottle N-2 Supplement (Thermo Fisher, Cat#17502048), and 0.4 µM 1-Thioglycerol (Sigma, Cat#M1753) was added to each well in the plate underneath the membrane. One explant per membrane was placed directly on the center of the membrane. Media was changed every 2 days.

Lung Explants: Infection with Adenovirus

Immediately following placement of the explants on the Nucleopore Track-Etched Membranes, 10^{10} pfu of control or LGR5 ECD adenovirus previously described in Yan *et al.* (Yan *et al.*, 2017) was pipetted directly on top of each explant under a stereo microscope using a p10 pipette. A maximum of 2 µL adenovirus was added to each explant at a time to prevent the adenovirus from running off the explant. The explants were re-infected every 2 – 3 days.

Infection of the tissue was confirmed by immunofluorescence for FLAG and the murine IgG2a Fc fragment for the LGR5 ECD adenovirus and control adenovirus respectively.

Bud Tip Organoids: Culture Establishment

Human fetal lung bud tip organoids were derived as previously reported (Miller *et al.*, 2018). In short, the lung was placed in a petri dish in ice-cold 1X PBS, and approximately 1cm^2 pieces of tissue were cut from the most distal edge of the lung under a stereomicroscope using forceps and a scalpel. The tissue was enzymatically digested using dispase (Corning, Cat#354235) for 30 minutes, then was placed in 100% FBS (Sigma, Cat#12103C) for 15 minutes. In DMEM/F-12 (Corning, Cat#10-092-CV) supplemented with 10% FBS and 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122), the tissue was vigorously pipetted with a P1000 and subsequently a P200 to dissociate the epithelium from the mesenchyme, then was washed multiple times in the media described to obtain as pure a population of epithelial bud tips as possible. Bud tips were plated in ~20 µL 8 mg/mL Matrigel (Corning, Cat#354234) droplets in 24-well tissue culture plates (Thermo Fisher, Cat#12565163) and fed every 3 – 4 days in previously published bud tip media and passaged by shearing through a 27-gauge needle as previously described (Miller *et al.* 2018). All experiments involving bud tip cultures were derived from lungs 8 – 15 weeks post-conception for experiments involving culture with RSPO and WNT ligands or 16.5 – 18 weeks post-conception for co-cultures.

Bud Tip Organoids: Growth Factor Experiments

Bud tips were established for a minimum of 1 passage to ensure epithelial-only cultures before experiments began. Bud tip organoids were passaged 3 days prior to the start of the experiment. On the day of the experiment, media conditions were changed from the original, published bud tip media (Miller *et al.* 2018) consisting of DMEM/F-12 (Corning, Cat#10-092-

CV), 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122), 1 bottle B-27 supplement (Thermo Fisher, Cat#17504044), 1 bottle N-2 supplement (Thermo Fisher, Cat#17502048), 0.05% BSA (Sigma, Cat#A9647) and final concentrations of 50 µg/mL L-ascorbic acid (Sigma, Cat#A4544), 0.4 µM 1-Thioglycerol (Sigma, Cat#M1753), 50 nM all trans retinoic acid (Sigma, Cat#R2625), 10 ng/mL recombinant human FGF7 (R&D Systems, Cat#251-KG), and 3 µM CHIR99021 (APExBIO, Cat#A3011). Positive control bud tips were given the described media and negative control bud tips had CHIR99021 removed. Experimental conditions were given either 500 ng/mL recombinant human R-Spondin 2 Protein (R&D Systems, Cat#3266-RS) and/or 1X Afamin/WNT3a conditioned media (MBL International Corporation, Cat#J2-001) in place of CHIR99021. Cultures were fed again on day 3 and collected on day 4 for qRT-PCR analysis. Three unique biological specimens with 1 technical replicate per condition each were used for qRT-PCR analysis.

Bud Tip Organoids: LGR5 Knock-down

3 wells of 20 µL-size Matrigel drops of bud tip organoids (approximately 200,000 cells/well), 5 – 10 days post-harvest, were placed into a microcentrifuge tube and removed from Matrigel by pipetting with a P1000. Bud tip organoids were digested to small fragments by incubation with TrypLE (Thermo Fisher, Cat#12605010) at 37ºC for 7 – 10 minutes, with mechanical digestion with a P1000 pipette at the end. Trypsinization was quenched with DMEM/F-12 (Corning, Cat#10-092-CV). Cells were centrifuged at 300g for 5 minutes and resuspended in approximately 10^6 TU/mL lentivirus carrying shRNA pLKO.1 plasmid vectors expressing either a non-target scrambled sequence (Sigma, Cat#SHC016) or a sequence against human *LGR5* (Sigma, #11586) with 6 µM polybrene (Sigma, Cat#TR-1003-G) and 10 µM Y-27632 (APExBIO, Cat#A3008). Cells were placed at 37ºC for 6 hours with gentle agitation

every 30 – 60 minutes. Cells were then centrifuged at 300g for 5 minutes, resuspended in 60 µL Matrigel (Corning, Cat#354234), and plated in 20 µL droplets. Cells were grown in bud tip organoid maintenance media previously described (Miller *et al.* 2018) plus 10 µM Y-27632 for the first 24-hours. Cells were given 72 hours to reform organoids prior to antibiotic selection using 1 µM puromycin (Sigma, Cat#P9620) for 3 days. After 3 days, bud tip organoids were switched to media containing 0.25 µM puromycin for the remainder of the experiment. Bud tip organoids were passaged as described above 2 – 4 days post-antibiotic selection. 3 – 4 days after passaging, experimental media conditions were added. Conditions are as described above (positive control: 3 µM CHIR, negative control: no CHIR, RSPO2 only: 500 ng/mL recombinant human RSPO2 in place of CHIR). Cultures were fed again on day 3 and collected on day 4 for qRT-PCR analysis. Two unique biological specimens with 1 – 3 technical replicates per condition each were used for qRT-PCR analysis.

FACS of Mesenchymal Cells

The lung (10 – 11.5 weeks post-conception) was placed in a petri dish and approximately 1 gram was cut from the most distal edge using forceps and a scalpel. The tissue was minced as much as possible using dissecting scissors, then was placed into a 15 mL conical tube containing 9 mL 0.1% (w/v) filter-sterilized Collagenase Type II (Thermo Fisher, Cat#17101015) in 1X PBS and 1 mL filter-sterilized 2.5 units/mL dispase (Thermo Fisher, Cat#17105041) in 1X PBS. The tube was placed at 37ºC for 60 minutes with mechanical dissociation using a serological pipette every 10 minutes. After 30 minutes, 75 µL DNase I was added to the tube. 5 mL isolation media containing 78% RPMI 1640 (Thermo Fisher, Cat#11875093), 20% FBS (Sigma, Cat#12103C), and 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122) was added. Cells were passed through 100 µm and 70 µm filters, pre-coated with isolation media, and

centrifuged at 400g for 5 minutes at 4ºC. 1 – 2 mL Red Blood Cell Lysis Buffer (Sigma, Cat#11814389001) and 0.5 – 1 mL FACS buffer (2% BSA, 10 μM Y-27632 (APExBIO, Cat#A3008), 100 U/mL penicillin-streptomycin) was added to the tube, and the tube was rocked for 15 minutes at 4ºC. The cells were centrifuged at 500g for 5 minutes at 4ºC, washed twice in 2mL FACS buffer, re-suspended in FACS buffer, and counted. 10^6 cells were placed into FACS tubes (Corning, Cat#352063) for all control tubes (no antibody, DAPI only, isotype controls, individual antibodies/fluorophores) and 8×10^6 cells were placed into a FACS tube for cell sorting. Primary antibodies were added at room temperature (30 minutes for LIFR and corresponding isotype, 10 minutes for CD324 and CD31 and corresponding isotypes) (see Table 2-1 for antibody dilutions). 3 mL FACS buffer was added to each tube, then tubes were centrifuged at 300g for 5 minutes at 4ºC. Cells were washed twice with 3 mL FACS buffer, centrifuging at 300g for 5 minutes at 4ºC between washes. Cells were resuspended in FACS buffer and 0.2 μg/mL DAPI was added to appropriate tubes. FACS was performed using a Sony MA900 cell sorter and accompanying software. $LIFR^{HI}/CD324^{-}/CD31^{-}$ cells and $LIFR^{-}/CD324^{-}/CD31^{-}$cells were collected in 1mL isolation media. $LIFR^{HI}$ cells were gated highest 30% of LIFR expression.

Bud Tip Organoid and Mesenchyme Co-cultures

Established bud tip organoids (at least 1 passage post-derivation) were placed into a microcentrifuge tube and removed from Matrigel by pipetting with a P1000. Bud tips that had not been passaged within 10 days were also passed 1x through a 27-guage needle. The bud tips were then centrifuged for ~10 seconds in a microcentrifuge and the media and Matrigel was removed under a stereomicroscope. Freshly FACS-isolated mesenchymal cells were immediately counted using a hemocytometer and enough cells were pelleted to reach approximately 150,000

mesenchymal cells per well. On ice, atrigel (Corning, Cat#354234) was added to the tubes containing bud tip organoids. For co-cultures, the bud tips in Matrigel were transferred to the tubes containing the mesenchymal cell pellets. The bud tip organoids and mesenchyme were thoroughly mixed in the Matrigel by pipetting and swirling a P200 with the tip cut off. ~20 µL droplets of Matrigel with bud tip organoids +/- mesenchyme were placed into the center of wells of a 24-well tissue culture plate (Thermo Fisher, Cat#12565163). The plate was inverted and placed in an incubator at 37ºC for 20 minutes. For co-culture and negative control wells, 0.5 mL media consisting of DMEM/F-12 (Corning, Cat#10-092-CV), 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122), 1 bottle B-27 supplement (Thermo Fisher, Cat#17504044), 1 bottle N-2 supplement (Thermo Fisher, Cat#17502048), 0.05% BSA (Sigma, Cat#A9647) and final concentrations of 50 µg/mL L-ascorbic acid (Sigma, Cat#A4544), 0.4 µM 1-Thioglycerol (Sigma, Cat#M1753), 50 nM all trans retinoic acid (Sigma, Cat#R2625), and 10 ng/mL recombinant human FGF7 (R&D Systems, Cat#251-KG) were added to each well. For positive control wells, 3 µM CHIR99021 (APExBIO, Cat#A3011) was added to the above media. The cultures were fed every 3 – 4 days and were cultured for a total of 10 – 11 days. Each biological replicate shown is from a unique bud tip line co-cultured with mesenchymal cells from a unique human tissue specimen, and all three experiments were performed with four technical replicates.

Mesenchyme 2D Cultures

After FACS of $LIFR^{HI}$ and $LIFR^{-}$ mesenchymal cells as described above, 100,000 cells/well (for 6-hour timepoint) or 10,000 cells/well (for 7-day timepoint) were added to 24-well tissue culture plates (Thermo Fisher, Cat#12565163) pre-coated with 5 µg/cm^2 human fibronectin (Corning, Cat#356008). Cells were cultured in serum-free media consisting of 50% IMDM (Thermo Fisher Cat#12440-053), 50% F12 (Thermo Fisher, Cat#11765-054), 1% lipid

mixture (Sigma, Cat#L0288), 1X insulin-tranfserrin-selenium-ethanolamine (Thermo Fisher, Cat#51500-056), and 3% BSA (Sigma, Cat#A9647). Culture media was changed every 3 days.

2.4.3 Method Details

scRNA-seq: Tissue Processing

All tubes and pipette tips were pre-washed with 1% BSA in 1X HBSS (all HBSS in protocol is with Mg^{2+} and Ca^{2+}) to prevent cell adhesion to the plastic. The tissue was placed in a petri dish in ice-cold 1X HBSS, and the tissue was minced under a stereomicroscope using scissors. For uncultured lung tissue, roughly 1 cm^2 of the most distal portion of the lung was isolated, and for lung explants, 4 explants were collected per condition. The whole explants were used for sequencing. The minced tissue was transferred to a 15 mL conical tube with the HBSS, centrifuged at 500g for 5 minutes at 10ºC, and the HBSS was removed. Mix 1 from the Neural Tissue Dissociation Kit (Miltenyi, Cat#130-092-628) was added to each tube, and the tube was placed at 37ºC for 15 minutes, then Mix 2 was added and the cells were incubated for 10 minutes at 37ºC. The cells were agitated by harshly pipetting with a P1000. The incubation/agitation step was repeated every 10 minutes until the cells looked to be a single-cell suspension, approximately 30 minutes. The cells were then filtered through a 70 µm filter, pre-coated with 1% BSA in 1X HBSS, into a 15 mL conical tube. The filter was rinsed 3x with 1 mL 1% BSA in 1X HBSS. The cells were centrifuged at 500g for 5 minutes at 10ºC, and the supernatant was removed. 1 mL Red Blood Cell Lysis Buffer (Sigma, Cat#11814389001) and 0.5 mL 1% BSA in 1X HBSS was added to the tube, and the tube was rocked for 15 minutes at 4ºC. The cells were centrifuged at 500g for 5 minutes at 10ºC, washed twice in 2 mL 1% BSA in 1X HBSS and centrifuged again. The cells were resuspended in 200 µL 1% BSA in 1X HBSS, counted using a hemocytometer, centrifuged at 500g for 5 minutes at 10ºC, and resuspended to reach a

concentration of 1,000 cells/μL. Approximately 100,000 cells were put on ice and single-cell libraries were immediately prepared on the 10x Chromium at the University of Michigan Sequencing Core with a target of 10,000 cells.

scRNA-seq: Analysis Overview

To visualize distinct cell populations within the single cell RNA sequencing dataset, we employed the general workflow outlined by the Scanpy Python package (Wolf, Angerer and Theis, 2018). This pipeline includes the following steps: filtering cells for quality control, log normalization of counts per cell, extraction of highly variable genes, regressing out specified variables, scaling, reducing dimensionality with principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) (McInnes, Healy and Melville, 2018), and clustering by the Louvain algorithm (Blondel *et al.*, 2008).

scRNA-seq: Sequencing Data and Processing FASTQ Reads into Gene Expression Matrices

All single-cell RNA sequencing was performed at the University of Michigan Advanced Genomics Core with an Illumina Novaseq 6000. The 10x Genomics Cell Ranger v3 pipeline was used to process raw Illumina base calls (BCLs) into gene expression matrices. BCL files were demultiplexed to trim adaptor sequences and unique molecular identifiers (UMIs) from reads. Each sample was then aligned to the human reference genome (hg19) to create a filtered feature bar code matrix that contains only the detectable genes for each sample.

scRNA-seq: Quality Control

To ensure quality of the data, all samples were filtered to remove cells expressing too few or too many genes (Figure 2-1, 2-2, 2-10 - <750, >3000; Figure 2-5, 2-6, 2-7, 2-8 - <750, >10000) with high UMI counts (Figure 2-1, 2-2, 2-10 - >15000; Figure 2-5, 2-6, 2-7, 2-8 -

>50000), or a fraction of mitochondrial genes greater than (Figure 2-1, 2-2, 2-10 - >0.05, Figure 2-5, 2-6, 2-7, 2-8 - >0.1)

scRNA-seq: Normalization and Scaling

Data matrix read counts per cell were log-normalized, and highly variable genes were extracted. Using Scanpy's simple linear regression functionality, the effects of total reads per cell and mitochondrial transcript fraction were removed. The output was then scaled by a z-transformation.

scRNA-seq: Variable Gene Selection

Highly variable genes were selected by splitting genes into 20 equal-width bins based on log normalized mean expression. Normalized variance-to-mean dispersion values were calculated for each bin. Genes with log normalized mean expression levels between 0.125 and 3 and normalized dispersion values above 0.5 were considered highly variable and extracted for downstream analysis.

scRNA-seq: Batch Correction

We have noticed batch effects when clustering data due to technical artifacts such as timing of data acquisition or differences in dissociation protocol. To mitigate these effects, we used the Python package BBKNN (batch balanced k nearest neighbors) (Polański *et al.*, 2019). BBKNN was selected over other batch correction algorithms due to its compatibility with Scanpy and optimal scaling with large datasets. This tool was used in place of Scanpy's nearest neighbor embedding functionality. BBKNN uses a modified procedure to the *k*-nearest neighbors' algorithm by first splitting the dataset into batches defined by technical artifacts. For each cell, the nearest neighbors are then computed independently per batch rather than finding

the nearest neighbors for each cell in the entire dataset. This helps to form connections between similar cells in different batches without altering the PCA space. After completion of batch correction, cell clustering should no longer be driven by technical artifacts.

scRNA-seq: Dimension Reduction and Clustering

Principal component analysis (PCA) was conducted on the filtered expression matrix followed. Using the top principal components (Figure 2-1, 2-5, 2-10 – 20; Figure 2-5, 2-7, 2-8 – 15; Figure 2-2 – 30), a neighborhood graph was calculated for the nearest neighbors (Figure 2-1, 2-5, 2-10 – 30; Figure 2-5, 2-7, 2-8 – 20; Figure 2-2 – 50). BBKNN was implemented when necessary and calculated using the top 50 principal components with 3 neighbors per batch. The UMAP algorithm was then applied for visualization on 2 dimensions. Using the Louvain algorithm, clusters were identified at set resolutions (Figure 2-1, 2-10 – 0.55; Figure 2-5, 2-6, 2-7, S-8 – 0.4, Figure 2-2 – 0.225). 43,079 cells were included in Figur2-2, 10,751 cells in Figure 2-6A – C, and 14,900 cells in Figure 2-6D – F.

scRNA-seq: Sub-clustering

After annotating clusters within the UMAP embedding, specific clusters of interest were identified for further sub-clustering and analysis. The corresponding cells were extracted from the original filtered but unnormalized data matrix to include 30,841 cells in Figure 2-1 and 2-10, 2,279 cells for Figure 2-5A – C, and 859 cells for Figure 2-5D – G, 2-7, and 2-8. The extracted cell matrix then underwent log-normalization, variable gene extraction, linear regression, z-transformation, and dimension reduction to obtain a 2-dimensional UMAP embedding for visualization.

scRNA-seq: Cluster Annotation

Using canonically expressed gene markers, each cluster's general cell identity was manually annotated. The list of genes can be found in Figure 2-2B.

scRNA-seq: Diffusion Pseudotime

Diffusion pseudotime analysis was used to predict temporal order of cells (Figure 2-10F) (Haghverdi *et al*., 2016). Using the Scanpy implementation, a diffusion map with 20 components was calculated using log-normalized data. Pseudotime analysis was then performed with cluster 0 set as the root cell type.

scRNA-seq: Cell Scoring

Cells were scored based on expression of a set of 22 – 50 marker genes per cell type. Gene lists were compiled based on the previously-published top 22 – 50 most differentially expressed genes from *in vivo* cells of the cell type of interest (Miller *et al.*, 2020). Gene lists are published in Hein *et al.* 2022 (Renee F.C. Hein *et al.*, 2022). After obtaining the log-normalized and scaled expression values for the data set, scores for each cell were calculated as the average z-score within each set of selected genes.

Tissue Processing for IF and FISH

Whole tissue and organoids (removed from Matrigel by gentle pipetting) were immediately fixed in 10% Neutral Buffered Formalin for 24 hours at room temperature on a rocker, washed 3x, for 15 minutes each, with UltraPure DNase/RNase-Free Distilled Water (Thermo Fisher, Cat#10977015), and dehydrated for 1 hour in each of the following alcohol series diluted in UltraPure DNase/RNase-Free Distilled Water: 25% MeOH, 50% MeOH, 75% MeOH, 100% MeOH, 100% EtOH, 70% EtOH. Tissue was processed into paraffin blocks in an

automated tissue processor with 1-hour solution changes in the following series: 70% EtOH, 80% EtOH, 95% EtOH x2, 100% EtOH x3, Xylene x3, Paraffin x3. Tissue was then embedded into paraffin wax blocks. For FISH, all equipment was sprayed with RNase AWAY (Thermo Fisher, Cat#700511) prior to sectioning. Paraffin blocks were sectioned into 4 µm-thick sections for FISH (no longer than one week prior to performing FISH) or 4 – 7 µm-thick sections for IF onto charged glass slides. Slides were baked for 1 hour in a 60ºC dry oven (within 24 hours of performing FISH). Slides were stored at room temperature in a slide box containing a silicone desiccator packet and with the seams sealed with parafilm wrap.

IF Protein Staining

Tissue slides were rehydrated in Histo-Clear II (National Diagnostics, Cat#HS-202) 2x for 5 minutes each, then put through the following solutions for 2x for 2 minutes each: 100% EtOH, 95% EtOH, 70% EtOH, 30% EtOH. Then, slides were put in double-distilled water (ddH20) 2x for 5 minutes each. Antigen retrieval was performed by steaming slides in 1X Sodium Citrate Buffer (100 mM trisodium citrate (Sigma, Cat#S1804), 0.5% Tween 20 (Thermo Fisher, Cat#BP337), pH 6.0) for 20 minutes and subsequently cooling and washing quickly (moving slides up and down 5x) 2x in ddH20 and 2x in 1X PBS. Slides were incubated in a humidified chamber at room temperature for 1 hour with blocking solution (5% normal donkey serum (Sigma, Cat#D9663) in PBS with 0.1% Tween 20). Slides were then incubated in primary antibodies diluted in blocking solution in a humidified chamber at 4ºC overnight. Slides were washed 3x in 1X PBS for 10 minutes each. Slides were incubated with secondary antibodies and DAPI (1 µg/mL) diluted in blocking solution and placed in a humidified chamber at room temperature for 1 hour, then were washed 3x in 1X PBS for 10 minutes each. Slides were mounted in ProLong Gold (Thermo Fisher, Cat#P369300) and imaged within 2 weeks. Stained

slides were stored in the dark a 4ºC. All primary antibody concentrations are listed in Table 2-1. Secondary antibodies were raised in donkey, purchased from Jackson Immuno, and were used at a dilution of 1:500.

Mesenchymal cells in 2D tissue culture plates were immediately fixed in 4% paraformaldehyde for 10 minutes at room temperature, then washed 3x with 1X PBS for 10 minutes prior to blocking, primary antibody incubations, and secondary antibody incubations as described above. 1X PBS was added to the plates for imaging, and the plates were sealed with paraffin.

FISH

The FISH protocol was performed according to the manufacturer's instructions (ACD Bio; RNAscope multiplex fluorescent manual protocol) with a 6-minute protease treatment and 15-minute antigen retrieval in a steamer. For IF protein co-stains, the last step of the FISH protocol (DAPI) was skipped. Instead, the slides were washed 1x in PBS followed by the IF protocol above, beginning with the blocking step.

Quantification of IF and FISH Images

All FISH images for quantification were taken at 40x magnification. For IF, 20x or 40x magnification was used. Nuclear stains and punctate FISH stains were analyzed using unbiased automated signal detection and quantification using CellProfiler (Lamprecht, Sabatini and Carpenter, 2007; Jones *et al.*, 2008; Erben *et al.*, 2018). Punctate per image, number of cells per image, punctate associated with specific nuclear stains, and numbers of specific positive nuclear stains were quantified using CellProfiler. Statistical analysis was performed using ordinary one-way ANOVA or unpaired Welch's one-tailed t-test using the GraphPad Prism software.

RNA Extraction and qRT-PCR

Three biological replicates as well three technical replicates from the same biological specimen were included in each analysis. mRNA was isolated using the MagMAX-96 Total RNA Isolation Kit (Thermo Fisher, Cat#AM1830) or the PicoPure RNA Isolation Kit (Thermo Fisher, Cat#KIT0204) for FACS-sorted cells and *LGR5* knock-down bud tip organoids, and RNA quality and yield was measured on a Nanodrop 2000 spectrophometer just prior to cDNA synthesis. cDNA synthesis was performed using 100 ng RNA from each sample and using the SuperScript VILO cDNA Kit (Thermo Fisher, Cat#11754250). qRT-PCR was performed on a Step One Plus Real-Time PCR System (Thermo Fisher, Cat#43765592R) using QuantiTect SYBR Green PCR Kit (Qiagen, Cat# 204145). Primer sequences can be found in Table 2-1. Expression of genes in the measurement of arbitrary units was calculated relative to GAPDH or ACTB (for bud tip organoid experiments involving exogenous RSPO2) using the following equation: $2^{(Housekeeping_{Ct} - Gene_{Ct})} \times 10{,}000$

2.5 Acknowledgments

2.5.1 Financial Support

This project has been made possible in part by grant number CZF2019-002440 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation, and in part by the NIH-NHLBI (R01HL119215) funding to J.R.S. R.F.C.H. was supported by a NIH Tissue Engineering and Regenerative Medicine Training Grant (NIH-NIDCR T32DE007057) and by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHLBI F31HL152531). A.J.M. was supported by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHLBI F31HL142197). E.M.H. was supported by a

Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHBLI F31HL146162). T.F. was supported by a NIH Tissue Engineering and Regenerative Medicine Training Grant (NIH-NIDCR T32DE007057). A.S.C. was supported by the T32 Michigan Medical Scientist Training Program (5T32GM007863-40). I.G. and the University of Washington Laboratory of Developmental Biology was supported by NIH award number 5R24HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

2.5.2 Other Acknowledgements

We would like to thank Judy Opp and the University of Michigan Advanced Genomics core for the operation of the 10X Chromium single cell capture platform, the University of Michigan Microscopy core for providing access to confocal microscopes, the Flow Cytometry core for providing access to flow cytometers, and the Vector core for providing lentivirus production and adenovirus purification and expansion services. We would also like to thank the University of Washington Laboratory of Developmental Biology. Lastly, we would like to thank Michael Dame and the Translational Tissue Modeling Laboratory (TTML) for providing helpful suggestions regarding the use of WNT3a Afamin conditioned media as well as Lindy K. Brastrom for her careful technical editing of the manuscript.

2.6 Chapter 2 Tables

Table 2-1: Antibody dilutions and primer sequences.

Antibody	**Dilution**
TAGLN (SM22) (Abcam)	1:500
SOX9 (R&D Systems) (used in combination with FISH)	1:500
SOX9 (Millipore) (used for IF only)	1:500
SOX9 (Santa Cruz) (used for IF only)	1:150
ECAD (BD Biosciences)	1:500
SOX2 (R&D Systems) (used for IF)	1:500
SOX2 (Seven Hills Bioreagents) (used in combination with FISH)	1:1000
TP63 (R&D Systems)	1:500
CHGA (Abcam)	1:300
RAGE (Abcam)	1:500
ABCA3 (Seven Hills Bioreagents)	1:500
FOXJ1 (Seven Hills Bioreagents)	1:250
FLAG (Sigma)	1:250
KI67 (Thermo Fisher)	1:250
CCAS3 (Cell Signaling)	1:500
Pro-SP-C (Seven Hills Bioreagents)	1:500

SP-B (Seven Hills Bioreagents)	1:500
SCGB3A2 (Abcam)	1:500
HTII-280 (Terrace Biotech)	1:100
Alpha smooth muscle actin (Sigma)	1:500, 1 hour
MUC5AC (Abcam)	1:500
LIFR alpha PE (R&D Systems)	10µL per 10^6 cells
IgG1 PE isotype control (R&D Systems)	10µL per 10^6 cells
CD326 (EPCAM) FITC (Miltenyi)	1:50 per 10^6 cells
REA control IgG1 FITC (Miltenyi)	1:50 per 10^6 cells
CD31 APC (Miltenyi)	1:50 per 10^6 cells
REA control IgG1 APC (Miltenyi)	0.6µL in 100µL for 10^6 cells
CD31 647 (BD Biosciences)	1:200 per 10^6 cells
IgG2a FC control APC (Thermo Fisher)	1:500 per 10^6 cells
Primer	**Sequence**
SOX9	F: GTACCCGCACTTGCACAAC R: GTGGTCCTTCTTGTGCTGC
ETV5	F: GAAGAGGTTCGCAGGGATAAG R: TAAGGCCCAATCTACAGGTTTAC
LGR5 (Figure S6)	F: CCTCAGCGTCTTCACCTCCT R: TTTCTTTCCCAGGGAGTGGAT
LGR5 (Figure 5)	F: GTTTCCCGCAAGACGTAACT R: CAGCGTCTTCACCTCCTACC

NPC2	F: CGGTTCTGTGGATGGAGTTAT R: GGTGACATTGACGCTGTAAGA
TESC	F: GCCTGGCTGATGAGATCAAT R: GAGTCGTACATGTGGAACAGAA
SCGB3A2	F: GGGGCTAAGGAAGTGTGTAAATG R: CACCAAGTGTGATAGCGCCTC
TP63	F: CCACAGTACACGAACCTGGG R: CCGTTCTGAATCTGCTGGTCC
MUC5B	F: GTACCTGGACTGCAGCAACA R: CTTGTAGGTGGCCTCGTTGT
FOXJ1	F: CAACTTCTGCTACTTCCGCC R: CGAGGCACTTTGATGAAGC
SFTPC	F: AGCAAAGAGGTCCTGATGGA R: CGATAAGAAGGCGTTTCAGG
ABCA3	F: TGCAGCGCCTACTTGAACTT R: CTGAGCACAGCCATCGTCT
RAGE	F: TCCGGGCTGTGATGTTTTGA R: CTCACCCACCCTGTCCAAAT
RSPO2	F: TGAATGCCAGATGGTTTTGC R: ATCTGCCGTGTTCTGGTTTC
FGFR4	F: GAACCGCATTGGAGGCATT R: TTCTCTACCAGGCAGGTGTATGTG
SM22	F: AACAGCCTGTACCCTGATGG R: CGGTAGTGCCCATCATTCTT
FOXF1	F: AGTCCCCAATGCAAAGACAC R: TCAGCAGAATTCCTGTGTGG
EPCAM	F: CTGCCAAATGTTTGGTGATG R: CTTCTGACCCCAGCAGTGTT
CDH5	F: CTTCACCCAGACCAAGTACACA R: AATGGTGAAAGCGTCCTGGT
AXIN2	F: AGTGTGAGGTCCACGGAAAC R: CTGGTGCAAAGACATAGCCA
KI67	F: TGGGTCTGTTATTGATGAGCC R: TGACTTCCTTCCATTCTGAAGAC
RNF43	F: ATACCAGCCAGTGGGTTCAG R: TTAGTGCTCAGGGCTCACCT

2.7 Chapter 2 Figures

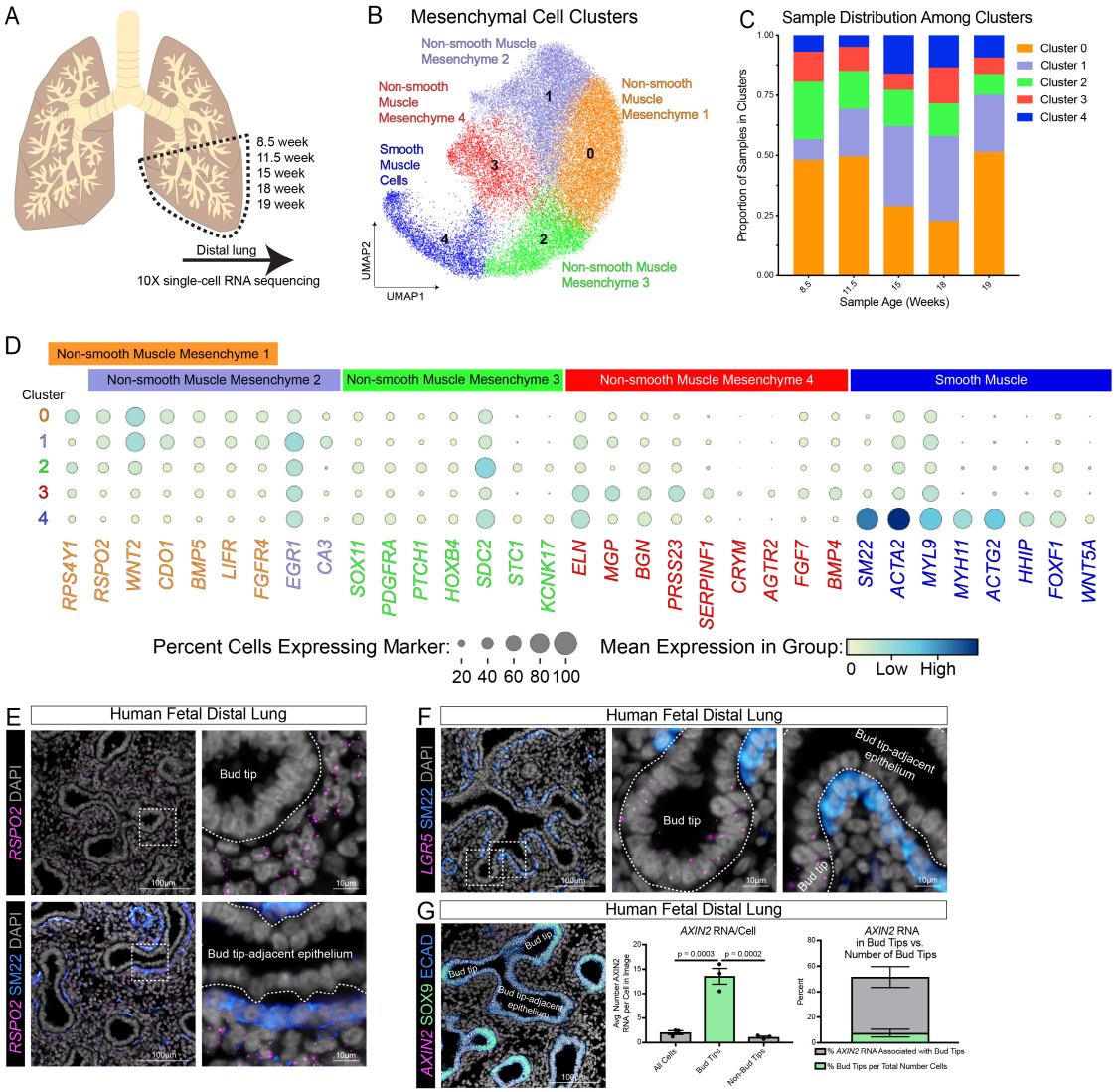


Figure 2-1: Identification of Bud Tip-associated Mesenchymal Populations and the Close Association of *RSPO2*$^{+}$ Mesenchymal Cells with Bud Tip Progenitors. (A) Schematic showing sample ages (days post-conception) and the general location of the distal lung where samples were taken from for single-cell RNA sequencing. **(B)** Cluster plot of non-vascular smooth muscle mesenchymal cells (see Figure 2-2A for all cells) generated from single-cell RNA sequencing data. Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cell-type identity of the cells composing each cluster. Cell-type labels for each cluster are based on expression of canonical smooth muscle or novel cell-type markers displayed in the dot plot in Figure 2-1D. **(C)** Stacked bar graph displaying the normalized proportion of cells from each sample in each cluster of the cluster plot in Figure 2-1B. **(D)** Dot plot of genes enriched in each cluster of the UMAP plot in Figure 2-1B. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized and z-transformed expression level of the gene. Clusters are colored corresponding to the cluster plot in Figure 2-1B. **(E)** Fluorescence *in situ* hybridization (FISH) of *RSPO2* and co-immunofluorescence (IF) for SM22 on 12-week human fetal distal lung tissue sections. Boxes on the left mark the inset location shown on the right. Dotted lines outline epithelium. **(F)** FISH of *LGR5* and co-IF for SM22 on 17.5-week human fetal distal lung tissue sections. Boxes on the left mark the inset locations shown on the right. Dotted lines outline epithelium. **(G)** The leftmost image shows FISH for *AXIN2* and co-IF for bud tip marker SOX9 on 13.5-week human fetal distal lung tissue sections. DAPI is shown in gray. The middle graph shows that bud tips have an average of 13.5 *AXIN2* RNA molecules per cell compared to an average of 1.1 *AXIN2* RNA molecules in non-bud tip cells ($p = 0.0002$, ordinary one-way ANOVA followed by Turkey's multiple comparison test) and an average of 2.0 *AXIN2* molecules in all cells ($p = 0.0003$, ordinary one-way ANOVA followed

by Turkey's multiple comparison test) in a single-plane image of a 4μm tissue section. The right-most graph shows that bud tips make up approximately 7.6% of cells in the distal lung but approximately 43.9% of *AXIN2* RNA molecules are associated with bud tip progenitor cells. Error bars show standard error of the mean. Each data point represents a separate image field from the same tissue specimen.

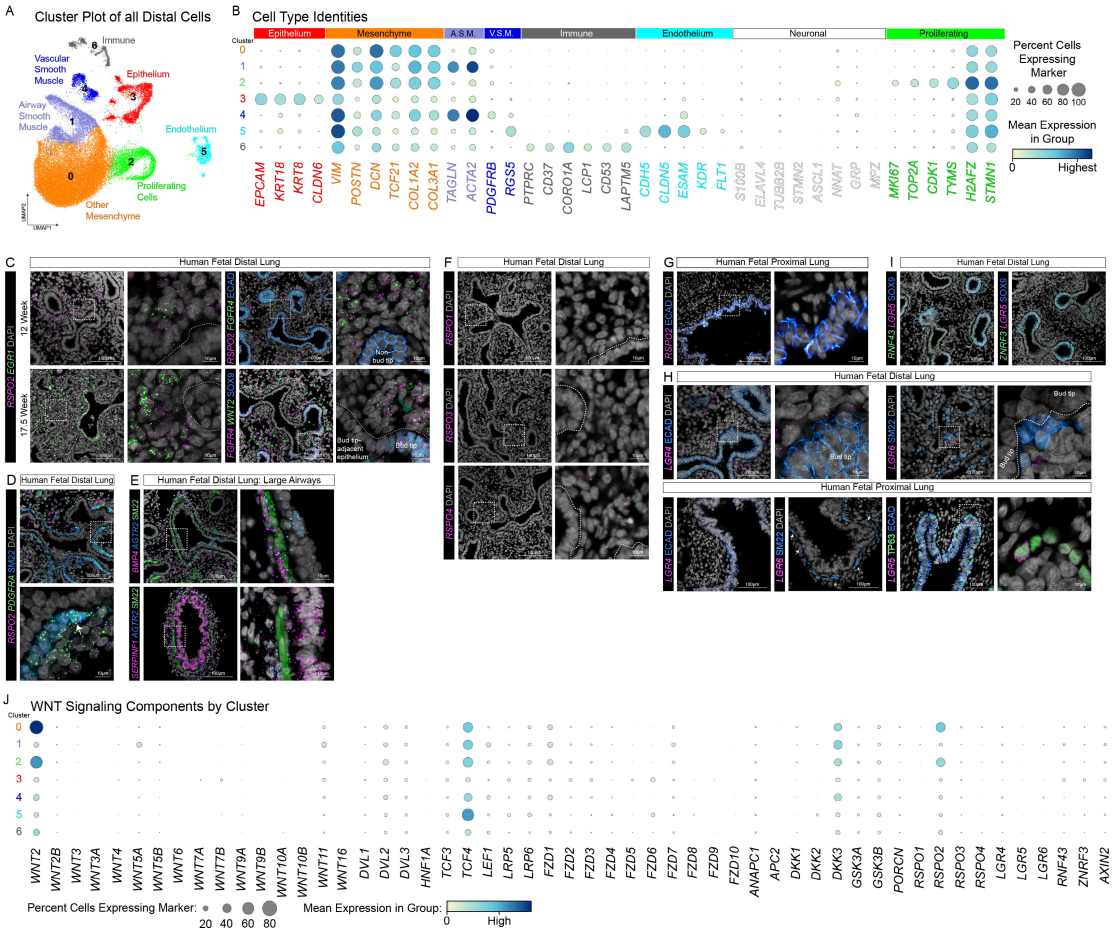


Figure 2-2: Identification of Major Cell Classes in the Distal Lung, Characterization of the Distal Lung Mesenchymal Populations, and *RSPO* and *LGR* Expression patterns, Related to Figure 2-1. (A) Cluster plot of cells from 8.5 – 19-week post-conception human distal lung specimens using single-cell RNA sequencing data. Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cell-type identity of the cells composing each cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the dot plot in Figure 2-2B. **(B)** Dot plot of genes enriched in each cluster shown in Figure 2-2A. The dot size represents the percentage of cells expressing the gene in the corresponding cluster and the color indicates log-normalized and z-transformed expression level of the gene. Clusters are colored corresponding to the cluster plot in Figure 2-2A. Each cluster is associated with a major cell class: epithelium, mesenchyme (including airway smooth muscle [A.S.M.] and vascular smooth muscle [V.S.M] cells), endothelium, immune, and proliferating cells. **(C)** The two leftmost panels show multiplexed fluorescence *in situ* hybridization (FISH) of *RSPO2* and *EGR1* on 12- and 17.5-week human fetal distal lung tissue sections. DAPI is shown in gray. Dotted lines outline epithelium. The two rightmost panels show multiplexed FISH of *FGFR4* and *WNT2* combined with immunofluorescence (IF) for the pan-epithelial marker ECAD or bud tip marker SOX9 on 12- to 13-week human fetal distal lung tissue sections. Boxes on the left mark the inset location shown on the right. **(D)** Multiplexed FISH of *PDGFRα* and *RSPO2* and co-IF for SM22 on 12-week human fetal distal lung tissue sections. Arrow points to the region of bud tip-adjacent smooth muscle where *PDGFRα* appears enriched. **(E)** Multiplexed FISH of *AGTR2* and *BMP4* (top) or *AGTR2* and *SERPINF1* (bottom) and co-IF for SM22 on 11-week human fetal distal lung tissue sections. DAPI is shown in gray. Scale bars represent 100µm or 10µm for insets. Enriched co-expression of *AGTR2* with *BMP4* or *SERPINF1* is evident in mesenchymal cells adjacent to SM22^{+} smooth muscle cells surrounding large airways in the distal lung. Box on the top mark the inset location shown on the right. **(F)** FISH of *RSPO1*, *RSPO3*, and *RSPO4* on 15-week human fetal distal lung tissue sections. Dotted lines outline epithelium. Boxes on the left mark the inset location shown on the right. **(G)** FISH of *RSPO2* and co-IF for the pan-epithelial marker ECAD on 17.5-week human distal proximal lung tissue sections. Box on the left mark the inset location shown on the right. **(H)** FISH of *LGR5* in the human fetal proximal lung and *LGR4* and *LGR6* in the distal and proximal lung and co-IF for the pan-epithelial marker ECAD, smooth muscle marker SM22, and/or the basal cell marker TP63 on 12-17-week human fetal distal lung tissue sections. DAPI is

shown in gray. Box on the left mark the inset location shown on the right. Dotted lines outline epithelium. Arrows point to *LGR6* expression in $SM22^{+}$ cells. **(I)** Multiplexed FISH of *LGR5* with *RNF43* or *ZNRF3* and co-IF for the bud tip marker SOX9 on 11-week human fetal distal lung tissue sections respectively. DAPI is shown in gray. **(J)** Dot plot of WNT signaling-related genes in each cluster shown in Figure 2-1A. The dot size represents the percentage of cells expressing the gene in the corresponding cluster and the color indicates log-normalized and z-transformed expression level of the gene.

