UNIVERSITY OF CALIFORNIA

Los Angeles

Lung Organoids for Personalized Disease Modeling

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Materials Science and Engineering

by

Dan Charles Wilkinson Jr.

2017

ABSTRACT OF THE DISSERTATION

Lung Organoids for Personalized Disease Modeling

by

Dan Charles Wilkinson Jr.

Doctor of Philosophy in Materials Science and Engineering

University of California, Los Angeles, 2017

Professor Bruce S. Dunn, Chair

Stem cell technologies, especially patient-specific, induced stem cell pluripotency and directed differentiation, hold great promise for changing the landscape of medical therapies, and will usher in a new era of personalized medicine. Induced pluripotent stem cells (iPSCs) are patient derived cells that may be expanded indefinitely and differentiate into every known cell type in the body providing the potential basis for personalized organ transplants and disease models. In order to meet these challenges, organoids and 3D tissue engineering approaches are being developed though there is still a large technical gap between promise and current technological expertise.

The work presented in this dissertation is founded on the development of a personalized medicine process flow for modeling Idiopathic Pulmonary Fibrosis (IPF). The basis of this technology is the development of the lung organoid, a 3D cell/hydrogel composite that mimics the alveolar geometry of human distal lung. We generated a model of IPF by culturing organoids with TGF-β1 and showed the resulting scarring in a dish was phenotypically similar to that seen in IPF histology. By optimizing the lung organoid process flow for producing large numbers of uniform organoids and inclusion of IPF patient derived mesenchymal cells we demonstrated how this

method could be used for high throughput drug discovery. Finally, we developed an artificial neural network for the classification of high throughput drug screening data and showed its applicability in classifying the complex phenotypic patters organoids demonstrate when treating organoids with dimethyl sulfoxide. In total, this work provides a blueprint for 3D phenotypic drug discovery in the context of the lung organoid model of IPF.

The dissertation of Dan Charles Wilkinson Jr. is approved.

Brigitte Gomperts

Dwight Streit

Qibing Pei

April Pyle

Bruce S. Dunn, Committee Chair

University of California, Los Angeles

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis is the culmination of years of highly collaborative work that would have been impossible if not for the kindness and generosity of a number of individuals and organizations. It was amazing to experience the scientific process firsthand and I have gained a new level of appreciation for the complex interconnected processes that must all work together in order to enable science and discovery. All of this was facilitated by my mentor, Dr. Bruce Dunn. Dr. Dunn opened up the way to the fellowships and collaborations that allowed for my projects and ideas to come to fruition. He gave me the freedom to pursue my own ideas while providing the guidance necessary for me to stay on track.

Science does not happen without funding and without the various fellowships I have been awarded over the years I would have not been able to accomplish the work outlined here. First and foremost, I have to thank the National Science Foundation for awarding me the NSF Graduate Student Fellowship. This fellowship, more than anything else, allowed me the freedom to pursue this work with the confidence of a guaranteed three years of financial support. I was also awarded the Broad Stem Cell Research Training Grant and the Nippon Sheet Glass Fellowship during my time at UCLA, each of which I am also very grateful for.

Collaboration is the backbone of scientific endeavor and I have had the pleasure of working with some of the most talented biologists at UCLA. Nearly all of this work was carried out in Dr. Brigitte Gomperts lab. Brigitte has been the catalyst of many of the ideas and concepts presented here and the support that both she and her lab has shown me has been remarkable. More importantly, she trained me to think like a clinician and focus my research on helping people. I could not be more thankful for her mentorship; she made me feel like a member of her family. In Brigitte's lab I have to thank a number of people including Jackie Alva-Ornelas, Preethi Vijayaraj,

Jennifer Sucre, Sarah Ahadome, Cody Aros, Abdo Durra, Manash Paul, and Tammy Rickabaugh. Jackie was there from the start, she taught me biological technique, cultured my cells and helped with almost every step of the organoid generation and analysis process. Without Jackie, there would have been no lung organoid. Preethi provided me great insight into fibrosis and helped develop the fundamental ideas that lead to the generation of the IPF model. Furthermore, she was an essential sounding board for all of my ideas and without her help the organoid would never have become a model of IPF. Jen was the first to want to use the lung organoid for actually performing disease modeling. She showed that the organoid could act as a substrate for modeling BPD and helped to get me thinking about how these techniques could and should be used. Sarah and I worked together to expand the usefulness of the lung organoid, specifically in its ability to provide a better model of small cell lung cancer, although none of that work has been included here. Abdo was a great help in the lab and really streamlined the hydroxyproline analysis. Manash helped me with confocal but more importantly kept me grounded by providing a continued stream of critical analysis as the project grew and evolved. Tammy came into the project a little later but her support and help was critical in keeping everything going and allowing me to focus on science. I wish all of you the best of luck and am eternally grateful for all you have taught me.

There are also a number of other people at UCLA that I need to thank. My first mentor at UCLA was Wade Richardson. Wade was just finishing up when I started but although our time together was brief he taught me a wide range of scientific techniques and introduced me to machine learning. He was also fundamental in initiating the lung organoid project and the lessons we learned together put me in the right direction for continuing and improving the organoid formation methodology. Ultimately, Wade provided me the tools I needed to do great science and interface with biologists. I also need to thank Steve Jonas. Steve was never directly involved in any of the

projects described here though we have turned a couple brainstorming sessions into a working transfection-oriented project. I also need to thank Robert Damoiseaux and Bobby Tofig at the Molecular Screening Shared Research core in CNSI. They helped me scale up lung organoid production for drug screening. I also need to thank April Pyle and her lab. While we never really got any projects off the ground April introduced me to stem cell culture and analysis for which I am very thankful. Finally, I had the chance to work with several highly talented undergraduate students during my time at UCLA including Michael Mellody and Luisa Meneses. Both Michael and Luisa have shown much promise and I expect to see them accomplish great things over the next few years.

I also need to thank my family and friends for the help and support over the years. Science can be an emotionally straining and highly stressful process and it was great to have people that cared for me cheering me on. My parents, Dan and Annette Wilkinson, have been anchors in a sea of uncertainty and our frequent phone conversations really helped me to stay focused and motivated. I have made many great friends while at UCLA and I would like to thank Owen Liang, Pedro Arias, Andrew Siordia, Max Kuciej, Xavier Petrissans, Mac Daley, Saran Karumbayaram, El Hadi Caoui, Doug Chen, Leland Smith, Dean Cheikh, Daniel Sivalingam, Drake Smith, Thomas Young, Nathan Majernik, Briant Vo, and Majib Jan.

The following thesis, excluding chapter 4, is based on the publication: “Wilkinson, D. C. *et al.* Development of a Three-Dimensional Bioengineering Technology to Generate Lung Tissue for Personalized Disease Modeling. *Stem Cells Transl. Med.* **6,** 622-633 (2016),” DOI: 10.5966/sctm.2016-0192.

VITA

EDUCATION

University of California, Los Angeles 2012- present
- Pursuing Ph.D. in Materials Science

University of New Mexico 2007-2012
- Bachelor of Science in Physics
- Bachelor of Science in Mathematics
- Magna Cum Laude

ACADEMIC AWARDS AND DISTINCTIONS

Broad Stem Cell Research Center Training Grant – University of California, Los Angeles 2016-current

NSF Graduate Student Fellowship - University of California, Los Angeles 2013-2016

Nippon Sheet Glass Fellowship Recipient - University of California, Los Angeles 2012-2013

Presidential Scholarship - University of New Mexico 2007-2011

RESEARCH ACTIVITY AND PRESENTATIONS

Next Gen Stem Cell Conference – Invited Talk 2017
"3D Bioengineered Lung Tissue for Personalized Disease Modeling"

International Society for Stem Cell Research Conference - Poster 2016
"Three Dimensional Bioengineered Lung Tissue for Personalized Disease Modeling"

UCLA Annual Stem Cell Symposium - Poster 2016, 2017
"Three Dimensional Bioengineered Lung Tissue for Personalized Disease Modeling"

ArtCenter College of Design Lecture 2015, 2016
"Biotechnology, Nanomedicine and Tissue Engineering"

UCLA Stem Cell Club Presentation 2016
"3D Bioengineered Lung Tissue for Personalized Disease Modeling"

UCLA Tech Forum – Best in Poster Design 2014
"Bioengineering Personalized Functional Lungs using Induced Pluripotent Stem Cell Technology"

PUBLICATIONS

1. **Wilkinson D. C.** *et al.* Development of a Three-Dimensional Bioengineering Technology to Generate Lung Tissue for Personalized Disease Modeling. *Stem Cells Transl. Med.* 1–12 (2016). doi:10.5966/sctm.2013-0145
2. Danielle C., Wilkinson D., *et al.* High Temperature Structural Stability of Ceria Based Inverse Opals. *J. Am. Ceram. Soc.* (2017). (Accepted for Publication)
3. Sucre J. M. S., Wilkinson D., *et al.* Posttranslational modification of β-catenin is associated with pathogenic fibroblastic changes in bronchopulmonary dysplasia. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **312,** 186–195 (2017).
4. Sucre J. M. S., Wilkinson D., *et al.* A three-dimensional human model of the fibroblast activation that accompanies bronchopulmonary dysplasia identifies Notch-mediated pathophysiology. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **310,** 889–898 (2016).

5. Richardson W., Wilkinson D., *et al.* Ensemble multivariate analysis to improve identification of articular cartilage disease in noisy Raman spectra. *J. Biophotonics* **8,** 555–566 (2015).
6. Xiong S., Wilkinson D., *et al.* Revealing the interfacial self-assembly pathway of large-scale, highly-ordered, nanoparticle/polymer monolayer arrays at an air/water interface. *Nano Lett.* **13,** (2013).
7. Epler K., Wilkinson D., *et al.* Delivery of Ricin Toxin A-Chain by Peptide-Targeted Mesoporous Silica Nanoparticle-Supported Lipid Bilayers. *Adv. Healthc. Mater.* **1,** 348–53 (2012).
8. Ashley C. E., Wilkinson D., *et al.* Delivery of Small Interfering RNA by Peptide-Targeted Mesoporous Silica Nanoparticle-Supported Lipid Bilayers. *ACS Nano* **6,** 2174–88 (2012).

Chapter 1. Introduction and Background

Chapter 1.1. Overview

In 2006, Shinya Yamanaka was the first to report that primary, terminally differentiated cells may be transformed into pluripotent stem cells[1]. This discovery has since reshaped the landscape of biology, inspired new routes of scientific inquiry, and heralded a new age of personalized medicine. These resulting cells were dubbed induced pluripotent stem cells (iPSCs) and showed similar properties to the traditionally studied and well characterized embryonic stems cells (ESCs). Specifically, iPSCs have the capacity to differentiate into all three germ layers and, subsequently, any cell type in the body. Furthermore, these cells contain the DNA specific to the individual who initially donated the primary cells for reprogramming. By combining these two properties of iPSCs, their pluripotency and patient-specific DNA, it is now possible to begin replicating, studying, and expanding an individual's own cells *ex vivo* without the need to immortalize them. The potential applications of iPSC technology range from personalized disease modeling to the generation of non-immunogenic transplantable tissues and organs. Though there is a deep chasm between the technical possibility of an iPSC-based personalized medicine and its widespread implementation. The challenges remaining to the field are numerous and range from the basic biology governing efficient iPSC generation, to the intricacies of growing and vascularizing organs from single cells. While some of these challenges are daunting, the promise of iPSCs has marshaled a fast growing, interdisciplinary research community to bring personalized medicine to the public.

The goal of the research presented here is to describe the materials and methods we have developed in an effort to exploit the clinical potential of the burgeoning field of iPSC-based disease modeling. The root of the technological advancement outlined here is the development of the lung

organoid. The lung organoid is a multicellular hydrogel composite that seeks to mimic the alveolar sac geometry of the human distal lung. Furthermore, the methodologies that allow for the production of such organoids is highly reproducible, scalable, and flexible to the inclusion of a wide range of cell types. Once developed, these lung organoids were used to generate a model if Idiopathic Pulmonary Fibrosis (IPF), replicating the phenotypic signatures of IPF and bronchopulmonary dysplasia (BPD) in a dish[2,3,4]. Finally, we show our preliminary efforts in the application of supervised machine learning to the classification of organoids generated for high throughput drug screening. When combined, these materials and techniques form a platform for an iPSC-driven, organoid-based high throughput screening and drug discovery.

Chapter 1.2. Lung Disease: Unmet Clinical Needs

Chapter 1.2.1. Lung Anatomy and Physiology

In order to appreciate the initiation, progression and other nuances of lung diseases it is essential to understand the anatomy and physiology of healthy lung tissue. The lungs are two fairly symmetric organs that reside in the chest cavity whose main function is to provide a respiratory membrane for gas exchange enabling sustained aerobic metabolism(**Fig. 1**). The lung may be broken down into two general regions, proximal and distal lung. Proximal lung provides mechanical support and air-conducting channels whereas distal lung provides a large respiratory membrane for gas exchange between the blood and the alveolar void space.

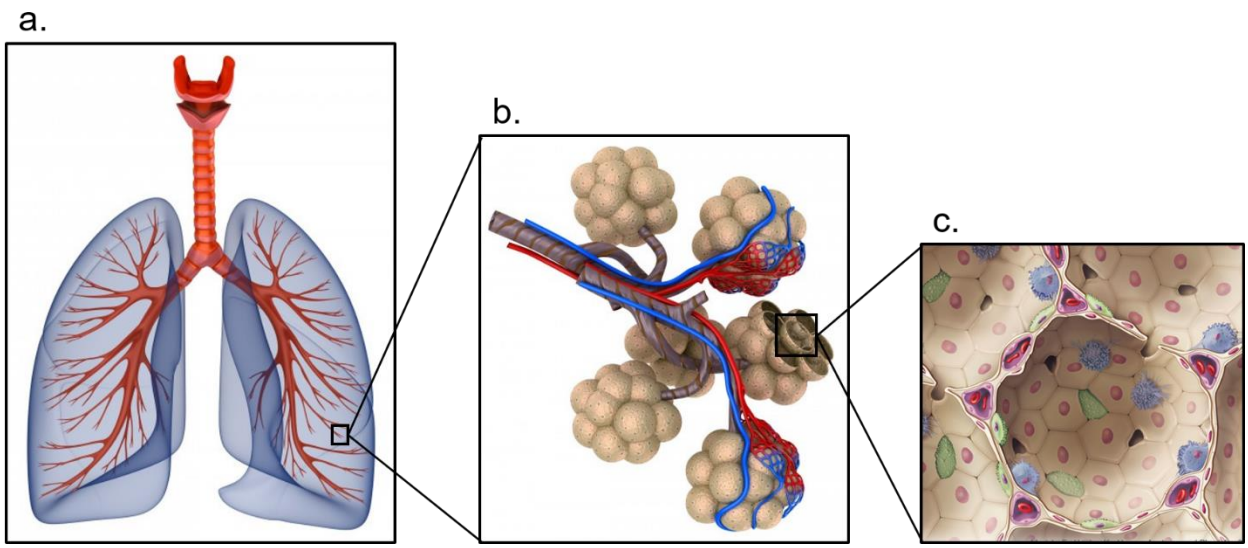


Figure 1: Overview of lung anatomy and cellular distributions. (**a**) The lungs are two symmetric lobes whose function is to provide a respiratory membrane for gas exchange. (**b**) Bronchial branching terminates into highly vascularized alveolar sacs. (**c**) The respiratory membrane is formed by the close opposition of type 1 epithelial cells and microvascular endothelial cells. Type 2 epithelial cell, macrophages, fibroblasts and a number of other minority cell types occupy the distal lung tissue.

The lungs originate as a series of bifurcating channels that conduct air from outside the body toward the lower, distal regions of the lung. Initiating at the base of the trachea, the left main stem and right main stem bronchus branch into bronchi, bronchioles, alveolar ducts and finally the alveolar sacs. Mechanically, the trachea and bronchi are supported and held open by periodically spaced cartilaginous rings though as the airways continue to narrow into terminal bronchioles these rings disappear. On average, each bronchus will undergo 23 divisions/branchings of which the bulk of the initial generations do not participate in actual gas exchange and account for what is known as the ‘dead volume’ of the lung. Each of these branching terminates into clusters of tightly packed alveolar sacs. The lungs are ventilated by the cyclic action of the diaphragm wherein upon expansion of the lungs a transient pressure differential causes atmospheric air to rush in and then out upon the subsequent contraction. The lung spaces maintain an open state in part due to the

continuous pressure exerted by the lymph system continually pumping fluid out of the plural space (space between the lung and inner chest wall).

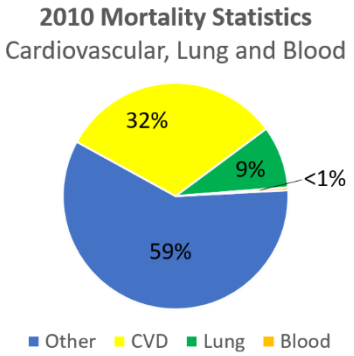
The cellular distribution of the lung varies gradually from proximal to distal lung. Proximal lung is lined with a combination of ciliated, secretory, basal, and undifferentiated cell types whose functions are to produce the mucus and tractional forces that clear the airways of inhaled debris and particulate matter. Moving to the distal lung, the bronchioles terminate into alveolar channels and finally networks of interconnected, highly vascularized alveolar sacs containing a combination of type I and II epithelial cells and interpenetrating endothelial capillaries. Fibroblasts occupy the intestinal spaces and are important for maintaining the extracellular matrix (ECM) and assist in wound healing. The alveolar sacs form the respiratory membrane, the portion of the lung that enables gas exchange between the blood and the atmosphere contained in the alveolar void space. The geometry of the alveolar sacs imparts a large surface area for gas exchange equaling between 50-75m^2. Furthermore, the lungs contain nearly 100mL of capillary blood in close proximity to the void space at an average distance of ~0.5μm between the alveolar capillaries and the atmospheric void space. This high surface area and close proximity to the capillary blood is necessitated by the properties of gas transport and diffusion into the lung tissue. Specifically, gasses diffuse across the alveolar membrane due to the difference in partial pressures of oxygen and carbon dioxide between the blood and alveolar atmosphere.

The lung's architecture and physiology allow for efficient gas exchange that under normal conditions easily meet metabolic demand. Though, the same design criteria that imparts these characterizes also makes the lung an easily damaged organ. The dynamics of gas diffusion require an exceedingly thin respiratory membrane that, due to a lack of mechanical integrity, is easily destroyed. Indeed, lung function will decline with age regardless of environment factors or any

level of preventative care measures. This decline is marked by the loss of alveolar sacs and functional lung volume. There is no known mechanism for the regeneration or repair of a damaged respiratory membrane.

Chapter 1.2.2. Lung Disease Statistics and Prevalence

According to the National Institute of Health (NIH) lung diseases cause nearly 10% of all deaths in the United States totaling over 200,000 people every year[5] (**Fig. 2**). The economic impact of these diseases were over $100 billion in 2009 with costs expected to increase as the population grows and ages. Furthermore, the clinical options for treating lung disease are primarily palliative and only provide temporary symptomatic relief[6,7,8]. These diseases include chronic obstructive pulmonary disease (COPD), lung disease due to external agents, interstitial lung disease, and



Cause of Death (2010)	Percentage of Total Lung Related Deaths
Chronic Obstructive Pulmonary Disease	58%
Influenza and Pneumonia	22%
Lung Disease due to External Agents	9%
Interstitial Lung Diseases	4%
Neonatal Pulmonary Disorders	3%
Asthma	2%
Other Lung Diseases	2%

Figure 2: Lung disease statistics according to the 2012 NIH factbook. CVD: cardiovascular disease.

neonatal pulmonary disorders[5]. Combinations of steroids, bronchodilators, oxygen therapy and antibiotics are commonly prescribed to patients suffering from these diseases but there are currently are no cures. In addition, there have been few new classes of drug therapies introduced in the past 40 years for treatment of lung disease. These factors highlight that lung diseases represent a major unmet clinical need.

Chapter 1.2.3. Idiopathic Pulmonary Fibrosis

Idiopathic pulmonary fibrosis (IPF) is a chronic lung disease characterized by the progressive, irreversible scarring of the lung. While IPF is rare in comparison with other lung conditions it is the most common cause of death from progressive lung disease having no effective therapies other than a lung transplantation and accounts for as many as 17,00 deaths per year in the United States. An IPF diagnosis will invariably result in death with nearly 50% of patients dying within 3 years of diagnosis. Along with this high mortality rate IPF imposes a significant economic burden of over $4 billion every year. Lung transplantation is the gold standard for treating terminal IPF patients though its efficacy at significantly extending life is limited to a 50% 5 year survival rate. While this is a significant improvement from the 20% 5 year survival rate of IPF patients without lung transplant, there are still some major problems with this treatment option.

Organ transplantation, in general, is fraught with challenges including scarcity of organs and the need for immunosuppression even with properly matched donor-receiver pairs.

