Investigation of Epithelial Cell Fate Decisions in the Developing Human Lung Guides the Generation of Human Lung Organoid Model Systems

by

Alyssa Jane Miller

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Jason R. Spence, Chair
Associate Professor Ben Allen
Assistant Professor Sue Hammoud
Associate Professor Ariella Shikanov

Alyssa Jane Miller

ajanemil@umich.edu

ORCID: 0000-0002-4069-1550

© Alyssa J. Miller 2019

Dedication

To my family, my husband, and my friends, for always encouraging me to take joy in the journey and to never stop exploring. Thank you.

Acknowledgements

I would first like to thank my thesis advisor Dr. Jason Spence. You listened patiently and provided critical feedback as I shaped the experiments that gave rise to this thesis project, and encouraged me to “follow the data,” wherever it led. Thank you for sharing your expertise in experimental design, data interpretation, and scientific writing. More importantly, thank you for teaching me the nuances of good science: how to keep good spirits in the face of failure. How to lead and mentor others by example and with respect. How to give and receive criticism constructively. I count myself lucky to have had these 5 years to learn from you, and to share the thrill of getting the most exciting data Fridays at 5 pm. I look forward to bothering you with very cool data for years to come, this time as a colleague.

Thank you to my committee members, Dr. Ben Allen, Dr. Sue Hammoud, and Dr. Ariella Shikanov. You asked the hard questions at the right time. I’m grateful for your constructive feedback and expertise, which shaped this project in unanticipated and fun ways. Thank you to my program, including Dr. Bob Fuller, for providing me with so many opportunities to share my science and build community with my peers.

Thank you to each of my former mentors, both in science and in music and art. Thank you to Dr. Vern Carruthers for giving me the opportunity to rotate in your lab. Working with you and your team was one of my first experiences in learning (or inventing!) entirely new techniques, like slicing mouse brains to quantify parasite

numbers. Your support gave me the confidence to truly “own” my experiments. Thank you to my former mentors at the Massachusetts General Hospital, Dr. Melissa Suter, for showing me how to develop a research project and build a great, supportive, and hilarious team, including Dr. Lida Hariri and Dr. David Adams, and to Dr. Robert Brown and Dr. Leo Ginns, for providing the true reason behind why I studied the lung in the first place: to help the patients we can’t yet help. Thank you to my former mentors in the arts, Kami Rowan, Mark Freundt, Patrick Lui and KT Scherman, for teaching me how to feed my soul and cultivate my own creativity and confidence, and for teaching me that genius is worthless without practice, but practice without genius is enough.

Thank you to my friends here at the University of Michigan for the coffee dates, happy hours and volleyball games that kept us sane. Thank you especially to my lab mates, former and present, for their kindness and support. Showing up to lab every day has been a joy because of you all. Thank you to Dr. Briana Dye, Dr. Alana Chin, Dr. Daysha Ferrer-Torres, Dr. Mike Czerwinski, Dr. Tristan Frum, Dr. David Hill, Emily Holloway, Renee Conway, Meghan Capeling, Ansley Semack, Yuhwai Tsai, Fei Fei Wu, Sha Huang, Josh Wu, Mike Dame and the TTML, Katelyn Green, Noah Steinfeld, Mark Painter, and Dr. Amanda Leone for all the wonderful science conversations. I can’t wait to see the amazing things you will all accomplish.

Thank you most of all to my family: Mark, Jane and Carolyn Miller, my grandparents, and my husband William Caddick, for all the support, love and listening.

And to each and every teacher, educator, parent, nanny, teammate, colleague and friend who told me, “yes you can.” Thank you.

Table of Contents

List of Figures

List of Tables

Abstract

1 in 10 U.S. infants are born prematurely (www.cdc.gov), before the lungs have fully formed. The primary cause of mortality in these patients is underdeveloped lungs, but even surviving patients have a 40% chance of developing life-long lung problems. Our ability to effectively treat premature babies by speeding up lung development is hindered by the severe lack of knowledge about normal human lung development that is exacerbated by a need for better human cell-based developmental model systems. Mouse studies have shown that the lung develops as a network of epithelial tubes undergo repeated rounds of bifurcation within surrounding mesenchyme. This process, called branching morphogenesis, gives rise to the conducting airways and alveoli, where gas exchange takes place. Bud tip progenitor cells reside at the tips of branching buds during development and give rise to all adult lung epithelial cell types, including conducting airway-specific cells and alveolar cells. Bud tip progenitor cells terminally differentiate into alveolar cells at the very end of development and are not found in adult lungs. However, bud tip progenitors remain immature in premature infants, impeding gas exchange. Despite the importance of bud tip progenitors in premature lungs, it is unknown how this population of specialized progenitors is maintained during development, or what signals prompt these cells to differentiate into specific lung lineages. To address these questions, I first characterized bud tip progenitors in the native developing lung using bulk RNA sequencing and protein staining, and then

developed a fetal bud tip progenitor organoid system to show that the growth factors Fibroblast Growth Factor 7 (FGF7), WNT and All-Trans Retinoic Acid (ATRA) are sufficient to maintain the bud tip progenitor cell state in culture. To determine how bud tip progenitors are instructed to give rise to differentiated epithelial cell types, I next employed single-cell RNA sequencing of the lung epithelium to identify signaling pathways active during development. Testing of these pathways in vitro using the bud tip progenitor organoids revealed that transient activation of the SMAD signaling pathway is sufficient to generate a population of cells with transcriptomes and functional properties mirroring adult lung basal stem cells, the resident stem cell of the adult airway. Finally, recapitulation of these developmental events using human pluripotent stem cells (hPSCs) in a dish led to the development of hPSC-derived lung bud tip organoids and efficient derivation of lung basal stem cells from hPSCs, which have many applications for future study of human lung development and disease without the use of donated human tissue.

Chapter 1: Introduction

Portions of this chapter have been published: Miller, A.J.; Spence, J.R. In Vitro Models to Study Lung Development, Disease and Homeostasis. *Physiology (Bethesda)*. **2017**, 32(3):246-260. Doi: 10.1152/physiol.00041.2016. PMID: 28404740.

1.1 Lung Overview

The main function of the lung is to provide oxygen to the body's tissues. Defects in lung development or diseases affecting the structure and function of the lung can have fatal consequences. Most of what we currently understand about human lung development and disease has come from animal models. However, animal models are not always fully able to recapitulate human lung development and disease, highlighting an area where *in vitro* models of the human lung can complement animal models to further understanding of critical developmental and pathological mechanisms. This overview will discuss current advances in generating *in vitro* human lung models using primary human tissue, cell lines, and human pluripotent stem cell derived lung tissue, and will discuss crucial next steps in the field.

Lungs have evolved to serve the essential function of extracting oxygen from the air for use in aerobic metabolism and to remove the gaseous waste of this process from the body. Large, multi-cellular animals like humans require an enormous quantity of oxygen to maintain baseline energy levels needed for survival (1, 2). To meet these

energy demands, lungs have evolved to maximize the surface area available for gas exchange by forming a complex network of tube-like epithelial branches known as the conducting airway, which consists of the trachea, bronchi and bronchioles (Figure 1-1). The tubes in this branched network get progressively smaller until they terminate with thin distal air sacs, called alveoli, which are closely associated with the capillary network to allow diffusion of oxygen into the bloodstream (Figure 1-1). When lung function is compromised, whether due to developmental defects or disease, the consequences can be severe and are often fatal.

Animal models have been instrumental in our understanding of lung development and disease (3-5). However, given many differences between animal and human lung physiology, there is an unmet need for human lung model systems that can complement animal models to improve our understanding of human lung physiology. Indeed, genetic gain- and loss-of-function studies in mice have improved our understanding of signaling networks in the developing lung and have shed significant light on physiological processes such as branching morphogenesis, which generates the arborized tree-like network of the functional lung (4-6). Similarly, rodent lung injury models can reproduce some aspects of complex human diseases such as lung fibrosis and emphysema (7, 8) and have helped illuminate multiple disease mechanisms. However, significant differences in lung physiology between mice and humans exist, which complicates inferring how animal studies can be applied to human development and disease (9). For example, mouse lungs develop quickly and do not begin forming alveoli until after birth, whereas human lungs undergo many additional rounds of branching before beginning alveolarization (10). Human lungs also have anatomical

differences in the localization of stem and progenitor cells when compared to rodent lungs. Thus, it is not surprising that animal models are not able to recapitulate the full spectrum of human pathophysiology, (11, 12) which is further highlighted by the fact that up to 80% of drugs that pass pre-clinical animal tests in rodents fail to effectively treat human disease during clinical trials (13).

Human model systems that reliably and accurately recapitulate the complexities of human biology *in vitro* will help improve our understanding and ability to treat human lung disease. *In vitro* models of the human lung are also likely to provide powerful platforms for large-scale screens, molecular level analysis of cell-cell interactions and the potential to study or treat lung disorders with personalized medicine.

In this introduction, we will discuss the current state-of-the-art for systems to model the human lung *in vitro,* and set the stage for the work outlined in this thesis to better understand human lung development and lung epithelial cell fate decisions. Since human systems are being developed to mimic human lung development and homeostasis/disease in the adult, this overview will briefly discuss lung development, adult homeostasis and injury repair followed by a discussion of current *in vitro* models of the human lung using immortalized cell lines, primary patient-derived tissue and pluripotent stem cell-derived human lung tissue. We will further discuss how these systems can be implemented to better understand human pathology, evaluate benefits and drawbacks of the current models and explore critical next steps for the field.

1.2 Human and mouse lung development

Proper lung formation during development is critical to survival after birth. Lung organogenesis begins around embryonic day (E) 9 in the mouse and around week 3

during human development as the primitive lungs bud from the foregut endoderm (5, 14, 15). As the lungs develop, they undergo branching morphogenesis, which generates a stereotyped network of epithelial tubes surrounded by mesenchymal tissue (6). During this time, the epithelial tubes are patterned along the proximal-distal axis of the lung (Figure 1-1; (5)). Work using mouse lineage tracing has described a unique population of lung epithelial progenitor cells, called bud tip progenitors, that reside in the distal tips of the branching lung epithelium during development (16, 17), and which give rise to all epithelial cell types of the adult lung, including airway cell types and, later in development, alveolar cell types.

The location of bud tip progenitor cells within the growing organ changes over time, such that the bud tip progenitors reside in the future trachea and primary bronchi during very early development, when they give rise to tracheal and bronchial cells. As development progresses, these bud tip progenitors are carried along with the growing airway. While bud tip progenitors leave differentiated progeny in their wake, a population of progenitors remains at the tips of branching buds, where they give rise to progressively more distal epithelial cell types. The bud tip progenitors that remain in late development ultimately give rise to alveolar cell types (16, 18). Once alveolarization is complete, no bud tip progenitors remain, and the lung relies on other resident stem and progenitor populations to support homeostasis and repair in adulthood. More recent work, including work reported here, has shown that a population of bud tip progenitors also exists in the developing human lung and has multilineage potential (19-21).

Multiple pathways and transcription factors are critical for establishing proximal-distal patterning, controlling progenitor states and regulating branching morphogenesis.

Regulation of branching morphogenesis has been reviewed extensively in numerous excellent reviews (4, 5, 14, 22-24). After the branching program, which lasts from E12.5-E17.5 in mice and from weeks 5-26 in humans, the distal airways undergo sacculation (10, 23, 25). During this process, distal epithelial progenitors begin to differentiate into the specialized cell types of the alveoli and undergo morphological changes to take on a sac-like structure. Full maturation of the alveolar sacs occurs from birth through postnatal week 2 in mice (5, 26), but in humans alveolarization begins in the third trimester and persists to up to three years postnatally (Figure 1; (10, 27). Recent work on the molecular mechanisms of postnatal alveologenesis has identified a specific subset of PDGFRa+ mesenchymal cells as critical to postnatal alveologenesis in mice(28).

Although the development and cellular components of the lung epithelium have been well characterized, the developing lung also requires proper formation of extensive branched vascular and peripheral nervous networks in addition to numerous mesenchymal cell types. While the diversity and function of all mesenchymal cell types in the lung is still poorly understood, (29, 30) we are beginning to understand the importance of vascular and neuronal networks during development. When normal vascular development is disrupted, for example, due to congenital diaphragmatic hernia or due to ventilator induced lung injury in pre-term infants, severe pulmonary hypertension and oxygen insufficiency can develop, significantly increasing morality rates (9, 31, 32). Similarly, during branching morphogenesis, neural crest cells migrate throughout the branching lung buds where they form neurons that are closely

associated with the developing bronchial tree. Disruption of these bronchial neurons leads to severe defects in branching morphogenesis (33-35).

Another critical aspect of lung development that has been difficult to study *in vivo* is the role of biomechanical forces and the extra cellular matrix (ECM) (36, 37). Though little is known in humans, mouse studies have shown that the basement membranes surrounding branching epithelial structures have dynamic ECM remodeling that is essential for proper lung development (38-40), and time-lapse imaging of precision cut mouse lung sections showed that cell migration and clustering are critical drivers of alveologenesis(41). Further, it has recently been shown that *Hox* genes regulate the elastin network during alveologenesis in mice(42). Similarly, 3-dimensional (3D) imaging of developing mouse lungs suggests collagen and elastin networks are deposited and extensively remodeled by mesenchymal cells in close association with alveolar epithelial cells as structurally mature alveoli form (43).

1.3 Adult lung homeostasis and disease

Proper lung homeostasis and response to injury is critical for adult survival. The lungs are continuously exposed to the external environment, and are vulnerable to damage from environmental toxins or contact with numerous microbes, including pathogens (44). In order to handle these external stimuli the adult lung must be able to clear debris, mount immune/inflammatory responses, regulate cell turnover and facilitate proper repair after damage without disrupting oxygen absorption. In order to carry out these diverse functions, the lung contains specialized epithelial cell types in the proximal airway, distal alveoli, and mesenchymal compartments that are critical to lung homeostasis and reaction to injury (Figure 1-1).

In both mice and humans, the proximal airway epithelium contains numerous specialized cells that participate in normal function and cell turnover as well as repair after injury. The proximal airway is primarily comprised of specialized cells that secrete mucous to trap environmental debris (Goblet cells; (45)), or that aid in innate immunity by secreting antimicrobial peptides (Club cells; (46)) and cells with multiple cilia that beat in unison to remove mucous from the airway. Disruption of mucociliary function, from genetic or acquired defects, can lead to repeated chest infections and airway remodeling in humans (47-49). Mucociliary clearance and the antimicrobial properties of the lung are severely impaired in patients with Cystic Fibrosis (CF) (50-52). Additionally, roughly 1% of cells in the proximal airway are neuroendocrine cells, which can occur alone or in small clusters of neurons and peptide-producing cells called neuroendocrine bodies (NEBs) (53). Disregulation of NEBs has long been noted in many infant and adult pulmonary diseases (54-56), and recent evidence in mice has shown that disruption of NEB formation results in a large immune response and irreversibly simplifies alveoli (57). In addition to these cell types, the most proximal airways (the trachea only in mouse and the trachea and primary and secondary bronchi in humans) contain basal stem cells on the basolateral side of the epithelium. These stem cells can give rise to any of the cell types of the proximal epithelium (58-60), and will be discussed further in the context of lung regeneration.

In contrast to the proximal epithelium, the adult alveoli contain specialized cell types that are optimized to facilitate gas exchange. The alveolus is a balloon-like structure lined by long, thin alveolar epithelial type I cells (AECI), which provide a thin barrier between the air and the surrounding microcapillary network to allow efficient

diffusion of gas. Each alveolar sac also possesses alveolar epithelial type II cells (AECII) that secrete surfactant proteins to reduce surface tension, preventing airway collapse during exhalation. Diseases affecting the alveoli are especially lethal, as they directly interfere with gas exchange. While specific cell types of the alveoli can be regenerated (25, 61, 62), so far there is little evidence that, once the structure of the alveolus is lost, it can be reinstated. In the case of Chronic Obstructive Pulmonary Disease (COPD) and Emphysema, an overactive immune response leads to increased activity of enzymes that destroy the ECM of the alveoli (63). In Interstitial Lung Disease (ILD), oxygen diffusion is disrupted by an increased deposition of ECM proteins between alveolar cells and the surrounding capillary network, increasing the distance gas needs to diffuse to enter the bloodstream (64-66).

In addition to the epithelial cells of the lung, diverse populations of mesenchymal cells surround the bronchial tree. The properties and functions of these populations are still an area of intensive study (29). Smooth muscle fibers surround the most proximal airways and function to regulate airway constriction in response to allergens. Thickening and hyper-responsiveness of airway smooth muscle are classic hallmarks of asthma (67). The lung mesenchyme also contains extensive vasculature, which branches in concordance with the bronchial tree and supplies and maintains the capillary network that is essential for gas exchange. Compromised endothelial cell function can lead to increased vasoconstriction and inflammation, subsequently leading to pulmonary arterial hypertension, a common and currently incurable disease which can result in heart failure and death (68).

1.4 Adult lung regeneration

The adult lung is largely quiescent (69) but has a remarkable capacity to regenerate after injury (70). To date, both animal models and some *in vitro* models using mouse and human-derived tissue have begun to uncover the mechanisms regulating the proliferation of lung progenitor and stem cells in the upper airways and alveoli (22, 25, 71-77), and have begun to uncover critical roles for mesenchymal cells in regulating repair after injury (78, 79).

The proximal airway contains basal stem cells which are able to self renew and give rise to secretory, ciliated and neuroendocrine cells (58, 59, 73). In addition to basal stem cells, significant cellular plasticity allows other cell types in the epithelium to act as progenitor cells and contribute to homeostasis by proliferating and repopulating lost cells after injury. For example, upon ablation of mouse basal stem cells, luminal secretory cells are able to de-differentiate into stable basal stem cells (80). Similarly, secretory club cells are able to generate new ciliated cells after injury, a process referred to as transdifferentiation (81).

Similar to the proximal airway, the distal airways and alveolar regions contain both putative, dedicated stem cells and a population of cells that are able to transdifferentiate to regenerate the epithelial layers after injury. In the mouse, bronchio-alveolar stem cells (BASC) reside at the junction between the distal airway and the alveolus and can give rise to multiple cell types upon injury (75). However, it is still unclear whether human lungs contain BASCs. Additionally, in both the mouse and the human, AECII cells in the alveoli have been shown to act as an alveolar progenitor, proliferating to replenish lost AECII cells after injury, and giving rise to AECI cells in

mice (25, 61). Interestingly, AECI cells have also been reported to give rise to AECII cells in the mouse after injury (82). Recent evidence suggests that there are also rare stem/progenitor cells in the distal lung that do not express mature lineage markers (Lineage Negative Epithelial Progenitors; LNEP) and that become active in response to severe distal lung injury in mice (62). Tissue sections from patients with severe lung fibrosis were phenotypically similar to those of mice in which LNEPs were induced during injury repair, suggesting a similar population of cells and similar repair mechanisms may be at play in human disease or injury repair.

The homeostasis and regenerative capacity of diverse mesenchymal populations of the lung remains poorly characterized, although we are beginning to understand the complex crosstalk between the mesenchyme and epithelium during lung regeneration. For example, both lipofibrobasts and myofibroblasts proliferate after injury to repopulate lost cells. More recent evidence suggests that lipo- and myofibroblasts also play a role in repairing the alveolar compartments after injury (78).

1.5 A need for human lung models

Most of what we know of adult human lung homeostasis is inferred from rodent studies or is gleaned from tissue samples obtained from human lungs. Studies utilizing human samples for analysis are inherently difficult to interpret, as it is difficult to control for lifestyle and genetic variability in different subjects. Similarly, despite the fact that rodents and human lungs contain most of the same cell types, the anatomy of the lung varies. For example, as elaborated above, human lungs contain basal cells throughout the trachea and bronchi, whereas mouse basal cells are restricted to the trachea; there

is no evidence that human lungs possess a BASC population; the majority of cells in the human proximal airway are multiciliated cells, whereas club cells are more abundant in rodents (83); and, while goblet cells are prevalent in the proximal human airway, they primarily appear in the mouse after injury (84). *In vitro* models of lung homeostasis will be particularly powerful for increasing our understanding of how the human lung maintains homeostasis and how disregulation of specific cellular processes lead to disease, especially in hard to study lung regions like the human alveolus. It may also allow for tightly controlled experiments exploring the direct roles of mechanical force, immune interactions, and the ECM on progenitor differentiation and response to injury. Additionally, leveraging *in vitro* models could aid the discovery of novel therapeutic targets, may provide powerful, scalable screening platforms to test pharmaceuticals, and can act as an important pre-clinical step that bridges the gap between drug testing in rodent models and human clinical trials, which are expensive and have a high failure rate.

What makes a good lung model?

The human lung undergoes a complex developmental program, maintains homeostasis in the adult to allow efficient gas exchange, and defends and repairs itself in response to continuous outside assaults, including inhaled environmental and infectious agents. Mechanistically, influences such as biomechanical forces, morphogen signaling and molecular-level regulation of gene and protein expression work together to regulate lung development, homeostasis and regeneration (4, 5, 22, 36, 43, 85). Currently, *in vitro* models utilizing human cells from cell lines, donors, or tissue derived

from human pluripotent stem cells (hPSCs) can reproduce some of these developmental events; however, these models may represent only a partial picture of processes that are more complex in the native lung. On the other hand, current *in vitro* systems offer tightly controlled cellular environments that can be evaluated in real time, and easily scaled and reproduced, offering significant advantages over animal models or clinical studies. As the field moves forward, layers of complexity can be added to these reductionist systems to model increasingly complex biological interactions, for example, by co-culturing epithelium with vascular, neural or immune cells, or using engineering approaches to develop more complex matrix environments (86, 87), which may begin to more closely mirror the environment of the native lung.

1.6 *In vitro* models utilizing cells derived from the human lung

Immortalized human cell lines

Human cell lines have been a simple and powerful tool for studying lung cell biology for decades, and remain popular especially in the CF field (88, 89), where access to primary tissue from patients can be limited. Cell lines have been cultured at a an air liquid interface (ALI) to induce differentiation of cells with mucociliary function, and have been used to show that CFTR mutations do not directly affect neutrophil migration (90), and to study the effect of CFTR mutations on cell permeability and cell-cell junction integrity (91, 92). In addition, cell lines with mRNA expression patterns similar to AECII cells have been used to probe the mechanisms governing fluid transport across the alveolar epithelium (93). Transwell co-culture systems, utilizing a lower chamber in which human vascular cells are grown on one side and cell lines derived from human airway or alveolar epithelium are grown on a permeable

membrane, have been used to study epithelial resistance and barrier function (94). Further, a cell line composed primarily of basal-like cells has been used to form self-assembling organoids in a 3D ECM environment (95). Co-culture of these cells with human vascular endothelial cells (HUVECs) in a 3D environment promoted morphological rearrangements similar to branching morphogenesis, leading to domains with both proximal and distal cell types. This work suggested that proximal airway cells may have the capacity to generate distal cell types given the correct environment in culture.

Since human cell lines are immortalized or transformed, it is unclear how well these cells maintain the physiology of normal airway cells, and more recent work has focused on the use of non-transformed primary lung tissue from human donors to model lung physiology. In this light, many approaches have been taken to grow human primary lung tissue in culture. Adult lung epithelium has been grown in monolayers, on an ALI, in self-organizing 3-dimensional environments (organoids) and in a variety of bioengineered or decellularized biological scaffolds to provide models of the upper and lower airways.

Monolayer and air liquid interface culture systems

Some of the earliest work growing primary human lung cells in culture employed a 2 dimensional (2D) approach, where primary epithelial cells taken from human nasal or bronchial brushings were cultured on a plastic plate submerged in media (96). 2D monolayer cultures are simple to generate, and are used clinically to diagnose ciliary dysfunction (97, 98). The development of the ALI method in the late 1980s was a major

advancement in the field (99). In the ALI system, cells are grown on a porous filter that physically separates the lung epithelial tissue from the underlying media, allowing the surface of the cells to interface with the surrounding air, while the basal surface of the cells have access to nutrients and other media additives via diffusion through the porous membrane (Figure 2). When human nasal or bronchial epithelial cells are grown on the ALI, ciliated and mucous producing cells become polarized and take on the proper apical-basal morphology, including functional apical cilia and/or mucous secretion from goblet cells (99-101). Additional advances such as the addition of Rho-associated protein kinase (ROCK) inhibitors, 'dual smad' inhibition, or the co-culture with fibroblast feeder cells has improved the ability to culture basal stem cells long-term and for many passages (60, 102-105). ALI cultures of primary human epithelial tissue have been particularly instrumental in understanding cellular mechanisms of cystic fibrosis (CF), where these systems have been used to study the cellular effects of CFTR mutations or defects in epithelial barrier function (91, 106-108) and have been instrumental in drug discovery to treat certain subsets of CF (109, 110). ALI cultures have also been employed to evaluate the effect of external toxins such as cigarette smoke, environmental particles or infectious agents such as HIV and influenza on bronchial epithelial cells (111-113), and to assess the role of soluble signals from healthy or diseased feeder fibroblasts on the gene expression and behavior of epithelial cells (114).

Self-assembling 3-dimensional cultures

More recently, primary adult human lung tissue has been cultured in 3D environments, which leads to self-assembly of the tissue into a variety of forms that partially reflect the structure of the lung (58, 115-117).

The generation of human “bronchospheres”, 3D spheres derived from primary human bronchiolar epithelial cells grown in a 3D matrix, was a huge innovation in the field (58). This method involves isolating basal stem cells from mouse or human epithelial tissue. When these cells are embedded in a 3D ECM gel, spherical clonal colonies form after a short time in culture. These initial experiments provided a platform to conduct functional experiments in human tissue to show that human basal stem cells are able to self-renew and give rise to proximal secretory and ciliated cells (58). More recently, human bronchospheres were used as a screening platform to probe the epithelial response to specific stimuli, identifying a number of proteins that bias basal stem cell differentiation toward a secretory cell fate (115, 118). While several reports have shown the generation of epithelium-only 3D bronchospheres containing cell types of the proximal airway, bronchiolar epithelium has also been co-cultured with lung fibroblasts and endothelial cells (116). Interestingly, these cultures rapidly formed cell condensations that had properties of both proximal and distal airway compartments, and similar to basal cell-like cell lines, suggested that proximal cell types maintain a level of plasticity that can be reactivated in organoid culture, perhaps due to specific cues from the microenvironment provided by co-cultured cells (116). Similarly, culture systems utilizing human fetal tissue have been reported – specifically fetal explant cultures and generation of bud tip progenitor organoids that can be expanded, and will be described in later chapters.

3D culture systems utilizing primary human cells are especially amenable to studying the proliferative and stem-like properties of cell populations in colony formation assays and for studying patient specific disease states. For example, organoids generated from colon biopsies of CF patients have been used in a forskolin-induced swelling assay to study functionality of the CFTR in individual patients (117).

More recently, some reports, including one described here, have generated 3D organoid model systems utilizing human fetal lung tissue or employed whole tissue explants of the native fetal lung (19, 20, 119, 120), which have been used to identify signaling networks important for cell fate specification during lung development.

Biological scaffolds

While isolated human lung cells are able to survive in 2D on plastic and in 3D Matrigel or collagen environments, the ECM in the native lung is intricate and complex, and shapes the cell microenvironment by providing structural cues, regulating access to soluble growth factors, and controlling cell adhesion and migration (37, 121-123). The ECM is extensively remodeled in disease states such as pulmonary fibrosis, COPD and emphysema (124). The ability to model cell-ECM interactions to better understand disease, and the promise of manufacturing a new lung for patients with end stage lung disease, has led researchers to develop methods for decellularizing rodent and human lungs or lung sections and reseeding human lung cells onto the matrix. Multiple techniques have been developed to decellularize lungs (125-130), though the ability to decellularize human lungs currently far exceeds the success of recellularizing them to generate mature, functional cell types seen in the native lung.

Early successes at recellularizing scaffolds relied on seeding scaffolds with primary mesenchymal stem cells (131), which are able to engraft and proliferate in the scaffold. More recently, basal stem cells isolated from donated human lungs and expanded in culture were seeded onto decellularized rat lungs where they maintained a basal cell identity (132). Despite advances in the creation of decellularized scaffolds, significant advances improving seeding cells on the scaffolds, and improving cell survival and differentiation/maturation will be critical for creating functional recellularized lungs for organ transplant, or for generating functional models of the lung to study basic lung biology.

Bioengineered scaffolds and niches

While decellularized scaffolds are a powerful tool, the stiffness, protein/carbohydrate composition and residual growth factors within scaffolds can vary between species. Further, the composition of scaffolds is dependent on the decellularization method, which can lead to significant differences in the properties of the native matrices and in cell seeding efficiency (133). In an effort to create a welcoming environment for human lung cells to thrive, researchers have turned to bioengineering to create a variety of structural scaffolds and microenvironments to mimic the lung. Several studies have attempted to seed primary lung cells from mice or rats onto scaffolds comprised of naturally occurring ECM, such as Matrigel (134, 135) and collagen (136), or on synthetic materials including polyglycolic acid (PGA), polylactic-coglycolic acid (PLGA) and pluronic F-127 (137-139). For example, novel manufacturing techniques have been reported to create a thin wafer composed of

nanofibres of varying diameter, which structurally resemble the basement membrane of the bronchial epithelium. The authors were able to successfully co-culture human bronchial cells and fibroblasts at the ALI on these wafers, showing that the fibroblast cells invaded the structure while bronchial cells grew and matured at the ALI (140). Additionally, lung mesenchymal cell types taken from fetal lung tissue or derived from hPSCs have been seeded onto alginate beads coated in collagen and cultured in a bioreactor, which resulted in the formation of mesenchymal organoids that have a structural phenotype similar to the alveolar regions of the native lung (141).While seeding cells on scaffolds is an excellent tool for probing questions related to the effect of the ECM on cell differentiation and repair, these models may not recapitulate structural or mechanical forces that are constantly at play in the breathing lung.

In order to model mechanical forces of the lung such as stretch, bioengineers have developed innovative microsystems, called "lungs-on-a-chip," in which the location and mechanical force placed on each cell are tightly defined to model both the small airway and the stretch experienced in the alveolar compartment (142, 143). In the alveolar lung-on-a-chip, structural, functional and mechanical elements of the interface between the alveoli and the capillary network were modeled by microfabricating a microfluidic device that contains two channels separated by a thin flexible membrane coated with ECM proteins. Vascular endothelial cells were cultured in one chamber and alveolar epithelial cells in the other, which could also be exposed to air. Controlled regulation of stretching of the membrane and fluid flow through the vascular chamber allowed study of the passage of materials and nutrients from the alveolar region to the vascular compartment (143). Follow up work with this device has shown that

mechanical forces generated during breathing can play a role in the development of vascular leakage associated with pulmonary edema in response to interleukin-2 (IL-2) treatment (144). More recently, a "small-airway-on-a-chip" has been developed, which houses a bronchial epithelium with mucociliary differentiation in one chamber exposed to air, and with the vascular endothelial layer in the opposite, fluid filled, chamber. Proof of concept experiments with this device showed that treating cells with IL-13 resulted in goblet cell hyperplasia and decreased cilliary function. Additionally, when cells from patients with COPD were used, the system was able to recapitulate some aspects of the disease state (142). To date, these model systems are the only systems in incorporate controlled mechanical stretching and liquid flow control to recapitulate physiological function at the organ level.

Benefits and Drawbacks of lung models utilizing primary human cells

Systems to culture primary human lung cells have promoted investigation into mechanisms regulating homeostasis, repair and disease in the upper airway and to some extent, the alveolus. *In vitro* models employing primary tissue provide certainty that the cell types obtained are biologically relevant and are patient specific. Recent advances in co-culture conditions with mesenchymal cells and with various types of ECM components have made these systems more attractive as models for studying complex interactions between cell types in chronic lung disease, such as COPD, ILD or cystic fibrosis. It would also be possible in this context to generate models from individuals with chronic diseases such as CF that include elements of the diseased microenvironment of the lung, including the general microbiota and/or multiple strains of

antibiotic resistant bacteria, which may have drastic implications for how well that patient responds to certain drug therapies.

Despite advantages, these systems also have caveats and drawbacks. For example, *in vitro* grown cells are supported by an artificial niche, which may not closely model the *in vivo* environment. Availability of primary human tissue is limited, and, in most cases, each laboratory will have unique cell lines. Thus, the use of non-standardized cell lines makes it is difficult or impossible to control for donor heterogeneity as a biological variable, and also renders these systems ill suited for high throughput studies like drug screens. Recent advances in differentiating human pluripotent stem cells to specific lung lineages, discussed below, have alleviated some of the concerns over donor heterogeneity and tissue availability; however, the use of primary human lung cells will no doubt continue to play an important role in evaluating disease and patient specific physiology.

In the future, building complexity within *in vitro* lung models in an attempt to provide a more comprehensive picture of lung function will increase the utility of these systems. For example, developing systems that include epithelial, neural, mesenchymal and human immune cells will provide increased biological complexity in a defined and experimentally tractable environment. This may be particularly useful for diseases like IPF and COPD, were better understanding epithelial-mesenchymal interactions may help improve our understanding of disease.

1.7 Directed differentiation of human pluripotent stem cells to generate *in vitro* lung models

Another area of scientific inquiry where there has been a flurry of progress developing *in vitro* models of the human lung has been in the field of human pluripotent stem cells (hPSCs), which include both embryonic and induced pluripotent stem cells (145-153). Most approaches to differentiate hPSCs into lung lineages follow an experimental paradigm called directed differentiation (9, 154).Directed differentiation aims to mimic *in vitro* the signaling cues that control cell fate decisions in the embryo in a step-by-step fashion. Although slightly different methods exist, most of the studies differentiating hPSCs into lung lineages share a general experimental framework. That is, hPSCs are coaxed to differentiate in a stepwise process through several cellular states that mimic gastrulation, patterning of the lung field, and specification of a lung lineage. In the case of the lung epithelium, this includes induction of the definitive endoderm (gastrulation), followed by anterior foregut endoderm (patterning), and finally induction of a lung identity (specification) (reviewed in (154-156). To date, researchers have created hPSC-derived lung cell models in many of the same contexts as primary human epithelial tissues discussed above, including 2D culture models, ALI cultures, 3D environments, and seeding cells onto biological or synthetic matrices.

Monolayer cultures

The ability to differentiate lung lineages from hPSCs was built on seminal work demonstrating how these cells could be instructed to generate endoderm (mimicking gastrulation), followed by foregut specification (mimicking embryonic patterning) (157,

158). Thus, the first reports of lung lineage differentiation from hPSCs took place in a 2D environment, and gave rise to progenitor cells that were similar to those induced early during lung development in the embryo (147, 149). In order to mature these hPSC-derived progenitors, culture conditions were identified that allowed many cells within the culture to stochastically differentiate into lung-epithelium specific cell types, including those from both the airway and the alveoli (145). More recently, this process has been optimized by borrowing from developmental cues in Xenopus, showing that a pulse of Retinoic Acid can drastically affect the expression of lung specific marker NKX2.1 (159, 160).However, 2D cultures generally show a lack of organization, with different cell types randomly distributed throughout the culture. To achieve an enriched airway-like cell population of lung progenitors, hPSC-derived lung progenitors were cultured at the ALI for many weeks. This promoted differentiation of multiciliated, goblet and basal cells, including some functional beating ciliated cells (148). While the majority of studies generating lung lineages from hPSCs have shown success generating proximal airway cell types, differentiation of AECI-like and AECII-like cells has also been reported. This required a step to enrich AECII-like cells from 2D iPSC-derived lung cultures followed by subsequent culturing in a bioreactor that alternately exposed cells to the air and (152).

The use of patient specific iPSCs has been a boon to research on genetic diseases affecting the lung, particularly CF. Mature lung cells from CF iPSCs were recently generated (149). Multiple groups have used CF patient specific iPSC-derived lung cells to show that genome engineering approaches can be used for functional gene correction to the CFTR in a patient specific manner (161, 162), potentially paving the

way for the use of these systems in identifying novel therapeutic approaches such as gene corrected patient-specific tissue replacement.

Self-assembly

While 2D cultures are well suited for studying epithelial barrier function and understanding biochemical components of disease, 3D culture systems have the added advantage of beginning to recapitulate some of the complex structural environment of the native lung. By modifying the differentiation conditions, endoderm cultures can be directed to differentiate into lineage restricted 3-dimensional structures, including foregut 'spheroids' (146, 163-166), which can be grown in a 3D ECM droplet such as Matrigel. Spheroids can be expanded into "human lung organoids" (HLOs) by adding growth factors such as FGF10, which are important for embryonic lung development (5, 21, 146, 167). HLOs possessed proximal airway-like structures, including basal cells and poorly differentiated ciliated cells, but lacked secretory-like cells. HLOs included regions containing cells expressing markers of AECI and AECII cells, but generally lacked alveolar structure. Additionally, proximal epithelial structures within HLOs were surrounded by lung mesenchymal cell types including smooth muscle cells. (146). Follow up work in this system showed that HLOs were able to give rise to more mature, adult-like structures when transplanted into mice (Dye, 2016). Generation of hPSC-derived bronchospheres has also been reported (150), where hPSC-derived lung progenitor cells were purified from 2D cultures and placed in a 3D ECM which led to the formation of epithelium-only cysts. Treatment of the cysts with DAPT, a Notch inhibitor, induced the formation of neuroendocrine cells, mucous producing goblet cells and

functional multiciliated cells (150). More recent work has shown that regulation of Wnt signaling is critical for regulating proximal-distal fate specification from NKX2.1+ hPSC-derived lung progenitors(168, 169).

While the formation of proximal airway-like structures and cell types has been reported in a number of contexts, there are fewer reports of the differentiation of alveolar cell types (145, 153), and only two reports have demonstrated clear evidence of alveolar cell-type differentiation in a 3D context (151, 170). In one report, hPSC-derived lung progenitors were purified using a cell-sorting strategy and co-cultured in a 3D environment with primary fetal human lung fibroblasts, which resulted in the generation of AECI-like and AECII-like alveolar epithelial cell types (151). In another, NKX2.1+ cells were isolated from a hPSC-derived lung progenitor culture using a reporter cell line, and cultured in a high Wnt environment. These cultures gave rise to functional AECII cells, with mature lamellar bodies that secreted and took up surfactant proteins (170).

Biological and Synthetic Scaffolds

Many researchers have turned to biological and synthetic scaffolds to attempt to mature hPSC-derived lung progenitors into functional adult cell types. For example, hPSC-derived AECII-like cells can engraft and proliferate on decellularized rat and human lung scaffolds, and these cells expressed markers of mature epithelial cell types after seeding (171). Similarly, hPSC-derived lung progenitors have been shown to seed onto decellularized lung slices or whole organs, (172) which is a first step in the effort to generate functional whole organs from patient specific iPSCs for transplantation. A

bioengineered synthetic scaffold was used to enhance engraftment and growth of hPSC-derived lung spheroids in mice, leading to tube-like airway structures with basal cells, functional ciliated cells, secretory cells, and many mature mesenchymal cell types surrounding the epithelial tubes (173).

Benefits and Drawbacks of stem cell derived in vitro lung model systems

The use of hPSCs to model human lung development and disease represents an enormous potential for generating patient specific models, and for overcoming barriers of tissue availability, donor heterogeneity and tissue quality. An ideal stem-cell derived system would be robust and reproducible with the potential for scale-up for cell therapy or large drug screens. Despite these advantages, the field has a number of areas that need to be improved in order to maximize the usefulness of these model systems. Although many reports use similar strategies to generate lung progenitor cells, no consensus has been reached for a standardized method and therefore, small differences persist. For example, subtle difference in the growth factors used can lead to drastic changes on experimental outcomes. Additionally, hPSC-derived tissues tend to remain immature, similar to fetal tissue, and don't generally mature to an adult-like state *in vitro.* Thus, more work needs to be done to understand how to mature hPSC-derived lung tissue and to demonstrate definitively that stem-cell derived lung cells have the same functionality as well as similar chromatin states and gene expression profiles as their *in vivo* counterparts before meaningful conclusions can be drawn from *in vitro* studies of adult cell behaviors.

1.8 The future of *in vitro* lung model systems

Significant progress has been made in the past decade to develop new *in vitro* models of the human lung. However, many major biological questions and technical advances remain to be conquered. While the majority of the systems reviewed here attempt to model either proximal (airway) or distal (alveolar) cell types, few systems are able to integrate both compartments of the lung into one model. Moreover, understanding how to maintain long-term growth, and control differentiation state of all human lung epithelial cell types *in vitro* is still a major challenge. Developing a deep and clear understanding of how the physical and chemical environment controls how progenitor cells make cell fate decisions, become mature and functional and are maintained both *in vitro* and *in vivo* is also a major challenge. In this light it should be pointed out that, despite the power of *in vitro* human models, *in vivo* animal models will be an integral part of forwarding our understanding of lung biology.

For this thesis work, I have focused on the regulation of bud tip progenitors during human lung development. The elucidation of how this population of fetal progenitors is maintained or the mechanisms regulating their differentiation may open doors to therapeutic targets to treat lungs after premature birth, and generates a new model system to study human lung development in vitro.

1.9 Figures

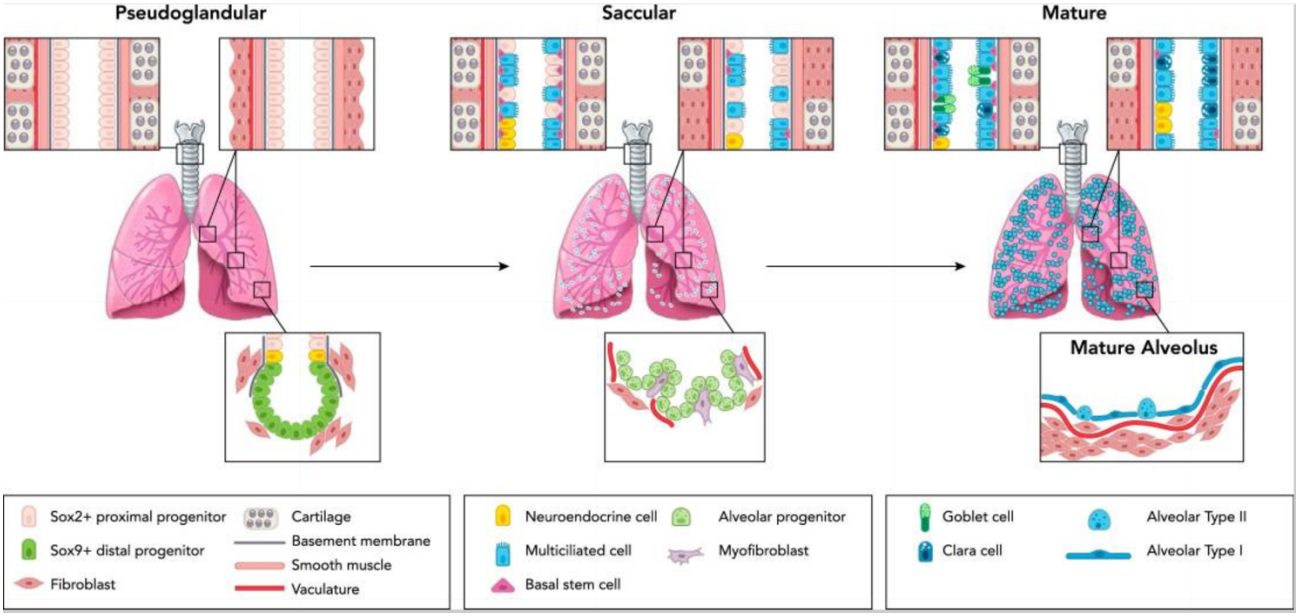


Figure 1-1. Development and anatomy of the human lung

(Left Panel) Branching morphogenesis occurs during the pseudoglandular stage of lung development. At this time point, the proximal epithelial regions contain Sox2+ progenitor cells, whereas distal epithelial bud tips contain Sox9+ progenitors. As development progresses, Sox9+ distal bud tip progenitors give rise to alveolar specific progenitor cells, and rudimentary saccules form (Middle Panel). Additionally, the proximal airway begins to form mature ciliated and secretory cells at this time. As the lung matures, saccules become mature alveolar structures comprised of AECI and AECII cells (Right Panel).

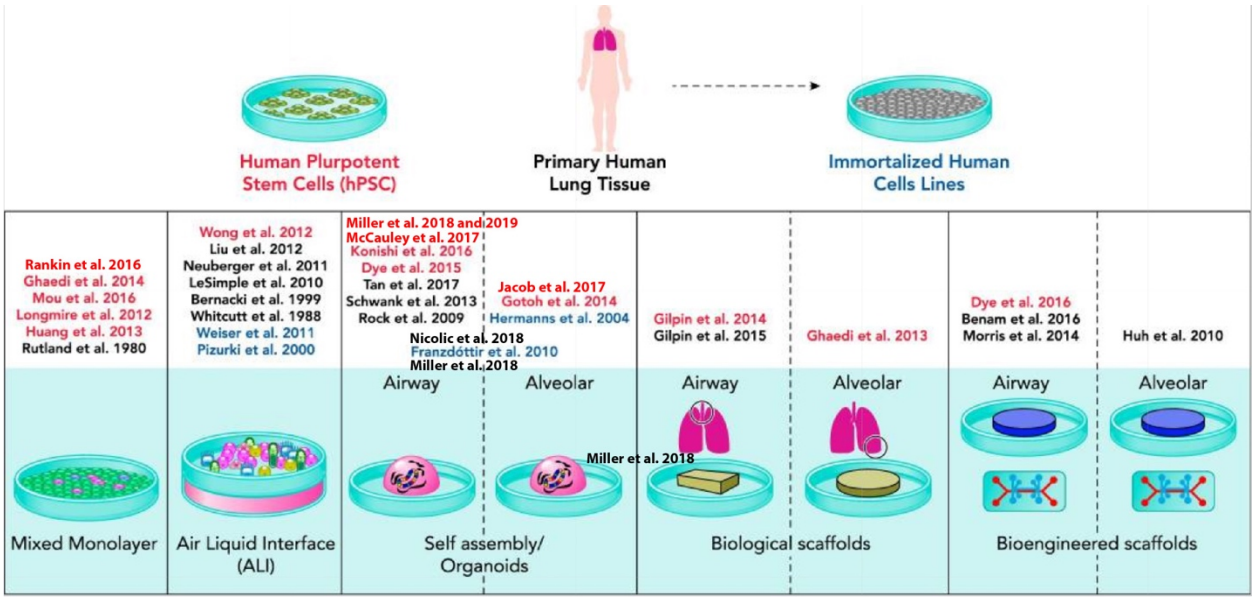


Figure 1-2 *In vitro* model systems of the human lung

A variety of *in vitro* human model systems have been developed using primary lung tissue (examples denoted in black text), lung cells from immortalized human cell lines (examples denoted in blue text), and lung specific cells derived from hPSCs (examples denoted in pink text). These systems have been developed in a variety of contexts, including monolayer, ALI, or in a variety of 3D and engineered environments in order to recapitulate elements of the native human lung.

1.10 References

1. Alvarado A, Arce I (2016) Metabolic Functions of the Lung, Disorders and Associated Pathologies. *J Clin Med Res* 8(10):689–700.

2. Bernhard W (2016) Lung surfactant: Function and composition in the context of development and respiratory physiology. *Ann Anat* 208:146–150.

3. Matute-Bello G, Frevert CW, Martin TR (2008) Animal models of acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 295(3):L379–99.

4. Cardoso WV (2001) Molecular regulation of lung development. *Annu Rev Physiol* 63:471–494.

5. Morrisey EE, Hogan BLM (2010) Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. *Dev Cell* 18(1):8–23.

6. Metzger RJ, Klein OD, Martin GR, Krasnow MA (2008) The branching programme of mouse lung development. *Nature* 453(7196):745–750.

7. Moore BB, Hogaboam CM (2008) Murine models of pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 294(2):L152–60.

8. Ruggeri BA, Camp F, Miknyoczki S (2014) Animal models of disease: pre-clinical animal models of cancer and their applications and utility in drug discovery. *Biochem Pharmacol* 87(1):150–161.

9. Aurora M, Spence JR (2016) hPSC-derived lung and intestinal organoids as models of human fetal tissue. *Developmental Biology* 420(2):230–238.

10. Rackley CR, Stripp BR (2012) Building and maintaining the epithelium of the lung. *J Clin Invest* 122(8):2724–2730.

11. Mak IW, Evaniew N, Ghert M (2014) Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res* 6(2):114–118.

12. Myllärniemi M, Kaarteenaho R (2015) Pharmacological treatment of idiopathic pulmonary fibrosis - preclinical and clinical studies of pirfenidone, nintedanib, and N-acetylcysteine. *Eur Clin Respir J* 2. doi:10.3402/ecrj.v2.26385.

13. Perrin S (2014) Preclinical research: Make mouse studies work. *Nature*

507(7493):423–425.

14. Domyan ET, Sun X (2010) Patterning and plasticity in development of the respiratory lineage. *Dev Dyn* 240(3):477–485.

15. de Bakker BS, et al. (2016) An interactive three-dimensional digital atlas and quantitative database of human development. *Science* 354(6315). doi:10.1126/science.aag0053.

16. Rawlins EL, Clark CP, Xue Y, Hogan BLM (2009) The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development* 136(22):3741–3745.

17. Yang Y, et al. (2018) Spatial-Temporal Lineage Restrictions of Embryonic p63+ Progenitors Establish Distinct Stem Cell Pools in Adult Airways. *Dev Cell* 44(6):752–761.e4.

18. Chang DR, et al. (2013) Lung epithelial branching program antagonizes alveolar differentiation. *Proceedings of the National Academy of Sciences* 110(45):18042–18051.

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

20. Miller AJ, et al. (2017) In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. *Stem Cell Reports*. doi:10.1016/j.stemcr.2017.11.012.

21. Miller AJ, et al. (2019) Generation of lung organoids from human pluripotent stem cells in vitro. *Nat Protoc*.

22. Herriges M, Morrisey EE (2014) Lung development: orchestrating the generation and regeneration of a complex organ. *Development* 141(3):502–513.

23. Rawlins EL (2010) The building blocks of mammalian lung development. *Dev Dyn* 240(3):463–476.

24. Rawlins EL, Ostrowski LE, Randell SH, Hogan BLM Lung development and repair: Contribution of the ciliated lineage. *pnasorg*.

25. Desai TJ, Brownfield DG, Krasnow MA (2014) Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature*. doi:10.1038/nature12930.

26. Chao C-M, Moiseenko A, Zimmer K-P, Bellusci S (2016) Alveologenesis: key cellular players and fibroblast growth factor 10 signaling. *Mol Cell Pediatr* 3(1):371.

27. Gonzalez RF, Allen L, Gonzales L, Ballard PL, Dobbs LG (2010) HTII-280, a biomarker specific to the apical plasma membrane of human lung alveolar type II cells. *J Histochem Cytochem* 58(10):891–901.

28. Li C, et al. (2019) Secondary crest myofibroblast PDGFRα controls the elastogenesis pathway via a secondary tier of signaling networks during alveologenesis. *Development* 146(15):dev176354.

29. Green J, Endale M, Auer H, Perl A-KT Diversity of Interstitial Lung Fibroblasts Is Regulated by Platelet-Derived Growth Factor Receptor α Kinase Activity. *American Journal of Respiratory Cell and Molecular Biology*.

30. McCulley D, Wienhold M, Sun X (2015) The pulmonary mesenchyme directs lung development. *Curr Opin Genet Dev* 32:98–105.

31. Tovar JA (2012) Congenital Diaphragmatic Hernia. *Orphanet J Rare Dis* 7(1):1.

32. BHATT AJ, et al. Disrupted Pulmonary Vasculature and Decreased Vascular Endothelial Growth Factor, Flt-1, and TIE-2 in Human Infants Dying with Bronchopulmonary Dysplasia. *Am J Respir Crit Care Med*.

33. Bower DV, et al. (2014) Airway branching has conserved needs for local parasympathetic innervation but not neurotransmission. *BMC Biol* 12:92.

34. Freem LJ, et al. (2010) The intrinsic innervation of the lung is derived from neural crest cells as shown by optical projection tomography in Wnt1-Cre;YFP reporter mice. *J Anat* 217(6):651–664.

35. Burns AJ, Delalande J-MM, Le Douarin NM (2002) In ovo transplantation of enteric nervous system precursors from vagal to sacral neural crest results in extensive hindgut colonisation. *Development* 129(12):2785–2796.

36. Varner VD, Nelson CM (2014) Cellular and physical mechanisms of branching morphogenesis. *Development* 141(14):2750–2759.

37. Patel VN, Pineda DL, Hoffman MP (2016) The function of heparan sulfate during branching morphogenesis. *Matrix Biol*. doi:10.1016/j.matbio.2016.09.004.

38. Rockich BE, et al. (2013) Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. *Proceedings of the National Academy of Sciences* 110(47):E4456–64.

39. Harunaga JS, Doyle AD, Yamada KM (2014) Local and global dynamics of the basement membrane during branching morphogenesis require protease activity and actomyosin contractility. *Developmental Biology* 394(2):197–205.

40. 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):2711–2722.

41. Akram KM, et al. (2019) Live imaging of alveologenesis in precision-cut lung slices reveals dynamic epithelial cell behaviour. *Nat Commun*:1–16.

42. Hrycaj SM, et al. (2018) Hox5genes direct elastin network formation during alveologenesis by regulating myofibroblast adhesion. *Proceedings of the National Academy of Sciences* 115(45):E10605–E10614.

43. Branchfield K, et al. (2016) A three-dimensional study of alveologenesis in mouse lung. *Developmental Biology* 409(2):429–441.

44. Holden VI, Breen P, Houle S, Dozois CM, Bachman MA (2016) Klebsiella pneumoniae Siderophores Induce Inflammation, Bacterial Dissemination, and HIF-1α Stabilization during Pneumonia. *mBio* 7(5). doi:10.1128/mBio.01397-16.

45. Hackett T-L, et al. (2008) Characterization of side population cells from human airway epithelium. *STEM CELLS* 26(10):2576–2585.

46. Boers JE, Ambergen AW, Thunnissen FB (1999) Number and proliferation of clara cells in normal human airway epithelium. *Am J Respir Crit Care Med* 159(5 Pt 1):1585–1591.

47. Wolff RK (1986) Effects of airborne pollutants on mucociliary clearance. *Environ Health Perspect* 66:223–237.

48. Tilley AE, Walters MS, Shaykhiev R, Crystal RG (2015) Cilia dysfunction in lung disease. *Annu Rev Physiol* 77:379–406.

49. Knowles MR, Boucher RC (2002) Mucus clearance as a primary innate defense mechanism for mammalian airways. *J Clin Invest* 109(5):571–577.

50. Hoegger MJ, et al. (2014) Impaired mucus detachment disrupts mucociliary transport in a piglet model of cystic fibrosis. *Science* 345(6198):818–822.

51. Shah VS, et al. (2016) Airway acidification initiates host defense abnormalities in cystic fibrosis mice. *Science* 351(6272):503–507.

52. Cutting GR (2015) Cystic fibrosis genetics: from molecular understanding to clinical application. *Nat Rev Genet* 16(1):45–56.

53. Boers JE, Brok den JL, Koudstaal J, Arends JW, Thunnissen FB (1996) Number and proliferation of neuroendocrine cells in normal human airway epithelium. *Am J Respir Crit Care Med* 154(3 Pt 1):758–763.

54. Cutz E, Perrin DG, Pan J, Haas EA, Krous HF (2007) Pulmonary neuroendocrine cells and neuroepithelial bodies in sudden infant death syndrome: potential markers of airway chemoreceptor dysfunction. *Pediatr Dev*

Pathol 10(2):106–116.

55. Perrin DG, McDonald TJ, Cutz E (1991) Hyperplasia of bombesin-immunoreactive pulmonary neuroendocrine cells and neuroepithelial bodies in sudden infant death syndrome. *Pediatr Pathol* 11(3):431–447.

56. Gu X, et al. (2014) Chemosensory functions for pulmonary neuroendocrine cells. *American Journal of Respiratory Cell and Molecular Biology* 50(3):637–646.

57. Branchfield K, et al. (2016) Pulmonary neuroendocrine cells function as airway sensors to control lung immune response. *Science*. doi:10.1126/science.aad7969.

58. Rock JR, et al. (2009) Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proceedings of the National Academy of Sciences* 106(31):12771–12775.

59. Hong KU, Reynolds SD, Watkins S, Fuchs E, Stripp BR (2004) Basal cells are a multipotent progenitor capable of renewing the bronchial epithelium. *AJPA* 164(2):577–588.

60. Mou H, et al. (2016) Dual SMAD Signaling Inhibition Enables Long-Term Expansion of Diverse Epithelial Basal Cells. *Cell Stem Cell*. doi:10.1016/j.stem.2016.05.012.

61. Barkauskas CE, et al. (2013) Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* 123(7):3025–3036.

62. Vaughan AE, et al. (2015) Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. *Nature* 517(7536):621–625.

63. Barnes PJ, Shapiro SD, Pauwels RA (2003) Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J* 22(4):672–688.

64. Bateman ED, Turner-Warwick M, Adelmann-Grill BC (1981) Immunohistochemical study of collagen types in human foetal lung and fibrotic lung disease. *Thorax* 36(9):645–653.

65. Snider GL (1981) Collagen concentration and rates of synthesis in idiopathic pulmonary fibrosis. *Am Rev Respir Dis* 124(3):341–342.

66. Glasser SW, et al. (2016) Mechanisms of Lung Fibrosis Resolution. *Am J Pathol* 186(5):1066–1077.

67. Ebina M, Takahashi T, Chiba T, Motomiya M (1993) Cellular Hypertrophy and Hyperplasia of Airway Smooth Muscles Underlying Bronchial Asthma: A 3-D Morphometric Study. *Am Rev Respir Dis* 148(3):720–726.

68. Sakao S, Tatsumi K, Voelkel NF (2009) Endothelial cells and pulmonary arterial hypertension: apoptosis, proliferation, interaction and transdifferentiation. *Respir Res* 10:95.

69. Rawlins EL, Ostrowski LE, Randell SH, Hogan BLM (2007) Lung development and repair: contribution of the ciliated lineage. *Proc Natl Acad Sci USA* 104(2):410–417.

70. Beers MF, Morrisey EE (2011) The three R's of lung health and disease: repair, remodeling, and regeneration. *J Clin Invest* 121(6):2065–2073.

71. Whitsett JA, Kalinichenko VV (2011) Notch and basal cells take center stage during airway epithelial regeneration. *Cell Stem Cell* 8(6):597–598.

72. Kotton DN, Morrisey EE (2014) Lung regeneration: mechanisms, applications and emerging stem cell populations. *Nature Medicine* 20(8):822–832.

73. Rock JR, et al. (2011) Notch-dependent differentiation of adult airway basal stem cells. *Cell Stem Cell* 8(6):639–648.

74. Pardo-Saganta A, et al. (2015) Injury induces direct lineage segregation of functionally distinct airway basal stem/progenitor cell subpopulations. *Cell Stem Cell* 16(2):184–197.

75. Kim CFB, et al. (2005) Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* 121(6):823–835.

76. Zacharias WJ, et al. (2018) Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* 555(7695):251–255.

77. Frank DB, et al. (2019) Early lineage specification defines alveolar epithelial ontogeny in the murine lung. *Proceedings of the National Academy of Sciences* 116(10):4362–4371.

78. Chen L, Acciani T, Le Cras T, Lutzko C, Perl A-KT (2012) Dynamic regulation of platelet-derived growth factor receptor α expression in alveolar fibroblasts during realveolarization. *American Journal of Respiratory Cell and Molecular Biology* 47(4):517–527.

79. Peng T, et al. (2015) Hedgehog actively maintains adult lung quiescence and regulates repair and regeneration. *Nature* 526(7574):578–582.

80. Tata PR, et al. (2013) Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature* 503(7475):218–223.

81. Rawlins EL, et al. (2009) The Role of Scgb1a1+ Clara Cells in the Long-Term Maintenance and Repair of Lung Airway, but Not Alveolar, Epithelium. *Cell Stem Cell* 4(6):525–534.

82. Jain R, et al. (2015) Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. *Nat Commun* 6:6727.

83. Mercer RR, Russell ML, Roggli VL, Crapo JD (1994) Cell number and distribution in human and rat airways. *American Journal of Respiratory Cell and Molecular Biology* 10(6):613–624.

84. Pardo-Saganta A, Law BM, Gonzalez-Celeiro M, Vinarsky V, Rajagopal J (2013) Ciliated cells of pseudostratified airway epithelium do not become mucous cells after ovalbumin challenge. *American Journal of Respiratory Cell and Molecular Biology* 48(3):364–373.

85. Kim HY, Nelson CM (2012) Extracellular matrix and cytoskeletal dynamics during branching morphogenesis. *organogenesis* 8(2):56–64.

86. Gjorevski N, et al. (2016) Designer matrices for intestinal stem cell and organoid culture. *Nature* 539(7630):560–564.

87. Workman MJ, et al. (2016) Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. *Nature Medicine*. doi:10.1038/nm.4233.

88. Gruenert DC, Willems M, Cassiman JJ, Frizzell RA (2004) Established cell lines used in cystic fibrosis research. *J Cyst Fibros* 3 Suppl 2:191–196.

89. Fulcher ML, et al. (2009) Novel human bronchial epithelial cell lines for cystic fibrosis research. *Am J Physiol Lung Cell Mol Physiol* 296(1):L82–91.

90. Pizurki L, et al. (2000) Cystic fibrosis transmembrane conductance regulator does not affect neutrophil migration across cystic fibrosis airway epithelial monolayers. *AJPA* 156(4):1407–1416.

91. Molina SA, et al. (2015) Junctional abnormalities in human airway epithelial cells expressing F508del CFTR. *AJP: Lung Cellular and Molecular Physiology* 309(5):L475–87.

92. Weiser N, et al. (2011) Paracellular permeability of bronchial epithelium is controlled by CFTR. *Cell Physiol Biochem* 28(2):289–296.

93. Ren H, Birch NP, Suresh V (2016) An Optimised Human Cell Culture Model for Alveolar Epithelial Transport. *PLoS ONE* 11(10):e0165225.

94. Hermanns MI, Unger RE, Kehe K, Peters K, Kirkpatrick CJ (2004) Lung epithelial cell lines in coculture with human pulmonary microvascular endothelial cells: development of an alveolo-capillary barrier in vitro. *Lab Invest* 84(6):736–752.

95. Franzdóttir SR, et al. (2010) Airway branching morphogenesis in three

dimensional culture. *Respir Res* 11:162.

96. Rutland J, Cole PJ (1980) Non-invasive sampling of nasal cilia for measurement of beat frequency and study of ultrastructure. *Lancet* 2(8194):564–565.

97. Toskala E, Haataja J, Shirasaki H, Rautiainen M (2005) Culture of cells harvested with nasal brushing: a method for evaluating ciliary function. *Rhinology* 43(2):121–124.

98. Pifferi M, et al. (2009) Simplified cell culture method for the diagnosis of atypical primary ciliary dyskinesia. *Thorax* 64(12):1077–1081.

99. Whitcutt MJ, Adler KB, Wu R (1988) A biphasic chamber system for maintaining polarity of differentiation of cultured respiratory tract epithelial cells. *In Vitro Cell Dev Biol* 24(5):420–428.

100. Bals R, Beisswenger C, Blouquit S, Chinet T (2004) Isolation and air-liquid interface culture of human large airway and bronchiolar epithelial cells. *J Cyst Fibros* 3 Suppl 2:49–51.

101. Yamaya M, Finkbeiner WE, Chun SY, Widdicombe JH (1992) Differentiated structure and function of cultures from human tracheal epithelium. *Am J Physiol* 262(6 Pt 1):L713–24.

102. Liu X, et al. (2012) ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. *Am J Pathol* 180(2):599–607.

103. Bernacki SH, et al. (1999) Mucin gene expression during differentiation of human airway epithelia in vitro. Muc4 and muc5b are strongly induced. *American Journal of Respiratory Cell and Molecular Biology* 20(4):595–604.

104. Fulcher ML, Randell SH (2013) Human nasal and tracheo-bronchial respiratory epithelial cell culture. *Methods Mol Biol* 945:109–121.

105. Tadokoro T, Gao X, Hong CC, Hotten D, Hogan BLM (2016) BMP signaling and cellular dynamics during regeneration of airway epithelium from basal progenitors. *Development* 143(5):764–773.

106. LeSimple P, Liao J, Robert R, Gruenert DC, Hanrahan JW (2010) Cystic fibrosis transmembrane conductance regulator trafficking modulates the barrier function of airway epithelial cell monolayers. *J Physiol (Lond)* 588(Pt 8):1195–1209.

107. Neuberger T, Burton B, Clark H, Van Goor F (2011) Use of primary cultures of human bronchial epithelial cells isolated from cystic fibrosis patients for the pre-clinical testing of CFTR modulators. *Methods Mol Biol* 741:39–54.

108. Randell SH, Fulcher ML, O'Neal W, Olsen JC (2011) Primary epithelial cell models for cystic fibrosis research. *Methods Mol Biol* 742:285–310.

109. Davis PB, Yasothan U, Kirkpatrick P (2012) Ivacaftor. *Nat Rev Drug Discov* 11(5):349–350.

110. Van Goor F, et al. (2009) Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proceedings of the National Academy of Sciences* 106(44):18825–18830.

111. Schamberger AC, Staab-Weijnitz CA, Mise-Racek N, Eickelberg O (2015) Cigarette smoke alters primary human bronchial epithelial cell differentiation at the air-liquid interface. *Sci Rep* 5:8163.

112. Brune KA, et al. (2016) HIV Impairs Lung Epithelial Integrity and Enters the Epithelium to Promote Chronic Lung Inflammation. *PLoS ONE* 11(3):e0149679.

113. Creager HM, et al. (2017) In vitro exposure system for study of aerosolized influenza virus. *Virology* 500:62–70.

114. Pageau SC, Sazonova OV, Wong JY, Soto AM, Sonnenschein C (2011) The effect of stromal components on the modulation of the phenotype of human bronchial epithelial cells in 3D culture. *Biomaterials* 32(29):7169–7180.

115. Danahay H, et al. (2015) Notch2 is required for inflammatory cytokine-driven goblet cell metaplasia in the lung. *Cell Rep* 10(2):239–252.

116. Tan Q, Choi KM, Sicard D, Tschumperlin DJ (2017) Human airway organoid engineering as a step toward lung regeneration and disease modeling. *Biomaterials* 113:118–132.

117. Schwank G, et al. (2013) Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13(6):653–658.

118. Hild M, Jaffe AB (2016) Production of 3-D Airway Organoids From Primary Human Airway Basal Cells and Their Use in High-Throughput Screening. *Curr Protoc Stem Cell Biol* 37:IE.9.1–IE.9.15.

119. Nikolić MZ, Sun D, Rawlins EL (2018) Human lung development: recent progress and new challenges. *Development* 145(16):dev163485.

120. Danopoulos S, et al. (2019) Discordant roles for FGF ligands in lung branching morphogenesis between human and mouse. *The Journal of pathology* 247(2):254–265.

121. White ES (2015) Lung Extracellular Matrix and Fibroblast Function. *Ann Am Thorac Soc* 12(Supplement 1):S30–S33.

122. McGowan SE (1992) Extracellular matrix and the regulation of lung development and repair. *FASEB J* 6(11):2895–2904.

123. Vorotnikova E, et al. (2010) Extracellular matrix-derived products modulate endothelial and progenitor cell migration and proliferation in vitro and stimulate regenerative healing in vivo. *Matrix Biol* 29(8):690–700.

124. Bonnans C, Chou J, Werb Z (2014) Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol* 15(12):786–801.

125. Booth AJ, et al. (2012) Acellular normal and fibrotic human lung matrices as a culture system for in vitro investigation. *Am J Respir Crit Care Med* 186(9):866–876.

126. Guyette JP, et al. (2014) Perfusion decellularization of whole organs. *Nat Protoc* 9(6):1451–1468.

127. Ott HC, et al. (2010) Regeneration and orthotopic transplantation of a bioartificial lung. *Nature Medicine* 16(8):927–933.

128. Petersen TH, et al. (2010) Tissue-engineered lungs for in vivo implantation. *Science* 329(5991):538–541.

129. Gilpin SE, Ott HC (2015) Using nature's platform to engineer bio-artificial lungs. *Ann Am Thorac Soc* 12 Suppl 1:S45–9.

130. Wagner DE, et al. (2014) Comparative decellularization and recellularization of normal versus emphysematous human lungs. *Biomaterials* 35(10):3281–3297.

131. Bonvillain RW, et al. (2012) A nonhuman primate model of lung regeneration: detergent-mediated decellularization and initial in vitro recellularization with mesenchymal stem cells. *Tissue Eng Part A* 18(23-24):2437–2452.

132. Gilpin SE, Charest JM, Ren X, Ott HC (2016) Bioengineering Lungs for Transplantation. *Thorac Surg Clin* 26(2):163–171.

133. Balestrini JL, et al. (2016) Comparative biology of decellularized lung matrix: Implications of species mismatch in regenerative medicine. *Biomaterials* 102:220–230.

134. Andrade CF, Wong AP, Waddell TK, Keshavjee S, Liu M (2007) Cell-based tissue engineering for lung regeneration. *Am J Physiol Lung Cell Mol Physiol* 292(2):L510–8.

135. Mondrinos MJ, Koutzaki S, Lelkes PI, Finck CM (2007) A tissue-engineered model of fetal distal lung tissue. *Am J Physiol Lung Cell Mol Physiol* 293(3):L639–50.

136. O'Leary C, et al. (2016) The development of a tissue-engineered tracheobronchial epithelial model using a bilayered collagen-hyaluronate scaffold. *Biomaterials* 85:111–127.

137. Cortiella J, et al. (2006) Tissue-engineered lung: an in vivo and in vitro comparison of polyglycolic acid and pluronic F-127 hydrogel/somatic lung progenitor cell constructs to support tissue growth. *Tissue Eng* 12(5):1213–1225.

138. Prakash YS, Tschumperlin DJ, Stenmark KR (2015) Coming to terms with tissue engineering and regenerative medicine in the lung. *AJP: Lung Cellular and Molecular Physiology* 309(7):L625–38.

139. Trecartin A, et al. (2016) Establishing Proximal and Distal Regional Identities in Murine and Human Tissue-Engineered Lung and Trachea. *Tissue Eng Part C Methods* 22(11):1049–1057.

140. Morris GE, et al. (2014) A novel electrospun biphasic scaffold provides optimal three-dimensional topography for in vitro co-culture of airway epithelial and fibroblast cells. *Biofabrication* 6(3):035014.

141. Wilkinson DC, et al. (2016) Development of a Three-Dimensional Bioengineering Technology to Generate Lung Tissue for Personalized Disease Modeling. *Stem Cells Translational Medicine*. doi:10.5966/sctm.2016-0192.

142. Benam KH, et al. (2016) Small airway-on-a-chip enables analysis of human lung inflammation and drug responses in vitro. *Nature Publishing Group* 13(2):151–157.

143. Huh D, et al. (2010) Reconstituting organ-level lung functions on a chip. *Science* 328(5986):1662–1668.

144. Huh D, et al. (2012) A human disease model of drug toxicity-induced pulmonary edema in a lung-on-a-chip microdevice. *Sci Transl Med* 4(159):159ra147.

145. Huang SXL, et al. (2013) Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nature Biotechnology*. doi:10.1038/nbt.2754.

146. Dye BR, et al. (2015) In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* 4. doi:10.7554/eLife.05098.

147. Longmire TA, et al. (2012) Efficient Derivation of Purified Lung and Thyroid Progenitors from Embryonic Stem Cells. *Stem Cell* 10(4):398–411.

148. Wong AP, et al. (2012) Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. *Nature Biotechnology* 30(9):875–881.

149. Mou H, et al. (2012) Generation of Multipotent Lung and Airway Progenitors from Mouse ESCs and Patient-Specific Cystic Fibrosis iPSCs. *Stem Cell* 10(4):385–397.

150. 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* 6(1):18–25.

151. Gotoh S, et al. (2014) Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. *Stem Cell Reports* 3(3):394–403.

152. Ghaedi M, et al. (2014) Alveolar epithelial differentiation of human induced pluripotent stem cells in a rotating bioreactor. *Biomaterials* 35(2):699–710.

153. Hannan NRF, Sampaziotis F, Segeritz C-P, Hanley NA, Vallier L (2015) Generation of Distal Airway Epithelium from Multipotent Human Foregut Stem Cells. *Stem Cells Dev* 24(14):1680–1690.

154. Dye BR, Miller AJ, Spence JR (2016) How to Grow a Lung: Applying Principles of Developmental Biology to Generate Lung Lineages from Human Pluripotent Stem Cells. *Curr Pathobiol Rep* 4:47–57.

155. Snoeck H-W (2015) Modeling human lung development and disease using pluripotent stem cells. *Development* 142(1):13–16.

156. Ghaedi M, Niklason LE, Williams J (2015) Development of Lung Epithelium from Induced Pluripotent Stem Cells. *Curr Transplant Rep* 2(1):81–89.

157. Green MD, et al. (2011) Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nature Biotechnology* 29(3):267–272.

158. D'Amour KA, et al. (2005) Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnology* 23(12):1534–1541.

159. Rankin SA, et al. (2016) A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification. *Cell Rep* 16(1):66–78.

160. Rankin SA, Zorn AM (2014) Gene regulatory networks governing lung specification. *J Cell Biochem* 115(8):1343–1350.

161. Firth AL, et al. (2015) Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep* 12(9):1385–1390.

162. Crane AM, et al. (2015) Targeted correction and restored function of the CFTR gene in cystic fibrosis induced pluripotent stem cells. *Stem Cell Reports* 4(4):569–577.

163. Spence JR, et al. (2011) Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* 469(7332):105–109.

164. McCracken KW, Howell JC, Wells JM, Spence JR (2011) Generating human intestinal tissue from pluripotent stem cells in vitro. *Nat Protoc* 6(12):1920–1928.

165. McCracken KW, et al. (2014) Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* 516(7531):400–404.

166. Hannan NRF, et al. (2013) Generation of multipotent foregut stem cells from human pluripotent stem cells. *Stem Cell Reports* 1(4):293–306.

167. Sekine K, et al. (1999) Fgf10 is essential for limb and lung formation. *Nat Genet* 21(1):138–141.

168. McCauley KB, 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):844–857.e6.

169. Hawkins F, et al. (2017) Prospective isolation of NKX2-1-expressing human lung progenitors derived from pluripotent stem cells. *J Clin Invest* 127(6):2277–2294.

170. Jacob A, et al. (2017) Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. *Cell Stem Cell* 21(4):472–488.e10.

171. Ghaedi M, et al. (2013) Human iPS cell-derived alveolar epithelium repopulates lung extracellular matrix. *J Clin Invest* 123(11):4950–4962.

172. Gilpin SE, et al. (2014) Enhanced lung epithelial specification of human induced pluripotent stem cells on decellularized lung matrix. *Ann Thorac Surg* 98(5):1721–9– discussion 1729.

173. Dye BR, et al. (2016) A bioengineered niche promotes in vivo engraftment and maturation of pluripotent stem cell derived human lung organoids. *Elife* 5. doi:10.7554/eLife.19732.

174. Nduka O Enemchukwu RC-ATBCTJJRGTSAJG (2016) Synthetic matrices reveal contributions of ECM biophysical and biochemical properties to epithelial morphogenesis. *J Cell Biol* 212(1):113.

Chapter 2. *In vitro* induction and *in vivo* engraftment of lung bud tip progenitor cells derived from human pluripotent stem cells

Portions of this chapter have been published:

Miller, A.J.; Hill, D.R.; Nagy, M.S.; Aoki, Y.; Dye, B.R.; Chin, A.M.; Huang, S.; Zhu, F.; White, E.S.; Lama, V.; Spence, J.R. *In Vitro* Induction and *In Vivo* Engraftment of Lung Bud Tip Progenitor Cells Derived From Human Pluripotent Stem Cells. *Stem Cell Reports*. **2018**, 10(1):101-119. Doi: 10.1016/j.stemcr.2017.11.012. PMID: 29249664.

2.1 Introduction

During development, the lung undergoes branching morphogenesis, where a series of stereotyped epithelial bifurcations give rise to the branched, tree-like architecture of the adult lung (1). A population of rapidly proliferating progenitor cells resides at the tips of the epithelium throughout the branching process ('bud tip progenitors') (2). This population, which expresses *Id2* and *Sox9* in mice, has the capability to differentiate into both mature airway and alveolar cell types. At early stages of branching morphogenesis, this population of progenitors gives rise to proximal airway cells, while at later time points these progenitors give rise to alveolar cells (2).

Studies utilizing genetic mouse models have shown that lung branching morphogenesis and proximal-distal patterning are regulated by a series of complex mesenchymal-epithelial interactions that involve multiple signaling events, transcription factors, and dynamic regulation of the physical environment (3,4). These studies have

identified major roles for several signaling pathways in these processes, including Wnt, Fibroblast Growth Factor (FGF), Bone Morphogenic Protein (BMP), Sonic Hedgehog (SHH), Retinoic Acid (RA) and Hippo signaling, among others. However, due to the complex and intertwined nature of these signaling networks, perturbations in one pathway often affect signaling activity of others (3).

These developmental principles, learned from studying model organism development, have been used as a guide to successfully direct differentiation of human pluripotent stem cells into differentiated lung lineages and 3-dimensional lung organoids (5) (6). However, specifically inducing and maintaining the epithelial bud tip progenitor cell population from hPSCs has remained elusive. For example, our own studies have shown that hPSCs can be differentiated into human lung organoids (HLOs) that possess airway-like epithelial structures and alveolar cell types; however, it was not clear if HLOs passed through a bud tip progenitor-like stage, mimicking all stages of normal development *in vivo* (7). More recent evidence from others has demonstrated that putative bud tip progenitor cells may be induced from hPSCs; however, these cells were rare and were not assessed in detail (8). Thus, generation of a robust population of bud tip progenitor cells from hPSCs would shed additional mechanistic light on how these cells are regulated, would provide a platform for further investigation into mechanisms of lung lineage cell fate specification, and would add a layer of control to existing directed differentiation protocols allowing them to pass through this developmentally important progenitor transition.

In the current study, we conducted a low-throughput screen using isolated mouse epithelial bud tip cultures to identify factors that maintained epithelial bud tip progenitors

in vitro. These conditions were also tested using isolated human fetal epithelial bud tip progenitors cultured *in vitro* and for the ability to induce a bud tip like population from hPSCs. We determined that FGF7 promoted an initial expansion of human epithelial bud tip progenitors, and that the addition of CHIR-99021 (a GSK3β inhibitor that acts to stabilize β-catenin) and All-trans Retinoic Acid (RA) (3-Factor conditions, herein referred to as '3F') were required for growth/expansion of human fetal bud tips as epithelial progenitor organoids that maintained their identity *in vitro.*

When applied to hPSC-derived foregut spheroid cultures, we observed that 3F conditions promoted growth into larger organoid structures with a patterned epithelium that had airway-like and bud tip-like domains. Bud tip-like domains could be preferentially expanded into 'bud tip organoids' using serial passaging. hPSC-derived bud tip organoids had a protein expression and transcriptional profile similar to human fetal progenitor organoids. Finally, we demonstrated that hPSC-derived epithelial bud tip organoids can engraft into an injured mouse airway and undergo multilineage differentiation. Taken together, these studies provide an improved mechanistic understanding of human lung bud tip progenitor cell regulation and establish a platform for studying the maintenance and differentiation of lung bud tip progenitor cells.

2.2 Results

Characterizing human fetal lung bud tip progenitors

In mice, epithelial bud tip progenitor cells express several transcription factors, including *Sox9*, *Nmyc* and *Id2* (2,9-12). However, recent studies have suggested that

significant differences exist between murine and human fetal bud tip progenitor cells (13,14). To confirm and extend these recent findings, we carried out an immunohistochemical analysis using well-established protein markers that are present during mouse lung development (Figure 2-1A-C, Figure 2-2) on human lungs between 10-20 weeks of gestation. We also conducted RNAsequencing on freshly isolated epithelial lung bud tips, which were dissected to remove mesenchymal cells, to identify genes that were enriched in epithelial progenitors (Figure 2-1D-E). We note that our approach using manual and enzymatic dissection techniques were unlikely to yield pure epithelial cells, and possessed a small population of associated mesenchyme. Consistent with the developing mouse lung (10,11), we observed that SOX9 is expressed in bud tip domains of the branching epithelium (Figure 2-1A, Figure 2-2A). In contrast to the developing murine lung, we observed SOX2 expression in these bud tip progenitor domains until 16 weeks of gestation, at which time SOX2 expression was lost from this population (Figure 2-1A, Figure 2-2A). We also observed expression of *ID2* by *in situ* hybridization (Figure 2-1B, Figure 2-2F), with expression becoming increasingly intense in the bud tips as branching progressed, up through 20 weeks gestation (Figure 2-2F). Bud tips expressed nearly undetectable levels of Pro-SFTPC at 10 and 12 weeks, with low levels of expression detected by 14 weeks (Figure 2-2D). Pro-SFTPC expression became more robust by 15 and 16 weeks and continued to increase in the bud tips until 20 weeks (Figure 2-1C; Figure 2-2D). SOX9+ bud tip cells were negative for several other lung epithelial markers including SFTPB, PDPN, RAGE and HOPX (Figure 2-1C, Figure 2-2C-E). We also examined expression of several proximal airway markers, including P63, acetylated-Tubulin (AC-TUB), FOXJ1, SCGB1A1 and MUC5AC

and noted that expression was absent from the epithelial bud tip progenitors (negative staining data not shown). We did observe that PDPN and HOPX were expressed in the transition zone/small airways directly adjacent to the SOX9+ bud tip domain at all time points examined (10-20 weeks of gestation) but that this region did not begin to express the AECI maker RAGE until 16 weeks of gestation (Figure 2-1C; Figure 2-2 B,C,E). RNA-sequencing (RNAseq) of isolated, uncultured human fetal epithelial lung bud tips (n=2; 59 days, 89 days gestation) supported protein staining analysis of human fetal buds. Differential expression analysis to identify genes enriched in the human fetal bud tips (isolated human fetal epithelial bud tips vs. whole adult lung) identified 7,166 genes that were differentially expressed (adjusted P-value < 0.01; Figure 2-1D). A curated heatmap highlights genes corresponding to Figure 2-1A-C and previously established markers of lung epithelial cells (Figure 2-1E). Human fetal bud tips have recently been shown to have enrichment for 37 transcription factors (13). In total, 20 of these 37 transcription factors were also enriched in our analysis (Figure 2-1F), and gene set enrichment analysis (GSEA) confirmed that this enrichment was highly significant (NES = -1.8, adjusted P-value=9.1e-5). Combined, this data provided a profile of the protein and gene expression in human fetal lung buds prior to 15 weeks gestation (summarized in Figure 2-1G).

Murine bud tip growth in vitro

In order to establish an experimental framework that would allow us to efficiently work with rare/precious human tissue samples, we first conducted a low-throughput screen using mouse epithelial bud tips to identify factors that promoted tissue expansion

and maintenance of SOX9 expression. Epithelial bud tips were isolated from lungs of embryonic day (E) 13.5 *Sox9*-eGFP mice and cultured in a Matrigel droplet (Figure 2-3A). Isolated *Sox9*-eGFP lung bud tips were shown to express GFP and to have the same level of *Sox9* mRNA as their wild type (WT) counterparts by QRT-PCR analysis (Figure 2-3B,C). Treating isolated E13.5 mouse bud tips with no growth factors (basal medium control) or individual factors identified from the literature as important for lung development showed that FGF7 robustly promoted growth, expansion and survival of isolated buds for up to two weeks (Figure 2-3D). Interestingly, the same concentration of FGF7 and FGF10 had different effects on lung bud outgrowth, an observation that could not be overcome even when buds were exposed to a 50-fold excess of FGF10 (Figure 2-3E).

FGF7, CHIR99021 and ATRA are sufficient to maintain the expression of SOX9 *in vitro*

A lineage trace utilizing isolated epithelial bud tips from *Sox9-Cre*^{ER};*Rosa26*^{Tomato} mice showed that FGF7 promoted outgrowth of *Sox9*+ distal progenitors cells (Figure 2-3F). However, *Sox9* mRNA and protein expression were significantly reduced over time (Figure 2-3G, M).

We therefore sought to identify additional growth factors that could maintain expression of SOX9 *in vitro.* To do this, we grew bud tips in media with all 5 factors included in our initial screen (‘5F’ media: FGF7, FGF10, CHIR-99021, RA), removed one growth factor at a time, and examined the effect on expression of *Sox9* and *Sox2* (Figure S2H-J). Bud tips were grown in 5F minus one factor for two weeks in culture. Removing any individual factor, with the exception of FGF7, did not affect the ability of

isolated buds to expand (Figure 2-3H). QRT-PCR analysis showed that removing BMP4 led to a statistically significant increase in *Sox9* mRNA expression levels when compared to other culture conditions (Figure 2-3I). Removing any other single factor did not lead to statically significant changes in *Sox9* expression (Figure 2-3I). *Sox2* gene expression was generally low in all culture conditions (Figure 2-3J).

Our data demonstrated that FGF7 is critical for *in vitro* expansion of isolated murine bud tips and removing BMP4 enhanced *Sox9* expression. We therefore screened combinations of the remaining factors to determine a minimal set of factors that could maintain high *Sox9* expression (Figure S2K-N). Cultured buds treated with 4-factor '4F'('4F'; FGF7, FGF10, CHIR-99021, RA) or 3-Factor conditions ('3F'; FGF7, CHIR-99021, RA) supported the most robust SOX9 protein and mRNA expression (Figure 2-3L-N)., with no significant differences between *Sox9* expression between these two groups. Therefore, a minimum set of 3 factors (FGF7, CHIR-99021, RA) are sufficient to allow growth of mouse epithelial bud tip progenitor cells and to maintain *Sox9* expression *in vitro.*

In vitro growth and maintenance of human fetal distal epithelial lung progenitors

We asked if conditions supporting mouse bud tip progenitors also supported growth and expansion of human bud tip progenitors *in vitro*. Distal epithelial lung buds were enzymatically and mechanically isolated from the lungs of 3 different biological specimens at 12 weeks of gestation (84-87 days; n=3) and cultured in a Matrigel droplet (Figure 2-4A-B). When human bud tips were cultured *in vitro*, we observed that FGF7 promoted an initial expansion of tissue *in vitro* after 2 and 4 weeks, but structures began

to collapse by 6 weeks in culture (Figure 2-5A). All other groups tested permitted expansion and survival of buds as “fetal progenitor organoids” for 6 weeks or longer (Figure 2-4C; Figure 2-5A). A description of the nomenclature for different tissues/samples/organoids used in this manuscript can be found in Table 1 of the Methods. Human fetal progenitor organoids exposed to 3F or 4F media supported robust expression of the distal progenitor markers *SOX9, SOX2, ID2 and NMYC* (Figure 2-5B-C). In contrast, culture in only 2 factors (FGF7+CHIR-99021 or FGF7+RA) did not support robust bud tip progenitor marker expression (Figure 2-5B-C). QRT-PCR also showed that fetal progenitor organoids cultured in 3F or 4F media expressed low levels of the proximal airway markers *P63, FOXJ1* and *SCGB1A1* when compared to FGF7-only conditions (Figure 2-5D-E).

Collectively, our experiments suggested that FGF7, CHIR99021 and RA represent a minimal set of additives required to maintain SOX9+/SOX2+ fetal progenitor organoids *in vitro.* Supporting mRNA expression data, we observed robust SOX9 and SOX2 protein expression as demonstrated by immunofluorescence in sections or by whole mount staining after 4 weeks in culture (Figure 2-4D-E). Consistent with expression data in lung sections prior to 16 weeks gestation, we observed very weak pro-SFTPC protein expression in human bud tip progenitor organoids. Low levels of *ID2* mRNA were also detected using *in situ* hybridization (Figure 2-4F-G). Protein staining for several markers was not detected in fetal progenitor organoids treated for 4 weeks in 3F medium, including P63, FOXJ1, SCGB1A1, MUC5AC, HOPX, RAGE, and SFTPB (n=8; negative data not shown), consistent with human fetal epithelial bud tips prior to 16 weeks gestation (Figure 2-2A-E).

We also performed RNA-sequencing on tissue from the distal portion of fetal lungs (epithelium plus mesenchyme) (n=3; 8, 12 and 18 weeks), on freshly isolated human epithelial bud tips (n=3; 8, 12 and 12 weeks) and on human fetal bud tip progenitor organoids grown in 3F media for 4 weeks in culture (n=2; 12 week biological samples, run with technical triplicates). Analysis revealed a high degree of similarity across samples when comparing epithelial gene expression (Figure 2-4H). Additionally, we identified genes highly enriched in isolated fetal epithelial bud tips by conducting differential expression analysis on RNAseq data. For this analysis, we compared whole human adult lung versus uncultured isolated lung buds (12 weeks gestation) and versus cultured fetal progenitor organoids (12 weeks gestation, cultured for 2 weeks). In these analyses, whole human lung was used as a 'baseline', with the assumption that average expression of any given gene would be low due to the heterogeneous mixture of cells pooled together in the sample, providing a good comparator to identify genes that were enriched in fetal tissue and organoid samples. The top 1000 upregulated genes in bud tips and in fetal progenitor organoids were identified (log2FoldChange < 0; adjusted p-value < 0.05). When comparing upregulated genes, we observed that 431 (27.5%) of the genes were commonly upregulated in both fetal bud tips and cultured fetal organoids, which was highly statistically significant (p-value = 1.4e-278 determined by a hypergeometric test)(Figure 2-4I).

FGF7, CHIR-99021, RA induce a bud tip progenitor-like population of cells from hPSC-derived foregut spheroids

Given the robustness by which 3F medium (FGF7, CHIR-99021, RA) supported mouse and human epithelial lung bud tip growth and maintenance *in vitro*, we sought to determine whether these culture conditions could promote an epithelial lung tip progenitor-like population from hPSCs. NKX2.1+ ventral foregut spheroids were generated as previously described (7,15), and were cultured in a droplet of Matrigel and overlaid with 3F medium. Spheroids were considered to be “day 0” on the day they were placed in Matrigel (Figure 2-6A-B). Foregut spheroids cultured in 3F medium generated patterned organoids (hereafter referred to as 'hPSC-derived patterned lung organoids'; PLOs) that grew robustly (Figure 2-7A-B), demonstrating predictable growth patterns (Figure 2-7C-D) and gene expression reproducible across several different embryonic and induced pluripotent stem cell lines (n=4) (Figure 2-6B; Figure 2-7G-H). PLOs survived for over 16 weeks in culture, and could be frozen, thawed and re-cultured (Figure 2-7E, E', F).

At 2 and 6 weeks, PLOs co-expressed NKX2.1 and SOX2 in >99% of all cells (99.41 +/- 0.82% and 99.06 +/- 0.83%, respectively), while 16 week old PLOs possessed 93.7 +/- 4.6% NKX2.1+/SOX2+ cells (Figure 2-6C-D). However, PLOs at this later time point appeared less healthy (Figure 2-6B). Interestingly, no mesenchymal cell types were detected in patterned lung organoid cultures at 2, 6 or 16 weeks by immunofluorescence (VIM, α-SMA, PDGFRα; negative immunostaining data not shown), and QRT-PCR analysis confirms very low expression of mesenchymal markers in PLOs generated from n=3 hPSC lines (Figure 2-7H).

PLOs also exhibited regionalized proximal-like and bud tip-like domains (Figure 2-6E-F). In 100% of analyzed PLOs (n=8) Peripheral budded regions contained cells

that co-stained for SOX9 and SOX2, whereas interior regions of the PLOs contained cells that were positive only for SOX2 (Figure 2-6E-F). PLOs expressed similar levels of *SOX2* and *SOX9* when compared to fetal progenitor lung organoids (Figure 2-6G). Budded regions of PLOs also contained SOX9+ cells that weakly co-expressed pro-SFTPC, and *ID2* based on *in situ* hybridization and QRT-PCR (Figure 2-6H-J).

Interior regions of PLOs contained a small number of cells that showed positive immunostaining for the club cell marker SCGB1A1 (9%) and the goblet cell marker MUC5AC (1%), with similar morphology to adult human proximal airway secretory cells (Figure 2-6K-N). No multiciliated cells were present, as Acetylated-Tubulin was not localized to cilia (Figure 2-6K) and FOXJ1 staining was absent (data not shown). Additionally, the basal cell marker P63 was absent from PLOs (Figure 2-6K), as was staining for markers of lung epithelial cell types including HOPX, RAGE, PDPN, ABCA3, SFTPB, and CHGA (negative data not shown).