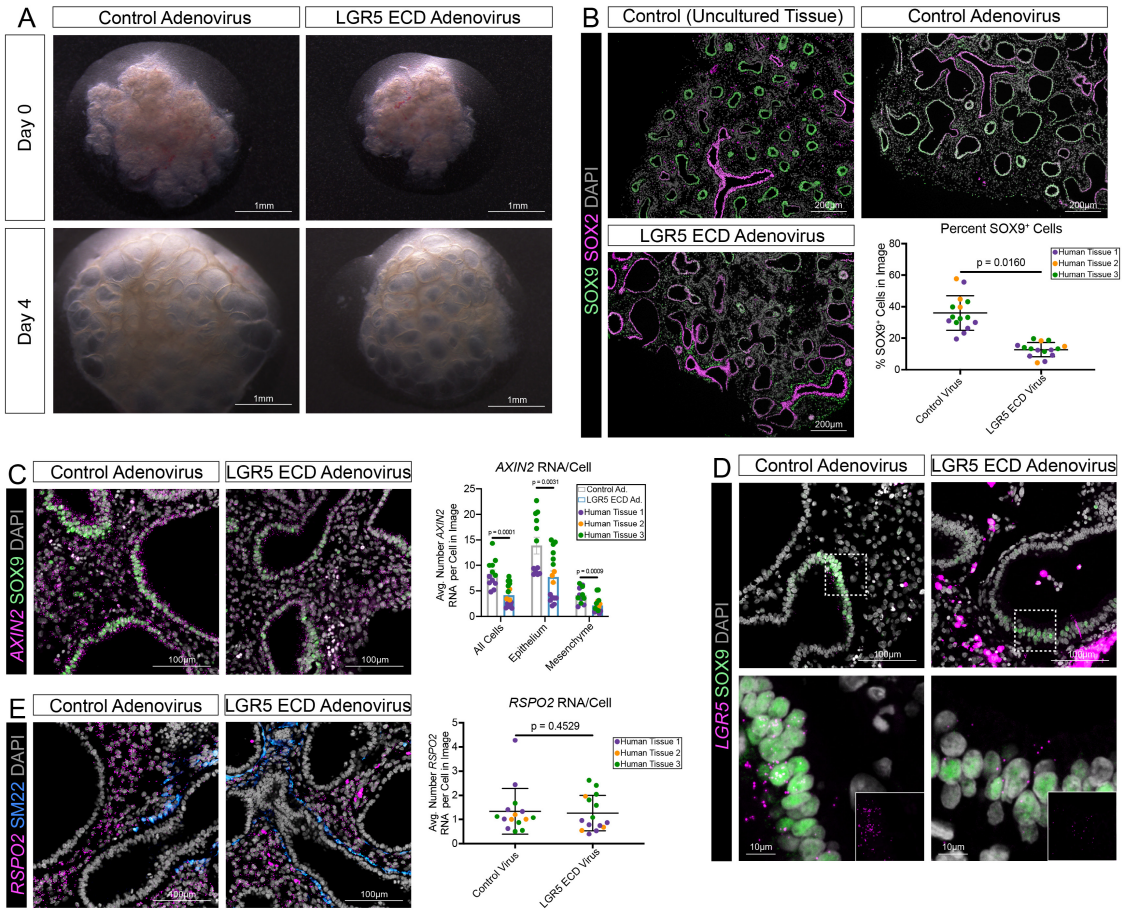


Figure 2-3: Inhibition of RSPO2-mediated WNT Signaling in Lung Explants Disrupts Proximal-distal Patterning. (A) Inverted microscope images of 11.5-week human fetal lung explants at the start of air-liquid interface culture (day 0) and after 4 days of culture (day 4) with 2 infections of either a control or LGR5 ECD adenovirus (ad). **(B)** Immunofluorescence (IF) for bud tip/distal marker SOX9 and airway/proximal epithelial marker SOX2 on sections from 12-week uncultured human fetal lung tissue, control ad-infected explants, and LGR5 ECD ad-infected explants. Quantification of the SOX9 stain is shown in the bottom right. At the end of the 4-day culture period, $SOX9^+$ cells in the control ad-infected explants comprised approximately 36.0% of cells while they only comprised 12.8% in LGR5 ECD ad-infected explants (unpaired Welch's one-tailed t-test). This quantification was performed in three unique biological specimens with one to three technical replicates and a minimum of three image fields for each specimen. **(C)** Fluorescence *in situ* hybridization (FISH) and quantification of *AXIN2* and co-IF for bud tip/distal marker SOX9 on control ad-infected explants and LGR5 ECD ad-infected explants. The average number of *AXIN2* molecules per cell in each image was calculated for epithelial and mesenchymal compartments separately by co-immunofluorescence for the nuclear lung epithelial marker TTF1 and counting *AXIN2* molecules per $TTF1^+$ cells (epithelium) and $TTF1^-$ cells (mesenchyme). This quantification was performed in three unique biological specimens for LGR5 ECD ad-infected explants and two unique biological specimens for control ad-infected explants with one to three technical replicates and a minimum of three image fields for each specimen. Statistical tests were carried out using unpaired Welch's one-tailed t-tests. **(D)** FISH of *LGR5* and co-IF for bud tip/distal marker SOX9 on sections from control ad-infected explants and LGR5 ECD ad-infected explants. Boxes on the top mark the inset location shown on the bottom. Insets on the bottom right on the bottom images show a single-channel image for *LGR5*. **(E)** FISH and quantification of *RSPO2* and co-IF for smooth muscle marker SM22 on sections from control ad-infected explants and LGR5 ECD ad-infected explants. The level of *RSPO2* expression in LGR5 ECD ad-infected explants is not significantly different than *RSPO2* expression in control ad-infected explants (unpaired Welch's one-tailed t-test). This quantification was performed in three unique biological specimens with one to three technical replicates and a minimum of three image fields for each specimen.

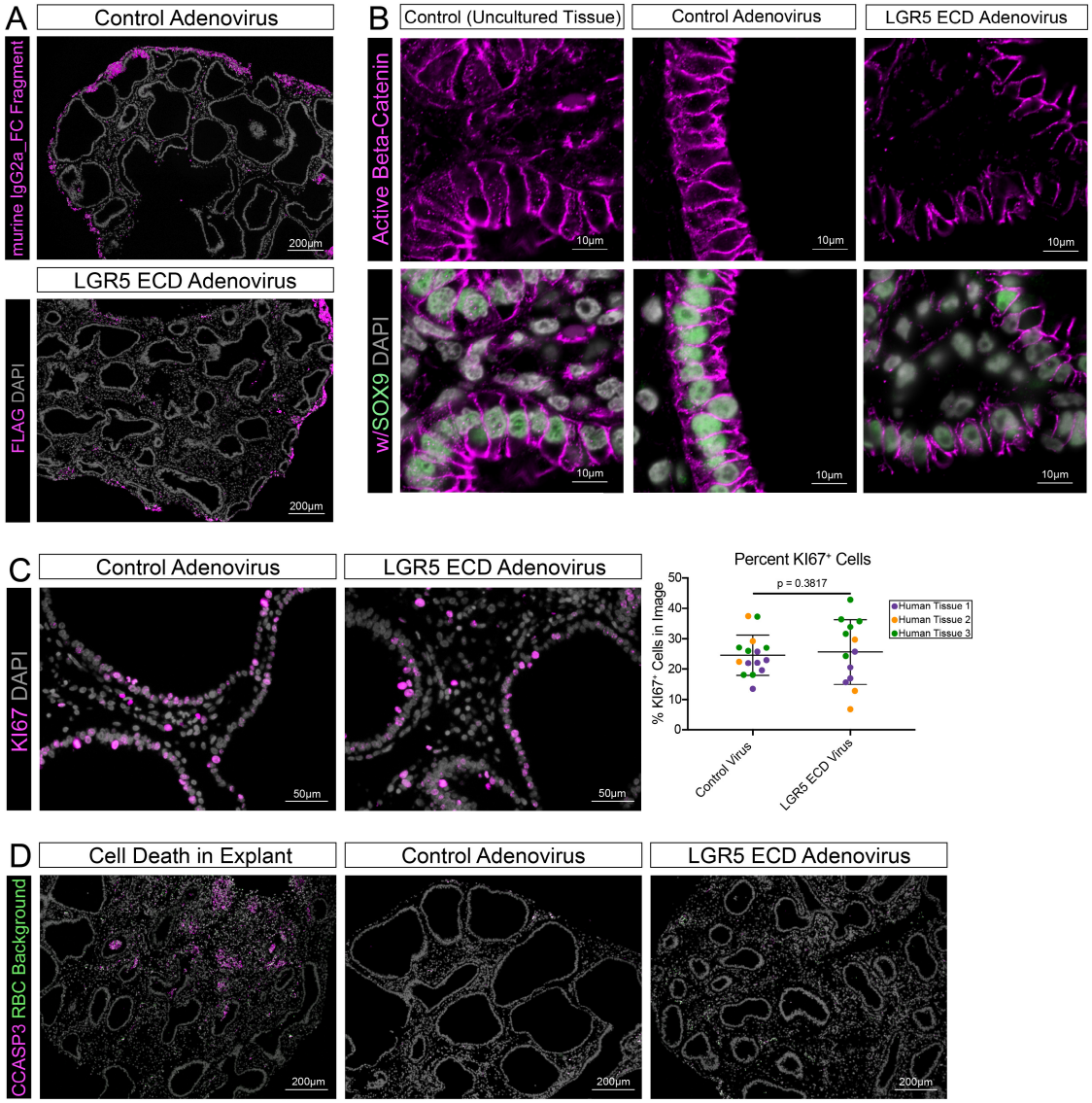


Figure 2-4: Confirmation of Adenovirus Infection and Proliferation & Apoptosis in Lung Explants, Related to Figure 2-3. **(A)** Immunofluorescence (IF) staining for murine IgG2a_FC on the control adenovirus (ad)-infected explants detects the presence of the control ad in explants and IF staining for FLAG on the LGR5 ECD ad-infected explants detects the presence of the LGR5 ECD ad in explants. DAPI is shown in gray. **(B)** IF for non-phosphorylated/active β-Catenin with or without distal/bud tip marker SOX9 on sections from 16-week uncultured human fetal lung tissue, control ad-infected explants, and LGR5 ECD ad-infected explants. **(C)** KI67 IF staining on and quantification of sections from control ad-infected explants and LGR5 ECD ad-infected explants. 24.6% of cells in control ad-infected explants are $KI67^+$ and 25.6% of cells in LGR5 ECD ad-infected explants are $KI67^+$ (Welch's one-tailed t-test). This quantification was performed in three unique biological specimens with one to three technical replicates and a minimum of three image fields for each specimen. **(D)** CCASP3 IF staining on sections from an explant that grew abnormally large as well as typical-sized control and LGR5 ECD ad-infected explants. DAPI is shown in gray. Background from red blood cells (RBC) is shown in green in order to identify true CCASP3 staining.

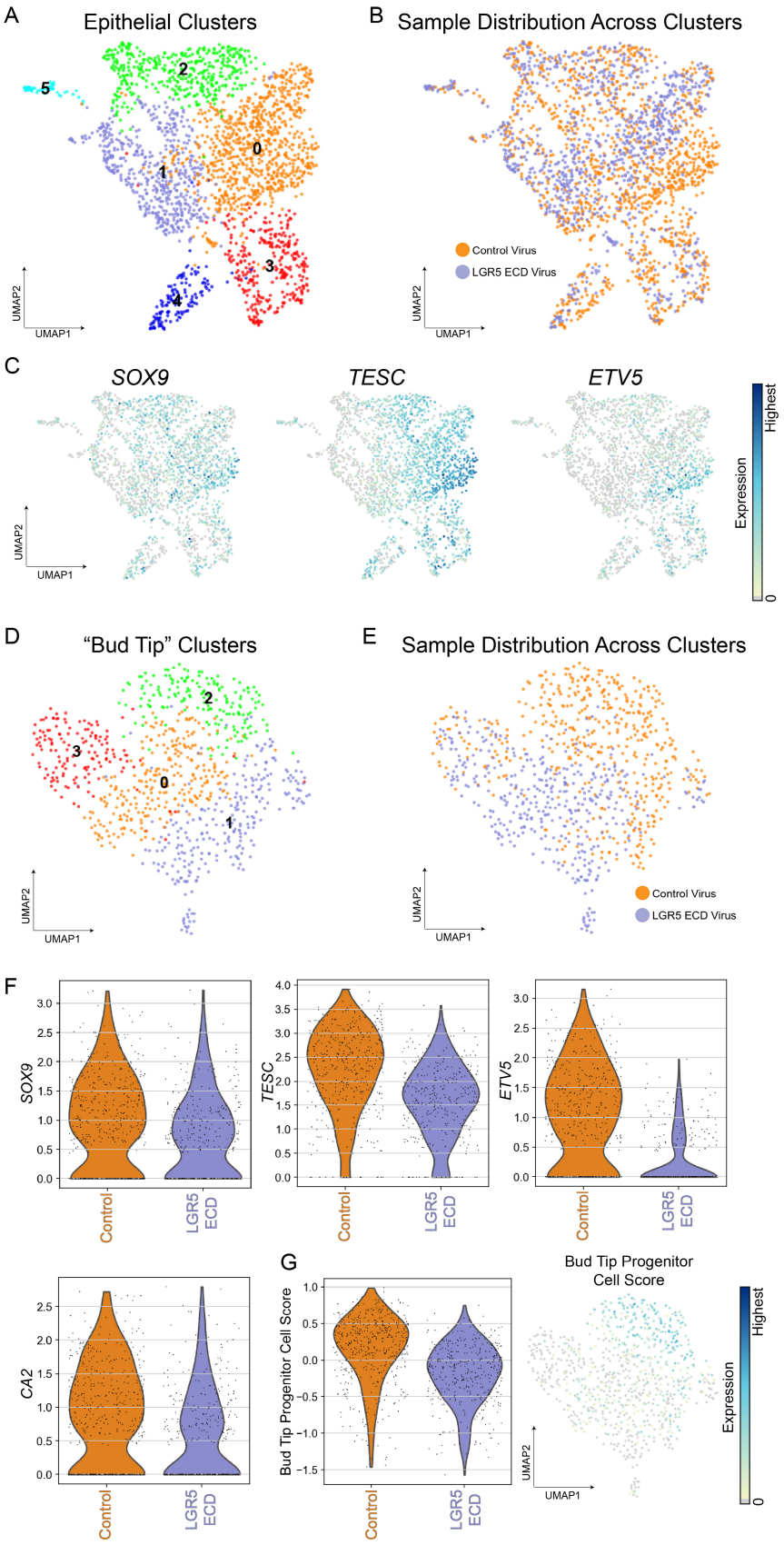

A
Epithelial Clusters
5
2
0
1
3
4
UMAP2
UMAP1
B
Sample Distribution Across Clusters
Control Virus
LGR5 ECD Virus
UMAP2
UMAP1
C
SOX9
TESC
ETV5
UMAP2
UMAP1
Expression
Highest
0
D
"Bud Tip" Clusters
2
3
0
1
UMAP2
UMAP1
E
Sample Distribution Across Clusters
Control Virus
LGR5 ECD Virus
UMAP2
UMAP1
F
SOX9
TESC
ETV5
Control
LGR5 ECD
CA2
G
Bud Tip Progenitor Cell Score
Bud Tip Progenitor Cell Score
UMAP2
UMAP1
Expression
Highest
0

Figure 2-5: The Bud Tip Progenitor Transcriptional Profile is Dependent on RSPO2-mediated WNT Signaling in Bud Tips. (A) Cluster plot of the epithelial cells (*EPCAM*$^{+}$/*KRT18*$^{+}$/*KRT8*$^{+}$/*CLDN6*$^{+}$) computationally extracted from LGR5 ECD adenovirus (ad)-infected explants and control ad-infected explants sequenced using single-cell RNA sequencing (scRNA-seq) (re-cluster of cluster 1 from Figure 2-6A). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. **(B)** UMAP plot corresponding to Figure 2-5A. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(C)** UMAP feature plots corresponding to the cluster plot in Figure 2-5A and displaying expression levels of the known bud tip progenitor markers *SOX9*, *TESC*, and *ETV5*. The color of each dot indicates log-normalized and z-transformed expression level of the given gene in the represented cell. **(D)** Cluster plot of bud tip-like cells (re-cluster of cluster 0 from Figure 2-5A). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. **(E)** UMAP plot corresponding to Figure 2-5D. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(F)** Violin plots corresponding to the cluster plot in Figure 2-5D and displaying expression known of bud tip progenitor markers *SOX9*, *TESC*, *ETV5*, and *CA2* in control ad- and LGR5 ECD ad-infected cells. **(G)** Violin plot and UMAP feature plot corresponding to the cluster plot in Figure 2-5D and displaying bud tip progenitor cell score, calculated as the average expression of the top 22 enriched genes in *in vivo* bud tip progenitor cells (see methods), for cells in LGR5 ECD ad-infected explants and control ad-infected explants. The color of each dot in the feature plot indicates log-normalized and z-transformed expression level of the set of bud tip genes in the represented cell.

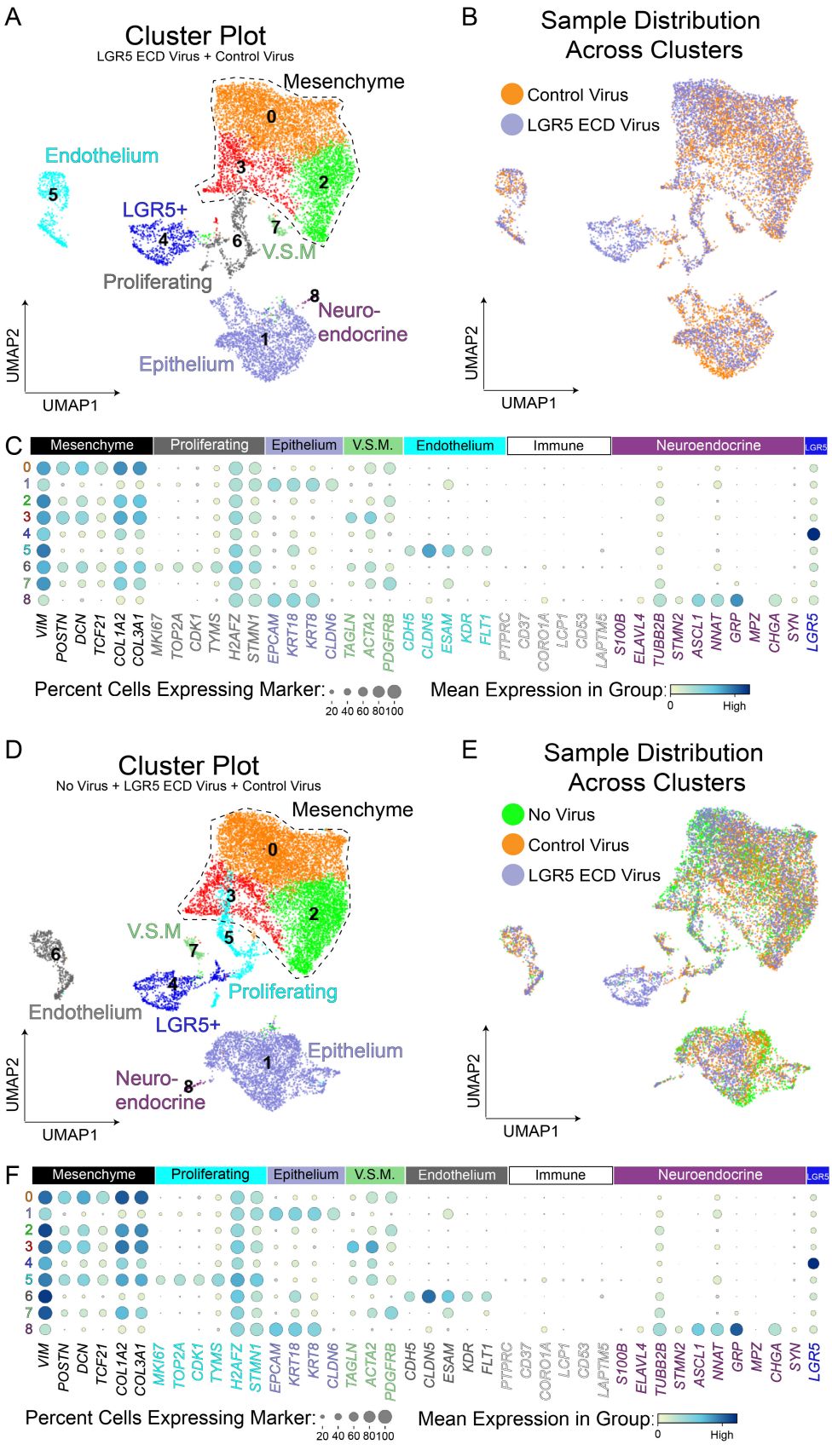

A
Cluster Plot
LGR5 ECD Virus + Control Virus
Mesenchyme
Endothelium
LGR5+
V.S.M
Proliferating
Neuro-
endocrine
Epithelium
UMAP2
UMAP1
B
Sample Distribution
Across Clusters
Control Virus
LGR5 ECD Virus
C
Mesenchyme
Proliferating
Epithelium
V.S.M.
Endothelium
Immune
Neuroendocrine
LGR5
VIM POSTN DCN TCF21 COL1A2 COL3A1 MKI67 TOP2A CDK1 TYMS H2AFZ STMN1 EPCAM KRT18 KRT8 CLDN6 TAGLN ACTA2 PDGFRB CDH5 CLDN5 ESAM KDR FLT1 PTPRC CD37 CORO1A LCP1 CD53 LAPTM5 S100B ELAVL4 TUBB2B STMN2 ASCL1 NNAT GRP MPZ CHGA SYN LGR5
Percent Cells Expressing Marker:
20 40 60 80100
Mean Expression in Group:
0
High
D
Cluster Plot
No Virus + LGR5 ECD Virus + Control Virus
Mesenchyme
V.S.M
Proliferating
Endothelium
LGR5+
Epithelium
Neuro-
endocrine
UMAP2
UMAP1
E
Sample Distribution
Across Clusters
No Virus
Control Virus
LGR5 ECD Virus
F
Mesenchyme
Proliferating
Epithelium
V.S.M.
Endothelium
Immune
Neuroendocrine
LGR5
VIM POSTN DCN TCF21 COL1A2 COL3A1 MKI67 TOP2A CDK1 TYMS H2AFZ STMN1 EPCAM KRT18 KRT8 CLDN6 TAGLN ACTA2 PDGFRB CDH5 CLDN5 ESAM KDR FLT1 PTPRC CD37 CORO1A LCP1 CD53 LAPTM5 S100B ELAVL4 TUBB2B STMN2 ASCL1 NNAT GRP MPZ CHGA SYN LGR5
Percent Cells Expressing Marker:
20 40 60 80100
Mean Expression in Group:
0
High

Figure 2-6: Identification of Major Cell Classes in Lung Explants, Related to Figure 2-5. (A) Cluster plot of cells from LGR5 ECD adenovirus (ad)-infected explants and control ad-infected explants sequenced using single-cell RNA sequencing (scRNA-seq). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cell-type identity of the cells composing each cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the dot plot in Figure 2-6C. Cells from the epithelium, mesenchyme, and endothelium were identified. Also identified was a cluster of neuro-endocrine cells (positive expression for epithelial markers, neuronal markers, and neuro-endocrine markers *CHGA* and *SYN*), a cluster of cycling cells, and a cluster of cells defined by high expression of *LGR5*. **(B)** Cluster plot corresponding to Figure 2-6A. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(C)** Dot plot of genes enriched in each cluster shown in Figure 2-6A. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized and z-transformed expression level of the gene. Clusters are colored corresponding to the cluster plot in Figure 2-6A. **(D)** Cluster plot of cells from non-viral infected explants, LGR5 ECD ad-infected explants, and control ad-infected explants sequenced using scRNA-seq. Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cell-type identity of the cells composing each cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the dot plot in Figure 2-6F. The same general cell type clusters found in the cluster plot in Figure 2-6A (not including non-infected cells) were found when non-infected cells were included in the clustering analysis. **(E)** UMAP plot corresponding to the cluster plot in Figure 2-6D. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(F)** Dot plot of genes enriched in each cluster shown in Figure 2-6D. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized and z-transformed expression level of the gene. Clusters are colored corresponding to the UMAP plot in Figure 2-6D.

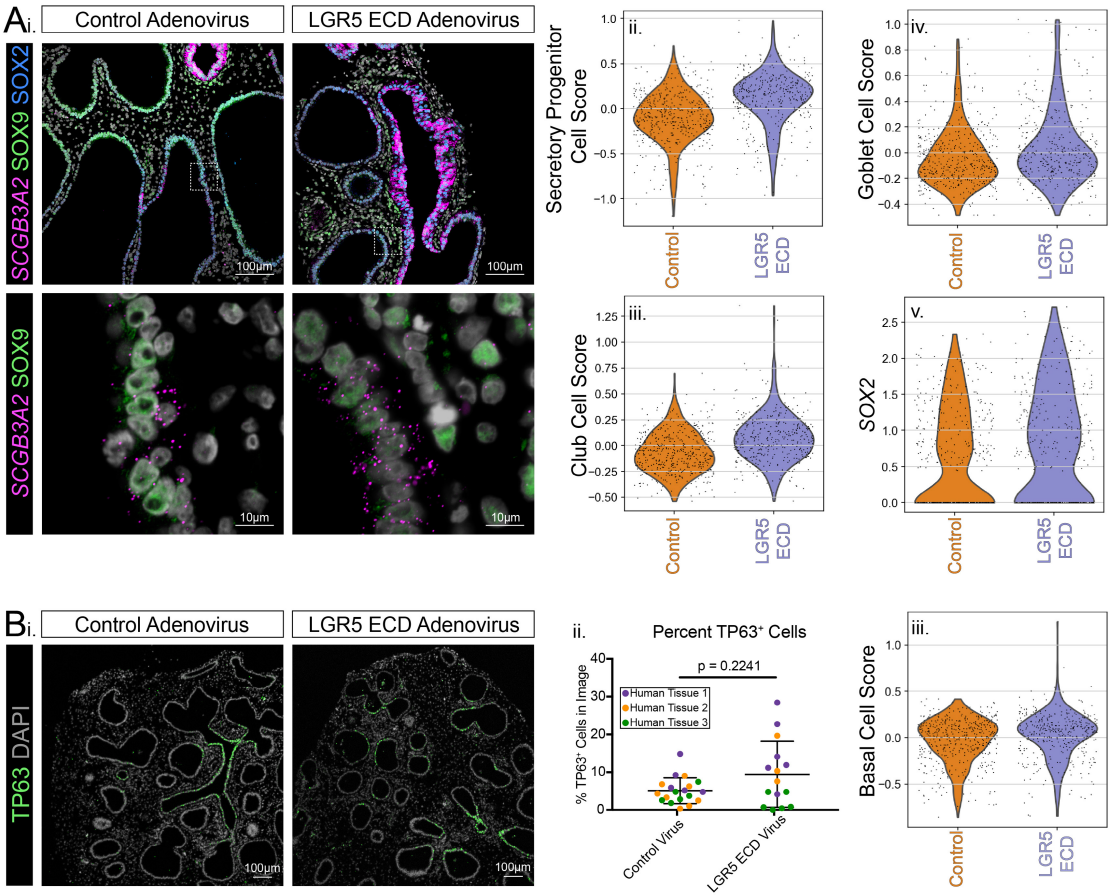


Figure 2-7: Inhibition of RSPO2-mediated WNT Signaling in Lung Explants Results in Bud Tip Differentiation into Airway Secretory and Basal Cell Types. (A) Expression of airway secretory cell type markers in LGR5 ECD adenovirus (ad)-infected explants and control ad-infected explants. **(i.)** Fluorescent *in situ* hybridization staining of the secretory cell marker *SCGB3A2* and co-immunofluorescence (IF) for the bud tip marker SOX9 and airway marker SOX2 on sections from control ad-infected explants and LGR5 ECD ad-infected explants. DAPI is shown in gray. Boxes on the top mark the inset location shown on the bottom. **(ii. - iv.)** Violin plots displaying the cell score for the listed airway secretory cell type, calculated from single-cell RNA sequencing (scRNA-seq) data as the average expression of the top 50 enriched genes in *in vivo* fetal secretory cell types (see methods), for cells in LGR5 ECD ad-infected explants and control ad-infected explants. **(v.)** Log-normalized and z-transformed expression level of the airway marker *SOX2* in scRNA-seq data from LGR5 ECD ad-infected explants and control ad-infected explants. **(B)** Expression of airway basal cell type markers in LGR5 ECD ad-infected explants and control ad-infected explants. **(i.)** IF staining for the airway progenitor/basal cell marker TP63 on sections from control ad-infected explants and LGR5 ECD ad-infected explants. **(ii.)** Quantification of basal cell marker TP63 on sections from control ad-infected explants and LGR5 ECD ad-infected explants. The number of $TP63^+$ cells in LGR5 ECD ad-infected explants is increased but not significantly different than the number of $TP63^+$ cells in control ad-infected explants (unpaired Welch's one-tailed t-test). This quantification was performed in three unique biological samples with one to three technical replicates and a minimum of three image fields for each sample. **(iii.)** Violin plot displaying the basal cell score for cells in LGR5 ECD and control ad-infected explants, calculated from scRNA-seq data as the average expression of the top 50 enriched genes in *in vivo* fetal basal cells (see methods), for cells in LGR5 ECD ad-infected explants and control ad-infected explants.

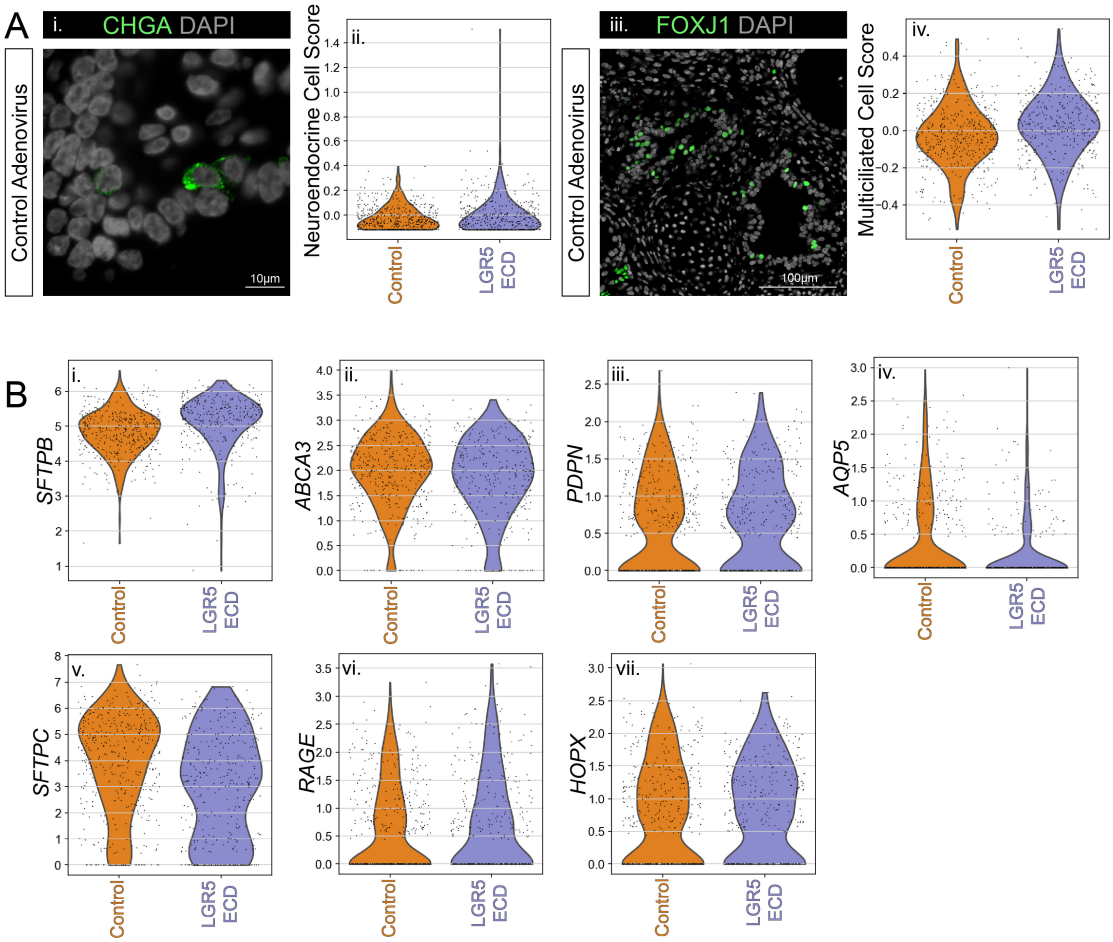


Figure 2-8: Inhibition of RSPO2-potentiated WNT Signaling in Lung Explants Results in Minimal Bud Tip Differentiation into Neuroendocrine and Multicilated Cell Types and Distal Cell Types, Related to Figure 2-7. (A) Expression of neuroendocrine and multiciliated cell type markers in LGR5 ECD ad-infected explants and control ad-infected explants. For violin plots, cell score for the listed airway cell type was calculated from single-cell RNA sequencing (scRNA-seq) data as the average expression of the top 50 enriched genes in *in vivo* cells of the listed cell type (see methods). **(i.)** Immunofluorescence (IF) staining for the neuroendocrine cell marker CHGA on a section from a control ad-infected explant. **(ii.)**. Neuorendocrine cell score for cells in LGR5 ECD ad-infected explants and control ad-infected explants. **(iii.)** IF staining for the multiciliated cell marker FOXJ1 on a section from a control ad-infected explant. **(iv.)**. Multiciliated cell score for cells in LGR5 ECD ad-infected explants and control ad-infected explants. **(B)** Violin plots displaying log-normalized and z-transformed expression level of a single distal/alveolar cell type genes in scRNA-seq data from LGR5 ECD ad-infected explants and control ad-infected explants.

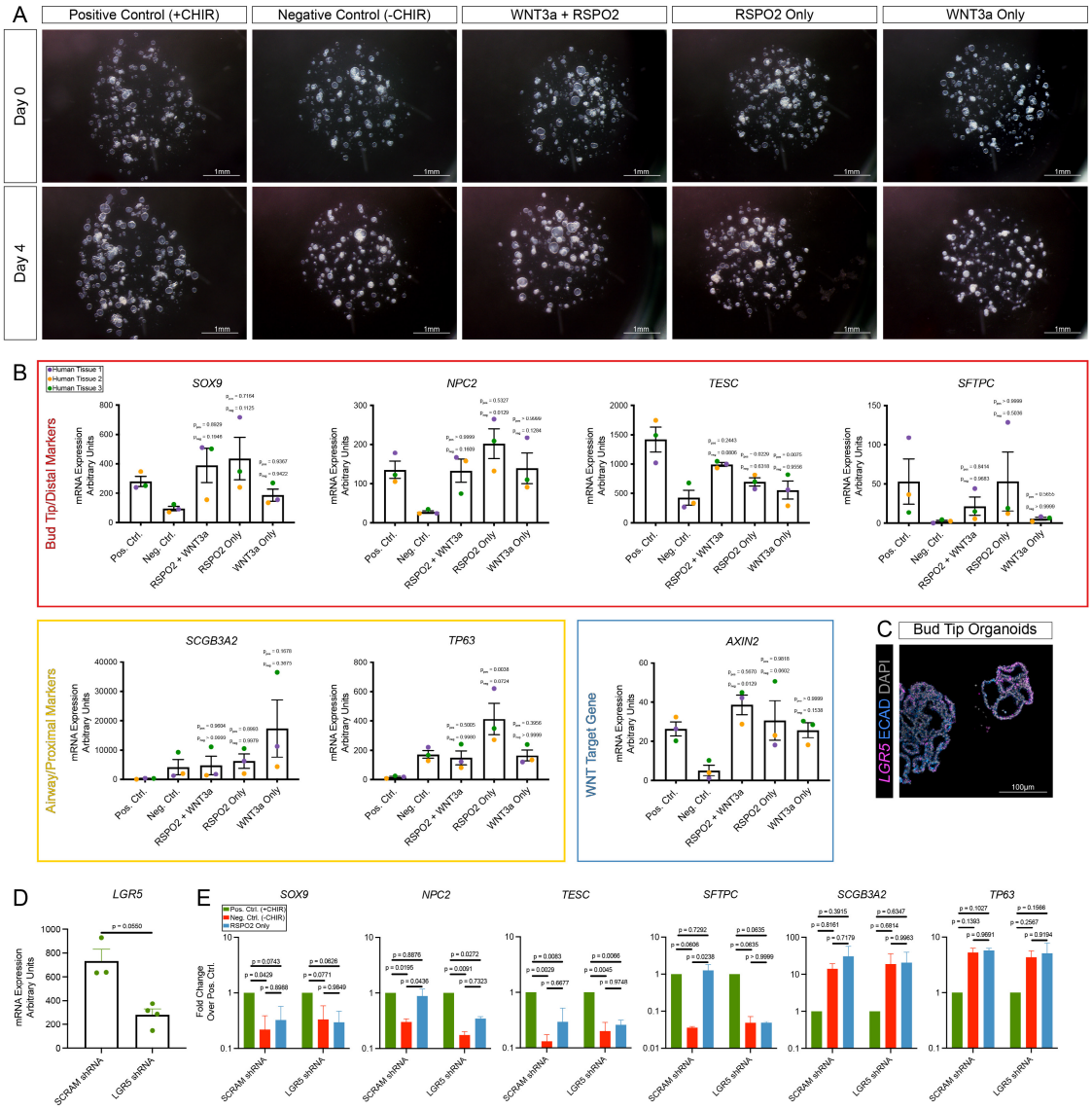


Figure 2-9: LGR5 can Respond to RSPO2 to Maintain Bud Tip Cell Fate. (A) Inverted microscope images of human fetal lung-derived bud tip organoids (BTOs) cultured with 3 μM CHIR99021 (CHIR) (positive control), no CHIR or any other WNT activator (negative control), 1X WNT3a afamin conditioned media (CM) with 500 μg/mL recombinant human RSPO2, RSPO2 only, or WNT3a only at the start of culture (day 0) and after 4 days (day 4). Media for every condition also contained 10 ng/mL FGF7 and 50 nM ATRA. **(B)** qRT-PCR for bud tip/distal (*SOX9*, *NPC2*, *TESC*, *SFTPC*), proximal/airway (*SCGB3A2*, *TP63*), and WNT target gene (*AXIN2*) markers on BTO cultures in each condition explained in Figure 2-9A. Each color represents bud tips from a unique specimen. Error bars represent standard error of the mean. Statistical tests were performed by ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each experimental group to the positive and negative control. **(C)** Fluorescence *in situ* hybridization of *LGR5* and co-immunofluorescence (IF) for the pan-epithelial marker ECAD on tissue sections from positive control (+CHIR) BTOs derived from 11.5-week distal lung tissue sections. **(D)** qRT-PCR for *LGR5* on positive control (+CHIR) BTO cultures given a scrambled or LGR5 shRNA. Error bars represent standard error of the mean. Statistical test was performed by unpaired Welch's one-tailed t-test. This experiment was performed on two unique biological specimens with 1 – 3 technical replicates per specimen. **(E)** qRT-PCR for bud tip/distal (*SOX9*, *NPC2*, *TESC*, *SFTPC*) and proximal/airway (*SCGB3A2*, *TP63*) markers on BTO cultures given a scrambled or LGR5 shRNA and cultured in BTO media with 3 μM CHIR (positive control: green), without CHIR (negative control: red), or 500 μg/mL recombinant human RSPO2 in place of CHIR (RSPO2 only: blue). Error bars represent standard error of the mean. Statistical test was performed by ordinary two-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group within the scrambled and shRNA sets. This experiment was performed on two unique biological specimens with 1-3 technical replicates per specimen.