IPF clinically presents as a dry cough and shortness of breath that worsens over months to years. Upon examination, a fine crackling sound known as Velcro rales are present in >80% of

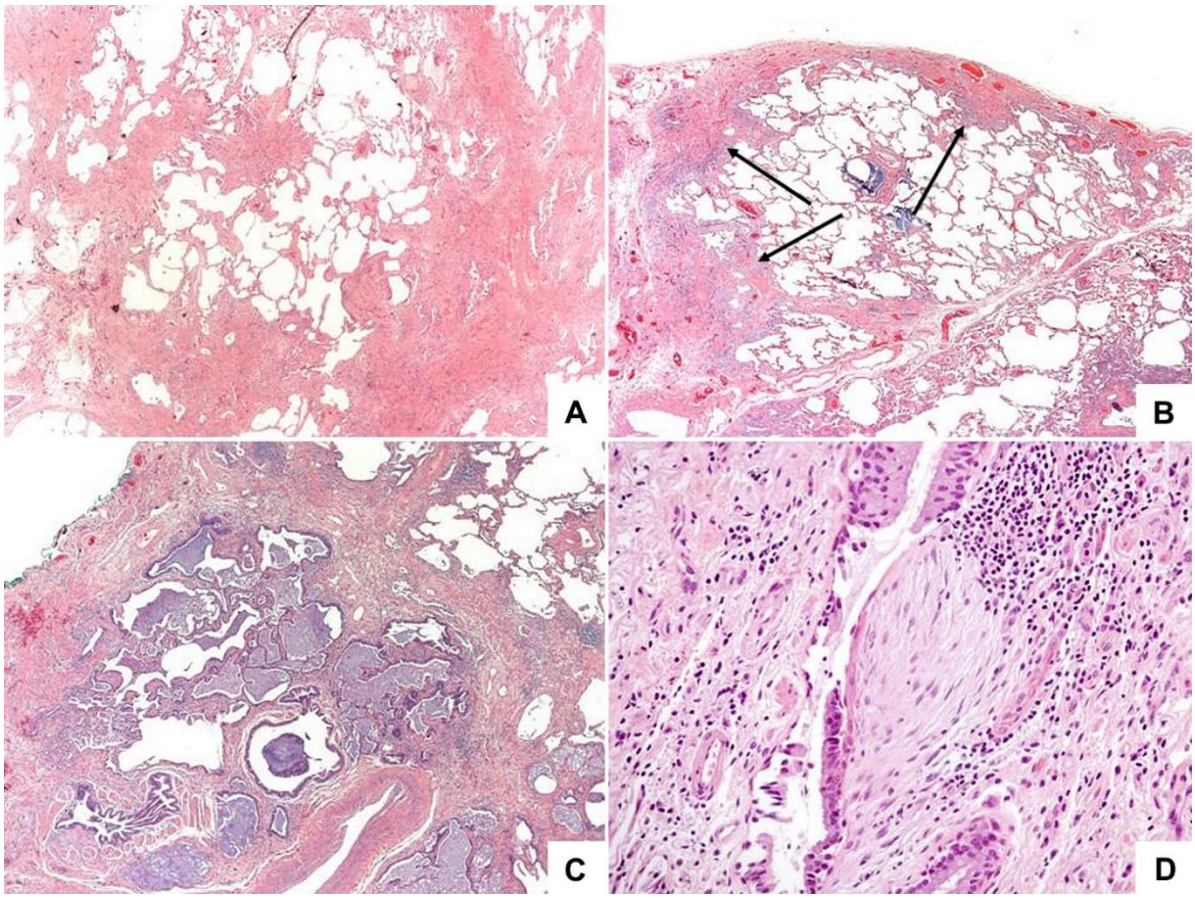


Figure 3: Histology of usual interstitial pneumonia. A) At low magnification the diagnostic key is the abrupt alternating of scarred and normal lung (patchwork pattern: scar-normal-scar-normal). In the scarred areas the alveolar architecture is obliterated (haematoxylin–eosin × 20). B) The fibrosis frequently prevails at the periphery of the lobule in the subpleural–paraseptal regions (arrows), with relative sparing of the centrolobule. This is a useful diagnostic clue, particularly in early cases like here (haematoxylin–eosin × 20). C) Honeycomb consists of enlarged airspaces lined by bronchiolar epithelium, frequently filled by mucus and surrounded by dense scars. Note the architectural distortion and the abrupt transition with residual normal lung seen in the right upper corner (haematoxylin–eosin × 20). D) A fibroblastic focus consisting of a dome-shaped proliferation of myofibroblasts immersed in a myxoid matrix. Fibroblastic foci can be covered by bronchiolar epithelium, as here, or by hyperplasic pneumocytes (haematoxylin–eosin × 100).[10]

cases. Characteristic high resolution cross-sectional tomographic features include a patchy assortment of course reticular and linear opacities, honeycomb cysts, and increased mucosal buildup in the bronchial tree[9]. Histological findings include the destruction of the alveolar membrane, accumulation of scar tissue and fibrotic foci characterized by large numbers of fibroblasts and collagen I, and associated widespread remodeling of the alveolar areas[10] (**Fig. 3**). Importantly, IPF either thickens or destroys the respiratory membrane limiting or completely inhibiting efficient gas exchange. Genetic analysis of IPF lungs has shown significant enrichment in lung development-associated genes and activation of the Wnt/β-catenin pathway[11,12]. There are also a significant number of IPF cases that are familial accounting for between 2-20% of total IPF cases[13]. While the clinical presentation and histology of IPF are well characterized, little is known about the disease etiology. Given the complex nature of this disease there have been a number of hypotheses that seek to explain the disease initiation and progression. One hypothesis proposes that IPF may be a result of the effects of chronic inflammation, namely repeated damage to the lung tissue initiates an aberrant wound healing response that culminates in the presentation if scarring and alveolar destruction characteristic of IPF. Though, clinical measurements of inflammation are not commensurate with what would be necessary to account for the large degree of scar formation nor is anti-inflammatory intervention effective at modulating this process. Another hypothesis proposes that IPF may be the result of many individual, repeated microinjuries. These repeated injuries initiate an epithelial to mesenchymal transition[14] that drives an accelerated wound healing process and account for the patchy nature of the disease. Finally, more recent work done on the analysis of the genetic anomalies in familial IPF cases has found variants in surfactant protein C (*SFTPC*), surfactant protein A2 (*SFTPA2*), a mucin constituent (*MUC5B*), and components of the telomerase complex (*TERT* and *TERC*) to be associated with IPF. Genome-

wide association studies of these genetic loci have shown that they are highly predictive of a pulmonary fibrosis diagnosis. These results may underline the importance of the barrier functions of the lung in maintaining homeostasis[13,15].

Despite the relative rarity of IPF compared with other lung diseases there have been 13 clinical trials for IPF-related therapies in the past 10 years. Yet in spite of the large amount of time and resources invested into these projects few have shown any clinical efficacy. The most significant of these include Phase III trials of Pirfenidone and Nintedanib. These compounds showed signs of slowing the progression of fibrosis though had no effect on the overall survival rate of IPF patients. Despite these shortcomings these compounds have become standard of care in several countries due to the lack of effective therapies.

Chapter 1.3. Pulmonary Tissue Engineering Fundamentals

Chapter 1.3.1. Introduction to Tissue Engineering

Tissue engineering is a growing field of bioengineering that encompasses a wide range of technologies and applications. These include the generation of stem cell-derived organoids, cellularized microfluidic systems, and scaffold-based tissue manipulation and generation. The underlying commonality between these methods is the goal to better approximate *in vivo* conditions and cellular microenvironment. There is a growing body of evidence that cells grown in 3D show genetic expression patterns that are more similar with organs than those grown in 2D[16] and therefore may be better models for studying complex biological processes and providing substrates of transplantation. In addition, these cultures will likely serve as better disease modeling platforms[17] and give key insights into embryology and development[18].

Chapter 1.3.2. Stem Cell-derived Lung Organoids

Stem cell-derived lung organoids are becoming a popular research topic for disease modeling[19], embryology[20], and regenerative medicine[21]. Organoids are multicellular organ mimics that seek to recapitulate the cellular distribution, structure, and function of fully formed organs. In general, these organoids are derived from stem cells or immature pulmonary progenitors in a multistep process. First stem cells are differentiated into definitive endoderm using activin A and other factors. This process is performed concurrently with spheroid formation in a low adhesion plate. The spheroids are then transferred to a Matrigel coated dish and treated with a combination of TGF-β, BMP, and Wnt inhibitors to reach anterior foregut endoderm. Next the cells are treated with a cocktail of Wnt, BMP, FGF, RA, and other factors to differentiate the cells into ventral airway progenitors. Finally, lung organoids are generated by encapsulating the ventral airway progenitor cells in Matrigel and treating with Wnt, FGF, cAMP, and glucocorticoids[21]. This process has shown promising results as the constituent cultures have been observed to undergo expansion, branching morphogenesis, and express many of the developmental markers observed *in vivo*[22,23,24,25,26,27].

While there are a wide variety of protocols that have been developed for producing lung organoids one commonality in all of these previously cited methods is the use of a product known as Matrigel. Matrigel is a gelatinous protein secretion derived from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells[28] and has effectively taken over the organoid-generation field due to its advantageous properties. Matrigel contains a broad variety of ECM proteins, growth factors, proteoglycans, and other extracellular materials that are important for cellular adhesion and 3D culture maintenance. Most importantly, Matrigel provides a microenvironment similar to that encountered *in vivo*. Though, due to the complexity and batch to batch variance of Matrigel, as it

is a cell-derived product and not well defined, there are worries that results derived from Matrigel-based assays may not be completely reproducible. In addition, cells/organoids derived from Matrigel culture may not be suitable for human transplantation due to the fact that Matrigel is animal-derived. Though, despite these potential pitfalls, Matrigel continues to dominate organoid formation methodologies.

Chapter 1.3.3. Scaffold-based Lung Tissues

Scaffold-based approaches to tissue generation represent another way to form the complex structures and cellular distributions observed in tissues and organs *ex vivo*. Typically, these scaffold-based approaches seek to provide an engineered extracellular matrix for cellular adhesion and templating. Organs and tissues function due to the synergistic interactions between many cell types assembled in a specific environment. Cells maintain their geometry and phenotype, in part, due to adhesions and interactions with the ECM, the non-living structural component of tissues. In general, scaffold based methodologies seek to provide cells with a biomimetic ECM that the cells are able to adhere to, spread upon, and migrate into. Therefore, the overall tissue architecture is dictated by the geometry of the scaffold material and the cells' ability to reorganize that material.

Common scaffolding materials for pulmonary tissue engineering include biodegradable polymers, hydrogels, and biologically-derived proteins. These materials may be assembled into foams[29], beads[30,2], or rings[31] depending on the region of lung being emulated. Foam and bead based approaches are common for reproducing a distal lung architecture while rings are used for providing a trachea-like substrate. Often accompanying these methods is the use of a rotational bioreactors[32,33,34,35]. Bioreactors produce a microgravity-like environment by rotating a cylindrical volume of liquid about a central axis. This provides a flow condition wherein small objects will travel within the bioreactor vessel in roughly circular orbits[36] (assuming the objects are sufficiently

small comparted to the vessel and have a density similar to that of water). These bioreactor based methods allow for the uniform coating of scaffold materials with cells and, in certain cases, aggregation and tissue formation of cell coated constructs[2].

Chapter 1.4. Pulmonary Disease Modeling and High Throughput Drug Screening

Chapter 1.4.1. Disease Modeling Methods

The goal of any disease modeling strategy is to produce a copy of the disease that can be experimentally interrogated as experimenting on humans directly is morally objectionable. In order to fill the need for useful disease models there have been a number of methods published for simulating human lung disease *ex vivo*. These methods may be broken up into two general categories: human cell culture and animal models of lung disease. Each of these methods has its own distinct benefits and drawbacks. Traditionally, human cells are grown on flat, tissue culture plastic. These methods benefit from the fact that 2D cell culture has been long established, well characterized, reproducible, high throughput and exceedingly easy to image. Though, cells grown in 2D lack many of the niche specific cues that exist *in vivo* and have limited usefulness in modeling disease unless there is a known molecular target. On the other hand, animal models, especially mice, provide complete biological systems for experimentation and have relevance as there is a high degree of genetic similarity between mammals. Yet there are many examples of both methods failing to provide mechanistic insights into the underlying disease pathophysiology[37].

Chapter 1.4.2. Modeling Idiopathic Pulmonary Fibrosis

The primary methods for studying and modeling IPF may be grouped into human cell culture and animal models. It is common practice to collect primary fibroblasts from IPF patients and culture them in standard, 2D tissue culture conditions. These studies have granted insight into

the differences between pathological and wild type fibroblasts. Specifically, IPF patient fibroblasts are more contractile[38], exhibit aberrant migratory abilities[39] and proliferate in normally inhibitory conditions[40]. On the other hand, the leading animal model of IPF is the bleomycin mouse model. This model has proven to be controversial though as there have been conflicting reports as to its usefulness as a IPF analogue[41,42,43].

Chapter 1.4.3. High Throughput Drug Screening

Research and development productivity has been on the decline in the pharmaceutical industry for a number of years[44]. Large companies are being financially burdened with increasing research costs while suffering from revenue loss due to patent expirations and generic competition. Though, even accounting for these industry-wide challenges, new drugs for lung diseases have lagged even further behind. Many of the new treatments to hit the market are not fundamentally novel; they are simply improvements on older therapies and existing classes of drugs[45]. While there are benefits to optimizing existing drugs and therapy strategies[7] this lack of innovative solutions is indicative of the fact that the methods used for drug discovery have become ineffective in providing novel solutions to lung disease. Though this lack of success is not due to lack of effort as these diseases are complex and heterogenous, characterized by a number of nonlinear, dynamic interactions between environmental forces and genetic traits[46]. The current drug development pipeline is simply not tuned for handling this level of complexity.

The main challenge for pulmonary drug discovery lies at the heart of the methodology used to study lung disease. To enable drug discovery, disease models must recapitulate the underlying human disease pathophysiology such that compounds found to affect the model are effective in a clinical setting. A good example of this process working successfully is in the discovery and development of Ivacaftor, a new drug for treatment of Cystic Fibrosis (CF). CF is characterized

by increased susceptibility to chronic airway infection (particularly with *Pseudomonas aeruginosa*) that will lead to premature decline in airway function and increased mortality[47]. CF is genetically heritable condition wherein the mutation of the cystic fibrosis transmembrane conductance regulator (CFTR) gene is the underlying genetic cause[48]. This mutation and the associated dysfunctional protein then became the target for high throughput drug screening. In a fluorescence-based, cellular screen of ~230,000 small molecules Ivacaftor was identified to potentiate CFTR activity at the cell surface and is now a state-of-the-art treatment for cystic fibrosis. The example of Ivacaftor is instrumental in understanding how a known molecular mechanism may be targeted and ultimately altered to provide a clinical benefit for CF patients. Though many lung diseases do not have known molecular drivers making it unclear how to proceed with drug discovery while knowing very little about the underlying disease mechanisms.

iPSC-based disease modeling has become an exciting avenue of inquiry for studying a wide range of lung diseases and conditions. Though, perhaps the most promising application of this technology is to provide new approaches for studying complex, heterogenous diseases wherein the patient's own genetic background may be crucial in finding a treatment. These kinds of idiopathic diseases are of prime interest because their origin and mechanism(s) of action are not known or poorly understood, greatly limiting both hypothesis-based scientific investigation and clinical treatment options. In other words, the traditional methods of studying disease have failed to provide the insight necessary to discover new drugs and advance the clinical options available to patients afflicted with such diseases.

One of the most powerful methods for discovering new drugs, as in the case of Ivacaftor, is high throughput drug screening (HTS). HTS is enabled by the continued aggregation of large compound libraries, small molecules that have potential to serve as new drugs. There are various

types of libraries, some containing the roughly 6,000 FDA approved drugs, others containing up to 500,000 various compounds. It has been estimated that there are over 10,000 biological targets with a small molecule space that ranges from 10^{14}-10^{30} molecules that may interact with the biological targets. Therefore, the challenge of any HTS is one of combinatorics, to find a method for identifying compounds that positively affect patients suffering from the disease. With such a large variable space the process of matching compounds with targets, especially within the context of disease, is daunting and a brute force approach is not a feasible option. Toward this aim, researchers have developed two distinct methods for performing a HTS: targeted and phenotypic. A targeted drug screen is geared toward effecting a known molecular target that is suspected to be mechanistically important for either causing disease or driving the spread of the disease state. The targeted screen benefits from the *a priori* knowledge of mechanism allowing for the focus of the screen to be centered around a specific molecular interaction greatly reducing the number of possible small molecule/ligand interactions. On the other hand, phenotypic screening takes a more unbiased approach toward drug discovery. Here, cellular models of the disease are used to mimic a known disease phenotype with the hope that the molecular mechanisms that govern the disease in a dish are homologous to the disease *in vivo*. These models are then administered experimental compounds and the phenotype observed and monitored. Compounds that cause the cellular model to revert to a healthy phenotype become candidates for further study. While the lack of a known molecular target may seem to detract from the relevance of such screens the phenotypic methodology may be the only option toward discovering drugs for such diseases wherein the phenotype is well characterized but the mechanism is unknown, i.e. idiopathic diseases.

The recent resurgence of phenotypic screening has highlighted the fact that targeted screens have some severe limitations. For example, the molecular target may not be known or the

suspected target may not be mechanistically relevant on a clinical level. In other words, the drug may positively affect the cellular model while failing subsequent clinical trials. While phenotypic screening does not assume a known molecular target, there are still significant challenges for successful implementation of this approach. Namely, the disease in a dish implemented in the phenotypic screen must adequately phenocopy the disease state *in vivo*. This game of imitation may necessitate the use of 3D organoids and cellular constructs in order to provide a disease model whose phenotypic similarity to the disease state is founded on the same molecular mechanism as its *in vivo* counterpart.

Chapter 1.5. Machine Learning Methods for Phenotypic Classification

Chapter 1.5.1. Artificial Intelligence and Big Data

Artificial intelligence (AI) is a general catchphrase that represents a wide variety of machine learning methods and processes. Recently, AI has become the state of the art in many information industries due to its ability to learn exceedingly complex tasks. These include voice recognition, image recognition, written character recognition, and stock market prediction. These networks have also been successful in learning and outcompeting human champions at the board game Go[49], a major accomplishment given that the lower bound on possible number of Go games is ~$10^{10^{48}}$. The success of artificial neural networks is due to two factors. Fist, these networks mimic the structure of the animal brain insofar as they contain many nodes whose vast interconnections enable a wide range of configurations. Second, these networks are trained using a technique call backpropagation[50] that enables quick learning of complex tasks. Though, in order to unlock the potential of this machine learning architecture large amounts of curated data must be assembled and curated.

Chapter 1.5.2. Convolutional Neural Networks for Image Classification

Artificial neural networks are universal approximators[51] and if used correctly may learn a wide range of complex tasks. Given the specific task of image recognition, a subclass of artificial neural network, the convolutional neural network has become the most popular approach. Convolutional neural networks exploit the convolution of a filter on sequential subregions of an image to achieve pattern matching. When the features of the filter overlap with the features in the image then the convolution will produce a large number and, therefore, a significant activation for future propagation through the network (**Fig. 4**). These convolutional layers each act on the image transforming it for further convolution. The sequential layering of the convolution process has enabled robust image classification.

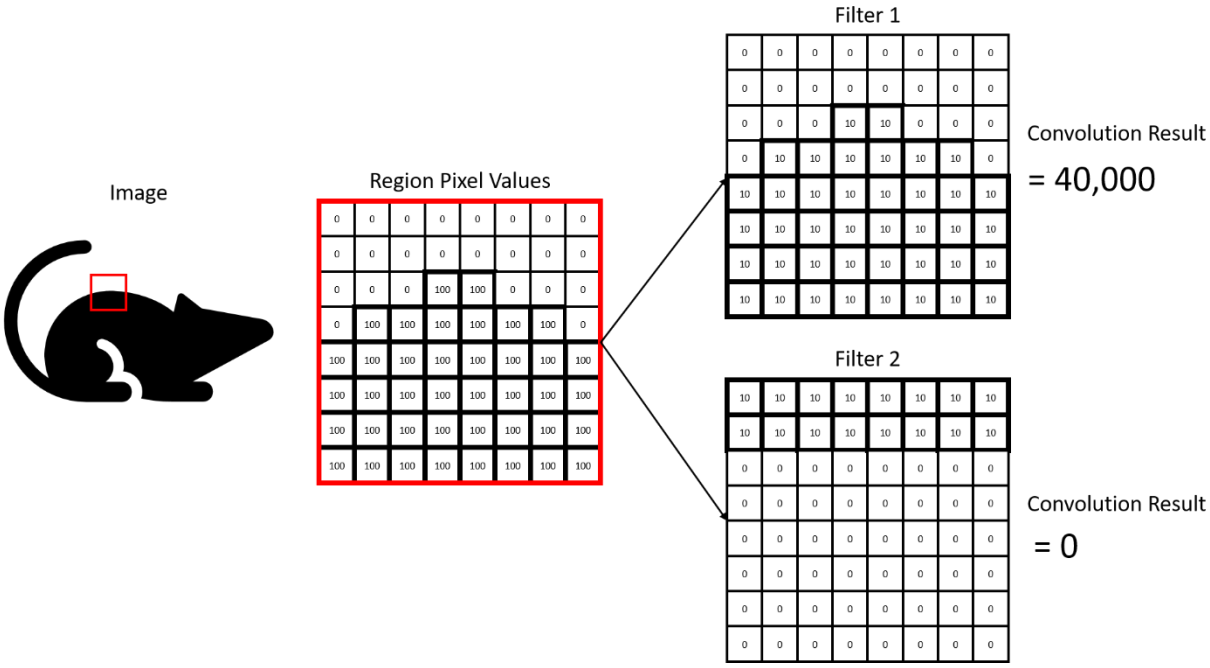


Figure 4: Convolutional Neural Networks compare image subregions with filters to allow for pattern matching. This process is known as convolution. When the filter features match with the image subregion then there is a large activation as observed with Filter 1. On the other hand, Filter 2 has a line feature that does not line up with the image subregion and therefore the result is 0.

Image classification is the process of identifying unknown images after a regimen of supervised learning. The best in class image classifiers include Google's Inception framework and Microsoft's ResNet[52]. The general architecture of these networks is inspired by the visual cortex whose underlying information functioning system was first described in the late 1950s[53] allowing for computer vision that may someday rival human capabilities.

Chapter 2. Lung Organoid Formation and Characterization

Chapter 2.1. Introduction and Motivation

One of the most promising avenues of research in the field of personalized medicine is the development of organoid formation techniques for pulmonary disease modeling. While the term 'organoid' describes a wide range of multicellular tissues and bio-engineered cellular composites, it is typically a small, stem cell-derived, 3D piece of human tissue that mimics aspects of an organ's anatomy and physiology. These properties make organoids a valuable tool for disease modeling methodologies as they may be used to produce a simulacrum of the patient's own diseased organ *ex vivo*. This is necessary as current disease modeling strategies often fail to provide mechanistic, actionable solutions for a range of idiopathic and degenerative diseases. Organ-level biological function is complex. Lung function is a synergy that evolves from the coordinated interactions between many millions of cells representing dozens of cell types all acting and responding to local cues. These cues include interactions with surrounding cells, soluble molecules and the properties of the extracellular matrix in a given cell's immediate vicinity. These various factors are known as the cell's microenvironment and act as key driving factors that govern cellular response, phenotype, and ultimately sum to result in pulmonary ventilation, gas exchange, and mucocilliary debris clearance. In other words, organ-level biological function cannot be truly understood without a detailed knowledge of the individual interactions and feedback loops present at the cellular and intercellular level. Current disease modeling strategies are centered around systems that lack the microenvironmental factors that are likely necessary to replicate the diseased-state phenotype and therefore lack the potential for the discovery of actionable mechanisms. Organoids, on the other hand, are tailor made for this application since organoids produce organ mimics

wherein multiple cell types exist in microenvironmental niches that are similar to those experienced in an organ *in vivo*.

Several methods for the generation of engineered lung have been developed previously. There are, in general, two approaches for accomplishing an engineered lung, mainly differing in the way the method seeks to develop cell phenotype and 3D structure. On one hand, scaffold-centered methods focus on seeding pulmonary cells onto a 3D structure that is chemically and structurally similar to that of native lung extracellular matrix (ECM). Scaffold materials range from decellularized whole-lung[54,55] to biodegradable foams[56,57], gels[58], and beads[59]. So far, the most developed of these methods is the decellularization and reseeding of whole lung. While decellularized lung may offer the most anatomically correct scaffold it is hampered by a shortage of donated lungs and scaffold immunogenicity. In addition, this method is not amenable to high throughput applications thereby hindering drug screening efforts. On the other hand, 3D structure may be developed by inducing developmental programs in immature pulmonary cell cultures. This is accomplished by a combination of 3D cell culture, growth factor supplementation and co-culture techniques[26,60,22,61,62,63,64], which results in 3D branching morphogenesis characteristic of early lung development.

In order to generate a scalable 3D system for pulmonary disease modeling we focused on mimicking the architecture of the distal lung. The lungs, in conjunction with the circulatory system, are primarily responsible for maintaining appropriate levels of oxygen and carbon dioxide in the blood. These gasses enter and leave the blood by passively diffusing across the respiratory membrane of the distal lung. To meet metabolic need, the lung provides a large surface area (50-75m^2) and thin respiratory membrane (~0.5μm) for adequate gas diffusion. This large surface area is packaged into tight clusters of alveolar sacs occurring at the terminal end of each branch of the

bronchial tree. As many pulmonary diseases directly affect the alveolar sacs, either by destruction and/or thickening of the respiratory membrane, we focused the design of the lung organoid to directly mimic the anatomy of these sacs. Alveolar sacs are hollow, roughly spherical (diameter ~ $150\mu m$) chambers whose geometry is reminiscent of the inverse opal. These sacs are lined with primarily Type 1 epithelial cells and a minority of Type 2 cells that act to separate the capillaries from the alveolar void space with fibroblasts occupying the interstitial spaces. These three cell types, epithelial, endothelial and mesenchymal cells account for nearly all the cell types in the lung and are roughly equal in number throughout the organ. Additionally, the lung contains a small number of over 3 dozen more supporting cell types. Inspired by the biology of the lung the lung organoid was designed to mimic the geometry of the alveolar sacs of the distal lung by organizing human pulmonary cells into the interstitial spaces between closely packed, functionalized alginate beads.