Expansion of epithelial tip progenitor-like cells from patterned lung organoids

PLOs gave rise to epithelial cysts after being passaged by mechanical shearing through a 27-gauge needle, followed by embedding in fresh Matrigel and growth in 3F medium (Figure 2-8A-B). PLOs were successfully needle passaged as early as 2 weeks and as late as 10 weeks with similar results. Needle passaged cysts can be generated from hPSCs in as little as 24 days (9 days to generate foregut spheroids, 14 days to expand patterned lung organoids, 1 day to establish cysts from needle passaged organoids)(Figure 2-8A). Needle passaged epithelium re-established small cysts within 24 hours and could be serially passaged every 7-10 days (Figure 2-8C). The cysts that

formed after needle passaging were NKX2.1+ (Figure 2-9A) and cells co-expressed SOX2 and SOX9 (Figure 2-8D-E). Based on these protein staining patterns, we refer to these cysts as 'bud tip organoids'. When compared to PLOs, bud tip organoids possessed a much higher proportion of SOX9+ cells (42.5% +/- 6.5%, n=5; vs. 88.3% +/- 2.3%, n=9; Figure 2-8E) and proliferating cells assed by KI67 immunostaining (38.24% +/- 4.7%, n=9 for bud tip organoids vs. 4.9% +/- 0.6%, n=5 for PLOs; Figure 2-8F-I). In PLOs, we noted that proliferation was largely restricted to SOX9+ budded regions, but only a small proportion of SOX9+ cells were proliferating (8.1% +/- 0.9%, n=5), whereas bud tip organoids had a much higher proportion of proliferative SOX9+ cells (40.2% +/- 4.3%, n=9) (Figure 2-8I). Together, this data suggests that needle passaging enriches the highly proliferative bud tip regions of PLOs.

Bud tip organoids were further characterized using *in situ* hybridization and immunofluorescence (Figure 2-9B-D). Bud tip organoids exhibited protein staining patterns consistent with 12-14 week fetal lungs, including the absence of HOPX, RAGE, PDPN, and ABCA3, while we did observe low expression of *ID2* (Figure 2-9B-D). Furthermore, no positive protein staining was detected for the proximal airway markers P63, FOXJ1, AC-TUB, MUC5AC, SCGB1A1 or CHGA (negative data not shown).

We then conducted unbiased analysis of bud tip organoids using RNA-sequencing to compare: i) hPSC-derived bud tip organoids; ii) whole peripheral (distal) human fetal lung tissue; iii) freshly isolated fetal bud tips; human fetal lung progenitor organoids; iv) undifferentiated hPSCs; v) hPSC-derived lung spheroids. Principal component analysis (PCA) and Spearman's correlation clustering revealed the highest degree of similarity between hPSC-derived bud tip organoids, patterned lung organoids

and human fetal organoids (Figure 2-8J; Figure 2-9E). Interestingly, freshly isolated bud tips and whole distal human fetal lungs clustered together, while all cultured tissues (bud tip organoids, PLOs, fetal progenitor organoids) clustered together, likely reflecting gene expression similarities driven by the tissue culture environment. This analysis highlights a high degree of molecular similarity between human fetal and hPSC-derived organoids grown *in vitro* (Figure 2-8J). A curated heatmap shows a high degree of similarity in gene expression of lung epithelial markers between the whole fetal lung, fetal progenitor organoids, patterned lung organoids and bud tip organoids (Figure 2-8K).

Differential expression of RNAseq data was also used to interrogate the relationship between tissues. We obtained upregulated genes from the following comparisons: i) uncultured, freshly isolated lung buds (8 and 12 weeks gestation) versus whole adult lung tissue; ii) cultured fetal progenitor organoids (12 weeks gestation, cultured for 2 weeks) versus whole adult lung tissue; iii) hPSC-derived bud tip organoids versus whole adult lung tissue. The top 1000 upregulated genes in each of the three groups were identified (log2FoldChange < 0; adjusted p-value < 0.05), and overlapping genes were identified (Figure 2-8L-M). A hypergeometric means test found that overlap of enriched genes from isolated human bud tips and hPSC-derived bud tip organoids were highly significant, with 377 (23.2%) overlapping genes (p-value=9.3e-901; Figure 2-8L). Overlap between cultured fetal progenitor organoids and hPSC-derived bud tip organoids was also highly significant, with 477 (31.3%) overlapping genes (p-value=1.2e-1021) (Figure 2-8M). Of note, when all three groups were compared, a core group of 285 common upregulated genes were identified representing

14.3% of genes. Of these 285, several have been associated with human or mouse fetal bud tips (11,13) (Figure 2-8N; *COL2A1, ETV4, E2F8, FGF20, HMGA2, MYBL2, RFX6, SALL4, SOX9, SOX11*).

hPSC-derived bud tip organoids maintain multilineage potential in vitro and in vivo.

In order to demonstrate that hPSC-derived organoids have *bona fide* progenitor potential, we took approaches to differentiate them *in vitro* (Figure 2-10) and by transplanting them into injured mouse lungs *in vivo* (Figure 2-11 and Figure 2-12). For these experiments, we used induced pluripotent stem cells (iPSC) derived foregut spheroids to generate bud tip organoids (iPSC line 20-1 (iPSC20-1); (16)).

In vitro differentiation

Since we were interested in investigating the potential of bud tip progenitor organoids to give rise to any lung epithelial lineage, we reasoned that withdrawing CHIR-99021 and RA may allow the cells to stochastically differentiate. Therefore, bud tip organoids were split into two treatment groups: 3F medium or FGF7-only for 23 days (Figure 2-10A). At the end of the experiment, control bud tip organoids maintained a clear, thin epithelium with a visible lumen, whereas FGF7-differentiated organoids appeared as denser cysts with mucous-like material inside of some organoids (Figure 2-10B). FGF7-differentiated organoids remained NKX2.1+/ECAD+ and SOX2+, signifying that they retained a lung epithelial identity, but showed decreased SOX9 protein expression (Figure 2-10C,D). Gene expression analysis using qRT-PCR to compare 3F (undifferentiated) vs. FGF7-differentiated organoids showed significant increases of

genes expressed in differentiated airway and alveolar epithelium (Figure 2-10E). When further compared to the whole distal portion of human fetal lung (positive control) and cultured human fetal duodenum (negative control), differentiated bud tip organoids showed a large increase relative to 3F in many genes associated with differentiated epithelial cells, with many genes trending towards levels seen in the whole fetal lung (Figure 2-10F).

Protein staining for airway and alveolar markers was carried out on human lung tissue samples as a reference and in FGF7-differentiated organoids (Figure 2-10G,I). We observed that cells within FGF7-differentiated organoids expressed proximal airway markers, including the goblet cell markers MUC5AC and MUC5B, the club cell markers SCGB1A1, and the neuroendocrine cell markers SNY and CHGA (Figure 2-10F). Of note, we observed that SCGB1A1 and PLUNC were co-expressed in only a small subset of club cells in the adult lung, whereas many club cells were marked only with SCGB1A1 (Figure 2-10G). We did not observe P63, FOXJ1 or cilia with Acetylated-Tubulin staining in FGF7-differentiated organoids (negative data not shown). Transmission electron microscopy (TEM) also revealed cells with mucous containing vesicles characteristic of goblet cells within FGF7-differentiated organoids (Figure 2-10H). Immunostaining for alveolar markers revealed cells positive for AECI markers PDPN and HOPX, as well as cells double positive for AECII markers Pro-SFTPC and SFTPB and Pro-SFTPC and ABCA3 (Figure 2-10I). Of note, TEM also revealed putative lamellar bodies that were similar to those seen in 13 week human fetal lungs (Figure 2-10J). In both tissues, putative lamellar bodies were characteristically surrounded by what appears to be monoparticulate glycogen. Both iPSC-derived and human fetal

tissue lamellar bodies appear to be immature, and did not possess the typical concentric lamellae of mature structures (Figure 2-10J). Based on both IF and TEM data, it is most likely that the alveolar cell types observed in FGF7-differentiated organoids are reminiscent of an immature cell, and may require more specific conditions to further mature the primitive AECI and AECII cells.

Cells expressing cell type-specific lung epithelial markers were quantitated within FGF7-differentiated organoids, showing that organoids possessed cells expressing the alveolar markers: PDPN - 8.6% of total cells; HOPX - 5.9% of total cells; SFTPB ~45% of total cells; pro-SFTPC - ~ 45% of total cells; ABCA3 – 6.3% of total cells (Figure 2-10K-L). Many cells expressed airway markers: 43.2% of cells stained positive for nuclear SOX2, but not SOX9 (Figure 2-10K-L), 26.8% of cells stained positive for MUC5AC, while 7% of cells were positive for the club cell marker SCGB1A1, while FOXJ1 and P63 were absent (Figure 2-10K-L). Additionally, 2.0% of cells expressed very clear co-staining for Chromagranin A (CHGA) and Synaptophysin (SYN), markers for neuroendocrine cells (Figure 2-10G, K-L).

In vivo engraftment

We next sought to explore the differentiation potential of hPSC-derived bud tip progenitor organoids by engrafting them into the airways of injured mouse lungs (Figure 2-11 and Figure 2-12). Recent studies have shown that damaging the lung epithelium promotes engraftment of adult human lung epithelial cells (32).

Interestingly, while we observed engrafted cells in the mouse airway 7 days post injury (Figure 2-12A-E), 79% of engrafted human cells were still co-labeled by nuclear SOX2 and nuclear SOX9 and remained highly proliferative (Figure 2-12F-J). Immunofluorescent staining showed that of the small proportion of cells positive for only SOX2, 42.8% of engrafted human cells expressed the club cell marker SCGB1A1, and 25% expressed the goblet cell marker MUC5AC (Figure 2-12K-N), a large increase when compared to bud tip organoid controls (Figure 2-12O-P). Multiciliated cells were not observed (negative FOXJ1 staining in Figure 2-12K) nor were basal cells nor alveolar cell markers (basal cells: P63, KRT5; AECI: HOPX, PDPN, RAGE; AECII: Pro-SFTPC, SFTPB, ABCA3; negative staining not shown). Collectively, our data suggested that hPSC-derived lung bud tip progenitor-like cells can engraft into the injured airway; however, the majority of engrafted cells (79%, Figure 2-12F-G) retained expression of SOX9 after 7 days.

To determine whether engrafted cells could differentiate into multiple lung cell lineages, we carried out a long-term engraftment experiment where lungs were harvested 6 weeks after injection of cells (Figure 2-11). Mice received Naphthalene injury 24 hours prior to intratracheal injection of 600,000 dissociated bud tip organoid cells derived from iPSCs (iPSC20.1 tet-O GFP) (Figure 2-11A). Doxycycline was given on the final week of the experiment to induce expression from the tet-O-GFP transgene in the iPSC-derived cells. Experimental cohorts included: injury plus no cell injection (n=8 surviving animals); injury plus undifferentiated hPSC injection (n=4 surviving animals); injury plus bud tip organoid cell injection (n=15 surviving animals) (Figure 2-11B). Of the 15 surviving mice that received bud tip organoid injections, 8 of them

showed patches of NUMA+/GFP+ cells that engrafted into the airway (Figure 2-11C-D, 2-11F-G). The number of engrafted cell patches (Figure 2-11C) and the number of cells per patch (Figure 2-11D) varied across individual mice. Lungs from animals in all experimental groups successfully recovered from the injury (Figure 2-11E). Engrafted human cells were not observed in the “no cell injection” nor the “undifferentiated hPSC injection” cohorts (Figure 2-11B) as determined by NUMA+/GFP+ staining (Figure 2-11F-G). While 2 mice that showed successful engraftment of bud tip organoid cells within the airway exhibited a total of 4 small patches of engrafted cells in the bronchioles, the overwhelming majority of engrafted cells were found in the trachea and primary/secondary bronchi of mice. The rare engrafted cells within the bronchioles stained positive for proximal cell markers SOX2, MUC5AC or FOXJ1/AcTUB (data not shown).

We conducted immunostaining for several lung epithelial markers to determine if engrafted tissue had differentiated, and we compared staining patterns to adult human lung tissue (Figure 2-11H). Immunostaining of the injury group that did not receive cells was also examined (Figure 2-12Q-R). We noted that engrafted human cells expressed low levels of NKX2.1, consistent with the human airway where the majority of cells expressed very low levels of NKX2.1 (Figure 2-12S-T). Engrafted human cells (as determined by GFP or NuMA expression) also expressed SOX2, but not SOX9, suggesting these cells had differentiated into an airway fate (Figure 2-11H). Consistent with this observation, we observed engrafted human cells that possessed multiple cilia and co-expressed NuMA and multiciliated cell markers, AcTub and FOXJ1 (Figure 2-11H) and as determined in histological sections (Figure 2-11I). We also observed

human cells expressing goblet cell markers MUC5AC and MUC5B. Finally, we observed human cells expressing the club cell marker SCGB1A1, but not PLUNC. Interestingly, we noted that PLUNC only marked a small subset of club cells in the human airway, suggesting that there is an underappreciated heterogeneity within this population in the human lung (Figure 2-11H). Quantitation of immunostaining confirmed that 100% of engrafted human cells were SOX2+, with roughly 75% of cells adopting a mucous-producing phenotype, ~13% adopting a ciliated cell profile and ~0.5% exhibiting a neuroendocrine cell profile (Figure 2-11H). We did not find evidence that engrafted human cells expressed the basal cell marker P63 nor alveolar cell specific markers (negative data not shown).

2.3 Discussion

The ability to study human lung development, homeostasis and disease is limited by our ability to carry out functional experiments in human tissues. This has led to the development of many different *in vitro* model systems using primary human tissue, and using cells and tissues derived from hPSCs (5,6). Current approaches to differentiate hPSCs have used many techniques, including the stochastic differentiation of lung-specified cultures into many different lung cell lineages (8,17,18), FACS based sorting methods to purify lung progenitors from mixed cultures followed by subsequent differentiation (19-23), and by expanding 3-dimensional ventral foregut spheroids into lung organoids (7,15). Many or all, of these approaches rely on the differentiation of a primitive NKX2.1+ lung progenitor cell population, followed by subsequent differentiation into different cell lineages. However, during normal development, early NKX2.1-

specified lung progenitors transition through a SOX9+ bud tip progenitor cell state on their way to terminal differentiation into both alveolar and airway cell fates (2,24). Despite the developmental importance of this bud tip progenitor, differentiation of a similar progenitor from hPSCs has remained elusive. For example, several studies have identified robust methods to sort and purify lung epithelial progenitor cells from a mixed population (19,20,22); however, whether or not these populations represent or transition through an epithelial bud tip-like state is unknown. Thus, our findings demonstrating that hPSCs can be induced into a bud tip-like progenitor state represents an important step towards faithfully mimicking embryonic lung development *in vitro*. It is important to point out that while the current study suggests that bud tip-like progenitors can give rise to alveolar and airway cell types, the functional capacity of these differentiated cells have not been interrogated and will be an important avenue for future studies.

Inducing and maintaining the bud tip progenitor state in hPSC-derivatives *in vitro* has remained elusive, in part, because the complex signaling networks required controlling this population are not well understood. A recent study demonstrated the ability to culture human fetal bud tips *in vitro*, supporting the findings of the current study (13), however *bona fide* SOX9+/SOX2+ bud tip progenitor-like cells derived from hPSCs have not been described nor well characterized prior to this work. The ability to induce, *de novo*, robust populations of cells from hPSCs that do not rely on specialized sorting or purification protocols suggest that biologically robust experimental findings can be used in a manner that predicts how a naïve cell will behave in the tissue culture dish with a high degree of accuracy, and across multiple cell lines (16,25,26).

Our studies also identified significant species-specific differences between the human and fetal mouse lung. Differences included both gene/protein expression differences, as well as functional differences when comparing how cells responded to diverse signaling environments *in vitro.* These mouse-human differences highlight the importance of validating observations made in hPSC-derived tissues by making direct comparisons with human tissues, as predictions or conclusions about human cell behavior based on results generated from the developing mouse lung may be misleading.

Our experimental findings, in combination with previously published work, have also raised new questions that may point to interesting avenues for future mechanistic studies to determine how specific cell types of the lung are generated. Previously, we have shown that lung organoids grown in high concentrations of FGF10 predominantly give rise to airway-like tissues, with a small number of alveolar cell types and a lack of alveolar structure, and these organoids also possess abundant mesenchymal cell populations (7,15). Here, our results suggest that high concentrations of FGF10 alone do not play a major role in supporting robust growth of epithelial bud tip progenitor cells. We also note that lung organoids grown in high FGF10 possess P63+ basal-like cells (7), whereas bud tip organoids grown in 3F medium lack this population. These findings suggest that we still do not fully appreciate how various signaling pathways interact to control cell fate decisions or expansion of mesenchymal populations, and lay the groundwork for many future studies.

Taken together, our current work has identified a signaling network required for the induction, expansion and maintenance of hPSC-derived lung epithelial bud tip

progenitors. Simple needle passaging allowed us to expand a nearly homogenous population of proliferative bud tip-like progenitor cells for over 16 weeks in culture, which remained multipotent *in vitro* and which were able to engraft into injured mouse lungs, terminally differentiate, and respond to systemic factors. The current thus study offers a robust and reproducible method to generate and maintain epithelial bud tip progenitors, which will facilitate future studies aimed at elucidating fundamental developmental mechanisms regulating human lung progenitor cells, and which may have applicability to regenerative medicine in the future.

2.4 Experimental Procedures

Mouse models

All animal research was approved by the University of Michigan Committee on Use and Care of Animals. Lungs from Sox9-eGFP (MGI ID:3844824), Sox9CreER;Rosa$^{Tomato/Tomato}$ (MGI ID:5009223 and 3809523)(27), or wild type lungs from CD1 embryos (Charles River) were dissected at embryonic day (E) 13.5. 8-10 week old Immunocompromised NSG mice (Jackson laboratories strain #0005557) were used for engraftment studies. Pilot studies identified that females were more sensitive to Naphthalene and died at a higher rate, therefore male mice were used for engraftment experiments.

Human lung tissue

All research utilizing human tissue was approved by the University of Michigan institutional review board. Additional information can be found in the Supplemental Detailed Methods.

Cell lines and culture conditions

Mouse and human primary cultures:

Epithelial bud tips were isolated as essentially as previously described (28), and see Supplemental Detailed Methods. Isolated mouse bud tips were cultured in 4-6 µl droplets of matrigel, covered with media, and kept at 37 degrees Celsius with 5% CO2. Isolated human fetal lung bud tips were cultured in 25-50 µl droplets of matrigel, covered with media, and kept at 37 degrees Celsius with 5% CO2. Cultures media was changed every 2-4 days.

Generation and culture of hPSC-derived lung organoids:

The University of Michigan Human Pluripotent Stem Cell Research Oversight (hPSCRO) Committee approved all experiments using human embryonic and induced pluripotent stem cell (hESC, iPSC) lines. Patterned lung organoids were generated from 4 independent pluripotent cell lines in this study: hESC line UM63-1 (NIH registry #0277) was obtained from the University of Michigan and hESC lines H9 and H1 (NIH registry #0062 and #0043, respectively) were obtained from the WiCell Research Institute. iPSC20.1 was previously described (16). ES cell lines were routinely karyotyped to ensure normal karyotype and ensure the sex of each line (H9 - XX, UM63-1 – XX, H1 - XY). All cell lines are routinely monitored for mycoplasma infection

monthly using the MycoAlert Mycoplasma Detection Kit (Lonza). Stem cells were maintained as previously described (16) and ventral foregut spheroids were generated as previously described (7,15). Following differentiation, free-floating foregut spheroids were collected from differentiated stem cell cultures and plated in a matrigel droplet on a 24-well tissue culture grade plate. A summary of different tissue/sample types, nomenclature and descriptions can be found in Table 2-1.

Culture media, growth factors and small molecules

Information for Serum-free basal media, growth factors and small molecules and the in-house generation and Isolation of human recombinant FGF10, including company information and catalogue numbers can be found in the Detailed Supplemental Methods.

RNA-sequencing and Bioinformatics Analysis

RNA sequencing and alignment was carried out as previously published (29). Data analysis was performed using the R statistical programming language (http://www.R-project.org/) as previously described (7,30). The complete sequence alignment, expression analysis and all corresponding scripts can be found at https://github.com/hilldr/Miller_Lung_Organoids_2017.

Accession Numbers

All raw RNA-sequencing data files have been deposited to the EMBL-EBI ArrayExpress database (accession ID: E-MTAB-6023). Adult human RNA-sequencing samples

representing bulk sequencing of whole lung homogenates were obtained from the EMBL-EBI ArrayExpress repository (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-1733/) (Fagerberg et al., 2014).

Naphthaline injury

Naphthaline (Sigma #147141) was dissolved in corn oil at a concentration of 40 mg/ml. Adult male NSG mice were chosen for these experiments because we observed improved recovery and survival following injury compared to female mice. Mice that were 8-10 weeks of age were given i.p. injections at a dose of 300 mg/kg weight.

Intratracheal injection of fetal progenitor organoids and hPSC-derived bud tip organoid cells into immunocompromised mouse lungs

Generating single cells from organoid tissues:

2-3 matrigel droplets containing organoid tissues were removed from the culture plate and combined in a 1.5mL eppendorf tube with 1mL of room temperature Accutase (Sigma #A6964). The tube was laid on its side to prevent organoids from settling to the bottom. Tissue was pipetted up and down 15-20 times with a 1mL tip every 5 minutes for a total of 20 minutes. Single cell status was determined by microscopic observation using a hemocytometer. Cells were diluted to a concentration of 500,000-600,000 cells per 0.03 mL in sterile PBS.

Intratracheal injection of cells:

Injection of cells into the mouse trachea was performed. Briefly, animals were anesthetized and intubated. Animals were given 500,000-600,000 single cells in 30-35 µL of sterile PBS through the intubation cannula.

RNA extraction and quantitative RT-PCR analysis

These procedures were carried out as previously described (30), and see Supplemental Detailed Methods. A list of primer sequences used can be found in Table 2-2.

Tissue preparation, Immunohistochemistry, Electron Microscopy and imaging

These procedures were carried out as previously described (7,11,31), and see Supplemental Detailed Methods.

Quantification and Statistical Analysis

All plots and statistical analysis were done using Prism 6 Software (GraphPad Software, Inc.). For statistical analysis of qRT-PCR results, at least 3 biological replicates for each experimental group were analyzed and plotted with the standard error of the mean. If only two groups were being compared, a two-sided student's T-test was performed. In assessing the effect of length of culture with FGF7 on gene expression in mouse buds (Figure S2G), a one-way, unpaired Analysis of Variance (ANOVA) was performed for each individual gene over time. The mean of each time point was compared to the mean of the expression level for that gene at day 0 of culture. If more than two groups were being compared within a single experiment, an unpaired one-way analysis of variance was performed followed by Tukey's multiple comparison test to compare the

mean of each group to the mean of every other group within the experiment. For all statistical tests, a significance value of 0.05 was used. For every analysis, the strength of p values is reported in the figures according the following: $P > 0.05$ ns, $P \leq 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***, $P \leq 0.0001$ ****. Details of statistical tests can be found in the figure legends.

Acknowledgements:

JRS is supported by the NIH-NHLBI (R01 HL119215). AJM is supported by the NIH Cellular and Molecular Biology training grant at Michigan (T32 GM007315), and by the Tissue Engineering and Regeneration Training Grant (DE00007057-40). 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 & Human Development. We thank Michael Ferguson for his preliminary experiments isolating and culturing human fetal epithelial bud tips. We apologize to those whose work we were unable to cite due to space limitations.

2.5 Supplemental methods

Culture media, growth factors and small molecules

Serum-free basal media:

All mouse bud, human fetal bud, and hPSC-derived human lung organoids were grown in serum-free basal media (basal media) with added growth factors. Basal media consisted of Dulbecco's Modified Eagle Medium: Nutrient Mixture F12 (DMEM/F12, ThermoFisher SKU# 12634-010) supplemented with 1X N2 supplement (ThermoFisher Catalog# 17502048) and 1X B27 supplement (ThermoFisher Catalog# 17504044),

along with 2mM Glutamax (ThermoFisher Catalog# 35050061), 1x Pennicillin-Streptomycin (ThermoFisher Cat# 15140122) and 0.05% Bovine Serum Albumin (BSA; Sigma product# A9647). BSA was weighed and dissolved in DMEM F/12 media before being passed through a SteriFlip 0.22 uM filter (Millipore Item# EW-29969-24) and being added to basal media. Media was stored at 4 degrees Celsius for up to 1 month. On the day of use, basal media was aliquoted and 50ug/mL Ascorbic acid and 0.4 uM Monothioglycerol was added. Once ascorbic acid and monothioglycerol had been added, media was used within one week.

Growth factors and small molecules:

Recombinant Human Fibroblast Growth Factor 7 (FGF7) was obtained from R&D systems (R&D #251-KG/CF) and used at a concentration of 10 ng/mL unless otherwise noted. Recombinant Human Fibroblast Growth Factor 10 (FGF10) was obtained either from R&D systems (R&D #345-FG) or generated in house (see below), and used at a concentration of 10 ng/mL (low) or 500 ng/mL (high) unless otherwise noted. Recombinant Human Bone Morphogenic Protein 4 (BMP4) was purchased from R&D systems (R&D Catalog # 314-BP) and used at a concentration of 10 ng/mL. All-trans Retinoic Acid (RA) was obtained from Stemgent (Stemgent Catalog# 04-0021) and used at a concentration of 50 nM. CHIR-99021, a GSK3β inhibitor that stabilizes β-CATENIN, was obtained from STEM CELL technologies (STEM CELL Technologies Catalog# 72054) and used at a concentration of 3 uM. Y27632, a ROCK inhibitor (APExBIO Cat# A3008) was used at a concentration of 10uM.

Generation and Isolation of human recombinant FGF10:

Recombinant human FGF10 was produced in-house, as previously described (Bagai et al., 2002). The recombinant human FGF10 (rhFGF10) expression plasmid pET21d-FGF10 in *E. coli* strain BL21*trxB*(DE3) was a gift from James A. Bassuk at the University of Washington School of Medicine. *E. coli* was grown in standard LB media with peptone derived from meat, carbenicillin and glucose. 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, 17040601) with step gradients of 0.5M to 2M LiCl. From a 200 ml culture, 3 mg of at least 98% pure rFGF-10 (evaluated by SDS-PAGE stained with Coomassie Blue R-250) was purified. rFGF10 was compared to commercially purchased human FGF10 (R&D Systems) to test/validate activity based on the ability to phosphorylate ERK1/2 in an A549 alveolar epithelial cell line (ATCC Cat#CCL-185) as assessed by western blot analysis.

Human fetal tissue samples:

All tissues were shipped overnight in Belzer's solution at 4 degrees Celsius and were processed and cultured within 4 hours of obtaining the specimen. Experiments to evaluate the effect on progenitor maintenance in culture by media conditions were repeated using tissues from 3 individual lung specimens; (1) 84 day post fertilization of unknown gender, (2), 87 day post fertilization male, and (3), 87 day post fertilization of unknown gender. RNAseq experiments of whole lung homogenate utilized tissue from 3 additional individual lungs; (4) 59 day male (5) 87 day of unknown gender and (6) 125 day of unknown gender. Use of normal adult human lung tissue was approved by

University of Michigan institutional review board, and was obtained from warm autopsy organ donors arranged by the Gift of Life, Michigan.

Isolation and culture of mouse epithelial bud tips:

Mouse buds were dissected from E13.5 embryos. For experiments using Sox9CreER;Rosa$^{Tomato/Tomato}$ mice, 50 ug/g of tamoxifen was dissolved in corn oil and given by oral gavage on E12.5, 24 hours prior to dissection. Briefly, in a sterile environment, whole lungs were placed in 100% dispase (Corning Cat# 354235) on ice for 30 minutes. Lungs were then transferred using a 25uL wiretrol tool (Drummond Scientific Cat# 5-000-2050) to 100% FBS (Corning Cat#35-010-CV) on ice for 15 minutes, and then transferred to a solution of Dulbecco's Modified Eagle Medium: Nutrient Mixture F12 (DMEM/F12, ThermoFisher SKU# 12634-010) with 10% FBS and 1x Pennicillin-Streptomycin (ThermoFisher Cat# 15140122) on ice. To dissect buds, a single lung or lung lobe was transferred by wiretrol within a droplet of media to a 100mm sterile petri dish. Under a dissecting microscope, the mesenchyme was carefully removed and epithelial bud tips were torn away from the bronchial tree using tungsten needles (Point Technologies, Inc.). Care was taken to remove the trachea and any connective tissue from dissected lungs. Isolated bud tips were picked up using a p20 pipette and added to an eppendorf tube with cold Matrigel (Corning Ref# 354248) on ice. The buds were picked up in a p20 pipette with 4-6 uL of Matrigel and plated on a 24-well tissue culture well (ThermoFisher Cat# 142475). The plate was moved to a tissue culture incubator and incubated for 5 minutes at 37 degrees Celsius and 5% CO2

to allow the Matrigel droplet to solidify. 500uL of media was then added to the dish in a laminar flow hood. Media was changed every 2-3 days.

Isolation and culture of human fetal lung epithelial bud tips

Distal regions of 12 week fetal lungs were cut into ~2mm sections and incubated with dispase, 100% FBS and then 10% FBS as described above, and moved to a sterile petri dish. Mesenchyme was removed by repeated pipetting of distal lung pieces after dispase treatment. Buds were washed with DMEM until mesenchymal cells were no longer visible. Buds were then moved to a 1.5 mL eppendorf tube containing 200 µL of Matrigel, mixed using a p200 pipette, and plated in ~20 µL droplets in a 24 well tissue culture plate. Plates were placed at 37 degrees Celsius with 5% CO2 for 20 minutes while droplets solidified. 500uL of media was added to each well containing a droplet. Media was changed every 2-4 days.

RNA extraction and quantitative RT-PCR analysis:

RNA was extracted using the MagMAX-96 Total RNA Isolation System (Life Technologies). RNA quality and concentration was determined on a Nanodrop 2000 spectrophotometer (Thermo Scientific). 100 ng of RNA was used to generate a cDNA library using the SuperScript VILO cDNA master mix kit (Invitrogen) according to manufacturer's instructions. qRT-PCR analysis was conducted using Quantitect SYBR Green Master Mix (Qiagen) on a Step One Plus Real-Time PCR system (Life Technologies). Expression was calculated as a change relative to GAPDH expression using arbitrary units, which were calculated by the following equation: [2^(GAPDH Ct -

Gene Ct)] x 10,000. A Ct value of 40 or greater was considered not detectable. A list of primer sequences used can be found in Table 2-2.

RNA-sequencing and Bioinformatic Analysis:

was isolated using the mirVana RNA isolation kit, following the "Total RNA" isolation protocol (Thermo-Fisher Scientific, Waltham MA). RNA sequencing library preparation and sequencing was carried out by the University of Michigan DNA Sequencing Core and Genomic Analysis Services (https://seqcore.brcf.med.umich.edu/). 50bp single end cDNA libraries were prepared using the Truseq RNA Sample Prep Kit v2 (Illumina), and samples were sequenced on an Illumina HiSeq 2500. Transcriptional quantitation analysis was conducted using 64-bit Debian Linux stable version 7.10 ("Wheezy"). Pseudoalignment of RNA-sequencing reads was computed using kallisto v0.43.0 and a normalized data matrix of pseudoaligned sequences (Transcripts Per Million, TPM) and differential expression was calculated using the R package DEseq2 (Bray et al., 2016; Love et al., 2014).

Tissue preparation, Immunohistochemistry, Electron Microscopy and imaging:

Paraffin sectioning and staining:

Mouse bud, human bud, and HLO tissue was fixed in 4% Paraformaldehyde (Sigma) for 2 hours and rinsed in PBS overnight. Tissue was dehydrated in an alcohol series, with 30 minutes each in 25%, 50%, 75% Methanol:PBS/0.05% Tween-20, followed by 100% Methanol, and then 100% Ethanol. Tissue was processed into paraffin using an

automated tissue processor (Leica ASP300). Paraffin blocks were sectioned 5-7 uM thick, and immunohistochemical staining was performed as previously described (Spence et al., 2009). A list of antibodies, antibody information and concentrations used can be found in Table 2-3. PAS Alcian blue staining was performed using the Newcomer supply Alcian Blue/PAS Stain kit (Newcomer Supply, Inc.) according to manufacturer's instructions.

Whole mount staining:

For whole mount staining tissue was placed in a 1.5mL eppendorf tube and fixed in 4% paraformaldehyde (Sigma) for 30 minutes. Tissue was then washed with PBS/0.05% Tween-20 (Sigma) for 5 hours, followed by a 2.5-hour incubation with blocking serum (PBS-Tween-20 plus 5% normal donkey serum). Primary antibodies were added to blocking serum and tissue was incubated for at least 24 hours at 4 degrees Celcius. Tissue was then washed for 5 hours with several changes of fresh PBS-Tween-20. Secondary antibodies were added to fresh blocking solution and tissue was incubated for 12-24 hours, followed by 5 hours of PBS-Tween-20 washes. Tissue was then dehydrated to 100% methanol and carefully moved to the center of a single-well EISCO concave microscope slide (ThermoFisher Cat# S99368) using a glass transfer pipette. 5-7 drops of Murray's clear (2 parts Benzyl alcohol, 1 part Benzyl benzoate [Sigma]) were added to the center of the slide, and slides were coverslipped and sealed with clear nail polish.

In situ hybridization:

In situ hybridization for ID2 was performed using the RNAscope 2.5 HD manual assay with brown chromogenic detection (Advanced Cell Diagnostics, Inc.) according to manufacturers instructions. The human 20 base pair ID2 probe was generated by Advanced Cell Diagnostics targeting 121-1301 of ID2 (gene accession NM_002166.4) and is commercially available.

Imaging and image processing:

Images of fluorescently stained slides were taken on a Nikon A-1 confocal microscope. When comparing groups within a single experiment, exposure times and laser power were kept consistent across all images. All Z-stack imaging was done on a Nikon A-1 confocal microscope and Z-stacks were 3-D rendered using Imaris software. Brightness and contrast adjustments were carried out using Adobe Photoshop Creative Suite 6 and adjustments were made uniformly across images. Low magnification brightfield images of live cultures were taken using an Olympus S2X16 dissecting microscope. Image brightness and contrast was enhanced equally for all images within a single experiment using Adobe Photoshop. Images were cropped where noted in figure legends to remove blank space surrounding buds or cultures.

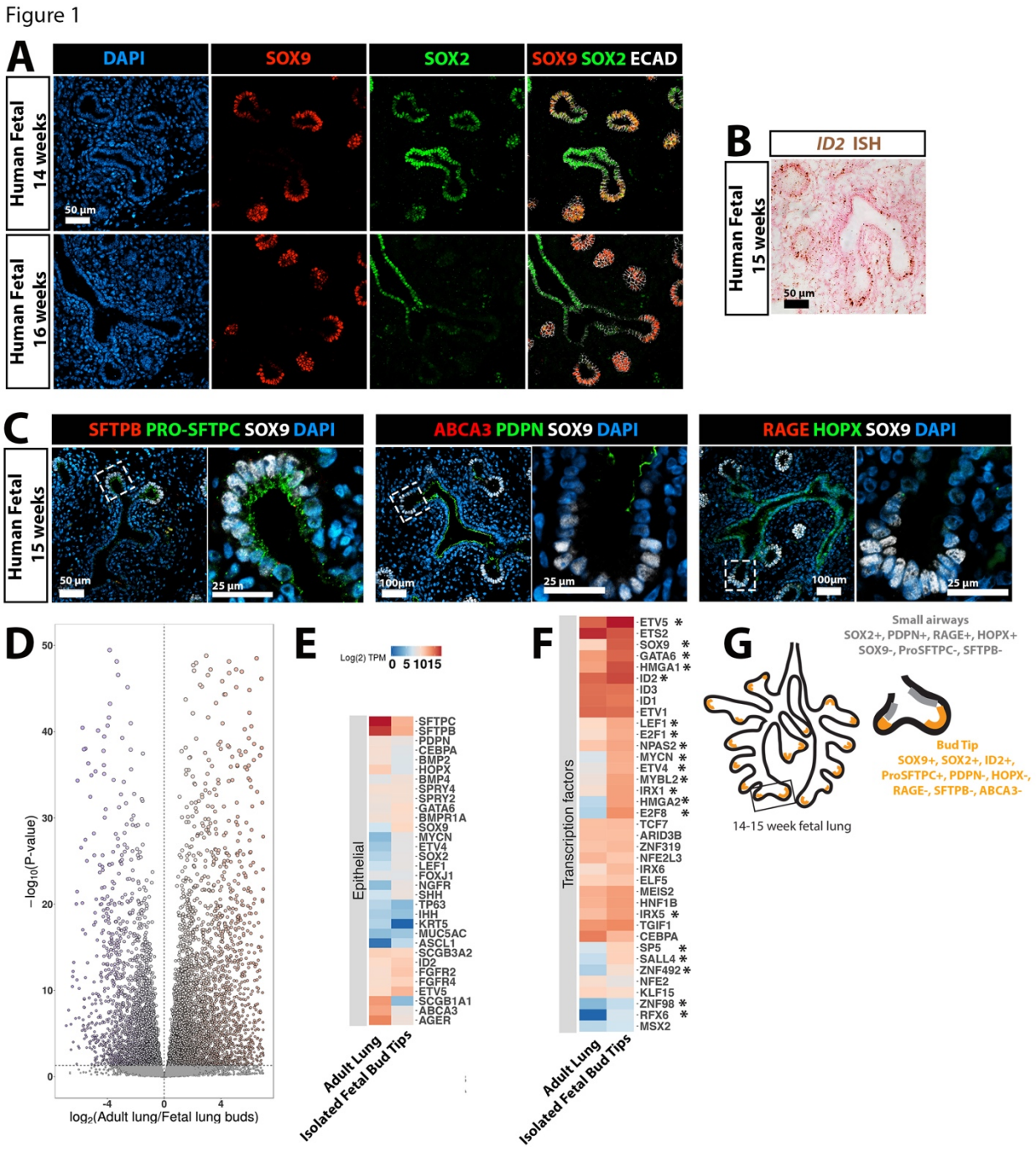
Transmission Electron Microscopy:

Undifferentiated and differentiated iPSC-derived bud tip organoids and 13 week human fetal lung tissue was fixed overnight in 2.5% glutaraldehyde in Sorensen's buffer (0.1M). The following day tissue was washed with Cacodylate buffer (0.05M) then post fixed for one hour in 1% osmium tetroxide. After fixation, the tissue was rinsed again in

Cacodylate buffer (0.05M). The tissue was then dehydrated in an ethanol gradient, with tissue incubating for 15 minutes in solutions of 25%, 50%, 75% and 100% ethanol. Tissue was then cleared in propylene oxide. Epon resin was infiltrated into the tissue by mixing propylene oxide and Epon resin (3:1, 1:1, 1:3), tissues were incubated in each mixture for 24 hours on a rocker. Tissues were then submerged into full strength Epon resin and rocked for two more days at room temperature, changing to fresh resin each day. Using fresh Epon resin, the tissues were arranged in molds and allowed to polymerize for 24 hours in a 60ºC oven. Sections were cut using a Leica EM UC7 ultramicrotome and imaged using a JEOL 1400-Plus Transmission Electron Microscope.

2.6 Figures

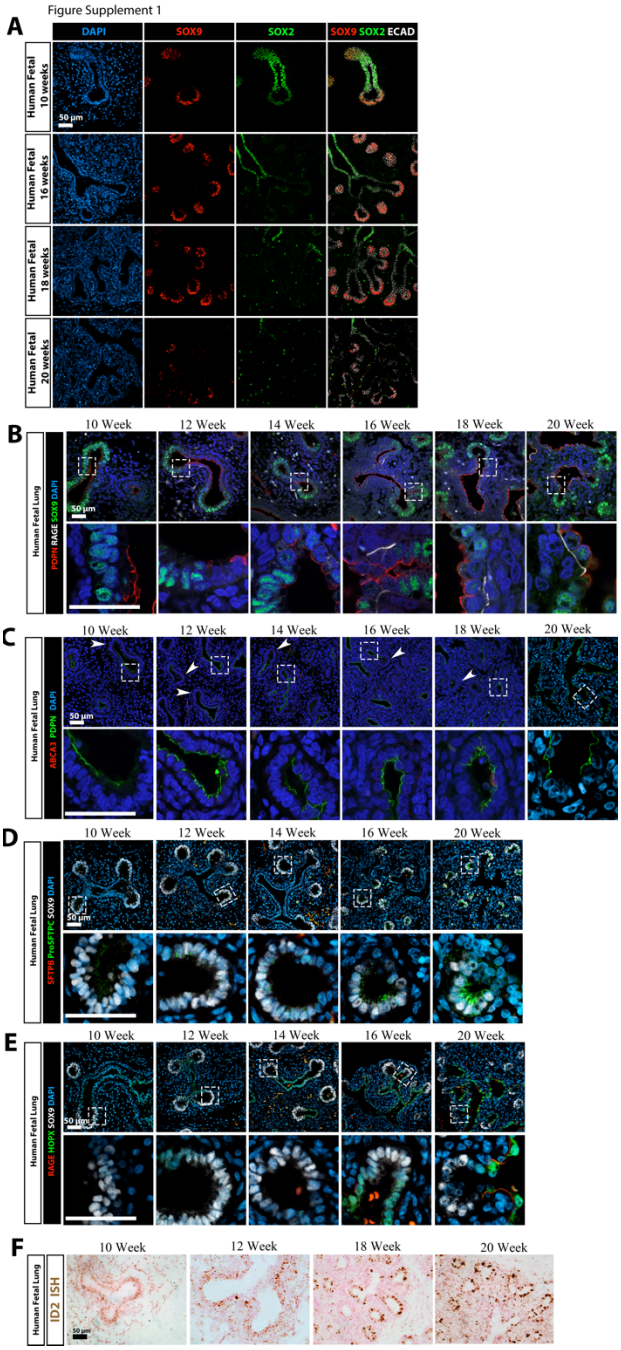
Figure 2-1 Characterization of bud tip progenitors from 14-16 weeks of human fetal lung development.



(A) Expression of SOX2 and SOX9 in human fetal lungs at 14 and 16 weeks of gestation. Scale bar represents 50 μm. **(B)** Expression of ID2 in human fetal lungs at 15

weeks gestation as identified by *in situ* hybridization. Scale bar represents 50 μm. **(C)** Expression of SOX9 along with Pro-SFTPC, SFTPB, PDPN,HOPX, ABCA3 or RAGE in human fetal lung at 15 weeks. Scale bars represent 50μm (low magnification images) and 25 μm (high magnification images). **(D)** Volcano plot of differentially expressed genes identified by comparing isolated, uncultured human fetal bud tips (n= 2 biological replicates; 8 weeks and 12 weeks gestation) with whole adult human lung (publically available dataset). A total of 7,166 genes were differentially expressed (adjusted P-value < 0.01). **(E)** Heatmap showing expression of genes known to be expressed lung epithelial cells in isolated human fetal bud tips and whole adult human lung. **(F)** Heatmap showing expression of 37 human bud-tip enriched transcription factors (13) in isolated human fetal bud tips and whole adult human lung. 20 of the 37 transcription factors (marked with asterisk "*") were statistically significantly enriched in isolated fetal bud tips. **(G)** Summary of markers expressed by bud tip cells in regions adjacent to the bud tips at 14-15 weeks gestation as identified by protein staining and *in situ* hybridization.

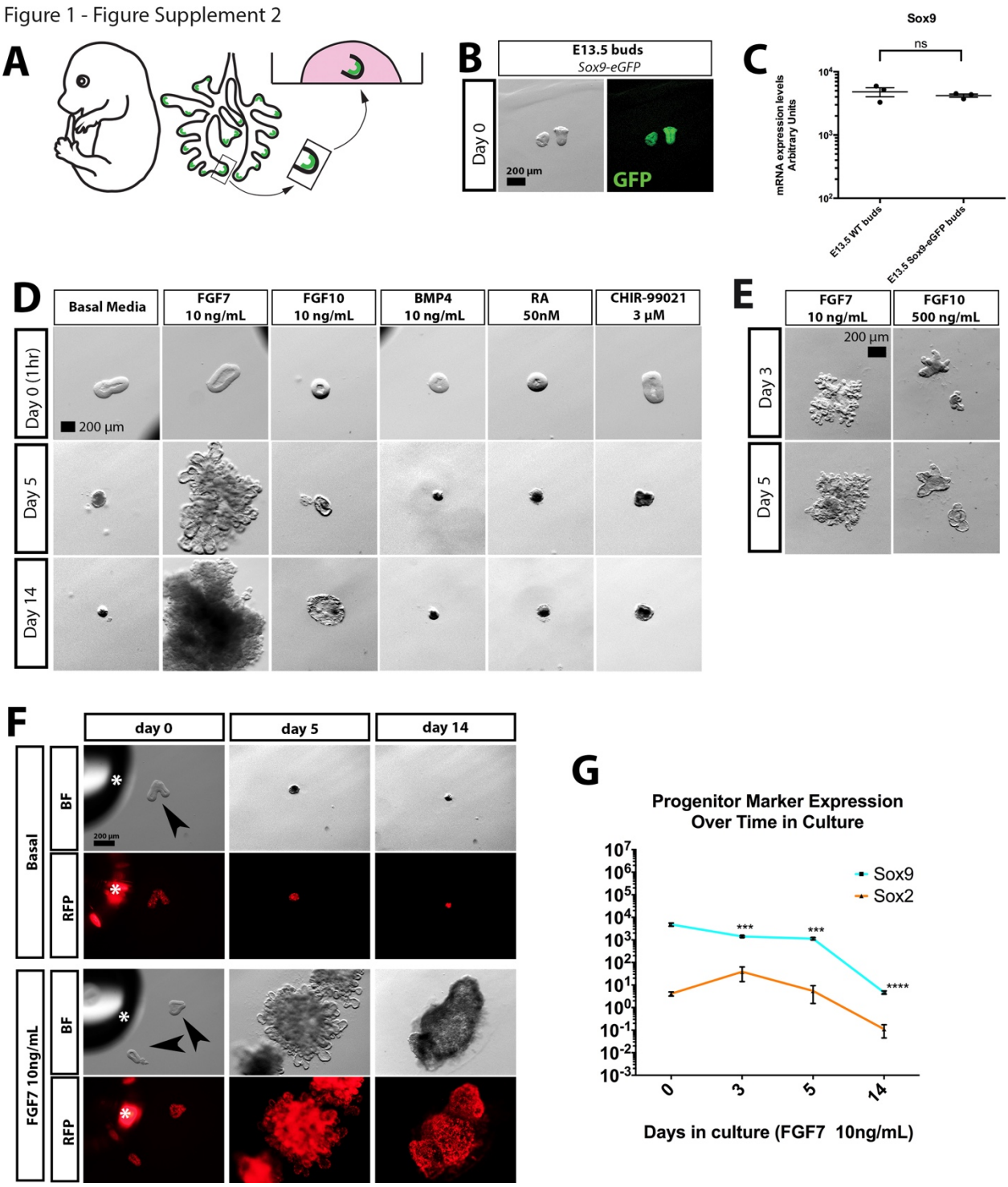
Figure 2-2 Characterization of human fetal bud tip cells throughout multiple time points during branching morphogenesis



(A) Immunofluorescence staining for SOX9 and SOX2 in human fetal lung specimens from 10-20 weeks gestation. Scale bar represents 50 μm. **(B)** Immunofluorescence

staining for SOX9 and AECI markers PDPN and RAGE in human fetal lung specimens from 10-20 weeks gestation. Scale bars represent 50 μm. **(C)** Immunofluorescence staining for AECII marker ABCA3 and AECI marker PDPN in human fetal lung specimens from 10-20 weeks gestation. Arrowheads point to bud tips, which are negative for PDPN. No positive staining for ABCA3 was detected. Scale bars represent 50 μm. **(D)** Immunofluorescence staining for AECII markers Pro-SFTPC and SFTPB with progenitor marker SOX9 in human fetal lung specimens from 10-20 weeks gestation. Weak Pro-SFTPC is not detected in the SOX9+ bud tip cells until 12 weeks, and staining within the bud tip cells increases over time. SFTPB was not detected at any time point analyzed. Scale bar represents 50 μm. **(E)** Immunofluorescence staining for AECI markers RAGE and HOPX in human fetal lung specimens from 10-20 weeks gestation. As seen in **(B)**, RAGE does not appear in the transition zone until 16 weeks gestation. Like PDPN, HOPX is excluded from the SOX9+ bud tip cells at all time points evaluated, but is expressed in the nuclei of airway cells as early as 10 weeks. Scale bar represents 50 μm. **(F)** *In situ* hybridization (brown dots indicate positive staining) for *ID2* shows that bud tip cells in the developing human fetal lung express ID2 from 10-20 weeks gestation. Intensity of signal increases with time, and the strongest positive signal in the bud tips occurs at 20 weeks. Scale bar represents 50 μm.

Figure 2-3 Identification of growth factors that promote growth of mouse SOX9+ bud tip progenitors *ex vivo.*

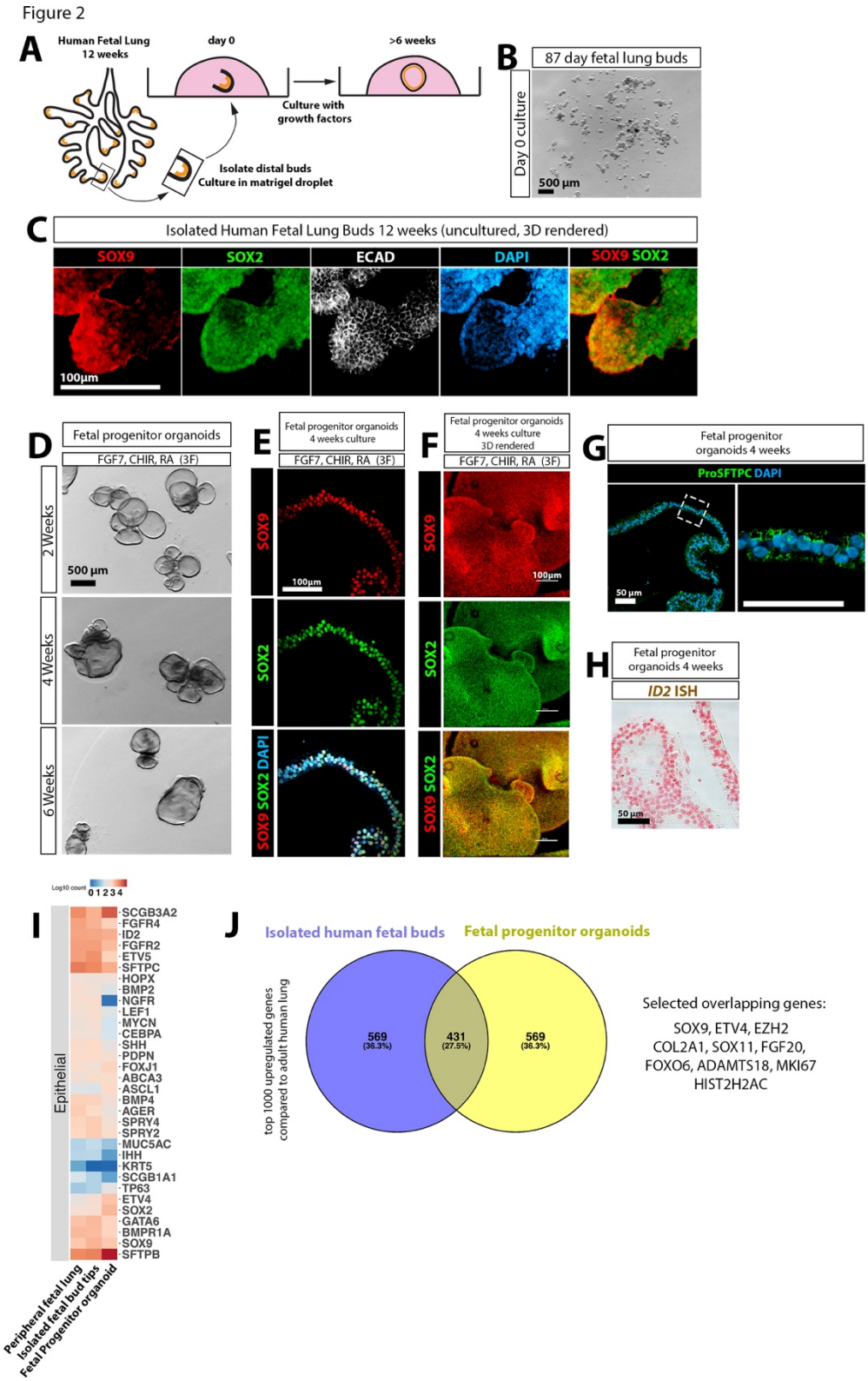


(A) Schematic of *Sox9*-eGFP distal lung buds dissected at E13.5 and cultured in a Matrigel droplet. N=3 **(B)** Low magnification images of isolated lung buds under brightfield (left) or showing *Sox9*-eGFP expression (right). Scale bar represents 200 μm. **(C)** QRT-PCR of wild type versus Sox9-eGFP isolated epithelial bud tips. Each data point represents a pool of 6-7 isolated bud tips from an independent biological replicate (littermates). Error bars plot the standard error of the mean. The mean of each group was compared using a two-sided student's T-test with a significance level of 0.05. **(D)** Isolated bud tips were cultured in basal media or individually with different factors (10ng/mL FGF7, 10ng/mL FGF10, 10ng/mL BMP4, 50nM RA, 3uM CHIR-99021) and imaged at Day 0, Day 5 and Day 14 in culture. Scale bar represents 200 μm.**(E)** Brightfield images of mouse lung bud tips grown with FGF7 or with a 50-fold increase in the concentration of FGF10 (500 ng/ml). Scale bar represents 200 μm. **(F)** *Sox9*-Cre^{ER};$Rosa26^{Tomato}$ lungs were induced with Tamoxifen 24 hours prior to isolation of the bud tips at E12.5, and bud tips were isolated and cultured at E13.5. Lineage labeled buds demonstrated that labeled cells expanded in culture over the course of two weeks only in the presence of FGF7. Asterisks (*) mark air bubbles within Matrigel droplets that were auto-fluorescent, and arrowheads point to day 0 isolated lung buds. Scale bar represents 200μm. **(G)** QRT-PCR analysis of bud tip expanded for 0, 3, 5, or 14 days in culture with 10ng/mL FGF7. Each data point represents the mean (+/- SEM) of 3 independent biological replicates (n=3; littermate pups). Statistical significance was determined by a one-way, unpaired Analysis of Variance (ANOVA) for each individual gene over time. The mean of each time point was compared to the mean of the expression level for that gene at day 0 of culture. **(H)** E13.5 isolated mouse lung bud

tips cultured with a combination of 5 growth factors in serum free media (10ng/mL FGF7, 10ng/mL FGF10, 10ng/mL BMP4, 50nM RA, 3uM CHIR-99021) with single growth factors removed. Scale bar represents 200 μm. **(I-J)** *Sox9 and Sox2 expression* measured by QRT-PCR is shown for all groups. One-way Analysis of Variance was performed followed by Tukey's multiple comparison to compare the mean of each group to the mean of every other group within the experiment. Each data point represents an independent biological replicate (littermate pups) and graphs indicate the mean +/- the standard error of the mean for each experimental group. **(K)** E13.5 mouse bud tips grown with different combinations of growth factors. Scale bar represents 500 μm. **(L-M)** Section (L) and whole mount (M) immunohistochemical staining for SOX9 on various growth conditions. Asterisks mark areas within cultures that contained debris with non-specific staining for DAPI. Scale bar represents 50 μm in D. **(N)** *Sox9* and *Sox2* mRNA expression levels from E13.5 isolated bud tips and from various growth conditions as assessed by QRT-PCR. Each data point in (N) represents an independent biological replicate (littermate pups). One-way Analysis of Variance was performed followed by Tukey's multiple comparison to compare the mean of each group to the mean of every other group within the experiment. Each data point represents an independent biological replicate and graphs indicate the mean +/- the standard error of the mean for each experimental group.

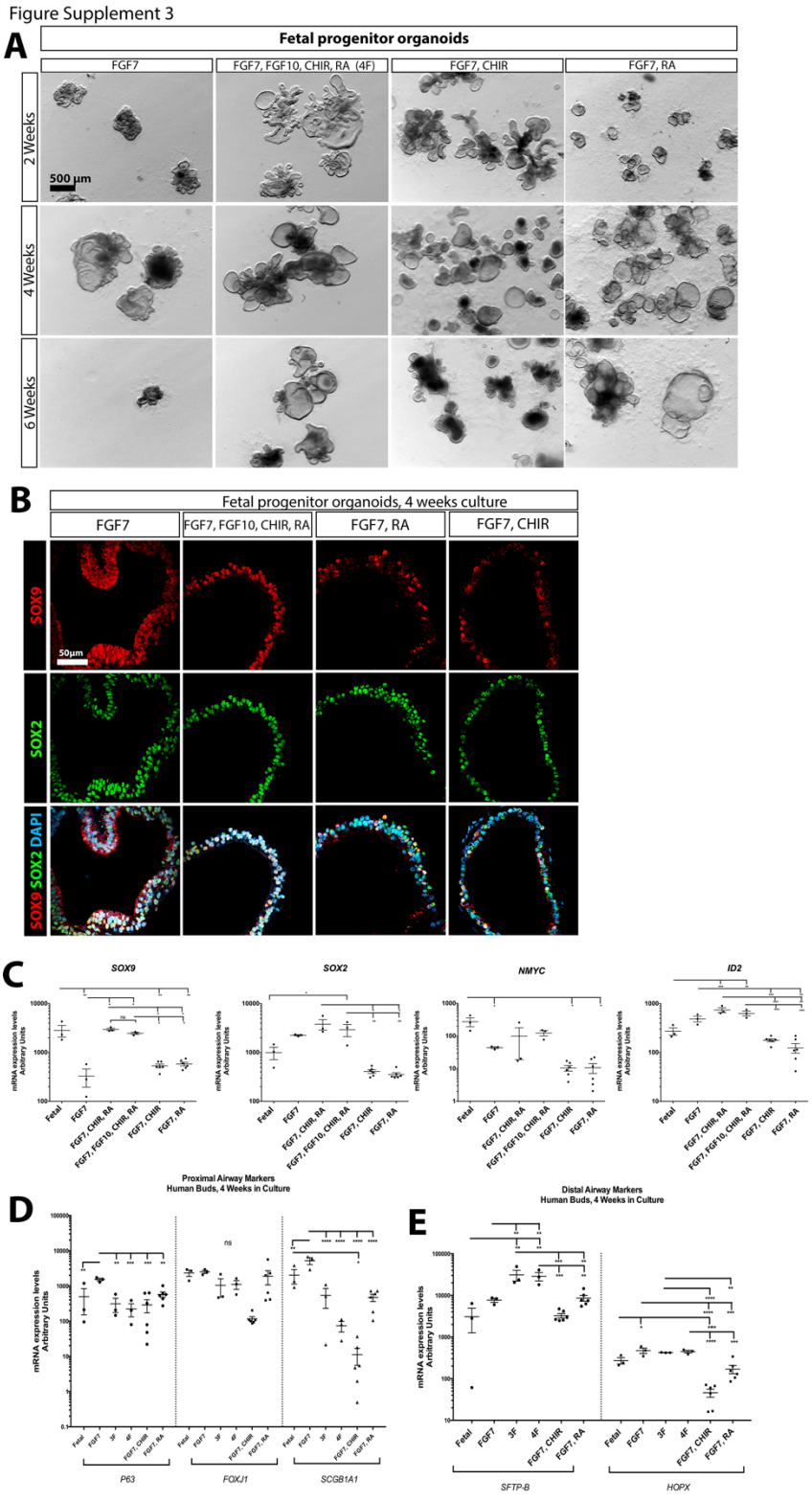
(C, G, I, J, N) $P > 0.05$ ns, $P \leq 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***, $P \leq 0.0001$ ****.

Figure 2-4 FGF7, CHIR-99021 and RA are sufficient to maintain isolated human fetal bud tip progenitor cells *in vitro.*



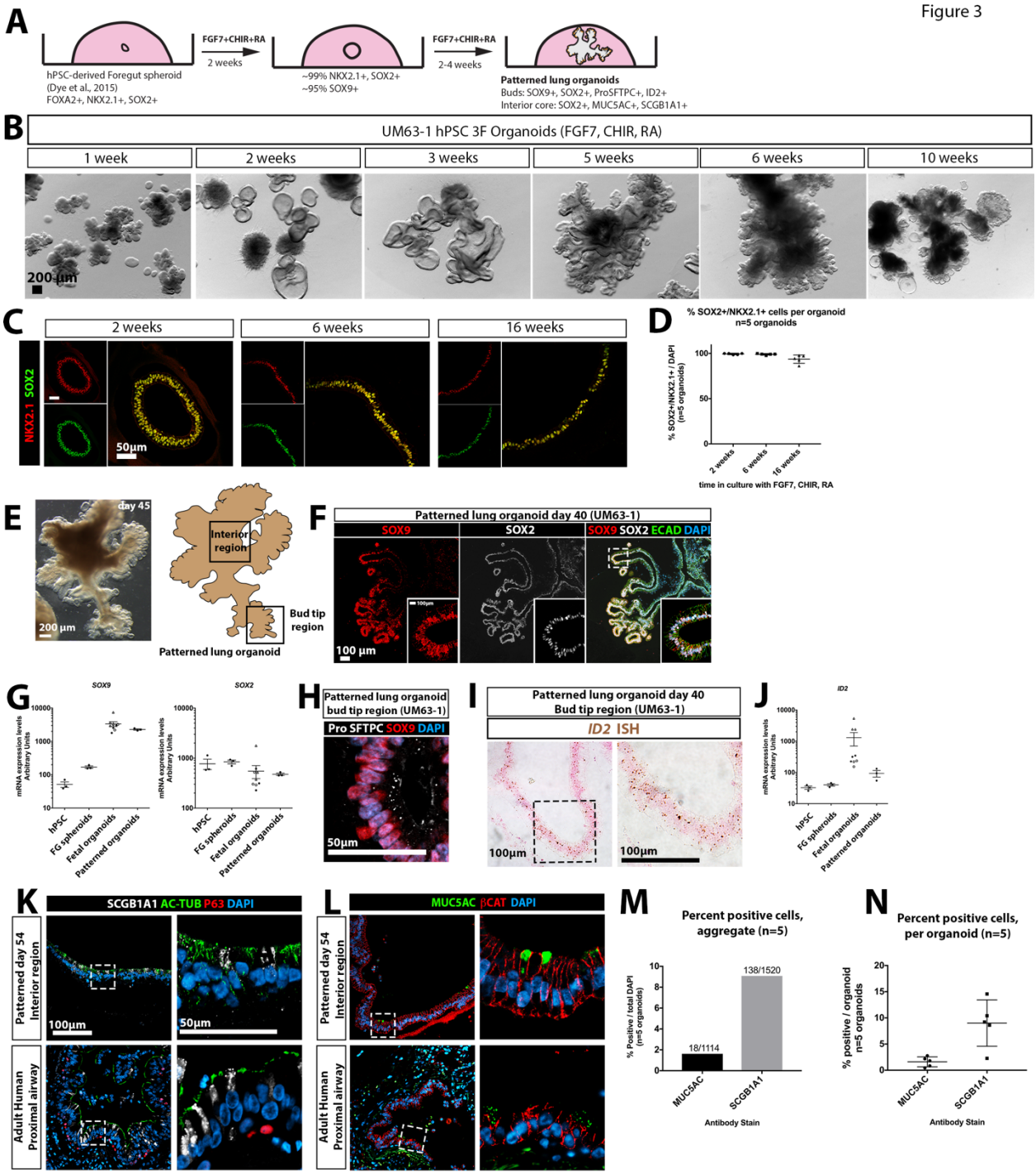
(A-B) Distal epithelial lung bud tips were isolated and cultured in Matrigel droplets. Scale bar in (B) represents 500μm. **(C).** Wholemount brightfield image of human fetal organoids grown in '3F' medium (FGF7, CHIR-99021, RA) at 2, 4 and 6 weeks. Scalebar represents 500 μm. **(D)** Protein staining of SOX2 and SOX9 in sections of fetal progenitor organoids grown in 3F medium. Scale bar represents 100μm. **(E)** Wholemount staining, confocal Z-stack imaging and 3D rendering of SOX2 and SOX9 in fetal progenitor organoids grown in 3F medium. Scale bar represents 100μm. **(F)** Pro-SFTPC after 4 weeks in culture in fetal progenitor organoids. Scale bars represent 50μm. **(G)** *ID2* expression in fetal progenitor organoids after 4 weeks in culture as determined by *in situ* hybridization. Scale bar represents 50μm. **(H)** Heatmap showing expression of genes known to be expressed lung epithelial cells in the whole adult human lung (publically available dataset), in isolated human fetal bud tips (n= 2; 8 weeks and 12 weeks gestation) and in fetal progenitor organoids (n=2; both 12 weeks gestation, cultured for 2 weeks). **(I)** Differential expression analysis (isolated fetal epithelial bud tips vs. whole adult lung; fetal progenitor organoids vs. whole adult lung) was used to identify the top 1000 most highly upregulated genes from each comparison (log2FoldChange < 0; adjusted p-value < 0.05). Gene overlap was identified using a Venn diagram. 27.5% of genes were common to both groups. A hypergeometric means test showed the number of overlapping genes was highly significant (overlapping p-value=1.4e-278).

Figure 2-5 Maintenance of cultured human fetal bud tips *in vitro.*



(A) Bright field image of human fetal bud tips cultured in a droplet of matrigel, overlayed with different media combinations, and examined at 2, 4 and 6 weeks. Scale bar represents 500 μm. **(B)** Immunofluorescent co-expression of SOX9 and SOX2 was examined in different media combinations tested. Scalebar represents 50 μm. **(C)** QRT-PCR analysis of human fetal bud tips after 4 weeks of *in vitro* growth examining expression of *SOX9*, *SOX2, NMYC* and *ID2*. Each data point represents an independent biological replicate (n=3; each 12 week gestation, cultured for 4 weeks) and graphs indicate the mean +/- the standard error of the mean for each experimental group. **(D-E)** QRT-PCR analysis of human fetal bud tips after 4 weeks of *in vitro* growth examining expression of airway cell markers *P63, FOXJ1, SCGB1A1* (D) and alveolar markers *SFTPB* and *HOPX (*E). **(C-E)** Each data point represents an independent biological replicate (n=3; each 12 week gestation, cultured for 4 weeks) and graphs indicate the mean +/- the standard error of the mean for each experimental group. An unpaired, one-way analysis of variance was performed for each experiment followed by Tukey's multiple comparison to compare the mean of each group to the mean of every other group within the experiment. A significance level of 0.05 was used. Significance is shown on graphs according to the following: $P > 0.05$ ns, $P \leq 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***, $P \leq 0.0001$ ****.

Figure 2-6 FGF7, CHIR-99021 and RA generate patterned epithelial lung organoids from hPSCs



(A) Schematic of approach to derive patterned lung organoids (PLOs) from human pluripotent stem cells. **(B)** Brightfield images of hPSC-derived foregut spheroids

cultured in 3F medium (FGF7, CHIR-99021 and RA) and grown *in vitro.* Images taken at 2, 3, 5, 6 and 10 weeks. Images from a single experiment, representative of 6 experiments across 4 hPSC cell lines. Scale bars represent 200 μm. **(C)** Immunostaining for NKX2.1 and SOX2 in PLOs at 2, 6 and 16 weeks. Scale bars represent 50μm. **(D)** Quantitative analysis of cells co-expressing NKX2.1 and SOX2 in PLOs at 2, 6 and 16 weeks, as shown in (B). Each data point (n=5) represents a single organoid from one experiment, and the mean +/- the standard error of the mean is shown for each group. Data is representative of 6 experiments with 4 different cell hPSC lines.**(E)** Brightfield image of a patterned lung organoid after 6 weeks (45 days) in culture, showing distal budded regions and interior regions (Scale bar represents 200μm) and a schematic representing a patterned lung organoid, highlighting bud tip region and interior regions. **(F)** PLOs co-stained for SOX9 and SOX2 protein expression by immunofluorescence. Inset shows high magnification of boxed region. Scale bar represents 100 μm. **(G)** QRT-PCR analysis of *SOX9* and *SOX2* in undifferentiated hPSCs (H9 hESC line, n=3 independent samples from one experiment), in foregut spheroids (FG; n=3 independent samples obtained from one experiment; each sample contains ≥25 pooled spheroids), fetal progenitor organoids (n=3 biological replicates [12 weeks, 12 weeks, 13 weeks] and 3 technical replicates each) and patterned lung organoids (n=3 independent samples obtained from one experiment; each sample contained ≥3 pooled organoids). Data is representative of 6 experiments with 4 different cell hPSC lines. **(H)** Pro-SFTPC and SOX9 co-expression in bud tip region of a PLO. Scale bars represent 50 μm. **(I)** *ID2* expression in PLOs after 40 days *in vitro* as determined by *in situ* hybridization. Scale bar represents 100 μm. **(J)** QRT-PCR

analysis of *ID2* in undifferentiated hPSCs in undifferentiated hPSCs (H9 hESC line, n=3 independent samples from one experiment), in foregut spheroids (FG; n=3 independent samples obtained from one experiment; each sample contains ≥25 pooled spheroids), fetal progenitor organoids (n=3 biological replicates [12 weeks, 12 weeks, 13 weeks] with 3 technical replicates each) and patterned lung organoids (n=3 independent samples obtained from one experiment; each sample contained ≥3 pooled organoids). Data is representative of 6 experiments with 4 different cell hPSC lines. **(K)** Interior regions of patterned lung organoids (top) and adult human airway (bottom) co-stained for SCGB1A1, Acetylated Tubulin (AC-TUB) and P63. Scale bars represent 100 μm (left panels, low magnification) or 50μm (right panels, high magnification). **(L)** Interior regions of patterned lung organoids (top) and adult human airway (bottom) co-stained for MUC5AC and the epithelial marker β–catenin (βCAT). Scale bars represent 100 μm (left panels, low magnification) or 50 μm (right panels, high magnification). **(M-N)** Percent of cells expressing MUC5AC or SCGB1A1, plotted as aggregate data ((M); # cells positive in all organoids/total cells counted in all organoids), and for each individual patterned lung organoid counted ((N); # cells positive in individual organoid/all cells counted in individual organoid). N=5 organoids for each graph, from a single experiment of organoids derived from UM63-1 hPSC line. Data is representative of 6 experiments across 4 hPSC lines.

Figure 2-7 Patterned lung organoids grown from hPSCs exhibit robust and stereotyped growth

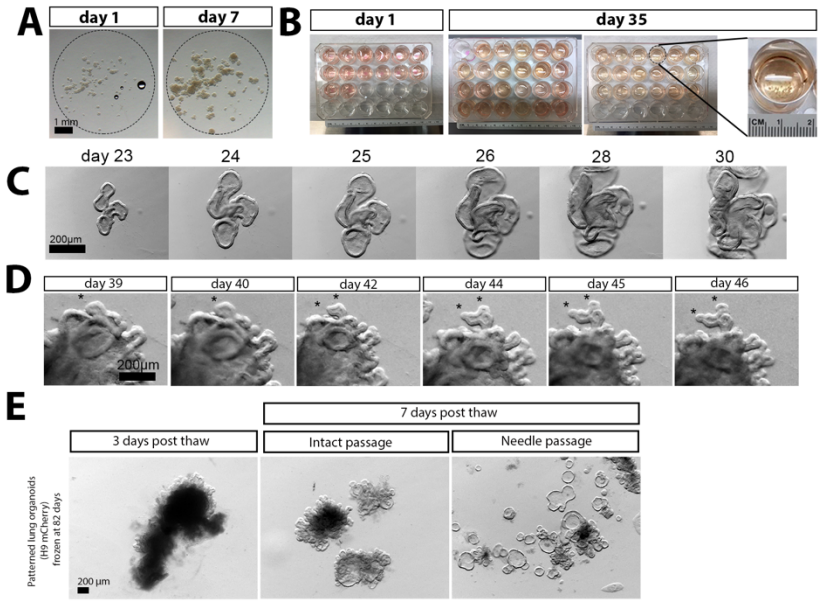
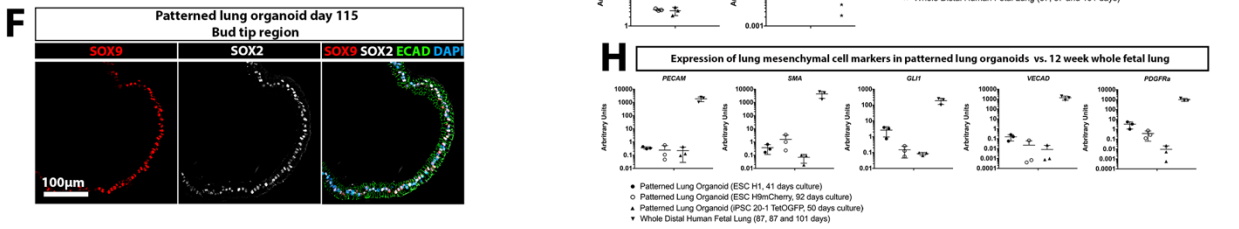


Figure Supplement 4

E′

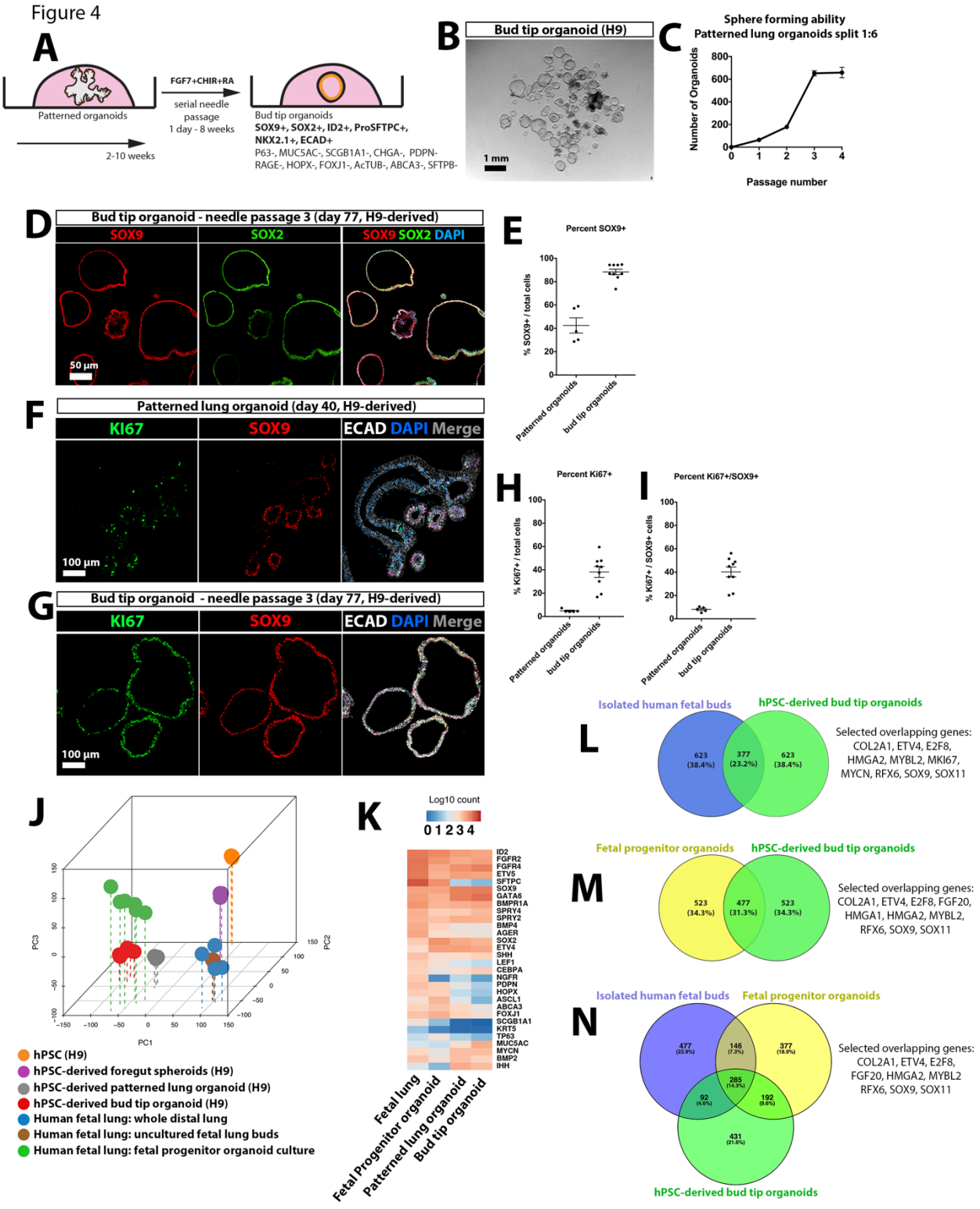
Organoid Cell Line	Total number of organoids thawed	% Surviving Organoids [(Number Alive After 5 Days/ Total Plated) *100, +/- S.D)	Vials Successfully Thawed (Successful/Attempted, (% Success))	Days *In Vitro* Prior to Freezing
hPSC H9 mCherry	95	**79.497%** (+/- 21.492)	3/3 (100%)	82 days
hPSC H1	297	**54.955%** (+/- 25.925)	4/4 (100%)	31 days
iPSC 20-1 tetO-GFP	125	**60.766%** (+/- 19.035)	3/3 (100%)	42 days
Total	517	**63.438%** (+/- 23.465)	10/10 (100%)	



(A) Representative low magnification image of foregut spheroids on day 1 plated in a Matrigel droplet, and cultured in 3F medium for 7 days. Dashed line outlines the Matrigel droplet. Scale bar represents 1 mm. **(B)** Each experiment yielded many wells of spheroids plated in a 24 well plate on day 1, which were subsequently split into multiple 24 well plates to allow sufficient space for growth by day 35. Each well contained many organoids. **(C)** During early stages of growth (day 23-30 shown), epithelial organoids

undergo epithelial 'folding', which typically occurs from weeks 3-5. Scale bar represents 200 μm. **(D)** The same organoid was imaged from day 39 through day 46 and gave rise to epithelial buds (asterisk) at the periphery. Scale bar represents 200 μm. **(E-E')** Patterned lung organoids were frozen and stored in liquid nitrogen, then thawed and grown in 3F medium. After recovering, patterned lung organoids were passaged through a 27-gauge needle to generate cystic bud tip organoids. Freeze/thaw was demonstrated for patterned lung organoids generated from 3 different hPSC lines and survival rate was determined by thawing of multiple vials for each line. **(F)** SOX9 and SOX2 co-expression in bud-tip region of patterned lung organoids after more than 16 weeks (115 days) *in vitro*. Scale bar represents 100 μm. **(G)** Lung epithelial gene expression was examined in patterned lung organoids derived from 4 different hPSC cell lines (H1, H9, UM63-1, iPSC 20-1) and was compared to whole distal human fetal lung by QRT-PCR. **(H)** Mesenchymal gene expression was examined in patterned lung organoids derived from 3 different hPSC cell lines (H1, H9, iPSC 20-1) and was compared to whole distal human fetal lung by QRT-PCR. **(G-H)** Each data point represents n=3 independent replicates. Each replicate contained n≥3 pooled organoids. Graphs indicate the mean +/- the standard error of the mean for each experimental group.

Figure 2-8 Proliferative SOX9+/SOX2+ progenitors can be expanded from patterned lung organoids



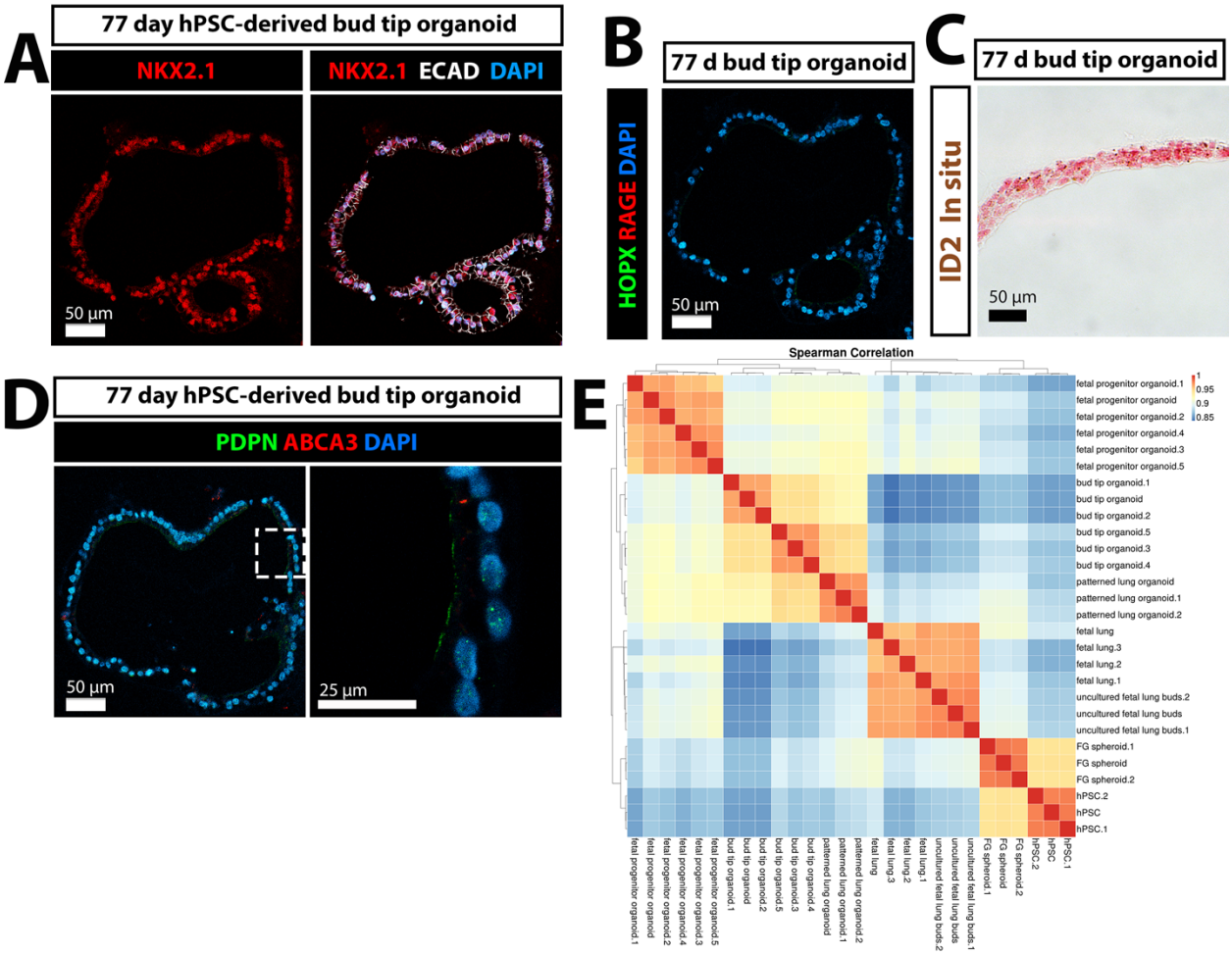
(A) Schematic of approach passage PLOs and expand bud tip organoids. **(B)** Needle sheared epithelial fragments were replated in a fresh Matrigel droplet that subsequently formed cystic structures, called 'bud tip organoids'. Scale bar represents 1mm. **(C)** Quantitative assessment of organoid passaging and expansion. One single patterned lung organoid was needle passaged into 6 wells (passage 1), generating 75 new bud tip organoids in total (average 12.5 per well). A single well of the resulting bud tip organoids were then passaged into 6 new wells after 2 weeks in culture (1:6 split ratio), generating 200 new organoids in total (average 33 per well). This 1:6 passaging was carried out two additional times, every 1-2 weeks for up to 4 passages before growth plateaued. 3 technical replicates were performed for expansion experiments using patterned lung organoids generated from UM63-1 hPSC line; graph plots the mean +/- the SEM. Data is representative of 6 experiments across 4 hPSC lines. **(D)** Immunostaining for SOX9 and SOX2 in bud tip organoids. Scale bar represents 50 μm. **(E)** Quantitation of the percent of SOX9+ cells in PLOs and bud tip organoids (# SOX9+ cells/total cells). Each data point represents an independent organoid (n=5 organoids) derived from a single experiment using H9 hPSC line and graphs indicate the mean +/- the standard error of the mean for each experimental group. Data is representative of 6 experiments across 4 hPSC lines. **(F-G)** Immunostaining for KI67 and SOX9 in patterned lung organoids (F) and bud tip organoids (G). Scale bar represents 100um. **(H)** Quantitation of the percent of all cells that were KI67+ in patterned and bud tip organoids (# KI67+ cells /total cells). **(I)** Quantitation of the percent of proliferating SOX9+ cells in patterned and bud tip organoids (# KI67+/SOX9+ cells/total cells. **(H-I)** Each data point represents an independent organoid (n=5 patterned lung organoids and

n=9 bud tip organoids) derived from a single experiment using H9 hPSC line and graphs indicate the mean +/- the standard error of the mean for each experimental group. Data is representative of 6 experiments across 4 hPSC lines. **(J)** Principal component analysis of RNA-sequencing data to compare the global transcriptome of hPSCs (n=3 independent replicates of H9 hPSCs), foregut spheroids (n=3 independent replicates of H9 hPSCs-derived spheroids. Each replicate contained ≥25 pooled spheroids), hPSC-derived patterned lung organoids (n=3 independent replicates of H9 hPSCs-derived patterned lung organoids; each replicate contained n≥3 pooled organoids), hPSC-derived bud tip organoids (n=3 independent replicates of H9 hPSCs-derived bud tip organoids; each replicate contained n≥3 organoids), whole peripheral (distal) fetal lung (n=3; 8, 12 and 18 weeks), freshly isolated (uncultured) fetal lung buds (n= 2; 8 weeks and 12 weeks gestation) and fetal progenitor organoids (n=2; both 12 weeks gestation and cultured for 2 weeks). All biological replicates were run in technical triplicate. **(K)** Heatmap showing expression of genes known to be expressed in lung epithelial cells in freshly isolated (uncultured) fetal lung buds and fetal progenitor organoids cultured for 2 weeks. **(L-N)** Differential expression analysis of: 1. isolated fetal bud tips vs. whole adult lung; 2. fetal progenitor organoids vs whole adult lung; 3. bud tip organoids vs. whole adult lung was used to identify the top 1000 most highly upregulated genes from each comparison (log2FoldChange < 0; adjusted p-value < 0.05). A Venn diagram illustrates common upregulated genes in **(L)** isolated bud tips and bud tip organoids; **(M)** fetal progenitor organoids and bud tip organoids **(N)** isolated bud tips, fetal progenitor organoids, bud tip organoids. **(L)** hypergeometric means test found that the shared gene overlap was highly significant (overlapping p-value=9.3e-901); **(M)** hypergeometric

means test found that the shared gene overlap was highly significant (overlapping p-value=1.2e-1021). **(N)** 285 overlapping genes were shared between the three groups. These genes represented 14.3% of all genes included in the comparison. A small subset of genes previously associated with human or mouse bud-tip progenitor cells are highlighted as “Selected overlapping genes” (L-N).

Figure 2-9 Characterization of hPSC-derived bud tip organoids.

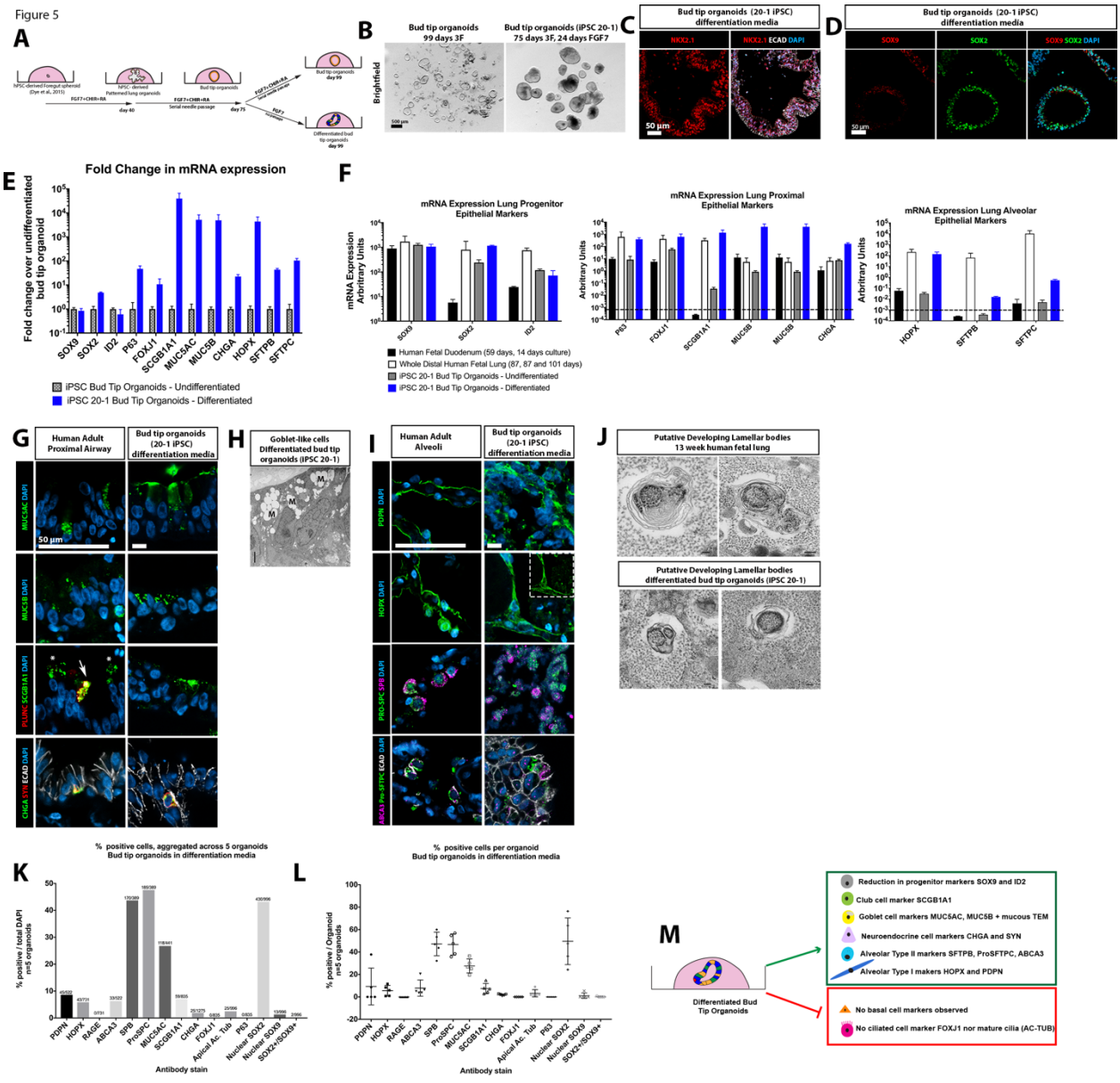
Figure Supplement 5



(A) ECAD and NKX2.1 immunofluorescence in bud tip organoids shows that bud tip organoids express NKX2.1. Scale bar represents 50 µm. **(B)** Immunofluorescence analysis of HOPX and RAGE in bud tip organoids show that neither marker is expressed, similar to what is seen in 15 week human fetal bud tip cells (Figure 1). Scale bar represents 50 µm. **(C)** *In situ* hybridization for *ID2* in bud tip organoids (brown dots indicate positive staining). Scale bar represents 50 µm. **(D)** Immunofluorescence analysis of PDPN and ABCA3 in bud tip organoids shows that neither PDPN nor ABCA3 are expressed in organoids, similar to what is seen in 15 week human fetal bud

tip cells (Figure 1). Scale bar represents 50 μm. **(E)** Spearman correlation matrix of RNAsequencing datasets for undifferentiated H9 hESCs (n=3 independent replicates of H9 hPSCs), Foregut (FG) spheroids derived from H9 hESCs (n=3 independent replicates of H9 hPSCs-derived spheroids. Each replicate contained ≥25 pooled spheroids), whole distal fetal lung (n=3 independent biological replicates from 8, 12 and 18 weeks gestation samples, each run in technical triplicate), freshly isolated bud tips (n=2 biological replicates from 8 and 12 weeks gestation; each sample run with 3 independent technical replicates), fetal progenitor organoids grown in 3F medium (n=2; both 12 weeks gestation cultured for 2 weeks and run in technical triplicate), hPSC-derived patterned lung organoids (n=3 independent replicates of H9 hPSCs-derived patterned lung organoids; each replicate contained n≥3 pooled organoids) and hPSC-derived bud tip organoids (each replicate is an independent sample of n≥3 pooled organoids. 3 samples were derived from iPSC20-1 line and 3 from H9 hPSC line).