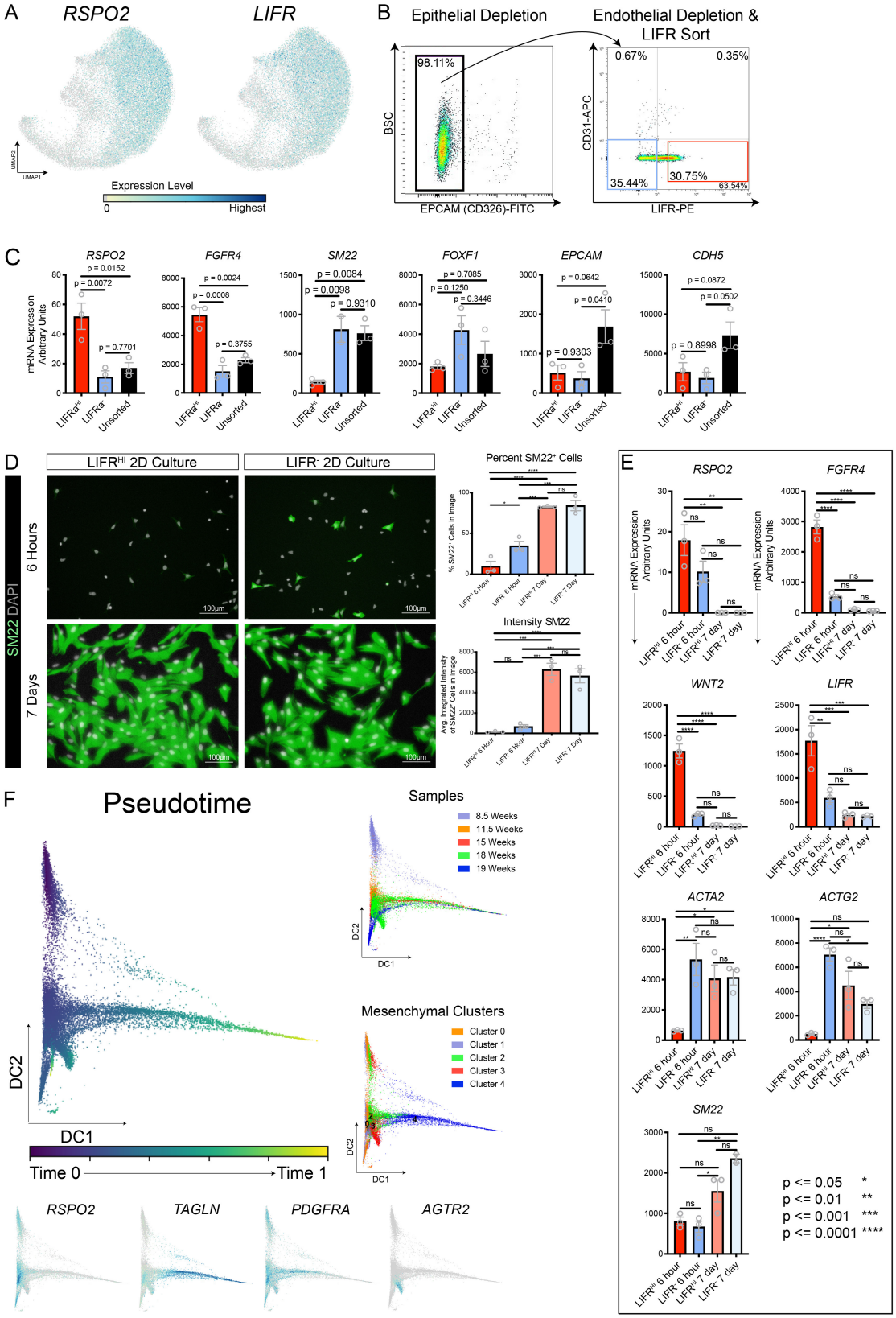

A
RSPO2
LIFR
UMAP2
UMAP1
Expression Level
0
Highest
B
Epithelial Depletion
Endothelial Depletion & LIFR Sort
98.11%
BSC
EPCAM (CD326)-FITC
0.67%
0.35%
CD31-APC
35.44%
30.75%
63.54%
LIFR-PE
C
RSPO2
FGFR4
SM22
FOXF1
EPCAM
CDH5
mRNA Expression Arbitrary Units
p = 0.0152
p = 0.0072
p = 0.7701
p = 0.0024
p = 0.0008
p = 0.3755
p = 0.0084
p = 0.0098
p = 0.9310
p = 0.7085
p = 0.1250
p = 0.3446
p = 0.0642
p = 0.0410
p = 0.9303
p = 0.0872
p = 0.0502
p = 0.8998
LIFRHI
LIFRa
Unsorted
D
LIFRHI 2D Culture
LIFR- 2D Culture
SM22 DAPI
6 Hours
7 Days
100μm
Percent SM22+ Cells
% SM22+ Cells in Image
Intensity SM22
Avg. Integrated Intensity of SM22+ Cells in Image
LIFRHI 6 Hour
LIFR- 6 Hour
LIFRHI 7 Day
LIFR- 7 Day
E
RSPO2
FGFR4
WNT2
LIFR
ACTA2
ACTG2
SM22
mRNA Expression Arbitrary Units
LIFRHI 6 hour
LIFR- 6 hour
LIFRHI 7 day
LIFR- 7 day
ns
p <= 0.05 *
p <= 0.01 **
p <= 0.001 ***
p <= 0.0001 ****
F
Pseudotime
DC2
DC1
Time 0
Time 1
Samples
8.5 Weeks
11.5 Weeks
15 Weeks
18 Weeks
19 Weeks
Mesenchymal Clusters
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
RSPO2
TAGLN
PDGFRA
AGTR2

Figure 2-10: FACS and 2D Culture of RSPO2$^+$ and SM22$^+$ Mesenchymal Cell Populations, Related to Figure 2-11. (A) UMAP feature plots corresponding to the cluster plot in Figure 2-1B and displaying expression levels *RSOP2* and *LIFR*. Each dot represents a single cell, and the color of each dot indicates log-normalized and z-transformed expression level of the given gene in the represented cell. **(B)** Flow cytometric analysis and pipeline for fluorescence activated cell sorting (FACS) of RSPO2$^+$ and SM22$^+$ mesenchymal cell populations by LIFR. Representative flow cytometry plots for EPCAM (CD326) (left plot) as well as CD31 and LIFR (right plot) from one tissue specimen. After gating of live, single cells, epithelial cells were negatively selected for via EPCAM. After negative selection of EPCAM$^+$ cells, endothelial cells were separated from mesenchyme via CD31. The CD326$^-$/CD31$^-$/LIFR^{HI} cells (Q4, red box) and CD326$^-$/CD31$^-$/LIFR$^-$ cells (Q3, blue box) were collected for co-culture experiments. LIFR^{HI} cells were defined as the top 30% of cells expressing LIFR. Isotype controls were used to set quadrants, and individual antibody/color channel controls were used to set compensation. **(C)** qRT-PCR for *RSPO2*$^+$ mesenchymal cell markers (*RSPO2*, *FGFR4*), airway smooth muscle cell markers (*SM22*, *FOXF1*), the epithelial cell marker *EPCAM*, and the endothelial cell marker *CDH5* from cells sorted using the FACS pipeline shown in Figure 2-10B. Error bars represent standard error of the mean. p-values were calculated using ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group. Each data point represents a unique biological replicate from three separate experiments. **(D)** Immunofluorescence for the smooth muscle marker SM22 on sorted LIFR^{HI} cells and LIFR$^-$ cells cultured in 2D with serum-free media for 6 hours or 7 days post-sorting. SM22$^+$ cell numbers and integrated staining intensity of SM22 was quantified for 3 technical replicates for each condition. Statistical tests were performed by ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group. '*' represents a p-value less than 0.05, '**' represents a p-value less than 0.01, '***' represents a p-value less than 0.001, '****' represents a p-value less than 0.0001, and 'ns' represents a p-value above 0.05. **(E)** qRT-PCR for *RSPO2*$^+$ mesenchymal cell markers (*RSPO2*, *FGFR4*, *WNT2*, *LIFR*) and airway smooth muscle cell markers (*SM22*, *ACTA2*, *ACTG2*) from sorted LIFR^{HI} cells and LIFR$^-$ cells cultured in 2D with serum-free media for 6 hours or 7 days post-sorting. Error bars represent standard error of the mean. Statistical tests were performed by ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group. Each data point represents a unique technical replicate from the same biological specimen. **(F)** Pseudotime analyses on single-cell RNA sequencing of non-vascular smooth muscle mesenchymal cells from 8.5 – 19-week post-conception human distal lungs (related to Figure 2-1B). Each dot represents a single cell. Plots are either colored by cluster or sample. Diffusion pseudotime (DPT) prediction of the temporal order of differentiating cells was carried out with cluster 0 set as the root cell (Haghverdi *et al.*, 2016). Feature plots show log-normalized and z-transformed expression level, with expression increasing with darker blue.

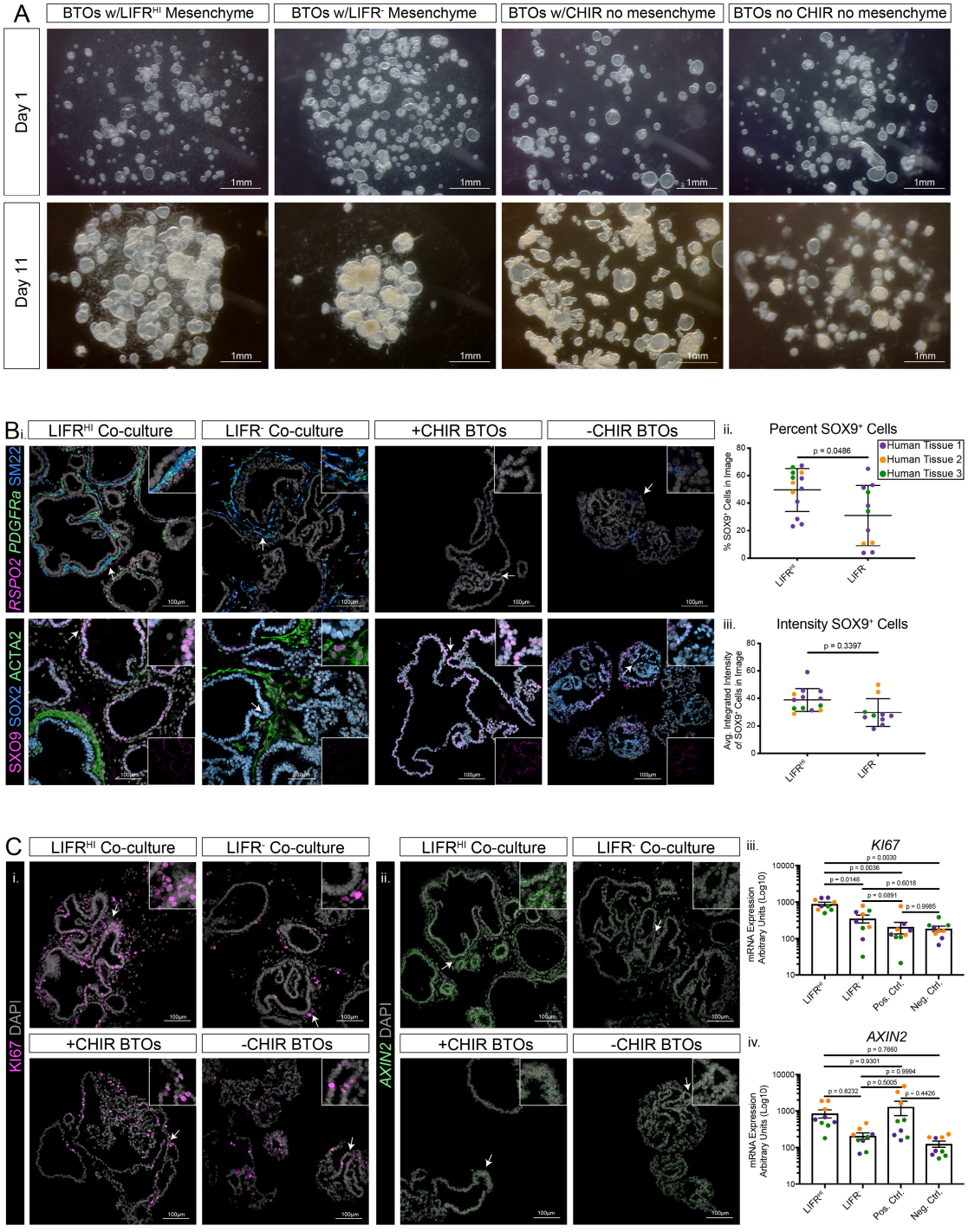


Figure 2-11: RSPO2$^+$ Mesenchymal Cells Support a Proximal and Distal Phenotype in Bud Tip Organoid Co-cultures. (A) Inverted microscope images of human fetal lung-derived bud tip organoids (BTOs) co-cultured with LIFR^{HI} mesenchyme, LIFR$^-$ mesenchyme, with previously-established BTO media (Miller *et al.*, 2018), or with BTO media where CHIR99021 (CHIR) was removed (same media co-cultures were grown in) the day after the start of the culture (day 1) and the day of collection (day 11). **(B)** Proximal/distal epithelial patterning in relation to mesenchymal cell type localization. **(i.)** Multiplexed fluorescence *in situ* (FISH) hybridization of *RSPO2* and *PDGFRα* and co-immunofluorescence (IF) for smooth muscle marker SM22 (top) or IF for the distal/bud tip epithelial marker SOX9, proximal/airway epithelial marker SOX2, and smooth muscle marker ACTA2 (bottom) on sections from LIFR^{HI} co-cultures, LIFR$^-$ co-cultures, positive control (+CHIR) BTOs, and negative control (-CHIR) BTOs after 11 days of culture. DAPI is shown in gray. Insets on the bottom right show a single channel image of SOX9. Insets on the

top right are zoomed in on the area denoted by the arrows. **(ii – iii.)** Quantification of $SOX9^{+}$ cell numbers and integrated staining intensity of SOX9 from IF stains. $LIFR^{HI}$ co-cultures retained higher numbers of $SOX9^{+}$ cells (unpaired Welch's one-tailed t-test) at higher levels (Welch's one-tailed t-test) compared to $LIFR^{-}$ co-cultures. The quantification in ii. and iii. was performed in three unique biological samples (mesenchyme and bud tips from unique specimens) with one to three technical replicates and a minimum of two image fields for each sample. **(C)** KI67 and *AXIN2* staining and quantification. **(i. – ii.)** IF for KI67 (left) and FISH for *AXIN2* (right) on sections from $LIFR^{HI}$ co-cultures, $LIFR^{-}$ co-cultures, positive control (+CHIR) BTOs, and negative control (-CHIR) BTOs after 11 days of culture. Insets on the top right are zoomed in on the area denoted by the arrows. **(iii. – iv.)** qRT-PCR for *KI67* and *AXIN2* on $LIFR^{HI}$ co-cultures, $LIFR^{-}$ co-cultures, positive control (+CHIR) BTOs, and negative control (-CHIR) BTOs after 11 days of culture from three independent experiments. Each color represents an independent experiment using bud tips and mesenchyme from unique specimens. Each data point of the same color represents a technical replicate from the same set of tissue specimens. Error bars represent standard error of the mean. Statistical tests were performed by ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group.

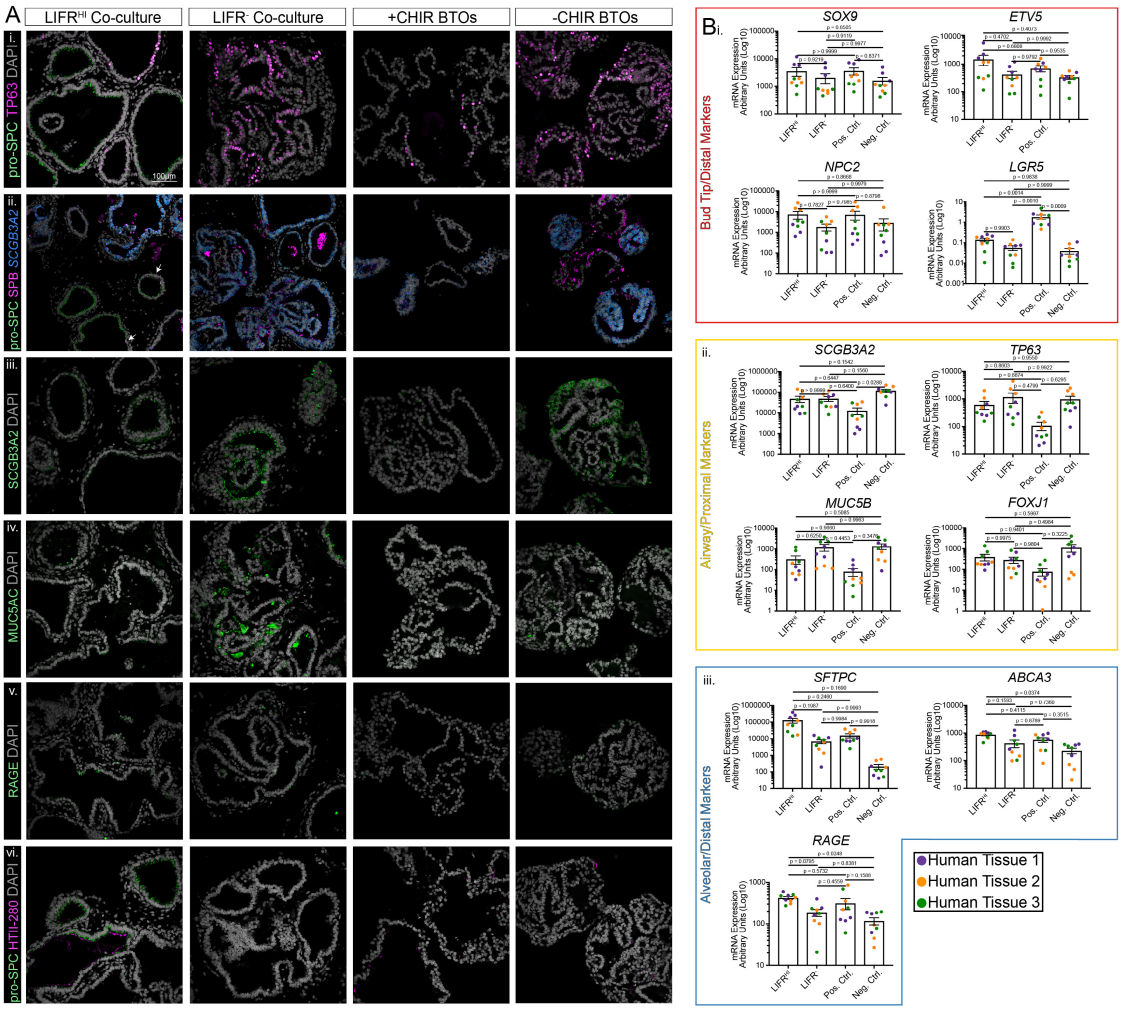


Figure 2-12: Differentiated airway and alveolar cell type markers in bud tip & mesenchyme co-cultures, Related to Figure 2-11. (A) Expression of proximal (airway) and distal (alveolar) differentiated cell type markers on tissue sections from LIFR^{HI} co-cultures, LIFR^{-} co-cultures, positive control (+CHIR99021; CHIR) BTOS, and negative control (-CHIR) BTOS after 11 days of culture from three independent experiments. Scale bar show on first image applies to all images. **(i.)** Immunofluorescence (IF) for alveolar marker pro-SFTPC and basal cell marker TP63. **(ii.)** Fluorescence *in situ* hybridization for secretory cell marker *SCGB3A2* with co-IF for alveolar markers pro-SFTPC and SFTPB. Arrows point to cells co-expressing pro-SFTPC and SFTPB. **(iii.)** IF for secretory cell marker SCGB3A2. **(iv.)** IF for goblet cell marker MUC5AC. **(v.)** IF for alveolar type I cell marker RAGE. **(vi.)** IF for alveolar type II markers pro-SFTPC and HTII-280. **(B)** qRT-PCR for distal/bud tip (*SOX9*, *ETV5*, *NPC2*, *LGR5*), proximal/airway (*SCGB3A2*, *TP63*, *MUC5B*, *FOXJ1*), and distal/alveolar (*SFTPC*, *ABCA3*, *RAGE*) markers on LIFR^{HI} co-cultures, LIFR^{-} co-cultures, positive control (+CHIR) BTOS, and negative control (-CHIR) BTOS after 11 days of culture from three independent experiments. Each color represents an independent experiment using bud tips and mesenchyme from unique specimens. Each data point of the same color represents a technical replicate from the same set of tissue specimens. Error bars represent standard error of the mean. Statistical tests were performed by ordinary one-way ANOVA followed by Turkey's multiple comparison test to compare the mean of each group with the mean of every other group.

Chapter 3 Stable iPSC-derived NKX2-1$^+$ Lung Bud Tip Progenitor Organoids Give Rise to Airway and Alveolar Cell Types[3]

3.1 Introduction

Advances in directed differentiation methods have led to the development of numerous embryonic or induced pluripotent stem cell (iPSC)-derived cell and organoid models of the airway and alveoli, which have enhanced our ability to model human lung development and disease (Wang *et al.*, 2007; Mou *et al.*, 2012; Gotoh *et al.*, 2014; Huang *et al.*, 2014; Dye *et al.*, 2015, 2016; Konishi *et al.*, 2016; Chen *et al.*, 2017; Yamamoto *et al.*, 2017, 2020; Hawkins *et al.*, 2017; Katherine B. McCauley *et al.*, 2017b; Miller *et al.*, 2018, 2019; Tamò *et al.*, 2018; de Carvalho *et al.*, 2019; Jacob *et al.*, 2019; Leibel *et al.*, 2020). Airway and alveolar cell types in mice and humans are derived from a common, developmentally transient progenitor population, called bud tip progenitors, which reside at the tips of the branching tree-like network of tubes that make up the lung epithelium (Serra, Pelton and Moses, 1994; Perl *et al.*, 2005; Abler, Mansour and Sun, 2008; Goss *et al.*, 2009; Rawlins *et al.*, 2009; Rockich *et al.*, 2013; Alanis *et al.*, 2014; Nikolić *et al.*, 2017; Miller *et al.*, 2018).

Bud tip progenitors obtained from the human fetal lung and grown as organoids serve as a useful tool to study the mechanisms responsible for bud tip progenitor cell maintenance and differentiation into airway and alveolar cell types (Nikolić *et al.*, 2017; Miller *et al.*, 2018, 2020; Conway *et al.*, 2020; Sun *et al.*, 2021). Despite this progress, organoids derived from fetal tissue

[3] This chapter has been published: Hein, Renee F. C., Ansley S. Conchola, Alexis S. Fine, Zhiwei Xiao, Tristan Frum, Lindy K. Brastrom, Mayowa A. Akinwale, et al. 2022. "Stable IPSC-Derived NKX2-1+ Lung Bud Tip Progenitor Organoids Give Rise to Airway and Alveolar Cell Types." *Development* 149 (20). https://doi.org/10.1242/DEV.200693.

are not broadly accessible to the research community and are associated with ethical and regulatory challenges, emphasizing the importance of iPSC-derived lung models. While we and others have made some progress in developing bud tip progenitor cell-like models from iPSCs (Chen *et al.*, 2017; Miller *et al.*, 2018), new technologies such as single cell RNA sequencing (scRNA-seq) and single cell lineage tracing have highlighted off-target cell types and unexpected plasticity in iPSC-derived cultures, where cells appear to be committed to a specific cell type or lineage but subsequently change fate (Little, Gerner-Mauro, Flodby, Crandall, Borok, Akiyama, *et al.*, 2019; Hurley *et al.*, 2020). This concept of cellular plasticity within the lung has also been demonstrated *in vivo,* where hyperactive WNT signaling in lung progenitors was shown to cause differentiation of intestinal cells in transgenic mouse embryos (Okubo and Hogan, 2004). Therefore, a challenge in the field, addressed in the current work, is to develop a long-lived and transcriptionally stable bud tip progenitor-like model from iPSCs.

Single cell RNA sequencing technologies have also made it possible to benchmark iPSC-derived cultures against primary tissue to compare transcriptional similarity and to accurately catalogue the diversity of on-target or off-target cell types observed *in vitro* (Hawkins *et al.*, 2017, 2021; McCauley *et al.*, 2018; Holloway, Wu, *et al.*, 2020; Hor *et al.*, 2020; Yu *et al.*, 2021). Since iPSC-derived cultures are known to be plastic, and the fact that iPSC differentiation is not 100% efficient, benchmarking has become an important step towards understanding the full complement of cells present in a culture. The current study therefore also sought to benchmark iPSC-derived bud tip progenitor organoids to interrogate the diversity of cell types in culture and similarity to primary bud tip progenitor organoids from the fetal lung.

This chapter demonstrates our optimized iPSC directed differentiation paradigm to generate self-organizing 3D spheroids with robust NKX2-1 expression. Expansion of NKX2-1^{+}

cells in bud tip progenitor media over 3 – 17 weeks gave rise to heterogenous organoids that contained NKX2-1$^+$ bud tip progenitor-like cells co-expressing markers of human bud tip progenitors, including SOX9, SOX2, and the cell surface marker CPM (Yamamoto *et al.*, 2020). Using an NKX2-1 reporter iPSC line along with CPM to quantitatively assess cultures via flow cytometry, we observed that bud tip progenitor-like cells expanded over subsequent weeks in culture. FACS isolation and further culture allowed for the expansion of NKX2-1$^+$/CPM$^+$ cells as bud tip progenitor-like organoids (iBTOs) that maintained ~80% NKX2-1$^+$/CPM$^+$ cells for at least 8 weeks. scRNA-seq analysis of bud tip progenitor cells from unsorted organoids or following FACS-enrichment revealed a high degree of transcriptional similarity to primary bud tip progenitor organoids as well as a shared transcriptional signature with *in vivo* bud tip progenitors. In addition, scRNA-seq from iBTOs that have spent less (3 weeks) or more (10 weeks) time in culture suggest that induced bud tip progenitors become more transcriptionally similar to native bud tip progenitors as they age. Finally, we used well-established methods to direct differentiation of iBTOs into organoids composed of airway epithelium (including basal, secretory, ciliated, goblet, and neuroendocrine cells) or alveolar type II (AT2) cells (Jacob *et al.*, 2017; Miller *et al.*, 2020). Collectively, this study describes a robust method to generate bud tip progenitor-like cells from iPSCs that closely resemble organoids derived from primary tissue. This model can be readily used to study lung development and illustrates a proof-of-concept for cellular engineering and cell therapy.

3.2 Results

3.2.1 Lung Spheroids are Optimized for NKX2-1 Expression but Remain Heterogenous

NKX2-1 is the earliest marker during lung epithelial specification (Lazzaro *et al.*, 1991) and loss of NKX2-1 leads to lung agenesis (Minoo *et al.*, 1995, 1999; Little, Gerner-Mauro,

Flodby, Crandall, Borok, Akiyama, *et al.*, 2019; Kuwahara *et al.*, 2020). We therefore sought to build on a previously-published method to generate iPSC-derived foregut spheroids (Dye *et al.*, 2015) by optimizing endoderm induction efficiency and foregut spheroid development in order to improve NKX2-1 expression.

We began by testing conditions to improve definitive endoderm (DE) differentiation efficiency and reproducibility. DE induction of varying efficiencies is achieved by using Activin A (ACTA) ligand (Kubo *et al.*, 2004; D'Amour *et al.*, 2006; Spence *et al.*, 2011) with WNT and BMP signaling playing a synergistic role during the initial stages DE specification (Gadue *et al.*, 2006; Green *et al.*, 2011; Loh *et al.*, 2014; Matsuno *et al.*, 2016; Rankin *et al.*, 2018; Heemskerk *et al.*, 2019). Therefore, we tested combinations of ACTA alongside the small molecule WNT activator CHIR99021 (CHIR) or BMP4 on the first day of a three-day ACTA differentiation culture. Using flow cytometry with a SOX17-tdTomato and SOX2-mCITRINE hESC reporter line (Martyn *et al.*, 2018) to quantitate cell composition, we observed that ACTA alone induced 48% $SOX17^+$ cells, while the addition of CHIR or BMP4 both enhanced DE differentiation, leading to 96% $SOX17^+$ cells or 87% $SOX17^+$ cells respectively (Figure 3-1A). Addition of CHIR and BMP4 together led to 71% $SOX17^+$ induction (Figure 3-1A). DE cultures from an additional cell line were co-stained with SOX17 and FOXA2 to confirm definitive endoderm cell identity (Figure 3-1B). We observed that near-pure $SOX17^+$ cultures obtained via ACTA and CHIR failed to give rise to self-organizing 3D spheroids, consistent with published data showing that self-organizing foregut and hindgut organoids consist of both epithelium and mesenchymal lineages (Spence *et al.*, 2011; Dye *et al.*, 2015, 2016). Therefore, ACTA and BMP4 were used for subsequent experiments.

Following DE specification, monolayers were directed into 3D foregut endoderm spheroids by combining a method that efficiently induces ventral foregut endoderm competent to be specified as lung (Rankin *et al.*, 2016) alongside methods that induce 3D self-organization (Spence *et al.*, 2011; Dye *et al.*, 2015). This included BMP inhibition via NOGGIN (NOG), FGF4, and CHIR (required for 3D spheroid formation) for 3 days plus all trans retinoic acid (ATRA) on the last day (Figures 3-1C, D). After 3 days, spheroids were collected and suspended in Matrigel and treated for 3 additional days with low-BMP4 as well as WNT3A and RSPO1, which together stimulate WNT signaling (Shu *et al.*, 2005; Goss *et al.*, 2009; Harris-Johnson *et al.*, 2009; Domyan *et al.*, 2011; Jacobs, Ku and Que, 2012; Miller *et al.*, 2012; Serra *et al.*, 2017). WNT activation through RSPO and WNT or CHIR resulted in comparable expression levels of foregut and hindgut markers (Figure 3-1E); however, CHIR is a GSK3b inhibitor and can have non-WNT mediated effects, so RSPO and WNT were used as more specific activators of WNT signaling.

At the end of the 9-day directed differentiation (Figure 3-2A), spheroids were analyzed for *NKX2-1* expression via qRT-PCR (on day 10) (Figure 3-2B). When this optimized method was directly compared to spheroids generated using previously published foregut/lung spheroid protocols (Dye *et al.*, 2015; Miller *et al.*, 2019), NKX2-1-optimized foregut spheroids expressed approximately 100-fold more *NKX2-1* (Figure 3-2B). Undifferentiated iPSCs and hindgut spheroids were included as controls, both of which had very low *NKX2-1,* and fetal lung was used as a positive control. Through analyzing EGFP expression with a NKX2-1-EGFP reporter cell line, EGFP was not detected at day 7, whereas a low-level of ubiquitous expression with scattered NKX2-1^{HI} cells could be detected starting on day 10, and expression was localized to specific regions by day 13 (Figure 3-2C). *NKX2-1* induction between day 7 and day 10 was

confirmed by qRT-PCR (Figure 3-2D), and whole mount immunofluorescence (IF) on day 10 spheroids correlated with reporter expression, with individual cells expressing high levels of NKX2-1 protein (Figure 3-2E, arrowheads).

When day 10 spheroids were analyzed by scRNA-seq, *NKX2-1* was also observed in a subset of cells (Figures 3-2F, G), revealing heterogeneity within the foregut spheroids and supporting EGFP reporter expression. Spheroids contained two clusters (Clusters 1 and 2) of lung-fated ventral foregut endoderm-like cells expressing *NKX2-1, FOXA2*, *FOXP1*, *FOXP4* and *HHEX* (Bogue *et al.*, 1998; Shu *et al.*, 2007; Spence *et al.*, 2009; Kearns *et al.*, 2013; Davenport *et al.*, 2016; S. Li *et al.*, 2016). Clusters and cells were largely negative for other foregut lineage markers including *TP63*, *TBX1*, *PDX1*, *PAX8*, and *ALB*; however, there was a clear population of cells that express *FOXA2*, *FOXA3*, *SOX17*, and *CDX2* (Cluster 3), which we refer to as "hindgut-primed" endoderm. Additionally, we observed a cluster of *CDX2*$^+$/*HAND1*$^+$/*ISL1*$^+$/*BMP4*$^+$/*FOXF1*$^+$/*LEF1*$^+$ (Cluster 4) foregut mesoderm-like cells (Han *et al.*, 2020) and a small population of cells (Cluster 5) expressing markers indicative of primordial germ cells, including *POU5F1*, *NANOG*, *T* (*TBXT*), *SOX17*, *NANOS3*, and *TFAP2C* (Davenport *et al.*, 2016; Jo *et al.*, 2022) (Figures 3-2G, 3-2G). Comparing optimized foregut spheroids to previously published methods (Dye *et al.*, 2015; Miller *et al.*, 2019) by qRT-PCR, we observed that the hindgut endoderm marker *CDX2* was not statistically different, but the early foregut and dorsal foregut endoderm marker *SOX2* (Que *et al.*, 2007) was reduced with the optimized method (Figure 3-2F), which suggests previous methods are competent to generate foregut that is relatively immature or more dorsal rather than ventral. Taken together, our data shows that optimized foregut spheroids have much higher levels of NKX2-1 when compared to prior

methods, but they are still heterogeneous with distinct populations of lung-fated cells and hindgut-fated cells.

3.2.2 iPSC-derived Bud Tip Progenitors Emerge Over Time

Once NKX2-1$^+$ spheroids were formed, we asked how efficiently these spheroids would give rise to bud tip progenitor (BTP)-like cells as spheroids expanded into larger organoid structures. Our previous studies have shown that "3 Factor (3F) media" possessing FGF7, CHIR99021, and ATRA expands primary BTPs derived from the fetal lung and SOX2$^+$/SOX9$^+$ iPSC-derived BTP-like cells; however, efficiency and culture over prolonged periods of time were not assessed (Miller *et al.*, 2018). After inducing NKX2-1-optimized spheroids on day 10, media was switched to 3F BTP media (Figure 3-3A). Spheroids were maintained in 3F for several weeks, where they expanded into complex, branching structures termed "lung progenitor organoids (LPOs)," similar to what we have observed previously (Miller *et al.*, 2018) (Figure 3-3B). LPOs were passaged every 2 – 3 weeks as either intact organoids with minimum fragmentation (whole passaged) or were sheared by being drawn through a hypodermic needle or pipette, which is a standard method for passaging BTP organoids derived from primary tissue (Miller *et al.*, 2018, 2020; Renee F.C. Hein *et al.*, 2022). Using an NKX2-1-EGFP reporter iPSC line to compare retention of lung identity after various passaging methods, whole passaged organoids maintained robust EGFP$^+$ reporter expression while fragmenting organoids ultimately led to a loss of NKX2-1-EGFP$^+$ cells (Figure 3-3C, quantified in Figure 3-3F). Based on the maintenance of NKX2-1-EGFP expression, we therefore chose to use whole passaged LPOs for our remaining experiments.

We evaluated the presence of induced bud tip progenitor (iBTP) cells within whole passaged LPOs by IF of paraffin sections and whole mount (Figures 3-3D, 3-4A). LPOs

possessed regions of NKX2-1-expressing cells, which co-expressed the BTP markers CPM, SOX9 and SOX2 (Nikolić *et al.*, 2017; Miller *et al.*, 2018; Yamamoto *et al.*, 2020) (Figures 3-3D, 3-4A). LPOs also contained sub-regions of distinct NKX2-1^{-}/CDX2^{+} cells, with many expressing MUC2, indicative of goblet-like intestinal cells, and smaller regions of NKX2-1^{-}/CDX2^{-} cells (Figures 3-3E, 3-4B).

We quantitatively assessed EGFP^{+}/CPM^{+} cells in LPOs grown for 3 – 17 weeks in 3F media using fluorescence activated cell sorting (FACS) (Figures 3-3F, 3-4C). We observed an increase in EGFP^{+}/CPM^{+} cells in culture over time (Figure 3-3F, purple bars). At 3.5 weeks, 17% of cells expressed CPM and EGFP, in contrast to 17 weeks, where 51% of cells in culture were EGFP^{+}/CPM^{+} (Figure 3-3F). A small portion (<10%) of cells were singly positive for EGFP (green) or CPM (blue) at any timepoint, and a population of double-negative cells was observed at all times (Figure 3-3F), suggesting that some heterogeneity is maintained in LPOs. We also investigated LPOs derived from two additional non-reporter iPSC lines, using only CPM to quantify iBTPs. We observed a similar increase of iBTPs in culture over time with 19% CPM^{+} cells at 3.5 weeks and 42% CPM^{+} cells by 17 weeks (Figure 3-4D, left). Lastly, organoids derived from previously-published lung organoid protocols (Dye *et al.*, 2015; Miller *et al.*, 2019) contained approximately 0.5 – 6% CPM^{+} cells at any timepoint examined (Figure 3-4D, right). The increase in CPM^{+} cells over time suggests that optimized culture conditions promote the emergence and selection of iBTPs.

To further interrogate the heterogeneity and complexity of the LPOs, we performed scRNA-seq on whole passaged LPOs at 3, 6, and 10 weeks (Figure 3-3G). We identified one cluster with robust levels of *NKX2-1* and BTP gene expression (Cluster 2), several clusters enriched for hindgut markers (Clusters 0, 5, 6), a mesenchymal cluster (Cluster 3), a

neuroendocrine-like cluster (Cluster 4), and a cluster of unknown/uncommitted cells (Cluster 1) (Figures 3-3G, 3-4G, 3-5). The proportion of cells in the BTP cluster (Cluster 2) increased over time while mesenchyme (Cluster 3) was depleted over time and hindgut cells (Clusters 0, 5, 6) were persistent (Figures 3-4E, F). Together this data supports FACS data suggesting that iBTPs continue to expand within LPOs over time and identifies contaminating lineages that persist.

To assess the proliferation of the cultures, we quantified *KI67* expression within each cluster and sample from the LPO scRNA-seq data (Figure 3-4H). *KI67*$^{+}$ cells were present in every cluster; however, the hindgut (Clusters 0, 5, 6), BTP (Cluster 2), and mesenchymal (Cluster 3) clusters accounted for 92.24% of *KI67*$^{+}$ cells within the culture compared to 7.76% from the uncommitted (Cluster 1) and neuroendocrine (Cluster 4) clusters (Figure 3-4H, top). Contribution of *KI67*$^{+}$ cells to each cluster by sample was also calculated and normalized to the number of cells within each sample (Figure 3-4H, bottom). The contribution of samples to the *KI67*$^{+}$ cells within clusters was varied. Between the 3-week samples, there were proliferating cells that contributed to all clusters; however, there was a significant contribution of this early timepoint to the uncommitted cluster (Cluster 1), comprising 87.11% of *KI67*$^{+}$ cells within this cluster. The 6-week samples also contributed *KI67*$^{+}$ cells to all the clusters, with the most significant contributions to hindgut (Clusters 0, 5, 6) and the mesenchyme (Cluster 3). The data suggests that by 6 weeks, any remaining contaminating cell types (i.e., *CDX2*$^{+}$ intestinal cells) are most proliferative, while the earlier cultures are altogether generally proliferative. Finally, the 10-week sample contributed a significant proportion of proliferating cells to the BTP cluster (Cluster 2; 39.14% of *KI67*$^{+}$ cells), the neuroendocrine cluster (Cluster 4), and the two smaller hindgut clusters (Cluster 5 & 6) (Figure 3-4H, top). Importantly, all samples contributed to the BTP cluster, and with a robust contribution from the 10-week sample (Figure 3-4H, top).

Cell death in LPOs, evaluated by H&E and cleaved caspase 3 (CCAS3) stains, was evident at all stages of growth and did not appear to increase with time or localize to contaminating $CDX2^+$ regions; however, localization of CCAS3 staining transitioned from individual cells at 3 weeks to luminal regions by 17 weeks in culture (Figure 3-4I). This data suggests that the increasing number of iBTP cells observed is not likely due to death of contaminating cell types.

3.2.3 iPSC-derived Bud Tip Progenitors Can be Isolated, Expanded, and Maintained Long-term

Although LPOs at every time point contain a proliferating population of iBTPs, given the large population and persistence of hindgut lineages, we aimed to isolate NKX2-1-$EGFP^+$/CPM^+ cells to generate higher purities of induced bud tip progenitor organoid (iBTO) cultures. iBTPs were isolated via FACS with CPM and NKX2-1-EGFP (or CPM only for non-reporter cell lines) and were replated in Matrigel at 5,000 cells/μL in 3F media to form iBTOs (Figure 3-6). After sorted cells formed into organoids, iBTOs were passaged by whole passaging approximately every two weeks. The NKX2-1-EGFP reporter line showed uniform expression of NKX2-1-EGFP in iBTOs (Figure 3-6B), and IF of paraffin sections showed iBTOs contain a near-homogenous population of cells expressing BTP markers CPM, SOX9, and SOX2 (Figure 3-6C). Interestingly, the age of LPOs at the time of sorting was a critical determinant of the ability of iBTOs to maintain NKX2-1 and CPM expression. For example, iBTOs generated from <6-week LPOs maintained 32% NKX2-1-$EGFP^+$/CPM^+ cells when resorted 7 weeks later (Figure 3-6D). In contrast, iBTOs from >6-week LPOs maintained 80% NKX2-1-$EGFP^+$/CPM^+ cells when resorted 8 weeks later (Figure 3-6D). This suggests that iBTPs undergo increasing commitment to a bud tip progenitor identity as they are maintained in 3F media at the LPO stage.

To assess iBTO expansion, we evaluated the growth rate of iBTOs from the day iBTPs were isolated from LPOs until 6-weeks of growth as iBTOs. Small cysts could be detected by 1 week of culture, which grew in circumference by 2 weeks (Figure 3-6E). By 4 weeks of culture, cysts took on a more complex, branched phenotype, which increased in complexity by 6 weeks (Figure 3-6E). Growth of iBTOs sorted from 4 – 6-week LPOs (early timepoints) or 10 – 11-week LPOs (late timepoints) was quantified by counting the number of cells in organoids generated from 40,000 to 75,000 original cells (depending on organoid batch) at 2-, 4-, and 6-weeks in culture. Overall, a steady increase in cell number was observed for iBTOs sorted from early or late LPOs, with rare batches showing low growth efficiency (Figure 3-6E, right). We also measured organoid forming efficiency of iBTOs by dissociating iBTOs to single cells and replating iBTPs to determine how many cysts formed after 2 weeks. iBTOs from later sorts (10 – 11-week LPOs) had higher organoid forming efficiency than iBTOs from earlier sorts (4 – 6-week LPOs) (Figure 3-6F).

3.2.4 iPSC-derived Bud Tip Progenitors are Transcriptionally Similar to Primary Bud Tip Progenitor Cultures

To further interrogate iBTO cellular composition, and to directly compare cells within iBTOs to primary (*in vivo*) bud tip progenitors and primary bud tip progenitor organoids derived from fetal tissue (Miller *et al.*, 2018), we performed scRNA-seq on iBTOs derived from 4- and 10-week LPOs. iBTOs were expanded for 4 weeks in culture post-sorting before scRNA-seq was carried out. Integrated analysis of both data sets resulted in 4 epithelial clusters (Figure 3-7A). Cells in cluster 4 expressed hindgut/intestinal markers (i.e., *CDX2*), which could also be identified by immunofluorescence (Figure 3-8A) and represented a small fraction of cells that were predominantly derived from the 4-week sample (Figures 3-7B, C). The remaining 3 clusters

(Clusters 1, 2, 3) have enriched expression of *NKX2-1* as well as BTP markers (*SOX9, CPM, ETV5, TESC*, *FGF20*, *SOX11*, *HMGB2*, *NPC2*, *LGR5*, *ETV4*) (Miller *et al.*, 2018, 2020; Renee F.C. Hein *et al.*, 2022); however, expression of proliferation genes was variable, suggesting there is some heterogeneity representing different iBTP cell states, that is likely driven by the expression of proliferation genes (Figures 3-7D, 3-8B). Similar to the analysis done on LPOs, we quantified *KI67* expressing cells in each cluster and sample in the iBTO scRNA-seq data (Figure 4-8C). All clusters had similar normalized percentages of *KI67*$^+$ cells (between 15 and 34%) (Figure 3-8C, left panel); however, the contribution of *KI67*$^+$ cells to each cluster varied by sample (i.e., timepoint). iBTOs from the 10-week sort contributed more *KI67*$^+$ cells to BTP Clusters 0 and 1 while iBTOs from the 4-week sort contributed more to the hindgut cluster (Cluster 4). Both samples contributed similarly to BTP Cluster 3 (Figure 3-8C, right panel). Together with the analysis on LPOs (Figure 3-3H), this suggests that although iBTPs are proliferative, the highly proliferative contaminating lineages (i.e., hindgut) can persist in culture despite the sorting strategy if the cultures are sorted too early. This could be attributed to plasticity within cultures during early time points.

To further benchmark transcriptional similarities of iBTPs to *in vivo* BTPs, we re-analyzed published scRNA-seq data from the primary human fetal lung (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022) and identified a panel of the top 100 enriched genes expressed in *in vivo* BTPs relative to all other epithelial cell types in the developing lung (see 'Materials and Methods'). We used this gene list as a reference to assign an "*in vivo* bud tip progenitor cell score" to each individual cell from the LPOs and iBTOs using published computational methods (Holloway, Wu, *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). For the purposes of this comparison, both whole LPOs and extracted *NKX2-1*$^+$ cells from the LPO data sets were included in the

analysis. As additional control comparisons, we analyzed published data and included scores for primary BTOs, *in vivo* basal cells, and *in vivo* neuroendocrine cells (Miller *et al.*, 2018, 2020; Renee F.C. Hein *et al.*, 2022) (Figure 3-7E). As expected, *in vivo* BTPs had the highest score (median score = 0.80) while *in vivo* neuroendocrine cells had the lowest score (0.37), followed by whole LPOs (0.47). A noteworthy observation is that human primary BTOs had a mean BTP score of 0.62, which was much lower than the *in vivo* BTP score of 0.80, indicating that the *in vitro* culture conditions significantly influence gene expression, as has been observed previously (Miller *et al.*, 2020). Relative to primary BTOs, both *NKX2-1*$^{+}$ cells from LPOs and iBTOs showed high transcriptional similarity; their scores improved with longer time in culture at the LPO stage (0.51 to 0.57 for *NKX2-1*$^{+}$ cells from LPOs; 0.48 to 0.54 for iBTOs) (Figure 3-7E). To provide additional confidence in this comparison method, we generated a similar *in vivo* basal cell score and *in vivo* neuroendocrine cell score and observed that the expected populations (i.e., *in vivo* basal cells and *in vivo* neuroendocrine cells, respectively) scored the highest relative to all the other data sets analyzed (Figure 3-8D).