The lung organoid generation process is directly inspired by the approach used to produce inverse opals. Inverse opal preparation typically occurs in a 3-step process. First, sacrificial beads are sedimented into a close-packed geometry. Next, the inverse opal material is introduced into the interstitial spaces between the sacrificial beads and, depending on the material, sintered, polymerized, or crosslinked into place. Finally, the sacrificial beads are removed leaving a free-standing, porous, interconnected network. This general workflow was adapted to accommodate human pulmonary cells as the inverse opal material and modified alginate microbeads as the sacrificial beads. Though there are several key differences, most notably the cells actively participate in organoid formation creating a close-packed network instead of simply being inserted into previously sedimented beads. Also, to maintain the organoid structure, the alginate beads are never removed during organoid formation and subsequent disease modeling.

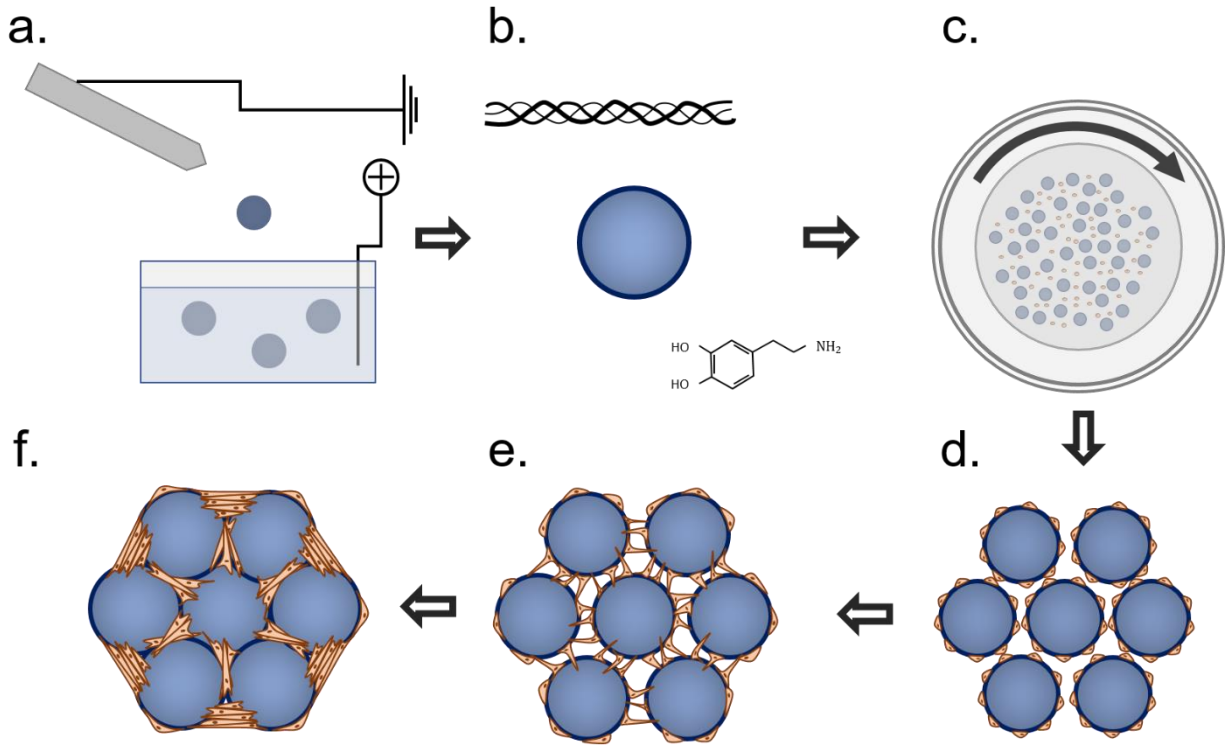


Figure 5: Organoid formation graphic. (**a**) Alginate beads are generated using an electrostatic droplet generator. (**b**) Alginate beads are functionalized for cellular adhesion by performing a two-step process of collagen I precipitation and dopamine polymerization (**c**) Cells and beads are combined in a high aspect ratio bioreactor vessel and rotated. This provides the flow condition necessary to achieve a uniform coating of cells onto the alginate beads. (**d**) Cell coated beads are compacted together in the bottom of a standard tissue culture well plate. (**e**) Fibroblasts that are included in the seeding process for bead-bead bridges and establish a tension that acts to pull the beads together into a single unit. (**f**) A mature, fully formed organoid.

The organoid generation process can be summarized in four steps **(Fig 5)**. First, alginate beads are generated **(Fig. 5a)** and functionalized **(Fig. 5b)** to allow for cell adhesion. Next, beads and a combination of epithelial, endothelial, and mesenchymal cells are combined in a high aspect ratio rotational (HARV) bioreactor and rotated at a speed conducive for uniform cellular adhesion to the bead surface **(Fig. 5c)**. Although several different cell types were included into the organoid seeding process, it was determined that fibroblasts were critical for organoid formation and must be included in the seeding process. Specifically, at least 1/3 of the cells must be fibroblasts for

organoid condensation to occur. Finally, the beads are compacted allowing for organoid condensation **(Fig. 5d-f)**. Bead compaction and organoid formation may occur either inside or outside the bioreactor chamber allowing for easy scalability in organoid size or in number of organoid generated. Large organoid formation is achieved by simply leaving the cell-coated beads inside the bioreactor vessel and rotating at a low rotational velocity. If a large number of organoids are desired, then the cell-coated beads are removed from the bioreactor vessel and partitioned into the wells of a 96 or 384-well tissue culture plate. The plate is then centrifuged, sedimenting and compacting the beads together. Organoid condensation is directly driven by fibroblasts, the mesenchymal cell type primarily responsible for wound closure and scar formation. In summary, we have designed and executed a strategy for generation of engineered lung-like tissues known as lung organoids. Furthermore, we have identified the key mechanisms necessary for organoid formation and have incorporated a wide range of pulmonary cells into the tissue generation process.

Chapter 2.2. Experimental Methods

Chapter 2.2.1. Alginate Bead Synthesis by Electrostatic Droplet Generation

Alginate beads were generated using an electrostatic droplet generator (custom) operated on a 3% medium viscosity alginate solution (Sigma-Aldrich) at 9000V over a bath of $100mM$ $BaCl_2$ (Sigma-Aldrich) solution. Bead size distribution was determined using a custom built Matlab algorithm. White light images of the beads under 5x magnification were displayed and the user defined each bead diameter by clicking on opposite bead edges. A total of 359 beads were imaged resulting in a size distribution of $(161 \pm 80\mu m)$. The beads (2.5mL sedimented) were rinsed and allowed to soak in 1mL of high concentration (9.37 mg/mL) rat tail collagen I solution (Corning) for 6 days at 4℃. After soaking, 2.5mL of beads were pipetted into a 35mm petri dish

(Corning), the excess collagen I aspirated, and 8mL of 2mg/mL dopamine hydrochloride in 50mM Tris buffer (Sigma-Aldrich), pH8.5 was added. The dish was sealed with parafilm (Sigma-Aldrich), and rotated at 16.5 rpm on a lab rotisserie (Labquake) for 1 hour at room temperature. Beads were then rinsed in the above mentioned Tris buffer and then soaked in experimentally-relevant, serum-free media.

Chapter 2.2.2. Alginate Bead Functionalization by Collagen I/poly(Dopamine)

Alginate beads require functionalization to allow for cellular adhesion. The beads (2.5mL sedimented) were rinsed and allowed to soak in 1mL of high concentration (9.37 mg/mL) rat tail collagen I solution (Corning) for 6 days at 4℃. After soaking, 2.5mL of beads were pipetted into a 35mm petri dish (Corning), the excess collagen I aspirated, and 8mL of 2mg/mL dopamine hydrochloride in 50mM Tris buffer (Sigma-Aldrich), pH8.5 was added. The dish was sealed with parafilm (Sigma-Aldrich), and rotated at 16.5 rpm on a lab rotisserie (Labquake) for 1 hour at room temperature. Beads were then rinsed in the above mentioned Tris buffer and then soaked in experimentally-relevant, serum-free media.

Chapter 2.2.3. Hydroxyproline Assay to Determine Collagen I Content

Hydroxyproline content was used to quantify the amount of collagen on alginate beads. It was measured colorimetrically by a method described previously[65] with modification. On day 0 sample aliquots of 100μL of alginate beads were combined with 40μL of 9.37 mg/mL rat tail collagen I solution in a microcentrifuge tube and stored at 4℃. A single sample aliquot was removed daily, over the course of 6 days, and the excess collagen I solution separated from the beads by pipetting and stored in a separate tube. At the end of the experiment, the beads and excess collagen were resuspended in 50μL of ddH_2O, after which 100μL of 12N HCl was added and the sample was hydrolyzed for 24 hours at 110°C. 10μL of samples were transferred to a 96-

well plate in triplicate along with a hydroxyproline standard and evaporated to dryness under vacuum. Samples and standards were rehydrated in 10µL of ddH_2O, then 20µL of isopropanol was added. 10µL of oxidation buffer (one part 7% chloramine T (Sigma-Aldrich) and four parts acetate citrate buffer (pH 6.0 per 100mL: 5.7g sodium acetate, 3.75g tri-sodium citrate, and 0.55g citric acid in 35.5mL of isopropyl alcohol and distilled water)) were added to the samples and incubated at room temperature for 4 minutes. Then 130µL of analytical isopropanol solution (3 parts Ehrlich's reagent (2g para-dimethylaminobenzaldehyde in 3mL of 60% (v/v) perchloric acid) to 13 parts isopropanol) was added to each well and placed at 60°C for 25 minutes. Samples were then cooled to room temperature for 5 minutes after which 100µL isopropyl alcohol was added. Absorbance was measured at 558nm using a spectrophotometer.

Chapter 2.2.4. Human Fetal Lung Cell Isolation and Cell Culture

Human fetal lung fibroblasts (FLFs) were isolated from 18-20 week old fetal lungs (Advanced Bioscience Resources, Inc.). Tissues were finely minced and dissociated using 1 mg/mL collagenase/dispase (Roche) and 0.1 mg/mL DNase (Sigma-Aldrich) with rotation for 45 minutes at 37°C. After washing using media containing 1% fetal bovine serum (FBS), a single cell suspension was generated using 100 and 40µm cell strainers. To remove red blood cells, the suspension was incubated in RBC lysis buffer (BD Pharmingen) for 15 minutes at room temperature. Cells were then plated in 6-well tissue culture plates and cultured in DMEM/F12 medium containing 10% FBS. Human Umbilical Vein Endothelial Cells (HUVECS) and Small Airway Epithelial Cells (SAECS) were maintained according to the manufacturer's recommendations (Lonza) in EGM-2 medium and SAGM medium, respectively.

Chapter 2.2.5. Bioreactor Loading and Mesenchymal Organoid Formation- HARV Bioreactor

1mL of functionalized alginate beads and 4 million FLFs in media were added to the 4mL HARV bioreactor vessel (Synthecon) using the built-in syringe/valve system. The vessel was screwed into the bioreactor base and the beads allowed to settle, without rotation for 10 minutes. After sedimentation, the bioreactor was powered on to 4rpm. Organoids were allowed to form and mature over the course of 2 days.

Chapter 2.2.6. Timelapse Imaging and Analysis of Organoid Formation

Organoid timelapse imaging was accomplished by mounting a GoPro Hero 3 camera onto the 4mL HARV bioreactor. The mount was custom built using polycaprolactone and included a macro lens and two white light emitting diodes. A high-capacity battery allowed for extended data acquisition. All timelapse experiments had the same initial seeding conditions of 1mL functionalized beads and 4 million FLFs. The GoPro camera was set to take images at a frequency of 2 Hz. Controlling bead flow in the HARV bioreactor was accomplished by changing the rotational speed of the vessel. In order to quantify bead flow characteristics, timelapse imaging and video were implemented. A GoPro camera and a custom-built mount were fixed onto the 4mL HARV and different rotational speeds were recorded. At a bead volume:vessel volume ratio of (1:4), the rotational speed was varied from 4rpm to 30rpm. Bead trajectories were broken down into two distinct flow regimes, mildly correlated flow and highly correlated flow. Highly correlated flow was dominant at low rpm wherein beads traveled as a group, maximizing bead-bead interaction. Mildly correlated flow occurred at higher rpm wherein beads orbited in epi-circular paths forming transient clumps. The bead flow transitions from highly correlated to mildly correlated flow around 12rpm. Organoid formation was monitored under the highly correlated

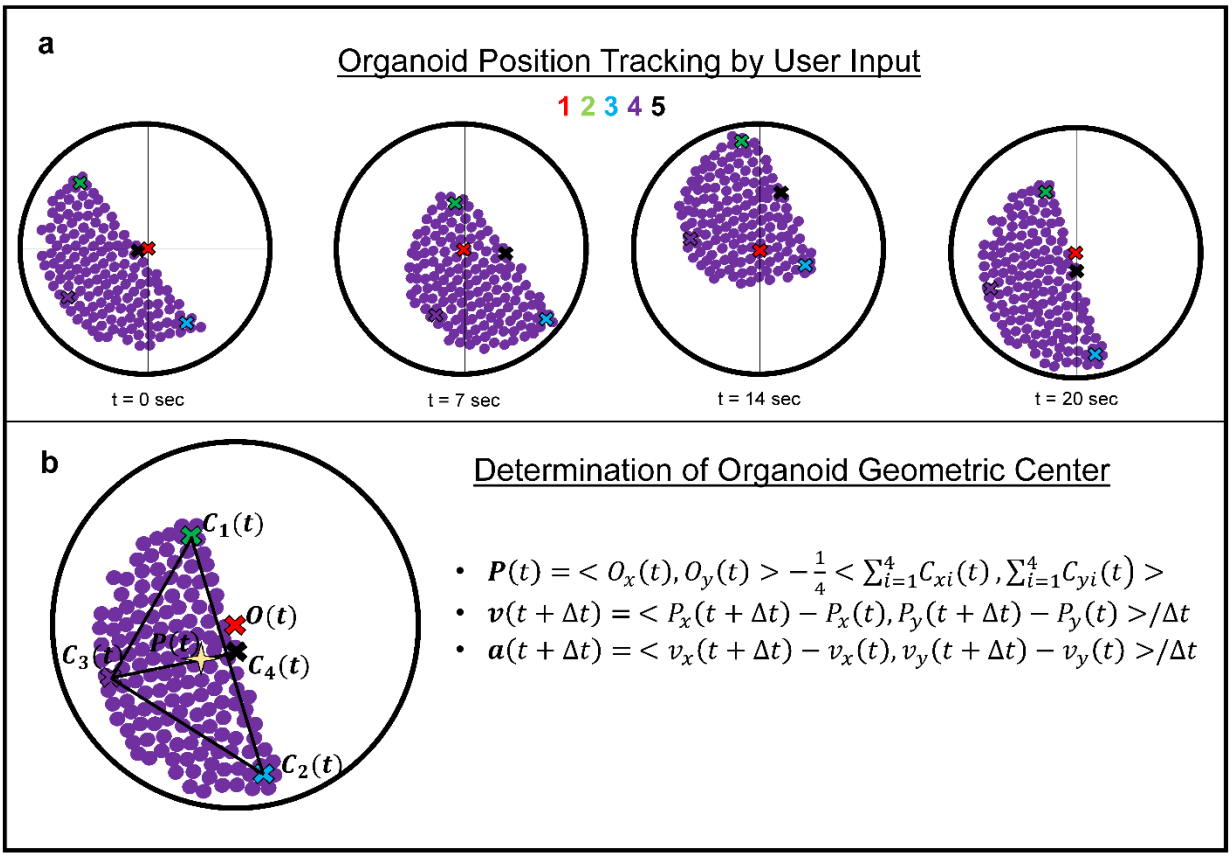


Figure 6: Graphic explanation of the method for organoid tracking. (**a**) Images acquired at a frequency of 2Hz displayed using a custom Matlab program. The five positions marked per image indicate the organoid edges and the center of the bioreactor vessel. This process was repeated for 100 images (50 seconds) for each data set. (**b**) Organoid geometric center was determined by averaging the position data and subtracting that average from the vessel center.

flow regime using timelapse imaging analysis. In order to characterize the kinetics of organoid formation, images taken at different time points were analyzed to determine the organoid geometric center as a function of time in addition to measuring organoid 'height' and 'width'. Analysis of timelapse data allowed for the tracking of organoid dimensions and quantification of organoid condensation. Analysis was accomplished by identifying organoid corners and edges in each frame, averaging these points, and referencing that position to the center of the bioreactor vessel (**Fig. 6a**). By overlaying the velocity vectors onto an image of the organoid from that time set it was possible to plot out the bulk organoid trajectory at different time points. During organoid

formation, the initial bolus of beads began to condense and stiffen. Velocity fluctuations in the organoid increased with time in culture as the bead-bead bridges contracted. We performed another finite difference derivative to allow for the derivation of the organoid acceleration (**Fig. 6b**). If organoid acceleration is coupled with mass, the sum total of forces acting upon the organoid may be arrived at. The main contributor to the organoid mass comes from the alginate beads, cells contribute very little to overall mass. Organoid stiffness was determined by dividing the displacement in organoid 'height' into the magnitude of the force necessary to generate that displacement. Stiffness increased due the fibroblast contraction of the bead-bead cellular bridges formed as a consequence of organoid compression bringing beads into direct contact with each other.

Chapter 2.2.7. 96-well Plate Lung Organoid Formation

1mL of functionalized alginate beads and 4 million cells in media were added to the 4mL HARV bioreactor vessel (Synthecon) using the built-in syringe/valve system. Mesenchmyal organoids only contained fibroblasts whereas multicellular organoids contained a 1:1:1 ratio (1.5×10^5 cells each) of small airway epithelial cells (SAECs)(Lonza), fetal lung fibroblasts, and HUVECs(Lonza). The vessel was screwed into the bioreactor base and rotated at 16.5rpm for 1 hour to allow for cellular adhesion to the bead surface. Cell-coated beads were removed from the HARV bioreactor and aliquots of 100μL were partitioned into the wells of a 96-well plate. The plate was then centrifuged at 1000g for 5 minutes to further sediment and pack the beads. 150μL of media was then added and changed daily. Organoid formation occurred over the course of 2 days.

Chapter 2.2.8. Immunofluorescence Staining

For whole-mount staining, organoids were fixed using 4% paraformaldehyde (Fisher) in TBS for 1 hour at room temperature and permeabilized using 0.1% Triton X-100 (Sigma-Aldrich) in TBS for 30 minutes. After blocking in 10% normal goat serum (Life Technologies) for 1 hour, organoids were incubated with primary antibodies for 24 hours at 4°C. After washing, organoids were incubated in secondary antibodies (Life Technologies) for 2 hours prior to the addition of DAPI. For immunofluorescence staining of organoid and lung sections, fixed samples were mounted in Histogel (Thermo Scientific), embedded in paraffin, and sectioned to 4μm. After deparaffinization and rehydration, antigen retrieval was performed using 1 mM EDTA in a pressure cooker for 10 minutes. After cooling, slides were permeabilized using 0.2% Triton-X 100 in PBS, washed in 0.1% Tween-20 (Sigma-Aldrich) in TBS and blocked with Protein Block (Dako) for 1 hour. After washing, sections were incubated in secondary antibodies and DAPI for 1 hour at room temperature, and mounted in Vectashield (Vector Laboratories). The following primary antibodies were used: rabbit anti-vimentin (Bioss), mouse anti-αSMA (Sigma), mouse anti-CD31 (Dako), rabbit anti-pro-SPB and pro-SPC (Seven Hills), mouse anti-Pro-collagen Type I (Developmental Studies Hybridoma Bank), rabbit anti-T1a (Abcam) and rabbit anti-cytokeratin (wide-spectrum, Abcam). Confocal imaging was performed using a Zeiss LSM 700. Human adult lung tissues were obtained from healthy donors and procured under Institutional Review Board–approved protocols at UCLA.

Chapter 2.2.9. Inhibition of Contraction in Organoid formation using Blebbistatin

Mesenchymal organoids were prepared in the 96-well plate as previously mentioned and allowed to mature over 3 days (with daily media changes). On day 4, organoids were given

media supplemented with 25μM, 5μM blebbistatin, or DMSO control. Supplemented media was changed daily over the next 9 days. Organoids were imaged daily.

Chapter 2.3 Results

Chapter 2.3.1. Scaffold Generation and Characterization

Alginate beads **(Fig. 7a)** were chosen as the alveolar template due to their advantageous material properties including biocompatibility and ionotropic crosslinking[66,67]. Alginate beads were generated using an electrostatic droplet generator operated on a 3% medium viscosity alginate solution at 9000V over a bath of $100mM\ BaCl_2$ solution. Bead size was determined using a custom built Matlab algorithm, resulting in a size distribution of $(161 \pm 80\mu m)$.

Native alginate does not offer an adhesive surface for cellular adhesion and thus requires additional processing to grant this functionality. A mussel-inspired, collagen I adlayer was then applied to allow for cellular adhesion[68]. This process exploits the self-polymerization of dopamine in the presence of ECM protein to create a stable, uniform, cell-adhesive coating[69]. Specifically, a technique was developed for coating alginate beads in a collagen I/poly(dopamine) adlayer to support cell adhesion and survival **(Fig. 7b)**[70]. This was accomplished by first soaking this beads in a high concentration collagen I solution allowing for the precipitation of collagen I onto the

bead surface (**Fig. 7b i**)[65], after which the beads were added to a freshly prepared dopamine solution to polymerize and anchor the collagen into place (**Fig. 7b ii**).