Figure 2-10 Bud tip organoids retain multilineage potential *in vitro*

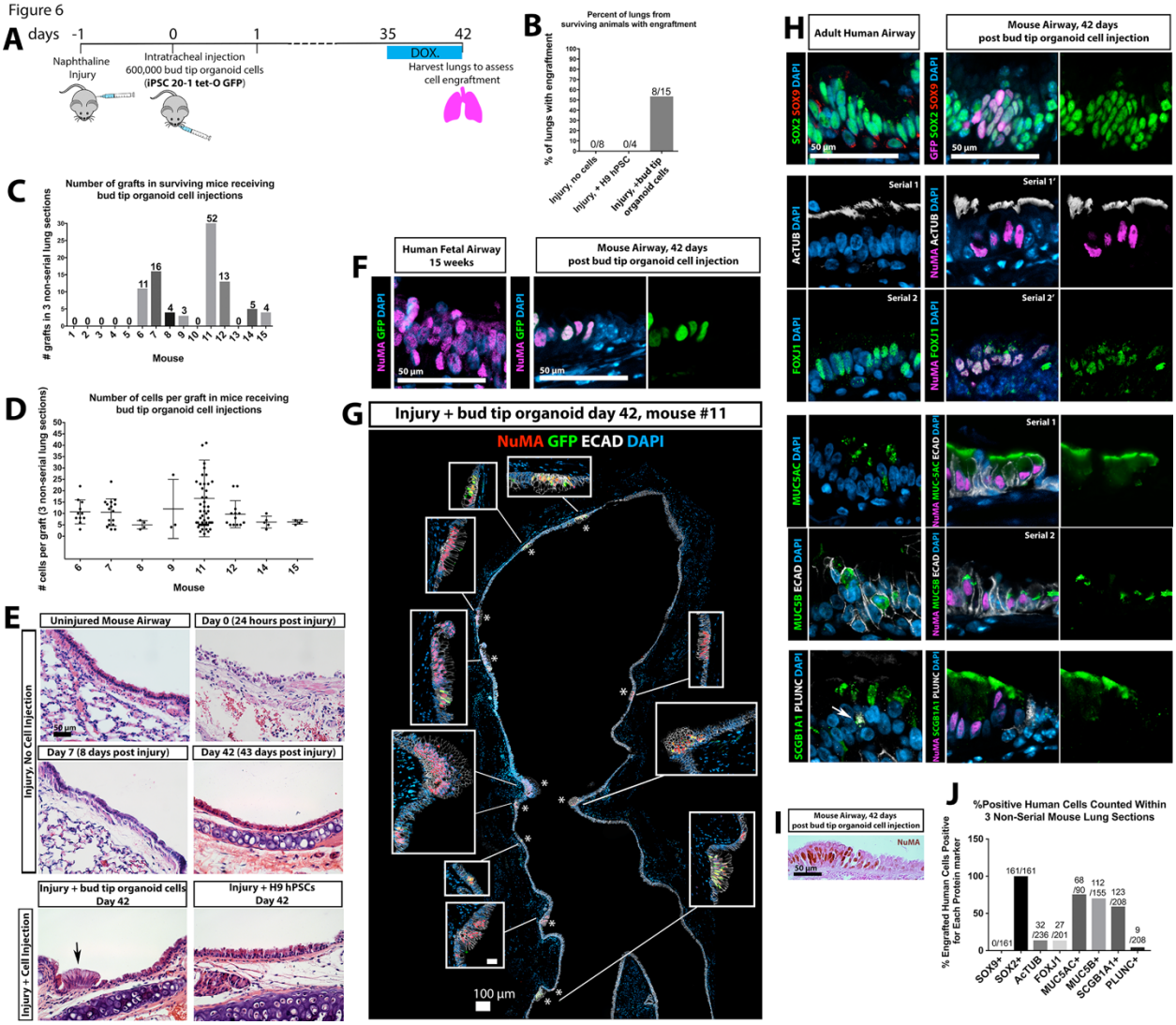


(A) Schematic of experimental setup. iPSC20-1 bud tip organoids were initially cultured in 3F medium and subsequently grown in 3F medium or medium containing FGF7 alone ('differentiation media') for 24 days. **(B)** Brightfield images of bud tip organoids growing in 3F media (left) or FGF7-differentiation media (right). Scale bar represents 500 µm.

(C) NKX2.1 immunofluorescence of bud tip organoids grown in FGF7-differentiation media for 24 days. Scale bar represents 50 μm. **(D)** SOX9 and SOX2 immunofluorescence of bud tip organoids grown in FGF7-differentiation media for 24 days. Scale bar represents 50 μm. **(E)** QRT-PCR analysis of bud tip organoids grown in 3F or FGF7-differentiation media showing expression of several genes expressed in the lung epithelium. Data is shown as Fold Change relative to 3F-grown undifferentiated bud tip organoids. Each data point represents an independent sample (n=3) obtained from one experiment derived from iPSC 20-1 hPSC line; each sample contained ≥3 pooled organoids. Data is representative of 5 experiments across 2 hPSC lines (iPSC20-1 and H9). **(F)** QRT-PCR analysis of bud tip organoids grown in 3F or FGF7-differentiation media, along with whole distal fetal lung and cultured whole fetal intestine as reference tissues, showing expression of several genes expressed in the lung epithelium. Data is shown as Arbitrary Units. Values lower than 10^{-3} were considered undetected. Each data point in bud tip organoid groups represents an independent sample (n=3) obtained from one experiment derived from iPSC 20-1 hPSC line; each sample contained ≥3 pooled organoids. Data is representative of 5 experiments across 2 hPSC lines (iPSC20-1 and H9). Whole human distal lung samples (n=3) were taken from 87, 87 and 101 day gestation, and 3 independent samples of human fetal duodenum tissue were used as a comparison. **(G)** Immunostaining for airway markers in the adult human lung, and in bud tip organoids grown in FGF7-differentiation media. Markers are shown for goblet cells (MUC5AC, MUC5B), club cells (SCGB1A1, PLUNC) and neuroendocrine cells (SYN, CHGA). Scale bars represents 50 μm. **(H)** Transmission Electron Microscopy through a bud tip organoid grown in FGF7-

differentiation media showing mucus filled vacuoles characteristic of goblet cells. Scale bar represents 100 nm. **(I)** Immunostaining for alveolar markers in the adult human lung, and in bud tip organoids grown in FGF7-differentiation media. Markers are shown for AECI cells (PDPN, HOPX) and AECII cells (Pro-SFTPC, SFTPB, ABCA3). Scale bars represent 50 μm. **(J)** Transmission Electron Microscopy of a human fetal lung at 13 weeks of gestation, and of a bud tip organoid grown in FGF7-differentiation media showing immature lamellar bodies surrounded by monoparticulate glycogen, characteristic of immature AECII cells. Scale bars represent 100 nm. **(K-L)** Quantitation of cell type markers in bud tip organoids grow in differentiation media plotted as aggregate data ((H) numbers at top of bars represent positive cells/total cells counted across 5 individual organoids), and as individual data per organoid ((I) number of positive cells per organoid). Graphs indicate the mean +/- the standard error of the mean. Data is from a single differentiation experiment derived from iPSC 20-1 hPSC line and each data point represents an individual organoid (n=5 organoids). Data is representative of 5 experiments across 2 hPSC lines (iPSC20-1 and H9). **(M)** Summary of putative differentiated lung epithelial cell types generated *in vitro*.

Figure 2-11 Engraftment of hPSC-derived bud tip progenitor organoid cells into the injured mouse airway

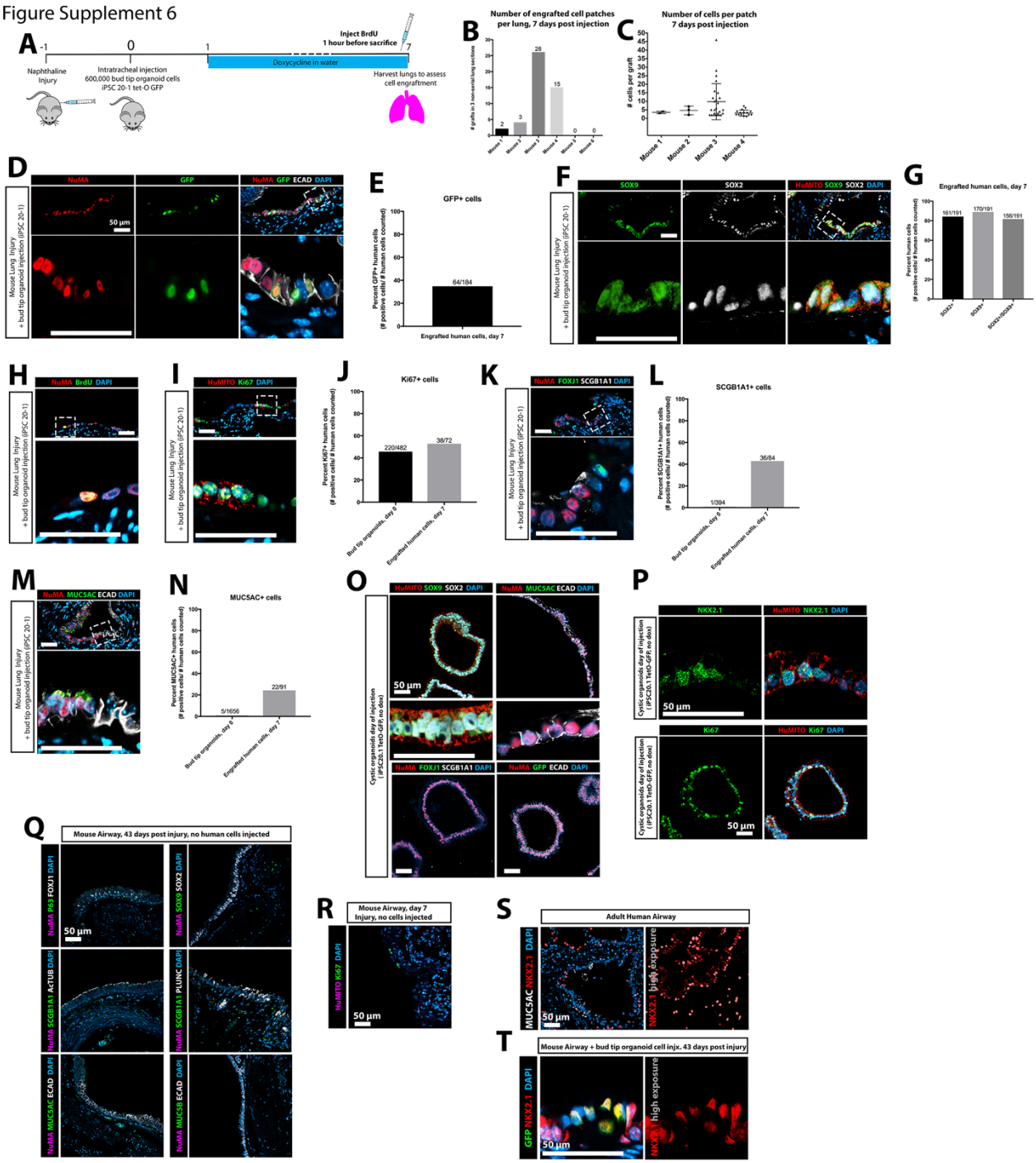


(A) Schematic of experimental design. Immunocompromised NSG male mice were injected with 300mg/kg of Naphthalene. 24 hours post-injury, mice were randomly assigned to receive an intratracheal injection of 600,000 single cells isolated from bud tip organoids generated from the iPSC 20-1 tet-O GFP line, undifferentiated H9 hPSCs, or no injection of cells. Doxycycline (1mg/ml) was added to the drinking water for the

final week to induce expression of the tet-O GFP construct. Lungs were analyzed 6 weeks after cell injection. **(B)** Percent of lungs from surviving animals in each group exhibiting engraftment of human cells after 6 weeks, as determined by NuMA and GFP protein staining. **(C-D)** Engraftment was assessed based on human specific expression on NuMA and GFP in 3 independent histological sections from each surviving mouse. (C) The number of engraftment cell patches observed in 15 surviving animals in the group receiving bud tip organoid cells. (D) Quantitation of the number of human cells in each engrafted cell patch, in each mouse. Every data point represents the number of cells in a single patch of cells. **(E)** Hematoxylin & Eosin staining documenting the lung epithelial airway injury in control mice (injury, no cell injection), and in mice that received bud tip organoids, or undifferentiated iPSC injections. Engrafted patches of human cells were obvious (arrow), and were confirmed in serial sections using human specific antibodies. Scale bar for all images represents 50 μm. **(F)** Immunostaining for NuMA and GFP in human fetal lungs and in bud tip organoid transplanted lungs. Scale bar represents 50 μm. **(G)** Bud tip organoid transplanted lung showing several images stiched together to generate a large panel demonstrating multiple sites of engraftment, marked by NuMA and/or GFP, in the upper airway. High magnification insets are shown. Scale bar represents 100μm. **(H)** Immunostaining of adult human lung tissue and bud tip organoid transplanted lung tissue showing immunostaining for several lung epithelial markers, including SOX9 and SOX2, multiciliated cell markers Ac-TUB and FOXJ1, goblet cell markers MUC5AC and MUC5B, club cell markers SCGB1A1 and PLUNC. Scale bar represents 50 μm. **(I)** Bright field microscopic image showing immunohistochemistry for NuMA in a patch of engrafted cells, counterstained using

eosin to visualize tufts of multiple cilia. Scale bar represents 50 μm. **(J)** Quantification of cell type markers in bud tip organoid transplanted lungs after 6 weeks. Data is plotted as aggregate data (numbers at top of bars represent positive cells/total cells counted across 3 engrafted lungs). Aggregated data from 3 non-serial sections for mouse 7, 11 and 12 is plotted.

Figure 2-12 Short-term engraftment of bud tip organoids into injured mouse lungs



(A) Schematic of experimental design. **(B-C)** Characterization of surviving cell-injected mice 7 days post cell injection. Engraftment was assessed based on human specific

antibodies (NuMA, huMITO) in 3 independent histological sections from each surviving mouse. (B) Engraftment was observed in 4/6 animals. (C) Quantitation of the number of human cells in each engrafted cell 'patch'. Each data point represents the number of cells in a single patch of cells. **(D-E)** Immunostaining for NuMA, ECAD and GFP in transplanted lungs (D). Low magnification (top) and high magnification (bottom) images are shown. Scale bars represent 50 μm. (E) shows quantification of GFP+ cells following *in vivo* engraftmet and dox induction after 7 days. Aggregate data from 3 non-serial sections from 3 mice is plotted for day 7. **(F-G)** Immunostaining for HuMITO, SOX9 and SOX2 in transplanted lungs(F). Low magnification (top) and high magnification (bottom) images are shown. Scale bars represent 50 μm. (G) shows quantification of immunostaining, comparing *in vitro* bud tip organoids (day 0) and following *in vivo* engraftment (day 7). Aggregate data from 3 non-serial sections from 3 mice is plotted for day 7. **(H)** Immunostaining for human nuclear specific antibody (NuMA) and BrdU in transplanted lungs. Low magnification (top) and high magnification (bottom) images are shown. Scale bars represent 50 μm. **(I-J)** Immunostaining for HuMito and KI67 demonstrates proliferating human cells in transplanted lungs (I). Low magnification (top) and high magnification (bottom) images are shown. Scale bars represent 50 μm. (J) Quantitation of human cells that co-express the proliferation marker KI67 comparing proliferation *in vitro* bud tip organoids (day 0) and following *in vivo* engraftment (day 7). Aggregate data is plotted showing total number of cells positive for KI67, and the total number of cells counted from 3 non-serial sections from 3 different mice is plotted for day 7. **(K-L)** Immunostaining for NuMA, FOXJ1 and SCGB1A1 in transplanted lungs (L). Low magnification (top) and high magnification

(bottom) images are shown. Scale bars represent 50 μm. Note that the FOXJ1+ cell shown is assumed to be of murine origin, since it does not express NuMA. (L) shows quantification of immunostaining, comparing *in vitro* bud tip organoids (day 0) and following *in vivo* engraftment (day 7). Aggregate data from 3 non-serial sections from 3 mice is plotted for day 7. **(M-N)** Immunostaining for NuMA and MUC5AC in transplanted lungs. (M) Low magnification (top) and high magnification (bottom) images are shown. Scale bars represent 50um. (N) shows quantification of immunostaining, comparing *in vitro* bud tip organoids (day 0) and following *in vivo* engraftment (day 7). Aggregate data from 3 non-serial sections from 3 mice is plotted for day 7. **(O-P**) Characterization of bud tip organoids generated from iPSC 20-1 on the day of transplantation into the injured mouse airway.**(O)** Bud tip organoids immunostained for HuMITO, SOX9 and SOX2 at low magnification (top) and high magnification (bottom). Negative staining in bud tip organoids is shown for MUC5AC, FOXJ1 and SCGB1A1. Scale bar for all images represents 50 μm. **(P)** Bud tip organoids immunostained for NKX2.1 and HuMITO (top), or for KI67 and HuMITO (bottom). Scale bar represents 50 um.Scale bar represents 50 μm. **(Q)** Staining of mouse lungs that were injured, but that did not receive human cell injections 43 days post injury, to demonstrate that the human specific antibody, NuMA, is highly specific to human cells and is not detected in the murine airway. Moreover, immunostaining for several markers shows positive staining in the mouse airway, including P63, FOXJ1, SOX2, AcTUB, PLUNC and ECAD. Note that MUC5AC, MUC5B and SCGB1A1 antibodies are not reactive to mouse airway cells. Scale bare represents 50 μm. **(S)** Positive control staining for NKX2.1 in the adult human lungs, demonstrating that only a subset of airway epithelium expresses high

levels of NKX2.1 (left, normal exposure), whereas most epithelial cells require over-exposure during imaging in order to visualize NKX2.1 (right image). MUC5AC appears to predominantly be expressed in NKX2.1-low expressing cells (right image). Scale bar represents 50 μm. **(T)** Immunofluorescence staining for NKX2.1 in GFP+ cells engrafted into the mouse airway 43 days post injury. Overexposure of NKX2.1 shows positive staining in cells of human origin. Scale bar represents 50 μm.

2.7 Tables

Table 2-1. Description of the nomenclature used in this manuscript

Sample nomenclature	Description
Peripheral fetal lung	The distal/peripheral portion of the fetal lung (i.e. distal 0.5 cm) was excised from the rest of the lung using a scalpel. This includes all components of the lung (e.g. epithelial, mesenchymal, vascular).
Isolated fetal bud tip	The bud peripheral portion of the fetal lung was excised with a scalpel and subjected to enzymatic digestion and microdissection. The epithelium was dissected and separated from the mesenchyme, but a small amount of associated mesenchyme likely remained.
Fetal progenitor organoid	3D organoid structures that arose from culturing isolated fetal epithelial bud tips.
Foregut spheroid	3D foregut endoderm structure as described in Dye et al., 2015. Gives

	rise to PLO when grown in 3F media.
Patterned lung organoid (PLO)	Lung organoids that were generated by differentiating hPSCs, as described throughout the manuscript.
Bud tip organoid	Organoids derived from PLOs, enriched for SOX9/SOX9 co-expressing cells, and grown/passaged in 3F medium.

Table 2-2 QRT-PCR Primer sequences

Species	Gene Target	Forward Primer Sequence	Reverse Primer Sequence
Mouse	*Aqp5*	TAGAAGATGGCTCGGAGCAG	CTGGGACCTGTGAGTGGTG
Mouse	*Foxj1*	TGTTCAAGGACAGGTTGTGG	GATCACTCTGTCGGCCATCT
Mouse	*Gapdh*	TGTCAGCAATGCATCCTGCA	CCGTTCAGCTCTGGGATGAC
Mouse	*Id2*	AGAAAAGAAAAAGTCCCCAAATG	GTCCTTGCAGGCATCTGAAT
Mouse	*Nmyc*	AGCACCTCCGGAGAGGATA	TCTCTACGGTGACCACATCG
Mouse	*P63*	AGCTTCTTCAGTTCGGTGGA	CCTCCAACACAGATTACCCG
Mouse	*Scgb1a1*	ACTTGAAGAAATCCTGGGCA	CAAAGCCTCCAACCTCTACC
Mouse	*Sftpb*	ACAGCCAGCACACCCTTG	TTCTCTGAGCAACAGCTCCC
Mouse	*Sox2*	AAAGCGTTAATTTGGATGGG	ACAAGAGAATTGGGAGGGGT
Mouse	*Sox9*	TCCACGAAGGGTCTCTTCTC	AGGAAGCTGGCAGACCAGTA
Human	*FOXJ1*	CAACTTCTGCTACTTCCGCC	CGAGGCACTTTGATGAAGC
Human	*GAPDH*	AATGAAGGGGTCATTGATGG	AAGGTGAAGGTCGGAGTCAA
Human	*HOPX*	GCCTTTCCGAGGAGGAGAC	TCTGTGACGGATCTGCACTC

Human	*ID2*	GACAGCAAAGCACTGTGTGG	TCAGCACTTAAAAGATTCCGTG
Human	*MUC5AC**	GCACCAACGACAGGAAGGATGAG	CACGTTCCAGAGCCGGACAT
Human	*NKX2.1*	CTCATGTTCATGCCGCTC	GACACCATGAGGAACAGCG
Human	*NMYC*	CACAGTGACCACGTCGATTT	CACAAGGCCCTCAGTACCTC
Human	*P63*	CCACAGTACACGAACCTGGG	CCGTTCTGAATCTGCTGGTCC
Human	*SCGB1A1*	ATGAAACTCGCTGTCACCCT	GTTTCGATGACACGCTGAAA
Human	*SFTPB*	CAGCACTTTAAAGGACGGTGT	GGGTGTGTGGGACCATGT
Human	*SOX2*	GCTTAGCCTCGTCGATGAAC	AACCCCAAGATGCACAACTC
Human	*SOX9*	GTACCCGCACTTGCACAAC	ATTCCACTTTGCGTTCAAGG
Human	*SFTPC*	AGCAAAGAGGTCCTGATGGA	CGATAAGAAGGCGTTTCAGG

Note: All primer sequences were obtained from http://primerdepot.nci.nih.gov (human) or http://mouseprimerdepot.nci.nih.gov (mouse) unless otherwise noted. All annealing temperatures are near 60°C.

*MUC5AC Huang, SX et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. Nature Biotechnol. 1–11 (2013). doi:10.1038/nbt.2754

Table 2-3 Antibody information

Primary Antibody	Source	Catalog #	Used for Species	Dilution (Sections)	Dilution (Whole mount)	Clone
Goat anti-CC10 (SCGB1A1)	Santa Cruz Biotechnology	sc-9770	Mouse, Human	1:200		C-20
Goat anti-Chromogranin A (CHGA)	Santa Cruz Biotechnology	sc-1488	Human	1:100		C-20
Goat anti-SOX2	Santa Cruz Biotechnology	Sc-17320	Mouse, Human	1:200	1:100	polyclonal
Mouse anti-ABCA3	Seven Hills Bioreagents	WMAB-17G524	Human	1:500		17-H5-24
Mouse anti-Acetylated Tubulin (ACTTUB)	Sigma-Aldrich	T7451	Mouse, Human	1:1000		6-11B-1
Mouse anti-E-Cadherin (ECAD)	BD Transduction Laboratories	610181	Mouse, Human	1:500		36/E-Cadherin
Mouse anti- human mitochondria	Millipore	MAB1273	Human	1:500		113-1
Mouse anti-human nuclear matrix protein-22 (NuMA)	Thermofisher	PA5-22285	Human	1:500		polyclonal
Mouse anti-Surfactant Protein B (SFTPB)	Seven Hills Bioreagents	Wmab-1B9	Mouse, Human	1:250		monoclonal
Rabbit anti-Aquaporin 5 (Aqp5)	Abcam	Ab78486	Mouse	1:500		polyclonal
Rabbit anti-Clara Cell Secretory Protein (CCSP; SCGB1A1)	Seven Hills Bioreagents	Wrab-3950	Mouse, Human	1:250		polyclonal
Rabbit anti-HOPX	Santa Cruz Biotechnology	Sc-30216	Human	1:250		polyclonal
Mouse anti-MUC5B	Abcam	AB77995	Human	1:250		monoclonal
Rabbit anti-NKX2.1	Abcam	ab76013	Human	1:200		EP1584Y
Rabbit anti-PDPN	Santa Cruz Biotechnology	Sc-134482	Human	1:500		polyclonal
Mouse anti-PLUNC	R&D systems	MAP1897	Mouse, Human	1:500		monoclonal
Rabbit anti-Pro-Surfactant protein C (Pro-SFTPC)	Seven Hills Bioreagents	Wrab-9337	Human, Mouse	1:500		polyclonal
Rabbit anti-P63	Santa Cruz	sc-8344	Mouse, Human	1:200		H-129

	Biotechnology					
Rabbit anti-SOX9	Millipore	AB5535	Mouse, Human	1:500	1:250	polyclonal
Rabbit anti-Synaptophysin	Abcam	AB32127	Human	1:500		monoclonal
Rat anti-KI67	Biolegend	652402	Mouse	1:100		16A8
*Biotin-Mouse anti MUC5AC	Abcam	ab79082	Human	1:500		Monoclonal
Secondary Antibody	**Source**	**Catalog #**		**Dilution**		
Donkey anti-goat 488	Jackson Immuno	705-545-147		1:500		
Donkey anti-goat 647	Jackson Immuno	705-605-147		1:500		
Donkey anti-goat Cy3	Jackson Immuno	705-165-147		1:500		
Donkey anti-mouse 488	Jackson Immuno	715-545-150		1:500		
Donkey anti-mouse 647	Jackson Immuno	415-605-350		1:500		
Donkey anti-mouse Cy3	Jackson Immuno	715-165-150		1:500		
Donkey anti-rabbit 488	Jackson Immuno	711-545-152		1:500		
Donkey anti-rabbit 647	Jackson Immuno	711-605-152		1:500		
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102		1:500		
Donkey anti-goat 488	Jackson Immuno	705-545-147		1:500		
Donkey anti-goat 647	Jackson Immuno	705-605-147		1:500		
Donkey anti-goat Cy3	Jackson Immuno	705-165-147		1:500		
Donkey anti-mouse 488	Jackson Immuno	715-545-150		1:500		
Donkey anti-mouse 647	Jackson Immuno	415-605-350		1:500		
Donkey anti-mouse Cy3	Jackson Immuno	715-165-150		1:500		
Donkey anti-rabbit 488	Jackson Immuno	711-545-152		1:500		
Donkey anti-rabbit 647	Jackson Immuno	711-605-152		1:500		
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102		1:500		
Streptavidin 488	Jackson Immuno	016-540-084		1:500		

2.8 References

1. Metzger RJ, Klein OD, Martin GR, Krasnow MA. The branching programme of mouse lung development. Nature. 2008 Jun 5;453(7196):745–50.

2. Rawlins EL, Clark CP, Xue Y, Hogan BLM. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development. 2009 Nov;136(22):3741–5.

3. Morrisey EE, Hogan BLM. Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. Dev Cell. 2010 Jan;18(1):8–23.

4. Varner VD, Nelson CM. Cellular and physical mechanisms of branching morphogenesis. Development. 2014 Jul;141(14):2750–9.

5. Miller AJ, Spence JR. In Vitro Models to Study Human Lung Development, Disease and Homeostasis. Physiology (Bethesda). 2017 May;32(3):246–60.

6. Dye BR, Miller AJ, Spence JR. How to Grow a Lung: Applying Principles of Developmental Biology to Generate Lung Lineages from Human Pluripotent Stem Cells. Curr Pathobiol Rep. 2016;4:47–57.

7. Dye BR, Hill DR, Ferguson MA, Tsai Y-H, Nagy MS, Dyal R, et al. In vitro generation of human pluripotent stem cell derived lung organoids. Elife. 2015;4.

8. Chen Y-W, Huang SX, de Carvalho ALRT, Ho S-H, Islam MN, Volpi S, et al. A three-dimensional model of human lung development and disease from pluripotent stem cells. Nat Cell Biol. 2017 Apr 24;19(5):542–9.

9. Okubo T. Nmyc plays an essential role during lung development as a dosage-sensitive regulator of progenitor cell proliferation and differentiation. Development. 2005 Feb 9;132(6):1363–74.

10. Perl A-KT, Kist R, Shan Z, Scherer G, Whitsett JA. Normal lung development and function afterSox9 inactivation in the respiratory epithelium. genesis. 2005;41(1):23–32.

11. Rockich BE, Hrycaj SM, Shih HP, Nagy MS, Ferguson MAH, Kopp JL, et al. Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proceedings of the National Academy of Sciences. 2013 Nov 19;110(47):E4456–64.

12. Chang DR, Alanis DM, Miller RK, Ji H, Akiyama H, McCrea PD, et al. Lung epithelial branching program antagonizes alveolar differentiation. Proceedings of the National Academy of Sciences. 2013 Nov 5;110(45):18042–51.

13. Nikolić MZ, Caritg O, Jeng Q, Johnson J-A, Sun D, Howell KJ, et al. Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. Elife. 2017 Jun 30;6.

14. Danopoulos S, Alonso I, Thornton ME, Grubbs BH, Bellusci S, Warburton D, et al. Human lung branching morphogenesis is orchestrated by the spatiotemporal

distribution of ACTA2, SOX2, and SOX9. AJP: Lung Cellular and Molecular Physiology. 2018 Jan 1;314(1):L144–9.

15. Dye BR, Dedhia PH, Miller AJ, Nagy MS, White ES, Shea LD, et al. A bioengineered niche promotes in vivo engraftment and maturation of pluripotent stem cell derived human lung organoids. Elife. 2016 Sep 28;5.

16. Spence JR, Mayhew CN, Rankin SA, Kuhar MF, Vallance JE, Tolle K, et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature. 2011 Feb 3;470(7332):105–9.

17. Huang SXL, Islam MN, O'Neill J, Hu Z, Yang Y-G, Chen Y-W, et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. Nature Biotechnology. 2013 Dec 1.

18. Wong AP, Bear CE, Chin S, Pasceri P, Thompson TO, Huan L-J, et al. Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. Nature Biotechnology. Nature Publishing Group; 2012 Aug 26;30(9):875–81.

19. Gotoh S, Ito I, Nagasaki T, Yamamoto Y, Konishi S, Korogi Y, et al. Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. Stem Cell Reports. 2014 Sep 9;3(3):394–403.

20. Konishi S, Gotoh S, Tateishi K, Yamamoto Y, Korogi Y, Nagasaki T, et al. Directed Induction of Functional Multi-ciliated Cells in Proximal Airway Epithelial

Spheroids from Human Pluripotent Stem Cells. Stem Cell Reports. 2016 Jan 12;6(1):18–25.

21. Longmire TA, Ikonomou L, Hawkins F, Christodoulou C, Cao Y, Jean JC, et al. Efficient Derivation of Purified Lung and Thyroid Progenitors from Embryonic Stem Cells. Stem Cell. Elsevier Inc; 2012 Apr 6;10(4):398–411.

22. McCauley KB, Hawkins F, Serra M, Thomas DC, Jacob A, Kotton DN. Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. Cell Stem Cell. 2017 Jun 1;20(6):844–6.

23. Jacob A, Morley M, Hawkins F, McCauley KB, Jean JC, Heins H, et al. Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. Cell Stem Cell. 2017 Oct 5;21(4):472–488.e10.

24. Branchfield K, Li R, Lungova V, Verheyden JM, McCulley D, Sun X. A three-dimensional study of alveologenesis in mouse lung. Developmental Biology. 2016 Jan 15;409(2):429–41.

25. D'Amour KA, Agulnick AD, Eliazer S, Kelly OG, Kroon E, Baetge EE. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nature Biotechnology. 2005 Dec;23(12):1534–41.

26. Green MD, Chen A, Nostro M-C, d'Souza SL, Schaniel C, Lemischka IR, et al. Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nature Biotechnology. 2011 Feb 27;29(3):267–72.

27. Kopp JL, Dubois CL, Schaffer AE, Hao E, Shih HP, Seymour PA, et al. Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. Development. 2011 Jan 25;138(4):653–65.

28. Del Moral P-M, Warburton D. Explant culture of mouse embryonic whole lung, isolated epithelium, or mesenchyme under chemically defined conditions as a system to evaluate the molecular mechanism of branching morphogenesis and cellular differentiation. Methods Mol Biol. 2010;633:71–9.

29. Hill DR, Huang S, Nagy MS, Yadagiri VK, Fields C, Mukherjee D, et al. Bacterial colonization stimulates a complex physiological response in the immature human intestinal epithelium. Elife. 2017 Nov 7;6.

30. Finkbeiner SR, Hill DR, Altheim CH, Dedhia PH, Taylor MJ, Tsai Y-H, et al. Transcriptome-wide Analysis Reveals Hallmarks of Human Intestine Development and Maturation In Vitro and In Vivo. Stem Cell Reports. 2015 Jun 3.

31. Spence JR, Lange AW, Lin S-CJ, Kaestner KH, Lowy AM, Kim I, et al. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell. 2009 Jul;17(1):62–74.

32. Ghosh M, Ahmad S, White CW, Reynolds SD. Transplantation of Airway Epithelial Stem/Progenitor Cells: A Future for Cell-Based Therapy. Am J Respir Cell Mol Biol. 2017 Jan;56(1):1-10.

Chapter 3 Basal stem cell fate specification is mediated by SMAD signaling in the developing human lung

Portions of this chapter have been published:

Miller, A.J.; Yu, Q.; Czerwinski M.; Tsai, YH; Conway, R.F.; Wu, A.; Holloway E.M.; Walker, T.; Glass, I.A.; Treutlein, B.; Camp, J.G.; Spence, J.R. Basal stem cell fate specification is mediated by SMAD signaling in the developing human lung. *Biorxiv.* **2018**. Doi: 10.1101/461103

3.1 Introduction

Basal stem cells (basal cells), located in the bronchi and trachea of the human lung epithelium, play a critical role in normal airway homeostasis and repair, and have been implicated in the development of diseases such as cancer(1-4). Additionally, basal-like cells contribute to alveolar regeneration and fibrosis following severe injury(5-8). However, the developmental origin of basal cells in humans is unclear. Previous work has shown that specialized progenitor cells exist at the tips of epithelial tubes during lung branching morphogenesis, and in mice, give rise to all alveolar and airway lineages(9,10). These ‘bud tip progenitor cells’ have also been described in the developing human lung(11-13), but the mechanisms controlling bud tip differentiation into specific cell lineages, including basal cells, are unknown. Here, we interrogated the bud tip-to-basal cell transition using human tissue specimens, bud tip progenitor organoid cultures(11), and single-cell transcriptomics. We used single-cell mRNA

sequencing (scRNA-seq) of developing human lung specimens from 11.5-21 weeks gestation to identify molecular signatures and cell states in the developing human airway epithelium, identifying two uncharacterized cell states, termed ‘hub cells’ and ‘bud tip adjacent’ cells, which we localized spatially and temporally using immunohistochemistry and in situ mRNA hybridization. We identified differentially expressed genes between bud tip progenitors and basal cells and used bud tip progenitor organoids to screen for signaling factors that promoted a bud tip to basal cell differentiation. This analysis implicated SMAD signaling as a regulator of the bud tip-to-basal cell transition, and bud tip progenitor organoid experiments showed that TGFβ1 and BMP4 mediated SMAD signaling robustly induced TP63, ultimately leading to differentiation of functional basal-like cells *in vitro*. These basal-like cells exhibited clonal expansion, self-renewal and multilineage differentiation capacity. This work provides a framework for deducing and validating key regulators of cell fate decisions using single cell transcriptomics and human organoid models. Further, the identification of SMAD signaling as a critical regulator of newly born basal cells in the lung has implications for regenerative medicine, basal cell development in other organs, and understanding basal cell misregulation in disease.

Basal stem cells serve as epithelial stem cells in multiple organ systems, including lung, skin, esophagus, breast and prostate, where they contribute to organ homeostasis and repair after injury(1,14-20) and their misregulation has been implicated in diseases, including cancer(4,14,21). In the lung, basal stem cells reside on the basolateral surface of the pseudostratified epithelium and are marked by the transcription factor TP63, which is essential for basal cell fate determination and

function(22,23). In mice, they are restricted to the trachea, whereas in humans, basal stem cells extend deep into the bronchial tree(1,2). Resident airway basal cells routinely give rise to secretory and multiciliated cells of the airway during homeostasis and repair. In addition to basal stem cells lining the airway epithelium, basal-like cells appear in the alveolar regions of the lung after severe injury in mice. These ectopic basal-like cells express KRT5 and TP63 and contribute to injury repair(5-8), and TP63+ basal-like cells have been identified in the distal airways of human patients with pulmonary fibrosis, suggesting these distal basal-like cells may contribute to repair and disease in humans(21). However, the mechanisms governing basal cell fate specification during development or during disease remain unclear. Lineage tracing experiments in mice have shown that bud tip progenitor cells give rise to both proximal and distal lung epithelial cell types during development(9), and mouse bud tip progenitors were shown to give rise to basal cells early during lung development, prior to embryonic day (E) 9.5(10). Similarly, human fetal bud tip progenitor cells have been shown to give rise to putative TP63+ basal stem cells upon transplantation *in vivo*(12); however, our understanding of basal cell specification during human development has been hindered by the inherent challenges of performing mechanistic studies in human tissue and by limited tissue availability. Here, we interrogated the developmental mechanisms controlling the bud tip-to basal cell transition during human lung development. We first used single-cell transcriptomics to identify molecularly distinct cell states, followed by immunohistochemistry and *in situ* mRNA hybridization to locate these cell states within the airway over developmental time, identifying signaling pathway candidates that could potentially induce a basal cell state from a bud tip

progenitor. We then investigated the ability of cell signaling pathways to induce a basal cell transcriptome and functional fate using bud tip progenitor organoids *in vitro*.

3.2 Results

Single-cell transcriptomics defines bud tip progenitor and basal stem cell signatures in the developing human lung

We used single-cell transcriptomics to identify the cell types and epithelial cell states present during human lung development. We performed scRNA-seq on dissociated whole distal lung, small airway, and scraped tracheal epithelial cells (Fig. 3-1a, Fig. 3-2a) from human fetal lung tissue ranging from 11.5 to 21 weeks gestation, and identified diverse immune, endothelial, mesenchymal, and epithelial populations (Fig. 3-2b, c). We computationally extracted 8,443 epithelial cells based on clusters expressing canonical epithelial markers (e.g. *EPCAM*; Fig. 3-2c; see Table 3-1 for numbers of cells per sample), then re-clustered these cells, and visualized the heterogeneity using t-distributed Stochastic Neighbor Embedding (tSNE) (Fig. 3-1b). Based on this analysis, we identified 12 epithelial cell types/states, and cluster identities were assigned based on known markers, where possible (Fig. 3-1c; Fig. 3-3c, Table 3-5). We identified multiple known groups of lung epithelial progenitors including bud tip progenitors (BP - cluster 5), basal cells (BC - cluster 7) and differentiated cell types including multiciliated cells (MC - cluster 2), neuroendocrine cells (NE - cluster 3), and club-like and goblet-like secretory cells (CS, GS - clusters 10 and 12, respectively). We also identified two previously undescribed epithelial cell clusters, which we term ‘bud tip

adjacent' (BA - cluster 6) and 'hub cells' (HC - cluster 4). Cluster 6 bud tip adjacent cells co-express a unique combination of genes including bud tip progenitor markers (low *SFTPC*, *ID2*) as well as canonical AECI markers such as *HOPX*, *PDPN* and *AGER* (Fig. 3-1c-d). Protein staining of the human fetal lung at 16 weeks shows that these cells are physically located adjacent to the SOX9+ bud tips and are PDPN+/AGER+/SOX9- (Fig. 3-1e; n=3 biological replicates). Cells in cluster 4 express high levels of *SCGB3A2*, *SFTPB* and *CFTR*, and are molecularly distinct from other secretory cell populations; for example, they do not express high levels of the club cell marker *SCGB1A1* (Fig. 3-1c). Additional analyses of bud tip adjacent and hub cell populations are provided in Figure 2. Interestingly, we did not observe any clearly identifiable clusters for the rare cell types tuft cells, marked by DCLK1, or ionocytes, marked by FOXI1 (Fig. 3-2d) (24,25). Based on the percentage of these cell types reported in the mouse airway(24) and in human culture systems (25), and based on the number of cells we analyzed here, we would expect to see clusters of these cells in our scRNA-seq. The absence of these cells in our dataset supports the idea that these cells are not yet present in the developing human lung. Although it is possible these cells have a lower abundance in the fetal lung than in adult mouse and human air liquid interface culture systems, and that we did not profile enough cells to identify a tuft or ionocyte cluster.

We next sought to identify unique gene signatures and to validate protein markers for bud tip progenitors and basal stem cells. A comparison of genes expressed in the bud tip progenitor cell cluster relative to expression in all other clusters allowed us to identify marker genes most highly enriched in bud tip progenitors, thereby defining a transcriptional signature of human fetal bud tip progenitors at 11.5-18 weeks gestation

(Fig. 3-1c, Fig. 3-3b, Table 3-6). This analysis shows that cells in cluster 5 exhibit the strongest expression of canonical markers of human bud tip progenitors, including *SOX9*, *SFTPC, ETV5* and *ID2* (Fig. 3-1c-d; Table 3-5)(5,7,11,12,26). Immunofluorescent staining in fetal lung tissue confirms protein expression of bud tip marker genes (Fig. 3-1e; n=3 biological replicates).

To define a basal stem cell signature, we first identified basal cells as cells within cluster 7 based on expression of canonical basal cell markers *TP63* and *KRT5* (Fig. 3-1c, f) (1,2,9,10,12,27). In addition, this cluster highly expresses a set of basal cell-enriched genes including *KRT15, IL33*, *S100A2*, *F3, EGFR*, and *PDPN* (Fig. 3-1c, f; Table 3-5), and protein staining validates that KRT5, IL33, F3, EGFR, KRT15 and PDPN are expressed within TP63+ cells in the trachea at 17 weeks gestation (Fig. 3-1g; Fig. 3-3a; n=3 biological replicates). We noted that some genes that were highly enriched in the basal cell cluster, such as KRT5 and KRT15, were also expressed in clusters 8, 9 and 12. We identified a number of genes in these clusters that corresponded to known markers of club cells (CS - cluster 12) or to cells within the submucosal glands (SMG, SMB - clusters 8 and 9). This interesting pattern of gene expression could represent the differentiation of some basal cells into a secretory lineage or a submucosal lineage, or it could reflect the fact that many submucosal gland cells express multiple basal cell markers, including KRT5 and KRT15, in addition to known submucosal markers SOX9 and ACTA2(28,29). Comparison of the average gene expression levels in basal cells (cluster 7) to the average gene expression in all other clusters allowed us to identify a list of the most highly enriched genes in basal cells, defining a transcriptional signature for this cell population (Fig. 3-1c, Fig. 3-3b;

Table 3-5). We interrogated basal cell enriched genes to identify cell surface markers uniquely co-expressed in basal cells in order to develop fluorescence activated cell sorting (FACS) strategies to purify human basal cells. Using this approach, we identify and validate EGFR and F3 as markers that are co-expressed in basal cells (Fig. 3-1c, g), and that can be used to isolate TP63+ basal cells from primary fetal lung tissue (Fig. 3-3d-e). Together, these data provide a reference atlas of epithelial cell states with defined molecular signatures that arise in the human developing lung epithelium.

Characterizing human basal cells across developmental time and lung region.

Next, we used analysis of scRNA-seq data, protein staining and mRNA fluorescent *in situ* hybridization (FISH) to gain insights into cell states and potential cell fate transitions in the developing human lung epithelium. Recent work using *TP63* lineage tracing in the mouse showed that P63+/KRT5- cells appeared throughout the airway at early stages of lung development and that these cells acted as progenitors that could give rise to both airway and alveolar lung epithelial lineages. In contrast, TP63+/KRT5+ cells in the trachea at later developmental time points only gave rise to airway lineages within the most proximal airway(10). To interrogate TP63 expression across human lung development, we performed protein staining and/or FISH on fetal lungs from 10-20 weeks gestation, and imaged regions along the bud tip-to-trachea (distal-proximal) axis (Fig. 3-4a-c; Fig. 3-5a-f). Tracheal basal cells at 12 weeks gestation stain positively for TP63, KRT5, PDPN, EGFR, and F3 (Fig. 3-4b, Fig. 3-5a, e), similar to tracheal basal cells in older lungs at 17 weeks of gestation (Fig. 3-1g). Interestingly, IL33 marks fewer basal cells at 12 weeks gestation than at 17 weeks (Fig.

3-5c compare to Fig. 3-1g). In the small cartilaginous airways, the majority of TP63+ cells are KRT5-, while KRT5+/TP63+ cells are present in raised patches of cells (Fig. 3-4b). In more distal regions of the airway, TP63 is detected and is co-localized with EGFR, while KRT5 is not detectable, F3 levels decline, and PDPN expression becomes less specific to TP63+ cells (Fig. 3-4b, Fig. 3-5e). Protein localization of TP63 and KRT5 was confirmed by FISH for *TP63* and *KRT5* in 12 week fetal lungs, and mirrors patterns of protein expression along the proximal-distal axis (Fig. 3-5a, b). The percentage of TP63+ cells at 10 weeks is as high as 80% in the middle airways (Fig. 3-4c). These numbers steadily decline as development progresses, and the number of TP63+ cells in the middle airways is reduced to 40% at 20 weeks gestation (Fig. 3-4c). By 17 weeks gestation, the overwhelming majority of TP63+ basal cells throughout the trachea and cartilaginous airways exhibit a mature TP63+/KRT5+ basal cell profile, and TP63+/KRT5- cells throughout the airway become less common, though rare TP63+/KRT5- cells can be found in the distal airways (Fig. 3-4c, Fig. 3-5d, f).

We performed similar analyses to quantify the location and abundance of Hub cells over time and space. Previous work has identified a *Scgb3a2+*/*Upk3a+* airway progenitor in mice(30), but a *SCGB3A2+/CFTR+/SFTPB+* cell state has not previously been described in mice or humans, and it is unclear if this population is related to the Scgb3a2+ proximal progenitor in mice. Hub cells co-express high levels of *SCGB3A2*, *SFTPB* and *CFTR* (Table 3-5; Fig. 3-4e), but lack expression of canonical markers for club cells (e.g. *SCGB3A1*) or ionocytes (e.g. *FOXI1*). Expression patterns and anatomical location of hub cells were evaluated by co-expression of SCGB3A2 and SFTPB protein staining across developmental time and lung region (Fig. 3-4b-c) Fig. 3-

6a). Anatomically, SCGB3A2+/SFTPB+ cells are most abundant in the non-cartilaginous and small to large cartilaginous airways and are physically located at the apical epithelium, close to the airway lumen (Fig. 3-6a). Quantification of hub cells (SCGB3A2+/SFTPB+) confirms that this population is more abundant in the middle airways and reveals that this population appears to decrease over developmental time, but is still detectable at 19-20 weeks gestation (Fig. 3-4c).

Continuous gene expression topologies reflect a physically restricted lung epithelial hierarchy in the fetal lung

We next used a force-directed graph (implemented by SPRING(31)) to visualize continuous gene expression topologies in the fetal lung data, and to identify cell populations that share high degrees of transcriptional similarity (Fig. 3-4d, e). Cells from distal lung samples populate the bud tip progenitor and bud tip adjacent populations, whereas cells from the airways and tracheal scrapings populate several cell clusters, including hub cells, basal cells and differentiated airway cell types (Fig. 3-4d). This analysis reveals that bud tip progenitors and bud tip adjacent cells, and bud tip adjacent and hub cells, share a high degree of transcriptional similarity and sit adjacent to one another on the SPRING plot, mirroring their *in vivo* spatial proximity. To test possible lineage relationships between bud tip progenitor, bud tip adjacent, hub and basal cells, we isolated bud tip adjacent cells from human fetal lungs (n=2; 10.5 and 11.5 week gestation) using FACS for EPCAM, AGER and PDPN (EPCAM+/AGER+/PDPN+) (Fig. 3-4f; Fig. 3-6b, c) and asked whether they preferentially give rise to hub and/or basal cells in media that is optimized to support basal cell expansion and airway organoid

growth (see Fig. 3-13 and 3-14). Isolated bud tip adjacent cells were cultured in a Matrigel droplet for 4 weeks and gave rise to small organoids (Fig. 3-4g). Analysis of these small organoids demonstrates that isolated EPCAM+/AGER+/PDPN+ bud tip adjacent cells primarily give rise to AECII-like cells, which co-stain for AECII markers SFTPC and SFTPB, but do not give rise to SCGB3A2+/SFTPB+ hub cells, and fewer than 10% of all cells express the basal cell marker TP63 (Fig. 3-4h, i; Fig. 3-6d-f). These results are consistent with recent lineage tracing experiments in the mouse, which showed that HOPX+ and SFTPC+ epithelial progenitor cells are committed to an AECI or AECII cell fate, respectively, as early as E15.5, prior to sacculation(32). This early lineage commitment suggests alveolar fate specification occurs early in development and reinforces the idea that fate specification is temporally restricted. Based on this evidence, our data showing bud tip adjacent cells giving rise to AECII-like cells, even in airway supportive media, suggests that bud tip adjacent cells are intrinsically primed to give rise to alveolar cell types. However, it is also possible that culture conditions were not optimal for proximal differentiation or differentiation into AECI cells, and lineage tracing experiments will be needed to resolve bud tip adjacent capabilities during human development. Together, these results show that the transcriptional similarity between cell states mirrors the close physical proximity of adjacent domains in developing human lungs.

SMAD activation induces TP63 expression in bud tip progenitors in vitro

In order to interrogate possible mechanisms by which bud tip progenitors are instructed to differentiate into basal stem cells, we performed a screen to identify

candidates that induce *TP63* in human fetal bud tip progenitor organoids isolated from 12 week fetal lungs (Fig. 3-7a, b). We tested the activity of several signaling pathways known to be important for lung development to induce *TP63* gene expression in bud tip progenitor organoids. As expected, bud tip progenitor organoids grown in the bud tip progenitor maintenance media containing CHIR99021/ATRA/FGF7 (3 factor media; 3F) have low *TP63.* Conversely, DMSO controls or FGF7-only or FGF10-only treated organoids showed some increase in *TP63*, due to stochastic differentiation upon withdrawal of growth factors required for progenitor maintenance and consistent with previous literature that FGF signaling plays a role in basal cell homeostasis(11,33). Many treatment groups are permissive for modest *TP63* expression, which are comparable to the levels observed in DMSO controls; however, TGFβ1 treatment significantly induces *TP63* above the control (Fig. 3-7b). To test whether inhibition of SMAD signaling could block *TP63* expression in a permissive environment, we treated bud tip progenitor organoids for 10 days with progenitor maintenance media ('3F'), FGF7-only to allow *TP63* expression, and FGF7 plus SMAD inhibitors A8301 and NOGGIN (Fig. 3-7c). While treatment with FGF7 results in significantly higher expression of *TP63* relative to 3F media, the addition of SMAD inhibitors blocks *TP63* induction, suggesting that activation of SMAD signaling is important for *TP63* expression in this context. Interestingly, inhibition of SMAD-mediated TGFβ and BMP signaling is important for maintaining basal cells in the adult lung(4,14,21,27,34), whereas our results suggest that SMAD-activation is required to induce *TP63* expression. We next sought to identify genes showing significantly higher expression levels in basal cells than bud tip progenitors (basal cell upregulated genes, Table 3-6),

and examined enrichment of functional gene sets based on differentially expressed genes (Fig. 3-7d; Table 3-2). In total, 642 genes showed significantly higher expression levels in basal cells compared to bud tip progenitors (see methods). Out of the 50 hallmark gene sets annotated in MSigDB(35), 28 were significantly enriched in basal cell upregulated genes (one-sided Fisher's exact test, Bonferroni-corrected $P < 0.05$). Many of the pathways that we tested in Fig. 3-7b, including IL6, IL2 and IFNγ, were identified by the DE analysis, but did not have any inductive effect on *TP63* expression in bud tip progenitor organoids. Consistent with our findings (Fig. 3-7b), TGFβ signaling was also identified in the DE analysis. Based on our results and the previously described role of SMAD signaling in the adult, we hypothesized that transient activation of SMAD signaling induces airway lineages, including basal stem cells, from fetal bud tip progenitors.

To test the hypothesis that transient SMAD activation promotes differentiation of bud tip progenitors towards a proximal airway and/or basal cell lineages in human tissue, we tested if activators or inhibitors of the TGFβ/BMP signaling pathways alone, or in combination, influenced *TP63* expression in bud tip progenitor organoids (Fig. 3-7e, Fig. 3-8a-c). Supplementing normal progenitor organoid growth/maintenance medium ('3F') with SMAD activators or inhibitors for 3 days reveals that TGFβ1 induces *TP63*, as expected from our previous result (Fig. 3-7b, e); however, we found that TGFβ1 plus BMP4 (herein referred to as 'dual SMAD activation'; DSA) cooperatively result in the most significant increase in *TP63* expression by QRT-PCR (Fig. 3-7e). 3 days of DSA also leads to non-significant increases of *KRT5* and *KRT14* but shows no increase in markers for other lung epithelial cell types (Fig. 3-7f). Protein staining further

reveals that 60.13% (+/-13.04%) of DSA treated bud tip organoid cells expressed TP63 after 3 days, compared to 0% in controls (Fig. 3-7g, h). DSA treatment leads to more dense epithelial structures (Fig. 3-7g; Fig. 3-8c, d, f) with significantly reduced proliferation as measured by KI67 staining (Fig. 3-8d, e) and increased apoptosis (Fig. 3-8f, g). DSA significantly increases *TP63* expression both in the presence and absence of CHIR99021 (Fig. 3-8h), a GSK3β inhibitor that is required to maintain bud tip organoids in their undifferentiated state(5-8,11). Consistent with bud tip progenitor organoid experiments, treatment of explanted pieces of whole distal lung tissue (n=2, 10 and 11 weeks gestation) with DSA leads to significant increases in TP63 protein and mRNA expression (Fig. 3-8i-l), suggesting that DSA is a potent inducer of TP63 in multiple contexts.

Previous studies in the adult trachea(1,2,27) have shown that phosphorylated SMAD (pSMAD) is low or non-detectable in basal cells, consistent with the need to inhibit SMAD to maintain mature basal cells in culture(21,27,34). In contrast, protein staining in P63+/KRT5- cells within the non-cartilaginous and small-cartilaginous airways exhibits strong nuclear staining of phospho-SMAD2 (pSMAD2; Fig. 3-7i) and phospho-SMAD1,5,8 (pSMAD1,5,8; Fig. 3-9a; Isotype controls shown in Fig. 3-9d) at 12 weeks gestation, whereas pSMAD2 and pSMAD1,5,8 staining is weak in both bud tip progenitors and nearly undetectable in tracheal TP63+ cells (Fig. 3-7i; Fig. 3-9a). Mature TP63+ basal cells in the 12 week trachea exhibit very weak nuclear pSMAD2 and pSMAD1,5,8 staining, but adjacent luminal tracheal epithelial cells have strong nuclear pSMAD expression (Fig. 3-7i, Fig. 3-9a). TP63+ basal cells in 17 week fetal lungs lack strong nuclear pSMAD staining irrespective of anatomical location (Fig. 3-9b,

c). These *in vivo* observations suggest that pSMADs may be active in newly born or immature TP63+/KRT5- cells, whereas active SMAD signaling is excluded from mature TP63+/KRT5+ basal cells.

In vitro-derived TP63+ cells can be expanded in culture and are transcriptionally similar to fetal basal cells

Our findings show that *TP63* is robustly induced in bud tip progenitor cells exposed to DSA but that prolonged treatment also blocks proliferation and induces cell death (Fig. 3-8d-g), creating technical challenges to isolating and studying this population. Previous work in mice and humans has shown that inhibition of TGFβ/BMP (dual SMAD inhibition; DSI) is required for expansion of mature adult basal stem cells in culture(21,27,34). Based on our data, we reasoned that while DSA is sufficient to induce TP63 expression, it is detrimental for long-term growth and expansion. Therefore, we screened for growth factor conditions that allowed expansion of DSA-induced cell populations in culture (Fig. 3-10a; Fig. 3-11a-e). We found that supplementation of serum-free basal medium with FGF10 and Y27632, a RHO kinase inhibitor, in addition to inhibitors of TGFβ and BMP ('DSI expansion medium': FGF10, A8308, NOGGIN, Y27632), allows organoids to recover following DSA induction, although many organoids do not survive DSA treatment (Fig. 3-11f, g). Further, treatment of bud tip progenitor organoids with DSA followed by DSI versus organoids treated immediately with DSI reveals that pre-treatment with DSA significantly increases *TP63* expression 10 days post-treatment (Fig. 3-11h).

Organoids grown in DSI expansion medium demonstrate increased mRNA expression of markers canonically associated with multiciliated and club cell lineages (*FOXJ1, SCGB1A1,* respectively (Fig. 3-11e) and protein expression for markers of hub cells (SCGB3A2, SFTPB) multiciliated cells (FOXJ1, ACTUB), club cells (SCGB1A1), goblet cells (MUC5AC) and neuroendocrine cells (CHGA, SYN; Fig. 4d). After several days in culture, many beating multiciliated cells are present within organoids and the proteinaceous luminal contents within organoids appeared to swirl with directionality, suggesting that multiciliated cells are functional and able to propel luminal contents. To more closely evaluate when *TP63* is upregulated and whether DSA-induced TP63+ cells may have passed through other cell states (i.e. hub cells), we performed a time course analysis of the DSA-DSI protocol (Fig. 3-10e) and found that *TP63* is upregulated as early as 30 minutes after DSA treatment and remains high throughout the protocol. In contrast, hub cell markers (SFTPB, SCGB3A2) and markers of differentiated ciliated cells (FOXJ1) decrease during DSA treatment and rise again once organoids are moved to DSI. These results indicate that bud tip progenitors can robustly induce *TP63* without inducing bud tip adjacent or hub cell markers.

In order to dissect cell composition in cultured bud tip organoids, we analyzed the transcriptomes of 2,106 cells prior to differentiation experiments (‘day 0’), 9,400 cells immediately after 3 days of DSA treatment (‘day 3’) and 3,755 cells from bud tip progenitor organoids treated for 3 days with DSA followed by 18 days of DSI expansion medium (‘day 21’ organoids) using scRNA-seq. tSNE projections of the data from each time point separately allow visualization of cell heterogeneity and the expression patterns of cell markers for bud tip progenitors (SFTPC and SOX9), Hub cells

(SCGB3A2, SFTPB) and Basal Cells (TP63, KRT5, KRT15), and other cell types (Fig. 3-10f). This analysis shows that day 0 cells are relatively homogeneous and robustly express bud tip progenitor markers SFTPC and SOX9, while the expression of SFTPC is greatly reduced by day 3, but increases again in specific populations by day 21 (Fig. 3-10f, Fig. 3-12a). The Hub cell marker SCGB3A2 is low at day 0 and day 3, but expressed highly in many cells by day 21, further supporting data that cells do not pass through a hub cell state prior to giving rise to TP63+ cells. *TP63* is absent from day 0 cells, but is detected in 26% of cells by day 3, and is detected in a more restricted subset cells (13%) by day 21 (Fig. 3-10f). For a general characterization of the acquisition of basal cell transcriptome features throughout treatment, we summarized the expression patterns of 331 fetal basal cell cluster marker genes ('basal cell signature') detected in *in vitro* cells. Compared to day 0, cells of day 3 and day 21 show general upregulation of basal cell marker genes. Meanwhile, compared to day 3, day 21 organoids show even further upregulation of the basal cell signature in specific cells (Fig. 3-10f, bottom row).

To evaluate the transcriptomic similarity between *in vitro* derived cells and *in vivo* fetal lung epithelial cells, we quantified the transcriptome similarity of each cell to the *in vivo* epithelial cell cluster gene expression signatures as identified in Fig. 1c using Pearson's correlation coefficients (Fig. 3-10g). *In vitro* maintained bud tip progenitor cells (day 0) share the highest degree of gene expression similarity with *in vivo* bud tip progenitor cells (Fig. 3-10g). The vast majority of cells from organoids treated with 3 days of DSA (day 3) are most similar to bud tip progenitors, while some of them show marginally higher similarity to bud tip adjacent, sub mucosal basal cells, or basal cells

(Fig. 3-10g). To identify possible biological processes accompanying induction of *TP63* under DSA treatment, we performed clustering for day 3 cells using genes showing highly variable expression levels in this sample and characterized the expression patterns and functional enrichment of markers of different clusters. We found that genes related to the extracellular region tend to be co-expressed with *TP63*, while immune related genes tend to be upregulated in other subsets of cells (Fig. 3-12b-d), suggesting cells at day 3 are diversified and are associated with *TP63* induction in a complex way. Consistent with the protein staining, organoids at day 21 show divergent transcriptomes, with specific subsets of cells most similar to multiciliated cells, neuroendocrine cells, hub cells, bud tip adjacent cells, or basal cells (Fig. 3-10g). Expression of individual marker genes for each population show consistent patterns (Fig. 3-12a). Interestingly, while certain cells express markers of mature goblet (MUC5B) and club (SCGB1A1) cells (Fig. 3-12a), which were also observed via protein staining (Fig. 3-10d), the transcriptomic profile of these cells do not fully match to *in vivo* signatures of human fetal goblet and club cells (Fig. 3-10g). Instead, these secretory cells match closest to hub cells, suggesting they remain immature at this time point. *In vitro*-derived *TP63*+ cells also express *KRT15*, but low levels of *KRT5*, suggesting these cells have phenotypes more similar to the TP63+/KRT5- cells that are abundant in the 10-15 week smaller airways (Fig. 3-12a, Fig. 3-4c).

In vitro-derived TP63+ cells can be isolated and exhibit basal cell properties of clonal expansion, self-renewal and multilineage differentiation capacity

We asked whether *in vitro* derived TP63+ cells exhibited functional hallmarks of basal stem cells, including clonal expansion, self-renewal, and the ability to undergo multilineage differentiation. In order to determine if DSA-induced TP63+ cells have the capability for self-renewal and multi-lineage differentiation, we isolated TP63+ cells with FACS using the basal cell enriched cell surface proteins EGFR and F3, which are expressed in TP63+ cells in basal cells of the fetal lung, and in TP63+ cells from DSA-DSI expansion treated organoids (Fig. 3-1c, Fig. 3-13a-c, Fig. 3-14a). Bud tip organoids were treated with DSA for 3 days to induce *TP63* expression and expanded for 18 days in DSI expansion medium (Fig. 3-10a). Organoids were subsequently dissociated and FACS was used to isolate EGFR+/F3+ cells (Fig. 3-13c, Fig. 3-14a-c). The percentage of EGFR+/F3+ double positive cells isolated for each biological replicate varied from 5.84% to 25.9%, which may reflect biological differences in replicates. FACS isolated cells from all 3 biological replicates were immediately affixed to glass slides via cytospin and were stained for TP63 protein expression. This analysis shows that 92.09 (+/-1.66)% of EGFR+/F3+ cells co-express nuclear TP63, whereas EGFR+/F3-, EGFR-/F3+ or EGFR-/F3- fractions have a far lower proportion of cells that express TP63 (Fig. 3-10d, e; n=3 biological replicates). After sorting, EGFR+/F3+ cells were re-plated and grown *in vitro* in DSI expansion medium. Organoids derived from a single cell suspension of EGFR+/F3+ cells contain TP63+ basal-like, MUC5AC+ goblet-like, SCGB1A1+ club-like and AcTUB+/FOXJ1+ multiciliated cells as shown by protein staining (Fig. 3-13f, g, Fig. 3-14g). These cell types are not present in untreated controls (Fig. 3-14e). Interestingly, no neuroendocrine cells exist in EGFR+/F3+ clonally expanded basal cell-derived organoids, although they are present in unsorted treated

organoids (Fig. 3-13f). Further, clonally expanded organoids exhibit functional multiciliated cells that beat, along with swirling luminal contents.

To test whether isolated EGFR+/F3+ cells exhibited clonal expansion, bud tip progenitor organoids were treated with 3 days of DSA followed by expansion with DSI. After 2 weeks of expansion in DSI, separate batches of organoids were infected with a lentivirus driving expression of GFP or with a lentivirus driving expression of mCherry, in order to track and visualize these cells over time. Organoids were allowed to expand for an additional 2 weeks in DSI, and were then sorted for EGFR and F3 (Fig. 3-14i). Sorted cells were plated either as GFP-EGFR/F3 cells alone, mCherry-EGFR/F3 cells alone, or as a mixture of GFP-EGFR/F3 and mCherry-EGFR/F3 cells to evaluate whether clonal populations formed. Of the resulting organoids, whole organoids are composed either entirely of GFP+, RFP+, or GFP/RFP- cells, but no organoids contain a mixture of GFP+ and RFP+ cells, nor do any organoids contain a mixture of fluorescence+ and fluorescence- cells, suggesting that individual organoids are derived from clonal expanded cells, rather than from cell aggregation (Fig. 3-13h). Together, this data supports the hypothesis that DSA induces a population of functional basal cells from bud tip progenitor organoids *in vitro*.

3.3 Discussion

Taken together, this study provides a descriptive atlas of the transcriptional signatures present in different cell populations within the epithelium during human lung development, which can serve as a reference for organoid and cell engineering. Functionally, we further identify a role for BMP/TGFβ in the transition from an undifferentiated bud tip progenitor to a proximal airway cell fate. Although future studies

will be needed to fully define the functional roles of hub and bud tip adjacent cell populations, our results here demonstrate that BMP/TGFβ potently induces *TP63* expression in bud tip progenitors. These SMAD-induced TP63+ cells likely represent an early TP63+/KRT5- cell type, similar to that described at E9.5 in the primordial mouse lung bud(10), but follow-up and conformational studies in human tissue at these early stages will be challenging. By defining lung epithelial cell profiles throughout human development, this work serves as an important benchmark for studies aimed at differentiating lung epithelial cell types from human pluripotent stem cells (hPSCs) for drug screening, personalized medicine or regenerative medicine purposes. Similarly, the identification of SMAD signaling as a key regulator of basal cell induction may be relevant to the regulation of airway basal stem cells during injury repair across multiple organ systems, or to the appearance of basal-like populations after acute injury or in illnesses such as idiopathic pulmonary fibrosis (IPF) or basal cell-derived cancers.

3.4 Methods

Human Lung Tissue

Human tissue research was reviewed and approved by The University of Michigan Institutional Review Board (IRB). Human lung tissue was obtained from the University of Washington Laboratory of Developmental Biology. Tissue was shipped overnight in UW-Belzer's solution on ice.

Paraffin processing, tissue preparation, protein staining and imaging

For all protein staining experiments, analysis was carried out on n=3 independent biological specimens, and representative images are shown in the figures. For protein analysis, tissue was immediately fixed in 4% Paraformaldehyde for 24 hours at 4ºC. Tissue was washed in 3 washes of 1X PBS for a total of 2 hours, and then dehydrated by an alcohol series of each concentration diluted in 1x PBS, 30 minutes in each solution: 25% Methanol, 50% Methanol, 75% Methanol, 100% Methanol, 100% Ethanol, 70% Ethanol. Tissue was processed into paraffin blocks in an automated tissue processor (Leica ASP300) with 1 hour solution changes overnight. 7 µm-thick sections were cut from paraffin blocks and immunohistochemical staining was performed as previously described(10,36). A list of antibodies and concentrations can be found in Table 3-3. All images were taken on a NIKON A1 confocal and assembled using Photoshop Creative Suite 6. Imaging parameters were kept consistent for all images of the same experiment and any post-imaging manipulations were performed equally on all images from a single experiment.

In situ hybridization

For all *in situ* hybridization staining experiments, analysis was carried out on n=3 independent biological specimens, and representative images are shown in the figures. Human fetal lung tissue was fixed for 24 hours at room temperature in 10% Neutral Buffered Formalin (NBF), washed with DNAse/RNAse free water (Gibco) for 3 changes for a total of 2 hours. Tissue was dehydrated by an alcohol series diluted in DNAse/RNAse free sterile water for 30 minutes in each solution: 25% Methanol, 50% Methanol, 75% Methanol, 100% Methanol. Tissue was stored long-term in 100%

Methanol at 4°C. Prior to paraffin embedding, tissue was equilibrated in 100% Ethanol, and then 70% Ethanol. Tissue was processed into paraffin blocks in an automated tissue processor (Leica ASP300) with 1 hour changes overnight. Paraffin blocks were sectioned to generate 7 µm-thick sections. All materials, including the microtome and blade, were sprayed with RNAse-away solution prior to use. Slides were sectioned freshly the night before the *in situ* hybridization procedure, baked for 1 hour in a 60ºC dry oven, and stored overnight at room temperature in a slide box with a silicone desiccator packet, and with seams sealed using parafilm. The *in situ* hybridization protocol was performed according to the manufacturer's instructions (ACDbio; RNAscope multiplex fluorescent manual protocol). The human *TP63* probe was generated by ACDbio targeting 4309-1404 of TP63 (accession NM_001114982.1) and is commercially available (acdbio.com, catalog number 601891-C2). The human KRT5 probe was generated by ACDbio targeting 78-2053 of KRT5 (accession NM_000424.3) and is commercially available (acdbio.com, catalog number 310241).

Bud Tip Progenitor Organoids

All experiments were carried out using tissue from 3 different biological specimens at approximately 12 weeks gestation. The peripheral portion of the lungs were enzymatically and mechanically disrupted to isolate bud tip epithelial cells, which were subsequently cultured in 3-dimensional Matrigel droplets with media conditions optimized to expand and maintain bud tip progenitor organoids, as previously described(11,23). Briefly, 1 cm^2 segments of distal lung tissue were cut from the lung with a scalpel, tissue was dissociated using lung dispase (Corning) on ice for 30

minutes, followed by incubation in 100% Fetal Bovine Serum (ThermoFisher Scientific cat. no. 16000044; FBS) for 15 minutes. Tissue was then vigorously pipetted up and down with a p200 to separate the epithelial bud tips from the mesenchyme. Tissue was washed multiple times in sterile 1x PBS to remove mesenchymal cells and epithelium-enriched bud tips were plated in a Matrigel droplet. Buds tips could be frozen down immediately after isolation or at any time after culture in 10% DMSO, 10% FBS and 80% DMEM F12.

Culture Media, Growth Factors and Small Molecules

All experiments utilized serum-free basal medium that has been previously described[11,33]. Briefly, serum-free basal medium consists of DMEM F12 (ThermoFisher Scientific cat. no. 21331020 or 21331-020) supplemented with 1X N2 supplement (ThermoFisher Scientific cat. no. 17502048), 1X B27 supplement (ThermoFisher Scientific cat. no. 17504044), 1X L Glutamine (200 mM), 1X Penicilin-Streptomycin (5000 U/mL, ThermoFisher Scientific cat. no. 15140122) and 0.05% Bovine Serum Albumin (Sigma-Aldrich cat. no. A9647). On the day of use, medium is supplemented with 0.4 µM Monothio-glycerol (Sigma-Aldrich, cat. no. M6145) and 50 µg/mL Ascorbic Acid (L-Ascorbic Acid, Sigma-Aldrich cat. no. A4544, CAS Number 50-81-7). To maintain bud tip progenitor organoids in a progenitor state, serum-free basal medium was further supplemented with FGF7 (10 ng/mL, Recombinant Human Fibroblast Growth Factor 7; R&D Systems cat. no. 251-KG/CF), CHIR99021 (3 µM, Stem Cell Technologies cat. no. 72054), and All Trans Retinoic Acid (ATRA; 50 nM, Stemgent cat. no. 04-0021, CAS Number 302-79-4). Basal cell expansion medium used the same

basal medium, but was supplemented with FGF10, A8301, NOGGIN and Y27632. Growth factors and small molecules were used at the following concentrations: FGF10 (500 ng/mL, made in-house as previously described(11)), A8301 (1 μM, Stem Cell Technologies cat. no. 72024), NOGGIN (100 ng/mL, R&D Systems, cat. no. 6057), Y27632 (APExBIO cat. no. A30008), LDN212854 (200 nM, R&D Systems cat. no. 6151/10), SB431542 (10 μM, Stemgent cat. no. 04-0010), TGFβ1(100 ng/mL, R&D systems cat. no. 240-B-002), BMP4 (100 ng/mL, R&D systems cat. no. 314-BP-050) Dexamethasone (25 ng/mL, Stem Cell Technologies cat. no. 72092), IL6 (10 ng/mL, R&D Systems, cat. no. 206-IL-010), IL2 (50 U/mL, R&D Systems, cat. no. 202-IL-010), Smoothened Agonist (SAG; 500 nM, R&D Systems, cat. no. 4366/1), EGF (100 ng/mL, R&D Systems cat. no. 236-EG-200), IFNγ (10 ng/mL, R&D Systems cat. no. CAA31639), DAPT (10 μM, R&D Systems cat. no. 2634/10), Hydrocortisone (100 ng/mL, Stem Cell Technologies cat. no. 74142).

RNA extraction and qRT-PCR analysis

At least 1 well, containing 20-50 organoids, for each biological replicate was collected and RNA was extracted for QRT-PCR analysis. More than 1 well of organoids was collected per biological replicate and served as a technical replicate when available. mRNA for QRT-PCR was isolated using the MagMAX-96 Total RNA Isolation Kit (Life Technologies). RNA quality and concentration was determined on a Nanodrop 2000 spectrophotometer (Thermo Scientific). 100 ng of RNA for each sample was used to generate a cDNA library using the VILO cDNA kit (Invitrogen). QRT-PCR was performed on a Step One Plus Real-Time PCR System (Life technologies) using SYBR

Green Master Mix (Qiagen). Expression was calculated as a change relative to GAPDH expression using arbitrary units, calculated using the following equation: [2^(GAPDH Ct – Gene Ct)] X 10000. Some data were plotted as fold change of arbitrary expression value over a control. For this analysis, expression values for each gene for each sample, including controls, were divided by the average expression of that gene for the control group. Fold change was calculated as follows: [ExpressionGene / AverageExpressionControls] A list of QRT-PCR primers used can be found in Table 3-4.

Fluorescence Activated Cell Sorting (FACS)

To dissociate organoids into single cell suspension, organoids were first removed from the Matrigel droplet by vigorous p200 pipetting in a small petri dish filled with basal medium. Whole organoids or epithelial fragments were then transferred to a 15 mL conical tube containing 8 mL of TrypLE Express (ThermoFisher Scientific cat. no. 12605036) and placed in a tissue culture incubator at 37ºC, 5% CO_2 on a rocker for no longer than 30 minutes. Every 5 minutes, the tube was removed from the incubator and cells were agitated by pipetting the solution up and down with a P1000 pipette. Cells were confirmed to be in single cell suspension by visualization under an inverted microscope. After cells were in single cell suspension, they were spun down at 300g for 5 minutes at 4ºC, resuspended in 1% BSA in HBSS, filtered through a 70 μm filter to remove any cell clusters (similar to Fisher Scientific cat. no. 087712) and spun down again at 300g for 5 minutes at 4ºC. All tubes used to handle cell suspensions were pre-washed with 1% BSA in HBSS to prevent adhesion of cells to the plastic walls of tubes.

After the second wash, cells were resuspended in sorting buffer and evenly distributed in to several tubes. Cells were incubated with isotype antibody controls that were used to set FACS gates, or incubated with Anti-EGFR-APC, human (Milteny cat. no. 130-110-587, 1:50 dilution), Anti-EGFR-PE, human (Milteny cat. no. 130-110-528, 1:50 dilution), Anti CD142 (F3)-PE, human (Milteny cat. no. 130-098-743, 1:11 dilution), Anti CD142 (F3)-APC, human (Milteny cat. no. 130-115-685, 1:11 dilution), REA control IgG1-APC (Milteny 130-113-434) or mouse IgG1-PE (Milteny cat. no. 130-113-762) and sorted using a Sony Synergy SY3200 system. Data was analyzed using Winlist 8.0 and FlowJo version 10.5.3 for Mac. Cells were sorted into 1mL of 1%BSA in HBSS.

To dissociate primary lung tissue for FACS isolation of bud tip adjacent cells, 1cm2 regions of lung were mechanically removed from the lung and cut in to small pieces using a scalpel until nearly homogeneous. This tissue was then transferred to a 15mL conical tube containing 15 mL TrypLE Express, and the protocol continued as described above for organoid dissociation.

Cytospin analysis

20% of cells isolated from each group for FACS were used to evaluate the percentage of sorted cells expressing TP63. 200 µL of cell suspension (20% of 1 mL) was isolated in a separate 1.5 mL microcentrifuge tube and FBS was added to each aliquot to a final concentration of 5% vol/vol. Cells were placed in clean cytospin cones and spun at 600g for 5 minutes on a Shandon Scientific Cytospin on to charged glass slides. Slides were allowed to air dry for 5 minutes before being fixed in 100% ice cold Methanol for 10 minutes. Slides were air dried for another 10 minutes and were then washed with 2 changes of PBS for a total of 10 minutes on a rocker. The regular

staining protocol for immunofluorescence was followed (blocking, primary antibody incubation overnight at 4ºC, wash, secondary antibody incubation for 1 hour at room temperature, wash, coverslip, image).

Infection of organoids with GFP-lentivirus or mCherry-lentivirus

Lentiviral particles were generated from a construct expressing GFP under the control of a PGK promoter with puromycin selection (Addgene plasmid #19070), or by a construct expressing mCherry under the control of a PGK promoter (Addgene plasmid # 21217) by the University of Michigan Viral Vector Core. Under a dissecting microscope in a sterile hood, organoids were removed from Matrigel droplets by vigorous pipetting with a p200 pipette. Organoids were transferred to a 1.5 mL Eppendorf snap-cap tube with ~250 µL of regular culture medium. Cells were then passaged through a 27-gauge needle attached to a 1 mL syringe 2 times in order to shear the epithelial organoids into small fragments. Cells were spun down for 5 seconds at full speed using a mini centrifuge (similar to ThermoScientific mySPIN 6, cat. no. 75004061), and the remaining floating Matrigel and culture medium were removed from the tube using the needle and syringe under a dissecting microscope. 1 mL of regular culture medium was added to the Eppendorf tube containing the organoid cell fragments, and the cells were transferred to 1 well of a 12-well tissue culture plate. 10 µM of Y27630 was added to improve cell survival and 0.5 mL of high titer virus was added to the well with cell fragments. The plate was placed in a tissue culture incubator (37ºC, 0.5% CO_2) on a rocker for 6 hours. After 6 hours, the cells suspension was moved to a 15 mL conical tube, spun down at 300g for 5 minutes at 4ºC, the supernatant was removed and

treated with bleach solution, and the cells were washed and spun down (300g, 5 minutes, 4ºC) 3X with DMEM. On the final wash, cells were resuspended in 100% Matrigel and plated as a 3-dimensional droplet, allowed to solidify, and then overlaid with bud tip progenitor maintenance medium.

Quantification and Statistical Analysis

All statistical analysis (quantification of immunofluorescent images and QRT-PCR data) was performed in GraphPad Prism 6 software. For quantification of protein stains in organoids, at least 3 independent organoids were counted (technical replicates) from n=3 separate biological specimens (biological replicates). All quantification of protein staining was done in a blinded fashion by an independent researcher. Statistical comparisons of data between two groups (e.g. control versus experimental condition) were made using unpaired two-sided Mann-Whitney rank-sum tests. A p-value of less than 0.05 was considered significant. For QRT-PCR analysis, n=3 biological replicates were used. For each biological replicate, 1-3 well of organoids containing 20-50 organoids per well (technical replicates) was collected for analysis. To determine significance differences across multiple groups, a one-way Analysis of Variance (ANOVA) was performed followed by Tukey's multiple comparisons analysis comparing the mean of each group to the mean of every other group. A p-value of less than 0.05 was considered significant. On graphs, p-values for multiple comparisons after ANOVAs are reported as follows: * $p<0.05$; ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

Preparation of tissue for single cell RNA sequencing

Human Fetal Tissue

To dissociate human fetal tissue to single cells, tissue was first dissected into regions (trachea/bronchi, small airways, distal lung) using forceps and a scalpel in a petri dish filled with ice-cold 1X HBSS (with Mg^{2+}, Ca^{2+}). For harvesting cells from the trachea and bronchi, the airways were transferred to a fresh petri dish filled with ice-cold 1X HBSS and were opened longitudinally with spring-loaded microscissors. The epithelium was scraped with a scalpel and a p200 pipette tip, prewashed with 1% BSA in HBSS to reduce cells from sticking, was used to collect epithelial cells and place them in a 15 mL conical tube. For distal lung tissue, a section roughly $1cm^2$ was cut with a scalpel from the most distal lung regions and minced using a scalpel and forceps. This tissue was then transferred to a 15 mL conical tube.

Dissociation enzymes and reagents from the Neural Tissue Dissociation Kit (Miltenyi, cat. no. 130-092-628) were used, and all incubation steps were carried out in a refrigerated centrifuge pre-chilled to 10ºC unless otherwise stated. All tubes and pipette tips used to handle cell suspensions were pre-washed with 1% BSA in HBSS to prevent adhesion of cells to the plastic. Tissue was treated for 15 minutes at 10ºC with Mix 1 and then incubated for 10 minute increments at 10ºC with Mix 2 interrupted by agitation by pipetting with a P200 pipette until fully dissociated. Cells were filtered through a 70 µm filter coated with 1% BSA in 1X HBSS, spun down at 500g for 5 minutes at 10ºC and resuspended in 500µl 1X HBSS (with Mg^{2+}, Ca^{2+}). 1 mL Red Blood Cell Lysis buffer was then added to the tube and the cell mixture was placed on a rocker for 15 minutes in the cold room (4ºC). Cells were spun down (500g for 5 minutes at 10ºC), and washed twice by suspension in 2 mLs of HBSS + 1% BSA followed by

centrifugation. Cells were then resuspended in 1% BSA in HBSS with 0.5 units/μL of RNAseI (ThermoFisher cat. no. AM2294) in order to reduce RNA present in the media/buffer. Cells were counted using a hemocytometer (ThermoFisher), then spun down and resuspended (if necessary) to reach a concentration of 700-1000 cells/μL and kept on ice. Single cell libraries were immediately prepared on the 10x Chromium at the University of Michigan Sequencing Core facility with a target of 10,000 cells. A full, detailed protocol of tissue dissociation for single cell RNA sequencing can be found at www.jasonspencelab.com/protocols .

Organoids

To dissociate organoids to single cell suspensions, organoids were removed from the Matrigel droplet by vigorous pipetting with a p200 in a small petri dish, then transferred to a 15 mL conical tube containing 8 mL of TrypLE express (ThermoFisher Scientific cat. no. 12605036) and placed in a tissue culture incubator at 37ºC, 5% CO_2 on a rocker for no longer than 30 minutes. Every 5 minutes, the tube was removed from the incubator and cells were agitated by pipetting the solution up and down with a P1000 pipette. Cells were confirmed to be at single cell suspension by microscope. After most cells were in a single cell suspension, cells were spun down at 300g for 5 minutes at 4ºC, resuspended in 1% BSA in HBSS, filtered through a 70 μm filter (similar to Fisher Scientific cat. no. 087712) and spun down again at 300g for 5 minutes at 4ºC. All tubes used to handle cell suspensions were pre-washed with 1% BSA in HBSS to prevent adhesion of cells to the plastic walls of tubes. Cells were then resuspended in 100 μL of 1% BSA in HBSS with 0.5 units/μL of RNAseI (ThermoFisher cat. no.

AM2294), counted using a hemocytometer (ThermoFisher), then spun down and resuspended (if necessary) to reach a concentration of 700-1000 cells/μL and kept on ice. Single cell libraries were immediately prepared on the 10x Chromium at the University of Michigan Sequencing Core facility with a target of 10,000 cells.

Single Cell RNA Sequencing Computational Analysis

Data preprocessing and cluster identification

All single-cell RNA-sequencing was performed with an Illumina Hiseq 4000 by the University of Michigan DNA Sequencing core. Reads alignment to human reference genome and generation of gene expression matrix were done using 10x Genomics Cell Ranger v2.1.1-2.2.1 software with provided hg19 reference genome and transcriptome annotation. In each sample, to ensure high data quality for further analysis, cells with more than 20,000 or less than 1,500 genes or mitochondrial transcript fraction over 10% and genes expressed in less than 3 cells were excluded for further analysis. All fetal tissue samples were then combined. Gene expression levels were log-normalized by the total number of unique molecular identifier (UMI) per cell. Cell cycle phase was scored by expression levels of cell cycle related genes. Cellular variance of total number of UMI, mitochondrial transcript fraction and cell cycle phase was regressed out using linear regression. Non-cell cycle related genes showing highly variable expression levels (highly variable genes) were identified in each fetal tissue sample separately. For the combined fetal data, principle component analysis (PCA) was based on z-transformed expression levels of genes that were identified as highly variable genes in

at least two samples. Graph-based clustering approach and tSNE dimension reduction were performed based on Euclidean distance between the top 20 PCs. Clusters were classified into epithelial, mesenchymal, endothelial, neuronal, immune cell lineages and red blood cell cluster based on expression patterns of canonical marker genes. To better resolve cellular heterogeneity of epithelial cells, cells with low expression levels of EPCAM in the epithelial lineage were further removed, sub-clustering on the remaining epithelial cells was performed. Identification of positive markers for each cluster was performed with Wilcoxon Rank Sum test. Hierarchical clustering of clusters based on average expression levels of cluster markers was performed. Combined with canonical cell type marker expression patterns, clusters with shared cell type identity were merged. Analyses mentioned in this section were mainly performed with Seurat(37).

kNN network construction and visualization

Correlation distance between cells was calculated using z-transformed expression levels of genes identified as top 50 marker ranked by log-transformed fold change of expression levels in any of the epithelial sub-clusters. KNN-network (k=50) was constructed. SRPING(31) was used for network visualization.

Differential expression and functional gene set enrichment analysis

We performed differential gene expression analysis between bud tip progenitor and basal cells using Wilcoxon Rank Sum test. Basal cell upregulated genes were defined as gene with Bonferroni-corrected $P < 0.05$, log-transformed fold change of

cluster average expression levels in basal cells compared to bud tip progenitors > 0.1 and expressed in over 25% of cells in basal cell cluster.

To infer the biological process associated with transition from bud tip progenitor to basal cell, we examined the enrichment of hall mark gene sets from MSigDB (35,38,39) (v6.1) in basal cell upregulated genes with one-sided Fisher's exact test. Significant enrichment was defined as Bonferroni-corrected $P < 0.05$.

Characterization of cellular transcriptome heterogeneity in *in vitro* samples

Each *in vitro* sample was analyzed separately. Data preprocessing of *in vitro* data followed the same procedure as fetal tissue data except that minimum number of detected genes in cells of day 0 is 1,000 and cellular variance of cell cycle phase was not regressed out. To quantify the transcriptome similarity between *in vitro* cells and fetal epithelial sub-clusters, Pearson's correlation coefficients of log-normalized gene expression levels between *in vitro* cells and fetal epithelial cell sub-clusters were calculated using the genes for fetal tissue kNN network construction. Clustering and positive marker gene expression identification followed the same procedure as for fetal data. Gene ontology enrichment in cluster markers was performed using DAVID (40,41) (v6.8). Significant enrichment was defined as Benjamini-corrected $P < 0.05$.

Detailed methods, including code used to process raw data, can be found at https://github.com/qianhuiyu/miller_lung.

Accession numbers

Raw scRNA-seq data associated with this study will be deposited in the EMBL-EBI ArrayExpress database (Accession number: **pending**). For data access ahead of final publication, contact the corresponding authors.

Acknowledgements

This work has been supported by the NIH-NHLBI R01HL119215 and by the Cystic Fibrosis Foundation Therapeutics Epithelial Stem Cell Consortium funding to JRS; AJM was supported by the Tissue Engineering and Regeneration Training Grant NIH-NIDCR T32DE007057, Cell and Molecular Biology Training Grant NIH-T-32-GM007315, and the Ruth L. Kirschstein Predoctoral Individual National Research Service Award NIH-NHLBI F31HL142197; EMH was supported by the Training Program in Basic and Translational Digestive Sciences NIH-NIDDK T32DK094775 and Cellular and Biotechnology Training Program NIH-NIGMS T32GM008353; MC was supported by the Training Program in Organogenesis Fellowship NIH-NICHD T32HD007505. The Laboratory of Developmental Biology, University of Washington, Seattle, WA, United States is supported by NIH-NICHD R24HD000836 to Ian Glass. We would like to thank Judy Opp and the University of Michigan Biomedical Research Core Facility DNA Sequencing Core. We apologize to those whose work we were unable to cite due to space limitations.

Author contributions

AJM and JRS conceived the study. BT, GC and JRS supervised the research. AJM designed, performed and interpreted studies to characterize the native human fetal

lung and *in vitro* studies utilizing fetal-derived bud tip progenitor organoids. QY, BT and GC designed and performed computational analysis on scRNA-seq datasets. QY, BT, GC, AJM and JRS interpreted computational results. AW, YHT, AJM and MC collected, dissociated and submitted tissue for scRNA-seq. AW and EH optimized tissue dissociation protocols for scRNA-seq. MC maintained the scRNA-seq database and performed quality control checks on all scRNA-seq data. RC designed and performed experiments utilizing whole fetal lung explants. IG facilitated transfer of human tissue samples. AJM and TW analyzed and quantified cellular data from *in vitro* studies. YHT performed FACS experiments. AJM and JRS wrote the manuscript. QY, MC, BT, and GC provided critical feedback on the manuscript. All authors read and approved the manuscript.

Competing interests

JRS and AJM are co-inventors on patents filed by the Regents Of The University of Michigan relating to the isolation and maintenance of lung bud tip progenitor cells, and the differentiation of lung epithelial progenitor cells into basal cells.

3.5 Figures

Figure 3-1 Defining cell signatures of human fetal lung epithelial cells

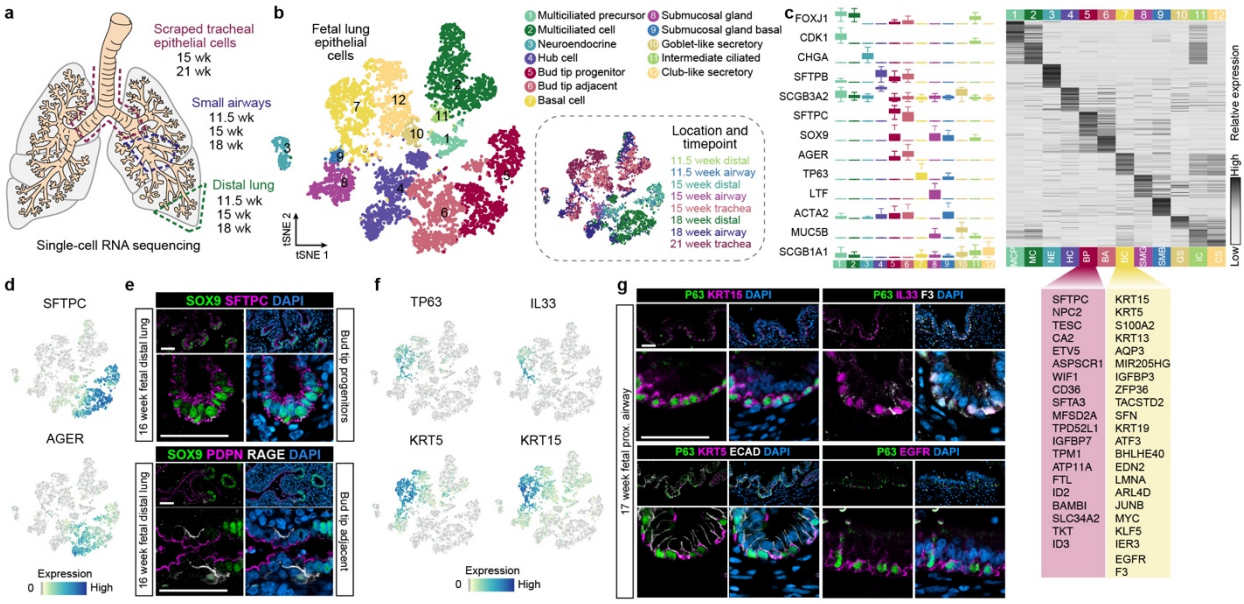


Figure 3-1. Defining cell signatures of human fetal lung epithelial cells. a) Schematic of the experimental setup and samples included in the analysis. **b**) A total of 8,443 EPCAM+ cells were computationally isolated and cellular transcriptome heterogeneity was visualized using a tSNE plot revealing 12 clusters of cells. **c**) Left panel: boxplots (interquartile range with minimum and maximum, outliers removed from plot) show expression level distributions of canonical cell type markers in each cluster. Right panel: heatmap shows z-transformed cluster average expression levels of top 50 cluster markers ranked by log-transformed fold change in expression levels in each cluster compared to other clusters in any of the clusters. Clusters were identified by expression of markers canonically associated with cell populations based on published literature, or given a new name (e.g. hub, bud tip adjacent). A list of the upregulated genes in the bud tip progenitor cluster (BP) and the basal cell cluster (BC) with top fold change in expression levels relative to all other clusters is shown and provides a gene signature for these cell populations. **d**) Feature plots show that cluster 5 contains high levels of expression of canonical bud tip progenitor marker *SFTPC* and cluster 6

contains cells that express low levels of bud tip progenitor marker *SFTPC* and also express *AGER,* a canonical marker of alveolar epithelial type 1 cells. **e**) Protein staining of a 16 week fetal lung specimen shows that bud tip progenitors stain positive for SOX9 (green) and SFTPC (pink). Scale bars represent 50 μm. Protein staining of the cells immediately adjacent to the bud tip progenitors reveals a population that is SOX9 negative (green) but which express PDPN (pink) and AGER (white). Scale bars represent 50 μm. **f**) Feature plots showing cells in cluster 7 express canonical basal cell markers *TP63* and *KRT5*, as well as other makers such as *KRT15* and *IL33* which are highly specific to cluster 7. **g)** Protein staining confirms that tracheal TP63+ basal cells in 17 week fetal lungs express basal cell markers KRT15 (pink), EGFR (pink), IL33 (pink), F3 (white) and PDPN (pink), markers that were strongly expressed in cluster 7 by scRNA-seq.