As an additional analysis, we utilized published scRNA-seq data from primary human fetal lung epithelium and performed 'label transfer' which mapped iBTOs to the most similar cell type within the human fetal lung reference data set (Stuart *et al.*, 2019a; Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). After clustering the primary epithelial cells (Figure 3-8E), iBTOs from LPOs sorted at 4-weeks or 10-weeks largely mapped to four different clusters (Figures 3-8F, G). Upon closer investigation into these clusters, three of the clusters (Clusters 2, 7 and 9) scored highly for BTP genes (using the BTP cell score) and expressed known markers of BTP cells; the fourth cluster (Cluster 6) resembled neuroendocrine progenitor cells (Figure 3-8F). Quantification of this mapping revealed that the majority of iBTO cells mapped to two of the

three BTP clusters (Clusters 2 & 7) while only a small sub-set mapped to the third BTP cluster (Cluster 9) and the neuroendocrine cluster (Cluster 6) (Figure 3-8G). Taken together, these data supports the conclusion that iBTOs contain a stable bud tip progenitor cell that shares a high degree of transcriptional similarity to primary bud tip progenitors.

3.2.5 iBTOs Can Give Rise to Airway and Alveolar Fates

Given that BTPs in the developing lung give rise to airway and alveolar fates, we hypothesized that iBTOs could be guided into airway and alveolar lineages. To test this possibility, we used methods previously developed to efficiently induce lung progenitors into airway (Miller *et al.*, 2020) or alveolar (Jacob *et al.*, 2017) lineages (Figure 3-9A). Airway induction involved 3 days of dual-SMAD activation (DSA) followed by 18 days of dual-SMAD inhibition (DSI) and resulted in condensed structures that maintained NKX2-1-EGFP expression and expressed mCherry driven by the TP63 promoter (TP63-mCherry) (Figure 3-10A, left column). By immunofluorescence, we observed that TP63 expression was highly induced after 3 days of DSA, as expected, and after 18 days of DSI, $TP63^+$ cells organized around the perimeter of the organoids (Figures 3-9B, C).

We performed airway induction on unsorted LPOs as well, hypothesizing that airway differentiation may select for expansion of lung lineages within the LPOs, while suppressing non-lung lineages. NKX2-1-$EGFP^+$ and TP63-$mCherry^+$ cells were quantified using flow cytometry on both LPOs and iBTOs following the 21-day DSA/DSI protocol or untreated controls (Figure 3-10B). Both untreated LPOs and iBTOs contained <1% NKX2-1-$EGFP^+$/TP63-$mCherry^+$ cells. Only upon treatment with the 21-day DSA/DSI protocol were significant numbers of $EGFP^+$/$mCherry^+$ cells detected (7% in LPOs, 5% in iBTOs) (Figure 3-10B). DSA/DSI-treated iBTOs also maintained a large portion (71%) of NKX2-1-$EGFP^+$/TP63-

mCherry$^-$ cells, which represented the spectrum of differentiated airway epithelial cell types identified by immunofluorescence and qRT-PCR, including multiciliated, neuroendocrine, goblet, and secretory cells (Figures 3-8C – D, 3-10B). Airway organoids contained AcTUB$^+$/*MUC16*$^+$ ciliated cells (Carraro *et al.*, 2021) (Figure 3-9D), and SYN$^+$ cells were double positive for ASCL1 (early neuroendocrine marker) or CHGA (late neuroendocrine marker) (Figure 3-9D). Treated LPOs contained a small proportion of cells only expressing TP63-mCherry (5%) and a large portion were negative for both markers (58%) (Figures 3-10B, D), suggesting non-lung lineages continue to expand in these conditions.

Alveolar induction consisted of 7 days of treatment with cyclic AMP and dexamethasone (CK + DCI) as has been previously described with the exception of an alternate composition of the base media (see 'Materials and Methods') (Jacob *et al.*, 2017). Treatment of iBTOs with CK + DCI resulted in expanded budded structures that maintained NXK2.1-EGFP expression (Figure 3-10A, right column). Alveolar type II markers *SFTPB*, *SFTPC*, and *ABCA3* were low to undetectable in iBTOs but were highly expressed upon treatment with CK + DCI, as determined by qRT-PCR, and co-expressed within the same cells as shown by immunofluorescence for (pro-)SFTPC, SFTPC, SFTPB, and NKX2-1 (Figures 3-9E, F). The lung-specific epithelial marker NKX2-1 was robust in both airway and alveolar organoids while expression of the intestinal epithelial marker CDX2 was not regularly detected (Figures 3-9C – F, 3-10C). Taken together, differentiation of iBTOs into airway and alveolar cell types with minimal contaminating non-lung cell types supports the observation that iBTOs have similar developmental potential to *in vivo* bud tip progenitors.

3.3 Discussion

With the current work, we demonstrate an optimized *in vitro* model system for differentiating iPSCs into an NKX2-1-expressing bud tip progenitor lineage, which we show can give rise to both airway and alveolar cell types. Several groups, including ours (Dye *et al.*, 2015, 2016; Miller *et al.*, 2018, 2019), have previously characterized protocols that yield lung-like cell types from iPSCs (Wang *et al.*, 2007; Mou *et al.*, 2012; Gotoh *et al.*, 2014; Huang *et al.*, 2014; Konishi *et al.*, 2016; Abler *et al.*, 2017; Chen *et al.*, 2017; Hawkins *et al.*, 2017; Katherine B. McCauley *et al.*, 2017b; Tamò *et al.*, 2018; de Carvalho *et al.*, 2019; Jacob *et al.*, 2019; Yamamoto *et al.*, 2020; Leibel *et al.*, 2020). While many of these protocols capture a transient bud tip progenitor-like stage, they focus primarily on deriving more mature lung epithelial lineages, and the induction and maintenance of NKX2-1^{+} bud tip progenitor-like cells from iPSCs has not been previously reported. Furthermore, benchmarking of induced cells using scRNA-seq in order to compare cell types to a human reference atlas has only recently become commonplace (Holloway, Wu, *et al.*, 2020; Hor *et al.*, 2020; Hawkins *et al.*, 2021; Yu *et al.*, 2021). The method reported here expands upon an improved understanding of critical signaling events that regulate bud tip progenitor maintenance in the human lung to better mirror this process in a dish (Conway *et al.*, 2020).

Although we demonstrate that optimizing NKX2-1 expression leads to a 100-fold increase when compared to previously reported methods (Dye *et al.*, 2015; Miller *et al.*, 2019), scRNA-seq and whole mount immunofluorescence analysis of day 10 spheroids revealed that low levels of NKX2-1 is expressed in many cells, while a small number of cells expressed high levels of NKX2-1, highlighting the heterogeneity that still exists at this early timepoint. We also observed non-lung lineages that co-emerge in spheroids, possibly due to the fact that the same

signaling pathways play a role during lineage commitment for multiple lineages while also reflecting the inherent plasticity of newly committed cells. We observed that non-lung lineages, and particularly hindgut lineages, persisted over time in culture. Based on this observation, there is an opportunity to further refine the signaling pathways that are manipulated to expand and maintain NKX2-1$^+$ lung lineages; however, this may prove challenging given the redundant use of some pathways (i.e., WNT) to maintain stem/progenitor cells from both lineages (Pinto and Clevers, 2005; Kapoor, Li and Leiter, 2007; Volckaert and De Langhe, 2015; Chin *et al.*, 2016; Nikolić and Rawlins, 2017; Yan *et al.*, 2017; Miller *et al.*, 2018; Rabata *et al.*, 2020; Aros, Pantoja and Gomperts, 2021). Additionally, since mesenchyme was observed by scRNA-seq at the spheroid and LPO stages, how these cells influence epithelial cell fate should be a topic of further exploration.

We observed that hindgut lineages expand after passaging organoids using the standard approach of mechanically shearing them into small fragments. It is currently unclear if this phenomenon occurs because shearing increases selection for highly proliferative CDX2$^+$ cells, if the mechanical stress results in higher levels of apoptosis in one lineage (i.e., lung) compared to the other, or if shearing disrupts the bud tip niche. Thus, while we have optimized the current methods, to generate a high proportion of NKX2-1$^+$ cells, LPO cultures still exhibited high plasticity that could be triggered through perturbations such as organoid dissociation.

One important observation from our current study is that induced bud tip progenitors exhibited increased lineage commitment with longer time in culture. Isolating NKX2-1-EGFP$^+$/CPM$^+$ cells prior to 6 weeks resulted in cultures replete with hindgut cells. While follow-up experiments are required to determine the origin of contaminating hindgut cells, the fact that we still see these cells in purified iBTO cultures suggests a fate switch from a lung bud tip

progenitor to a hindgut identity. On the other hand, purifying cells at 6 – 10 weeks reliably led to the establishment of NKX2-1-EGFP^{+}/CPM^{+} iBTOs that expanded and maintained their fate. Whether or not the mechanism underlying the loss of bud tip progenitors caused by sorting LPO cultures too early is the same as that underlying bud tip progenitor loss caused by shearing cultures is an avenue for future research.

We attempted airway differentiation on unsorted LPOs, hypothesizing that the process of differentiation itself may select against gut lineages and expand only lung lineages; unfortunately, non-lung lineages persisted. Given this result, sorting remains an important step in establishing iBTO cultures that can be differentiated into airway or alveolar cell types in the absence of non-lung cell types. Finally, the experiments carried out here relied heavily on an NKX2-1-EGFP reporter; however, our results suggest that this should not limit studies with non-reporter iPSC lines, since we also observed that sorting with CPM alone enriched for induced bud tip progenitors. As we detected some NKX2-1-EGFP^{-}/CPM^{+} cells, identifying additional bud tip progenitor cell surface markers to further improve purity of cultures may be valuable for enhancing purity in non-reporter iPSC cultures.

The utilization of emerging technologies such as scRNA-seq has provided critical insights into the heterogeneity and complexity of human tissues that were once difficult to study (Miller *et al.*, 2020; Yu *et al.*, 2021). New information from human tissue has provided both an atlas against which *in vitro* model systems can be benchmarked and a roadmap that can be used to infer transcriptional and signaling mechanisms that control cellular transitions. Here, we used scRNA-seq to benchmark *in vitro* cultures at several stages during differentiation to catalog the induction, emergence, and maintenance of lung-fated cells as they acquire a bud tip progenitor fate. One interesting observation from the benchmarking carried out in this study is a significant

shift in the transcriptome when comparing primary *in vivo* tissue to *in vitro* grown primary organoids, even when the source of the cells is the same. For example, here, we compared iBTOs to both primary BTOs and *in vivo* fetal bud tip progenitors. We observed a significant shift when comparing fetal lung bud tip progenitors to primary BTOs, indicating that the *in vitro* environment significantly changes the transcriptome of the cell, as has been recently reported (Miller *et al.*, 2020; Alysandratos *et al.*, 2022). With this caveat in mind, iBTOs shared a very similar transcriptome to primary BTOs and a high degree of similarity to *in vivo* primary bud tip progenitors. The cell scoring metric used to compare similarity across samples also suggested that iBTOs became more transcriptionally similar to *in vivo* bud tip progenitors as they spent more time in culture, supporting the idea that induced bud tip progenitors undergo continued differentiation towards a bud tip progenitor identity as they are maintained in culture.

Functionally, iBTOs also behave like bud tip progenitors. Using methods to differentiate lung progenitor cells into airway or alveolar cell types (Jacob *et al.*, 2017; Miller *et al.*, 2020), we demonstrated that iBTOs robustly differentiate into airway and alveolar cell types. Taken together, the current work describes a robust iPSC-derived bud tip progenitor model to study human lung epithelial development and differentiation and uses scRNA-seq to benchmark both on- and off-target cell types present in the cultures. Practically, this model presents a novel opportunity to expand the downstream efforts of lung organoid studies, where iBTOs can be shared to reduce the expertise needed for lung cell generation and ultimately facilitate faster turn-around through iBTO banking. Overall, this study enhances the utility of iPSC-derived lung organoids to interrogate lung development and to study disease and regeneration.

3.4 Materials and Methods

3.4.1 Data and Code Availability

Sequencing data used in this study is deposited at EMBL-EBI ArrayExpress. Single-cell RNA sequencing of human fetal lung and human fetal bud tip progenitor organoids: human fetal lung (ArrayExpress: E-MTAB-8221, ArrayExpress: E-MTAB-10662) (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022), spheroids, LPOs and iBTOs (ArrayExpress: E-MTAB-11953) (this study). Code used to process data can be found at: https://github.com/jason-spence-lab/Hein_Conchola_2022.

3.4.2 Cell Lines and Culture Conditions

hPSC Lines and Culture Conditions

LPOs were generated from three human induced pluripotent stem cell (iPSC) lines: human WTC11 (RRID: CVCL_Y803) was obtained from Bruce Conklin at the University of California San Francisco (Kreitzer *et al.*, 2013), human iPSC line 72.3 was obtained from Cincinnati Children's Hospital Medical Center (McCracken, Catá, Crawford, Sinagoga, Schumacher, Rockich, Y. H. Tsai, *et al.*, 2014) and human iPSC17 WT 7B2, expressing NKX2-1-EGFP and TP63-mCherry, was obtained from the Cystic Fibrosis Foundation (Crane *et al.*, 2015). In addition to iPSC lines, human embryonic stem cell (hESC) lines were used for definitive endoderm and spheroid optimizations: hESC line H9 (NIH registry no. 0062) was obtained from the WiCell Research Institute and hESC line RUES2-GLR (NIH registry no. 0013), expressing SOX2-mCitrine, BRA-mCerulean and SOX17-tdTomato, was obtained from The Rockefeller University (Martyn *et al.*, 2018). The University of Michigan Human Pluripotent Stem Cell Research Oversight Committee approved all experiments using hESC and iPSC lines. Stem cells were maintained as previously described (Spence *et al.*, 2011) and grown in mTeSR Plus media (StemCell Technologies, Cat#100-0276).

NKX2-1-optimized Spheroid Differentiation Protocol

Generation of definitive endoderm (DE) from hPSCs (differentiation days 1 – 3) was carried out as previously described with slight modifications (D'Amour *et al.*, 2006; McCracken *et al.*, 2011; Spence *et al.*, 2011; Dye *et al.*, 2015). Briefly, 100 ng/mL ActivinA (R&D Systems, Cat#338-AC) was added in RPMI 1640 media (Thermo Fisher, Cat#21875034) with increasing concentrations of HyClone defined FBS (dFBS) (Thermo Fisher, Cat#SH3007002) on subsequent days (0% day 1, 0.2% day 2, 2% day 3). 50 ng/mL BMP4 (R&D Systems, Cat#314-BP) was added on day 1. Following DE specification, anterior foregut spheroids were generated (differentiation days 4 – 6) by a 3-day treatment with 500 ng/mL FGF4 (lab purified – see supplemental methods), 200 ng/mL NOGGIN (R&D Systems, Cat#6057-NG), and 2 μM CHIR99021 (APExBIO, Cat#A3011). 2 μM all trans retinoic acid (Sigma, Cat#R2625) was added on the final day of anterior foregut spheroid generation (day 6). On day 7, self-organizing 3D spheroids that had detached from the tissue culture dish were collected with a P200 pipette and were transferred into Matrigel (Corning, Cat#354234) as previously described (McCracken *et al.*, 2011; Miller *et al.*, 2019). After the Matrigel had solidified, encapsulating spheroids, they were cultured in media that contained 250 ng/mL WNT3a (R&D Systems, Cat#5036-WN), 500 ng/mL RSPO1 (lab purified – see supplemental methods), and 10 ng/mL BMP4 for 3 days to induce NKX2-1$^+$ cells (differentiation days 7 – 9). Following DE specification (after day 3), RPMI 1640 media + 2% dFBS was used on all days to dilute growth factors. Fresh media and growth factors were added each day.

Growth and Maintenance of LPOs

NKX2-1-optimized spheroids were transferred to "3 Factor" (3F) bud tip maintenance media as previously described (Miller *et al.*, 2018, 2019), including 50 nM all trans retinoic acid

(Sigma, Cat#R2625), 10 ng/mL FGF7 (R&D Systems, Cat#251-KG), and 3 μM CHIR99021(APExBIO, Cat#A3011) in serum-free basal media. Serum-free basal media consists of DMEM/F12 containing HEPES and L-Glutamine (Corning, Cat#10-092-CV), 100 U/mL penicillin-streptomycin (Thermo Fisher, Cat#15140122), 1x B-27 supplement (Thermo Fisher, Cat#17504044), 1x N-2 supplement (Thermo Fisher, Cat#17502048), 0.05% BSA (Sigma, Cat#A9647), 50 μg/mL L-ascorbic acid (Sigma, Cat#A4544), and 0.4 μM 1-Thioglycerol (Sigma, Cat#M1753). LPOs were grown for 3 weeks, then whole passaged or needle or pipette sheared every 2 – 4 weeks. Whole passaging was achieved by collecting LPOs into a 1.5mL microcentrifuge tube and gently releasing them from Matrigel by a P1000 cut pipette tip (P200 tip was used for young/small LPOs). LPOs were spun in a microcentrifuge tube, residual media and Matrigel was removed, then LPOs were re-suspended in Matrigel (Corning, Cat#354234) with a P1000 cut pipette tip. ~35 μL droplets of Matrigel were placed into the center of wells of a 24-well tissue culture plate (Thermo Fisher, Cat#12565163), and the plate was inverted and placed in an incubator at 37ºC for 20 minutes. Whole passaging was performed approximately every 2 – 3 weeks at a ratio of 1:2 (individual LPOs were kept whole but each were given more space, i.e., n spheroids were given 2x the space to grow) for up to 17-weeks. For LPOs passaged by needle or pipette shearing, LPOs were passed through a 27-gauge needle or P200 pipette, respectively, and embedded in fresh Matrigel as previously described (Miller *et al.*, 2018, 2019, 2020). LPOs were fed with 3F media every 2 – 4 days.

Growth and Maintenance of iBTOs

iBTPs in LPOs grown in 3F media for 3 – 17 weeks were isolated by FACS (see below). After collection, iBTPs were centrifuged at 300g for 5 minutes at 4ºC and media was removed. Matrigel (Corning, Cat#354234) was added to the cells at a concentration of 5,000 cells/μL

Matrigel and 3 – 20 µL droplets of Matrigel were placed into the center of wells of a 24-well tissue culture plate (Thermo Fisher, Cat#12565163). The plate was inverted and placed in an incubator at 37ºC for 20 minutes. iBTPs were fed every 2 – 4 days with 3F media and allowed to reform organoids (iBTOs) for up to 4 weeks. iBTOs were whole passaged as described above every 2 – 3 weeks following organoid formation.

Airway and Alveolar Differentiations

Airway differentiation was carried out as previously described (Miller *et al.*, 2020). Briefly, iBTOs were exposed to dual-SMAD activation via 100 ng/mL BMP4 (R&D Systems, Cat#314-BP-050) and 100 ng/mL TGFβ1 (R&D Systems, Cat#240-B-002) in 3F media (described above) for 3 days. On the fourth day, iBTOs were exposed to dual-SMAD inactivation via 1 µM A8301 (APExBIO, Cat#3133), 100 ng/mL NOGGIN (R&D Systems, Cat#6057), 10 µM Y-27632 (APExBIO, Cat#B1293), and 500 ng/mL FGF10 (lab purified – see supplemental methods) in serum-free basal media for 18 days (media changed every 3 – 4 days) with whole passaging as necessary.

Alveolar differentiation was carried out as previously described (Jacob *et al.*, 2017). Briefly, iBTOs were transitioned to alveolar differentiation media for 7 days (media changed on days 3 and 6). Alveolar differentiation media consists of a modified serum-free basal media (SFB-VA) with DMEM/F12 containing HEPES and L-Glutamine (Corning, Cat#10-092-CV) supplemented with 1X N-2 supplement (Thermo Fisher, Cat#17502048), 1X B-27 supplement without vitamin A (Thermo Fisher, Cat#12587010), 0.05% BSA (Sigma, Cat#A9647), 100 U/mL Penicillin-Streptomycin (Thermo Fisher, Cat#15140122), 50 µg/mL L-Ascorbic Acid (Sigma, Cat#A9647) and 0.4 µM 1-Thioglycerol (Sigma, Cat#M6145). On the first day of alveolar differentiation, SFB-VA is supplemented with 10 ng/mL FGF7 (R&D Systems,

Cat#251-KG/CF), 3 μM CHIR99021 (APExBIO, Cat#A3011), 100 μM 3-Isobutyl-1-methylxanthine (Sigma, Cat#I5879), 100 μM 8-bromoadenosine 3', 5' -cyclic monophosphate sodium salt (Sigma, Cat#B7780) and 50 nM Dexamethasone (Sigma, Cat#D4802).

iBTO Growth Rate Experiments

iBTPs in LPOs grown in 3F media for 4 – 6 weeks (early timepoints) or 10 – 11 weeks (late timepoints) were isolated by FACS (see below). After collection, iBTPs were centrifuged at 300g for 5 minutes at 4ºC and media was removed. Matrigel (Corning, Cat#354234) was added to the cells at a concentration of 5,000 cells/μL Matrigel and 8 – 15 μL droplets (droplet size was kept consistent per batch) of Matrigel were placed into the center of wells of a 24-well tissue culture plate (Thermo Fisher, Cat#12565163). The plate was inverted and placed in an incubator at 37ºC for 20 minutes. iBTPs were fed every 2 – 4 days with 3F. At timepoints 2, 4, and 6-weeks post-FACS, iBTOs from one original well of plated iBTPs were collected. iBTOs were removed from Matrigel using a P1000 and/or a P200 pipette tip and vigorously pipetted in a 1.5 mL tube to remove as much Matrigel as possible. Tissue was centrifuged at 300g for 5 minutes at 4ºC, then excess media and Matrigel was removed. Tissue was digested to single cells using 250 – 500 μL TrypLE (Invitrogen, Cat#12605010), depending on pellet size, and incubated at 37ºC for 5 – 15 minutes, adding mechanical digestion with a pipette every 5 minutes, until a single cell suspension was reached. Trypsinization was quenched with DMEM/F-12 (Corning, Cat#10-092-CV) + 10 μM Y-27632 (APExBIO, Cat#B1293). Cells were centrifuged at 300g for 5 minutes at 4ºC, then liquid was removed. Cells were resuspended in 100 – 500 μL 1X PBS + 10 μM Y-27632 and counted using a hemocytometer.

Organoid Forming Efficiency Assays

Following dissociation for iBTO growth rate experiments described above, at the 4-week timepoint, iBTPs were centrifuged at 300g for 5 minutes at 4ºC, then media was removed. Matrigel (Corning, Cat#354234) was added to the cells at a concentration of 1,000 cells/µL for late timepoints or 2,5000 cells/µL for early timepoints (organoids did not form at 1,000 cells/µL for early timepoints), and 20 µL droplets of Matrigel were placed into the center of wells of a 24-well tissue culture plate (Thermo Fisher, Cat#12565163). The plate was inverted and placed in an incubator at 37ºC for 20 minutes. iBTPs were fed every 2 – 4 days with 3F. Two weeks later, cultures were imaged for counting number of cysts. Counting was carried out using the count tool in Adobe Photoshop CC 2022.

Expression and Purification of Human Recombinant FGF10

The recombinant human FGF10 (rhFGF10) expression plasmid pET21d-FGF10 was a gift from James A. Bassuk (Bagai *et al.*, 2002) at the University of Washington School of Medicine. This plasmid was transformed to Novagen's Rosetta™ 2(DE3)pLysS competent cells (Millipore Sigma, Cat#71403-3) for rhFGF10 expression. In brief, *E.coli* strain Rosetta™ 2(DE3)pLysS bearing pET21d-FGF10 was grown in 2x YT medium (BD Biosciences, Cat#244020) with Carbenicillin (50 µg/mL) and Chloramphenicol (17 µg/mL). rhFGF10 expression was induced by addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG). rhFGF10 was purified by using a HiTrap-Heparin HP column (GE Healthcare, Cat#17040601) with step gradients of 0.2 M to 0.92 M NaCl. From a 200 mL culture, 3 – 4 mg of at least 98% pure rhFGF-10 (evaluated by SDS-PAGE stained with Coomassie Blue R-250) was purified. In-house purified rhFGF10 was confirmed by western blot analysis using anti-FGF10 antibody and compared to commercially purchased rhFGF10 (R&D Systems, Cat#345-FG) to test/validate

activity based on the efficiency to phosphorylate ERK1/2 in an A549 alveolar epithelial cell line (ATCC, Cat#CCL-185) as assessed by western blot analysis.

Expression and Purification of Human Recombinant FGF4

E.coli strain Rosetta^{TM}(DE3)pLysS bearing pET28/MhFGF4L was a gift from Masayuki Kobayashi (Sugawara *et al.*, 2014) at Akita Prefectural University, Japan. Culture of the *E.coli* strain was in 2X YT medium with Kanamycin (50 µg/mL) and Chloramphenicol (17 µg/mL). Expression and purification of rhFGF4 was performed as described above for rhFGF10. In-house purified rhFGF4 activity was compared to commercially purchased rhFGF4 (R&D Systems, Cat#235-F4) based on the efficiency to differentiate definitive endoderm cells from hPSCs into $CDX2^{+}$ hindgut cells by qRT-PCR and immunofluorescence (McCracken *et al.*, 2011).

Expression and Purification of Human Recombinant R-Spondin1

The stable 293T cell line expressing HA-R-Spondin1-Fc (Cultrex, Cat#3710-001-01) was cultured and passaged according to the manufacturer's protocol. For expression and purification, after at least five days in selection growth medium (0.3 mg/mL Zeocine), the cells were passaged to Triple Flasks (Nunc, Cat#132913) with 150-200 mL growth medium per flask without selection. The cells were cultured until they reached 90% confluency (~5 days), then medium was changed to CD293 medium (Gibco, Cat#11913-019) containing L-glutamine and Pen/Strep. Confluent cells were maintained in CD293 medium for 7 – 14 days, then the supernatant was collected and filtered for HA-R-Spondin1-Fc purification. The activity of the conditioned medium was determined by performing a TopFlash Luciferase Assay (Biotium, Cat#30085-2) by its ability to induce activation of Wnt/β-catenin. HA-R-Spondin1-Fc was purified using the Fc tag via protein A agarose purification kit (KPL, Cat#553-50-00) per manufacturer's instructions. From 1 L of conditioned medium, ~1.5 mg of at least 95% pure HA-R-Spondin1-Fc was

purified. In-house purified HA-R-Spondin1-Fc activity was compared to commercially purchased rhR-Spondin1 (R&D Systems, Cat#4645-RS) based on the efficiency to differentiate $CDX2^+$ hindgut spheroids from hPSCs to human intestinal organoids (HIO) by qRT-PCR and HIO morphology (McCracken *et al.*, 2011).

3.4.3 Method Details

scRNA-seq: Tissue Processing

All tubes and pipette tips were pre-washed in 1X HBSS with 1% BSA to prevent cell adhesion to the plastic. 3D cultures (spheroids, LPOs, iBTOs) were removed from Matrigel using a P1000 pipette tip and vigorously pipetted in a 1.5 mL Eppendorf tube to remove as much Matrigel as possible. Tissue was centrifuged at 300g for 3 minutes at 4ºC, then excess media and Matrigel was removed. Tissue was digested to single cells using 0.5 mL TrypLE (Invitrogen, Cat#12605010) and incubated at 37ºC for 30 minutes, adding mechanical digestion with pipette every 10 minutes. After 30 minutes, trypsinization was quenched with 1X HBSS + 1% BSA. Cells were passed through a 40 µm filter (Bel-Art Flowmi, Cat#136800040) and centrifuged at 300g for 3 minutes at 4ºC. Cells were resuspended in 1 mL 1X HBSS + 1% BSA and counted using a hemocytometer, centrifuged at 300g for 3 minutes at 4ºC, and resuspended to a final concentration of 1,100 cells/µL. If samples were planned for combined submission, cells would be cryopreserved in ice-cold CryoStor CS10 (Biolife Solutions, Cat#210102). If cryopreservation was used, cold CryoStor CS10 solution was added to the cells and mixed thoroughly and transferred to a cryovial. The cells were incubated at 2 – 8ºC for 10 minutes, then cryopreserved in using an isopropanol freezing container for 24 hours, then transferred to liquid nitrogen. Cells were thawed in a 37ºC water bath, washed in DMEM/F12 (Corning, Cat#10-092-CV) with 10% FBS, centrifuged at 300g for 5 minutes, and supernatant was removed. The pellet was then

washed in HBSS + 1% BSA, then resuspended in 1 mL fresh HBSS + 1% BSA, passed through a 40 μm filter (Bel-Art Flowmi, Cat#136800040), and counted using a hemocytometer. Cells were centrifuged at 300g for 3 minutes at 4ºC and resuspended to a final concentration of 1,100 cells/μL. Approximately 100,000 fresh or thawed cells were put on ice and single cell libraries were immediately prepared at the 10X Chromium at the University of Michigan Sequencing Core with a target of 10,000 cells per sample.

scRNA-seq: Analysis Overview

To visualize distinct cell populations within the single-cell RNA sequencing dataset, we employed the recommended workflow outlined by the Seurat 4.0 R package (Hao *et al.*, 2021). This pipeline includes the following steps: filtering cells for quality control by applying the SCTransform technique (Hafemeister and Satija, 2019) in place of traditional log normalization, variable gene selection, and scaling, identifying anchors and integrating if multiple single-cell RNA samples are involved (Stuart *et al.*, 2019b), reducing dimensionality with principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) (McInnes, Healy and Melville, 2018; Becht *et al.*, 2019), clustering by either the Louvain algorithm (Blondel *et al.*, 2008) or the Leiden algorithm (Traag, Waltman and van Eck, 2019); and log normalization on RNA assay for final visualization and for differential gene expression analysis.

scRNA-seq: Sequencing Data and Processing FASTQ Reads into Gene Expression Matrices

All single-cell RNA sequencing was performed at the University of Michigan Advanced Genomics Core with an Illumina Novaseq 6000. The 10X Genomics Cell Ranger pipeline was used to process raw Illumina base calls (BCLs) into gene expression matrices. BCL files were demultiplexed to trim adaptor sequences and unique molecular identifiers (UMIs) from reads.

Each sample was then aligned to the human reference genome (hg19) to create a filtered feature bar code matrix that contains only the detectable genes for each sample.

scRNA-seq: Quality Control

To ensure quality of the data, all samples were filtered to remove cells expressing too few or too many genes (Figure 3-1G/Figure 3-2F – G/Figure 3-8E – G - <500, >8000; Figure 3-3G/Figures 3-4E – H/Figure 3-5A - <500, >7000; Figure 3-7A – D/Figure 3-8B – C - <200, >9500), with too low or too high UMI counts (Figure 3-1G/Figure 3-2F – G – <500, >50000; Figure 3-3G/Figure 3-4E – H/Figure 3-5A – <500, >50000; Figure 3-7A – D/Figure 3-8B – C - <200, >50000; Figure 3-8E – G - <500, >60000), or a fraction of mitochondrial genes greater than 0.1. Following the above steps, a total of (Figure 3-1G/Figure 3-2F – G – 1067 cells, 36601 genes; Figure 3-3G/Figure 3-4E – H/Figure 3-5A – 9133 cells, 36601 genes; Figure 3-7A – D/Figure 3-8B – C – 4334 cells, 36602 genes; Figure 3-8E – G – 10888 cells, 34738 genes) were kept for downstream analysis and visualization.

scRNA-seq: ScTransform and Integration

Seurat's SCTransform method allows efficient pre-processing, normalization, and variance stabilization of molecular count data from scRNA-seq samples. Running this algorithm will reveal a model of technical noise in the scRNA-seq data through "regularized negative binomial regression", whose residuals are returned as the SCTransform-normalized values that can be used for further downstream analysis such as dimension reduction. During the SCTransform process, we also chose to regress out a confounding source of variation – mitochondrial mapping percentage. When dealing with one sample (Figure 3-1G/Figure 3-2F – G), there's no batch effect. But when multiple samples are present (Figure 3-3G/Figure 3-4E – H/Figure 3-5A/Figure 3-7A – D/Figure 3-8B – C/Figure 3-84E – G), we have noticed certain

amount of batch effects when clustering data due to technical artifacts such as timing of data acquisition or differences in dissociation protocol. To mitigate these effects, we chose to follow Seurat's integration workflow due to its optimal efficiency in harmonizing large datasets. The two methods used are integration on SCTransform-normalized datasets (Figure 3-3G/Figure 3-4E – H/Figure 3-5A/Figure 3-7A – D/Figure 3-8B – C and integration on Log-normalized datasets (Figure 3-8E – G). After completion of such batch correction, cell clustering should no longer be driven by technical artifacts.

scRNA-seq: Dimension Reduction and Clustering

Principal component analysis (PCA) was conducted on the corrected expression matrix as follows. Using the top principal components, a neighborhood graph was calculated for the 20 nearest neighbors (Figure 3-1G/Figure 3-2F – G – 20 principal components) or the 30 nearest neighbors (Figure 3-3G/Figure 3-4E – H/Figure 3-5A/ Figure 3-7A – D/Figure 3-8B – C/Figure 3-8E – G – 30 principal components). The UMAP algorithm was then applied for visualization on 2 dimensions. Using the Leiden algorithm, clusters were identified with a resolution of (Figure 3-1G/Figure 3-2G/Figure 3-7A – D/Figure 3-8B – C/Figure 3-8E – G – 0.2). Using the Louvain algorithm, clusters were identified with a resolution of (Figure 3-3G/Figure 3-4E – H/Figure 3-A – 0.08).

scRNA-seq: Cluster Annotation

Using canonically expressed gene markers, each cluster's general cell identity was annotated. Cell identities (with markers) include epithelial (*EPCAM, KRT18, KRT8, CLDN6*), mesenchymal (*POSTN, DCN, COL1A2, COL3A1*), neuronal (*S100B, STMN2, ELAVL4, ASCL1*), endothelial (*ESAM, CDH5, CLDN5, KDR*), proliferative (*MKI67, TOP2A, CDK1*), primordial

germ cell (*POU5F1, NANOG, TBXT, NANOS3, TFAP2C*), and foregut mesodermal (*ISL1, HAND1, BMP4, FOXF1, LEF1*).

scRNA-seq: Cell Scoring

Gene lists for cell scoring (Figure 3-7E, 3-8D) are published in Hein and Conchola *et al.* 2022 (Renee F. C. Hein *et al.*, 2022), and application of cell scoring strategy is as previously described (Holloway, Wu, *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). Briefly, cells were scored based on expression of a set of 100 marker genes per cell type. Gene lists were compiled by analyzing previously-published data from human fetal lung (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). Clusters were first identified by major cell classes (epithelium, mesenchyme, neuronal, endothelium, immune), then the epithelium was sub-clustered to identify bud tip progenitor, basal, and neuroendocrine cell clusters by visualizing canonical marker gene expression for each respective cell type. In the case of bud tip progenitor and basal cells, clusters were again sub-clustered to identify the clusters with enriched bud tip progenitor or basal cell marker expression, respectively. Setting the rest of the epithelial cell population as the comparison group, the top 100 differentially expressed genes from bud tip progenitor, basal or neuroendocrine cell clusters were defined as the gene sets for cell scoring. After obtaining the scaled expression values for the data set, scores for each cell were calculated with the AddModuleScore function of Seurat. Cell scores were visualized by violin plots or feature plots.

scRNA-seq: Normalization for Visualization and Differential Gene Expression

As recommended by Seurat developers, we employed the method of log normalization on the standard RNA assay for graphing dot plots, feature plots, and conducting DGEs. Expression matrix read counts per cell were normalized by the total expression, multiplied by a scale factor of 10000, and finally log-transformed. For the differential gene expression testing, we only tested

features that are first, detected in a minimum fraction of 0.25 in either of the two cell populations, and second, show at least 0.25-fold difference in log-scale between the two cell populations on average.

scRNA-seq: Quantification of KI67$^+$ Cells

First, the cells in our LPO and iBTO data were grouped by their sample origin or cluster assignment. Then within each group, we established the frequency table of *KI67* expression based on individual cell's *KI67* RNA count, and all cells expressing > 0 were quantified as *KI67*$^+$. To calculate each cluster's contribution to *KI67*$^+$ cells in the data, the number of *KI67*$^+$ cells in each cluster was divided by the total number of cells in the given cluster ('normalized value') and then calculated as a percentage of the total normalized values of all the clusters. To calculate each sample's contribution to *KI67*$^+$ cells per cluster, a 'normalization ratio' for each cluster per sample was created by dividing the number of cells each sample contributed to each cluster by the total number of cells in the given sample. The normalization ratio was then used to recalculate a 'normalized' number of *KI67*$^+$ cells in each cluster by multiplying the ratio by the number of *KI67*$^+$ cells in each cluster by sample. The percentage of each's sample's contribution to a given cluster's *KI67*$^+$ cells was then calculated by dividing the number of normalized *KI67*$^+$ cells from each sample in each cluster by the total number of normalized *KI67*$^+$ cells in the given cluster.

scRNA-seq: Label Transfer and UMAP Projection

Seurat's single-cell reference mapping pipeline allows efficient projection between scRNA-seq datasets (Stuart *et al.*, 2019a). The Log-normalized Integration-prepared primary fetal lung epithelial data (Figure 3-8E) was set as reference while the Log-normalized PCA-prepared iBTOs samples being the query data in our analysis. Subsequently, a set of anchors

between the reference and query objects were obtained based on PCA projection and were later utilized to transfer the reference UMAP structure and cluster annotation. Finally, visualization of query data was achieved based on its predicted UMAP representation and cluster labels (Figure 3-8F – G).

Tissue Processing for IF and FISH

Tissue was immediately fixed in 10% Neutral Buffered Formalin (NBF) for 24 hours at room temperature on a rocker. Tissue was then washed 3x in UltraPure DNase/RNase-Free Distilled Water (Thermo Fisher, Cat#10977015) for 15 minutes each and then dehydrated in an alcohol series of concentrations diluted in UltraPure DNase/RNase-Free Distilled Water for 1 hour per solution: 25% Methanol, 50% Methanol, 75% Methanol, 100% Methanol, 100% Ethanol, 70% Ethanol. Dehydrated tissue was then processed into paraffin blocks in an automated tissue processor (Leica ASP300) with 1 hour solution changes. Sections at 4 (FISH) or 7 (IF) µm-thick were cut from paraffin blocks onto charged glass slides (within one week of performing FISH). For FISH, microtome and slides were sprayed with RNase Away (Thermo Fisher, Cat#700511) prior to sectioning. Slides were baked for 1 hour in 60ºC dry oven (within 24 hours of performing FISH). Slides were stored at room temperature in a slide box containing a silica desiccator packet and the seams sealed with parafilm wrap.

IF Protein Staining on Paraffin Sections

Tissue slides were rehydrated in Histo-Clear II (National Diagnostics, Cat#HS-202) 2x for 5 minutes each, followed by serial rinses through the following solutions 2x for 2 – 3 minutes each: 100% EtOH, 95% EtOH, 70% EtOH, 30% EtOH, and finally in double-distilled water (ddH2O) 2x for 5 minutes each. Antigen retrieval was performed by steaming slides in 1X Sodium Citrate Buffer (100 mM trisodium citrate (Sigma, Cat#S1804), 0.5% Tween 20 (Thermo

Fisher, Cat#BP337), pH 6.0) for 20 minutes, followed by cooling and washing quickly 2x in ddH2O and 2x in 1X PBS. Slides were incubated in a humidified chamber at RT for 1 hour with blocking solution (5% normal donkey serum (Sigma, Cat#D9663) in PBS with 0.1% Tween 20). Slides were then incubated in primary antibody diluted in blocking solution at 4ºC overnight in a humidified chamber. Next, slides were washed 3x in 1X PBS for 5 minutes each and incubated with secondary antibody with DAPI (1 µg/mL) diluted in blocking solution for 1 hour at RT in a humidified chamber. Slides were then washed 3x in 1X PBS for 5 minutes each and mounted with ProLong Gold (Thermo Fisher, Cat#P36930) and imaged within 2 weeks. Stained slides were stored in the dark at 4ºC. All primary antibody concentrations are listed in Table 3-1. Secondary antibodies, raised in donkey, were purchased from Jackson Immuno and used at a dilution of 1:500.

FISH

The FISH protocol was performed according to the manufacturer's instructions (ACDbio, RNAscope multiplex fluorescent manual) with a 5-minute protease treatment and 15-minute antigen retrieval. For IF co-staining with antibodies, the last step of the FISH protocol was skipped and instead the slides were washed 1x in PBS followed by the IF protocol above, starting at the blocking step. A list of probes and reagents can be found in Table 3-1.

Whole Mount IF Protein Staining

All tips and tubes were coated with 1% BSA in PBS to prevent tissue sticking. 3D cultures (spheroids, LPOs) were dislodged from Matrigel using a P1000 cut tip and transferred to a 1.5 mL Eppendorf tube. 500 µL of Cell Recovery Solution (Corning, Cat#354253) was added to the tube, and the tube was placed on a rocker at 4°C for 45 minutes to completely dissolve Matrigel. Tube was spun at 100g for 5 minutes, and solution and remaining Matrigel was then

removed. Tissue was fixed in 10% NBF overnight at RT on a rocker. Tissue then was washed 3x for 2 hours with 1 mL room temperature Organoid Wash Buffer (OWB) (0.1% Triton, 0.2% BSA in 1X PBS), at RT on a rocker. Wash times vary (30 minutes – 2 hours) depending on tissue size. 1 mL CUBIC-L (TCI Chemicals, Cat#T3740) was added to the tube, and the tube was placed on a rocker for 24 hours at 37°C. Tissue was then permeabilized for 24 hours at 4°C on a rocker with 1 mL permeabilization solution (5% Normal Donkey Serum, 0.5% Triton in 1X PBS). After 24 hours, permeabilization solution was removed and 500 µL primary antibody (diluted in OWB) was added overnight at 4°C on a rocker. The next day, tissue was washed 3x with 1 mL of OWB, 2 hours each at RT. 500 µL of secondary antibody (diluted in OWB) was added and incubated overnight at 4°C, wrapped in foil. Tissue was washed again 3x with 1 mL OWB at RT, first wash for 2 hours, then 30 minutes for the remaining washes. Samples were transferred to imaging plate (ThermoFisher Cat#12-566-70) and then cleared and mounted with 50 µL CUBIC-R (just enough to cover tissue) (TCI Chemicals Cat#T3741).