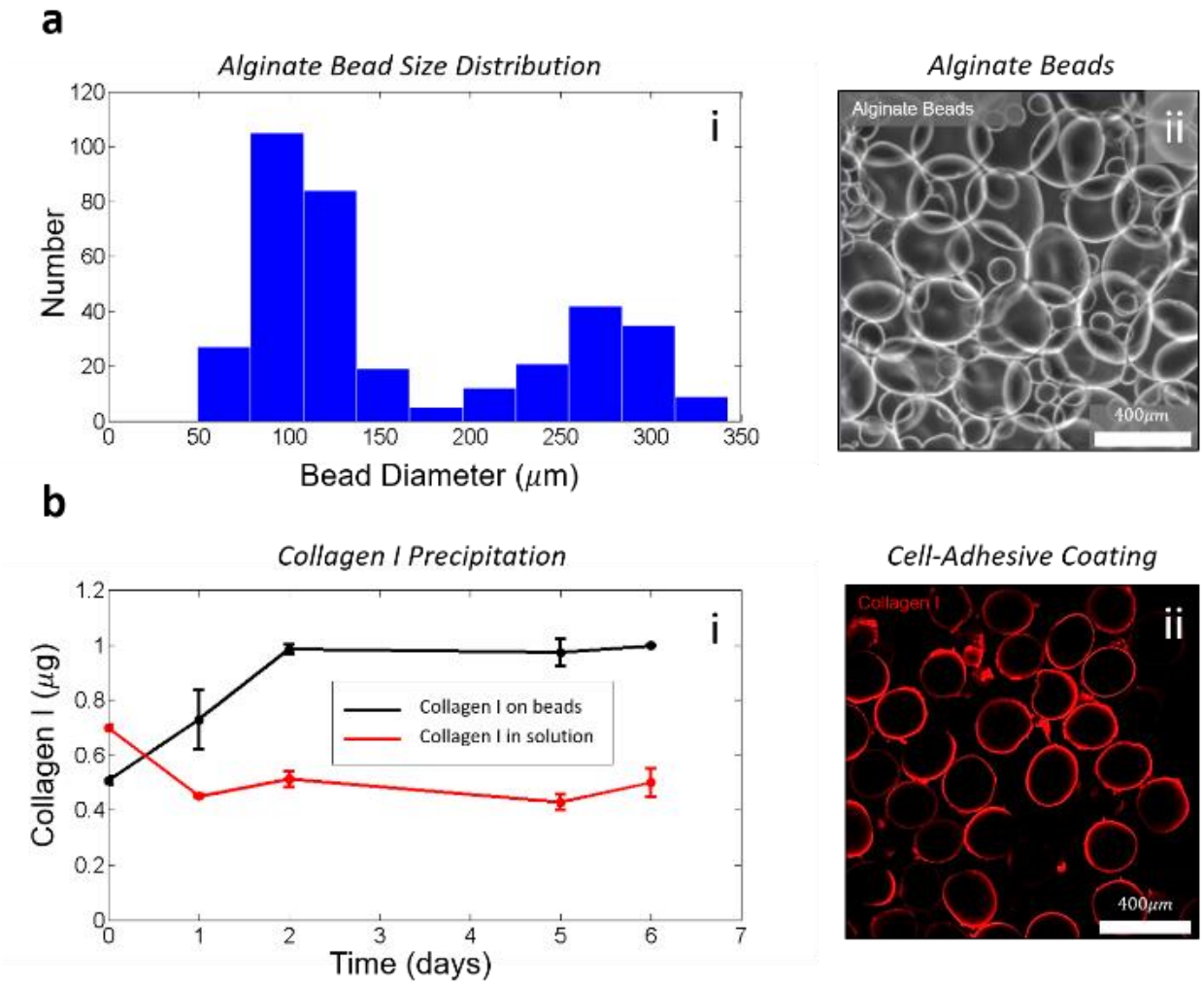


Figure 7: Bead size distribution and collagen I deposition characterization. (**a i**) Bead size distribution was quantified by image analysis of white light micrographs using a custom Matlab algorithm. (**a ii**) White light micrograph of alginate beads. (Scale bar = 400μm). (**b i**) Collagen I precipitation during bead incubation was quantified by the colorometric hydroxyproline assay. (**b ii**) Collagen I immunofluorescence showing conformal coating of collagen I on the bead surface.

Chapter 2.3.2 Coating Functionalized Alginate Beads with Cells

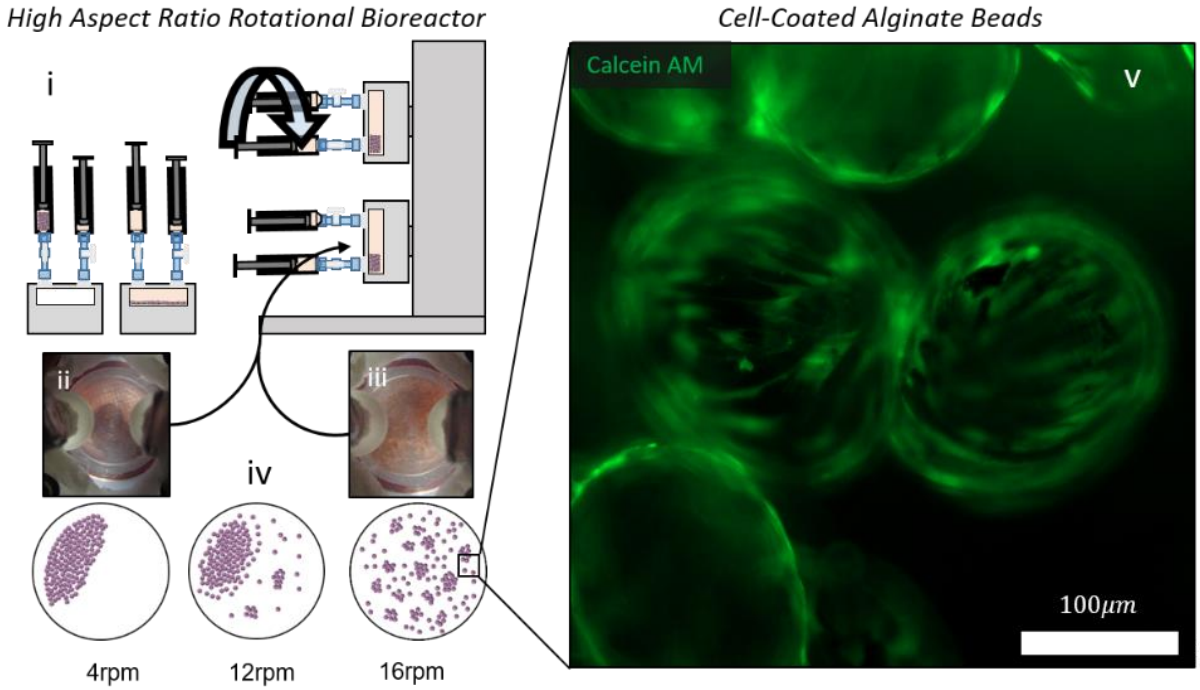


Figure 8: High-aspect-ratio-vessel-bioreactor (HARV) flow regimes and resulting bead coating. (**i**) Loading and function of HARV bioreactor. 1mL of sedimented, functionalized alginate beads were loaded into a 4mL vessel. 4 million fibroblasts were seeded into the vessel. The vessel was attached to the rotary base and rotation initiated. (**ii**) Timelapse image of beads moving together in the 4mL HARV bioreactor as a single unit at 4rpm. (**iii**) Image of beads moving independently in the 4mL HARV bioreactor at 16rpm. (**iv**) Graphical summary of bead flow patterns over several rpm values. (**ii**) Fluorescence micrograph of calcein AM (viability dye) showing labeled fetal lung fibroblasts (FLFs) evenly coating functionalized beads. (Scale bar = 100μm).

Alginate bead cell coating was achieved using the HARV bioreactor system (**Fig 8i**). The HARV is a truncated cylindrical vessel that is rotated along the central axis in order to achieve simulated microgravity for small particles present in the chamber. At a sufficiently high rotational speed particles in the HARV bioreactor will remain suspended in the liquid, rather than settling to the bottom of the well, due to the frictional forces imparted on the particles by the rotating liquid. This effect may be achieved as long as the particles are much smaller than the dimensions of the vessel and have a density comparable to the liquid present in the vessel. In addition, the vessel

must be fully filled with liquid to minimize turbulence caused by the presence of bubbles near the top of the vessel. Since both alginate beads and human cells have a density comparable to that of water. this allows for a stable suspension of these components to be achieved upon rotation in the HARV system.

Specifically, alginate beads and cells are combined in the HARV vessel and rotated at 16.5 rpm. The beads are added to the bioreactor at a volumetric ratio of (1:4) (sedimented bead volume: bioreactor volume). Cells are added at a ratio 1 million cells for every 0.25mL of beads present. Uniform coating of the alginate beads is achieved by adding beads and cells to the HARV vessel at previously specified ratios and rotating the vessel at 16.5 rpm for 1 hour. At 16.5rpm (**Fig. 8iii,iv**) the bead flow is mildly correlated wherein beads move in roughly circular orbits about the central axis of rotation. This rotational speed provides a semi-laminar flow condition allowing for interaction times between cells and beads that are long enough to ensure cellular adhesion to the bead surface (**Fig. 8v**).

Further characterization of the bead flow dynamics was performed by mounting a GoPro camera to the 4mL bioreactor vessel and imaging the chamber over time (**Fig. 8ii,iii**). All timelapse experiments had the same initial seeding conditions of 1mL functionalized beads and 4 million fetal lung fibroblasts (FLFs). The GoPro camera was set to take images at a frequency of 2 Hz. Controlling bead flow in the HARV bioreactor was accomplished by changing the rotational speed of the vessel. At a bead volume:vessel volume ratio of (1:4), the rotational speed was varied from 4rpm to 30rpm. Bead trajectories were broken down into two distinct flow regimes, mildly correlated flow and highly correlated flow. Highly correlated flow was dominant at low rpm wherein beads traveled as a group, maximizing bead-bead interaction. Mildly correlated flow

occurred at higher rpm wherein beads orbited in epi-circular paths forming transient clumps. The bead flow transitions from highly correlated to mildly correlated flow around 12rpm (**Fig. 8iv**).

Chapter 2.3.3. Large Organoid Formation in the HARV Bioreactor

Large organoid formation is possible if the cell coated beads are allowed to remain the HARV vessel and rotate under a highly-correlated flow condition at 4rpm. Specifically, 1mL of functionalized alginate beads and 4 million fibroblasts were added to the 4mL bioreactor vessel

and rotated at 4rpm. Timelapse imaging was performed using the previously mentioned GoPro camera.

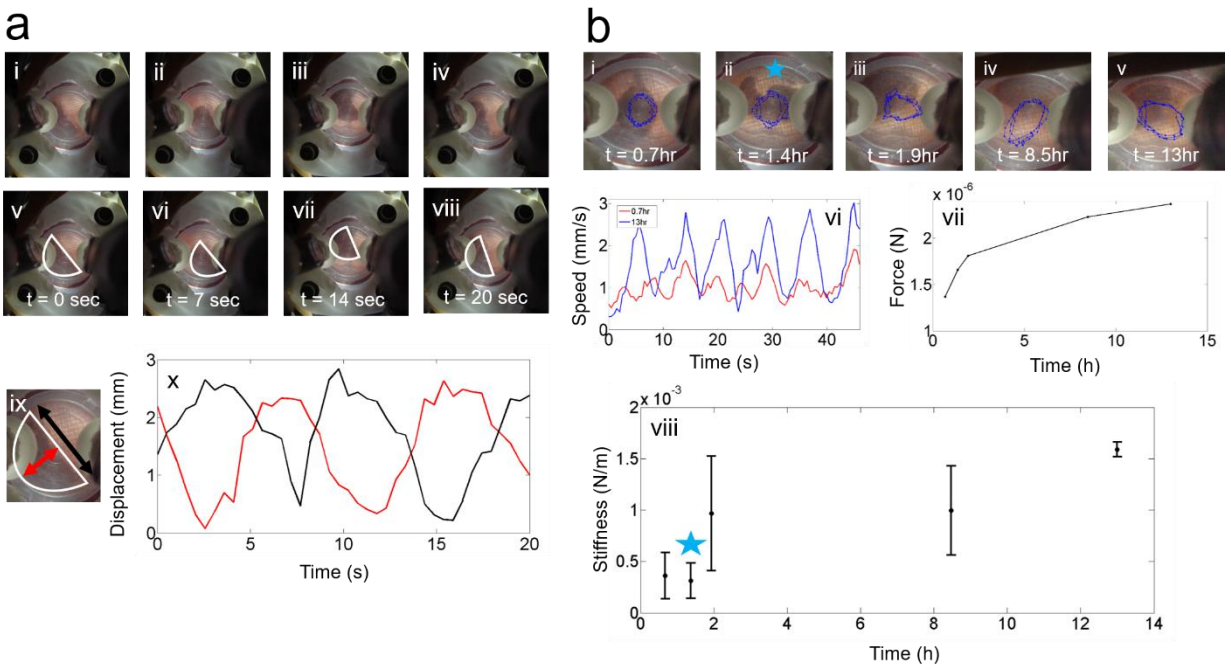


Figure 9: Characterization of the mechanism of large organoid formation. (**a i-iv**) Representative images taken of organoid formation after 0.7 hours. (**a v-viii**) Organoid position over time is highlighted indicating cyclic deformation with a period of transit of 0.05Hz (**a ix**) Red and black arrows indicate user-tracked dimensions of organoid. (**a x**) Measured strain vs. time plot of indicated organoid dimensions at the 0.7 hour time point (colors coordinate to dimensions specified in a ix). This process increases bead-bead interactions aiding organoid formation. (**b i-v**) Images of organoids at various time points during organoid formation. Superimposed is a blue track comprised of velocity vectors arrived at during the tracking process. At the 1.5 hour time point the organoid developed a defect, artificially increasing the measured strain for that time sequence. This rip was repaired shortly after, indicating the active role fibroblasts play in organoid formation. (**b vi**) Plot of organoid speed over 50 seconds at two different time series. (**b vii**) Plot of observed force applied to organoid over the 13 hour period. This increase in force is due to increased organoid elasticity; as the organoid stiffens less energy is dissipated by bead-bead friction and the organoid speeds up. (**b viii**) Organoid stiffness vs time plot.

In order to characterize the formation of HARV bioreactor-generated mesenchymal organoids we sought to image and quantify the bead flow characteristics and organoid formation kinetics (**Fig. 9**). Video imaging was used to identify flow regimes that maximized bead-bead

interactions. Timelapse imaging and subsequent image analysis was performed under the identified flow regime in order to identify organoid position and geometry (**Fig. 9a ix**) over representative 50 second periods of the total 13 hours of recorded organoid formation. The resulting analysis allowed for the characterization of cyclic organoid deformation (**Fig. 9a x**). In addition, it was possible to plot out organoid trajectory and resulting speed over time (**Fig. 9b i-v**). As the organoid matures, its average speed through the vessel increases by nearly two-fold (**Fig. 9b vi**). These measurements were then used to compute the average force experienced by the organoid (**Fig. 9b vii**) and, finally, the increase in organoid stiffness over time (**Fig. 9b viii**).

Chapter 2.3.4. High Throughput Organoid Formation- The 96-Well Plate System

Small, uniformly sized organoids are desirable for disease modeling applications as initial uniformity between organoid samples is critical for proper experimental control. Towards this end the large organoid generation process was modified to accommodate the formation of many organoids. This was achieved by first coating the functionalized alginate beads with cells then

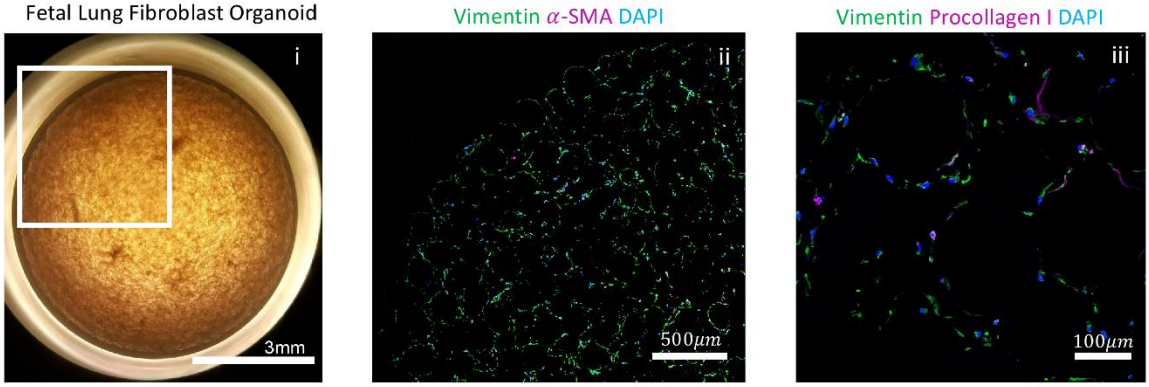


Figure 10. Characterization of the fetal lung fibroblast 96-well organoid model. (**i**) Representative organoid generated using fetal lung fibroblasts. (**ii-iii**) Confocal immunofluorescence micrographs of fetal lung fibroblast organoid sections for vimentin, collagen I, α-SMA and DAPI.

portioning the cell-coated beads into the wells of a 96-well plate and then centrifuging the plate to

sediment and compact the beads. This method was also exploited for the formation of multicellular organoids.

Organoid formation was performed with either single cellular (**Fig. 10**) or multicellular seeding conditions (**Fig. 11**). Single cell type seeding conditions are the same as those used for large organoid formation whereas multicellular seeding conditions represented a 1:1:1 combination of endothelial, epithelial, and mesenchymal cell types. Organoid condensation was never observed if fibroblasts were not included. Organoid formation is confirmed when the beads are observed to contract inward leaving a gap between the newly formed organoid and the edge of the well.

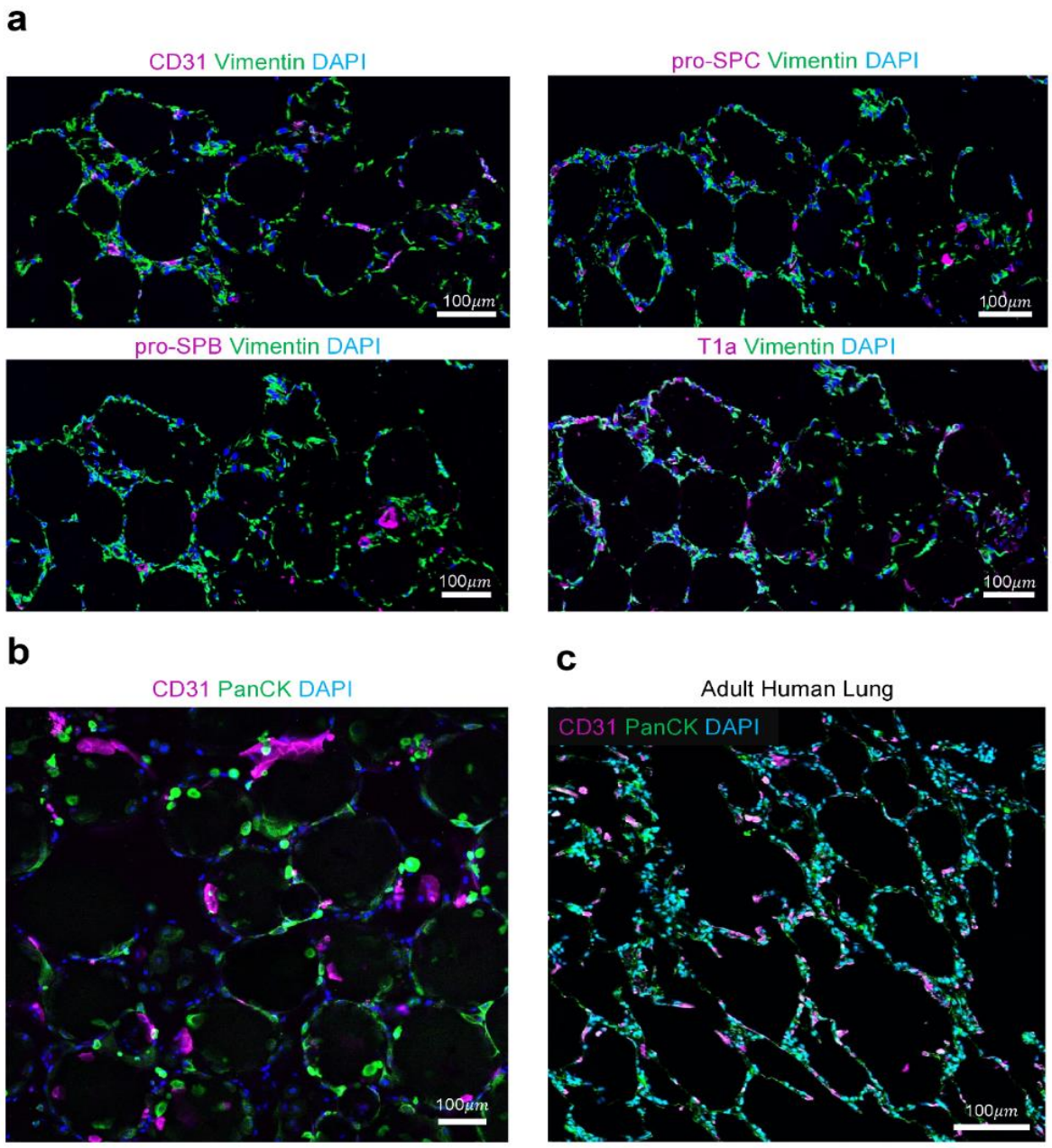


Figure 11. Immunostaining of 3D, multicellular organoids compared to adult human distal lung. (**a**) Confocal micrograph of cross sections of 3D multicellular lung organoids with immunofluorescence for CD31 (HUVECs), vimentin (FLFs) and pro-SPB and pro-SPC (Type II alveolar epithelial cells) and T1a (Type I alveolar epithelial cells). (Scale bar = 100μm). (**b**) Confocal micrograph of multicellular 3D lung organoids with immunofluorescence for CD31 (HUVECs) and PanCK (SAECs). FLFs were also seeded. (**c**) Confocal micrograph of a cross section of normal adult human lung with immunofluorescence for CD31 (HUVECs) and PanCK (SAECs). (Scale bar = 100μm).

Chapter 2.3.5. Myosin II Phosphorylation Drives Organoid Condensation

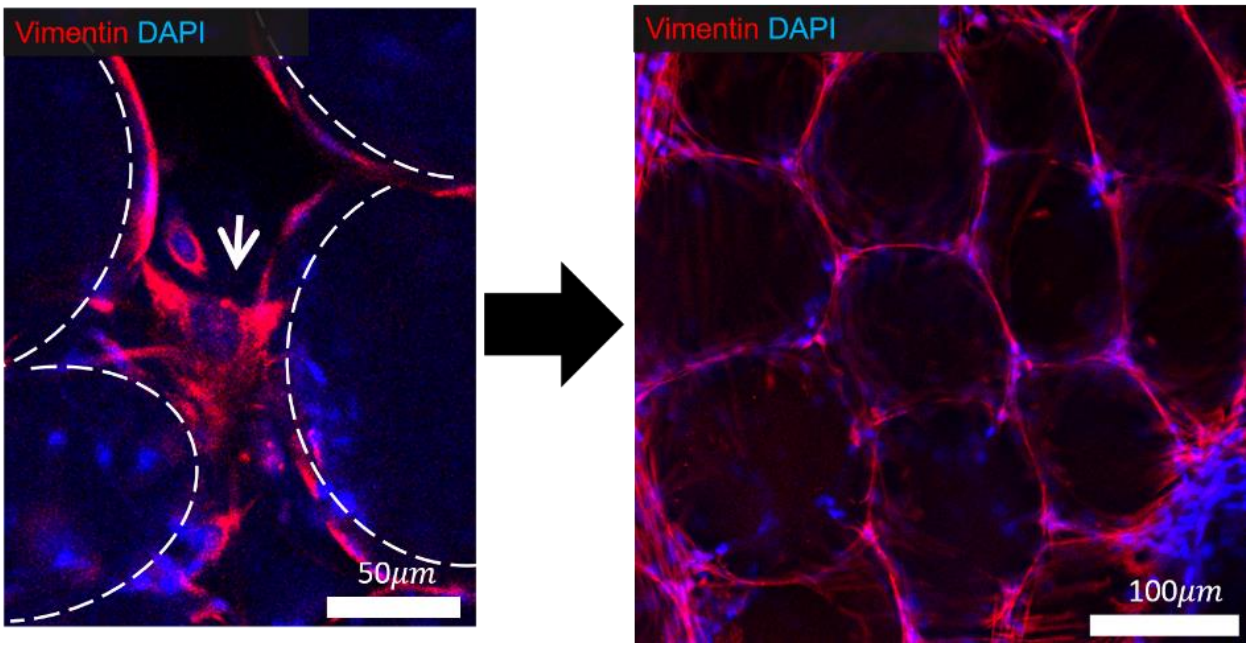


Figure 12. Confocal high power micrograph of vimentin immunostained mesenchymal cells showing cellular bead-bead bridge formation leading to bead compaction and organoid formation.