Figure 3-2 Fetal lung characterization by scRNA-seq

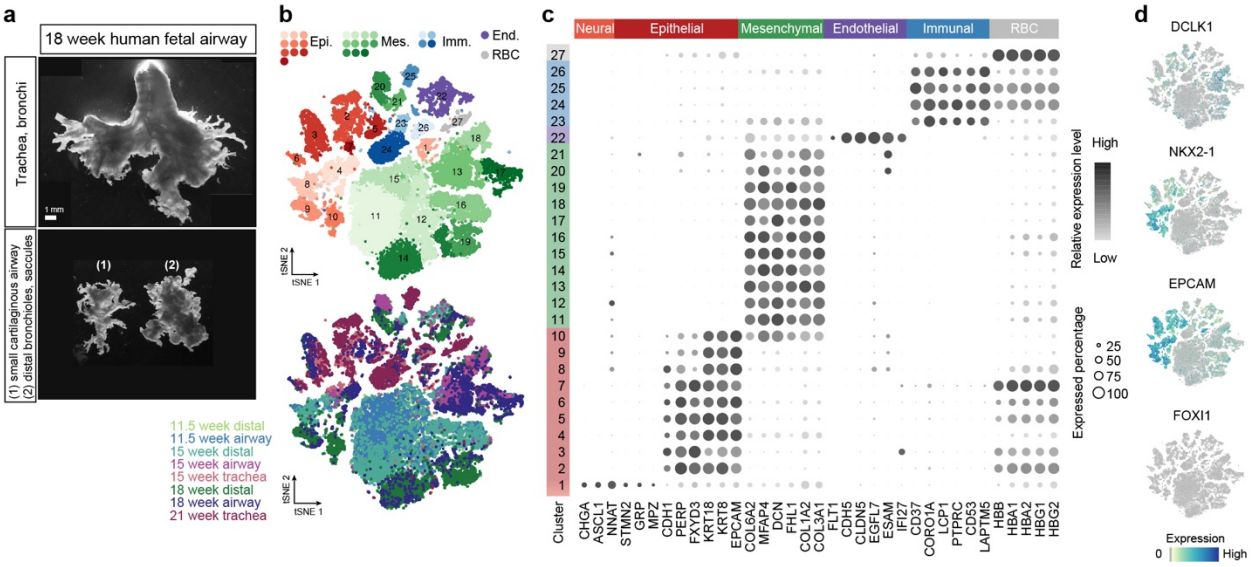


Figure 3-2. Fetal lung characterization by scRNA-seq. **a**) Brightfield images of 18 week dissected human fetal lung tissue. Top panel shows the trachea and main bronchi. To collect cells for scRNA-seq from the trachea, the tracheal tube was cut lengthwise, and epithelial cells were scraped with a scalpel and processed for scRNA-seq. Bottom panel shows the small cartilaginous airways (1) and the distal lung (2). Pieces of tissue this size were homogenized in full to single cells and subjected to scRNA-seq. For all groups, EPCAM+ cells were later isolated from the dataset for downstream analysis. Scale bar represents 1mm. **b**) Cellular transcriptome heterogeneity visualized by tSNE, with cells colored by identity of cell lineages (top; Epi: epithelial clusters; Mes: mesenchymal clusters; Imm: immune clusters; End: endothelial cluster; RBC: red blood cell cluster) and location of specimen (bottom), respectively. Numbers on the top panel tSNE indicate identity of *de novo* identified clusters. **c**) Dot plot showing z-transformed cluster average expression levels and proportion of cells with detectable expression levels in each cluster for canonical cell lineage markers. Colors of cluster index indicate inferred cell lineages. Red: epithelial lineage. Green:

mesenchymal lineage. Purple: endothelial cells. Blue: immune lineage. Gray: red blood cells. Cell lineage identity of marker genes are represented by side bar on top, following the same color scheme as that for clusters. Orange: neuronal lineage. **d**) Feature plots show the expression patterns of lung epithelial marker NKX2.1, pan-epithelial marker EPCAM, and markers for rare cell types Tuft cells (DCLK1) and Ionocytes (FOXI1).

Figure 3-3 Bud tip progenitor and basal cell profiles

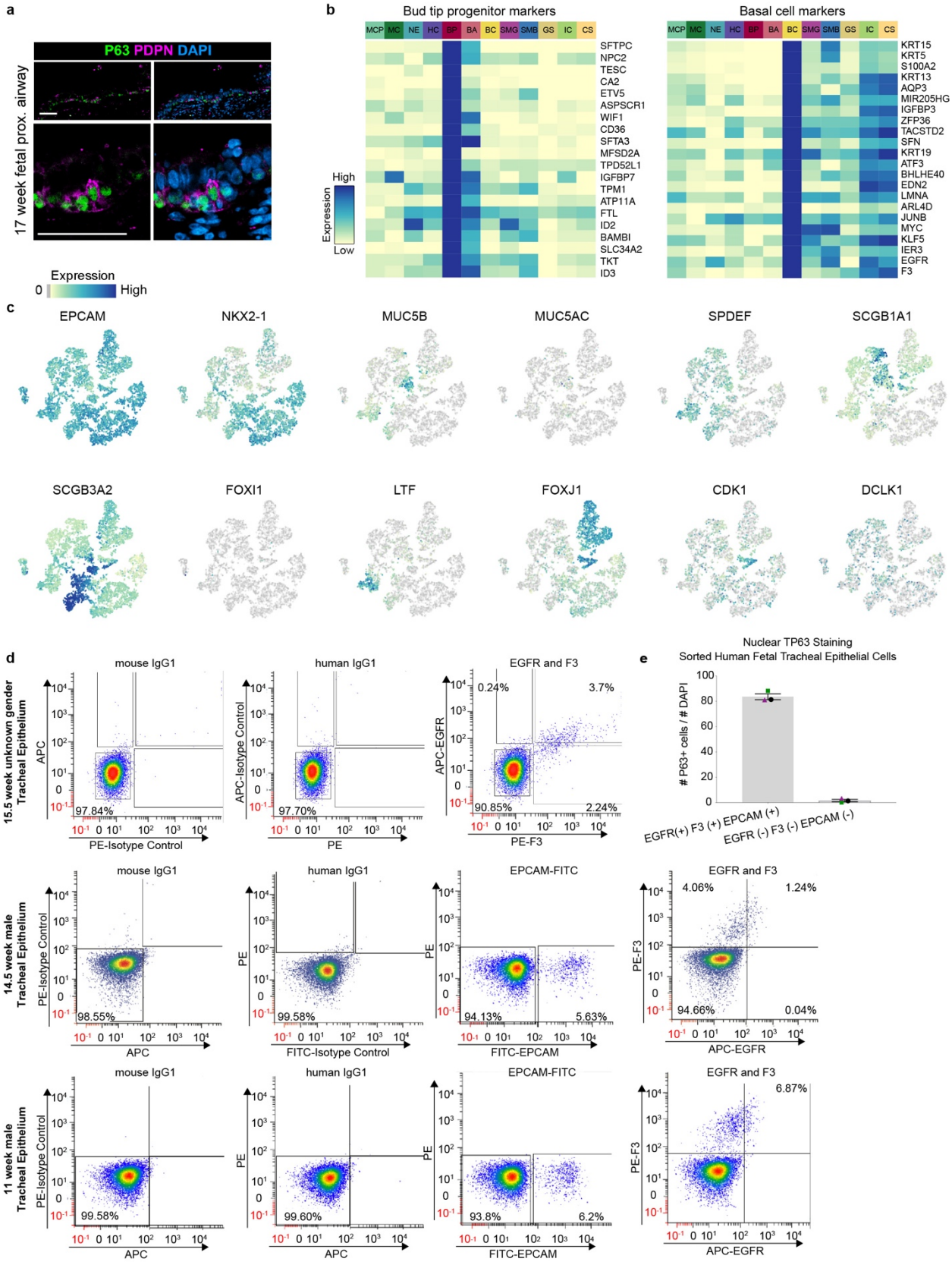


Figure 3-3. Bud tip progenitor and basal cell profiles. a) TP63+ basal cells (green) in the trachea of 17 week fetal lungs (n=3 biological replicates) stain positive for surface marker PDPN (pink). Scale bar represents 50 μm. **b**) Heatmaps show quantiles of cluster average expression levels normalized to expression range of each gene in the scRNA-seq data of 11.5-21 week gestation human fetal lung specimens. Only genes with top fold change in bud tip progenitors (BT, left) or basal cell (BC, right) are presented. **c**) Feature plots on tSNE embeddings show expression of markers for lung epithelial cell types, including goblet cell markers *MUC5B* and *MUC5AC*, secretory marker *SPDEF*, club cell marker *SCGB1A1,* hub marker *SCGB3A2*, submucosal gland cell marker LTF, multiciliated cell marker FOXJ1, proliferation markers MKI67 and CDK1, lung epithelial marker NKX2.1, pan-epithelial marker EPCAM, and markers for rare cell types Tuft cells (DCLK1) and Ionocytes (FOXI1). **d**) Fluorescence Activated Cell Sorting (FACS) data for epithelial cells scraped from fetal tracheas of N=3 biological replicates and stained with EGFR and F3 to isolate native basal cells. Gates were set using IgG negative controls. **e**) Quantification of the number of cells from F3+/EPCAM+ and F3-/EGFR- populations that stained positive for nuclear TP63 after cytospin.

Figure 3-4 A distal- to-proximal cellular hierarchy is reflected by cell type stratification and described by continuous gene expression analysis

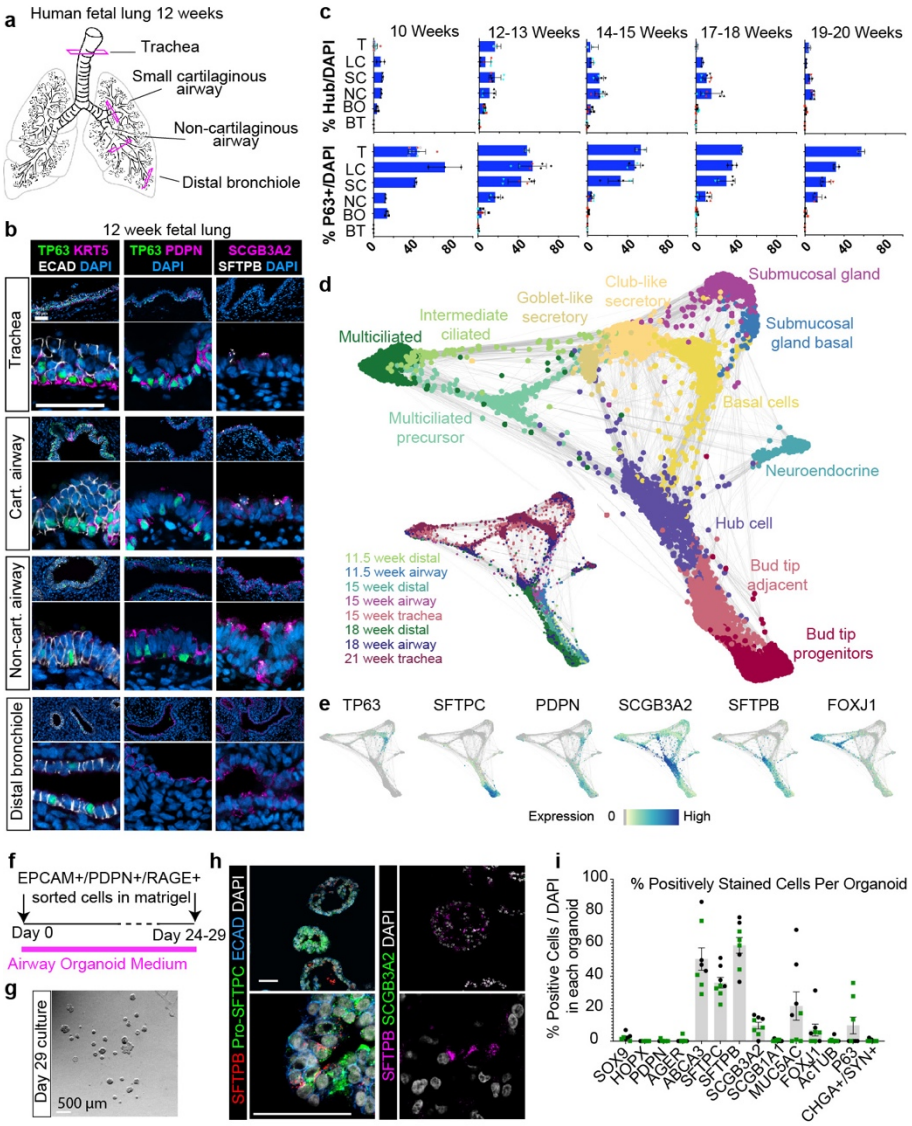


Figure 3-4. A distal- to-proximal cellular hierarchy is reflected by cell type stratification and described by continuous gene expression analysis a) Schematic of a 12 week human fetal lung showing the areas of analysis along the proximal-distal axis of the airway epithelium. **b**) Left panel: protein staining of 12 week fetal lungs along the proximal-distal axis for TP63 (green) and KRT5 (pink). Middle panel: co-staining of TP63 (green) with PDPN (pink). Right panel: co-staining for hub cell markers SCGB3A2

(pink) and SFTPB (white). Note that SCGB3A2 staining is diffuse within the cytoplasm whereas SFTPB staining is punctate and bright. Representative images shown from n=3 independent specimens. Scale bars represent 50 μm. **c)** Top row: quantificiation of hub cells, defined as SCGB3A2+/SFTPB+ cells within an airway region divided by the number of DAPI+ cells within that airway region, throughout the bronchial tree from 10-20 weeks gestation. Bottom row: quantificiation of TP63+ cells, defined as TP63+ cells within an airway region divided by the number of DAPI+ cells within that airway region, throughout the bronchial tree from 10-20 weeks gestation. Biological replicates are plotted using cyan, red, and black dots. Trachea (T), Large Cartilaginous Airway (LC), Small Cartilaginous Airway (SC), Non-Cartilaginous Airway (NC), Bronchiole (BO), Bud Tip (BT). **d)** k-nearest neighbor network construction was performed on scRNA-seq data from the human fetal lung and visualized with SPRING. Cell colors indicate inferred cell type identity presented in Fig. 1 (center) or locations of cells from each specimen (left). **e**) Expression of individual genes were overlaid on the SPRING plot. *TP63* marks basal cells, *SFTPC* marks the bud tip progenitor cluster, *PDPN* marks the bud tip adjacent population as well as basal cells. Hub cells express high levels of *SCGB3A2* and *SFTPB.* Markers of more differentiated cells, including *FOXJ1* (ciliated cells) define the other clusters. **f)** Experimental schematic for the isolation and culture of bud tip adjacent cells. Distal fetal lung tissue from 2 biological replicates (10.5 and 11.5 weeks gestation) was digested and stained for EPCAM, PDPN and RAGE. FACS was used to isolate EPCAM+ cells, then to isolate PDPN+/RAGE+ cells **g)** Brightfield image showing organoids grown from isolated bud tip adjacent cells after 29 days in culture with airway organoid media. Scale bar represents 500 μm. **h**) Organoids derived from isolated bud

tip adjacent cells contain a majority of cells that co-express alveolar epithelial cell type II (AECII) markers SFTPB (green) and Pro-SFTPC (red), first column. Many cells expressed the hub/AECII marker SFTPB (pink), but do not express the hub marker SCGB3A2 (green), second column. Scale bar represents 50 μm. **i**) Quantification of the percentage of positively stained cells for each marker in individual organoids. Green and black dots represent separate biological replicates.

Figure 3-5 Fetal characterization by protein staining and mRNA *in situ* hybridization

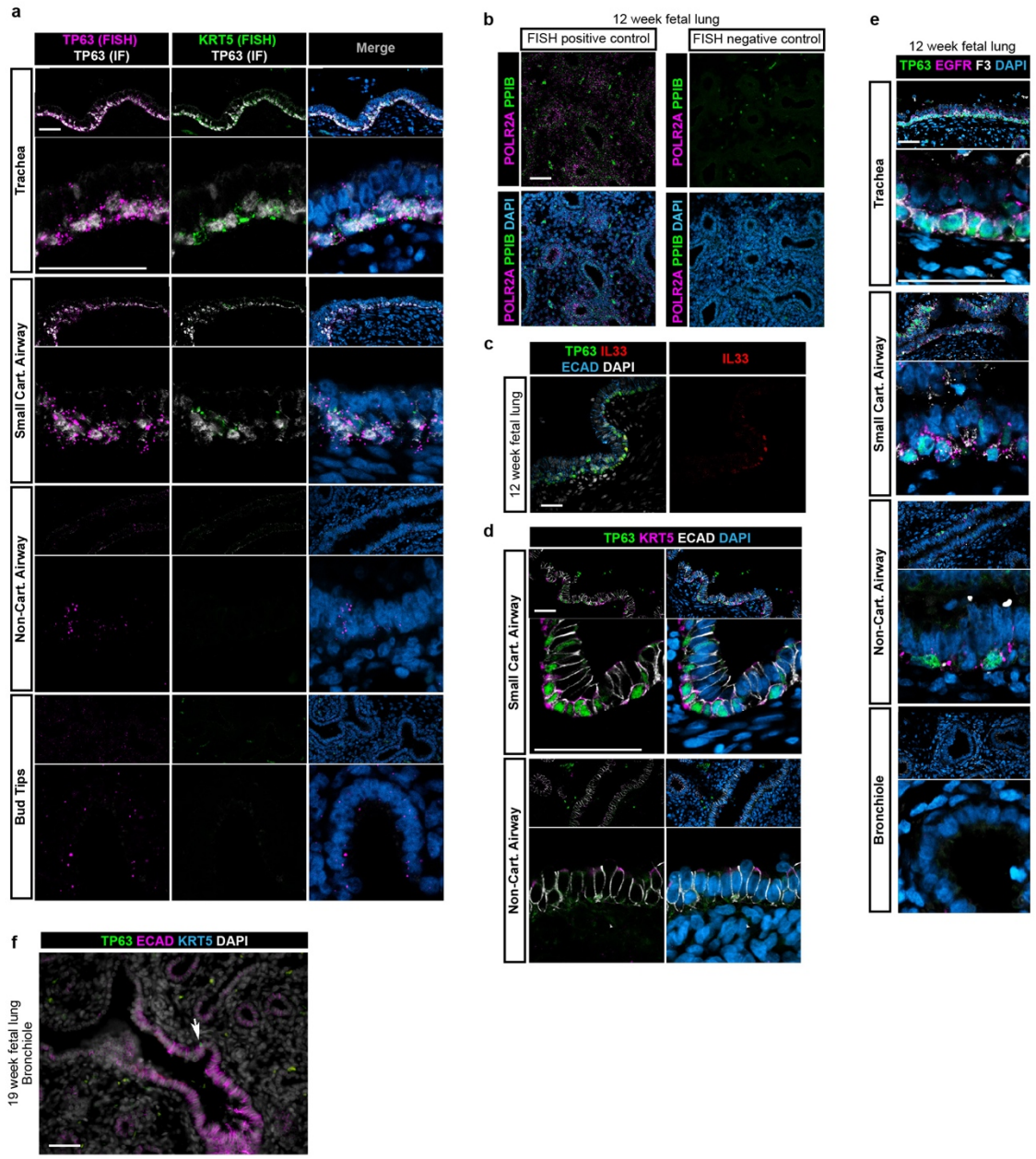


Figure 3-5. Fetal characterization by protein staining and mRNA *in situ* hybridization. **a**) mRNA *in situ* hybridization for TP63 (pink) and *KRT5* (green), combined with protein staining for TP63 (white) confirms the patterns seen in protein staining across the proximal-distal axis of the lung. Tracheal TP63+ basal cells exhibit strong *TP63* and *KRT5* mRNA expression. *KRT5* mRNA expression drops dramatically

in the small cartilaginous airways and is absent in all *TP63*+ cells in the lower airways. Some cells exhibit low levels of *TP63* mRNA expression in the distal lung, including in the bud tip region. Scale bars represent 50 μm. **b**) Positive (POLR2A [pink] and PPIB [green]) and negative control (Scrambled probes 1 [pink] and 2 [green]) stains for *in situ* hybridization. Scale bar represents 50 μm. **c**) Only a portion of TP63+ (green) cells in the 12 week fetal trachea (n=3 biological replicates) stain positively for IL33 (red). Scale bar represents 50 μm. **d**) By 17 weeks gestation (n=3 biological replicates), TP63+ cells are primarily found in the trachea and large cartilaginous airways, and almost all TP63+ (green) cells exhibit KRT5 (pink) staining. **e**) To evaluate the phenotypic differences in protein expression of basal cell markers identified in Figure 1 across the proximal-distal axis in the developing 12 week fetal lung we stained for TP63 (green), EGFR (pink) and F3 (white). Similar to other markers, we observed strong co-staining for all 3 markers in tracheal basal cells, with a decrease in staining intensity for F3 to progressively more distal airways. Interestingly, EGFR appeared to mark TP63+ cells even in the non-cartilaginous airways, suggesting EGFR is an early marker for basal cells. Scale bars represent 50 μm. **f**) Despite low abundance, rare TP63+/KRT5- cells were observed in the distal airways (arrow) in 19-20 fetal lungs. Image is representative of n=3 biological replicates between 19 and 20 weeks gestation. Scale bar represents 50 μm.

Figure 3-6 Characterization of hub progenitor cells and bud tip adjacent cells

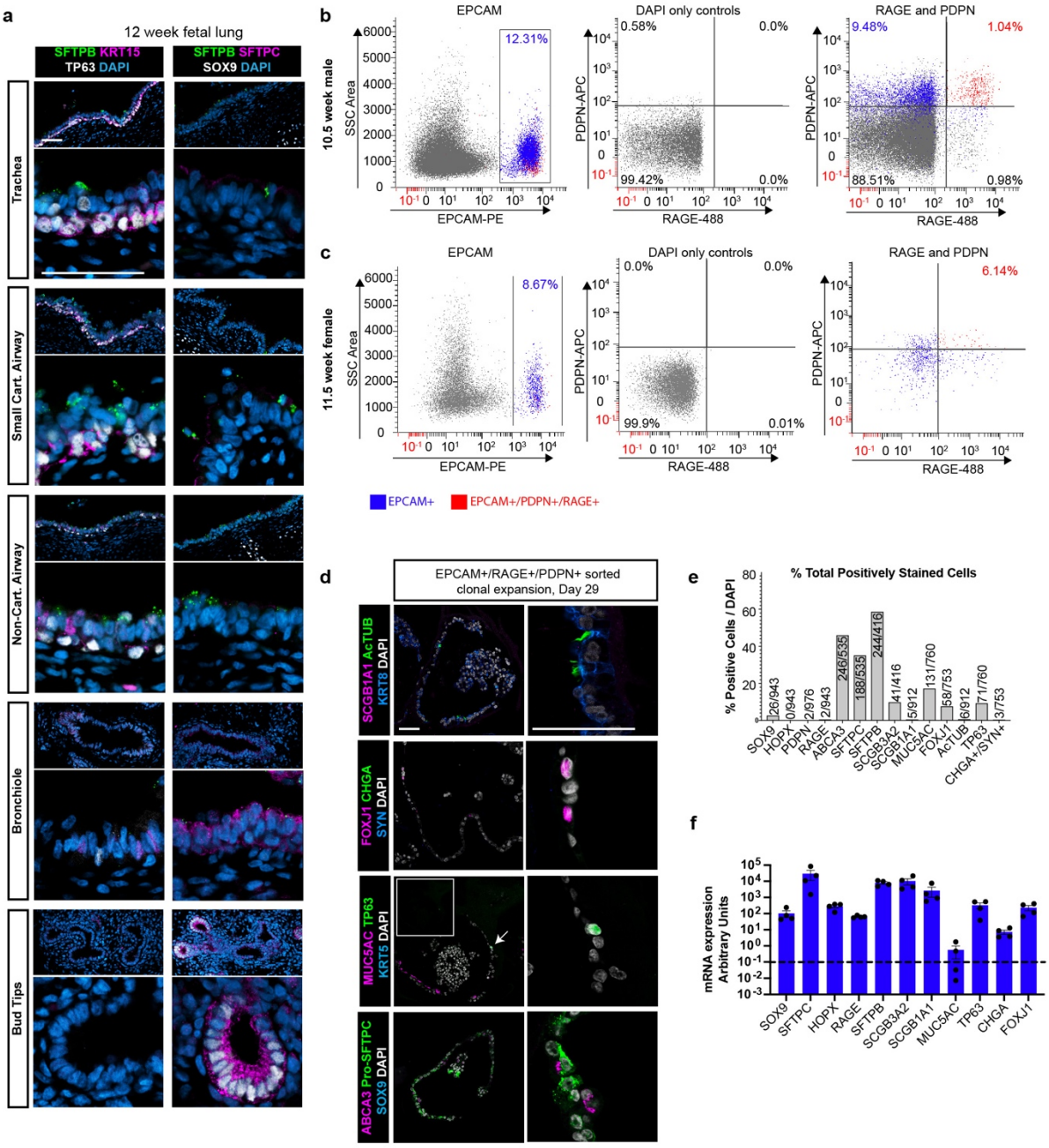


Figure 3-6. Characterization of hub progenitor cells and bud tip adjacent cells. a) Protein staining characterization of hub progenitors throughout the proximal-distal axis of the 12 week fetal lung (n=3 biological replicates) shows that SFTPB+ (green) cells,

which appear as bright punctate marks, strongly mark hub progenitors by scRNAseq, are most prominent in the small and large cartilaginous airways at 12 weeks gestation. They are often apically localized, with some cells appearing above the epithelium (large cartilaginous airway left and second to left panels). Some hub progenitor cells appeared to contain weak nuclear TP63 (white) staining (left panel, small cartilaginous airway). Scale bar represents 50 μm. **b-c)** Bud tip adjacent cells were isolated using Fluorescence Activated Cell Sorting (FACS). The distal lung regions of n=2 biological replicates (10.5 and 11.5 weeks gestation) were dissociated to single cells, and stained for the epithelial marker EPCAM along with bud tip adjacent markers PDPN and RAGE. Cells were sorted first on EPCAM, then PDPN+/RAGE+ cells were isolated. The percentage of triple positive cells was 1.04% in the first sample and 6.14% in the second sample. Gates were set from DAPI only negative controls. **d)** EPCAM+/RAGE+/PDPN+ cells were isolated and grown in matrigel droplets in airway organoid medium for 29 days. Protein staining analysis revealed rare positive staining for multiciliated cell markers (Acetylated Tubulin, AcTUB, green, and FOXJ1, pink), and abundant staining for alveolar epithelial cell type II (AECII) markers such as STFPB and ABCA3. **e)** Quantification of (d), showing the percent total positively stained cells out of all cells counted across 2 biological replicates and 5-7 organoids per replicate. **f)** QRT-PCR of canonical lung epithelial cell type markers in organoids derived from sorted bud tip adjacent cells. Expression is reported as arbitrary units. Error bars represent the mean +/- the standard error of the mean. N=2 biological replicates with N=2 technical replicates each.

Figure 3-7 SMAD activation robustly induces TP63 expression in bud tip progenitors

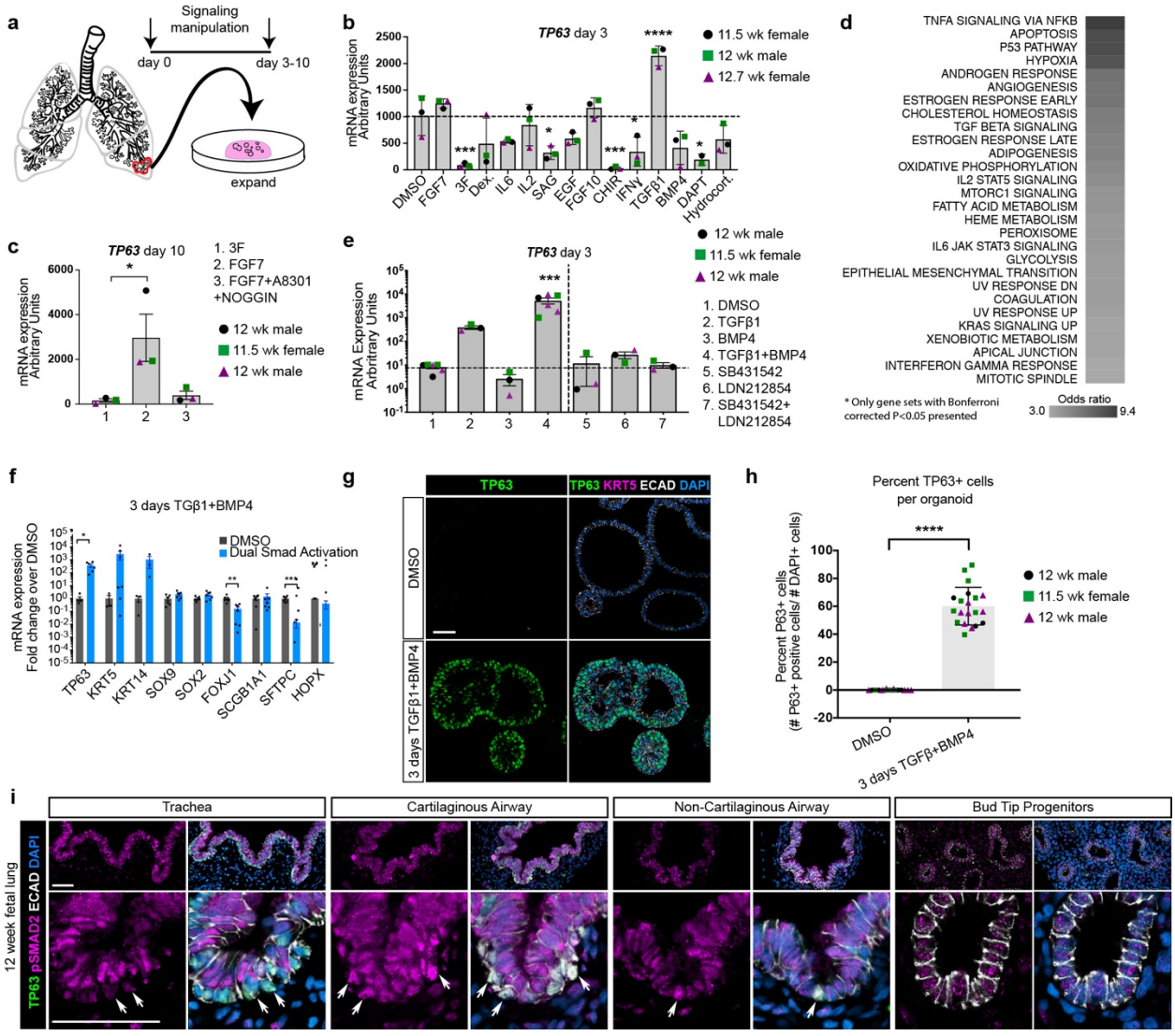


Figure 3-7. SMAD activation robustly induces TP63 expression in bud tip progenitors **a**) Schematic of experimental design. All experiments were carried out using n=3 12 week human fetal lung-derived bud tip progenitor organoids grown and maintained in a 3-dimensional environment within a Matrigel droplet. **b**) Bud tip progenitor organoids were treated for 3 days with serum-free basal medium supplemented with DMSO (control) or with signaling factors known to be important for lung development and cellular differentiation. DMSO (1:1,000 dilution), FGF7 (10

ng/mL), ‘3F’ (FGF7 10ng/mL, CHIR99021 3μM, ATRA 50 nM), Dexamethasone (25 ng/mL), IL6 (10 ng/mL), IL2 (50 U/mL), Smoothened Agonist (SAG; 500 nM), EGF (100 ng/mL), FGF10 (500 ng/mL), CHIR99021 (2 μM), IFNγ (10 ng/mL), TGFβ1 (100 ng/mL), BMP4 (100 ng/mL), DAPT (10 μM), Hydrocortisone (100 ng/mL). QRT-PCR analysis after 3 days for the basal cell marker gene *TP63* is shown, gene expression is reported as arbitrary units. Treatment with TGFβ1 led to a significant increase in the expression of *TP63* (one-way Analysis of Variance (ANOVA) (alpha=0.05, $p<0.0001$, $F=14.7$. Dunnett’s test of multiple comparison’s compared the mean of each group to the mean of the DMSO control group.) Estimated p values are shown on the graph. Error bars are plotted to show mean +/- the standard error of the mean. N=3 independent biological specimens. Data is from a single experiment. **c**) Bud tip progenitor organoids were treated with FGF7 (10 ng/mL), a permissive environment for TP63 expression compared to maintenance in bud tip progenitor medium (‘3F’), or with FGF7 (10 ng/mL) plus factors to inhibit SMAD signaling (A8301 [1 μM] and NOGGIN [100 ng/mL]) and *TP63* gene expression was evaluated by QRT-PCR after 10 days in culture. Treatment with FGF7 led to significantly higher expression of *TP63* over 3F control, and there was no significant difference between *TP63* expression levels in 3F maintenance medium versus treatment with FGF7+SMAD inhibitors A8301 and NOGGIN (one-way ANOVA (alpha=0.05, $p=0.0336$, $F=6.297$). Tukey’s multiple comparison test compared the means of each group to the mean of every other group). Estimated p values are reported on the graph. Error bars are plotted to show mean +/- the standard error of the mean. N=3 independent biological specimens. Data is from a single experiment. **d**) Heatmap shows odds ratio representing enrichment of pathway hallmark gene sets

annotated in MSigDB in basal cell upregulated genes. Only gene sets with Bonferroni-corrected P < 0.05 in one-sided Fisher's exact test are presented. **e)** Bud tip progenitor organoids were treated for 3 days with SMAD activation or inhibition conditions and expression of *TP63* was evaluated by QRT-PCR for all treatment groups. All media contained 3F components (FGF7 10ng/mL, CHIR99021 3μM, ATRA 50 nM), with individual groups containing combinations of: DMSO (1:1000 dilution), TGFβ1 (100 ng/mL), BMP4 (100 ng/mL), SB431542 (10 μM), LDN212854 (200 nM). Treatment for 3 days with both TGFβ1 (100 ng/mL) and BMP4 (100 ng/mL) led to a significant increase in *TP63* expression relative to all other groups (One-way ANOVA alpha=0.05, F=21.19, p<0.0001; Tukey's multiple comparisons of the mean of each group versus the mean in all other groups, estimated p values are reported on the graph). 3 days TGFβ1 and BMP4 is referred to as 'dual SMAD activation', or 'DSA'. Data is plotted as arbitrary units. Error bars are plotted to show mean +/- the standard error of the mean. N=3 independent biological specimens. Data is from a single experiment and is representative of n=3 experiments. **f)** QRT-PCR for markers of canonical differentiated lung epithelial cell types showing DMSO (gray bars) and DSA treated (blue bars) organoids after 3 total days of treatment. Data is plotted as fold change over DMSO controls. DSA treated organoids exhibited a 370-fold increase over DMSO controls in mean *TP63* expression, a 2780-fold increase in mean KRT5 expression and a 1045-fold increase in mean *KRT14* expression, all basal canonical cell markers. *TP63* expression was statistically significantly higher than DMSO controls (Two-sided Mann-Whitney Test, p=0.0238), but the increases in KRT5 and KRT14 were not statistically significant. No other markers exhibited increases in expression after 3 days of DSA treatment.

Some markers exhibited a significant reduction in expression, including the ciliated cell marker FOXJ1 (Two-sided Mann-Whitney Test, p=0.02) and bud tip progenitor marker/AECII marker SFTPC (Two-sided Mann-Whitney Test, p=0.002). Error bars represent the mean +/- the standard error of the mean. n=3 independent biological specimens, and data is from a single experiment and is representative of n=3 experiments. **g**) Protein staining of DMSO treated (control) fetal bud tip progenitor organoids (top row) and 3 days of DSA treatment (bottom row) for TP63+ protein (green), KRT5 (pink) and DAPI (blue). Scale bar represents 50 μm. **h**) Quantification of (g). Total number of TP63+ cells were counted for 3-9 individual organoids across 3 biological replicates. DMSO treated organoid controls contained 0.125% (+/- 0.08%) TP63+ cells, whereas 60.13% (+/- 3.035%) of cells within organoids treated with 3 days of TGFβ1 and BMP4 showed positive TP63 staining. n=3 independent biological specimens. **i)** Protein staining of 12 week fetal lungs showed strong nuclear phospho-SMAD2 staining (pSMAD2; pink) in TP63+ cells (green) in the small- and non-cartilaginous airways, whereas TP63+ cells within the trachea show very low to no nuclear pSMAD2 staining. Scale bars represent 50 μm. Representative images shown from n=3 independent specimens. Estimated p values are reported as follows: * $p<0.05$; ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

Figure 3-8 Screen for growth factor combinations that induce *TP63* expression in fetal bud tip progenitor organoids

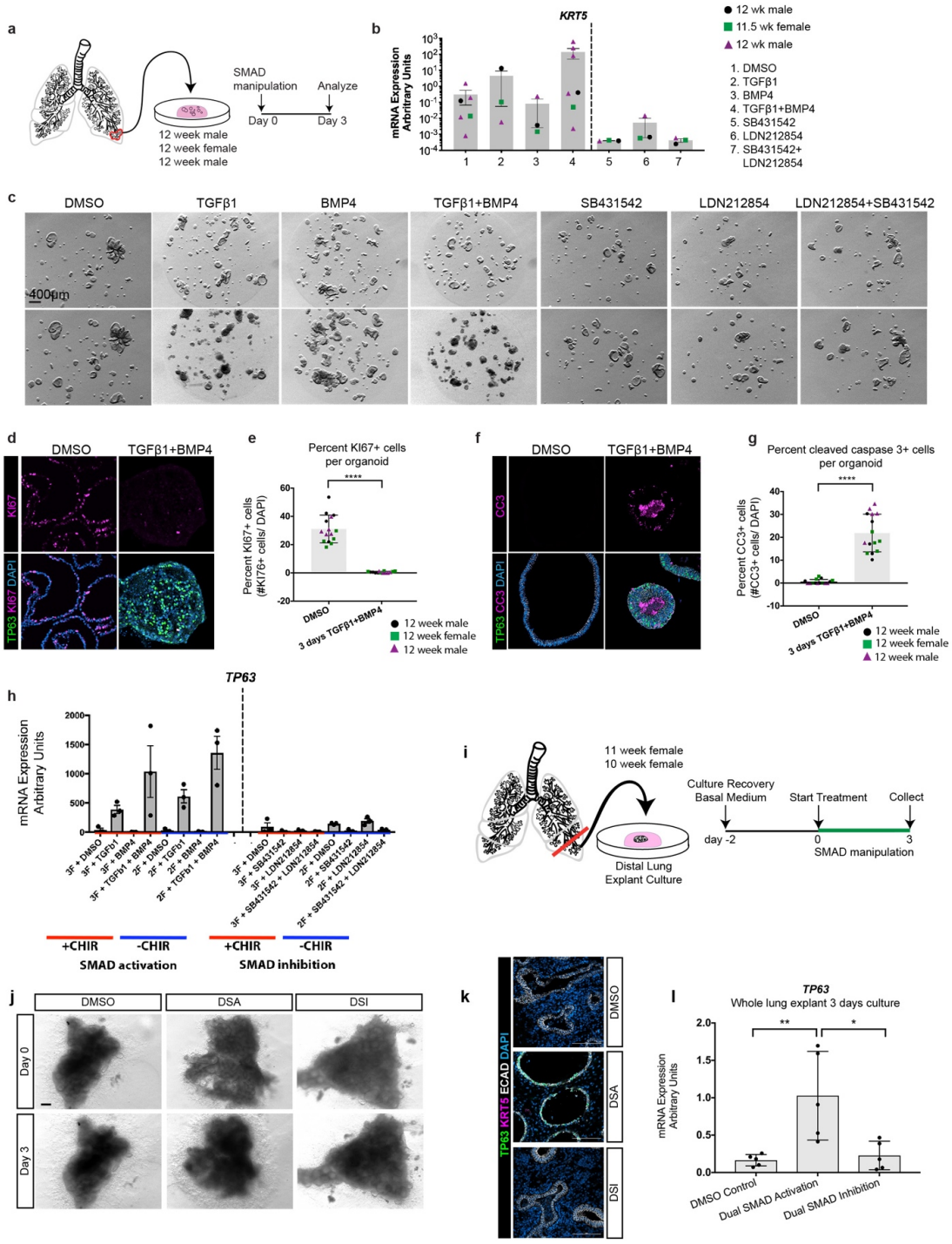


Figure 3-8. Screen for growth factor combinations that induce *TP63* expression in fetal bud tip progenitor organoids. a) Schematic of experimental design. Bud tip progenitors were enzymatically and mechanically isolated from n=3 independent 12 week human fetal lung specimens for all experiments and expanded as 3-dimensional organoids in Matrigel droplets. Once cultures were established, they were treated with serum-free progenitor maintenance medium supplemented with combinations of activators and/or inhibitors of TGFβ and BMP signaling. **b)** QRT-PCR for *KRT5* showed no significant changes in expression for any group (one-way ANOVA, $p>0.05$). Data is presented as mean +/- the standard error of the mean. N=3 biological replicates. **c)** Brightfield images of a single well of organoids for each group at day 0 and after 3 days of treatment. Treatment with TGFβ1 alone or TGFβ and BMP4 led to a reduction in organoid growth, an increase in apparent density of the organoids, and the appearance of many dead cells and debris around the organoids. **d)** Protein staining of proliferation marker KI67 (pink) and TP63 (green) in organoids treated with DMSO or with 3 days of DSA shows that DMSO treated organoids are very proliferative and do not contain any TP63+ cells, whereas organoids treated with 3 days of DSA medium exhibit very few proliferating cells and many TP63+ cells. DSA treated organoids also exhibit a denser morphology. Scale bars represent 50 μm. **e)** Quantification of e. Total number of KI67+ cells were counted for 4-5 individual organoids across 3 biological replicates. DMSO treated controls exhibited 31.04% (+/- 9.79%) KI67+ cells, whereas 0.54% (+/- 2.53%) of cells within organoids treated with 3 days of DSA showed positive KI67 staining. DSA treated organoids exhibited a statistically significant decrease in the number of KI67+ cells (Two-sided Mann-Whitney test, $P<0.0001$). **f)** Protein staining of apoptosis marker

Cleaved Caspase 3 (CC3; pink) and TP63 (green) in organoids treated with DMSO or with 3 days of DSA shows that DMSO treated organoids do not exhibit any apoptosis staining and do not contain any TP63+ cells, whereas organoids treated with 3 days of DSA medium exhibit a dramatic increase in CC3+ cells and many TP63+ cells. DSA treated organoids also exhibit a denser morphology. Scale bar represent 50 μm. **g)** Quantification of g. Total number of CC3+ cells were counted for 5 individual organoids across 3 biological replicates. DMSO treated controls exhibited 0.64% (+/- 0.25%) CC3+ cells, whereas 21.86% (+/-2.12%) of cells within organoids treated with 3 days of DSA showed positive CC3 staining. DSA treated organoids exhibited a statistically significant increase in the number of CC3+ cells (Two-sided Mann-Whitney test, $p<0.0001$). **h)** We tested whether the presence of CHIR99021, a ROCK inhibitor and potent activator of the WNT pathway, affected the outcome of SMAD manipulation on *TP63* expression. No statistical differences were observed between groups treated in the presence or absence of CHIR99021, though we note that variability within groups is high (one-way ANOVA, $p>0.05$). For this work, we continued to include CHIR99021 in the dual SMAD activation (DSA) medium because it seemed to improve survival of organoid cultures. Data is reported as arbitrary units and error bars represent the mean +/- the standard error of the mean. **i**) Schematic of experimental design. The most distal portions of a 10 week and 11 week fetal lung were cut with a scalpel to $1cm^2$ units and laid on top of a fresh Matrigel droplet prior to solidification (see methods). They were cultured in serum-free, growth factor-free medium supplemented with 1% Fetal Bovine Serum for 2 days to facilitate recovery. They were then treated with serum-free medium (no addition of bud tip progenitor maintenance growth factors) supplemented with TGFβ

and BMP pathway activators and inhibitors. **j**) Bright field images of lung explants on day 0 (2 days post plating) and day 3 of treatment. All explants survived and expanded in culture. DSA appeared to reduce the amount of epithelium and allow expansion of the mesenchyme, whereas DSI appeared to allow expansion of the epithelium and the mesenchyme. Scale bar represents 1 mm. **k**) Protein staining for TP63 (green) and KRT5 (pink) reveals a drastic increase in the number of TP63+ cells in the DSA tissue compared to the DMSO and DSI groups. DSA treatment also led to a reduction in the total amount of epithelium and a morphologically more thin epithelium compared to other groups. KRT5 was not detected in any group. Scale bars represent 100 μm. **l**) 2 explants from the 10 week lung (technical replicates) and 3 explants from the 11 week lung (technical replicates) were processed for RNA extraction. QRT-PCR of *TP63* showed that 3 days of DSA treatment led to a significant increase in *TP63* expression relative to both DMSO controls and to DSI treated organoids (one-way ANOVA, alpha=0.05, $p<0.001$, followed by Tukey's multiple comparisons comparing the mean of each group to the mean of every other group. Estimated p values are reported on the graph. There was no statistical difference in *TP63* expression between DMSO controls and DSI treated organoids. Error bars represent the mean +/- the standard error of the mean. Estimated p values are reported as follows: * $p<0.05$; ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

Figure 3-9 Characterization of SMAD signaling in the developing fetal lung

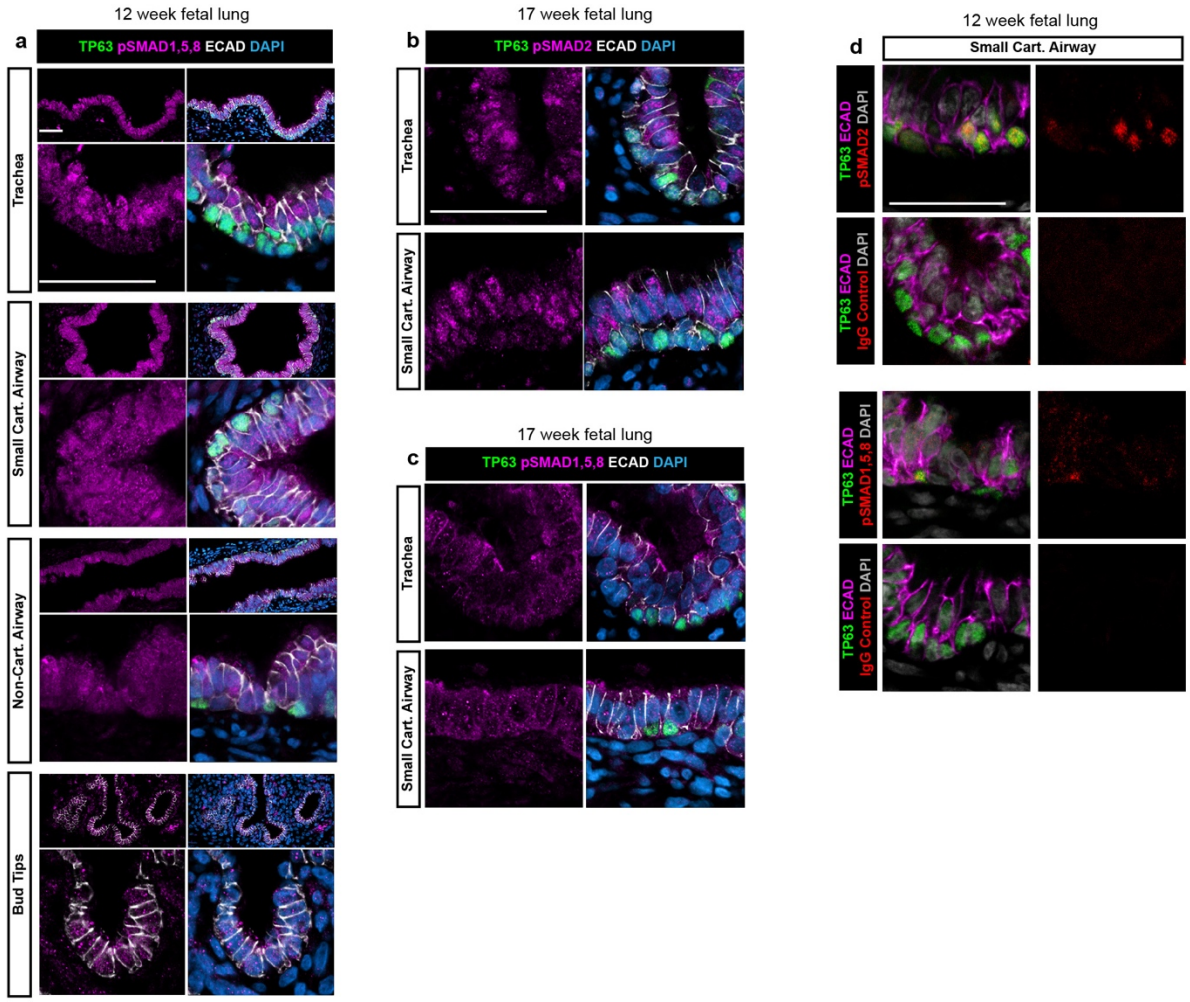


Figure 3-9. Characterization of SMAD signaling in the developing fetal lung. **a**) pSMAD1,5,8 (pink; middle panels) with TP63 (green) shows that nuclear SMAD staining is low in the bud tip progenitors, high in TP63+ cells in the middle airways and low in tracheal TP63+ cells. ECAD staining marks the cell membranes (white). Scale bars represent 50 μm. **b**) Staining for pSMAD2 (pink; left panels), and **c**) pSMAD1,5,8 (pink; middle panels) with TP63 (green) show that, by 17 weeks gestation, all TP63+ cells exhibit little to no nuclear pSMAD staining. ECAD staining marks the cell membranes (white). Scale bars represent 50 μm. **d**) Adjacent tissue sections of 12 week fetal lung

were stained with pSMAD2 or pSMAD1,5,8 and the rabbit IgG isotype control. Slides were imaged with identical laser power and microscope settings. Scale bar represents 50 μm.

Figure 3-10 *In vitro*-derived TP63+ cells can be expanded in culture and are transcriptionally similar to human fetal basal cells

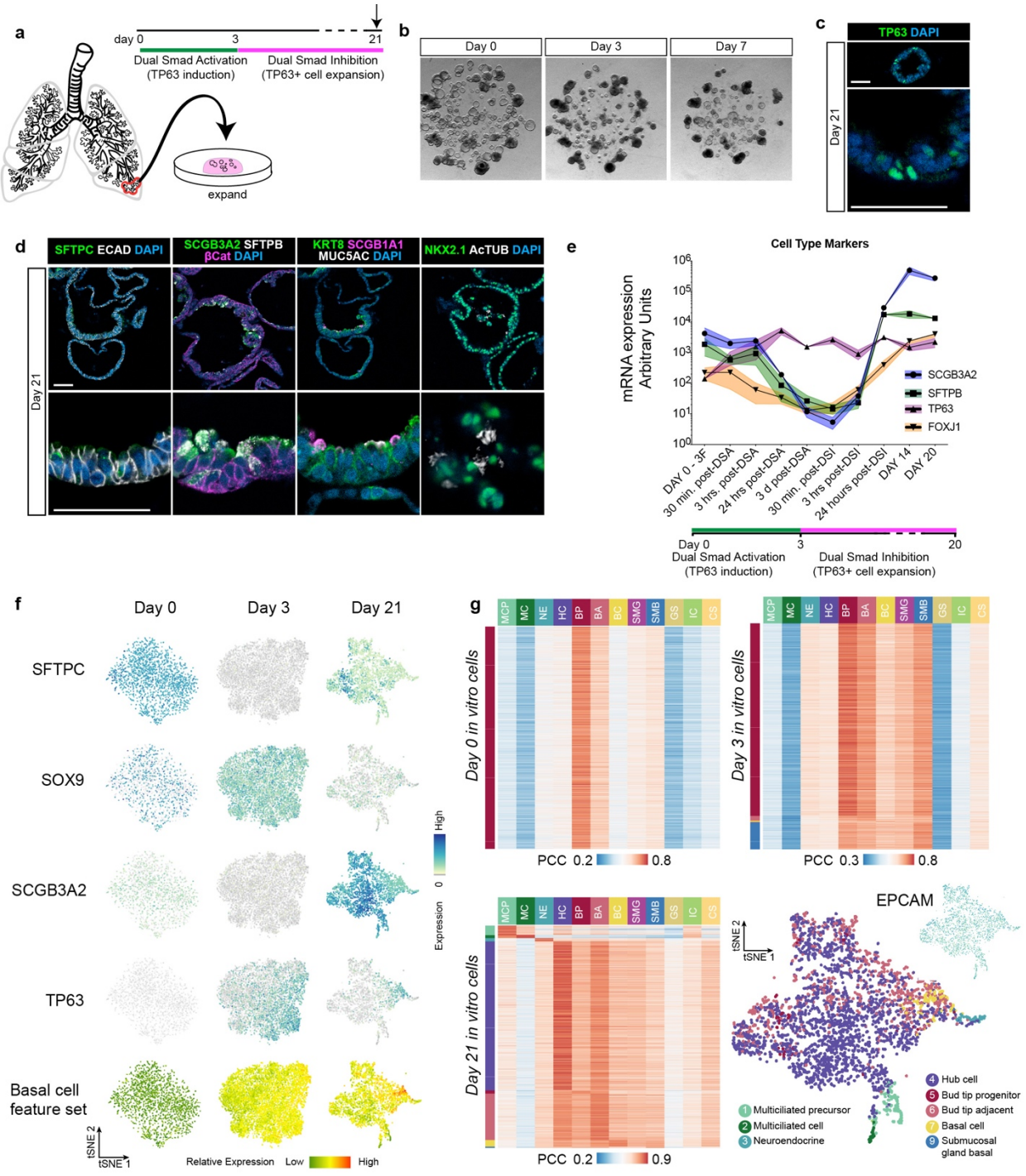


Figure 3-10. ***In vitro*-derived TP63+ cells can be expanded in culture and are transcriptionally similar to human fetal basal cells a**) Experimental overview of

experimental design. Bud tip progenitor organoids were treated with 3 days of DSA and then treated with Dual Smad Inhibition (DSI) to attempt to expand the TP63+ cell population. **b**) Brightfield images of bud tip progenitor organoids at day 0, day 3 of DSA treatment, and after 3 days of DSA and 7 days of DSI (day 10). DSI led to recovery of organoid size after DSA treatment. Scale bar represents 200 μm. **c**) DSA induced organoids expanded for an additional 18 days contained cells positive for TP63 protein. Scale bars represent 50 μm. **d**) After 21 days in culture, organoids were interrogated for protein staining of the bud tip progenitor marker SFTPC (green), hub progenitor cells (SCGB3A2 [green], SFTPB [white]), club cells (SCGB1A1 [pink]), goblet cells (MUC5AC [white]) and multiciliated cells (AcTUB [white]). **e**) Gene expression was evaluated for hub cell markers SCGB3A2, SFTPB, basal cell marker TP63 and multiciliated cell marker FOXJ1 using QRT-PCR at multiple timepoints throughout the DSA-DSI protocol. Data points represent averages of N=3 biological replicates, shaded color bands represent +/- the standard error of the mean. **f)** Feature plots showing expression of bud tip progenitor markers SFTPC and SOX9, hub cell marker SCGB3A2 and basal cell marker TP63 from scRNA-seq of *in vitro* cells at day 0 (bud tip progenitor maintenance), day 3 (after 3 days of DSA) and day 21 (3 days DSA, 18 days DSI). The bottom row shows the sum of z-transformed expression levels of 331 fetal basal cell cluster markers, as identified from human fetal basal cells in Fig. 1, in day 0, day 3 and day 21 *in vitro* samples. **g)** Heatmap showing Pearson's correlation coefficients (PCCs) of transcriptome between each of 2,000 randomly selected cell from each time point and 12 fetal epithelial sub-clusters. PCCs were calculated using log-normalized expression levels of genes that are top 50 markers in any of the fetal epithelial sub-clusters and

detected in *in vitro* cells. Colors represent range of PCCs for all cells in each time point. Bottom right center: 21 day organoids were subjected to scRNA-seq and visualized by tSNE. Colors represent the fetal cell type identity as described in Fig 1 with the highest PCC to each individual 21 day cell. Bottom right corner: feature plot showing expression pattern of EPCAM across day 21 cells. Color scheme follows f).

Figure 3-11 Screen for factors that maintain growth and expansion of TP63+ cells in culture

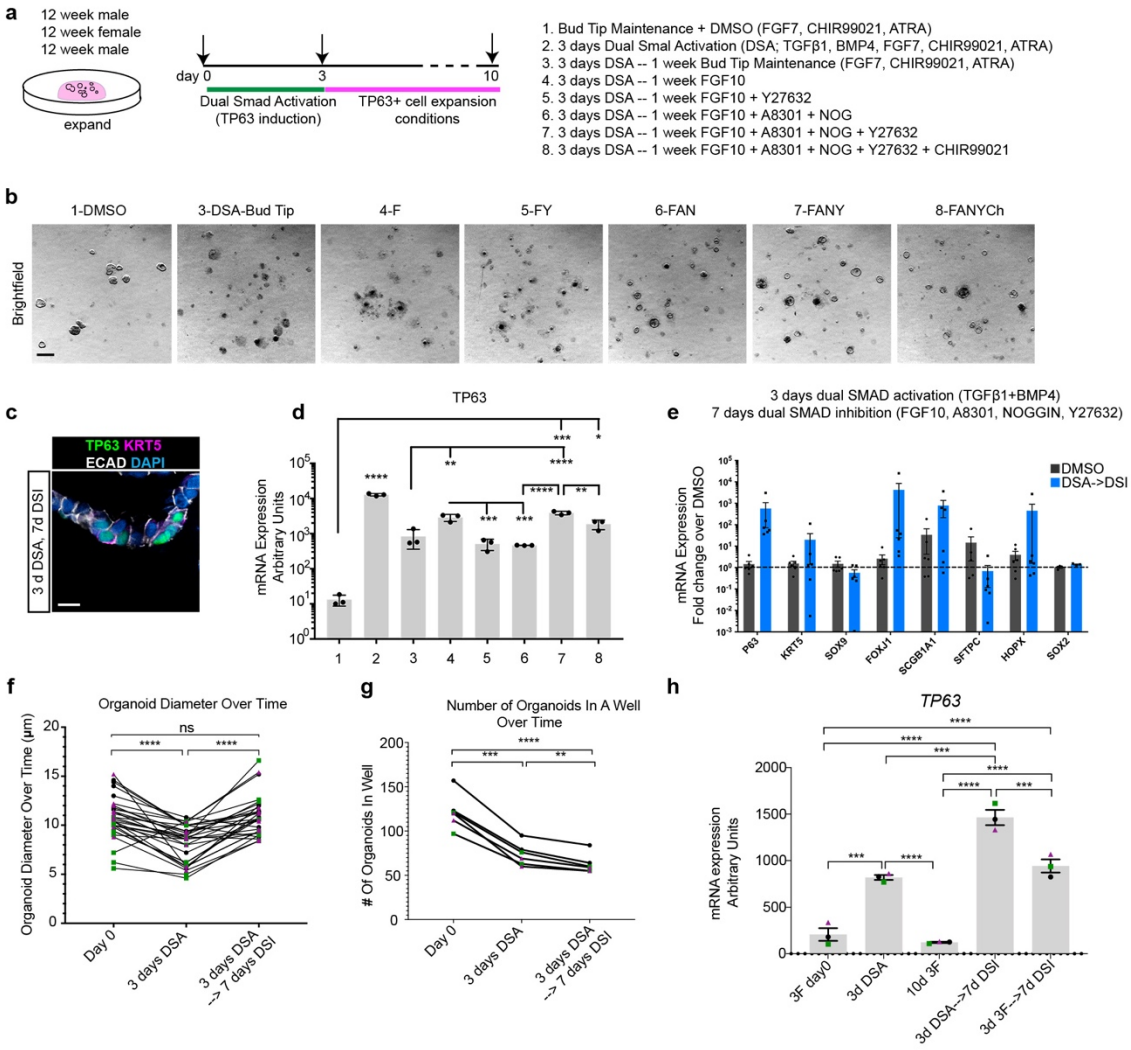


Figure 3-11. Screen for factors that maintain growth and expansion of TP63+ cells in culture. a) Experimental schematic. Fetal bud tip progenitor organoids were treated with DSA for 3 days and then treated with various combinations of growth factors for an additional 7 days to identify conditions that maintained and expanded TP63+ population of cells (10 days total). Experimental conditions are listed as groups 1-8. **b**) Brightfield images of organoids 10 days after treatment. Treatment with group 7 medium led to the best expansion and survival of organoids. Scale bar represents 500 μm. **c**) After 7 days

of expansion in group 7 medium (Dual SMAD Inhibition; DSI), organoids maintained cells that expressed TP63 (green), and a very few number of TP63+ cells also exhibited staining for mature basal cell marker KRT5 (pink). ECAD marks the cell membrane (white). Scale bar represents 50 μm. **d**) Expression of *TP63* was evaluated by QRT-PCR for all treatment groups. Treatment for 3 with DSA led to a statistically significant increase in *TP63* expression over all groups. Treatment with DSA followed by 7 days of treatment with medium in group 7 (FGF10, A8301, NOGGIN, Y2763) or group 8 (FGF10, A8301, NOGGIN, Y2763, CHIR99021) medium led to a significant increase in *TP63* expression compared with the DMSO control (One-way ANOVA, alpha=0.05, $F=191.9$, $p<0.0001$; Tukey's multiple comparisons of the mean of each group versus the mean in all other groups, estimated p values are reported on the graph. $p<0.05$= * ; $p<0.01$ = **, $p<0.001$= ***, $p<0.0001$= ****). N=3 biological replicates. Data is plotted as arbitrary units. Error bars are plotted to show mean +/- the standard error of the mean. Data is from a single experiment and is representative of n=3 experiments. **e)** QRT-PCR for markers of canonical differentiated lung epithelial cell types showing DMSO (gray bars) and DSA--DSI treated (blue bars) organoids after 3 days DSI and 7 days DSI treatment. Data is plotted as fold change over DMSO controls. DSA--DSI treated organoids exhibited a 394-fold increase over DMSO controls in mean*TP63* expression. No markers exhibited statistically significant increases or decreases in expression after treatment (Wilcox Rank-Sum tests), although trends suggest an increase in *TP63,* ciliated cell marker *FOXJ1* and club cell marker *SCGB1A1*. Error bars represent the mean +/- the standard error of the mean. **f)** Quantification of organoid diameter over time prior to culture with DSA (day 0), after 3 days of DSA treatment (day 3) and after 3

days of DSA treatment followed by 7 days of DSI treatment. Each data point represents a single organoid connected by a line to measurements of that organoid on a different day. Organoid size decreases significantly by day 3, and increases significantly from day 3 to day 10 (one-way paired ANOVA, alpha=0.05, $p<0.0001$, $f=25.3$, followed by Tukey's multiple comparison's test to compare the mean of each group to the mean of every other group). Error bars represent mean +/- the standard error of the mean. N=3 biological replicates denoted by color. Estimated p values are reported on the graph. $p<0.05$= * ; $p<0.01$ = **, $p<0.001$= ***, $p<0.0001$= ****. **g)** Quantification of the number of organoids in a well over time prior to culture with DSA (day 0), after 3 days of DSA treatment (day 3) and after 3 days of DSA treatment followed by 7 days of DSI treatment. Each data point represents a single well, connected by lines to measurements from that same well on a different day. The number of organoids decreased significantly after 3 days of DSA treatment and decreased significantly further after treatment with DSI (one-way paired ANOVA, alpha=0.05, $p<0.0001$, $f=143.1$, followed by Tukey's multiple comparison's test to compare the mean of each group to the mean of every other group). Error bars represent mean +/- the standard error of the mean. N=3 biological replicates denoted by color. Estimated p values are reported on the graph. $p<0.05$= * ; $p<0.01$ = **, $p<0.001$= ***, $p<0.0001$= ****. **h)** Bud tip progenitor organoids were treated with either 3 days of DSA followed by DSI treatment for 7 days, or moved directly into DSI treatment without any DSA treatment. Levels of the gene *TP63* were measured by QRT-PCR after 10 days in culture to evaluate whether DSA treatment is beneficial to long term expression levels of *TP63*. Treatment with DSA followed by DSI led to significantly higher levels of *TP63* compared to

treatment with DSI alone (one-way ANOVA, alpha=0.05, $p<0.0001$, $f=89.7$, followed by Tukey's multiple comparison's test to compare the mean of each group to the mean of every other group). Error bars represent mean +/- the standard error of the mean. N=3 biological replicates denoted by color. Estimated p values are reported on the graph. $p<0.05$= * ; $p<0.01$ = **, $p<0.001$= ***, $p<0.0001$= ****.

Figure 3-12 scRNA-seq analysis of *in vitro*-derived TP63+ cells and airway organoids

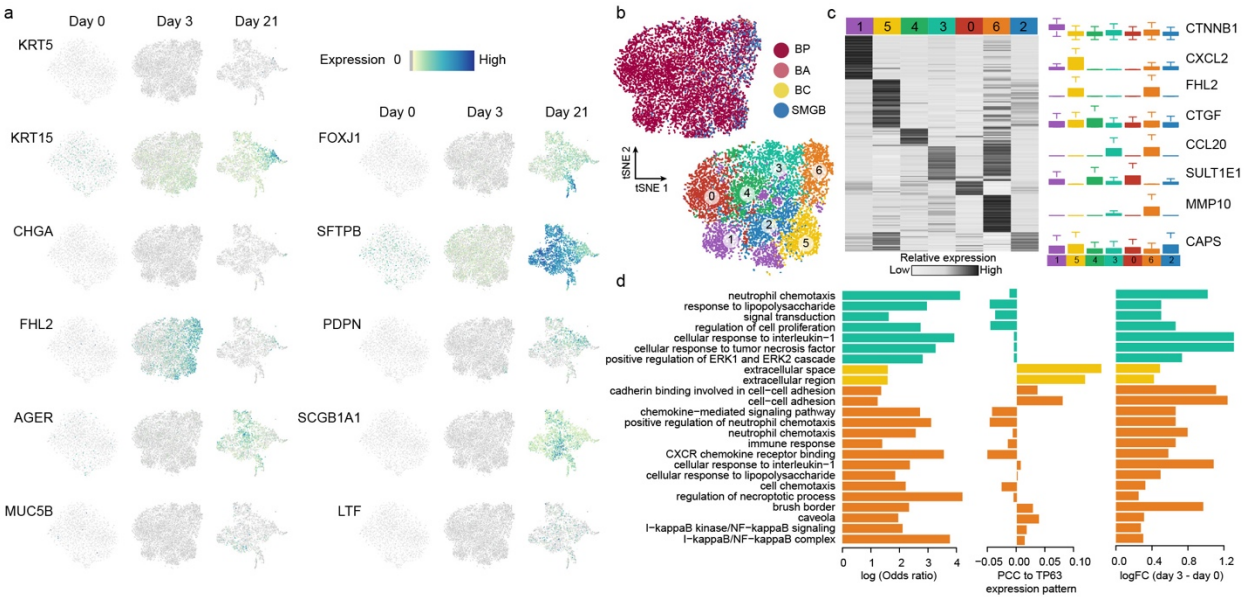


Figure 3-12. scRNA-seq analysis of *in vitro*-derived TP63+ cells and airway organoids

a) Feature plots of gene expression in day 0, day 3 and day 21 *in vitro* organoids from scRNA-seq. **b**) Day 3 organoids treated for 3 days with DSA were analyzed using scRNA-seq. Cellular transcriptome heterogeneity visualized by tSNE. Top panel: cells are colored by the fetal cell type identity to which cells showed highest transcriptome similarity, showing the majority of cells correlated highly with bud tip progenitor, bud tip adjacent and submucosal basal cell clusters. Color scheme follows Fig. 1. Bottom panel: cells are colored by identity of the seven *de novo* identified clusters. **c**) Left panel: heatmap showing z-transformed cluster average expression levels of top marker genes ranked by log-transformed fold change of cluster average expression levels compared to other clusters. Right panel: boxplots (interquartile range with minimum and maximum, outliers removed from plot) showing expression level distributions of selected cluster markers in each cluster. **d**) Characterization of gene ontologies significantly

enriched in marker genes of each cluster. Colors represent cluster identity as b). Significant enrichment was defined as Benjamini-corrected $P < 0.05$ in modified Fisher's exact test as implemented in DAVID. Barplots show the log-transformed odds ratio in modified Fisher's exact test (left), median Pearson's correlation coefficients of expression patterns across day 3 cells between cluster markers annotated to each enriched term and TP63 (middle), and mean log-transformed expression fold change of cluster markers annotated to each enriched term in all day 3 cells compared to all day 0 cells (right).

Figure 3-13 In vitro derived TP63+ cells can be isolated and expanded in culture and display clonal expansion and multilineage differentiation

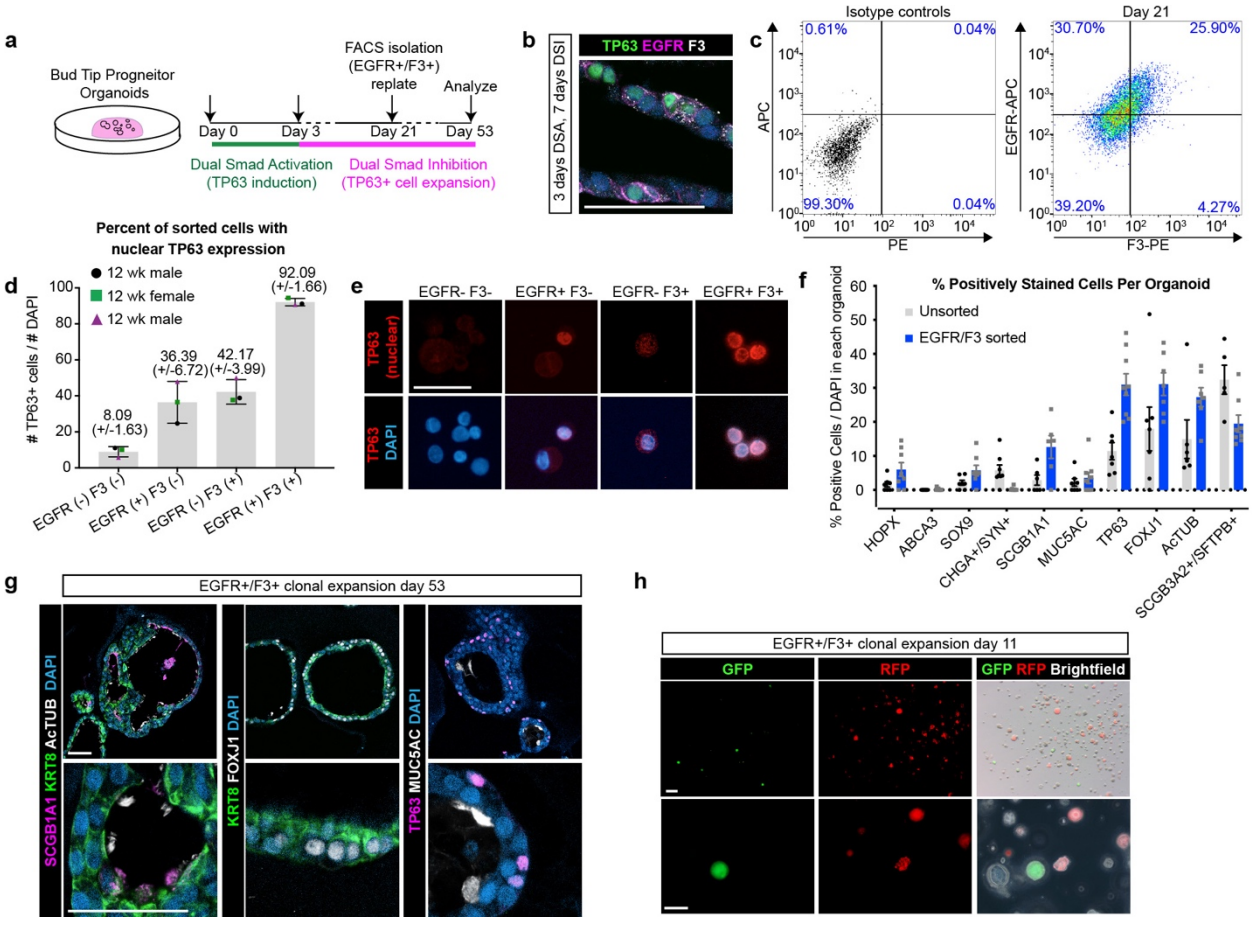


Figure 3-13. In vitro derived TP63+ cells can be isolated and expanded in culture and display clonal expansion and multilineage differentiation a) 12 week fetal bud tip progenitor organoids (n=3 biological replicates) were infected with a lentiviral construct to express constitutive GFP. Organoids were then treated with 3 days DSA to induce *TP63* expression, followed by 18-19 days of treatment with DSI expansion medium, Fluoresence Activated Cell Sorting (FACS) to isolate TP63+ cells, and replating for expansion in DSI. **b**) Organoids treated for 3 days of DSA followed by 7 days of DSI contain TP63+ cells that also stain positive for cell surface markers EGFR and F3, identified as markers of fetal basal cells from scRNAseq in Figure 1. **c**)

EGFR/F3 and control FACS plots for 1 representative biological replicate (n=3 biological replicates). Isotype controls showed negative staining in 99.30% of cells. EGFR-APC and F3-PE sorting segregated 25.90% of all cells to the double positive group. **d-e**) Immediately after cell sorting, cells were spun using a Cytospin on to a glass slide and stained for TP63 (red, nuclear) to identify the percentage of cells in each sorted group that were TP63+. 92.09% of double positive (EGFR+/F3+) cells exhibited nuclear TP63 staining, compared to 36.39% and 42.17% of EGFR-only and F3-only positive groups, respectively, and 8.09% of double negative (EGFR-/F3-) cells. Error bars represent the mean +/- the standard error of the mean. N=3 biological replicates, shown as 3 separate colors on the plot. **f**) Quantification of the percentage of cells in each organoid containing positive protein staining for canonical cell type markers in unsorted organoids versus organoids that were grown from isolated EGFR+/F3+ cells. N=3 biological replicates per group and n=3 technical replicates (individual organoids) per biological replicate. **g**) Protein staining of organoids derived from EGFR/F3 sorted cells revealed the presence of SCGB1A1 (pink) club-like cells and AcTUB+ and FOXJ1+ multiciliated cells (white), first two panels from left. KRT8 (green) stains the pseudo-stratified epithelium. Staining for TP63 was detected in some epithelial cells (pink, right panel), and MUC5AC+ (pink) marked goblet-like cells. N=3 biological replicates. Data shown from a single biological replicate and is representative of N=3 biological replicates. Scale bars represents 50 μm. **i**) GFP (green), RFP (red) and brightfield images of whole organoids 11 days after re-plating and mixing EGFR/F3 sorted cells from GFP and RFP expressing groups. Organoids were either entirely GFP+, entirely RFP+ or entirely fluorescence negative, suggesting clonal expansion from a single cell. N=1 biological

replicate with n=6 technical replicates (wells of mixed organoids). Data shown from a single experiment and is representative of n=3 experiments. Scale bar represents 200 μm (top row) and 100 μm (bottom row).

Figure 3-14 Sorting and clonal expansion of fetal bud tip progenitor organoids

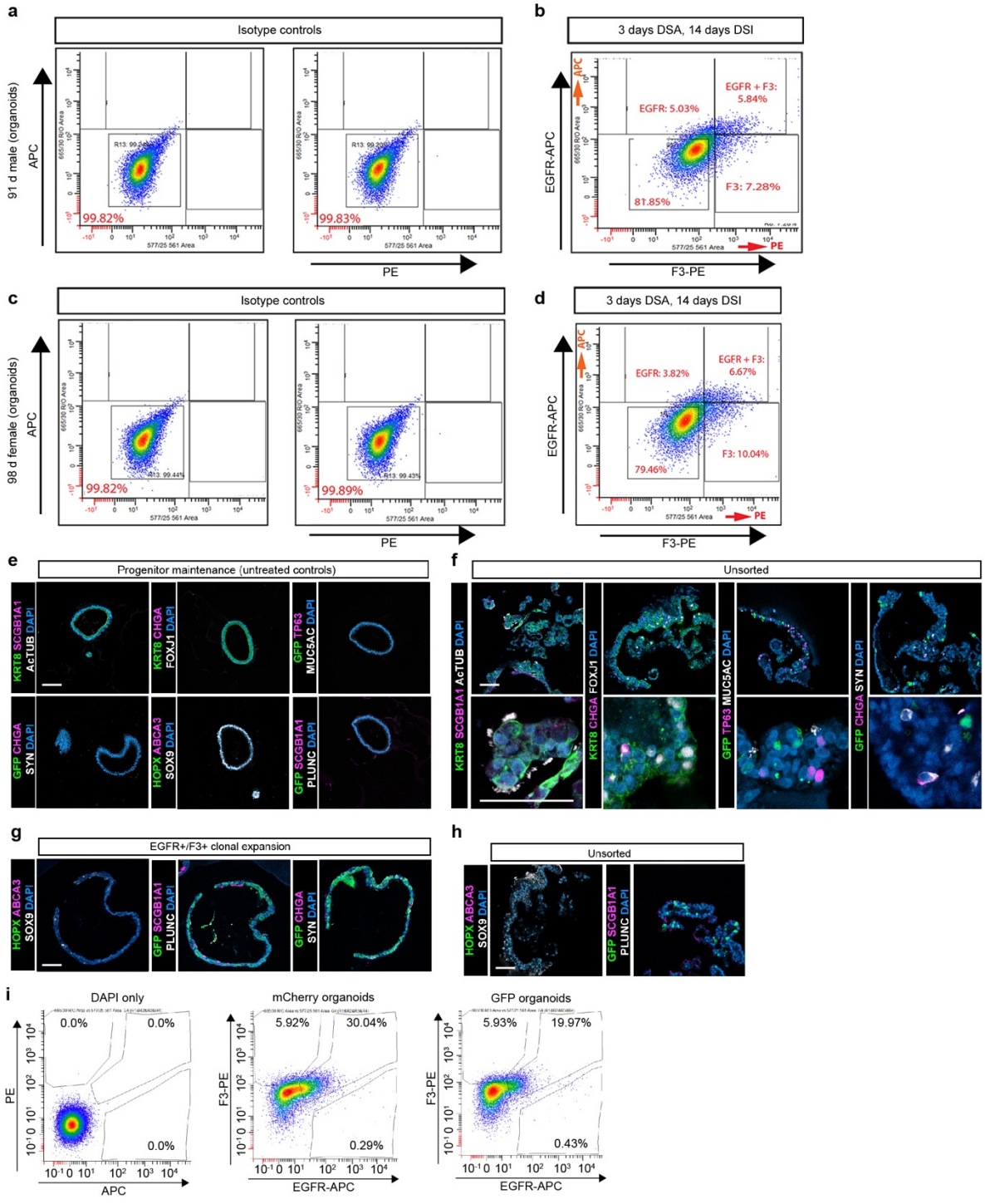


Figure 3-14. Sorting and clonal expansion of fetal bud tip progenitor organoids. a) Fluorescence Activated Cell Sorting (FACS) plots for organoids treated with DSA-DSI

and sorted to isolate EGFR+/F3+ TP63+ cells, biological replicate 2. Isotype controls for APC and PE show 99.82% of cells were negative for APC and 99.83% were negative for PE. **b**) Sorting on EGFR-APC and F3-PE sorted 5.84% of all cells to the double positive group. **c**) Fluorescence Activated Cell Sorting (FACS) plots for biological replicate 3. Isotype controls for APC and PE show 99.82% of cells were negative APC and 99.89 cells were negative for F3. **d**) Sorting on EGFR-APC and F3-PE sorted 6.67% of all cells to the double positive group. **e**) Bud tip progenitor organoids from n=3 biological replicates from 12 week fetal lungs were maintained in serum-free progenitor maintenance medium (FGF7, CHIR99021, ATRA) for 56 days. Staining for markers of differentiated lung epithelial cell types determined that bud tip progenitor organoids did not contain any differentiated cell types (club cell marker SCGB1A1 (pink), neuroendocrine markers Chromagranin A (CHGA; pink) and synaptophysin (SYN; white), ciliated cell marker FOXJ1 (white), basal cell marker TP63 (pink), goblet cell marker mucin 5AC (MUC5AC; white), AECII marker ABCA3 (pink), AECI and hub progenitor cell marker HOPX (green), secretory lineage marker PLUNC (white)). The majority of bud tip progenitor organoid cells were SOX9+ (white). Scale bar represents 50 μm. **f**) Organoids that had been infected with GFP lentivirus but not sorted and left to expand in basal cell expansion medium were collected after 56 days in culture and stained for differentiated epithelial cell markers. Unsorted organoids exhibited clear positive staining for club cell marker SCGB1A1 (pink) and multiciliated markers Acetylated Tubulin (white) and FOXJ1 (white). Neuroendocrine cells were clearly detected (CHGA, pink; SYN, white). Organoids also exhibited TP63+ cells (pink) and cells that stained positive for goblet cell marker MUC5AC (white). Scale bar represents

50 μm. **g**) Organoids treated for 3 days DSA and 18 days DSI-expansion, FACS sorted for EGFR/F3 and clonally expanded in DSI expansion medium were stained for differentiated cell markers. Staining for AECI and hub progenitor marker HOPX (green) was negative, as was staining for AECII marker ABCA3 (pink). Organoids grew clonally, with organoids being either entirely GFP negative or GFP positive (GFP, green, second panel). Many cells exhibited positive staining for club cell marker SCGB1A1 (pink), but staining for secretory lineage marker PLUNC (white) was undetected. In EGFR+/F3+ clonal organoids, no neuroendocrine cells were detected (CHGA, pink; SYN, white), though these cells were clearly detected in unsorted organoids. Scale bar represents 50 μm. **h**) Unsorted organoids also did not show any positive staining for AECI/hub marker HOPX (green), or AECII marker ABCA3 (pink). Consistent with results from sorted organoids, the secretory lineage marker PLUNC was not detected (white). Scale bar represents 50 μm. **i**) FACS isolation of EGFR+/F3+ cells from organoids that had been treated with DSA-DSI, expanded in DSI, and then infected with a lentivirus driving expression of either GFP or mCherry and expanded again. DAPI only controls show 100% of cells are negative for EGFR and F3. 30.04% of cells in the mCherry infected group were EGFR+/F3+, whereas 19.97% of cells from the GFP infected group were EGFR+/F3+.