RNA Extraction and qRT-PCR

Each analysis includes three biological replicates from three separate differentiation attempts as well as three technical replicates. mRNA was isolated using the MagMAX-96 Total RNA Isolation Kit (Thermo Fisher, Cat#AM1830) (airway and alveolar differentiations) or the PicoPure RNA Isolation Kit (Thermo Fisher, Cat#KIT0204) (spheroids). RNA quality and yield was measured on a Nanodrop 2000 spectrophotometer just prior to cDNA synthesis. cDNA synthesis was performed using 100ng RNA per sample with the SuperScript VILO cDNA Kit (Thermo Fisher, Cat#11754250). qRT-PCR was performed on a Step One Plus Real-Time PCR System (Thermo Fisher, Cat#42765592R) using QuantiTect SYBR Green PCR Kit (Qiagen, Cat#204145). Primer sequences can be found in Table 3-1. Gene expression as a measure of

arbitrary units was calculated relative to GAPDH using the following equation:

$$2^{(Housekeeping_{Ct} - Gene_{Ct})} \times 10{,}000$$

hPSC Flow Cytometry

1 mL accutase (Sigma, Cat#A6964) was added to each well of hPSC cultures in a 6-well plate and was incubated at 37ºC for 5 – 10 minutes, until cells detached. An equal volume of mTeSR Plus media (StemCell Technologies, Cat#100-0276) with 10 µM Y-27632 (APExBIO, Cat#B1293) was added to cells and cells were dissociated mechanically by pipetting with a P1000 pipette 2x, were centrifuged at 300g for 5 minutes at 4ºC, excess media was removed, and cells were resuspended in FACS Buffer (1X PBS containing 2% BSA, 10 µM Y-27632, 100 U/mL penicillin-streptomycin). Cells were passed through a 70 µm cell strainer, pre-coated with FACS buffer, and centrifuged at 300g for 5 minutes at 4ºC. Cells were resuspended in 1 mL FACS Buffer and transferred to 5 mL FACS tubes (Corning, Cat#352063). 0.2 µg/mL DAPI was added to respective tubes. Flow cytometry was performed using a Bio Rad Ze5#3 and accompanying software.

3D Culture Sorting (FACS)

3D cultures (LPOs, iBTOs, airway organoids) were removed from Matrigel using a P1000 pipette tip and vigorously pipetted in a 15 mL conical tube to remove as much Matrigel as possible. Tissue was centrifuged at 300g for 3 minutes at 4ºC, then excess media and Matrigel was removed. Tissue was digested to single cells using 2 – 4 mL TrypLE (Invitrogen, Cat#12605010), depending on pellet size, and incubated at 37ºC for 30 minutes, adding mechanical digestion with a pipette every 10 minutes. After 15 minutes, DNase I (Qiagen, Cat#79254) was added to the digestion at 7.5 µL/mL TrypLE. After 30 minutes, trypsinization was quenched with DMEM/F-12 (Corning, Cat#10-092-CV) + 10 µM Y-27632 (APExBIO,

Cat#B1293). Cells were passed through a 70 μm cell strainer, pre-coated with DMEM/F-12 +10 μM Y-27632 and centrifuged at 500g for 5 minutes at 4ºC. Cells were resuspended in 4 mL FACS Buffer (2% BSA, 10 μM Y-27632, 100 U/mL penicillin-streptomycin) and transferred into 5 mL FACS tubes (Corning, Cat#352063). Cells were centrifuged again at 300g for 3 minutes at 4ºC, then resuspended in 1 mL FACS buffer and counted. 10^5 cells were placed into new FACS tubes for all controls (no antibody, DAPI only, individual antibodies/fluorophores) and all remaining cells were centrifuged and resuspended in FACS buffer for a concentration of 10^6 cells/100μL. Primary antibodies were incubated for 30 minutes on ice. 3 mL FACS buffer was added to each tube after 30 minutes and tubes were centrifuged at 300g for 3 minutes at 4ºC. Cells were washed again with 3 mL FACS buffer and centrifuged at 300g for 3 minutes at 4ºC. Secondary antibodies were incubated for 30 minutes on ice. 3 mL FACS buffer was added to each tube after 30 minutes and tubes were centrifuged at 300g for 3 minutes at 4ºC. Cells were washed again with 3 mL FACS buffer and centrifuged at 300g for 3 minutes at 4ºC. Cells were resuspended in FACS buffer and 0.2 μg/mL DAPI was added to respective tubes. FACS was performed using a Sony MA900 cell sorter and accompanying software. Cells were collected in 1 mL 3F media + 10 μM Y-27632. All primary antibody concentrations are listed in Table 3-1. Secondary antibodies were purchased from Jackson Immuno and used at a dilution of 1:500.

3.5 Acknowledgements

3.5.1 Financial Support

JRS is supported by the Cystic Fibrosis Foundation Epithelial Stem Cell Consortium, by grant CZFF2019-002440 from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation, and by the National Heart, Lung, and Blood Institute (NHLBI; R01HL119215). RFCH was supported by a National Institutes of Health Tissue

Engineering and Regenerative Medicine Training Grant (NIH-NIDCR T32DE007057) and by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHLBI F31HL152531). ASC was supported by the T32 Michigan Medical Scientist Training Program (5T32GM007863-40) and by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHLBI F30HL156474). TF was supported by a National Institutes of Health Tissue Engineering and Regenerative Medicine Training Grant (NIH-NIDCR T32DE007057). EMH was supported by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award (NIH-NHBLI F31HL146162).

3.5.2 Other Acknowledgements

We would like to thank the University of Michigan Advanced Genomics core for the operation of the 10X Chromium single cell capture platform, the University of Michigan Microscopy core for providing access to confocal microscopes and the Flow Cytometry core for providing access to flow cytometers.

3.6 Chapter 3 Tables

Table 3-1: Antibody dilutions and primer sequences.

Antibody	**Dilution**
FOXA2 (Seven Hills Bioreagents, Cat#WRAB-1200)	1:500
SOX17 (R&D Systems, Cat#AF1924)	1:500
VIM (R&D Systems, Cat#AFf2105)	1:500
TTF1/NKX2.1 (Abcam, Cat#ab76013)	1:500
SOX9 (Millipore, Cat#AB5535)	1:500
CPM (Wako/FujiFilm, Cat#014-27501)	1:500 for IF; 1:300 for FACS
CDX2 (BioGenex, Cat#MU392A-UC)	1:500
CDX2 (Cell Signaling Technology, Cat#12306) (Used only for co-stain with MUC2)	1:50
MUC2 (Santa Cruz Biotechnology, Cat#sc-515032)	1:300
Cleaved Caspase 3 (CCAS3) (Cell Signaling Technology, Cat#9664)	1:500
ECAD (R&D, Cat#AF748)	1:500
ECAD (BD Biosciences, Cat#610181)	1:500
SOX2 (R&D Systems, Cat#AF2018)	1:500
TP63 (R&D Systems, Cat#BAF1916)	1:500
CHGA (Santa Cruz Biotechnology, Cat#sc-1488)	1:100
FOXJ1 (Seven Hills Bioreagents, Cat#WMAB-319)	1:250

MUC5AC (Abcam, Cat#ab79082)	1:500
CC10/SCGB1A1 (Santa Cruz Biotechnology, Cat#sc-365992)	1:250
Acetylated Tubulin (AcTUB) (Sigma-Aldrich, Cat#T7451)	1:1000
ASCL1 (Santa Cruz Biotechnology, Cat#sc-374104)	1:50
Synaptophysin (SYN) (Abcam, Cat#ab32127)	1:500
SP-C (Santa Cruz Biotechnology, Cat#sc-518029)	1:200
Pro-SP-C (Seven Hills Bioreagents)	1:500
SP-B (Seven Hills Bioreagents, Cat#WMAB-1B9)	1:500
Primer	**Sequence**
NKX2.1	F: CTCATGTTCATGCCGCTC R: GACACCATGAGGAACAGCG
CDX2	F: GGGCTCTCTGAGAGGCAGTT R: GGTGACGGTGGGGTTTAGCA
ECAD	F: TTGACGCCGAGAGCTACAC R: GACCGGTGCAATCTTCAAA
VIM	F: CTTCAGAGAGAGGAAGCCGA R: ATTCCACTTTGCGTTCAAGG
SOX2	F: GCTTAGCCTCGTCGATGAAC R: AACCCCAAGATGCACAACTC
TP63	F: CCACAGTACACGAACCTGGG R: CCGTTCTGAATCTGCTGGTCC
FOXJ1	F: CAACTTCTGCTACTTCCGCC R: CGAGGCACTTTGATGAAGC
CHGA	F: CTGTCCTGGCTCTTCTGCTC R: TGACCTCAACGATGCATTTC
MUC5AC	F: GCACCAACGACAGGAAGGATGAG R: CACGTTCCAGAGCCGGACAT

SCGB1A1	F: ATGAAACTCGCTGTCACCCT R: GTTTCGATGACACGCTGAAA
SCGB3A2	F: GGGGCTAAGGAAGTGTGTAAATG R: CACCAAGTGTGATAGCGCCTC
SFTPC	F: AGCAAAGAGGTCCTGATGGA R: CGATAAGAAGGCGTTTCAGG
SFTPB	F: GGGTGTGTGGGACCATGT R: CAGCACTTTAAAGGACGGTGT
ABCA3	F: TGCAGCGCCTACTTGAACTT R: CTGAGCACAGCCATCGTCT

3.7 Chapter 3 Figures

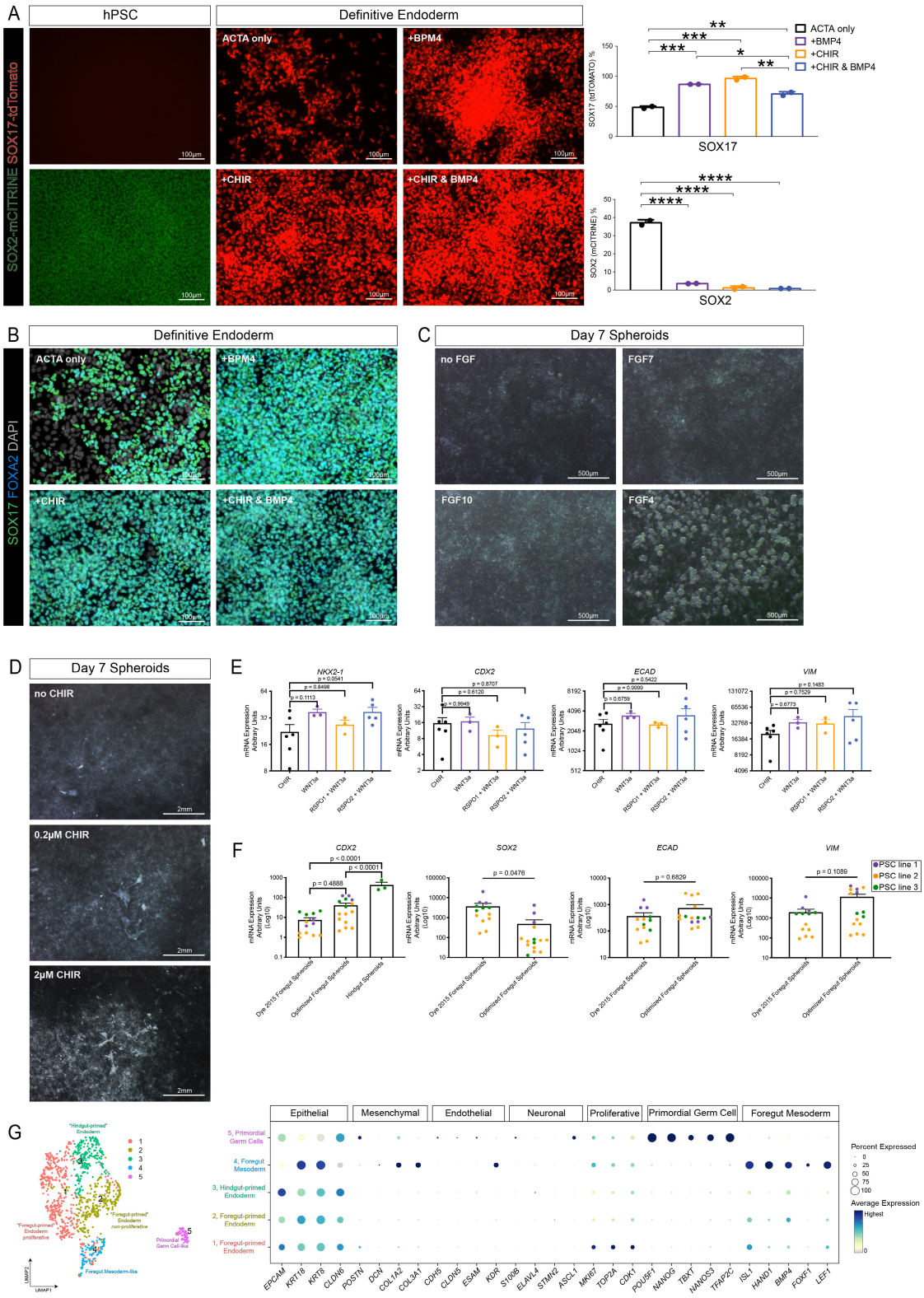


Figure 3-1: Optimization of Directed Differentiation of hPSCs at Definitive Endoderm, Foregut Spheroid and NKX2-1$^+$ Spheroid Stages, Related to Figure 3-2. (A) (Left) SOX2-mCitrine (pluripotent stem cell marker) and SOX17-tdTomato (definitive endoderm marker) reporter images of SOX2-mCITRINE and SOX17-tdTomato on hPSCs and hPSC-derived day 4 definitive endoderm (DE) where 50ng/mL BMP4 and/or 2μM CHIR were added on day 1 of a 3-day 100ng/mL Activin A (ACTA) treatment to induce DE (see Figure 3-2A) in a hPSC line expressing SOX2-mCitrine/SOX17-tdTomato hPSC line (Martyn *et al.*, 2018). (Right) Flow cytometry quantifications of day 4 DE cultures. Two technical replicates from the same experiment were performed for flow cytometry quantifications. Statistical tests were performed by ordinary one-way ANOVA. **(B)** Immunofluorescence staining for the DE markers SOX17 and FOXA2 on day 4 DE testing conditions as noted in S1A. This experiment was performed on two independent hPSC lines different than the reporter line in Figure 3-1A. Representative images are shown from H9 ESCs. **(C)** Brightfield images of day 7 spheroids evaluated for FGF-dependent spheroid formation. Conditions include no FGF, 10ng/mL FGF7, 500ng/mL FGF10, or 500ng/mL FGF4 on days 4 – 6. **(D)** Brightfield images of day 7 spheroids evaluated for WNT-dependent spheroid formation. Conditions include no CHIR99021, 0.2μM CHIR99021, or 2μM CHIR99021 on days 4 – 6. **(E)** qRT-PCR data comparing expression of the lung epithelial marker *NKX2-1*, the hindgut epithelial marker *CDX2*, the pan-epithelial marker *ECAD*, and the pan-mesenchymal marker *VIM* when CHIR99021, WNT3a, RSPO1 & WNT3a, or RSPO2 & WNT3a was used to activate WNT signaling on days 7 – 9 of the NKX2-1-optimized spheroid directed differentiation protocol. Data points for CHIR99021 and WNT3a + RSPO2 represent 2 independent experiments with 2 – 3 technical replicates and data points from other conditions represent one independent experiment with 3 technical replicates. All experiments were performed on H9 ESCs. Statistical tests compared all conditions to CHIR99021 and were performed by ordinary one-way ANOVA followed by Dunnett's multiple comparison test. **(F)** qRT-PCR data comparing expression of the hindgut epithelial marker *CDX2*, the posterior foregut marker *SOX2*, the pan-epithelial marker *ECAD*, and pan-mesenchymal marker *VIM* from previously published foregut spheroids (Dye *et al.*, 2015) to optimized foregut spheroids, and in the case of *CDX2*, hindgut spheroids. Each color represents an independent experiment with a unique iPSC line (purple: WTC11, orange: iPSC17 WT 7B2, green: iPSC line 72.3). Each data point of the same color represents a technical replicate from the same iPSC line in one or more independent experiments. Error bars represent standard error of the mean. Statistical test for *CDX2* was performed by ordinary one-way ANOVA followed by Tukey's multiple comparison test and statistical tests for the remaining markers was performed by unpaired Welch's one-tailed t-tests. **(G)** Cluster plot of scRNA-seq data from day 10 spheroids and dot plot of epithelial, mesenchymal, endothelial, and neuronal lineage genes as well as proliferation markers in each cluster of the UMAP. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized expression level of the gene.

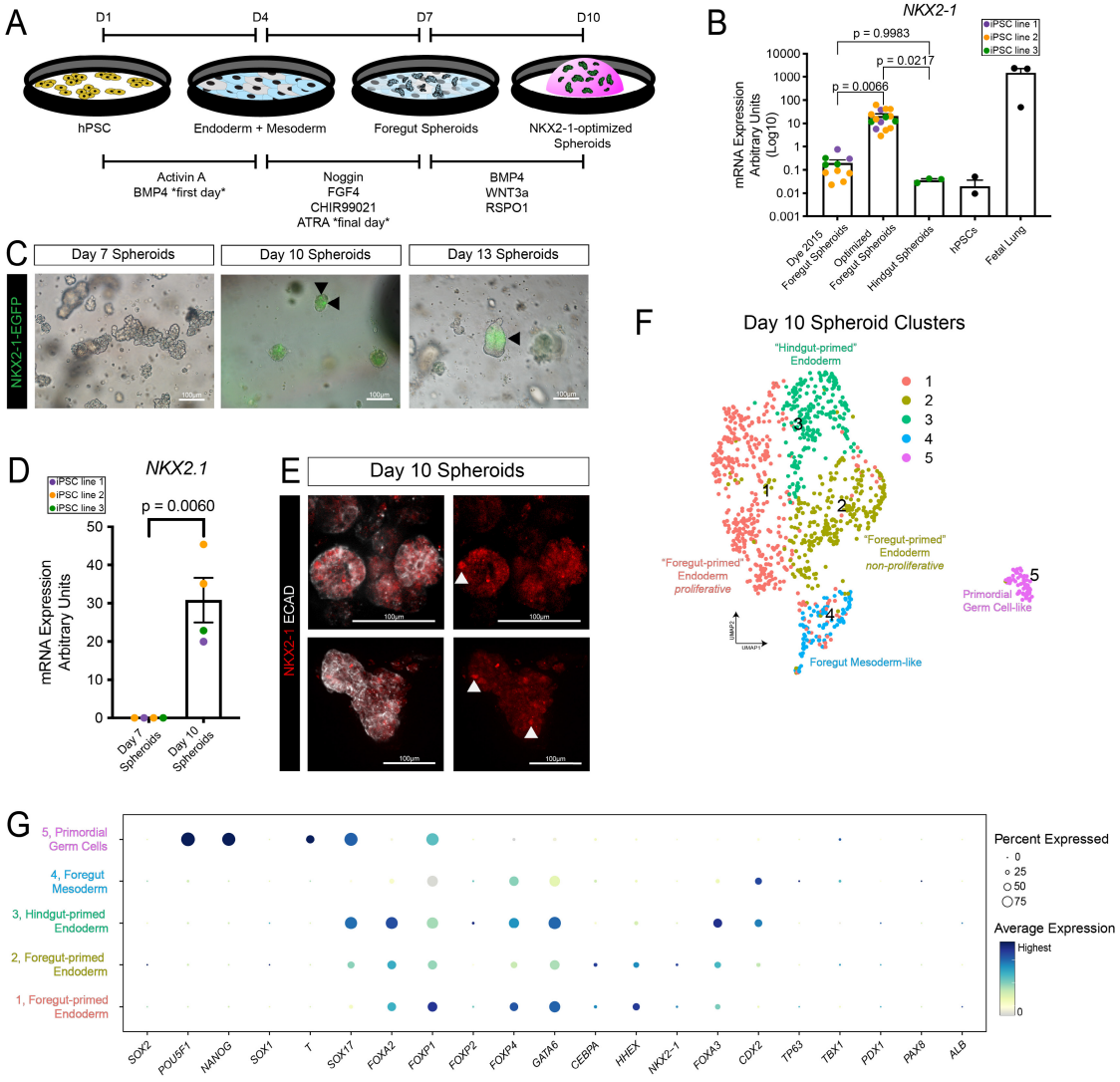


Figure 3-2: Optimization of Lung Spheroids for NKX2-1 Expression. (A) Schematic displaying the directed differentiation protocol from hPSC to NKX2-1-optimized spheroids. **(B)** qRT-PCR data comparing *NKX2-1* expression in previously published foregut spheroids (Dye *et al.*, 2015) to optimized foregut spheroids and hindgut spheroids (Spence *et al.*, 2011). hPSCs and whole fetal lung are also included as references. Each color represents an independent experiment with a unique iPSC line (purple: WTC11, orange: iPSC17 WT 7B2, green: iPSC line 72.3). Each data point of the same color represents a technical replicate from the same iPSC line in one or more independent experiments. Error bars represent standard error of the mean. Statistical tests were performed by ordinary one-way ANOVA followed by Tukey's multiple comparison test. **(C)** Representative reporter expression for NKX2-1-EGFP on day 7, 10, or 13 spheroids. **(D)** qRT-PCR data comparing *NKX2-1* expression in spheroids collected on day 7 or day 10 (see Figure 3-2A). Each color represents an independent experiment with a unique iPSC line (purple: WTC11, orange: iPSC17 WT 7B2, green: iPSC line 72.3). Error bars represent standard error of the mean. Statistical tests were performed by unpaired Welch's one-tailed t-tests. **(E)** Maximum intensity projection of a whole mount immunofluorescence confocal Z-series staining for the pan-epithelial marker ECAD and lung epithelial marker NKX2-1 on day 10 spheroids. **(F)** UMAP cluster plot of scRNA-seq data from day 10 spheroids (n = 1 batch, ~100 spheroids). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the dot plot in Figure 3-2G or Figure 3-1G. **(G)** Dot plot of cell lineage genes in each cluster of the UMAP plot in Figure 3-2F. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized expression level of the gene.

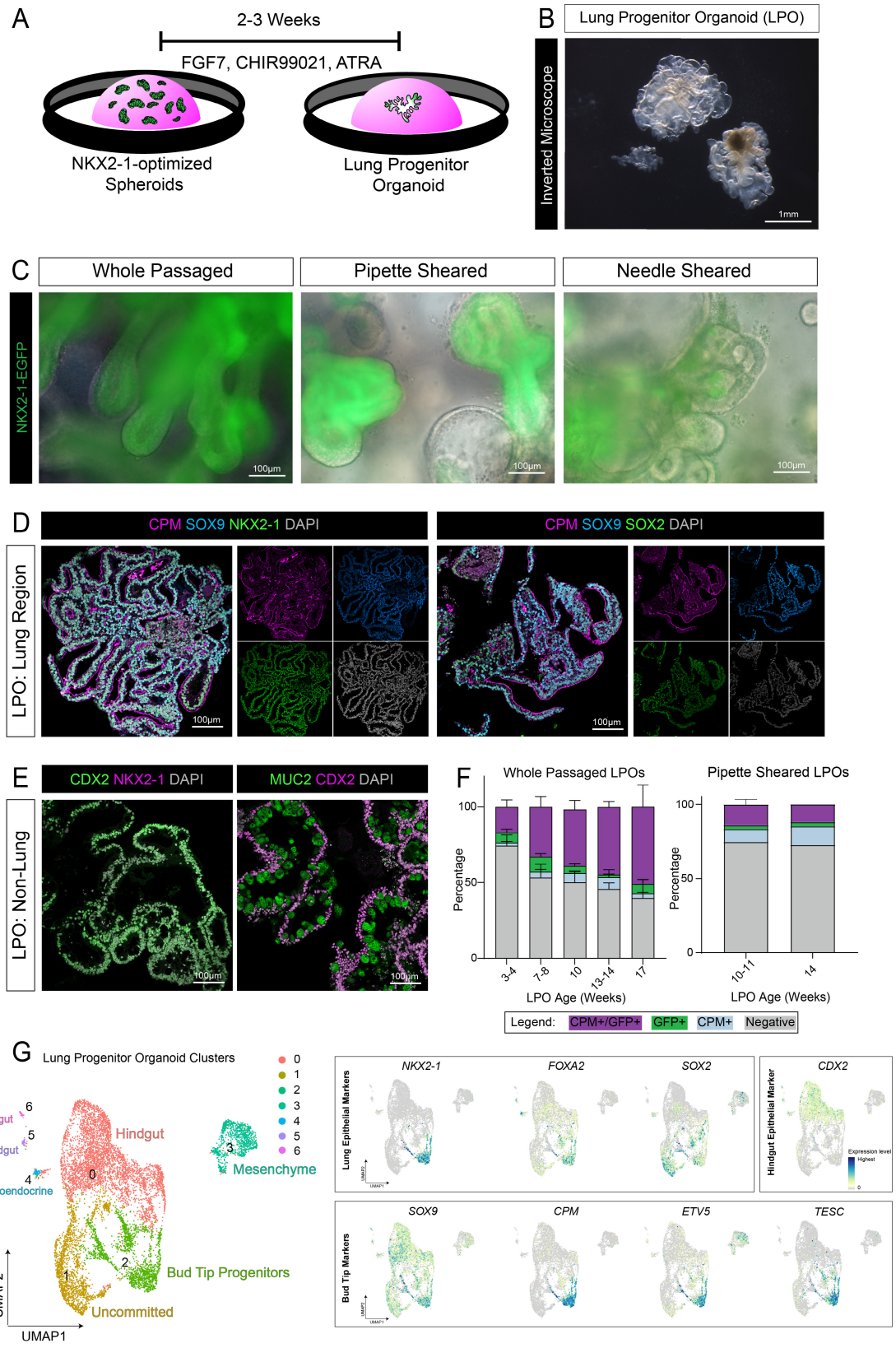


Figure 3-3: iPSC-derived Bud Tip Progenitors Emerge Over Time in LPOs. (A) Schematic displaying the Lung Progenitor Organoid (LPO) expansion protocol from NKX2-1-optimized spheroids. LPOs form after 2 – 3 weeks in culture. **(B)** Brightfield

image of 6-week LPO on inverted microscope. **(C)** Representative NKX2-1-EGFP reporter images of 10 – 11-week LPOs, passaged whole or sheared (pipette and needle). **(D)** Immunofluorescence staining on paraffin sections of the lung-like regions of 12-week LPOs for bud tip progenitor markers CPM and SOX9 and either lung epithelial markers NKX2-1 (left panels) or SOX2 (right panels). **(E)** Immunofluorescence staining on paraffin sections for intestinal epithelial marker CDX2, intestinal goblet cell marker MUC2, and lung epithelial marker NKX2-1 on non-lung regions of 12-week LPOs. **(F)** FACS quantification of NKX2-1-$EGFP^+/CPM^+$ cells in whole passaged or pipette sheared LPOs in aggregate time course (3 – 17 weeks) from the NKX2-1-EGFP reporter cell line (iPSC17 WT 7B2). Percentages of live cells expressing neither marker (Negative, grey), each separate marker (CPM^+-only, blue; $EGFP^+$-only, green), or dual-expressing cells ($CPM^+/EGFP^+$, purple) are reported as mean ± SEM for 3 – 7 replicates per time point. **(G)** UMAP cluster plot of scRNA-seq data from LPOs (n=2 biological replicates for 3- and 6-week timepoints, n=1 for 10-week timepoint). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the heat map in Figure 3-5 and the dot plot and feature plots in Figure 3-3G and Figure 3-4G. Feature plots corresponding to the LPO cluster plot and displaying canonical bud tip progenitor markers (*SOX9, CPM, ETV5, TESC*) (Miller *et al.*, 2018, 2020; Yamamoto *et al.*, 2020), lung epithelial markers (*NKX2-1, FOXA2, SOX2*), and hindgut epithelial marker (*CDX2*). The color of each dot in the feature plot indicates log-normalized expression level of the labeled gene in the represented cell.

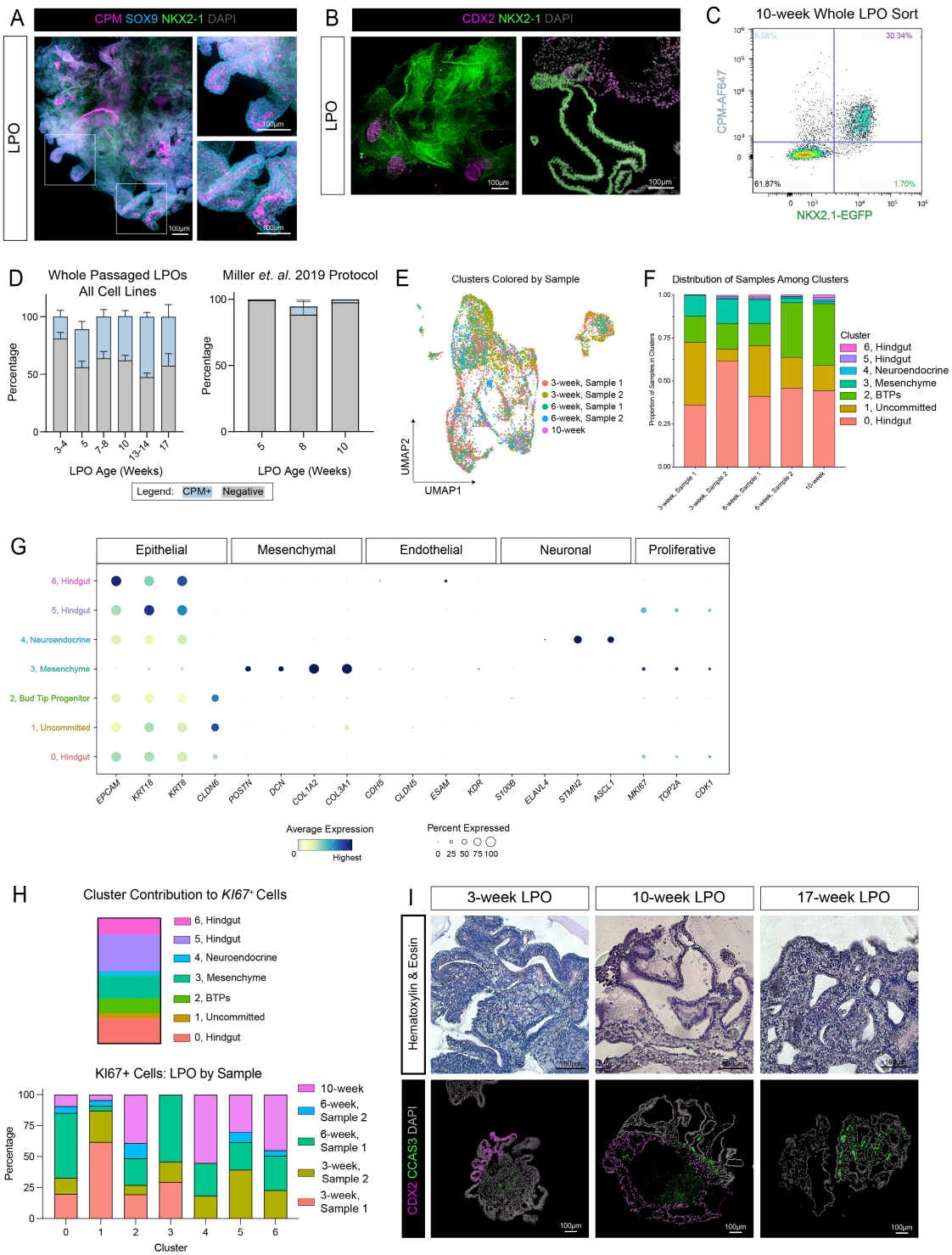


Figure 3-4: Bud Tip Progenitors and Contaminating Lineages Emerge and Expand in LPOs Over Time, Related to Figure 3-3. (A) Maximum intensity projection of a whole mount immunofluorescence confocal Z-series staining of a 10-week LPO for bud tip progenitor markers CPM and SOX9 and lung epithelial marker NKX2-1. **(B)** Maximum intensity projection of a whole mount immunofluorescence confocal Z-series staining of a 10-week LPO (left) or immunofluorescence staining on a paraffin section of 4-week LPOs (right) for intestinal epithelial marker CDX2 and lung epithelial marker NKX2-1. **(C)** Representative flow cytometry plot for iBTP selection based on positive CPM and NKX2-1-EGFP selection on 10-week LPOs. **(D)** (Left) FACS quantification of CPM^+ cells in LPOs in aggregate time course from three separate cell lines, including the NKX2-1-EGFP reporter line (iPSC17 WT 7B2), iPSC line 72.3 and WTC11. LPOs were sorted using CPM from 3 – 17 weeks in

3F media. Percentages of live cells expressing CPM (blue) or negative (grey) are reported as mean ± SEM for 3 – 9 replicates per time point. (Right) FACS quantification of CPM^+ cells in LPOs generated from spheroids using a previously published protocol (Miller *et al.*, 2019) from two separate cell lines, including the NKX2-1-EGFP reporter line and iPSC line 72.3 (WTC11-derived LPOs did not survive to 10 weeks) (Miller *et al.*, 2019). **(E)** UMAP plot corresponding to the LPO cluster plot in Figure 3-3G. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(F)** Stacked bar graph displaying the proportion of cells from each sample in each cluster of the LPO cluster plot in Figure 3-3G. **(G)** Dot plot of epithelial, mesenchymal, endothelial, and neuronal lineage genes as well as proliferation markers in each cluster of the LPO cluster plot in Figure 3-3G. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized expression level of the gene. **(H)** (Top) Stacked bar graph showing each cluster's normalized contribution to $KI67^+$ cells, calculated from the scRNA-seq data. (Bottom) Stacked bar graph showing each sample's normalized contribution to $KI67^+$ cells per cluster, calculated from the scRNA-seq data. **(I)** (Top) Hematoxylin & eosin or (Bottom) immunofluorescence staining on paraffin sections of 3-, 10- and 17-week LPOs. IF stains for cell death marker cleaved caspase 3 (CCAS3) and intestinal epithelial marker CDX2.

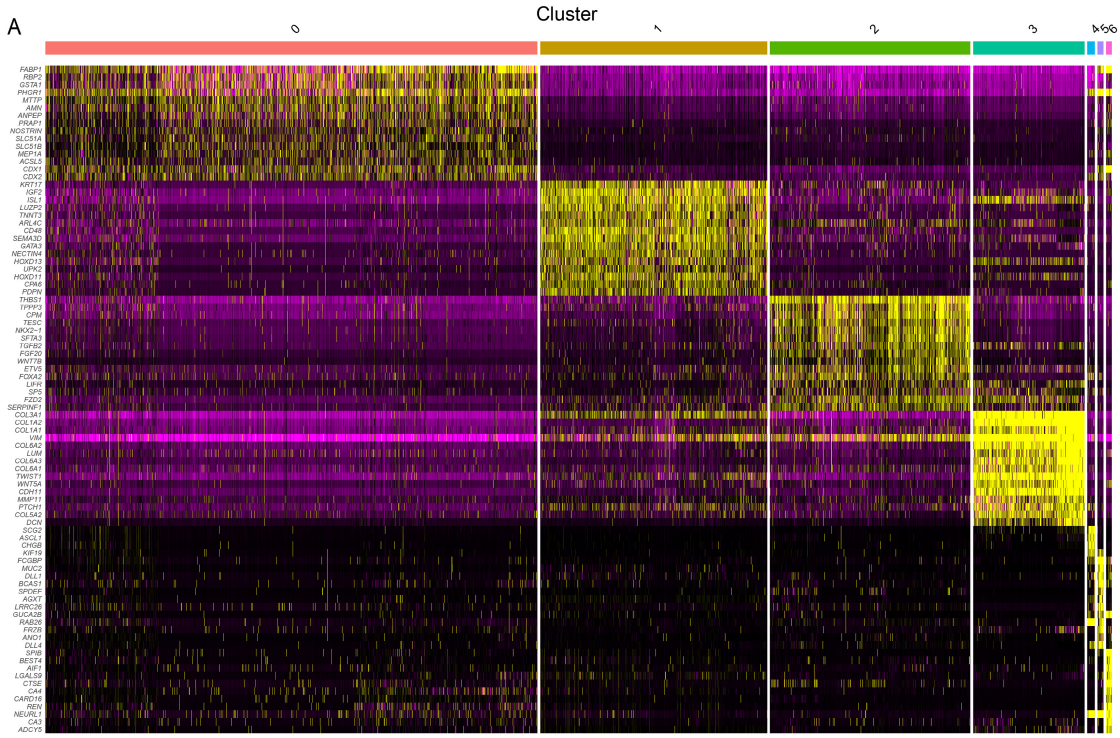


Figure 3-5: LPOs are Transcriptionally Heterogenous, Related to Figure 3-3. (A) Heat map including 12 genes per cluster picked from the top 50 highly enriched genes in each cluster corresponding to the UMAP cluster plot in Figure 3-3G. Top-50 genes were based on log-fold change of each gene.

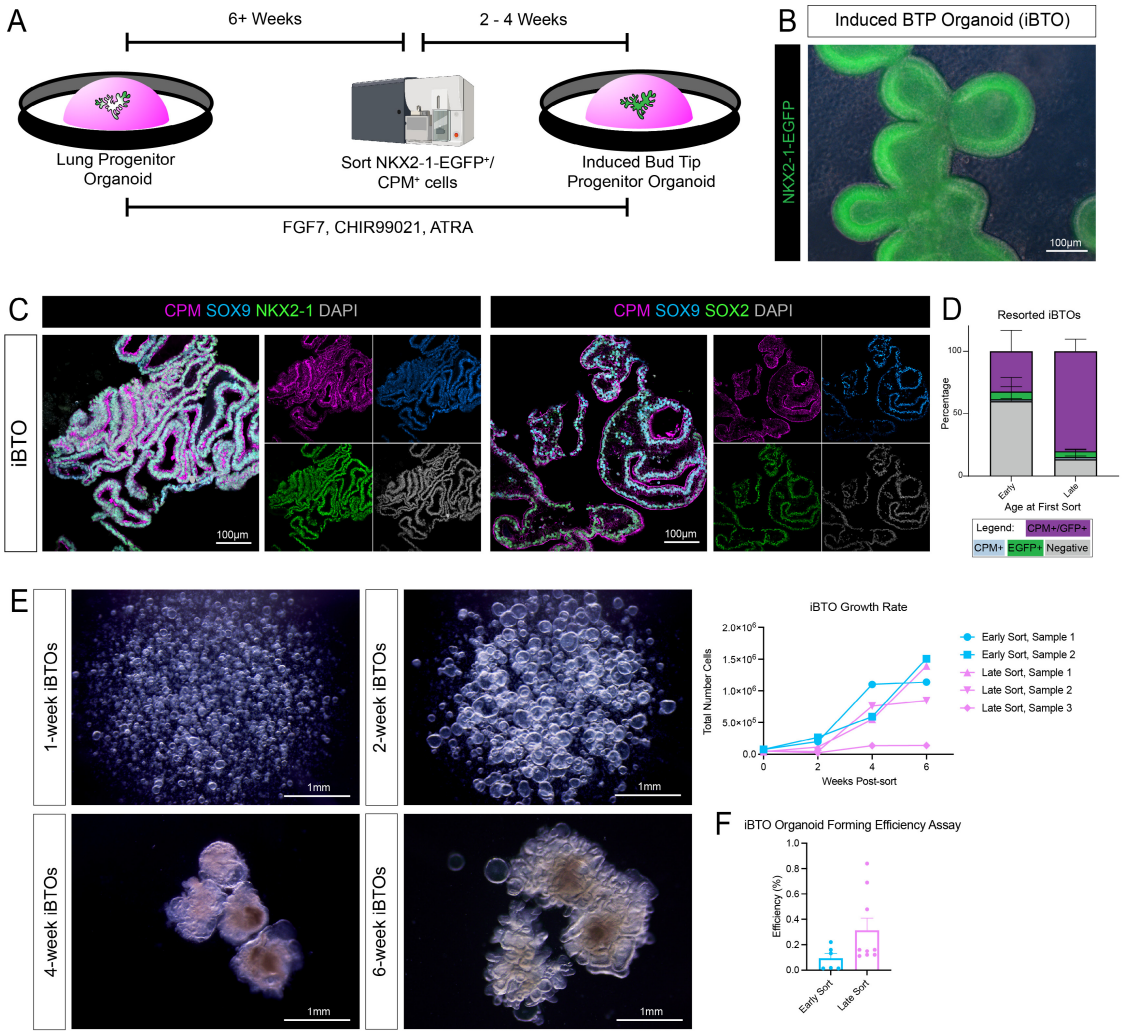


Figure 3-6: iPSC-derived Bud Tip Progenitors can be Isolated, Expanded Long-term. (A) Schematic displaying isolation and expansion of induced bud tip progenitor organoids (iBTOs). LPOs are maintained in 3F media and whole passaged for at least 6 weeks, then are dissociated for FACS. iBTPs are isolated using CPM^+ expression +/- $NKX2\text{-}1EGFP^+$ reporter expression and replated as isolated iBTPs. iBTPs re-form to iBTOs over 2 – 4 weeks, are maintained in 3F media, and whole passaged. **(B)** Representative NKX2-1-EGFP reporter image of 3-week iBTO. **(C)** Immunofluorescence staining on paraffin sections of nearly homogenous 4-week iBTOs for bud tip progenitor markers CPM and SOX9 and either lung epithelial markers NKX2-1 (left panels) or SOX2 (right panels). **(D)** FACS quantification of $NKX2\text{-}1\text{-}EGFP^+/CPM^+$ cells in iBTOs from 3-week sorted LPOs (early) or 8 – 17-week sorted LPOs (late) from the NKX2-1-EGFP reporter cell line (iPSC17 WT 7B2). iBTOs were analyzed 7 – 8 weeks after iBTP purification from LPOs. Percentages of live cells expressing neither marker (Negative, grey), each separate marker (CPM^+-only, blue; $EGFP^+$-only, green), or dual-expressing cells ($CPM^+/EGFP^+$, purple) are reported as mean ± SEM for 3 replicates per time point. **(E)** (Left) Brightfield images of iBTOs 1-, 2-, 4-, and 6-weeks post-sort from 11-week LPOs on inverted microscope. (Right) Quantification of iBTO growth from the day of sort to 6-weeks post-sort as the number of cells in organoids generated from the same number of starting cells from iBTOs sorted from 4 – 6-week LPOs (early sorts) or 10 – 11-week LPOs (late sorts). **(F)** Organoid forming efficiency assay of iBTOs sorted from 4 – 6-week LPOs (early sort) or 10 – 11-week LPOs (late sort). Organoid forming efficiency was measured as the number of cysts formed after 2 weeks from when iBTOs were digested to a single cell suspension and re-plated in Matrigel at 1,000 or 2,500 cells per µL.