Mesenchymal cells, specifically fibroblasts, are necessary for organoid formation due to their ability to transdifferentiate into the myofibroblast phenotype. Myofibroblasts, *in vivo*, are critical to the wound healing process as these cells actively produce ECM and contract to achieve wound closure. When fibroblasts are added into the organoid generation process they undertake a similar phenotype wherein the cells will spread from one bead to another and subsequently contract thereby compacting the beads into an organoid (**Fig. 12**). Myosin II is the key driver of cellular contraction and thereby organoid formation. This was demonstrated in both the large organoid and 96-well systems by showing that inhibition of myosin II phosphorylation using blebbistatin subsequently results in the lack of organoid formation.

Large organoid formation under the influence of blebbistatin was observed using timelapse imaging. Initially, fibroblasts and functionalized beads were added to the 4mL bioreactor. The

bioreactor was then rotated under the mildly correlated flow regime (16rpm) for 24 hours. A sample removed after 24 hours was incubated with 2mM calcien AM. Beads were observed to be coated with cells along with bead-bead bridges and clumping (data not shown). $5\mu M$ blebbistatin was added to the culture and maintained for 2 days. Over the course of the first few hours beads formed into a lacey, loose structure though no contraction was observed until day 6, 3 days after blebbistatin was removed from culture. Stiffness measurement were not possible on blebbistatin treated organoids during blebbistatin culture as the organoid was not observed to move once all beads had clumped into a single unit. Therefore, blebbistatin suppressed organoid formation and

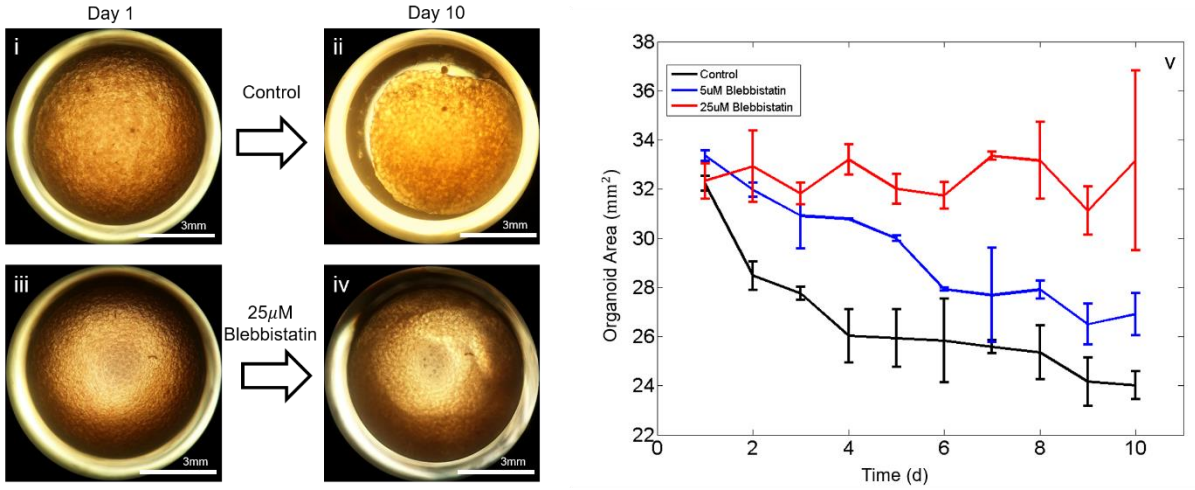


Figure 13: Organoid contraction is inhibited by the addition of blebbisatin, a myosin II heavy chain phosphorylation inhibitor. (**i-iv**) Effect of blebbistatin on organoid contraction. Organoid contraction either slowed or was completely inhibited by adding increasing amounts of blebbistatin to culture media. (Scale bar = $3mm$). (**v**) Plot of organoid area vs. time at different concentrations of blebbistatin.

upon its removal from the culture media was organoid contraction observed.

96-well generated organoids were also cultured under the influence of blebbistatin. Mesenchymal organoids were prepared in the 96-well plate as previously mentioned and allowed to mature over 3 days (with daily media changes). On day 4, organoids were given media

supplemented with 25μM, 5μM blebbistatin, or DMSO control. Supplemented media was changed daily over the next 9 days. Organoids were imaged daily. Here the use of two different concentrations of blebbistatin showed a dose-dependent effect on organoid contraction (**Fig. 13**).

Chapter 2.4. Discussion

The future of tissue engineering lies in the exploitation of the third dimension. Traditional, two dimensional cell culture lacks the dynamic complexity and hierarchy of even the simplest of *in vivo* tissues. This inherently affects the interaction, function and phenotype of the cultured cells and is demonstrated by the fact that most primary cells are difficult to maintain in regular culture. Three dimensional systems, on the other hand, mimic the interactions between cells and between cells and their tissue microenvironment by influencing surrounding cell types, scaffold stiffness and degradability, cell-cell and cell-scaffold adhesions, and by establishing cytokine and growth factor gradients. Along these lines we have developed a method for the generation of distal lung-like 3D organoids. These organoids were designed to emulate both the architecture and cellular composition of distal lung by scaffolding multiple cells around functionalized hydrogel beads and allowing these beads to interact and condense in a rotational bioreactor. The resulting organoids have a low cellular density with their structure being maintained by the presence of the indwelling alginate beads. These beads serve as a 3D alveolar template, maintaining an opposing force that keeps the interpenetrating cells from contracting into a tight clump. We speculate that the modularity of the beads and cell patterning could be used to model other 3D tissues which consist of repetitive patterns.

We discovered that organoid formation was not possible without the inclusion of fibroblasts, as organoids that were seeded without fibroblasts failed to contract and lacked the structural integrity of 3D tissue. This observation suggested that organoid formation is analogous

to aspects of *in vivo* wound healing, specifically, the tendency for fibroblasts to infiltrate a wound site, lay down collagen I and contract. Given their mechanism of formation, we hypothesized that these organoids would be prime candidates for modeling fibrotic lung diseases such as IPF.

This lung organoid generation method differs from other 3D culture methods in the sense that it exploits the aggregation of many individual cell-coated scaffold units, alveolar units in this case, in order to form the extended tissue network. Other scaffold-based methods require cells to either be perfused into a decellularized lung[54] or to migrate into the interior of a biodegradable foam[71]. Our method, on the other hand, ensures that the seeded cells are initially distributed throughout the entire scaffold due to the fact that each bead is first individually coated with cells. In this case, individual cell coated beads form interbead adhesions and subsequent contraction leads to organoid formation. In addition, the geometry imparted by the agglomeration of the alginate beads introduces a more physiologically relevant scaffold to study the individual cellular mechanics of contraction. Specifically, the curvature of the beads is similar to that of alveolar sacs *in vivo*. Finally, in order to fully densify, the organoid fibroblasts must grow into the interstitial spaces between beads. This process of bridging a void, instead of proliferating in a gel, is likely more analogous to the fibrotic processes occurring *in vivo*.

Takebe et al. have reported that the contraction of mesenchymal stem cells on a soft matrix in the presence of other endothelial and pancreatic cells resulted in a self-organized organoid that, when transplanted in type 1 diabetic mice, was able to regulate blood glucose levels. Our organoid system is a 3D analogue of the work done by Takebe [25] as it also relies on mesenchymal contraction for organoid formation. Unfortunately, our attempts at multicellular culture did not show signs of capillary organization though it was clear that these cells were able to survive the seeding process and remain viable for 2 weeks in culture.

Chapter 2.5. Conclusions

In summary, we generated a 3D model system of human distal lung, the lung organoid. One of the major advances of our organoid is the fact that it is a high throughput, multicellular system tailored for generating many experimental replicates. This will allow disease modeling in a highly reproducible way enabling high throughput drug screening. Though, we did not observe any higher order cellular self-assembly, such as vascularization. Ultimately, the goal is to develop functional lung organoids that preform gas exchange and could replace damaged patient lungs in an autologous fashion. In the future, these lung organoids may hold great potential to be used as platforms for developing a respiratory membrane with functional vasculature.

Chapter 3. Engineering a Model of Idiopathic Pulmonary Fibrosis

Chapter 3.1. Introduction and Motivation

Lung diseases are among the leading causes of morbidity and mortality worldwide and account for many billions of dollars of healthcare expenditure[5]. As the population ages, the burden from chronic lung diseases is expected to increase, with higher morbidity and mortality rates attributable to them[72]. Lung diseases include idiopathic pulmonary fibrosis (IPF), chronic obstructive pulmonary disease (COPD), acute lung injury (ALI), acute respiratory distress

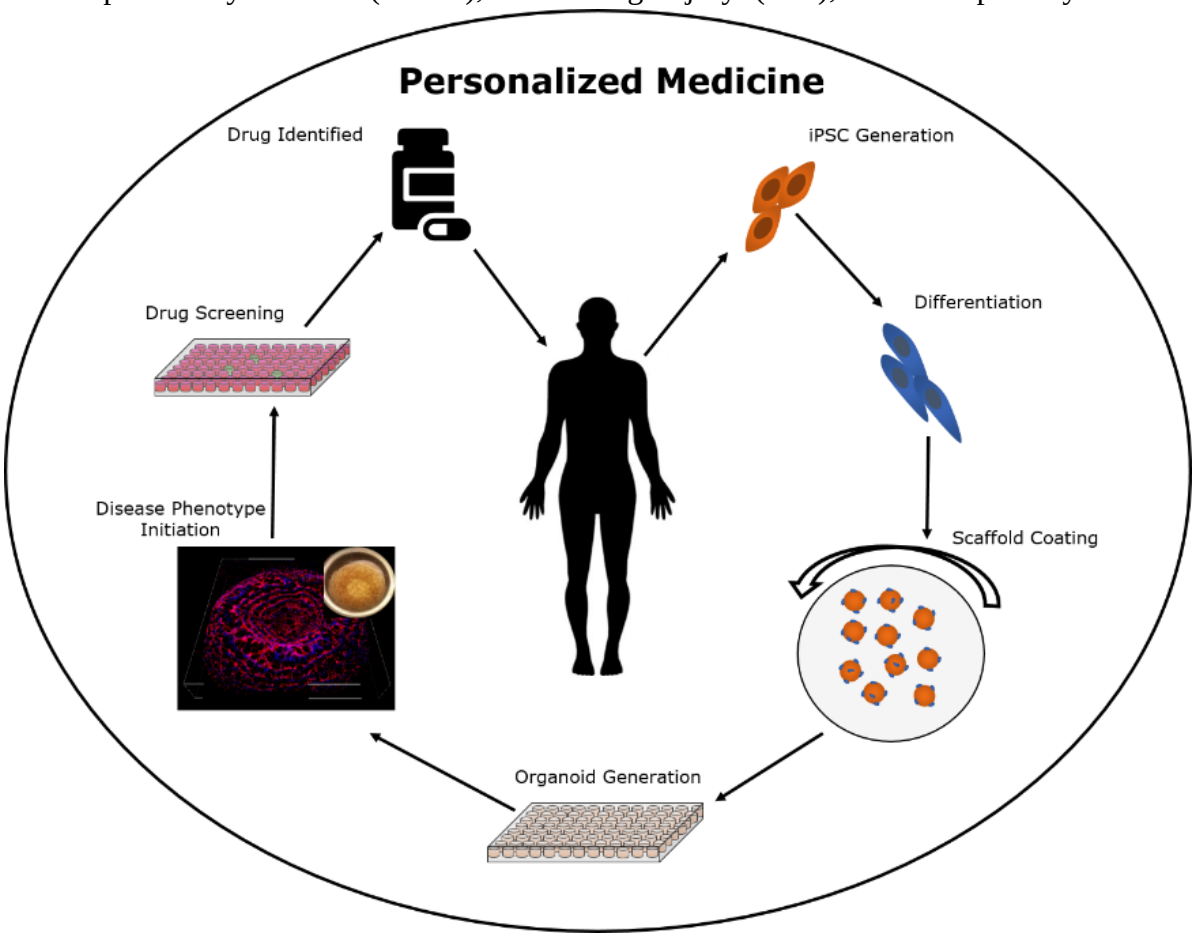


Figure 14. The personalized medicine loop. The process beings when primary cells are collected from the patient. These cells are inducted into a pluripotent state and differentiated into the cells necessary for disease recapitulation. These cells are then used in generating a disease-in-a-dish and then screened for compounds that would inhibit or reverse disease progression. Finally, drugs that prove successful are administered to the patient.

syndrome (ARDS) and bronchopulmonary dysplasia (BPD). These conditions are all characterized by abnormalities of the respiratory membrane that limit gas exchange and for which there are no cures[73]. Thus, there is a critical need to develop more effective therapies for these respiratory diseases. A major obstacle for the development of therapies for IPF has been a lack of human models, as *in vitro* and animal *in vivo* studies do not faithfully reproduce the pathophysiology of the disease[43,74]. There is also a well-recognized heterogeneity of these lung diseases with wide variability from patient to patient, though no individualized models are available to screen for the best patient-specific therapy.

This lack of effective therapies for IPF is likely do to the fact that the etiology of the disease is unknown. While IPF is poorly understood, it is believed to be a fibroblast-driven, complex, multifactorial disease arising from an abnormal wound healing response to multiple microscopic injuries[74] and asymmetric stress distributions due to heterogeneous alveolar geometries[75]. In addition, analysis of IPF patient samples shows transcriptional similarities between IPF and fetal lung development[12] indicating that studies employing immature pulmonary cells may yield relevant results. Regular 2D cultures of lung fibroblasts derived from IPF patients do not demonstrate the morphological fibroblastic foci that are classic for the disease[39]. In addition, animal models have failed to recapitulate many of the features seen in patients with IPF. Thus, compounds identified to reverse fibrosis in animal models and 2D tissue culture screens have failed to cure humans[43], underscoring the need for a relevant human disease model.

The complex, heterogenous nature of IPF makes it a prime candidate for personalized medicine-based drug discovery (**Fig. 14**). This goal of personalized medicine requires the confluence of several key factors. First patient cells must be collected, induced into a pluripotent state and then expanded. The cells must then be differentiated down the lineage/lineages necessary

for disease recapitulation. Next, these cells must be induced into the disease state and a disease readout determined. Finally, this process must be adapted to high throughput drug screening such that the disease model, readout, and overall process flow are easily automated to allow for large number of drug candidates to be assayed and then administered back to the patient. We hypothesized that the fibroblast-only lung organoid model would provide a disease modeling platform amenable to the necessities of high throughput drug screening.

Prior to drug screening for compounds that would resolve the progressive fibrosis observed in IPF patients, the disease model must be generated, characterized, and validated. Toward this aim we treated fetal lung fibroblast, 96-well organoids with TGF-β1, a known fibrogenic initiator. We observed that TGF-β1 treated organoids showed signs of the fibrotic phenotype observed in IPF *in vivo*. Namely, these organoids expressed higher quantities of collagen 1, α-smooth muscle actin (α-SMA), and contracted significantly more than controls. In addition, there were local areas of mesenchymal proliferation reminiscent of the fibrotic foci that are the hallmark of IPF. Furthermore, we used patient-derived cells to engineer an α-SMA reporter wherein cells would fluoresce in proportion to the amount of α-SMA being produced, removing the need for time consuming fixation and labor intensive immunostaining procedures.

Finally, we adapted our disease model to the process flow of high throughput drug screening. We used a large particle sorter to partition out individual cell-coated beads into the wells for a 384-well plate. This new methodology ensured the same number of beads, and therefore cells, were aliquoted into each well thereby improving well to well uniformity and dramatically reducing the total number material and cells necessary per organoid.

Chapter 3.2. Experimental Methods

Chapter 3.2.1. Generation of induced pluripotent stem cells from healthy adult lung samples and their spontaneous differentiation along the mesenchymal lineage.

Collection of adult lung biopsies was procured according to UCLA IRB protocol (#08-09-038-01), from the UCLA Medical Center at the time of lung transplantation. Lung biopsies were obtained from 5 healthy adults. The iPSCs were generated as per established protocols by Karumbayaram S. et al., 2011. Briefly, the punch biopsy samples were rinsed twice in HBSS and chopped into 1-mm pieces in 2% animal origin free collagenase solution. After 90 minutes incubation at 37°C in a 5% CO_2 incubator, the tissue was collected and centrifuged at 300g for 5 minutes. The supernatant was aspirated, and the pellet was washed once with 10 mL of MSCGM-CD and centrifuged as described above. The pellets containing the dissociated cells and tissue clumps were collected in 2 ml of MSCGM-CD medium and plated on a CELLstart-coated dish. Media was changed once every 72 hours until the cell monolayer was 70% confluent. Cells were passaged using TrypLE. and cryopreserved in ProFreeze CDM as per the manufacturer's protocol. For the generation of iPSC's, 1×10^5 fibroblast cells were plated in a CELLstart-coated well of a six-well plate in MSCGM-CD medium and transduced with STEMCAA (kind gift from Dr. Darrell Kotton, Boston University, MA) vector concentrate (7×10^6 TU/ml) in 1 ml of MSCGM-CD medium containing 10 μg/ml polybrene and incubated overnight at 37°C in 5% CO_2 incubator. The next day, media was aspirated, cells rinsed three times with MSCGM-CD and cultured for an additional 3 days in the same medium. On the 5th day, cells were replated in 50:50 TeSR2/Nutristem containing 10 ng/ml bFGF in two 6-cm dishes coated with CELLstart and cultured until iPSC-like colonies appeared. The colonies were picked mechanically and cultured in CELLstart-coated dishes, passaged mechanically using the EZPassage tool as per the manufacturer's protocol. The

colonies were collected by gentle pipetting and transferred to a 15-ml tube, and passaged at the dilution of 1:6 into a new CELLstart-coated plate. Three independent iPSC lines per lung sample were generated from lung biopsy.

To induce differentiation of iPSCs along the mesenchymal (osteogenic and adipogenic) lineage, iPSCs were dissociated using 1mg/ml of dispase for 10 minutes and gently scrapped to collect the colonies. The colonies were rinsed twice DMEM/F12 medium and then cultured in non-adherent dishes in DMEM/F12 medium supplemented with 10% FBS, 1x Glutamax, 10nM Non-essential amino acids and 0.1 mM monothioglycerol (MTG) for the generation of embryoid bodies. After 4 days, the embryoid bodies were collected and plated on gelatinized dishes to allow to adhere and cultured in media containing DMEM/F12 medium supplemented with 10% FBS, 1x Glutamax and 10nM non-essential amino acids resulting cells were cultured in DMEM with 10% FBS and additives for 3 weeks.

Chapter 3.2.2. ACTA2-mCherry iPSC-derived Mesenchymal Cell Line Derivation

Lentiviral particles that express mcherry under the control of the ACTA2 (α-SMA) promoter were purchased from Genecopoeia (Cat# LPP-HPRM14109-LvPM02). iPSC derived mesenchymal cells were plated in a 35mm dish at a density of 1×10^5 cells. Cells were about 80% confluent the next day and were transduced with 8μL lentivirus (1.15×10^8 TU/ml) in the presence of 2.0μL Polybrene transfection reagent (10mg/ml Millipore) in 1.5mL DMEM/F12. After 3h, cells were supplemented with 10% fetal calf serum. Stable clones were selected with puromycin (1.0mg/mL; Invitrogen). Selected cells were expanded in a T25 flask until 80% confluent.

Chapter 3.2.3. Lineage Dependent Characterization of iPSC-derived Mesenchymal Cells

For osteogenic and adipogenic differentiation, iPSC derived mesenchymal cells were plated at densities of 4.2×10^3 cells/cm^2 and 2.1×10^4 cells/cm^{2,} respectively in 8-chamber slides. After 2

days of incubation at which the cells reached 100% and 70% confluency, the cells were cultured for an additional 14 days in osteogenic differentiation medium (R&D Cat# CCM008) or adipogenic differentiation medium (R&D Cat# CCM011). Medium was changed every two days. Staining with osteocalcin for the presence of calcium deposits was used to assess the osteogenic differentiation of the cells. Similarly, the presence of lipid vacuoles and a positive stain for FABP4 was used to assess adipogenic differentiation of the iPSC derived mesenchymal cells.

Chapter 3.2.4. Generation and Quantification of the FLF Organoid Fibrosis Model

FLF mesenchymal organoids were prepared in the 96-well plate bioreactor as previously mentioned and allowed to mature for 4 days with daily media changes. On day 4, serum free media was introduced and maintained for a total of 2 days with daily changes. On days 6 and 7, organoids were given low-serum control media (1% FBS) or media supplemented with (10ng/mL TGF-β1) and incubated. Organoids were imaged daily. On day 8 hours, the organoids were processed for immunostaining or RNA analysis.

Chapter 3.2.5 Real-time PCR (qPCR)

Organoids were processed for RNA using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. An on column DNase (Qiagen) digestion step was included. cDNA was generated using the TaqMan Reverse Transcription Kit (Applied Biosystems) according to manufacturer's instructions. qPCR was performed using Taq Universal SYBR Green Supermix (Bio-Rad) on a StepOnePlus PCR system (Applied Biosystems). The following primer sequences were used, αSMA: Fwd: AAAAGACAGCTACGTGGGTGA, Rev: GCCATGTTCTATCGGGTACTTC; Col1A2: Fwd: GAGCGGTAACAAGGGTGAGC, Rev: CTTCCCCATTAGGGCCTCTC; and vimentin: Fwd: AGTCCACTGAGTACCGGAGAC, Rev: CATTTCACGCATCTGGCGTTC.

Chapter 3.2.6. 384-well Organoid Generation and Imaging

384-well organoid generation differs from 96-well organoid generation in only one respect, the method by which cell coated beads are added to the well plate following the HARV bioreactor process. We employed the use of a large particle sorter (Biosorter-Union Biometrica) to enable this process. The Biosorter functions similar to the FACS process; a continuous flow system capable of analyzing, computationally sorting, and dispensing objects ranging in size from 10-1500μm.

First cell coated beads are prepared by combining 2 million reporter line iPSC mesenchymal cells and 0.5mL of functionalized alginate beads into a 2mL HARV bioreactor. The bioreactor is then rotated at 16.5rpm for 1 hour in a cell culture incubator. Next, cell coated beads are diluted into 200mL of DMEM/F12 with 10% FBS media and added to the 750mL sample cup of the Biosorter using a 500μm mesh strainer(Pluriselect). Straining is necessary to break up any clumps that may have aggregated during the cell coating process. In order to ensure organoid viability serum free DMEM/F12 was used as the sheath fluid to minimize problems due to sheath-based sample dilution. The bead size distribution for sorting was gated to include beads in the range of 200-400μm. Given this sorting criteria, 25 beads were sorted into each well of a round bottom, low-adhesion 384-well plate(Corning). The process usually takes ~30 minutes to fill an entire 384-well plate. After filling the entire plate the plate is then centrifuged at 1000g for 5 minutes to sediment and compact the cell coated beads and then allowed to mature in a standard tissue culture incubator.

Fluorescence-based organoid imaging was performed on a ImageXpress Confocal High-Content Imaging System(Molecular Devices) at days 1 and 5 of the organoid formation process. Each plate was imaged using a 20x air objective under confocal mode. The microscope

automatically moves the plate using an automated stage to each well wherein 10 images are acquired with a 50μm z-step size. These images are then saved for future analysis.

Chapter 3.2.7. Generating a Model of Idiopathic Pulmonary Fibrosis

FLF mesenchymal organoids were prepared in the 96-well plate bioreactor as previously described and allowed to mature for 4 days with daily media changes. On day 4, serum free media was introduced and maintained for a total of 2 days with daily changes. On days 6 and 7, organoids were given low-serum control media (1% FBS) or media supplemented with (10ng/mL TGF-β1) and incubated. Organoids were imaged daily. On day 8 hours, the organoids were processed for immunostaining or RNA analysis.