3.6 Tables

Table 3-1 Cell numbers for scRNAseq in vitro and in vivo data sets

Sample_index	Cell_Ranger_estimated_cell_number	Cell_number_after_filtering	Epithelial_cell_number_1	Epithelial_cell_number_2
Distal_11.5	7,389	6,781	327	326
Airway_11.5	4,530	4,020	415	394
Distal_15	6,780	4,695	686	683
Airway_15	6,475	3,972	455	353
Trachea_15	4,011	3,211	2,329	2,216
Distal_18	22,947	6,954	1,880	1,869
Airway_18	9,613	3,837	1,369	1,291
Trachea_21	4,169	3,337	1,338	1,311
Total	65914	36807	8799	8443
			* extract epithelial lineage	* further remove EpCAM lowly expressed cells within epithelial lineage

	Cell_Ranger_estimated_cell_number	Cell_number_after_filtering
D0	55,828	2,106
D3	9,511	9,400
D21	3,799	3,755

Table 3-2 Enrichment analysis of MSigDB hallmark genes in genes with significantly higher expression levels in basal cells compared to bud tip progenitors

Gene_set_name	Nonimal_P_value	Odds_ratio	Hit_count	Adjusted_P_value
HALLMARK_TNFA_SIGNALING_VIA_NFKB	1.96E-28	9.371647269	50	9.79E-27
HALLMARK_P53_PATHWAY	4.86E-25	8.622227468	46	2.43E-23
HALLMARK_HYPOXIA	8.70E-23	8.184307133	43	4.35E-21
HALLMARK_APOPTOSIS	1.47E-20	8.636349057	37	7.34E-19
HALLMARK_ESTROGEN_RESPONSE_EARLY	8.69E-15	6.154711204	33	4.34E-13
HALLMARK_ESTROGEN_RESPONSE_LATE	2.38E-11	5.248630921	28	1.19E-09
HALLMARK_ADIPOGENESIS	5.91E-11	5.219361002	27	2.95E-09
HALLMARK_OXIDATIVE_PHOSPHORYLATION	1.76E-09	4.734339003	25	8.81E-08
HALLMARK_IL2_STAT5_SIGNALING	9.71E-09	4.476360209	24	4.86E-07
HALLMARK_ANDROGEN_RESPONSE	1.34E-08	6.373167665	17	6.71E-07
HALLMARK_MTORC1_SIGNALING	1.53E-07	4.103428666	22	7.63E-06
HALLMARK_HEME_METABOLISM	4.57E-07	3.976920426	21	2.28E-05
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	1.93E-06	3.749544614	20	9.63E-05
HALLMARK_GLYCOLYSIS	1.93E-06	3.749544614	20	9.63E-05
HALLMARK_FATTY_ACID_METABOLISM	4.65E-06	4.01911169	17	0.00023267
HALLMARK_CHOLESTEROL_HOMEOSTASIS	1.23E-05	5.671088971	11	0.000614424
HALLMARK_KRAS_SIGNALING_UP	5.97E-05	3.235925509	17	0.002983435
HALLMARK_XENOBIOTIC_METABOLISM	6.32E-05	3.219415688	17	0.003161954
HALLMARK_APICAL_JUNCTION	6.70E-05	3.203073452	17	0.003349811
HALLMARK_UV_RESPONSE_DN	8.89E-05	3.608457759	14	0.00444547
HALLMARK_COAGULATION	0.00015862	3.600789585	13	0.007930985
HALLMARK_UV_RESPONSE_UP	0.000170088	3.374357024	14	0.008504393
HALLMARK_INTERFERON_GAMMA_RESPONSE	0.000203992	3.014679346	16	0.010199623
HALLMARK_TGF_BETA_SIGNALING	0.00022272	5.498076651	8	0.011135982
HALLMARK_PEROXISOME	0.00023087	3.964057689	11	0.011543482
HALLMARK_MITOTIC_SPINDLE	0.000238774	2.969465751	16	0.011938681
HALLMARK_ANGIOGENESIS	0.000777437	6.185741424	6	0.038871871
HALLMARK_IL6_JAK_STAT3_SIGNALING	0.000875162	3.930168699	9	0.043758089
HALLMARK_MYOGENESIS	0.001351666	2.692444575	14	0.067583315
HALLMARK_PROTEIN_SECRETION	0.00178641	3.516426139	9	0.089320493
HALLMARK_REACTIVE_OXIGEN_SPECIES_PATHWAY	0.002657855	4.738363097	6	0.13289273
HALLMARK_INFLAMMATORY_RESPONSE	0.003858856	2.474488019	13	0.192942795
HALLMARK_KRAS_SIGNALING_DN	0.004730397	2.516684242	12	0.236519827
HALLMARK_COMPLEMENT	0.008805187	2.307950374	12	0.440259347
HALLMARK_INTERFERON_ALPHA_RESPONSE	0.017660478	2.764157716	7	0.883023896
HALLMARK_WNT_BETA_CATENIN_SIGNALING	0.103374721	2.783896045	3	1
HALLMARK_DNA_REPAIR	0.10244942	1.816998539	7	1
HALLMARK_G2M_CHECKPOINT	0.174859859	1.507412592	8	1
HALLMARK_NOTCH_SIGNALING	0.063790277	3.479794705	3	1
HALLMARK_APICAL_SURFACE	0.120198116	2.589707436	3	1
HALLMARK_HEDGEHOG_SIGNALING	0.020627488	4.124136156	4	1
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	0.038508025	2.31996703	7	1
HALLMARK_PI3K_AKT_MTOR_SIGNALING	0.027889688	2.498584336	7	1
HALLMARK_E2F_TARGETS	0.622187793	0.932703307	5	1
HALLMARK_MYC_TARGETS_V1	0.024573607	2.051854079	11	1
HALLMARK_MYC_TARGETS_V2	0.469656173	1.280034824	2	1
HALLMARK_BILE_ACID_METABOLISM	0.08655248	2.024727069	6	1
HALLMARK_ALLOGRAFT_REJECTION	0.021725529	2.09393471	11	1
HALLMARK_SPERMATOGENESIS	0.807254631	0.65702555	2	1
HALLMARK_PANCREAS_BETA_CELLS	0.616238193	1.060577473	1	1

Table 3-3 Antibodies

Primary Antibody	Source	Catalog #	Dilution (Sections)	Clone
*Biotin-Goat anti-TP63	R&D systems	BAF1916	1:500	
*Biotin-Mouse anti MUC5AC	Abcam	ab79082	1:500	Monoclonal
Goat anti-CC10 (SCGB1A1)	Santa Cruz Biotechnology	sc-9770	1:200	C-20
Goat anti-Chromogranin A (CHGA)	Santa Cruz Biotechnology	sc-1488	1:100	C-20
Goat anti-SOX2	Santa Cruz Biotechnology	Sc-17320	1:200	polyclonal
Mouse anti-ABCA3	Seven Hills Bioreagents	WMAB-17G524	1:500	17-H5-24
Mouse anti-Acetylated Tubulin (ACTUB)	Sigma-Aldrich	T7451	1:1000	6-11B-1
Mouse anti-E-Cadherin (ECAD)	BD Transduction Laboratories	610181	1:500	36/E-Cadherin
Mouse anti-F3	Sigma-Aldrich	AMAb91236	1:250	Monoclonal
Mouse anti-MUC5B	Abcam	AB77995	1:250	monoclonal
Mouse anti-PLUNC	R&D systems	MAP1897	1:500	monoclonal
Mouse anti-Surfactant Protein B (SFTPB)	Seven Hills Bioreagents	Wmab-1B9	1:250	monoclonal
Rabbit anti-Clara Cell Secretory Protein (CCSP; SCGB1A1)	Seven Hills Bioreagents	Wrab-3950	1:250	polyclonal
Rabbit anti-Cleaved Caspase 3	Cell Signaling Technology	9664	1:500	Monoclonal
Rabbit anti-EGFR	Sigma-Aldrich	HPA018530	1:500	Polyclonal
Rabbit anti-HOPX	Santa Cruz Biotechnology	Sc-30216	1:250	polyclonal
Rabbit anti-IL33	Atlas Antibodies	HPA024426	1:500	
Rabbit anti-KRT15	Atlas Antibodies	HPA024554	1:500	Polyclonal
Rabbit anti-KRT5	Sigma-Aldrich	HPA059479	1:500	Polyclonal
Rabbit anti-NKX2.1	Abcam	ab76013	1:200	EP1584Y
Rabbit anti-PDPN	Santa Cruz Biotechnology	Sc-134482	1:500	polyclonal
Rabbit anti-phospho-SMAD1,5,8	Millipore	AB3848	1:500	

Rabbit anti-phospho-SMAD2	Abcam	AB188334	1:250	
Rabbit anti-Pro-Surfactant protein C (Pro-SFTPC)	Seven Hills Bioreagents	Wrab-9337	1:500	polyclonal
Rabbit anti-SCGB3A2	Abcam	AB181853	1:500	
Rabbit anti-SMAD4	Abcam	AB40759	1:250	Polyclonal
Rabbit anti-SOX9	Millipore	AB5535	1:500	polyclonal
Rabbit anti-Synaptophysin	Abcam	AB32127	1:500	monoclonal
Rat anti-KI67	Biolegend	652402	1:100	16A8
Secondary Antibody	**Source**	**Catalog #**	**Dilution**	
Donkey anti-goat 488	Jackson Immuno	705-545-147	1:500	
Donkey anti-goat 647	Jackson Immuno	705-605-147	1:500	
Donkey anti-goat Cy3	Jackson Immuno	705-165-147	1:500	
Donkey anti-mouse 488	Jackson Immuno	715-545-150	1:500	
Donkey anti-mouse 647	Jackson Immuno	415-605-350	1:500	
Donkey anti-mouse Cy3	Jackson Immuno	715-165-150	1:500	
Donkey anti-rabbit 488	Jackson Immuno	711-545-152	1:500	
Donkey anti-rabbit 647	Jackson Immuno	711-605-152	1:500	
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102	1:500	
Donkey anti-goat 488	Jackson Immuno	705-545-147	1:500	
Donkey anti-goat 647	Jackson Immuno	705-605-147	1:500	
Donkey anti-goat Cy3	Jackson Immuno	705-165-147	1:500	
Donkey anti-mouse 488	Jackson Immuno	715-545-150	1:500	
Donkey anti-mouse 647	Jackson Immuno	415-605-350	1:500	
Donkey anti-mouse Cy3	Jackson Immuno	715-165-150	1:500	
Donkey anti-rabbit 488	Jackson Immuno	711-545-152	1:500	
Donkey anti-rabbit 647	Jackson Immuno	711-605-152	1:500	
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102	1:500	
Streptavidin 488	Jackson Immuno	016-540-084	1:500	

Table 3-4 QRT-PCR Primers

Species	Gene Target	Forward Primer Sequence	Reverse Primer Sequence
Human	*FOXJ1*	CAACTTCTGCTACTTCCGCC	CGAGGCACTTTGATGAAGC
Human	*GAPDH*	AATGAAGGGGTCATTGATGG	AAGGTGAAGGTCGGAGTCAA
Human	*HOPX*	GCCTTTCCGAGGAGGAGAC	TCTGTGACGGATCTGCACTC
Human	*KRT5*	CTGGTCCAACTCCTTCTCCA	GGAGCTCATGAACACCAAGC
Human	*KRT14*	TCTGCAGAAGGACATTGGC	GGCCTGCTGAGATCAAAGAC
Human	*MUC5AC**	GCACCAACGACAGGAAGGATGAG	CACGTTCCAGAGCCGGACAT
Human	*TP63*	CCACAGTACACGAACCTGGG	CCGTTCTGAATCTGCTGGTCC
Human	*SCGB1A1*	ATGAAACTCGCTGTCACCCT	GTTTCGATGACACGCTGAAA
Human	*SOX2*	GCTTAGCCTCGTCGATGAAC	AACCCCAAGATGCACAACTC
Human	*SOX9*	GTACCCGCACTTGCACAAC	ATTCCACTTTGCGTTCAAGG
Human	*SFTPC*	AGCAAAGAGGTCCTGATGGA	CGATAAGAAGGCGTTTCAGG

Table 3-5 Gene markers of clusters from single cell RNA sequencing of human lung epithelium or in vitro cultured lung epithelial cells

Index	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene_name	fetal_SPRING_selected	vitro_vivo_comparison_selected
CDC20B_1	0	2.33998684	0.935	0.045	0	1	CDC20B	TRUE	FALSE
HES6_1	0	1.82466563	0.932	0.109	0	1	HES6	TRUE	TRUE
LRRC26_1	0	1.76562439	0.945	0.156	0	1	LRRC26	TRUE	FALSE
CCDC67_1	0	1.29457936	0.831	0.056	0	1	CCDC67	TRUE	TRUE
HYLS1_1	0	1.28918037	0.739	0.074	0	1	HYLS1	TRUE	TRUE
NEK2_1	0	1.26048992	0.811	0.047	0	1	NEK2	FALSE	FALSE
CDC20_1	0	1.19439483	0.756	0.069	0	1	CDC20	FALSE	FALSE
PLK4_1	0	1.13305439	0.811	0.049	0	1	PLK4	TRUE	TRUE
ANLN_1	0	0.97606661	0.691	0.052	0	1	ANLN	FALSE	FALSE
POC1A_1	0	0.9142292	0.7	0.054	0	1	POC1A	FALSE	FALSE
HIST1H2BJ_1	0	0.8329366	0.577	0.03	0	1	HIST1H2BJ	FALSE	FALSE
E2F7_1	0	0.6178903	0.557	0.028	0	1	E2F7	FALSE	FALSE
FOXN4_1	0	0.60735541	0.518	0.01	0	1	FOXN4	FALSE	FALSE
MCIDAS_1	0	0.53876482	0.508	0.025	0	1	MCIDAS	FALSE	FALSE
SYT5_1	0	0.536571	0.436	0.015	0	1	SYT5	FALSE	FALSE
MELK_1	0	0.51508479	0.57	0.044	0	1	MELK	FALSE	FALSE
SAG_1	0	0.34670539	0.293	0.003	0	1	SAG	FALSE	FALSE
RIBC2_1	2.18E-300	1.1222468	0.821	0.112	5.19E-296	1	RIBC2	TRUE	FALSE
CDK1_1	4.96E-299	1.39711823	0.866	0.132	1.18E-294	1	CDK1	FALSE	FALSE
STIL_1	5.52E-290	0.47339855	0.469	0.03	1.32E-285	1	STIL	FALSE	FALSE
MUC12_1	1.31E-282	0.74873686	0.436	0.026	3.12E-278	1	MUC12	FALSE	FALSE
SGOL2_1	1.20E-266	0.74794646	0.635	0.069	2.87E-262	1	SGOL2	FALSE	FALSE
KDELC2_1	5.90E-255	1.05270004	0.814	0.142	1.41E-250	1	KDELC2	TRUE	TRUE
ZMYND10_1	8.98E-252	1.63521016	0.99	0.227	2.14E-247	1	ZMYND10	TRUE	FALSE
CCDC74A_1	6.45E-241	1.36244282	0.932	0.193	1.54E-236	1	CCDC74A	TRUE	TRUE
TOP2A_1	3.82E-240	0.74188897	0.717	0.098	9.10E-236	1	TOP2A	FALSE	FALSE
CEP152_1	7.94E-234	0.54887983	0.537	0.055	1.89E-229	1	CEP152	FALSE	FALSE
CCDC19_1	5.25E-232	0.97556854	0.896	0.16	1.25E-227	1	CCDC19	TRUE	TRUE
CEP78_1	6.56E-224	0.96265303	0.834	0.172	1.56E-219	1	CEP78	FALSE	FALSE
ASF1B_1	1.15E-217	0.67633949	0.57	0.067	2.74E-213	1	ASF1B	FALSE	FALSE
CCNO_1	3.49E-210	2.54765071	0.941	0.334	8.32E-206	1	CCNO	TRUE	TRUE
RP11-263K19.4_1	1.72E-209	0.70937619	0.726	0.112	4.10E-205	1	RP11-263K19.4	FALSE	FALSE
C1orf192_1	2.22E-207	1.09936893	0.954	0.198	5.29E-203	1	C1orf192	TRUE	FALSE
CENPW_1	5.75E-205	0.81376836	0.779	0.147	1.37E-200	1	CENPW	FALSE	FALSE
SPAG6_1	1.60E-201	1.04650713	0.893	0.182	3.81E-197	1	SPAG6	TRUE	TRUE
CEP41_1	1.57E-198	0.95994618	0.824	0.181	3.74E-194	1	CEP41	FALSE	FALSE
PITPNM1_1	3.43E-197	0.86364627	0.808	0.164	8.18E-193	1	PITPNM1	FALSE	FALSE
C21orf128_1	9.44E-197	0.66262546	0.495	0.054	2.25E-192	1	C21orf128	FALSE	FALSE
C11orf88_1	4.76E-196	1.30128113	0.886	0.192	1.13E-191	1	C11orf88	TRUE	FALSE
USP13_1	3.93E-195	0.68040216	0.651	0.105	9.37E-191	1	USP13	FALSE	FALSE
TEKT2_1	2.08E-193	0.95576712	0.873	0.184	4.95E-189	1	TEKT2	FALSE	FALSE
FOXJ1_1	7.16E-193	1.48731283	0.987	0.333	1.71E-188	1	FOXJ1	TRUE	TRUE
RSPH9_1	1.85E-192	1.23555617	0.86	0.184	4.42E-188	1	RSPH9	TRUE	TRUE
C7orf57_1	2.53E-191	0.73500336	0.775	0.136	6.03E-187	1	C7orf57	FALSE	FALSE
C11orf16_1	7.49E-190	0.92030139	0.788	0.149	1.79E-185	1	C11orf16	FALSE	FALSE
DEPDC1_1	1.04E-187	0.35913881	0.394	0.034	2.47E-183	1	DEPDC1	FALSE	FALSE

TMEM106C_1	8.33E-187	1.320376	0.935	0.334	1.98E-182	1	TMEM106C	TRUE	TRUE
MAD2L1_1	7.40E-186	0.89091094	0.772	0.157	1.76E-181	1	MAD2L1	FALSE	FALSE
NUF2_1	8.37E-185	0.44190594	0.463	0.05	1.99E-180	1	NUF2	FALSE	FALSE
TSPAN2_1	3.82E-184	0.46224279	0.453	0.048	9.11E-180	1	TSPAN2	FALSE	FALSE
TMEM190_2	0	2.97629168	0.979	0.106	0	2	TMEM190	TRUE	FALSE
CAPS_2	0	2.87079309	1	0.488	0	2	CAPS	TRUE	TRUE
C9orf24_2	0	2.82341963	0.998	0.164	0	2	C9orf24	TRUE	FALSE
C20orf85_2	0	2.67791231	0.999	0.148	0	2	C20orf85	TRUE	TRUE
C1orf194_2	0	2.65794922	0.99	0.142	0	2	C1orf194	TRUE	FALSE
FAM183A_2	0	2.5792285	0.996	0.149	0	2	FAM183A	TRUE	TRUE
OMG_2	0	2.40168269	0.902	0.09	0	2	OMG	TRUE	FALSE
RP11-356K23.1_2	0	2.38755189	0.948	0.092	0	2	RP11-356K23.1	TRUE	FALSE
PIFO_2	0	2.30720547	0.993	0.162	0	2	PIFO	TRUE	TRUE
AGR3_2	0	2.28685737	0.996	0.455	0	2	AGR3	TRUE	TRUE
RSPH1_2	0	2.28033912	0.988	0.147	0	2	RSPH1	TRUE	TRUE
ODF3B_2	0	2.25309138	0.989	0.152	0	2	ODF3B	TRUE	TRUE
SNTN_2	0	2.25257526	0.973	0.066	0	2	SNTN	TRUE	FALSE
DYNLRB2_2	0	2.12294904	0.977	0.113	0	2	DYNLRB2	TRUE	FALSE
FABP6_2	0	2.11125913	0.904	0.053	0	2	FABP6	TRUE	TRUE
DNAAF1_2	0	2.10589292	0.959	0.105	0	2	DNAAF1	TRUE	TRUE
CCDC78_2	0	2.10471881	0.976	0.112	0	2	CCDC78	TRUE	TRUE
TSPAN1_2	0	2.06975242	0.987	0.307	0	2	TSPAN1	TRUE	TRUE
TPPP3_2	0	2.05687028	1	0.451	0	2	TPPP3	TRUE	TRUE
C5orf49_2	0	2.00471083	0.976	0.14	0	2	C5orf49	TRUE	TRUE
MS4A8_2	0	2.00071039	0.974	0.11	0	2	MS4A8	TRUE	FALSE
C9orf116_2	0	1.97760347	0.976	0.186	0	2	C9orf116	TRUE	TRUE
CAPSL_2	0	1.93462351	0.954	0.104	0	2	CAPSL	TRUE	FALSE
C9orf117_2	0	1.91044654	0.959	0.119	0	2	C9orf117	TRUE	FALSE
CETN2_2	0	1.91015377	0.996	0.595	0	2	CETN2	TRUE	TRUE
SMIM22_2	0	1.90935383	0.967	0.34	0	2	SMIM22	TRUE	TRUE
AC013264.2_2	0	1.89751966	0.887	0.054	0	2	AC013264.2	TRUE	FALSE
MORN5_2	0	1.86440237	0.936	0.085	0	2	MORN5	TRUE	FALSE
CRIP1_2	0	1.85515876	0.995	0.69	0	2	CRIP1	TRUE	TRUE
TSPAN19_2	0	1.84529837	0.846	0.044	0	2	TSPAN19	TRUE	FALSE
TMC5_2	0	1.83626803	0.932	0.207	0	2	TMC5	TRUE	TRUE
MRPS31_2	0	1.82411963	0.955	0.292	0	2	MRPS31	TRUE	TRUE
CDHR3_2	0	1.79976496	0.922	0.093	0	2	CDHR3	TRUE	FALSE
FAM92B_2	0	1.79890798	0.924	0.071	0	2	FAM92B	TRUE	FALSE
MORN2_2	0	1.7979764	0.964	0.31	0	2	MORN2	TRUE	TRUE
TUBB4B_2	0	1.7924057	0.993	0.756	0	2	TUBB4B	FALSE	FALSE
FAM216B_2	0	1.79044925	0.918	0.059	0	2	FAM216B	TRUE	FALSE
CCDC17_2	0	1.76702138	0.889	0.083	0	2	CCDC17	TRUE	TRUE
C1orf192_2	0	1.76682417	0.936	0.09	0	2	C1orf192	TRUE	FALSE
C21orf58_2	0	1.72431372	0.947	0.149	0	2	C21orf58	TRUE	TRUE
PSENEN_2	0	1.72294253	0.996	0.546	0	2	PSENEN	TRUE	TRUE
ZMYND10_2	0	1.70361241	0.954	0.122	0	2	ZMYND10	TRUE	FALSE
C22orf15_2	0	1.69880533	0.903	0.081	0	2	C22orf15	TRUE	FALSE
LRRIQ1_2	0	1.69268795	0.94	0.176	0	2	LRRIQ1	TRUE	TRUE
EFHC1_2	0	1.68783651	0.951	0.33	0	2	EFHC1	TRUE	TRUE
SLC44A4_2	0	1.66453274	0.93	0.236	0	2	SLC44A4	TRUE	TRUE
C11orf88_2	0	1.65273323	0.908	0.085	0	2	C11orf88	TRUE	FALSE
LRRC46_2	0	1.64694474	0.911	0.091	0	2	LRRC46	TRUE	TRUE
DYNLL1_2	0	1.63350928	0.999	0.926	0	2	DYNLL1	TRUE	TRUE
CCDC11_2	0	1.63313982	0.902	0.084	0	2	CCDC11	TRUE	TRUE
DNAH12_2	0	1.62577348	0.905	0.076	0	2	DNAH12	FALSE	FALSE
GRP_3	0	5.12254436	0.802	0.054	0	3	GRP	TRUE	TRUE
CHGA_3	0	2.2460644	0.797	0.008	0	3	CHGA	TRUE	FALSE
ASCL1_3	0	2.20947633	0.722	0.012	0	3	ASCL1	TRUE	FALSE
RPRM_3	0	1.88753288	0.608	0.029	0	3	RPRM	TRUE	TRUE
BEX1_3	0	1.82119532	0.894	0.067	0	3	BEX1	TRUE	TRUE
CALCA_3	0	1.75188975	0.366	0.003	0	3	CALCA	TRUE	FALSE
SCGN_3	0	1.7246277	0.784	0.008	0	3	SCGN	TRUE	FALSE

MIAT_3	0	1.63377791	0.555	0.011	0	3	MIAT	TRUE	FALSE
SST_3	0	1.63044426	0.291	0.002	0	3	SST	TRUE	FALSE
HEPACAM2_3	0	1.58853959	0.683	0.022	0	3	HEPACAM2	TRUE	FALSE
TMEM176B_3	0	1.57636055	0.722	0.058	0	3	TMEM176B	TRUE	FALSE
TMEM176A_3	0	1.55453463	0.736	0.061	0	3	TMEM176A	TRUE	FALSE
SCG5_3	0	1.46024755	0.696	0.04	0	3	SCG5	TRUE	TRUE
UCHL1_3	0	1.42106617	0.731	0.052	0	3	UCHL1	TRUE	TRUE
SLC35D3_3	0	1.29640086	0.533	0.002	0	3	SLC35D3	TRUE	FALSE
CHGB_3	0	1.25389937	0.648	0.004	0	3	CHGB	TRUE	TRUE
DDC_3	0	1.20606402	0.577	0.027	0	3	DDC	TRUE	FALSE
SCG2_3	0	1.16409129	0.537	0.003	0	3	SCG2	TRUE	FALSE
SCG3_3	0	1.08681345	0.674	0.004	0	3	SCG3	TRUE	FALSE
APLP1_3	0	1.05126163	0.599	0.017	0	3	APLP1	TRUE	TRUE
SYT1_3	0	0.92997719	0.476	0.012	0	3	SYT1	FALSE	FALSE
INSM1_3	0	0.9028185	0.502	0.004	0	3	INSM1	FALSE	FALSE
PROX1_3	0	0.8855678	0.533	0.006	0	3	PROX1	FALSE	FALSE
CRMP1_3	0	0.86140559	0.498	0.016	0	3	CRMP1	FALSE	FALSE
PHGR1_3	0	0.81953713	0.37	0.004	0	3	PHGR1	FALSE	FALSE
KCNH6_3	0	0.79195028	0.493	0.018	0	3	KCNH6	FALSE	FALSE
DUSP26_3	0	0.78443713	0.48	0.021	0	3	DUSP26	FALSE	FALSE
PGAM2_3	0	0.78196155	0.432	0.012	0	3	PGAM2	FALSE	FALSE
SNAP25_3	0	0.7415359	0.48	0.008	0	3	SNAP25	FALSE	FALSE
CPLX2_3	0	0.65443638	0.339	0.002	0	3	CPLX2	FALSE	FALSE
RGS4_3	0	0.61027419	0.278	0.002	0	3	RGS4	FALSE	FALSE
PCSK2_3	0	0.60208333	0.352	0.002	0	3	PCSK2	FALSE	FALSE
NRXN1_3	0	0.57320811	0.317	0.007	0	3	NRXN1	FALSE	FALSE
SYT4_3	0	0.53180788	0.322	0.001	0	3	SYT4	FALSE	FALSE
CADPS_3	0	0.50905891	0.33	0.003	0	3	CADPS	FALSE	FALSE
SCN3A_3	0	0.48666541	0.291	0.002	0	3	SCN3A	FALSE	FALSE
PTPRN_3	0	0.40832256	0.251	0.002	0	3	PTPRN	FALSE	FALSE
SRRM4_3	0	0.40348161	0.269	0.002	0	3	SRRM4	FALSE	FALSE
GFRA3_3	2.25E-294	1.08080704	0.498	0.027	5.36E-290	3	GFRA3	TRUE	TRUE
RUNDC3A_3	7.10E-293	0.44322078	0.295	0.006	1.69E-288	3	RUNDC3A	FALSE	FALSE
TCERG1L_3	1.93E-286	0.44772773	0.269	0.005	4.60E-282	3	TCERG1L	FALSE	FALSE
KIF1A_3	5.13E-279	0.48462255	0.304	0.008	1.22E-274	3	KIF1A	FALSE	FALSE
ELAVL4_3	1.93E-278	0.46358186	0.304	0.008	4.60E-274	3	ELAVL4	FALSE	FALSE
DNAJC12_3	2.38E-269	0.83436989	0.542	0.037	5.67E-265	3	DNAJC12	FALSE	FALSE
RAB3B_3	4.79E-262	0.55849152	0.339	0.012	1.14E-257	3	RAB3B	FALSE	FALSE
NOL4_3	8.19E-259	0.57671492	0.352	0.013	1.95E-254	3	NOL4	FALSE	FALSE
TAGLN3_3	2.66E-255	1.20717791	0.648	0.061	6.33E-251	3	TAGLN3	TRUE	TRUE
VWA5B2_3	4.64E-250	0.39992624	0.264	0.006	1.11E-245	3	VWA5B2	FALSE	FALSE
NEURL1_3	6.88E-250	0.49181759	0.335	0.012	1.64E-245	3	NEURL1	FALSE	FALSE
PCSK1N_3	4.57E-248	3.28073033	1	0.229	1.09E-243	3	PCSK1N	TRUE	TRUE
FIBCD1_3	2.35E-237	0.54734827	0.269	0.007	5.59E-233	3	FIBCD1	FALSE	FALSE
ESPL1_3	4.00E-235	0.70866469	0.357	0.016	9.52E-231	3	ESPL1	FALSE	FALSE
FAM105A_3	1.70E-234	1.00087773	0.577	0.052	4.05E-230	3	FAM105A	TRUE	TRUE
SCGB3A2_4	0	3.58373631	1	0.943	0	4	SCGB3A2	TRUE	TRUE
SFTPB_4	0	2.26107455	0.893	0.367	0	4	SFTPB	TRUE	TRUE
CYB5A_4	0	1.49788416	0.997	0.87	0	4	CYB5A	TRUE	TRUE
VIM_4	0	1.48206839	0.98	0.616	0	4	VIM	TRUE	TRUE
LMO3_4	0	0.70327641	0.593	0.123	0	4	LMO3	TRUE	TRUE
FMO2_4	0	0.648367	0.431	0.059	0	4	FMO2	TRUE	FALSE
AARD_4	0	0.58196317	0.34	0.027	0	4	AARD	TRUE	TRUE
MGST1_4	3.55E-303	0.87898279	0.979	0.807	8.45E-299	4	MGST1	TRUE	TRUE
CFH_4	3.53E-300	0.83167136	0.765	0.287	8.42E-296	4	CFH	TRUE	TRUE
HOPX_4	3.23E-290	0.57736132	0.832	0.345	7.70E-286	4	HOPX	TRUE	TRUE

TMSB4X_4	1.32E-272	0.74959897	1	0.999	3.14E-268	4	TMSB4X	TRUE	TRUE
CFHR1_4	7.37E-264	0.79943662	0.392	0.058	1.76E-259	4	CFHR1	TRUE	TRUE
C16orf89_4	3.42E-253	1.02186514	0.809	0.415	8.14E-249	4	C16orf89	TRUE	TRUE
CYTL1_4	2.50E-249	1.00143142	0.422	0.08	5.97E-245	4	CYTL1	TRUE	TRUE
CFTR_4	9.69E-241	0.46924698	0.414	0.077	2.31E-236	4	CFTR	TRUE	TRUE
HNMT_4	5.62E-236	0.68229595	0.796	0.385	1.34E-231	4	HNMT	TRUE	TRUE
CXCL17_4	2.62E-229	0.82479473	0.916	0.508	6.24E-225	4	CXCL17	TRUE	TRUE
CLU_4	5.47E-224	1.32071636	0.917	0.675	1.30E-219	4	CLU	TRUE	TRUE
SSR4_4	4.92E-216	0.78459699	0.976	0.917	1.17E-211	4	SSR4	TRUE	TRUE
C3_4	3.25E-213	0.6947927	0.357	0.066	7.76E-209	4	C3	TRUE	TRUE
PTP4A1_4	2.26E-204	0.72219938	0.842	0.511	5.38E-200	4	PTP4A1	TRUE	TRUE
SERPINF1_4	2.12E-202	0.69767624	0.755	0.357	5.06E-198	4	SERPINF1	TRUE	TRUE
GDF15_4	1.80E-199	0.71570517	0.685	0.247	4.30E-195	4	GDF15	TRUE	TRUE
RNU12_4	6.43E-179	0.90295912	0.726	0.349	1.53E-174	4	RNU12	TRUE	TRUE
KLHL24_4	3.23E-175	0.5434368	0.622	0.265	7.70E-171	4	KLHL24	TRUE	TRUE
SPINK1_4	6.76E-174	1.15797602	0.257	0.036	1.61E-169	4	SPINK1	TRUE	TRUE
DSTN_4	1.82E-172	0.47025915	0.993	0.968	4.34E-168	4	DSTN	TRUE	TRUE
SFTA2_4	2.06E-168	0.56440449	0.603	0.219	4.91E-164	4	SFTA2	TRUE	TRUE
RNASE1_4	4.40E-167	0.70427323	0.885	0.543	1.05E-162	4	RNASE1	TRUE	TRUE
PTP4A2_4	3.46E-164	0.5879702	0.767	0.466	8.24E-160	4	PTP4A2	TRUE	TRUE
XBP1_4	1.58E-161	0.72876469	0.906	0.687	3.77E-157	4	XBP1	TRUE	TRUE
STEAP4_4	6.03E-161	0.57689654	0.421	0.127	1.44E-156	4	STEAP4	TRUE	TRUE
IFITM3_4	8.87E-156	0.62381359	0.926	0.715	2.11E-151	4	IFITM3	TRUE	TRUE
ERP27_4	1.26E-135	0.32360054	0.319	0.076	2.99E-131	4	ERP27	FALSE	FALSE
RAMP2_4	2.40E-130	0.50246633	0.556	0.249	5.72E-126	4	RAMP2	TRUE	TRUE
IFITM2_4	1.60E-129	0.55800062	0.672	0.353	3.81E-125	4	IFITM2	TRUE	TRUE
SLC4A4_4	4.52E-129	0.43883901	0.383	0.121	1.08E-124	4	SLC4A4	TRUE	TRUE
PCDH20_4	9.42E-123	0.29314398	0.315	0.082	2.24E-118	4	PCDH20	FALSE	FALSE
SERPINI1_4	1.29E-120	0.53633143	0.327	0.09	3.06E-116	4	SERPINI1	TRUE	TRUE
SCUBE2_4	1.54E-119	0.29549298	0.28	0.067	3.66E-115	4	SCUBE2	FALSE	FALSE
HPGD_4	3.40E-119	0.2894507	0.266	0.061	8.11E-115	4	HPGD	FALSE	FALSE
NDRG1_4	9.20E-119	0.503458	0.745	0.418	2.19E-114	4	NDRG1	TRUE	TRUE
CP_4	3.44E-116	0.64547262	0.555	0.26	8.20E-112	4	CP	TRUE	TRUE
MET_4	3.96E-116	0.3349332	0.397	0.138	9.44E-112	4	MET	FALSE	FALSE
CHN1_4	5.48E-109	0.32081956	0.38	0.127	1.31E-104	4	CHN1	FALSE	FALSE
ASAH1_4	1.83E-107	0.4261494	0.921	0.8	4.37E-103	4	ASAH1	FALSE	FALSE
BTG1_4	7.45E-103	0.39818478	0.974	0.873	1.78E-98	4	BTG1	TRUE	TRUE
KIAA1324_4	1.07E-99	0.35851957	0.449	0.183	2.55E-95	4	KIAA1324	TRUE	TRUE
TSPAN4_4	1.56E-99	0.39150224	0.618	0.321	3.71E-95	4	TSPAN4	FALSE	FALSE
SERPINA1_4	2.36E-98	0.51185794	0.488	0.222	5.63E-94	4	SERPINA1	TRUE	TRUE
TMEM45A_4	6.58E-98	0.64403419	0.416	0.173	1.57E-93	4	TMEM45A	TRUE	TRUE
SFTPC_5	0	2.89124833	0.968	0.192	0	5	SFTPC	TRUE	TRUE
NPC2_5	0	1.85636605	1	0.867	0	5	NPC2	TRUE	TRUE
TESC_5	0	1.47438499	0.792	0.044	0	5	TESC	TRUE	TRUE
CA2_5	0	1.38980322	0.829	0.051	0	5	CA2	TRUE	TRUE
ETV5_5	0	1.20192973	0.836	0.128	0	5	ETV5	TRUE	TRUE
ASPSCR1_5	0	1.16845964	0.811	0.273	0	5	ASPSCR1	TRUE	TRUE

WIF1_5	0	1.11264274	0.931	0.231	0	5	WIF1	TRUE	TRUE
CD36_5	0	1.09121912	0.67	0.11	0	5	CD36	TRUE	TRUE
SFTA3_5	0	1.05536674	0.978	0.466	0	5	SFTA3	TRUE	TRUE
CPM_5	0	1.00907657	0.949	0.321	0	5	CPM	TRUE	TRUE
MFSD2A_5	0	0.99532541	0.623	0.037	0	5	MFSD2A	TRUE	TRUE
TPD52L1_5	0	0.99122817	0.804	0.201	0	5	TPD52L1	TRUE	TRUE
IGFBP7_5	0	0.97800304	0.997	0.494	0	5	IGFBP7	TRUE	TRUE
TPM1_5	0	0.96990191	0.979	0.639	0	5	TPM1	TRUE	TRUE
ATP11A_5	0	0.93339063	0.73	0.172	0	5	ATP11A	TRUE	TRUE
FTL_5	0	0.92654905	1	0.988	0	5	FTL	TRUE	TRUE
ID2_5	0	0.91441823	0.981	0.664	0	5	ID2	TRUE	TRUE
COL9A3_5	0	0.89930173	0.719	0.183	0	5	COL9A3	TRUE	TRUE
BAMBI_5	0	0.89780043	0.8	0.242	0	5	BAMBI	TRUE	TRUE
SLC34A2_5	0	0.87097145	0.711	0.082	0	5	SLC34A2	TRUE	TRUE
TMSB10_5	0	0.77461864	1	0.993	0	5	TMSB10	TRUE	TRUE
ADAMTS1_5	0	0.76571759	0.61	0.084	0	5	ADAMTS1	TRUE	TRUE
MEST_5	0	0.75382921	0.924	0.504	0	5	MEST	TRUE	TRUE
MARCKSL1_5	0	0.75368053	0.989	0.704	0	5	MARCKSL1	TRUE	TRUE
NREP_5	0	0.7534242	0.927	0.455	0	5	NREP	TRUE	TRUE
CLDN6_5	0	0.75318713	0.846	0.217	0	5	CLDN6	TRUE	TRUE
HNRNPA1_5	0	0.73216936	1	0.979	0	5	HNRNPA1	TRUE	TRUE
NOTUM_5	0	0.71839089	0.546	0.025	0	5	NOTUM	TRUE	TRUE
PIK3C2G_5	0	0.70956749	0.552	0.021	0	5	PIK3C2G	TRUE	TRUE
NPM1_5	0	0.66724688	1	0.959	0	5	NPM1	TRUE	TRUE
GSTM3_5	0	0.64606589	0.73	0.202	0	5	GSTM3	TRUE	TRUE
EDNRB_5	0	0.64180624	0.576	0.058	0	5	EDNRB	TRUE	TRUE
CLDN2_5	0	0.64170962	0.472	0.008	0	5	CLDN2	TRUE	TRUE
RPL5_5	0	0.62897469	1	0.989	0	5	RPL5	TRUE	TRUE
RPL6_5	0	0.62304757	1	0.994	0	5	RPL6	TRUE	TRUE
RPS23_5	0	0.61896649	1	0.995	0	5	RPS23	TRUE	TRUE
RPL3_5	0	0.61679242	1	0.999	0	5	RPL3	TRUE	TRUE
RPS3A_5	0	0.60593388	1	0.998	0	5	RPS3A	TRUE	TRUE
HAS3_5	0	0.59385326	0.508	0.053	0	5	HAS3	FALSE	FALSE
RPLP0_5	0	0.59279884	1	0.991	0	5	RPLP0	FALSE	FALSE
EEF1A1_5	0	0.58789751	1	1	0	5	EEF1A1	FALSE	FALSE
RPL11_5	0	0.5818313	1	0.999	0	5	RPL11	FALSE	FALSE
RPS6_5	0	0.55798144	1	0.999	0	5	RPS6	FALSE	FALSE
NAPSA_5	0	0.55454101	0.442	0.055	0	5	NAPSA	FALSE	FALSE
RPL7A_5	0	0.55321381	1	0.994	0	5	RPL7A	FALSE	FALSE
RPS26_5	0	0.5464097	1	0.986	0	5	RPS26	FALSE	FALSE
RPS3_5	0	0.53986449	1	0.995	0	5	RPS3	FALSE	FALSE
RPS7_5	0	0.53930048	1	0.997	0	5	RPS7	FALSE	FALSE
RPSA_5	0	0.53519461	0.999	0.985	0	5	RPSA	FALSE	FALSE
GNB2L1_5	0	0.5326438	1	0.99	0	5	GNB2L1	FALSE	FALSE
RNU12_5	2.56E-44	0.26217499	0.567	0.37	6.09E-40	5	RNU12	TRUE	TRUE
MYL9_6	0	1.56395547	0.965	0.37	0	6	MYL9	TRUE	TRUE
AGER_6	0	1.53124963	0.829	0.117	0	6	AGER	TRUE	TRUE
CLIC3_6	0	1.47903681	0.923	0.256	0	6	CLIC3	TRUE	TRUE
TNNC1_6	0	1.33375539	0.78	0.092	0	6	TNNC1	TRUE	TRUE
CAV1_6	0	1.31579053	0.857	0.226	0	6	CAV1	TRUE	TRUE
SMARCA5_6	0	1.30856222	0.926	0.485	0	6	SMARCA5	TRUE	TRUE
CPM_6	0	1.22737445	0.947	0.327	0	6	CPM	TRUE	TRUE
S100A10_6	0	1.1727792	0.995	0.807	0	6	S100A10	TRUE	TRUE
CLDN6_6	0	1.12082694	0.88	0.217	0	6	CLDN6	TRUE	TRUE
SFTA3_6	0	1.04629292	0.976	0.471	0	6	SFTA3	TRUE	TRUE
ICAM4_6	0	1.03451614	0.648	0.115	0	6	ICAM4	TRUE	TRUE
CLDN18_6	0	0.96864439	0.6	0.055	0	6	CLDN18	TRUE	TRUE
FOLR1_6	0	0.96171091	0.865	0.388	0	6	FOLR1	TRUE	TRUE
SPARC_6	0	0.93249513	0.919	0.415	0	6	SPARC	TRUE	TRUE
AQP4_6	0	0.91397618	0.669	0.14	0	6	AQP4	TRUE	FALSE
EDN3_6	0	0.86685901	0.517	0.039	0	6	EDN3	TRUE	FALSE
NGFRAP1_6	0	0.82048626	0.998	0.842	0	6	NGFRAP1	TRUE	TRUE
NDNF_6	0	0.79289681	0.595	0.076	0	6	NDNF	TRUE	TRUE
PDPN_6	0	0.75436795	0.612	0.128	0	6	PDPN	TRUE	FALSE

TMSB10_6	0	0.72750714	1	0.993	0	6	TMSB10	TRUE	TRUE
PCP4_6	0	0.66690401	0.353	0.033	0	6	PCP4	TRUE	FALSE
ANKRD29_6	0	0.65554806	0.493	0.081	0	6	ANKRD29	TRUE	TRUE
RPS20_6	0	0.55126643	1	0.991	0	6	RPS20	FALSE	FALSE
FTL_6	2.89E-300	0.61467715	1	0.988	6.89E-296	6	FTL	TRUE	TRUE
NREP_6	2.28E-288	0.77294417	0.901	0.464	5.43E-284	6	NREP	TRUE	TRUE
WIF1_6	9.10E-285	0.76374982	0.785	0.263	2.17E-280	6	WIF1	TRUE	TRUE
MXRA8_6	8.62E-281	0.67974875	0.593	0.164	2.06E-276	6	MXRA8	TRUE	TRUE
RPS3_6	3.33E-278	0.50692975	1	0.995	7.94E-274	6	RPS3	FALSE	FALSE
ANXA3_6	3.80E-278	0.81725152	0.784	0.349	9.06E-274	6	ANXA3	TRUE	TRUE
NKX2-1_6	1.88E-277	0.7845488	0.881	0.469	4.48E-273	6	NKX2-1	TRUE	TRUE
MBIP_6	9.34E-271	0.81325936	0.782	0.357	2.23E-266	6	MBIP	TRUE	TRUE
TSPAN13_6	8.45E-264	0.77879959	0.932	0.727	2.01E-259	6	TSPAN13	TRUE	TRUE
TUBB_6	1.43E-263	0.7072166	0.986	0.806	3.41E-259	6	TUBB	TRUE	TRUE
RPL5_6	1.03E-261	0.49396765	1	0.989	2.45E-257	6	RPL5	TRUE	TRUE
CD47_6	1.48E-259	0.83768065	0.888	0.592	3.54E-255	6	CD47	TRUE	TRUE
ZFAS1_6	2.72E-256	0.60205628	0.998	0.899	6.47E-252	6	ZFAS1	FALSE	FALSE
HOPX_6	9.59E-249	0.51696861	0.789	0.347	2.29E-244	6	HOPX	TRUE	TRUE
PTMA_6	1.54E-244	0.40420598	1	0.999	3.67E-240	6	PTMA	FALSE	FALSE
SFTPB_6	5.94E-244	0.40521685	0.867	0.366	1.41E-239	6	SFTPB	TRUE	TRUE
KCNJ15_6	1.86E-243	0.58441984	0.487	0.116	4.44E-239	6	KCNJ15	FALSE	FALSE
NPM1_6	1.24E-237	0.51694443	0.999	0.959	2.96E-233	6	NPM1	TRUE	TRUE
RPL31_6	2.95E-233	0.38891766	1	1	7.03E-229	6	RPL31	FALSE	FALSE
GSTM3_6	1.28E-231	0.64913209	0.64	0.222	3.05E-227	6	GSTM3	TRUE	TRUE
PRSS8_6	2.28E-231	0.65259044	0.822	0.427	5.45E-227	6	PRSS8	TRUE	TRUE
RNASE1_6	6.32E-231	0.70860409	0.919	0.533	1.51E-226	6	RNASE1	TRUE	TRUE
PLD3_6	3.91E-224	0.62746464	0.971	0.786	9.31E-220	6	PLD3	TRUE	TRUE
RPL11_6	4.77E-224	0.3904643	1	0.999	1.14E-219	6	RPL11	FALSE	FALSE
RPL24_6	7.40E-224	0.40910478	1	0.993	1.76E-219	6	RPL24	FALSE	FALSE
RPS8_6	1.13E-222	0.40567751	1	0.999	2.68E-218	6	RPS8	FALSE	FALSE
RPL23A_6	4.18E-222	0.40286342	1	0.998	9.97E-218	6	RPL23A	FALSE	FALSE
KRT15_7	0	2.78622435	0.94	0.264	0	7	KRT15	TRUE	TRUE
KRT5_7	0	2.41139953	0.887	0.145	0	7	KRT5	TRUE	FALSE
S100A2_7	0	2.10777013	0.833	0.141	0	7	S100A2	TRUE	TRUE
AQP3_7	0	1.70026303	0.911	0.479	0	7	AQP3	TRUE	TRUE
MIR205HG_7	0	1.6486529	0.992	0.551	0	7	MIR205HG	TRUE	TRUE
ZFP36_7	0	1.43061227	0.976	0.783	0	7	ZFP36	TRUE	TRUE
TACSTD2_7	0	1.41664233	0.99	0.747	0	7	TACSTD2	TRUE	TRUE
SFN_7	0	1.40318276	0.738	0.148	0	7	SFN	TRUE	TRUE
KRT19_7	0	1.39837285	0.996	0.935	0	7	KRT19	TRUE	TRUE
BHLHE40_7	0	1.32446046	0.79	0.272	0	7	BHLHE40	TRUE	TRUE
JUNB_7	0	1.18686066	0.998	0.944	0	7	JUNB	TRUE	TRUE
KLF5_7	0	1.13600947	0.947	0.549	0	7	KLF5	TRUE	TRUE
CD9_7	0	1.07551327	0.994	0.883	0	7	CD9	TRUE	TRUE
F3_7	0	1.07546729	0.727	0.217	0	7	F3	TRUE	TRUE
HCAR3_7	0	1.0729668	0.46	0.049	0	7	HCAR3	TRUE	FALSE
PERP_7	0	1.06145797	0.996	0.814	0	7	PERP	TRUE	TRUE
CSTA_7	0	1.0360238	0.686	0.132	0	7	CSTA	TRUE	FALSE
HSPB1_7	0	1.03332699	0.991	0.91	0	7	HSPB1	TRUE	TRUE
JUN_7	0	0.97629601	0.996	0.959	0	7	JUN	TRUE	TRUE
HCAR2_7	0	0.9392252	0.557	0.108	0	7	HCAR2	TRUE	FALSE
FOS_7	0	0.89729556	0.999	0.975	0	7	FOS	TRUE	TRUE

CALML3_7	0	0.84576609	0.4	0.049	0	7	CALML3	FALSE	FALSE
ADH7_7	0	0.75662406	0.523	0.038	0	7	ADH7	FALSE	FALSE
IL33_7	0	0.72841561	0.32	0.024	0	7	IL33	FALSE	FALSE
CAPNS2_7	0	0.70780239	0.547	0.077	0	7	CAPNS2	FALSE	FALSE
SERPINB5_7	0	0.66824131	0.467	0.065	0	7	SERPINB5	FALSE	FALSE
PKP1_7	0	0.62912536	0.531	0.081	0	7	PKP1	FALSE	FALSE
TP63_7	0	0.57125379	0.426	0.036	0	7	TP63	FALSE	FALSE
PNCK_7	0	0.54261584	0.435	0.042	0	7	PNCK	FALSE	FALSE
RPLP1_7	0	0.51945165	1	1	0	7	RPLP1	FALSE	FALSE
LSP1_7	1.69E-306	0.90230393	0.683	0.193	4.02E-302	7	LSP1	TRUE	FALSE
IGFBP3_7	5.78E-304	1.53357014	0.675	0.199	1.38E-299	7	IGFBP3	TRUE	TRUE
ARL4D_7	4.38E-294	1.22199073	0.519	0.109	1.04E-289	7	ARL4D	TRUE	TRUE
KRT13_7	2.06E-289	1.78728104	0.656	0.194	4.90E-285	7	KRT13	TRUE	FALSE
EGR1_7	2.57E-289	0.91085906	0.987	0.856	6.13E-285	7	EGR1	TRUE	TRUE
GJB2_7	1.35E-269	0.66611993	0.427	0.07	3.23E-265	7	GJB2	FALSE	FALSE
FOSB_7	1.39E-268	0.92720729	0.978	0.772	3.32E-264	7	FOSB	TRUE	TRUE
DLK2_7	6.42E-266	0.63624818	0.298	0.027	1.53E-261	7	DLK2	FALSE	FALSE
PLP2_7	4.34E-265	0.84701779	0.904	0.568	1.03E-260	7	PLP2	FALSE	FALSE
LMNA_7	1.05E-258	1.2303195	0.931	0.648	2.51E-254	7	LMNA	TRUE	TRUE
IGFBP2_7	1.68E-258	0.88704964	0.998	0.828	4.00E-254	7	IGFBP2	TRUE	TRUE
MYC_7	2.74E-258	1.17973982	0.655	0.21	6.52E-254	7	MYC	TRUE	TRUE
ATF3_7	2.89E-253	1.35875415	0.836	0.461	6.89E-249	7	ATF3	TRUE	TRUE
MPZL2_7	8.94E-248	0.74316162	0.807	0.386	2.13E-243	7	MPZL2	FALSE	FALSE
ZFP36L1_7	6.28E-243	0.80037165	0.987	0.855	1.50E-238	7	ZFP36L1	TRUE	TRUE
IGF2_7	7.41E-234	1.23277282	0.812	0.384	1.77E-229	7	IGF2	TRUE	FALSE
ACKR3_7	1.65E-233	0.63382334	0.487	0.113	3.94E-229	7	ACKR3	FALSE	FALSE
PVRL1_7	1.40E-230	0.5131878	0.496	0.121	3.34E-226	7	PVRL1	FALSE	FALSE
MEG3_7	2.93E-223	0.81566005	0.896	0.465	6.99E-219	7	MEG3	TRUE	TRUE
DUSP1_7	4.57E-222	0.96714536	0.984	0.918	1.09E-217	7	DUSP1	TRUE	TRUE
LTF_8	0	2.2723415	0.743	0.048	0	8	LTF	TRUE	FALSE
TM4SF1_8	0	1.91854803	0.922	0.256	0	8	TM4SF1	TRUE	TRUE
AZGP1_8	0	1.49292964	0.433	0.019	0	8	AZGP1	TRUE	FALSE
TTYH1_8	0	1.34007832	0.771	0.066	0	8	TTYH1	TRUE	FALSE
CALML5_8	0	1.04198615	0.484	0.033	0	8	CALML5	TRUE	FALSE
MIA_8	0	0.9659213	0.619	0.048	0	8	MIA	TRUE	TRUE
CCL28_8	0	0.8741445	0.533	0.037	0	8	CCL28	TRUE	TRUE
KCNN4_8	0	0.69712129	0.569	0.024	0	8	KCNN4	TRUE	TRUE
FOXC1_8	0	0.66960712	0.588	0.073	0	8	FOXC1	TRUE	TRUE
GLYATL2_8	0	0.66734213	0.51	0.037	0	8	GLYATL2	TRUE	FALSE
INHBA_8	0	0.52330288	0.36	0.018	0	8	INHBA	TRUE	FALSE
SLCO1A2_8	0	0.43928552	0.369	0.005	0	8	SLCO1A2	FALSE	FALSE
NOXO1_8	0	0.38165967	0.395	0.027	0	8	NOXO1	FALSE	FALSE
SLC5A1_8	0	0.347502	0.324	0.006	0	8	SLC5A1	FALSE	FALSE
ALDH1A3_8	1.44E-285	0.64587472	0.486	0.058	3.42E-281	8	ALDH1A3	TRUE	FALSE
NDRG2_8	3.42E-283	1.29589666	0.974	0.507	8.16E-279	8	NDRG2	TRUE	TRUE
KRT7_8	4.14E-272	1.53881258	0.952	0.395	9.86E-268	8	KRT7	TRUE	TRUE
TSPAN8_8	6.76E-234	0.60671191	0.526	0.084	1.61E-229	8	TSPAN8	TRUE	TRUE
FABP7_8	1.70E-226	0.55626811	0.366	0.037	4.05E-222	8	FABP7	TRUE	FALSE
PHLDA1_8	2.66E-219	0.97431607	0.598	0.129	6.33E-215	8	PHLDA1	TRUE	TRUE
GMDS_8	3.22E-218	0.79971383	0.716	0.199	7.68E-214	8	GMDS	TRUE	TRUE
SLPI_8	3.32E-207	1.6038844	0.976	0.523	7.90E-203	8	SLPI	TRUE	TRUE

RPL36_8	2.41E-206	0.63377418	1	1	5.74E-202	8	RPL36	TRUE	TRUE
EHF_8	1.54E-203	0.84463812	0.84	0.308	3.67E-199	8	EHF	TRUE	TRUE
BARX2_8	1.64E-202	0.49112921	0.519	0.097	3.91E-198	8	BARX2	FALSE	FALSE
ZG16B_8	3.31E-201	0.64451858	0.498	0.089	7.88E-197	8	ZG16B	TRUE	TRUE
DSC2_8	1.32E-198	0.3980133	0.434	0.066	3.14E-194	8	DSC2	FALSE	FALSE
KRT23_8	4.19E-195	0.46801438	0.316	0.033	9.99E-191	8	KRT23	FALSE	FALSE
SCGB3A1_8	2.02E-192	1.25468424	1	0.621	4.82E-188	8	SCGB3A1	TRUE	TRUE
APIP_8	1.30E-189	1.20473588	0.826	0.352	3.10E-185	8	APIP	TRUE	TRUE
MUC5B_8	9.01E-189	0.49627012	0.552	0.109	2.15E-184	8	MUC5B	TRUE	FALSE
WFDC2_8	7.72E-184	1.08744068	0.991	0.732	1.84E-179	8	WFDC2	TRUE	TRUE
FAM3D_8	1.38E-181	0.91323387	0.769	0.25	3.29E-177	8	FAM3D	TRUE	FALSE
SOX9_8	2.72E-175	0.61743079	0.729	0.219	6.49E-171	8	SOX9	TRUE	TRUE
KIAA1324_8	1.14E-170	0.55713144	0.669	0.186	2.71E-166	8	KIAA1324	TRUE	TRUE
EPHB3_8	9.57E-161	0.49236373	0.529	0.124	2.28E-156	8	EPHB3	TRUE	TRUE
LA16c-380H5.4_8	1.04E-154	0.31980939	0.29	0.035	2.49E-150	8	LA16c-380H5.4	FALSE	FALSE
GABRP_8	2.10E-154	0.50577356	0.471	0.1	5.00E-150	8	GABRP	TRUE	FALSE
FAM46B_8	2.32E-152	0.34141154	0.35	0.054	5.52E-148	8	FAM46B	FALSE	FALSE
SPDEF_8	1.37E-150	0.48802083	0.531	0.133	3.26E-146	8	SPDEF	FALSE	FALSE
RPL34_8	2.77E-140	0.44329294	1	1	6.60E-136	8	RPL34	FALSE	FALSE
SLC12A2_8	3.55E-138	0.9935829	0.767	0.354	8.46E-134	8	SLC12A2	TRUE	TRUE
TCN1_8	6.85E-126	0.86210948	0.362	0.073	1.63E-121	8	TCN1	TRUE	FALSE
PTN_8	1.32E-125	0.63652937	0.769	0.337	3.15E-121	8	PTN	TRUE	TRUE
DEPTOR_8	3.99E-125	0.26006117	0.309	0.051	9.52E-121	8	DEPTOR	FALSE	FALSE
RPS19_8	1.08E-124	0.46546883	1	1	2.57E-120	8	RPS19	FALSE	FALSE
RPL18A_8	2.10E-124	0.46223695	1	0.999	5.01E-120	8	RPL18A	FALSE	FALSE
MYC_8	6.85E-121	0.72026639	0.648	0.236	1.63E-116	8	MYC	TRUE	TRUE
KRT18_8	3.64E-120	0.58297782	0.988	0.862	8.67E-116	8	KRT18	TRUE	TRUE
RPL13_8	7.78E-120	0.40669516	1	1	1.85E-115	8	RPL13	FALSE	FALSE
NPDC1_8	4.78E-119	0.4642124	0.667	0.247	1.14E-114	8	NPDC1	TRUE	TRUE
NTRK2_9	0	1.51874109	0.705	0.024	0	9	NTRK2	TRUE	FALSE
CD70_9	0	0.40879561	0.337	0.003	0	9	CD70	FALSE	FALSE
SAA1_9	8.75E-286	1.35105595	0.305	0.003	2.08E-281	9	SAA1	TRUE	FALSE
BCL2A1_9	2.35E-273	0.88649326	0.411	0.008	5.61E-269	9	BCL2A1	TRUE	FALSE
KREMEN2_9	9.03E-231	1.46945899	0.663	0.032	2.15E-226	9	KREMEN2	TRUE	FALSE
STRA6_9	1.51E-206	0.89593007	0.6	0.028	3.59E-202	9	STRA6	TRUE	TRUE
SOX10_9	2.12E-201	0.48375751	0.442	0.014	5.05E-197	9	SOX10	FALSE	FALSE
BCAN_9	3.38E-199	0.373665	0.274	0.005	8.06E-195	9	BCAN	FALSE	FALSE
LGR6_9	3.10E-198	0.6613304	0.432	0.014	7.38E-194	9	LGR6	FALSE	FALSE
SYNPO2_9	1.08E-188	0.69180573	0.495	0.02	2.57E-184	9	SYNPO2	FALSE	FALSE
CSPG4_9	1.17E-176	0.42629774	0.4	0.014	2.79E-172	9	CSPG4	FALSE	FALSE
MIA_9	6.08E-173	1.78101128	0.853	0.078	1.45E-168	9	MIA	TRUE	TRUE
SERPINA11_9	1.13E-166	0.30773937	0.253	0.005	2.69E-162	9	SERPINA11	FALSE	FALSE
C1QTNF1_9	1.79E-161	0.30665359	0.253	0.005	4.27E-157	9	C1QTNF1	FALSE	FALSE
CXCR4_9	3.54E-153	1.05823399	0.705	0.058	8.43E-149	9	CXCR4	TRUE	TRUE
LAMP5_9	1.25E-	0.26814219	0.263	0.006	2.98E-	9	LAMP5	FALSE	FALSE

	151				147				
EDAR_9	7.02E-151	0.32215202	0.284	0.008	1.67E-146	9	EDAR	FALSE	FALSE
ANO1_9	1.91E-140	0.4805431	0.421	0.02	4.55E-136	9	ANO1	FALSE	FALSE
ABTB2_9	7.57E-136	0.38761687	0.326	0.012	1.81E-131	9	ABTB2	FALSE	FALSE
NPPC_9	1.38E-127	1.37821228	0.642	0.055	3.28E-123	9	NPPC	TRUE	TRUE
SERPINE2_9	1.15E-118	1.64066932	0.884	0.135	2.74E-114	9	SERPINE2	TRUE	TRUE
TMEM158_9	1.88E-118	0.36834685	0.337	0.015	4.49E-114	9	TMEM158	FALSE	FALSE
PDE4B_9	4.91E-115	0.44390578	0.453	0.029	1.17E-110	9	PDE4B	FALSE	FALSE
SPHK1_9	3.56E-113	0.50574021	0.516	0.039	8.48E-109	9	SPHK1	FALSE	FALSE
THY1_9	3.68E-109	0.83471406	0.484	0.036	8.76E-105	9	THY1	TRUE	FALSE
COL4A2_9	1.99E-105	1.2134917	0.863	0.133	4.75E-101	9	COL4A2	TRUE	TRUE
SHISA2_9	4.91E-105	1.70269305	0.821	0.126	1.17E-100	9	SHISA2	TRUE	TRUE
COL4A1_9	1.37E-100	1.07823424	0.716	0.092	3.26E-96	9	COL4A1	TRUE	TRUE
TFAP2A_9	6.61E-97	0.97871009	0.8	0.122	1.57E-92	9	TFAP2A	TRUE	FALSE
GFRA1_9	7.48E-93	0.43915823	0.442	0.034	1.78E-88	9	GFRA1	FALSE	FALSE
FGFR1_9	7.90E-93	1.29009509	0.758	0.115	1.88E-88	9	FGFR1	TRUE	TRUE
RP11-161M6.2_9	1.57E-90	0.34187641	0.389	0.027	3.75E-86	9	RP11-161M6.2	FALSE	FALSE
IL27RA_9	1.03E-88	0.47857303	0.547	0.056	2.45E-84	9	IL27RA	FALSE	FALSE
FABP7_9	1.02E-86	1.22948953	0.526	0.054	2.43E-82	9	FABP7	TRUE	FALSE
TGFA_9	1.71E-79	0.63776545	0.547	0.064	4.08E-75	9	TGFA	FALSE	FALSE
FRZB_9	3.32E-78	0.76834429	0.642	0.086	7.91E-74	9	FRZB	FALSE	FALSE
ADAMTS18_9	7.90E-76	0.34810308	0.305	0.02	1.88E-71	9	ADAMTS18	FALSE	FALSE
IGFBP4_9	1.55E-75	1.33743299	0.895	0.23	3.70E-71	9	IGFBP4	TRUE	TRUE
ETV4_9	2.76E-75	0.36861512	0.442	0.043	6.57E-71	9	ETV4	FALSE	FALSE
NGFR_9	1.17E-73	0.40339171	0.253	0.014	2.79E-69	9	NGFR	FALSE	FALSE
COL9A1_9	1.64E-73	0.39298307	0.442	0.043	3.91E-69	9	COL9A1	FALSE	FALSE
SPOCK2_9	2.06E-72	0.30945599	0.337	0.026	4.91E-68	9	SPOCK2	FALSE	FALSE
WNT10A_9	4.49E-70	0.55499248	0.284	0.019	1.07E-65	9	WNT10A	FALSE	FALSE
FAM101B_9	9.23E-68	0.26971122	0.284	0.02	2.20E-63	9	FAM101B	FALSE	FALSE
LAMA1_9	1.74E-67	0.3076036	0.263	0.017	4.14E-63	9	LAMA1	FALSE	FALSE
NFE2L3_9	3.64E-58	0.45065171	0.484	0.066	8.68E-54	9	NFE2L3	FALSE	FALSE
FRMD4A_9	1.97E-57	0.72885149	0.621	0.116	4.69E-53	9	FRMD4A	FALSE	FALSE
EGR2_9	1.12E-56	0.7034992	0.463	0.062	2.67E-52	9	EGR2	FALSE	FALSE
PTRF_9	1.14E-56	0.79841652	0.747	0.169	2.73E-52	9	PTRF	TRUE	FALSE
COL7A1_9	7.97E-56	0.37494342	0.411	0.048	1.90E-51	9	COL7A1	FALSE	FALSE
PIGR_10	0	2.30973366	0.989	0.103	0	10	PIGR	TRUE	TRUE
LYZ_10	3.27E-304	1.13443724	0.855	0.073	7.80E-300	10	LYZ	TRUE	TRUE
MUC5B_10	7.34E-295	2.37639842	0.994	0.121	1.75E-290	10	MUC5B	TRUE	FALSE
BPIFB1_10	1.55E-239	2.73027206	0.933	0.123	3.69E-235	10	BPIFB1	TRUE	FALSE
FCGBP_10	6.70E-235	3.04488124	0.989	0.172	1.60E-230	10	FCGBP	TRUE	TRUE
CEACAM5_10	1.65E-177	0.4850052	0.369	0.02	3.94E-173	10	CEACAM5	TRUE	FALSE
ZG16B_10	2.81E-172	0.80400303	0.771	0.103	6.71E-168	10	ZG16B	TRUE	TRUE
HCK_10	6.48E-170	0.28789968	0.419	0.027	1.55E-165	10	HCK	FALSE	FALSE
RP11-1143G9.4_10	1.17E-169	0.3276097	0.363	0.02	2.79E-165	10	RP11-1143G9.4	TRUE	TRUE
TSPAN8_10	1.31E-163	0.71356318	0.76	0.1	3.12E-159	10	TSPAN8	TRUE	TRUE
TGM2_10	2.75E-153	0.39530659	0.659	0.078	6.55E-149	10	TGM2	TRUE	FALSE
CAPN8_10	3.14E-150	0.28168538	0.469	0.039	7.48E-146	10	CAPN8	FALSE	FALSE
LCN2_10	1.20E-136	1.27137413	0.911	0.2	2.86E-132	10	LCN2	TRUE	TRUE
CREB3L1_10	4.19E-128	0.51341169	0.715	0.112	9.99E-124	10	CREB3L1	TRUE	TRUE
SFTPA2_10	1.64E-127	2.29510312	0.57	0.075	3.92E-123	10	SFTPA2	TRUE	TRUE
SCGB3A1_10	3.45E-120	3.49864973	1	0.64	8.22E-116	10	SCGB3A1	TRUE	TRUE

CEACAM6_10	8.20E-120	1.15275712	0.793	0.157	1.95E-115	10	CEACAM6	TRUE	TRUE
FAM3D_10	4.03E-111	1.02669613	0.961	0.271	9.61E-107	10	FAM3D	TRUE	FALSE
TFF3_10	4.38E-104	2.55541297	0.994	0.613	1.04E-99	10	TFF3	TRUE	TRUE
SPINK5_10	9.21E-102	0.97255887	0.749	0.159	2.20E-97	10	SPINK5	TRUE	TRUE
BMPR1B_10	3.14E-85	0.28261092	0.609	0.107	7.48E-81	10	BMPR1B	FALSE	FALSE
SLPI_10	6.15E-85	1.37444777	1	0.545	1.47E-80	10	SLPI	TRUE	TRUE
DEFB1_10	1.68E-84	0.45656622	0.486	0.074	4.00E-80	10	DEFB1	TRUE	FALSE
SLC4A4_10	9.67E-84	0.41742343	0.698	0.145	2.30E-79	10	SLC4A4	TRUE	TRUE
MT3_10	1.00E-83	0.46425444	0.654	0.129	2.39E-79	10	MT3	TRUE	FALSE
KLK11_10	2.60E-79	0.69752392	0.905	0.274	6.20E-75	10	KLK11	TRUE	FALSE
ITLN1_10	1.12E-77	0.60229274	0.346	0.042	2.67E-73	10	ITLN1	TRUE	FALSE
LRMP_10	5.76E-69	0.39940109	0.559	0.114	1.37E-64	10	LRMP	TRUE	TRUE
GSTA2_10	6.52E-66	0.33734876	0.413	0.066	1.55E-61	10	GSTA2	TRUE	FALSE
XBP1_10	1.12E-65	0.90335452	1	0.711	2.66E-61	10	XBP1	TRUE	TRUE
CP_10	7.53E-65	0.52435049	0.899	0.288	1.80E-60	10	CP	TRUE	TRUE
S100P_10	1.65E-63	0.47988341	0.816	0.242	3.94E-59	10	S100P	TRUE	FALSE
RP11-294O2.2_10	1.30E-56	0.25330495	0.57	0.128	3.09E-52	10	RP11-294O2.2	FALSE	FALSE
RARRES1_10	1.71E-53	0.38735331	0.508	0.117	4.09E-49	10	RARRES1	TRUE	TRUE
HBM_10	3.79E-53	0.34599084	0.715	0.202	9.03E-49	10	HBM	TRUE	FALSE
HLA-DRA_10	1.69E-52	0.29274622	0.525	0.122	4.04E-48	10	HLA-DRA	FALSE	FALSE
SPDEF_10	4.23E-51	0.2792417	0.598	0.151	1.01E-46	10	SPDEF	FALSE	FALSE
GSTA1_10	4.92E-51	0.45936001	0.799	0.265	1.17E-46	10	GSTA1	TRUE	TRUE
FAM46C_10	1.08E-50	0.26464033	0.659	0.182	2.56E-46	10	FAM46C	FALSE	FALSE
GALNT12_10	5.80E-50	0.29037335	0.62	0.165	1.38E-45	10	GALNT12	FALSE	FALSE
FKBP11_10	4.90E-47	0.3171165	0.676	0.2	1.17E-42	10	FKBP11	FALSE	FALSE
HBA1_10	1.29E-46	0.81365188	0.972	0.531	3.07E-42	10	HBA1	FALSE	FALSE
AHSP_10	1.87E-46	0.3544169	0.654	0.19	4.46E-42	10	AHSP	TRUE	FALSE
WFDC2_10	4.62E-46	0.64676738	1	0.745	1.10E-41	10	WFDC2	TRUE	TRUE
TIMP1_10	9.34E-44	0.54655878	0.905	0.419	2.23E-39	10	TIMP1	TRUE	TRUE
HBA2_10	1.08E-43	0.82430269	0.978	0.585	2.58E-39	10	HBA2	FALSE	FALSE
KIAA1244_10	8.42E-42	0.26634806	0.777	0.259	2.01E-37	10	KIAA1244	FALSE	FALSE
SCGB1A1_10	2.32E-40	1.82199459	0.888	0.378	5.53E-36	10	SCGB1A1	TRUE	FALSE
ANKUB1_11	3.81E-126	0.3604386	0.689	0.063	9.08E-122	11	ANKUB1	FALSE	FALSE
HS3ST6_11	7.18E-114	0.51537689	0.84	0.106	1.71E-109	11	HS3ST6	FALSE	FALSE
C1orf110_11	9.83E-104	0.39103018	0.632	0.065	2.34E-99	11	C1orf110	FALSE	FALSE
MAP6_11	1.31E-94	0.33004876	0.66	0.076	3.13E-90	11	MAP6	FALSE	FALSE
LPPR3_11	1.01E-90	0.31599451	0.604	0.067	2.40E-86	11	LPPR3	FALSE	FALSE
MUC16_11	3.03E-90	0.92704653	0.953	0.195	7.22E-86	11	MUC16	TRUE	FALSE
GSTA2_11	6.11E-90	0.74642823	0.585	0.067	1.46E-85	11	GSTA2	TRUE	FALSE
FRMPD2_11	1.28E-88	0.36175487	0.755	0.1	3.06E-84	11	FRMPD2	FALSE	FALSE
LRRC10B_11	1.85E-84	0.51229831	0.868	0.138	4.40E-80	11	LRRC10B	FALSE	FALSE
OMG_11	8.26E-84	1.08720502	1	0.21	1.97E-79	11	OMG	TRUE	FALSE
SLC7A2_11	1.04E-83	0.60557839	0.906	0.164	2.49E-79	11	SLC7A2	FALSE	FALSE
AC013264.2_11	8.52E-81	0.70275932	0.972	0.178	2.03E-76	11	AC013264.2	TRUE	FALSE
CCDC17_11	1.01E-80	0.76809381	1	0.202	2.40E-76	11	CCDC17	TRUE	TRUE
PSCA_11	1.35E-80	0.41747698	0.67	0.095	3.21E-76	11	PSCA	FALSE	FALSE
DHRS9_11	2.51E-80	0.42400882	0.604	0.078	5.99E-76	11	DHRS9	FALSE	FALSE
ABCA13_11	6.12E-79	0.49007485	0.736	0.112	1.46E-74	11	ABCA13	FALSE	FALSE
IFI27_11	1.32E-78	0.83305761	0.774	0.13	3.16E-74	11	IFI27	TRUE	TRUE
C1orf173_11	5.96E-78	0.59134328	0.934	0.169	1.42E-73	11	C1orf173	FALSE	FALSE
EPB41L4B_11	8.99E-78	0.45763185	0.868	0.157	2.14E-73	11	EPB41L4B	FALSE	FALSE
MB_11	9.88E-78	0.71247038	0.925	0.19	2.35E-73	11	MB	TRUE	FALSE
AK7_11	2.19E-77	0.33654812	0.726	0.105	5.22E-73	11	AK7	FALSE	FALSE
NWD1_11	5.79E-77	0.40936133	0.726	0.109	1.38E-72	11	NWD1	FALSE	FALSE
HRASLS2_11	6.64E-77	0.2615721	0.594	0.075	1.58E-72	11	HRASLS2	FALSE	FALSE
KRT4_11	8.30E-77	1.39860029	0.934	0.226	1.98E-72	11	KRT4	TRUE	FALSE
WI2-1959D15.1_11	1.28E-76	0.37734929	0.717	0.106	3.05E-72	11	WI2-1959D15.1	FALSE	FALSE
TSPAN19_11	1.80E-76	0.39804391	0.934	0.163	4.30E-72	11	TSPAN19	TRUE	FALSE
CCDC113_11	1.97E-76	0.5039917	0.877	0.153	4.69E-72	11	CCDC113	FALSE	FALSE
TCTEX1D1_11	3.45E-76	0.39605323	0.868	0.142	8.23E-72	11	TCTEX1D1	FALSE	FALSE

TEX26_11	4.89E-76	0.44934292	0.811	0.13	1.17E-71	11	TEX26	FALSE	FALSE
CDHR3_11	8.47E-76	0.86957445	0.981	0.216	2.02E-71	11	CDHR3	TRUE	FALSE
DNAH2_11	1.54E-75	0.33425318	0.726	0.107	3.67E-71	11	DNAH2	FALSE	FALSE
CCDC170_11	1.84E-75	0.71906107	0.981	0.191	4.38E-71	11	CCDC170	TRUE	TRUE
C17orf72_11	1.93E-74	0.6167665	0.962	0.193	4.61E-70	11	C17orf72	FALSE	FALSE
FAM154B_11	2.37E-74	0.6860531	0.991	0.202	5.64E-70	11	FAM154B	TRUE	TRUE
KCNE1_11	2.79E-74	0.29713147	0.708	0.102	6.66E-70	11	KCNE1	FALSE	FALSE
C1orf222_11	4.45E-74	0.36647372	0.755	0.117	1.06E-69	11	C1orf222	FALSE	FALSE
EPPIN_11	1.28E-73	0.26889309	0.538	0.064	3.06E-69	11	EPPIN	FALSE	FALSE
KLHL6_11	2.02E-73	0.35903773	0.575	0.074	4.82E-69	11	KLHL6	FALSE	FALSE
FHAD1_11	2.32E-73	0.32005081	0.774	0.121	5.52E-69	11	FHAD1	FALSE	FALSE
VWA3B_11	1.26E-72	0.39482246	0.811	0.134	3.01E-68	11	VWA3B	FALSE	FALSE
CTXN1_11	8.50E-72	0.46786808	0.887	0.169	2.03E-67	11	CTXN1	FALSE	FALSE
DNAH5_11	9.85E-71	0.62037527	0.972	0.205	2.35E-66	11	DNAH5	FALSE	FALSE
FAM179A_11	1.15E-70	0.29440887	0.736	0.114	2.74E-66	11	FAM179A	FALSE	FALSE
DNAH11_11	2.80E-70	0.41495757	0.811	0.139	6.67E-66	11	DNAH11	FALSE	FALSE
TTLL10_11	5.28E-70	0.3754723	0.84	0.146	1.26E-65	11	TTLL10	FALSE	FALSE
FAM216B_11	6.78E-70	0.5527875	0.962	0.187	1.62E-65	11	FAM216B	TRUE	FALSE
KIF19_11	9.90E-70	0.3946941	0.821	0.14	2.36E-65	11	KIF19	FALSE	FALSE
RP11-193M21.1_11	1.06E-69	0.28810368	0.604	0.083	2.52E-65	11	RP11-193M21.1	FALSE	FALSE
CCDC114_11	9.20E-69	0.49343661	0.887	0.17	2.19E-64	11	CCDC114	FALSE	FALSE
VWA3A_11	1.75E-68	0.28752278	0.774	0.127	4.18E-64	11	VWA3A	FALSE	FALSE
SCGB1A1_12	0	2.99153482	0.897	0.329	0	12	SCGB1A1	TRUE	FALSE
KRT4_12	0	2.17887645	0.884	0.159	0	12	KRT4	TRUE	FALSE
WFDC2_12	0	1.9985975	0.974	0.724	0	12	WFDC2	TRUE	TRUE
MSLN_12	0	1.81432583	0.926	0.265	0	12	MSLN	TRUE	FALSE
SERPINB3_12	0	1.66395863	0.518	0.062	0	12	SERPINB3	TRUE	FALSE
IGF2_12	0	1.64468	0.959	0.376	0	12	IGF2	TRUE	FALSE
S100A9_12	0	1.56754837	0.612	0.084	0	12	S100A9	TRUE	FALSE
SLPI_12	0	1.5237345	0.995	0.502	0	12	SLPI	TRUE	TRUE
S100P_12	0	1.41802392	0.771	0.193	0	12	S100P	TRUE	FALSE
ANXA1_12	0	1.31600577	0.999	0.618	0	12	ANXA1	TRUE	TRUE
MUC4_12	0	1.30742658	0.937	0.24	0	12	MUC4	TRUE	FALSE
KRT19_12	0	1.19241685	1	0.936	0	12	KRT19	TRUE	TRUE
UPK1B_12	0	1.17870972	0.818	0.122	0	12	UPK1B	TRUE	TRUE
CYP2F1_12	0	1.16499103	0.782	0.09	0	12	CYP2F1	TRUE	FALSE
S100A6_12	0	1.14580715	0.998	0.959	0	12	S100A6	TRUE	TRUE
CEACAM6_12	0	1.13423241	0.703	0.108	0	12	CEACAM6	TRUE	TRUE
RHOV_12	0	1.10909036	0.859	0.212	0	12	RHOV	TRUE	TRUE
LCN2_12	0	1.10057766	0.725	0.155	0	12	LCN2	TRUE	TRUE
C19orf33_12	0	1.0757202	0.986	0.566	0	12	C19orf33	TRUE	TRUE
APOBEC3A_12	0	1.04508298	0.347	0.028	0	12	APOBEC3A	TRUE	TRUE
FAM3D_12	0	1.03448692	0.911	0.212	0	12	FAM3D	TRUE	FALSE
TNNT3_12	0	0.94680086	0.837	0.232	0	12	TNNT3	TRUE	FALSE
LYPD2_12	0	0.85318093	0.359	0.018	0	12	LYPD2	FALSE	FALSE
LSP1_12	0	0.84675879	0.801	0.189	0	12	LSP1	TRUE	FALSE
ASS1_12	0	0.74108122	0.71	0.148	0	12	ASS1	FALSE	FALSE
SERPINB2_12	0	0.73992253	0.393	0.036	0	12	SERPINB2	FALSE	FALSE
GPR110_12	0	0.71839917	0.678	0.105	0	12	GPR110	FALSE	FALSE
PDE4C_12	0	0.44442377	0.437	0.05	0	12	PDE4C	FALSE	FALSE
TCN1_12	1.33E-306	0.52497604	0.437	0.052	3.17E-302	12	TCN1	TRUE	FALSE
TFF3_12	1.00E-303	1.9364315	0.959	0.582	2.39E-299	12	TFF3	TRUE	TRUE
MUC20_12	6.69E-303	0.65698059	0.739	0.192	1.60E-298	12	MUC20	FALSE	FALSE
TACSTD2_12	1.81E-302	1.08871591	0.999	0.751	4.32E-298	12	TACSTD2	TRUE	TRUE
ELF3_12	1.11E-301	0.9482772	0.999	0.887	2.64E-297	12	ELF3	TRUE	TRUE
GABRP_12	1.96E-296	0.4687569	0.511	0.081	4.68E-292	12	GABRP	TRUE	FALSE
SLC16A9_12	1.52E-290	0.75548512	0.749	0.222	3.61E-286	12	SLC16A9	FALSE	FALSE
KRT7_12	4.12E-290	0.98316438	0.902	0.378	9.83E-286	12	KRT7	TRUE	TRUE
LGALS3_12	6.59E-286	1.09356255	0.975	0.63	1.57E-281	12	LGALS3	TRUE	TRUE
FUT3_12	3.39E-	0.4771	0.429	0.057	8.09E-	12	FUT3	FALSE	FALSE

	283				279				
PLAC8_12	1.21E-279	0.89249802	0.885	0.3	2.89E-275	12	PLAC8	TRUE	TRUE
VMO1_12	7.40E-277	0.9055322	0.367	0.04	1.76E-272	12	VMO1	TRUE	FALSE
C15orf48_12	9.14E-277	0.79577833	0.35	0.035	2.18E-272	12	C15orf48	FALSE	FALSE
AQP3_12	1.79E-276	0.93822094	0.967	0.481	4.26E-272	12	AQP3	TRUE	TRUE
KLF5_12	7.70E-275	0.90605901	0.973	0.554	1.83E-270	12	KLF5	TRUE	TRUE
FABP5_12	5.07E-274	1.35219313	0.898	0.435	1.21E-269	12	FABP5	TRUE	TRUE
CLDN4_12	7.82E-268	1.00879056	0.999	0.948	1.86E-263	12	CLDN4	TRUE	TRUE
A4GALT_12	6.54E-262	0.47242022	0.546	0.108	1.56E-257	12	A4GALT	FALSE	FALSE
CXCL17_12	4.04E-261	0.89358405	0.972	0.516	9.63E-257	12	CXCL17	TRUE	TRUE
CSTA_12	7.76E-261	0.68807422	0.646	0.148	1.85E-256	12	CSTA	TRUE	FALSE
F3_12	4.47E-253	0.75182651	0.745	0.225	1.06E-248	12	F3	TRUE	TRUE
IGFBP3_12	3.84E-252	0.98049788	0.71	0.205	9.15E-248	12	IGFBP3	TRUE	TRUE
SCGB3A1_12	2.82E-248	0.90694372	0.998	0.606	6.73E-244	12	SCGB3A1	TRUE	TRUE
ERN2_12	1.03E-246	0.3956324	0.453	0.076	2.45E-242	12	ERN2	FALSE	FALSE
NDRG1_12	4.80E-246	0.82106999	0.889	0.412	1.14E-241	12	NDRG1	TRUE	TRUE
AMN_12	7.57E-245	0.42291369	0.451	0.077	1.80E-240	12	AMN	FALSE	FALSE

Table 3-6 Differential gene expression analysis between basal cells and bud tip progenitor cells from single cell RNA-sequencing of the developing human lung

Gene_name	P_value	logFC	Prop_1	Prop_2	padj
TACSTD2	0	2.954226067	0.793209877	0.793209877	0
S100A6	0	1.993279634	0.797067901	0.797067901	0
S100A2	0	1.804545845	0.667438272	0.667438272	0
MIR205HG	0	2.241606886	0.794753086	0.794753086	0
TMSB10	0	-1.362325342	0.797067901	0.797067901	0
IGFBP7	0	-1.975431227	0.201388889	0.201388889	0
SCGB3A1	0	2.757911556	0.752314815	0.752314815	0
PERP	0	2.189373409	0.797839506	0.797839506	0
SFTPC	0	-4.180507787	0.015432099	0.015432099	0
AQP3	0	2.212112364	0.729938272	0.729938272	0
ANXA1	0	2.383607533	0.721450617	0.721450617	0
CD9	0	1.870542177	0.796296296	0.796296296	0
KRT5	0	2.403998295	0.710648148	0.710648148	0
WIF1	0	-1.512899438	0.011574074	0.011574074	0
CPM	0	-1.699148131	0.036265432	0.036265432	0
KLF5	0	1.931030881	0.758487654	0.758487654	0
SFTA3	0	-1.769717889	0.18132716	0.18132716	0
NPC2	0	-2.28293615	0.669753086	0.669753086	0
TPPP3	0	-2.395237478	0.188271605	0.188271605	0
KRT15	0	2.864758409	0.75308642	0.75308642	0
KRT19	0	2.64215892	0.797839506	0.797839506	0
SLPI	0	2.189742754	0.702932099	0.702932099	0
FXYD3	0	1.665822911	0.767746914	0.767746914	0
FTL	0	-1.097215616	0.798611111	0.798611111	0
CLU	1.92E-306	-1.547409177	0.223765432	0.223765432	4.10E-302
CLDN6	3.44E-304	-1.136538985	0.017746914	0.017746914	7.32E-300
BHLHE40	1.58E-302	1.518102889	0.632716049	0.632716049	3.37E-298
CA2	7.01E-300	-1.206248651	0.016975309	0.016975309	1.49E-295
C16orf89	7.46E-299	-1.09624034	0.053240741	0.053240741	1.59E-294
CLIC3	6.30E-298	-1.204263192	0.033179012	0.033179012	1.34E-293
JUN	1.85E-295	1.435794623	0.797839506	0.797839506	3.94E-291
IGF2	2.34E-290	1.913454396	0.650462963	0.650462963	4.98E-286
MEG3	2.56E-287	1.318913258	0.717592593	0.717592593	5.46E-283
SFN	2.71E-284	1.221900631	0.591049383	0.591049383	5.77E-280
NGFRAP1	8.50E-283	-1.17903167	0.655092593	0.655092593	1.81E-278

S100A11	1.30E-275	1.156227427	0.797839506	0.797839506	2.77E-271
FOSB	2.47E-275	1.507097843	0.783179012	0.783179012	5.27E-271
ETV5	4.72E-274	-1.139674664	0.064814815	0.064814815	1.01E-269
NEAT1	6.13E-274	1.993822028	0.682098765	0.682098765	1.30E-269
TESC	4.06E-272	-1.227030772	0.033950617	0.033950617	8.65E-268
MYL9	3.61E-271	-1.356866711	0.175925926	0.175925926	7.68E-267
KLF6	6.27E-269	1.575001313	0.729166667	0.729166667	1.33E-264
LMNA	2.47E-268	1.525895963	0.74537037	0.74537037	5.25E-264
SOX2	2.50E-268	1.066210456	0.650462963	0.650462963	5.32E-264
CSTA	3.18E-267	0.996765113	0.549382716	0.549382716	6.76E-263
IGFBP2	8.69E-267	1.226898274	0.799382716	0.799382716	1.85E-262
NREP	6.60E-265	-1.20730377	0.221450617	0.221450617	1.41E-260
TNNT3	6.55E-264	0.999012085	0.547067901	0.547067901	1.39E-259
LSP1	2.16E-260	0.979897428	0.547067901	0.547067901	4.59E-256
CHST9	3.37E-259	0.780130009	0.544753086	0.544753086	7.17E-255
AGR2	5.19E-259	1.387451267	0.74537037	0.74537037	1.11E-254
MARCKSL1	3.09E-254	-1.27348853	0.492283951	0.492283951	6.59E-250
F3	3.59E-254	1.04959529	0.582561728	0.582561728	7.65E-250
RPL11	9.83E-254	-0.431086543	0.800925926	0.800925926	2.09E-249
KRT13	6.31E-251	1.746403707	0.525462963	0.525462963	1.34E-246
ZFP36L1	1.33E-248	1.250479318	0.790895062	0.790895062	2.84E-244
RPL6	2.39E-247	-0.473356356	0.800925926	0.800925926	5.09E-243
RNASE1	7.65E-247	-1.118428786	0.13117284	0.13117284	1.63E-242
ID2	1.20E-246	-1.467167277	0.43904321	0.43904321	2.56E-242
HBA1	5.83E-245	1.990740628	0.652006173	0.652006173	1.24E-240
SLC34A2	3.34E-244	-0.830124271	0.00308642	0.00308642	7.12E-240
BAMBI	2.63E-243	-0.967734279	0.084876543	0.084876543	5.61E-239
TPM1	1.02E-241	-1.246549038	0.494598765	0.494598765	2.18E-237
JUNB	4.61E-239	1.47103199	0.799382716	0.799382716	9.81E-235
RPS23	2.68E-238	-0.451383543	0.800925926	0.800925926	5.70E-234
CLDN4	7.67E-238	1.750402669	0.779320988	0.779320988	1.63E-233
FOS	2.06E-237	1.201191254	0.800154321	0.800154321	4.38E-233
TNFSF10	3.91E-237	0.769097143	0.514660494	0.514660494	8.33E-233
SFTPB	1.80E-236	-1.151318339	0.054012346	0.054012346	3.84E-232
KRT8	2.84E-236	1.172920277	0.790895062	0.790895062	6.05E-232
RPL5	6.70E-236	-0.487754359	0.800925926	0.800925926	1.43E-231
MEST	1.23E-234	-1.123886944	0.293209877	0.293209877	2.61E-230
SCGB1A1	6.29E-234	0.909117854	0.520833333	0.520833333	1.34E-229
CST3	7.21E-233	-1.056825004	0.627314815	0.627314815	1.53E-228
HBG1	8.97E-233	1.984319685	0.649691358	0.649691358	1.91E-228

TPD52L1	2.78E-231	-0.916151157	0.116512346	0.116512346	5.92E-227
RPS26	5.00E-231	-1.01947815	0.778549383	0.778549383	1.06E-226
ZFP36	7.63E-231	1.781828555	0.781635802	0.781635802	1.63E-226
HBA2	3.66E-229	2.259483405	0.665123457	0.665123457	7.80E-225
ATF3	3.05E-227	1.562437853	0.669753086	0.669753086	6.49E-223
CXCL17	1.40E-226	0.944761749	0.564814815	0.564814815	2.99E-222
HBG2	3.65E-226	2.470810002	0.688271605	0.688271605	7.77E-222
FOLR1	3.42E-225	-0.795506457	0.038580247	0.038580247	7.28E-221
EEF1A1	3.86E-225	-0.443800271	0.800925926	0.800925926	8.22E-221
IGFBP3	3.06E-224	1.409667248	0.540895062	0.540895062	6.51E-220
RPL3	2.10E-222	-0.388064854	0.800925926	0.800925926	4.47E-218
ATP11A	5.99E-222	-0.900106537	0.06404321	0.06404321	1.28E-217
MT-CO1	6.88E-221	1.242145542	0.800925926	0.800925926	1.46E-216
LMO4	1.42E-220	1.060452264	0.722222222	0.722222222	3.02E-216
HBB	3.74E-220	1.004570371	0.535493827	0.535493827	7.96E-216
SSR2	1.93E-219	-0.817028209	0.696759259	0.696759259	4.10E-215
PLP2	8.63E-218	1.009753316	0.723765432	0.723765432	1.84E-213
MYC	6.17E-214	1.044268053	0.524691358	0.524691358	1.31E-209
ASPSCR1	5.71E-213	-1.017551503	0.174382716	0.174382716	1.22E-208
TFF3	8.49E-211	0.989149434	0.558641975	0.558641975	1.81E-206
ELF3	1.39E-207	1.406107855	0.768518519	0.768518519	2.96E-203
KLK11	3.52E-206	0.618306247	0.453703704	0.453703704	7.48E-202
HCAR2	2.80E-204	0.717905679	0.445987654	0.445987654	5.95E-200
MT-CO2	1.08E-203	0.815814608	0.800154321	0.800154321	2.29E-199
EGR1	2.09E-203	1.036991591	0.790123457	0.790123457	4.44E-199
GSTM3	4.64E-203	-0.747922131	0.067901235	0.067901235	9.88E-199
COL9A3	5.78E-203	-0.898159334	0.067901235	0.067901235	1.23E-198
PPP1R15A	1.11E-202	1.057084648	0.735339506	0.735339506	2.37E-198
RPS20	1.30E-202	-0.49551234	0.799382716	0.799382716	2.76E-198
EVL	2.77E-202	-0.877919882	0.233024691	0.233024691	5.89E-198
CAPNS2	5.69E-199	0.58900242	0.438271605	0.438271605	1.21E-194
MDK	4.13E-197	-0.900688914	0.763888889	0.763888889	8.78E-193
MUC4	4.71E-197	0.706123974	0.43595679	0.43595679	1.00E-192
PPAP2C	4.86E-197	0.656815147	0.517746914	0.517746914	1.04E-192
CLDN10	1.90E-195	0.60887356	0.43904321	0.43904321	4.05E-191
RPLP1	7.88E-194	0.348149426	0.800925926	0.800925926	1.68E-189
RPS3A	1.04E-192	-0.375682009	0.800925926	0.800925926	2.21E-188
HSP90AB1	1.14E-191	-0.768362741	0.765432099	0.765432099	2.43E-187
PKP1	1.22E-191	0.545113809	0.425154321	0.425154321	2.59E-187
MFSD2A	7.80E-191	-0.760492869	0.020061728	0.020061728	1.66E-186

ADH7	1.17E-189	0.573099979	0.418981481	0.418981481	2.50E-185
RPL13A	8.81E-189	-0.304285437	0.800925926	0.800925926	1.88E-184
PLAC8	2.01E-188	0.704602321	0.429012346	0.429012346	4.27E-184
HES4	9.55E-188	0.794233	0.554012346	0.554012346	2.03E-183
MT-CO3	2.54E-186	0.945281445	0.800925926	0.800925926	5.40E-182
C19orf33	3.13E-185	1.033543062	0.617283951	0.617283951	6.67E-181
MT-ND3	6.51E-185	0.774241869	0.798611111	0.798611111	1.39E-180
ANXA2	7.89E-185	0.881433511	0.793981481	0.793981481	1.68E-180
PLD3	1.29E-184	-0.865535176	0.604166667	0.604166667	2.75E-180
FABP5	3.50E-184	1.214890092	0.635802469	0.635802469	7.46E-180
LDHB	2.55E-183	-0.747671513	0.770061728	0.770061728	5.43E-179
RPL31	1.78E-182	-0.375866376	0.800925926	0.800925926	3.79E-178
EDNRB	6.11E-182	-0.571247893	0.002314815	0.002314815	1.30E-177
HS3ST1	4.58E-181	0.639924022	0.452932099	0.452932099	9.76E-177
RPL35A	5.55E-181	-0.349362373	0.800925926	0.800925926	1.18E-176
TSPO	8.31E-181	0.802057869	0.75308642	0.75308642	1.77E-176
NPM1	4.20E-180	-0.547765278	0.796296296	0.796296296	8.94E-176
NDRG1	5.46E-180	0.821415596	0.577160494	0.577160494	1.16E-175
TSC22D1	7.24E-180	0.990787667	0.763888889	0.763888889	1.54E-175
RPLP0	1.72E-178	-0.374047293	0.800925926	0.800925926	3.67E-174
EHF	1.08E-177	0.583934141	0.449074074	0.449074074	2.30E-173
SFTA2	2.78E-177	-0.611829574	0.019290123	0.019290123	5.92E-173
CD36	5.96E-177	-0.863121013	0.084104938	0.084104938	1.27E-172
CDKN1A	7.11E-177	0.613776579	0.422067901	0.422067901	1.51E-172
HPCAL1	1.51E-176	-0.630116486	0.071759259	0.071759259	3.22E-172
H3F3A	6.50E-175	-0.448224172	0.800154321	0.800154321	1.38E-170
RPL23A	2.73E-174	-0.344579535	0.800925926	0.800925926	5.82E-170
HNRNPA1	7.96E-174	-0.523039507	0.800925926	0.800925926	1.70E-169
CPE	2.29E-173	-0.80079926	0.212962963	0.212962963	4.87E-169
MT-ND4	1.04E-171	0.772773441	0.800925926	0.800925926	2.22E-167
RHOV	1.91E-171	0.660592203	0.397376543	0.397376543	4.06E-167
ALDOA	4.58E-171	0.769505602	0.777777778	0.777777778	9.75E-167
SIX1	7.57E-171	0.499244711	0.395061728	0.395061728	1.61E-166
DUSP1	1.46E-170	1.235343407	0.787808642	0.787808642	3.11E-166
PIK3C2G	1.28E-169	-0.539247753	0.00462963	0.00462963	2.73E-165
NOTUM	3.15E-169	-0.564937893	0.00308642	0.00308642	6.70E-165
HSPB1	3.97E-169	0.768434479	0.793981481	0.793981481	8.46E-165
TRIM29	2.32E-168	0.474830432	0.38117284	0.38117284	4.93E-164
MT-ATP6	3.60E-168	0.859538603	0.799382716	0.799382716	7.66E-164
RPS3	1.86E-167	-0.32117962	0.800925926	0.800925926	3.97E-163

NCOA7	6.89E-167	1.036205345	0.533950617	0.533950617	1.47E-162
ADAMTS1	2.62E-166	-0.63877012	0.036265432	0.036265432	5.57E-162
PITX1	2.96E-166	0.475501815	0.378858025	0.378858025	6.30E-162
ACKR3	7.06E-166	0.547270911	0.390432099	0.390432099	1.50E-161
LGALS3	8.32E-166	0.847957702	0.611882716	0.611882716	1.77E-161
ENO1	1.36E-165	-0.796175734	0.689814815	0.689814815	2.90E-161
UBC	2.54E-165	0.887108584	0.794753086	0.794753086	5.40E-161
SERPINB5	1.15E-164	0.520202034	0.374228395	0.374228395	2.44E-160
SYTL1	6.52E-163	0.648139411	0.539351852	0.539351852	1.39E-158
RPL7A	1.33E-162	-0.406015167	0.800154321	0.800154321	2.83E-158
SOCS3	1.67E-162	1.245549823	0.578703704	0.578703704	3.55E-158
RND3	1.96E-162	0.830876856	0.570987654	0.570987654	4.16E-158
EPAS1	3.96E-162	0.541846841	0.4375	0.4375	8.44E-158
KLF4	1.53E-161	0.762840933	0.49691358	0.49691358	3.26E-157
HCAR3	9.89E-161	0.648531274	0.368055556	0.368055556	2.10E-156
TKT	1.59E-160	-0.822148211	0.472222222	0.472222222	3.38E-156
RPS27A	3.24E-160	-0.338447409	0.800925926	0.800925926	6.91E-156
C5orf38	6.92E-160	-0.580955507	0.06712963	0.06712963	1.47E-155
SPARC	2.24E-158	-0.841350167	0.202160494	0.202160494	4.76E-154
YBX3	3.57E-158	0.752940355	0.678240741	0.678240741	7.60E-154
UPK1B	1.79E-157	0.586365711	0.364969136	0.364969136	3.81E-153
FAM84A	1.99E-157	0.514799789	0.400462963	0.400462963	4.24E-153
VMP1	6.93E-157	0.95278733	0.644290123	0.644290123	1.47E-152
NHSL1	7.10E-157	-0.667057173	0.11882716	0.11882716	1.51E-152
RPL24	4.53E-156	-0.387741544	0.800925926	0.800925926	9.65E-152
RPS7	3.01E-155	-0.339915838	0.800925926	0.800925926	6.40E-151
RUNX1	3.83E-155	0.523009628	0.456790123	0.456790123	8.16E-151
CKB	5.85E-155	-0.876177959	0.452932099	0.452932099	1.24E-150
SAT1	5.16E-153	1.011764253	0.787037037	0.787037037	1.10E-148
ID3	5.66E-153	-0.960545512	0.249228395	0.249228395	1.21E-148
RPS6	4.05E-152	-0.307417681	0.800925926	0.800925926	8.62E-148
PAX9	5.59E-152	0.424329288	0.353395062	0.353395062	1.19E-147
CTSH	8.62E-152	-0.72415463	0.253858025	0.253858025	1.83E-147
RPL18	1.16E-150	-0.339111009	0.800925926	0.800925926	2.46E-146
ODC1	1.69E-150	-0.760009178	0.283950617	0.283950617	3.60E-146
BTG2	3.84E-150	1.095141683	0.691358025	0.691358025	8.17E-146
ADAM28	6.78E-150	0.468225234	0.361111111	0.361111111	1.44E-145
CXCL1	1.12E-149	0.851979649	0.359567901	0.359567901	2.38E-145
PNCK	3.42E-148	0.418849978	0.348765432	0.348765432	7.29E-144
RPL23	4.60E-148	-0.420938237	0.797067901	0.797067901	9.79E-144

HBM	1.64E-147	0.464601415	0.367283951	0.367283951	3.49E-143
MIR24-2	2.05E-147	0.595843952	0.37962963	0.37962963	4.37E-143
LYPD6B	4.41E-147	0.451977119	0.394290123	0.394290123	9.38E-143
PRSS8	1.13E-145	-0.690626948	0.252314815	0.252314815	2.41E-141
TP63	1.44E-145	0.420611694	0.341049383	0.341049383	3.06E-141
MBIP	6.58E-145	-0.672117563	0.236882716	0.236882716	1.40E-140
GSN	4.92E-144	0.640681515	0.522376543	0.522376543	1.05E-139
EPCAM	3.59E-142	-0.65679932	0.689814815	0.689814815	7.65E-138
SERPINF1	5.09E-142	0.795841555	0.537808642	0.537808642	1.08E-137
CD63	1.01E-141	-0.60402585	0.765432099	0.765432099	2.15E-137
PMAIP1	6.72E-141	0.517463013	0.369598765	0.369598765	1.43E-136
RPL10	9.64E-141	-0.271188531	0.800925926	0.800925926	2.05E-136
CLDN2	1.22E-140	-0.469987461	0.00154321	0.00154321	2.60E-136
NUPR1	1.82E-140	0.841535106	0.525462963	0.525462963	3.87E-136
TMEM123	2.10E-140	0.689883566	0.733024691	0.733024691	4.47E-136
RPS11	2.84E-140	-0.437438518	0.795524691	0.795524691	6.05E-136
TSPAN1	4.92E-140	0.471024119	0.351851852	0.351851852	1.05E-135
MCL1	5.07E-140	0.803941421	0.723765432	0.723765432	1.08E-135
SRP14	5.70E-139	-0.474344886	0.787037037	0.787037037	1.21E-134
DPYSL3	7.08E-139	0.463481685	0.377314815	0.377314815	1.51E-134
NKX2-1	8.35E-138	-0.753158625	0.302469136	0.302469136	1.78E-133
PDLIM1	3.61E-137	0.723691996	0.696759259	0.696759259	7.68E-133
EMP1	8.49E-137	0.633096187	0.362654321	0.362654321	1.81E-132
RPS15A	3.27E-136	-0.300541882	0.800925926	0.800925926	6.96E-132
TUBA1A	5.23E-136	-0.766446374	0.403549383	0.403549383	1.11E-131
MUC1	6.54E-136	-0.599081337	0.135030864	0.135030864	1.39E-131
SNHG7	2.13E-135	-0.706471625	0.445987654	0.445987654	4.54E-131
NME4	2.70E-135	-0.689521903	0.341049383	0.341049383	5.75E-131
FCGRT	1.36E-134	-0.682747159	0.579475309	0.579475309	2.89E-130
AHSP	2.01E-133	0.429805928	0.345679012	0.345679012	4.27E-129
SOX9	3.63E-133	-0.553147318	0.074845679	0.074845679	7.73E-129
TGFB2	7.14E-133	-0.432849817	0.012345679	0.012345679	1.52E-128
A4GALT	8.85E-133	0.358662286	0.317901235	0.317901235	1.88E-128
GJB2	1.05E-132	0.474810467	0.341820988	0.341820988	2.24E-128
MAFB	2.86E-132	0.513199182	0.400462963	0.400462963	6.08E-128
KRT17	1.77E-131	0.585042346	0.347993827	0.347993827	3.76E-127
HSD11B2	2.11E-131	-0.413951667	0.010030864	0.010030864	4.50E-127
PON2	2.65E-131	-0.66341963	0.368055556	0.368055556	5.65E-127
STEAP1	8.81E-131	0.381823813	0.333333333	0.333333333	1.88E-126
ZNF750	3.69E-129	0.367264731	0.30632716	0.30632716	7.85E-125

IRF6	6.34E-129	0.519379274	0.466820988	0.466820988	1.35E-124
ALCAM	1.41E-128	0.649387403	0.621141975	0.621141975	3.00E-124
WFDC2	2.14E-128	0.915044104	0.682098765	0.682098765	4.56E-124
NFE2L2	5.80E-128	0.688470064	0.62191358	0.62191358	1.23E-123
AK1	1.57E-127	-0.524966215	0.098765432	0.098765432	3.34E-123
TMEM59L	3.88E-127	-0.423545004	0.032407407	0.032407407	8.26E-123
DLC1	3.99E-127	-0.4666063	0.05941358	0.05941358	8.49E-123
GLUD1	8.84E-127	0.660706378	0.597222222	0.597222222	1.88E-122
PVRL1	1.11E-126	0.436680566	0.397376543	0.397376543	2.37E-122
RPS8	1.51E-126	-0.291379902	0.800925926	0.800925926	3.21E-122
TNFAIP8	1.56E-126	0.455161849	0.412808642	0.412808642	3.32E-122
SMARCA5	1.58E-126	-0.681978145	0.283179012	0.283179012	3.37E-122
MT-CYB	3.99E-126	0.682710885	0.800925926	0.800925926	8.49E-122
CCNO	4.78E-126	0.396893572	0.319444444	0.319444444	1.02E-121
CRLF1	2.72E-125	-0.450334459	0.027006173	0.027006173	5.80E-121
COTL1	5.11E-125	-0.535830381	0.151234568	0.151234568	1.09E-120
RPL7	1.18E-124	-0.294463973	0.800925926	0.800925926	2.51E-120
NID2	6.69E-123	-0.370905404	0.00462963	0.00462963	1.42E-118
NAPSA	1.86E-122	-0.439030915	0.010030864	0.010030864	3.96E-118
CFI	3.56E-122	-0.43702043	0.050154321	0.050154321	7.57E-118
C9orf3	4.75E-122	0.545598155	0.499228395	0.499228395	1.01E-117
LAMP3	2.56E-121	-0.398467012	0.020833333	0.020833333	5.46E-117
LRRC36	6.97E-121	-0.385973313	0.006944444	0.006944444	1.48E-116
MPZL2	3.66E-120	0.624587589	0.646604938	0.646604938	7.78E-116
GPRC5A	2.72E-119	0.687589171	0.388888889	0.388888889	5.79E-115
CRIP2	2.75E-119	-0.592705425	0.235339506	0.235339506	5.86E-115
SDC2	2.14E-118	-0.506932148	0.102623457	0.102623457	4.57E-114
VIM	2.29E-118	-0.888265417	0.358796296	0.358796296	4.89E-114
MT-ND5	2.87E-118	0.730888632	0.779320988	0.779320988	6.10E-114
STMN1	8.31E-118	-0.817291536	0.633487654	0.633487654	1.77E-113
EDN2	1.22E-117	0.731272031	0.304012346	0.304012346	2.59E-113
GRHL1	3.78E-117	0.326044749	0.287808642	0.287808642	8.06E-113
ITM2B	1.25E-116	0.471972384	0.800925926	0.800925926	2.67E-112
DPYSL2	2.25E-116	-0.431071006	0.06095679	0.06095679	4.78E-112
ST3GAL5	2.28E-116	-0.426777317	0.061728395	0.061728395	4.86E-112
PART1	2.85E-116	0.487684595	0.451388889	0.451388889	6.07E-112
KRT4	3.49E-116	0.478899929	0.295524691	0.295524691	7.43E-112
ALB	7.41E-116	0.368442312	0.322530864	0.322530864	1.58E-111
GATA6-AS1	8.42E-116	-0.34534241	0	0	1.79E-111
BAALC	3.65E-115	0.340488036	0.290123457	0.290123457	7.76E-111

TSHZ2	1.57E-114	0.448028547	0.384259259	0.384259259	3.34E-110
SLC2A1	3.24E-114	0.595298927	0.514660494	0.514660494	6.89E-110
HAS3	5.70E-114	-0.438026244	0.054012346	0.054012346	1.21E-109
RHOC	9.74E-114	0.591276015	0.672839506	0.672839506	2.07E-109
TXNIP	2.29E-113	0.690874393	0.689814815	0.689814815	4.88E-109
SELENBP1	8.54E-113	-0.520888395	0.162037037	0.162037037	1.82E-108
NXPH4	1.25E-112	0.327800755	0.282407407	0.282407407	2.67E-108
LIMCH1	2.34E-112	-0.3594372	0.029320988	0.029320988	4.99E-108
PRSS35	4.59E-112	-0.368987583	0.018518519	0.018518519	9.77E-108
APOA2	1.26E-111	0.331737927	0.297067901	0.297067901	2.68E-107
MT1X	4.32E-111	0.934799078	0.580246914	0.580246914	9.20E-107
FOSL2	8.00E-111	0.340262132	0.294753086	0.294753086	1.70E-106
TACC2	8.77E-111	0.42491595	0.334876543	0.334876543	1.87E-106
TMEM125	1.72E-110	-0.394087526	0.053240741	0.053240741	3.67E-106
RPL21	2.10E-110	-0.257506785	0.800925926	0.800925926	4.46E-106
S100A10	2.39E-110	-0.747799464	0.611882716	0.611882716	5.09E-106
PPIB	2.54E-110	-0.588983564	0.650462963	0.650462963	5.41E-106
PVRL3	2.60E-110	-0.381693849	0.035493827	0.035493827	5.54E-106
PCBD1	4.72E-110	-0.571723037	0.500771605	0.500771605	1.00E-105
MT1G	2.71E-109	0.367352356	0.324074074	0.324074074	5.76E-105
ARL4D	4.28E-109	0.711836255	0.415895062	0.415895062	9.11E-105
SEL1L3	4.66E-109	-0.435331117	0.064814815	0.064814815	9.91E-105
BTG1	8.23E-109	-0.579714709	0.756944444	0.756944444	1.75E-104
MGST1	8.84E-109	0.653289048	0.715277778	0.715277778	1.88E-104

3.7 References

1. Rock JR, Onaitis MW, Rawlins EL, Lu Y, Clark CP, Xue Y, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. Proceedings of the National Academy of Sciences. 2009 Aug 4;106(31):12771–5.