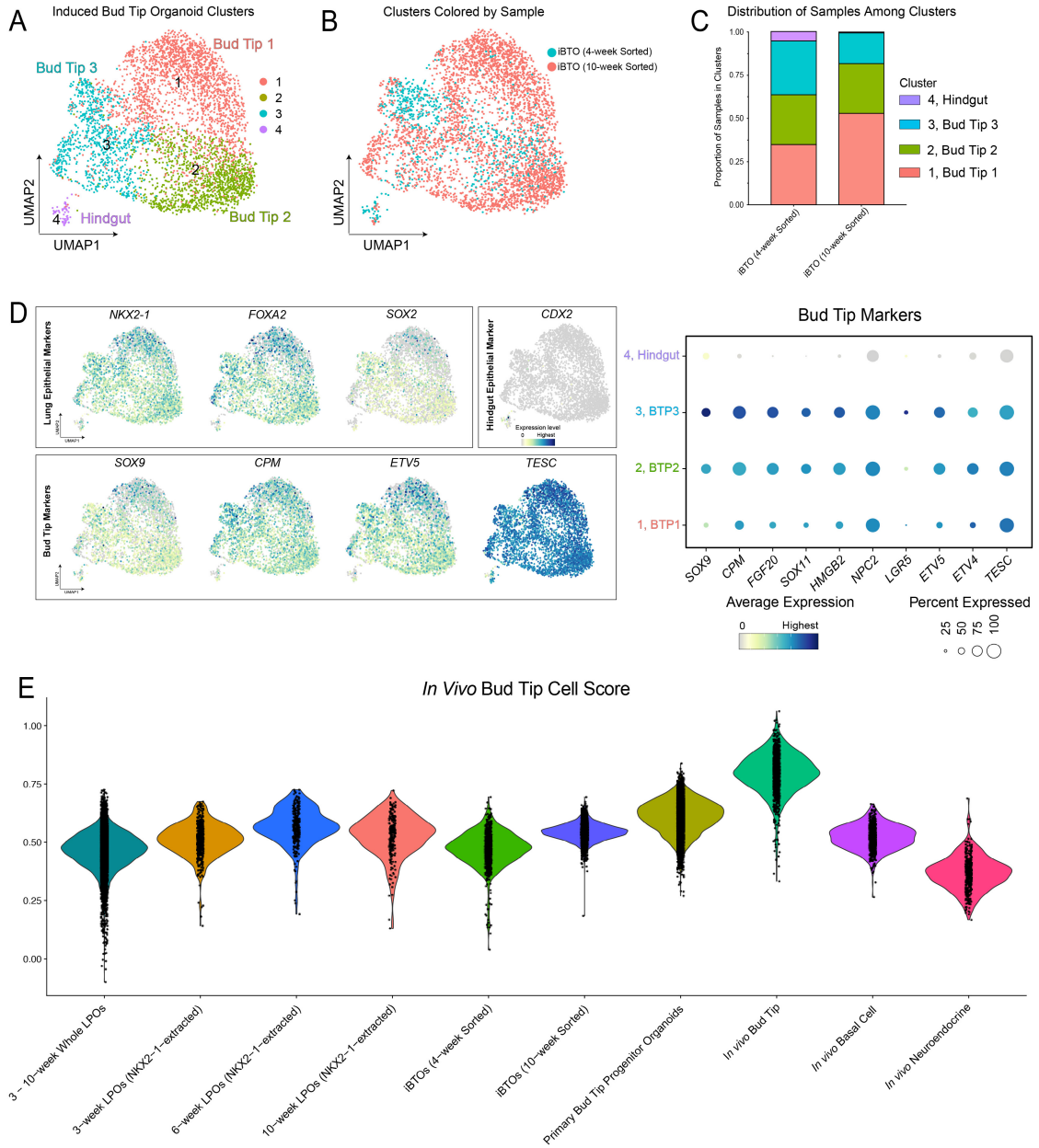


Figure 3-7: iPSC-derived Bud Tip Progenitors are Transcriptionally Similar to Human Fetal Bud Tip Progenitor Cultures. **(A)** UMAP cluster plot of scRNA-seq data from iBTOs (n=1 biological replicates for iBTOs from 4- and 10-week LPOs). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. Cell-type labels for each cluster are based on expression of canonical cell-type markers displayed in the dot plots and feature plots in Figure 3-7D and Figure 3-8B. **(B)** UMAP plot corresponding to the iBTO cluster plot in Figure 3-7A. Each dot represents a single cell and dots/cells are colored by the sample from which they came from. **(C)** Stacked bar graph displaying the proportion of cells from each sample in each cluster of the iBTO cluster plot in Figure 3-7A. **(D)** Feature plots and dot plot corresponding to the iBTO cluster plot in Figure 3-7A and displaying canonical bud tip progenitor markers (*SOX9, CPM, ETV5, TESC, FGF20, SOX11, HGMB2, NPC2, LGR5, ETV4*) (Miller *et al.*, 2018, 2020; Yamamoto *et al.*, 2020), lung epithelial markers (*NKX2-1, FOXA2, SOX2*), and hindgut epithelial marker (*CDX2*). The color of each dot in the feature plot indicates log-normalized expression level of the labeled gene in the represented cell. For the dot plot, the dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized expression level of the gene. **(E)** Violin plot displaying an *in vivo* bud tip progenitor cell score, calculated as the average expression of the top 100 enriched genes in *in vivo* bud tip progenitor cells, for each sample. Samples include whole LPOs (2x3-week LPOs, 2x6-week LPOs, 1x10-week LPOs), NXK2-1-extracted cells from 3-, 6-, and 10-week LPOs (n = 2 for 3- and 6-week LPOs, n = 1 for 10-week LPOs), whole iBTOs (derived from LPOs sorted for NKX2-1^{+}/CPM^{+} cells at 4 or 10 weeks, n=1

of each), human fetal-derived (primary) bud tip progenitor organoids (14-weeks post-conception) (Miller *et al.*, 2020), and primary *in vivo* tissue (8.5 – 19-weeks post-conception), including computationally-extracted bud tip, basal, and neuroendocrine cells (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022).

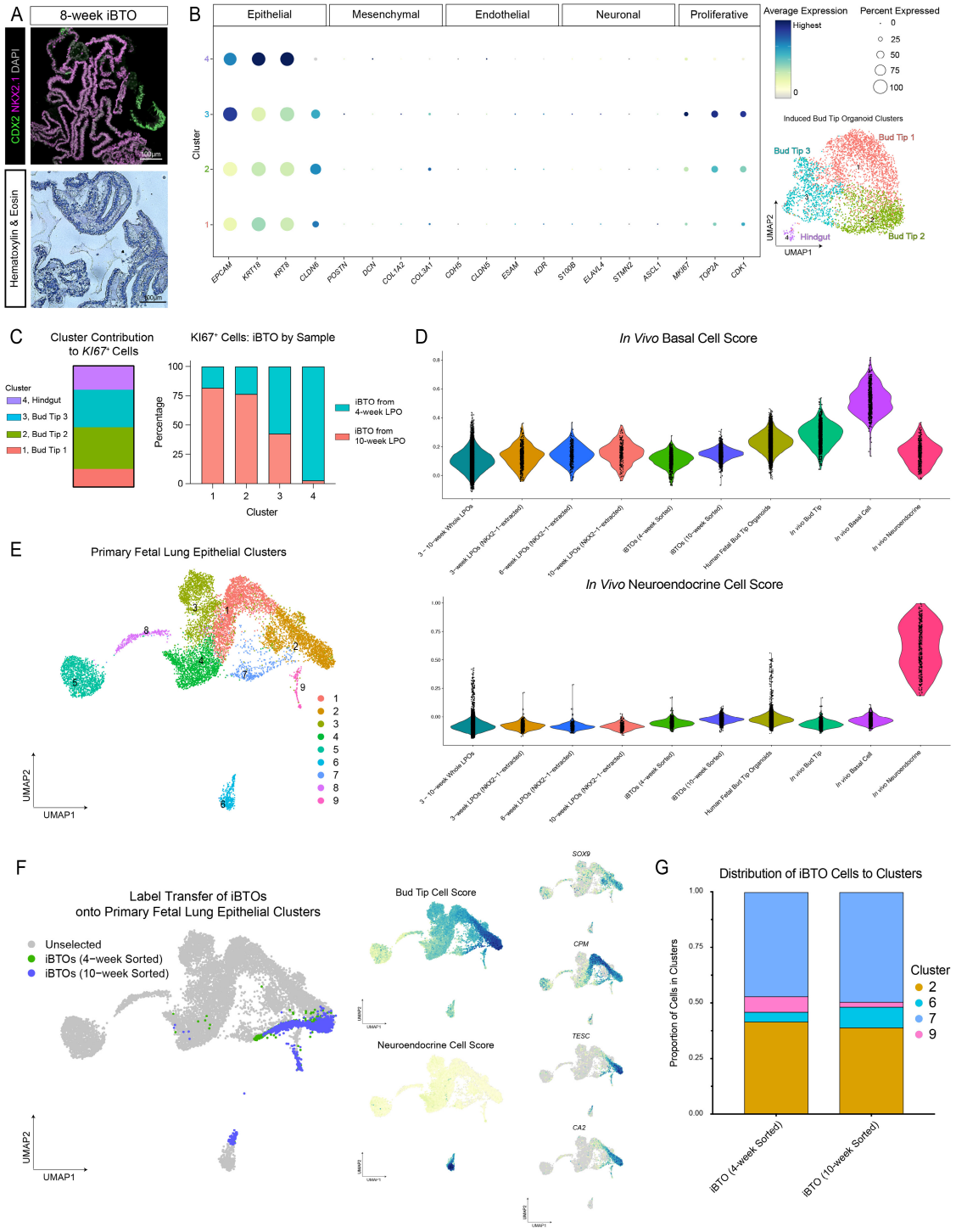


Figure 3-8: Induced Bud Tip Organoids are Enriched for Bud Tip Progenitor Cells, Related to Figures 3-6 and 3-7. (A) (Top) Immunofluorescence staining on paraffin sections for lung epithelial marker NKX2-1 and intestinal epithelial marker CDX2 or (Bottom) hematoxylin & eosin stain on 8-week iBTOs. **(B)** Dot plot of epithelial, mesenchymal, endothelial, and neuronal lineage genes as well as proliferation markers in each cluster of the iBTO cluster plot in Figure 3-7A. The dot size represents the percentage of cells expressing the gene in the corresponding cluster, and the dot color indicates log-normalized expression level of the gene. **(C)** (Left) Stacked bar graph showing each cluster's normalized contribution to KI67^{+} cells, calculated from the scRNA-seq data. (Right) Stacked bar graph showing each sample's normalized contribution to KI67^{+} cells per cluster, calculated from the scRNA-seq data. **(D)** Violin plots displaying an *in vivo* basal (top) or neuroendocrine (bottom) cell score, calculated as the average expression of the top 100 enriched genes in *in vivo* basal or neuroendocrine cells, for each

sample. Samples include whole LPOs (n = 2 for 3- and 6-week LPOs, n = 1 for 10-week LPOs), NXK2-1-extracted cells from 3-, 6-, and 10-week LPOs whole iBTOs (derived from LPOs sorted for $NKX2\text{-}1^{+}/CPM^{+}$ cells at 4 or 10 weeks, n=1 of each), human fetal-derived (primary) bud tip organoids, and primary *in vivo* tissue including computationally-extracted bud tip, basal, and neuroendocrine cells (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). **(E)** UMAP cluster plot of previously-published scRNA-seq data from primary fetal lung epithelium spanning 8.5 – 21 weeks post-conception (n = 11 samples from trachea, airway, and distal bud tips) (Miller *et al.*, 2020; Renee F.C. Hein *et al.*, 2022). Each dot represents a single cell and cells were computationally clustered based on transcriptional similarities. The plot is colored and numbered by cluster. **(F)** Label transfer (Stuart *et al.*, 2019a) of iBTO cells (derived from 4- or 10-week) sequenced via scRNA-seq onto the primary fetal lung epithelial UMAP embedding shown in Figure 3-8E. Feature plots display either an *in vivo* bud tip progenitor or *in vivo* neuroendocrine cell score, calculated as the average expression of the top 100 enriched genes in *in vivo* bud tip progenitor or neuroendocrine cells respectively, or single bud tip progenitor cell markers (*SOX9*, *CPM*, *TESC*, *CA2*). The color of each dot in the feature plot indicates log-normalized expression level of the individual genes or set of genes in the represented cell. **(G)** Stacked bar graph displaying the percentage of iBTO cells from each timepoint that mapped to clusters 2, 6, 7, or 9 in the UMAP embedding shown in Figure 3-8E.

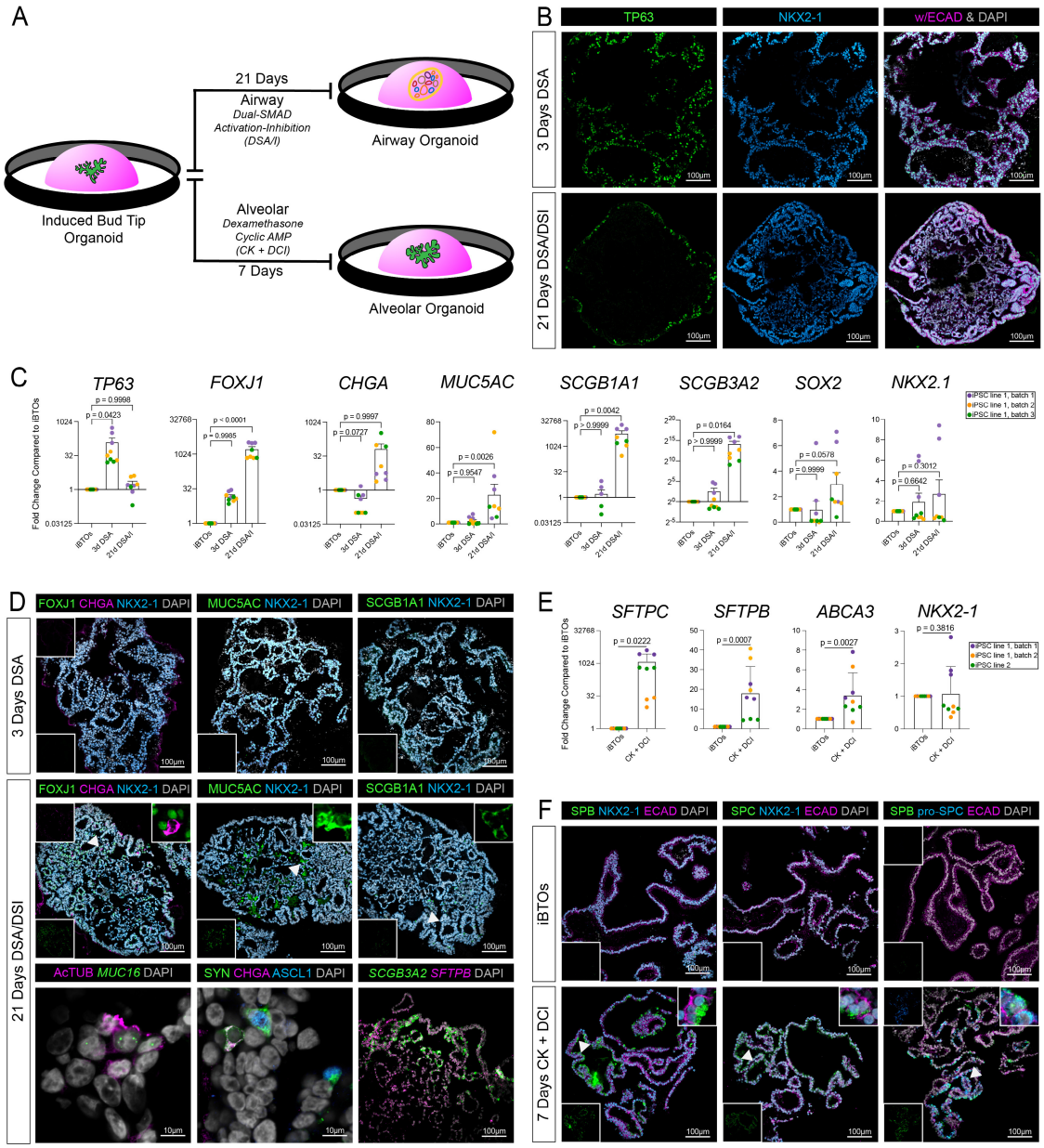


Figure 3-9: iBTOs are Competent for Proximal Airway and Distal Alveolar Differentiation. (A) Schematic displaying the airway (Miller *et al.*, 2020) and alveolar (Jacob *et al.*, 2017) induction protocol from iBTOs. **(B)** Immunofluorescence staining on paraffin sections for the airway progenitor marker TP63 and lung epithelial marker NKX2-1 on iBTOs undergone 3 days of dual-SMAD activation (DSA) or 3 days DSA followed by 18 days of dual-SMAD inactivation (DSI) in the airway induction protocol. **(C)** qRT-PCR data comparing expression of airway markers *TP63, FOXJ1, CHGA*, *MUC5AC*, *SCGB1A1*, *SCGB3A2*, and *SOX2* and lung epithelial marker *NKX2-1* in untreated iBTOs or iBTOs undergone 3 days DSA or 3 days DSA followed by 18 days of DSI of the airway induction protocol. Each color represents an independent experiment using the iPSC17 WT 7B2 line. Error bars represent standard error of the mean. Statistical tests were performed by one-way ANOVA followed by Dunnett's multiple comparison test. **(D)** Fluorescence *in situ* hybridization and/or immunofluorescence staining on paraffin sections for differentiated airway epithelial markers (multiciliated: FOXJ1/AcTUB/*MUC16*, neuroendocrine: CHGA, SYN, ASCL1, goblet: MUC5AC, secretory: SCGB1A1, *SCGB3A2*, *SFTPB*) and lung epithelial marker NKX2-1 on iBTOs undergone 3 days DSA or 3 days DSA followed by 18 days of DSI of the airway induction protocol. Insets on the top right are zoomed in on the area denoted by the arrows (DAPI and NKX2-1 channels removed for clarity). Insets in the bottom left or top left corners are single channel images of non-DAPI, non-NKX2-1 channels. **(E)** qRT-PCR data comparing expression of alveolar markers *SFTPC, SFTPB,* and *ABCA3* and lung epithelial marker *NKX2-1* in untreated iBTOs (in 3F media) or after the alveolar differentiation protocol (7 days CK + DCI). Each color represents an independent experiment with a unique iPSC line or independent experiment (purple/orange: iPSC17 WT 7B2 – LPOs sorted by CPM and NKX2-1-EGFP, green: iPSC line 72.3 – LPOs sorted by CPM only). Error bars

represent standard error of the mean. Statistical tests were performed by unpaired Welch's one-tailed t-tests. **(F)** Immunofluorescence staining on paraffin sections for differentiated type II alveolar epithelial markers SPB, pro-SPC, and SPC and lung epithelial marker NKX2-1 or general epithelial marker ECAD on iBTOs and iBTOs undergone the alveolar differentiation protocol (7 days CK + DCI). Insets on the top right are zoomed in on the area denoted by the arrows. Insets in the bottom left or top left corners are single channel images of non-DAPI, non-NKX2-1, non-ECAD channels.

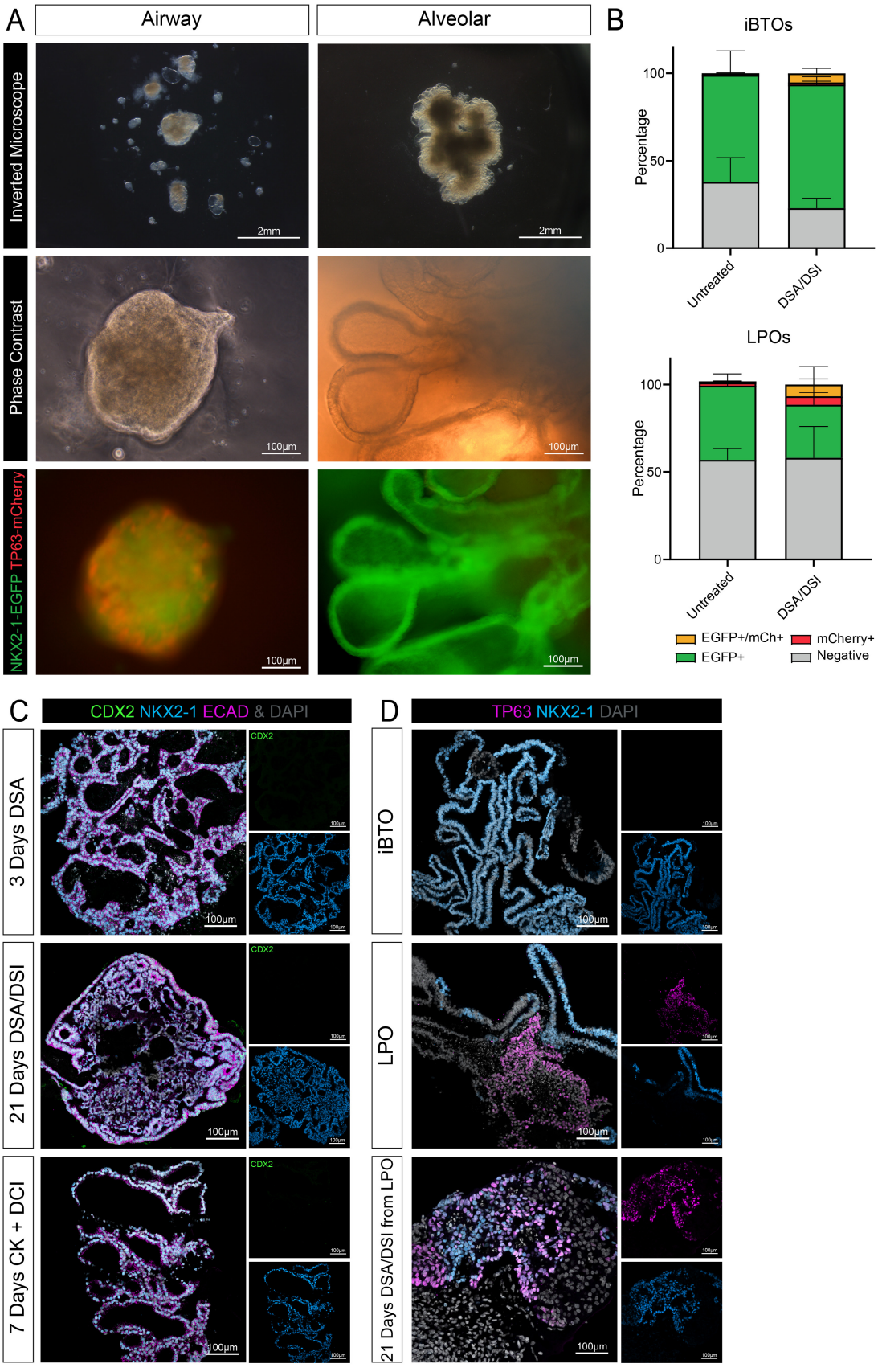

A
Airway
Alveolar
Inverted Microscope
Phase Contrast
NKX2-1-EGFP TP63-mCherry
2mm
100µm
B
iBTOs
LPOs
Percentage
100
50
0
Untreated
DSA/DSI
EGFP+/mCh+
EGFP+
mCherry+
Negative
C
CDX2 NKX2-1 ECAD & DAPI
3 Days DSA
21 Days DSA/DSI
7 Days CK + DCI
D
TP63 NKX2-1 DAPI
iBTO
LPO
21 Days DSA/DSI from LPO

Figure 3-10: iBTOs Maintain NKX2-1 Expression after Differentiation into Proximal Airway and Distal Alveolar Organoids, Related to Figure 3-9. (A) Representative brightfield, phase contrast, and reporter expression for NKX2-1-EGFP and TP63-mCherry images of iBTOs undergone the airway or alveolar induction protocol shown in Figure 3-9A. Phase contrast and reporter images are from the same field of view. **(B)** FACS quantification of NKX2-1-EGFP^{+} and mCherry-TP63^{+} cells from iBTOs and LPOs from the NKX2-1-EGFP/TP63-mCherry reporter cell line, (iPSC 17 WT 7B2) either in 3F media (untreated) or after 21 days of the airway differentiation protocol (DSA/DSI treatment). Percentages of live cells expressing neither reporter, each separate reporter, or dual-expressing are reported as mean ± SEM for 3 – 10 replicates per time point. **(C)** Immunofluorescence staining on paraffin sections for the intestinal epithelial marker CDX2, lung epithelial marker NKX2-1, and general epithelial cell-type marker ECAD on 12-week iBTOs undergone the airway differentiation protocol (DSA/DSI treatment) or 3-week iBTOs undergone the alveolar differentiation protocol (7 days CK + DCI) (see Figure 3-9A). **(D)** Immunofluorescence staining on paraffin sections for the airway progenitor/basal cell marker TP63 and lung epithelial marker NKX2-1 on 12-week iBTOs, 4-week LPOs, and 3-week LPOs after 21 days of the airway differentiation protocol (DSA/DSI treatment).

Chapter 4 Future Directions

This chapter discusses outstanding questions related to the thesis content and explores the next steps to begin addressing these unknown questions. Additionally, future therapeutic potential related to iPSC-based technologies, and particularly to the bud tip organoid model presented in Chapter 3, is discussed.

4.1 Mesenchymal Cell Heterogeneity During Human Lung Development

This thesis has just begun to scratch the surface of characterizing the specific mesenchymal cell types present in the distal bud tip domain during branching morphogenesis as well as their role in regulating bud tip progenitor cells. The analysis performed here was limited to a specific window of time during development (branching morphogenesis) as well as to a specific region of the lung (the distal bud tip domain); this leaves much room for further investigation across developmental time and anatomical region.

4.1.1 Mesenchymal Cell Specification in Early Embryonic Development

How mesenchymal cells are specified and how they influence the formation of NKX2-1^{+} anterior foregut endoderm during the earliest, embryonic stage of respiratory development are not well understood. Recent instrumental work has shown that the WNT, SMAD, RA, FGF, and SHH signaling pathways are involved in specifying respiratory mesenchyme from hPSCs (Kishimoto *et al.*, 2022). Nevertheless, confirmation of how this correlates to signaling in an *in vivo* context, the complete signaling networks between epithelial and mesenchymal cells that

governs their differentiation process, and the mechanisms responsible for the formation of proximal versus distal mesenchymal cell types are incomplete pictures. Lastly, how mechanical ques play a role in the morphogenesis of the lung mesoderm is a fascinating question.

For the time being, animal models, and in particular murine models, will be of extreme value in answering these questions as mimicking foregut morphogenesis with the correct cellular organization from hPSCs remains a critical obstacle. Nevertheless, the recent description of hPSC-derived respiratory mesoderm progenitors (Kishimoto *et al.*, 2022) opens the opportunity to culture these progenitors with different factors to determine which signaling pathways are involved in the specification of specific lung mesenchymal lineages, such as $RSPO2^{+}$ mesenchymal cells.

Current hPSC-derived models generally lack co-differentiation of endoderm-derived cells and tissue-specific mesoderm cells (Yu *et al.*, 2021). The development of more complex hPSC-derived models through co-differentiation or co-culture will be important for answering questions about epithelial-mesenchymal crosstalk during respiratory specification as well as foregut morphogenesis in a human-specific context (Holloway, Capeling and Spence, 2019; Frum and Spence, 2021; Hofer and Lutolf, 2021).

hPSC models of human lung development that include a diversity of lung-specific epithelial and mesenchymal cell types would open the doors to many exciting questions related to epithelial-mesenchymal crosstalk during lung specification. For example, mutations in *RSPO2* that render RSPO2 non-functional (Szenker-Ravi *et al.*, 2018) could be introduced to hPSCs, which could subsequently be used in hPSC-derived lung models to study if RSPO2 has a role in lung endoderm specification.

4.1.2 Changes in Mesenchymal Cell Populations over Developmental Time

Tracking how the mesenchymal cells identified in this thesis emerge before branching morphogenesis and how they evolve into the canalicular, saccular, and alveolar stages of lung development is a space for future exploration. It will be valuable to determine which mesenchymal populations remain constant versus which populations change or are depleted over time. Changes may occur at a small level, with minor changes in the expression of gene sets, or there could be changes in entire cell identity and function. Literature from mice suggests that drastically different mesenchymal cell types are present during alveolar development compared to those identified in this study (during human branching morphogenesis) (Boström *et al.*, 1996; McGowan and Torday, 1997; McGowan *et al.*, 2008; Ahlbrecht and McGowan, 2014; El Agha *et al.*, 2014; A. Li *et al.*, 2016; Green *et al.*, 2016; Endale *et al.*, 2017), and coupled with the extreme morphological changes that occur with development of the alveolar region of the lung (Bal and Ghoshal, 1988; Alanis *et al.*, 2014; Yang *et al.*, 2016), it is likely that the dynamics of the mesenchyme also change drastically.

Of particular interest to the findings in this thesis is the potential of $RSPO2^+$ mesenchymal cells to differentiate into smooth muscle cells. Coupling literature from mice with data from this thesis points strongly towards $RSPO2^+$ mesenchymal cells as lung mesenchymal progenitors. The expression pattern of *Fgf10* with *Wnt2* in the murine lung mesenchyme suggests these cells might be the murine equivalent to $RSPO2^+/WNT2^+$ cells in the human (Goss *et al.*, 2011). Since $Fgf10^+$ cells have been shown to give rise to matrix fibroblasts, myofibroblasts, and lipofibroblasts in mice (El Agha *et al.*, 2014), it is logical to predict $RSPO2^+$ mesenchymal cells may be a progenitor cell in humans. The lineage relation of $RSPO2^+$ mesenchymal cells to other cell types, such as PDGFRa^+ mesenchymal cells, will provide a more complete picture of the lineage dynamics in the human lung mesenchyme.

Confirmation of RSPO2$^+$ mesenchymal cell lineage dynamics in humans is challenging question to address in full *in vitro*; however, the advent of lineage tracing in organoid models (Kong *et al.*, 2020) opens an unprecedented opportunity to answer this question using complex 3-dimensional organoid co-cultures that retain cellular organization, such as the one described here. Additionally, finding culture conditions to isolate and culture RSPO2$^+$ mesenchymal cells in 2D long-term will both inform us about how RSPO2$^+$ mesenchymal cells are maintained and allow us the ability to determine which mechanisms are involved in RSPO2$^+$ mesenchymal cell differentiation.

Because access to human tissue past 20 weeks post-conception is limited or nonexistent due to ethical and legal barriers, studying the mesenchyme during the canalicular through alveolar stages of development using human tissue presents a roadblock. Although animal models may play an important role in answering these questions, developing ways to culture and differentiate mesenchyme from pseudoglandular or earlier stages of development and/or developing complex hESC-derived models that include tissue-specific mesenchyme may hold the answers to these questions.

Through determining the dynamics of the lung mesenchyme during development, we may be able to better understand disease pathogenesis related to the lung mesenchyme that occurs in pediatric or adult patients. It's possible that diseases involving the lung mesenchyme, including induced pulmonary fibrosis and cancer, regain developmental programs and/or that mesenchymal cell types revert to transient populations normally present during development. Regardless, efforts to build pediatric and adult cell atlases, culture mesenchymal cells such that they retain their identity and function, and build methods to study mesenchymal cell dynamics will be instrumental in investigating these questions.

4.1.3 Regional Changes in Mesenchymal Cell Populations

The work in this thesis uncovered heterogeneity among mesenchyme that exists in the distal lung near bud tips; however, whether the same patterns exist in distal (non-cartilaginous) airways, proximal (cartilaginous) airways, and the trachea is unexplored territory. The work in this thesis has shown that $RSPO2^+$ mesenchyme is non-existent in the very proximal lung/trachea, and it is known that smooth muscle cells line the conducting airways along the proximal-distal axis while cartilage only exists in the trachea and larger, most proximal airways. However, there is room to explore potential unknown mesenchymal cell types that exist proximal to bud tip regions as well as how these mesenchymal cells crosstalk with epithelial cells to influence epithelial cell fate and function and/or tissue morphogenesis. Moreover, the specific characteristics of smooth muscle cells that line more distal airways versus more proximal airways is yet to be investigated.

4.1.4 Sub-populations of Major Mesenchymal Cell Classes

This thesis revealed five transcriptionally and spatially distinct populations of mesenchyme in the bud tip domain during branching morphogenesis. The function of $RSPO2^+$ mesenchymal cells was attributed to maintaining bud tip progenitor fate; however, these cells are not only prevalent in regions within the niche directly adjacent to bud tip progenitors, suggesting that $RSPO2^+$ cells could play roles beyond bud tip progenitor maintenance. Coupled with the expression pattern of *LGR5* in subsets of proximal lung basal cells, *LGR4* expression broadly throughout the mesenchyme, and *LGR6* expression in airway smooth muscle cells, how RSPO2 influences these other cells is a fascinating avenue for future exploration. This also opens the question about differences between $LGR5^+$ and $LGR5^-$ basal cells and the role of RSPO2 in maintaining important functional or cell fate differences they may have.

Current *in vitro* models of human airway differentiation and basal cell isolation (Katherine B McCauley *et al.*, 2017; Miller *et al.*, 2020) could be implemented to test questions about how RSPO2 influences basal cells. Deeper investigation into the scRNA-seq data presented here and ligand-receptor analysis has the potential to uncover sub-populations of $RSPO2^+$ mesenchyme that might inspire hypotheses about crosstalk between $RSPO2^+$ mesenchyme and other cell types.

4.2 Further Interrogating the Role of RSPO2 During Human Lung Development

This thesis displayed an important role for RSPO2 in maintaining the bud tip progenitor cell niche during human lung development. The role of other R-Spondin proteins as well as roles of RSPO2 beyond bud tip progenitor maintenance are areas for further research.

4.2.1 The Specificity of RSPO2 Versus Other R-Spondin Proteins

Although RSPO2 is the main R-Spondin expressed in the developing human lung, whether or not the other R-Spondin proteins (RSPO1, RSPO3, RSPO4) can serve the same functions as RSPO2 in the developing human lung is unknown. Answering this could be important for determining differences in mouse-human biology and may explain the differences seen in the severity of lung defects in mice with *Rspo2* mutations versus humans with *RSPO2* mutations.

The advent of homogenous cultures of bud tip progenitor cells for both mice and humans (Nichane *et al.*, 2017; Nikolić *et al.*, 2017; Miller *et al.*, 2018) opens up the possibility for experiments to test this question. The research in this thesis has shown that bud tip progenitor organoids can be cultured with recombinant RSPO2 protein in place of the small molecule WNT

activator CHIR99021. By culturing bud tip progenitor organoids with the other R-Spondin proteins and determining the maintenance of bud tip progenitor cells would answer this question.

4.2.2 The Larger Signaling Network Regulating Bud Tip Progenitors

Experiments primarily using mouse genetics have uncovered many of the signaling pathways active and inactive at the bud tips (Morrisey and Hogan, 2010); however, the intricacies of the signaling pathways that intersect to create a signaling network regulating bud tip progenitor cell fate, the downstream targets of these networks, and the mechanisms either conserved or specific to humans is unknown. Additionally, it would be interesting to uncover which cell types are facilitating these signaling networks.

For example, human bud tips can be maintained in a minimal media containing FGF7, WNT, and ATRA, but the upstream signaling pathways that converge on these factors and the transcriptional targets downstream of these factors would be exciting to explore. CHIP-sequencing experiments are currently underway to analyze the DNA binding sites of β-catenin in native and cultured bud tips. In conjunction with this, CHIP-sequencing experiments to determine how the DNA binding sites of SOX9 and SOX2 change when the levels of WNT signaling change (and bud tip progenitors differentiate to airway) in cultured bud tips is underway. Coupling CHIP-seq experiments with ATAC-seq experiments may lead to interesting hypotheses about how the signaling pathways we know to be at play in bud tips relate to SOX9 and SOX2 expression and function. Lastly, investigating these dynamics across time to determine if WNT target genes change from lung specification to branching morphogenesis to alveologenesis and into adult tissue homeostasis and regeneration, and whether this relates to which upstream WNT components are at play, would provide the field with detailed information

about how the intricacies of the WNT signaling pathway can lead to differential downstream results.

4.3 Increasing Complexity and Accuracy of hPSC-derived Lung Models

hPSC-derived lung models offer more accurate and complex representations of native cells than immortalized cell lines, mitigate the shortcomings of tissue availability that primary-derived *in vitro* models carry, and don't come with the risk of missing human-specific biology that comes with using non-human animal models in studies (Kim, Koo and Knoblich, 2020). However, there are still many limitations of current hPSC-derived lung models.

4.3.1 Co-differentiation and Co-culture of Non-epithelial Cell Types

First, most hPSC-derived lung models are epithelial-only and therefore lack neural, immune, endothelial, and mesenchymal cell types, although work is underway to incorporate these cell types into organoids (Holloway, Capeling and Spence, 2019). Until hPSC-derived lung models advance in complexity to include these cell types, it will be impossible to study epithelial to non-epithelial cell interactions without the use of primary tissue or animal models. Another critical step forward will be to organizing these cells in a way that mimics native tissue architecture and leads to functional neural, immune, and vascular systems within organoids or other models, such as organ-on-a-chip systems.

4.3.2 Benchmarking hPSC-derived Lung Models

As hPSC-derived lung models increase in complexity, it will be important to determine similarity between hPSC-derived cells and related primary cells. Studies, such as the one presented here, have started this effort as scRNA-seq technologies have become more commonplace (Bhaduri *et al.*, 2020; De Los Angeles and Tunbridge, 2020). The creation of cell

atlases has been and will be an important benchmarking resource for scientists to compare hPSC-derived cultures to primary cells and cultures. Although transcriptomic analyses like scRNA-seq provide a unique blueprint of individual cell types, they do not provide a complete picture of a cell. Additional ways to benchmark cells, such as through single nucleus ATAC sequencing to determine if the epigenome of hPSC-derived cells has been modified to that of its target cell type, can help paint a more complete picture of individual cell types and maturation status. Moreover, functional analyses will always be important in determining if hPSC-generated cells have the same functional capacity as their *in vivo* counterparts.

As the number of single cell datasets increases, a better appreciation of the heterogeneity among cell types and within what was previously thought to be a single cell type will result. For example, within the human lung, basal cells are located in the trachea, large cartilaginous airways, and non-cartilaginous airways (Rock and Hogan, 2011). It is unknown how the characteristics of basal cells change based on anatomical location within the lung. Determining which subset of basal cells are most similar to the basal cells induced from iBTOs via the DSA/I protocol in this study will be critical for deciding if to use this model for specific experiments and for determining if additional protocols need to be developed to induce basal cells that better resemble basal cells from a different anatomical location.

4.3.3 Cellular Plasticity and Maturation

As is evident through the work in this thesis and other recent work (Little, Gerner-Mauro, Flodby, Crandall, Borok, Akiyama, *et al.*, 2019; Hurley *et al.*, 2020), hPSC-derived cultures exhibit enormous amounts of cellular plasticity, which is likely due to the naive and immature nature of these cultures. Cellular plasticity results in cultures that may not maintain their intended cell fate and could therefore lead to inaccurate data from studies using these models;

therefore, overcoming this hurdle will be critical if the use of hPSC-derived models to answer biological questions is going to become commonplace.

Currently, the method to mature hPSC-derived cultures away from resembling early, fetal-like cells is transplantation into adult mice (Dye *et al.*, 2016). This is not practical for the design of many experiments and prevents scientists from moving away from the use of animal models. The advent of non-transplantation methods to mature hPSC-derived cultures and limit cellular plasticity will be a huge step forward for accurately modeling biological phenomenon *in vitro*.

4.3.4 Simple Steps to Improve the Accuracy of hPSC-derived Lung Models

Simple but important steps forward in generating cells that better mimic native cells will be improving media conditions through small molecule and growth factor screens and biochemical analyses to determine ideal growth factor concentrations. An example relevant to the current work is the concentration of the small molecule WNT activator CHIR99021 used in iBTO culture, which is currently 3 μM. Studies have suggested that overactive WNT signaling causes cell fate changes from lung to gut (Okubo and Hogan, 2004; Hurley *et al.*, 2020). Whether something as simple as decreasing the concentration of CHIR99021 in iBTO cultures will reduce gut contamination could be an important future direction of this work.

4.4 Therapeutic Potential of iPSC-derived Lung Tissue

The invention of reprogramming somatic cells into pluripotent-like stem cells (i.e., iPSCs) both reduces the need for human embryonic stem cells in scientific research and presents the possibility to use patient-specific cells to make iPSCs for cell therapy and tissue regeneration.

Although transforming patient cells into iPSCs opens a door for "personalized" medicines and therapies, there are still many challenges left to overcome.

4.4.1 Disease Modeling

Many human diseases do not exist naturally in non-human animals or arise via different mechanisms; therefore, established animal models to study these diseases often do not begin and progress how they would in humans (Hau, 2008; Picher-Martel *et al.*, 2016; Pound and Ritskes-Hoitinga, 2018; Onaciu *et al.*, 2020). iPSC-derived models offer an extremely valuable tool for studying human-specific diseases. Cells from patients with a particular disease of interest can be used to make iPSCs and subsequently *in vitro* models of the disease by applying directed differentiation approaches. This is particularly relevant to genetic-related diseases, such that iPSCs carrying specific, real-life mutations in a gene can be used to study disease onset, progression, and outcome resulting from a particular mutation. For example, although cystic fibrosis is caused by a mutation in a single gene *CFTR*, there are thousands of unique mutations in *CFTR* that lead to the disease (Bobadilla *et al.*, 2002; Wang *et al.*, 2014). The iBTOs undergone airway differentiation presented in this thesis could serve as a model to study how different mutations in *CFTR* lead to specific effects on lung cells and lung function.

4.4.2 Personalized Medicine and Drug Discovery

Perhaps the most promising use of iPSC-derived cells and tissues is within drug discovery and personalized medicine applications (Corbett and Duncan, 2019; Plummer *et al.*, 2019; Rowe and Daley, 2019; Vatine *et al.*, 2019; Sharma *et al.*, 2020; Jiang *et al.*, 2021; Cholon and Gentzsch, 2022; Moreira *et al.*, 2022). In general, iPSC-derived organoids mimic native tissue much better than immortalized cell lines do (Kim, Koo and Knoblich, 2020), and drugs

developed in *in vivo* animal models often fail clinically due to inherent differences between humans and other animals (Shanks, Greek and Greek, 2009; Bart van der Worp *et al.*, 2010; Mak, Evaniew and Ghert, 2014; McGonigle and Ruggeri, 2014; Akhtar, 2015; Pound and Ritskes-Hoitinga, 2018). Beyond just anatomical differences between animal models and humans, animal models do not represent the genetic diversity, lifespan, comorbidities, and lifelong environmental exposures seen in humans. Using iPSC-derived organoids in drug discovery efforts are more likely to result in the identification of target drugs that will be effective when moved into clinical trials when compared to drugs discovered using cell lines and animal models.