Chapter 3.3. Results

Chapter 3.3.1 IPF Model Generation and Characterization

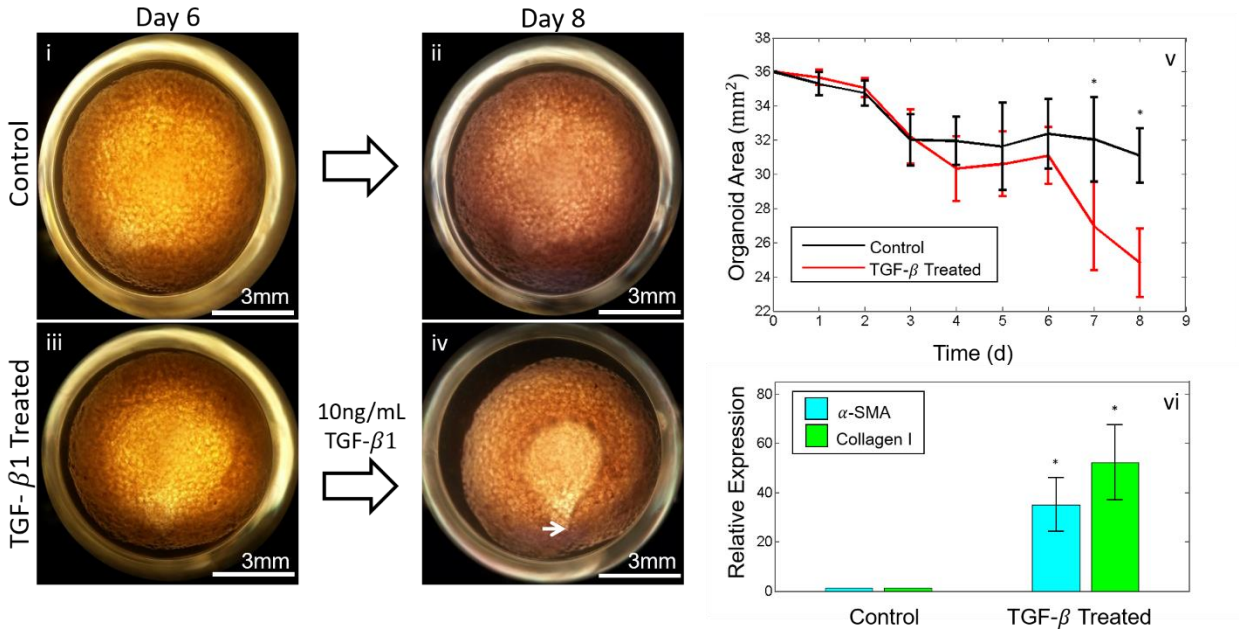


Figure 15. Effect of TGF-β1 on organoid contraction and development of a fibrotic phenotype. (**i-ii**) Representative control organoid imaged on days 6 and 8 post seeding. (Scale bar = 3mm). (**iii-iv**) Representative organoid treated with TGF-β1 during the same 2-day period. The organoid contracted forming a saddle-like geometry with the focal point near the bottom of the image indicated by an arrow. (Scale bar = 3mm). (**v**) Aggregate analysis of 20 organoids (10 experimental, 10 control) analyzed over the 8 day experiment. TGF-β1 was administered on day 6 thereafter a clear separation between experimental and control organoid contraction was observed. (*) P < 0.05. (**vi**) Expression levels of two key genes involved in fibrosis, collagen I and α-SMA, on treatment with 10ng/mL TGF-β1 by qPCR. (*) P < 0.05.

We generated a model of scarring by treating fetal lung organoids with TGF-β1 and showed that the resulting tissues were phenotypically similar to that of IPF lung. Fetal lung fibroblasts were chosen for these studies due to the transcriptional homologies between IPF and fetal lung samples[12]. TGF-β1 is known to play a central role in the development of tissue fibrosis as it causes fibroblasts to differentiate into a myofibroblast phenotype and synthesize and contact ECM[76]. Mature fetal lung organoids were treated with exogenous TGF-β1. Organoid size was

monitored daily revealing that treated organoids contracted at a higher rate when compared to untreated organoids (**Fig. 15**). Hematoxylin and eosin (H&E) micrographs indicate that this contraction leads to increased bead packing and overall denser, smaller organoids (data not shown). Real-time PCR (qPCR) identified an increase in expression of collagen I and α-smooth muscle actin (α-SMA) in the TGF-β1 treated samples when compared to controls (**Fig. 15**). Immunofluorescence of sectioned organoids indicated higher levels of collagen I and local patches of α-SMA demonstrating activated myofibroblasts which morphologically resembled the fibroblastic foci that are the hallmarks of IPF (**Fig. 16**).

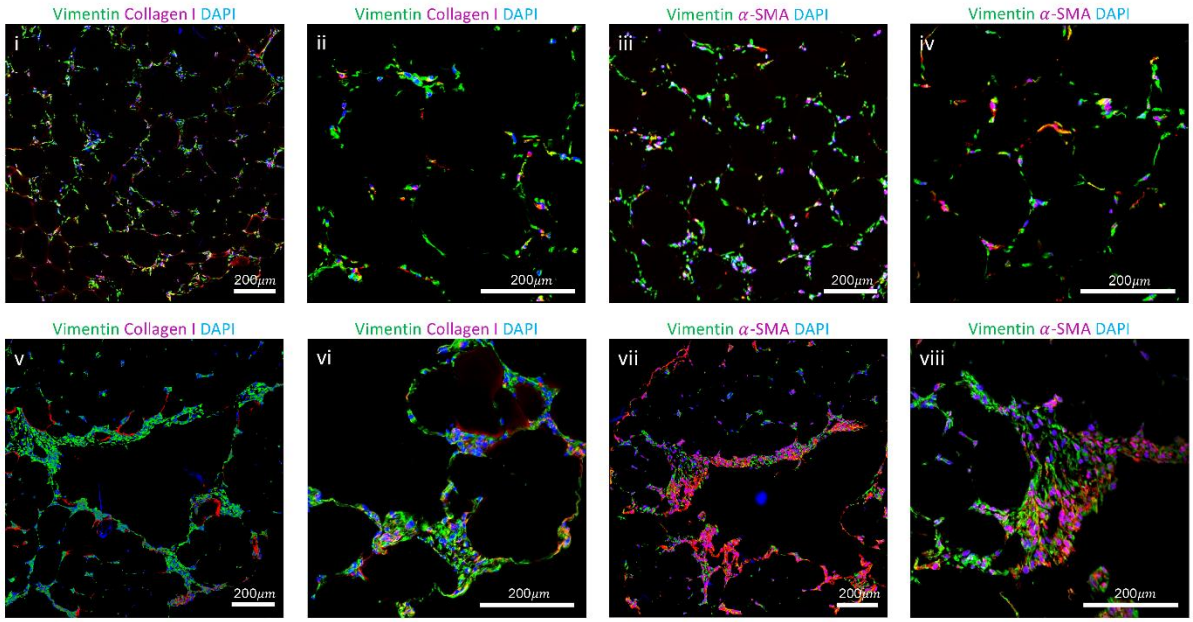


Figure 16. Characterization of the fibrotic phenotype established upon exposure to TFG-β1 (**i-ii**) Confocal immunofluorescence micrographs of representative control organoid sections for vimentin, collagen I, α-SMA and DAPI. (**iii-iv**) Confocal immunofluorescence micrograph of representative TGF-β1 treated organoid sections for vimentin, collagen I, α-SMA and DAPI. Fibrotic areas, indicated by the white arrows, show increased accumulation of cells that stain positive for collagen I and α-SMA resembling fibrotic foci, the hallmark of IPF. (Scale bar i,iii,v,vii = 400μm). (Scale bar ii,iv,vi,viii = 200μm).

Chapter 3.3.2. iPSC-based α-SMA Reporter Organoid Generation and Characterization

In order to perform a successful high throughput screen, it is necessary to develop an assay of a disease model that both faithfully mimics the disease and is reproducible over the large number of culture samples. In addition, the characterization and analysis of the assay needs to be done efficiently and reproducibly. With this in mind, we realized that the process of fixation, sectioning and immunostaining of thousands of organoids would not be technically feasible. We addressed this problem by generating a fluorescent reporter line wherein live cells would produce fluorescent molecules in proportion to the production of the α-SMA. This means that the degree of α-SMA expression, and therefore fibrosis, is directly reflected in the fluorescent signal of the organoids allowing for immediate, live cell imaging. This was accomplished by using a lentivirus to transduce our patient specific, iPSC-derived mesenchymal cells to express mCherry under the control of an ACTA2 (α2 actin, smooth muscle or α-SMA) promoter(**Fig. 17**).

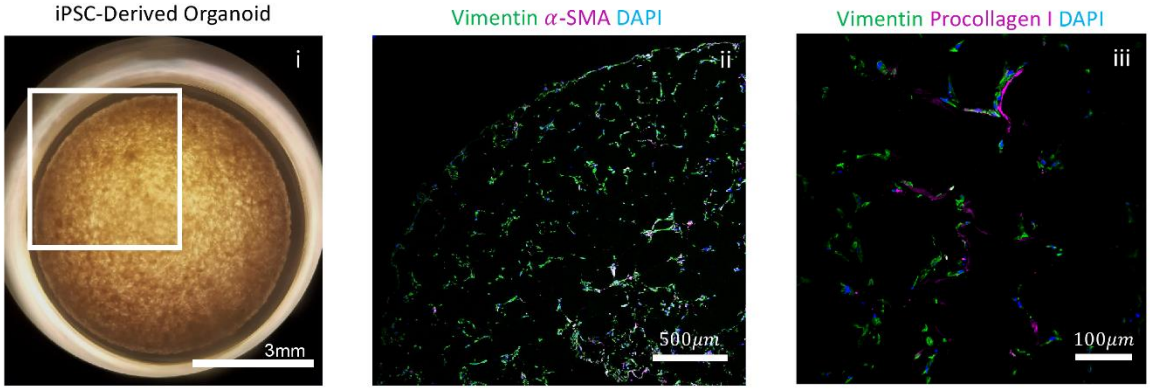


Figure 17: Characterization of the iPSC-derived 96-well organoid model. (**i**) Representative organoid generated using iPSC-derived fibroblasts. (**ii-iii**) Confocal immunofluorescence micrographs of iPSC-derived fibroblast organoid sections for vimentin, collagen I, α-SMA and DAPI.

These cells were used to form organoids and were put through the previously described TGF-β1 treatment. Live cell, confocal imaging was used to quantify the fluorescence from both control and treated organoids (**Fig. 18**). The fluorescence signal was summed over the z-stacks resulting in a 2.5x increase in mCherry in the TGF-β1 treated samples when compared to control.

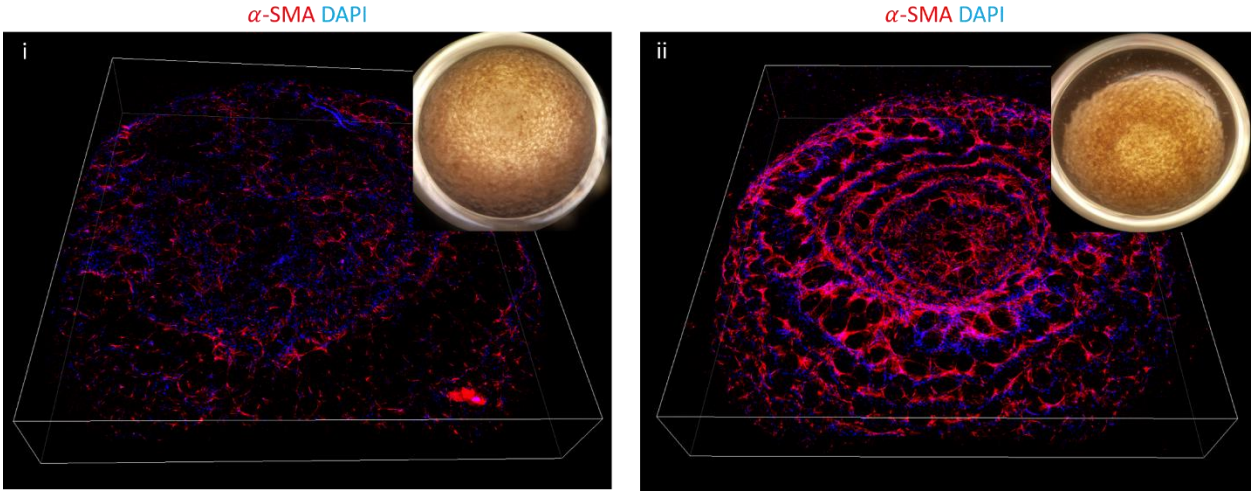


Figure 18: Development of fibrotic phenotype organoid for patient-specific, high throughput drug screening. (**i**) Merged, rotated confocal z-stack of patient, iPSC-derived, α-SMA reporter line control organoid. Inset, white light image of organoid. (**ii**) Merged confocal z-stack of patient, iPSC-derived, α-SMA reporter line organoid treated with TGF-β1. Inset, white light image of organoid showing high degree of contraction.

In addition, a difference in organoid contraction commensurate to that measured in fetal lung organoids was observed. We believe that the combination of these two metrics will provide a robust platform for future TGF-β1 based drug screening.

Chapter 3.3.3. 384-well Organoid Formation for High Throughput Screening

Uniformity of organoid size and cellular content is essential for high throughput drug screening. We were able to meet this need by adapting the lung organoid methodology to accommodate a large particle sorter system wherein 25 cell coated beads are sorted into each well of a 384-well plate (**Fig. 19**). Organoid imaging was performed using an automated confocal system and summed to produce a single image projection (**Fig. 19 b**). This was necessary due to

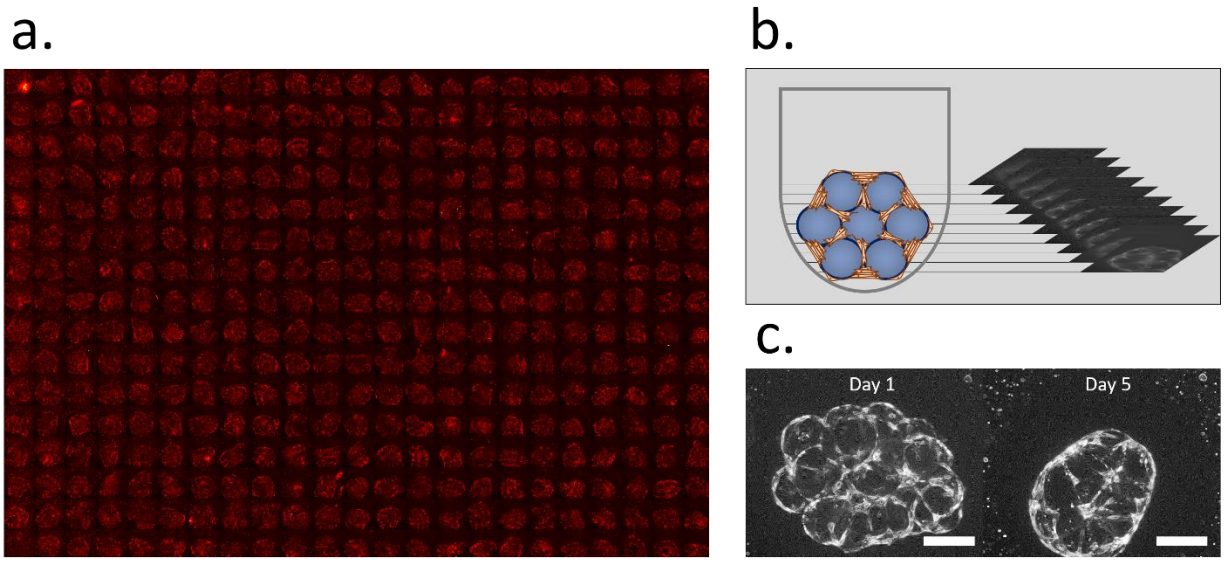


Figure 19: Organoid formation in a 384-well plate format. (**a**) Confocal projections of α-SMA-reporter lung organoids. (**b**) Confocal microscopy is used to image lung organoids. (**c**) 384-well organoids form normally by forming bead-bead bridges and contracting. Scale bar (200μm).

scaling constraints inherent in making enough organoids for even the most limited of compound screens, on the order of 10^3. To put this into perspective the 384-well system reduces the number of cells per organoid by a factor of almost 100x. Given the fact that we were able to collect roughly 1.5 million cells per each T182 tissue culture flask, a screen of 6,000 compounds would require 20 plates for the 384-well and 2,000 plates for the 96-well system.

The critical aspect in maintaining organoid viability during the particle sorting process was the use of serum-free cell culture media as the sheath fluid. This is due to the fact that the particle

sorting process inherently dilutes the particle solution. 384-well organoids were observed over several days and were seen to contract and reorganize the beads indicating the particle sorting process maintained cell viability and function(**Fig. 19 c**).

Chapter 3.4. Discussion

A major obstacle for finding new therapies for patients suffering from IPF is the lack of effective human models for studying the disease in a dish. In addition, it is uncertain if there will be one solution that will be effective in treating all IPF patients as the disease is heterogenous and may have a complex genetic basis[13]. These considerations make IPF a prime candidate for personalized disease modeling. The current gold standard for modeling IPF is currently the bleomycin mouse model though despite widespread usage it has yet to provide any clinically relevant advances[43]. Despite the inherent advantages of such models, the presence of fully formed and functional lungs to experiment on, there are major differences genetic response to stimuli between humans and mice[77] potentially negating the very benefits that makes animal models so attractive.

We have shown that using our established organoid generation methodology that we can generate a model of scarring in a dish. The lung organoid model grants several distinct advantages for disease modeling of fibrosis. Lung organoid formation begins when individually cell-coated beads are compacted together and the included fibroblasts begin to form cellular bead-bead bridges and contract. The process of proliferation, elongation, and contraction is a direct result of the fibroblasts entering a myofibroblast-like state similar to what is observed during the normal wound healing process. Since IPF is a disease characterized by aberrant wound healing we believe that the lung organoid formation process should be a good match for modeling the progression of IPF in patients. Furthermore, the remodeling observed in the organoid samples is taking place in a 3D

environment that is structurally similar to that of human distal lung further capturing any geometric factors that may play into the progression of the disease scarring process. Finally, the lung organoid formation process is amenable to the inclusion of a wide range of cell types including iPSC-derived patient cells. The ability to personalized this process and preform high throughput biology may provide a better disease modeling platform for personalized drug discovery.

This model of IPF in a dish was generated using TGF-β1 as the fibrogenic initiator and was successful at reproducing the hallmark fibrotic foci observed in IPF histological samples. α-SMA was chosen as an indicator of the fibrotic phenotype as it is commonly observed to be upregulated in patient samples. Though, there are some potential pitfalls in using these methods and indicators. First, while it is true that TGF-β1 initiates a fibrotic phenotype it is uncertain whether this initiation process will induce genetic expression patterns that are relevant to the *in vivo* disease process, especially given the genetic abnormalities of IPF patients (mutations in mucus and surfactant producing genes). Furthermore, α-SMA may not be the best indicator of fibrosis for modeling IPF. It may be more fruitful to consider metal metalloprotease (MMP) activity or enzymatic initiators of collagen crosslinking as these factors are fundamental in scar formation and resolution.

Chapter 3.5 Conclusions

In summary, we generated a patient-specific human model of IPF in a dish and showed that the methods may be adapted to high throughput drug screening. We believe that by using patient-derived cells in a 3D, high throughput drug screen we may be able to find new therapeutics for patients suffering from IPF and other fibrotic diseases of the lung.

Chapter 4. Convolutional Neural Network Analysis for High Throughput Drug Screening Classification

Chapter 4.1. Introduction and Motivation

The basis of phenotypic drug screening is the assertion that a cell's phenotype is a direct consequence of the cell's underlying genetic expression profile[78]. A phenotype contains those factors that are directly observable characteristics, rather than features that are only indirectly inferred. These phenotypic characteristics include cell size, shape, stiffness, behavior, and genetic expression profiles as assayed by reporter fluorescence, immunostaining, or other optical or mechanical means. Furthermore, it is assumed that changes in phenotypic expression are due to underlying changes in genetic or epigenetic profiles and therefore changes in the various RNA and protein landscapes within the cell. Therefore, the basis of phenotypic drug screening is to observe the effect of experimental small molecule supplementation on a disease state phenotype. This information can then be used to narrow down the list of bioactive compounds for further investigation and trials.

The success of any high throughput phenotypic drug discovery program relies on the successful characterization and classification of large quantities of data. These data are typically too large for a human to analyze and require an automated analysis system to evaluate the dataset. As new methods for generating 3D models of disease become mainstream this task becomes ever more difficult for the current generation of automated data analysis systems. Specifically, while there are a wide range of analysis packages available there is no generalized system for quickly classifying experimental datasets given known controls. Therefore, each new experiment requires its own specialized analysis package whose design and implementation are complicated and typically out of the range of expertise of the scientist performing the screen[79]. Furthermore, the

automated data analysis algorithms that are commercially available and commonly used employ predetermined, static rule sets for quantifying the data. It is unclear if these methods will be sufficient for analyzing the more complicated datasets attained from new classes of 3D disease models.

Images comprise the bulk of data collected during such screens and therefore encode a wealth of experimental information. These images reflect underlying cellular phenotypes given various experimental conditions, typically the supplementation of a small molecule compound. For 2D, monolayer cultures cellular phenotype is defined by a limited set of parameters. These include the number of viable cells before and after experimental treatment, cell monolayer confluency, cellular morphology and aspect ratio, cellular adhesions either between adjacent cells or the underlying substrate, fluorophore localization, and, finally, optical intensity readouts that correlate with a molecular expression pattern of interest. While this list is not exhaustive it does highlight some of the main metrics that are considered for most 2D screens and therefore the underlying challenges for automated image analysis. Namely, cell counting, edge detection, image segmentation and cellular identification, and intensity summation after thresholding[80,81]. While, these methods are well developed and commonly used each must be tailored to accommodate new datasets. Moreover, because the rule set employed in these algorithms are static they do not offer much flexibility and demand a highly uniform dataset for meaningful results.

Three-dimensional disease models offer an entire new set of challenges for automated image analysis. While, fundamentally, the same basic metrics previously outline are still of interest the mathematical complexity of computing these metrics is dramatically increased. For example, counting the number of cells in a 2D monolayer is straightforward and there are many algorithms that have shown a high degree of accuracy when performing this task. Counting cells in 3D

however requires the deconvolution of signals from multiple, overlapping cells. Confocal sectioning and z-stack formation may allow for pseudo-2D approaches to be attempted but it imposes new challenges of continuity. Specifically, how to determine if a single cell is present in multiple z-stack images to reduce the chance of counting the same cell multiple times. To be clear, these problems are fundamentally not mathematically intractable yet tend to require a high degree of computational skill and customization[82].