2. Hong KU, Reynolds SD, Watkins S, Fuchs E, Stripp BR. Basal cells are a multipotent progenitor capable of renewing the bronchial epithelium. AJPA. 2004 Feb;164(2):577–88.

3. Weeden CE, Chen Y, Ma SB, Hu Y, Ramm G, Sutherland KD, et al. Lung Basal Stem Cells Rapidly Repair DNA Damage Using the Error-Prone Nonhomologous End-Joining Pathway. PLoS Biol. 2017 Jan;15(1):e2000731.

4. Emily Van de Laar MCSHBRKDWSLJPNMVTKWSKM-STLANM. Cell surface marker profiling of human tracheal basal cells reveals distinct subpopulations, identifies MST1/MSP as a mitogenic signal, and identifies new biomarkers for lung squamous cell carcinomas. Respir Res. BioMed Central; 2014;15(1).

5. Vaughan AE, Brumwell AN, Xi Y, Gotts JE, Brownfield DG, Treutlein B, et al. Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature. 2015 Jan 29;517(7536):621–5.

6. Xi Y, Kim T, Brumwell AN, Driver IH, Wei Y, Tan V, et al. Local lung hypoxia determines epithelial fate decisions during alveolar regeneration. Nat Cell Biol. 2017 Jul 24;503(8):904–14.

7. Zuo W, Zhang T, Wu DZ, Guan SP, Liew A-A, Yamamoto Y, et al. distal airway stem cells are essential for lung regeneration. Nature. Nature Publishing Group; 2015 Jan 20;517(7536):616–20.

8. Ray S, Chiba N, Yao C, Guan X, McConnell AM, Brockway B, et al. Rare SOX2+ Airway Progenitor Cells Generate KRT5+ Cells that Repopulate Damaged Alveolar Parenchyma following Influenza Virus Infection. Stem Cell Reports. 2016 Nov 8;7(5):817–25.

9. Rawlins EL, Clark CP, Xue Y, Hogan BLM. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development. 2009 Nov;136(22):3741–5.

10. Yang Y, Riccio P, Schotsaert M, Mori M, Lü J, Lee D-K, et al. Spatial-Temporal Lineage Restrictions of Embryonic p63+ Progenitors Establish Distinct Stem Cell Pools in Adult Airways. Dev Cell. 2018 Mar 26;44(6):752–4.

11. Miller AJ, Hill DR, Nagy MS, Aoki Y, Dye BR, Chin AM, et al. In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. Stem Cell Reports. 2018 Jan 9;10(1):101–19.

12. Nikolić MZ, Caritg O, Jeng Q, Johnson J-A, Sun D, Howell KJ, et al. Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. Elife. 2017 Jun 30;6.

13. Danopoulos S, Alonso I, Thornton ME, Grubbs BH, Bellusci S, Warburton D, et al. Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9. AJP: Lung Cellular and Molecular Physiology. 2018 Jan 1;314(1):L144–9.

14. Injury Induces Direct Lineage Segregation of Functionally Distinct Airway Basal Stem/Progenitor Cell Subpopulations. Stem Cell. Cell Press; 16(2):184–97.

15. Schnittke N, Herrick DB, Lin B, Peterson J, Coleman JH, Packard AI, et al. Transcription factor p63 controls the reserve status but not the stemness of horizontal basal cells in the olfactory epithelium. Proceedings of the National Academy of Sciences. 2015 Sep 8;112(36):E5068–77.

16. Fletcher RB, Prasol MS, Estrada J, Baudhuin A, Vranizan K, Choi YG, et al. p63 regulates olfactory stem cell self-renewal and differentiation. Neuron. 2011 Dec 8;72(5):748–59.

17. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat M-L, et al. Generation of a functional mammary gland from a single stem cell. Nature. Nature Publishing Group; 2006 Jan 5;439(7072):84.

18. Tumbar T, Guasch G, Greco V, Blanpain C, Lowry WE, Rendl M, et al. Defining the epithelial stem cell niche in skin. Science. 2004 Jan 16;303(5656):359–63.

19. Lu CP, Polak L, Rocha AS, Pasolli HA, Chen S-C, Sharma N, et al. Identification of stem cell populations in sweat glands and ducts reveals roles in homeostasis and wound repair. Cell. 2012 Jul 6;150(1):136–50.

20. Seery JP. Stem cells of the oesophageal epithelium. J Cell Sci. The Company of Biologists Ltd; 2002 May 1;115(9):1783–9.

21. Murata K, Ota S, Niki T, Goto A, Li C-P, Ruriko UMR, et al. p63 - Key molecule in the early phase of epithelial abnormality in idiopathic pulmonary fibrosis. Exp Mol Pathol. 2007 Dec;83(3):367–76.

22. Daniely Y. Critical role of p63 in the development of a normal esophageal and tracheobronchial epithelium. AJP: Cell Physiology. 2004 Feb 25;287(1):C171–81.

23. Arason AJ, Jonsdottir HR, Halldorsson S, Benediktsdottir BE, Bergthorsson JT, Ingthorsson S, et al. deltaNp63 Has a Role in Maintaining Epithelial Integrity in Airway Epithelium. PLoS ONE. Public Library of Science; 2014 Feb 12;9(2):e88683.

24. Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature. Springer US; 2018 Jul 25;:1–30.

25. Plasschaert LW, Žilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature. Nature Publishing Group; 2018 Aug;560(7718):377–81.

26. Rockich BE, Hrycaj SM, Shih HP, Nagy MS, Ferguson MAH, Kopp JL, et al. Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proceedings of the National Academy of Sciences. 2013 Nov 19;110(47):E4456–64.

27. Mou H, Vinarsky V, Tata PR, Brazauskas K, Choi SH, Crooke AK, et al. Dual SMAD Signaling Inhibition Enables Long-Term Expansion of Diverse Epithelial Basal Cells. Cell Stem Cell. 2016 Jun 15.

28. Fischer AJ, Goss KL, Scheetz TE, Wohlford-Lenane CL, Snyder JM, McCray PB. Differential gene expression in human conducting airway surface epithelia and submucosal glands. American Journal of Respiratory Cell and Molecular Biology. American Thoracic Society; 2009 Feb;40(2):189–99.

29. Tata A, Kobayashi Y, Chow RD-W, Tran J, Desai A, Massri AJ, et al. Myoepithelial cells of submucosal glands can function as reserve stem cells to regenerate airways after injury. Cell Stem Cell. NIH Public Access; 2018 May 3;22(5):668–683.e6.

30. Guha A, Vasconcelos M, Cai Y, Yoneda M, Hinds A, Qian J, et al. Neuroepithelial body microenvironment is a niche for a distinct subset of Clara-like precursors in the developing airways. Proceedings of the National Academy of Sciences. 2012 Jul 31;109(31):12592–7.

31. Weinreb C, Wolock S, Klein AM. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. Bioinformatics. 2018 Apr 1;34(7):1246–8.

32. Frank DB, Penkala IJ, Zepp JA, Sivakumar A, Linares-Saldana R, Zacharias WJ, et al. Early lineage specification defines alveolar epithelial ontogeny in the murine lung. Proc Natl Acad Sci USA. National Academy of Sciences; 2019 Feb 19;116(10):4362–71.

33. Balasooriya GI, Goschorska M, Piddini E, Rawlins EL. FGFR2 is required for airway basal cell self-renewal and terminal differentiation. Development. 2017 May 2;144(9):1600–6.

34. Tadokoro T, Gao X, Hong CC, Hotten D, Hogan BLM. BMP signaling and cellular dynamics during regeneration of airway epithelium from basal progenitors. Development. 2016 Mar 1;143(5):764–73.

35. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005 Oct 25;102(43):15545–50.

36. Spence JR, Lange AW, Lin S-CJ, Kaestner KH, Lowy AM, Kim I, et al. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell. 2009 Jul;17(1):62–74.

37. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology. Nature Publishing Group; 2018 Apr 2;36(5):411–7.

38. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011 Jun 1;27(12):1739–40.

39. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems. 2015 Dec;1(6):417–25.

40. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. Nature Publishing Group; 2009;4(1):44–57.

41. Da Wei Huang, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research. Oxford University Press; 2009 Jan 1;37(1):1–13.

Chapter 4 Generation of human pluripotent stem cell-derived bud tip progenitor organoids

Portions of this chapter have been published:

Miller, A.J.; Yu, Q.; Dye B.R..; Ferrer-Torres, D.; Hill, D.R.; Overeem, A.W., Shea, L.D., Spence, J.R. Generation of lung organoids from human pluripotent stem cells in vitro. *Nature Protocols.* **2019**. Doi: 10.1038/s41596-018-0104-8

4.1 Statement of contribution for this chapter within the dissertation

RELATED LINKS

Key reference(s) using this protocol

Miller, A. J. et al. Stem Cell Reports. 10(1), 101-119 (2018)

https://doi.org/10.1016/j.stemcr.2017.11.012

Dye, B. R. et al. eLIFE. 5:e19732 (2016) https://doi.org/10.7554/eLife.19732

Dye, B. R. et al. eLIFE. 4:e05098 (2015) https://doi.10.7554/eLife.05098

This chapter outlines the state of the field for human pluripotent stem-cell (hPSC) derived lung organoids, and provides a detailed protocol for methods to derive lung organoid model systems from hPSCs. hPSC-derived organoids overcome many

challenges in the field. They provide a human cell-based system that may recapitulate human biology better than model organisms such as mice, they do not require donated human tissue, they can be generated from multiple cell lines to provide a wide range of biological diversity, and they can be generated from individual patient cells for use in personalized medicine applications, for example, to test responsiveness to certain drug therapies.

This protocol describes methods to generate two varieties of lung organoids. The first variety, referred to simply as “human lung organoids,” was developed by our lab and published in 2015 (Dye et al. 205 and 2016). The second variety of organoids, called “bud tip progenitor organoids,” built upon the original work by Dye et al. and was developed by me (see chapter 2, Miller et al. 2018) as a part of my thesis work. The publication outlining this protocol was written by me (Miller et al. 2019) and allows hPSC-derived lung organoid model systems to be easily implemented in labs wishing to use them to study lung development and disease.

4.2 Introduction

The lung epithelium is derived from the endodermal germ layer, which undergoes a complex series of endoderm-mesoderm mediated signaling events to generate the final arborized network of conducting airways (bronchi, bronchioles) and gas-exchanging units (alveoli). These stages include endoderm induction, anterior-posterior and dorsal-ventral patterning, lung specification, lung budding, branching morphogenesis and finally, maturation. Here we describe a protocol that recapitulates several of these milestones in order to differentiate human pluripotent stem cells

(hPSCs) into ventral-anterior foregut spheroids and further into two distinct types of organoids; human lung organoids and bud tip progenitor organoids. The resulting human lung organoids possess cell types and structures that resemble the bronchi/bronchioles of the developing human airway surrounded by lung mesenchyme and cells expressing alveolar-cell markers. The bud tip progenitor organoids possess a population of highly proliferative multipotent cells with *in vitro* multi-lineage differentiation potential and *in vivo* engraftment potential. Human lung organoids can be generated from hPSCs in 50-85 days, and bud tip progenitor organoids can be generated in 22 days. The two hPSC-derived models presented here have been benchmarked using human fetal tissue and found to be representative of human fetal-like tissue. The bud tip progenitor organoids are thus ideal for exploring epithelial fate decisions, whilst the human lung organoids can be used to model epithelial-mesenchymal cross-talk during human lung development. In addition to their applications in developmental biology, human lung organoids and bud tip progenitor organoids can be used in research into regenerative medicine, tissue engineering and for pharmaceutical safety and efficacy testing.

4.3 Development of the protocol

During development, the endodermal germ layer gives rise to the lining of the gut tube, which is patterned along the anterior-posterior axis of the embryo into distinct morphological and molecular domains(1). The lung is specified in the ventral anterior foregut endoderm region of the gut tube, and development begins as two primordial lung buds emerge from this region. The lung buds possess a population of multipotent

epithelial progenitors at the tips ('bud tip progenitors'), and are surrounded by mesenchyme. As the lung grows, the epithelium undergoes repeated rounds of bifurcation in a process, known as branching morphogenesis, in order to establish the arborized architecture of the adult lung. During the branching process, bud tip progenitors are maintained, continuously proliferate, and give rise to all lung epithelial cell types. Early during development, branching establishes the network of tubes that will conduct air (bronchi, bronchioles). Later during development, when the branching program is completed, bud tip progenitors that remain at the end of the airways give rise to alveolar epithelial cells, (2,3) as confirmed by lineage tracing experiments in mice(4).

Several studies have demonstrated that recapitulating key stages of embryonic development through a series of steps *in vitro*, known as directed differentiation, is an effective method to generate cell and tissue lineages of interest from human pluripotent stem cells (hPSCs). Directed differentiation has previously been used to generate 3-dimesnional (3D) mid- and hindgut spheroids that give rise to small and large human intestinal organoids (5-8), and has been used to generate 2-dimensional (2D) monolayers of ventral foregut cells and lung epithelial cell types (9-15). Here, we describe protocols based on published work to generate 3D ventral anterior foregut spheroids from hPSCs(15,16), and to differentiate foregut spheroids into two distinct types of lung organoids; human lung organoids(15), and, bud tip progenitor organoids(17) (**Figure 4-1**). Human lung organoids are generated if foregut spheroids are cultured with high levels of FGF10 and 1% FBS and possess airway-like epithelium surrounded by a diffuse network of mesenchymal cells and epithelial cells expressing alveolar cell type markers. The transcriptional profile of these organoids is highly similar

to fetal lung. The presence of mesenchyme and organized airway structures are a strength of this system making it ideal for studying mesenchymal epithelial cross-talk during fetal lung development. Bud tip progenitor organoids are generated when foregut spheroids are cultured in a serum-free environment with FGF7, CHIR-99021 and all trans retinoic acid (ATRA). After 22 days in culture, bud tip progenitor organoids contain a highly enriched and proliferative population of SOX2+/SOX9+/ID2+/NKX2.1+ cells that are transcriptionally similar to human fetal bud tip progenitors. These cells can be expanded in culture for over 16 weeks. Given that bud tip progenitors are a precursor cell to all epithelial cell types during development, bud tip progenitor organoids are uniquely suited to study the mechanisms involved in epithelial cell fate decisions in the developing human lung.

4.4 Overview of the experimental design

Generation of ventral-anterior foregut spheroids takes approximately 10 days. Subsequent differentiation into human lung organoids takes 50-85 days while differentiation into bud tip progenitor organoids takes approximately 22 days (**Figure 4-1**). The protocol begins with hPSCs grown in a monolayer, and differentiation is directed through several stages, including definitive endoderm (4 days) and ventral-anterior foregut spheroids (5-6 days). While cells begin as 2D monolayers of undifferentiated hPSCs, during the specification into ventral anterior foregut tissue, cells self-assemble into small clusters called spheroids and form free-floating structures that detach from the cell monolayer. Foregut spheroids are then encapsulated in a 3D extracellular matrix (ECM) or synthetic hydrogel(18), and are overlaid with specific growth factors

and small molecules to generate human lung organoids (50-85 days) (15), or to generate bud tip progenitor organoids (22 days)(17).

Foregut spheroids are differentiated in serum-free media. If subsequently embedded in a droplet of Matrigel and grown in high levels of FGF10 and 1% FBS, they give rise to human lung organoids, which form airway-like structures and cell types surrounded by mesenchymal populations. Further, some cells expressing AECI and AECII markers are present within the organoids. These organoids contain optimal airway-like tissue between 50-85 days of culture, but airway-like tissue has been observed for over 100 days in the organoids.

In contrast, foregut spheroids embedded in Matrigel and grown in serum-free medium containing FGF7, CHIR-99021 and RA give rise to bud tip progenitor organoids, which are composed of SOX9+/SOX2+ bud tip progenitor-like cells and have a transcriptome similar to bud tip progenitors found in the native human fetal lung during branching morphogenesis. If left unperturbed in Matrigel droplets for several weeks, spheroids treated with FGF7/CHIR-99021/RA will form patterned branch-like structures with interior regions expressing proximal airway markers and distal bud-tip regions. Bud tip progenitors are highly proliferative and can be easily needle-passaged to expand the progenitors as an enriched population. Bud tip progenitor organoids can be serially passaged every 2 weeks and maintain a relatively uniform multipotent population of progenitor cells *in vitro*. Manipulating the *in vitro* growth factor milieu can promote differentiation of bud tip progenitor cells into differentiated lung epithelial cells expressing both airway/bronchiolar cell markers and alveolar cell markers(17).

4.5 Experimental design

A schematic of the protocol is shown in Figure 4-1. The first 9 days of the protocol are the same for both types of organoids described here; hPSCs are first directed to definitive endoderm using Activin A, then directed to anterior foregut endoderm by inhibiting Transforming Growth Factor (TGFβ) and Bone Morphogenic Protein (BMP) signaling using SB-431542 and NOGGIN, respectively, in addition to simultaneously activating WNT (CHIR-99021), Hedgehog (SAG) and Fibroblast Growth Factor 4 (FGF4) pathways. Foregut spheroids will self-assemble and can be transferred into a droplet of Matrigel. After plating in Matrigel, differentiation into different lung lineages is controlled by the growth factor signaling milieu (Figure 4-1). Treating foregut spheroids with serum-free media containing FGF7, CHIR-99021 and All-Trans Retinoic Acid (ATRA) will result in bud tip progenitor organoids that possess bud tip progenitors after 14 days in culture. This population can be easily expanded and maintained by needle passaging, or, if left unpassaged, will give rise to lung epithelial structures with proximal-distal patterning and a population of bud tip progenitors at budded tips after ~40 days in culture. Conversely, treating foregut spheroids with 1% Fetal Bovine Serum and a high concentration of FGF10 for ~50 days will result in human lung organoids which contain fetal airway-like structures with surrounding lung mesenchyme and alveolar progenitor-like cells. Specific mRNA and protein markers identified from human lung development are used to determine whether cells have differentiated as expected.

4.6 Application of the method

Generating self-aggregating foregut spheroids from hPSCs is the first step in this protocol for either type of organoid (Figure 4-1). Depending on the signaling milieu applied to foregut spheroids, the researcher can choose to grow foregut spheroids that will give rise either to human lung organoids or to bud tip progenitor organoids. Human lung organoids and bud tip progenitor organoids have been compared to both human adult and fetal lung, and possess cells that are highly similar to the developing human lung. Therefore, both systems are ideal for studying developmental biology and tissue engineering.

Human lung organoids: These contain airway-like structures surrounded by lung mesenchymal cells and cells that stain positive for AECI and AECII markers. Bulk RNAseq analysis shows that these organoids are more similar to the native human fetal lung than to adult lung(15). Due to the presence of mesenchyme, airway structures and immature alveolar-like epithelial cells, human lung organoids are ideal for studying mesenchymal-epithelial interactions during human lung development and for modeling human fetal lung malformations using patient-specific cell lines or mutations. For example, ongoing work with these organoids is focused on epithelial mesenchymal cross-talk during pulmonary fibrosis and congenital malformations(19) related to Wnt signaling (unpublished). Human lung organoids can be microinjected(20,21) to study the effect of drugs or bacteria in the context of development or disease (Figure 4-2A), can be transplanted on to a bioengineered scaffold to form mature airway structures and cell types (Figure 4-5B-D), or seeded on to a decellularized lung scaffold to generate mature airway cell types (Figure 4-2E). These organoids are ideal for use after ~50-85 days in

culture but are long-lived and stable *in vitro*, so experiments that last several weeks can easily be carried out. 1) Human lung organoids(15): Based on these unique features, the human lung organoid model is ideal for studying epithelial-mesenchymal interactions during lung development, including maturation of alveolar cell types, and may be ideal for modeling infections common to the immature lung, such as RSV.

Bud tip progenitor organoids. These can be generated from hPSCs in 22 days(17). By 22 days, organoids contain a highly enriched population of epithelial bud tip progenitor cells that are similar to bud tip progenitors residing at the tips of branching buds in the native human fetal lung based on transcriptional profiling, protein expression, and functional characterization. These cells are ideally suited to study mechanisms of epithelial cell fate decisions in the developing lung, as bud tip progenitors are obligate precursors to all lung epithelial cell types. Our work has shown that bud tip progenitor organoids have multilineage potential *in vitro*, where they can give rise to airway-like and alveolar-like cells (Figure 4-2G-H), and ongoing studies have used this model system to study lineage decisions as bud tip progenitors give rise to multiple epithelial cell lineages (unpublished). We speculate that this model may also be useful as a predictive pre-clinical model for many human lung diseases(22,23), or for testing pharmaceutical safety during pregnancy(24). Because they are highly expandable, bud tip progenitor organoids can be expanded to accommodate large screens to test pharmaceuticals for potential effects on proliferation, apoptosis, metabolic function and changes in lineage decisions. Large, easily identifiable lumens within these organoids allows for drugs or pathogens to be delivered by microinjection(20,21) or within the surrounding medium (Figure 4-2F). One unique

feature of bud tip progenitor organoid-derived cells is their ability to successfully engraft, proliferate, and repopulate differentiated airway cells in the injured mouse airway(17). Thus the high expandability and multilineage potential of these organoids may result in them being suitable for cell-based regenerative medicine applications, such as reseeding a damaged airway epithelium *in vivo* (Figure 4-2I).

Both human lung organoids and bud tip progenitor organoids. Both organoids can engraft into immunocompromised mice. Human lung organoids are able to engraft into ectopic vascular beds (i.e. epididymal fat pad; Figure 4-2B-D), while bud tip progenitor organoids have been shown to engraft into the injured mouse airway (Figure 4-2I), suggesting potential in regenerative medicine. However, a deeper understanding of these cells is required before they could be used in a clinical setting.

4.7 Comparison to other methods

The protocol described here outlines methods to generate 3D foregut spheroids, which can be further cultured in a 3D environment such as Matrigel to generate: 1) human lung organoids containing lung epithelium and mesenchyme, or 2) a highly expandable population of bud tip progenitor organoids, containing SOX9+/SOX2+ bud tip progenitors that are similar to bud tip progenitors in the native developing fetal lung. Other methods to direct the differentiation of hPSCs to endoderm(25), foregut(13), and lung lineages on a monolayer(11,12,14) and in 3D culture have been described(15,17,26-31). Each of the different 2D and 3D systems described have unique features, and each have strengths and limitations that should be considered to ensure the model system being used matches the scientific question at hand.

The first reports of directed differentiation of hPSCs into a lung lineage utilized monolayer cultures(11,12,32-34) in which hPSCs are directed first to definitive endoderm(25), then to anterior foregut endoderm(13) , and finally to an NKX2.1+ lung lineage.

More recent reports, including our own, have focused on generating 3D lung tissue in attempts to recapitulate more complex elements of a native organ(15,17,26-31). The organoid models described here and from other reports are summarized in Table 4-1. A recent report by Chen *et al.* directed hPSCs differentiation on a monolayer to early NKX2.1+ lung progenitors, which formed clumps of cells that could be dissociated from the monolayer and placed in a droplet of Matrigel with FGF7, FGF10, CHIR-99021 and ATRA, giving rise to 3D structures, termed 'Lung Bud Organoids' (LBOs)(28). A subset of LBOs cultured in Matrigel quickly formed branched structures that could be manipulated by plating density. Epithelial cells expressed markers of goblet cells and AECII cells, although many other cell types (P63+ cells, SCGB3A2+ secretory, and multiciliated cells) were not detected. While mesenchymal cells were present in these cultures, the proportion declined to less than 2% of total cells after prolonged culture in Matrigel. Transplantation of LBOs into the mouse kidney capsule yielded NKX2.1+ epithelial tubes that formed branch-like structures surrounded by lung mesenchyme. Tips of these branch-like structures did not express SOX2 but were positive for the AECII markers SFTPB and SFTPC, suggesting these structures were patterned into alveolar-like structures. The stalks of these tubules contained cells expressing multiciliated and mucous-producing cell markers. Mucous was found within lumen, suggesting functionality of mucous-producing cells. Of note, when LBOs grown

in vitro in Matrigel were treated with RSV virus they exhibited epithelial shedding, characteristic of the physiological lung response(28). This LBO model is thus well suited for experiments exploring the effect of infections or diseases affecting the mucous producing cells or AECIIs of the lung epithelium, such as flu. Further, LBOs could be valuable tools to further explore mechanisms regulating epithelial branching events.

Following directed differentiation to foregut endoderm/lung lineages, several reports have used cell sorting techniques to isolate early lung-lineage progenitors for subsequent 3D culture(26,27,29-31). Gotoh et al., identified Carboxypeptidase M (CPM) as a marker of ventral foregut endoderm(29). 3D culture of sorted CPM+ cells in co-culture with human fetal fibroblasts gave rise to organoids with cells expressing alveolar epithelial cell markers and lamellar bodies. This technique was later built upon by Konishi et al. who reported that CPM+ organoids could be cultured in a commercially available medium specialized for air-liquid interface culture supplemented with DAPT(30). After 42-56 days in culture, the CPM+ cells gave rise to airway-like organoids with functional multiciliated and mucous producing cells. Given the presence of functional multiciliated and mucous-producing cells, this model is ideal for interrogating diseases affecting the airway, such as ciliopathies, cystic fibrosis, or the process of goblet cell hyperplasia following infection. More recently, Yamamoto et al. (2017) demonstrated that CPM+ cells could be differentiated into AECII-like organoids using co-culture with fetal fibroblasts, or in defined medium(35).

Hawkins et al. (2017) described the generation of an NKX2.1 reporter hPSC line (RUES2), which allowed isolation of NKX2.1+ cells. NKX2.1+ cells plated in 3D formed epithelial-only organoid structures(31). Transcriptomic profiling of these cells identified

cell surface markers ($CD47^{hi}CD26^{lo}$) that allow the prospective isolation of NKX2.1+ cells without using reporter cell lines. This method was built upon by Jacob et al. and McCauley et al., showing that modulation of growth factor signaling in these cultures leads to generation of distal (alveolar-like)(26) or proximal (airway-like)(27) organoids that are epithelial only, respectively. Proximal airway organoids have functional multiciliated and mucous-producing cells and have been shown to be able to model some hallmarks of cystic fibrosis using CF-iPSC cell lines. These systems are ideal system for studying airway diseases affecting functional epithelial cells, such as cystic fibrosis and ciliopathies, and are also highly expandable, making them potential candidates for large drug screening platforms.

In addition to hPSC-derived lung organoid models, it should be noted that many more *in vitro* human lung models have been generated from primary human adult or fetal lung tissue. These systems have been thoroughly reviewed elsewhere(36-38) and are the ideal for interrogating lung diseases as samples can be patient specific, and may be critical for personalized-medicine and screening for therapeutic responsiveness.

Several of the models described above have excelled at generating functional and differentiated epithelial cell types that may be uniquely suited to interrogating disease phenotypes. On the other hand the protocols described here generate fetal-like lung tissue that is ideal for studying human lung development. One additional difference between other methods and those described here is that self-aggregating 3D foregut spheroids form spontaneously and do not require cell dissociation, cell sorting or other specialized techniques. Foregut spheroids can be cultured in a Matrigel droplet and further differentiated into two distinct 3D lung organoid models (Figure 4-1).

4.8 Limitations of our protocol

The human lung organoid model (Figure 4-1A) generates fetal-like airway epithelium and surrounding mesenchyme, as well as a population of cells expressing alveolar cell markers. However, like all other 3D organoid models described, transplantation into mouse is required to generate fully mature lung-like structures and cell types. The largest limitation to using this model to study disease and development is that HLOs take ~60 days to form and are not highly expandable, limiting the amount of material available for study. For both protocols, the use of Matrigel or other animal-derived products such as serum (i.e. used during endoderm differentiation) remains a barrier to translation. However given serum free alternatives have been developed for other organoid models(28,39), the protocol should be adaptable to the use of defined serum (dFBS), serum free protocols, or fully synthetic hydrogels which will aid in progress towards translation(18,40). For example, human intestinal organoid protocols(5) , which were initially developed using serum now have commercially available serum free alternatives (i.e. STEMdiff Intestinal Organoid Kit, StemCell Technologies), suggesting that these systems are fairly flexible and can be modified depending on the specific needs of the end-user. In some cases, such as those where maintaining a mesenchymal population is desirable (e.g. human lung organoids(15)) serum may be required since lung organoids plus mesenchyme grown in serum free medium was shown to lead to mesenchymal loss over time(28).

The bud tip progenitor organoid model (Figure 4-1B) generates a highly enriched population of lung bud tip progenitors from hPSCs in 22 days in a serum-free

environment, but this population is nevertheless composed largely of undifferentiated progenitor cells. Thus, this system is not ideal for studying differentiated cell function (i.e. airway function or alveolar function). Understanding the mechanisms by which bud tip progenitor cells differentiate into a specific functional epithelial cell type remains a significant gap in knowledge in the field, which the bud tip progenitor organoid model will help to close. Moreover, our initial attempts to differentiate bud tip progenitors has used a stochastic differentiation strategy where cells were provided with a supportive growth environment containing only FGF7, and were allowed to spontaneously differentiate. In this context, bud tip progenitors gave rise to immature alveolar and airway cell-types. However, we noted that these cells exhibited a bias towards generating mucous-producing cells *in vitro.* This result is not well understood. However, much further progress will be required before truly controlled directed differentiation is achieved and becomes a viable option for regenerative medicine.

4.9 Expertise needed to implement the protocol

Any student or post-doctoral fellow with human pluripotent stem cell culture experience can perform this protocol. No core facilities are required. All equipment is standard to most cell culture facilities and reagents can be purchased from standard scientific vendors.

4.10 Materials

Reagents

Biological Material

hESCs or hiPSCs. This protocol can be performed with any hESC or human iPSC line. We have successfully used hESC lines H1, H9, UM63-1 and UM77-1 (NIH registry nos. 0277, 0062, 0043 and 0278, respectively). H9 and H1 cell lines were obtained from the WiCell Research Institute and UM 63-1 and 77-2 lines were obtained from the University of Michigan. We have also successfully used iPSC line 20-1(5). In our experiences, genetically modified reporter lines derived from parental cell lines performed equally well as unmodified counterparts (e.g. H9 mCherry and iPSC 20-1 TetO-GFP).

Caution: research utilizing human pluripotent hPSCs must be performed in accordance with federal, state, local and institutional ethical guidelines and regulations.

Caution: cell lines should be regularly checked for mycoplasma contamination and tested using cell line authentication methods (i.e. STR profiling) to ensure proper cell line identity.

Cell growth media and supplements

- 1, thioglycerol (Sigma-Aldrich, cat. no. M6145)
- Activin A (R&D Systems, cat. no. 338-AC)
- L-Ascorbic Acid (Sigma-Aldrich, cat. no. A4403).
- Advanced DMEM (ThermoFisher Scientific cat. no. 12491015)
- All-Trans Retinoic Acid (Stemgent cat. no. 04-0021, CAS Number 302-79-4)

- B27 supplement (ThermoFisher Scientific cat. no. 17504044)
- Bovine Serum Albumin (Sigma-Aldrich cat. no. A9647)
- CHIR-99021 (Stem Cell Technologies cat. no. 72054)
- DMEM/F12, No Glutamine (DMEM/F12) (ThermoFisher Scientific cat. no. 21331020 or 21331-020)
- Fetal Bovine Serum (ThermoFisher Scientific cat. no. 16000044)
- FGF10 (Recombinant Human Fibroblast Growth Factor 10; R&D Systems cat. no. 345-FG OR made in house as previously described (17)
- FGF4 (Recombinant Human Fibroblast Growth Factor 4; R&D Systems cat. no. 7460-F4 OR made in house as previously described (17)).
- FGF7 (Recombinant Human Fibroblast Growth Factor 7; R&D Systems cat. no. 251-KG/CF)
- Glutamax 100x (ThermoFisher Scientific cat. no. 35050061)
- HEPES buffer (1 M ThermoFisher Scientific cat. no. 15630080)
- HyClone defined Fetal Bovine Serum (dFBS, ThermoFisher Scientific cat. no. 10437010)
- L-Ascorbic Acid (Sigma-Aldrich cat. no. A4544, CAS Number 50-81-7)
- mTeSR™1 (Stem Cell Technologies cat. no. 85850)
- N-2 supplement (ThermoFisher Scientific cat. no. 17502048)
- NOGGIN (R&D Systems, cat. no. 6057)
- Penicillin-Streptomycin 100x (ThermoFisher Scientific cat. no. 15140122)
- RPMI 1640 Medium (ThermoFisher Scientific cat. no. 11875119)
- SB431542 (Stemgent cat. no. 04-0010)

- Smoothened Agonist (SAG; Enzo Life Sciences cat. no. ALX-270-426-M001)
- Y27632 (APExBIO cat. no. A30008)

Matrices, enzymes and other reagents

- Dispase (ThermoFisher Scientific cat. no. 17105041)
- Dimethyl Sulfoxide (DMSO; Sterile Filtered; Sigma-Aldrich cat. no. D2650, CAS number 67-68-5)
- Matrigel Basement Membrane Matrix Growth Factor Reduced (Corning cat. no. 354230). Aliquot the volume of Matrigel that will yield 100 µg/mL when diluted in 12 mL of liquid. A protein concentration is provided with the product specification sheet.
    - CRITICAL For coating plates to maintain hPSCs.
- Matrigel Basement Membrane Matrix (Corning cat. no. 354234)
    - **CRITICAL** For culturing organoids in droplets.
    - **CRITICAL**: Lot to lot variation occurs. Ensure Matrigel has a protein concentration above 8.0 mg/mL. If protein concentration is higher, dilute to 8.0 mg/mL using DMEM/F12. If protein concentration is too low, organoids will sink to the bottom of the droplet and adhere to the plastic culture dish, losing their 3D structure and likely changing the fate of the cells.
- Parafilm M™ Wrapping Film (ThermoFisher Scientific cat. no. S37440)
- Sterile Phosphate Buffered Saline (PBS, pH 7.4; ThermoFisher Scientific cat. no. 10010049)
- Sterile Distilled Water (ThermoFisher Scientific cat. no. 15230170)

Equipment

- Tissue Culture hood
- Cell Culture Incubator
- Dissecting microscope
- Centrifuge
- Pipet aid (like FisherScientific cat. no. 1368106)
- Cell Scraper (VWR cat. no. 76036-004)
- Steri-Flip (Millipore item no. EW-29969-24)
- 6 well plates (tissue culture treated; ThermoFisher Scientific, Thermo Scientific™ Nunc™ Cell-Culture Treated Multidishes 6 well cat. no. 140675)
- 24 well plates (tissue culture treated; ThermoFisher Scientific, Thermo Scientific™ Nunc™ Cell-Culture Treated Multidishes 24 well cat. no. 142475)
- 10 mL serological pipettes (ThermoFisher Scientific cat. no. 07200574, Corning™ Stripette™ Plastic-Wrapped Disposable Polystyrene Serological Pipettes)
- 1.5 mL tubes (ThermoFisher Scientific cat. no. 02682000)
- 15 mL conical centrifuge tubes (Falcon™; ThermoFisher Scientific cat. no. 1495953A)
- 50 mL conical centrifuge tubes (Falcon™; ThermoFisher Scientific cat. no. 1443222)
- 27 gauge needles (BD PrecisionGlide™ Single-use Needles: Regular Bevel - Regular Wall; ThermoFisher Scientific cat. no. 1482113B)

- 1 mL syringes (BD Disposable Syringes with Luer-Lok™ Tips; ThermoFisher Scientific cat. no. 1482330)

4.11 Reagent setup

Human Pluripotent Stem Cell Culture

Culture of hPSCs has been described elsewhere, e.g. Harvard Stem Cell Institute StemBook Protocols http://www.stembook.org/protocols/pluripotent-cells. Detailed instructions are provided here to ensure the successful completion of this protocol. More technical information regarding maintenance of hPSCs using mTeSR™1 can be found at the Stem Cell Technologies website, https://www.stemcell.com/guide-to-passaging-human-pluripotent-stem-cells-using-mTeSR1.html.

- Aliquoting matrigel for coating plates to maintain hPSCs: thaw Matrix Growth Factor Reduced (BM GF-reduced Matrigel) Matrigel at 4ºC overnight. Aliquot BM GF-reduced Matrigel as indicated on the manufacturer product sheet (lot dependent). The product specification sheet will provide a protein concentration for each lot of Matrigel. Aliquot the volume of Matrigel that will lead to a 100 µg/mL of hESC qualified GF-reduced Matrigel when it is diluted in to 12 mL of cold DMEM F/12. 12 mL of 100 µg/mL of Matrigel is enough to coat one 24-well plate or two 6-well plates. Aliquots can be stored at -80 for up to 6 months.

- Coating plates with basement membrane growth factor reduced Matrigel. **Timing**: 15-30 minutes to set up. 1 hour-to-overnight incubation time.

6 well tissue culture plates are coated with a thin layer of Matrigel to improve adhesion of stem cells. For routine maintenance of hPSCs, 6 well plates are used. For differentiation, 24 well plates are used, and the number of wells can be scaled according to the needs of an individual experiment. The following procedure provides instructions to prepare two 6-well plates or one 24-well plate. Plates can be prepared 1 hour prior to passaging cells and left to coat at room temperature in a sterile hood, or plates can be coated up to a week before cells are split and stored at 4ºC with the plate wrapped in parafilm to prevent evaporation. When you are ready to coat plates, remove one aliquot of basement membrane growth factor reduced Matrigel from -80ºC and leave on ice until completely thawed. This aliquot corresponds to a volume of Matrigel that will yield a concentration of 100 µg/mL of protein when diluted into 12 mL of liquid. Each lot of hESC qualified Matrigel comes with a protein concentration supplied by the manufacturer. Gently pipette the thawed basement membrane growth factor-reduced Matrigel into 12 mL of cold DMEM/F12. Use immediately or store at 4ºC for up to 1 week. Do not agitate Matrigel while thawing to avoid premature polymerization. Add 1 mL of diluted Matrigel to each well of a 6-well plate or 0.5 mL of thawed Matrigel to each well of a 24-well plate. Gently swirl or tap the plate to ensure the entire bottom of each well is covered. Let

plates sit for 1 hour at room temperature prior to use, or, wrap edges of plate with Parafilm and store at 4ºC for up to 1 week.

Preparing matrigel before making matrigel droplets to culture organoids:

- Basement Membrane Matrigel should be stored at -80ºC and thawed overnight at 4ºC prior to use. Dilute basement membrane Matrigel with DEME/F12 to a concentration of 8.2 mg/mL of protein.
- **CRITICAL**: When working with Matrigel, keep on ice at all times and work quickly. Matrix will begin to solidify as it warms up.

- Dispase:
  - Prepare a 5 mg/mL solution of dispase powder with DMEM/F12. Make 2 mL aliquots in 15 mL conical tubes and store at -20ºC for up to 1 year. When it is time to use, add 8 mL of DMEM/F12 to a 2 mL aliquot of dispase for a working concentration of 1 mg/mL. Can store at 4 degrees for up to 1 month.
- Growth Factors and Small Molecules:
  - Dilute, aliquot and store according to manufacturers' instructions.
  - Once thawed, growth factors and small molecules can be kept at 4ºC for up to 2 weeks.

Ascorbic Acid

To make a 50 mg/mL solution, dissolve 500 mg of L-Ascorbic Acid into 10 mL of tissue culture grade sterile water in a 50 mL conical tube. In a tissue culture hood, filter this solution using a 0.22 µM Steriflip filter. Make aliquots of 50 µL and store at -20ºC for up to 1 year. Use a fresh aliquot each time; do not store a thawed aliquot for reuse.

- Definitive endoderm differentiation medium for use on Day 2

  Dilute HyClone FBS in RPMI basal medium to 0.2% concentration by volume for day 2 endoderm differentiation. To make day 2 endoderm medium for all wells of a single 24-well plate, add 24 µL of HyClone FBS to 12 mL of RPMI medium. Detail also included in **Table 4-2.**

- Definitive endoderm differentiation medium for use on Days 3 and 4:
  - Dilute HyClone FBS in RPMI basal medium to a 2% concentration for day 3 and 4 of endoderm differentiation. To make basal medium for 2 days for a single 24-well plate, add 480 µL of HyClone FBS to 24 mL of RPMI medium. Store at 4ºC overnight between uses. Detail also included in Table 4-2.

- Foregut differentiation medium

  Make foregut basal medium by adding 1X N2, 1X B27, 10mM HEPES, 1X GlutaMAX and 1X Penn/Strep to Advanced DMEM/F12 in a sterile hood. Store at 4ºC for up to 2 months. On day of use, add 10 µM SB431542, 200 ng/mL Noggin, 1 µM Smoothened Agonist (SAG), 500 ng/mL FGF4 and 2 µM CHIR99021. Once growth factors and small molecules have been added, use same day. Do not store. Detail also included in Table 4-2.

- Human lung organoid medium:

- On day of use, add 1% FBS and 500 ng/mL of FGF10 to foregut differentiation medium. 50 mL is enough to feed a 24-well plate for roughly 2 weeks. Store at 4ºC for up to 2 weeks. Detail also included in Table 4-2.

- Bud tip progenitor organoid basal medium:
  - First, prepare Bovine Serum Albumin (BSA) by taking a 500 mL bottle of DMEM/F12 and transferring 10 mL of DMEM/F12 from the 500mL bottle into a 50 mL conical tube. Weigh out 0.25 g of BSA and add into your DMEM/F12 aliquot. Warm in a water bath for 15 minutes with occasional gentle agitation to dissolve BSA. In a tissue culture hood, filter the dissolved solution through a Steri-Flip filter. Add BSA solution back to the DMEM/F12 bottle in a sterile hood.

    Add 1X N2, 1X B27, 1X Glutamine and 1X Penn/Strep to the DMEM/F12+BSA in a sterile hood to complete the bud tip progenitor organoid basal medium. Store at 4 degrees up to 2 months. Detail also included in Table 4-2.

- Bud tip progenitor organoid complete medium for generation and maintenance of bud tip progenitor organoids:
  - When one is ready to generate or maintain bud tip progenitor organoids, add appropriate supplements to the bud tip progenitor organoid basal medium:

To make a 50 mL volume of ‘bud tip progenitor organoid complete medium’, place 50 mL of bud tip progenitor organoid basal medium in a 50mL conical tube and add 50 µg/mL ascorbic acid and 0.4 µM monothioglycerol. This is the complete medium.

- Then add 3 µM CHIR99021, 10 ng/mL FGF7 and 50 nM All-Trans Retinoic Acid (ATRA). 50 mL of medium is enough to feed a single 24-well plate for ~2 weeks. Can store at 4ºC for up to 2 weeks.

4.12 Procedure

Splitting of Human Pluripotent Stem Cell Culture

Timing: 30 minutes to 1 hour.

CRITICAL: Proper maintenance of hPSCs is essential to successfully complete these protocols. Seeding density and colony size must be closely monitored to ensure no premature differentiation occurs. See Figures 4-3 and 4-4 for visual guides. Note: we do not use feeder layers to maintain hPSCs.

CRITICAL: This procedure has been developed and optimized for hPSCs maintained in 6-well tissue culture dishes coated with 100 µg/mL of basement membrane growth factor reduced Matrigel and fed daily with mTeSR™1. For routine maintenance, cells are kept in a tissue culture grade incubator with 95% air and 5% CO_2 and are maintained in a sterile environment.

CRITICAL: To begin the differentiation protocol, split 4 wells of a 6-well plate into all wells of one 24-well plate. Alternatively, 1 well of a 6-well plate can be split into 6 wells of a 6-well plate to maintain hPSCs.

1. Undifferentiated hPSC colonies that are ~75-85% confluent will appear rounded with well-defined smooth edges (Figure 4-3A and 4-3G). Colonies with jagged edges, many bright white raised spots or 'donut-holes' have likely begun to differentiate, and should not be used for this protocol (Figure 4-4A). If there are only a few differentiated areas, scrape differentiated colonies off with a pipette tip under a dissecting microscope before splitting (Figure 4-4A, arrow).
2. Warm Dispase, mTeSR™1 and DMEM/F12 in a 37ºC water bath for ~10 minutes.
3. Add 1 mL of 1 mg/mL Dispase solution to each well that will be split. Return to incubator for 5-10 minutes. Observe colonies under a microscope, and proceed when the edges of the colonies begin to lift off of the tissue culture dish.
4. Carefully aspirate dispase. Colonies will still be attached to the dish.
5. Gently wash cells 3 times with prewarmed DMEM/F12.
6. Add 3.5 mL of mTeSR™1 medium to each well that will be split. This includes an extra 0.5 mL of medium per well to accommodate for volume that may be lost during trituration and transfer to new plates.
7. Use a cell scraper to gently scrape all cells off the bottom of the dish.

8. Using a 10 mL serological pipette and automatic dispenser, triturate colonies into small clumps by doing the following: pull all cells up into the pipette, place the pipette firmly on the bottom of the dish and pipette down to shear the cell colonies apart. Check under a dissecting scope under low magnification (images in Figure 4-2B' shown at 0.7x magnification) to assess the size of the sheared colonies. Repeat this process until colonies are small enough that they do not immediately sink to the bottom of the dish. Figure 4-4B shows sheared colonies in suspension under a dissecting microscope that are too large. Aim to shear colonies until their size is similar to what is shown in Figure 4-3H. Cell clumps should be small enough to remain floating in suspension, but should not be triturated all the way to single cells. **CRITICAL**: if sheared cell colonies are too large when splitting, proper differentiation will be difficult to achieve during this protocol (Figure 4-4D).
9. Aspirate the diluted Matrigel media that was used to coat each well of the 24 well dish. If splitting into a 6-well dish, add 1 mL of fresh, prewarmed mTeSR™1 plus 0.5 mL of triturated hPSC clumps. If splitting into a 24-well dish for differentiation, add 0.5 mL of triturated hPSCs to each well.
10. Under the dissecting microscope, aggressively shake the plate in a side to side motion to ensure that hPSCs are evenly distributed throughout the entire well, as seen in Figure 4-2C'). This can also be achieved by tapping the side of the 24 well plate several times. **CRITICAL**: If all of the hPSCs

congregate to the center of the well (as shown in Figure 4-4C), spheroids will not form.

Directed differentiation into endoderm and anterior foregut spheroids.

Timing: 9-12 days. 4 days definitive endoderm differentiation, 4-5 days anterior foregut endoderm differentiation.

11. Allow passaged cells to become 45-75% confluent before starting the protocol. This may take 1-2 days after passaging. **CRITICAL**: If confluency is too low, insufficient cells will remain following endoderm differentiation. If confluency is too high, differentiation within the center of the colonies may be incomplete and result in reduced spheroid formation See Figure 4-4D for images of improper seeding density or colonies that are too large. Proper confluency of cells in a 24-well plate ready to proceed to endoderm differentiation is shown in Figure 4-3I.
12.
    - Feed cells for 4 days with Definitive Endoderm Differentiation medium (Each day the medium needs to be changed by removing existing medium and replacing with 0.5mL medium per well of the 24-well plate. The appropriate media to use each day are listed below and in Table 4-2:

Day	Medium
1	RPMI + 0% FBS + 100 ng/mL Activin A
2	RPMI + 0.2% FBS + 100 ng/mL Activin A
3	RPMI + 2% FBS + 100 ng/mL Activin A
4	RPMI + 2% FBS + 100 ng/mL Activin A

Also check that cell morphology is as expected by comparing with the representative images shown in 3J.

13. On day 5, replace medium with anterior foregut endoderm differentiation medium (0.5mLs of foregut basal media + 10 µM SB431542 + 200 ng/mL Noggin + 1 µM Smoothened Agonist (SAG) + 500 ng/mL FGF4 + 2 µM CHIR99021, as listed in Table 4-2). Replace medium on days 6, 7, 8 and 9.
    - Day 5-9: Change media daily. Also check that cell morphology is as expected by comparing with the representative images shown in Figure 4-3K.
14. Observe that detached, floating spheroids begin to form around day 8 and 9 of culture (representative images are shown in Figure 4-3K-L).

Collection of floating anterior foregut spheroids Timing: 15-60 minutes.

- **CRITICAL**: Ensure sterile technique and a sterile environment throughout the process. This procedure is routinely performed in a laminar flow hood equipped with a stereomicroscope, or in a semi-sterile hood (e.g. Labconco 3970321) equipped with a stereomicroscope.
- **CRITICAL** A video presentation of generating Matrigel droplets for intestinal organoid culture has been published and can be used as a helpful reference for setting up lung organoid cultures(41) (Found at the 4:15 minute mark of the video https://www.jove.com/video/53359/organoids-as-model-for-infectious-diseases-culture-human-murine).

15. On day X or Y, place Matrigel on ice in the sterile hood.
16. Cut the tip off of a sterile p200 pipette filter tip (~5 mm) using sterile scissors or a scalpel and prepare an empty 1.5 mL snap-cap tube, a new (uncoated) 24-well plate in a hood with a dissecting microscope. **CRITICAL**: some 24 well tissue culture dishes will allow Matrigel to form 3-dimensional droplets, whereas Matrigel disperses across the surface of other dishes. Nunc™ Cell-Culture Treated Multidishes perform well for allowing droplets to form. Other brands should be tested prior to transfer of Matrigel plus spheroids.

17. Place the 24 well culture plate that contains the foregut spheroids under a dissecting microscope. Observe one well under the stereomicroscope to identify free-floating spheroids. Without disturbing the monolayer, gently remove the free-floating spheroids by pipetting with a cut p200 pipette tip and transfer in to an empty 1.5 mL snap-cap tube. Spheroids from multiple wells can be pooled together in a single tube. We typically pool one entire 24 well plate into one 1.5mL tube. Allow 5-10 minutes for the spheroids to settle to the bottom of the tube. Although spheroids should settle by gravity, they can be gently spun down using a micro-centrifuge on the lowest setting for ~10 seconds if necessary.
18. Plate spheroids in basement membrane Matrigel (8.0 mg/mL protein concentration). To do this gently remove as much media as possible from the tube under the dissecting microscope using a fresh p200 pipette. Once media has been removed, cut the tip off a fresh p200 tip. Draw up ~200 µL of Matrigel. Place Matrigel in the tube with the spheroids and gently pipette up and down several times to mix the spheroids in the Matrigel. Work quickly so the Matrigel does not congeal. Pipette with care to avoid making bubbles. Draw up all the Matrigel plus spheroids, and pipette ~25 µL of the Matrigel-spheroid mixture into the center of a single well of the 24 well plate. 200 µL of starting mixture will be sufficient to make 8 individual wells with a Matrigel droplet containing spheroids.

- **CRITICAL** STEP While we routinely draw up the entire contents of Matrigel+spheroids into a single cut p200 tip and estimate 25 µL droplets, it is

also acceptable to pipette individual 25 µL droplets if greater accuracy is desired.

- **CRITICAL** STEP Keep the Matrigel on ice throughout the protocol to ensure it does not polymerize inside the bottle.

19. Place the plate in the incubator for ~10 minutes until Matrigel droplets have solidified. Once droplets have solidified, add 0.5mL of desired media to generate either human lung orgaoinds or epithelial bud tip progenitor organoids (see next step for details).

Generation of human lung organoids or bud tip progenitor organoids.

20. Either generate human lung organoids (option A), which contain airway-like structures surrounded by lung mesenchyme and a population of cells with alveolar markers. or bud tip progenitor organoids (options B and C). Use option B if doing needle passaging. Use option C if using bulk passaging.

A. Generation and maintenance of human lung organoids. **Timing**: Organoids with airway-like structures and surrounding mesenchyme will take ~50 days to grow, and are optimal for use from 50-85 days, as this is when the organoids have the maxium amount of airway-like structures. They need to be re-embedded in fresh matrigel every 2-3 weeks.

I. Add 1% FBS and 500 ng/mL of FGF10 to Foregut Basal medium. Add 500 µL of media to each well of spheroids, ensuring that the media completely covers Matrigel droplets. Change media every 3-5 days. See **Table 4-2 for a summary of the media components**.

II. ***Re-embedding of organoids.*** Every 2-3 weeks, or sooner if organoids appear to be accumulating cellular debris within lumens or sinking to the bottom of the Matrigel droplet, cut the tip off a P1000 tip and gently scrape the bottom of the well to dislodge the Matrigel droplet(s) containing organoids.

III. Pick up the entire droplet and surrounding media with the pipette tip and move to a petri dish. Under a dissecting scope in a sterile environment, use a scalpel and 27-gauge needle to gently cut and remove the old Matrigel away from the organoids. Be cautious not to cut the tissue. Once finished, move the now cleaned organoids to a 1.5 mL snap cap tube and remove any media.

IV. Cut the tip off a p200 tip and transfer 200 µL of fresh Matrigel (kept on ice) to the tube with the organoids, mix, and draw up all Matrigel into the p200 tip. Serially pipette ~50 µL of the Matrigel-organoid mixture into single wells in a fresh 24 well plate. Aim to put 1-3 organoids in each droplet.

V. Place plate in the incubator for ~10 minutes, or until Matrigel droplets have solidified. Once droplets have solidified, add 500 µL of media. vi) Maintain organoids for the desired time period, repeating steps ii-v to

re-embedd organoids every 2-3 weeks. Once foregut spheroids are placed in Matrigel, it takes roughly 50 days to obtain pseudostratified epithelial structures that possess basal cells, mucous producing cells and ciliated cells, as well as surrounding lung mesenchymal cell types and primitive alveolar cell-types. The organoids can be maintained for over 100 days; however, we find that the epithelium is gradually lost over time, and utilization between 60-85 days is optimal.

B. Generation and maintenance of bud tip progenitor organoids using serial needle passaging. **Timing**: Epithelium-only cysts will form after ~2 weeks which will contain a nearly homogenous population of bud tip progenitor cells. This population can be maintained and expanded by serial needle passaging for over 120 days in culture.

I. On the day of use, add 50 µg/mL L-Ascorbic acid and 0.04 µl/mL Monothioglycerol, 10 ng/mL FGF7, 50 nM ATRA and 3 µM CHIR-99021 to basal medium composed of DMEM/F12 + N-2 supplement + B-27 supplement + 1X L Glutamine + 1X Penn/Strep to make bud tip progenitor organoid complete medium (as also detailed in Table 4-2) . Feed foregut spheroids with 500 µL of bud tip progenitor organoid complete medium.
II. Replace media every 3-5 days.
III. ***Needle passage.*** After 2-3 weeks, or sooner if organoids appear to be accumulating cellular debris within lumens or sinking to the bottom of

the Matrigel droplet, use a P1000 pipette tip to dislodge Matrigel droplets containing organoids. Place the dislodged droplets together into a 1.5 mL snap-cap tube.

IV. Using a 1 mL syringe and a 27-gauge needle, draw the Matrigel and cells into the syringe and forcefully expel the contents back into the tube. Repeat 2-3 times.

V. Spin down the contents of the tube in a benchtop mini centrifuge (like ThermoFisher Scientific mySPIN^{TM} 6 Mini Centrifuge cat. no. 75004061) at full speed (6,300 g) for ~3-5 seconds. It is also acceptable to spin cells down at 300g for 3 minutes at 4ºC. Epithelial fragments will settle to the bottom and Matrigel will remain suspended in the media.

VI. Under a dissecting scope in a sterile hood, use the needle and syringe to remove the media and Matrigel from the tube, leaving behind the cell pellet.

VII. Add ~200 µL of fresh ice cold Matrigel to the tube using a p200 with the tip cut off. Mix the cells and Matrigel, avoiding bubbles. Draw up all 200 µL of the Matrigel mixture and serially pipette ~25 µL of the mixture into a single well of a 24-well plate.

VIII. Place plate in the incubator for ~10 minutes or until Matrigel droplets have solidified. Add 750 µL of bud tip progenitor organoid complete medium.

IX. Repeat the needle-passage (steps iii-viii) every 2-3 weeks or at any time to select for the bud tip progenitor population.

C. Generation and maintenance of bud tip progenitor organoids using bulk passaging.

Timing: Epithelium-only cysts will form after ~2 weeks which will contain a nearly homogenous population of bud tip progenitor cells. Bud tip progenitor organoids can be maintained by keeping their structure intact. Bud tip progenitor organoids that are not needle passaged will develop budded epithelial structures and elements of proximal-distal patterning, as well as functional mucous producing cells on the interior of the structure.

CRITICAL Bud tip progenitor organoids that are not needle passaged will secrete mucous into the lumen of the organoid, causing a dense appearance and leading to increased cell death, therefore these organoids must be cleaned out regularly.

I. On the day of use, add 50 µg/mL L-Ascorbic acid and 0.04 µl/mL Monothioglycerol, 10 ng/mL FGF7, 50 nM ATRA and 3 µM CHIR-99021 to basal medium composed of DMEM/F12 + N-2 supplement + B-27 supplement + 1X L Glutamine + 1X Penn/Strep to make bud tip progenitor organoid complete medium (as also detailed in Table 4-2) . Feed foregut spheroids with 500 µL of bud tip progenitor organoid complete medium.
II. Replace media every 3-5 days.

III. ***Passage of bud tip progenitor organoids without a needle***: After 2-3 weeks, or sooner if organoids appear to be accumulating cellular debris within lumens or sinking to the bottom of the Matrigel droplet, scrape Matrigel droplets with bud tip progenitor organoids using a cut P1000 and transfer both the media within the well along with the matrigel droplet to a petri dish.

IV. Use a scalpel to cut away old Matrigel without disturbing the tissue structure, and re-embed in fresh Matrigel droplets as described in steps X-Y above. These organoids contain bud tip progenitor cells in budded structures at the periphery of the organoid. Interior regions will predominantly show SOX2+ airway-like structures with differentiated mucous producing cells. They can be needle-passaged as described in option B steps iii-viii at any time to select for the bud tip progenitor population.

V. ***Clearing mucous from non-needle passaged organoids:*** Mucous or cellular debris will often accumulate within the lumens of organoids and will appear as an optically dense, dark and granulated substance within the lumen after ~2 weeks in culture without passage. To clear this debris, get a sterile scalpel, a clean 27-gauge needle, a 1 mL syringe and a fresh aliquot of prewarmed DMEM/F12. Transfer organoids plus Matrigel to a petri dish with prewarmed DMEM/F12 and transfer to a sterile hood with a stereomicroscope.

VI. Under the stereomicroscope, manually cut organoids in half with the sterile scalpel. Using the 1 mL syringe with the 27-gauge needle, draw up clean pre-warmed DMEM/F12 into the syringe. Gently expel the media through the needle to wash the mucous away from the organoids.

Timing: See **Figure 4-1** for an overview of the timeline of the protocol. Generating foregut spheroids from hPSCs takes 9 days (procedure steps 11-14). It takes roughly 1 hour to collect foregut spheroids and culture them in a Matrigel droplet (procedure steps 15-20). From there, human lung organoids with airway-like structures and multiple immature epithelial and mesenchymal cell types can be generated after roughly 50 days in culture (41 days from foregut spheroid stage; procedure steps A I-V), whereas bud tip progenitor organoids can be generated within 22 days (14 days from foregut spheroids stage; procedure steps B I-II), and can be maintained as proliferating progenitors (procedure steps B III-IX) or as budded organoids with proximal-distal patterning (procedure steps C I-VI).

4.13 Troubleshooting

The most common problem we have identified is a failure of foregut spheroids to form during the directed differentiation of hPSCs (procedure steps 11-14). Common causes of failure to form spheroids are: 1) the initial seeding density of hPSCs is too dense; 2) hPSCs colonies are allowed to grow too large in the 24 well plate; 3) growth factors have been stored at 4^{o}C for too long and have lost potency; 4) hPSC lines were

not maintained well (they contained spiky, unhealthy colonies) or differentiation was started at too high a passage number. Generally, hPSCs with a passage number less than 90 should be used. We have also noticed that high passage (over 70 passages) hPSCs have reduced spheroid forming ability. If spheroids fail to form, the experiment must be attempted again from the beginning.

To ensure that the seeding density and colony sizes of hPSC starting material is correct, we recommend comparing results to those shown in Figure 4-3 H-I and Figure 4-4B-D. Preliminary and unpublished experiments in our lab have suggested that stem cells can be broken into colonies as small as single cells and plated at a density of 125,000 cells per single well of a 24-well plate to generate foregut spheroids - this could be attempted if mechanical trituration is not working well. Our experience is that small colonies differentiate far more robustly than large colonies. Review Figures 4-3 and 4-4 as a guide for appropriate colony sizes.

We also recommend that you record how long growth factors have been stored at 4^{o}C. When in doubt, throw them out.

Once spheroids have been plated in Matrigel, another common problem is that the spheroids and resulting organoids will appear to drop to the bottom of the Matrigel droplet and adhere to the plastic culture dish (procedure steps AI and BI). If this happens, it is crucial to replate the organoids in a fresh Matrigel droplet as soon as possible – any organoids that have adhered to the plate and have a “flat” morphology should be discarded. Possible reasons for organoids “bottoming out” are 1) it has been longer than 2-3 weeks since the last passage, 2) the Matrigel has been at 4ºC for too long and the protein concentration has decreased due to adhesion to the container or 3)

the Matrigel droplets were too large and do not have a high dome shape that allows ample room for 3D growth.

4.14 Anticipated Results

Foregut spheroids should self-aggregate, detach from the monolayer, and float freely in the media by the 4th day of foregut differentiation (Figure 4-3K-L). The number of spheroids formed in each well can vary. Generally, a single well will produce roughly 100-400 spheroids by day 5 of foregut differentiation. They appear as small irregular cell clumps (Figure 4-5A) that express NKX2.1 and SOX2 (Figure 4-5B-C).

Human lung organoids: Spheroids that are treated with FGF10 and 1% FBS will give rise to human lung organoids. Spheroids grown in these conditions will generate clear epithelial structures by day 12 in culture. As the organoids grow, mesenchymal populations will be visible surrounding epithelial structures (Figure 4-5D). Epithelial structures will persist until ~100 days in culture as seen by Hemotoxylin and Eosin staining (Figure 4-5E). By 65 days in culture, Human Lung Organoids will have pseudostratified epithelial structures with P63+ basal stem cell-like cells lining the basal surface (Figure 4-5F), as well as a population of cells with positive staining for the ciliated cell marker FOXJ1 (~5% of all cells; Figure 4-5G) and alveolar cell markers SFTPC and HOPX (~5% of all cells; Figure 4-5I-J) (15,16). Cells surrounding the epithelial structures stain positive for multiple mesenchymal cell markers, including smooth muscle actin (SMA; Figure 4-5G), PDGFRα and Vimentin (VIM; Figure 4-5H).

Bud tip progenitor organoids: For spheroids that are cultured in serum-free medium with FGF7, CHIR99021 and ATRA, epithelial structures will form clearly by day

12 of culture that are SOX2 and NKX2.1 positive (Figure 4-5K, L). To maintain a highly enriched population of bud tip progenitors, organoids can be needle passaged. This can be done from 2 weeks of culture, or at any time thereafter. Needle passaged organoids form round cysts (Figure 4-5M [day 77-needle passage]), and over 88% of these cells are SOX2+/SOX9+ (Figure 4-5M)(17), and share similar molecular profiles with bud tip progenitors in the human fetal lung prior to 16 weeks gestation(17). If cultures are kept intact and not broken up, they will form organoids with clear budded structures by day 45 (Figure 4-5N). Epithelial bud tip structures will undergo clear bifurcation events (Figure 4-5O). These 'patterned lung organoids' exhibit bud tip-like and airway-like regions (Figure 4-5P). Interior regions of the organoid contain SOX2+ airway-like cells and mucous-producing cells. Roughly 5% of cells in these organoids stain positive for Club cell marker SCGB1A1, whereas roughly 2% of cells stain positive for Goblet cell marker MUC5AC (Figure 4-5Q). Bud tip regions of these organoids will maintain SOX9+, Pro-Surfactant Protein C (ProSFTPC)+ cells, indicating maintenance of bud tip progenitors (Figure 4-5R). Mucous will build up within these organoids, causing the interior to become dense and optically opaque. Mucous can be cleared during passaging to a fresh matrigel droplet.

Author Contribution Statement: AJM, BRD and JRS conceived the studies. AJM, BRD, DFT, DRH and AWO performed experiments. AJM, BRD, DFT, AWO, LDS and JRS analyzed results. AJM and JRS wrote the manuscript. AJM, BRD, DFT, AWO, DRH, LDS and JRS read and edited the manuscript.

Acknowledgements

Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL119215 to JRS. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability statement

Some of the data presented in the current study was generated for and shown in previously published reports. Figure 4-2 B-D was reproduced with permission from Dye *et al*. 2016 and licensed under CC BY-NC-ND 4.0. The original data can be viewed at https://doi.10.7554/eLife.05098. Figure 4-2 E and Figure 4-5 B-C, G-J were reproduced with permission from Dye *et al*. 2015 and licensed under CC BY-NC-ND 4.0. The original data can be viewed at https://doi.10.7554/eLife.05098. Figure 4-2 G-I and Figure 4-5 L-R were reproduced with permission from Miller *et al*. 2018 and licensed under CC BY-NC-ND 4.0. The original data can be viewed at https://doi.org/10.1016/j.stemcr.2017.11.012

Competing/Financial interests: JRS, BRD and AJM hold patents pertaining to the lung organoid technologies described.

4.15 Figures

Figure 4-1 Overview of the protocol

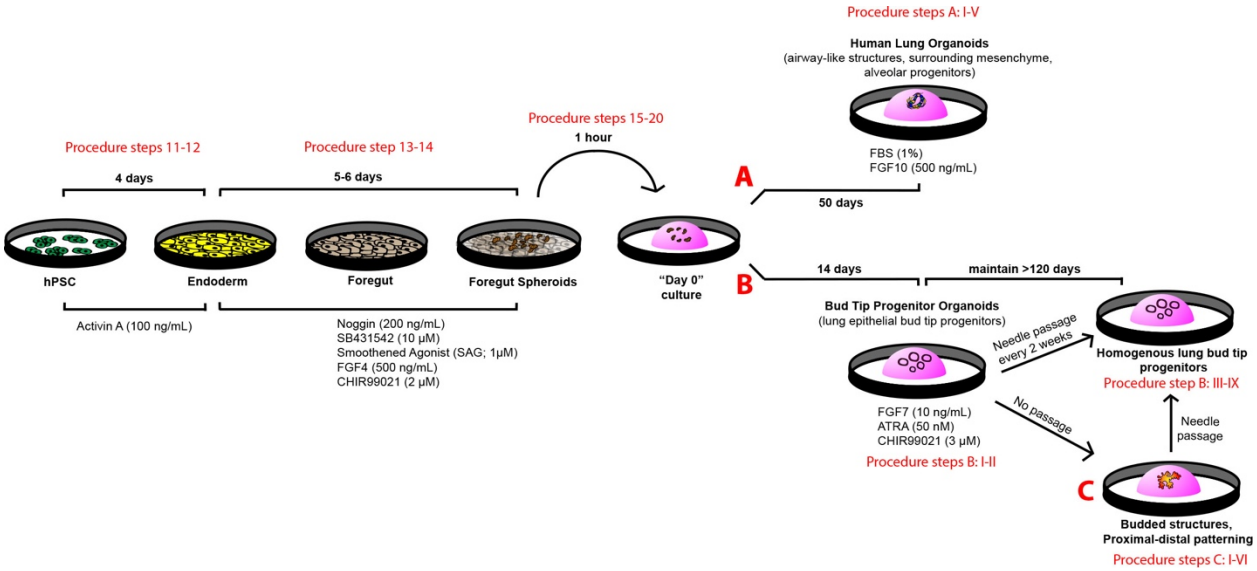


Figure 4-1: Schematic of protocol and timeline. hPSCs are directed in a monolayer to endoderm and then anterior foregut endoderm over the course of 9-10 days. Foregut spheroids self-aggregate and lift away from the monolayer to float in the media. These spheroids are then cultured in a 3D Matrigel droplet, where they can be directed to become airway-like human lung organoids or bud tip progenitor organoids containing bud tip progenitors.

Figure 4-2 Applications of the protocol

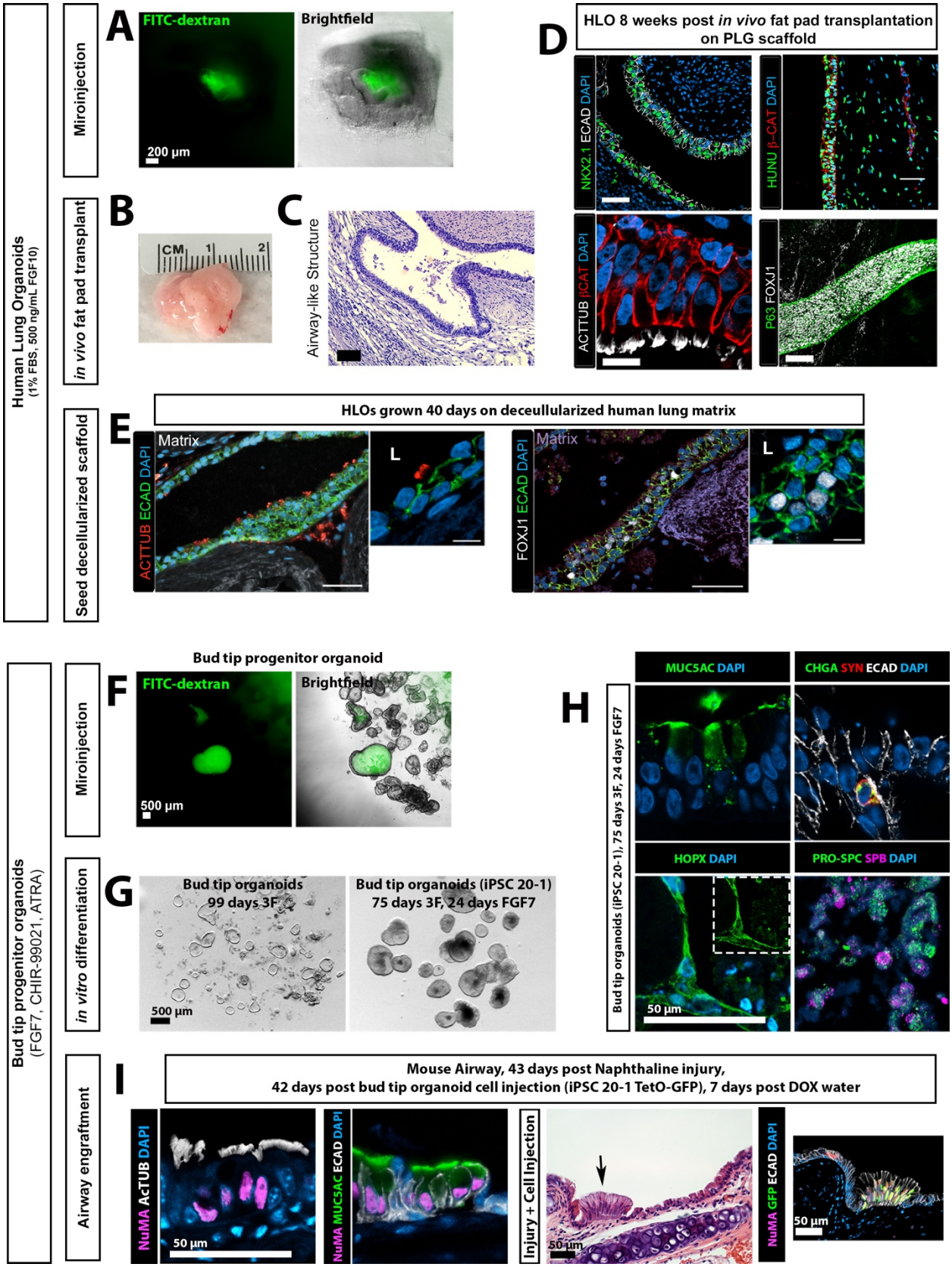


Figure 4-2. Applications of the protocol. A-E) Potential applications for human lung organoids. A) Organoid lumens can be microinjected, as shown here with a microinjection of 1 µL of 4 kDa FITC-Dextran suspended in PBS at a concentration of 2 mg/ml (green)(20). Images show 1 hour after organoid injection. Scale bar represents 200 µm. B) Human lung organoids can be seeded on a bioengineered poly(lactic-co-glycolic acid) (PLG) scaffold and transplanted into the mouse fat pad (full methods can be found in Dye et al. 2016 (16)). After 8 weeks this structure can grow to over 1 cm in diameter. C) Transplanted human lung organoids contain airway-like structures as seen by Hematoxylin and Eosin staining after 8 weeks(16). Scale bar represents 200 µm. D) After 8 weeks of transplantation, human lung organoids contain NKX2.1+ airway epithelial structures that stain positive for the human nuclear antigen (HUNU), exhibit multiciliated cells that stain positive for acetylated tubulin (a video showing the resulting functional beating ciliated cells is available in Dye et al. 2016(16)). Further, these structures contain epithelial tubes with P63+ basal-like cells lining the basolateral surface of the tube and FOXJ1+ ciliated cells lining the interior, luminal side of the pseudostratified epithelium. Scale bars represent 50 µm. E) Roughly 50 Human lung organoids can were seeded onto a decellularized human lung matrix punch in a 96 well plate. Human lung organoid seeded cells gave rise to epithelial structures expressing multiciliated cell markers Acetylated Tubulin and FOXJ1 after 40 days of culture. Scale bars represent 100 µm. F-I) Applications for bud tip progenitor organoids. F) Bud tip progenitor organoids can be microinjected with fluids, as shown by injection of 1 µL of 4 kDa FITC-Dextran suspended in PBS at a concentration of 2 mg/ml (green)(20). 5 organoids were injected. Images show organoids 1 hour post injection. Scale bar

represents 500 μm. G) Bud tip progenitor cells within bud tip progenitor organoids are capable of undergoing multilineage differentiation into airway and alveolar-like cells, and this system can be used to interrogate mechanisms of lineage decisions *in vitro*. Bud tip progenitor organoids treated for 24 days with FGF7 alone generated cells expressing AECI marker HOPX, AECII markers Pro-Surfactant Protein C (PRO-SPC) and Surfactant Protein B (SFTPB). Alevolar-like cells made up roughly 50% of these organoids. Airway-like cells were also generated, including cells staining positive for Goblet cell marker MUC5AC and neuroendocrine markers Chromagranin A (CHGA) and Synaptophysin (SYN). Scale bar represents 50 μm. I) Bud tip progenitor organoids can be expanded in culture and injected directly into the injured mouse airway, where they can engraft. In this experiment, bud tip progenitor organoids were generated from iPSC 20-1 tet-O-GFP cell line to provide Doxycycline (DOX) -inducible GFP expression. Organoids were grown and expanded, dissociated to single cells, and roughly 500,000 bud tip progenitor organoid cells were injected into the mouse airway 24 hours after airway injury with Naphthaline. Mice were allowed to recover for 6 weeks. During the 5th week after cell injection, mice were given DOX water – surviving human cells with access to the host blood stream could then begin expressing GFP upon exposure to DOX.(17) Engrafted cells express human nuclear marker NuMA, and exhibit markers of differentiation including multiciliated cells (Acetylated Tubulin) and Goblet cell marker MUC5AC. These cells integrate seamlessly with the host epithelium. Further, cells were able to receive signals from the host bloodstream, as bud tip progenitor cells derived from a tet-O GFP iPSC line were able to turn on GFP expression when the host mouse was given Doxycycline in drinking water. Scale bars represent 50 μm.

Detailed information about antibodies used for staining can be found in Table 4-3.

Figure 4-2 B-D, reproduced with permission from Dye et al. 2016 and licensed under CC BY-NC-ND 4.0. Figure 4-2 E, reproduced with permission from Dye et al. 2015 and licensed under CC BY-NC-ND 4.0. Figure 4-2 G-I reproduced with permission from Miller *et al*. 2018 and licensed under CC BY-NC-ND 4.0.

Figure 4-3 Human pluripotent stem cell splitting and directed differentiation

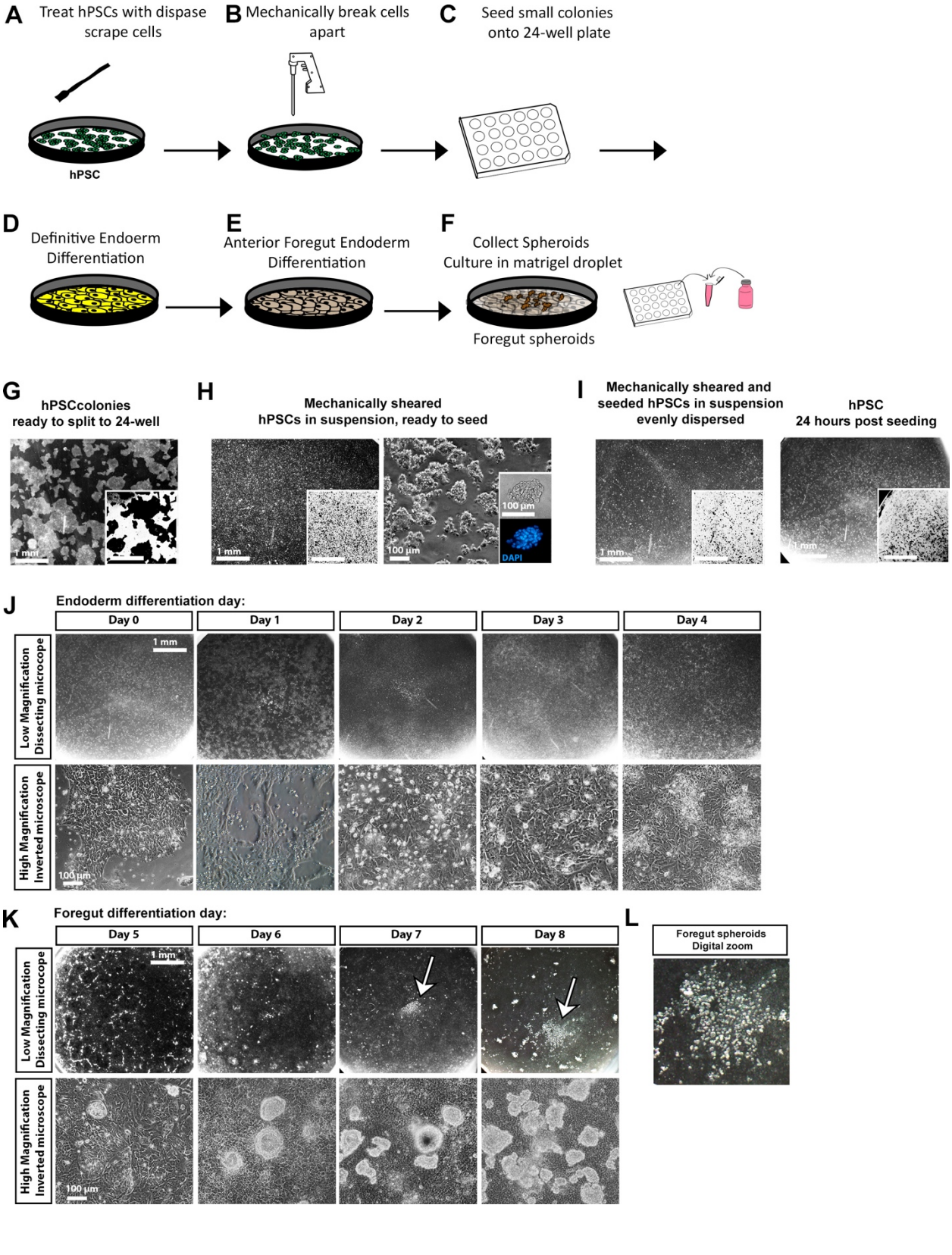


Figure 4-3: Human pluripotent stem cell splitting and directed differentiation. A-F) Schematic of timeline and experimental procedures. G-I) Bright field images of hPSCs during growth and splitting procedures. Insets are inverted images taken at the same magnification to help visualize cell density and colony shape. Scale bars represent 1 mm or 100 μm as noted. A) First, treat hPSCs with dispase. G) Brightfield image indicates proper confluency and healthy colony shape. These cells are ready to split. B) Wash cells, scrape them up with a cell scraper, and then use a serological pipette to mechanically break apart colonies. H) Bright-field image shows the proper size of broken up colonies in suspension. These cells are ready to be seeded in to a 24-well plate. C) Seed cells in to a 24 well plate. I) Bright field image showing floating colonies. These colonies are evenly dispersed throughout the well and can be placed in the incubator overnight. 24 hours after seeding, colonies should cover roughly 50% of the bottom of the well and should be evenly dispersed. D) Begin endoderm differentiation protocol. J-L) Bright field images showing typical morphology in the cell monolayer throughout endoderm differentiation at low magnification through a dissecting scope (J, top row; scale bars represent 1 mm) or higher magnification under an inverted microscope (J, bottom row; scale bars represent 100 μm). E) Begin anterior foregut differentiation protocol. K) Bright field images showing typical morphology in the cell monolayer throughout anterior foregut differentiation. Foregut spheroids begin to form around day 3-4 of anterior foregut differentiation (K, top panel, white arrows). Shown at low magnification through a dissecting scope (K, top row; scale bars represent 1 mm) or higher magnification under an inverted microscope (K, bottom row; scale bars represent 100 μm). I) A digitally zoomed image shows what foregut spheroids look like on the

culture plate on day 8 of the differentiation protocol. F) Collect floating spheroids, mix them with matrigel, and put them into droplets on a fresh 24 well plate.