Moreover, although many diseases have one or few accepted therapies that are applied to all patients with that disease, the same disease can present itself with different phenotypes. Therefore, different patients may benefit from personalized drug regimens that interplay with the unique aspects of their disease (Goyette *et al.*, 2009; Agusti *et al.*, 2016; Kaur and Chupp, 2019; Leopold, Maron and Loscalzo, 2020; Matthay *et al.*, 2020). This is true for cystic fibrosis, where different genetic mutations and environmental factors result in diverse manifestations of the disease (Marson, Bertuzzo and Ribeiro, 2017). Using patient-specific iPSCs with directed differentiation paradigms can serve as invaluable models to predict how a patient will respond to a drug or set of drugs.

4.4.3 Regenerative Medicine

A potentially groundbreaking therapeutic area of iPSC work is in regenerative medicine applications. If new cells or tissues can be made from iPSCs through directed differentiation, this opens the possibility for those cells or tissues to be transplanted into human bodies to replace damaged or diseased tissue. Using cystic fibrosis as an example, ongoing efforts are underway to

reprogram patient cells into iPSCs, genetically correct *CFTR* mutations in patient-derived iPSCs, and direct their differentiation into functional basal cells that could be transplanted back into the patient to repopulate their lungs with healthy cells and restore lung function (*Stem Cells for Cystic Fibrosis Therapy | Cystic Fibrosis Foundation*, no date). By using patient-specific cells, the chance of immune rejection that is common in whole organ transplants is greatly reduced (Madrid *et al.*, 2021). Regenerative medicine using iPSCs is being investigated in clinical trials for many diseases, including heart failure, cancer, Parkinson's disease, macular degeneration and COVID-19 (Kim *et al.*, 2022).

4.4.4 Current Limitations

Although there is great promise in the therapeutic use of iPSC-derived cells and tissues, there are still many challenges at bay (Doss and Sachinidis, 2019; Yamanaka, 2020; Madrid *et al.*, 2021). First off, iPSC-derived cultures are generally immature *in vitro* and mimic fetal rather than adult tissue (Dye *et al.*, 2015, 2016; Holloway, Capeling and Spence, 2019; Frum and Spence, 2021). The iBTOs presented in this thesis require a 10+ weeks in culture to mature to cells that closely resemble fetal bud tip progenitors. In many cases, it isn't until organoids are transplanted *in vivo* that they mature. However, this suggests that an *in vivo* environment provides the necessary support for iPSC-derived tissue maturation, and if the end goal of a therapy is a transplantation, this may not be an issue. It may, however, be ideal to execute experiments that determine what the maturation factors an *in vivo* environment provides such that iPSC-derived cultures can me matured *in vitro*. This could be especially important because the immature and plastic nature of iPSC-derived cells increases that chance for tumor growth (Lee *et al.*, 2013).

Another important consideration of the clinical use of iPSC-derived cultures is manufacturing of the cells and tissues (e.g. scale up, cryopreservation). Manufacturing iPSCs from many individuals comes with huge manufacturing challenges that has prompted the investigation into allogenic iPSCs; however, allogenic iPSCs reintroduce issues with immune rejection (Petrus-Reurer *et al.*, 2021). Additionally, using FDA-approved/non-harmful reagents will be especially critical if cells are to be transplanted into patients.

Nevertheless, ongoing research regarding these challenges may remove the current roadblocks, paving a way for groundbreaking progress in personalized medicine and regenerative medicine using iPSC technology.

4.5 Concluding Remarks

Overall, this thesis has increased the field's depth of knowledge about bud tip progenitor cell fate decisions by identifying an important mesenchymal niche-cell that provides a high WNT environment for bud tips. This knowledge and vast amounts of previous knowledge from the field was then used to develop an optimized model of iPSC-derived bud tip progenitors, which was benchmarked to native tissue. This research is important for better understanding how human-specific development occurs, what mechanisms are at play when development goes awry, and for being able to model human lung development using iPSCs.

Bibliography

Abler, L. L. *et al.* (2017) 'Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids', *Development*. American Physiological Society, 4(1), p. e26575. doi: 10.7554/eLife.26575.

Abler, L. L., Mansour, S. L. and Sun, X. (2008) 'Conditional gene inactivation reveals roles for Fgf10 and Fgfr2 in establishing a normal pattern of epithelial branching in the mouse lung', *Developmental Dynamics*, 238(8), pp. 1999–2013. doi: 10.1002/dvdy.22032.

Adamson, I. Y. and Bowden, D. H. (1975) 'Derivation of type 1 epithelium from type 2 cells in the developing rat lung.', *Laboratory investigation; a journal of technical methods and pathology*, 32(6), pp. 736–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1171339 (Accessed: 3 December 2019).

El Agha, E. *et al.* (2014) 'Fgf10-positive cells represent a progenitor cell population during lung development and postnatally.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 141(2), pp. 296–306. doi: 10.1242/dev.099747.

El Agha, E. and Bellusci, S. (2014) 'Walking along the Fibroblast Growth Factor 10 Route: A Key Pathway to Understand the Control and Regulation of Epithelial and Mesenchymal Cell-Lineage Formation during Lung Development and Repair after Injury.', *Scientifica*. Hindawi, 2014, p. 538379. doi: 10.1155/2014/538379.

Agusti, A. *et al.* (2016) 'Treatable traits: toward precision medicine of chronic airway diseases', *European Respiratory Journal*. European Respiratory Society, 47(2), pp. 410–419. doi: 10.1183/13993003.01359-2015.

Ahlbrecht, K. and McGowan, S. E. (2014) 'In search of the elusive lipofibroblast in human lungs', *American Journal of Physiology-Lung Cellular and Molecular Physiology*. American Physiological Society Bethesda, MD, 307(8), pp. L605–L608. doi: 10.1152/ajplung.00230.2014.

Akhtar, A. (2015) 'The Flaws and Human Harms of Animal Experimentation', *Cambridge Quarterly of Healthcare Ethics*. Cambridge University Press, 24(4), p. 407. doi: 10.1017/S0963180115000079.

Al Alam, D. *et al.* (2015) 'Evidence for the involvement of fibroblast growth factor 10 in lipofibroblast formation during embryonic lung development.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 142(23), pp. 4139–50. doi: 10.1242/dev.109173.

Alanis, D. M. *et al.* (2014) 'Two nested developmental waves demarcate a compartment boundary in the mouse lung', *Nature Communications*. Nature Publishing Group, 5(1), p. 3923. doi: 10.1038/ncomms4923.

Alejandre-Alcázar, M. A. *et al.* (2007) 'Temporal and spatial regulation of bone morphogenetic protein signaling in late lung development', *Developmental Dynamics*. John Wiley & Sons, Ltd, 236(10), pp. 2825–2835. doi: 10.1002/dvdy.21293.

Alescio, T. and Cassini, A. (1962) 'Inductionin vitro of tracheal buds by pulmonary mesenchyme grafted on tracheal epithelium', *Journal of Experimental Zoology*. John Wiley & Sons, Ltd, 150(2), pp. 83–94. doi: 10.1002/jez.1401500202.

Alysandratos, K.-D. *et al.* (2022) 'Impact of cell culture on the transcriptomic programs of primary and iPSC-derived human alveolar type 2 cells', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2022.02.08.479591. doi: 10.1101/2022.02.08.479591.

Ardini-Poleske, M. E. *et al.* (2017) 'LungMAP: The Molecular Atlas of Lung Development Program', *American Journal of Physiology-Lung Cellular and Molecular Physiology*. American Physiological Society Bethesda, MD, 313(5), pp. L733–L740. doi: 10.1152/ajplung.00139.2017.

Arora, R., Metzger, R. J. and Papaioannou, V. E. (2012) 'Multiple Roles and Interactions of Tbx4 and Tbx5 in Development of the Respiratory System', *PLoS Genetics*. Edited by G. S. Barsh. Public Library of Science, 8(8), p. e1002866. doi: 10.1371/journal.pgen.1002866.

Aros, C. J., Pantoja, C. J. and Gomperts, B. N. (2021) 'Wnt signaling in lung development, regeneration, and disease progression', *Communications Biology 2021 4:1*. Nature Publishing Group, 4(1), pp. 1–13. doi: 10.1038/s42003-021-02118-w.

Bader, B. L. *et al.* (2005) ‘Compound genetic ablation of nidogen 1 and 2 causes basement membrane defects and perinatal lethality in mice.’, *Molecular and Cellular Biology*, 25(15), pp. 6846–6856.

Bagai, S. *et al.* (2002) ‘Fibroblast growth factor-10 is a mitogen for urothelial cells’, *The Journal of biological chemistry*. J Biol Chem, 277(26), pp. 23828–23837. doi: 10.1074/JBC.M201658200.

Bal, H. S. and Ghoshal, N. G. (1988) ‘Morphology of the terminal bronchiolar region of common laboratory mammals’, *Laboratory Animals*. SAGE PublicationsSage UK: London, England, 22(1), pp. 76–82. doi: 10.1258/002367788780746539.

Barker, N. *et al.* (2007) ‘Identification of stem cells in small intestine and colon by marker gene Lgr5’, *Nature*. Nature Publishing Group, 449(7165), pp. 1003–1007. doi: 10.1038/nature06196.

Bart van der Worp, H. *et al.* (2010) ‘Can Animal Models of Disease Reliably Inform Human Studies?’, *PLOS Medicine*. Public Library of Science, 7(3), p. e1000245. doi: 10.1371/JOURNAL.PMED.1000245.

Baud, O. *et al.* (2016) ‘Effect of early low-dose hydrocortisone on survival without bronchopulmonary dysplasia in extremely preterm infants (PREMILOC): a double-blind, placebo-controlled, multicentre, randomised trial’, *The Lancet*. Elsevier, 387(10030), pp. 1827–1836. doi: 10.1016/S0140-6736(16)00202-6.

Baulies, A., Angelis, N. and Li, V. S. W. (2020) ‘Hallmarks of intestinal stem cells’, *Development (Cambridge, England)*. NLM (Medline). doi: 10.1242/dev.182675.

Becht, E. *et al.* (2019) ‘Dimensionality reduction for visualizing single-cell data using UMAP’, *Nature Biotechnology*. Nature Publishing Group, 37(1), pp. 38–44. doi: 10.1038/nbt.4314.

Bell, S. M. *et al.* (2008) ‘R-spondin 2 is required for normal laryngeal-tracheal, lung and limb morphogenesis.’, *Development (Cambridge, England)*. The Company of Biologists Ltd, 135(6), pp. 1049–58. doi: 10.1242/dev.013359.

Bellusci, S. *et al.* (1997) ‘Fibroblast growth factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung.’, *Development (Cambridge, England)*, 124(23), pp. 4867–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9428423 (Accessed: 16 August 2019).

Bhaduri, A. *et al.* (2020) 'Are Organoids Ready for Prime Time?', *Cell stem cell*. Cell Stem Cell, 27(3), pp. 361–365. doi: 10.1016/J.STEM.2020.08.013.

Blondel, V. D. *et al.* (2008) 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10). doi: 10.1088/1742-5468/2008/10/P10008.

Bobadilla, J. L. *et al.* (2002) 'Cystic fibrosis: A worldwide analysis of CFTR mutations—correlation with incidence data and application to screening', *Human Mutation*. John Wiley & Sons, Ltd, 19(6), pp. 575–606. doi: 10.1002/HUMU.10041.

Bogue, C. W. *et al.* (1998) 'Hex Expression During Development Suggests an Important Role in Both Gastrulation and Organogenesis • 251', *Pediatric Research 1998 43:4*. Nature Publishing Group, 43(4), pp. 45–45. doi: 10.1203/00006450-199804001-00272.

Boström, H. *et al.* (1996) 'PDGF-A Signaling Is a Critical Event in Lung Alveolar Myofibroblast Development and Alveogenesis', *Cell*. Cell Press, 85(6), pp. 863–873. doi: 10.1016/S0092-8674(00)81270-2.

Branchfield, K. *et al.* (2016) 'A three-dimensional study of alveologenesis in mouse lung', *Developmental Biology*. Academic Press, 409(2), pp. 429–441. doi: 10.1016/J.YDBIO.2015.11.017.

Brazovskaja, A., Treutlein, B. and Camp, J. G. (2019) 'High-throughput single-cell transcriptomics on organoids', *Current Opinion in Biotechnology*. Elsevier Current Trends, 55, pp. 167–171. doi: 10.1016/J.COPBIO.2018.11.002.

Bridges, J. P. *et al.* (2020) 'Glucocorticoid Regulates Mesenchymal Cell Differentiation Required for Perinatal Lung Morphogenesis and Function', *American Journal of Physiology-Lung Cellular and Molecular Physiology*. American Physiological Society Rockville, MD , p. ajplung.00459.2019. doi: 10.1152/ajplung.00459.2019.

Cardoso, W. V *et al.* (2006) 'Regulation of early lung morphogenesis: questions, facts and controversies.', *Development (Cambridge, England)*. The Company of Biologists Ltd, 133(9), pp. 1611–24. doi: 10.1242/dev.02310.

Carraro, G. *et al.* (2021) 'Transcriptional analysis of cystic fibrosis airways at single-cell

resolution reveals altered epithelial cell states and composition', *Nature Medicine 2021 27:5*. Nature Publishing Group, 27(5), pp. 806–814. doi: 10.1038/s41591-021-01332-7.

de Carvalho, A. L. R. T. *et al.* (2019) 'Glycogen synthase kinase 3 induces multilineage maturation of human pluripotent stem cell-derived lung progenitors in 3D culture.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 146(2), p. dev171652. doi: 10.1242/dev.171652.

Chakraborty, M. and Kotecha, S. (2013) 'Pulmonary surfactant in newborn infants and children', *Breathe*. European Respiratory Society, 9(6), pp. 476–488. doi: 10.1183/20734735.006513.

Chang, D. R. *et al.* (2013) 'Lung epithelial branching program antagonizes alveolar differentiation.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(45), pp. 18042–51. doi: 10.1073/pnas.1311760110.

Chen, F. *et al.* (2010) 'A retinoic acid–dependent network in the foregut controls formation of the mouse lung primordium', *Journal of Clinical Investigation*, 120(6), pp. 2040–2048. doi: 10.1172/JCI40253.

Chen, L. *et al.* (2012) 'Dynamic regulation of platelet-derived growth factor receptor α expression in alveolar fibroblasts during realveolarization.', *American journal of respiratory cell and molecular biology*. American Thoracic Society, 47(4), pp. 517–27. doi: 10.1165/rcmb.2012-0030OC.

Chen, P. H. *et al.* (2013) 'The structural basis of R-spondin recognition by LGR5 and RNF43', *Genes and Development*. Genes Dev, 27(12), pp. 1345–1350. doi: 10.1101/gad.219915.113.

Chen, Y.-W. *et al.* (2017) 'A three-dimensional model of human lung development and disease from pluripotent stem cells', *Nature Cell Biology*. Nature Publishing Group, 19(5), pp. 542–549. doi: 10.1038/ncb3510.

Chin, A. M. *et al.* (2016) 'A Dynamic WNT/β-CATENIN Signaling Environment Leads to WNT-Independent and WNT-Dependent Proliferation of Embryonic Intestinal Progenitor Cells', *Stem Cell Reports*. Cell Press, 7(5), pp. 826–839. doi: 10.1016/J.STEMCR.2016.09.004/ATTACHMENT/29E9B580-5C50-4DF3-B3B1-3B16564B4215/MMC1.PDF.

Cholon, D. M. and Gentzsch, M. (2022) 'Established and novel human translational models to advance cystic fibrosis research, drug discovery, and optimize CFTR-targeting therapeutics', *Current Opinion in Pharmacology*. Elsevier, 64, p. 102210. doi: 10.1016/J.COPH.2022.102210.

Cole, T. J. *et al.* (2004) 'Altered Epithelial Cell Proportions in the Fetal Lung of Glucocorticoid Receptor Null Mice', *American Journal of Respiratory Cell and Molecular Biology*, 30(5), pp. 613–619. doi: 10.1165/rcmb.2003-0236OC.

Conchola, A. S. *et al.* (2022) 'Distinct airway progenitor cells drive epithelial heterogeneity in the developing human lung', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2022.06.13.495813. doi: 10.1101/2022.06.13.495813.

Conway, R. F. *et al.* (2020) 'Understanding Human Lung Development through In Vitro Model Systems', *BioEssays*. John Wiley and Sons Inc., 42(6). doi: 10.1002/bies.202000006.

Corbett, J. L. and Duncan, S. A. (2019) 'iPSC-Derived Hepatocytes as a Platform for Disease Modeling and Drug Discovery', *Frontiers in Medicine*. Frontiers Media S.A., 6, p. 265. doi: 10.3389/FMED.2019.00265/BIBTEX.

Crane, A. M. *et al.* (2015) 'Targeted correction and restored function of the CFTR gene in cystic fibrosis induced pluripotent stem cells', *Stem Cell Reports*. Cell Press, 4(4), pp. 569–577. doi: 10.1016/J.STEMCR.2015.02.005/ATTACHMENT/B791AC47-2A8A-4089-885A-233B1EE7974B/MMC1.PDF.

D'Amour, K. A. *et al.* (2006) 'Production of pancreatic hormone–expressing endocrine cells from human embryonic stem cells', *Nature Biotechnology*. Nature Publishing Group, 24(11), pp. 1392–1401. doi: 10.1038/nbt1259.

Dame, M. K. *et al.* (2018) 'Identification, isolation and characterization of human LGR5-positive colon adenoma cells', *Development (Cambridge)*. Company of Biologists Ltd, 145(6). doi: 10.1242/dev.153049.

Danopoulos, S. *et al.* (2016) 'Rac1 modulates mammalian lung branching morphogenesis in part through canonical Wnt signaling.', *American journal of physiology. Lung cellular and molecular physiology*, 311(6), pp. L1036--L1049. doi: 10.1152/ajplung.00274.2016.

Danopoulos, S., Thornton, M. E., *et al.* (2018) 'Discordant roles for FGF ligands in lung

branching morphogenesis between human and mouse', *The Journal of Pathology*. John Wiley & Sons, Ltd, 247(2), p. path.5188. doi: 10.1002/path.5188.

Danopoulos, S., Alonso, I., *et al.* (2018) 'Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9', *American Journal of Physiology-Lung Cellular and Molecular Physiology*. American Physiological Society Bethesda, MD , 314(1), pp. L144–L149. doi: 10.1152/ajplung.00379.2017.

Danopoulos, S. *et al.* (2019) 'Transcriptional Characterisation of Human Lung Cells Identifies Novel Mesenchymal Lineage Markers', *European Respiratory Journal*, p. 1900746. doi: 10.1183/13993003.00746-2019.

Danopoulos, S. *et al.* (2020) 'Transcriptional characterisation of human lung cells identifies novel mesenchymal lineage markers', *European Respiratory Society*, 55(1), p. 1900746. doi: 10.1183/13993003.00746-2019.

Danopoulos, S., Shiosaki, J. and Al Alam, D. (2019) 'FGF Signaling in Lung Development and Disease: Human Versus Mouse', *Frontiers in Genetics*, 10, p. 170. doi: 10.3389/fgene.2019.00170.

Davenport, C. *et al.* (2016) 'Anterior-Posterior Patterning of Definitive Endoderm Generated from Human Embryonic Stem Cells Depends on the Differential Signaling of Retinoic Acid, Wnt-, and BMP-Signaling', *STEM CELLS*. John Wiley & Sons, Ltd, 34(11), pp. 2635–2647. doi: 10.1002/stem.2428.

Desai, T. J., Brownfield, D. G. and Krasnow, M. A. (2014) 'Alveolar progenitor and stem cells in lung development, renewal and cancer', *Nature*, 507(7491), pp. 190–194. doi: 10.1038/nature12930.

Deterding, R. R. *et al.* (1994) 'Regulation of surfactant protein D expression by glucocorticoids in vitro and in vivo.', *American Journal of Respiratory Cell and Molecular Biology*, 10(1), pp. 30–37. doi: 10.1165/ajrcmb.10.1.8292379.

Dluholucký, S., Babic, J. and Taufer, I. (1976) 'Reduction of incidence and mortality of respiratory distress syndrome by adminstration of hydrocortisone to mother.', *Archives of disease in childhood*, 51(6), pp. 420–423.

Domyan, E. T. *et al.* (2011) 'Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2.', *Development (Cambridge, England)*. Company of Biologists, 138(5), pp. 971–81. doi: 10.1242/dev.053694.

Doss, M. X. and Sachinidis, A. (2019) 'Current Challenges of iPSC-Based Disease Modeling and Therapeutic Implications', *Cells 2019, Vol. 8, Page 403*. Multidisciplinary Digital Publishing Institute, 8(5), p. 403. doi: 10.3390/CELLS8050403.

Du, Y. *et al.* (2019) 'Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung', *American journal of physiology. Lung cellular and molecular physiology*. NLM (Medline), 317(3), pp. L347–L360. doi: 10.1152/ajplung.00475.2018.

Dubey, R. *et al.* (2020) 'R-spondins engage heparan sulfate proteoglycans to potentiate WNT signaling.', *eLife*. Elife, 9. doi: 10.7554/eLife.54469.

Dye, B. R. *et al.* (2015) 'In vitro generation of human pluripotent stem cell derived lung organoids', *eLife*. eLife Sciences Publications Limited, 4, p. e05098. doi: 10.7554/eLife.05098.

Dye, B. R. *et al.* (2016) 'A bioengineered niche promotes in vivo engraftment and maturation of pluripotent stem cell derived human lung organoids', *eLife*. eLife Sciences Publications Limited, 5, p. e19732. doi: 10.7554/eLife.19732.

Dye, B. R., Miller, A. J. and Spence, J. R. (2016) 'How to Grow a Lung: Applying Principles of Developmental Biology to Generate Lung Lineages from Human Pluripotent Stem Cells', *Current Pathobiology Reports*, 4(2), pp. 47–57. doi: 10.1007/s40139-016-0102-x.

Endale, M. *et al.* (2017) 'Temporal, spatial, and phenotypical changes of PDGFRα expressing fibroblasts during late lung development', *Developmental Biology*. Academic Press, 425(2), pp. 161–175. doi: 10.1016/J.YDBIO.2017.03.020.

Erben, L. *et al.* (2018) 'A Novel Ultrasensitive In Situ Hybridization Approach to Detect Short Sequences and Splice Variants with Cellular Resolution', *Molecular Neurobiology*. Humana Press Inc., 55(7), pp. 6169–6181. doi: 10.1007/s12035-017-0834-6.

Ferkol, T. and Schraufnagel, D. (2014) 'The Global Burden of Respiratory Disease', *Annals of the American Thoracic Society*, 11(3), pp. 404–406. doi: 10.1513/AnnalsATS.201311-405PS.

Frank, D. B. *et al.* (2016) 'Emergence of a Wave of Wnt Signaling that Regulates Lung Alveologenesis by Controlling Epithelial Self-Renewal and Differentiation', *Cell Reports*. Cell Press, 17(9), pp. 2312–2325. doi: 10.1016/J.CELREP.2016.11.001.

Frank, D. B. *et al.* (2019) 'Early lineage specification defines alveolar epithelial ontogeny in the murine lung', *Proceedings of the National Academy of Sciences*, 116(10), pp. 4362–4371. doi: 10.1073/pnas.1813952116.

Frum, T. and Spence, J. R. (2021) 'hPSC-derived organoids: models of human development and disease', *Journal of Molecular Medicine (Berlin, Germany)*. Springer, 99(4), p. 463. doi: 10.1007/S00109-020-01969-W.

Fumoto, K. *et al.* (2019) 'Mark1 regulates distal airspace expansion through type I pneumocyte flattening in lung development.', *Journal of Cell Science*, 132(24), p. jcs235556.

Gadue, P. *et al.* (2006) 'Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(45), pp. 16806–11. doi: 10.1073/pnas.0603916103.

Ghaedi, M. *et al.* (2013) 'Human iPS cellâ€"derived alveolar epithelium repopulates lung extracellular matrix', *Journal of Clinical Investigation*, 123(11), pp. 4950–4962. doi: 10.1172/JCI68793.

Gilpin, S. E. *et al.* (2014) 'Enhanced lung epithelial specification of human induced pluripotent stem cells on decellularized lung matrix.', *The Annals of thoracic surgery*, 98(5), pp. 1721--9--discussion 1729.

Gontan, C. *et al.* (2008) 'Sox2 is important for two crucial processes in lung development: Branching morphogenesis and epithelial cell differentiation', *Developmental Biology*, 317(1), pp. 296–309. doi: 10.1016/j.ydbio.2008.02.035.

Goodwin, K. *et al.* (2022) 'Patterning the embryonic pulmonary mesenchyme', *iScience*. Elsevier, 25(3), p. 103838. doi: 10.1016/J.ISCI.2022.103838.

Goss, A. M. *et al.* (2009) 'Wnt2/2b and β-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut', *Developmental cell*. NIH Public Access, 17(2), p. 290.

doi: 10.1016/J.DEVCEL.2009.06.005.

Goss, A. M. *et al.* (2011) 'Wnt2 signaling is necessary and sufficient to activate the airway smooth muscle program in the lung by regulating myocardin/Mrtf-B and Fgf10 expression', *Developmental Biology*. NIH Public Access, 356(2), p. 541. doi: 10.1016/J.YDBIO.2011.06.011.

Gotoh, S. *et al.* (2014) 'Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells.', *Stem cell reports*. Elsevier, 3(3), pp. 394–403. doi: 10.1016/j.stemcr.2014.07.005.

Gouveia, L. *et al.* (2020) 'Lung developmental arrest caused by PDGF-A deletion: consequences for the adult mouse lung', *American Journal of Physiology-Lung Cellular and Molecular Physiology*, p. ajplung.00295.2019. doi: 10.1152/ajplung.00295.2019.

Goyette, P. *et al.* (2009) 'Molecular pathogenesis of inflammatory bowel disease: Genotypes, phenotypes and personalized medicine', *https://doi.org/10.1080/07853890701197615*. Taylor & Francis, 39(3), pp. 177–199. doi: 10.1080/07853890701197615.

Green, J. *et al.* (2016) 'Diversity of Interstitial Lung Fibroblasts Is Regulated by Platelet-Derived Growth Factor Receptor α Kinase Activity.', *American journal of respiratory cell and molecular biology*. American Thoracic Society, 54(4), pp. 532–45. doi: 10.1165/rcmb.2015-0095OC.

Green, M. D. *et al.* (2011) 'Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells.', *Nature biotechnology*. NIH Public Access, 29(3), pp. 267–72. doi: 10.1038/nbt.1788.

Güney, T. G. *et al.* (2021) 'Epithelial-stromal cell interactions and extracellular matrix mechanics drive the formation of airway-mimetic tubular morphology in lung organoids', *iScience*. iScience, 24(9). doi: 10.1016/J.ISCI.2021.103061.

Guo, M. *et al.* (2019) 'Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth', *Nature Communications*. Nature Publishing Group, 10(1), p. 37. doi: 10.1038/s41467-018-07770-1.

Hafemeister, C. and Satija, R. (2019) 'Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression', *Genome Biology*. BioMed Central Ltd., 20(1), pp. 1–15. doi: 10.1186/S13059-019-1874-1/FIGURES/6.

Haghverdi, L. *et al.* (2016) 'Diffusion pseudotime robustly reconstructs lineage branching', *Nature Methods*. Nature Publishing Group, 13(10), pp. 845–848. doi: 10.1038/nmeth.3971.

Han, L. *et al.* (2020) 'Single cell transcriptomics identifies a signaling network coordinating endoderm and mesoderm diversification during foregut organogenesis', *Nature Communications 2020 11:1*. Nature Publishing Group, 11(1), pp. 1–16. doi: 10.1038/s41467-020-17968-x.

Hao, Y. *et al.* (2021) 'Integrated analysis of multimodal single-cell data', *Cell*. Cell, 184(13), pp. 3573-3587.e29. doi: 10.1016/J.CELL.2021.04.048.

Harris-Johnson, K. S. *et al.* (2009) 'beta-Catenin promotes respiratory progenitor identity in mouse foregut.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(38), pp. 16287–92. doi: 10.1073/pnas.0902274106.

Hau, J. (2008) 'Animal models for human diseases: An overview', *Source Book of Models for Biomedical Research*. Humana Press, pp. 3–8. doi: 10.1007/978-1-59745-285-4_1/COVER/.

Hawkins, F. *et al.* (2017) 'Prospective isolation of NKX2-1–expressing human lung progenitors derived from pluripotent stem cells', *The Journal of Clinical Investigation*. American Society for Clinical Investigation, 127(6), pp. 2277–2294. doi: 10.1172/JCI89950.

Hawkins, F. J. *et al.* (2021) 'Derivation of Airway Basal Stem Cells from Human Pluripotent Stem Cells', *Cell stem cell*. NIH Public Access, 28(1), p. 79. doi: 10.1016/J.STEM.2020.09.017.

He, H. *et al.* (2021) 'IGF1R controls mechanosignaling in myofibroblasts required for pulmonary alveologenesis', *JCI insight*. JCI Insight, 6(6). doi: 10.1172/JCI.INSIGHT.144863.

Heemskerk, I. *et al.* (2019) 'Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells', *eLife*. eLife Sciences Publications Ltd, 8. doi: 10.7554/ELIFE.40526.

Hein, Renee F.C. *et al.* (2022) 'R-SPONDIN2+ mesenchymal cells form the bud tip progenitor niche during human lung development', *Developmental Cell*. Cell Press. doi: 10.1016/J.DEVCEL.2022.05.010.

Hein, Renee F. C. *et al.* (2022) 'Stable iPSC-derived NKX2-1+ lung bud tip progenitor organoids give rise to airway and alveolar cell types', *Development*. The Company of Biologists, 149(20). doi: 10.1242/DEV.200693.

Heo, H.-R. *et al.* (2019) 'Human pluripotent stem cell-derived alveolar epithelial cells are alternatives for in vitro pulmotoxicity assessment.', *Scientific Reports*, 9(1), pp. 505–511.

Hines, E. A. and Sun, X. (2014) 'Tissue Crosstalk in Lung Development', *Journal of Cellular Biochemistry*. John Wiley & Sons, Ltd, 115(9), pp. 1469–1477. doi: 10.1002/jcb.24811.

Hirashima, T., Iwasa, Y. and Morishita, Y. (2009) 'Mechanisms for split localization of *Fgf10* expression in early lung development', *Developmental Dynamics*. John Wiley & Sons, Ltd, 238(11), pp. 2813–2822. doi: 10.1002/dvdy.22108.

Hofer, M. and Lutolf, M. P. (2021) 'Engineering organoids', *Nature Reviews Materials 2021 6:5*. Nature Publishing Group, 6(5), pp. 402–420. doi: 10.1038/s41578-021-00279-y.

Hogan, B. L. M. and Yinalina, J. M. (1998) *Epithelial/mesenchymal interactions and branching morphogenesis of the lung*, *Current Opinion in Genetics & Development*. Available at: http://biomednet.com/elecref/0959437X00800481.

Holloway, E. M., Wu, J. H., *et al.* (2020) 'Differentiation of Human Intestinal Organoids with Endogenous Vascular Endothelial Cells', *Developmental Cell*. Cell Press, 54(4), pp. 516-528.e7. doi: 10.1016/j.devcel.2020.07.023.

Holloway, E. M., Czerwinski, M., *et al.* (2020) 'Mapping Development of the Human Intestinal Niche at Single-Cell Resolution', *Cell Stem Cell*. Cell Press. doi: 10.1016/j.stem.2020.11.008.

Holloway, E. M., Capeling, M. M. and Spence, J. R. (2019) 'Biologically inspired approaches to enhance human organoid complexity.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 146(8), p. dev166173. doi: 10.1242/dev.166173.

Hor, P. *et al.* (2020) 'Efficient Generation and Transcriptomic Profiling of Human iPSC-Derived Pulmonary Neuroendocrine Cells', *iScience*. iScience, 23(5). doi: 10.1016/J.ISCI.2020.101083.

Howden, S. E. *et al.* (2019) 'Reporter-based fate mapping in human kidney organoids confirms nephron lineage relationships and reveals synchronous nephron formation', *EMBO reports*. EMBO Press, p. e47483. doi: 10.15252/embr.201847483.

Hrycaj, S. M. *et al.* (2015) 'Hox5 Genes Regulate the Wnt2/2b-Bmp4-Signaling Axis during Lung Development', *Cell Reports*. Cell Press, 12(6), pp. 903–912. doi: 10.1016/J.CELREP.2015.07.020.

Huang, S. X. L. *et al.* (2014) 'Efficient generation of lung and airway epithelial cells from human pluripotent stem cells', *Nature Biotechnology*. Nature Publishing Group, 32(1), pp. 84–91. doi: 10.1038/nbt.2754.

Huang, S. X. L. *et al.* (2015) 'The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells', *Nature Protocols*. Nature Publishing Group, 10(3), pp. 413–425. doi: 10.1038/nprot.2015.023.

Hurley, K. *et al.* (2020) 'Reconstructed Single-Cell Fate Trajectories Define Lineage Plasticity Windows during Differentiation of Human PSC-Derived Distal Lung Progenitors', *Cell Stem Cell*. Cell Press. doi: 10.1016/J.STEM.2019.12.009.

Husain, A. N., Siddiqui, N. H. and Stocker, J. T. (1998) 'Pathology of arrested acinar development in postsurfactant bronchopulmonary dysplasia.', *Human pathology*, 29(7), pp. 710–717.

Irvin, C. G. and Bates, J. H. T. (2003) 'Measuring the lung function in the mouse: The challenge of size', *Respiratory Research*. doi: 10.1186/rr199.

Isago, H. *et al.* (2019) 'The Epithelial Expressions of YAP and TAZ are Sequentially Required in Lung Development.', *American Journal of Respiratory Cell and Molecular Biology*.

Jacob, A. *et al.* (2017) 'Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells', *Cell Stem Cell*. Cell Press, 21(4), pp. 472-488.e10. doi: 10.1016/J.STEM.2017.08.014.

Jacob, A. *et al.* (2019) 'Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells', *Nature Protocols 2019 14:12*. Nature Publishing Group, 14(12), pp. 3303–3332. doi: 10.1038/s41596-019-0220-0.

Jacobs, I. J., Ku, W.-Y. and Que, J. (2012) 'Genetic and cellular mechanisms regulating anterior foregut and esophageal development', *Developmental Biology*. Academic Press, 369(1), pp. 54–64. doi: 10.1016/J.YDBIO.2012.06.016.

Jaks, V. *et al.* (2008) 'Lgr5 marks cycling, yet long-lived, hair follicle stem cells', *Nature Genetics*. Nat Genet, 40(11), pp. 1291–1299. doi: 10.1038/ng.239.

Jha, R. *et al.* (2017) 'Downregulation of LGR5 Expression Inhibits Cardiomyocyte Differentiation and Potentiates Endothelial Differentiation from Human Pluripotent Stem Cells', *Stem cell reports*. Stem Cell Reports, 9(2), pp. 513–527. doi: 10.1016/J.STEMCR.2017.07.006.

Jiang, J. X. *et al.* (2021) 'A new platform for high-throughput therapy testing on iPSC-derived lung progenitor cells from cystic fibrosis patients', *Stem Cell Reports*. Cell Press, 16(11), pp. 2825–2837. doi: 10.1016/J.STEMCR.2021.09.020.

Jo, K. *et al.* (2022) 'Efficient differentiation of human primordial germ cells through geometric control reveals a key role for Nodal signaling', *eLife*. eLife Sciences Publications, Ltd, 11. doi: 10.7554/ELIFE.72811.

Jobe, A. J. (1999) 'The new BPD: an arrest of lung development.', *Pediatric research*, 46(6), pp. 641–643.

Jones, M. R. *et al.* (2019) 'A Comprehensive Analysis of Fibroblast Growth Factor Receptor 2b Signaling on Epithelial Tip Progenitor Cells During Early Mouse Lung Branching Morphogenesis', *Frontiers in Genetics*. Frontiers, 9, p. 746. doi: 10.3389/fgene.2018.00746.

Jones, T. R. *et al.* (2008) 'CellProfiler Analyst: Data exploration and analysis software for complex image-based screens', *BMC Bioinformatics*. BioMed Central, 9(1), p. 482. doi: 10.1186/1471-2105-9-482.

Kalucka, J. *et al.* (2020) 'Single-Cell Transcriptome Atlas of Murine Endothelial Cells', *Cell*, 180(4), pp. 764-779.e20. doi: 10.1016/j.cell.2020.01.015.

Kapoor, A., Li, H. J. and Leiter, A. B. (2007) 'Intestinal Development: The Many Faces of Wnt Signaling', *Gastroenterology*. Elsevier, 133(2), pp. 710–712. doi: 10.1053/J.GASTRO.2007.06.029.

Kaur, R. and Chupp, G. (2019) 'Phenotypes and endotypes of adult asthma: Moving toward precision medicine', *Journal of Allergy and Clinical Immunology*. Mosby, 144(1), pp. 1–12. doi: 10.1016/J.JACI.2019.05.031.

Kearns, N. A. *et al.* (2013) 'Generation of organized anterior foregut epithelia from pluripotent stem cells using small molecules', *Stem Cell Research*. Elsevier, 11(3), pp. 1003–1012. doi: 10.1016/J.SCR.2013.06.007.

Kida, K. and Thurlbeck, W. M. (1980) 'The effects of beta-aminopropionitrile on the growing rat lung.', *The American Journal of Pathology*, 101(3), pp. 693–710.

Kim, H. Y. *et al.* (2015) 'Localized Smooth Muscle Differentiation Is Essential for Epithelial Bifurcation during Branching Morphogenesis of the Mammalian Lung', *Developmental Cell*. Cell Press, 34(6), pp. 719–726. doi: 10.1016/J.DEVCEL.2015.08.012.

Kim, J., Koo, B. K. and Knoblich, J. A. (2020) 'Human organoids: model systems for human biology and medicine', *Nature Reviews Molecular Cell Biology 2020 21:10*. Nature Publishing Group, 21(10), pp. 571–584. doi: 10.1038/s41580-020-0259-3.

Kim, J. Y. *et al.* (2022) 'Review of the Current Trends in Clinical Trials Involving Induced Pluripotent Stem Cells', *Stem Cell Reviews and Reports*. Springer, 18(1), pp. 142–154. doi: 10.1007/S12015-021-10262-3/FIGURES/4.

Kishimoto, K. *et al.* (2018) 'Synchronized mesenchymal cell polarization and differentiation shape the formation of the murine trachea and esophagus', *Nature Communications*. Nature Publishing Group, 9(1), pp. 1–13. doi: 10.1038/s41467-018-05189-2.

Kishimoto, K. *et al.* (2020) 'Bidirectional Wnt signaling between endoderm and mesoderm confers tracheal identity in mouse and human cells', *Nature Communications 2020 11:1*. Nature Publishing Group, 11(1), pp. 1–12. doi: 10.1038/s41467-020-17969-w.

Kishimoto, K. *et al.* (2022) 'Directed differentiation of human pluripotent stem cells into diverse organ-specific mesenchyme of the digestive and respiratory systems', *Nature Protocols 2022*. Nature Publishing Group, pp. 1–21. doi: 10.1038/s41596-022-00733-3.

Kong, W. *et al.* (2020) 'CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution', *Nature Protocols 2020 15:3*. Nature Publishing Group, 15(3),

pp. 750–772. doi: 10.1038/s41596-019-0247-2.

Konishi, S. *et al.* (2016) 'Directed Induction of Functional Multi-ciliated Cells in Proximal Airway Epithelial Spheroids from Human Pluripotent Stem Cells', *Stem Cell Reports*. The Authors, 6(1), pp. 18–25. doi: 10.1016/j.stemcr.2015.11.010.

Korogi, Y. *et al.* (2019) 'In Vitro Disease Modeling of Hermansky-Pudlak Syndrome Type 2 Using Human Induced Pluripotent Stem Cell-Derived Alveolar Organoids', *Stem Cell Reports*. Cell Press, 12(3), pp. 431–440. doi: 10.1016/J.STEMCR.2019.01.014.

Kreitzer, F. R. *et al.* (2013) 'A robust method to derive functional neural crest cells from human pluripotent stem cells', *American Journal of Stem Cells*. e-Century Publishing Corporation, 2(2), p. 119. Available at: /pmc/articles/PMC3708511/ (Accessed: 15 February 2022).

Kubo, A. *et al.* (2004) 'Development of definitive endoderm from embryonic stem cells in culture.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 131(7), pp. 1651–62. doi: 10.1242/dev.01044.

Kuwahara, A. *et al.* (2020) 'Delineating the early transcriptional specification of the mammalian trachea and esophagus', *eLife*. eLife Sciences Publications Ltd, 9, pp. 1–23. doi: 10.7554/eLife.55526.

Lamprecht, M. R., Sabatini, D. M. and Carpenter, A. E. (2007) 'CellProfiler™: Free, versatile software for automated biological image analysis', *BioTechniques*. Biotechniques, 42(1), pp. 71–75. doi: 10.2144/000112257.

Laresgoiti, U. *et al.* (2016) 'Lung epithelial tip progenitors integrate glucocorticoid- and STAT3-mediated signals to control progeny fate', *Development*, 143(20), pp. 3686–3699. doi: 10.1242/dev.134023.

de Lau, W. B. M., Snel, B. and Clevers, H. C. (2012) 'The R-spondin protein family.', *Genome biology*. BioMed Central, 13(3), p. 242. doi: 10.1186/gb-2012-13-3-242.

de Lau, W. *et al.* (2014) 'The R-spondin/Lgr5/Rnf43 module: Regulator of Wnt signal strength', *Genes and Development*. Genes Dev, pp. 305–316. doi: 10.1101/gad.235473.113.

Lazarus, A. *et al.* (2011) 'A perfusion-independent role of blood vessels in determining branching stereotypy of lung airways', *Development*, 138(11), pp. 2359–2368. doi: 10.1242/dev.060723.

Lazzaro, D. *et al.* (1991) 'The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain.', *Development (Cambridge, England)*, 113(4), pp. 1093–104. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1811929 (Accessed: 10 September 2019).

Lee, A. S. *et al.* (2013) 'Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies', *Nature Medicine 2013 19:8*. Nature Publishing Group, 19(8), pp. 998–1004. doi: 10.1038/nm.3267.

Lee, H. *et al.* (2020) 'R-spondins are BMP receptor antagonists in Xenopus early embryonic development', *Nature Communications*. Nature Research, 11(1). doi: 10.1038/s41467-020-19373-w.

Leeman, K. T. *et al.* (2019) 'Mesenchymal Stem Cells Increase Alveolar Differentiation in Lung Progenitor Organoid Cultures', *Scientific Reports*. Nature Publishing Group, 9(1), p. 6479. doi: 10.1038/s41598-019-42819-1.

Leibel, S. L. *et al.* (2019) 'Reversal of Surfactant Protein~B Deficiency in Patient Specific Human Induced Pluripotent Stem Cell Derived Lung Organoids by Gene Therapy.', *Scientific Reports*, 9(1), pp. 13415–13450.