While the standard automated image analysis programs may struggle to quantify and categorize the phenotypes contained in images, to the trained expert, the differences between various experimental categories is obvious[83]. In other words, there is a great gap between the expert's ability to discern phenotype and the algorithm's ability to calculate metrics that correlate well with the expert's analysis. In order to bridge this gap, it may be necessary to employ machine learning approaches to analyze and classify phenotypically complex high content data. State of the art image analysis algorithms rely on machine learning techniques for their success[52]. Machine

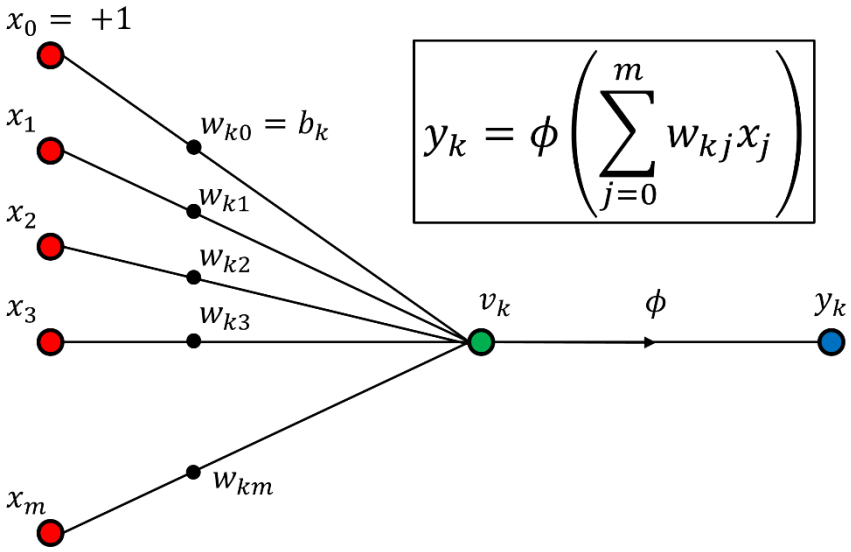


Figure 20: Schematic of an artificial neuron. The output of the neuron y_k is equal to the sum of weighted inputs from previous artificial neurons $w_{kj} \cdot x_m$ multiplied onto the activation function ϕ.

learning differs from traditional image analysis insofar as these systems are mathematically malleable and require training for proper function. While there are a wide range of machine learning methods, and likewise training strategies, the basic necessity for any supervised learning technique is user input in the form of dataset curation and labeling. For example, Imagenet[52] is an open project focused on compiling lists of image labels and URLs available online. This list currently includes over 14 million images grouped into over 21 thousand categories. To date, there are a host of algorithms that, given the task of assigning a known label to an unknown image found in the Imagenet database, have achieved error rates of less than %10 with the very best rivaling human error rates. The secret to the success of these algorithms is the use of Deep Convolutional Neural Networks. Convolutional neural networks are mathematical structures that were originally inspired by the biological processes at work for object detection at a cellular level of the animal visual cortex, first described by pioneering work by Hubel and Wiesel in the late 1950s[53]. In addition to generalized image classification, convolutional neural networks have also shown promise in computer aided diagnostics[83] and single cell phenotypic classification[84]. Due to the fact that these networks have been shown to be universal approximators[51] (mathematical structures that can fit any functional mapping given some weak requirements) they will likely be applicable for any image classification based application.

Convolutional neural networks are a class of multilayer feed-forward neural networks comprised of an input layer, output layer and multiple hidden layers of fully or partially interconnected artificial neurons. These layers are typically convolutional, pooling, and fully connected layers wherein the calculations performed on a previous layer are then fed into the next layer with this process repeating sequentially from the first to the final layer[85]. Linear multipliers, known as weights, are multiplied onto each nodal operation providing an adjustable parameter

necessary for network training. Finally, an activation function is applied to each neuron's output (**Fig. 20**). The final output layer is typically a vector of numbers whose values represent the categories of data used to train the network. For example, the network may be trained on two categories labeled 'healthy' and 'diseased.' The output of this network would contain a vector of two numbers each representing the network's probability assessment of that category. The network is trained by providing a set of training and validation data. Training data is fed into the network and classification probabilities are calculated. Next a technique known as back-propagation is performed wherein gradients between layers are calculated and the weights adjusted such that the classification error is minimized. Validation data (data the network is not explicitly trained with) is also used for error calculation to minimize overfitting.

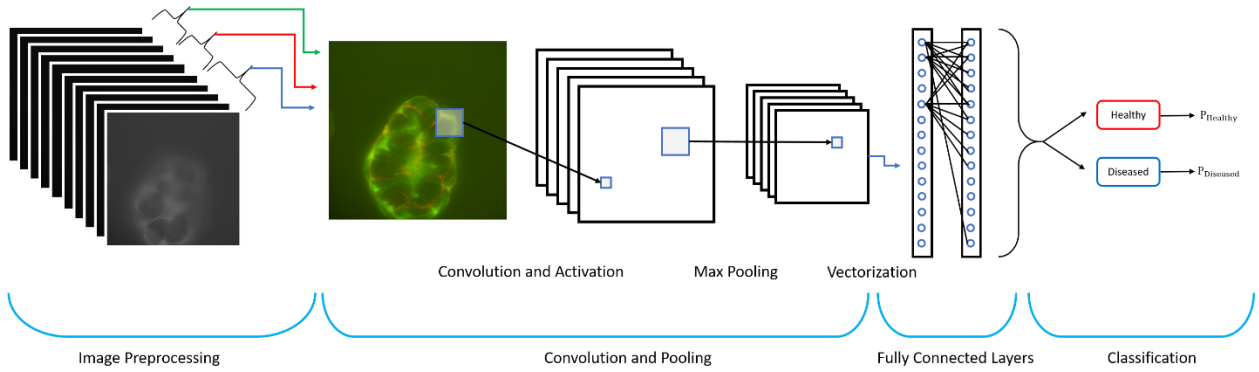


Figure 21: Generalized schematic for multilayered neural network classification system. Z-stack images from α-SMA reporter organoids are sequentially summed and merged into an RGB images. These images are then fed forward into the network through a sequence of convolution, activation, and pooling layers. The data is then passed through fully connected layers to arrive at the final classification probabilities.

Herein, we describe a method for using a convolutional neural network to classify 3D, high content screening data based on the α-SMA reporter IPF model. The goal of this organoid based

model is to use the phenotypic scarring observed upon supplementation of TGF-β1 as an indicator of fibrosis for high throughput screening. Prior to screening for antifibrotic compounds, we applied this method to categorize confocal stacks of α-SMA reporter line organoids that had been treated with increasing concentrations of DMSO as a proof of principal exposition on network training and classification. DMSO was chosen due to its widespread use as a solvent for small molecules in drug screening settings and its known cytotoxic effects at high concentrations. In addition, DMSO is known to alter collagen I deposition and generally effect the wound healing process[86]. This is of particular concern due to the fact that the organoid formation process itself is driven by fibroblasts, the cell type primarily responsible for collagen I production and scar formation during wound healing. Therefore, without proper DMSO control any fibrosis assay performed using the lung organoid model may be compromised by unforeseen DMSO-fibroblast interactions making phenotypic characterization of these interactions critical. We employ a simple, multilayer convolutional neural network to output the probability of the inclusion in one of two datasets labeled 'healthy' and 'diseased' (**Fig. 21**). Training data were taken from the no-DMSO and high concentration DMSO datasets and labeled 'healthy' and 'diseased' accordingly. It should be noted that these labels do not necessarily imply a strict biological significance but are simply used to denote the positive and negative experimental controls for this example.

Overall, this approach is structured to provide a generalized platform for quickly assessing large datasets of images by first training the network on curated control data followed by employing the network to classify unknown, experimental data. We found that the network classification provided a non-biased metric for assessing DMSO's effects on organoid formation. Specifically, organoids treated with DMSO up to a concentration of 2% vol. grouped with the 0% DMSO controls labeled 'healthy'.

Chapter 4.2. Experimental Methods

Chapter 4.2.1. Convolutional Neural Network Design and Training

Convolutional neural network construction, training, and implementation were executed in the Python v3.5 programming language and used Google's TensorFlow API r1.2. TensorFlow is an open source software library for enabling rapid development of machine learning and deep neural networks. These software packages are completely open source and may be downloaded and installed via these websites (https://www.python.org/), (https://www.tensorflow.org/).

The network outlined here is designed to classify images into two categories. The architecture is outlined in (**Table 1.**). After each convolutional layer a rectified linear unit (ReLU) activation function is applied. The final output layer has a sigmoidal activation function. Dropout is also applied to the fully connected layers to minimize overfitting. Binary cross entropy was used to compute the loss during training[50] and optimized using RMSProp. Training data is fed into the network in batches of 14 images and training continues until the calculated classification error reaches a minimum value.

Table 1. Architecture of the Convolutional Neural Network

Layer Description	No. of Images/Feature Maps x Their Dimensions	No. of Weights
Input	$3 \times 300 \times 300$	
Convolution (3x3)	$64 \times 300 \times 300$	$64 \times 9 \times 3$
Max Pooling (2x2)	$64 \times 150 \times 150$	
Convolution (3x3)	$64 \times 150 \times 150$	$64 \times 9 \times 64$
Max Pooling (2x2)	$64 \times 75 \times 75$	
Convolution (3x3)	$128 \times 75 \times 75$	$128 \times 9 \times 64$
Max Pooling (2x2)	$128 \times 38 \times 38$	
Fully Connected	128	$128 \times 128 \times 38 \times 38$
Fully Connected	1	1×128
Output	2	2×1

Table 1. The first layer consists of 3 channel images each with 300 x 300 pixels each. The hidden layers consist of feature maps that reduce by size due to max pooling layers, the number of feature maps are listed along with their weights. Dropout was applied to the fully connected layers to prevent overfitting.

Chapter 4.2.2. 384-well DMSO Stress Test Generation and Image Preprocessing

A 384-well plate of α-SMA reporter organoids was formed and imaged using the methods described in Chapter 3. Each two rows of the plate were supplemented with DMSO, starting at 0% and increasing in concentration by 0.5% for every subsequent grouping of two rows. Z-stack images, 10 per well, were processed into 3 channel RGB JPEGs each consisting of 300x300 pixels. Training and validation data were taken from the 0% and 3.5% DMSO conditions. 34 images were

taken, per condition, for training while the remaining 14 were left for validation. After training the intermediate DMSO concentrations were then categorized.

Chapter 4.3. Results

A multilayered convolutional neural network was designed, trained and implemented for the classification of a DMSO stress test dataset. The network was designed to import images consisting of 3 channels representing the summed signals from the acquired z-stack data. The DMSO dataset was chosen for this task because of the variety of DMSO concentrations that were included in the experimental setup. Upon inspection, the observed organoid phenotypes were observed to change gradually with increasing DMSO though there was no clear cutoff in terms of DMSO's effects on organoid formation (**Fig. 22**).

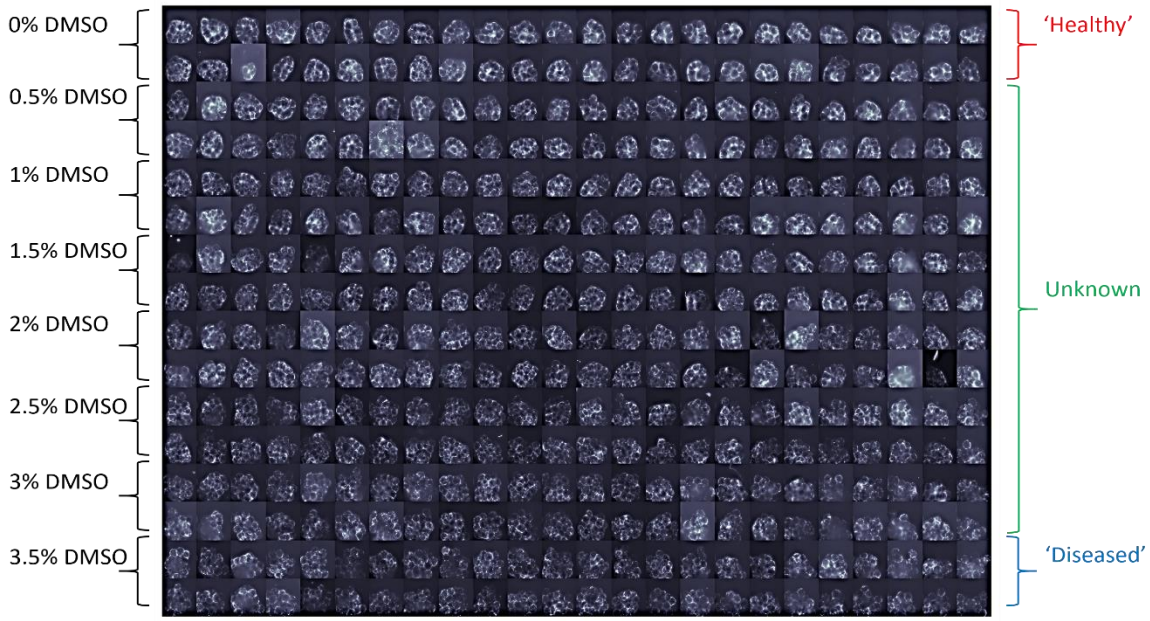


Figure 22: Overview of the DMSO stress test 384-well plate. Images are the maximum projections of the acquired z-stack, α-SMA reporter fluorescence. The first and final two rows are used as training data and labeled 'healthy' and 'diseased' respectively.

Network training was performed on the 0% DMSO and 3.5% DMSO datasets as these datasets showed distinct morphological features. The 0% DMSO set show signs of significant cell spreading, bead-bead bridge formation, and geometrical reorganization indicative of normal organoid formation. On the other hand, the 3.5% DMSO dataset included malformed organoids whose cells showed signs of minimal spreading and a lack of condensation (**Fig. 23**).

These conditions were used as training data to optimize the network for organoid classification. 34 of the 48 images in each class were used for training while the remainder were used for validation. Network training was accomplished in 2 minutes with a training accuracy of

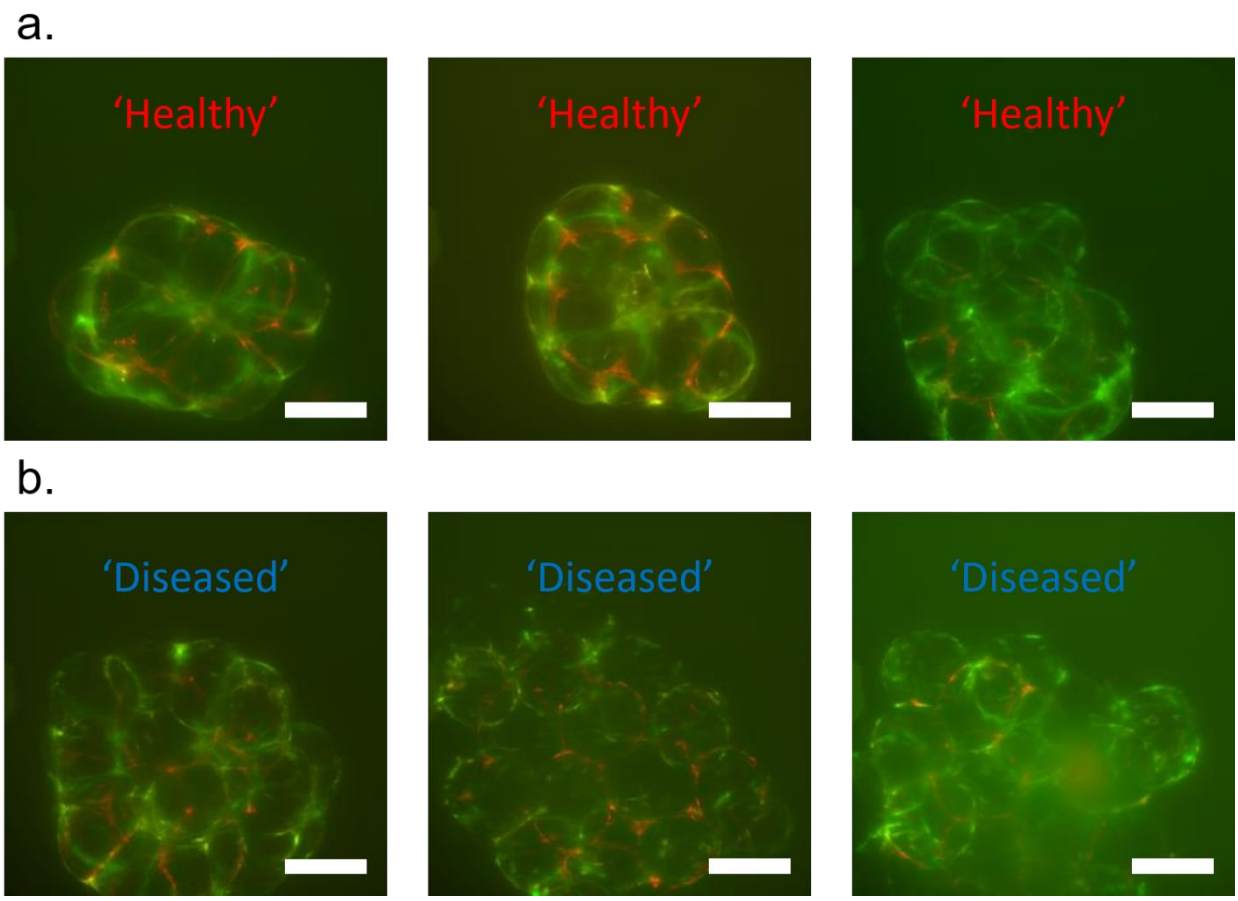


Figure 23: Representative organoids used as data for training the neural network. (**a**) Training data selected from the 0% DMSO condition, labeled 'healthy.' (**b**) Training data selected from the 3.5% DMSO condition, labeled 'diseased.'

99% and a validation accuracy of 100%. The binary cross entropy loss was calculated to be 2.9×10^{-3} for the training data and 1.1×10^{-7} for the validation data. Given the general uniformity of the data, especially when compared with ImageNet-like datasets, the accuracies observed here are not surprising though may be indicative of overfitting. When viewed on a plate-wide level the classification results show that most of the 'healthy' classed organoids have lower DMSO concentrations whereas the 'disease' classed organoids were given higher concentrations of DMSO(**Fig. 24a**). Furthermore, when the average values are calculated along every row there is a clear transition point wherein, on average, organoids ceased being classed as 'healthy' and begin grouping with the 'diseased' category(**Fig. 24b**). This may then be used as the neural network's approximation for how much DMSO may be administered to an organoid while still maintaining it 'healthy' categorization.

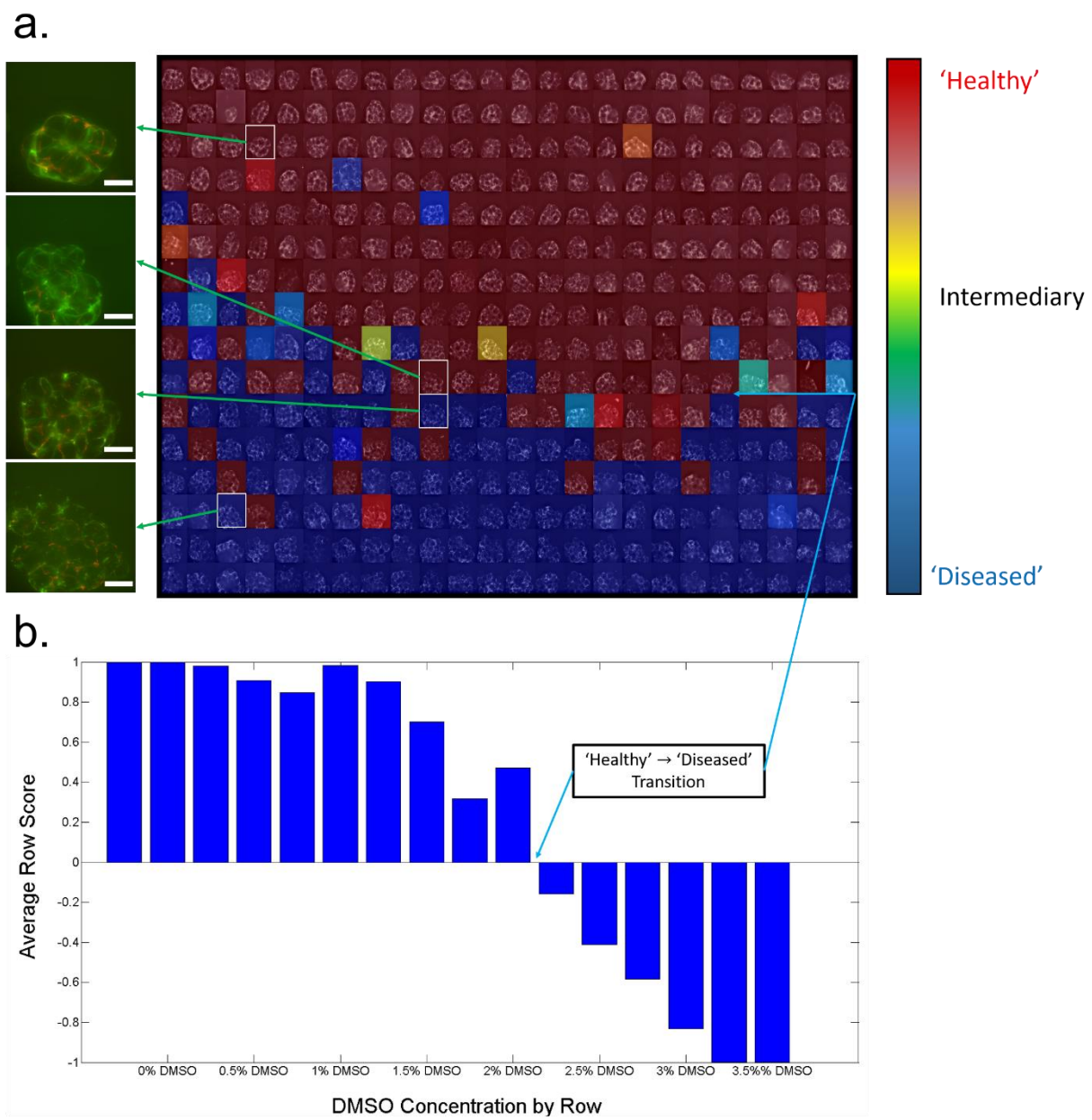


Figure 24: Results of convolutional network classification on the DMSO stress test dataset. (**a**) Well plate images are color coded to indicate classification wherein red denotes 'healthy' and blue 'diseased.' Select images are highlighted for observation. (**b**) Bar graph of averaged row scores indicates a clear shift in categorization when transitioning from 2% to 2.5% DMSO.

Chapter 4.4. Discussion

We employed a convolutional neural network to characterize and classify the effects of DMSO on organoid formation. In the context of an IPF drug screen, organoid formation and

collagen deposition are critical parameters for scar formation and DMSO may confound experimental results when used as a solvent. We found that DMSO dramatically altered the organoid phenotype at higher concentrations and, after network classification, found that organoids treated with DMSO up to 2% vol. grouped with the 0% DMSO control.

The use of artificial neural networks for classification of phenotypic, high throughput data shows great promise. These networks are at the forefront of machine learning and are current state of the art technologies in a wide range of industries. Though it is important to understand the limitations of what this type of classification can, and cannot do. Due to network complexity and the learning process it is difficult to extract biologically relevant answers using the learning and classification architecture provided here. For example, the network does not provide a cell count, or a quantification of cellular aspect ratio or marker expression profile, though one may envision a method for training a network to do just that. At a very basic level this process is one of pattern matching and therefore requires good, curated training data for the method to yield relevant results. In other words, if the training data i.e. experimental controls do not represent biologically relevant phenotypes then this method will fail to provide meaningful categorization.

Though, with the strengths and weaknesses of artificial neural networks in mind it is now possible to begin modifying standard drug screening protocols to be amenable with machine learning. Specifically, with the proper acquisition and labeling of training data a researcher can train a simple neural network to find those experimental instances of interest without the laborious and time-consuming process of designing a custom algorithm for every study. Furthermore, instances of the same neural network can be trained to analyze and categorize a wide, disparate range of data making neural networks a universal solution to phenotypic categorization.

The neural network described here is tailored for binary categorization and performed well given the severely limited amount of training data available. As this dataset was quite uniform it is unclear if there will be problems with overfitting (generalization of this classification technique to similar lung organoid datasets). Minimizing overfitting and maximizing generalization is of key importance for using these techniques successfully and can be accomplished by increasing the amount and dimensionality of the training data.