Figure 4-4 Common errors and troubleshooting

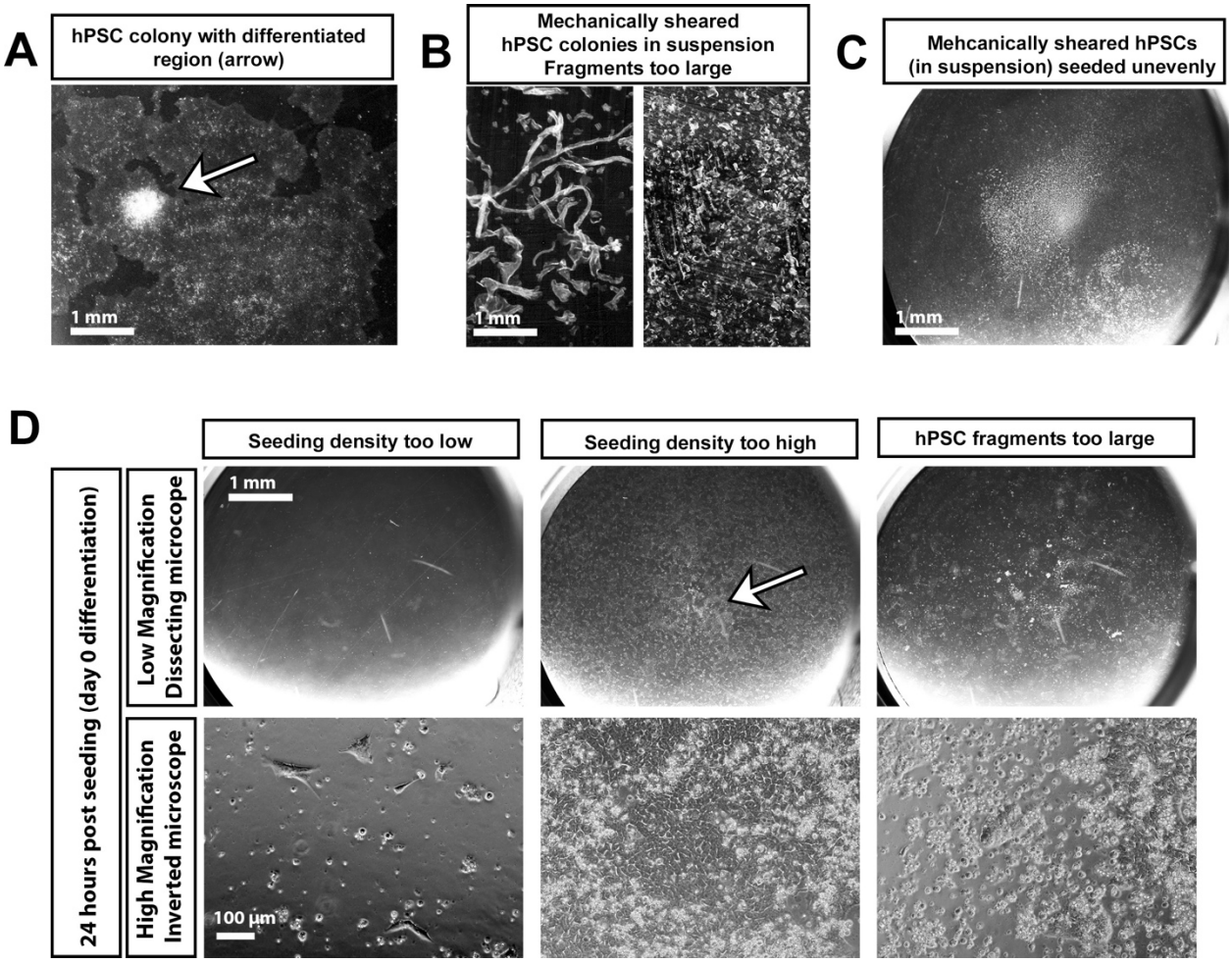


Figure 4. Common errors and troubleshooting. A) Before splitting hPSCs, check that there are no obvious differentiated colonies on the plate. Scrape away any abnormal colonies using a sterile pipette tip before splitting hPSCs into 24-well plates. The arrow indicates an area of hPSC differentiation that is a raised white structure, which should be scraped away before splitting. Scale bar represents 1 mm. B) When breaking up hPSC colonies for splitting, ensure the colonies are broken up to a small enough size. These images show colonies that are still too large. Continue breaking up colonies before proceeding with splitting cells. Scale bar represents 1 mm. C) After hPSC colonies have been placed in a 24-well dish, shake the plate vigorously under a

dissecting microscope to ensure that the cells are evenly dispersed. This image shows the majority of cells are still clustered in the center of the well. This will cause an uneven growth pattern and the differentiation will not yield spheroids. Shake cells before proceeding with differentiation. Scale bar represents 1 mm. D) Bright field images at low (top row, scale bar represents 1 mm) and high (bottom row, scale bar represents 100 μm) magnification showing culture conditions on day 0 of differentiation that are likely to fail to form spheroids. If the seeding density is too low, too high, or if colony sizes are too big because the hPSC colonies were not sufficiently sheared upon splitting, the differentiation protocol will not be successful. The arrow in the high density seeding column indicates that the cells have become confluent in the center of the well, which will lead to uneven differentiation and a likely failure of the protocol to yield foregut spheroids.

Figure 4-5 Expected outcomes of the protocol

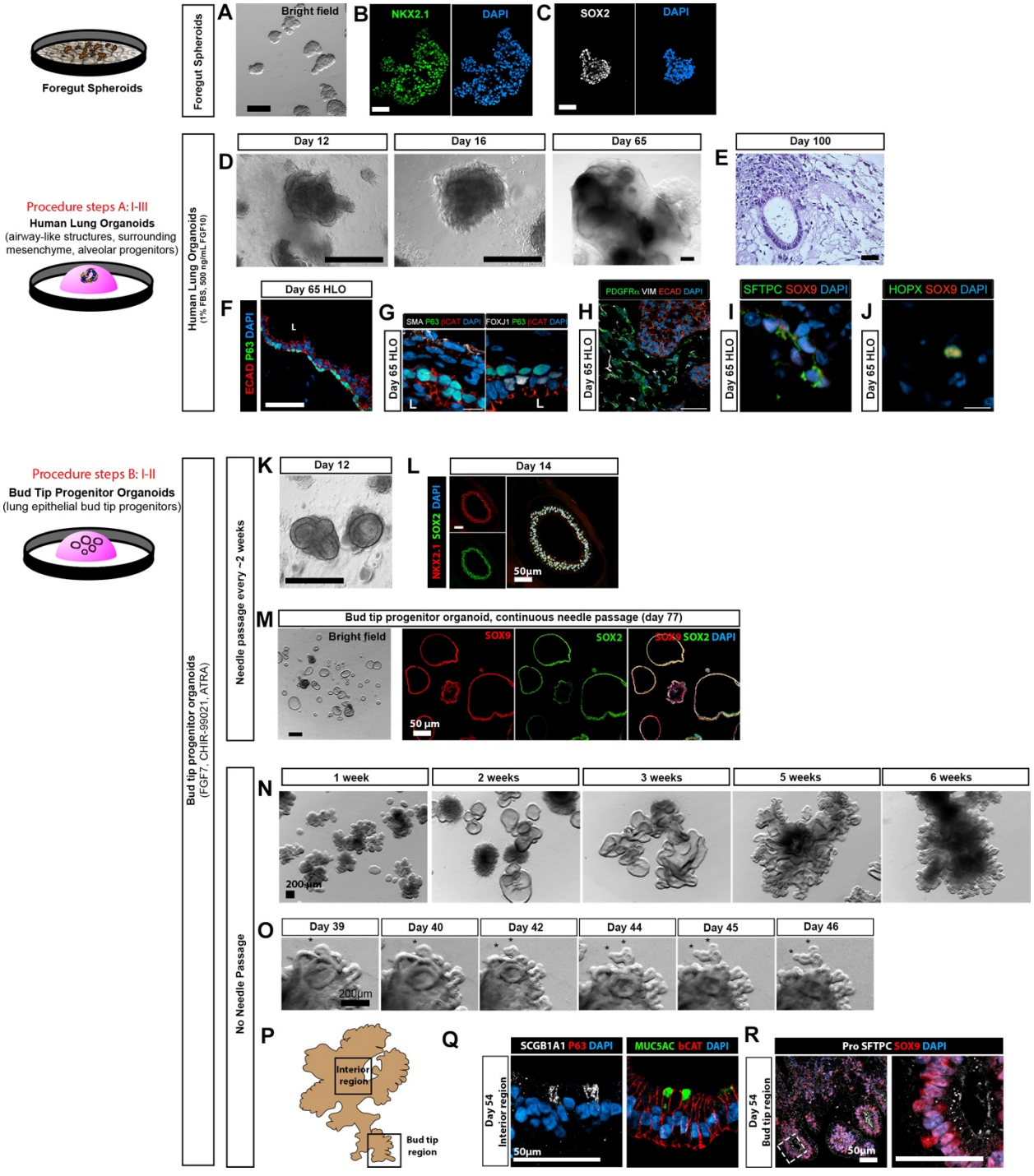


Figure 5: Expected outcomes of the protocol. A) Bright-field image of foregut spheroids on the day of collection cultured in a Matrigel droplet. Scale bar represents 200 μm. B) The overwhelming majority of foregut spheroid cells are NKX2.1+ and C) SOX2+. Scale

bars represent 50 μm. D-J) Expected outcomes of the human lung organoid protocol. D) Bright-field images showing expected growth patterns for human lung organoids. Scale bars represent 500 μm. E) Human lung organoids have persisting epithelial structures and surrounding mesenchyme as shown by Hematoxylin and Eosin staining at day 100 culture. Scale bars represent 500 μm. F) Epithelial structures contain P63+ basal-like cells and an organized airway-like epithelium. Scale bars represent 50 μm. G) Surrounding mesenchyme is positive for smooth muscle marker smooth muscle actin (SMA), and epithelial structures contain P63 positive cells on the basolateral surface with FOXJ1+ cells on the luminal side of the epithelium. P63+ cells composed, on average, roughly 40% of all cells within organoids. Though FOXJ1+ cells composed ~5% of the total cells in the HLOs, mature AcTUB structures were not seen at day 65 and were rare at 85 days (data not shown). Scale bar represents 50 μm. H-J) Scale bar represents 50 μm. H) Cells surrounding day 65 HLO ECAD positive epithelial structures were positive for multiple mesenchymal cell type markers, including PDGFRα and Vimentin. I) Rare cells within HLOs express SOX9 and markers of AECII marker surfactant protein C (SFTPC). These cells make up roughly 5% of all cells within an organoid. J) Rare cells within HLOs express SOX9 and markers of AECI marker HOPX. These cells make up roughly 5% of all cells within an organoid. K-M) Expected outcomes of the bud tip progenitor organoid protocol if organoids are maintained by regular needle passaging. K) Epithelial structures form and can be isolated by day 12 in culture. Scale bar represents 500 μm. L) By day 14 of culture, over 95% of epithelial cells are NKX2.1/SOX2 double positive. Scale bar represents 50 μm. M) Needle passaged Bud tip progenitor organoids have a simple epithelial structure composed of

SOX9+/SOX2+ cells similar to that observed in the native human lung bud tip progenitors prior to 16 weeks gestation. Scale bars represent 50 μm. N-R) Expected outcomes of the bud tip progenitor organoid protocol if organoids are not maintained by needle passage, and are instead allowed to grow intact to complex epithelial structures. N) Expected growth patterns over time of bud tip progenitor organoids. Epithelial structures predictably begin to fold around 3 weeks, and form branched-like structures at 5-6 weeks. Scale bar represents 200 μm, all images taken at the same magnification. O) Brightfield images over 5 days in culture showing bifurcation of an epithelial bud structure. Scale bar represents 200 μm, all images taken at the same magnification. P) Cartoon representation of the interior and bud tip regions of non-needle passaged bud tip progenitor organoids. Q) Interior regions of these structures contain roughly 5% of cells that exhibit SCGB1A1 staining and roughly 2% of cells that stain positive for Goblet cell marker MUC5AC. Scale bar represents 50 μm. R) The bud tip regions of these structures maintain the expression of bud tip progenitor markers SOX9 and pro Surfactant Protein C (SFTPC). Scale bars represent 50 μm. Detailed information about antibodies used for staining can be found in Table 4-3.

Figure 4-5 G-J reproduced with permission from Dye *et al*. 2015 and licensed under CC BY-NC-ND 4.0.

Figure 4-5 L-R reproduced with permission from Miller *et al*. 2018 and licensed under CC BY-NC-ND 4.0.

4.16 Tables

Table 4-1 Comparison and contrast of methods to derive lung organoid model systems

		Bud Tip Progenitor Organoid (routine passaging) **Miller et al. 2018**(17)	**Patterned Lung Organoid (intact structure, no passage)** **Miller et al. 2018**(17)	**Human Lung Organoid** **Dye et al. 2015 and Dye et al. 2016**(15,16)	**Lung Bud Organoid** **Chen et al. 2017**(28)	**Proximal Airway Organoids** **McCauley et al. 2017**(27)	**Alveolospheres** **Jacob et al. 2017**(26)	**Alveolar Organoids** **Gotoh et al. 2014 and Yamamoto et al. 2017**(29,35)	**Airway Epithelial Spheroids** **Konishi et al. 2016**(30)
In vitro	Time required to generate from hPSC	~3 weeks	~6 weeks	~8 weeks	~10 weeks	~3 weeks	~4 weeks	~4 weeks	~4 weeks
	Cell sorting step required	No	No	No	No	Yes	Yes	Yes	Yes
	Mesenchyme present	No	No	Yes	Minimal	No	No	No	No
	Airway epithelial structures	No	Yes	Yes	Yes	Yes	No	No	Yes
	Airway-like cells	No	Yes SCGB1A1+ Club-like and MUC5AC/B+ Goblet-like	Yes FOXJ1+ ACTTUB multiciliated-like cells, SCGB1A1+ club-like cells, P63+	Yes, MUC5AC+ MUC1+ Goblet-like cells	Yes FOXJ1+ ACTTUB+ multiciliated- like cells SCGB3A2 club-like	No	No	Yes FOXJ1+ ACTTUB+ mutli-ciliated cells, SCGB3A2 club-like

				basal-like cells.		cells, P63+ KRT5+ basal-like cells, MUC5AC+ goblet-like cells			cells, KRT5 basal-like cells MUC5AC+ goblet-like cells
	Functional airway cells	No	Yes, mucous secreted into organoid lumens	Not tested	Yes, mucous secreted into organoid lumens	Yes, mucous secreted into organoid lumen, beating multi-ciliated cells, and forskolin-induced swelling	No	No	Yes, mucous secreted into organoid lumen, beating multi-ciliated cells
	Alveolar-like cells	No	Yes AECI and AECII-like cells (FGF7 alone media)	Yes AECI and AECII-like cells	Yes AECII-like cells	Yes AECII-like cells	Yes AECII-like SFTPC+ cells	Yes AECI and AECII-like cells	No
	Functional alveolar cells	No	Not tested	Not tested	Yes, cells uptake tagged SFTPB	Not tested	Yes Functional surfactant production	Not tested	No
	Epithelial budded/branch-like structures	No	Yes	No	Yes	No	No	No	No
	Functional (self-	Yes	Yes	No	Not tested	Not tested	No	No	No

	renewing, multilineage potential) bud tip progenitors								
	Expandable Organoids	Yes	No	No	No	No	Yes	Yes	No
In vivo maturation	Time required for maturation in vivo	6 weeks	Not tested	8 weeks	6 to 18 weeks	Not tested	Not tested	Not tested	Not tested
	Transplantation into mouse fat pat/kidney capsule	Not tested	Not tested	Yes, maturation of adult airway structures and airway cell types including beating cilia and mucous secretion	Yes, maturation of airway structures and cell types including beating cilia and mucous secretion; maturation of Alveolar Type I and II cell types	Not tested	Not tested	Not tested	Not tested
	Transplantation into injured mouse airway	Yes, successful engraftment and differentiation	Not tested	Not tested	Not tested	Not tested	Not tested	Not tested	Not tested
	Best suited for studies exploring:	• Epithelial lineage fate decisions • Fetal lung development • High through put drug/	• Epithelial morphogenesis • Fetal lung development • Pathogen infection • Mucous metaplasia	• Mesenchymal-epithelial interactions during development and disease (e.g.	• Epithelial morphogenesis • Pathogen infection • Mucous metaplasia	• Adult diseases involving mucociliary function (CF, asthma) • Pathoge	• Genetic alveolar disease • High through put drug/ toxicology testing • Pathog	• Genetic alveolar disease • High through put drug/ toxicology testing • Pathogen infection	• Adult diseases involving mucociliary function (CF, asthma) • Pathogen infection

		toxicology testing • Pathogen infection		pulmonary fibrosis) • Pathogen infection		n infection	en infectio n		
	Not well suited for studies exploring:	• Adult epithelial disease modeling (e.g. CF, asthma) • Mesenchy mal-epithelial interactions	• Some adult disease modeling (e.g. CF, asthma) • Mesenchym al-epithelial interactions	• Adult diseases involving mucociliary function (CF, asthma) • High through put drug/toxicol ogy testing	• Adult epithelial disease modeling (e.g. CF, asthma) • Mesenchy mal-epithelial interactions	• Mesenc hymal-epithelial interacti ons • Alveolar diseases • Fetal lung develop ment	• Mesenc hymal-epithelial interacti ons • Airway diseases • Fetal lung develop ment	Mesenchymal- epithelial interactions; Airway diseases	• Mesenchy mal-epithelial interaction s • Alveolar diseases • Fetal lung developm ent

Table 4-2 Media components

Media	Definitive Endoderm day 1	Definitive Endoderm day 2	Definitive Endoderm day 3	Definitive Endoderm day 4	Anterior Foregut Endoderm days 5-9	Human lung organoid maintenance	Bud tip progenitor organoid medium
Basal Media	RPMI 1640	RPMI 1640 0.2% Hyclone FBS	RPMI 1640 2% Hyclone FBS	RPMI 1640 2% Hyclone FBS	(Foregut basal) Advanced DMEM/F12 1X N-2 1X B27 10 mM HEPES buffer 1X L Glutamine (2 mM) 1X Penicillin-Streptomycin (5000 U/mL)	(Foregut basal) Advanced DMEM/F12 1X N-2 1X B27 10 mM HEPES buffer 1X L Glutamine (2 mM) 1X Penicillin-Streptomycin (5000 U/mL)	(Bud tip basal) DMEM F12 1X N-2 1X B27 1X L Glutamine (200 mM) 1X Penicillin-Streptomycin (5000 U/mL) 0.05% Bovine Serum Albumin

Add day-of-use	Activin A, 100 ng/mL	Activin A, 100 ng/mL	Activin A, 100 ng/mL	Activin A, 100 ng/mL	10 µM SB431542 200 ng/mL Noggin 1 µM Smoothened Agonist (SAG) 500 ng/mL FGF4 2 µM CHIR-99021	500 ng/mL FGF10 1% Fetal Bovine Serum	0.4 µM Monothio-glycerol 50 µg/mL Ascorbic Acid 10 ng/mL FGF7 50 nM ATRA 3 µM CHIR-99021

Table 4-3 Antibody information

Primary Antibody	Source	Catalog #	Dilution (Sections)	Clone
Chicken anti-GFP	Abcam	Ab13970	1:500	polyclonal
*Biotin-Goat anti-TP63	R&D systems	BAF1916	1:500	
*Biotin-Mouse anti MUC5AC	Abcam	ab79082	1:500	Monoclonal
Goat anti-CC10 (SCGB1A1)	Santa Cruz Biotechnology	sc-9770	1:200	C-20
Goat anti-Chromogranin A (CHGA)	Santa Cruz Biotechnology	sc-1488	1:100	C-20
Goat anti-SOX2	Santa Cruz Biotechnology	Sc-17320	1:200	polyclonal
Goat anti-VIMENTIN (VIM)	Santa Cruz Biotechnology	sc-7558	1:100	S-20
Mouse anti-Acetylated Tubulin (ACTUB)	Sigma-Aldrich	T7451	1:1000	6-11B-1
Mouse anti-alpha smooth muscle actin (SMA)*Cy3 conjugated	Sigma	C6198	1:400	MonoClonal
Mouse anti-E-Cadherin (ECAD)	BD Transduction Laboratories	610181	1:500	36/E-Cadherin
Mouse anti-FOXJ1	eBioscience	14-9965-82	1:500	2A5
Mouse anti-Human Nuclear Antigen (HuNu)	Abcam	ab191181	1:250	Monoclonal
Mouse anti-Surfactant Protein B (SFTPB)	Seven Hills Bioreagents	Wmab-1B9	1:250	monoclonal
Rabbit anti-HOPX	Santa Cruz Biotechnology	Sc-30216	1:250	polyclonal
Rabbit anti-NKX2.1	Abcam	ab76013	1:200	EP1584Y
Rabbit anti-Nuclear Mitotic Aparatus – Human (NuMa)	Thermo Scientific	PA5-22285	1:500	Polyclonal
Rabbit anti-PDGFRalpha	Santa Cruz Biotechnology	sc-338	1:100	C-20

Rabbit anti-Pro-Surfactant protein C (Pro-SFTPC)	Seven Hills Bioreagents	Wrab-9337	1:500	polyclonal
Rabbit anti-SOX9	Millipore	AB5535	1:500	polyclonal
Rabbit anti-Synaptophysin	Abcam	AB32127	1:500	monoclonal
Secondary Antibody	**Source**	**Catalog #**	**Dilution**	
Donkey anti-goat 488	Jackson Immuno	705-545-147	1:500	
Donkey anti-goat 647	Jackson Immuno	705-605-147	1:500	
Donkey anti-goat Cy3	Jackson Immuno	705-165-147	1:500	
Donkey anti-mouse 488	Jackson Immuno	715-545-150	1:500	
Donkey anti-mouse 647	Jackson Immuno	415-605-350	1:500	
Donkey anti-mouse Cy3	Jackson Immuno	715-165-150	1:500	
Donkey anti-rabbit 488	Jackson Immuno	711-545-152	1:500	
Donkey anti-rabbit 647	Jackson Immuno	711-605-152	1:500	
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102	1:500	
Donkey anti-goat 488	Jackson Immuno	705-545-147	1:500	
Donkey anti-goat 647	Jackson Immuno	705-605-147	1:500	
Donkey anti-goat Cy3	Jackson Immuno	705-165-147	1:500	
Donkey anti-mouse 488	Jackson Immuno	715-545-150	1:500	
Donkey anti-mouse 647	Jackson Immuno	415-605-350	1:500	

Donkey anti-mouse Cy3	Jackson Immuno	715-165-150	1:500	
Donkey anti-rabbit 488	Jackson Immuno	711-545-152	1:500	
Donkey anti-rabbit 647	Jackson Immuno	711-605-152	1:500	
Donkey anti-rabbit Cy3	Jackson Immuno	711-165-102	1:500	
Streptavidin 488	Jackson Immuno	016-540-084	1:500	

4.17 References

1. Zorn AM, Wells JM. Vertebrate Endoderm Development and Organ Formation. Annu Rev Cell Dev Biol. NIH Public Access; 2009 Nov;25(1):221–51.

2. Chang DR, Alanis DM, Miller RK, Ji H, Akiyama H, McCrea PD, et al. Lung epithelial branching program antagonizes alveolar differentiation. Proceedings of the National Academy of Sciences. 2013 Nov 5;110(45):18042–51.

3. Rawlins EL. Lung Epithelial Progenitor Cells: Lessons from Development. Proceedings of the American Thoracic Society. 2008 Aug 15;5(6):675–81.

4. Rawlins EL, Clark CP, Xue Y, Hogan BLM. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development. 2009 Nov;136(22):3741–5.

5. Spence JR, Mayhew CN, Rankin SA, Kuhar MF, Vallance JE, Tolle K, et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature. 2011 Feb 3;470(7332):105–9.

6. McCracken KW, Howell JC, Wells JM, Spence JR. Generating human intestinal tissue from pluripotent stem cells in vitro. Nat Protoc. 2011 Dec;6(12):1920–8.

7. Múnera JO, Sundaram N, Rankin SA, Hill D, Watson C, Mahe M, et al. Differentiation of Human Pluripotent Stem Cells into Colonic Organoids via Transient Activation of BMP Signaling. Cell Stem Cell. 2017 Jul 6;21(1):51–6.

8. Tsai Y-H, Nattiv R, Dedhia PH, Nagy MS, Chin AM, Thomson M, et al. In vitro patterning of pluripotent stem cell-derived intestine recapitulates in vivo human development. Development. Oxford University Press for The Company of Biologists Limited; 2017 Mar 15;144(6):1045–55.

9. Hannan NRF, Sampaziotis F, Segeritz C-P, Hanley NA, Vallier L. Generation of Distal Airway Epithelium from Multipotent Human Foregut Stem Cells. Stem Cells Dev. 2015 Jul 15;24(14):1680–90.

10. Hannan NRF, Fordham RP, Syed YA, Moignard V, Berry A, Bautista R, et al. Generation of multipotent foregut stem cells from human pluripotent stem cells. Stem Cell Reports. 2013;1(4):293–306.

11. Longmire TA, Ikonomou L, Hawkins F, Christodoulou C, Cao Y, Jean JC, et al. Efficient Derivation of Purified Lung and Thyroid Progenitors from Embryonic Stem Cells. Stem Cell. Elsevier Inc; 2012 Apr 6;10(4):398–411.

12. Huang SXL, Islam MN, O'Neill J, Hu Z, Yang Y-G, Chen Y-W, et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. Nature Biotechnology. 2013 Dec 1.

13. Green MD, Chen A, Nostro M-C, d'Souza SL, Schaniel C, Lemischka IR, et al. Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nature Biotechnology. 2011 Feb 27;29(3):267–72.

14. Huang SXL, Green MD, de Carvalho AT, Mumau M, Chen Y-W, d'Souza SL, et al. The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells. Nat Protoc. 2015 Mar;10(3):413–25.

15. Dye BR, Hill DR, Ferguson MA, Tsai Y-H, Nagy MS, Dyal R, et al. In vitro generation of human pluripotent stem cell derived lung organoids. Elife. 2015;4.

16. Dye BR, Dedhia PH, Miller AJ, Nagy MS, White ES, Shea LD, et al. A bioengineered niche promotes in vivo engraftment and maturation of pluripotent stem cell derived human lung organoids. Elife. 2016 Sep 28;5.

17. Miller AJ, Hill DR, Nagy MS, Aoki Y, Dye BR, Chin AM, et al. In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. Stem Cell Reports. 2018 Jan 9;10(1):101–19.

18. Cruz-Acuña R, Quirós M, Farkas AE, Dedhia PH, Huang S, Siuda D, et al. Synthetic hydrogels for human intestinal organoid generation and colonic wound repair. Nat Cell Biol. 2017 Nov;19(11):1326–35.

19. Szenker-Ravi E, Altunoglu U, Leushacke M, Bosso-Lefèvre C, Khatoo M, Tran HT, et al. RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6. Nature. Nature Publishing Group; 2018 May 16;557(7706):564.

20. Hill DR, Huang S, Nagy MS, Yadagiri VK, Fields C, Mukherjee D, et al. Bacterial colonization stimulates a complex physiological response in the immature human intestinal epithelium. Elife. 2017 Nov 7;6.

21. Hill DR, Huang S, Tsai Y-H, Spence JR, Young VB. Real-time Measurement of Epithelial Barrier Permeability in Human Intestinal Organoids. JoVE. 2017 Dec 18;(130).

22. Perrin S. Preclinical research: Make mouse studies work. Nature. 2014 Mar 27;507(7493):423–5.

23. Ruggeri BA, Camp F, Miknyoczki S. Animal models of disease: pre-clinical animal models of cancer and their applications and utility in drug discovery. Biochem Pharmacol. 2014 Jan 1;87(1):150–61.

24. van der Laan JW, Chapin RE, Haenen B, Jacobs AC, Piersma A. Regulatory Toxicology and Pharmacology. Regulatory Toxicology and Pharmacology. Elsevier Inc; 2012 Jun 1;63(1):115–23.

25. D'Amour KA, Agulnick AD, Eliazer S, Kelly OG, Kroon E, Baetge EE. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nature Biotechnology. 2005 Dec;23(12):1534–41.

26. Jacob A, Morley M, Hawkins F, McCauley KB, Jean JC, Heins H, et al. Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. Cell Stem Cell. 2017 Oct 5;21(4):472–488.e10.

27. McCauley KB, Hawkins F, Serra M, Thomas DC, Jacob A, Kotton DN. Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. Cell Stem Cell. 2017 Jun 1;20(6):844–6.

28. Chen Y-W, Huang SX, de Carvalho ALRT, Ho S-H, Islam MN, Volpi S, et al. A three-dimensional model of human lung development and disease from pluripotent stem cells. Nat Cell Biol. 2017 Apr 24;19(5):542–9.

29. Gotoh S, Ito I, Nagasaki T, Yamamoto Y, Konishi S, Korogi Y, et al. Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. Stem Cell Reports. 2014 Sep 9;3(3):394–403.

30. Konishi S, Gotoh S, Tateishi K, Yamamoto Y, Korogi Y, Nagasaki T, et al. Directed Induction of Functional Multi-ciliated Cells in Proximal Airway Epithelial Spheroids from Human Pluripotent Stem Cells. Stem Cell Reports. 2016 Jan 12;6(1):18–25.

31. Hawkins F, Kramer P, Jacob A, Driver I, Thomas DC, McCauley KB, et al. Prospective isolation of NKX2-1-expressing human lung progenitors derived from pluripotent stem cells. J Clin Invest. 2017 Jun 1;127(6):2277–94.

32. Wang D, Haviland DL, Burns AR, Zsigmond E, Wetsel RA. A pure population of lung alveolar epithelial type II cells derived from human embryonic stem cells. Proc Natl Acad Sci USA. 2007 Mar 13;104(11):4449–54.

33. Firth AL, Dargitz CT, Qualls SJ, Menon T, Wright R, Singer O, et al. Generation of multiciliated cells in functional airway epithelia from human induced pluripotent stem cells. Proceedings of the National Academy of Sciences. 2014 Apr 29;111(17):E1723–30.

34. Van Haute L, De Block G, Liebaers I, Sermon K, De Rycke M. Generation of lung epithelial-like tissue from human embryonic stem cells. Respir Res. 2009 Nov 5;10:105.

35. Yamamoto Y, Gotoh S, Korogi Y, Seki M, Konishi S, Ikeo S, et al. Long-term expansion of alveolar stem cells derived from human iPS cells in organoids. Nature Publishing Group. 2017 Nov;14(11):1097–106.

36. Lung Organoids and Their Use To Study Cell-Cell Interaction. 2017;5(2):223–31. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=28596933&retmode=ref&cmd=prlinks

37. Miller AJ, Spence JR. In Vitro Models to Study Human Lung Development, Disease and Homeostasis. Physiology (Bethesda). 2017 May;32(3):246–60.

38. Clevers H. Modeling Development and Disease with Organoids. Cell. 2016 Jun 16;165(7):1586–97.

39. Hannan NRF, Segeritz C-P, Touboul T, Vallier L. Production of hepatocyte-like cells from human pluripotent stem cells. Nat Protoc. 2013 Jan 31;8(2):430–7.

40. Gjorevski N, Sachs N, Manfrin A, Giger S, Bragina ME, Ordóñez-Morán P, et al. Designer matrices for intestinal stem cell and organoid culture. Nature. 2016 Nov 16;539(7630):560–4.

41. Bartfeld S, Clevers H. Organoids as Model for Infectious Diseases: Culture of Human and Murine Stomach Organoids and Microinjection of Helicobacter Pylori. JoVE. 2015;(105):e53359–9.

Chapter 5 Discussion and future directions

5.1 Introduction

In this chapter, I outline the current limitations and big questions in the field of lung developmental biology, review my contributions to the field described within this dissertation, and identify intriguing new directions and propose future experiments based on this work that will help fill some of the most pressing current gaps in the field.

5.2 Overview of gaps of knowledge in the field of lung biology

There are three primary gaps in knowledge that are a major hinderance to furthering our understanding of human lung development and disease. These include a failure of mouse and rodent models to predict human development and disease leading to a need for human cell-based lung model systems, existing problems with human cell-based model systems, and a lack of understanding of basic human lung developmental biology.

First, it has become increasing evident that differences between mouse and human biology make it difficult to draw conclusions about human biological events from mouse studies. This is highlighted by the fact that 80% of drugs shown to be efficacious in mouse models of disease fail during clinical trial(1). This number is even more abysmal for lung specific diseases. For example, only 2 drugs have entered the clinic to

treat idiopathic pulmonary fibrosis, out of over 2,000 that have been shown to be efficacious in mice, leaving a success rate of 0.1%. Some of the primary differences in mouse and human lung development include size and timing of developmental events. While the mouse lung completes the growth of an arborized network of airways in roughly 12 days, human lungs begin branching around 7 weeks gestation and the airways and not complete until 1-2 years post birth. This incredible increase in the number of branches in humans leaves highly complex small airway networks that are susceptible to small airway diseases such as asthma, emphysema and chronic obstructive pulmonary disease (COPD). It also creates a window of vulnerability for premature infants prior to full lung maturation. Furthermore, my work in Chapters 2 and 3 has shown that biological differences exist in the cell states and gene expression patterns in bud tip progenitors and intermediate cells within the lung epithelium, and previous work has shown differences in the localization and identity of stem cells in the adult lung between mice and humans(2-4). All of these factors contribute for a need for human cell-based systems of the human lung to study lung development and disease.

Second, while human cell-based model systems allow for study of human-specific cell behaviors, the in vitro lung model systems as they currently exist have major limitations and are not yet able to fully recapitulate many aspects of lung biology. Some of these limitations include: access to tissue, models that faithfully represent whole organ cell types and function, difficult (e.g. transplantation into a mouse fat pad) protocols for reaching full tissue maturation, reproducibility of experiments due to large variability in donor health and age, and a lack of genetic-based tools to manipulate these systems. For example, one major hurdle in the field is the ability to grow or

maintain alveolar cell types in a system that recapitulates the morphology, structure, and function of alveolar cell types. The inability to do this currently inhibits our ability to study alveolar development and alveolar-specific diseases.

Third, one of the major hinderances to developing better human in-vitro lung model systems is a lack of understanding of human developmental biology based on human fetal lung tissue. To date, only a handful of studies exist that have been carried out using donated human fetal tissue. While this practice is politically fraught, it is necessary in order to provide a benchmark or a 'gold standard' of cell states and transcriptomes present in the native developing lung that can be used in developing hPSC-derived systems. This practice of comparing hPSC-derived tissues to native fetal lung cells will ensure we are creating model systems that have the highest degree of biological similarity to developing human lungs. Generation of more representative human lung models could help identify drug candidates for many fatal illnesses that will pass clinical trial, and may contribute to the identification and testing of many therapies to treat premature infants.

5.3 Contributions of this dissertation work to the field

The goals of this work were, first, to understand the basic biology behind human lung epithelial development, including what factors maintain bud tip progenitors and what environments push bud tip progenitors to differentiate into specific lineages, including basal cells, the stem cells of the adult airway. Second, I sought to generate 3 dimensional in vitro models to study human lung development.

In this dissertation, I've shown data that extensively characterizes the developing human lung (Chapters 2 and 3) (5,6), identified FGF7, WNT, and ATRA as sufficient to maintain bud tip progenitors as progenitors (Chapter 2), shown that SMAD signaling regulates the cell fate transition from a bud tip progenitor to an adult basal stem cell (Chapter 3), and developed two new 3D in vitro models of human bud tip progenitors (Chapter 2 and Chapter 4) (7).

Extensive characterization of the developing human lung epithelium is a major gap in the field of lung biology, and is needed for the progression of a multitude of research questions. I first characterized the developing human lung epithelium using bulk RNA sequencing (Chapter 2), protein and in situ mRNA staining in fetal lungs from 10-20 weeks gestation (Chapters 2 and 3), and single cell RNA sequencing from 11.5-21 weeks gestation (Chapter 3). These data sets, especially the single cell RNA sequencing dataset described in Chapter 3, will be a major resource in the field and will help our lab and other labs benchmark hPSC-derived tissue and formulate new hypotheses about lung development as the field moves forward.

This characterization of the fetal lung led to the identification of two previously unidentified cell states that appear to be specific to the human fetal lung: bud tip adjacent and hub cells. Bud tip adjacent are likely alveolar precursors, as shown in Chapter 3, and hub cells are likely airway secretory precursors. However, additional work is still needed to define the function of these cell states, which is a major future direction described below. This characterization also served to provide transcriptomic and proteomic signatures of bud tip progenitor cells, which I used to help identify major signaling pathways essential for bud tip progenitor maintenance and differentiation. In

Chapter 2 I show that bud tip progenitors (both mouse and human) are maintained primarily by a high Wnt environment that is supported by FGF7 and ATRA. Careful characterization also aided in the discovery, outlined in Chapter 3, that transient activation of SMAD signaling is sufficient to drive bud tip progenitors to a proximal cell fate, by first patterning TP63+/NKX2.1+/SOX2+ progenitors. Further inhibition of SMAD signaling leads to the development of functional basal-like cells and other functional airway cell types such as secretory and ciliated cells.

Together, this progress in understanding the developmental cell states and signaling pathways that give rise to the mature lung has led to the development and improvement of hPSC stem cell-derived lung organoid models. In Chapters 2 and 4 I show that bud tip progenitor organoids can be grown from human fetal bud tip progenitors in a 3D environment in vitro by supplementing them with FGF7, WNT and ATRA. I also showed in Chapter 2 and 4 that treating foregut spheroids(8) with these same growth factors led to the generation of bud tip progenitor organoids from hPSCs. Bud tip progenitor organoids can be generated in ~2 weeks, are highly expandable, and can be maintained long-term in culture, making them ideal for studying what causes bud tip progenitors to differentiate into specific lineages. I further showed in Chapter 2 that these hPSC-derived bud tip progenitor cells could be injected and successfully engrafted into the airways of an injured mice. This is one of the first reports to show the use of hPSC-derived cells for lung repair and regeneration, if not the first, and marks a step forward in the fight against epithelial lung diseases such as Cystic Fibrosis, lung rejection after transplant (bronchiolitis obliterans), and severe lung damage due to flu infection, ventilator-induced injury, or smoke inhalation.

In addition to studying bud tip progenitor organoids in the progenitor state, bud tip progenitor organoids can be differentiated by non-passage, which leads to budded organoids that exhibit proximal-distal patterning, or by adding different growth factors such as dual SMAD activation, which leads to the generation of TP63+ cells and other airway cell types.

Taken together, this body of work has made major contributions to the field. Data sets and characterization of precious human fetal tissue samples will serve as guides for hPSC-derived tissue and benefit developmental biologists in multiple fields. Understanding of bud tip progenitor maintenance and differentiation cues could aid development of therapies to speed up lung development in premature infants, for example, by manipulating wnt signaling in the lung epithelium. Publication of a detailed protocol will enable other labs to easily adopt lung organoid technologies in their labs. And lastly, generation of fetal-derived and hPSC-derived bud tip progenitor organoid model systems provide 3D in vitro platforms to study human-specific lung development. hPSC-derived model system bypasses tissue scarcity issues and increases access to genetic tools that have been developed for hSPCs.

5.4 Future directions

Functions of newly discovered fetal specific cell states and mechanisms guiding their differentiation trajectories

In Chapter 3, I describe the discovery of two previously uncharacterized cell states, hub cells and bud tip adjacent cells. While the description of these cell states by scRNAseq and by protein staining analysis is interesting, learning more about the functions and mechanisms for the generation of these cells and their subsequent differentiation prior to adulthood will add critical new knowledge to the field that may aid in developing better in vitro lung models in the future.

Hub cells are characterized by high expression of the genes SCGB3A2, SFTPB and CFTR. Their location and gene expression suggest they may be progenitor cells that give rise to other secretory cells in the adult airway, and unpublished scRNAseq data from our collaborators has shown that hub cells are absent from the healthy adult lung, bud clearly detectable and highly abundant in diseased lungs. High upregulation of the CFTR gene in this population, along with the other observations, leads to the hypothesis that this cell state might have functional consequences during development and may be reactivated in disease.

Moving forward, it will be critical to develop methods to isolate or genetically target these cells, either by using cell surface antibodies, by generating lentiviral reporters, or by lineage tracing these cells in vitro or in whole lung explants to evaluate their differentiation trajectories. Further, it will be valuable to interrogate the mechanisms that lead to differentiation into hub cells and the mechanisms leading hub cells to differentiate into mature lung cell types. Some of my preliminary evidence shows that NOTCH signaling may play a critical role either in the differentiation into hub cells or in the maintenance of this population in the fetal lung, as activated NOTCH signaling co-localizes with hub cell markers in the developing small airways (Figure 5-1A), and

inhibition of NOTCH in whole explant cultures using DAPT led to a severe reduction in the number of hub cells in explants (Figure 5-1b-c). The role of NOTCH signaling in hub cell maintenance or differentiation could be studied by further developing in vitro whole lung culture methods to permit better expansion and differentiation of the bud tip region. It could also be explored using the hPSC- or fetal-derived bud tip progenitor organoid systems. Barcoded lineage tracing systems have recently been developed(9,10) that allow for unbiased lineage tracing of cells, where lineage relationships can be untangled during post analysis using single cell RNAsequencing. This approach is ideal for studying potential relationships between bud tip progenitors, hub cells, and more differentiated cell types using in vitro model systems of bud tip progenitor organoids. Specifically, bud tip progenitor organoids could be infected with barcoded constructs, treated with 3 days of dual SMAD activation followed by 3-4 weeks of dual SMAD inhibition, and then subjected to single cell RNA-seq. Differentiation trajectories could be reconstructed to evaluate what hub cells give rise to in this culture system. This same approach could be used with whole fetal lung explant culture systems to evaluate the role of hub cells in a more in vivo-like environment, taking into account cross-talk with the mesenchyme or residual immune/endothelial/nervous cells.

Bud tip adjacent cells are another cell state we identified in Chapters 2 and 3. These cells sit directly adjacent to bud tip progenitors in the native developing human lung and have a unique gene and protein expression pattern that is SOX9-/SFTPC-/HOPX+/PDPN+. Isolation of these cells, as shown in Chapter 3, and their subsequent growth in a 3D environment in culture showed that they gave rise primarily to alveolar type II cells. However, generation of alveolar cell types and their appropriate growth in

culture remains a large outstanding question in the field. This finding that bud tip adjacent cells might be alveolar-specific progenitors needs to be supported further by lineage tracing experiments to show the true differentiation potentials of bud tip adjacent cells. Further, culture conditions need to be identified to determine whether these can give rise to AECI and AECIIs both or just AECII.

Specifically, it will be critical to repeat the experiment where bud tip adjacent cells are isolated from fetal lungs, and culture them in a variety of conditions: in airway media (dual SMAD inhibition), commercially available airway media (e.g. Pneumacult), and alveolar support media (e.g. high Wnt), then subjecting the resulting organoids to single cell RNA-seq. The transcriptomes of the in vitro grown organoids can be compared to a combined data set of fetal lung epithelial cells and adult lung epithelial cells, in order to ensure that populations of mature alveolar cell types are present. This experiment will determine whether bud tip adjacent cells always give rise to alveolar-like cells, or whether they retain differentiation potential to give rise to airway cell types under the right conditions. The single cell RNA-seq analysis will further determine whether the resulting cell types share molecular features with mature alveolar cells. In future iterations of this work, bud tip adjacent cells could be isolated from the fetal lung, infected with distinct molecular barcodes, grown in alveolar vs. airway media and then subjected to single cell RNA-seq. This experiment will allow us to elucidate whether bud tip adjacent cells have multilineage potential, or whether they pass through a type II cell-like state before giving rise to type I-like cells.

Generation of human lung in vitro models that mimic the structure and function of a human alveolus

One of the primary hurdles faced by premature infants is severely underdeveloped alveoli. If the physically bulky bud tip progenitors, which grow in a simple columnar epithelium, have not yet differentiated into AECI and AECII cells organized into thin, balloon like structures, gas exchange will be severely hindered. Furthermore, many deadly adult diseases affect the alveolar regions including emphysema/COPD and idiopathic pulmonary fibrosis. Current alveolar lung models either have some alveolar type II cell functionality(11), or recapitulate some aspects of the air-liquid interface seen at the alveolus(12,13). However, neither of these models faithfully recapitulates AECI cell morphology or function, and neither combines functionality with the appropriate cell types.

In order to study differentiation of bud tips into alveolar cell types, we will first need to develop methods to culture and support human alveolar type I and II cells in a way that can faithfully recapitulate both form and function of this critical organ component. One component of a culture system that could be explored is better materials to scaffold the delicate structure of an alveolus. Current methods rely on plastic ‘chips’ or on Matrigel embedded epithelial structures. Matrigel poses many problems to these types of cultures, including the fact that they are fully submerged in gel/liquid, and that Matrigel is an undefined animal-by product that contains varying degrees of bioactive molecules and growth factors. Identifying novel bioengineered materials and methods for culturing alveolar cell types will be a critical next step in the field. For example, bud tip progenitor organoids or isolated bud tip adjacent cells could

be needle passaged and seeded onto porous biological scaffolds (8), and cultured either submerged or on top of a filter at the air liquid interface with differentiation media to promote differentiation and functional maturation of alveolar cell types.

Generation of hPSC-derived human lung in vitro models that can generate functionally mature cells for cell therapies and regenerative medicine

One promising path forward to treat lung epithelial diseases such as Cystic Fibrosis, ciliopathies, bronchiolitis obliterans, flu infection or ventilator induced lung injury is to generate large quantities of human lung-specific stem cells and inject these healthy cells into a damaged lung, where they could engraft and repopulate a damaged airway. My work in Chapter 2 showed that hPSC alone are not sufficient to engraft in an injured mouse lung, and that injecting bud tip progenitor cells led to engraftment but also to goblet-cell fate by many engrafted cells, suggesting a more properly patterned cell may yield better engraftment and repair results. To this end, I tested the ability of DSA-DSI treatment, outlined in Chapter 3, on hPSC-derived bud tip progenitor organoids to elucidate whether bud tip progenitor organoids could give rise to TP63+ basal stem cells (Figures 5-2 and 5-3). Organoids were treated for 3 days with DSA and then expanded for 56 days in DSI (Figure 5-2 a-b). Basal cell marker and early lung epithelial progenitor marker TP63 was highly upregulated after 3 days of DSA treatment (Figure 5-2 c). QRT-PCR was used to evaluate the expression of lung marker NKX2.1 and epithelial marker ECAD over time throughout the protocol (Figure 5-2 d), finding a rise in NKX2.1 over time in culture and a reduction in ECAD expression. Gene expression of other lung epithelial markers over time revealed an increase in TP63,

FOXJ1, SCGB1A1 and HOPX over time in culture (Figure 5-2 e). After 45 days in DSI expansion medium, organoids were found that exhibited TP63+ cells organized along the basolateral surface of a lung-like pseudostratified epithelium, along with expression of lung marker NKX2.1. Positive protein staining was also found for lung epithelial cell markers for multiciliated cells (AcTUB) and secretory cells (SCGB1A1) (Figure 5-2 f).

We then performed scRNA-seq on these organoids to determine the composition and probable cell types. Cells from a single experiment were submitted for sequencing, derived from the H9 cell line. scRNA-seq library preparation, pre-processing, hierarchical clustering, and tSNE visualization were performed as described in Chapter 3 methods. Interestingly, this analysis revealed a population of lung epithelial and mesenchymal cells, but also revealed clusters of other endoderm-derived organs, including the esophagus, stomach, thyroid, and liver (Figure 5-3 a-b).

This preliminary data provides a proof of principle that lung TP63+ cells can be generated from hPSCs, although this method was inefficient and also produced cells of other endoderm lineages. Before we can generate large quantities of organ-specific stem cells for transplant, we must first as a field make progress on a number of critical questions, including:

- Is a transcriptomic signature sufficient to identify a functional cell type?
- During normal human or mammalian development, what mechanisms control the silencing of other same-germ-layer lineages?
- How can we re-create these events in a dish to generate specific cell types from hPSCs? Are growth factors alone enough?
- What is the best way to bench mark these hPSC-generated cells?

Although the above questions lack the glamor of generating cell types for regenerative medicine, they are necessary first-steps towards this goal.

Another critical component of improving airway epithelial cell transplants is to optimize transplantation methods. This could be done using isolated human adult or fetal cell types, including bud tip progenitors, basal cells, secretory cells, or hub cells. Isolated cell types can be injected into naphthalene injured mouse lungs at varying concentrations and with multiple schedules of one-time or repeated rounds of injection. 6-8 weeks after injected, mice could be sacrificed and their airways evaluated for engraftment of human cells, recovery of the injured airway, and the differentiation state of human cells. If a certain cell type (e.g. basal cell or secretory cell or hub cell) appears to perform better than other cell types during these experiments, it will be good preliminary evidence to focus attention on generating that cell type from iPSCs for cell based therapies. It will also be of great interest to the field to generate cell types that are good candidates for cell transplant therapies from iPSC lines that have genetic corrections for genetic defects associated with human disease, such as cystic fibrosis or ciliopathies.

5.5 Concluding remarks

Taken together, this body of work defines cell states, characteristics and timing of human fetal lung development, identifies requirements for bud tip progenitor maintenance and differentiation into a basal cell lineage, and shows novel in vitro models of the human lung from primary tissue and hPSCs. This work opens doors to study previously unidentified cell states within the fetal lung, including the hub cells and bud tip adjacent cells, and to move towards the development of in vitro alveolar models

and improved understanding of directed differentiation techniques to generate specific cell types from hPSC for regenerative medicine.

5.6 Figures

Figure 5-1 NOTCH signaling plays a role in hub cell differentiation or maintenance in the developing human lung

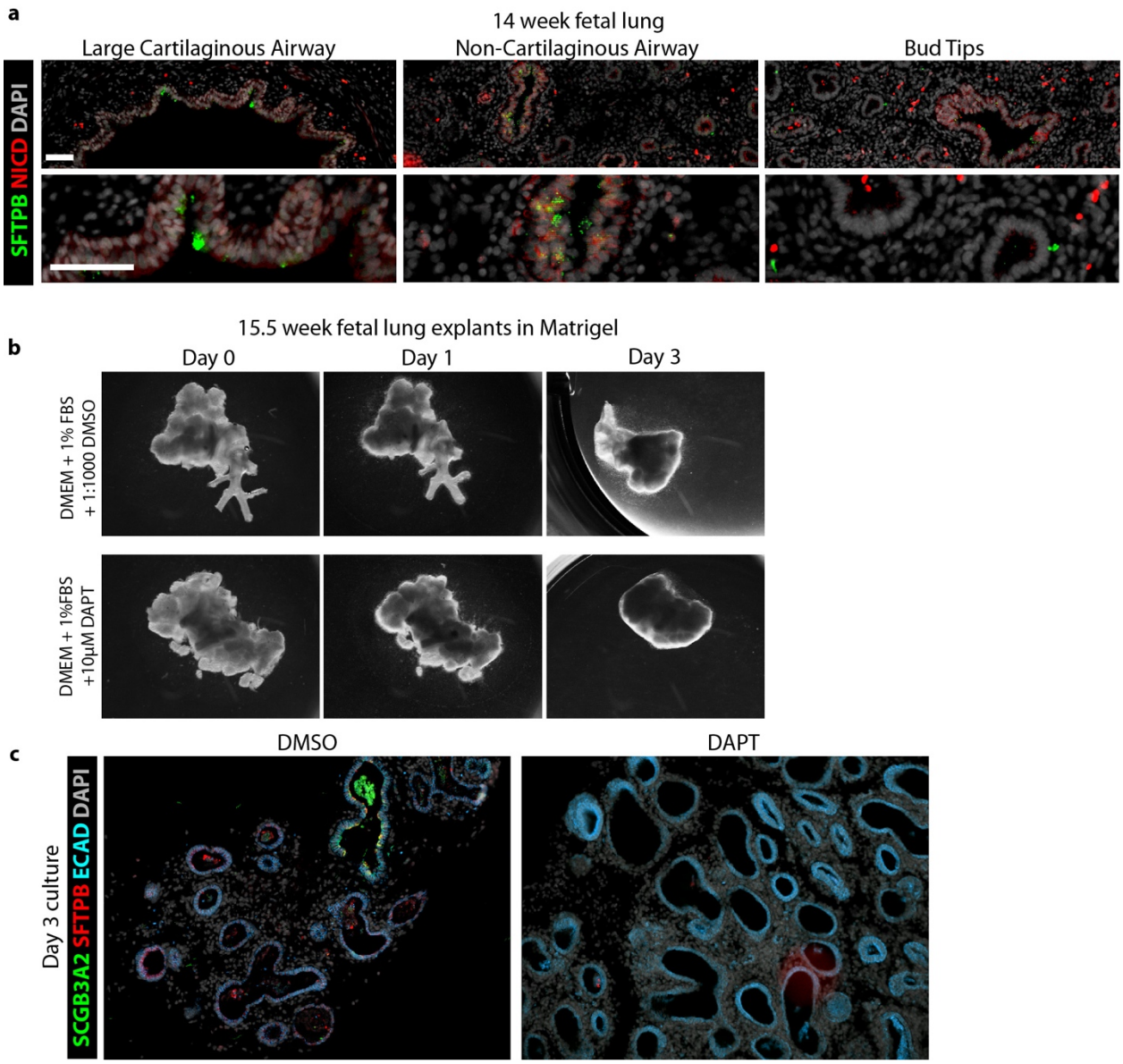


Figure 5-1. a) Protein stain of 14 week fetal lungs showing activated Notch signaling (Notch Intracellular Domain; NICD, red) co-localizes with hub markers SFTPB (green) in

the Non-Cartilaginous airways, where hub cells are abundant, but does not co-localize with SFTPB+ cells in the upper cartilaginous airways. Panel A generated in collaboration with Ansley Semack. Scale bars represent 50μm. B) 15.5 week fetal lung explants roughly 1cm^2 in size were cultured for 3 days in a Matrigel droplet with either control media or NOTCH inhibition conditions (DAPT). C) Protein stain of cultured lung explants for hub cells (SCGB3A2, green/SFTPB, red).

Figure 5-2 Generation of basal stem cell-like cells and airway organoids from hPSCs by following development events

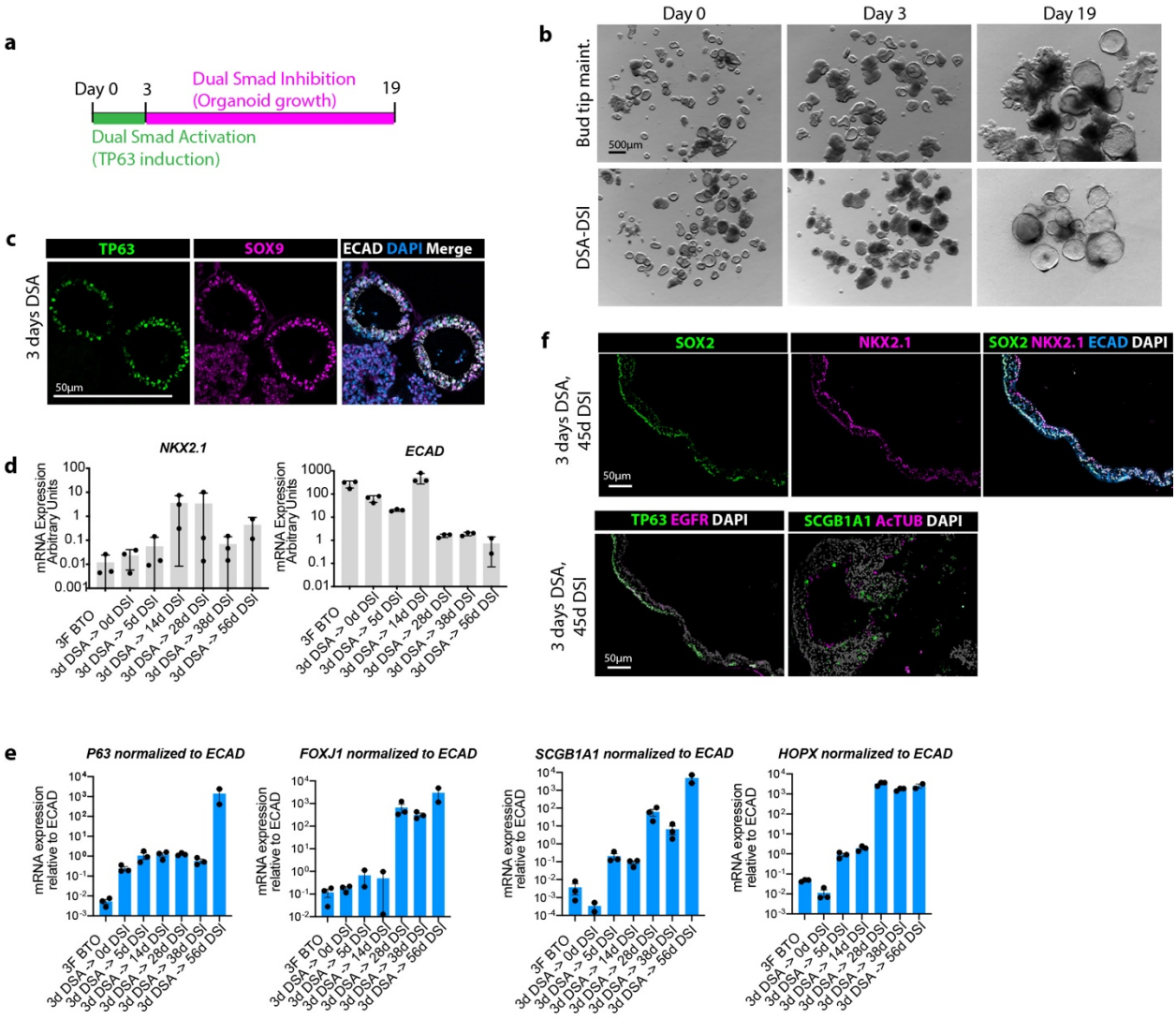


Figure 5-2. a) Schematic of experimental design. b) Brightfield images of hPSC-derived bud tip progenitor organoids cultured in bud tip protenitor maintenance conditions (control) or with dual SMAD activation for 3 days followed by expansion in dual SMAD inhibition conditions. c) Protein staining after 3 days for basal cell marker and proximal progenitor marker TP63 (green) and progenitor marker SOX9 (pink). d) QRT-PCR of

organoids maintained in 3F as a control, and throughout the experimental protocol of DSI and expansion in DSI for lung epithelial marker NKX2.1 and epithelial marker ECAD. Each dot represents a pooled well of 10-25 organoids from a single experiment. Error bars show the mean +/- the SEM. e) QRT-PCR of organoids maintained in 3F (control) and throughout the experimental protocol. Expression of each gene is normalized to the expression of ECAD to control for the relative reduction of ECAD expression and epithelial structures over time in the cultures. f) Protein staining after 45 days in expansion culture medium for proximal lung marker SOX2 (green) and lung epithelial marker NKX2.1 (pink), top panel. Bottom panel, protein stain after 45 days in organoid expansion conditions showing TP63+/EGFR+ basal cells lining the basolateral surface of a pseudostratified epithelium (left panel, TP63 green), and the presence of other airway specific cell types such as multiciliated cells (AcTUB, pink) and secretory cells (SCGB1A1, green).

Figure 5-3 scRNA-seq of hPSC-derived DSA-DSI treated lung organoids reveals cell fate confusion

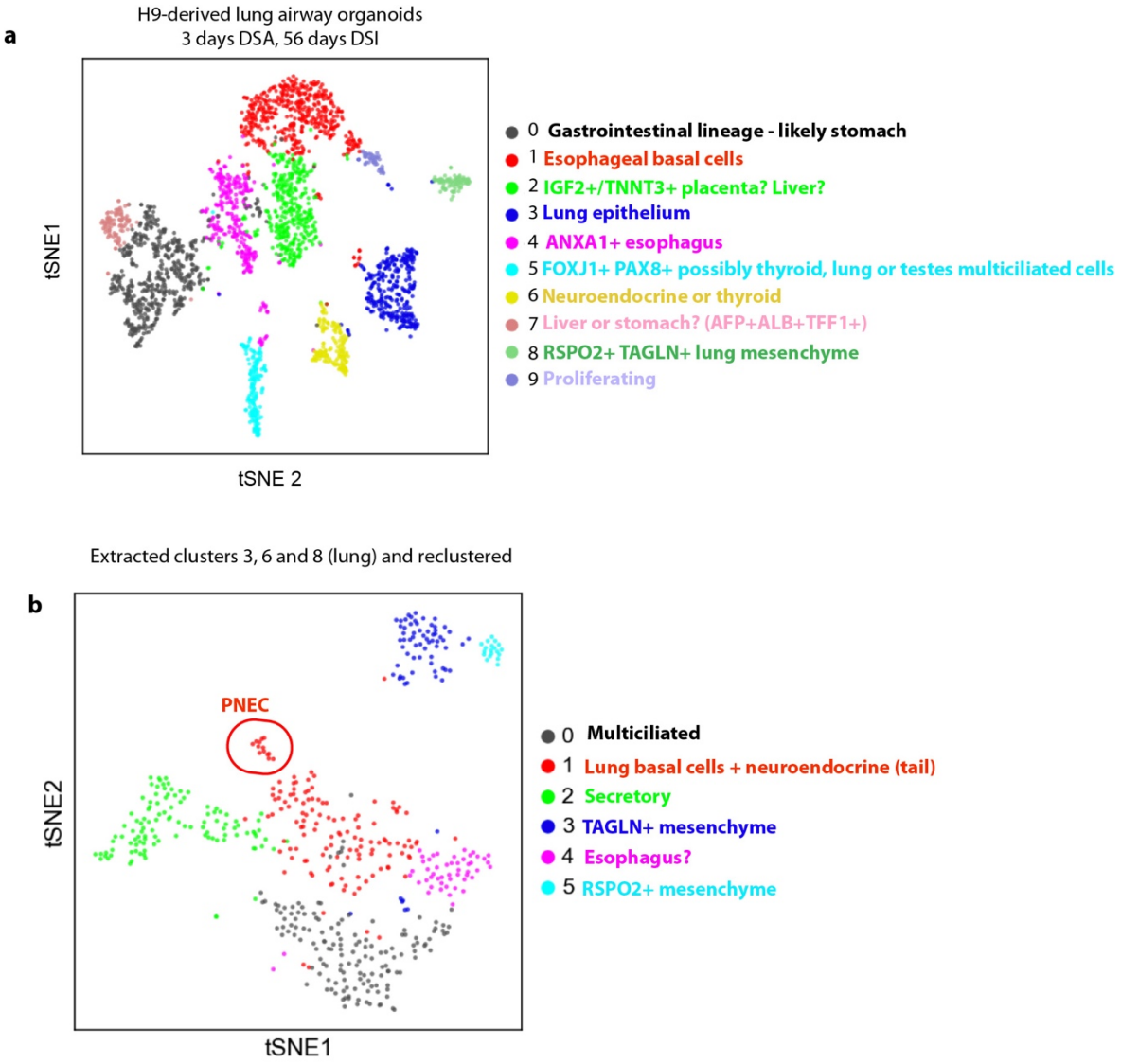


Figure 5-3. a) scRNA-seq of hPSC-derived bud tip progenitor organoids treated for 3 days with dual SMAD activation (DSA) followed by 56 days of expansion in dual SMAD inhibition (DSI). Hierarchical clustering revealed 10 clusters visualized here by t Stochastic Neighbor Embedding (tSNE). Cluster identities were estimated based on the top gene markers for each cluster (table 5-1). Many clusters exhibited gene expression patterns found in lung, as well as other endoderm lineage organs including esophagus,

thyroid, liver and stomach. b) Clusters with a high likelihood of representing lung were extracted and re-clustered. The resulting 6 clusters were identified based on cluster marker genes shown in table 5-2. Note the presence of both lung epithelial cells (clusters 0, 1, 2) and lung mesenchyme (clusters 3, 5).

5.7 Tables

Table 5-1 cluster marker genes for scRNA-seq of hPSC-derived organoids after 3 days DSA and 56 days DSI treatment

names	scores	logfoldchanges	pvals	pvals_adj	cluster #
SPINK1	60.868465	1.955386	2e-323	6.4683e-319	0
FXYD3	40.925602	0.9806598	1.45E-246	1.58E-242	0
TFF2	40.894436	1.5545392	6.26E-192	4.10E-188	0
AGR2	39.737614	1.4154602	2.86E-215	2.34E-211	0
TSPAN8	38.609875	3.2378898	3.82E-176	1.39E-172	0
CST3	38.136463	0.97215015	5.81E-263	9.50E-259	0
TFF1	37.559914	1.7385668	3.82E-174	1.14E-170	0
CTSE	36.53988	3.6249735	6.71E-164	1.37E-160	0
C12orf75	36.334476	2.0795434	2.01E-185	9.38E-182	0
LYZ	35.99354	2.861548	6.18E-164	1.37E-160	0
STARD10	34.030025	1.1779437	2.18E-186	1.19E-182	0
CYSTM1	34.025303	1.6794171	6.69E-164	1.37E-160	0
TPT1	33.403137	0.22728154	1	1	0
B2M	32.674034	0.6118988	9.19E-171	2.51E-167	0
GOLM1	31.844599	1.7103369	4.43E-155	8.05E-152	0
EPCAM	31.509975	1.2060524	3.80E-157	7.32E-154	0
VSIG2	31.40716	2.484752	1.05E-139	1.71E-136	0
PRSS1	30.451296	2.146877	6.34E-132	8.65E-129	0
CHP1	30.338882	1.6356393	2.56E-144	4.41E-141	0
C19orf33	29.694433	0.81076515	2.99E-174	9.78E-171	0
CLDN18	29.62956	3.257709	7.88E-126	1.03E-122	0
SPINT2	29.622885	0.60903215	2.13E-180	8.72E-177	0
SLC44A4	29.07487	2.1531863	5.09E-135	7.57E-132	0
TFF3	28.597244	0.80537504	1.55E-133	2.21E-130	0
AGR3	28.122614	1.0775981	1.01E-166	2.55E-	0

				163	
				3.90E-	
QSOX1	27.405365	1.8550988	3.58E-119	116	0
				2.14E-	
CD63	27.319187	0.6949524	1.37E-136	133	0
				7.44E-	
AGPAT2	27.305687	1.8228623	6.59E-120	117	0
				3.09E-	
PRSS3	27.220472	2.580264	3.68E-112	109	0
				2.82E-	
MSMB	27.160896	2.8772151	3.45E-111	108	0
				3.72E-	
PRSS8	26.964294	1.9136342	3.64E-118	115	0
				2.16E-	
LIPH	26.94587	2.1125653	2.31E-116	113	0
				2.04E-	
ANG	26.696259	2.0672073	2.12E-116	113	0
				7.25E-	
PHGR1	26.68107	3.163172	9.97E-108	105	0
				1.01E-	
FCGRT	25.85456	1.4006978	1.14E-114	111	0
				6.03E-	
CD3G	25.785624	2.0310705	8.11E-109	106	0
				3.40E-	
FTL	25.69716	0.30042756	2.80E-124	121	0
				3.42E-	
AMN	25.600765	1.8857266	4.49E-109	106	0
				1.88E-	
ERBB3	25.501196	1.83594	2.41E-109	106	0
				6.35E-	
PLAC8	25.491734	1.5064222	6.41E-118	115	0
				2.40E-	
TSPAN13	25.448864	1.1933595	2.27E-118	115	0
				7.20E-	
NDRG1	25.330303	1.14375	8.36E-113	110	0
				1.67E-	
SPINT1	24.952257	1.4507645	2.10E-109	106	0
SDCBP2	24.901562	2.5972176	1.01E-100	6.78E-98	0
S100P	24.752327	1.9298226	2.09E-99	1.34E-96	0
				4.62E-	
SMIM22	24.660912	1.2670702	3.67E-125	122	0
MUC1	24.659029	2.757759	2.57E-97	1.59E-94	0
HNMT	23.941402	2.787084	3.91E-94	2.29E-91	0
				1.55E-	
S100A10	23.765013	0.70838165	2.19E-105	102	0
TMEM45B	23.488667	1.9432753	5.78E-95	3.44E-92	0
				5.01E-	
APLP2	23.44181	0.9973869	7.20E-105	102	0
CDH17	23.36512	4.615108	1.60E-88	8.29E-86	0
				4.99E-	
TMSB4X	23.281704	0.17563313	5.49E-115	112	0
GRN	23.225668	1.2049319	2.14E-98	1.35E-95	0
CEACAM5	22.962847	3.1910925	4.32E-87	2.18E-84	0
VSIG1	22.859476	3.8886094	4.33E-86	2.05E-83	0
TSPAN1	22.479475	1.2139803	2.27E-95	1.38E-92	0
SLC44A1	22.440033	1.7763096	2.00E-89	1.07E-86	0

ELF3	22.272484	0.70872855	9.79E-102	6.67E-99	0
IMPA2	22.123539	1.9378104	3.91E-86	1.88E-83	0
KRT8	22.067606	0.6304932	5.67E-100	3.71E-97	0
MUC5B	22.009441	2.9955275	8.06E-82	3.61E-79	0
TMEM176A	21.781069	1.556786	2.80E-86	1.39E-83	0
CD68	21.732553	1.952971	4.47E-84	2.06E-81	0
LAMB3	21.649803	1.6885933	8.00E-85	3.74E-82	0
RBM47	21.608437	1.3406014	2.95E-86	1.44E-83	0
DUOX2	21.399078	3.2809935	5.34E-78	2.08E-75	0
PPARG	21.39581	3.549954	1.83E-78	7.40E-76	0
TMEM92	21.322447	2.7355044	7.46E-79	3.09E-76	0
FOSL2	21.235205	1.5974839	1.57E-81	6.84E-79	0
CD164	21.23488	0.67937875	6.36E-93	3.59E-90	0
GATA6	21.165451	3.2903287	1.37E-77	5.17E-75	0
PPIC	21.053778	2.29078	2.23E-78	8.88E-76	0
CAPN8	21.015408	3.633058	4.30E-76	1.54E-73	0
LINC01133	21.004604	4.185357	1.33E-75	4.59E-73	0
PRR15L	20.988916	3.2482708	3.14E-76	1.14E-73	0
FZD5	20.871634	3.1226447	6.12E-76	2.16E-73	0
TIMP1	20.862408	1.1920884	1.06E-81	4.69E-79	0
GDF15	20.631714	1.6600097	2.50E-77	9.21E-75	0
SERINC2	20.594765	1.7946671	2.43E-77	9.05E-75	0
TMBIM1	20.57651	1.6905441	1.08E-77	4.17E-75	0
APOA1	20.545605	3.1465821	2.19E-73	7.16E-71	0
HPGD	20.525803	2.959559	2.22E-73	7.18E-71	0
RHOB	20.444239	1.6713315	4.48E-76	1.59E-73	0
ANXA10	20.360329	3.916801	3.08E-72	9.60E-70	0
PRR13	20.343473	0.9379578	7.89E-81	3.40E-78	0
TM9SF3	20.20244	0.9800435	3.96E-79	1.66E-76	0
GPR160	20.196264	2.525202	5.97E-73	1.88E-70	0
RNF128	20.175575	3.8952477	2.02E-71	6.13E-69	0
SLC39A5	19.970924	3.7685006	2.14E-70	6.30E-68	0
CLDN4	19.93819	0.7867111	5.67E-80	2.41E-77	0
BMP2	19.840946	3.5214005	9.35E-70	2.68E-67	0
TMEM141	19.83679	0.9477685	1.67E-75	5.64E-73	0
POLD4	19.797104	1.0676298	1.36E-75	4.65E-73	0
LGALS2	19.560791	4.284257	1.19E-67	3.16E-65	0
FABP1	19.433098	2.706332	4.00E-67	1.04E-64	0
COX7C	19.253113	0.15871415	2.79E-75	9.31E-73	0
SERPINB1	19.246944	1.4859607	4.85E-70	1.41E-67	0
CDH1	19.24304	1.098609	1.16E-71	3.58E-69	0
UPK1B	19.212341	2.315458	1.18E-67	3.16E-65	0
MIR205HG	51.859547	1.5574914	1.27E-236	4.15E-232	1
RPLP1	48.008224	0.22652665	3.60E-196	5.90E-192	1
RPL10A	42.361282	0.31694883	7.66E-148	3.58E-144	1
RPS7	39.576347	0.27468514	6.52E-148	3.56E-144	1
RPS9	38.981834	0.26542294	4.44E-146	1.82E-142	1
RPS4X	38.254436	0.24460407	1.58E-138	4.71E-135	1
KRT19	37.6086	0.63922864	9.64E-159	1.05E-154	1

RPS3	37.579407	0.22148676	9.51E-136	2.22E-132	1
RPL15	37.220886	0.1943507	1.15E-132	2.22E-129	1
NPM1	36.33778	0.45319268	1.33E-157	1.09E-153	1
RPS14	35.69846	0.16490912	1.40E-135	3.05E-132	1
KRT15	35.3334	2.4532447	1.29E-143	4.71E-140	1
RPS15A	35.33226	0.1803123	7.60E-131	1.32E-127	1
RPL7	35.114563	0.21709876	1.93E-135	3.94E-132	1
PDLIM1	34.68335	1.0644586	1.42E-152	9.32E-149	1
GNB2L1	33.92066	0.26764604	7.66E-131	1.32E-127	1
RPS15	33.852047	0.18146518	4.32E-136	1.09E-132	1
RPS2	33.36002	0.19905199	3.35E-129	5.23E-126	1
RPS8	32.986565	0.17730035	9.65E-123	1.37E-119	1
BTF3	32.643017	0.4102606	8.89E-141	2.91E-137	1
RPL34	32.379368	0.161298	1.81E-123	2.69E-120	1
SNCG	32.376945	2.4547274	2.06E-130	3.36E-127	1
RPL6	32.238873	0.24424823	7.43E-117	8.10E-114	1
RPL13	32.201935	0.14679193	1.10E-121	1.38E-118	1
RPL13A	31.994755	0.12362921	1.49E-113	1.44E-110	1
HNRNPA1	31.940905	0.4475977	1.01E-136	2.76E-133	1
RPS3A	31.465023	0.15678297	1.85E-113	1.68E-110	1
RPL35A	31.410727	0.16142498	3.18E-122	4.16E-119	1
RPL12	31.398794	0.22249961	7.30E-116	7.71E-113	1
RPL32	31.17384	0.16403091	1.60E-113	1.50E-110	1
SERPINF1	30.813314	2.3186328	1.17E-122	1.59E-119	1
RPL19	30.776241	0.17314051	2.71E-119	3.17E-116	1
RPL10	30.683144	0.16172518	1.76E-104	1.23E-101	1
RPL3	30.653055	0.19144736	2.08E-121	2.53E-118	1
RPS23	30.419827	0.16361514	7.73E-110	6.03E-107	1

RPL11	30.246517	0.15844299	2.43E-112	2.15E-109	1
RPS24	29.847511	0.1601002	9.30E-108	6.92E-105	1
RPL5	29.824116	0.21094063	8.86E-109	6.75E-106	1
RPS18	29.788174	0.15895796	4.09E-105	2.91E-102	1
RPL18A	29.784204	0.16561979	1.12E-111	9.66E-109	1
RPL23	29.652674	0.24244756	4.73E-118	5.34E-115	1
EIF3E	29.354055	0.42470267	2.74E-114	2.72E-111	1
RPS25	29.151497	0.1597642	2.74E-111	2.30E-108	1
RPS12	28.875679	0.17494683	5.49E-101	3.60E-98	1
PTMA	28.292604	0.23187959	6.58E-115	6.73E-112	1
S100A2	28.139038	2.1229749	6.59E-111	5.26E-108	1
RPS27A	27.47188	0.13623892	1.83E-101	1.25E-98	1
RPS5	27.405775	0.1985189	4.72E-101	3.15E-98	1
RPL37	26.87756	0.16470344	1.47E-96	8.59E-94	1
RPL21	26.689766	0.12703598	5.05E-96	2.90E-93	1
TNNT3	26.631138	1.2306689	1.46E-106	1.06E-103	1
RPL30	26.62824	0.16382743	1.49E-97	8.87E-95	1
RPL26	26.318243	0.14274351	1.39E-90	6.79E-88	1
RPL31	26.085361	0.1444435	1.12E-95	6.31E-93	1
BCAM	26.021317	2.0624852	3.60E-98	2.18E-95	1
KRT5	25.861088	2.2265892	2.72E-98	1.68E-95	1
RPL39	25.736227	0.1355417	1.04E-91	5.43E-89	1
RPL18	25.325642	0.15959221	2.87E-91	1.43E-88	1
EGR1	25.289118	0.6856466	5.28E-100	3.32E-97	1
RPL37A	25.265217	0.1352248	2.28E-91	1.15E-88	1
RPS28	24.884295	0.14293876	5.88E-90	2.79E-87	1
RPL7A	24.758865	0.18201816	1.75E-95	9.71E-93	1
RPSA	24.741327	0.23911518	7.82E-94	4.27E-91	1
RPL9	24.653875	0.14236481	2.16E-90	1.04E-87	1
DLK1	24.651857	1.7284662	1.85E-93	9.92E-91	1
RPS6	24.574713	0.13833745	5.00E-81	2.02E-78	1
RPL14	24.095985	0.17126688	6.04E-83	2.67E-80	1
RPS19	24.09591	0.11873612	1.47E-82	6.34E-80	1
PITX1	23.084267	1.266424	1.15E-86	5.22E-84	1
PLP2	23.060823	0.6902583	2.13E-87	9.80E-85	1
RPL35	23.053574	0.14478882	1.11E-82	4.85E-80	1
AQP3	22.969856	1.3045702	4.52E-86	2.03E-83	1
IGF2	22.926361	0.36725375	6.17E-79	2.38E-76	1
RPL22	22.707895	0.17458427	8.43E-81	3.36E-78	1
NACA	22.652113	0.17254752	6.90E-82	2.90E-79	1
RPL27A	22.551743	0.10679802	1.50E-77	5.57E-75	1
RPL24	22.480137	0.14863543	2.99E-81	1.22E-78	1
RPL29	22.44016	0.12984528	1.54E-79	6.06E-77	1
RPS13	22.122498	0.13322431	1.17E-76	4.20E-74	1
CIRBP	22.059977	0.4188855	9.13E-82	3.78E-79	1

RPL28	21.717024	0.12586768	7.20E-77	2.65E-74	1
RPL23A	21.70353	0.12383228	3.24E-73	1.08E-70	1
RPL41	21.686234	0.07838144	6.28E-74	2.16E-71	1
IER2	21.625975	0.6063618	3.79E-79	1.48E-76	1
RPS20	21.138182	0.17848457	1.66E-73	5.61E-71	1
RPS27	20.96412	0.08986468	7.79E-70	2.47E-67	1
LMO4	20.71028	0.86379075	1.13E-73	3.85E-71	1
S100A11	20.5597	0.35949075	6.67E-71	2.18E-68	1
EBPL	19.601198	0.9868929	3.85E-67	1.17E-64	1
H3F3B	19.58953	0.2080286	4.84E-68	1.51E-65	1
NBEAL1	19.466911	0.38900205	2.00E-67	6.18E-65	1
RPL36A	19.319988	0.22952254	1.58E-65	4.54E-63	1
SNHG8	19.1682	0.4722381	8.14E-66	2.38E-63	1
RP11-65J3.1	19.062225	1.8884945	9.66E-63	2.59E-60	1
RPL8	18.820454	0.11386246	5.33E-62	1.41E-59	1
HINT1	18.783785	0.19595586	1.31E-63	3.59E-61	1
S100A14	18.368793	0.84349895	3.69E-61	9.45E-59	1
KRT4	18.081264	2.0008466	9.83E-58	2.30E-55	1
GPC3	18.065624	1.4942791	1.29E-58	3.16E-56	1
FILIP1	17.792292	2.7466047	1.62E-55	3.46E-53	1
ZFAS1	25.032148	0.3333819	5.64E-99	1.85E-94	2
RPL14	22.51412	0.20089677	1.36E-93	2.23E-89	2
RPL10	21.995943	0.16182698	1.90E-88	2.08E-84	2
PERP	21.093712	0.52780694	9.10E-78	4.26E-74	2
RPL31	20.95007	0.14981309	3.00E-80	1.96E-76	2
ALDOA	20.422478	0.5098063	4.37E-68	1.43E-64	2
ANXA1	20.369385	0.8922307	1.69E-68	6.16E-65	2
EIF1	19.254107	0.2799377	1.78E-61	3.23E-58	2
RPL18A	18.747042	0.14158711	1.41E-66	4.18E-63	2
RPL18	18.70908	0.16002701	1.56E-65	3.92E-62	2
NACA	18.474066	0.18914002	2.94E-62	5.65E-59	2
RPS18	18.458202	0.14604855	1.01E-63	2.21E-60	2
RPL27A	17.827694	0.11289673	6.15E-61	1.01E-57	2
RPS13	17.801937	0.15220203	9.70E-59	1.32E-55	2
RPL41	17.674896	0.08894496	8.60E-59	1.22E-55	2
IGF2	17.604702	0.38836935	7.42E-59	1.10E-55	2
FAU	17.53433	0.16635634	2.75E-55	3.00E-52	2
VDAC2	17.528526	0.6619072	3.58E-53	3.25E-50	2
RPL8	17.113922	0.13443418	1.34E-55	1.56E-52	2
RPL24	17.10664	0.15479468	1.09E-54	1.08E-51	2
RPS28	16.925625	0.12777156	3.45E-56	4.19E-53	2
ERO1L	16.875303	1.1597977	3.31E-50	2.41E-47	2
RPL22	16.869534	0.17893176	5.04E-54	4.71E-51	2
RPL34	16.673492	0.11758925	3.03E-54	2.92E-51	2
RPL35A	16.506395	0.12292723	2.35E-52	2.08E-49	2
EIF3E	16.383385	0.36271003	1.93E-50	1.47E-47	2
RPS25	16.307604	0.12565422	1.05E-51	8.62E-49	2
HILPDA	16.172829	1.5370139	5.56E-46	3.25E-43	2
RPL37	16.143229	0.13726917	1.22E-51	9.71E-49	2
SLC2A1	15.918136	1.1540555	1.08E-45	6.20E-43	2
RPL32	15.747694	0.12313322	6.24E-49	4.25E-46	2
ANKRD37	15.676654	2.1104293	2.66E-43	1.30E-40	2
RPL26	15.504665	0.12941974	4.34E-47	2.68E-44	2
RPS24	15.500697	0.12530604	2.33E-47	1.46E-44	2
RPS5	15.473938	0.15401658	5.74E-48	3.76E-45	2
C4orf3	15.162422	0.37289023	5.69E-44	2.91E-41	2

RPL30	14.988114	0.1280057	3.36E-45	1.80E-42	2
RPL9	14.803163	0.12258774	2.60E-43	1.29E-40	2
RPS20	14.598103	0.16331655	8.53E-43	4.10E-40	2
RPS14	14.424401	0.10919704	3.58E-41	1.56E-38	2
LDHA	14.41216	0.66302675	2.49E-39	9.83E-37	2
ANXA2	14.404663	0.55324966	3.49E-39	1.36E-36	2
UBA52	14.390763	0.15305479	5.72E-41	2.43E-38	2
RPLP2	14.358998	0.10104397	2.27E-41	1.00E-38	2
WDR45B	13.956704	0.6825644	1.83E-37	6.16E-35	2
POLR1D	13.93676	0.5811875	2.07E-37	6.84E-35	2
RP11-841O20.2	13.757456	1.4366165	6.52E-36	1.92E-33	2
EIF3H	13.752802	0.4293812	4.81E-37	1.53E-34	2
RPS27	13.692906	0.08577059	4.27E-38	1.52E-35	2
PIM3	13.63135	1.3886935	3.80E-35	1.07E-32	2
ENO1	13.495923	0.62355924	1.61E-35	4.68E-33	2
RPL15	13.472493	0.11094958	1.65E-37	5.62E-35	2
RPL35	13.364556	0.11975403	2.03E-36	6.28E-34	2
MIF	13.298162	0.25871307	1.85E-35	5.31E-33	2
RPS3A	13.248332	0.10430426	4.76E-36	1.43E-33	2
TPT1	13.186407	0.10268214	2.21E-37	7.23E-35	2
SRCAP	13.18621	1.5095544	2.30E-33	5.87E-31	2
RPL39	13.174682	0.10459886	1.15E-35	3.37E-33	2
NDRG1	13.128437	0.6817836	1.42E-34	3.97E-32	2
RPL36	12.963941	0.10440496	7.73E-34	2.09E-31	2
PRSS22	12.866274	1.6992006	6.92E-32	1.63E-29	2
RPL29	12.844959	0.10505793	4.23E-34	1.15E-31	2
RAC1	12.819444	0.35974482	1.45E-32	3.56E-30	2
RPS15A	12.757884	0.10467491	9.49E-34	2.53E-31	2
RPL11	12.651891	0.10368048	3.61E-33	8.96E-31	2
RPS4X	12.608639	0.12883176	2.28E-33	5.87E-31	2
RPL13A	12.55133	0.06991839	1.10E-33	2.90E-31	2
GAPDH	12.490434	0.28919563	8.36E-31	1.78E-28	2
MKNK2	12.169249	1.1275598	1.63E-29	3.27E-27	2
EGLN3	12.159132	1.0888512	1.26E-29	2.55E-27	2
GPI	12.007775	1.041947	6.72E-29	1.28E-26	2
RPL21	11.807971	0.08926513	2.51E-29	4.92E-27	2
RPS27A	11.672969	0.08521469	4.61E-29	8.83E-27	2
EEF1A1	11.660203	0.08654461	2.58E-28	4.60E-26	2
SH3YL1	11.621527	0.6425189	1.18E-27	1.98E-25	2
RPL7	11.55354	0.10757878	1.14E-28	2.13E-26	2
CLDN4	11.482251	0.6207902	3.58E-27	5.66E-25	2
RPS12	11.421098	0.10359316	3.04E-28	5.36E-26	2
RPL17	11.391446	0.16590382	1.41E-27	2.34E-25	2
RPL38	11.216695	0.09081364	1.20E-26	1.84E-24	2
NFKBIA	11.132274	0.86995894	1.53E-25	2.16E-23	2
RPS29	11.114081	0.08045404	1.12E-26	1.73E-24	2
RPL13	11.100404	0.07192532	5.73E-27	8.93E-25	2
RPS6	11.079342	0.09110085	6.64E-27	1.03E-24	2
BTF3	11.052241	0.20008716	5.50E-26	7.99E-24	2
TAF1D	11.006735	0.70602137	3.47E-25	4.77E-23	2
NPM1	10.803779	0.21215768	5.44E-25	7.27E-23	2
ADM	10.766749	1.5666052	6.33E-24	8.07E-22	2
RPL4	10.730817	0.14441492	5.12E-25	6.90E-23	2
RPS23	10.685013	0.08330998	3.14E-25	4.35E-23	2
RPL36AL	10.642993	0.15937085	2.32E-24	3.06E-22	2

PGK1	10.566235	0.5441838	1.11E-23	1.37E-21	2
ATP5G2	10.481979	0.18593703	2.11E-23	2.55E-21	2
SNX3	10.340002	0.29721156	6.79E-23	7.96E-21	2
GNB2L1	10.261668	0.12560317	3.35E-23	4.00E-21	2
MAL2	10.165412	0.71258825	5.45E-22	5.91E-20	2
RPL36A	10.138413	0.17064987	1.92E-22	2.16E-20	2
PLP2	10.126367	0.45858935	4.82E-22	5.27E-20	2
RPS16	10.059672	0.09188272	4.05E-22	4.51E-20	2
YBX3	10.007866	0.5486172	1.43E-21	1.51E-19	2
IGF2	37.075832	0.57631063	8.87E-287	2.90E-282	3
MT-CO3	32.90985	0.3279145	1	1	3
MT-CO1	29.968157	0.24612048	1	1	3
S100A6	27.103432	0.35591814	3.09E-139	3.37E-135	3
S100A11	25.867397	0.48819774	1.59E-115	7.44E-112	3
MT-ATP6	25.348377	0.37658355	2.00E-139	3.27E-135	3
MT-CYB	25.159685	0.32456583	1	1	3
MT-ND3	24.861345	0.32782063	1	1	3
TNNT3	24.345633	1.1292766	1.24E-90	3.68E-87	3
RPS28	24.344793	0.15275514	2.30E-120	1.88E-116	3
RPS18	24.32018	0.15734228	9.83E-117	5.36E-113	3
RPL10A	24.153755	0.21871099	1	1	3
RPL10	23.434782	0.13825978	1	1	3
AQP3	23.395634	1.4490108	5.44E-80	9.89E-77	3
RPS4X	23.34361	0.18032773	3.78E-117	2.47E-113	3
MT-CO2	23.268122	0.22153473	1	1	3
MT-ND1	23.144342	0.26601523	1	1	3
RPL13	22.678722	0.11912875	2.44E-107	9.99E-104	3
HSPB1	22.660395	0.5981727	2.81E-81	5.41E-78	3
RPS12	22.569736	0.154113	1	1	3
SLC25A6	22.124174	0.36743638	4.68E-86	1.09E-82	3
RPL34	21.811216	0.12832715	2.72E-96	9.91E-93	3
RPL26	21.087706	0.12968092	1	1	3
RPL41	21.076471	0.08692154	6.33E-89	1.73E-85	3
RPL32	20.820408	0.1264337	6.56E-95	2.15E-91	3
RPL37	20.560196	0.14581883	2.89E-88	7.28E-85	3
MT-ND4	20.51605	0.22316052	1	1	3
RPS14	20.484703	0.11696507	5.65E-85	1.23E-81	3
MT-ND2	20.316088	0.2282362	1	1	3
ANXA1	19.579248	0.74706626	2.71E-67	2.68E-64	3
PITX1	19.577175	1.1603166	7.49E-64	6.45E-61	3
RPL28	19.38999	0.12071689	9.02E-77	1.34E-73	3
RPS24	19.101707	0.12005581	4.88E-79	7.99E-76	3
RPL37A	19.057661	0.11749616	4.99E-75	7.11E-72	3
TACSTD2	18.906166	0.8802939	6.32E-61	4.59E-58	3
S100A9	18.899216	1.9809206	9.88E-56	6.07E-53	3
CSTA	18.864384	2.1423752	1.86E-56	1.17E-53	3
PERP	18.688835	0.4101955	5.49E-70	6.65E-67	3
KRT13	18.523252	2.102131	2.62E-54	1.48E-51	3

RPL29	18.489128	0.12083836	1.99E-69	2.33E-66	3
IER2	18.396164	0.53895414	1.09E-60	7.72E-58	3
RPL15	18.156656	0.11415596	1	1	3
RPL12	18.057539	0.15544347	8.46E-69	8.93E-66	3
RPS23	17.829973	0.11037721	1	1	3
RPS29	17.706491	0.10604781	1.18E-63	9.65E-61	3
PPDPF	17.698032	0.27051017	1.64E-58	1.10E-55	3
KRT19	17.495298	0.39118236	1.25E-59	8.69E-57	3
RPL14	17.412046	0.14362775	2.30E-65	2.09E-62	3
RPL30	17.360554	0.1197174	7.54E-66	7.05E-63	3
RPL3	17.279812	0.12132371	1.03E-63	8.65E-61	3
RPL7	17.239695	0.12627845	1.49E-64	1.32E-61	3
RPL18A	17.236948	0.11566082	4.27E-62	3.33E-59	3
RPS9	17.189915	0.15153491	5.34E-62	4.06E-59	3
RPL8	17.123774	0.1107287	2.38E-61	1.77E-58	3
RPL27A	16.923847	0.08865181	5.94E-63	4.74E-60	3
RPSA	16.901604	0.17811993	2.58E-59	1.76E-56	3
RPL31	16.195217	0.10049503	2.15E-57	1.38E-54	3
RPL13A	16.145395	0.0758704	1.83E-57	1.20E-54	3
RPL19	16.090597	0.11514269	1.31E-52	7.03E-50	3
FAM3B	16.090424	1.3662786	1.73E-45	7.55E-43	3
EIF3E	16.047804	0.27740914	7.41E-54	4.11E-51	3
RPS2	16.003489	0.11646055	2.19E-54	1.26E-51	3
ZFP36L1	15.973936	0.57329476	5.83E-48	2.73E-45	3
RPS8	15.955069	0.10585899	1.18E-54	6.92E-52	3
SDC1	15.934042	1.4438888	6.75E-45	2.83E-42	3
RPL23A	15.927393	0.10399801	7.30E-55	4.35E-52	3
RPL35	15.888945	0.11755107	4.75E-52	2.47E-49	3
NACA	15.643723	0.12937132	3.02E-52	1.59E-49	3
RPS15A	15.640723	0.10380917	1.26E-51	6.42E-49	3
TMSB4X	15.049631	0.11477038	4.97E-47	2.23E-44	3
TPT1	14.955504	0.10140127	1	1	3
RPS5	14.928145	0.13022988	9.61E-48	4.43E-45	3
RPLP2	14.909063	0.08348084	2.78E-48	1.32E-45	3
RPS3	14.839389	0.10984667	5.71E-49	2.75E-46	3
RPL11	14.827107	0.09427802	1.15E-47	5.25E-45	3
RPS20	14.790689	0.14141867	2.47E-46	1.09E-43	3
RPS13	14.647524	0.10502353	1.99E-45	8.58E-43	3
DEFB1	14.094306	2.0658972	3.29E-36	9.98E-34	3
RAB25	14.034076	0.6939525	9.10E-38	3.04E-35	3
RPL9	13.807816	0.09108438	1.47E-41	5.54E-39	3
RPL5	13.738741	0.11733373	6.19E-42	2.41E-39	3
RPS3A	13.700582	0.08403686	1.05E-41	3.99E-39	3
ALCAM	13.575503	0.855655	1.36E-35	3.98E-33	3
RPS15	13.506649	0.09414771	1.89E-38	6.43E-36	3
RPS7	13.441924	0.1290601	9.32E-39	3.24E-36	3
LINC01088	13.316772	0.93723094	2.04E-34	5.75E-32	3
GNB2L1	13.228925	0.12584108	4.53E-39	1.59E-36	3
RPL39	13.165408	0.08761119	2.19E-37	7.09E-35	3
SMAGP	13.151391	1.3671789	4.32E-33	1.13E-30	3
RPL27	13.093229	0.10031558	3.68E-36	1.10E-33	3
KLF4	13.001174	0.76976436	3.02E-33	7.97E-31	3
RPS25	12.990775	0.08910248	2.29E-36	6.99E-34	3
KRT7	12.964136	2.0158215	6.59E-32	1.63E-29	3
CLDN10	12.880778	1.6052285	9.43E-32	2.29E-29	3
RPL36A	12.840249	0.17704023	7.50E-35	2.15E-32	3