Leibel, S. L. *et al.* (2020) 'Generation of Complete Multi-Cell Type Lung Organoids From Human Embryonic and Patient-Specific Induced Pluripotent Stem Cells for Infectious Disease Modeling and Therapeutics Validation.', *Current protocols in stem cell biology*. Curr Protoc Stem Cell Biol, 54(1), p. e118. doi: 10.1002/cpsc.118.

Leopold, J. A., Maron, B. A. and Loscalzo, J. (2020) 'The application of big data to cardiovascular disease: paths to precision medicine', *The Journal of Clinical Investigation*. American Society for Clinical Investigation, 130(1), pp. 29–38. doi: 10.1172/JCI129203.

Li, A. *et al.* (2016) 'Mesodermal ALK5 controls lung myofibroblast versus lipofibroblast cell fate', *BMC Biology*. BioMed Central, 14(1), p. 19. doi: 10.1186/s12915-016-0242-9.

Li, C. *et al.* (2015) ‘Progenitors of Secondary Crest Myofibroblasts Are Developmentally Committed in Early Lung Mesoderm’, *STEM CELLS*, 33(3), pp. 999–1012. doi: 10.1002/stem.1911.

Li, M. *et al.* (2008) ‘Mesodermal Deletion of Transforming Growth Factor-β Receptor II Disrupts Lung Epithelial Morphogenesis: CROSS-TALK BETWEEN TGF-β AND SONIC HEDGEHOG PATHWAYS*’, *The Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 283(52), p. 36257. doi: 10.1074/JBC.M806786200.

Li, R. *et al.* (2018a) ‘Pdgfra marks a cellular lineage with distinct contributions to myofibroblasts in lung maturation and injury response.’, *eLife*. Elife, 7. doi: 10.7554/eLife.36865.

Li, R. *et al.* (2018b) ‘Pdgfra marks a cellular lineage with distinct contributions to myofibroblasts in lung maturation and injury response’, *eLife*, 7. doi: 10.7554/eLife.36865.

Li, R. *et al.* (2020) ‘Myofibroblast contraction is essential for generating and regenerating the gas-exchange surface’, *Journal of Clinical Investigation*. doi: 10.1172/JCI132189.

Li, S. *et al.* (2016) ‘Foxp transcription factors suppress a non-pulmonary gene expression program to permit proper lung development’, *Developmental biology*. NIH Public Access, 416(2), p. 338. doi: 10.1016/J.YDBIO.2016.06.020.

Liggins, G. C. and Howie, R. N. (1972) ‘A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants.’, *Pediatrics*, 50(4), pp. 515–25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/4561295 (Accessed: 4 January 2020).

Lim, K. *et al.* (2021) ‘Acquisition of alveolar fate and differentiation competence by human fetal lung epithelial progenitor cells’, *bioRxiv*. Cold Spring Harbor Laboratory, p. 2021.06.30.450501. doi: 10.1101/2021.06.30.450501.

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 (Cambridge, England)*, 124(20), pp. 3943–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9374392 (Accessed: 22 August 2019).

Litingtung, Y. *et al.* (1998) ‘Sonic hedgehog is essential to foregut development’, *Nature*

Genetics. Nature Publishing Group, 20(1), pp. 58–61. doi: 10.1038/1717.

Little, D. R., Gerner-Mauro, K. N., Flodby, P., Crandall, E. D., Borok, Z., Spence, J. R., *et al.* (2019) 'Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids', *eLife*. eLife Sciences Publications Limited, 59(1), p. e26575. doi: 10.7554/eLife.26575.

Little, D. R., Gerner-Mauro, K. N., Flodby, P., Crandall, E. D., Borok, Z., Akiyama, H., *et al.* (2019) 'Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1', *Proceedings of the National Academy of Sciences*, 116(41), pp. 20545–20555. doi: 10.1073/pnas.1906663116.

Little, D. R. *et al.* (2021) 'Differential chromatin binding of the lung lineage transcription factor NKX2-1 resolves opposing murine alveolar cell fates in vivo', *Nature Communications 2021 12:1*. Nature Publishing Group, 12(1), pp. 1–18. doi: 10.1038/s41467-021-22817-6.

Loh, K. M. *et al.* (2014) 'Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations.', *Cell stem cell*. NIH Public Access, 14(2), pp. 237–52. doi: 10.1016/j.stem.2013.12.007.

De Los Angeles, A. and Tunbridge, E. M. (2020) 'Benchmarking pluripotent stem cell-derived organoid models', *Experimental neurology*. Exp Neurol, 330. doi: 10.1016/J.EXPNEUROL.2020.113333.

Loscertales, M. *et al.* (2016) 'Type IV collagen drives alveolar epithelial–endothelial association and the morphogenetic movements of septation', *BMC Biology*. BioMed Central, 14(1), p. 59. doi: 10.1186/s12915-016-0281-2.

Luo, Y. *et al.* (2018) 'Spatial and temporal changes in extracellular elastin and laminin distribution during lung alveolar development', *Scientific Reports*. Nature Publishing Group, 8(1), p. 8334. doi: 10.1038/s41598-018-26673-1.

Madrid, M. *et al.* (2021) 'Autologous Induced Pluripotent Stem Cell–Based Cell Therapies: Promise, Progress, and Challenges', *Current Protocols*. John Wiley & Sons, Ltd, 1(3), p. e88. doi: 10.1002/CPZ1.88.

Mahoney, J. E. *et al.* (2014) 'The hippo pathway effector Yap controls patterning and

differentiation of airway epithelial progenitors.', *Developmental cell*. Elsevier, 30(2), pp. 137–50. doi: 10.1016/j.devcel.2014.06.003.

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), pp. 2157–2166. doi: 10.1242/dev.01795.

Mak, I. W. Y., Evaniew, N. and Ghert, M. (2014) 'Lost in translation: animal models and clinical trials in cancer treatment', *American Journal of Translational Research*. e-Century Publishing Corporation, 6(2), p. 114. Available at: /pmc/articles/PMC3902221/ (Accessed: 20 July 2022).

Makita, R. *et al.* (2008) 'Multiple renal cysts, urinary concentration defects, and pulmonary emphysematous changes in mice lacking TAZ.', *American journal of physiology. Renal physiology*, 294(3), pp. F542--53.

Manwani, N. *et al.* (2010) 'Reduced Viability of Mice with Lung Epithelial-Specific Knockout of Glucocorticoid Receptor', *American Journal of Respiratory Cell and Molecular Biology*, 43(5), pp. 599–606. doi: 10.1165/rcmb.2009-0263OC.

Marson, F. A. L., Bertuzzo, C. S. and Ribeiro, J. D. (2017) 'Personalized or precision medicine? The example of cystic fibrosis', *Frontiers in Pharmacology*. Frontiers Media S.A., 8(JUN), p. 390. doi: 10.3389/FPHAR.2017.00390/BIBTEX.

Martyn, I. *et al.* (2018) 'Self-organization of a human organizer by combined Wnt and Nodal signalling', *Nature*. Nature Publishing Group, 558(7708), pp. 132–135. doi: 10.1038/s41586-018-0150-y.

Matsuno, K. *et al.* (2016) 'Redefining definitive endoderm subtypes by robust induction of human induced pluripotent stem cells', *Differentiation*. Elsevier, 92(5), pp. 281–290. doi: 10.1016/J.DIFF.2016.04.002.

Matthay, M. A. *et al.* (2020) 'Phenotypes and personalized medicine in the acute respiratory distress syndrome', *Intensive Care Medicine*. Springer Science and Business Media Deutschland GmbH, 46(12), pp. 2136–2152. doi: 10.1007/S00134-020-06296-9/FIGURES/4.

McCauley, Katherine B. *et al.* (2017a) 'Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling', *Cell Stem*

Cell. Cell Press, 20(6), pp. 844-857.e6. doi: 10.1016/J.STEM.2017.03.001.

McCauley, Katherine B. *et al.* (2017b) 'Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling', *Cell stem cell*. Cell Stem Cell, 20(6), pp. 844-857.e6. doi: 10.1016/J.STEM.2017.03.001.

McCauley, Katherine B *et al.* (2017) 'Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling', *Cell Stem Cell*, 20(6), pp. 844--857.e6. doi: 10.1016/j.stem.2017.03.001.

McCauley, K. B. *et al.* (2018) 'Single-Cell Transcriptomic Profiling of Pluripotent Stem Cell-Derived SCGB3A2+ Airway Epithelium', *Stem Cell Reports*, 10(5), pp. 1579–1595. doi: 10.1016/j.stemcr.2018.03.013.

McCracken, K. W. *et al.* (2011) 'Generating human intestinal tissue from pluripotent stem cells in vitro', *Nature Protocols*. Nature Publishing Group, 6(12), pp. 1920–1928. doi: 10.1038/nprot.2011.410.

McCracken, K. W., Catá, E. M., Crawford, C. M., Sinagoga, K. L., Schumacher, M., Rockich, B. E., Tsai, Y. H., *et al.* (2014) 'Modeling human development and disease in pluripotent stem cell-derived gastric organoids', *Nature*. NIH Public Access, 516(7531), p. 400. doi: 10.1038/NATURE13863.

McCracken, K. W., Catá, E. M., Crawford, C. M., Sinagoga, K. L., Schumacher, M., Rockich, B. E., Tsai, Y.-H., *et al.* (2014) 'Modelling human development and disease in pluripotent stem-cell-derived gastric organoids', *Nature*. Nature Publishing Group, 516(7531), pp. 400–404. doi: 10.1038/nature13863.

McCulley, D., Wienhold, M. and Sun, X. (2015) 'The Pulmonary Mesenchyme Directs Lung Development', *Current opinion in genetics & development*. NIH Public Access, 32, p. 98. doi: 10.1016/J.GDE.2015.01.011.

McGonigle, P. and Ruggeri, B. (2014) 'Animal models of human disease: Challenges in enabling translation', *Biochemical Pharmacology*. Elsevier, 87(1), pp. 162–171. doi: 10.1016/J.BCP.2013.08.006.

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', *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 291(12), pp. 1649–1661. doi: 10.1002/ar.20764.

McGowan, S. E. and Torday, J. S. (1997) 'THE PULMONARY LIPOFIBROBLAST (LIPID INTERSTITIAL CELL) AND ITS CONTRIBUTIONS TO ALVEOLAR DEVELOPMENT', *Annual Review of Physiology*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA , 59(1), pp. 43–62. doi: 10.1146/annurev.physiol.59.1.43.

McInnes, L., Healy, J. and Melville, J. (2018) 'UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction', *arXiv*. arXiv. Available at: http://arxiv.org/abs/1802.03426 (Accessed: 13 January 2021).

McQualter, J. L. *et al.* (2013) 'TGF-β signaling in stromal cells acts upstream of FGF-10 to regulate epithelial stem cell growth in the adult lung', *Stem Cell Research*. Elsevier, 11(3), pp. 1222–1233. doi: 10.1016/J.SCR.2013.08.007.

Metzger, R. J. *et al.* (2008) 'The branching programme of mouse lung development', *Nature*. Nature Publishing Group, 453(7196), pp. 745–750. doi: 10.1038/nature07005.

Mihara, E. *et al.* (2016) 'Active and water-soluble form of lipidated wnt protein is maintained by a serum glycoprotein afamin/α-albumin', *eLife*. eLife Sciences Publications Ltd, 5(FEBRUARY2016). doi: 10.7554/ELIFE.11621.

Miller, A. J. *et al.* (2018) 'In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells.', *Stem cell reports*. Elsevier, 10(1), pp. 101–119. doi: 10.1016/j.stemcr.2017.11.012.

Miller, A. J. *et al.* (2019) 'Generation of lung organoids from human pluripotent stem cells in vitro', *Nature Protocols*. Nature Publishing Group, 14(2), pp. 518–540. doi: 10.1038/s41596-018-0104-8.

Miller, A. J. *et al.* (2020) 'In Vitro and In Vivo Development of the Human Airway at Single-Cell Resolution', *Developmental Cell*. Elsevier, 0(0). doi: 10.1016/j.devcel.2020.01.033.

Miller, A. J. and Spence, J. R. (2017) 'In vitro models to study human lung development, disease and homeostasis', *Physiology*. American Physiological Society, pp. 246–260. doi:

10.1152/physiol.00041.2016.

Miller, M. F. *et al.* (2012) 'Wnt ligands signal in a cooperative manner to promote foregut organogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 109(38), pp. 15348–15353. doi: 10.1073/pnas.1201583109.

Min, H. *et al.* (1998) 'Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to Drosophila branchless.', *Genes & development*. Cold Spring Harbor Laboratory Press, 12(20), pp. 3156–61. doi: 10.1101/gad.12.20.3156.

Minoo, P. *et al.* (1995) 'TTF-1 Regulates Lung Epithelial Morphogenesis', *Developmental Biology*, 172(2), pp. 694–698. doi: 10.1006/dbio.1995.8080.

Minoo, P. *et al.* (1999) 'Defects in Tracheoesophageal and Lung Morphogenesis inNkx2.1(−/−) Mouse Embryos', *Developmental Biology*. Academic Press, 209(1), pp. 60–71. doi: 10.1006/DBIO.1999.9234.

Mitani, A. *et al.* (2009) 'Transcriptional Coactivator with PDZ-binding Motif Is Essential for Normal Alveolarization in Mice', *American Journal of Respiratory and Critical Care Medicine*, 180(4), pp. 326–338. doi: 10.1164/rccm.200812-1827OC.

Montoro, D. T. *et al.* (2018) 'A revised airway epithelial hierarchy includes CFTR-expressing ionocytes', *Nature*. Nature Publishing Group, 560(7718), pp. 319–324. doi: 10.1038/s41586-018-0393-7.

Moreira, A. *et al.* (2022) 'Advanced In Vitro Lung Models for Drug and Toxicity Screening: The Promising Role of Induced Pluripotent Stem Cells', *Advanced Biology*. John Wiley & Sons, Ltd, 6(2), p. 2101139. doi: 10.1002/ADBI.202101139.

Morrisey, E. E. and Hogan, B. L. M. (2010) 'Preparing for the first breath: genetic and cellular mechanisms in lung development.', *Developmental cell*. NIH Public Access, 18(1), pp. 8–23. doi: 10.1016/j.devcel.2009.12.010.

Morrison, J. C. *et al.* (1978) 'Injection of corticosteroids into mother to prevent neonatal respiratory distress syndrome', *American Journal of Obstetrics and Gynecology*. Elsevier, 131(4), pp. 358–366. doi: 10.1016/0002-9378(78)90408-8.

Motoyama, J. *et al.* (1998) 'Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus', *Nature Genetics*. Nature Publishing Group, 20(1), pp. 54–57. doi: 10.1038/1711.

Mou, H. *et al.* (2012) 'Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs.', *Cell stem cell*. NIH Public Access, 10(4), pp. 385–97. doi: 10.1016/j.stem.2012.01.018.

Mou, H. *et al.* (2016) 'Dual SMAD Signaling Inhibition Enables Long-Term Expansion of Diverse Epithelial Basal Cells', *Cell Stem Cell*. Cell Press, 19(2), pp. 217–231. doi: 10.1016/J.STEM.2016.05.012.

Mucenski, M. L. *et al.* (2003) 'beta-Catenin is required for specification of proximal/distal cell fate during lung morphogenesis.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 278(41), pp. 40231–8. doi: 10.1074/jbc.M305892200.

Nanki, K. *et al.* (2018) 'Divergent Routes toward Wnt and R-spondin Niche Independency during Human Gastric Carcinogenesis', *Cell*. Cell Press, 174(4), pp. 856-869.e17. doi: 10.1016/J.CELL.2018.07.027.

Nantie, L. B. *et al.* (2018) 'Lats1/2 inactivation reveals Hippo function in alveolar type I cell differentiation during lung transition to air breathing.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 145(21), p. dev163105. doi: 10.1242/dev.163105.

Negretti, N. M. *et al.* (2021) 'A single-cell atlas of mouse lung development', *Development (Cambridge, England)*. Development, 148(24). doi: 10.1242/DEV.199512.

Nelson, C. M. *et al.* (2017) 'Microfluidic chest cavities reveal that transmural pressure controls the rate of lung development', *Development (Cambridge)*. Company of Biologists Ltd, 144(23), pp. 4328–4335. doi: 10.1242/dev.154823.

Ng, W. H. *et al.* (2022) 'Recapitulating Human Cardio-pulmonary Co-development Using Simultaneous Multilineage Differentiation of Pluripotent Stem Cells', *eLife*. eLife Sciences Publications Ltd, 11. doi: 10.7554/ELIFE.67872.

Niakan, K. K. *et al.* (2012) 'Human pre-implantation embryo development', *Development*. The

Company of Biologists, 139(5), pp. 829–841. doi: 10.1242/DEV.060426.

Nichane, M. *et al.* (2017) 'Isolation and 3D expansion of multipotent Sox9+ mouse lung progenitors', *Nature Methods*. Nature Publishing Group, 14(12), pp. 1205–1212. doi: 10.1038/nmeth.4498.

Niehrs, C. (2012) 'The complex world of WNT receptor signalling', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 13(12), pp. 767–779. doi: 10.1038/nrm3470.

Nikolić, M. Z. *et al.* (2017) 'Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids.', *eLife*. eLife Sciences Publications, Ltd, 6. doi: 10.7554/eLife.26575.

Nikolić, M. Z. and Rawlins, E. L. (2017) 'Lung Organoids and Their Use To Study Cell-Cell Interaction', *Current Pathobiology Reports*. Springer US, 5(2), pp. 223–231. doi: 10.1007/s40139-017-0137-7.

Noe, N. *et al.* (2019) 'Mesenchyme-specific deletion of Tgf-β1 in the embryonic lung disrupts branching morphogenesis and induces lung hypoplasia', *Laboratory Investigation*. doi: 10.1038/s41374-019-0256-3.

Okubo, T. (2005) 'Nmyc plays an essential role during lung development as a dosage-sensitive regulator of progenitor cell proliferation and differentiation', *Development*, 132(6), pp. 1363–1374. doi: 10.1242/dev.01678.

Okubo, T. and Hogan, B. L. M. (2004) 'Hyperactive Wnt signaling changes the developmental potential of embryonic lung endoderm', *Journal of biology*. J Biol, 3(3). doi: 10.1186/JBIOL3.

Onaciu, A. *et al.* (2020) 'Spontaneous and Induced Animal Models for Cancer Research', *Diagnostics*. Multidisciplinary Digital Publishing Institute (MDPI), 10(9), p. 660. doi: 10.3390/DIAGNOSTICS10090660.

Park, K.-S. *et al.* (2004) 'TAZ interacts with TTF-1 and regulates expression of surfactant protein-C.', *Journal of Biological Chemistry*, 279(17), pp. 17384–17390.

Park, S. *et al.* (2018) 'Differential activities and mechanisms of the four R-spondins in potentiating Wnt/β-catenin signaling.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 293(25), pp. 9759–9769. doi: 10.1074/jbc.RA118.002743.

Perl, A.-K. T. *et al.* (2002) 'Early restriction of peripheral and proximal cell lineages during formation of the lung.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(16), pp. 10482–7. doi: 10.1073/pnas.152238499.

Perl, A. K. T. *et al.* (2005) 'Normal lung development and function after Sox9 inactivation in the respiratory epithelium', *Genesis (New York, N.Y. : 2000)*. Genesis, 41(1), pp. 23–32. doi: 10.1002/GENE.20093.

Perrin, S. (2014) 'Preclinical research: Make mouse studies work', *Nature*. Nature Publishing Group, 507(7493), pp. 423–425. doi: 10.1038/507423a.

Petrus-Reurer, S. *et al.* (2021) 'Immunological considerations and challenges for regenerative cellular therapies', *Communications Biology 2021 4:1*. Nature Publishing Group, 4(1), pp. 1–16. doi: 10.1038/s42003-021-02237-4.

Picher-Martel, V. *et al.* (2016) 'From animal models to human disease: a genetic approach for personalized medicine in ALS', *Acta Neuropathologica Communications 2016 4:1*. BioMed Central, 4(1), pp. 1–29. doi: 10.1186/S40478-016-0340-5.

Pinto, D. and Clevers, H. (2005) 'Wnt control of stem cells and differentiation in the intestinal epithelium', *Experimental cell research*. Exp Cell Res, 306(2), pp. 357–363. doi: 10.1016/J.YEXCR.2005.02.022.

Pittman, J. E. and Ferkol, T. W. (2015) 'The Evolution of Cystic Fibrosis Care.', *Chest*. Elsevier, 148(2), pp. 533–542. doi: 10.1378/chest.14-1997.

Plummer, S. *et al.* (2019) 'A Human iPSC-derived 3D platform using primary brain cancer cells to study drug development and personalized medicine', *Scientific Reports 2019 9:1*. Nature Publishing Group, 9(1), pp. 1–11. doi: 10.1038/s41598-018-38130-0.

Polański, K. *et al.* (2019) 'BBKNN: fast batch alignment of single cell transcriptomes', *Bioinformatics*. Edited by B. Berger. Oxford University Press, 36(3), pp. 964–965. doi:

10.1093/bioinformatics/btz625.

Pound, P. and Ritskes-Hoitinga, M. (2018) 'Is it possible to overcome issues of external validity in preclinical animal research? Why most animal models are bound to fail', *Journal of Translational Medicine*. BioMed Central Ltd., 16(1), pp. 1–8. doi: 10.1186/S12967-018-1678-1/FIGURES/1.

Prince, L. S. (2018) 'FGF10 and Human Lung Disease Across the Life Spectrum', *Frontiers in Genetics*. Frontiers Media SA, 9. doi: 10.3389/fgene.2018.00517.

Que, J. *et al.* (2007) 'Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm.', *Development (Cambridge, England)*. The Company of Biologists Ltd, 134(13), pp. 2521–31. doi: 10.1242/dev.003855.

Rabata, A. *et al.* (2020) '3D Cell Culture Models Demonstrate a Role for FGF and WNT Signaling in Regulation of Lung Epithelial Cell Fate and Morphogenesis', *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A., 8, p. 574. doi: 10.3389/FCELL.2020.00574/BIBTEX.

Rajagopal, J. *et al.* (2008) 'Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme', *Development*, 135(9), pp. 1625–1634. doi: 10.1242/dev.015495.

Rankin, S. A. *et al.* (2016) 'A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification.', *Cell reports*. Elsevier, 16(1), pp. 66–78. doi: 10.1016/j.celrep.2016.05.060.

Rankin, S. A. *et al.* (2018) 'Timing is everything: Reiterative Wnt, BMP and RA signaling regulate developmental competence during endoderm organogenesis', *Developmental Biology*. Academic Press, 434(1), pp. 121–132. doi: 10.1016/J.YDBIO.2017.11.018.

Rankin, S. A. *et al.* (2021) 'Tbx5 drives aldh1a2 expression to regulate a ra-hedgehog-wnt gene regulatory network coordinating cardiopulmonary development', *eLife*. eLife Sciences Publications Ltd, 10. doi: 10.7554/ELIFE.69288.

Rankin, S. A. and Zorn, A. M. (2014) 'Gene Regulatory Networks Governing Lung Specification', *Journal of Cellular Biochemistry*. John Wiley & Sons, Ltd, 115(8), pp. 1343–

1350. doi: 10.1002/jcb.24810.

Raslan, A. A. and Yoon, J. K. (2019) ‘R-spondins: Multi-mode WNT signaling regulators in adult stem cells’, *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd, pp. 26–34. doi: 10.1016/j.biocel.2018.11.005.

Rawlins, E. L. (2008) ‘Lung Epithelial Progenitor Cells: Lessons from Development’, *Proceedings of the American Thoracic Society*, 5(6), pp. 675–681. doi: 10.1513/pats.200801-006AW.

Rawlins, E. L. *et al.* (2009) ‘The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells’, *Development*, 136(22), pp. 3741–3745. doi: 10.1242/dev.037317.

Rawlins, E. L. (2011) ‘The building blocks of mammalian lung development’, *Developmental Dynamics*. John Wiley & Sons, Ltd, 240(3), pp. 463–476. doi: 10.1002/dvdy.22482.

Rehan, V. K. *et al.* (2006) ‘EVIDENCE FOR THE PRESENCE OF LIPOFIBROBLASTS IN HUMAN LUNG’, *Experimental Lung Research*. Taylor & Francis, 32(8), pp. 379–393. doi: 10.1080/01902140600880257.

Riccetti, M. *et al.* (2020) ‘The elephant in the lung: Integrating lineage-tracing, molecular markers, and single cell sequencing data to identify distinct fibroblast populations during lung development and regeneration’, *Matrix Biology*. Elsevier B.V., 91–92, pp. 51–74. doi: 10.1016/j.matbio.2020.05.002.

Roberts, D. *et al.* (2017) ‘Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth.’, *The Cochrane database of systematic reviews*, 3(3), p. CD004454.

Rock, J. R. *et al.* (2009) ‘Basal cells as stem cells of the mouse trachea and human airway epithelium.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(31), pp. 12771–5. doi: 10.1073/pnas.0906850106.

Rock, J. R. and Hogan, B. L. M. (2011) ‘Epithelial Progenitor Cells in Lung Development, Maintenance, Repair, and Disease’, *Annual Review of Cell and Developmental Biology*, 27(1), pp. 493–512. doi: 10.1146/annurev-cellbio-100109-104040.

Rockich, B. E. *et al.* (2013) 'Sox9 plays multiple roles in the lung epithelium during branching morphogenesis', *Proceedings of the National Academy of Sciences*, 110(47), pp. E4456–E4464. doi: 10.1073/pnas.1311847110.

Rowe, R. G. and Daley, G. Q. (2019) 'Induced pluripotent stem cells in disease modelling and drug discovery', *Nature Reviews Genetics 2019 20:7*. Nature Publishing Group, 20(7), pp. 377–388. doi: 10.1038/s41576-019-0100-z.

Rubin, L. P. *et al.* (1997) 'The parathyroid hormone-related protein (PTHrP) knockout mouse shows delayed lung development. {\textbullet} 1583', *Pediatric research*, 41(4), p. 266.

Rubin, L. P. *et al.* (2004) 'Arrested pulmonary alveolar cytodifferentiation and defective surfactant synthesis in mice missing the gene for parathyroid hormone-related protein', *Developmental Dynamics*. John Wiley & Sons, Ltd, 230(2), pp. 278–289. doi: 10.1002/dvdy.20058.

Scadden, D. T. (2006) 'The stem-cell niche as an entity of action', *Nature 2006 441:7097*. Nature Publishing Group, 441(7097), pp. 1075–1079. doi: 10.1038/nature04957.

Schittny, J. C. (2017) 'Development of the lung.', *Cell and tissue research*. Springer, 367(3), pp. 427–444. doi: 10.1007/s00441-016-2545-0.

Sekine, K. *et al.* (1999) 'Fgf10 is essential for limb and lung formation', *Nature Genetics*, 21(1), pp. 138–141. doi: 10.1038/5096.

Serls, A. E. *et al.* (2005) 'Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung.', *Development (Cambridge, England)*. The Company of Biologists Ltd, 132(1), pp. 35–47. doi: 10.1242/dev.01570.

Serra, M. *et al.* (2017) 'Pluripotent stem cell differentiation reveals distinct developmental pathways regulating lung- versus thyroid-lineage specification.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 144(21), pp. 3879–3893. doi: 10.1242/dev.150193.

Serra, R., Pelton, R. W. and Moses, H. L. (1994) 'TGF beta 1 inhibits branching morphogenesis and N-myc expression in lung bud organ cultures', *Development*. The Company of Biologists, 120(8), pp. 2153–2161. doi: 10.1242/DEV.120.8.2153.

Shanks, N., Greek, R. and Greek, J. (2009) 'Are animal models predictive for humans?', *Philosophy, Ethics, and Humanities in Medicine*. BioMed Central, 4(1), pp. 1–20. doi: 10.1186/1747-5341-4-2/COMMENTS.

Shannon, J. M., Gebb, S. A. and Nielsen, L. D. (1999) 'Induction of alveolar type II cell differentiation in embryonic tracheal epithelium in mesenchyme-free culture.', *Development (Cambridge, England)*, 126(8), pp. 1675–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10079230 (Accessed: 28 October 2019).

Sharma, A. *et al.* (2020) 'Multi-lineage Human iPSC-Derived Platforms for Disease Modeling and Drug Discovery', *Cell Stem Cell*. Cell Press, 26(3), pp. 309–329. doi: 10.1016/J.STEM.2020.02.011.

Shiraishi, K., Nakajima, T., *et al.* (2019) 'In vitro expansion of endogenous human alveolar epithelial type II cells in fibroblast-free spheroid culture', *Biochemical and Biophysical Research Communications*. Academic Press, 515(4), pp. 579–585. doi: 10.1016/J.BBRC.2019.05.187.

Shiraishi, K., Shichino, S., *et al.* (2019) 'Mesenchymal-Epithelial Interactome Analysis Reveals Essential Factors Required for Fibroblast-Free Alveolosphere Formation', *iScience*. Elsevier, 11, pp. 318–333. doi: 10.1016/J.ISCI.2018.12.022.

Shu, W. *et al.* (2002) 'Wnt7b regulates mesenchymal proliferation and vascular development in the lung.', *Development (Cambridge, England)*, 129(20), pp. 4831–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12361974 (Accessed: 21 July 2019).

Shu, W. *et al.* (2005) 'Wnt/β-catenin signaling acts upstream of N-myc, BMP4, and FGF signaling to regulate proximal–distal patterning in the lung', *Developmental Biology*. Academic Press, 283(1), pp. 226–239. doi: 10.1016/J.YDBIO.2005.04.014.

Shu, W. *et al.* (2007) 'Foxp2 and Foxp1 cooperatively regulate lung and esophagus development', *Development*. The Company of Biologists, 134(10), pp. 1991–2000. doi: 10.1242/DEV.02846.

Spence, J. R. *et al.* (2009) 'Sox17 Regulates Organ Lineage Segregation of Ventral Foregut Progenitor Cells', *Developmental Cell*. Cell Press, 17(1), pp. 62–74. doi: 10.1016/J.DEVCEL.2009.05.012.

Spence, J. R. *et al.* (2011) 'Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro.', *Nature*. NIH Public Access, 470(7332), pp. 105–9. doi: 10.1038/nature09691.

Stem Cells for Cystic Fibrosis Therapy | Cystic Fibrosis Foundation (no date). Available at: https://www.cff.org/stem-cells-cystic-fibrosis-therapy (Accessed: 20 July 2022).

Stuart, T. *et al.* (2019a) 'Comprehensive Integration of Single-Cell Data', *Cell*, 177(7), pp. 1888-1902.e21. doi: 10.1016/j.cell.2019.05.031.

Stuart, T. *et al.* (2019b) 'Comprehensive Integration of Single-Cell Data', *Cell*. Cell Press, 177(7), pp. 1888-1902.e21. doi: 10.1016/J.CELL.2019.05.031/ATTACHMENT/2F8B9EBE-54E6-43EB-9EF2-949B6BDA8BA2/MMC3.PDF.

Stucki, J. D. *et al.* (2018) 'Medium throughput breathing human primary cell alveolus-on-chip model', *Scientific Reports*, 8(1), p. 14359. doi: 10.1038/s41598-018-32523-x.

Sucre, J. M. S. *et al.* (2018) 'Successful Establishment of Primary Type II Alveolar Epithelium with 3D Organotypic Coculture', *American Journal of Respiratory Cell and Molecular Biology*, 59(2), pp. 158–166. doi: 10.1165/rcmb.2017-0442MA.

Sugawara, S. *et al.* (2014) 'Production of an aminoterminally truncated, stable type of bioactive mouse fibroblast growth factor 4 in Escherichia coli', *Journal of bioscience and bioengineering*. J Biosci Bioeng, 117(5), pp. 525–530. doi: 10.1016/J.JBIOSC.2013.10.009.

Sun, D. *et al.* (2021) 'A functional genetic toolbox for human tissue-derived organoids', *eLife*. eLife Sciences Publications Ltd, 10. doi: 10.7554/ELIFE.67886.

Swarr, D. T. and Morrisey, E. E. (2015) 'Lung Endoderm Morphogenesis: Gasping for Form and Function', *Annual Review of Cell and Developmental Biology*. Annual Reviews , 31(1), pp. 553–573. doi: 10.1146/annurev-cellbio-100814-125249.

Szenker-Ravi, E. *et al.* (2018) 'RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6', *Nature*. Nature Publishing Group, p. 1. doi: 10.1038/s41586-018-0118-y.

Takahashi, K. and Yamanaka, S. (2006) 'Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors', *Cell*. Cell Press, 126(4), pp. 663–676. doi: 10.1016/J.CELL.2006.07.024.

Tamò, L. *et al.* (2018) 'Generation of an alveolar epithelial type II cell line from induced pluripotent stem cells', *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 315(6), pp. L921–L932. doi: 10.1152/ajplung.00357.2017.

Tan, Q. *et al.* (2017) 'Human airway organoid engineering as a step toward lung regeneration and disease modeling', *Biomaterials*. Elsevier, 113, pp. 118–132. doi: 10.1016/J.BIOMATERIALS.2016.10.046.

Tollet, J., Everett, A. W. and Sparrow, M. P. (2001) 'Spatial and temporal distribution of nerves, ganglia, and smooth muscle during the early pseudoglandular stage of fetal mouse lung development', *Developmental Dynamics*. John Wiley & Sons, Ltd, 221(1), pp. 48–60. doi: 10.1002/dvdy.1124.

Torday, J. S. *et al.* (2002) 'Leptin mediates the parathyroid hormone-related protein paracrine stimulation of fetal lung maturation.', *American journal of physiology. Lung cellular and molecular physiology*. NIH Public Access, 282(3), pp. L405-10. doi: 10.1152/ajplung.2002.282.3.L405.

Torday, J. S. and Rehan, V. K. (2006) 'Up-Regulation of Fetal Rat Lung Parathyroid Hormone-Related Protein Gene Regulatory Network Down-Regulates the Sonic Hedgehog/Wnt/βcatenin Gene Regulatory Network', *Pediatric Research*. Nature Publishing Group, 60(4), pp. 382–388. doi: 10.1203/01.pdr.0000238326.42590.03.

Torday, J. S., Torres, E. and Rehan, V. K. (2003) 'The role of fibroblast transdifferentiation in lung epithelial cell proliferation, differentiation, and repair in vitro.', *Pediatric pathology & molecular medicine*, 22(3), pp. 189–207. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12746170 (Accessed: 2 October 2019).

Traag, V. A., Waltman, L. and van Eck, N. J. (2019) 'From Louvain to Leiden: guaranteeing well-connected communities', *Scientific Reports 2019 9:1*. Nature Publishing Group, 9(1), pp. 1–12. doi: 10.1038/s41598-019-41695-z.

Travaglini, K. J. *et al.* (2020) 'A molecular cell atlas of the human lung from single-cell RNA sequencing', *Nature 2020 587:7835*. Nature Publishing Group, 587(7835), pp. 619–625. doi:

10.1038/s41586-020-2922-4.

Trejo, C. L. *et al.* (2017) 'Lgr5 is a marker for fetal mammary stem cells, but is not essential for stem cell activity or tumorigenesis', *npj Breast Cancer*. Nature Publishing Group, 3(1), p. 16. doi: 10.1038/s41523-017-0018-6.

Treutlein, B. *et al.* (2014) 'Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq', *Nature*. Nature Publishing Group, 509(7500), pp. 371–375. doi: 10.1038/nature13173.

Tsao, P.-N. *et al.* (2008) 'Gamma-secretase activation of notch signaling regulates the balance of proximal and distal fates in progenitor cells of the developing lung.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 283(43), pp. 29532–44. doi: 10.1074/jbc.M801565200.

Varner, V. D. *et al.* (2015) 'Mechanically patterning the embryonic airway epithelium', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(30), pp. 9230–9235. doi: 10.1073/pnas.1504102112.

Vatine, G. D. *et al.* (2019) 'Human iPSC-Derived Blood-Brain Barrier Chips Enable Disease Modeling and Personalized Medicine Applications', *Cell Stem Cell*. Cell Press, 24(6), pp. 995-1005.e6. doi: 10.1016/J.STEM.2019.05.011.

Volckaert, T. *et al.* (2013) 'Localized Fgf10 expression is not required for lung branching morphogenesis but prevents differentiation of epithelial progenitors', *Development*, 140(18), pp. 3731–3742. doi: 10.1242/dev.096560.

Volckaert, T. *et al.* (2017) 'Fgf10-Hippo Epithelial-Mesenchymal Crosstalk Maintains and Recruits Lung Basal Stem Cells', *Developmental Cell*. Cell Press, 43(1), pp. 48-59.e5. doi: 10.1016/J.DEVCEL.2017.09.003.

Volckaert, T. *et al.* (2019) 'Hippo signaling promotes lung epithelial lineage commitment by curbing Fgf10 and β-catenin signaling.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 146(2), p. dev166454. doi: 10.1242/dev.166454.

Volckaert, T. and De Langhe, S. P. (2015) 'Wnt and FGF mediated epithelial-mesenchymal crosstalk during lung development', *Developmental Dynamics*. Wiley-Blackwell, 244(3), pp.

342–366. doi: 10.1002/dvdy.24234.

Waas, M. *et al.* (2020) 'SurfaceGenie: A web-based application for prioritizing cell-type-specific marker candidates', *Bioinformatics*. Oxford University Press, 36(11), pp. 3447–3456. doi: 10.1093/bioinformatics/btaa092.

Wamaitha, S. E. and Niakan, K. K. (2018) 'Human Pre-gastrulation Development', *Current Topics in Developmental Biology*. Academic Press, 128, pp. 295–338. doi: 10.1016/BS.CTDB.2017.11.004.

Wang, D. *et al.* (2007) 'A pure population of lung alveolar epithelial type II cells derived from human embryonic stem cells.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 104(11), pp. 4449–54. doi: 10.1073/pnas.0700052104.

Wang, Y. *et al.* (2014) 'Understanding how cystic fibrosis mutations disrupt CFTR function: From single molecules to animal models', *The International Journal of Biochemistry & Cell Biology*. Pergamon, 52, pp. 47–57. doi: 10.1016/J.BIOCEL.2014.04.001.

Warburton, D. *et al.* (2005) 'Molecular Mechanisms of Early Lung Specification and Branching Morphogenesis', *Pediatric Research*. Nature Publishing Group, 57(5 Part 2), pp. 26R-37R. doi: 10.1203/01.PDR.0000159570.01327.ED.

Weatherbee, B. A. T., Cui, T. and Zernicka-Goetz, M. (2021) 'Modeling human embryo development with embryonic and extra-embryonic stem cells', *Developmental Biology*. Academic Press, 474, pp. 91–99. doi: 10.1016/J.YDBIO.2020.12.010.

Weaver, M., Batts, L. and Hogan, B. L. M. (2003) 'Tissue interactions pattern the mesenchyme of the embryonic mouse lung.', *Developmental biology*, 258(1), pp. 169–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12781691 (Accessed: 21 July 2019).

Willem, M. *et al.* (2002) 'Specific ablation of the nidogen-binding site in the laminin gamma1 chain interferes with kidney and lung development.', *Development*, 129(11), pp. 2711–2722.

Wolf, F. A., Angerer, P. and Theis, F. J. (2018) 'SCANPY: Large-scale single-cell gene expression data analysis', *Genome Biology*. BioMed Central Ltd., 19(1), p. 15. doi: 10.1186/s13059-017-1382-0.

Wong, A. P. *et al.* (2012) 'Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein', *Nature Biotechnology*. Nature Publishing Group, 30(9), pp. 876–882. doi: 10.1038/nbt.2328.

Wu, J. *et al.* (2018) 'Role of Fibroblast Growth Factor 10 in Mesenchymal Cell Differentiation During Lung Development and Disease', *Frontiers in Genetics*. Frontiers, 9, p. 545. doi: 10.3389/fgene.2018.00545.

Yamamoto, Y. *et al.* (2017) 'Long-term expansion of alveolar stem cells derived from human iPS cells in organoids.', *Nature methods*, 14(11), pp. 1097–1106. doi: 10.1038/nmeth.4448.

Yamamoto, Y. *et al.* (2020) 'A method of generating alveolar organoids using human pluripotent stem cells', *Methods in Cell Biology*. Academic Press, 159, pp. 115–141. doi: 10.1016/BS.MCB.2020.02.004.

Yamanaka, S. (2020) 'Pluripotent Stem Cell-Based Cell Therapy-Promise and Challenges', *Cell stem cell*. Cell Stem Cell, 27(4), pp. 523–531. doi: 10.1016/J.STEM.2020.09.014.

Yan, K. S. *et al.* (2017) 'Non-equivalence of Wnt and R-spondin ligands during Lgr5+ intestinal stem-cell self-renewal.', *Nature*. NIH Public Access, 545(7653), pp. 238–242. doi: 10.1038/nature22313.

Yang, J. *et al.* (2016) 'The development and plasticity of alveolar type 1 cells.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 143(1), pp. 54–65. doi: 10.1242/dev.130005.

Yang, Y. *et al.* (2018) 'Spatial-Temporal Lineage Restrictions of Embryonic p63+ Progenitors Establish Distinct Stem Cell Pools in Adult Airways', *Developmental Cell*. Cell Press, 44(6), pp. 752-761.e4. doi: 10.1016/J.DEVCEL.2018.03.001.

Yu, Q. *et al.* (2021) 'Charting human development using a multi-endodermal organ atlas and organoid models', *Cell*. Cell, 184(12), pp. 3281-3298.e22. doi: 10.1016/J.CELL.2021.04.028.

Yuan, T. *et al.* (2018) 'Fgf10 Signaling in Lung Development, Homeostasis, Disease, and Repair After Injury', *Frontiers in Genetics*. Frontiers Media S.A., 9, p. 418. doi: 10.3389/FGENE.2018.00418/BIBTEX.

Zakrzewski, W. *et al.* (2019) 'Stem cells: Past, present, and future', *Stem Cell Research and Therapy*. BioMed Central Ltd., 10(1), pp. 1–22. doi: 10.1186/S13287-019-1165-5/FIGURES/8.

Zaw-Tun, H. A. (1982) 'The tracheo-esophageal septum--fact or fantasy? Origin and development of the respiratory primordium and esophagus.', *Acta anatomica*, 114(1), pp. 1–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7148374 (Accessed: 29 October 2019).

Zepp, J. A. *et al.* (2017) 'Distinct Mesenchymal Lineages and Niches Promote Epithelial Self-Renewal and Myofibrogenesis in the Lung', *Cell*. Cell Press, 170(6), pp. 1134-1148.e10. doi: 10.1016/J.CELL.2017.07.034.

Zepp, J. A. and Morrisey, E. E. (2019) 'Cellular crosstalk in the development and regeneration of the respiratory system', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 20(9), pp. 551–566. doi: 10.1038/s41580-019-0141-3.

Zhai, J. *et al.* (2022) 'Human embryonic development: from peri-implantation to gastrulation', *Trends in Cell Biology*. Elsevier Current Trends, 32(1), pp. 18–29. doi: 10.1016/J.TCB.2021.07.008.