Chapter 4.5. Conclusions

Convolutional neural network based approaches may be used in a wide variety of phenotypic classification schemes. In order to further advance the technology, it is necessary to investigate different network architectures and their applicability to various high throughput data. In addition, it may be important to incorporate a user feedback system wherein expert users grade the output of trained networks and increase the total amount of training data to ensure good results[87].

Chapter 5. Conclusions

The future of disease modeling and drug discovery is likely to take place using multicellular, stem cell-derived organoid cultures. The advent of induced pluripotent stem cells and the continued progress in organoid formation techniques has changed the landscape of biological inquisition and is paving the way for new scientific insights and, hopefully, new therapies. These advances are important since the current, state of the art disease modeling methodologies are failing to find new treatments for idiopathic diseases. IPF, for example, is a multifactorial, heterogenous disease that has been difficult to model due to the many interconnected genetic and environmental variables that are likely the cause of the disease initiation and progression. Though, using patient-derived cells and modeling the disease in 3D there is a better chance at making true mechanistic insights, or, at least, finding individualized therapies that help patients.

The work presented here outlines methods designed for the implementation and analysis of a 3D, high throughput drug screen of a novel model of idiopathic pulmonary fibrosis. The basis of this technology is the development of the lung organoid, a multicellular tissue construct that replicates the geometry of human distal lung. We have shown that this lung organoid formulation is amenable to the inclusion of multiple cell types and is scalable in both organoid size and number of organoids generated. Lung organoids were treated with TGF-β1 in order to initiate a fibrotic phenotype and we observed scarring characteristic to that observed in histological samples of IPF lung. Furthermore, we showed that these methods are amenable with iPSC-derived cells making these techniques personalized. Finally, we developed an artificial neural network for the classification of the complex phenotypes present when treating organoids with DMSO.

It is unclear if this method for modeling IPF will bear fruit for patients. It is a minimalistic approach geared for high throughput drug discovery and may be missing some of factors important for finding relevant drugs for patients. Specifically, epithelial and immune cell types missing from the disease model may be necessary for true phenotypic, and underlying mechanistic, mimicry. In addition, while supplementation of TFG-β1 has shown to initiate the fibrotic phenotype observed in the dish, it is uncertain if compounds that inhibit TGF-β1 (likely hits from a high throughput screen) will be helpful to patients. Though, despite these uncertainties, we have pushed forward and hope that these methods will one day be helpful in identifying new theriputic compounds for those suffering from IPF.

The final section of this thesis outlined a machine learning method for characterizing and classifying the large datasets that are inherent to high throughput drug screening. We applied a convolutional neural network to the task of classifying the complex phenotypes observed in our organoid cultures upon the supplementation of DMSO. These results have shown promise at this task and will likely be useful in quickly categorizing large datasets, irregardless of the underlying biology and phenotypic presentation of the dataset in question.

Chapter 6. Future Directions

Chapter 6.1. Overview

There are two important areas of advancement of the lung organoid technology. First, the lung is a highly vascular organ and the inclusion of a vascular compartment may be necessary for modeling disease. Finally, the carrying out of a drug screen using the established IPF model may provide novel compounds for the inhibition of the progressive fibrosis observed in IPF patients.

Chapter 6.2. Organoid Vascularization

We have proposed to design and fabricate a fluidic perfusion system with the goal of vascularizing the lung organoid. This will be achieved using a syringe pump driven PDMS chamber that provides a flow of media around the organoid exterior (**Fig. 25**). In addition, this system will allow for *in situ* microscopy for the visualization and quantification of angiogenesis.

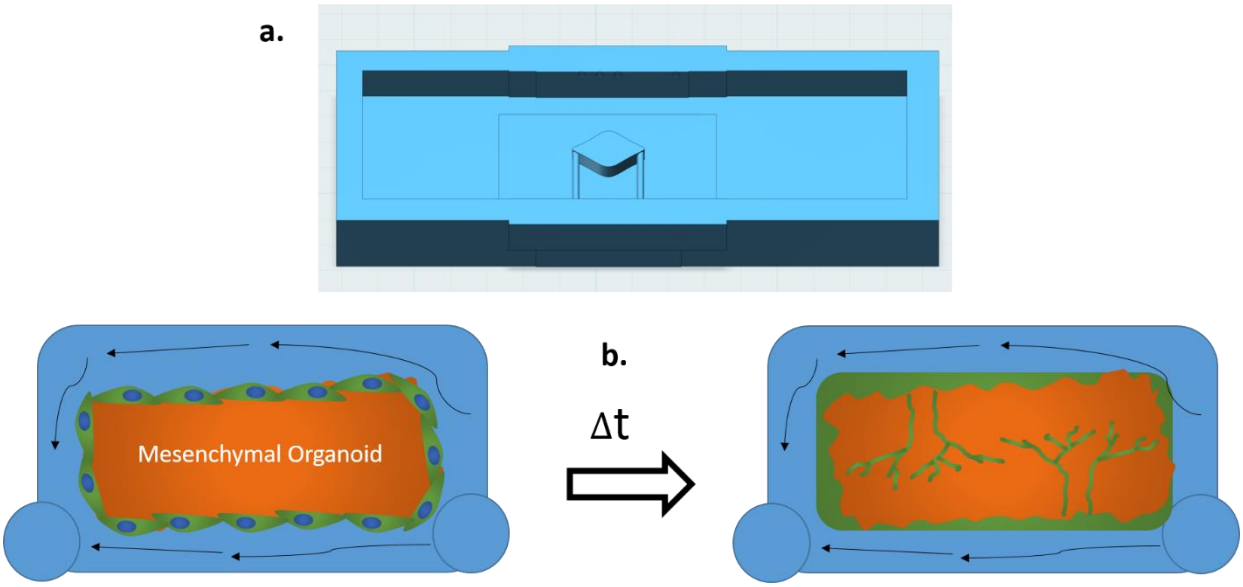


Figure 25: The proposed organoid perfusion system and hypothesized organoid angiogenesis. Endothelial cells will be coated onto the outside of a mesenchymal organoid using the HARV bioreactor and the resulting tissue placed in the chamber. Media will then be perfused through the chamber. (**a**) CAD model of the PDMS mold. (**b**) Hypothesized flow characteristics and organoid vascularization.

First, endothelial cells will be coated onto the organoid exterior by placing a mature mesenchymal organoid and a solution of endothelial cells into the HARV bioreactor and rotating at a rate tuned to allow for cellular adhesion. The resulting tissue will be place in the fluidic chamber and perfused with media. We hypothesize that the cells in the center of the organoid will release and establish an angiogenic growth factor gradient that drives endothelial cells on the organoid exterior to invade the organoid thereby forming vasculature. If cellular production of growth factor is not sufficient exogenous vascular endothelial growth factor may be added to the alginate beads[88] supplementing the system and driving angiogenesis. We will then verify the functionality of these newly formed endothelial tubes by perfusing a solution of fluorescently labeled beads into the chamber and verifying that these beads successfully enter the organoid interior.

Chapter 6.3. Drug Screening for Compounds that Inhibit Fibrosis

An overarching goal of the lung organoid system is to meet the promise of personalized medicine by executing an efficient strategy to accelerate the drug discovery process for IPF (Fig. 1). We propose to carry out a drug screen for FDA approved that inhibit the fibrotic metrics of contraction and α-SMA expression. Finding new uses for old drugs (drug repurposing) is the approach we aim to undertake[89]. These include drugs that may have (i) passed safety, but failed efficacy studies in late stage clinical trials; (ii) stalled in development for commercial reasons; (iii) crossed the point of patent expiry; or (iv) are being expanded into new geographical markets. Since the repurposing drugs has been previously fully evaluated, safety data is already established and hence these drugs can be brought to the market much quicker. Given the lack of effective therapies for IPF, we think this would be the most effective approach. We aim to use our 3D disease model

to screen 2560 high-value compounds from two libraries – The Prestwick and LOPAC®1280 libraries.

References

1. Yamanaka, S. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–72 (2007).

2. Wilkinson, D. C. *et al.* Development of a Three-Dimensional Bioengineering Technology to Generate Lung Tissue for Personalized Disease Modeling. *Stem Cells Transl. Med.* **1,** 1–12 (2016).

3. Sucre, J. M. S. *et al.* A 3-Dimensional Human Model of the Fibroblast Activation that Accompanies Bronchopulmonary Dysplasia Identifies Notch-Mediated Pathophysiology. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* (2016). doi:10.1152/ajplung.00446.2015

4. Sucre, J. M. S. *et al.* Post-Translational Modification of β-Catenin is Associated with Pathogenic Fibroblastic Changes in Bronchopulmonary Dysplasia. *Am. J. Physiol. Lung Cell. Mol. Physiol.* (2016). doi:10.1152/ajplung.00477.2016

5. NIH & NHLBI. Chapter 4 Disease Statistics. *NHLBI Fact Book, Fiscal Year 2012* 33–52 (2012). Available at: https://www.nhlbi.nih.gov/about/factbook/chapter4.htm#gr36%5Cnhttp://www.nhlbi.nih.gov/about/documents/factbook/2012.

6. Tashkin, D. P., Klein, G. L., Colman, S. S., Zayed, H. & Schonfeld, W. H. Comparing COPD Treatment: Nebulizer, Metered Dose Inhaler, and Concomitant Therapy. *Am. J. Med.* **120,** 435–441 (2007).

7. Stoller, J. K., Panos, R. J., Krachman, S., Doherty, D. E. & Make, B. Oxygen therapy for patients with COPD: Current evidence and the long-term oxygen treatment trial. *Chest*

138, 179–187 (2010).

8. Wallis, A. The diagnosis and management of interstitial lung disease. *Bmj* **350,** 555–556 (2015).

9. Selman, M. *et al.* Idiopathic pulmonary fibrosis. *Drugs* **67,** 405–430 (2004).

10. Cavazza, A. *et al.* The role of histology in idiopathic pulmonary fibrosis: An update. *Respir. Med.* **104,** S11–S22 (2010).

11. Studer, S. M. & Kaminski, N. Towards systems biology of human pulmonary fibrosis. *Proc. Am. Thorac. Soc.* **4,** 85–91 (2007).

12. Selman, M., Pardo, A. & Kaminski, N. Idiopathic pulmonary fibrosis: Aberrant recapitulation of developmental programs? *PLoS Med.* **5,** 0373–0380 (2008).

13. Gardet, A., Zheng, T. S. & Viney, J. L. Genetic architecture of human fibrotic diseases: Disease risk and disease progression. *Front. Pharmacol.* **4 DEC,** 1–12 (2013).

14. Fernandez, I. E. & Eickelberg, O. Interstitial Lung Disease 1 New cellular and molecular mechanisms of lung injury and fi brosis in idiopathic pulmonary fibrosis. *Lancet* **380,** 680–88 (2012).

15. Fingerlin, T. E. *et al.* Loci for Pulmonary Fibrosis. **45,** 613–620 (2013).

16. Birgersdotter, A., Sandberg, R. & Ernberg, I. Gene expression perturbation in vitro--a growing case for three-dimensional (3D) culture systems. *Semin. Cancer Biol.* **15,** 405–12 (2005).

17. Pampaloni, F., Reynaud, E. G. & Stelzer, E. H. K. The third dimension bridges the gap

between cell culture and live tissue. *Nat. Rev. Mol. Cell Biol.* **8,** 839–845 (2007).

18. Baharvand, H., Hashemi, S. M., Ashtiani, S. K. & Farrokhi, A. Differentiation of human embryonic stem cells into hepatocytes in 2D and 3D culture systems in vitro. *Int. J. Dev. Biol.* **50,** 645–652 (2006).

19. Dutta, D., Heo, I. & Clevers, H. Disease Modeling in Stem Cell-Derived 3D Organoid Systems. *Trends Mol. Med.* **23,** 393–410 (2017).

20. Stabler, C. T. & Morrisey, E. E. Developmental pathways in lung regeneration. *Cell Tissue Res.* **367,** 677–685 (2017).

21. Barkauskas, C. E. *et al.* Lung organoids: current uses and future promise. *Development* **144,** 986–997 (2017).

22. Dye, B. R. *et al.* In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* **4,** 1–25 (2015).

23. Nichols, J. E. & Cortiella, J. Engineering of a complex organ: progress toward development of a tissue-engineered lung. *Proc. Am. Thorac. Soc.* **5,** 723–30 (2008).

24. Nichols, J. E., Niles, J. A. & Cortiella, J. Design and development of tissue engineered lung: Progress and challenges. *Organogenesis* **5,** 57–61 (2009).

25. Takebe, T. *et al.* Vascularized and Complex Organ Buds from Diverse Tissues via Mesenchymal Cell-Driven Condensation. *Cell Stem Cell* **16,** 556–565 (2015).

26. Franzdóttir, S. R. *et al.* Airway branching morphogenesis in three dimensional culture. *Respir. Res.* **11,** 162 (2010).

27. Nadkarni, R. R., Abed, S. & Draper, J. S. Organoids as a model system for studying human lung development and disease. *Biochem. Biophys. Res. Commun.* **473,** 675–682 (2016).

28. Hughes, C. S., Postovit, L. M. & Lajoie, G. A. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* **10,** 1886–1890 (2010).

29. Mondrinos, M. J. *et al.* Engineering three-dimensional pulmonary tissue constructs. *Tissue Eng.* **12,** 717–728 (2006).

30. Barrila, J. *et al.* Organotypic 3D cell culture models: using the rotating wall vessel to study host-pathogen interactions. *Nat. Rev. Microbiol.* **8,** 791–801 (2010).

31. Kojima, K. & Vacanti, C. a. Tissue engineering in the trachea. *Anat. Rec. (Hoboken).* **297,** 44–50 (2014).

32. Mondrinos, M. J. *et al.* Engineering three-dimensional pulmonary tissue constructs. *Tissue Eng.* **12,** 717–28 (2006).

33. Weiss, D. J. Concise Review: current status of stem cells and regenerative medicine in lung biology and diseases. *Stem Cells* **32,** 16–25 (2014).

34. Dunn, J. C. Y. *et al.* Analysis of cell growth in three-dimensional scaffolds. *Tissue Eng.* **12,** 705–16 (2006).

35. Sam, M., Raredon, B., Calle, E. a & Niklason, L. E. A rotating bioreactor for scalable culture and differentiation of respiratory epithelium. 1–35 (2014). doi:10.3727/215517914X681794

36. Cummings, L. J., Sawyer, N. B. E., Morgan, S. P., Rose, F. R. a J. & Waters, S. L.

Tracking large solid constructs suspended in a rotating bioreactor: A combined experimental and theoretical study. *Biotechnol. Bioeng.* **104,** 1224–1234 (2009).

37. McGonigle, P. & Ruggeri, B. Animal models of human disease: Challenges in enabling translation. *Biochem. Pharmacol.* **87,** 162–171 (2014).

38. Miki, H. *et al.* Fibroblast Contractility. *Am. J. Respir. Crit. Care Med.* **162,** 2259–2264 (2000).

39. Pierce, E. M. *et al.* Idiopathic pulmonary fibrosis fibroblasts migrate and proliferate to CC chemokine ligand 21. *Eur. Respir. J.* **29,** 1082–1093 (2007).

40. Xia, H. *et al.* Pathological integrin signaling enhances proliferation of primary lung fibroblasts from patients with idiopathic pulmonary fibrosis. *J. Exp. Med.* **205,** 1659–1672 (2008).

41. Peng, R. *et al.* Bleomycin Induces Molecular Changes Directly Relevant to Idiopathic Pulmonary Fibrosis: A Model for 'Active' Disease. *PLoS One* **8,** (2013).

42. Moore, B. B. *et al.* Animal models of fibrotic lung disease. *Am. J. Respir. Cell Mol. Biol.* **49,** 167–179 (2013).

43. Moeller, A., Ask, K., Warburton, D., Gauldie, J. & Kolb, M. The bleomycin animal model: A useful tool to investigate treatment options for idiopathic pulmonary fibrosis? *Int. J. Biochem. Cell Biol.* **40,** 362–382 (2008).

44. Paul, S. M. *et al.* How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **9,** 203–14 (2010).

45. Barnes, P. J. *et al.* Barriers to new drug development in respiratory disease. *Eur. Respir. J.*

45, 1197–1207 (2015).

46. Agusti, A. *et al.* Treatable traits: Toward precision medicine of chronic airway diseases. *Eur. Respir. J.* **47,** 410–419 (2016).

47. Davis, P. B., Yasothan, U. & Kirkpatrick, P. Ivacaftor. *Nat. Rev. Drug Discov.* **11,** 349–350 (2012).

48. Cheng, S. H. *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **63,** 827–834 (1990).

49. Wang, F. Y. *et al.* Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA Journal of Automatica Sinica* **3,** 113–120 (2016).

50. Hecht-Nielsen, R. Theory of the Backpropagation Neural Network. *Proc. Int. Jt. Conf. Neural Networks* **1,** 593–605 (1989).

51. Chung Tsoi, A. & Scarselli, F. Universal Approximation Using Feedforward Neural Networks: A Survey of Some Existing Methods, and Some New Results. *Neural Networks* **11,** 15–37 (1998).

52. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *arXiv Prepr. arXiv1512.03385v1* **7,** 171–180 (2015).

53. Hubel, D. H. & Weisel, T. N. RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX By. *J. Physiol.* **148,** 574–591 (1959).

54. Calle, E. A. *et al.* Fate of Distal Lung Epithelium Cultured in a Decellularized Lung Extracellular Matrix. *Tissue Eng. Part A* **21,** 1916–1928 (2015).

55. Ren, X. *et al.* Engineering pulmonary vasculature in decellularized rat and human lungs. *Nat. Biotechnol.* **33,** 1097–1102 (2015).

56. Andrade, C. F., Wong, A. P., Waddell, T. K., Keshavjee, S. & Liu, M. Cell-based tissue engineering for lung regeneration. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **292,** L510–L518 (2007).

57. Singh, D., Zo, S. M., Kumar, A. & Han, S. S. Engineering three-dimensional macroporous hydroxyethyl methacrylate-alginate-gelatin cryogel for growth and proliferation of lung epithelial cells. *J. Biomater. Sci. Polym. Ed.* **24,** 1343–1359 (2013).

58. Gill, B. J. *et al.* A synthetic matrix with independently tunable biochemistry and mechanical properties to study epithelial morphogenesis and EMT in a lung adenocarcinoma model. *Cancer Res.* **72,** 6013–6023 (2012).

59. Zhang, W.-J. *et al.* The reconstruction of lung alveolus-like structure in collagen-matrigel/microcapsules scaffolds in vitro. *J. Cell. Mol. Med.* **15,** 1878–1886 (2011).

60. Kaisani, A. *et al.* Branching morphogenesis of immortalized human bronchial epithelial cells in three-dimensional culture. *Differentiation* **87,** 119–126 (2014).

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

62. Lee, J. *et al.* Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-Thrombospondin-1 axis. *Cell* **156,** 440–455 (2014).

63. Gotoh, S. *et al.* Generation of Alveolar Epithelial Spheroids via Isolated Progenitor Cells from Human Pluripotent Stem Cells. *Stem Cell Reports* **3,** 394–403 (2014).

64. Teisanu, R. M. *et al.* Functional Analysis of Two Distinct Bronchiolar Progenitors during Lung Injury and Repair. *Am. J. Respir. Cell Mol. Biol.* **44,** 794–803 (2011).

65. Samuel, C. S. in *Methods in Molecular Biology* **466,** 223–235 (2009).

66. Hunt, N. C. & Grover, L. M. Cell encapsulation using biopolymer gels for regenerative medicine. *Biotechnol. Lett.* **32,** 733–742 (2010).

67. Rowley, J. a, Madlambayan, G. & Mooney, D. J. Alginate hydrogels as synthetic extracellular matrix materials. *Biomaterials* **20,** 45–53 (1999).

68. Lee, H., Dellatore, S. M., Miller, W. M. & Messersmith, P. B. Mussel-inspired surface chemistry for multifunctional coatings. *Science* **318,** 426–30 (2007).

69. Yu, X., Wei, M. & Walsh, J. Covalent Immobilization of Collagen on Titanium through Polydopamine Coating to Improve Cellular Performances of MC3T3-E1 Cells. *RSC Adv.* **4,** 7185–7192 (2014).

70. Wilkinson, D. C. *et al.* Development of a Three-Dimensional Bioengineering Technology to Generate Lung Tissue for Personalized Disease Modeling. *Stem Cells Transl. Med.* (2016). doi:10.5966/sctm.2013-0145

71. Nichols, J. E. & Cortiella, J. Engineering of a complex organ: progress toward development of a tissue-engineered lung. *Proc. Am. Thorac. Soc.* **5,** 723–30 (2008).

72. Hubbard, R. The burden of lung disease. *Thorax* **61,** 558–559 (2006).

73. Rackley, C. R. & Stripp, B. R. Building and maintaining the epithelium of the lung. *J. Clin. Invest.* **122,** 2724–2730 (2012).

74. Selman, M., King, T. E. & Pardo, A. Review Idiopathic Pulmonary Fibrosis : Prevailing and Evolving Hypotheses about Its Pathogenesis and Implications for Therapy. *Ann. Intern. Med.* **134,** 136–151 (2001).

75. Carloni, a., Poletti, V., Fermo, L., Bellomo, N. & Chilosi, M. Heterogeneous distribution of mechanical stress in human lung: A mathematical approach to evaluate abnormal remodeling in IPF. *J. Theor. Biol.* **332,** 136–140 (2013).

76. Wynn, T. & Ramalingam, T. Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nat. Med.* **18,** 1028–1040 (2013).

77. Seok, J. *et al.* Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 3507–3512 (2013).

78. Ursu, A., Schöler, H. R. & Waldmann, H. Small-molecule phenotypic screening with stem cells. *Nat. Chem. Biol.* **13,** 560–563 (2017).

79. Sommer, C. & Gerlich, D. W. Machine learning in cell biology – teaching computers to recognize phenotypes. *J. Cell Sciience* **126,** 5529–5539 (2013).

80. Shariff, A., Kangas, J., Coelho, L. P., Quinn, S. & Murphy, R. F. Automated Image Analysis for High-Content Screening and Analysis. *J. Biomol. Screen.* **15,** 726–734 (2010).

81. Kraus, O. Z. & Frey, B. J. Computer vision for high content screening. *Crit. Rev. Biochem. Mol. Biol.* **9238,** 1–8 (2016).

82. Richardson, W. Materials Science Tools for Regenerative Medicine. (UCLA, 2013).

83. Bellare, M. & Rogaway, P. Information Processing in Medical Imaging. *Lect. Notes*

Comput. Sci. 3565 - 19th Int. Conf. **3565,** 399–416 (2005).

84. Dürr, O. & Sick, B. Single-Cell Phenotype Classification Using Deep Convolutional Neural Networks. *J. Biomol. Screen.* 3–8 (2016). doi:10.1177/1087057116631284

85. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* 1–9 (2012). doi:http://dx.doi.org/10.1016/j.protcy.2014.09.007

86. Sari, E., Bakar, B., Dincel, G. C. & Budak Yildiran, F. A. Effects of DMSO on a rabbit ear hypertrophic scar model: A controlled randomized experimental study. *J. Plast. Reconstr. Aesthetic Surg.* **70,** 509–517 (2017).

87. Jones, T. R. *et al.* Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc. Natl. Acad. Sci.* **106,** 1826–1831 (2009).

88. Jay, S. M. & Saltzman, W. M. Controlled delivery of VEGF via modulation of alginate microparticle ionic crosslinking. *J. Control. Release* **134,** 26–34 (2009).

89. Oprea, T. I. & Mestres, J. Drug Repurposing: Far Beyond New Targets for Old Drugs. *AAPS J.* **14,** 759–763 (2012).