MUC16	12.824875	1.4136411	7.81E-32	1.92E-29	3
RPS10	12.810772	0.11478685	6.59E-35	1.91E-32	3
SPRR3	12.783237	2.430473	4.40E-31	1.02E-28	3
HMGA1	12.690899	0.48674402	9.30E-33	2.41E-30	3
RPL7A	12.641447	0.10541653	3.31E-34	9.17E-32	3
VIM	38.151382	1.3474516	2.34E-149	7.66E-145	4
TUBA1A	30.202652	1.9245687	2.11E-103	1.38E-99	4
ID3	29.943733	2.2407393	4.59E-96	1.67E-92	4
TUBB2B	28.51339	2.8651443	1.51E-91	3.72E-88	4
PSAP	27.91542	1.3541609	1.59E-91	3.72E-88	4
GPM6B	27.075771	3.2731714	4.09E-83	6.08E-80	4
RBP1	26.860409	2.1052668	1.49E-86	2.56E-83	4
PCSK1N	26.037169	2.6652987	9.08E-83	1.29E-79	4
ID4	25.947107	2.1229584	5.47E-81	7.16E-78	4
IFI27L2	25.647701	1.8780738	7.39E-81	9.30E-78	4
TMEM47	24.692158	3.687744	2.75E-74	2.50E-71	4
PFN2	24.60546	1.4867201	1.01E-78	1.17E-75	4
IGFBP5	24.474136	1.6917413	3.23E-78	3.65E-75	4
SPARC	23.876019	2.3021295	1.55E-75	1.49E-72	4
CLU	23.366959	2.408972	7.17E-71	5.10E-68	4
NTRK2	22.971159	2.453985	3.26E-69	2.09E-66	4
CD99	22.902723	0.9220391	1.95E-72	1.59E-69	4
RFX4	21.977043	4.948031	4.40E-64	2.44E-61	4
RGMA	21.878096	3.527376	4.66E-64	2.54E-61	4
ATP1B2	21.721754	4.2538104	2.52E-63	1.31E-60	4
TTYH1	21.711718	4.2972827	3.72E-63	1.90E-60	4
SOX2	21.609613	1.4703971	3.50E-66	2.12E-63	4
METRN	21.10199	2.4855678	2.95E-62	1.46E-59	4
MAP1B	21.088896	2.1876419	6.44E-64	3.46E-61	4
C4orf48	21.08876	1.8959968	2.24E-64	1.29E-61	4
PDLIM3	21.035204	4.8882875	1.28E-60	5.81E-58	4
NLRP1	20.756016	3.8013868	8.18E-60	3.43E-57	4
DKK3	20.531368	3.509377	3.64E-59	1.49E-56	4
RCN2	20.120737	1.3512747	4.80E-60	2.10E-57	4
CALM2	19.800308	0.6696688	3.87E-62	1.89E-59	4
PEA15	19.737484	3.234751	3.44E-56	1.24E-53	4
SPARCL1	19.65642	4.0210876	1.08E-55	3.81E-53	4
QKI	19.562656	2.2107387	2.70E-56	9.81E-54	4
NOVA1	19.354202	2.7873611	4.08E-56	1.45E-53	4
MALAT1	18.814903	0.22019136	5.14E-58	2.00E-55	4
CRB2	18.77018	5.758175	3.78E-52	1.14E-49	4
NFIA	18.716417	2.2974398	3.99E-53	1.29E-50	4
MLF1	18.587545	1.76885	3.70E-54	1.22E-51	4
MEIS2	18.471394	1.4506651	5.57E-53	1.75E-50	4
HOTAIRM1	18.23033	1.6270924	1.81E-51	5.39E-49	4
TPPP3	18.120525	1.7050477	4.20E-53	1.35E-50	4
CNN3	18.092644	1.8289719	9.42E-51	2.78E-48	4
ID2	17.912712	1.8826343	2.38E-50	6.82E-48	4
NGFRAP1	17.864094	0.797612	5.12E-53	1.63E-50	4
NCAM1	17.468313	3.4664867	8.89E-48	2.44E-45	4
CDH2	17.279118	4.712829	1.01E-46	2.74E-44	4
MSI1	16.966686	3.4839694	1.00E-45	2.66E-43	4
MEST	16.690619	1.9305097	1.14E-45	3.01E-43	4
FABP7	16.608511	4.2498164	3.26E-44	8.09E-42	4
NDFIP1	16.603832	1.4985684	3.64E-45	9.46E-43	4

UCHL1	16.480627	2.9443004	2.72E-44	6.80E-42	4
NFIB	16.318123	2.104435	6.51E-44	1.58E-41	4
FEZ1	16.120295	3.2381244	1.14E-42	2.63E-40	4
AK1	16.053705	2.6195312	2.10E-42	4.85E-40	4
TCEAL3	15.466072	1.7927306	5.36E-41	1.18E-38	4
FAM213A	15.383076	1.5447793	1.29E-40	2.72E-38	4
TIMP2	15.224552	2.696628	2.51E-39	4.99E-37	4
TUBA1B	15.171136	0.7654182	5.45E-41	1.19E-38	4
EDNRB	15.080652	5.073821	1.96E-38	3.73E-36	4
TCEAL4	15.068275	0.98502165	3.14E-40	6.46E-38	4
SRP9	15.061904	0.50476736	5.79E-41	1.25E-38	4
SLC1A3	15.051258	2.2752502	6.69E-39	1.31E-36	4
LYPD1	15.010363	4.590542	3.16E-38	5.94E-36	4
B3GNT1	14.802601	2.9168828	1.13E-37	2.07E-35	4
TMEM66	14.716279	0.69705635	8.56E-39	1.65E-36	4
SOD1	14.604506	0.59108335	5.26E-39	1.04E-36	4
C1orf61	14.581018	3.5912971	1.08E-36	1.85E-34	4
CRIP2	14.545071	1.8449926	4.06E-37	7.26E-35	4
AP1S2	14.505245	2.651332	1.20E-36	2.05E-34	4
FXYD6	14.498737	2.1133373	6.69E-37	1.18E-34	4
HSBP1	14.490043	0.4702122	7.65E-39	1.48E-36	4
RHEB	14.386663	0.89943343	2.39E-37	4.37E-35	4
MLLT11	14.379721	1.7469821	3.72E-37	6.72E-35	4
HOXA5	14.357041	3.102457	4.15E-36	6.86E-34	4
GCSH	14.350936	1.4034543	1.22E-36	2.07E-34	4
PRMT2	14.343958	1.1249079	7.29E-37	1.27E-34	4
RCN1	14.34212	1.4274576	1.55E-36	2.63E-34	4
TIMP3	14.28681	3.2566469	1.07E-35	1.74E-33	4
PLTP	14.144022	2.231027	2.29E-35	3.66E-33	4
HOXA3	14.073186	2.4756904	4.14E-35	6.46E-33	4
CETN2	13.987148	1.1129097	4.36E-36	7.16E-34	4
FHL1	13.932909	3.2061353	1.92E-34	2.86E-32	4
GABARAPL2	13.930618	0.67724967	1.25E-35	2.02E-33	4
HTRA1	13.836823	4.3336997	6.83E-34	9.76E-32	4
WSB1	13.748079	0.5681405	7.65E-35	1.18E-32	4
RPS27L	13.684023	0.6605328	1.04E-34	1.60E-32	4
PLA2G16	13.673897	1.6870716	6.16E-34	8.84E-32	4
BTG3	13.658028	1.3316094	7.78E-34	1.10E-31	4
STRAP	13.654089	0.9888031	2.68E-34	3.95E-32	4
SOX11	13.653029	2.0691452	4.71E-34	6.83E-32	4
DCLK1	13.627069	3.9833379	3.43E-33	4.75E-31	4
NTN1	13.572859	2.7342553	5.25E-33	7.25E-31	4
OAZ2	13.532486	1.3333995	2.21E-33	3.08E-31	4
DDR1	13.40214	1.322851	5.46E-33	7.50E-31	4
KIF1A	13.384892	3.2905154	1.97E-32	2.69E-30	4
AEBP1	13.191062	3.8521173	1.39E-31	1.82E-29	4
CKB	13.114635	0.51527196	1.37E-33	1.93E-31	4
NSMF	13.042727	3.0384939	4.08E-31	5.16E-29	4
CDO1	13.040968	4.146623	5.18E-31	6.47E-29	4
NR2F1	12.997394	2.369801	2.96E-31	3.80E-29	4
FAM183A	72.69346	3.6752775	1.62E-152	2.65E-148	5
C20orf85	69.31349	3.9424129	5.00E-142	5.46E-138	5
CAPS	69.186516	2.5815058	8.42E-155	2.76E-150	5

PIFO	61.675274	3.3800669	1.34E-136	8.79E-133	5
C9orf116	59.72702	3.6495812	7.93E-131	4.33E-127	5
C9orf24	58.845528	3.399298	9.98E-131	4.67E-127	5
C5orf49	56.375828	3.4787152	6.47E-130	2.65E-126	5
DYNLL1	55.750324	1.3155217	1.94E-139	1.59E-135	5
TPPP3	53.168476	2.935012	2.66E-126	9.68E-123	5
DYNLRB2	53.04523	4.5526013	1.16E-116	2.92E-113	5
RSPH1	52.34778	3.5593379	1.37E-118	3.74E-115	5
FAM229B	50.698524	3.066431	1.60E-119	4.77E-116	5
CETN2	50.018135	2.248417	1.06E-124	3.47E-121	5
C1orf192	47.82076	3.6160958	7.65E-113	1.67E-109	5
RP11-356K23.1	46.400597	4.245063	2.50E-106	5.12E-103	5
ODF3B	45.42978	3.914027	1.72E-105	3.32E-102	5
SMIM22	43.257107	1.9803568	2.12E-114	4.96E-111	5
KIF9	42.721546	2.913805	2.64E-104	4.80E-101	5
ROPN1L	42.090733	4.6429625	1.73E-98	2.47E-95	5
MORN2	41.98575	2.939801	6.50E-101	1.06E-97	5
ZMYND10	41.597652	3.883678	1.05E-99	1.64E-96	5
EFCAB1	40.38676	4.0519176	1.85E-96	2.32E-93	5
C9orf135	39.55615	5.0182586	1.69E-93	1.85E-90	5
TUBB4B	39.330887	2.0124512	4.03E-98	5.50E-95	5
C1orf194	38.95979	4.424011	1.07E-92	1.10E-89	5
UCP2	38.610233	2.50598	5.52E-96	6.69E-93	5
CRIP1	38.268948	3.162516	7.21E-93	7.61E-90	5
CCDC170	38.162468	5.062359	1.17E-90	1.10E-87	5
MLF1	37.592396	2.439581	1.22E-97	1.60E-94	5
TMEM190	37.481632	3.9892318	7.92E-90	7.01E-87	5
CALM1	37.209717	1.1196948	1.50E-102	2.58E-99	5
PSENEN	37.175896	2.380774	1.23E-92	1.22E-89	5
DPY30	37.095894	1.885787	1.32E-95	1.54E-92	5
PPIL6	37.014923	3.5670755	2.61E-90	2.38E-87	5
LRRC23	36.93203	3.1642225	9.01E-91	8.67E-88	5
POLR2I	36.86356	1.7938464	1.22E-94	1.38E-91	5
CDHR3	36.566242	4.3561735	5.69E-88	4.71E-85	5
DNAAF1	36.344097	4.8958907	4.05E-87	3.16E-84	5
MB	36.280342	3.5760612	5.76E-88	4.71E-85	5
WDR38	36.16159	4.9244323	1.01E-86	7.32E-84	5
HMGN3	36.12996	1.0056942	9.46E-99	1.41E-95	5
SNTN	36.062996	4.99471	2.87E-86	2.00E-83	5
CCDC78	36.033913	4.3240695	6.77E-87	5.15E-84	5
C14orf142	35.577675	3.5520127	8.61E-87	6.40E-84	5

ENKUR	35.456276	4.5271	6.32E-86	4.31E-83	5
LRRIQ1	35.023643	3.7565122	2.37E-86	1.69E-83	5
FANK1	34.036804	3.7648425	6.81E-84	4.21E-81	5
TCTEX1D2	33.88285	2.9900959	3.35E-84	2.11E-81	5
SPAG6	33.685246	4.1074643	4.53E-83	2.56E-80	5
CCDC146	33.518284	4.3847857	1.42E-81	7.49E-79	5
FAM174A	33.2972	2.54398	1.00E-84	6.69E-82	5
FOXJ1	33.11042	3.6641517	4.89E-82	2.71E-79	5
CRNDE	33.08085	2.694046	9.16E-84	5.45E-81	5
CCDC176	33.02599	4.0113025	3.77E-81	1.96E-78	5
SPATA18	32.784843	4.349983	3.19E-80	1.58E-77	5
FAM216B	32.67957	5.2850995	2.67E-79	1.26E-76	5
PLAC8	32.46876	1.7946666	1.16E-88	9.96E-86	5
WDR16	32.237377	4.6074157	5.22E-79	2.41E-76	5
HDLBP	32.07686	1.4697157	1.16E-84	7.58E-82	5
CCDC74A	31.92095	4.1075945	1.55E-78	6.78E-76	5
TMEM14B	31.808422	1.3294034	3.13E-87	2.50E-84	5
MORN5	31.545547	4.275051	2.38E-77	9.74E-75	5
FAM81B	31.385593	3.7247438	2.11E-78	9.08E-76	5
DYNLT1	31.276022	1.2890301	3.59E-83	2.10E-80	5
C11orf88	31.20453	4.5148416	1.90E-76	7.26E-74	5
CCDC153	30.998962	4.131126	1.91E-76	7.26E-74	5
TSTD1	30.984663	1.6155105	9.09E-82	4.96E-79	5
CDS1	30.858137	2.3295865	7.79E-79	3.44E-76	5
DMKN	30.5992	1.870428	1.53E-79	7.35E-77	5
RABL5	30.512194	2.4608963	1.10E-77	4.62E-75	5
TEKT2	30.409014	4.1243353	4.40E-75	1.57E-72	5
CAPSL	30.180021	4.616875	4.47E-74	1.48E-71	5
CIB1	30.038742	1.4815493	3.96E-80	1.93E-77	5
TMEM231	29.953043	4.2231097	5.09E-74	1.67E-71	5
DNAH9	29.936022	4.668563	1.25E-73	3.99E-71	5
HN1L	29.872913	2.2281365	1.15E-76	4.60E-74	5
UQCR11-1	29.872852	0.66949284	7.11E-84	4.31E-81	5
UBL5	29.846233	0.6528807	2.76E-84	1.77E-81	5
FAM166B	29.680561	4.781959	4.77E-73	1.47E-70	5
RABL2B	29.676445	2.9164972	6.94E-75	2.42E-72	5
ENDOG	29.660166	3.782237	1.65E-73	5.19E-71	5
CCDC17	29.64002	4.714557	9.13E-73	2.79E-70	5
AGR3	29.634102	1.3485067	1.04E-81	5.59E-79	5
UFC1	29.347607	1.3782005	2.18E-77	9.05E-75	5
PRDX5	29.254898	1.0033934	5.53E-76	2.08E-73	5
CFDP1	29.123615	1.5479943	7.30E-76	2.71E-73	5
RP4-666F24.3	28.82588	5.2195144	8.57E-71	2.42E-68	5
IQCG	28.745516	3.7487533	1.53E-71	4.57E-69	5
FAM92B	28.61013	4.243594	8.49E-71	2.42E-68	5
TNFAIP8L1	28.452911	4.379506	2.61E-70	7.23E-68	5
CCDC11	28.278002	4.868351	1.33E-69	3.53E-67	5
DNALI1	27.92758	3.1912758	6.16E-70	1.68E-67	5
IK	27.710806	2.0701718	5.14E-72	1.54E-69	5
HSBP1	27.692215	0.85610384	1.84E-74	6.27E-72	5
AC013264.2	27.420715	5.4765825	2.52E-67	6.11E-65	5
CCDC19	27.39204	4.351949	8.02E-68	1.97E-65	5
ARL3	27.283678	2.022048	7.30E-71	2.11E-68	5
CCDC74B	26.96005	4.830542	1.81E-66	4.26E-64	5
EFCAB10	26.923574	4.8208537	2.07E-66	4.83E-64	5

NUDC	26.890816	1.772056	3.93E-70	1.08E-67	5
BEX1	23.446606	4.3343415	5.42E-53	1.77E-48	6
VAMP2	21.919428	1.6159265	1.45E-50	2.38E-46	6
MAP1B	21.873817	2.664323	4.23E-50	3.12E-46	6
TUBB2B	21.828108	2.7775536	4.07E-50	3.12E-46	6
PCSK1N	21.141232	2.7648976	2.52E-48	1.37E-44	6
MLLT11	20.759045	2.939668	5.82E-47	2.38E-43	6
TUBA1A	20.48868	1.8839046	1.68E-47	7.84E-44	6
NGFRAP1	19.362146	0.9693484	6.44E-46	2.34E-42	6
STMN1	19.330097	1.7641577	3.05E-44	8.32E-41	6
STMN2	18.85965	5.4237576	8.00E-42	1.75E-38	6
BEX2	18.29711	2.3174455	4.65E-41	9.50E-38	6
C4orf48	17.65753	2.03115	8.28E-40	1.51E-36	6
DCX	17.046614	5.9781833	3.14E-37	4.90E-34	6
HN1	16.434414	1.6410221	2.85E-36	3.89E-33	6
BASP1	15.65578	2.1611655	4.80E-34	5.42E-31	6
MARCKSL1	15.298321	1.0833428	1.75E-33	1.68E-30	6
STMN4	15.030055	5.3135576	5.57E-32	3.72E-29	6
MIAT	14.917024	3.9593737	9.25E-32	5.94E-29	6
TUBB	14.597002	0.98831135	1.28E-31	7.93E-29	6
TERF2IP	14.212248	1.9956018	3.95E-30	2.12E-27	6
TCEAL4	13.704204	1.1495091	3.68E-29	1.80E-26	6
GNG3	13.629695	5.672058	3.20E-28	1.26E-25	6
MORF4L2	13.578051	0.6379846	6.98E-29	3.05E-26	6
RAB3A	13.535487	4.016918	5.08E-28	1.96E-25	6
CAMK2N1	13.395716	2.6858492	8.67E-28	3.19E-25	6
CCNI	13.259678	0.4732313	6.73E-28	2.50E-25	6
BEX5	13.167342	3.6755686	4.93E-27	1.51E-24	6
SOX11	13.115094	2.9173682	5.84E-27	1.74E-24	6
ATP6V1G2	13.114184	4.9242086	7.61E-27	2.17E-24	6
RTN1	12.989186	1.206264	8.21E-27	2.30E-24	6
APLP1	12.97052	3.3119593	1.47E-26	3.89E-24	6
BEX4	12.792578	1.2056994	2.03E-26	5.24E-24	6
TTC3	12.78915	0.70437455	1.46E-26	3.89E-24	6
ELAVL3	12.532735	4.8886375	2.80E-25	6.06E-23	6
UCHL1	12.483789	2.871511	2.74E-25	5.99E-23	6
NOVA1	12.454932	2.8784025	3.55E-25	7.56E-23	6
NSG1	12.409797	4.7219453	6.13E-25	1.29E-22	6
CFL1	12.404735	0.5453498	1.28E-25	3.00E-23	6
EIF1B	12.388722	0.86495614	1.98E-25	4.41E-23	6
EID1	12.181234	0.9190144	6.60E-25	1.38E-22	6
PAK3	12.123525	4.932227	3.65E-24	7.20E-22	6
ZNF428	12.087504	1.2407407	1.69E-24	3.41E-22	6
TCEAL7	12.01713	3.3882537	6.20E-24	1.19E-21	6
GPM6A	11.818674	3.7161589	2.24E-23	4.10E-21	6
SEC11C	11.759984	2.1635509	2.52E-23	4.50E-21	6
PFDN2	11.630645	1.338213	4.07E-23	7.05E-21	6
TMEM66	11.617498	0.7153681	2.47E-23	4.44E-21	6
KCNQ1OT1	11.566927	1.1142226	5.04E-23	8.45E-21	6
ATP6V0B	11.463036	1.2335727	1.23E-22	2.00E-20	6
SCG5	11.461961	3.9785502	2.11E-22	3.31E-20	6
TCEAL3	11.251547	1.7579148	4.66E-22	7.06E-20	6
GABARAPL2	11.180929	0.6837051	3.90E-22	6.02E-20	6
PTMS	11.154782	1.0554028	6.05E-22	9.02E-20	6
TUBB2A	11.128235	1.9128187	1.26E-21	1.82E-19	6
PEBP1	11.070372	0.77629346	9.99E-22	1.45E-19	6

SEZ6L2	11.044107	2.806193	2.56E-21	3.57E-19	6
ATP6V0E2	11.029923	2.9069233	2.78E-21	3.86E-19	6
MAPT	11.0204	5.29036	3.58E-21	4.91E-19	6
TMSB15A	10.923989	2.871279	5.42E-21	7.18E-19	6
SYT1	10.917815	3.180445	5.95E-21	7.85E-19	6
RTN4	10.829506	0.7023911	3.94E-21	5.34E-19	6
RPAIN	10.794539	1.2110785	7.37E-21	9.54E-19	6
SNAP25	10.793788	5.0782733	1.44E-20	1.82E-18	6
HOXB5	10.78466	3.462786	1.41E-20	1.79E-18	6
CPE	10.774669	2.8620806	1.37E-20	1.75E-18	6
TAGLN3	10.738423	3.460078	1.84E-20	2.30E-18	6
MAP1LC3A	10.712715	1.9924332	1.66E-20	2.09E-18	6
MAP2	10.700789	3.7647662	2.38E-20	2.96E-18	6
TMEM59L	10.682014	3.3533041	2.53E-20	3.13E-18	6
SYP	10.661796	5.0156956	3.22E-20	3.95E-18	6
MAGEH1	10.644382	2.0529187	2.73E-20	3.38E-18	6
LINC00632	10.568884	3.548592	5.28E-20	6.30E-18	6
CRMP1	10.471423	4.484228	1.04E-19	1.20E-17	6
LUC7L3	10.410752	0.88100535	7.10E-20	8.33E-18	6
EEF1A2	10.400601	3.6388154	1.52E-19	1.70E-17	6
USP11	10.332975	1.5116383	1.62E-19	1.80E-17	6
NREP	10.315433	1.4833326	1.87E-19	2.06E-17	6
PODXL2	10.287995	3.5914514	3.00E-19	3.22E-17	6
ERH	10.287763	0.58123326	1.23E-19	1.41E-17	6
KMT2E	10.281467	1.0298641	1.79E-19	1.97E-17	6
CDKN2D	10.269892	2.9181278	3.27E-19	3.48E-17	6
FXYD6	10.198339	1.9080036	3.84E-19	4.04E-17	6
PGAM1	10.170321	1.0015751	3.61E-19	3.82E-17	6
PAIP2	10.16846	1.1055162	3.92E-19	4.11E-17	6
MORF4L1	10.150419	0.49639708	3.14E-19	3.35E-17	6
SYT11	10.084746	3.1928458	9.90E-19	9.76E-17	6
GAP43	10.047291	4.7941422	1.41E-18	1.37E-16	6
H2AFZ	9.987463	0.4102181	6.38E-19	6.44E-17	6
MED10	9.972824	1.1702031	1.43E-18	1.38E-16	6
RNF187	9.970903	1.1970041	1.27E-18	1.24E-16	6
PEG10	9.968439	1.6067775	1.63E-18	1.55E-16	6
DAAM1	9.805506	1.2977364	4.27E-18	3.77E-16	6
PKIA	9.803439	3.8106027	6.18E-18	5.33E-16	6
ARPP19	9.798968	1.0247746	3.79E-18	3.40E-16	6
CELF4	9.781482	4.1834583	7.17E-18	6.06E-16	6
CALM2	9.772552	0.48911774	2.73E-18	2.48E-16	6
SOX4	9.745742	0.96338123	6.21E-18	5.34E-16	6
DPYSL2	9.700539	2.1390874	9.43E-18	7.85E-16	6
KIF5C	9.695714	3.7747042	1.16E-17	9.56E-16	6
PSIP1	9.660627	1.6419567	1.07E-17	8.82E-16	6
COL3A1	45.16627	5.002575	3.65E-65	5.97E-61	7
COL1A1	39.95015	4.9998193	2.90E-60	1.90E-56	7
LGALS1	39.110542	4.6474543	7.84E-60	4.28E-56	7
IFITM3	37.175625	2.0332513	4.78E-61	3.91E-57	7
COL1A2	34.607	5.167108	8.31E-55	3.02E-51	7
SPARC	29.417236	2.96234	3.94E-50	1.07E-46	7
VIM	24.939322	1.1318125	3.74E-49	8.57E-46	7
PHLDA1	23.234652	3.5009604	7.42E-41	9.00E-38	7
IFITM2	22.565308	2.5853689	5.04E-40	5.69E-37	7
NR2F1	20.771614	3.5115657	6.09E-37	5.38E-34	7
COL6A3	20.329865	6.881423	8.00E-36	6.09E-33	7

SELM	19.924252	3.0072682	9.20E-36	6.69E-33	7
CALD1	19.232376	2.7920115	1.76E-34	1.13E-31	7
COL6A2	18.723478	4.5839815	2.90E-33	1.53E-30	7
PTN	18.680696	3.5085325	2.36E-33	1.29E-30	7
FN1	18.033207	5.1741424	5.27E-32	2.50E-29	7
RBP1	17.922398	1.7544878	2.49E-33	1.34E-30	7
FTL	17.762115	0.39952892	3.50E-33	1.82E-30	7
PLAC9	17.460535	6.792794	5.63E-31	2.46E-28	7
COL4A2	16.64554	5.021274	1.45E-29	5.92E-27	7
TGFBI	16.48943	5.2914205	2.92E-29	1.15E-26	7
SERPINH1	16.220701	2.4821644	4.45E-29	1.71E-26	7
MEST	16.123835	2.2498496	4.29E-29	1.67E-26	7
COL5A2	15.919869	4.3130646	2.68E-28	9.34E-26	7
TPM4	15.804683	1.5579185	1.13E-28	4.12E-26	7
COL4A1	14.915972	6.0072193	2.59E-26	7.12E-24	7
COL5A1	14.740219	3.9902742	4.71E-26	1.23E-23	7
DCN	14.636744	4.544484	7.67E-26	1.98E-23	7
CTSC	14.513163	3.7943823	1.30E-25	3.31E-23	7
CLDN11	14.442942	5.215036	2.02E-25	5.08E-23	7
LUM	13.832424	4.876903	3.11E-24	7.12E-22	7
A1BG	13.824708	2.2575665	1.37E-24	3.24E-22	7
OLFML3	13.726144	6.1973233	5.33E-24	1.19E-21	7
FBLN1	13.724037	3.9590592	4.82E-24	1.08E-21	7
EMP3	13.601906	1.7713875	2.49E-24	5.82E-22	7
TPM2	13.469235	4.027921	1.44E-23	3.08E-21	7
COL6A1	13.454184	2.3828502	1.16E-23	2.52E-21	7
OST4	13.27297	0.37951037	4.79E-24	1.08E-21	7
AKAP12	13.228676	3.5503662	4.13E-23	8.66E-21	7
PCOLCE	13.211601	4.5155425	5.27E-23	1.09E-20	7
MDK	13.184353	1.0401665	1.37E-23	2.94E-21	7
TMSB4X	13.123682	0.18136701	3.04E-24	7.01E-22	7
LHFP	13.091931	3.5411243	8.07E-23	1.61E-20	7
RHOBTB3	13.030703	2.1216624	6.75E-23	1.37E-20	7
EGR1	12.840446	0.5763249	9.94E-24	2.18E-21	7
NME4	12.569433	1.2979804	3.50E-22	6.70E-20	7
NAP1L1	12.408457	0.575266	3.14E-22	6.04E-20	7
CCL2	12.383357	5.420883	2.54E-21	4.48E-19	7
CEBPD	12.270819	1.4900626	2.12E-21	3.77E-19	7
PTGDS	12.25489	3.648865	4.29E-21	7.32E-19	7
FKBP10	12.146194	3.1856408	5.96E-21	1.00E-18	7
MDFI	12.027198	2.884507	1.02E-20	1.66E-18	7
ARHGAP29	12.012037	3.1863291	1.30E-20	2.08E-18	7
FSTL1	11.915882	2.200904	1.45E-20	2.32E-18	7
OGN	11.895268	7.4884815	2.60E-20	4.06E-18	7
NGFRAP1	11.888756	0.7133594	3.18E-21	5.50E-19	7
CALU	11.881723	1.2745538	9.96E-21	1.64E-18	7
CYP1B1	11.66213	3.7410154	7.13E-20	1.05E-17	7
NR2F2	11.53033	2.5953255	1.13E-19	1.64E-17	7
WBP5	11.528366	0.87604177	3.50E-20	5.33E-18	7
BGN	11.496466	6.3756547	1.70E-19	2.43E-17	7
NID1	11.404688	5.869412	2.60E-19	3.62E-17	7
A2M	11.317408	5.914242	3.98E-19	5.32E-17	7
GYPC	11.211776	3.4925852	5.60E-19	7.33E-17	7
LRRC17	11.187679	6.161305	7.36E-19	9.60E-17	7
C12orf57	11.172472	0.87482625	2.16E-19	3.08E-17	7
KCNQ1OT1	11.154731	1.1016583	3.50E-19	4.70E-17	7

PDGFRA	11.099793	8.039744	1.14E-18	1.45E-16	7
APOE	11.099486	4.2690682	1.09E-18	1.40E-16	7
LAMA4	11.077139	5.429612	1.23E-18	1.56E-16	7
COL4A5	10.969347	2.328899	1.50E-18	1.89E-16	7
FBN2	10.837775	4.637137	3.82E-18	4.63E-16	7
CTHRC1	10.806541	5.5757575	4.55E-18	5.48E-16	7
MYL9	10.769095	3.366778	4.82E-18	5.77E-16	7
SERPING1	10.746316	4.568564	5.86E-18	6.97E-16	7
MMP2	10.647179	6.1202765	9.89E-18	1.15E-15	7
NUPR1	10.465594	4.519283	2.28E-17	2.55E-15	7
IGFBP4	10.371827	2.6273367	3.01E-17	3.32E-15	7
IFITM1	10.269238	3.492791	5.79E-17	6.24E-15	7
ID3	10.24328	1.2632217	2.83E-17	3.13E-15	7
ISLR	10.222007	7.2416453	7.77E-17	8.17E-15	7
CNN3	10.205537	1.3482093	4.20E-17	4.55E-15	7
C11orf96	10.116972	5.986264	1.28E-16	1.30E-14	7
NID2	10.100693	6.8892293	1.39E-16	1.40E-14	7
PTMS	10.049725	1.0156959	8.12E-17	8.51E-15	7
CPE	10.039521	2.7763925	1.55E-16	1.55E-14	7
RARRES2	10.036538	3.311602	1.73E-16	1.72E-14	7
CDH11	10.020047	4.757746	1.97E-16	1.91E-14	7
EFEMP2	9.947658	2.3549638	2.28E-16	2.21E-14	7
EVA1B	9.888998	3.7840486	3.62E-16	3.39E-14	7
PEG10	9.826884	1.6741753	3.61E-16	3.39E-14	7
C6orf48	9.825465	0.49664629	1.88E-16	1.83E-14	7
RBMS3	9.823852	3.1551108	4.70E-16	4.36E-14	7
GMDS	9.817458	2.940336	5.05E-16	4.65E-14	7
CPXM1	9.778046	4.4375296	6.42E-16	5.84E-14	7
CKAP4	9.62336	1.6043593	9.30E-16	8.34E-14	7
CXCL12	9.585542	6.0020695	1.68E-15	1.46E-13	7
PDGFRB	9.581837	4.442584	1.66E-15	1.45E-13	7
FLRT2	9.552939	5.4061217	1.95E-15	1.69E-13	7
SPARCL1	9.532569	2.3159816	1.51E-15	1.32E-13	7
HMGB2	24.181173	2.6638143	5.49E-29	5.99E-25	8
H2AFZ	20.849926	0.9718027	9.62E-27	6.30E-23	8
TUBA1B	20.839823	1.3617413	9.32E-27	6.30E-23	8
HMGB1	20.731758	1.0850037	1.17E-26	6.36E-23	8
TUBB	18.847418	1.2308029	7.92E-25	2.88E-21	8
STMN1	18.474798	1.8961885	4.10E-24	1.22E-20	8
HMGN2	18.465553	1.1696095	3.37E-24	1.10E-20	8
H2AFV	16.634216	1.5114632	3.88E-22	9.76E-19	8
SMC4	15.388149	3.3584073	2.58E-20	4.69E-17	8
CKS1B	14.681474	1.9783486	1.32E-19	2.16E-16	8
CENPW	14.400428	3.4854846	3.79E-19	5.91E-16	8
BIRC5	14.361466	5.4884753	5.04E-19	7.50E-16	8
HNRNPA1	13.48716	0.4583875	8.64E-19	1.23E-15	8
CKS2	12.852003	2.348187	2.85E-17	3.89E-14	8
PTTG1	12.712687	2.3063445	4.39E-17	5.75E-14	8
NUSAP1	12.653502	5.7851033	6.47E-17	7.84E-14	8
NUCKS1	12.439096	0.8989926	5.24E-17	6.59E-14	8
DEK	12.269151	1.3095132	1.00E-16	1.17E-13	8
CENPF	11.725805	5.524009	1.05E-15	1.04E-12	8
TYMS	11.218962	3.2827733	4.47E-15	4.06E-12	8
AURKB	11.132109	6.6765337	6.66E-15	5.74E-12	8
RAN	11.077017	0.8888839	3.88E-15	3.63E-12	8
MKI67	10.913289	6.442764	1.32E-14	1.06E-11	8

H2AFX	10.893776	2.642463	1.20E-14	9.78E-12	8
TOP2A	10.870499	5.603913	1.51E-14	1.17E-11	8
SKA2	10.711636	2.0158792	1.78E-14	1.35E-11	8
MAD2L1	10.691129	3.0702393	2.39E-14	1.74E-11	8
TPX2	10.682462	4.871878	2.71E-14	1.93E-11	8
UBE2C	10.564225	5.417495	4.01E-14	2.73E-11	8
KIAA0101	10.540088	3.7982955	4.13E-14	2.76E-11	8
HMGN1	10.399348	0.46807414	3.33E-14	2.32E-11	8
YBX1	10.363112	0.3437525	2.37E-14	1.74E-11	8
HNRNPA2B1	10.30184	0.814712	5.04E-14	3.30E-11	8
HMGB3	9.838732	2.3481557	3.62E-13	2.19E-10	8
ZWINT	9.780556	4.189735	4.94E-13	2.74E-10	8
PTMA	9.640275	0.25366634	4.27E-13	2.45E-10	8
SNRPF	9.52661	0.851324	6.59E-13	3.59E-10	8
LSM4	9.4577265	1.0286552	9.75E-13	5.23E-10	8
DTYMK	9.429237	2.2005641	1.39E-12	7.11E-10	8
CCNA2	9.320925	5.9177256	2.38E-12	1.14E-09	8
SGOL1	9.31767	5.7078547	2.39E-12	1.14E-09	8
DBI	9.262149	0.59099275	1.18E-12	6.21E-10	8
CDK1	9.0605345	4.316832	5.61E-12	2.48E-09	8
PRC1	8.893026	4.761811	1.00E-11	4.20E-09	8
ANAPC11	8.889083	0.5349907	5.16E-12	2.31E-09	8
SNRPD1	8.863328	0.77066386	6.50E-12	2.80E-09	8
CDKN3	8.81329	5.2846904	1.32E-11	5.29E-09	8
HIST1H4C	8.794792	1.7329351	1.33E-11	5.29E-09	8
MZT2B	8.716303	0.57173413	1.10E-11	4.55E-09	8
LSM5	8.692708	0.9270951	1.42E-11	5.60E-09	8
ANP32E	8.690229	1.9752792	1.79E-11	6.88E-09	8
PBK	8.688662	6.5677	2.04E-11	7.75E-09	8
CENPK	8.642366	3.9121516	2.29E-11	8.51E-09	8
SRSF3	8.541302	0.69514096	2.24E-11	8.41E-09	8
MND1	8.49681	5.141748	3.88E-11	1.34E-08	8
TMEM106C	8.321904	1.5547224	5.83E-11	1.93E-08	8
LMNB1	8.158569	3.6264966	1.22E-10	3.87E-08	8
MIS18BP1	8.140392	2.5131428	1.23E-10	3.87E-08	8
HP1BP3	8.127857	1.3886675	1.17E-10	3.74E-08	8
CENPN	8.046926	3.681064	1.79E-10	5.35E-08	8
CENPA	8.045638	7.1195593	1.87E-10	5.48E-08	8
DLGAP5	8.014429	5.966372	2.08E-10	5.97E-08	8
ASPM	7.9807677	6.22277	2.34E-10	6.67E-08	8
COX8A	7.930208	0.3059146	1.58E-10	4.89E-08	8
CBX3	7.9276137	0.83189154	1.97E-10	5.70E-08	8
TROAP	7.8488584	5.4112854	3.68E-10	1.00E-07	8
CCNB1	7.841271	4.2651863	3.77E-10	1.02E-07	8
KPNA2	7.829449	3.353684	3.88E-10	1.03E-07	8
SGOL2	7.775512	4.6566277	4.75E-10	1.22E-07	8
SMC2	7.7653556	3.0271566	4.74E-10	1.22E-07	8
RBBP7	7.7573047	1.1575111	4.07E-10	1.07E-07	8
RDX	7.6573353	1.116325	5.64E-10	1.42E-07	8
TK1	7.655099	4.307225	7.19E-10	1.77E-07	8
NCL	7.620251	0.71337444	6.10E-10	1.51E-07	8
KIFC1	7.5881257	4.3700805	9.11E-10	2.18E-07	8
NUF2	7.572717	5.72098	9.73E-10	2.26E-07	8
MZT1	7.55176	1.7910818	9.48E-10	2.22E-07	8
ARHGAP11A	7.5265665	5.080023	1.14E-09	2.61E-07	8
PPIA	7.463433	0.346073	1.09E-09	2.50E-07	8

CDCA3	7.452936	5.144599	1.48E-09	3.36E-07	8
NUDT1	7.3980713	2.357799	1.68E-09	3.69E-07	8
RAD21	7.33447	1.4295236	1.98E-09	4.21E-07	8
TMPO	7.318236	1.8283767	2.13E-09	4.50E-07	8
SNRPG	7.27736	0.48320967	1.74E-09	3.79E-07	8
KIF20B	7.274791	3.2613392	2.69E-09	5.50E-07	8
UBE2S	7.2612534	3.0323236	2.84E-09	5.71E-07	8
CCDC34	7.254658	1.4881519	2.58E-09	5.34E-07	8
RPL39L	7.253254	2.6259599	2.80E-09	5.68E-07	8
RNASEH2A	7.1762695	2.6457996	3.69E-09	7.27E-07	8
KIF22	7.172087	2.5562031	3.77E-09	7.38E-07	8
MZT2A	7.171662	0.5190294	2.37E-09	4.98E-07	8
PTMS	7.1653194	1.1791116	3.61E-09	7.16E-07	8
CCNB2	7.132723	5.3898177	4.58E-09	8.66E-07	8
HJURP	7.1260805	6.5169106	4.69E-09	8.83E-07	8
RAD51AP1	7.081787	4.5974603	5.43E-09	1.02E-06	8
RRM2	7.0702252	5.0151	5.68E-09	1.06E-06	8
KNSTRN	7.0063143	4.0892544	7.08E-09	1.28E-06	8
CDCA5	6.937406	4.7013106	9.03E-09	1.60E-06	8
NAP1L1	6.925985	0.48149675	7.31E-09	1.31E-06	8
TUBB4B	6.9205127	0.9287383	7.76E-09	1.39E-06	8

Table 5-2 Cluster marker genes for scRNA-seq of hPSC-derived organoids after 3 days of DSA and 56 days DSI treatment. Lung clusters extracted and re-clustered.

names	scores	logfoldchanges	pvals	pvals_adj	Cluster #
SOX2	18.19199	1.454794	3.88E-51	1.27E-46	0
FABP7	16.57626	2.400911	3.46E-39	2.83E-35	0
TTYH1	16.12451	1.970226	2.09E-38	1.37E-34	0
SPARCL1	15.93433	1.885728	2.59E-38	1.41E-34	0
CD99	15.80638	0.885538	2.59E-42	4.23E-38	0
RGMA	15.78539	1.872493	4.49E-38	2.10E-34	0
EDNRB	15.75176	3.02079	2.32E-35	9.51E-32	0
ID4	15.64543	1.248437	7.90E-40	8.62E-36	0
C1orf61	14.31482	2.111412	3.98E-32	1.45E-28	0
NTRK2	13.14753	1.363064	6.53E-30	1.94E-26	0
PSAP	13.00353	0.784451	5.02E-31	1.64E-27	0
PEA15	12.92796	1.571211	1.71E-28	4.17E-25	0
PDLIM3	12.79017	1.852202	3.57E-28	7.79E-25	0
QKI	12.77309	1.326368	8.88E-29	2.42E-25	0
GPM6B	12.57298	1.178543	1.78E-28	4.17E-25	0
CNN3	12.0952	1.174841	3.70E-26	7.56E-23	0
ID2	11.5936	1.250606	4.07E-25	7.39E-22	0
PFN2	11.5726	0.757132	2.30E-25	4.43E-22	0
PTPRZ1	11.1487	2.667531	4.61E-22	6.86E-19	0
MGST1	10.88998	1.641306	8.39E-22	1.14E-18	0
SCD	10.79095	1.261921	4.83E-22	6.87E-19	0
SRI	10.76242	0.992341	1.54E-22	2.40E-19	0
ATP1B2	10.47471	1.330372	5.36E-21	6.75E-18	0
HOPX	10.10489	2.503717	5.01E-19	4.56E-16	0
WNT7A	9.936479	2.665078	1.45E-18	1.25E-15	0
ID3	9.680615	0.813804	1.92E-	1.79E-16	0

			19		
			5.58E-		
VIM	9.631726	0.48522	20	5.71E-17	0
			2.10E-		
GFAP	9.595495	3.211197	17	1.60E-14	0
			8.32E-		
LINC00461	9.553686	2.212019	18	6.64E-15	0
			3.24E-		
PSAT1	9.414024	2.306761	17	2.36E-14	0
			1.12E-		
AP1S2	9.347283	1.245314	17	8.77E-15	0
			3.35E-		
ANKRD45	9.205925	1.430618	17	2.38E-14	0
			1.17E-		
SOX9	9.141642	1.846933	16	7.85E-14	0
			1.17E-		
LYPD1	9.057364	1.822829	16	7.85E-14	0
			5.48E-		
MT3	8.95405	2.341236	16	3.26E-13	0
			8.00E-		
GATM	8.921925	2.327658	16	4.59E-13	0
			4.76E-		
CDH2	8.822533	1.494517	16	2.88E-13	0
			6.15E-		
DKK3	8.742237	1.177243	16	3.59E-13	0
			3.20E-		
FAM181B	8.69302	2.729228	15	1.69E-12	0
			4.81E-		
BCAN	8.63298	2.756757	15	2.35E-12	0
			2.57E-		
METRN	8.504575	0.989306	15	1.38E-12	0
			6.51E-		
FGFR3	8.497943	1.787505	15	2.97E-12	0
			9.81E-		
LRIG1	8.487337	2.220942	15	4.40E-12	0
			4.18E-		
CRB2	8.483152	1.534167	15	2.14E-12	0
			4.04E-		
PLA2G16	8.45046	1.175053	15	2.10E-12	0
			1.53E-		
JAM2	8.367828	1.985865	14	6.49E-12	0
			4.07E-		
HEY2	8.295833	2.996057	14	1.57E-11	0
			3.11E-		
SDC3	8.278572	2.17366	14	1.24E-11	0
			3.71E-		
HS6ST1	8.244093	2.091505	14	1.44E-11	0
			1.51E-		
CLU	8.155451	0.81956	14	6.49E-12	0
			2.97E-		
TMEM47	8.083627	1.006852	14	1.22E-11	0
			3.11E-		
NFIA	8.078782	1.020316	14	1.24E-11	0
			9.68E-		
PHYHIPL	8.046005	2.207465	14	3.60E-11	0
NPC2	8.002723	0.548564	2.81E-	1.16E-11	0

			14		
			6.16E-		
MLF1	7.948972	0.880748	14	2.32E-11	0
			2.01E-		
NKAIN4	7.920522	1.883525	13	6.87E-11	0
			1.16E-		
LINC00998	7.882422	0.833139	13	4.16E-11	0
			6.01E-		
MAN1C1	7.846684	3.264101	13	1.77E-10	0
			4.80E-		
KCNN3	7.825645	2.273323	13	1.48E-10	0
			4.83E-		
NSMF	7.758761	1.582286	13	1.48E-10	0
			4.27E-		
FAM84B	7.757122	1.523009	13	1.34E-10	0
			1.03E-		
IGFBP5	7.753534	0.68271	13	3.79E-11	0
			8.09E-		
PAX6	7.743047	2.538811	13	2.34E-10	0
			5.41E-		
FOXK1	7.713315	1.320139	13	1.63E-10	0
			1.36E-		
SLC4A4	7.674096	2.447393	12	3.78E-10	0
			7.42E-		
TIMP3	7.665924	1.361234	13	2.17E-10	0
			3.57E-		
RPS27L	7.637168	0.467668	13	1.15E-10	0
			1.83E-		
FAM181A	7.585322	2.224265	12	4.90E-10	0
			3.28E-		
TUBB2B	7.558613	0.504152	13	1.07E-10	0
			3.99E-		
CKB	7.558233	0.430382	13	1.27E-10	0
			3.92E-		
CXCR4	7.49177	2.604055	12	9.94E-10	0
			1.55E-		
FEZ1	7.490528	0.985994	12	4.18E-10	0
			4.58E-		
DOK5	7.447167	2.378744	12	1.15E-09	0
			3.65E-		
AK4	7.426387	1.562119	12	9.41E-10	0
			2.08E-		
HSPE1	7.383887	0.449096	12	5.49E-10	0
			3.79E-		
MSI1	7.356313	1.145763	12	9.69E-10	0
			1.04E-		
DDR1	7.17749	0.971322	11	2.49E-09	0
			1.87E-		
SULF2	7.161525	1.663876	11	4.26E-09	0
			2.17E-		
ADCYAP1R1	7.144482	2.095747	11	4.86E-09	0
			4.14E-		
FAM107A	7.132736	5.18442	11	8.81E-09	0
			1.29E-		
HLA-A	7.132283	0.783742	11	3.05E-09	0
RTN4	7.076475	0.4974	1.27E-	3.03E-09	0

			11		
			2.20E-		
RCN1	7.04209	0.743681	11	4.90E-09	0
			3.37E-		
TMEM170A	7.01206	1.291603	11	7.36E-09	0
			2.33E-		
CHCHD10	7.004945	0.90534	11	5.15E-09	0
			6.85E-		
RAB31	6.973491	2.543093	11	1.38E-08	0
			6.23E-		
GPT2	6.959883	1.978824	11	1.29E-08	0
			4.11E-		
GLUD1	6.959635	1.061291	11	8.79E-09	0
			9.91E-		
PSRC1	6.905767	2.330367	11	1.90E-08	0
			6.68E-		
ASAH1	6.899046	0.984871	11	1.36E-08	0
			4.72E-		
RFX4	6.867361	0.987297	11	9.96E-09	0
			1.33E-		
KAL1	6.861109	2.795447	10	2.42E-08	0
			8.47E-		
MSI2	6.825286	0.977408	11	1.68E-08	0
			9.83E-		
FAM213A	6.756251	0.682493	11	1.90E-08	0
			9.63E-		
ZFP36L1	6.741999	0.575541	11	1.88E-08	0
			1.46E-		
HES4	6.730419	0.890096	10	2.63E-08	0
			3.16E-		
PRUNE2	6.71464	3.35414	10	5.23E-08	0
			1.87E-		
TIMP2	6.676795	0.981415	10	3.31E-08	0
			1.61E-		
GCSH	6.670877	0.726454	10	2.88E-08	0
			1.64E-		
WSB1	6.652316	0.363437	10	2.91E-08	0
			2.64E-		
RPL41	13.28923	0.14357	32	7.30E-28	1
			4.46E-		
RPL31	13.22894	0.203809	32	7.30E-28	1
			1.33E-		
ZFAS1	13.17001	0.424374	30	9.90E-27	1
			8.29E-		
S100A6	13.01977	0.683506	29	4.52E-25	1
			4.20E-		
RPL23A	12.47976	0.208161	31	4.58E-27	1
			1.51E-		
RPS27	12.46022	0.146325	30	9.90E-27	1
			8.78E-		
RPL11	12.00771	0.183269	28	4.10E-24	1
			1.45E-		
RPS6	11.99166	0.244374	27	5.94E-24	1
			1.83E-		
RPS15A	11.85743	0.176442	27	6.66E-24	1
RPS4X	11.84488	0.241677	4.64E-	1.52E-23	1

			27		
			1.83E-		
RPS24	11.71035	0.177709	26	5.00E-23	1
			1.09E-		
RPL39	11.6043	0.18133	26	3.24E-23	1
			1.56E-		
RPS3	11.49655	0.194035	25	3.93E-22	1
			2.87E-		
RPS27A	11.2429	0.15053	25	6.72E-22	1
			9.75E-		
RPS14	11.21435	0.150951	25	2.13E-21	1
			5.49E-		
RPL10	11.17065	0.171989	24	9.98E-21	1
			2.38E-		
RPL13A	10.96219	0.119713	24	4.88E-21	1
			3.98E-		
RPS12	10.85938	0.170536	24	7.67E-21	1
			1.36E-		
RPS3A	10.81223	0.14436	23	2.34E-20	1
			1.97E-		
RPL26	10.79116	0.161512	23	3.22E-20	1
			4.80E-		
RPS28	10.72293	0.198174	23	6.84E-20	1
			3.76E-		
RPL34	10.68464	0.15863	23	5.59E-20	1
			2.66E-		
RPL27A	10.66991	0.140328	23	4.14E-20	1
			1.45E-		
RPL21	10.54903	0.129699	22	1.90E-19	1
			1.87E-		
RPS25	10.53093	0.161278	22	2.36E-19	1
			9.29E-		
RPS29	10.50569	0.158878	23	1.27E-19	1
			2.66E-		
RPL12	10.49825	0.265854	22	3.22E-19	1
			1.29E-		
RPL14	10.37101	0.231431	21	1.45E-18	1
			1.40E-		
S100A11	10.34043	0.944024	20	1.35E-17	1
			8.34E-		
RPL13	10.31196	0.137873	22	9.74E-19	1
			4.24E-		
RPL32	10.17295	0.149826	21	4.48E-18	1
			3.43E-		
RPL7	10.14359	0.177816	21	3.74E-18	1
			4.85E-		
RPL9	10.14293	0.172101	21	4.96E-18	1
			6.47E-		
RPL18A	10.12159	0.151048	21	6.42E-18	1
			5.66E-		
RPL36	9.885892	0.155246	20	5.29E-17	1
			6.99E-		
RPS18	9.860388	0.145957	20	6.35E-17	1
			7.52E-		
RPL30	9.752546	0.15631	20	6.65E-17	1
RPL35A	9.743472	0.152815	1.38E-	1.19E-16	1

			19		
			2.82E-		
GNB2L1	9.540092	0.2406	19	2.37E-16	1
			1.61E-		
RPL8	9.453907	0.156146	18	1.29E-15	1
			1.06E-		
RPL28	9.316669	0.157871	18	8.65E-16	1
			4.93E-		
EIF3E	9.313086	0.471603	18	3.76E-15	1
			4.57E-		
RPS23	9.231217	0.148896	18	3.56E-15	1
			3.64E-		
HSPB1	9.031297	0.596767	17	2.43E-14	1
			2.94E-		
RPL5	8.999932	0.153557	17	2.01E-14	1
			3.07E-		
PERP	8.993682	1.294566	16	1.86E-13	1
			1.79E-		
RPS8	8.982495	0.139457	17	1.30E-14	1
			2.23E-		
RPSA	8.958356	0.214864	17	1.59E-14	1
			2.66E-		
RPL19	8.88989	0.135202	17	1.86E-14	1
			8.97E-		
RPL29	8.868405	0.153785	17	5.87E-14	1
			1.11E-		
RPL3	8.775098	0.139995	16	7.01E-14	1
			9.51E-		
RPS9	8.761948	0.193436	17	6.10E-14	1
			1.56E-		
KRT19	8.756835	0.976168	15	8.66E-13	1
			2.38E-		
C19orf33	8.736829	1.72943	15	1.28E-12	1
			1.84E-		
PLP2	8.676952	1.356515	15	1.00E-12	1
			2.28E-		
RPS5	8.676296	0.164534	16	1.41E-13	1
			4.02E-		
RPLP1	8.612431	0.139563	16	2.39E-13	1
			7.51E-		
RPL10A	8.577909	0.215034	16	4.31E-13	1
			7.01E-		
RPL18	8.530478	0.149391	16	4.10E-13	1
			1.53E-		
RAB25	8.504952	2.587646	14	7.58E-12	1
			9.95E-		
RPL36A	8.433289	0.249833	16	5.61E-13	1
			2.87E-		
RPLP0	8.26928	0.153845	15	1.49E-12	1
			4.26E-		
KRT8	8.167848	1.121027	14	1.99E-11	1
			5.08E-		
ELF3	8.161208	1.444894	14	2.34E-11	1
			1.24E-		
RPS15	8.150404	0.127189	14	6.25E-12	1
FXYD3	8.138947	0.998983	7.34E-	3.29E-11	1

			14		
			1.92E-		
TPT1	8.048067	0.169647	14	9.38E-12	1
			5.93E-		
SPINT2	8.012937	0.756767	14	2.69E-11	1
			3.75E-		
RPS7	7.933375	0.162495	14	1.78E-11	1
			3.65E-		
RPS2	7.646295	0.124904	13	1.59E-10	1
			2.23E-		
KLF5	7.555424	1.57068	12	8.51E-10	1
			6.61E-		
RPL37	7.53517	0.157079	13	2.85E-10	1
			6.28E-		
S100A14	7.428708	1.847439	12	2.21E-09	1
			1.56E-		
RPL24	7.415851	0.122802	12	6.22E-10	1
			2.90E-		
IGF2	7.356839	0.536243	12	1.08E-09	1
			2.49E-		
RPL15	7.33514	0.108625	12	9.35E-10	1
			5.38E-		
LY6E	7.331647	0.940149	12	1.93E-09	1
			2.14E-		
RPS20	7.317449	0.186192	12	8.22E-10	1
			4.22E-		
RPS13	7.234441	0.118401	12	1.53E-09	1
			6.98E-		
RPS21	7.08499	0.172215	12	2.43E-09	1
			4.71E-		
MIR205HG	7.045761	1.674282	11	1.47E-08	1
			6.13E-		
DSP	6.984591	1.600692	11	1.86E-08	1
			3.03E-		
RPL22	6.968406	0.16519	11	1.00E-08	1
			6.33E-		
ANXA1	6.953962	1.142191	11	1.90E-08	1
			8.42E-		
PDLIM1	6.933531	1.434797	11	2.46E-08	1
			4.46E-		
POLR1D	6.895615	0.630698	11	1.40E-08	1
			4.37E-		
RPL35	6.883266	0.125237	11	1.40E-08	1
			2.99E-		
RPLP2	6.857367	0.105584	11	9.97E-09	1
			3.20E-		
RPS16	6.850988	0.113061	11	1.05E-08	1
			4.46E-		
UBA52	6.816187	0.179546	11	1.40E-08	1
			1.03E-		
BTF3	6.736257	0.264193	10	2.96E-08	1
			1.39E-		
NBEAL1	6.665387	0.430938	10	3.90E-08	1
			1.37E-		
RPL17	6.64316	0.203276	10	3.85E-08	1
TACSTD2	6.614042	1.305859	4.11E-	1.07E-07	1

			10		
			7.29E-		
AGR2	6.512719	1.105815	10	1.82E-07	1
			2.62E-		
EEF1A1	6.505839	0.082063	10	7.03E-08	1
			3.32E-		
RPL7A	6.500345	0.121352	10	8.84E-08	1
			3.58E-		
RPL37A	6.49878	0.093145	10	9.46E-08	1
			4.33E-		
ATP5G2	6.495469	0.231105	10	1.11E-07	1
			9.91E-		
BCAM	6.47751	1.597513	10	2.37E-07	1
			3.04E-		
BEX1	17.59288	2.083473	39	9.06E-36	2
			1.31E-		
STMN2	16.78858	3.279383	33	2.85E-30	2
			4.65E-		
DCX	15.29256	3.52593	30	6.08E-27	2
			1.15E-		
VAMP2	13.77551	1.032817	28	1.35E-25	2
			3.60E-		
STMN4	13.19984	3.152029	25	2.68E-22	2
			2.38E-		
HN1	13.02152	1.193461	26	2.05E-23	2
			6.70E-		
GNG3	12.774	3.7485	24	4.30E-21	2
			4.63E-		
MLLT11	12.24516	1.196674	24	3.03E-21	2
			1.22E-		
MAP1B	12.09221	0.881857	24	8.49E-22	2
			1.30E-		
BEX2	12.01299	1.523318	23	8.17E-21	2
			2.43E-		
NSG1	11.64564	3.16597	21	1.13E-18	2
			1.88E-		
RAB3A	11.55444	2.622445	21	9.07E-19	2
			7.53E-		
ATP6V1G2	11.37985	2.650045	21	3.20E-18	2
			1.46E-		
STMN1	11.20652	0.911178	21	7.23E-19	2
			1.82E-		
BEX5	11.19615	2.615336	20	6.99E-18	2
			1.65E-		
ELAVL3	11.16213	2.687616	20	6.61E-18	2
			1.17E-		
BASP1	10.89773	1.260113	20	4.78E-18	2
			7.18E-		
TERF2IP	10.57836	1.205889	20	2.56E-17	2
			3.79E-		
SEC11C	10.48551	1.882868	19	1.18E-16	2
			2.81E-		
RTN1	10.48023	1.049002	19	9.02E-17	2
			2.20E-		
TUBB	10.24956	0.683128	19	7.20E-17	2
CCNI	10.15095	0.427676	4.94E-	1.53E-16	2

			19		
			2.55E-		
COX7A2	9.802148	0.348014	18	7.19E-16	2
			4.96E-		
TCEAL7	9.65462	1.818731	17	1.18E-14	2
			7.28E-		
TUBB2B	9.602316	0.721075	18	1.99E-15	2
			2.60E-		
MAPT	9.537065	3.038497	16	5.56E-14	2
			8.32E-		
CAMK2N1	9.51276	1.536759	17	1.89E-14	2
			2.54E-		
CDKN2D	9.502693	2.57153	16	5.47E-14	2
			1.70E-		
ATP6V0B	9.336912	1.06556	16	3.72E-14	2
			8.23E-		
MARCKSL1	8.996672	0.592915	16	1.66E-13	2
			1.09E-		
CYCS	8.981799	0.775066	15	2.17E-13	2
			3.34E-		
MIAT	8.920246	1.708968	15	6.43E-13	2
			1.61E-		
TTC3	8.800306	0.458159	15	3.16E-13	2
			3.25E-		
NSG2	8.692225	4.229618	14	5.60E-12	2
			3.70E-		
PAK3	8.543448	2.305008	14	6.27E-12	2
			8.54E-		
YBX1	8.519643	0.296402	15	1.59E-12	2
			6.33E-		
SNAP25	8.496946	2.75222	14	1.05E-11	2
			1.30E-		
CFL1	8.427618	0.384698	14	2.38E-12	2
			1.32E-		
SYP	8.389491	3.276407	13	2.06E-11	2
			1.10E-		
SEZ6L2	8.344841	2.296895	13	1.74E-11	2
			5.62E-		
PFDN2	8.308831	0.933385	14	9.38E-12	2
			2.45E-		
PKIA	8.266267	2.843299	13	3.64E-11	2
			2.02E-		
GAP43	8.260351	2.567442	13	3.07E-11	2
			6.56E-		
MORF4L2	8.227035	0.430806	14	1.08E-11	2
			3.32E-		
FAM57B	8.226135	3.836013	13	4.83E-11	2
			2.05E-		
EEF1A2	8.222071	2.193839	13	3.10E-11	2
			3.68E-		
INA	8.218361	3.97344	13	5.29E-11	2
			1.66E-		
TSPAN13	8.217634	1.516299	13	2.54E-11	2
			4.36E-		
ELAVL4	8.189945	4.370355	13	6.12E-11	2
CELF4	8.059758	2.781397	6.70E-	9.10E-11	2

			13		
			5.83E-		
SYT1	8.002978	1.824554	13	7.99E-11	2
			3.39E-		
OCIAD2	8.000093	1.218046	13	4.91E-11	2
			1.02E-		
NGFRAP1	7.999276	0.347166	13	1.64E-11	2
			5.45E-		
SOX4	7.978378	0.94	13	7.52E-11	2
			5.01E-		
CUTA	7.927234	0.69503	13	6.95E-11	2
			1.03E-		
TAGLN3	7.873707	1.720155	12	1.35E-10	2
			2.29E-		
TTC9B	7.86841	4.232163	12	2.76E-10	2
			6.61E-		
BEX4	7.83059	0.669208	13	9.01E-11	2
			8.44E-		
TUBA1A	7.743577	0.509843	13	1.13E-10	2
			4.97E-		
PCSK1N	7.739162	0.64727	13	6.92E-11	2
			2.80E-		
MAP1LC3A	7.669214	1.2877	12	3.32E-10	2
			2.42E-		
MAP1LC3B	7.663591	0.967031	12	2.91E-10	2
			2.68E-		
ATP6V0E2	7.660034	1.3952	12	3.20E-10	2
			2.10E-		
C4orf48	7.563456	0.620204	12	2.58E-10	2
			8.49E-		
MAP2	7.497404	1.71927	12	9.39E-10	2
			5.82E-		
MED10	7.478306	0.921143	12	6.50E-10	2
			1.84E-		
MZT1	7.347036	1.380806	11	1.92E-09	2
			1.56E-		
PAFAH1B3	7.338198	1.132353	11	1.66E-09	2
			2.83E-		
LIN7B	7.314927	2.121174	11	2.82E-09	2
			2.55E-		
GPM6A	7.288104	1.554102	11	2.57E-09	2
			1.94E-		
PEBP1	7.18617	0.491886	11	2.02E-09	2
			9.22E-		
ATCAY	7.137292	3.803985	11	8.48E-09	2
			8.97E-		
CELF5	7.128315	3.179112	11	8.30E-09	2
			9.45E-		
CRMP1	7.080372	2.097416	11	8.67E-09	2
			1.24E-		
HOXB5	6.997617	1.357247	10	1.11E-08	2
			1.13E-		
APLP1	6.972839	1.232651	10	1.02E-08	2
			1.59E-		
TMSB15A	6.92975	1.445278	10	1.40E-08	2
RIC3	6.919869	2.015612	1.93E-	1.66E-08	2

			10		
			1.58E-		
ATP6V1F	6.835699	0.565674	10	1.40E-08	2
			2.77E-		
RNF5	6.807011	1.114555	10	2.32E-08	2
			1.99E-		
RPAIN	6.802578	0.759342	10	1.71E-08	2
			2.68E-		
TUBB2A	6.780778	0.925773	10	2.25E-08	2
			4.18E-		
C3orf14	6.748238	1.569444	10	3.42E-08	2
			2.02E-		
H2AFZ	6.721795	0.333707	10	1.73E-08	2
			4.80E-		
SOX11	6.717456	1.072838	10	3.89E-08	2
			8.18E-		
SCN3B	6.693927	3.324411	10	6.34E-08	2
			8.91E-		
SCG3	6.664262	3.098949	10	6.84E-08	2
			9.04E-		
DLL3	6.642696	2.381136	10	6.93E-08	2
			7.00E-		
YWHAH	6.551121	0.696715	10	5.53E-08	2
			1.42E-		
SBK1	6.536051	2.267136	09	1.05E-07	2
			1.32E-		
MAGEH1	6.496289	1.011905	09	9.90E-08	2
			1.97E-		
INPP5F	6.485771	2.211564	09	1.39E-07	2
			1.44E-		
TCP1	6.454723	0.929896	09	1.06E-07	2
			1.34E-		
USP11	6.452288	0.836512	09	1.00E-07	2
			3.58E-		
SCG2	6.375782	2.8943	09	2.37E-07	2
			2.78E-		
FAM177A1	6.359605	1.047916	09	1.89E-07	2
			3.95E-		
C10orf35	6.328913	1.909845	09	2.58E-07	2
			4.85E-		
ATP1A3	6.326269	3.310989	09	3.09E-07	2
			1.63E-		
PTMA	6.311715	0.170573	09	1.18E-07	2
			3.41E-		
TMEM59L	6.302464	1.22302	09	2.26E-07	2
			3.71E-		
COL1A1	21.75987	3.844498	34	1.19E-29	3
			7.26E-		
COL3A1	21.13053	3.423185	34	1.19E-29	3
			2.32E-		
COL1A2	19.65305	3.26777	31	1.52E-27	3
			5.10E-		
IFITM3	17.79524	1.662006	32	4.17E-28	3
			1.52E-		
FN1	16.82315	4.083968	26	3.83E-23	3
SPARC	16.80664	1.397022	3.01E-	1.64E-27	3

			31		
			8.64E-		
LGALS1	16.32668	3.000611	27	2.57E-23	3
			9.61E-		
COL6A3	14.88078	4.154115	24	1.27E-20	3
			1.70E-		
PTN	14.30265	1.82998	24	2.77E-21	3
			1.01E-		
CLDN11	14.20233	3.453712	22	1.23E-19	3
			2.70E-		
IFITM2	14.0598	1.966118	24	4.21E-21	3
			9.71E-		
MDK	13.67301	1.41967	24	1.27E-20	3
			2.17E-		
OST4	13.45084	0.476286	26	5.08E-23	3
			2.95E-		
PLAC9	12.89281	4.26562	20	2.36E-17	3
			3.86E-		
LHFP	12.85913	2.270729	21	3.66E-18	3
			6.29E-		
COL6A2	12.85497	2.458608	21	5.71E-18	3
			7.80E-		
TGFBI	12.66028	4.226897	20	5.68E-17	3
			1.15E-		
CYP1B1	12.63989	4.853955	19	8.15E-17	3
			4.35E-		
COL4A2	12.5334	2.942271	20	3.31E-17	3
			1.91E-		
TPM4	12.36754	1.867434	20	1.65E-17	3
			2.35E-		
NR2F1	12.11993	1.527674	20	1.97E-17	3
			5.63E-		
PHLDA1	11.99646	1.956238	20	4.19E-17	3
			2.26E-		
ARHGAP29	11.77061	3.474191	18	1.26E-15	3
			4.25E-		
COL5A1	11.49157	3.151294	18	2.28E-15	3
			6.76E-		
COL4A1	11.44216	3.452344	18	3.46E-15	3
			2.26E-		
DCN	11.41	2.753794	18	1.26E-15	3
			1.00E-		
PTGDS	11.14066	1.997069	17	4.96E-15	3
			2.67E-		
SERF2	10.95954	0.364455	19	1.71E-16	3
			1.34E-		
COL5A2	10.91657	2.193638	17	6.56E-15	3
			8.20E-		
PCOLCE	10.89172	3.483806	17	3.53E-14	3
			7.49E-		
LUM	10.84144	3.056897	17	3.31E-14	3
			1.02E-		
FBLN1	10.71632	3.155859	16	4.29E-14	3
			6.00E-		
EGR1	10.39282	0.840453	18	3.12E-15	3
CTSC	10.3914	2.909035	4.33E-	1.69E-13	3

			16		
			1.33E-		
MT-CO3	10.31723	0.323544	18	7.91E-16	3
			3.07E-		
SERPINH1	10.17951	1.651152	16	1.22E-13	3
			2.99E-		
OGN	10.08918	4.834265	15	1.03E-12	3
			2.56E-		
CTHRC1	10.08396	3.894068	15	9.01E-13	3
			1.01E-		
TMSB4X	10.03215	0.241476	16	4.29E-14	3
			7.90E-		
MT-CO1	10.02098	0.275799	18	3.98E-15	3
			4.27E-		
GMDS	9.956097	3.435536	15	1.43E-12	3
			1.89E-		
CALD1	9.794713	1.507762	15	6.78E-13	3
			1.84E-		
FOXC1	9.684173	5.854145	14	5.48E-12	3
			2.11E-		
OLFML3	9.576836	3.867126	14	6.22E-12	3
			3.54E-		
C7	9.526854	5.972115	14	9.90E-12	3
			3.62E-		
ADD3	9.513156	1.369918	15	1.22E-12	3
			2.22E-		
NID1	9.509205	3.329319	14	6.48E-12	3
			3.29E-		
IFITM1	9.478225	3.532731	14	9.28E-12	3
			3.04E-		
LAMA4	9.43771	3.432255	14	8.71E-12	3
			3.06E-		
CDC42EP5	9.403427	2.74668	14	8.71E-12	3
			8.23E-		
FBN2	9.246714	3.61977	14	2.17E-11	3
			1.08E-		
CPXM1	9.141879	2.791548	13	2.75E-11	3
			2.33E-		
ISLR	9.062643	4.877981	13	5.64E-11	3
			6.46E-		
TPM2	9.029902	2.001414	14	1.78E-11	3
			8.00E-		
SEPP1	8.916569	1.518518	14	2.13E-11	3
			3.59E-		
APOE	8.907168	2.984519	13	8.22E-11	3
			3.49E-		
RBMS3	8.853462	2.908474	13	8.03E-11	3
			6.11E-		
VIM	8.778646	0.406923	16	2.30E-13	3
			6.74E-		
LRRC17	8.772002	3.890897	13	1.47E-10	3
			7.03E-		
NID2	8.768668	4.412563	13	1.52E-10	3
			3.01E-		
SVIL	8.732174	1.794129	13	6.98E-11	3
NME4	8.722622	0.983376	6.74E-	1.82E-11	3

			14		
			1.60E-		
IGF2	8.710052	0.580612	14	4.79E-12	3
			2.44E-		
RHOBTB3	8.673872	1.441396	13	5.86E-11	3
			2.79E-		
FTL	8.662066	0.332771	13	6.58E-11	3
			2.02E-		
CEBPD	8.651706	1.393139	13	4.96E-11	3
			1.14E-		
ZIC1	8.562778	2.769041	12	2.38E-10	3
			1.42E-		
A2M	8.560263	3.386341	12	2.94E-10	3
			4.06E-		
MT-ND1	8.429554	0.386928	13	9.24E-11	3
			2.51E-		
PDGFRA	8.420974	3.840936	12	5.02E-10	3
			2.38E-		
SERPING1	8.405061	2.868908	12	4.79E-10	3
			1.04E-		
SELM	8.310821	1.199093	12	2.20E-10	3
			5.80E-		
BGN	8.258093	3.531859	12	1.10E-09	3
			2.31E-		
B2M	8.22094	0.380875	13	5.64E-11	3
			1.51E-		
GPC3	8.176388	1.333064	12	3.11E-10	3
			7.60E-		
HDDC2	8.17439	0.932805	13	1.64E-10	3
			6.23E-		
CCL2	8.163953	2.689397	12	1.17E-09	3
			5.24E-		
MT-ND3	8.133873	0.335069	13	1.17E-10	3
			4.17E-		
NR2F2	8.086003	1.483579	12	8.13E-10	3
			9.98E-		
UACA	8.077753	2.774861	12	1.81E-09	3
			8.87E-		
LRP1	7.902438	1.444546	12	1.64E-09	3
			2.38E-		
COLEC12	7.898678	3.2531	11	3.91E-09	3
			2.68E-		
FLRT2	7.895036	3.914958	11	4.37E-09	3
			1.14E-		
COL4A5	7.894229	1.671995	11	2.03E-09	3
			3.32E-		
OAF	7.845126	3.605274	11	5.26E-09	3
			3.78E-		
MMP2	7.812168	3.378187	11	5.83E-09	3
			1.97E-		
GYPC	7.76456	1.714021	11	3.30E-09	3
			6.41E-		
COL12A1	7.719585	4.555876	11	9.49E-09	3
			6.92E-		
LAMA2	7.704827	4.885502	11	1.02E-08	3
NFE2L2	7.680578	1.18678	2.26E-	3.77E-09	3

			11		
			3.49E-		
AKAP12	7.66318	1.396398	11	5.47E-09	3
			3.58E-		
TIMP1	7.62821	1.139792	11	5.59E-09	3
			9.23E-		
C11orf96	7.620945	3.660431	11	1.32E-08	3
			1.01E-		
PTH1R	7.613281	4.99255	10	1.44E-08	3
			8.82E-		
LAMB1	7.587922	3.075028	11	1.27E-08	3
			2.38E-		
CALU	7.584034	0.910027	11	3.91E-09	3
			1.26E-		
CD248	7.545947	4.115691	10	1.76E-08	3
			2.74E-		
TPT1	7.455735	0.19197	11	4.44E-09	3
CTD-			2.15E-		
2339F6.1	7.449546	6.278214	10	2.78E-08	3
			1.42E-		
NTN1	18.24527	2.724592	25	1.55E-21	4
			1.69E-		
AK1	13.96661	1.787753	20	6.13E-17	4
			1.89E-		
TPPP3	12.96088	1.511532	19	6.18E-16	4
			3.77E-		
C5orf49	11.57578	1.890433	17	6.50E-14	4
			7.31E-		
ANXA2	11.09514	0.987831	17	1.20E-13	4
			3.72E-		
CALM2	10.85914	0.425133	18	8.11E-15	4
			1.38E-		
OAZ2	10.11857	1.152594	14	1.50E-11	4
			5.98E-		
TMEM47	9.953992	1.211908	15	6.88E-12	4
			1.50E-		
ID3	9.796823	0.81022	15	1.89E-12	4
			1.36E-		
EFNB3	9.780455	2.389047	13	9.28E-11	4
			1.75E-		
CETN2	9.740498	1.008587	14	1.65E-11	4
			6.17E-		
NCAM1	9.687566	1.421027	14	4.92E-11	4
			2.29E-		
ALCAM	9.603195	2.09188	13	1.47E-10	4
			1.62E-		
DYNLL1	9.454106	0.435702	14	1.65E-11	4
			2.41E-		
VIM	9.086617	0.513686	14	2.14E-11	4
			1.41E-		
NLRP1	9.000818	1.371182	12	7.22E-10	4
			3.45E-		
IGFBP5	8.951391	0.922397	13	2.02E-10	4
			2.40E-		
C9orf24	8.911149	1.801921	12	1.16E-09	4
DSTN	8.885805	0.734921	2.48E-	1.18E-09	4

			12		
			2.70E-		
MEIS2	8.828055	0.84599	13	1.61E-10	4
			9.12E-		
HOTAIRM1	8.436363	1.016018	12	3.93E-09	4
			2.66E-		
FOXA1	8.384522	2.636913	11	1.01E-08	4
			3.16E-		
IGFBP7	8.128577	1.506852	11	1.11E-08	4
			7.75E-		
FSTL1	7.99484	1.586611	11	2.35E-08	4
			4.19E-		
S100A10	7.929429	0.994853	11	1.43E-08	4
			1.52E-		
FAM183A	7.847312	1.979927	10	4.15E-08	4
			1.20E-		
CRIP2	7.780401	1.319648	10	3.49E-08	4
			4.51E-		
SPARC	7.639067	0.774391	11	1.52E-08	4
			6.76E-		
TMEM190	7.446092	1.935608	10	1.61E-07	4
			5.31E-		
LY6H	7.418316	1.496326	10	1.30E-07	4
			7.79E-		
COL18A1	7.409063	1.896004	10	1.78E-07	4
			9.44E-		
AGT	7.39417	2.332941	10	2.09E-07	4
			1.46E-		
ARX	7.364155	4.340382	09	3.08E-07	4
			8.99E-		
CD47	7.327478	1.366039	10	2.02E-07	4
			8.63E-		
C1orf192	7.32324	1.671953	10	1.95E-07	4
			1.70E-		
SHH	7.316028	3.996883	09	3.54E-07	4
			1.36E-		
RSPH1	7.231162	1.677927	09	2.90E-07	4
			1.16E-		
TUBB4B	7.215828	1.001976	09	2.53E-07	4
			9.41E-		
METRN	7.19123	0.983802	10	2.09E-07	4
			1.26E-		
IFI27L2	7.122876	0.740032	09	2.71E-07	4
			2.38E-		
ZNF503	7.038138	1.409023	09	4.72E-07	4
			2.00E-		
CLU	7.01683	0.908186	09	4.12E-07	4
			4.25E-		
PON2	6.971555	1.721934	09	7.68E-07	4
			9.45E-		
TCTEX1D1	6.787803	2.479483	09	1.53E-06	4
			3.93E-		
GPM6B	6.773056	0.820948	09	7.26E-07	4
			3.95E-		
PSAP	6.72331	0.550708	09	7.26E-07	4
EGLN3	6.707788	1.730282	1.18E-	1.86E-06	4

			08		
DRD1	6.667789	3.800717	1.83E-08	2.66E-06	4
CALM1	6.606616	0.48536	7.09E-09	1.20E-06	4
EFCAB1	6.590205	2.071732	1.82E-08	2.66E-06	4
ADAM32	6.508233	3.37309	3.17E-08	4.20E-06	4
TTC6	6.504223	4.785626	3.45E-08	4.49E-06	4
LRRIQ1	6.456492	1.809516	2.88E-08	3.92E-06	4
PIFO	6.420805	1.400932	3.01E-08	4.06E-06	4
CCDC80	6.308489	1.757096	5.25E-08	6.20E-06	4
ATP1B1	6.304827	0.678284	2.73E-08	3.79E-06	4
RBP1	6.289139	0.437076	1.53E-08	2.29E-06	4
DYNLT1	6.220052	0.531025	4.58E-08	5.59E-06	4
CITED2	6.193944	1.176539	6.43E-08	7.26E-06	4
SLIT2	6.161768	2.720593	1.07E-07	1.09E-05	4
PLTP	6.142779	1.370353	8.80E-08	9.48E-06	4
CD99	6.044263	0.468942	5.12E-08	6.09E-06	4
RNF19A	6.038345	1.553605	1.30E-07	1.28E-05	4
PDLIM3	6.005267	1.074176	9.74E-08	1.03E-05	4
C2orf40	5.995108	2.658917	1.88E-07	1.75E-05	4
HES1	5.959014	1.053418	1.54E-07	1.48E-05	4
SERPINF1	5.953855	1.740695	1.98E-07	1.84E-05	4
MLF1	5.940629	0.895251	1.50E-07	1.44E-05	4
PPP1R14C	5.889889	1.935476	2.58E-07	2.27E-05	4
CD36	5.888593	3.784736	3.06E-07	2.62E-05	4
LIFR	5.875664	1.365396	2.41E-07	2.17E-05	4
CD9	5.852355	1.227613	2.48E-07	2.21E-05	4
CD276	5.846605	2.098003	3.18E-07	2.71E-05	4
LPPR1	5.792294	3.196433	4.14E-07	3.38E-05	4
RP11-	5.775599	0.947041	2.66E-	2.34E-05	4

620J15.3			07		
RFX4	5.749254	1.198605	3.61E-07	3.01E-05	4
CANX	5.730496	0.669158	2.95E-07	2.55E-05	4
KIF9	5.727543	1.166792	3.37E-07	2.85E-05	4
FAM81B	5.72645	1.712396	3.96E-07	3.25E-05	4
CMTM8	5.710558	2.629779	5.59E-07	4.40E-05	4
FAM92B	5.649282	2.45398	6.67E-07	5.17E-05	4
APLP2	5.645802	0.674427	3.82E-07	3.17E-05	4
TSPAN2	5.643997	2.362039	6.79E-07	5.24E-05	4
ODF3B	5.616808	2.000411	7.13E-07	5.47E-05	4
C9orf9	5.59906	1.888681	7.22E-07	5.49E-05	4
SRP14	5.556377	0.182476	3.39E-07	2.86E-05	4
TAGLN2	5.549316	1.139767	8.26E-07	6.13E-05	4
IQCG	5.54667	1.734337	8.39E-07	6.21E-05	4
CSPP1	5.53237	1.991009	9.55E-07	6.85E-05	4
CTC-503J8.6	5.527835	1.72005	9.16E-07	6.64E-05	4
DALRD3	5.523158	1.161624	7.63E-07	5.75E-05	4
HTRA1	5.498015	1.561638	9.82E-07	6.97E-05	4
CCDC74A	5.484398	2.117733	1.18E-06	8.11E-05	4
FOXJ1	5.473944	1.596771	1.07E-06	7.40E-05	4
DRC1	5.468164	2.503925	1.26E-06	8.63E-05	4
PCSK1N	5.45467	0.571163	7.46E-07	5.64E-05	4
STC2	5.438828	1.792687	1.33E-06	9.05E-05	4
FAM229B	5.438643	0.796897	9.75E-07	6.94E-05	4
SMDT1	5.416637	0.616218	9.98E-07	7.04E-05	4
ZMYND10	5.411716	1.818095	1.40E-06	9.50E-05	4
LGALS1	21.73431	2.536144	1.09E-17	9.49E-16	5
COL3A1	16.01	2.350916	4.43E-14	2.12E-12	5
TPT1	14.34428	0.327013	2.18E-	1.10E-12	5

			14		
			1.58E-		
COL1A1	12.73698	2.381921	11	5.21E-10	5
			1.96E-		
NAP1L1	12.661	0.643848	12	7.44E-11	5
			3.13E-		
IFITM2	12.43043	1.893137	11	9.82E-10	5
			1.17E-		
IFITM3	11.97955	1.285544	11	3.94E-10	5
			3.89E-		
TMSB4X	11.91724	0.3425	11	1.20E-09	5
			5.85E-		
COL1A2	11.77355	2.067076	11	1.75E-09	5
			1.53E-		
RPL15	10.74702	0.212478	10	4.26E-09	5
			2.01E-		
RPS7	10.06628	0.372627	09	4.63E-08	5
			7.92E-		
RPS12	9.707006	0.230375	10	1.95E-08	5
			1.07E-		
RPL23A	9.666986	0.2756	09	2.61E-08	5
			6.48E-		
RPL23	9.562958	0.281241	10	1.63E-08	5
			1.10E-		
CALD1	9.265505	1.615825	08	2.27E-07	5
			2.82E-		
SPRY1	9.131452	3.008743	08	5.48E-07	5
			6.22E-		
RPLP1	8.636672	0.186484	09	1.34E-07	5
			1.63E-		
EMP3	8.630044	1.314259	08	3.28E-07	5
			4.99E-		
PHLDA1	8.606553	2.028747	08	9.28E-07	5
			4.41E-		
RPL9	8.5953	0.187051	09	9.74E-08	5
			1.95E-		
RPS3	8.542188	0.240244	08	3.88E-07	5
			9.66E-		
RPL28	8.475125	0.211084	09	2.02E-07	5
			1.11E-		
RPL10	8.427394	0.174906	08	2.29E-07	5
			9.97E-		
RARRES2	8.400598	2.936508	08	1.75E-06	5
			2.15E-		
RPL37	8.370809	0.252265	08	4.26E-07	5
			2.87E-		
EGR1	8.305257	0.868814	08	5.57E-07	5
			3.70E-		
MYL12A	8.227308	0.95329	08	7.03E-07	5
			3.34E-		
RPL10A	8.165697	0.294037	08	6.40E-07	5
			3.18E-		
RPS8	8.125025	0.196866	08	6.11E-07	5
			3.66E-		
RPL19	8.017469	0.192025	08	6.97E-07	5
IGF2	7.987895	0.784997	7.12E-	1.28E-06	5

			08		
			1.03E-		
RPL27	7.774965	0.226439	07	1.79E-06	5
			1.08E-		
RPS20	7.469883	0.264007	07	1.89E-06	5
			5.48E-		
RPL27A	7.449187	0.13249	08	1.01E-06	5
			1.79E-		
RPL30	7.344737	0.183729	07	2.98E-06	5
			1.21E-		
RPS28	7.327758	0.190264	07	2.09E-06	5
			2.16E-		
RPL13	7.302952	0.160566	07	3.55E-06	5
			3.29E-		
RPS23	7.260891	0.205849	07	5.28E-06	5
			1.13E-		
EMILIN1	7.12999	4.123562	06	1.66E-05	5
			3.24E-		
RPL6	6.968545	0.165012	07	5.21E-06	5
			1.62E-		
PLAT	6.941542	3.648149	06	2.31E-05	5
			7.61E-		
ACTG1	6.940427	0.316112	07	1.14E-05	5
			5.07E-		
RPS9	6.860031	0.232018	07	7.85E-06	5
			4.86E-		
RPS6	6.837274	0.213485	07	7.56E-06	5
			7.22E-		
RPL24	6.831696	0.175085	07	1.09E-05	5
			6.59E-		
RPL32	6.749939	0.151855	07	1.00E-05	5
			1.18E-		
EEF1A1	6.659861	0.151818	06	1.73E-05	5
			3.07E-		
TPM4	6.457321	1.528599	06	4.15E-05	5
			2.17E-		
RPL7	6.392414	0.20372	06	3.00E-05	5
			4.33E-		
COL5A1	6.374739	2.45348	06	5.70E-05	5
			1.39E-		
B2M	6.371245	0.383169	06	2.01E-05	5
			3.43E-		
RPS29	6.254496	0.197605	06	4.60E-05	5
			1.37E-		
RPL39	6.242198	0.140289	06	1.98E-05	5
			6.15E-		
CEBPD	6.203433	1.779435	06	7.87E-05	5
			2.59E-		
RPL35A	6.195333	0.14939	06	3.54E-05	5
			4.71E-		
IER2	6.190475	0.914717	06	6.15E-05	5
			1.96E-		
RPLP2	6.183686	0.138103	06	2.75E-05	5
			8.99E-		
CCL2	6.044748	2.635301	06	0.000111	5
MFAP4	6.040723	3.92719	1.01E-	0.000123	5

			05		
RPL22	6.026644	0.192077	3.80E-06	5.06E-05	5
RPS15A	6.001408	0.165982	4.96E-06	6.46E-05	5
RPS27	5.993874	0.144641	5.28E-06	6.84E-05	5
LTBP4	5.969494	2.195638	1.05E-05	0.000128	5
RPS4X	5.950865	0.208624	5.15E-06	6.68E-05	5
RPL21	5.929594	0.135272	6.23E-06	7.95E-05	5
ATP5O	5.910908	0.485466	4.42E-06	5.81E-05	5
CTSC	5.878955	2.357586	1.24E-05	0.000149	5
RPL37A	5.874602	0.112584	4.39E-06	5.77E-05	5
RPL34	5.84981	0.150496	6.49E-06	8.23E-05	5
SLC25A6	5.836078	0.484323	8.58E-06	0.000107	5
RPL14	5.829963	0.242102	8.66E-06	0.000107	5
RPS27A	5.810854	0.145055	7.78E-06	9.73E-05	5
RPS18	5.763656	0.147499	9.09E-06	0.000112	5
RPL4	5.742888	0.289623	1.20E-05	0.000145	5
PFN1	5.672705	0.53113	1.28E-05	0.000154	5
RPL12	5.670303	0.221765	8.54E-06	0.000106	5
MYL9	5.579537	2.306845	2.37E-05	0.000271	5
THBS1	5.495329	3.495163	3.07E-05	0.000342	5
COL6A3	5.453186	2.559046	3.13E-05	0.000347	5
FHL2	5.423804	3.61917	3.63E-05	0.000397	5
COL6A2	5.356497	1.849028	3.69E-05	0.000403	5
CCNB1IP1	5.297158	0.862686	2.71E-05	0.000304	5
SSR2	5.283417	0.693309	3.36E-05	0.000371	5
RPL11	5.261445	0.16375	3.23E-05	0.000357	5
SPARC	5.258349	0.76311	3.00E-05	0.000335	5
HNRNPA1	5.25238	0.356639	3.79E-05	0.000412	5
CEBPB	5.233003	1.677013	4.98E-	0.000531	5

			05		
FBLN1	5.220429	2.754532	5.45E-05	0.000576	5
NPM3	5.191805	1.643186	5.24E-05	0.000556	5
NFKBIA	5.18505	1.538908	5.47E-05	0.000579	5
RPS3A	5.184479	0.116652	3.03E-05	0.000338	5
CXCL12	5.16465	3.260646	6.28E-05	0.000655	5
NUPR1	5.122091	2.802576	6.64E-05	0.000688	5
PCDH18	5.119086	2.965793	6.83E-05	0.000706	5
PLAC9	5.107369	2.312895	6.35E-05	0.000663	5
TCF21	5.064339	8.756227	8.08E-05	0.000824	5
TPM2	5.007429	2.043595	8.30E-05	0.000845	5
RPS16	5.002708	0.160356	5.84E-05	0.000615	5
PLK2	4.975311	2.110635	8.86E-05	0.000898	5
OLFML3	4.951478	2.600256	9.41E-05	0.000949	5

5.8 References

1. Perrin S. Preclinical research: Make mouse studies work. Nature. 2014 Mar 27;507(7493):423–5.

2. Kim CFB, Jackson EL, Woolfenden AE, Lawrence S, Babar I, Vogel S, et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell. 2005 Jun 17;121(6):823–35.

3. Rock JR, Onaitis MW, Rawlins EL, Lu Y, Clark CP, Xue Y, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. Proceedings of the National Academy of Sciences. 2009 Aug 4;106(31):12771–5.

4. Rock JR, Randell SH, Hogan BLM. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. Dis Model Mech. 2010 Sep 1;3(9-10):545–56.

5. Miller AJ, Hill DR, Nagy MS, Aoki Y, Dye BR, Chin AM, et al. In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. Stem Cell Reports. 2018 Jan 9;10(1):101–19.

6. Miller AJ, Yu Q, Czerwinski M, Tsai Y-H, Conway RF, Wu A, et al. Basal stem cell fate specification is mediated by SMAD signaling in the developing human lung.

7. Miller AJ, Dye BR, Ferrer-Torres D, Hill DR, Overeem AW, Shea LD, et al. Generation of lung organoids from human pluripotent stem cells in vitro. Nat Protoc. 2019.

8. Dye BR, Hill DR, Ferguson MA, Tsai Y-H, Nagy MS, Dyal R, et al. In vitro generation of human pluripotent stem cell derived lung organoids. Elife. 2015;4.

9. Kebschull JM, Zador AM. Cellular barcoding: lineage tracing, screening and beyond. Nature Methods 2018 15:11. Nature Publishing Group; 2018 Nov 1;15(11):871–9.

10. Hwang B, Lee W, Yum S-Y, Jeon Y, Cho N, Jang G, et al. Lineage tracing using a Cas9-deaminase barcoding system targeting endogenous L1 elements. Nat Commun. Nature Publishing Group; 2019 Mar 15;10(1):1234.

11. Jacob A, Morley M, Hawkins F, McCauley KB, Jean JC, Heins H, et al. Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. Cell Stem Cell. 2017 Oct 5;21(4):472–488.e10.

12. Benam KH, Villenave R, Lucchesi C, Varone A, Hubeau C, Lee H-H, et al. Small airway-on-a-chip enables analysis of human lung inflammation and drug responses in vitro. Nature Publishing Group. 2016 Feb;13(2):151–7.

13. Stucki JD, Hobi N, Galimov A, Stucki AO, Schneider-Daum N, Lehr C-M, et al. Medium throughput breathing human primary cell alveolus-on-chip model. Sci Rep. Nature Publishing Group; 2018 Sep 25;8(1):14359.